Jan-Christoph Otto
Richard Dikau

LECTURE NOTES IN EARTH SCIENCES

Landform - Structure, Evolution, Process Control

Proceedings of the International Symposium on Landform organised by the Research Training Group 437

Springer

Lecture Notes in Earth Sciences 115

Jan-Christoph Otto · Richard Dikau
Editors

Landform – Structure, Evolution, Process Control

Proceedings of the International Symposium on Landform Organised by the Research Training Group 437

Springer

Editors
Dr. Jan-Christoph Otto
Universität Salzburg
Fachbereich Geographie und Geologie
Hellbrunnerstr. 34
5020 Salzburg
Austria
jan-christoph.otto@sbg.ac.at

Prof. Dr. Richard Dikau
Universität Bonn
Geographische Institut
Meckenheimer Allee 172
53115 Bonn
Germany
rdikau@giub.uni-bonn.de

ISSN 0930-0317
ISBN 978-3-540-75760-3 e-ISBN 978-3-540-75761-0
DOI 10.1007/978-3-540-75761-0
Springer Heidelberg Dordrecht London New York

Library of Congress Control Number: 2009936209

Cover design: Integra Software Services Pvt. Ltd., Pondicherry

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Landforms constitute boundary surfaces between different components of the earth system (atmosphere, hydrosphere, biosphere, pedosphere, lithosphere). At these locations most of the human activity on earth takes place. This central position evokes a bi-directional interaction with the other spheres of the earth system. Spatial landform structures strongly affect processes of other earth system components. At the same time, the land-surface is shaped by the influence of these processes impacting geomorphologic processes and landform morphometry.

These interactions are the focus in the Research Training Group 437 "Landform – a structured and variable boundary layer" at the University of Bonn in Germany. Funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) the Research Training Group is a multidisciplinary research programme for postgraduate studies. Disciplines involved in this programme include: biology, climatology, computer sciences, geodynamics, geology, geomorphology, geophysics, hydrology, mathematics, meteorology, pedology, and remote sensing. These different disciplines offer various scientific approaches, theories, methods and data for the study of landforms within their specific paradigms. Over a period of ten years (1998–2008) more than 25 PhD projects have been completed.

Dedicated to ongoing and completed research activities of the Research Training Group an international symposium titled "Landform – structure, evolution, process control" was held at the Department of Geography, University of Bonn, in 2007. The meeting brought together young scientists and senior experts of the various disciplines involved. Five major topics structured the presentations, each opened by 2 invited keynote lectures of international experts including

- Landform structure in Earth System Science
- 3D-Modelling, processing and visualisation of landform
- Landform and fluxes in atmospheric and hydrospheric systems
- Sediment fluxes and storage in periglacial and hillslope environments
- Landform structure and resilience in ecosystems

One major aim of the symposium was the communication of new approaches, results and questions emerging from current research in relation to landforms and to its importance for earth system processes. More than 80 participants from various

countries attended the meeting creating a lively atmosphere of fruitful discussions and exciting presentations. Some of the contributions to this symposium have been summarised in this book, including both invited keynote papers and contributions by projects from the Research Training Group.

The contents of this book reveal the multidisciplinary of landform research. The scope of the papers spreads from landform modelling and cartographic visualisation, atmospheric and hydrologic interferences with landform to issues of sediment fluxes and ecology.

Acknowledgments

The symposium was funded by the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) that enabled the invitation of experts for keynote lectures. We are very grateful for this support. The compilation of this book would not have been possible without the kind help from various reviewers. We would like to thank the following colleagues for their valuable comments and suggestions that considerably contributed to the quality of the papers presented: Fabien Arnaud, Walter Bauer, Günter Blöschl, Hans-Jürgen Böhmer, Andreas Bott, Tobias Bolch, Ian S. Evans, Simone Giertz, Martin Hölzle, Anke Veronika Jentsch, Hermann Klug, Roger A. Pielke Sr., Jochen Schmidt, Herbert Weingartner.

Finally, many thanks to the authors of this book who showed great patience and understanding during the publication process of this book.

Salzburg, Austria — Jan-Christoph Otto
Bonn, Germany — Richard Dikau

Contents

Contributors

Isabel Alberts Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany, ialberts@uni-bonn.de

Felix Ament Meteorologisches Institut, Universität Hamburg, Bundesstraße 55, D-20146, felix.ament@zmaw.de

Günter Blöschl Institut für Wasserbau und Ingenieurhydrologie, Technische Universität Wien, Karlsplatz 13/222, A-1040 Vienna, Australia, bloeschl@hydro.tuwien.ac.at

Andreas Bott Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany, a.bott@uni-bonn.de

Benjamin Burkhard Ecology Centre, University of Kiel, Olshausenstrasse 75, 24118, Kiel, Germany, bburkhard@ecology-uni-kiel.de

Bernd Diekkrüger Department of Geography, University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany, b.diekkrueger@uni-bonn.de

Carsten Ebmeyer Mathematical Seminar, Meteorological Institute, Nußallee 15, D-53115 Bonn, Germany, ebmeyer.c@uni-bonn.de

René Graßelt Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany, grasselt@uni-bonn.de

Herwig Hölzel Department of Geography, University of Bonn, Meckenheimer Allee 166, 53115 Bonn, Germany, herwighoelzel@giub.uni-bonn.de

Lorenz Hurni Institut f. Kartographie, ETH-Hönggerberg, Wolfgang-Pauli-Str. 15, 8093 Zürich, Switzerland, lorenz.hurni@karto.baug.ethz.ch

Reinhard Klein Institute of Computer Science II, University of Bonn, Römerstraße 164, 53117 Bonn, Germany, rk@cs.uni-bonn.de

Michael Krautblatter Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany, michael.krautblatter@giub.uni-bonn.de

Franziska Kroll Ecology Centre, University of Kiel, Olshausenstrasse 75, 24118, Kiel, Germany, fkroll@ecology.uni-kiel.de

Jörg Löffler Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany, joerg.loeffler@uni-bonn.de

Marc-Oliver Löwner Institut für Geodäsie und Photogrammetrie University of Braunschweig, Gaußstraße 22, 38106 Braunschweig, Germany, m-o.loewner@tu-braunschweig.de

Matthieu Masbou Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany, mmasbou@uni-bonn.de

Ralf Merz Institute for Hydraulic and Water Resources Engineering, Vienna University of Technology, Karlsplatz 13/222, 1040 Vienna, Austria, merz@hydro.tuwien.ac.at

Felix Müller Department of Ecosystem Research, Ecology Centre, University of Kiel, Olshausenstr. 75, 24118 Kiel, Germany, fmueller@ecology.uni-kiel.de

Dev Niyogi Departments of Agronomy and Earth & Atmos. Sciences, Purdue University, 915 W. State Street, West Lafayette, IN 47907-2054, USA, climate@purdue.edu

Michael Nyenhuis Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany, michael.nyenhuis@uni-bonn.de

Roger Pielke Sr. CIRES, University of Colorado, Stadium 255-16, Boulder, CO 80309, USA, pielkesr@cires.colorado.edu

Isabelle Gärtner-Roer Department of Geography, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland, iroer@geo.uzh.ch

Ole Rößler Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany, o.roessler@giub.uni-bonn.de

Susanne Schmidt Südasien-Institut, University of Heidelberg, Im Neuenheimer Feld 330, D-69120 Heidelberg, Germany, s.schmidt@sai.uni-heidelberg.de

Martin Schneider Institute of Computer Science II, University of Bonn, Römerstraße 164, 53117 Bonn, Germany, ms@cs.uni-bonn.de

Clemens Simmer Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany, csimmer@uni-bonn.de

Jens Vogelgesang Institute for Applied Analysis, Endenicheralle 60, 53115 Bonn, Germany, vjens@uni-bonn.de

Jeff Warburton Department of Geography, Durham University, Science Laboratories, South Road, Durham DH1 3LE, UK, jeff.warburton@durham.ac.uk

Kirsten Warrach-Sagi Universitaet Hohenheim, Institut fuer Physik und Meteorologie (120), Garbenstr. 30, D-70599 Stuttgart, Germany, warrach@uni-hohenheim.de

Cartographic Relief Presentation Revisited – Forty Years after Eduard Imhof

Lorenz Hurni

Abstract Topographic maps represent in a symbolised way the main features of the Earth's surface shape as well as the major, mostly visible objects covering the topography. The clear and readable depiction of the relief is one of the main challenges in topographic cartography. The Swiss cartographer Eduard Imhof had a significant influence on the development of modern topographic cartography. He published his trend-setting findings 40 years ago in his textbook "Cartographic Relief Presentation". This overview paper reviews his major contributions to topographic cartography and presents today's state of cartographic relief depiction, illustrated by current map examples and projects elaborated at Imhof's Alma Mater, the Institute of Cartography of ETH Zurich.

Keywords Eduard Imhof · Cartography · Topographic maps · Relief presentation · Contour lines · Hill shading · Analytical shading · Swiss-style colour relief shading · Rock drawing

1 The Roots of Eduard Imhof's Scientific and Educational Work

Eduard Imhof (1895–1986, Fig. 1) was Lecturer and Professor of Cartography at the Swiss Federal Institute of Technology (ETH Zurich) from 1925 to 1965. After his diploma graduation in 1919, he was immediately appointed Lecturer to replace Prof. Fridolin Becker who was already severely ill at that time. After Becker's death in 1924, Imhof was promoted to Professor in 1925. One of his first actions was the foundation of the Institute of Cartography, thus becoming the first academic cartographic research institute world-wide. According to the well-cultivated legend (Imhof 1990), Imhof simply asked a painter to write the designation "Institute of Cartography" above the entrance door of his office!

L. Hurni (✉)
Institut f. Kartographie, ETH-Hönggerberg, Wolfgang-Pauli-Str. 15, 8093 Zürich, Switzerland
e-mail: lorenz.hurni@karto.baug.ethz.ch

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_1,

Fig. 1 Eduard Imhof (ca. 1925)

Imhof was educated as a as geodesist/land surveyor at ETH Zurich. At the beginning of the 20th century, the focus of the studies was much laid on geodetic research questions, especially on the set up of new National Reference Systems. The Federal Office of Topography, the Swiss National Mapping Authority, had defined a new geodetic reference system in 1903 which replaced an earlier system from the mid-19th century. Based on this framework, a new triangulation of the whole country was underway. These new reference points were first used for a new cadastral surveying campaign and the introduction of a land register which was demanded by a new civil legislation in 1903. The mapping of settled and built areas of the country

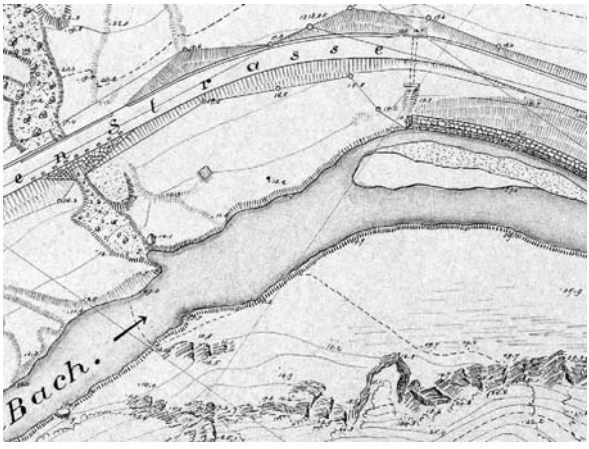

Fig. 2 Plane table compilation of the Klausenpass road. ETH diploma thesis by O. Giger, A. Weber and K. Tobler, 1903

at large scales was therefore just about to start. Figure 2 shows an example of a large scale map of the Klausenpass road in Central Switzerland, established in the framework of a diploma thesis by ETH students in 1903, thus shortly before Imhof's study period. Besides a very precise compilation of the man-made and topographic features applying triangulation and plane table techniques, the map already shows a high graphical quality in the depiction of the topographic and especially of the rock features.

At the time of Imhof's studies, tests were initiated for a new National Map Series (based on the new geodetic reference system) to replace the old Dufour and Siegfried maps from the 19th century. The introduction of the new National Map however started only in 1938, and ended 1978. Imhof had a major influence on the definition and the design of the new maps. He especially urged to introduce a metric map scale series consisting of scales of 1:25,000, 1:50,000, 1:100,000, 1:200,000, 1:500,000 and 1:1,000,000. A fortunately unsuccessful counterproposal postulated a scale of 1:33,333, replacing both the scales 1:25,000 and 1:50,000 (Imhof 1979). Military purposes were the main reason for producing these new maps, especially after the experience of World War I. However, all topographic maps are designed to fulfil multiple purposes such as for leisure, planning, scientific applications, etc. The choice of map elements and their design represent clearly a compromise in order to match as many needs as possible. For scientific applications, e.g., in geomorphology, classic topographic maps may serve as basic source, describing topographic features, and as base maps for field work and publications.

Besides geometrical aspects of topographic mapping Imhof also devoted a considerable part of his scientific work to cartographic design. Imhof's teacher and predecessor Fridolin Becker is considered one of the pioneers of relief depiction for topographical maps. The introduction of chromolithography in cartography in the second half of the 19th century allowed the introduction of continuous tone shadings. In Switzerland, the Alpine Club published the first multicoloured relief maps which were further developed by Becker. He especially focused on the development of natural colour depiction (Schertenleib 1997) by applying hypsometric tinting combined with shading and representation of sunlit faces. As the Siegfried Map, the precursor of the Swiss National Map Series, was published from the end of the 19th century (but without a shaded relief), the Alpine Club Maps were abandoned. Imhof's main contribution was the further refinement of Becker's colour scales, finally leading to his famous Swiss Style Colour Relief Shading (Jenny and Hurni 2006). He developed a hypsometric tinting starting with a grey-bluish green in the lowlands, then into an olive, brown-red, yellow and finally fading into a white in the highest area.

2 Combining Cartography and Art

Eduard Imhof combined in a unique way the skills of a land surveyor/topographer and an artist. Numerous pencil sketches, watercolours and lithographic prints testify to his abilities to three-dimensionally perceive and vividly depict the Earth's surface

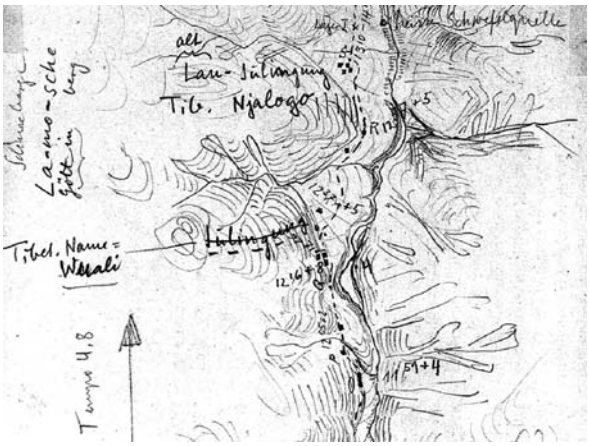

Fig. 3 "Horseback route sketch" of an area in Eastern Tibet based on compass azimuth measurements and distance estimations (Eduard Imhof, 1974)

not only on maps but also on topographic sketches. Figure 3 shows a "horseback route sketch" of an area in Eastern Tibet based on compass azimuth measurements and distance estimations which then served as a basis for a route map at the scale 1:200,000 of the region. Together with his sound knowledge in geodesy, precision surveying and mapping, Imhof had the opportunity to build-up a new and original cartographic edifice of teaching. He is widely seen as the founder of modern, academic cartography.

Besides elaborating countless maps and atlases as a kind o proof-of-concept of his findings, Imhof published his main findings and cartographic theories in articles and book chapters, but also in three important textbooks: "Gelände und Karte" ("Terrain and Map"), published in 1950 and 1968 (Imhof 1950, 1968), 1st edition also translated in French (Imhof, 1951) and English (by US Army, no ref. available); "Kartographische Geländedarstellung" ("Cartographic Relief Representation"), published in 1965, translated into English in 1982 and re-issued in English in 2007 (Imhof 1965, 1982, 2007); and "Thematische Kartographie" ("Thematic Cartography"), published 1972 (Imhof 1972), not translated. The first book "Gelände und Karte" was initially published by the Swiss Army and focused mainly on mapping techniques, terrain interpretation, topographic map content and map interpretation. In his second book "Kartographische Geländedarstellung", Imhof focused on the representation of terrain elements on topographic maps. It reflects his main findings and theories about terrain representation on maps, concentrating on topographic elements such as contour lines, spot heights, skeletal lines, shading, hachures, rock drawing and area colours. The opus uniquely shows the close connection between cartography and art. In all steps of the topographic map production process, artistic elements play a vital role. Not only are field sketches as shown above made by applying subjective depiction methods, but also map depictions of the terrain may contain individually designed elements. The following two chapters highlight some of the major relief presentation methods developed by Imhof and put them in relation to current topographic cartography.

3 Eduard Imhof's Legacy and Its Transfer to Modern Cartographic Applications

In his textbook "Kartographische Geländedarstellung" ("Cartographic Relief Presentation"), Imhof gives an introduction to topographic surveying and then systematically treats every group of graphic elements topographic maps:

1. Topographic surveying methods and accuracy considerations
2. Topographic terrain interpretation by field and aerial image observation
3. Colour theory
4. Spot heights and soundings
5. Contour lines
6. Shading
7. Hachures
8. Rock drawing
9. Symbols for small landforms
10. Area colours
11. Interplay of elements
12. Map reproduction

Some of the described methods and techniques are today out-of-date. Especially surveying and reproduction technologies have dramatically changed due to GPS- and Desktop Publishing (DTP) technology and some of the map elements no longer have the same importance as 30 years ago. The following chapter describes the major map element groups and their depiction. The major findings of Imhof concerning these features are pointed out and current cartographic applications and techniques related to those representation forms are presented.

4 Elements of Topographic Maps

4.1 Situation: Roads, Paths, Buildings

The situation elements are an important part of every topographic map. They represent most of the human constructional intervention in a landscape. Although Imhof did not concentrate on the situation features in his textbook "Cartographic Relief Presentation", he published many articles and book chapters concerning this topic. His main message was the retention of typical characteristics of settlement types also in smaller scales after the cartographic generalisation process. For geomorphological applications however, these features play only a minor role; therefore they are not further covered in this paper.

4.2 Area Features

Classic topographic maps contain only little symbolised area features, except for forest areas (vegetation), area patterns like a swamp (hydrography) and in some maps – like the USGS-series – settled areas. Modern digital cartographic systems allow to easily include such features. Areas can be defined as vector polygons. The filling and the patterning can be accomplished both in vector and in raster mode. Seamless area representations can be found on newer topographic maps like the new GIS-data derived German topographic maps (Grimm 1993) or rather on thematic maps such as land use maps or geologic maps. These examples represent qualitative, discrete features. Other area features like elevation models and shadings are continuous features.

An alternative definition can be found in Imhof (1965, 1982, 2007): Pragmatically, he differentiates between area tints to represent land cover, hypsometric tints and their combination. The design can be laid out to obtain an optimal natural resemblance, a symbolised choice of colours or again a combination. Strictly spoken, most topographic/cartographic representations are such compromises.

Perfect natural resemblance cannot be obtained in a map (actually, this must not be the goal of any symbolised, thus abstract map). Colour orthophotos are not an adequate substitute: They are not interpreted or they can be largely affected by the sun's position and atmospheric effects which create an unfavourable relief representation. Imhof stipulates that orthophotos are better suited for large scale maps. He warns of too coloured, saturated and rough area mosaics which can lead to misinterpretations. He therefore developed further the Swiss Style Colour Shading, built

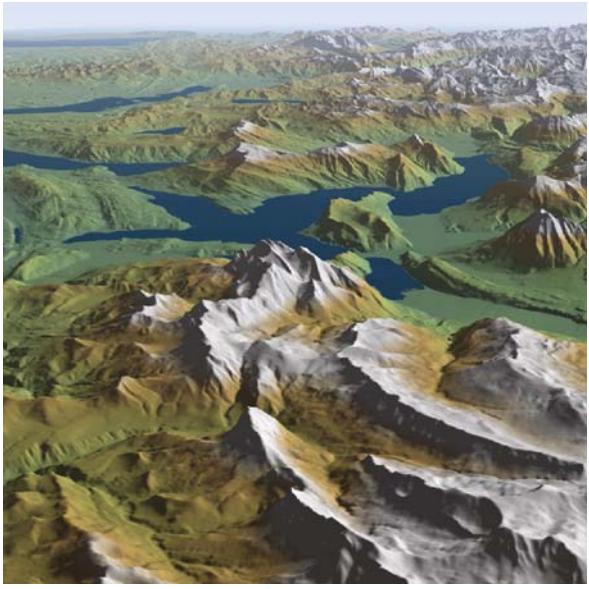

Fig. 4 Panorama visualisation of Lake of Lucerne in the "Atlas of Switzerland – interactive". Hypsometric tinting of the terrain model *Elevation data (DHM25): © swisstopo.*

up by combining relief shading, symbolic, and hypsometric area tints and aerial perspective (misty appearance of lower, more distant parts of the terrain). A modern 3D example of the coloured area tints can be seen in Fig. 4. Herrmann (1972) remarks that the choice of such "compromise tints" can lead to contradictions with the real land cover. He proposes the use of tints resembling to the natural colour of the terrain. They are chosen according to climatic zones, vegetation and land cover instead of hypsometry. Such a uniform legend is especially interesting to be applied in school atlases. The Swiss School Atlas ("Schweizer Weltatlas") for instance is built on a generally suitable legend based on the vegetation zones (Spiess 2006).

4.3 Hydrologic Features

Hydrologic objects on a map are represented either as point elements (springs, geysers, wells, etc.), line elements (rivers, creeks, canals, pipes, etc.) or area elements (sea, lakes, swamps, glaciers, etc.). They contain significant information about the morphology of a terrain. The sea and lakes are equi-potential surfaces defined by the earth's gravity field. Linear objects are usually directed according to the flow direction. They represent a break-line on the terrain such as lakes, too. Together with skeletal lines (e.g., ridgelines) Imhof (1965, 1982, 2007) treats those elements as a special case of the spot heights.

4.4 Contour Lines and Spot Heights

In a topographic map, the three-dimensional shape of the terrain has to be shown two-dimensionally, including the presentation of terrain elevations. The most common features used for this task are contour lines and additional spot heights for significant points. Other relevant terrain features which fall in between two contour lines can be represented by special symbols, e.g., slope hachures. Imhof (1965, 1982, 2007) devotes three chapters to terrain describing elevation features. He focuses on the adequate choice of the contour interval (equidistance) for different map scales and on the possibilities and limitations of contour lines for describing various morphological features for improved recognition. The following examples show some of the main problems which can occur in topographic mapping with respect to contour lines (see also Buckley et al. 2004).

The Figs. 5, 6, and 7 show three contour line representations of the same area in the Alps (Monte Leone). Figure 5 is a photogrammetric compilation with an equidistance of 20 m. Figure 6 shows its generalised representation (brown plate, glaciated areas are omitted). The contour lines in Fig. 7 have been generated by interpolation from photogrammetrically measured profiles (distance between profiles: 50 m). The advantages of direct measurements of contour lines can be seen in the example. Morphologic finesses represented by the compiled contour lines are almost entirely missing on the interpolated version. The generalised version

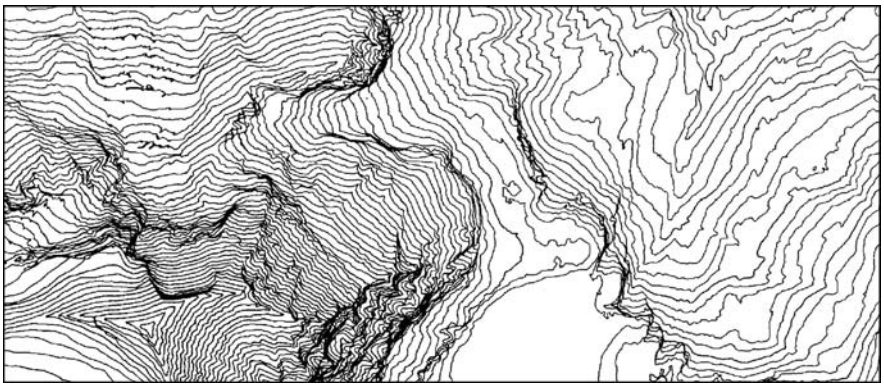

Fig. 5 Photogrammetric compilation of contour lines 1:25,000; Monte Leone area in the Swiss-Italian part of the Alps (after Hurni 1989, slightly reduced scale) (© *swisstopo*)

Fig. 6 Contour lines fro the Swiss national map 1:25,000 (© *swisstopo*)

emphasizes the important terrain features and relieves the map image from unnecessary details. The introduction of break-lines improves the interpolation of contour lines by interrupting that process along those predefined lines. Figure 8 shows a contour line representation of the Matterhorn which was interpolated from 100 m profiles (with photogrammetrically compiled characteristic points along the north-south-profiles). Figure 9 displays the same representation but with integration of photogrammetric break-lines along the major ridges. It can be seen that within extended rock areas countless additional break-lines should be compiled in order to precisely represent the morphologic detail structure. In such cases, it is still easier to directly restitute contour lines.

On the other hand, such contour lines must be compiled with extreme care, especially in mountainous areas. Figure 10 shows an enlarged detail of the south-eastern side of Monte Leone. This extremely steep rock wall shows many overhanging areas which must be retouched in the final contour line map. Besides, such inconsistent restitutions can occur due to parallax effects on the aerial stereo image pair (hidden

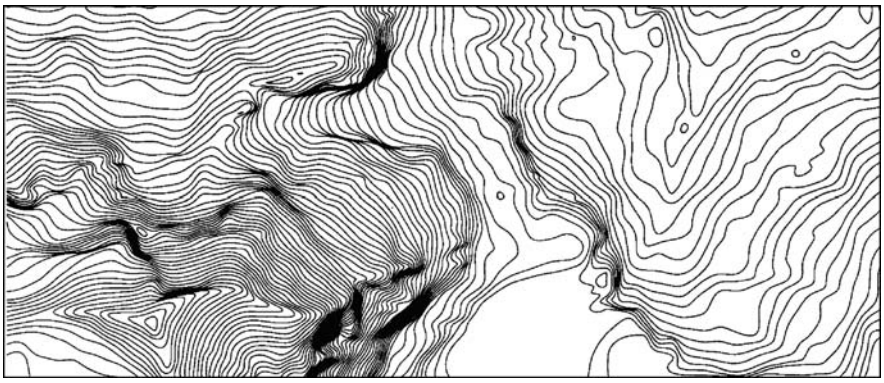

Fig. 7 Interpolated contour lines from photogrammetrically compiled profiles; original profile distance: 50 m (after Hurni 1989)

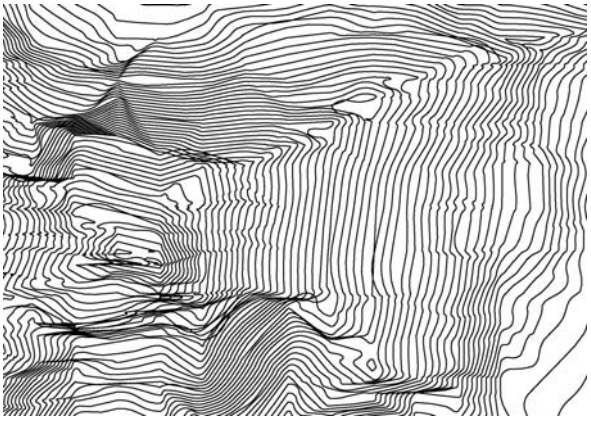

Fig. 8 Matterhorn: Interpolated contour lines. Equidistance 20 m. Generated from characteristic points along 100 m north-south-profiles
Data source: © swisstopo.

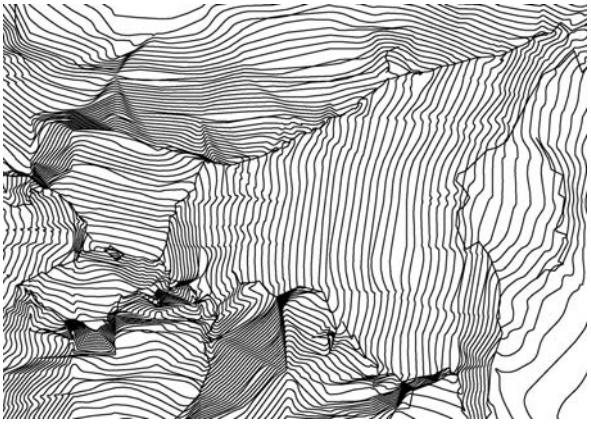

Fig. 9 Matterhorn: Interpolated contour lines. Equidistance 20 m. Interpolation as in Fig. 8, but with integration of major break-lines
Data source: © swisstopo.

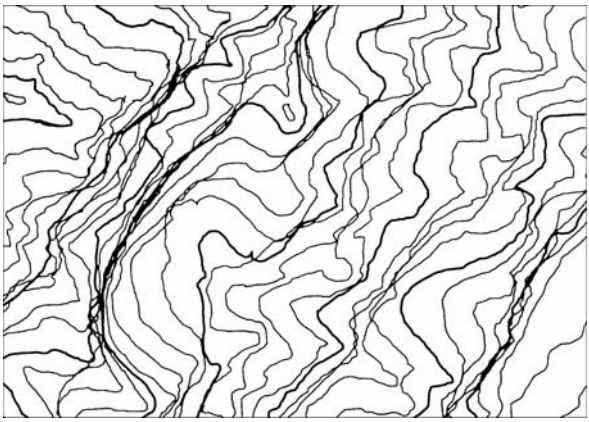

Fig. 10 Overlapping contour line images due to overhanging areas and parallax effects on the aerial stereo image pair (hidden areas). Monte Leone, Switzerland/Italy (after Hurni 1989)

areas). Figure 11 shows a contour line image with 20 m equidistance of a morphologically different area: This area in Greece (Methana peninsula, see also the corresponding example below) contains volcanic domes and many neighbouring singular points (depressions, hilltops). In order to represent the difficult topography, intermediate lines (10, 5 m) are included. During the photogrammetric compilation of the contour lines on an analytical plotter, a combined distance-angle criterium was applied: In case of high curvatures, the distance between the vertices is automatically reduced. The final curves are interpolated by a spline function offered by the cartographic software. Possible overlaps are resolved manually, since every curve is treated individually.

However, contour lines are in many cases visualisation elements of the terrain rather than the original "storage elements" for terrain information. Today, this function has been taken over by digital terrain models, although carefully compiled contour lines and spot heights remain even today still the clearest representation of the terrain on maps.

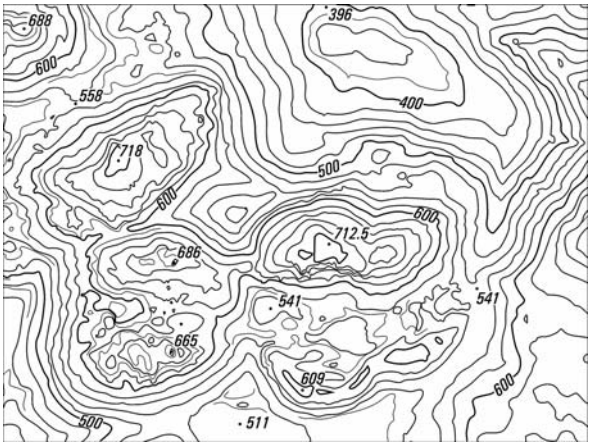

Fig. 11 Contour lines: Extract of map of Methana peninsula in Greece with volcanic morphology (after Hurni 1995)

4.5 Relief Depiction

Manual relief shading is a time consuming task of depicting a terrain model modulated according to a specific illumination model. Highly skilled cartographers are needed to clarify and simplify topography. The aim is not to depict the terrain in a geometrically correct form, but to create an image that is easy to interpret. According to Imhof (1965, 1982, 2007), this can be achieved with the following techniques: Locally, the light direction is slightly turned out of the main light direction in order to emphasize and clarify topographic features. Landforms that lay in the main light direction are thereby accentuated. Moreover, flat areas are filled with a bright grey tone to build a relationship between hillsides separated by flat lowland. The applied tint is brighter than the physically correct value, in order to avoid any darkening of these usually densely settled zones. Furthermore, a brightening of local shadows on the light side of hills, and a darkening on the shadow side, is used to emphasize large landforms, to structure the landscape and to accentuate characteristic forms. Aerial perspective is used by cartographers to depict differences between high mountain summits and lower, more distant lowlands. Hence, contrast is sharpened towards the topographic peaks and softened towards the lowlands.

The digital process of deriving shaded relief from a digital elevation model (DEM) is analytical relief shading. For the needs of cartographers a great variety of methods for analytical shading have been developed. Generally, grey values depend on slope and aspect (or exposition), both calculated from the DEM. The first to produce an analytical relief shading was Yoëli (1959, 1965, 1966, 1967, 1967a) by applying diffuse reflection on a DEM. This illumination model determines the grey value of each pixel by calculating the cosine of the angle between the surface normal and the light vector. Figure 12 shows an example of a shaded relief, calculated with diffuse reflection. Some efforts have been made to adapt the different shading algorithms to the specific needs of cartography. Yoëli (1967) and Brassel (1974) for instance made first experiments with local adjustments of the light direction, and many others followed with improved shading algorithms (see www.reliefshading.com). The figures in this chapter have been produced using special semi-interactive software developed at the Institute of Cartography at ETH Zurich (see also the detailed description in Jenny 2001).

In mountainous areas, comparisons of analytical and manually shaded reliefs show that the calculated versions often contain unwanted details, whereas the manual shadings accentuate vertical transition. Smooth vertical transition is used in manual hill shading to emphasize aerial perspective and to structure the topography. If the slope information is ignored and shading is based on aspect only, this manual style can be simulated. Such aspect-based shading is generated according to a modified cosine shading equation (Moellering and Kimerling 1990). Since aspect is undefined in flat areas, a bright grey tone has to cover these regions. This grey tone can be mixed with the aspect based shading in function of slope, using a mathematical function or an interactive control panel. In nearly flat, but slightly undulating areas, aspect based shading produces almost random values. In Fig. 13, this shortcoming is remedied by a bright grey tone covering the planes.

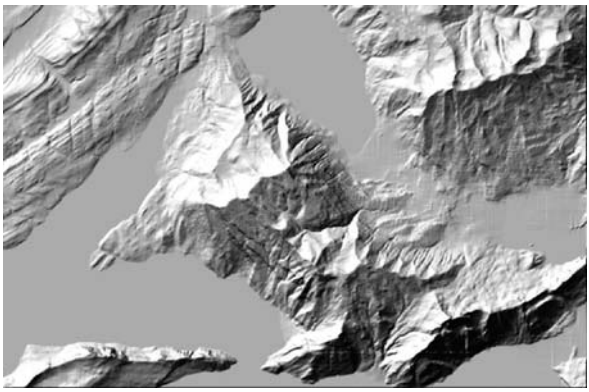

Fig. 12 Diffuse reflection
Elevation data DHM25: © swisstopo.

Fig. 13 Aspect based shading with a tone for flat areas
Elevation data DHM25: © swisstopo.

To simulate aerial perspective, three components are transformed into a weight for each pixel. The first weight is the relative elevation of the considered point. The second weight is based on the exposition towards the light direction (aspect) and the third weight is based on the relative position of the considered point within a hillside that is identified using slope lines. These weights correct the previously calculated grey value. The grey value is first reduced in contrast and then a definable constant value is multiplied by the three weights and added to the grey value. Figure 14 illustrates this effect.

For local adaptations, the user has the possibility to enclose sub-areas of the DEM and to provide them with adapted parameters for the calculation of the shaded relief. The following parameters can be adjusted in the software: Light direction, vertical exaggeration, brightness, elevation dependant contrast and interpolation between diffuse reflection and aspect based shading. Figure 15 illustrates the effect of a local adaptation of the light direction. The main light source from north-west is locally replaced by a light source from west. After finishing the digitisation of a polygon,

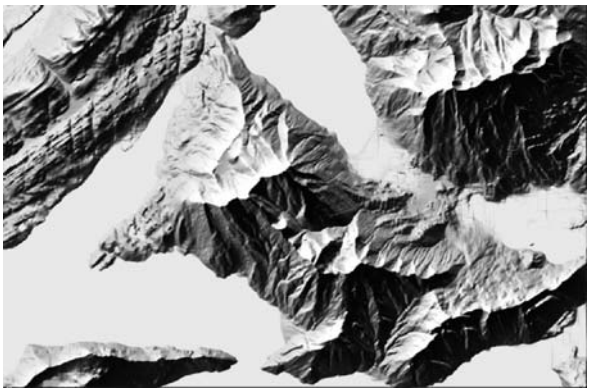

Fig. 14 Addition of aerial perspective
Elevation data DHM25: © swisstopo.

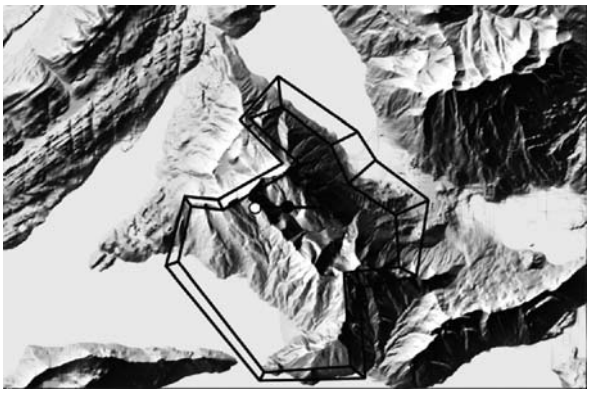

Fig. 15 Local adjustment of light direction
Elevation data DHM25: © swisstopo.

the program automatically generates a second polygon inside the first one and interpolates the grey values between them.

When comparing Figs. 12 and 13, one notes that aspect based shading results in clearer images with a higher contrast. Horizontal structures are suppressed by exaggerating vertical gradients. Mountain summits can be accentuated and large landforms are emphasized by applying the algorithm for aerial perspective. Generally, aspect based shading and aerial perspective turned out to be well suited for mountainous areas; diffuse reflection should be used for lowland and flat areas. At a global level, the two techniques can be combined in function of slope. Local adaptations with fences are intuitive to use as tests in Jenny (2000, 2001), and Hurni et al. (2001) have shown. With adjustments of light directions and brightness, important landforms can be easily emphasized. Important small details and characteristic structures of the terrain can be accentuated. Cartographers are able to transfer their experience and knowledge acquired in manual shading and to apply them to digital elevation models.

One of Imhof's major achievements was also the definition of a scale of tints for colour shaded reliefs. As mentioned in Sect. 4.2, a compromise must be found between natural resemblance and symbolic colours. According to Jenny and Hurni (2006), besides a pure hypsometric tinting, there is also a colouring possible which combines hypsometric tinting by elevation with a modulation according to exposure to illumination. Based on precursor work by Fridolin Becker, Imhof developed such coloured relief shading based on a hypsometric colour scale starting from a bluish-greyish green for lowlands by olive and brownish-yellowish tones for mid-altitude areas to even white tones for the highest, snow covered peaks. He also developed a photomechanical method to derive such a tinted depiction from one greyscale shaded relief. Several copies of the relief with different contrast levels were tinted in different colours and combined with a negative of the shaded image in yellow (= sun illumination) and with a hypsometrically graded plate. This Swiss Style Colour Relief Shading was applied in many maps of the Swiss World Atlas ("Schweizer Weltatlas", the official Swiss school atlas), in school maps and in small scale topographic maps. Figure 16 shows a manual watercolour painting with the combined shading and hypsometric tinting method.

Recently a new method and a programme for digital production of Swiss-style colour relief shading were developed at the Institute of Cartography (Jenny and Hurni 2006). The reason for this implementation was the production of a new edition of the school map of the Swiss Canton of Schaffhausen. Imhof's original relief from the 1950 was still available, but not the derived colour plates. The software allows to defining the colour tint at specific place in the terrain, e.g., yellow in higher areas of the illuminated north-western side of a hill or a blue-green-grey in lower shadowed areas. After having defined a number of significant points with their respective tints, the programme sets up a two-dimensional look-up table with the elevation on the y-axis and the greyscale values on the x-axis. Figure 17 shows the input parameters

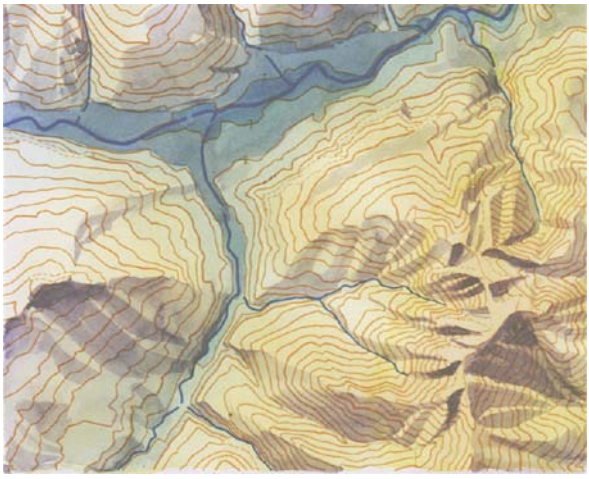

Fig. 16 Manual watercolour painting with the combined shading and hypsometric tinting method. Original painted by E. Spiess, 1951

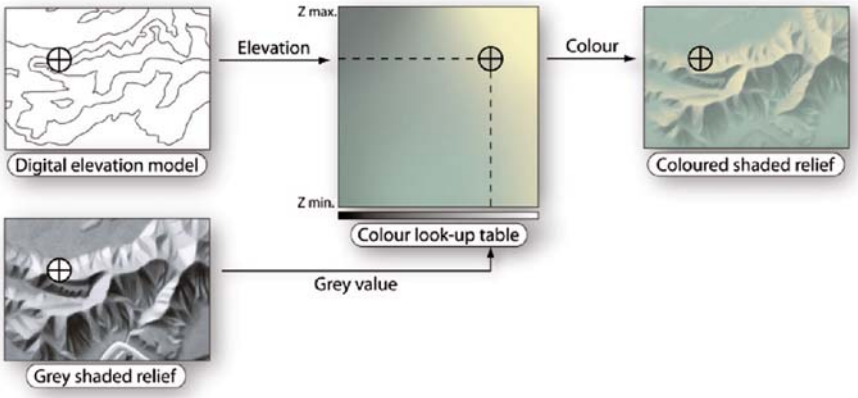

Fig. 17 Generating a coloured shaded relief by choosing the colour from a two-dimensional look-up table built up by elevation and exposition parameters (after Jenny and Hurni, 2006)

elevation (from a digital elevation model DEM) and grey value (from an existing shaded relief, representing the exposition to illumination). The resulting coloured shaded relief combines the greyscale value of the relief with the height dependent colour of the specific point.

4.6 Rock Drawing

Scree or solid rock cover extensive parts of mountains all over the world. Those areas have an important influence on different aspects of the mountain environment and human presence, like natural hazards or mountaineering. Human activities in those domains therefore require precise maps. Techniques for a clear and precise rock representation were developed in the Alpine countries in the 19th century. Mostly, methods derived from slope shading hachures were used. At the Federal Office of Topography the technique of shaded rock fill hachures has been developed further, especially after 1935 when the production of the new National Map Series began. Rock areas are divided into morphologically compact units using structure lines. Structure or shape lines as well as the vertical or horizontal fill hachures are modulated according to an illumination model. Figure 18 shows the modulation of rock hachures (swisstopo 1996). Eduard Imhof postulated a simplified version, the so-called ridge-line representation which displays only the major contours, ridges and drains. Figure 19 (upper image) shows an example of a small-scale map of Imhof with a ridge-line representation.

Unfortunately, the presented techniques are very time-consuming and require a costly production workflow. For instance, about 2000 working hours were required for the rock drawing of an average mountain sheet of the Swiss National Map Series, and thus caused costs of about 250,000 Swiss Francs per sheet. It does not surprise

Fig. 18 Modulation of shape lines (outlines) and fill hachures according to aspect © *swisstopo.*

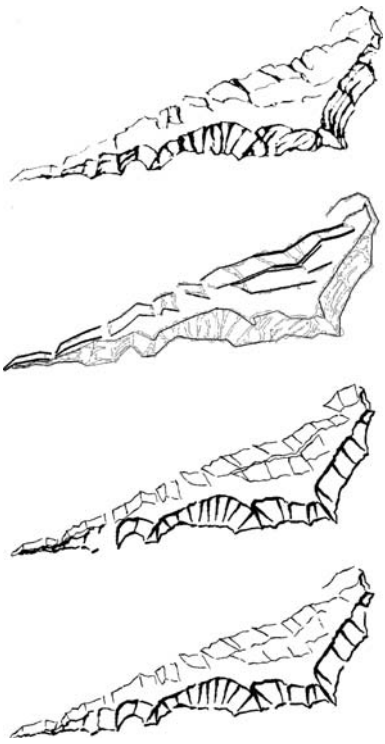

Fig. 19 Upper image: Conventional small scale rock drawing by Eduard Imhof. Second image: Digitisation of upper and lower edges of rocks (programme input). Third image: Raw programme output. Lower image: Manually edited output (raster mode). After Hurni et al. (2001)

that entirely new elaborations of rock plates can only be carried out in special cases today, like for instance the National Geographic map of Mount Everest. However, the question of how to represent rocky areas geometrically correct, well designed, inexpensive and computer-compatible is still relevant when carrying out new cartographic surveys.

In a pilot project at ETH Zurich, a programme which allows the semi-automatic generation of rock drawings has been developed for the first time (detailed description in Hurni 1995 and Hurni et al. 2001). In order to avoid the use of complex, hard-to-elaborate hachures, the ridge line representation has been chosen. The upper

and lower edges of a rock are digitised manually from field sketches. The software works according to the following rules (after Hurni et al. 2001):

- The representation is reduced to vertical ridges and drain lines ("form" lines) and upper and lower edge lines. There are no fill hachures.
- The line hachures should have a rough appearance.
- Aggregation of objects and high light/shadow contrasts are necessary.
- Light and shadow sides should be easily distinguishable by their overall brightness (line weights!).
- On the shadow side, upper edges are thicker than lower edges (cuneiform).
- On the light side, upper edges are thinner than lower edges.
- On the shadow side, vertical form lines are thicker on the upper side; on the light side they are thicker on the lower side.
- The angle of light is assumed from north-west.
- On the shadow side and with south-western to south-eastern aspects, edge lines are thicker on the left side; with north-eastern to south-eastern aspects, they are thicker on the right side. With pure south-eastern aspects, the lines are thicker on both sides and thinner in the centre.
- On the light side and with north-western to south-western aspects, edge lines are thicker on the left side; with south-western to north-eastern aspects, they are thicker on the right side. With pure north-western aspects, the lines are thinner on both sides and thicker in the centre.
- Due to erosive effects, very often concave and convex shapes of the rocks can be found.

Figure 19 shows a conventional rock drawing by Eduard Imhof with ridge-lines, representing the main structure lines (upper representation). Digitised upper and lower edges of the rock objects with an identical number of vertices serve as input data for the digital rock drawing (second representation). This leads to a box shaped ridge line image containing those edges and the vertical form lines. Therefore, the method is best used for long rock bands. The third representation in Fig. 19 shows the raw output (rasterised vectors), which is finally edited by hand using raster software (last representation). All edge and form hachures are cuneiform and modulated according to an illumination model, i.e., their overall line thickness is varied

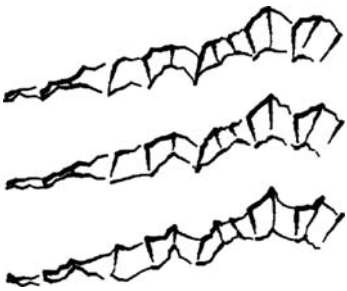

Fig. 20 Three examples (from *top* to *bottom*) with concave, without and with convex cavity and the rough appearance. After Hurni et al. (2001)

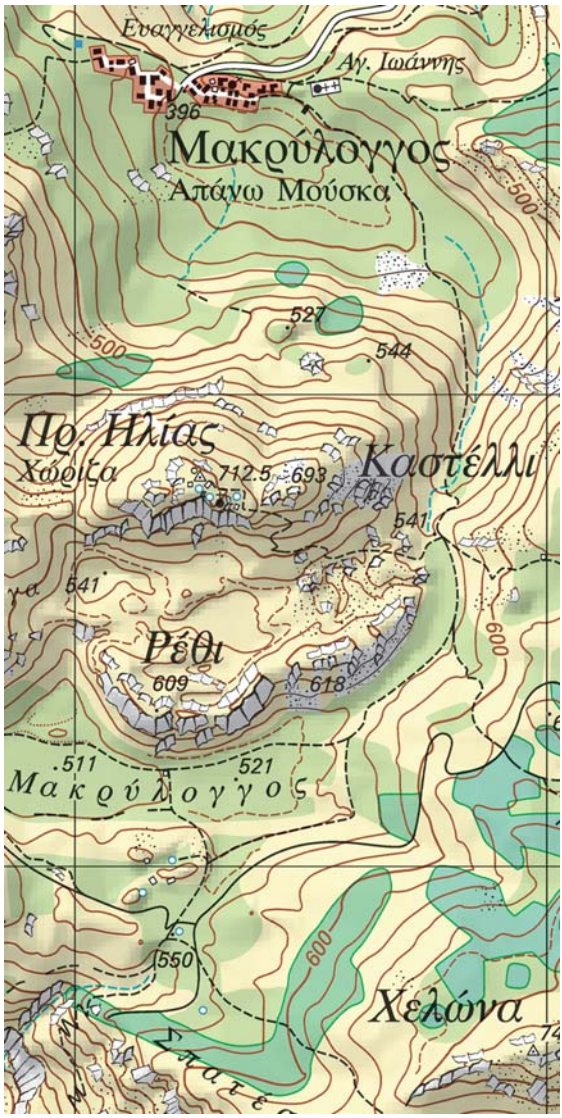

Fig. 21 Extract of map of Methana peninsula/Greece 1:25,000 with digital rock drawing, published by ETH Zurich

by aspect. The rough appearance is simulated by a local variation of the line widths and the line positions using a random function. The programme allows to fine-tune the line widths according to their position and aspect with about 250 parameters. Also concave and convex shapes between the vertices can be simulated with different degrees of cavity (Fig. 20). Manual corrections can easily be applied in an image manipulation program after rasterizing the vector output of the rock drawing software. Figure 21 shows an extract of a map of Methana peninsula/Greece 1:25,000 with digital rock drawing.

5 Conclusions

Since the beginning of high quality topographic cartography in the 19th century, the adequate representation of relief features is a major and demanding task. The work of Eduard Imhof which was a consequent continuation of his forerunners activities, laid the base for modern topographic mapping not only in Switzerland, but also on an international level. In his book "Cartographic Relief Presentation", he defined rules and gave examples how to depict various topographic using adequate map elements and cartographic techniques. It is not surprising that this standard textbook has recently been re-edited by a major GIS vendor. The proper application of this basic cartographic knowledge was for a long time neglected. In the last 20 years cartographers were to a large extent dealing with the change from analogous to digital cartography and with the connected technological challenges. However, today new developments as described in this paper allow simulating Imhof's techniques to a high degree. It is even possible to develop them further and to apply them to new fields of applications such as perspective representations of terrain features. We hope that in the future such functionality will be part of any GIS package and thereby accessible to a large number of users in cartography but also in other thematic fields like geomorphology depending on up-to-date mapping technology.

References

Brassel K (1974) Modelle und Versuche zur automatischen Schräglichtschattierung. PhD thesis, University of Zurich

Buckley A, Hurni L, Kriz K, Patterson T, Olsenholler J (2004) Cartography and visualisation in mountain geomorphology. In: Bishop MP, Shroder JF (eds.) Geographic Information Science and Mountain Geomorphology. Springer-Praxis, Chichester, 253–287

Grimm W (1993) Eine neue Kartengraphik für das digitale kartographische Modell "ATKIS-DKM 25". Kartographische Nachrichten 43:61–68

Herrmann C (1972) Studie zu einer naturähnlichen topographischen Karte 1:500,000. PhD thesis, University of Zurich

Hurni L (1989) Verschiedene Felsdarstellungen für Gebirgskarten 1:25,000. Thesis, ETHZ

Hurni L (1995) Modellhafte Arbeitsabläufe zur digitalen Erstellung von topographischen und geologischen Karten und dreidimensionalen Visualisierungen. PhD thesis, ETH Zurich

Hurni L, Dahinden T, Hutzler E (2001) Digital Cliff Drawing for Topographic Maps – Traditional Representations by Means of New Technologies. Cartographica 38:55–65

Imhof E (editions 1950 and 1968) Gelände und Karte. Eugen Rentsch Verlag, Erlenbach/Zurich

Imhof E (1951) Terrain et carte. Eugen Rentsch Verlag, Erlenbach/Zurich

Imhof E (1965) Kartographische Geländedarstellung. DeGruyter, Berlin

Imhof E (1972) Thematische Kartographie. DeGruyter, Berlin

Imhof E (1974) Die grossen kalten Berge von Szetschuan. Orell Füssli, Zurich

Imhof E (1979) Die Bemühungen um neue topographische Karten der Schweiz. Die Alpen 55 (I/79):17–19

Imhof E (1982) Cartographic Relief Presentation. DeGruyter, Berlin

Imhof V (1990) Eduard Imhof, ein Leben mit Landkarten. Schweizer Pioniere der Wirtschaft und Technik 50. Verein für wirtschaftshistorische Studien, Meilen

Imhof E (2007) Cartographic Relief Presentation (reprint). ESRI Press, Redmond

Jenny B (2000) Computergestützte Schattierung in der Kartografie. Thesis, ETHZ/EPFL

Jenny B (2001) An Interactive Approach to Analytical Relief Shading. Cartographica 38:67–75
Jenny B, Hurni L (2006) Swiss-Style Colour Relief Shading Modulated by Elevation and by Exposure to Illumination. The Cartographic Journal 43:198–207
Moellering H, Kimerling AJ (1990) A New Digital Slope-Aspect Display Process. Cartography and Geographic Information Science 31:67–77
Schertenleib U (1997) Fridolin Becker (1854–1922): Topograph, Kartograph, Innovator. Cartographica Helvetica 15:3–10
Spiess E (2006) Schweizer Weltatlas – Atlas Mondial Suisse – Atlante Mondiale Svizzera. Kantonaler Lehrmittelverlag, Zurich
swisstopo (1996) Richtlinien der TopoKarto für die kartographische Bearbeitung der Landeskarte 1:25,000, 1:50,000 und 1:100,000. swisstopo, Wabern
Yoëli P (1959) Relief Shading. Surveying and Mapping 19:229–232
Yoëli P (1965) Analytical Hill Shading. Surveying and Mapping 25:573–579
Yoëli P (1966) Analytical Hill Shading and Density. Surveying and Mapping 26:253–259
Yoëli P (1967) The Mechanisation of Analytical Hill Shading. The Cartographic Journal 4:82–88
Yoëli P (1967a) Die Richtung des Lichtes bei analytischer Schattierung. Kartographische Nachrichten 17:37–44

New GML-Based Application Schema for Landforms, Processes and Their Interaction

Marc-Oliver Löwner

Abstract Here we propose an application schema for features and processes of science of geomorphology based on international geoinformation standards. Using the Geography Markup Language and the Unified Modeling Language, this object oriented model is a precondition for data exchange without loss of semantical information.

Landforms and their evolution are determined by the surface, subsurface preconditions, and external forces, which result in erosion processes. The analysis of this complex process-form interaction is covered by the field of geomorphology, whose members work in various locations around the world.

The main problem of a synoptic approach is that data cannot be easily exchanged among different study groups. This is partially due to the fact that commercial GI-software is not adapted to the needs of the science of geomorphology. Another problem is the storage of data in so-called flat files without a documented data structure.

Geoinformatics has been dealing with the questions of data management and representation of 3-D objects for quite a while. The efforts of the International Organization for Standardization and the Open Geospatial Consortium deserve particular credit in this context. They are striving for standardization in order to achieve their main goal of interoperability of the different GIS and data formats that are being used. The development of formal semantic models by the community of geomorphologists is imperative to achieve these innovations.

Here we present an application model for geomorphic purposes that must fulfill the following requirements: First, an object-oriented view of landforms with a true 3D geometric data format has to be established. Second, the internal structure of landforms needs to be stored in an adequate way. Third, the interaction of process and a Geoobject must be represented. Fourth, the change of landforms over time must be considered.

M.-O. Löwner (✉)
Institut für Geodäsie und Photogrammetrie, University of Braunschweig, Gaußstrasse 22, 38106 Braunschweig, Germany
e-mail: m-o.loewner@tu-braunschweig.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_2,

The goal is to develop a framework for a Geomorphic Information System that will enable scientists to share data worldwide. Such a global data transfer is necessary to evaluate landscape evolution research results in areas such as the impact of climate change on the land surface on a larger scale.

Keywords Interoperability · Geomorphic information system · Data model · Geoobject

1 Introduction

Geomorphology as the science of the land's surface investigates landforms, their change, and the processes causing their change all over the world (Hugget 2003). The main problem in comparing results of observations and predictions is that landforms first have a complex 3 dimensional geometry, second have numerous internal parameters and third develop in a process-response system, sometimes over a very long time. To solve the latter problem geomorphology has adopted the ergodic principle. This is a space for time substitution which means after Paine that sampling across an ensemble is equivalent to sampling through time for a single system (Paine 1985, Chorley and Kennedy 1971, Dikau et al. 1998). Anyway, to describe landforms in a formal way with the objective of exchanging complex land surface of data of features and processes still poses a great challenge.

In geographical information science the exchange of data without loss of information from one application to another is called interoperability (Gröger and Kolbe 2005). It is achieved when datasets are heterogeneous which poses three questions (Bishr 1998). First, the semantically heterogeneity addresses the problem that a different perception of phenomena leads to different abstractions. Second, the schema heterogeneity refers to structural differences in modeling one and the same feature in different ways. Third, the syntactically heterogeneity, adverting more technical issues like the interchange format to transport the data. The first point can only be clarified by the involved scientists (Fonseca and Egenhofer 1999, Dehn et al. 2001), the second and third by the use of international standards for application modeling and data transfer.

In the field of geoinformation systems mainly two organisations work on standards and norms. These are the Technical Committee 211 (TC 211) of the International Organization for Standardization (ISO) and the Open Geospatial Consortium (OGC). For 3D geodata the most important standard is the ISO 19107 Spatial Schema, which specifies the representation of 0–3 dimensional geometrical and topological primitives. This is done as an abstract specification on the basis of the Unified Modelling Language (UML) (Booch et al. 1999). Further abstract specifications rule how application models have to be built or how annotations are formulated, like the ISO 19109 (2002) or ISO 19115 (2002), respectively. However, no implementation rules are defined.

Using the Geography Markup Language 3 (GML3) assures syntactically heterogeneity (Cox et al. 2005, Lake et al. 2004). GML3 is the realization of the abstract concepts of ISO 19107 and other standards mentioned above applying the widely

used Internet standard Extensible Markup Language (XML) (Yergeau et al. 2004, Hunter et al. 2004) as a computer and human readable language. Although GML3 is written in XML, representing GML3 concepts using UML class diagrams is quite established.

When heterogeneity is achieved, geographic information can be readily shared on the Internet today. A Web Feature Service (WFS) allocates standardised methods to retrieve and update geospatial data encoded in GML3 using the Internet standard Hypertext Transfer Protocol (HTTP) (Vretanos 2005, Lake et al. 2004). However, GML3 as a syntactical basis for request and response message does not solve the problems of formalising our scientific perception of phenomena. It remains in hands of the scientists to develop an application model covering universally accepted concepts of geomorphology.

Here we present such an application model for geomorphic purposes that fulfills the following requirements: First, an object-oriented view of landforms with a true 3D geometric data format. A representation of features by 2D tessellations is unable to cover the as-is state and volume of the material involved. Second, the internal structure of landforms are represented in an adequate way. Drillings and also geophysical data mining provide a lot more of information about substrate and subsurface shapes than a map can show. Third, the interaction of process and a Geoobject is modeled through a class concept of a geoprocess. This representation can be used to store a process-related accessibility (German: Prozessuale Erreichbarkeit). Some neighboring features come into contact with each other through the exchange of material, some do not. Fourth, the change of landforms over time is considered. While the shape and internal properties of features may change over time, their semantically identity will remain unless they are completely erased through erosion processes. Developing a geomorphological application model further aims a common definition of the basic entities, attributes, and relations of the land surface's entities.

We briefly describe the approach to develop a formal model in informational science (Sect. 2) as well as the formalism UML. Section 2.2 includes the used GML3 classes to represent the geoobject's geometry. In Sect. 3 we describe the application model developed in this study. Therefore, we focus on the essential concepts of a geoobject and the geoprocess. Section 3.2 comprises a more detailed model of a special type, a soil slope. At the end we will discuss the archived findings.

2 Methodological Approach to Formal Modeling of Geoobjects

Worldwide geographical data can be shared over the Internet using Web Feature Services (Vretanos 2005, Lake et al. 2004). The precondition is the development of a semantic model or ontology (Fonseca et al. 2002) based on the international standard GML3 as an implementation of the ISO 109107 and others.

Gruber defines an ontology as a formal, explicit specification of a shared conceptualisation (Gruber 1993). Knowing that our perception of the real world is influenced by our subjective knowledge and cultural background (Frank 2001, Fonseca et al. 2002, Burrough and Frank 1995), he defines five criteria for designing such an

ontology: First, clarity; a semantic model is supposed to be impartial using a documented formalism. Second, logical coherence, third expandability on the basis of the existing model, fourth minimal encoding bias and last minimal ontological commitment. The first two points cannot easily be proofed but falsified. To be conform to the third and the fifths criterion only basic concepts of geomorphological science are represented here on the basis of the least commitment. Thus, the semantic model formalized in this study is not meant to be complete but monotone expandable.

To achieve interoperability it is essential to follow international norms and standards as mentioned above. Here we use the Unified Modelling Language (UML) and the Geography Markup Language (GML3) as an implementation of the ISO 19107, ISO 19123, and others.

2.1 The Used Formalism UML

In this study we represent all application models using the Unified Modeling Language (UML) following the ISO 19109, rules for application schema (ISO 19109 2002). UML is an object oriented language to specify, visualize an document software and application schemas (Cranefield and Purvis 1999).

As a simplification of real world phenomenon in UML `Classes` are drawn in boxes (Fig. 1). A `Class` in UML is used to instantiate objects with the properties of its `Class`. Anyway, `_Classes` with a prefixed underline are called abstract and cannot directly be instantiated. Every `Class` may have additional *`attributes`* and *`methods()`* determining its behavior.

One main advantage of UML is the concept of inheritance. A `Subclass` has the relation of specialisation to a `Superclass`, i.e. it receives all the *`attributes`*, *`methods()`*, and *associations* to other classes from the `Superclass`, which might be overwritten. A specialisation in UML is drawn as a line with a white-filled triangle pointing at the `Superclass`. Associations to other `Classes` may be *named* or unnamed whereas multiplicities rule how many objects of one `Class` are allowed to be associated with that of another. A special association is the aggregation as a "part-of-association" drawn as a line with a white-filled diamond at the

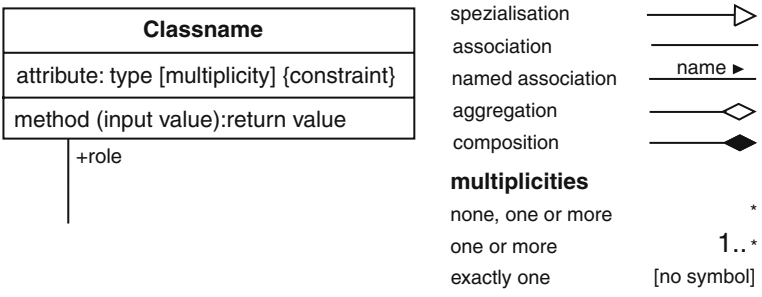

Fig. 1 Overview of the main UML symbols used

Class representing the "whole". A composition is semantically equivalent to a aggregation but with the added constraint that the whole is responsible for managing the lifetime of the part. It is drawn as a line with a black-filled diamond at the Class representing the "whole". Associated Classes may get *roles* to differentiate two or more special instances of a Class. See (Booch et al. 1999, Oesterreich 1998) for detailed description of UML.

2.2 Geometry Model

We represent the spatial properties of landforms by objects of GML3's geometry model. As the implementation of the standards ISO 19107, Spatial schema (Herring 2001), ISO 19123, Schema for coverage geometry and functions (ISO/DIS 19123 2004), and others this geometry model meets the claim of interoperability. Thus, data modeled with GML3 geometries can be exchanged from one application to another without loss of content (Gröger and Kolbe 2005).

GML3 as well as ISO 19109 represents 3D geometries according to the concept of boundary representation (Foley et al. 1995). That means that a Solid is represented by its bounding Surfaces which again are represented by their enveloping Polygons. The application model introduced here uses a subset of the GML3 geometry package only and is quite similar to that of CityGML, an OGC adopted Best Practice Paper for modelling 3D-Virtual-Cities (Gröger et al. 2005, 2007). The used profile of GML3 is depicted in Fig. 2.

The geometry model of GML3 allocates classes of geometrical primitives for each dimension. A zero-dimensional class Point, a one-dimensional class _Curve, a two-dimensional class _Surface, and a three-dimensional class _Solid. A _Solid is bounded by _Surfaces and _Surfaces by Curves. In this model a Curve is restricted to be a straight line, thus only the GML3 class LineString is used. Surfaces are represented by Polygons, which are defined as a planar geometry, i.e. all interior points and the boundary are required to be located in one single plain. A Polygone is associated with exactly one instance of the class _Ring representing the exterior boundary and zero or more, representing interior spaces within a Surface. Here only LinearRings are used.

A Surface is an UML composition of one ore more SurfacePatches, while the Surface only belongs to one particular SurfacePatch. One special SurfacePatch is the Triangle, which composes the TriangulatedSurface. Again, one Triangle can only be associated with one specific TriangulatedSurface.

All geometrical primitives may be combined to form aggregates, complexes, or composite geometries. There is no restriction on the spatial relationship between an aggregate's components. They may be disjoint, overlapping, touching or disconnected. GML3 provides a special aggregate for each dimension, a MultiPoint, a MultiCurve, a MultiSurface and a MultiSolid. By contrast a Complex is topologically structured, i.e. its parts must be disjoint, must not overlap but are allowed to touch at their boundaries or share parts of their boundaries. A

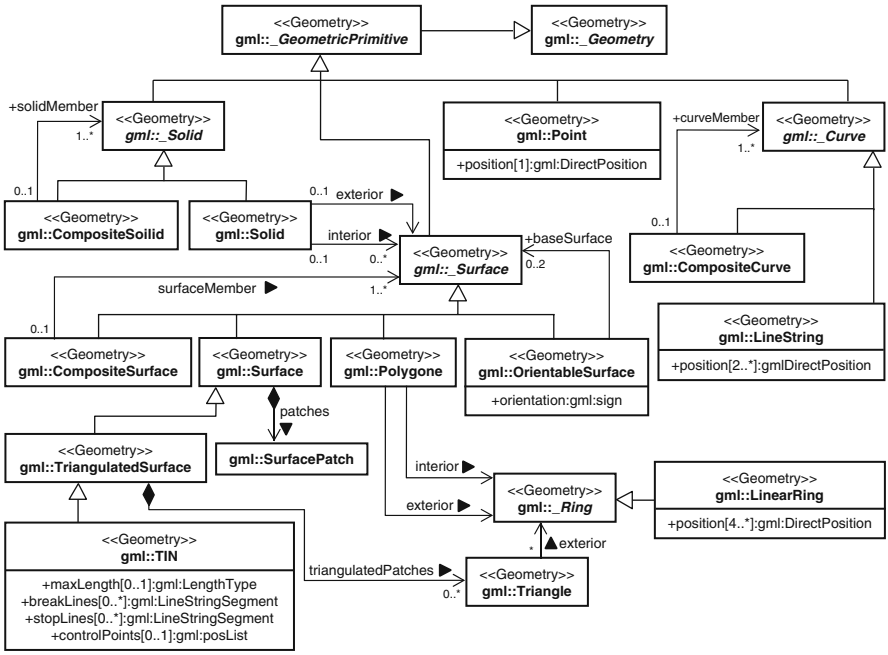

Fig. 2 UML diagram of the used geometry model (subset and profile of GML3)

`Composite` is a special complex provided by GML3 containing only elements of the same dimension. Its elements must be disjoint as that of a `Complex` but they must be topologically connected along their boundaries. A `Composite` can be specialised to a `CompositeSolid`, a `CompositeSurface`, or a `CompositeCurve`.

3 Application Model to Represent Geoobjects and Geoprocesses

Specific features are not directly represented in GML3. An application schema, which is a formal model of the world we like to describe, has to be developed. To do this for geoobjects and geoprocesses every entity class needs to be a specialisation of the GML3 class `_AbstractFeatureClass`. Geometry is then linked to the semantical features not by inheritance but by associating the GML3 geometry class needed (Lake et al. 2004).

3.1 Formal Representation of a `_Geoobject`

A `_Geoobject` in this model is a landform that is relevant in the process-response system (Fig. 3). The abstract class `_Geoobject` is modeled as a specialisation

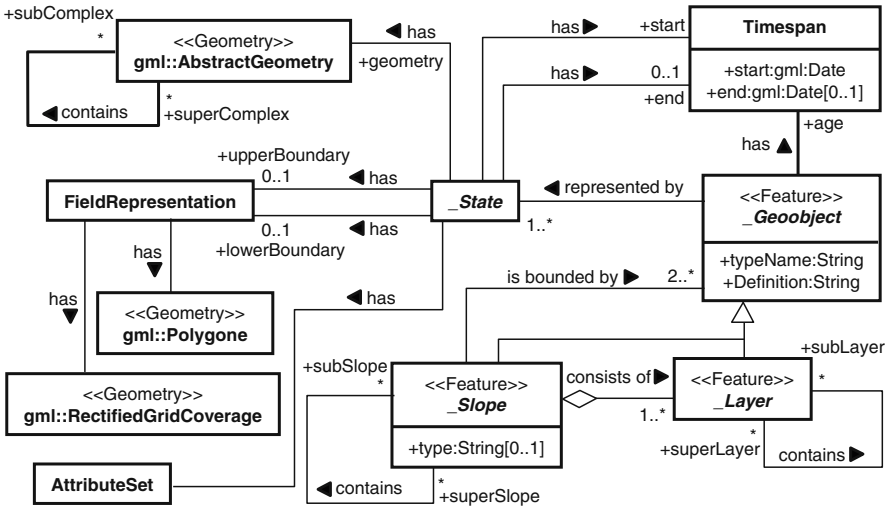

Fig. 3 UML diagram of the formal model of a Geoobject

of GML3's `AbstactFeatureClass`. It has one association to a `Timespan` qualifying its age, i.e. the date of origin. While the age of a `_Geoobject` normally cannot be assigned accurate to the nearest second, a `Timespan` has the attributes `start` and *`end`* to determine a range of time.[1] It is represented by one or more *`_States`* keeping the information of the `_Geoobjects's` characteristics at a given period of time. A `_State` is valid in between two `Timespans` whereas the *start* is mandatory. However, the *end* is optional because a state of a geoobject might still be valid. The modeling of a `_State` is meaningful since the characteristics of a `_Geoobject`, like geometry, can change but its semantical identity remains. The rockfall on 15 July 2003 at the Matterhorn is an obvious example (Rambauske 2003). Though the shape of the Hörnligrat was definitely changed, the identity of the Matterhorn is still the same. In other words, the representation of a `_Geoobject` by a `_State` enables us to keep records of its genesis.

A `State` of a `_Geoobject` is characterized by its geometry and material. Latter is modeled by an `AttributeSet` which contains all the attributes to be stored with a specialisation of a `_Geoobject`.

Geometry can be represented in two ways: First with a `FieldRepresentation` and second with an association of GML3's `AbstractGeometry`. As a `_Geoobject` is an abstract class and cannot be instantiation directly, there is no need to specify its geometry in more detail. Subclasses' geometries, of course, need a more detailed description. A `FieldRepresentation` holds an association to a `RectifiedGridCoverage`, which is a common raster dataset like ESRI's Grid format. Additionally it has a planar `LinearRing` to map the feature's shape,

[1] There is no restriction on the durability of a `_Geoobject's` creation. This might be the duration of a storm event for a linear erosional feature or tenth of years for a protalus rampart.

e.g. when creating a digital geomorphologial map (Otto and Dikau 2004). There is no need to store z-coordinates for the representing 2D shape. They can be derived from the values of the `RectifiedGridCoverage`. A `_State` has none, one, or two `FieldRepresentations`, one representing the *upperBoundary* and the other the *lowerBoundary* of the feature. This is to calculate volumes if both `FieldRepresentations` are available.

We want to stress here that the representation of a `_Geoobject's` geometry by a `FieldRepresentation` is not recommended. As a tesselation, i.e. a collection of plane figures that fills the plane with no overlaps and no gaps (Worboys 1995), it is not suitable to represent vertical walls or even overhangs. This restriction does not apply to GML3 geometry classes used here.

A `_Slope` is a specialisation of the abstract class `_Geoobject`. As such it inherits all the properties and associations. Referring to Dalrymple's et al. and Caine's slope model (Caine 1974, Dalrymple et al. 1968) a `_Slope` may again *contain* `_Slopes`. Therefore, it must be defined which of them is the hierarchically superordinately *superslope* and which one is the *subslope*. For example the valley side of the Turtmann valley would be the *superslope* regarding the side valleys cutting it, which are likewise composed of at least two slopes. The association *contains* thus represents the nested hierarchy of landforms. Smaller landforms sit on top of bigger ones and may cover them partly or in total (Ahnert 1988, Dikau 1989, Brunsden 1996). Hierarchy is a fundamental property of natural systems.

Scale itself is explicitly not modeled as an attribute of any `_Geoobject`. The purpose of doing so is threefold. First, there is no uniform definition in geomorphology how scales have to be appointed (Barsch 1978, Kugler 1974, Dikau 1989, Ahnert 1996). Second, the recommended definitions may cause confusions regarding other natural sciences. Normally Dimensions are named from yocto (10^{-24}) to yotta (10^{24}) in steps of 10^3. Third, scale as a definition of the science of geomorphology can be easily derived from geometry properties of a `_Geoobject` using geomorphometrical approaches discussed in (Pike 1995) or (Rasemann 2004).

A `_Slope` *is bounded by* two or more `_Geoobjects` or specialisations of this class. At the upper end that is a crest, at the lower end the depth contour (Leopold et al. 1964, Ahnert 1970, Dehn et al. 2001).[2] It *consists of* one or more abstract classes `_Layer`. A `_Layer` again may *contain subLayers*. Because the `_Layer` is derived from `_Geoobject`, it exhibits association to a `Timespan` representing its *age* and to a `_State` likewise.

3.2 Formal Representation of a Typical Soil Slope

Similar to the abstract class `_Geoobject` only specialised classes of `_Slope` can be instantiated. A `SoilSlope` (Fig. 4) is aggregated of

[2]These geoobjects can be modeled as linear features (Löwner 2005, 2008), what was not made here, however from space reasons.

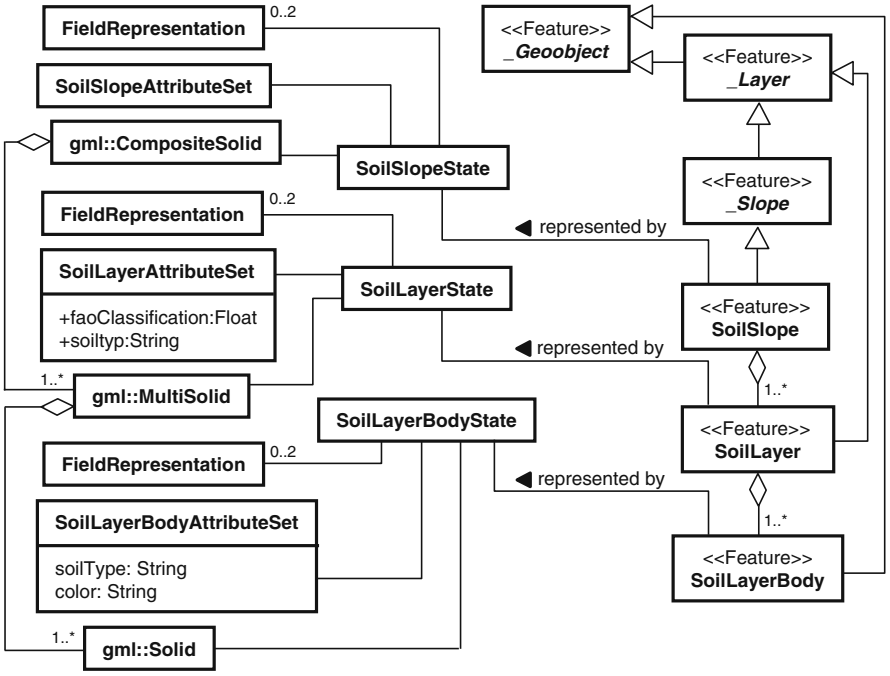

Fig. 4 UML diagram of the formal representation of a `SoilSlope`

one or more `SoilLayers` which are specialisations of a `_Layer`. As a `SoilLayer` may be dissected by erosional processes, it consists of one or more `SoilLayerBodies` which are subclasses of a `_Geoobject`. It is *represented by* a `SoilLayerBodyState` which keeps its associations to zero, one, or two `FieldRepresentations` and to exactly one `SoilLayerBodyAttributeSet`. This class contains all the attributes worth to be stored for a `SoilLayer`. This might be the `soilType` and the `colour` for instance.[3] Furthermore it *has* one GML3 geometry, a `Solid`. Thus, a `SoilLayerBody` cannot be further divided into smaller parts. A `Solid` consists of exactly one exterior `Surface` which is not depicted here (r.f. Fig. 2).

One or more `SoilLayerBodies` build up a `SoilLayer`. On the geometry side this is formalised by a aggregation association of one or more `Solids` to a `MultiSolid`. The advantage of a `MultiSolid` is that parts may be disjoint, touching, or disconnected. A disadvantage is the lack of topological structure as the parts of a `MultiSolid` may overlap. From both, the semantically and the geometrically point of view, this must be explicitly demanded.

[3] Note that this modeling approach does not mean to be complete neither in terms of classes that may be defined, nor in terms of attributes describing the characteristic of a modeled class.

The SoilSlope is aggregated by SoilLayers. One or more MultiySolids of one or more SoilSlopes build up a CompositeSolid. This GML3 geometry consists of Solids that must not overlap as well, but they must be topologically connected along their boundaries.

Modeling of other slopes can be done analogous to the SoilSlope discussed here. Hereby the number of subclasses of _Slope only depends on the different sets of attributes one may find for special slopes. Genesis is not a reason for a certain specialisation of _Slope. In this semantic model the development of every _Geoobject can be stored by an association to a _Geoprocess as discussed in Sect. 3.3. We state that only a SoilSlope, a DebrisSlope, and a RockSlope has to be modeled. This is due to different set of attributes one may define in order to represent the different kinds of materials.

3.3 Formal Representation of a Geoprocess

Landforms are results of processes that, on the one hand, alter their geometry by transportation of material and change their internal properties by weathering. On the other hand, a _Geoobject influences a _Geoprocess by its shape and internal resistance to erosion processes, for instance. Thus, this dichotomy is not mono directional. For a short time a process may alter landforms, but on a long time scale it is affected by a land surface's feature (Schumm and Lichty 1965). It depends on the internal properties of a landform, whether it is affected by a process or not (Schumm 1973).

In this formal representation of land surface features the modeling of a class _Geoprocess serves two goals: First, to store the interconnection of two or more _Geoobjects as a process-related accessibility; second, to represent the genesis of a _Geoobject.

A _Geoprocess holds two associations to a _Geoobject (Fig. 5). It *alters* one or more _Geoobjects while a _Geoobject *enables* one or more _Geoprocesses. It is *driven by* a _Processforce, which might be specialised but always stronger than the internal thresholds of the landform altered.

A _Geoprocess occurs during a given Timespan. This might be different to the Timespan of the corresponding _Processforce, again depending on the internal thresholds of the _Geoobject. Take gravity as an example. It is present at all times but only the weathering of a wall determines, whether and when a rock fall takes place.

After a _Geoobject was altered by a _Geoprocess, its _State has changed. This could be the change of one part of the geometry or of an attribute value of the corresponding AttributeSet. Therefore, a method actualize (Geoobject) is formulated that actualizes the associated _Geoobject. This method corresponds to the action part of a trigger used in database management

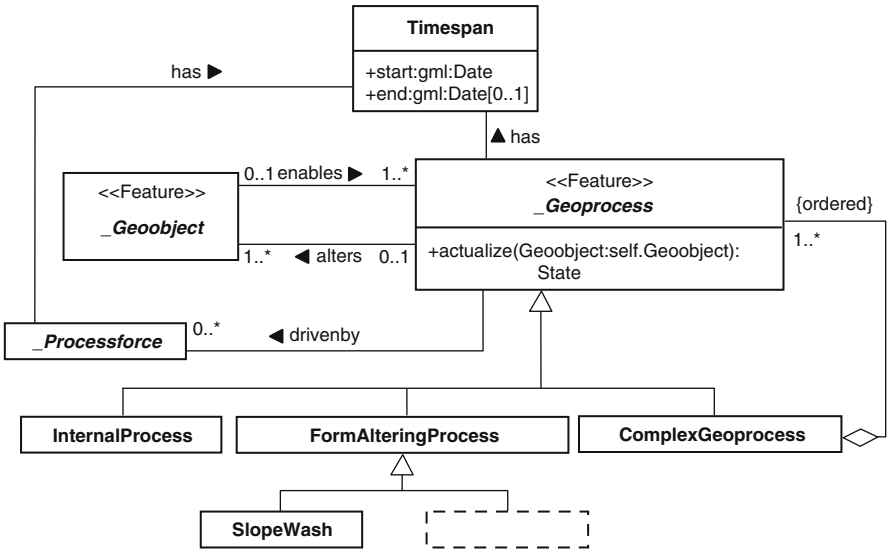

Fig. 5 UML diagram of the formal model of a Geoprocess

systems, if a change of one attribute has to entail a change of other datasets (Ullman 1988).

The change of a feature has to be propagated to that one it is composing. Thus, a change of a minor part's `_State` of a `_Geoobject` induces change of the major `_Geoobject's _State` as well. If, for instance, a `Surface` of a `_SoilLayerBody` is changed by a `_Geoprocess`, a new instance of its `SoilLayer's State` as well as a new instance of a `_State` of the `SoilSlope` has to be created. That does not mean to copy all the data associated with the `_Geoobject`. If aggregation of `_Geoobjects` viewed as tree in graph theory (Jungnickel 1991), only father knots have to be actualized but not the brothers. Due to XML's XLink syntax, every constant geometry can be reused (Fig. 6).

The abstract class `_Geoprocess` must be specialised for instantiation. Subclasses modeled here are the `InternalProcess`, the `FormAlteringProcess`, and the `ComplexGeoprocess`. First refers to changes of the internal state of a feature, i. e. the set of a `_Geoobject's` attributes, while the second refers to a process, changing the `-Geoobject's` geometry by the transport of material. The last one is an aggregation of `_Geoprocesses`. The parts have to be ordered by time of occurrence. This is possible because every subclass of a `_Geoprocess` has its own `Timespan`. The association of a `ComplexGeoprocess` is meant to store the genesis of a `_Geoobject`. Then the `_Geoobject` can be viewed as an integral of all processes over a given timespan.

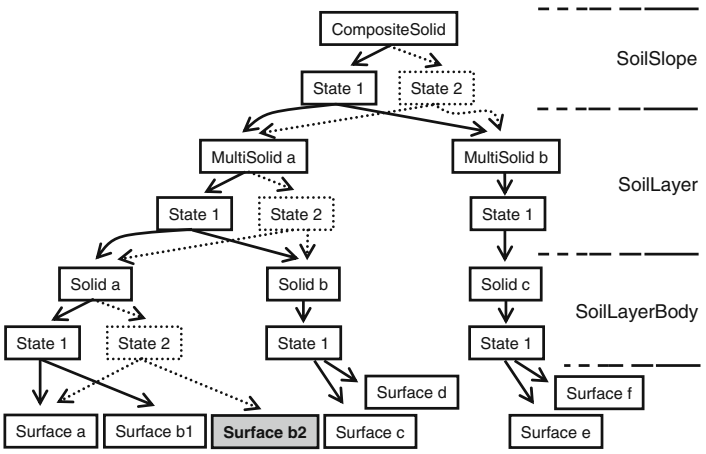

Fig. 6 A `SoilSlope` and its composing parts at two different states viewed as a tree in graph theory. The `SoilSlope` consists of a `CompositeSolid`, which again consists of two `MultiSolids` (*a, b*). `MultiSolid` *a* consists of two `Solids` (*a, b*) from which (*a*) is bounded by two `Surfaces` (*a, b1*) at State 1 (*solid arrows* and *boxes*). After a `Geoprocess` changed `Surface` *b1* to *b2*, the change is propagated to its father knot until the top geometry is actualized. On this path, new `_States` are initialized pointing at the valid geometries (*dotted arrows* and *boxes*). Only data of `Surface` *b2* will be additionally stored, none of the geometries will be erased

4 Discussion

In this article we presented an application schema for landforms, processes and their interaction based on the Geography Markup Language 3. Therefore, it is built up on international standards like the ISO 19107, ISO 19123, and others. While GML3 is implemented using the Internet standard Extensible Markup Language (XML), it fulfills all demands of heterogeneity (Sect. 1). Thus, it is capable to serve the lossless exchange of data about landforms which are extensive in geometry and internal properties.

To formalize some concepts of the science of geomorphology we followed a object oriented approach, despite to advancements of field based methods. Of course, geomorphometry has revealed reasonably findings in the filed of classification of land surface units (Dikau 1989, 1996, Dikau and Schmidt 1999), quantitative analysis of the surface (e.g. Zeverbergen and Thorne 1987, Evans 1972, Neil and Mark 1987) or object extraction (e.g. Brändli 1997, Löwner et al. 2003), partly with respect to hydrological evidence (e.g. Schmidt et al. 2000). Nevertheless, when defining an application schema for landforms, more than their boundary layer has to be considered. We adopted a normative approach to define a target format for landforms. This needs a definition of what we expect a landform to be in terms of objects of the real world, like Thornbury did in the 1950th (Thornbury 1956). Moreover, there are some doubts of the semantic evidence of land surface classification (Fisher

and Wood 1998, Dehn et al. 2001), which is almost limited to the 2D land surface (Chorley 1972), a limitation that should be overcome (Raper and Livingstone 1995).

The 3 dimensional geometry classes of the Geography Markup Language are used to represent the shape of the landforms modeled. This is of great advantage concerning interoperabilty while GML3 implements the international standards for spatial data. Using the GML3 boundary representation implies that geoobjects, like other objects in an object oriented view, have crisp boundaries (Burrough 1996). This seems to be a problem, because even on the surface boundaries of landforms are hard to determine. Often fiat boundaries are defined by the scientist (Smith 2000), like on geomorphological maps. When using a normative approach of semantic modeling, it is accidental whether boundaries can easily be determined in the field or not. It is more crucial that landforms actually have boundaries (Couclelis 1996).

In addition to geometry, representation the internal properties of geoobjects are modeled here. We did this using a set of attributes that has to be redefined for every type of geoobject. In this model the status of the set of attributes as well as the geometry are not directly linked with the geoobject. They are valid for a specific state of the geoobject and therefore, independent from the semantic identity of the landform. While geometry and attributes like, for instance, soil type might change over thousands of years, it is possible to represent still the same object in this formal model. Therefore, one has the ability to store the evolution of a geomorphic system. If desired, the state of a system thousand years ago may be queried in a database. Even scenarios of process modeling may be stored.

The presented application model allows the representation of processes. This is meant to be a representation within a database, not for empirical or physical based process modeling. Casually some formula may be stored as an attribute of a specific class. The main reason to model geoprocesses is a general association with a geoobject. Despite the taxonomy of different geomorphic processes is not very detailed in this approach, it is possible to represent a geoobject's genesis in terms of different processes affecting it during lifetime for different periods of time. The differentiated modeling of a `_Geoprocess` and a `ProcessForce` enables the representation of internal resistance or thresholds of a landform against external forces (Schumm 1973, 1979). For instance a rain fall event can be stored as long as it might take as well as the form altering process of overland flow. Moreover, this representation is capable to map cascading systems of material transportation and storage in more detail than other formalisms used in geomorphology (Löwner 2008).

The main advantage of the formal model presented here is that it is based on international standards. For this reason it is applicable to all informational technologies developed for data exchange. To archive this, it needs to do the following. First, the model needs to be extended in terms of more classes representing more geoobjects than in this study. Therefore, the model is expandable in a monotone way (Gruber 1993), meaning that the existing formalisation does not need to be altered when adding new concepts. This can be found in Löwner (2005) and Löwner (2008). However, the approach presented here strictly divides the shape and internal properties of a geoobject from it semantic identity. Second, a formulation of the developed model in a true GML3 schema needs to be done. While expressing a

model in UML is more usefully for discussion, a web feature service needs a XML-schema. Third, as our intension is the development of a target format to store the complexity of the interrelationship of landforms and processes, this model needs to be expanded in terms of field work. This alludes to the representation of meta data, meaning how the data is archived in the field. A basic approach for representing field and laboratory data is given in (Schmidt 2001). Nevertheless, meta data have to be represented using the ISO 19115 ISO/FDIS 19115 to archive interoperability and thus a possibility to transfer even complex data without loss of semantical information.

References

Ahnert F (1970) An approach towards a descriptive classification of slopes. Z. Geomorph. N. F. Suppl.-Bd. 9: 71–84

Ahnert F (1988) Modelling landform change. In: Anderson M G (ed) Modelling geomorphological systems. Wiley, Chichester, 375–400

Ahnert F (1996) Einführung in die Geomorphologie. Eugen Ulmer, Stuttgart

Barsch D (1978) Active rock glaciers as indicators for discontinous alpine permafrost. An example from the Swiss Alps. In: Proceedings of the third international conference on permafrost, 1. NRC-Ottawa 349–352.

Bishr Y A (1998) Overcoming the semantic and other barriers to GIS interoperability. Int. J. Geog. Inf. Sci. 12(4): 299–314

Booch G, Rumbaugh J, Jacobson I (1999) The unified modeling language guide. Addison-Wesley, London

Brändli M (1997) Modelle und Algorithmen für die Extraktion geomorphologischer und hydrologischer Objekte aus digitalen Geländemodellen, Geoprocesing 32

Brunsden D (1996) Geomorphological events and landform change. Z. Geomorph. N. F. Suppl.-Bd. 40: 273–288

Burrough P A (1996) Natural objects with indeterminate boundaries. In: Burrough P A, Frank A U (eds) Geographic objects with indeterminate boundaries, Gisdata 2, Taylor & Francis, London, 3–28

Burrough P A, Frank U A (1995) Concepts and paradigms in spatial information: are current geographic information systems truly generic? Int. J. Geog. Inf. Syst. 9(2): 101–116

Caine N (1974) The geomorphic processes of the alpine environment. In: Ives J D, Barry R G (eds) Arctic and alpine environments. Methuen, London, 721–748

Chorley R J (1972) Spatial analysis in geomorphology. In: Chorley R J (ed) Spatial analysis in geomorphology. Methuen, London, 1–16

Chorley R J, Kennedy B A (1971) Physical geography – a systems approach. Prentice Hall, London

Couclelis H (1996) Towards an operational typology of geographic entities with ill-defined boundaries. In: Burrough P A, Frank A U (eds) Geographic objects with indeterminate boundaries, Gisdata 2, Taylor & Francis, London, 45–55

Cox S, Daisey P, Lake R, Portele C, Whiteside A. (2005) (eds.) OpenGIS Implementation Specification \#02-023r4: OpenGIS. Geography Markup Language (GML) http://www.opengeospatial.org/standards/gml . Cited Nov 2007

Cranefield S, Purvis M (1999) UML as an ontology modelling language. In: Proceedings of the workshop on intelligent information integration of the 16th international joint conference on artificial intelligence (IJCAI-99), 230–239 http://sunsite.informatik.rwth-aachen.de/Publications/CEUR-WS/Vol-23/cranefield-ijcai99-iii.pdf . Cited 17 Jun 2007

Dalrymple J B, Blong R J, Conacher A J (1968) An hypothetical nine unit landsurface model. Z. Geomorph. N. F. 12: 60–76

Dikau R (1989) The application of a digital relief model to landform analysis in geomorphology. In: Raper J (ed) Three dimensional application in geographic information systems. Taylor & Francis, London, 51–77

Dikau R (1996) Geomorphologische Reliefklassifikation und -analyse. Heidelberger Geographische Arbeiten 104: 15–23

Dikau R (1998) The need for field evidence in modelling landform evolution. In: Hergarten S, Neugebauer H J (ed) Process modelling and landform evolution, Lecture notes in earth sciences. Springer, Heidelberg, 3–12

Dikau R, Schmidt J (1999) Georeliefklassifikation. In: Schneider-Sliwa R, Schaub D, Gerold G (eds) Angewandte Landschaftsökologie. Grundlagen und Methoden. Springer, Heidelberg, 217–244

Dehn M, Gärtner H, Dikau R (2001) Principles of semantic modelling of landform structures. Comput. and Geosci. 27: 1005–1010

Evans I S (1972) General geomorphometry, derivates of altitude, and descriptive statistics. In: Chorley R J (ed) Spatial analysis in geomorphology. Methuen, London, 17–90

Fisher P, Wood J. (1998) What is a mountain? Or the englishman who went up a boolean geographical concept and realized it was fuzzy. Geography 83: 247–256

Foley J, van Dam A, Feiner S, Hughes J (1995) Computer graphics: principles and practice. 2nd Ed. Addison Wesley Professional

Fonseca F T, Egenhofer M J (1999) Ontology-driven geographic information systems. In: Bauzer C (ed) Proceedings of fifteenth annual ACM, No7 in ACM symposium on advances in geographic information systems. ACM, Kansas 14–19

Fonseca F T, Egenhofer M J, Davis C, Câmara G (2002) Semantic granularity in ontology-driven geographic information systems. Annals of methematics and artificial intelligence – Special issue on spatial and temporal granularity 36(1–2): 121–151

Frank A U (2001) Tiers of ontology and consistency constraints in geographical information systems. Int. J. Geogr. Inf. Sci. 15: 667–678

Gröger G, Benner J, Dörschlag D, Drees R, Gruber U, Leinemann K, Löwner M-O (2005) Das interoperable 3D Stadtmodell der SIG 3D. Zeitschrift für Vermessungswesen 130(6): 343–353

Gröger G, Kolbe T H (2005) Normen und Standards für 3D-Geodaten. In: Corrs V, Zipf A (eds.) 3D-Geoinformationssysteme – Grundlagen und Anwendungen, Wichmann, Heidelberg, 56–70

Gröger G, Kolbe T H, Czerwinski A. (2007) City Geography Markup Language (CityGML) OGC Best Practices Document, Version 0.4.0, OGC Doc. No. 07-062, Open Geospatial Consortium

Gruber T R (1993) Towards principles for the design of ontologies used for knowledge sharing. Int. J. Hum. Comput. Stud. 43: 907–928

Herring J, (2001) The OpenGIS abstract specification, Topic 1: Feature Geometry (ISO 19107 Spatial Schema), Version 5. OGC Document 01-101

Hugget R J (2003) Fundamentals of geomorphology. Routledge, London

Hunter D, Rafter J, Fawcett J, van der Vlist E, Ayers D, Duckett J, Watt A, McKinnon L (2004): Beginning XML. Wiley, Indianapolis

ISO 19109 (2002) Geographic Information – Rules for application schema.

ISO/FDIS 19115 (2002): Geographic Information – Metadata. ISO Technical Committee 211, Final Draft International Standard

ISO/DIS 19123 (2004) Geographic Information – Schema for coverage geometry and functions. Draft International Standard

Jungnickel, D (1991) Graphs, networks and Algorithms. Springer, Berlin

Kugler H (1974) Das Georelief und seine kartographische Modellierung. Dissertation B, Martin-Luther-Universität Halle

Lake R, Burggraf D S, Trninić M, Rae L (2004) Geography mark-up language. Wiley, Chichester

Leopold L B, Wolmann M G, Miller J P (1964) Fluvial processes in geomorphology. Freeman & Company, London.

Löwner M-O (2005) Semantische Modellierung und Repräsentation geomorphologischer Objekte in einem geographischen Informationssystem (GIS) unter besonderer Berücksichtigung von

Wänden und steilen Hangbereichen (Semantic Modelling of geomorphic objects within a geoinformation system (GIS) with emphasis on free faces and steep slopes). PhD-thesis at the University of Bonn. 144 pp. urn: nbn:de:hbz:5 N-06839 http://hss.ulb.uni-bonn.de/diss_online/landw_fak/2005/loewner_marc-oliver . Cited 28 Nov 2007

Löwner M-O (2008) Formale semantische Modellierung von geomorphologischen Objekten und Prozessen des Hochgebirges zur Repräsentation in einem Geoinformationssystem (GIS). Bonner Geogr. Abh.

Löwner M-O, Dörschlag D Plümer L (2003) Interaktive Geoobjekterkennung in digitalen Höhenmodellen mittels Parametervisualisierung, Kartographische Schriften 7: 59–66

Neill M P, Mark D M (1987) On the Frequency Distribution of Land Slope. ESP&L 12: 127–136

Oesterreich B (1998) Objektorientierte Softwareentwicklung – Analyse und Design mit der Unified Modeling Language. Oldenbourg, München

Otto J-C, Dikau R (2004) Geomorphologic system analysis of a high mountain valley in the Swiss Alps. Z. Geomorph. N. F. 48/3: 323–341

Paine A D M (1985) Ergodic reasoning in geomorphology: time for a review of the term? Prog. Phys. Geogr. 9: 1–15

Pike R J (1995) Geomorphometry - progress, practice, and prospect. Z. Geomorph. N. F. Suppl. -Bd. 101: 221–238

Rambauske T (2003) Chronik 2003. http://www.bergnews.com/service/chronik2003.htm . Cited 21 Apr 2005.

Raper J, Livingstone D (1995) Development of a geomorphological spatial model using object-oriented design. Int. J. Geogr. Inf. Syst. 9(4): 359–383

Rasemann S (2004) Geomorphometrische Struktur eines mesoskaligen alpinen Geosystems Bonner Geographische Abhandlungen 111

Schmidt J (2001) The role of mass movements for slope evolution: conceptual approaches and model applications in the Bonn area. PhD-thesis, Rheinische Friedrich-Wilhelms-University Bonn, Institute of Geography. http://hss.ulb.uni-bonn.de/diss_online/math_nat_fak/2001/schmidt_jochen/index.htm

Schmidt J, Hennrich K, Dikau, R (2000) Scales and similarities in runoff processes with respect to geomorphometry. Hydrological Processes 14: 1963–1979

Schumm S A (1973) Geomorphic threshold and complex response of drainage systems. In: Morisawa M (ed.) Fluvial geomorphology, 299–310

Schumm S A (1979) Geomorphic thresholds: the concept and its application. Transactions of the Institute of British Geographers (New Series) 4: 485–515

Schumm S A, Lichty W (1965) Time, space, and causality in geomorphology. Am. J. Sci. 263: 110–119

Smith B (2000) Fiat objects. Philos. Phenomenol. Res. 60/2: 401–420

Thornbury W D (1956) Principles of geomorphology. Wiley & Sons, New York

Ullman J D (1988) Principles of database and knowledge-base systems, vol. 1. Computer Science Press, New York

Vretanos P A (2005) (ed) OpenGIS Implementation Specification \#04-094: OpenGIS: Web Feature Service http://www.opengeospatial.org/standards/wfs . Cited Nov 2007

Worboys M F (1995) Geographic Information Systems: A Computing Perspective. Taylor & Francis, London

Yergeau F, Bray T, Paoli J, Sperberg-McQueen C M, Maler E. (2004) Extensible Markup Language (XML) 1.0 (Third Edition): W3C Recommendation, 4. Februar 2004 www.w3.org/TR/REC-xml . Cited Nov 2007

Zeverbergen L W, Thorne C R (1987) Quantitative analysis of land surface topography. ESP&L 12(4): 47–56

Semi-Automatic Digital Landform Mapping

Martin Schneider and Reinhard Klein

Abstract In this paper a framework for landform mapping on digital aerial photos and elevation models is presented. The developed mapping tools are integrated in a real-time terrain visualization engine in order to improve the visual recovery and identification of objects. Moreover, semi-automatic image segmentation techniques are built into the mapping tools to make object specification faster and easier without reducing accuracy. Thus, the high level cognitive task of object identification is left to the user whereas the segmentation algorithm performs the low level task of capturing the fine details of the object boundary. In addition to that, the user is able to supply additional photos of regions of interest and to match them with the textured DEM. The matched photos do not only drastically increase the visual information content of the data set but also contribute to the mapping process. Using this additional information precise landform mapping becomes even possible at steep slopes although they are only insufficiently represented in aerial imagery. As proof of concept we mapped several geomorphological structures in a high alpine valley.

Keywords landform mapping · semi-automatic image segmentation · image registration · matching

1 Introduction

Landform mapping often serves as the basis for various kinds of geomorphological investigations. The resulting geomorphological map decomposes the land surface into structural patterns, landforms and landform elements and is the standard tool to perceive and investigate an area at focus in a complex and holistic way. Geomorphological maps compile knowledge on landforms, surface processes and surface materials that have a widespread application in land management practices, natural hazard assessments or landform evolution studies (Cooke and

Martin Schneider (✉)
Institute of Computer Science II, University of Bonn, Römerstraße 164, 53117 Bonn, Germany
e-mail: ms@cs.uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_3,

Doornkamp 1974, Otto and Dikau 2004, Seijmonsbergen and de Graff 2006). Traditionally, geomorphological mapping is based on field work supplemented by the interpretation of aerial imagery and literature research. The availability and usage of high resolution digital elevation models, satellite and aerial images permits to avoid or reduce the amount of the time-consuming and costly field work, especially in remote and highly dynamic regions like high mountains.

Recent developments in remote sensing techniques and GIS aim at an automated recognition of geomorphological objects on digital terrain data. Automatic landform recognition is performed using elevation data only (Schmidt and Hewitt 2004; Asselen and Seijmonsbergen 2006, Seijmonsbergen and de Graff 2006), or by combining elevation data and imagery information (Schneevoigt and Schrott 2006). However, so far automatic recognition suffers from land surface complexity and its continuum character, represented by fuzzy landform boundaries, overlapping landforms and a great variety of structural properties. Consequently, automatic landform recognition is restricted to landform classification on a large scale, while the boundary of single landforms cannot be identified exactly. A detailed geomorphological map still requires manual landform mapping, either transferred from previously acquired field data, or genuinely mapped from remote sensing data on screen. The accuracy of the mapping result depends on the resolution of the terrain data, the visual perception of the virtual land surface morphology and the diligence and knowledge of user. Unfortunately, manual landform mapping is a particularly tedious and time-consuming task for the user. Modern GIS tools facilitate the compilation, production and distribution of geomorphological maps. Enhanced mapping tools, data layers, data base functions, symbol creation, print and web publishing are some of the enhancements provided by GIS. However, most of the cartographic features of GIS software are limited to a 2d representation of the data. Typically, different kinds of relief shading based on the derivatives of the elevation data are used to emphasize morphology changes and break lines in the land surface to compensate for that. Nevertheless, a fixed 2d aerial perspective of imagery and elevation data heavily restricts the perception of landforms. As a consequence, the conventional stereoscopic aerial photo interpretation is more and more replaced by the combined 3d visualization of aerial imagery and elevation data. Although this representation of the data improves the perception of landforms, so far 3d visualization software is usually limited to simple data exploration without the opportunity to interact with the data. Another difficulty in the creation of a geomorphological map is caused by the irregular sampling of the land surface in aerial imagery. Even in very high resolution data sets steep slopes are only very sparsely sampled and hence lack information that would allow for a precise analysis and mapping.

In this paper we present a framework for landform mapping that aims at removing the aforementioned restrictions. The proposed mapping tools allow the mapping of objects within a 3d visualization environment. By navigating in the 3d environment landforms can be inspected from arbitrary views and directly marked on the 3d land surface. The built-in semi-automatic image segmentation algorithms assist the user in specifying landforms and lead to a faster mapping process with fewer user interactions necessary compared to manual mapping without reducing accuracy. Sparsely sampled areas contained the data set, like steep slopes, can be enhanced with additional photos that can be interactively matched with the data by the user.

The matched photos do not only drastically increase the visual information content of the data set but also contribute to the mapping process. As a result, landform mapping can even be performed at places where there is insufficient information available in the aerial photography provided that the user matches appropriate photos. The object boundaries produced by our system are completely georeferenced and can be exported in the popular shapefile format and as such imported by other GIS for further processing.

2 Previous Work

Semi-automatic segmentation techniques provide high efficiency and accuracy by allowing users to do the object recognition and letting computers capture the fine details. These methods can basically be divided into region-based and boundary-based approaches.

Boundary-based methods cut out an object by allowing the user to surround its boundary with an evolving curve. The user traces along the object boundary and the system optimizes the curve in a piecewise manner. One well-known group of boundary-based techniques are those based on Intelligent Scissors (Mortensen and Barrett 1995; Falcâo et al. 1998). Intelligent Scissors is a highly interactive tool which formulates the boundary detection problem in an image as a shortest path search in graph. By planting an initial seed point, a path map is constructed that provides the minimum-cost path from the seed to every pixel in the image. By interactively moving a cursor near the boundary of an object, the path is extended according to the path map to form a boundary segment. Whenever the path deviates from the true object boundary the user can insert an additional seed point thereby fixing the old boundary segment and creating a new one starting from the newly created seed point. While these tools provides highly interactive visual feedback once all shortest paths have been computed, it is time-consuming to recompute them especially when the image is large. Therefore, several attempts (Mortensen and Barrett 1999; Falcâo and Udupa 2000; Wong et al. 2000; Kang and Shin 2002) were presented that aim at increasing the efficiency of the boundary construction by restricting the search domain. An application of Intelligent Scissors for landform mapping based on aerial imagery and elevation data inside a 3d visualization environment was presented in (Schneider and Klein 2006).

Region-based methods on the other hand allow the user to give loose hints which parts of the image are foreground or background without the need to fully enclose regions. An underlying optimization algorithm extracts the actual object boundary based on the provided user input. In the seminal work (Boykov and Jolly 2001) a graph cut optimization was used for this purpose. Since then, many approaches were published that extended and improved the original method (Li et al. 2004; Rother et al. 2004), while others aimed at further accelerating the graph cut (Lombaert et al. 2005; Juan and Boykov 2006).

In (Schneider and Klein 2007) landform mapping tools based on Intelligent Scissors and graph cut were presented that allow for a detailed mapping at steep slopes by taking into account photos in addition to conventional aerial imagery.

3 Digital Landform Mapping

In our landform mapping framework the user can navigate freely within the virtual landscape and explore objects of interests from arbitrary views. By providing simple mouse gestures directly on the 3d terrain surface, objects can be extracted easily from the relief. To this end the user input is projected to the underlying aerial photography and elevation data on which the respective segmentation algorithm is then performed. Afterwards, the obtained object boundary is projected back onto the 3d terrain surface. Since the segmentation is designed to provide immediate feedback the user has the impression to work completely in 3d.

3.1 The Visualization Engine

A vital requirement for our method is to ensure interactive response even for very large data sets comprising several gigabytes of raw data. Consequently, an efficient high quality, real-time visualization of the landscape as well as interactive feedback by the segmentation algorithm is needed. We use the terrain rendering system presented by Wahl et al. 2004which is based on a quadtree data structure (see Fig. 1). The system has proven to be able to visualize very large data sets efficiently and with high quality, e.g. data sets with a resolution up to a few centimeters for the aerial photography together with elevation models of about 1m covering areas of hundreds of square kilometers have already been visualized with real-time frame rates. Due to the good scalability of the system the real-time visualization of upcoming data sets of even higher resolution as a result of improved acquisition methods can be expected to run at real-time frame rates as well.

The visualization engine represents the aerial photography as well as the elevation model in a quadtree data structure (see Fig. 1). The root of the quadtree holds the entire domain of the data set in a single tile. Up to four children partition their

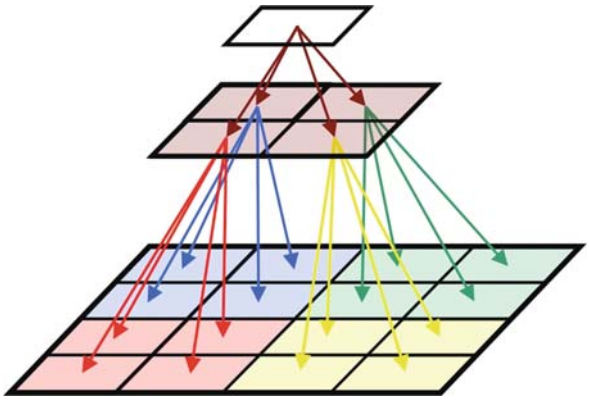

Fig. 1 Quadtree representation of the data used by the visualization engine as well as the segmentation algorithms

parents' domain into equally sized quarters, where each quarter has the same size as the parent tile, i.e. the four children represent the same domain but with twice the resolution. The quadtree is created by initially partitioning the air photo into equally sized square tiles of a given size representing the base level of the tree. The remaining levels of the quadtree are created by recursively downsampling the original air photo by a factor of two and merging 2×2 tiles from the lower level into one tile of the same size at the higher level. The segmentation tool operates on the same quadtree data structure used for rendering. This is advantageous in several respects: image data for rendering and segmentation has to be held in memory only once, the quadtree data structure allows fast access to spatial subparts of the terrain and different levels of detail of the data set are already available.

Typically, geomorphological maps are represented as vector data consisting of lines and polygons. In order to be able to overlay such vector data on the textured DEM two different methods for their visualization were developed and integrated in the visualization system. The first method (Schneider et al. 2005) is a texture-based approach that creates textures on-the fly in an offscreen buffer. A perspective reparameterization is applied taking into account the current point-of-view to optimize the texture utilization. The basic idea of the second approach (Schneider and Klein 2007) is to extrude the vector data to polyhedra and to compute their intersection with the terrain surface. Graphics hardware can be used to perform the intersection tests very efficiently.

3.2 Intelligent Scissors Based Landform Mapping

The Intelligent Scissors based segmentation algorithm formulates the boundary detection problem as an optimal path search in a graph. The objective is to find the optimal path from a seed node to a destination node where pixels in the image represent nodes with directed and weighted edges connecting its eight adjacent neighbours. An optimal path is defined by the minimum cost path, i.e.,a path with the smallest sum of edge costs. Since a shortest path in the graph should correspond to an object boundary in the image, pixels with strong edge features in the image should lead to low local costs in the graph and vice-versa. Hence, local costs are created as a weighted sum of the edge features.

The user starts the segmentation process by planting an initial seed point by simply clicking with the mouse (see Fig. 2) on the 3d terrain surface. Hereafter, the path map is constructed that provides the minimum cost path from the selected seed node to every other node. By moving the mouse, a path from the seed point to the current mouse position is interactively displayed snapping to nearby object boundaries. If the proposed boundary segment deviates from the desired object boundary, a new seed point can be established by the user fixing the current path segment and starting a new one that is extended from the new seed.

Despite the fact that the Intelligent Scissors technique provides a powerful tool for image segmentation its speed and memory consumption constrain its feasibility

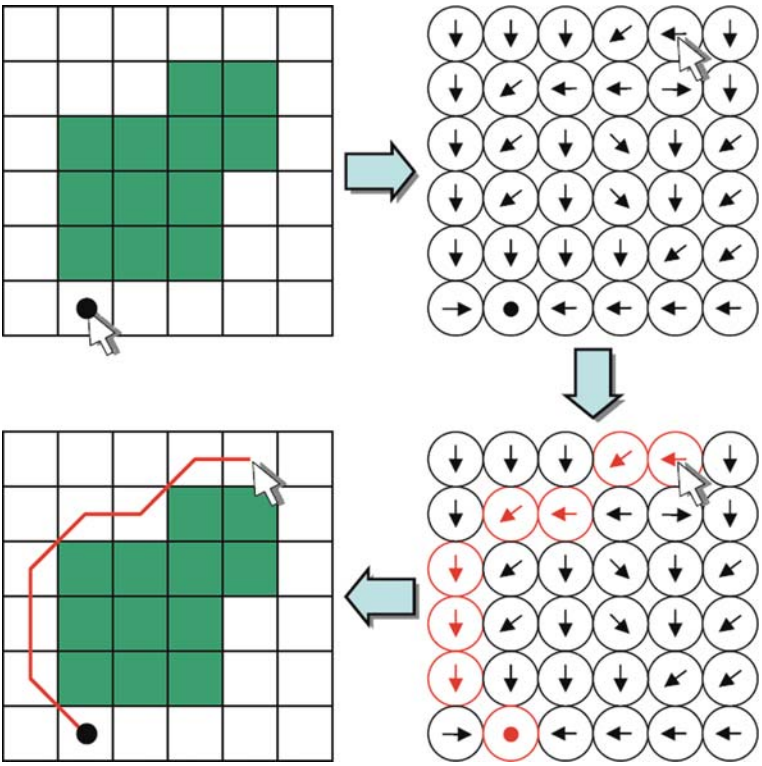

Fig. 2 Schematic workflow of Intelligent Scissors by means of a simple example image. On the *top left*an initial seed point is planted. In the next image the corresponding path map is shown. The shortest path with respect to a new mouse position is highlighted in the *bottom right*image and the resulting object border is shown in the *bottom left*image

when large data sets need to be processed. Since in the case of terrain data usually very large data sets must be processed, the direct application of the original Intelligent Scissors approach is not possible. In order to ensure interactive response even on very large data sets the quadtree data structure is exploited in two ways: Localizing the search domain within a quadtree level and employing a multilevel banded heuristic to exploit the hierarchical structure. Localizing the search domain means to search for a shortest path only in a restricted area around the user input. The search domain is incrementally extended when needed depending on the user input. The necessary parts of the data set are loaded from hard disk and the corresponding edge features are computed on-the-fly. As a result only a very small subset of the whole data set has to be held in memory while the majority resides on disk. The multilevel banded heuristic starts the segmentation on a coarser level and propagates the result to the next higher resolution level where the segmentation is performed only within a narrow band surrounding the projected result from the coarser level. This procedure is repeated until the highest resolution level is reached. A more detailed description of the algorithm can be found in Schneider and Klein (2006).

3.3 Graph Cut Based Landform Mapping

The graph cut based algorithm (see Fig. 3) formulates object extraction as a binary labeling problem which assigns to each node p in a graph, i.e. pixel, a unique label $x_p \epsilon$ {foreground, background}. The solution $X = \{x_p\}$ can be obtained by minimizing a Gibbs energy (Geman and Geman 1984)

$$G(X) = \sum_{p} R(x_p) + \lambda \sum_{(p,q)} B(x_p,x_q).$$

$R(x_p)$ is the likelihood energy that encodes the cost when the label of node p is x_p, i.e. it encodes the likelihood that the pixel belongs to the foreground or background. The likelihood is estimated as the color similarity of the pixel′s color to the color distribution of the areas marked by the user as foreground and background. The prior energy $B(x_p, x_q)$ denotes the cost when the labels of adjacent nodes p and q are x_p and x_q, respectively. B is defined as a function of the color gradient between the two nodes. In other words, B is a penalty term when adjacent nodes are assigned to different labels. The more similar the colors of the two nodes are, the larger is B, and thus the less likely the edge is on the object boundary. The influence of the region and boundary terms is controlled by the weighting factor λ. Decreasing λ leads to more complex and longer boundaries and generally increases the number of resulting objects, especially small ones. The energy function G(X) is minimized using the max-flow algorithm presented in (Boykov and Kolmogorov 2001), which is especially designed for vision problems. Since it is crucial to generate the object

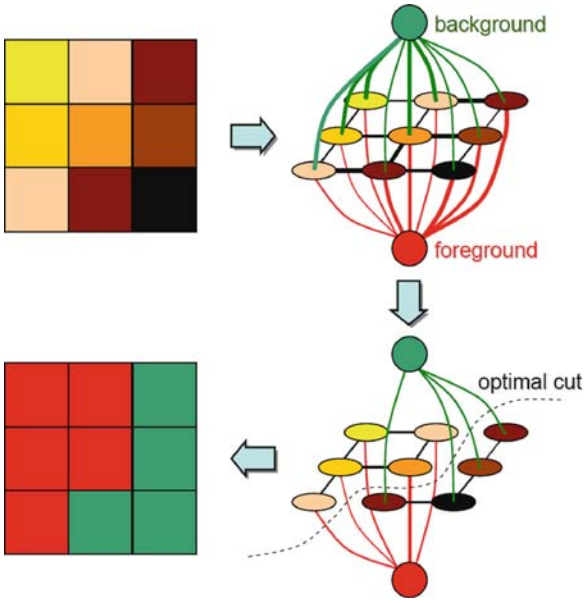

Fig. 3 Schematic illustration of graph cut based segmentation. The *top row*shows a simple example image on the *left* and the corresponding graph on the *right*. The edges between the pixels corresponds to $B(x_p, x_q)$ and the edges to the *top*and *bottom*node to $R(x_p)$. The thickness of the edges relates to the magnitude of the associated costs. The *bottom right*image shows the graph after the optimization with the optimal edges only. In the final image the segmentation of the image into foreground and background is visualized

boundary with very little delay, a multilevel banded heuristic (Lombaert et al. 2005) is applied to reduce both running time and memory consumption of the graph cut optimization. The procedure is similar to the heuristic proposed to accelerate the Intelligent Scissors based approach.

To specify an object, the user marks a few lines by dragging the mouse cursor while holding the left or right button to indicate foreground or background. This high-level painting user interface does not require very precise user inputs since marking lines do not need to be very close to the actual boundary. If, however, the result produced based on this user input is not satisfying, further user editing is possibly. By roughly marking additional lines in the incorrectly labelled areas the result can be updated appropriately.

4 Augmenting a Textured DEM with Additional Photos

As a result of improved acquisition devices, aerial photography and digital elevation models with a resolution up to a few meters or even centimeters have become available. However, the resolution specification is with respect to a surface perpendicular to the acquisition direction, whereas the effective resolution of surfaces at oblique angles is much lower. Hence, the representation of steep slopes, which are of great interest in many disciplines, suffers from a sparse sampling. As a consequence, even in recent high resolution data sets there is usually insufficient information available in these areas to perform a detailed analysis or precise mapping. To remove these restrictions, we allow the user to provide additional photos in order to increase resolution in areas of interest. The photos can be matched with the textured elevation model interactively by marking corresponding points in the photo and in the 3d environment, which allow solving for the camera matrix. The visualization of the photos is performed by means of projective texture mapping using the camera matrix resulting from the matching procedure. The associated visibility problem is solved by applying a shadow mapping algorithm. Different illumination conditions inherent in the photos are compensated using a histogram matching algorithm. In addition to that, the best available views for every surface point are determined and blended together appropriately (Fig. 4).

Fig. 4 The images show a screenshot of the textured elevation model (*left*), with a photo mapped on it (*middle*) and with applied histogram matching and blending (*right*)

4.1 Matching of Photos with the Textured DEM

The task of matching a photo to a textured digital elevation model can be formulated as estimating the camera projection matrix which is known as resectioning or the Perspective-n-Point (PnP) problem. Given sufficiently many correspondences between world and image points the camera matrix can be determined from them by using the Gold Standard algorithm (Hartley and Zisserman 2004). In our framework we let the user select appropriate correspondence points in the supplied photo and in the 3d the terrain through a simple point-and-click interface.

Given the point correspondences between 2d image points x_i and 3d world points X_i, a camera matrix P is estimated that maps the X_i onto the x_i, i.e. $PX_i=x_i$. Since P is a 3×4 matrix with 11 degrees of freedom, 11 equations are needed to solve for it. Because each point correspondence results in two equations at least 6 correspondences are needed. Given the minimum number of correspondences an exact solution can be found by organizing the equations in a matrix A and solving for $Ap=0$, where p contains the entries of the camera matrix. If 6 or more points are provided, there is no exact solution and the maximum likelihood estimate of P is determined. To this end, the image points x_i as well as the world points X_i are normalized using similarity transformations T and U, respectively. Then, a linear solution is computed using the Direct Linear Transformation (DLT) algorithm which computes the camera matrix as the unit singular vector of A corresponding to the smallest eigenvalue. This linear result is then used as input for the non-linear minimization of the geometric error $\sum_i \| x_i - PX_i \|^2$with the Levenberg-Marquardt algorithm. Lastly, the camera matrix P of the original points is computed in a denormalization step from the estimated camera matrix $\overline{P}$ of the normalized points as $P = T^{-1}\overline{P}U$.

4.2 Visualization of Photos Together with the Textured DEM

Rendering the photographs on the terrain is performed using projective texture mapping (Segal et al. 1992). Projective texture mapping is directly applicable to image-based rendering because it simulates the inverse projection of taking photographs with a camera and can be thought of as replacing the camera with a slide projector that projects the original photo back onto the terrain. In order to perform projective texture mapping, the estimated camera matrix is used to calculate the texture coordinates for each vertex by multiplying its 3d position with it. Since projective texture mapping does not automatically perform visibility tests, we apply an image-space shadow map algorithm (Williams 1978) to handle visibility.

In general, a photograph will only contain a part of the object of interest. Thus, it is usually necessary to combine multiple overlapping images in order to render the entire object at increased resolution. Consequently, some parts of the object are covered by only one while others might be covered by several photos. If a surface point is contained in multiple images, the renderer has to decide which image or combinations of them to use. What is more, the images will usually not agree perfectly in the overlapping areas due to different lighting conditions, non-lambertian

reflection or unmodeled geometric detail in the terrain surface. To account for the different color distributions in the data, we apply the histogram matching proposed in (Reinhard et al. 2001), which adapts the distribution of color values in the photo to that of the aerial photography. Although histogram matching helps to adapt the photos to the data set and each other, they still do not match perfectly in the overlapping regions and at the borders. Using only a single view at a pixel means that neighboring pixels may be sampled from different photos, which can cause visible seams in a rendering. To account for this, transitions are smoothed by performing a weighted averaging of the best available views.

5 Landform Mapping Using Additional Photos

A limitation of all mapping methods working solely on aerial photography and elevation data is that they cannot be used to perform a reasonable mapping at steep slopes due to lack of data in these areas. Therefore, we revise our segmentation algorithms to be able to use the matched photos as input in addition to the data set. To this end, we compute a map of the surface, and hence for the matched photos as well, into the plane. For an infinitesimal small surface patch a projection to its tangential plane defines a perfect mapping. In contrast to that, parameterizing the terrain surface as a whole introduces noticeable distortions. As a reasonable compromise we apply an adaptive, i.e. distortion-controlled, parameterization of the terrain surface in a preprocessing step. For this purpose, we start on the finest level of the quadtree and parameterize each patch in it. Then, we recursively merge and parameterize neighboring 2×2 tiles until the the distortion imposed by the parameterization exceeds a given threshold.

5.1 Parameterization

A parameterization of a surface can be viewed as a one-to-one mapping from a suitable domain to the surface. Given an orientable 2-manifold surface patch $S \subset \mathbb{R}^k$a parameterization is defined as a homeomorphism

$$\phi{:}\Omega \subset \mathbb{R}^2 \rightarrow S$$
$$(u,v) \mapsto \phi(u,v)$$

from the parameter space Ω into S. These surfaces are usually represented by triangular meshes and the mappings are linear over the triangles. Parameterizations almost always introduce distortion in either angles or areas. Most applications demand parameterizations that minimize these distortions in some sense. Many different ways of achieving this have been proposed in the literature. A comprehensive survey of local parameterization methods can be found in (Floater and Hormann 2005).

Several parameterization methods demand the boundary mapping to be fixed in advance and map to convex polygons which may be sufficient or even desirable for

some applications. In our case it would be advantageous because fixing the boundary vertices and mapping to a square would result in patches with coinciding borders. For the sake of quality, however, we use the parameterization presented in (Degener et al. 2003) that does not constrain the boundary. The algorithm quantifies angle and global area deformations simultaneously and lets the user control the relative importance between them through a parameter. We choose this parameter in order to obtain a parameterization that is optimized for a uniform sampling of the surface. This, however, comes at the cost of having to handle patch borders appropriately, since they do not match.

A 2d image of the patch can then be created by rendering the geometry of the patch where its 2d texture coordinates are used as vertex positions (see Fig. 5). The corresponding texture coordinates for accessing the photo are computed by multiplying the original vertex position with the estimated camera matrix. Since the surface geometry patches in our data set contain usually about up to a thousand vertices, the parameterization of a patch with the aforementioned algorithm might take up to several seconds. For that reason, we perform the parameterization in a preprocessing step, store the texture coordinates on hard disk and load them on demand. Further details of the described procedure can be found in (Schneider and Klein 2007).

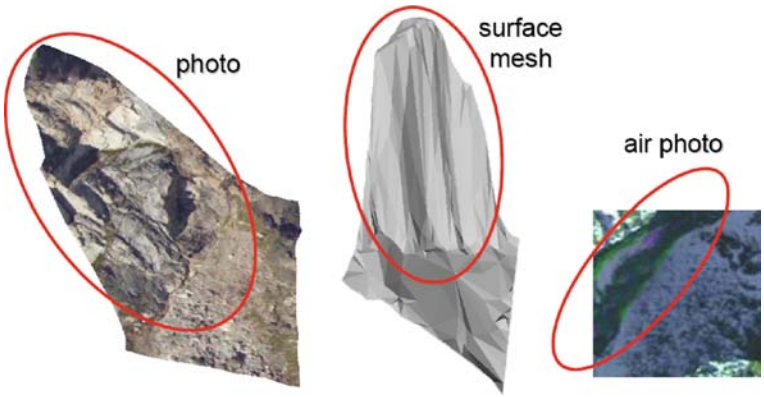

Fig. 5 An example of a surface geometry patch (*top*). The corresponding textured projections into the plane using the computed parameterization and a photo (*bottom left*) and using orthogonal projection and the aerial photography (*bottom right*). Corresponding areas are marked by ellipses demonstrating the irregular sampling of the surface in the air photo

6 Results and Conclusions

As a proof of concept we mapped meso-scale geomorphologic landforms in an HRSC-A data set of Turtmann valley in Switzerland. The data set covers an area of about 200 $km^2$and has a resolution of 1 m for aerial photography as well as elevation data. Unfortunately, HRSC-A data do only contain a red channel that is close

to infrared, which is why the data set′s color look, despite postprocessing, somewhat unnatural.

For the Turtmann Valley a detailed geomorphological map and a GIS database of landform polygons exists (Otto and Dikau 2004) that are based on field work data and manual mapping on the HRSC data using ArcGIS. In order to assess the applicability of the new mapping tools, different landforms were mapped again with our new tools and compared to the manually mapped objects. Fig. 6and 7show some mapping results obtained with our method. In conclusion, when compared to traditional manual mapping, semi-automatic segmentation in a 3d environment offers increased insight into the structure of the landscape and the objects it contains and provides quicker and more accurate mapping results at the same time.

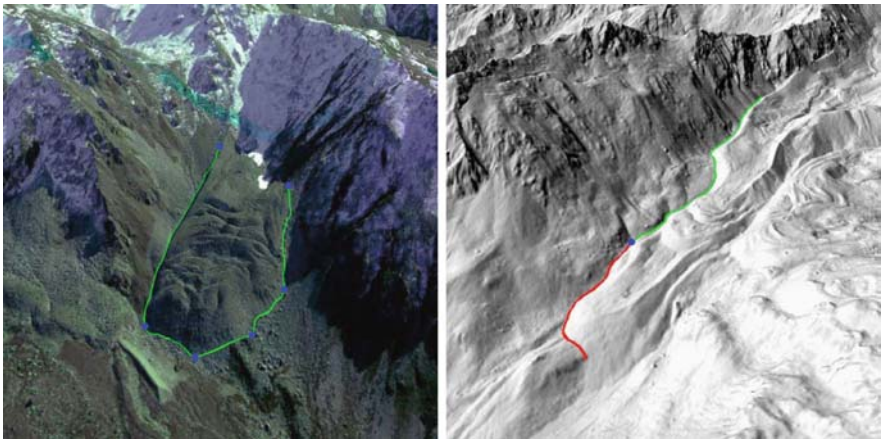

Fig. 6 Landform mapping based on intelligent scissors. The *left image*shows the mapping of a block glacier using the aerial photography as user input. The *blue points*are the positions where the user placed a seed point. On the *right image*mapping is performed on a shaded relief representation of the terrain. The *red part*of the boundary is currently active and changes depending on mouse movents while *green part*is already fixed

Fig. 7 Landform mapping based on a graph cut optimization. The *left*image shows the initial user input marking foreground and background regions. The *right*image shows the segmentation result automatically computed from the user input

In general boundary identification on aerial photography is influenced to a great extent by the lighting conditions during the acquisition process. Therefore mapping results might be biased, especially by shadows, causing the segmentation algorithm in the worst case to follow shadow boundaries instead of true object boundaries. One approach to handle these cases is to consider the elevation data in the boundary estimation procedure in addition to the aerial photography. We therefore allow the user not only to perform the mapping based on the aerial photography but also on the shaded relief or slope. The framework can easily be extendend to work on other data derived from the aerial imagery or elevation data and used as input for the segmentation.

In order to asses the mapping at steep slopes we first parameterized the terrain surface and then matched several photos with the data set. An additional difficulty here was that the used photos were acquired during summer when snow was only present at the highest mountains, while in the data set snow is also present in lower-lying areas. This presents a challenge for the histogram matching and blending because images acquired under completely different conditions have to be combined. The augmentation of the data with the additional photos drastically enhanced its visual quality and information content (see Fig. 4). Moreover, photos can not only increase resolution but contain information of parts of the surface that might be hidden or not clearly visible in the aerial photography due to snow or shadows, for example. Especially the in the aerial photography only sparsely sampled steep slopes benefit from the additional information and their resolution can be increased by orders of magnitude, which is a requirement for a detailed landform mapping in these areas. Figure 8 shows mapping results at steep slopes using information from matched photos.

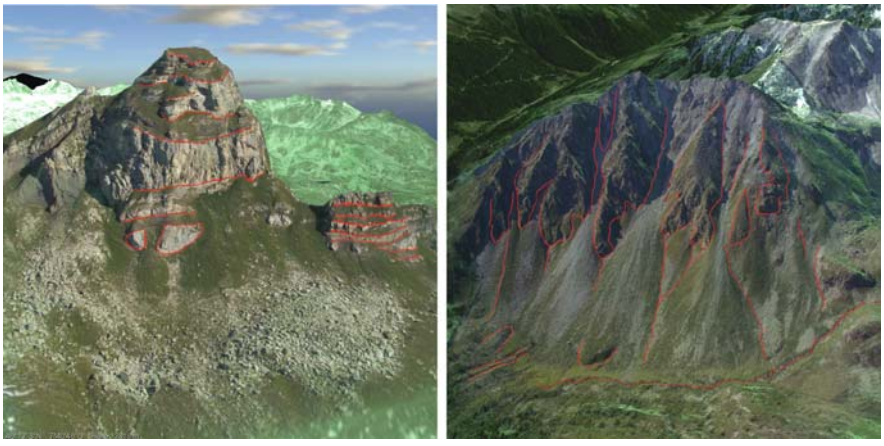

Fig. 8 Mapping results using previously matched high resolution photos

References

Boykov, Y., Jolly, M.P.: Interactive graph cuts for optimal boundary & region segmentation of objects in n-d images. Proceedings of ICCV 1(2001) 105–112

Boykov, Y., Kolmogorov, V.: An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. Energy Minimization Methods in Computer Vision and Pattern Recognition (2001) 359–374

Cooke, R.U., Doornkamp, J.C.: Geomorphology in environmental management. Clarendon Press, Oxford (1974) 352–397

Degener, P., Meseth, J., Klein, R.: An adaptable surface parametrization method. The 12th International Meshing Roundtable (2003)

Falcâo, A.X., Udupa, J.K.: A 3d generalization of the user-steered live-wire image segmentation. Medical Image Analysis 4(2000) 389–402

Falcâo, A.X., Udupa., J.K., Samarasekera, S., Sharma, S., Hirsh B.E., Lotufo, R.: User-steered image segmentation paradigms: live wire and live lane. Graphical Models and Image Processing 60(1998) 233–260

Floater, M.S., Hormann, K.: Advances in multiresolution for geometric modelling, Springer Verlag (2005) 157–186

Geman, S., Geman, D.: Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. IEEE Transactions on Pattern Analysis and Machine Intelligence 6(1984) 721–741

Hartley, R.I., Zisserman, A.: Multiple View Geometry in Computer Vision. Second edn. Cambridge University Press, ISBN: 0521540518 (2004)

Juan, O., Boykov, Y.: Active graph cuts. IEEE conference on Computer Vision and Pattern Recognition (2006)

Kang, H.W., Shin, S.Y.: Enhanced lane: interactive image segmentation by incremental path map construction. Graphical Models 64(5) (2002) 292–303

Li, Y., Sun, J., Tang, C.K., Shum, H.Y.: Lazy snapping. ACM Transaction on Graphics 23(3) (2004)

Lombaert, H., Sun, Y., Grady, L., Xu, C.: A multilevel banded graph cuts method for fast image segmentation. Proceedings of ICCV (2005) 259–265

Mortensen, E.N., Barrett, W.A.: Intelligent scissors for image composition. SIGGRAPH 95 Proceedings (1995) 191–198

Mortensen, E.N., Barrett, W.A.: Toboggan-based intelligent scissors with a four parameter edge model. Proceedings of IEEE Computer Vision and Pattern Recognition 2(1999) 452–458

Otto, J., Dikau, R.: Geomorphologic system analysis of a high mountain valley in the swiss alps. Zeitschrift für Geomorphologie 48(3) (2004) 323–341

Reinhard, E., Ashikhmin, M., Gooch, B., Shirley, P.: Color transfer between images. IEEE Computer Graphics and Applications 21(5) (2001) 34–41

Rother, C., Kolmogorov, V., Blake, A.: Grabcut - interactive foreground extraction using iterated graph cuts. ACM Transaction On Graphics 23(3) (2004) 309–314

Schmidt, J., Hewitt, A.: Fuzzy land element classification from dtms based on geometry and terrain position. Geoderma 121(3–4) (2004) 243–256

Schneevoigt, N., Schrott, L.: Fernerkundungsbasierte reliefformenerkennung im hochgebirge (reintal, bayerische alpen). Geographica Helvetica 3(2006)

Schneider, M., Guthe, M., Klein, R.: Real-time rendering of complex vector data on 3d terrain models. In: Proceedings of The 11th International Conference on Virtual Systems and Multimedia. (2005) 573–582

Schneider, M., Klein, R.: A multilevel banded intelligent scissors method for fast segmentation in large virtual terrains. II International Conference Remote Sensing Archaeology (2006)

Schneider, M., Klein, R.: Efficient and accurate rendering of vector data on virtual landscapes. Journal of WSCG 15(1–3) (2007) 59–64

Schneider, M., Klein, R.: Semi-automatic landform mapping at steep slopes. In: Proceedings of Joint Workshop ‘Visualization and Exploration of Geospatial Data’. (2007)

Segal, M., Korobkin, C., van Widenfelt, R., Foran, J., Haeberli, P.: Fast shadows and lighting effects using texture mapping. Proceedings of the 19th annual conference on Computer graphics and interactive techniques (1992) 249–252

Seijmonsbergen, A., de Graaff, L.: Geomorphological mapping and geophysical profiling for the evaluation of natural hazards in an alpine catchment. Natural Hazards and Earth System Science 6(2) (2006) 185–193

van Asselen, S., Seijmonsbergen, A.: Expert-driven semi-automated geomorphological mapping for a mountainous area using a laser dtm. Geomorphology 78(3–4) (2006) 309–320

Wahl, R., Massing, M., Degener, P., Guthe, M., Klein, R.: Scalable compression of textured terrain data. Journal of WSCG 12(3) (2004) 521–528

Williams, L.: Casting curved shadows on curved surfaces. In SIGGRAPH 78 (1978) 270–274

Wong, K.C., Heng, P.A., Wong, T.T.: Accelerating intelligent scissors using slimmed graphs. Journal of Graphic Tools 5(2) (2000) 1–13

A Perona-Malik Type Method in Shape Generalization of Digital Elevation Models

Carsten Ebmeyer and Jens Vogelgesang

Abstract During the last century, the amount and the complexity of land surface data have enormously increased and have shown the need for efficient automated analysis methods. In this paper we focus our interest on methods helping to analyze Digital Elevation Models. A finite element method in shape generalization of Digital Elevation Models is presented and numerical results are given. The finite element scheme is a fully discrete approximation of a diffusion equation of forward-backward Perona-Malik type. C^0-piecewise linear elements in space and the backward Euler difference scheme in time are used.

Keywords Data representation · Geoscientific modeling · Nonlinear diffusion methods · Scales · Finite element approximation

1 Introduction

Today landform data are mostly collected and stored in digital form. The most common form of digital representations of topographic surfaces are Digital Elevation Models (DEMs) consisting of points of elevation, sampled systematically at equally spaced grids. In recent years, however, the data amount collected and stored in a single DEM has enormously grown. This new situation makes the land surface analysis by hand challenging. One major facilitation to this problem is the automated preprocessing step of shape generalization. By shape generalization we mean the process of simplifying the elevation representation in conjunction with a preservation of landform characteristics.This step makes the DEM admissible to the automated extraction of specific landform information, like the slope, from a digital elevation representation of a surface. In this section we introduce a diffusion method originated in image processing and also used in shape generalization.

C. Ebmeyer (✉)
Mathematical Seminar, Meteorological Institute, Nußallee 15, D-53115 Bonn, Germany
e-mail: ebmeyer.c@uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_4,

The basic concept of using diffusion models for DEM analysis is that of a selective smoothing directed along a synthetic time axis, the scale space axis. Let us introduce the concept more precisely. Given a DEM u_0 we are asking for a selective simplified DEM u_1, containing only those details we are interested in. The trouble with this step is, that the surface characteristics we are interested in lie on different length scales. For instance, one is not interested in the roughness information caused by small stones, but the elevation changes at sharp cliffs of that length scale. Now the idea is to process the DEM in a continuous way, instead of directly moving from u_0 to u_1. One of the major benefits of this concept is the freedom of choosing the appropriate detail scale. This concept leads to an evolution problem and is well known and successfully used in image processing where the elevation points correspond to the intensity of the points of a gray valued image.

Let us consider the following parabolic partial differential equation,

$$u_t - \operatorname{div}(g(|\nabla u|)\,\nabla u) = 0, \tag{1}$$

where $\Omega \subset \mathbb{R}^2$ is a polygonal domain containing the grid points, and the initial value $u_0 : \Omega \to \mathbb{R}$ is a suitable approximation of the considered DEM. The initial value problem (1) has to be completed by some boundary conditions, and no-flux conditions being the most appropriate choice. Thus, $u_0(x) := u(x, 0)$ may be seen as a representation of the DEM. Calculating the solution $u(x, t)$ at different time points t_1,t_2, ... we obtain a sequence $u(x, t_1)$,$u(x, t_2)$, ... of new DEMs, which may be seen as generalizations of the original DEM represented by u_0.

The diffusion coefficient $g(\,\cdot\,)$ is designed to be very small near sharp elevation changes, that is, at points where $|\nabla u|$, i.e. the modulus of the spatial gradient, is large. Whereas, at points where $|\nabla u|$ is small, the diffusion coefficient $g(\,\cdot\,)$ is large and, therefore, strengthens the diffusion at points where the elevation u varies only slowly in space. Thus it is expected that small disturbances, represented by small values of $|\nabla u|$, are smoothed out and that sharp elevation changes, which are represented by a large modulus of the spatial gradient, are preserved.

A good candidate for this purpose is a function inversely related on the modulus, e.g.,

$$u_t - \operatorname{div}\left(\frac{\nabla u}{|\nabla u|}\right) = 0, \tag{2}$$

where $g = g_{TV}(|\nabla u|) := \frac{1}{|\nabla u|}$. This equation is called TV-flow model. It was proposed by Osher et al. (2003) and is frequently used in image processing; see Alter et al. (2005), Andreu et al. (2001, 2004), Ballettini et al. (2002), Feng and prohl (2003), Giga et al. (2004) and Moll (2005). In particular, it has the ability to preserve edges although, at a first glance, the equation has a quite simple structure. However, this equation has a drawback for our purpose, since it has a strong convexification effect. That is, convex objects will evolve and gaps between adjacent objects will fill up; cf. Bellettini et al. (2002), where some explicit examples are given. To give a more intuitive understanding of this convexification effect consider Fig. 1. Two snapshots of the evolution of three initially separated cylinders under the TV-flow model are

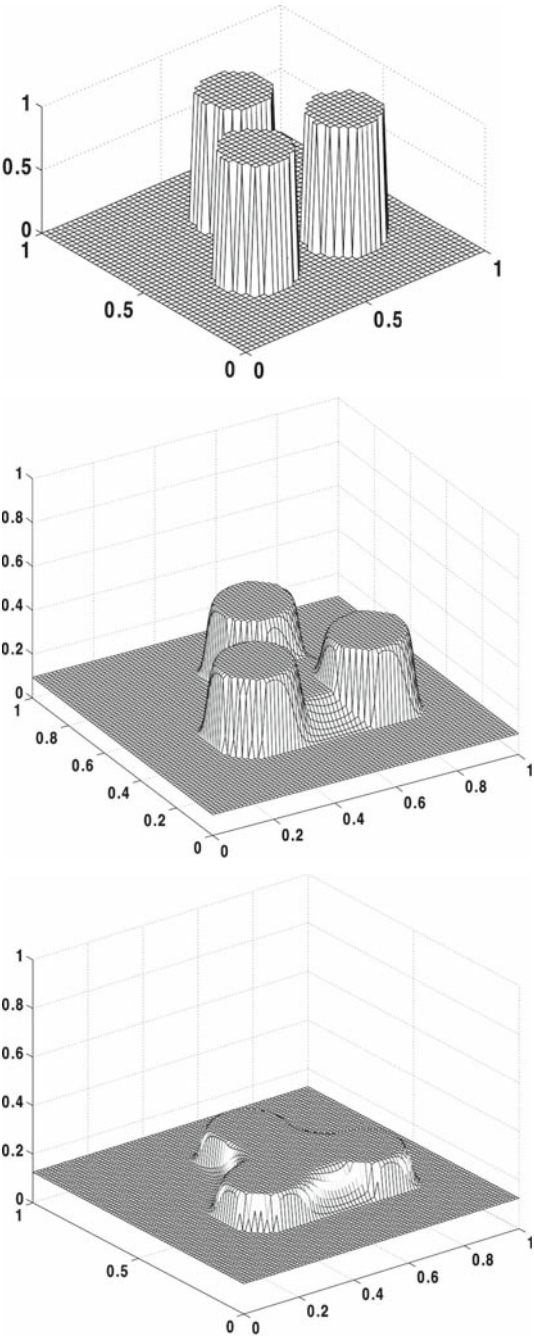

Fig. 1 Convexification effect. Displayed are snapshots of the TV-flow at the time points t=0, t=0.5, and t=1. Starting with an artificial image of three separated cylinders, the three objects fuse into one object. This final object has a convex support. The used finite element parameters are h=0.01 and τ=0.001

shown. In time the cylinders fuse together to one object as shown in the figure. The crucial result is that the final object has a convex support.

Further, let us note that under homogeneous Neumann boundary value conditions the TV-flow has a finite time extinction, which means that there exists a time t_{ex} where $\nabla u(t,x) = 0$ for all $x \in \Omega$ and all $t > t_{ex}$.

To overcome this shortcome of the TV-flow, Perona and Malik (1990) proposed a function $g = g_{PM}$ depending inversely on the squared modulus of the gradient, i.e.

$$g_{PM}(|\nabla u|) := \frac{1}{1 + |\nabla u|^2}.$$

Unfortunately, the diffusion problem with the diffusion coefficient g_{PM} is mathematically ill posed. The main problem is the forward-backward diffusion character of the equation. Backward diffusion is known to be highly unstable and even smooth initial conditions can lead to singularities after arbitrarily short times. Nevertheless, the numerical implementation of the diffusion equation yields striking results; cf. Aubert and Kornprobst (2002), Bänsch and Mikula (2001), Braunmandel et al. (2003), Esedoglu (2001), Kacur and Mikula (1995), Kawohl and Kuter (1998) and Perona and Malik (1990).

The good numerical results motivated different approaches to modify the equation (2). These modifications were done in a way such that they essentially preserve the numerical results and are accessible to rigorous foundations in terms of mathematical existence and uniqueness theorems. One of the most widely used replacement approaches was introduced by Catte et al. (1992). They replace the image information ∇u in the nonlinearity g_{PM} by a space regularized form. More precisely, ∇u is replaced by $\nabla G_\sigma \star \bar{u}$ where σ is the variance of the Gaussian kernel G_σ and $\bar{u}$ is the appropriate supplemented form of u extended over R^2.

In this paper we consider the forward-backward diffusion model introduced recently in Ebmeyer and Vogelgesang (2008). It enables us to preserve the essential features of the DEM and is accessible to the mathematical framework. In fact, our equation has the major benefit of being able to identify the morphology of features on the surface such as the slope, the shape, and the size. In detail, we are concerned with the nonlinear diffusion equation

$$u_t - \operatorname{div}\left(\frac{\lambda + \mu\,|\nabla u|}{\lambda + |\nabla u|^2}\nabla u\right) = 0 \tag{3}$$

completed by homogeneous Neumann boundary value conditions and the initial value u_0 representing the DEM.

Note that the limit case $\lambda = 1$ and $\mu = 0$ is the Perona-Malik equation. Thus, for $\mu > 0$ being sufficiently small equation (3) may be seen as a stabilized Perona-Malik model.

Moreover, let us note that the condition $\mu > 0$ implies that the diffusion model (3) satisfies a p-growth condition for p=1. That is, (3) is a model with linear growth, such as the TV-flow model. Whereas in the case of $\mu = 0$ the model would have

constant growth, corresponding to the case of p=0, such as the Perona-Malik model, and would be mathematically ill-posed.

The parameters $\lambda \in (0,1)$ and $\mu \in (0,1)$ have to be chosen according to the application. In fact, the choice of λ depends on the slope of the DEM and $\mu > 0$ has to be sufficiently small. To describe in detail, let us assume that we can find a regular solution $u(\cdot,t)$ to problem (3) at some time point $t > 0$ that provides a smooth level set $\Gamma(t)$. Let us introduce the notation $\xi = \xi(t)$ for the tangential direction along the level set $\Gamma(t)$ and $\eta = \eta(t)$ for the normal direction. Now we are able to rewrite equation (3) in the form

$$u_t - g_{EV}(|\nabla u|)\partial_\xi^2 u + b(|\nabla u|)\partial_\eta^2 u = 0,$$

where b and g_{EV} are given by

$$b(s) := g_{EV}(s) - 2sg'_{EV}(s) \text{ and } g_{EV}(s) := \frac{\lambda + \mu s}{\lambda + s^2}.$$

The key characteristic of the equation is that the second term b switches its sign and gives rise to a backward diffusion in the normal direction. Hence, we have forward diffusion along $\Gamma(t)$ and forward-backward diffusion across $\Gamma(t)$. In fact, only edges such that $|\nabla u| < \mu + \sqrt{\mu^2 + \lambda}$ are smoothed out. As the diffusion in η direction differs from that one in ξ direction we speak of anisotropic diffusion.

Thus, the choice of λ depends on the characteristic slope of the considered territory, and μ has to be much smaller than λ. A good choice could be, say, $\mu = \frac{1}{10}\lambda$.

2 Continuous Problem and the FEM Method

The aim of this section is to formulate the continuous and the fully discrete problem of the nonlinear equation (3) in a mathematical rigorous sense. After summarizing the results for the continuous solution we shall develop a fully discrete scheme using the finite element method.

2.1 The Continuous Problem

We now discuss equation (3). We adopt the standard notation, $L^q(0,T;L^p(\Omega))$ and $L^s(0,T;W^{q,p}(\Omega))$ denote the usual Lebesgue and Sobolev spaces, and $BV(\Omega)$ is the space of bounded variation. Further, $\Omega \subset \mathbb{R}^2$ is a polygonal domain and T>0.

Let η be the outward unit normal of $\partial\Omega$ and $\mu,\lambda > 0$ be two constants. For functions $u:\Omega \times [0,T] \to \mathbb{R}$ we consider the initial and boundary value problem

$$\begin{aligned} u_t - div\left(\frac{\lambda + \mu|\nabla u|}{\lambda + |\nabla u|^2}\nabla u\right) &= 0 \quad \text{in } \Omega \times (0,T], \\ \frac{\lambda + u|\nabla u|}{\lambda + |\nabla u|^2}\nabla u \cdot \eta &= 0 \quad \text{on } \partial\Omega \times (0,T], \\ u(\cdot,0) &= 0 \quad \text{in } \Omega. \end{aligned} \tag{4}$$

In the following we assume that $u_0 \in L^\infty(\Omega) \cap BV(\Omega)$. Moreover, for simplicity, we suppose that $\sup_{x\in\Omega} u_0(x) = 1$ and $\inf_{x\in\Omega} u_0(x) = 0$. This may be obtained by shifting our elevation profile, such that the lowest point is equal zero, and normalizing it.

A lot of interesting facts are known about problem (4); cf. Ebmeyer and vogelgesang (2008). One of the most striking effects is that the equation preserves sharp elevation changes and landform characteristics. However, there is still a need to characterize the singular set completely and little is known about the case of $\mu = 0$.

Remark 1. For $\mu > 0$ theinitial and boundary value problem (4) has a solution in the sense of Young measures. More precisely, a pair (u, ν) is called a measure solution in the sense of Young measures if $u \in L^\infty(0,T;BV(\Omega)) \cap L^\infty(0,T;L^\infty(\Omega))$, $u_t \in L^2(0,T;L^2(\Omega))$, and $\nu = (\nu_{x,t})_{(x,t)\in\Omega\times[0,T]}$ is a parametrized family of probability measures on R^2 such that

$$\int_0^T \int_\Omega \Big(\Big\langle \frac{\lambda + \mu|\gamma|}{\lambda + |\gamma|^2}\, \gamma, \nu_{x,t}(\gamma) \Big\rangle \cdot \nabla\phi + u_t\, \phi \Big) dx\, dt = 0 \qquad \forall \phi \in C^\infty(\Omega \times (0,T)),$$

$$\text{where } \Big\langle \tfrac{\lambda+\mu|\lambda|}{\lambda+|\gamma|^2}\, \gamma, \nu_{x,t}\,(\gamma) \Big\rangle = \textstyle\int_{R^2} \frac{\lambda+\mu|\gamma|}{\lambda+|\gamma|^2}\, \gamma\, d\nu_{x,t}\,(\gamma)\,,$$

$$\nabla u = \big\langle \gamma, \nu_{x,t}\,(\gamma) \big\rangle \qquad \text{a.e. in } \Omega \times (0,T]\,,$$

and $u(\cdot\,,0) = u_0$ in Ω.

The existence of Young measure valued solutions is shown in Yin and Wang (2003) for forward-backward type equations with linear growth under a homogeneous Dirichlet boundary value condition. Existence of Young measure valued solutions for forward-backward type equations with quadratic growth can be found in De moulini (1996) and Kinderlehrer and Pedregal (1992).

For sufficiently smooth solutions (u,ν) like $u \in L^1(0,T;W^{1,1}(\Omega))$ and ν being a Dirac measure it follows that

$$\int_\Omega \Big(\frac{\lambda + \mu|\nabla u|}{\lambda + |\nabla u|^2}\, \nabla u \cdot \nabla\phi + u_t\, \phi \Big) dx = 0 \qquad \text{a.e. } t > 0$$

for all $\phi \in W^{1,1}(\Omega) \cap L^2(\Omega)$. Below we shall discretize this weak formulation of the problem using finite elements.

2.2 The Finite Element Method

Before studying the finite element approximation of the problem, we introduce the finite element spaces. Let (T_h) be a regular triangulation of Ω into disjoint open regular triangles K, so that $\overline{\Omega} = \bigcup_{K\in(T_h)} \bar{K}$. Regular means that each element has at most one edge on $\partial\Omega$, and each pair of triangles $\bar{K}, \bar{K}' \in (T_h)$ has either only one common vertex, or a whole edge, or $\bar{K}$ and $\bar{K}'$ are disjoint. Let $\mathrm{h_K}$ denote the diameter of the element K in (T_h) and

$$h = \max_{K\in(T_h)} h_K$$

the mesh size. We assume that there is a regularity constant C of (T_h), independent of h, such that $1 \leq \max_{K\in(T_h)} (h_K/\rho_K) \leq C$, where ρ_K denotes the diameter of the largest ball contained in K. That means, there is a lower bound of the inner angles of the triangles. Furthermore, we assume that adjacent triangles are similar in size, that is, there is a $C^1(\overline{\Omega})$–function $h(x)$ such that

$$c' h_K \leq h(x) \leq h_K \qquad \text{in } K$$

for all simplices $K \in (T_h)$ and some constant $c' > 0$ independent of K.

Associated with (T_h) is a finite dimensional subspace S_h of $C^0(\bar{\Omega})$, such that $\chi|_K \in \mathcal{P}_1$ for all $\chi \in S_h$ and $K \in (T_h)$, where $\mathcal{P}_1$ is the linear functions space. We define

$$S_h := \{v \in C^0(\Omega) : v_{|K} \in \mathcal{P}_1 \text{ for all } K \in (T_h)\}.$$

Remark 2. If Ω is a rectangular domain containing an equally spaced grid of elevation points the mesh may be generated in a quite simple way choosing the grid points as the vertices of the triangles. More general, a triangulation of a polygonal domain could be generated with the help of more sophisticated mesh generators, like the Delaunay algorithm; see, e.g., Persson and Strang (2004).

Now we state the fully discrete finite element scheme. In time we discretize the equation using the backward Euler scheme. Let $N > 0$ be an integer and we define the size of each time step by

$$\tau := N^{-1} T.$$

Since it is of advantage to use a lumping mass technique we define

$$(f,g)_h := \int_\Omega \Pi_h(fg)\, dx.$$

.

Here, $\Pi_h v \in S_h$ is the C^0 -piecewise linear interpolant of the function v, that is, $\Pi_h v(y) = v(y)$ for all nodes y of the triangulation (T_h). Further, let us introduce the notation $(f,g) := \int_\Omega fg\, dx$.

The backward difference scheme consists of finding functions $U^n \in S_h$, $n \in \{1,\ldots,N\}$, which are solutions of the equations

$$\left(\frac{U^n - U^{n-1}}{\tau},\phi\right)_h + \left(\frac{\lambda + \mu|\nabla U^n|}{\lambda + |\nabla U^n|^2}\nabla U^n,\nabla\phi\right) = 0 \qquad \forall\phi \in S_h\,, \tag{5}$$

where $U^0 \in S_h$ is a given approximation of the initial DEM u_0.

Let us note that U^n may be represented as

$$U^n = \sum_{i=1}^{m} \alpha_i^n \psi_i(x)\,,$$

where $\{\psi_1, \ldots, \psi_m\}$ are some basis functions of S_h . Hence, the backward difference scheme provides a system of algebraic equations, as discussed in the next section.

To conclude this section we summarize some results of Ebmeyer and Vogelgesang (2008). The first result provides the existence of the finite element solutions and the crucial energy estimate.

Theorem 1. There exist solutions $U^n \in S_h$ for $1 \leq n \leq N$ of the equations *(5)* satisfying

$$\sup_{1\leq n\leq m} \left(U^n, U^n\right)_h + \mu\tau \sum_{n=1}^{N} || \nabla U^n ||_{L^1(\Omega)} \leq c$$

for a constant c independent of h, τ, λ and μ.

Next, under an additional acuteness assumption on the triangulation there holds the following discrete maximum principle. Let γ_K denote the largest angle of the triangle $K \in (T_h)$ and $\gamma_{\max} := \max_{K\in(T_h)} \gamma_K$.

Theorem 2. If

$$\gamma_{\max} \leq 90^0$$

then each solution U^n ($1 \leq n \leq N$) of the equations (5) satisfies

$$|| U^n ||_{L^\infty(\Omega)} \leq || U^0 ||_{L^\infty(\Omega).}$$

For a proof of these results we refer to Ebmeyer and Vogelge (2008).

3 Algorithm

In this section we will derive an algorithm for solving the equations (5) in an efficient way and state a scheme for actual computations.

3.1 The Scheme for Actual Computations

Let us reformulate the n-th equation of (5) using matrix notation. Let $\{\psi_i\}_{i=0,\ldots,m}$ be the chosen basis functions of S_h. Then U^n can be written in the form

$$U^n = \sum_{i=1}^{m} \alpha_i^n \psi_i(x)$$

and the n-th equation of (5) becomes

$$M\alpha^n + \tau K(\alpha^n)\alpha^n = M\alpha^{n-1}\,, \tag{6}$$

where $M \in \mathbb{R}^{m\times m}$ is the mass matrix with the elements $(M)_{ij} = (\psi_j,\psi_i)_h$. Indeed, M is the lumping mass matrix, since we use the mass lumping technique as mentioned before. Further, $\alpha^n \in \mathbb{R}^m$ is a vector whose elements are $(\alpha_i^n)_{i=1,\ldots,m}$, and $K(\alpha) \in \mathbb{R}^{m\times m}$ is the stiffness matrix depending in a nonlinear way on α, where $(K(\alpha))_{ij} = (g_{EV}(|\nabla U^n|)\nabla\psi_j,\nabla\psi_i)$ and $|\nabla U^n|^2 = \sum_{k=1}^2 (\sum_{i=1}^m \alpha_i^n \partial_k\psi_i)^2$.

Notice that the algebraic system (6) is nonlinear, due to the nonlinearity of the function g_{EV}. To solve these equations we use an iteration method calculating successively solutions of a linearized system.

Moreover, due to measurement errors the data of the DEM are degraded by some noise. We may suppose that the disturbance of the given data is small in comparison to the length scale of the elevations. For instance, there is a measurement error of a few meters, in comparison to a difference in elevation of more than 1000 m. Hence, it is desirable that the minimum and maximum value of our time step sequences of the DEM's $\{U^n\}_{n=0,\ldots,N}$ will stay constant. Therefore, we will use a post-processing projection step in order to normalize the finite element solutions such that their maximal and minimal elevations are identical with those of the initial value. Hence, the final DEM will differ from the original FEM solution by a time dependent scaling factor. Moreover, for simplicity, we will use a pre-processing projection step in order to normalize the initial DEM, such that $\sup_{x\in\Omega} u_0(x) = 1$ and $\inf_{x\in\Omega} u_0(x) = 0$. Notice that λ have to be chosen in accord with this projection step.

Let us now discuss the solution of the nonlinear algebraic systems (6) and the pre- and post-processing steps in more detail. For our purpose we introduce an iteration index ν and use the following fixed-point type iteration in order to solve the nonlinear system. We note that a sequence of such operations may be performed efficiently utilizing the l^2-norm as a stopping condition.

For a given $U^0 \in S_h$, where

$$U^0 = \sum_{i=1}^m \alpha_i^0 \psi_i(x)\,,$$

and given $\epsilon > 0$, we calculate

$$U^n = \sum_{i=1}^m \alpha_i^n \psi_i(x) \qquad \text{for } n = 1,\ldots,N$$

by successively performing the steps:

(1) Pre-processing step: Normalize the initial DEM, such that $\sup_{x\in\Omega} u_0(x) = 1$ and $\inf_{x\in\Omega} u_0(x) = 0$.
(2) Determine U^0 , and define $\bar{\alpha}^0 := \alpha^0$.
Set $n = 1$.
(3) Set $b := M\bar{\alpha}^{n-1}$ and $\beta_0 := \bar{\alpha}^{n-1}$.
(4) For $\nu = 1,2,\ldots$ solve

$$M\beta_\nu + \tau K(\beta_{\nu 1})\,\beta_\nu = b \tag{7}$$

until $|\beta_\nu - \beta_{\nu-1}|_{l_2} < \epsilon$.

(5) Set $\bar{\alpha}^n := \beta_\nu$.
If $n < N$ increase n by one and go back to step (3).

(6) Post-processing step: For every component $(\bar{\alpha}_i^n)$ of the final vectors $\bar{\alpha}^n := (\bar{\alpha}_1^n, \bar{\alpha}_2^n, \ldots, \bar{\alpha}_m^n)^T$ of step (5) use the normalization

$$(\alpha_i^n) = \frac{(\bar{\alpha}_i^n) - \min_i(\bar{\alpha}_i^n)}{\max_i\left(\bar{\alpha}_i^n - \min_i(\bar{\alpha}_i^n)\right)}\,. \tag{8}$$

(7) Reverse the normalization step (1).

Remark 3. Given the basis functions $\{\psi_i\}_{i=0,\ldots,m}$, the algorithm results in a sequence of vectors α^{n} , $n = 1, \ldots, N$, corresponding to the solutions $\{U^n\}_{n=1}^N$ of problem (5). Notice that the algorithm consists of two iteration sequences, an outer iteration sequence with index n and an inner sequence with index ν. For each time step t_n we solve the resulting nonlinear algebraic system (7) by the inner iteration sequence while freezing the right hand side b.

Remark 4. Clearly, we may skip the normalization step (1). Let us note that this must be taken into account for the choices of λ and μ.

3.2 The Convergence of the Fix-Point Scheme

Now we shall discuss the convergence of the inner iteration scheme of step (4) in more detail and show its convergence, due to Banachs′ fixed-point theorem (cf. Zeidler, 1993).

Reformulating equation (7) we obtain the following iterative scheme for actual computations,

$$(U_\nu^n, \chi^n)_h + \tau\,(g_{EV}(|\nabla U_{\nu-1}^n|)\nabla U_\nu^n, \nabla\chi^n) = (U^{n-1}, \chi^n)_h \qquad \forall \chi^n \in S_h, \tag{9}$$

successively for $n = 1, \ldots, N$, where

$$g_{EV}(|\nabla U_{\nu-1}^n|) = \frac{\lambda + \mu|\nabla U_{\nu-1}^n|}{\lambda + |\nabla U_{\nu-1}^n|^2}\,.$$

Now let n be fixed. Then U_ν^n admits the representation

$$U_\nu^n = \sum_{i=1}^{m} \beta_\nu^i \psi_i(x)\,,$$

where ψ_{i}, $1 \le i \le m$, are the basis functions of S_{h}. For each given function $U^{\mathrm{n\text{-}1}}$ we therefore obtain a sequence of functions $(U_\nu^n)_\nu \in S_h$. This sequence converges and it holds that

$$\lim_{\nu\to\infty} U_\nu^n = U^n.$$

Let us discuss the convergence of $U_\nu^n(\nu \to \infty)$ in more detail. Let β_ν be the vector with the elements $\beta_\nu^i, i = 1,\ldots,m$. Further, let A_ν be the matrix whose elements are defined by

$$(A_\nu)_{ij} = (\psi_j,\psi_i)_h + \tau\,(g_{EV}(|\nabla U_\nu^n|)\nabla\psi_j,\nabla\psi_i).$$

Hence, in matrix notation, (9) may be expressed as

$$A_{\nu-1}\beta_\nu = b$$

for some vector $b \in \mathbb{R}^m$. We have

$$\beta_{\nu+2} - \beta_{\nu+1} = (A_{\nu+1}^{-1} - A_\nu^{-1})\,b = A_{\nu+1}^{-1}(A_\nu - A_{\nu+1})A_\nu^{-1}b\,.$$

Let K_ν be the matrix with the elements $(K_\nu)_{ij} = (g_{EV}(|\nabla U_\nu^n|)\nabla\psi_j,\nabla\psi_i)$. We estimate

$$\begin{aligned}\left\|\beta_{\nu+2} - \beta_{\nu+1}\right\|_\infty &= \left\|A_{\nu+1}^{-1}\,(\tau\,K_\nu - \tau\,K_{\nu+1})\,A_\nu^{-1}\,b\right\|_\infty \\ &\le \left\|A_{\nu+1}^{-1}\right\|\,\tau\,\|K_\nu - K_{\nu+1}\|\,\|A_\nu^{-1}\|\,\|b\|_\infty\,,\end{aligned}$$

where $\|\cdot\|$ is the row-sum norm. Notice that $|g_{EV}(|s|)| \le c(\frac{\mu}{\sqrt{\lambda}} + 1)$. Thus, $\left\|A_{\nu+1}^{-1}\right\|$ is uniformly bounded in ν, if τ is sufficiently small. Further, utilizing the fact that $|g_{EV}(|s|) - g_{EV}(|\bar{s}|)| \le c_\lambda\,(\mu+1)\,|s-\bar{s}|$ for a constant c_λ depending only on λ we get

$$\left\|K_{\nu+1} - K_\nu\right\| \le c_h\,\left\|\beta_{\nu+1} - \beta_\nu\right\|_\infty.$$

Choosing τ sufficiently small we obtain a constant $\bar{c} < 1$ independent of ν (and μ) such that

$$\|\beta_{\nu+2} - \beta_{\nu+1}\|_\infty \le \bar{c}\,\left\|\beta_{\nu+1} - \beta_\nu\right\|_\infty.$$

Hence, due to Banachs′ fixed-point theorem, the sequence $(\beta^\nu)_\nu$ convergences.

4 Numerical Experiments

We present some practical performance of the algorithm introduced in the previous section. Figure 4 displays a DEM from a New Zealand section and two snapshots taken at different time points. The DEM was supplied by the National Institute of Water and Atmospheric Research of New Zealand. The implementation of the algorithm was performed in MATLAB and the linear system was solved using the backslash operator in MATLAB.

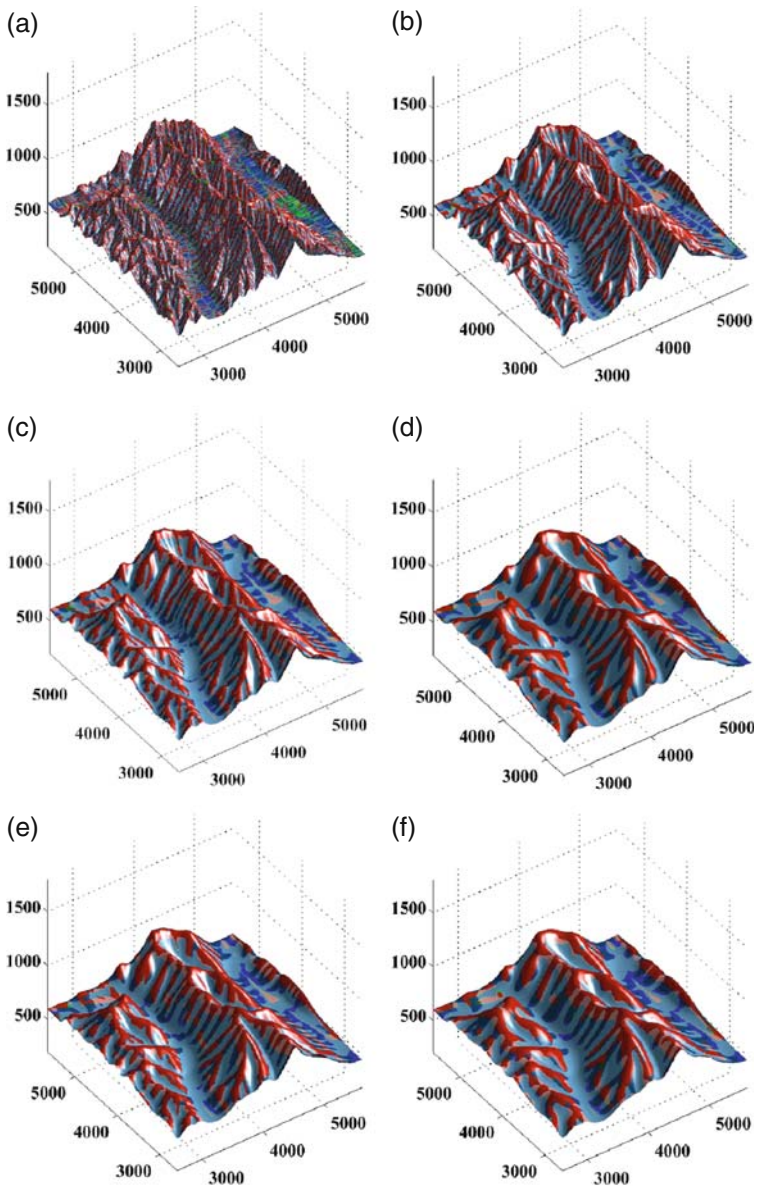

Fig. 2. Evolution of the EV-flow. Snapshots of a DEM from a New Zealand section are displayed at the time points $t = 0$, t=0.01, ..., $t = 0.05$. The used parameters are h=*0.01*, $\tau = 0.001$, $\lambda = 1$, and $\mu = 0.01$

References

F. Alter, V. Caselles and A. Chambolle, *Evolution of characteristic functions of convex sets in the plane by the minimizing total variation flow,* Interfaces Free Bound. 7 (2005), 29–53.

F. Andreu, C. Ballester, V. Caselles and J.M. Mazón, *Minimizing total variation flow*, Differ. Integral Equ. 14 (2001), 321–360.

F. Andreu-Vaillo, V. Caselles and J.M. Mazón, *Parabolic quasilinear equations minimizing linear growth functionals*, Prog. in Math 223, Birkhäuser, Basel, 2004.

G. Aubert and P. Kornprobst, *Mathematical problems in image processing, Partial differential equations and the calculus of variations*, Applied Mathematical Sciences 147, Springer, New York, 2002.

E. Bänsch and K. Mikula, *Adaptivity in 3D image processing*, Comput. Vis. Sci. 4 (2001), 21–30.

G. Bellettini, V. Caselles and M. Novaga, *The total variation flow in* R^N J. Differ. Equations 184 (2002), 475–525.

A. Braunmandel, T. Canarius and H.-P. Helfrich, *Diffusion methods for form generalisation* in Dynamics of Multiscale Earth Systems (eds: H. Neugebauer, C. Simmer), Springer Verlag New York, 2003, pp. 89–101.

F. Catté, P.-L. Lions, J.-M. Morel and T. Coll, *Image selective smoothing and edge detection by nonlinear diffusion*, SIAM J. Numer. Anal. 29 (1992), 182–193.

S. Demoulini, *Young measure solutions for a nonlinear parabolic equation of forward-backward type*, SIAM J. Math. Anal. 27 (1996), 376–403.

C. Ebmeyer and J. Vogelgesang, *Finite element approximation of a forward and backward anisotropic diffusion model in image denoising and form generalization*, Numer. Methods Partial Differ. Equations 24 (2008), 646–662.

S. Esedoglu, *An analysis of the Perona-Malik scheme* Commun. Pure Appl. Math. 54 (2001), 1442–1487.

X. Feng and A. Prohl, *Analysis of total variation flow and its finite element approximations*, Math. Model. Numer. Anal. 37 (2003), 533–556.

Y. Giga, Y. Kashima and N. Yamazaki, *Local solvability of a constrained gradient system of total variation*, Abstr. Appl. Anal. 2004 (8) (2004), 651–682. doi: 10.1155/S1085337504311048.

J. Kacur and K. Mikula, *Solution of nonlinear diffusion appearing in image smoothing and edge detection*, Appl. Numer. Math. 17 (1995), 47–59.

B. Kawohl and N. Kutev, *Maximum and comparison principle for one-dimensional anisotropic diffusion*, Math. Ann. 311 (1998), 107–123.

D. Kinderlehrer and P. Pedregal, *Weak convergence of integrands and the Young measure representation*, SIAM J. Math. Anal. 23 (1992), 1–19.

J.S. Moll, *The anisotropic total variation flow*, Math. Ann. 332 (2005), 177–218.

S. Osher, A. Solé and L. Vese, *Image decomposition and restoration using total variation minimization and the* H^{-1} *norm*, Multiscale Model. Simul. 1 (2003), 349–370.

P. Perona and J. Malik, *Scale-space and edge detection using anisotropic diffusion*, IEEE Trans. Pattern Anal. Mach. Intell. 12 (1990), 629–639.

P.-O. Persson and G. Strang, *A simple mesh generator in MATLAB*, SIAM Review 46(2) (2004), 329–345.

J. Yin and C. Wang, *Young measure solutions of a class of forward–backward diffusion equations*, J. Math. Anal. Appl. 279 (2003), 659–683.

E. Zeidler, *Nonlinear functional analysis and its applications, Volume I: Fixed-point theorems* Springer-Verlag, New York, 1993.

The Role of Landscape Processes within the Climate System

Roger A. Pielke Sr. and Dev Niyogi

1 Introduction

Land-surface processes form a dynamic boundary interface within the Earth system (Fig. 1). The multiscale impacts of land-surface processes in modifying regional weather and climate is noted from both the analysis of observations, as well as systematic experiments involving nonlinear, coupled modeling systems. Landscape processes and their interactions with the atmosphere are critical at different micro, regional, and global scales for weather, hydrological and other broad range environmental modeling studies (Alpert et al. 2006).

The land-surface characteristics determine the surface energy partitioning by assigning the distribution of incoming solar radiative energy (insolation) into sensible, latent, and ground heat fluxes. The change in the surface radiative energy affects regional- and larger-scale moisture and temperature. Modifications in the surface fluxes and the thermodynamic parameters lead to changes in regional wind fields and localized circulation patterns. The changes in wind and regional thermodynamic variables alter convective potential and interact with large-scale processes to affect the amount and distribution of clouds and rainfall (Pielke et al. 2007). At a larger scale, the systematic transformation of the land surface can alter regional flow patterns associated with developing persistent zones of moisture convergence, and localized pockets that lead to long-term regional warming or cooling.

The role of landscape processes within the climate assessments however has been mostly ignored except in terms of how carbon assimilation is affected. As summarized in the National Research Council (2005) report, the role of vegetation and soils is much more than that and includes effects on water and heat storage and their fluxes, as well as effects on a variety of other gases and aerosols. The climate system is an integration of physical, biological, and chemical effects associated with land, atmosphere, ocean, and continental ice interactions. The current IPCC (2007) focus on radiative forcing of well-mixed greenhouse gases is too limiting; a broadening in its perspective is overdue. The current view, unfortunately, does not

R.A. Pielke Sr. (✉)
CIRES, University of Colorado, Stadium 255-16, Boulder, CO 80309, USA
e-mail: pielkesr@cires.colorado.edu

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_5,

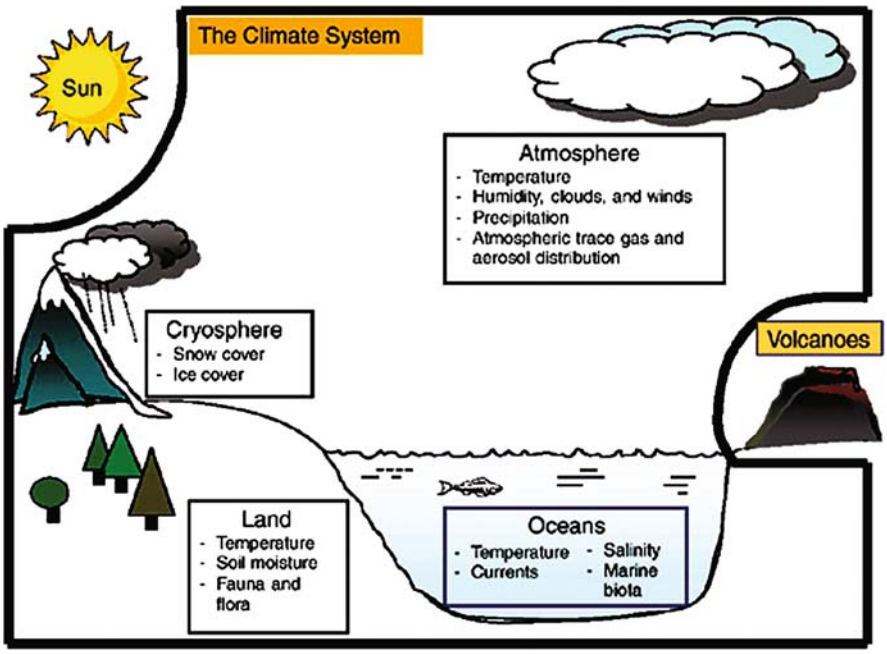

Fig. 1 The climate system, consisting of the atmosphere, oceans, land, and cryosphere. Important state variables for each sphere of the climate system are listed in the boxes. For the purposes of this paper, the Sun, volcanic emissions, and human-caused emissions of greenhouse gases and changes to the land surface are considered external to the climate system
[Source: National Research Council 2005]

properly address the diverse effect of the human disturbance of the climate system. The role of land-surface forcing and feedbacks within the climate system, as one important example of this need for broadening, provides a dynamical feedback that is required if regional-scale climate assessments are to become skillful.

The nonlinear interrelationship between the various surface, boundary layer, regional circulation, cloud and moisture, and biogeochemical factors hinders an accurate assessment of the role of land surface on regional and global climate. Hence, the need for integrated assessments is becoming more important in order to understand the uncertainty and variability of the climate system (Pielke et al. 2002, Marland et al. 2003). We show in this paper that land-surface processes, as part of the climate system, need to be considered not only for developing projections at regional and larger scales, but also for developing vulnerability assessments and mitigation strategies to satisfy the United National Framework Convention on Climate Change (UNFCCC 2007).

Recognizing the need to address the broader role of the land surface and to include nonradiative climate processes, the 2005 National Research Council report had the following priority recommendations:

- Test and improve the ability of climate models to reproduce the observed vertical structure of forcing for a variety of locations and forcing conditions.

- Undertake research to characterize the dependence of climate response on the vertical structure of radiative forcing.
- Report global mean radiative forcing at *both* the surface and the top of the atmosphere in climate change assessments.
- Use climate records to investigate relationships between regional radiative forcing (e.g., land use or aerosol changes) and climate response in the same region, other regions, and globally.
- Quantify and compare climate responses from regional radiative forcings in different climate models and on different timescales (e.g., seasonal, interannual), and report results in climate change assessments.
- Improve understanding and parameterizations of aerosol-cloud thermodynamic interactions and land-atmosphere interactions in climate models in order to quantify the impacts of these nonradiative forcings on both regional and global scales.
- Develop improved land-use and land-cover classifications at high resolution for the past and present, as well as scenarios for the future.
- Encourage policy analysts and integrated assessment modelers to move beyond simple climate models based entirely on global mean top of the atmosphere radiative forcing and incorporate new global and regional radiative and nonradiative forcing metrics as they become available.

Unfortunately, the timeline adopted by the IPCC in initiating the assessments, compile the material available for analysis, and develop a consensus view on the reports did not generally allow for recently published papers and reports to be properly considered. Thus the recommendations from the NRC report received little representation in the 4th assessment. It is not known if this would be modified in the next assessment.

The following sections provide examples of why landscape processes and the human role in altering them, need to be explicitly accounted for in future climate change studies and IPCC assessments.

2 Land Surface Processes

The land-surface feedback leads to an often significant forcing of regional and global climate through changes in the physical properties of the land surface. Traditionally, these changes are attributed to temporal and spatial inhomogeneity in surface albedo and evaporative fraction. The importance of land cover can be extracted from the review of the surface energy budget equations.

$$Q_N = Q_H + Q_{LE} + Q_G, \tag{1}$$

and

$$Q_N = Q_S^{\downarrow}\,(1-A) + (1-\varepsilon)\,Q_{LW}^{\downarrow} - \varepsilon\sigma T_S^4 \tag{2}$$

Here, Q_N is the net radiative flux, Q_H is the turbulent sensible heat flux, Q_{LE} is the turbulent latent heat flux (physical evaporation and transpiration), Q_G is the heat flux into the Earth's surface, $Q_S^{\downarrow}$ is the shortwave global irradiance, A is the surface albedo, $Q_{LW}^{\downarrow}$ is the downward atmospheric longwave radiation, and $\varepsilon\sigma T_S^4$ is the upward surface longwave radiation. Surface albedo, A, (fraction of the incoming shortwave radiative flux to that is reflected by the surface), varies with the land surface ranging from 5% (e.g., wet soil) to 90% (e.g., fresh snow).

Equation (2) suggests that a decrease in surface albedo under constant net longwave radiation, $(1-\varepsilon)\,Q_{LW}^{\downarrow} - \varepsilon\sigma T_S^4$, increases net radiation, Q_N. Thereby, more energy is available for the sensible and latent radiative heat fluxes, $Q_{LE} + Q_G$.

Considering humidity and moisture processes, over a bare soil, the landscape fluxes are governed by bare soil evaporation and diffusion within the soil layers depleting surface moisture availability. The moisture loss from the land surface then becomes a source for atmospheric humidity. When vegetation is present, the complex role of vegetation canopy and associated processes, such as from transpiration, dominate the energy and water vapor (and carbon) exchange. Additionally, the canopy can also intercept precipitation and part of the water that is intercepted by the leaves can be evaporated back from the leaf surface without reaching the ground.

The hydrologic balance over the vegetated landscape can be described as

$$\mathrm{Pr} = \mathrm{E} + \mathrm{T} + \mathrm{Ro} + \mathrm{I} \tag{3}$$

where Pr is the precipitation, E is bare ground evaporation, T is transpiration, Ro is surface runoff, and I is the interception. Processes such as ground water seepages, runoff, and the changes in the root dynamics are still poorly represented in current landform models and are included in most weather and climate modeling systems.

When a landform becomes vegetated, or has a transformation in its surface characteristics, additional energy partitions need to be considered. These would be related to the energy balance of the different mosaics (patches) of land surface, as well as the energetics that are dictated by the albedo and emissivity of the vegetation canopy. The transmittance, absorption, and reflectance of the vegetation canopy, the air space just above and within the canopy, the heat flux in the surface layer of the land boundary, and the atmospheric turbulent boundary layer each become means of complex energy exchange and storage.

Soil moisture is an important interactive and integrative factor relating the different processes within the land-surface system. Soil moisture regulates the evolution of various surface hydrological and energy balance processes. Changes in soil moisture affect surface albedo (Idso et al. 1975), the evaporative fraction (Kabat et al. 2004), and at the regional scale, the potential for cloud formation and precipitation and evaporation/transpiration recirculation ratio (Brubaker et al. 1993). Elevated soil moisture leads to lower albedo, higher emissivity, and higher evaporative fraction.

For example, Post et al. (2000) shows wet soil changes from wilting to field capacity can decrease surface albedo by up to 15%. Using NDVI datasets and a seasonal simulation using a coupled modeling system, Matsui et al. (2003) showed that soil moisture and evaporative fraction appear to be directly correlated. Soil moisture can lower the surface upward longwave radiation. Small and Kurc (2003) found that a volumetric water content increase of ≈5% yields an increase of 50 W m^{-2} of net radiation due to the decreased $\varepsilon\sigma T_S^4$.

The radiative, temperature, hydrological, and the biogeochemical processes over the land surface are inter-related. Analysis of the seasonal and annual values of CO_2 flux and water vapor exchange across global Fluxnet sites in forests, grasslands, crops, and tundra was pursued by Niyogi et al. (2004). They found that net carbon and net primary productivity uptake are greater under diffuse than under direct radiation conditions. The variability in seasonal plant phenological cycles modify total leaf area, canopy transmission/reflection, nutrient and carbon exchange rates, shading and canopy scaling, and root extent, all of which modify the capacity for the net photosynthesis, soil-moisture uptake and transpiration rate. Thus, changes in the biogeochemical cycles are also linked to the land-surface feedback and alterations of surface albedo and evaporative fraction.

Indeed, the presence of soil moisture feedback associated with landform changes has a strong impact on terrestrial ecosystem processes as well. Analysis by Niyogi and Xue (2006), using a coupled photosynthesis modeling system and resource allocation analogy, indicates that the soil moisture availability controls the participation of soil and vegetative processes as a unified system; and that under non-drought conditions, the transpiration and carbon assimilation responses are about 10–15% more than under moisture stress.

Figure 2 shows the partitioning of the sensible and latent heat flux under high and low vegetation, and for soil moisture availability. The results are based on a coupled land – atmospheric model (Alapaty et al. 1999, 2001) that was set up over the central United States for typical summertime environmental conditions. As shown in the figure, when the landform has high vegetation and abundant soil moisture, the incoming radiation is dominantly partitioned into the latent heat flux (combination of bare evaporation and canopy transpiration), and the residue is further partitioned into the sensible and ground heat flux in that order. Under low vegetation and low soil moisture conditions, the radiative energy is primarily partitioned into surface sensible heat flux and the residue is partitioned into ground heat flux and latent heat. Thus, the vegetated surface would have relatively lower air temperatures and higher specific humidity as compared to the low vegetation fraction and low soil moisture landform.

Figure 3 shows the surface air and specific humidity corresponding to the energy partitioning presented in Fig. 2. As shown in Fig. 3, there can be significant differences in the maximum daytime and minimum nighttime air temperatures and surface humidity as a response to the land atmosphere interactions. As shown in Pielke et al. (2009), changes in landform characteristics can significantly influence observations and can result in misinterpretations on the reasons for decadal surface temperature trends.

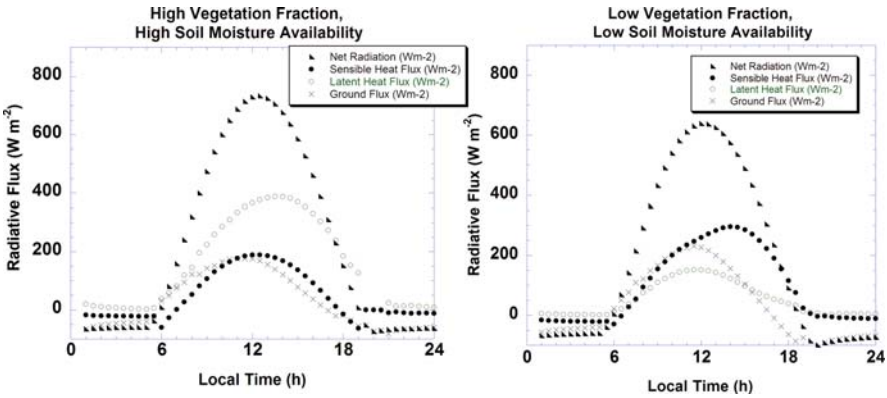

Fig. 2 Coupled land–atmosphere model-simulated diurnal surface radiative flux components representative of a vegetated high soil moisture landscape and a low vegetation fraction, low soil moisture landscape

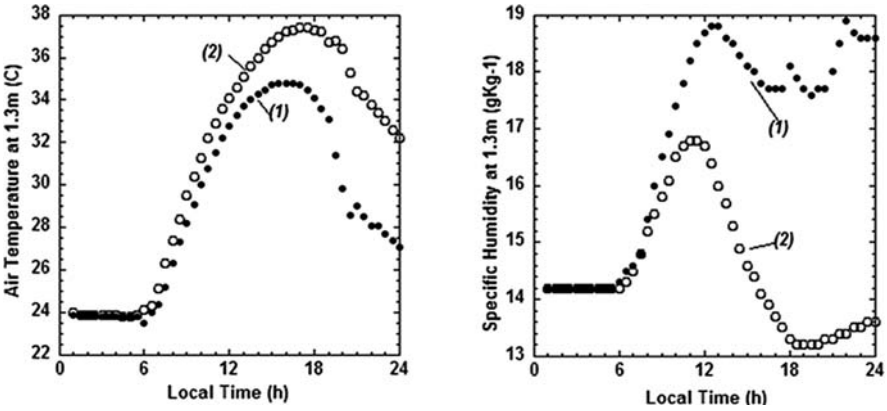

Fig. 3 Surface air temperature and specific humidity for (*1*): high vegetated/high soil moisture and (*2*): low vegetated/soil moisture landscape conditions and energy partitioning shown in Fig. 2

2.1 Boundary Layer Processes

The influence of land-surface characteristics is large on the atmospheric surface layer (≈100 m during daytime). The entire atmospheric boundary layer, of course, is also influenced by the land-surface feedback and surface turbulent heat fluxes (Pielke 2001a). The surface fluxes, particularly during daytime, typically have a maxima at the surface. The height at which the turbulence ceases to exist (or is a very small fraction of the surface value) is designated as the boundary layer height.

During daytime, from Eqs. (1) and (2), the overall effect of elevated soil moisture and vegetation availability will often result in an increase in surface net available energy and evaporative fraction so as to induce an unstable vertical distribution in

moisture and temperature in the lower atmosphere (Eltahir 1998). These fluxes can alter the surface temperature, provide more energy for evaporation from the bare soil, and transpiration from the vegetated surface. The moisture fluxes, in tandem with the winds and sensible heat can provide energy for convective clouds (e.g., Pielke 2001a). Transformation of increased surface albedo, (as a result of deforestation or desertification), leads to less availability of turbulent energy flux and thus a smaller likelihood of moist convection (e.g., Charney et al. 1977, Sud and Molod 1988, Xue and Shukla 1993). Similar to the surface albedo, the evaporative fraction: $Q_{LE}/(Q_H + Q_{LE})$, defined as a fraction of the latent turbulent heat flux over the available turbulent heat energy, varies with the landform.

Kabat et al. (2004) summarize, for example, that the evaporative fractions over a temperate forest is observed to be twice as large as that over a boreal forest. Consequently, the daytime boundary layer over the temperate forest is typically $\approx$1500 m, while it is closer to $\approx$3000 m for the boreal forest. Similarly the boundary layer heights over croplands and wooded forests are typically of the order of 1000 and 1500 m, respectively. A typical time evolution of the potential temperature sounding corresponding to the high and low vegetation and soil moisture conditions (corresponding to Figs. 2 and 3) is shown in Fig. 4. The boundary layer height (typically considered as the point of inversion on the temperature profile) is generally higher for low vegetation and low soil moisture conditions (which as shown in Fig. 2 are conditions of high surface sensible heating). Higher evaporative fraction (EF) lowers afternoon Planetary Boundary Layer (PBL) height and the Lifting Condensation Level (LCL) level, because of the suppressed sensible heat flux.

Alapaty et al. (1999) analyzed results from a 1D soil – vegetation – atmosphere – transfer model to assess the impact of different landform characteristics on the atmospheric boundary layer evolution. They concluded that soil texture changes have a

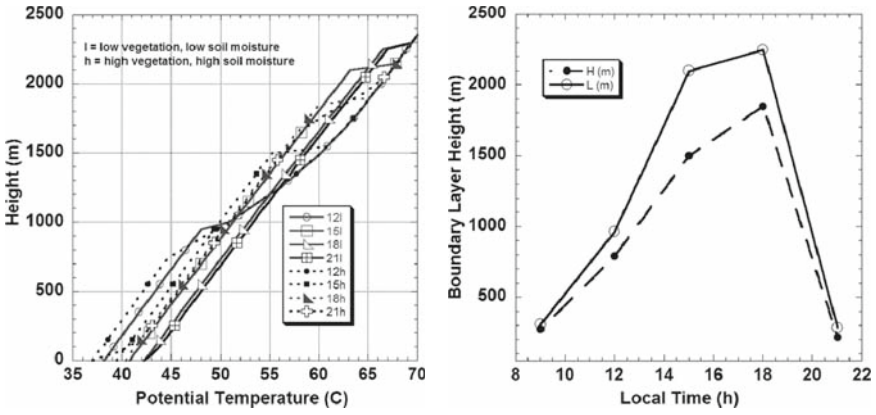

Fig. 4 (*left*) Potential temperature sounding and (*right*) the estimated atmospheric boundary layer height for high (h) and low (l) vegetation and soil moisture availability. The number in the legend for the plot on the *left* refers to the local time

dominant effect on surface fluxes, and boundary layer evolution because of its control of the hydraulic and thermal conductivity. A similar order of magnitude effect was noted when the vegetation characteristics such as transpiration/canopy resistance to water vapor exchange were modified. Depending on the land surface characteristics they found that the mid afternoon boundary layer height could vary from 1200 to 2400 m for otherwise similar initial atmospheric conditions.

Pielke (2001a) estimated the mean July convective available potential energy (CAPE) for North America using the 12 UTC rawinsonde observations. Typical values ranged between 300 and 600 J kg^{-1} over the Midwest and Great Plains and between 600 and 1000 for the southeast. If the dewpoints were increased by 1°C, an increase in the CAPE values by over 40% in some cases is obtained. The values tend to be increased by 200–400 J kg^{-1} for the Midwest and the Plains, and about 400–600 J kg^{-1} for the southeast. A 1°C warming of the surface layer temperature has a much smaller increase in the CAPE values $\approx$200 J kg^{-1} at most.

Changes in the boundary layer also affect the ventilation for atmospheric trace gases and aerosols. Shallower boundary layers generally have higher concentrations of aerosols and trace gases. The vertical gradients in the aerosols can cause further modification in diabatic heating rates within the boundary layer and thus alter the cloud and convective potential.

Thus, landscape processes and land use have a major effect on surface and vertical boundary layer fluxes of heat, water vapor, and other trace gases and aerosols.

2.2 Heterogeneous Mesoscale and Regional Landscapes

The mosaic of heterogeneous landforms or land-surface discontinuities lead to lateral gradients in surface fluxes. Often, the available solar radiative flux received at the surface is relatively unchanged at a regional scale. The energy that the surface and boundary layer receive via increased latent heat flux is compensated by a loss in sensible heat flux. The gradients in mesoscale fluxes (that is, over areas that are typically tens of km) lead to atmospheric boundaries that can trigger and organize regional circulation and convergences. For example, higher evaporative fraction contributes to moisture convergence (e.g., Brubaker et al. 1993). As a result, entropy (moist static energy) is concentrated in the lower atmosphere, and is more likely to trigger moist convection (e.g., Houston and Niyogi 2007). Indeed, the horizontal variation in evaporative fraction (or sensible heat flux) can induce local solenoidal wind circulations (Simpson 1994). Subsequently, the likelihood of deep cumulus convection is increased in response to boundary wind convergence associated with local wind circulations (as is also seen for example in the case of sea breeze circulations, Pielke 1974). The deep cumulonimbus clouds export heat, moisture and kinetic energy to the upper troposphere, and these variables are diverged to form stratiform clouds (e.g., Houze 1993). Thus, changes in surface albedo and evaporative fraction due to the landform characteristics can affect the likelihood of thunderstorms (Pielke 2001a), which would result in further alteration in surface fluxes elsewhere through nonlinear feedbacks within the atmosphere's regional atmospheric

circulations (Chase et al. 2000). We provide specific examples of such regional scale feedbacks here.

Pielke (1974) provided one of the first three dimensional modeling analyses of land-surface heterogeneities, using the southern Florida coastline as an example. The land – sea surface heterogeneity provides a classical case for the impact of landform heterogeneity and resulting impact on the regional convergence, mesoscale convection, rainfall occurrences, and the location of preferential zones for thunderstorms. Landform heterogeneities are ubiquitously present and can provide the setting for regional moisture convergences. Landscape heterogeneities tend to create similar mesoscale circulations as seen for land – sea breeze circulations (Segal et al. 1988).

Figure 5 shows an example of the impact of urban – rural land form heterogeneity and the land – sea heterogeneity on regional mesoscale circulation and convergence patterns. The domain covers central and eastern North Carolina, USA. As shown in Fig. 5 (left, top), the dark regions in the coastal periphery are indicative of vegetated landscape during late spring. The central North Carolina region is more urbanized and has different average landscape characteristics (for albedo and other vegetation/soil parameters). As a result of this landscape heterogeneity, the central region has relatively warmer surface temperatures as compared to the more vegetated regions (Fig. 5 right, top). The gradients in the surface temperature lead to sufficient gradients in the surface mesoscale boundaries so as to trigger an active mid to late afternoon sea breeze that penetrates about 50–70 km inland in the simulation (Fig. 5 left, bottom). Also interesting is the local circulation and convergence formed due to the urban – rural heterogeneity with the winds converging to central North Carolina. At night there are still remnants of some local circulations inland along with a significant land breeze in the coastal region (Fig. 5 right, bottom).

A traditional view related to the significance of the landform heterogeneity implied their significance only under calm weather conditions with weak synoptic forcing. Such conditions typically, for example, include summer afternoons without frontal passages. The landform heterogeneities are expected to provide significant surface forcing to cause modifications in mesoscale convergence and convective potential. Recent findings, however, provide consistent evidence that the landform heterogeneities and the land surface responses are generically important even under active synoptic conditions. As an example, Vidale et al. (1997) reviewed the fine/sub kilometer scale landscape heterogeneity during the Boreal forest experiment (BOREAS) and concluded that both the fine-scale turbulent and the mesoscale fluxes together contribute in generating realistic surface – atmosphere interactions under both the high and low synoptically active conditions. The landform heterogeneities exert an influence possibly because they tend to form coherent structures that have predictable features (Zeng and Pielke 1993).

Similarly, Pielke et al. (1997) used updated USGS land-use data to accurately simulate a dryline case over the Great Plains of the USA. With the correct (current) land surface representation, the model was able to reproduce an organized cumulonimbus system. With the natural land surface, only a disorganized line of towering cumulus were generated.

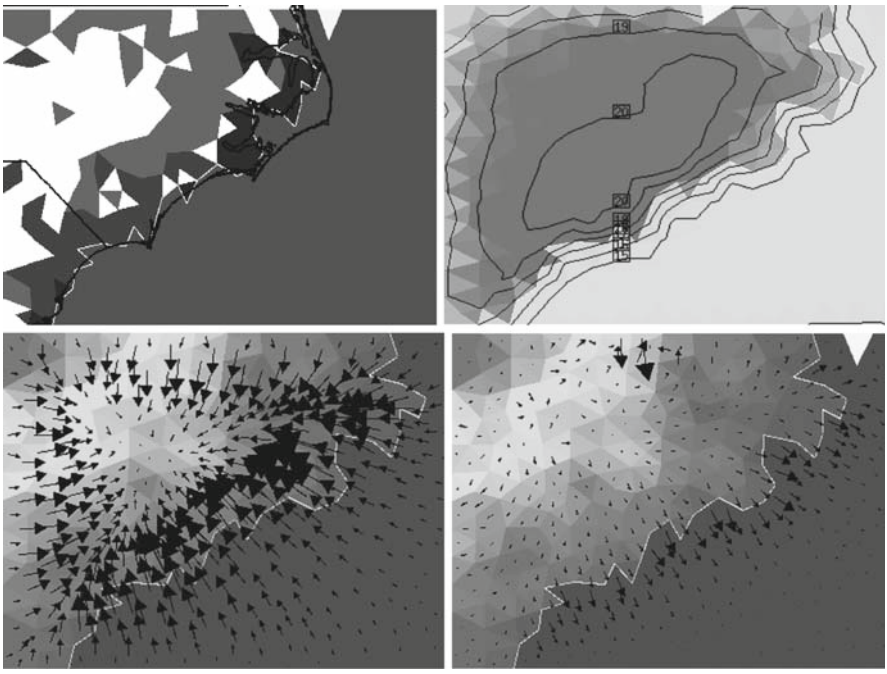

Fig. 5 Example of the impact of landform heterogeneity manifesting in regional temperature, and convergence/circulation patterns. The domain is for coastal and central North Carolina, USA. The panel shows, (*top left*) dark areas of high vegetation, and the white (light) areas of higher urban center; (*top right*) simulated midday surface temperature gradients for a typical early spring; (*bottom left*) simulated afternoon sea breeze and urban convergence due to heterogeneity; and (*bottom right*) nighttime land breeze and zones of local inland convergences

In another recent analysis, Holt et al. (2006) studied the impact of the representation of land surface heterogeneity on a synoptically active summer storm event over the Southern Great Plains that was observed during the International H2O Project (IHOP 2002). Based on the synthesis of different model configurations, the study concluded that only when an accurate representation of the landform characteristics is made, is the coupled atmospheric model able to simulate the storm characteristics of the synoptically forced convection event. Thus the land-surface feedback and heterogeneity significantly affect the timing, location, and intensity of storms and associated rainfall.

The Holt et al. study was extended further to study the case of urban – rural heterogeneity by Niyogi et al. (2006). Taking a thunderstorm case from the Joint Urban Experiment over Oklahoma City, Niyogi et al. showed that the urban landforms act in modifying the storm direction and intensity by creating changes in convergence zones. The feedback was diagnosed using a statistical – dynamical approach in which the "pure" effect of the urban area, the rural landscape, and the urban – rural heterogeneity interacted to affect the convection. The results indicate that the

rural areas can provide the moisture feedback to make the storms more convective, while the urban regions with their roughness and sensible heat flux gradients can induce zones of enhanced convection at the boundaries that can cause more intense storms. The interaction between the mesoscale boundaries due to the land gradients appears to cause the tilting of the storms that leads them into an enhanced convection region after splitting around the urban landscape (Shepherd 2005).

The majority of the studies that report on the impact of land-surface processes on the weather and climate patterns are often process based and rely on few select cases. However, the findings are generic enough so as to be applied to regional climate studies. For example, Marshall et al. (2004a) extended the sea breeze – land surface heterogeneity analysis of Pielke (1974) to assess the impact of agriculture -based land transformation on the regional climate over Florida. They concluded that the decreasing mean July–August rainfall and the increasing surface temperature from 1924 to 2000 over peninsular Florida can be explained by the massive land transformation that has occurred at the regional scale. The changes in the land cover have led to the modification of convection potential and hence the associated rainfall. The land cover transformation has also led to large changes in the surface energy balance, often resulting in warmer temperature trends. Interestingly, the land transformation could also explain the increasing freeze damage incidences for cold winter nights.

This role of land transformation affecting regional climate was also reported over the Indian monsoon region (Roy et al. 2007). The Indian monsoon region has undergone widespread agricultural intensification since its "green revolution" in the 1960s. Analyzing the temperature data in northwest and north central India, Roy et al. conclude that the agriculture and irrigation have caused a reduction in the daily temperature range over India primarily from the March to May period over the 20th century. The processes are similar to those elucidated via cases studies in Douglas et al. (2006, 2009) and are associated with the energy budget changes and associated boundary layer feedbacks (Niyogi et al. 2007).

Studies such as Gero et al. (2006) and Pyle et al. (2009) have conducted multi-year storm analyses around urban centers. These studies use high resolution radar datasets with mesoscale models in analyzing the changes in the temperature and rainfall characteristics. While the role of urbanization and the localized warming because of urban-heat island is well known, the impact on rainfall climatology is still evolving (Shepherd 2005). An analysis reported in Pyle et al. (2007) for the Indianapolis, Indiana urban region indicates that nearly 60% of storms changed composition as influenced by urban regions compared to only 38% over the rural regions. As a response to the urban – rural heterogeneity and the associated surface fluxes, daytime convection appeared to be the most likely to change with 70% changing composition and only 30% during nighttime hours. Coupled modeling -results confirm that the urban region causes distinct differences in the regional convection via the modification of the mesoscale boundaries.

Thus, land-surface processes and the land–atmosphere feedbacks that have been delineated from the mesoscale case studies for short-term weather can be extrapolated to explain broader climatic changes. Further, there is sustained evidence now

that heterogeneous landscapes and land management have a major effect on the weather and climate at the mesoscale and regional scale.

2.3 Global Climate Effects

There are a number of papers that document that landscape effects alter the climate on the global scale. These include Werth and Avissar (2005), Feddema et al. (2006), Chase et al. (2001), Marland et al. (2003), Pielke (2005), and Voldoire (2006). Figure 6 produced from the model output in Chase et al. (2000), illustrates that the hydrologic cycle is altered when human-caused land-use change occurs. Feddema et al. (2005) shows that a significant human disturbance of the climate system on the global scale is a robust conclusion (e.g., see Fig. 7 from Feddema et al. 2005).

The reason for this is straightforward as discussed in Pielke (2001a, b) and summarized as follows. As with ENSO events (Glantz 2001), land-use change is of a *large magnitude* with respect to its alteration of the energy fluxes, *persists for a long time*, and is *spatially coherent.* Since ENSO events, which involve alterations in the spatial pattern of energy fluxes in the tropical Pacific Ocean, cause significant global climate effects, land-use change should also have global scale consequences.

The role of heterogeneous climate forcings such as landscape change was further demonstrated by Matsui and Pielke (2006) with respect to the role of the atmospheric heating by aerosols, who found that the much larger spatial variations of the direct aerosol heating of the atmosphere, has a larger effect the pressure field (and thus circulation features) in the atmosphere than do the well-mixed greenhouse gases.

Feddema et al. (2005) found, as summarized in their abstract, that considering the effects of changes in land cover to the A2 and B1 transient climate simulations described in the Special Report on Emissions Scenarios (SRES) by the Intergovernmental Panel on Climate Change lead to significantly different regional climates in

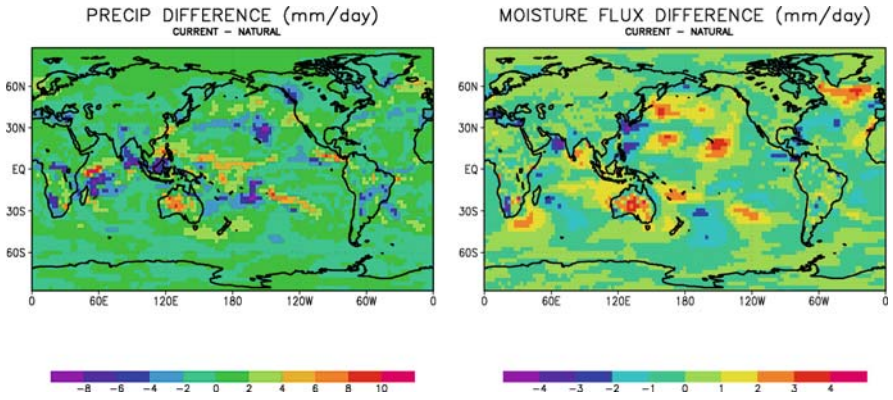

Fig. 6 Changes in the (*left*) precipitation (mm/day) and (*right*) flux changes (mm/day) for changes in the landscape. Absolute value of the global average change in the precipitation is 1.2 mm/day and the moisture flux is 0.6 mm/day. From Chase et al. 2001

2100 as compared with climates resulting from atmospheric SRES forcings alone. Agricultural expansion in the A2 scenario, for example, results in significant additional warming over the Amazon and cooling of the upper air column and nearby oceans. These and other influences on the Hadley and monsoon circulations affect extratropical climates. Agricultural expansion in the mid-latitudes produces cooling and decreases in the mean daily temperature range over many areas (e.g., see Fig. 2 in their paper).

2.4 The Need to Focus on Vulnerability

Within the climate system, the need to consider the broader role of land-surface feedback becomes important not only for assessing the impacts but also for developing regional vulnerability and mitigation strategies.

The IPCC fourth assessment second and third working groups deal with a range of issues targeted to these topics (Schneider et al. 2007). The IPCC identifies seven criteria for "key" vulnerabilities. They are: magnitude of impacts, timing of impacts, persistence and reversibility of impacts, likelihood (estimates of uncertainty) of impacts and vulnerabilities and confidence in those estimates, potential for adaptation, distributional aspects of impacts and vulnerabilities, and the importance of the system(s) at risk. While a number of potential vulnerabilities and uncertainties are considered (such as irreversible change in urbanization), the resulting feedback on the atmospheric processes due to such changes is still poorly understood or unaccounted for in these assessments. Indeed the UNFCCC Article 1 states: "'Adverse effects of climate change' means changes in the physical environment or biota resulting from climate change which have significant deleterious effects on the composition, resilience or productivity of natural and managed ecosystems or on the operation of socio-economic systems or on human health and welfare." Thus, while the role of landscape is inherent within the UNFCCC framework, the corresponding translation for the assessments still remains largely greenhouse gas driven.

Further, while the climate change projections have largely been at coarser resolution, the impacts and potential mitigation policies are often at local to regional scales. For example, climate models often project increasing drought at a regional scale. The resilience to such increased occurrence as well as changes in the intensity of droughts is, however, dependent on the local scale environmental conditions (such as moisture storage, and convective rainfall), and farming approaches (access to irrigation, timing of rain or stress, etc). As summarized in Adger (1996), an important issue for IPCC-like global assessments is to assess if the top-down approach can incorporate the "aggregation of individual decision-making in a realistic way, so that results of the modelling are applicable and policy relevant".

Therefore, as the community braces to develop resilience strategies it will becoming increasingly important to consider a bidirectional impact, i.e., not just the role of atmospheric changes (such as temperature and rainfall) on the physical environmental or biota, but also a feedback of the biota and other land-surface processes on further changes in the atmospheric processes – such as reviewed in this chapter.

Klein et al. (1999) sought to assess whether the IPCC guidelines for assessing climate change impacts as well as adapative strategies can be applied to one example of coastal adaptation. They recommend that a broader approach is needed which has more local-scale information and input for assessing as well as monitoring the options. Again the missing link between local-scale features with global scale projections become apparent. The expanded eight-step approach of Schroter et al. (2005), designed to assess vulnerability to climate change, states the need for considering multiple interacting stresses. They recognize that climate change can be a result of greenhouse gas changes which are coupled to socioeconomic developments, which in turn are coupled to land-use changes – and that all of these drivers are expected to interactively affect the human – environmental system (such as crop yields).

To extract the significance of the individual versus multiple stressors on crop yields, Mera et al. (2006) developed a crop modeling study with over 25 different climatic scenarios of temperature, rainfall, and radiation changes at a farm scale for both C3 and C4 types of crops (e.g., soybean and maize). As seen in many crop yield studies, the results suggested that yields were most sensitive to the amount of effective precipitation (estimated as rainfall minus physical evaporation/transpiration loss from the land surface). Changes in radiation had a nonlinear response with crops showing an increased productivity for some reduction in the radiation as a result of cloudiness and increased diffuse radiation and a decline in yield with further reduction in radiation amounts. The impact of temperature changes, which has been at the heart of many climate projections, however, was quite limited particularly if the soils did not have moisture stress. The analysis from the multiple climate change settings do not agree with those from individual changes, making a case for multivariable, ensemble approaches to identify the vulnerability and feedbacks in estimating climate-related impacts (cf. Turner et al. 2003).

Another issue is the coupled vulnerability of the land surface to socioeconomic and climate change processes. This question was addressed by Metzger et al. (2006). They concluded that most assessment studies cannot provide needed information on regions or on ecosystem goods that are vulnerable. To address this question, we can hypothesize that the vulnerability of landscape (V) change is a product of the probability of the landscape change (Lc) and the service (S) provided by the landscape:

$$V = \text{prob}\,(Lc)\ ^{*}\ S \tag{4}$$

The service provided is a broad term and could mean societal benefits (such as recreation), or economic benefits (such as timber and food), or physical feedbacks as in terms of the modulating impact a landscape may have on regional temperatures or precipitation. While a variety of studies on vulnerability have sought to look at the economic and societal feedbacks, the physical feedback of the fine-scale land heterogeneities have been critically missing in the literature. It is however important that land heterogeneity and transformation potential be considered at a finer scale because the landscape changes will in turn affect the regional and local vulnerability.

Current economical assessment studies (Stern 2007) conclude that controlling land-use change such as from deforestation provides an opportunity cost in excess of $5 billion per annum. This estimate however appears to only consider the land transformation impact of deforestation and the resulting greenhouse emissions. As summarized in this chapter, the dynamical effects such as changes in rainfall, evaporation, convection, and temperature patterns due to landform changes can cause additional vulnerability (or resilience in some cases) and needs to be considered in such assessments (Marland et al. 2003). Similarly, the UNFCCC Article 3 also seeks afforestation (reforestation minus deforestation) since 1990 as a country's commitment towards the greenhouse gas emission controls. Not considering the dynamical feedbacks due to such forest land transformation can lead to additional vulnerabilities as described in Pielke et al. (2001a, 2002).

3 Conclusions

Humans are significantly altering the global climate, but in a variety of diverse ways beyond the radiative effect of carbon dioxide (Cotton and Pielke 2007, Kabat et al. 2004, National Research Council 2005). The IPCC assessments have been too conservative in recognizing the importance of these human climate forcings as they alter regional and global climate. These assessments have also not communicated the inability of the models to accurately forecast the spread of possibilities of future climate. The forecasts, therefore, provide limited skill in quantifying the impact of different mitigation strategies on the actual climate response that would occur. In this paper, we discuss one of these issues, namely the role of land-surface processes within the climate system, including how these processes are altered by human activities.

The needed focus for the study of climate change and variability is, therefore, on the regional and local scales. Global and zonally-averaged climate metrics would only be important to the extent that they provide useful information on these space scales. Global and zonally-averaged surface temperature trend assessments, besides having major difficulties in terms of how this metric is diagnosed and analyzed, do not provide significant information on climate change and variability on the regional and local scales.

Global warming is also not equivalent to climate change. Significant, societally important climate change, due to both natural- and human – climate forcings, can occur without any global warming or cooling. Landscape management is one example of such a forcing. Thus attempts to significantly influence regional and local-scale climate based on controlling CO_2 emissions alone is an inadequate policy for this purpose.

We, therefore, propose that the assessment of vulnerability, focused on regional and local societal and environmental resources of importance, is a more inclusive, useful, and scientifically robust framework to interact with policymakers, than is the focus on global multi-decadal climate predictions which are downscaled to the

regional and local scales (see also Pielke Jr. et al. 2007). The vulnerability paradigm permits the evaluation of the entire spectrum of risks associated with different social and environmental threats, including climate variability and change.

Finally, unless there is a broadening of the current IPCC focus it will only lead to promote energy policy changes, and not provide an effective climate policy, which necessarily needs to include how humans are altering the climate system through land surface processes. Policymakers need to be informed of this very important distinction where a separation of climate policy from energy policy is essential.

Acknowledgements This work was supported in part by the Shortgrass Steppe Long Term Ecological Research project by funds from the NSF award DEB 0217631, NASA Grant NNX06AG74G and NNX07AG35G, NASA GWEC Grant NNG05GB41G, NOAA JCSDA Grant NA06NES4400013, DOE ARM 08ER64674, and NSF CAREER ATM-0847472 and NSF Grant ATM-0296159. Roger A. Pielke Sr. was also supported in part by the University of Colorado at Boulder (CIRES/ATOC). Dallas Staley completed the editing of this paper in her usual outstanding manner, and we gratefully acknowledge this very important contribution.

References

Adger WN (1996) Approaches to vulnerability to climate change. CSERGE Working Paper GEC 96-05, ISSN 0967-8875

Alapaty K, Raman S, Niyogi D (1999) Uncertainty in the specification of surface characteristics: A study of prediction errors in the boundary layer. Bound-Layer Meteor, 82: 475–502

Alapaty K, Seaman N, Niyogi D, Hanna A (2001) Assimilating surface data to improve the accuracy of atmospheric boundary layer simulations. Journal of Applied Meteorology, 40: 2068–2082

Alpert P, Niyogi D, Pielke RA Sr, Eastman JL, Xue YK, Raman S (2006) Evidence for carbon dioxide and moisture synergies from the leaf cell up to global scales: Implications to human-caused climate change. Global Planetary Change, Special Issue, 54: 202–208

Brubaker KL, Entekhabi D, Eagleson PS (1993) Estimation of continental precipitation recycling. Journal of Climate, 6: 1077–1089

Charney JG, Quirk WJ, Chow SH, Kornfield J (1977) A comparative study of the effects of albedo change on drought in semi-arid regions. Journal of the Atmospheric Sciences, 34: 1366–1385

Chase TN, Pielke RA, Kittel TGF, Nemani RR, Running SW (2000) Simulated impacts of historical land cover changes on global climate in northern winter. Climate Dynamics, 16: 93–105

Chase TN, Pielke RA Sr., Kittel TGF, Zhao M, Pitman AJ, Running SW, Nemani RR (2001) The relative climatic effects of landcover change and elevated carbon dioxide combined with aerosols: A comparison of model results and observations. Journal of Geophysical Research -Atmospheres 106, 31: 31685–31691

Cotton WR, Pielke RA Sr (2007) Human impacts on weather and climate. Cambridge University Press, 330 pp

Douglas E, Niyogi D, Frolking S, Yeluripati JB, Pielke RA Sr, Niyogi N, Vörösmarty CJ, Mohanty UC (2006) Changes in moisture and energy fluxes due to agricultural land use and irrigation in the Indian Monsoon belt, Geophysical Research Letters, 33: L14403, doi:10.1029/2006GL026550

Douglas E,M, Beltrán-Przekurat A, Niyogi D, Pielke RA Sr, Vörösmarty CJ (2009) The impact of agricultural intensification and irrigation on land-atmosphere interactions and Indian monsoon precipitation – a mesoscale modeling perspective. Global Planetary Change, doi:10.1016/j.gloplacha.2008.12.007 in press

Eltahir, EAB (1998) A soil moisture-rainfall feedback mechanism 1. Theory and observations, Water Resources Research, 34: 765–776

Feddema J, Oleson KW, Bonan GB, Mearns LO, Buja LE, Meehl GA, Washington WM (2005) The importance of land-cover change in simulating future climates. Science, 310: 1674–1678

Feddema J, Oleson KW, Bonan GB, Mearns LO, Meehl GA, Washington WM, Meehl G, Nychka D. (2006) A comparison of a GCM response to historical anthropogenic land cover change and model sensitivity to uncertainty in present-day land cover representations. Climate Dynamics, 25: 581–609

Gero AF, Pitman AJ, Narisma GT, Jacobson C, Pielke RA Sr (2006) The impact of land cover change on storms in the Sydney Basin, Australia. Global and Planetary Change, 54: 57–78

Glantz MH (2001) Currents of change: Impacts of El Niño and La Niña on climate and society. 2nd Edn, Cambridge University Press, 266 pp

Holt T, Niyogi D, Chen F, LeMone MA, Manning K, Qureshi AL (2006) Effect of land – atmosphere interactions on the IHOP 24–25 May 2002 convection case. Monthly Weather Review, 134: 113–133

Houston A, Niyogi D (2007) The sensitivity of convective initiation to the lapse rate of the active cloud-bearing layer, Monthly Weather Review, 135: 3013–3032 DOI: 10.1175/ MWR3449.1

Houze RA Jr (1993) Cloud Dynamics, Academic Press, San Diego, CA, 573 pp

Idso SB, Jackson RD, Reginato RJ, Kimball BA, Nakayama FS (1975) The dependence of bare soil albedo on soil water content. Journal of Applied Meteorology, 14: 109–113

IPCC (2007) Intergovernmental Panel for Climate Change. http://www.ipcc.ch accessed November 2007

Kabat P, Claussen M, Dirmeyer PA, Gash JHC, Bravo de Guenni L, Meybeck M, Pielke RA Sr, Vorosmarty CJ, Hutjes RWA, Lutkemeier S, (Eds.) (2004) Vegetation, water, humans and the climate: A new perspective on an interactive system. Springer, Berlin, Global Change – The IGBP Series, 566 pp

Klein RJT, Nicholls RJ, Mimura N (1999) Coastal adaptation to climate change: Can the IPCC technical guidelines be applied?Mitigation and Adaptation Strategies for Global Change, 4: 239–252

Marland G, Pielke RA Sr, Apps M, Avissar R, Betts RA, Davis KJ, Frumhoff PC, Jackson ST, Joyce L, Kauppi P, Katzenberger J, MacDicken KG, Neilson R, Niles JO, Niyogi D, Norby RJ, Pena N, Sampson N, Xue Y (2003) The climatic impacts of land surface change and carbon management, and the implications for climate-change mitigation policy. Climate Policy, 3: 149–157

Marshall Jr CH, Pielke RA Sr, Steyaert LT, Willard DA (2004a) The impact of anthropogenic land-cover change on the Florida peninsula sea breezes and warm season sensible weather. Monthly Weather Review, 132: 28–52

Marshall Jr CH, Pielke RA Sr, Steyaert LT, (2004b) Has the conversion of natural wetlands to agricultural land increased the incidence and severity of damaging freezes in south Florida? Monthly Weather Review, 132: 2243–2258

Matsui T, Lakshmi V, Small E (2003) Links between snow cover, surface skin temperature, and rainfall variability in the North American monsoon system. Journal of Climate, 16: 1821–1829

Matsui T, Pielke RA Sr (2006) Measurement-based estimation of the spatial gradient of aerosol radiative forcing. Geophysical Research Letters, 33: L11813, doi:10.1029/2006GL025974

Mera RJ, Niyogi D, Buol GS, Wilkerson GG, Semazzi F (2006) Potential individual versus simultaneous climate change effects on soybean (C3) and maize (C4) crops: An agrotechnology model based study, global and Planetary Change Special Issue, 54: 163–182

Metzger MJ, Rounsevell MDA, Acosta-Michlik L, Leemans R, Schroter D (2006) The vulnerability of ecosystem services to land-use change. Agriculture, Ecosystems and Environment, 114: 69–85

National Research Council (2005) Radiative forcing of climate change: Expanding the concept and addressing uncertainties. Committee on Radiative Forcing Effects on Climate Change, Climate Research Committee, Board on Atmospheric Sciences and Climate, Division on Earth and Life Studies, The National Academies Press, Washington, DC, 208 pp

Niyogi D, Chang H, Saxena VK, Holt T, Alapaty K, Booker F, Chen F, Davis KJ, Holben B, Matsui T, Meyers T, Oechel WC, Pielke RA Sr, Wells R, Wilson K, Xue YK (2004) Direct observations of the effects of aerosol loading on net ecosystem CO_2 exchanges over different landscapes. Geophysical Research Letters, 31: L20506, doi:10.1029/2004GL020915

Niyogi D, Chang H-I, Chen F, Gu L, Kumar A, Menon S, Pielke RA Sr (2007) Potential impacts of aerosol-land-atmosphere interaction on the Indian monsoonal rainfall characteristics. Natural Hazards, Monsoon Special Issue, DOI – 10.1007/s11069-006-9085-y

Niyogi D, Holt T, Zhong S, Pyle PC, Basara J (2006) Urban and land surface effects on the 30 July 2003 mesoscale convective system event observed in the Southern Great Plains. Journal of Geophysical Research, 111: D19107, doi:10.1029/2005JD006746

Niyogi D, Xue Y (2006) Soil moisture regulates the biological response of elevated atmospheric CO_2 concentrations in a coupled atmosphere biosphere model, global and Planetary Change Special Issue, 54: 94–108

Pielke Jr RA, Prins G, Rayner S, Sarewitz D (2007) Lifting the taboo on adaptation. Nature, 445: 597–598

Pielke Sr RA (1974) A three-dimensional numerical model of the sea breezes over south Florida. Monthly Weather Review, 102: 115–139

Pielke Sr RA (2001a) Influence of the spatial distribution of vegetation and soils on the prediction of cumulus convective rainfall. Reviews Geophysics, 39: 151–177

Pielke Sr RA (2001b) Carbon sequestration – The need for an integrated climate system approach. Bulletin of the American Meteorological Society, 82: 2021

Pielke Sr RA (2005) Land use and climate change. Science, 310: 1625–1626

Pielke Sr RA, Adegoke J, Beltran-Przekurat A, Hiemstra CA, Lin J, Nair US, Niyogi D, Nobis TE (2007) An overview of regional land use and land cover impacts on rainfall. Tellus B, 59B: 587–601

Pielke Sr RA, Davey C, Niyogi D, Fall S, Steinweg-Woods J, Hubbard K, Lin X, Cai M, Lim Y-K, Li H, Nielsen-Gammon J, Gallo K, Hale R, Angel J, Mahmood R, Foster S, McNider RT, Blanken P (2009) Unresolved issues with the assessment of multi-decadal global land surface temperature trends. Journal of Geophysical Research, 114, D05105, doi:10.1029/2008JD 010938

Pielke Sr RA, Lee TJ, Copeland JH, Eastman JL, Ziegler CL, Finley CA (1997) Use of USGS-provided data to improve weather and climate simulations. Ecological Applications, 7: 3–21

Pielke Sr RA, Marland G, Betts RA, Chase TN, Eastman JL, Niles JO, Niyogi D, Running S, (2002) The influence of land-use change and landscape dynamics on the climate system: Relevance to climate change policy beyond the radiative effect of greenhouse gases. The Philosophical Transactions of the Royal Society, London a Special Theme Issue, 360: 1705–1719

Post DF, Fimbres A, Matthias AD, Sano EE, Accioly L, Batchily AK, Ferreira LG (2000) Predicting soil albedo from soil color and spectral reflectance data. Soil Science Society of American Journal, 64: 1027–1034

Pyle P, Niyogi D, Arya SP, Shepherd M, Chen F, Wolfe B (2009) An observational and modeling-based storm climatology assessment for the Indianapolis, urban region. Journal of Applied Meteorological Climate, conditionally accepted

Roy SS, Mahmood R, Niyogi D, Lei M, Foster SA, Hubbard KG, Douglas E, Pielke RA Sr (2007) Impacts of the agricultural Green Revolution-induced land use changes on air temperatures in India. Journal of Geophysical Research, 112: D21108, doi:10.1029/2007JD008834

Schneider SH, Semenov S, Patwardhan A, Burton I, Magadza CHD, Oppenheimer M, Pittock AB, Rahman A, Smith JB, Suarez A, Yamin F (2007) Assessing key vulnerabilities and the risk from climate change. Climate Change 2007: Impacts, Adaptation and Vulnerability. Contribution of Working Group II to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, Parry ML, Canziani OF, Palutikof JP, van der Linden PJ, Hanson CE, (Eds.), Cambridge University Press, Cambridge, UK, 779–810

Schroter D, Polsky C, Patt AG (2005) Assessing vulnerabilities to the effects of global change: An eight step approach. Mitigation and Adaptation Strategies for Global Change 10: 573–596

Segal M, Avissar R, McCumber M, Pielke RA Sr (1988) Evaluation of vegetation effects on the generation and modification of mesoscale circulations. Journal of Atmospheric Sciences, 45: 2268–2293

Shepherd JM (2005) A review of current investigations of urban-induced rainfall and recommendations for the future. Earth Interactions 9: 1–27

Simpson JE (1994) Sea breeze and local winds. Cambridge University Press: Cambridge, UK. ISBN 0-521-45211-2. xi, 234 pp.

Small EE, Kurc SA (2003) Tight coupling between soil moisture and the surface radiation budget in semiarid environments: Implications for land-atmosphere interactions. Water Resources Research 39: 1278, doi:10.1029/2002WR001297

Stern N (2007) The economics of climate change: The Stern review. Cambridge University Press, Cambridge, UK, 712 pp

Sud YC, Molod A (1988) A GCM simulation study of the influence of Saharan evapotranspiration and surface-albedo anomalies on July circulation and rainfall. Monthly Weather Review 116: 2388–2400

Turner II BL, Kasperson RE, Matson PA, McCarthy JJ, Corell RW, Christensen L, Eckley N, Kasperson JX, Luers A, Martello ML, Polskya C, Pulsipher A, Schiller A (2003) A framework for vulnerability analysis in sustainability science. Proceedings of the Natural Academy of Sciences, 100(14): 8074–8079 www.pnas.orgycgiydoiy10.1073ypnas.1231335100

UNFCCC (2007) The United National Framework Convention on Climate Change, unfccc.int accessed November 2007

Vidale PL, Pielke RA, Barr A, Steyaert LT (1997) Case study modeling of turbulent and mesoscale fluxes over the BOREAS region. Journal of Geophysical Research 102: 29167–29188

Voldoire A (2006) Quantifying the impact of future land-use changes against increases in GHG concentrations. Geophysical Research Letters 33: L04701, doi:10.1029/2005GL024354

Werth D, Avissar R (2005) The local and global effects of Southeast Asian deforestation. Geophysical Research Letters, 32: L20702 doi:10.1029/2005GL022970

Xue Y, Shukla J (1993) The influence of land surface properties on Sahel climate. Part 1: Desertification. Journal of Climate, 6: 2232–2245

Zeng X, Pielke RA (1993) Error-growth dynamics and predictability of surface thermally-induced atmospheric flow. Journal Atmospheric Sciences, 50: 2817–2844

The Impact of Landform Structure on the Formation of Fog – Numerical Simulations with COSMO-FOG

Isabel Alberts, Matthieu Masbou, and Andreas Bott

Abstract The development of fog is very sensitive to the physical and thermodynamical structure of the surface layer. Especially turbulent heat and moisture fluxes at the earth's surface and in the atmosphere have an effect on duration and intensity of fog events. Spatio-temporal patterns of the surface properties control these fluxes which depend in a complex way on physical conditions such as local soil properties, soil moisture, radiation received at the surface, specific land use and orography. In the same way the soil-canopy-atmosphere interactions influenced the formation and dissipation of fog. Thus, the influence of the surface layer on the formation of fog is investigated. In order to take the high spatial variability of landform structure into account, numerical simulations are performed with the three-dimensional fog forecasting model COSMO-FOG.

Keywords Fog · Fog modeling · 3-D modeling · COSMO-FOG · Air–land interaction

1 Introduction

The necessity of realistic fog simulations is caused by a wide range of areas: first of all, fog events have high impact on various branches, mostly in connection with traffic safety, e.g., car accidents because of low visibility and disturbance of time-table of airports and shipping ports. Thus, the forecast of fog events is also a task of economic benefit. In literature, a wide range of numerical models of different complexity for the simulation of the main features of fog exists (e.g., Bott et al., 1990, 1996; Bott and Trautman, 2002; Bergot and Guedalia, 1994; Clark and Hopwood, 2001; Fisher and Caplan, 1963; Forkel et al., 1987; Masbou and Bott, 2005; Müller, 2006; Teixeira and Mirinda, 2001). But still, there are difficulties in accurately forecasting

I. Alberts (✉)
Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany
e-mail: ialberts@uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_6,

fog and representing the physical processes involved in fog formation and dissipation (Gulteppeet al., 2007).

These processes can vary nonlinearly over spatial scales due to subgrid spatial variability of soil and vegetation characteristics, topography, and near-surface water storage (Boone et al., 2004; Pielke, 1984). Therefore, detailed three-dimensional fog models are needed for better understanding issues related to fog. The model used in this study is the three-dimensional fog forecast model COSMO-FOG (during the development it was called LM-PAFOG, Masbou and Bott, 2005).

Since fog is a meteorological phenomenon of the boundary layer, surface conditions have a strong influence on the formation, evolution and dissipation of fog (Gulteppe et al., 2007). Hence, the correct description of interaction between surface and atmosphere and the processes within the soil are of special interest. In this study, the coupling between the atmosphere and soil is done by the SVAT (Soil-Vegetation-Atmosphere-Transfer) scheme TERRA-ML (TERRA-Multi Layer) whose main functions are to compartmentalize incoming energy into fluxes of heat and moisture.

Numerous field experiments exist helping to improve the understanding of the link between the earth surface and the atmosphere: e.g., *Hydrologicale – Atmospheric pilot Experiment-Modelisation du Bilan Hydrique, HAPEX-MOBILHY* (André et al., 1986); *First International Satellite Land surface Climatology Project (ISLCP) Field Exeriment, FIFE* (Sellers et al., 1992); *Boreal Ecosystem – Atmosphere Study, BOREAS* (Sellers et al., 1997); *Project for the Intercomparison of Land Surface Parameterization Schemes, PILPS (*Henderson-Sellers et al., 1993) and the Rhône-Aggregation (Rhône-AGG) Land Surface Scheme (LSS) intercomparison project (Boone, 2004). In the same way the work of Avissar and Pielke (1989) underline the importance of understanding the spatial heterogeneity of the surface.

But still, there is a lack of accurate modelling of the basic exchange mechanism particularly in connection with fog events. Several approaches on numerical modelling of fog events and the interaction with vegetation exist (Duynkerke, 1991; von Glasow and Bott, 1999; Winterrath, 2001). However, none of these include a three dimensional fog model with detailed description of the surface.

This paper is structured as follows: firstly the definition and a short explanation of the impacts on fog development is given and secondly, a description of the basic characteristics of the used model follows. Section 4 gives the results for a sensitivity study, followed by a conclusion and an outlook.

2 Fog Characteristics

2.1 Definition of Fog and Influences on the Formation

Fog can be defined as a stratiform cloud layer on the surface or next to it. A common definition of fog is that of the World Meteorological Organisation (WMO, 1992):

> Fog is a suspension of very small, usually microscopic water droplets in the air, generally reducing the horizontal visibility at the earth's surface to less than 1 km

As mentioned before, the life cycle of a fog event is strongly dependant on the impacts of the atmospheric boundary layer. The boundary layer is defined "as that part of the troposphere that is directly influenced by the presence of the earth's surface and responds to surface forcing with a timescale of about an hour or less" (Stull, 1988). Thus, a fog event is directly influenced from the surface by modifications of the profiles of wind, temperature and humidity. Local circulations are controlled by horizontal heterogeneities. In addition, the life cycle of fog events is indirectly influenced by modifications of radiative properties of the atmosphere by microphysical processes and varying aerosol spectra (Gulteppe et al., 2007). Incoming and outgoing radiation determines the surface temperature changes. Heat fluxes in the atmosphere, the canopy and the soil, surface albedo as well as emissivity influence the radiative balance. Hence, the state of the underlying surface and variations in surface composition, including soil types, vegetation, soil moisture etc. cause local variations in humidity and the rate of radiative cooling. Moist soils help to create favourable surface conditions for fog development. Different surfaces cool at different rates depending on surface type and thermal conductivity beneath the surface, and therefore surfaces cooling more rapidly reach saturation more quickly. The thermal conductivity of soil also depends on the soil moisture content.

Aside from saturated surface air, which is the most important ingredient to generate fog, the absence of cloud cover at night, relatively weak winds and high dew point temperatures play a dominant role.

Depending on the fog type, dissipation processes differ, but they generally occur due to local heating, increasing wind speed, air masses changes or precipitation occurrence (Stull, 1988).

3 Materials and Methods

3.1 Model Description

The need for a separate fog model can be explained by the spatio-temporal heterogeneities of the meteorological phenomenon fog. At least for appropriate simulations of fog with numerical weather forecast models the following two requirements have to be considered:

1. high horizontal and vertical model resolution to simulate the slow and steady growth of fog near the surface, and
2. in terms of cloud mircrophysics, a detailed description is needed because the amount of liquid water content of fog is in contrast to most of the clouds very low and varies between 0.01 and 0.5 g/m^3 (Pruppacher et al., 1997).

Therefore, the model used for this study is a coupling of the three dimensional forecast model COSMO-Model of the German Meteorological Service ("Deutscher Wetterdienst", DWD) with the one dimensional fog model PAFOG (Bott and

Trautmann, 2002), called COSMO-FOG. The microphysical part of COSMO-FOG is based on the one-dimensional fog model PAFOG. For a better description of the condensation processes, this parameterization scheme introduces the new prognostic variable **C**loud **C**ondensation **N**uclei (CNN). The model establishes a link between CNN and the specific cloud water content. This microphysic parameterization scheme replaces the condensation processes of the COSMO cloud microphysics which is based on the Kessler scheme (Kessler, 1969) in the lowest 2000 meters of the atmosphere. The mesoscale weather prediction model COSMO-Model (up to spring 2007 it was known as the "Lokal-Modell", LM [Steppeler et al., 2003]) is part of the operational weather prediction system of the DWD. The COSMO-Model is a fully compressible non-hydrostatic limited-area numerical weather prediction model where all relevant subgrid-scale processes are parameterized. Since the boundaries at the top and the sides are artificial, initial conditions and the fluxes at the lateral boundaries are provided by external data which have been calculated from the Globalmodel (GME, Majewski et al., 2002) respectively the COSMO-Model. A detailed description of the COSMO-Model design is given in Doms and Schättler (2002) and in the documentations provided at the COSMO-web page (http://cosmo-model.org).

3.2 Land Surface Scheme

In the COSMO-Model the link between the earth's surface and the atmosphere is established by the SVAT module TERRA-ML (Schrodin et al., 2001). The interaction between the earth's surface and the atmosphere is parameterized with transfer coefficients for momentum K_m, and heat and moisture K_h. For this purpose, a stability and roughness-length dependent surface flux formulation according to the Monin-Obukhov similarity theory is used (Louis, 1979). The exchange coefficients depend on the roughness length z_o and the bulk Richardson number Ri_b. The bulk formula is used for the evaluation of the sensible and latent heat flux. The sensible heat flux is computed as follows:

$$H = -pC_h^d|v|\left(\Theta \cdot \tau_{sfc} - T_{sfc}\right),$$

with the air density p, the bulk-aerodynamical transfer coefficient for turbulent heat exchange at the surface C_h^d, $|v|$ the absolute wind speed at the same level, Θ and τ_{sfc} are, respectively, the potential temperature at the lowest grid level above the earth's surface and the scaled pressure at the ground. T_{sfc} is the ground temperature.

The surface flux of latent heat is defined accordingly,

$$\lambda E = -pC_q^d|v|\left(q^v - q_{sfc}^v\right),$$

where C_q^d is the bulk-aerodynamical coefficient for turbulent moisture transfer at the surface, q^v is the specific humidity at the lowest grid level at the ground and q_{sfc}^v is the ground level specific humidity (Doms et al., 2005). Evaporation from bare

soil and plants is parameterized using the BATS (Biosphere Atmosphere Transfer) scheme of Dickinson (1984). Soil temperature is calculated by numerically solving the heat conduction equation. For this purpose and for calculating the soil water transport, the soil is subdivided into eight nonequidistant layers. A more detailed description of the calculation of hydrological processes in the soil is given in Graβelt et al. (2009, this issue).

3.3 Model Domain

The geographic domain for this study is the catchment area of the river Sieg, a tributary of the river Rhine in western Germany (German federal state North Rhine Westfalia [NRW]). This area is a typical german low mountain range, the height varies between 100 and 600 m (at the eastern border of the domain). The area is covered with meadows, forest, a variety of agricultural land cover and with urbanization in the Rhine valley (Seuffert et al., 2002)

4 Model Input Parameters

As a limited area model, just the boundary layer of the COSMO-Model is physical what means that some parameters which are required for the parameterization of physical processes cannot be derived by data assimilation or by interpolation from a driving model. These so-called external parameters are defined in additional data sets and describe the earth's surface. The following table lists the required parameters and the sources for the data, operationally used at the DWD:

Seasonal variations, e.g., of the leaf area index, are considered by varying the initial data, but not by the forecast model itself. Using a lookup-table (Schrodin, 1995), each soil type is assigned to various physical material constants like, e.g., heat conductivity or pore volume. In this study higher horizontal resolution is needed for the simulation of fog. Because no operational data set in 1 km resolution is available, the DWD provided a new data set with a resolution of 1 km via preprocessing with the same sources as for greater resolution (see Table 1). In general this process lead to a loss of information.

For the purpose of a more realistic description of the earth's surface a more detailed set of external surface parameters for the model domain has been implemented into the COSMO-FOG. The new implemented data set has a resolution of 1 km. The source for the newly modified soil data is the digital soil map 1:50,000 of the federal state NRW (Geologischer Dienst NRW, Krefeld, 2002) and Rhinland-Palatinate (Geologisches Landesamt Rheinland Pfalz, 2002). The origin of the land use data is the CORINE Land Cover 2000 data set (EEA, 2004), and for the topography, the SRTM-data are utilized (Shuttle Radar Topography Mision, DLR, 2000). With the newly generated data set in the framework of this study, the description of the surface is more realistic than the description with the DWD data. Instead of two soil classes (in the DWD-data set just loam

Table 1 External surface parameters

Parameters	References
Mean topographical height	GTOPO30 data set (Global 30 Arc second elevation data,1 km) from U.S. geological survey (USGS, 1997)
Roughness length	GTOPO30 and CORINE (CoORdination of INformation on the Environmnet, 250 m) data set
Soil type	DSMW data set (Digital Soil Map of the World, 10 km) of FAO (Food and Agricultural Organization of the United Nations)
Vegetation cover	CORINE data set of ETC/LC (European Topic Centre on Land Cover)
Land fraction	CORINE data set of ETC/LC
Root depth	CORINE data set of ETC/LC
Leaf area index	CORINE data set of ETC/LC

and loamy sand are distinguished) ten classes are in use (see figures in chapter 'Influence of Drainage Parameterization and Precipitation Analysis on Discharge Simulation in the Sieg River Catchment' by Graβelt et al.). This is much more similar to the heterogeneity of natural soil patterns. In the same way differences exist for the other parameters.

5 Results

5.1 Model Setup

In many model runs COSMO-FOG has been tested on its sensitivity to the new description of the surface layer. For this purpose, two different settings are distinguished:

1. the control run with the external data set generated by the DWD (Fog-DWD)
2. the test run with the new modified data set (Fog-Modified).

The experiments discussed in this paper were conducted under the following conditions: the COSMO-FOG (1 km grid spacing) is one-way nested into the COSMO-Model with a horizontal grid spacing of 2.8 km to provide the initialization and the lateral boundary values with a temporal resolution of 1 h. Simulations are performed on October, 5th 2005, the time step is 10 s and the model COSMO-FOG starts at 12 UTC and runs for 48 h. The synoptic situation was favorable for fog formation: a high pressure system over Europe with a flow of relative dry air from the eastern direction dominated Germany. Radiation fog is locally formed and at most places fog dissipates during the day.

6 Sensitivity Study

6.1 Surface Fluxes and Surface Temperatures

The spatial distribution of latent and sensible heat fluxes as well as the surface temperature is compared for the different model runs. Because of the importance of the correct description of the surface the results concerning the lower boundary layer are presented first. Figure 1 a–c illustrates a contour plot of the surface temperature in the model domain for the Fog-DWD run (Fig. 1a) and the Fog-Modified run (Fig. 1b) at 03 UTC. This time is chosen because the fog event calculated by the COSMO-FOG has the highest extension. In both cases, cooler temperatures between 0°C and 5°C are calculated within the valley of the river Rhine (see the circle in Fig.1a, b).The surface temperature calculated with

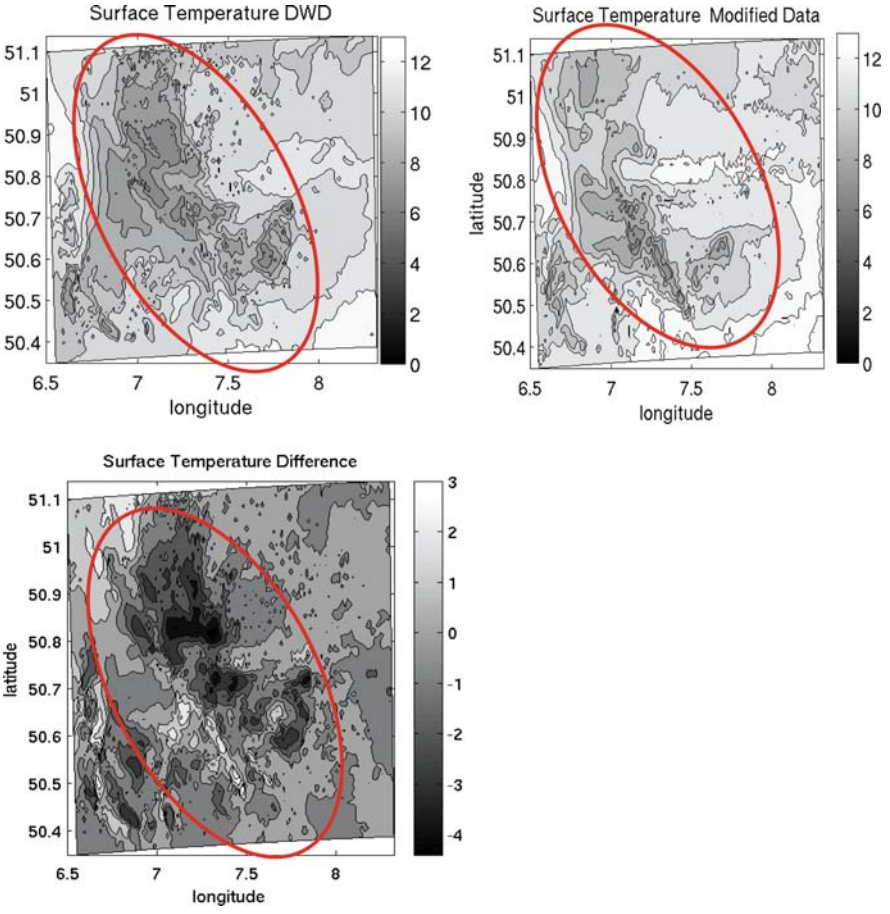

Fig. 1 **a–c**: Horizontal contour plot of the surface temperature in the whole model domain for (**a**) Fog-DWD run, (**b**) Fog-Modified run and (**c**) the calculated differences

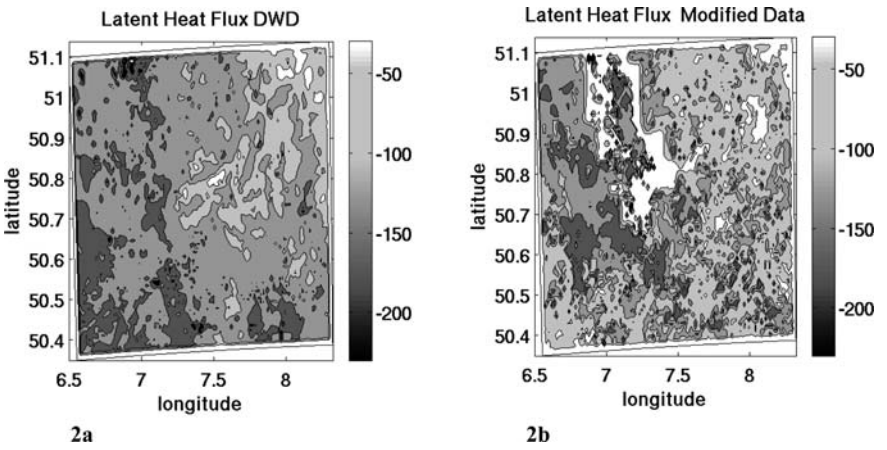

Fig. 2 **a–b**: Horizontal contour plot of latent heat flux in the whole model domain at 03 UTC. (**a**) Fog-DWD run, (**b**) Fog-modified run

the DWD-data is more cooler in this domain. In terms of quantification, the difference (Difference = Fog-DWD minus Fog-Modified) was calculated (Fig. 1c). The temperature varies in a range of – 4°C and + 3°C with the biggest discrepancy in the valley of the river Rhine.

In Fig. 2 the results for the calculated latent heat fluxes at 03 UTC are shown. The spatial patterns of the results for both model runs (Fig. 2 a–b) are different. The results for the simulation with the modified data (Fig. 2b) have a very patchy structure and thus, they are more heterogeneous. In both cases, the values of the fluxes have lower values in the western part of the model domain. The differences are calculated in the same way as for the temperature. The values differ in a range of – 100 W/m^2 and + 20 W/m^2. Figure 3 a–billustrates the results for sensible heat

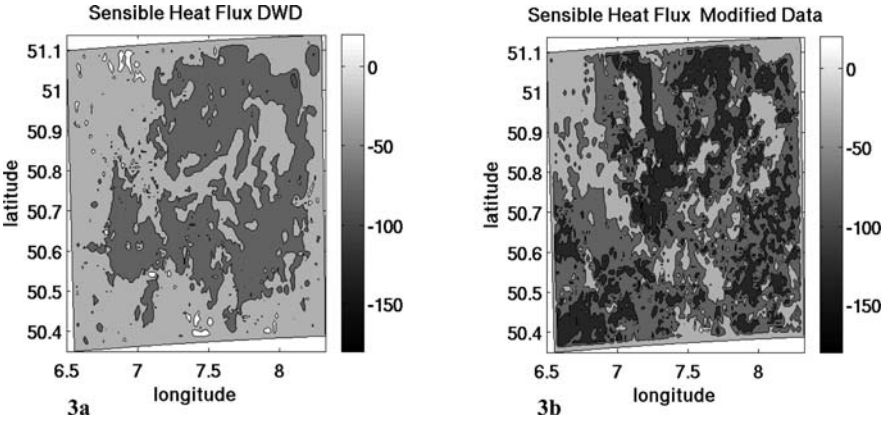

Fig. 3 **a–b**: Horizontal contour plot of sensible heat flux in the whole model domain. (**a**) Fog-DWD run, (**b**) Fog-modified run

fluxes. Similar characteristics as for the latent heat flux are observed. The structure of the results for the sensible heat fluxes calculated with FOG-DWD are more heterogeneous (Fig. 3b) and the values are lower than the values calculated with FOG-Modified. The differences vary in range between – 50 W/m^2 and + 100 W/m^2.

6.2 Simulation of Fog

The next step is to verify whether the observed differences in the results concerning the surface have an influence on the simulation of fog events. Figure 4a–c presents the calculated visibilities in 2 m height at 03 UTC. The variable visibility is chosen

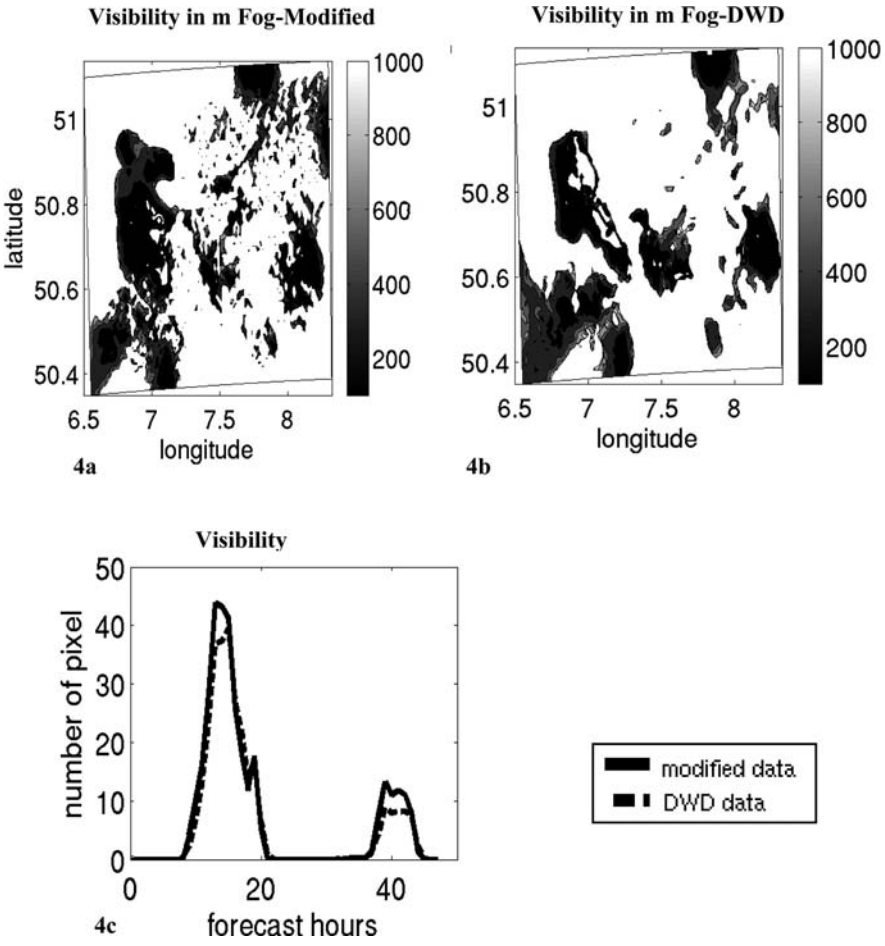

Fig. 4 **a–c**: Horizontal contour plot of visibility the whole model domain. (**a**) Fog-Modified run, (**b**) Fog-DWD run and (**c**) comparison of the number of pixel with visibility values less than 1000 m in percent

in accordance to the definition of fog where visibility values below 1000 m will be identified as foggy period. Parts of the model area are covered with fog and the heterogeneous structure of fog is also apparent. The lowest values of calculated visibility occur in both cases in the western part of the model domain. The same observations done by the surface fluxes match to the results of the visibility. Model results for the calculation with the modified data set is more heterogeneous, the fog is more dense and visibility values are lower. Most differences are observed in the center of the model domain (Fig. 4b). To quantify the differences, the number of pixels with visibility values less than 1000 m are calculated. Figure 4c depicts the time evolution of the identified pixels with visibility values below 1000 m in percent. Differences vary between 5 and 8%. Note, that just nearly 50% of the model domain is covered with fog. In Fig. 5a–b, a comparison with measurements is given. At the two weather stations Cologne and Bad Marienberg, fog were observed between 20 UTC and 06 UTC (correspond to the forecast hours 08–16). The curves of the model simulations done with the modified data agree quite well with the observations. In Cologne, the model run with the modified data reproduce the fog event. The minimum values of the observed visibility and the model results agree quite well, but the duration of the modelled fog event is too short. Even though the simulated duration by the run with DWD data fit quite well with the observations, the calculated visibility is much too high (Fig. 5a). Although in the case of Bad Marienberg both

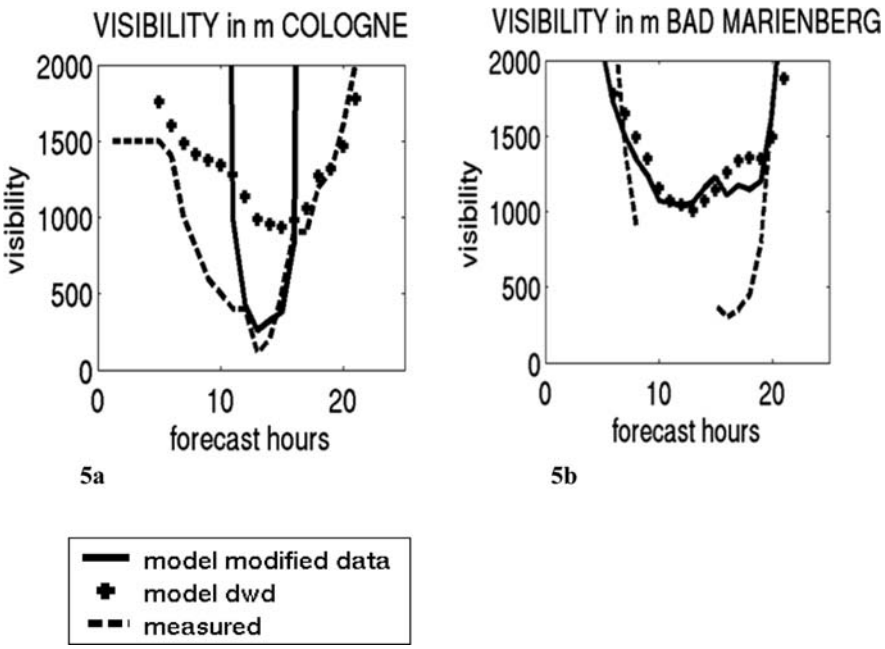

Fig. 5 **a–b**: Comparison of visibility observations and simulated visibility values. (**a**) weather station Cologne and (**b**) weather station Bad Marienberg

model runs Fog-DWD and Fog-Modified agree quite well with the observed fog duration, but the calculated visibility is too low.

Among the scientific community some debate exist if the soil or the vegetation cover has an higher affect on the atmospheric boundary layer (Braun and Schädler, 2005). For this purpose two more sensitivity studies are performed:

1. a study with homogeneous soil classes for the whole model domain, and
2. a study with homogeneous vegetation classes for the whole model domain.

First results of the study with the homogeneous soil types point out that the calculated results concerning the fog event differ not very much from each other. Although the calculated fluxes are sensitive to the different soil classes, the modelled liquid water content and the visibility differ nor very much.

The study with homogenous vegetation classes underlines the high impact of vegetation on the formation of fog. The calculated fluxes and the liquid water content together with the visibility show higher sensitivity concerning the land use types. As expected highest impact have the classes grass and farming in contrast to forest types.

7 Conclusion

This paper presents a study about the influence of the surface layer on the development of fog formation. Hence, a new detailed description of the earth surface has been developed and has been implemented into the numerical three dimensional fog model COSMO-FOG. For the purpose of comparison, sensitivity studies with two different data sets describing the surface have been performed. The new implemented surface parameters have an impact on both latent and sensible heat fluxes and the surface temperature. Differences of latent heat flux are in the range of + 100 W/m^2 and + 20 W/m^2 and of sensible heat flux in the range of – 50 W/m^2 and + 100 W/m^2. Differences of temperature are between – 3°C and + 4°C. As the changed surface fluxes, the fog formation is also influenced by the modified surface description. The characteristic spatial patterns are similar, but simulation results with the modified parameters are more heterogeneous. For the shown example run, the simulation with the new modified data set produces a more realistic fog period at the two available weather stations. One assumption is that the higher number of soil classes lead to the shown differences because the characteristics of soil, e.g., heat- and moisture capacity varies in connection to the soil classes. Thus, the exchange processes between surface and atmosphere leading to fog are influenced. But our sensitivity study with the homogenous soil and vegetation classes lead to the assumption that vegetation is more important concerning the life cycle of a fog event However, for final conclusions some more studies have to be performed.

References

André JC, Goutorbe JP, Perrier A (1986) HAPEX-MOBILHY: A hydrologic atmospheric experiment for the study of water budget and evaporation flux at climatic scale. Bulletin of the American Meteorological Society 67: 138–144.

Avissar R, Pielke RA (1989) A parameterization of heterogenous land surfaces for atmospheric numerical model and its impact on regional meteorology. Monthly Weather Review 117: 2113–2136.

Bergot T, Guedalia D (1994) Numerical forecasting of radiation Fog. Part I: Numerical model and sensitivity Tests, Monthly Weather Review 122: 1218–1230.

Boone A, Habets F, Noilhan J, Clark D, Dirmeyer P, Fox S, Gusev, Y, Hadeland I, Koster R, Lohmann D, Mahanama S, Mitchell K, Nasonova O, Niu G-Y, Pitman A, Polcher J, Shmakin, AB, Tanaka K, VAN DEN Hurk B, Verant S, Verseghy D, Viterbo P, Yang Z-L (2004) The Rhone-aggregation land surface scheme intercomparison project: An overview. Journal of Climate 17: 187–208.

Bott A, Sievers U, Zdunkowski W (1990) A radiation fog model with detailed treatment of the interaction between radiative transfer and fog microphsysics. Journal of the Atmospheric Sciences 47: 2153–2166.

Bott A, Sievers U, Zdunkowski W (1996) A numerical model of the cloud -topped planetary boundary layer: Radiation, turbulence and spectral microphysics in marine stratus. Q.J.R. Meteorological Society 122: 635–667.

Bott A, Trautmann T (2002) PAFOG – a new efficient forecast model of radiation fog and low-level stratiform clouds. Atmospheric Research 64: 191–203.

Braun FJ, Schädler G (2005) Comparison of soil hydraulic paramterizations for mesoscale meteorological models. Journal of Applied Meteorology 44: 1116–1132.

Clark P, Hopwood W (2001) One-dimensional site specific forecasting of radiation fog. Part 1: Model formulation and idealiuzed sensitivity studies, Meteorological Applications 8: 279–286.

Dickinson R (1984) Modelling evapotranspiration for the threedimensional global climate models. In: Climate Processes and Climate Sensitivity, Geophysical Monograph, vol. 29. Maurice Ewing Volume 5, pp 58–72.

Doms G, Förstner J, Heise E, Herzog H-J, Raschendorfer M, Schrodin R, Reinhardt T, Vogel G (2005) A description of the Nonhydrostatic Regional Model LM. Part II: Physical Parameterization.

Doms G, Schättler U (2002) The Non-Hydrostatic Limited-Area Modell LM (Lokal-Modell) of DWD – Part I: Scientific Documentation, DWD.

Duynkerke PG (1991) Radiation fog. A comparison of model simulation with detailed observations. Monthly Weather Review 119: 324–341.

EEA (European Environment Agency, Copenhagen (2004) CORINE Land Cover (CLC); Umweltbundesamt, DLR-DFD. http://dataservice.eea.eu.int/dataservice/.

Fisher EL, Caplan P (1963) An experiment in numerical prediction of fog and stratus, Journal of the Atmospheric Sciences 20: 425–437.

Forkel R, Sievers U; Zdunkowski W (1987) Fog modelling with a new treatment of the chemical equilibrium condition. Beitr. Phys. Atmos. 60: 340–360.

Geologischer Dienst NRW (2002) Bodenkarte BK 50. Blätter L4908, L4910, L4912, L4914, L5108,L5110, L5112, L5114, L5308.

Geologische Landesamt Rheinland-Pfalz (2002) Bodenkarte BK 50.

Gulteppe I, Tardi R, Michaelides SC, Cermak J, Bott A, Bendix J, Müller M, Pagwski M, Hansen B, Ellrod G, Jacobs W, Toth G, Cober SG (2007) Fog research: A review of past achievements and future perspectives. Journal of Pure and Applied Geophysics 10.

Henderson-Sellers A, Yang ZL, Dickinson RE (1993) The Project for Intercomparison of Land surface Parametrization schemes. Bulletin of the American Meteorological Society 74: 1335–1349.

Kessler E (1969) On the distribution and continuity of water substance in atmospheric circulation. Meteorological Monographs, No. 32, American Meteorological Society, pp 84.

Louis J-F (1979) A parametric model of vertical eddy fluxes in the atmosphere. Boundary-Layer Meteorological 17: 187–202.
Majewski D, Ritter B (2002) Das Globalmodell GME. Promet 27(3/4): 111–122.
Masbou M, Bott A (2005) Fog Forecasting at high resolution with the Lokal-Modell of German Weather Service. World Weather Research Programme's Symposium on Nowcasting and Very Short Range Forecasting. Toulouse, France, September 5–9. 6.24.
Müller M (2006) Numerical Simulation of fog and radiation in complex terrain, Ph.D. thesis, University of Basel.
Pielke RA (1984) Mesoscale Meteorological Modelling. Academic Press, Orlando
Pruppacher HR, Klett JD (1997) Microphysics of cloud and precipitation. Kluwer Academic Publishers, London, pp 954.
Schädler G (1990) Numerische Simulationen zur Wechselwirkung zwischen Landoberflächen und atmosphärischer Grenzschicht. Wissenschaftliche Berichte des Instituts für Meteorologie und Klimaforschung der Universität Karlsruhe (TH).
Schrodin R, Doms G, Edelmann W, Fischer E, Frey B, Gertz M, Gube-Lenhardt A, Hanisch T, Heise E, Link A, Majewski D, Prohl P, Ritter B, Schättler U (1995) German Weather Service, Research Department, Documentation of the EM/DM-System.
Schrodin R, Heise E (2001) The Multi-Layer Version of the DWD Soil Model TERRA_ML, Technical Report, abrufbar unter: http://www.cosmo-model.org
Sellers PJ et al. (1997) BOREAS: Experiment overview, scientific results, and future directions. Journal of Geophysical Research 102(D24): 28, 731–28, 769.
Sellers PJ, Hall FG, Asrar G, Strebel DE, Murphy RE (1992) An overview of the First International Satellite Land Surface Climatology Project (ISLSCP) Field Experiments (FIFE). Journal of Geophysical Research 97D: 18345–18373.
Seuffert G, Gross P, Simmer C, Wood EF (2002) The influence of hydrologic modeling on predicted local weather: Two-way coupling of a mesoscale weather prediction model and a land surface hydrology model, Journal of Hydrometerology 3: 505–523.
Shuttle Radar Topography Mission (SRTM): DLR abrufbar unter (2000) http://www.dlr.de/caf.
Steppeler J, Doms, Schättler U, Blitzer HW, Gassman A, Damrath U, Gregoric G (2003) Meso-gamma scale forecast using the non-hydrostatic model LM. Meteorological and Atmospheric Physics 82: 75–76.
Stull RB (1988) An Introduction to Boundary Layer Meteorology, Kluwer Academic Publisher.
Teixeira and Mirinda (2001) Fog prediction at Lisbon airport using a one dimensional boundary layer model, Meteorological Applications 8, 497–505.
USGS (1997) GTOPO30 Digital elevation model, Eros data center, US Geological Survey, Soiux Falls, SD. http://edcaac.usgs.gov./gtopo30.asp.
von Glasow R, Bott A (1999) Interaction of radiation fog with tall vegetation. Atmospheric Environment 33: 1333–1346.
Winterrath T, Bott A (2001) Chemifog_V-A model to simulate Radiation Fogs and their interaction with vegetation and chemistry. Water, Air, and Soil Pollution: Focus1: 373–380.
WMO (1992) International Meteorological Vocabulary, World Meteorological Organization, 182, Geneva, Switzerland, 2nd edn.

Influence of Drainage Parameterization and Precipitation Analysis on Discharge Simulation in the Sieg River Catchment

René Graßelt, Kirsten Warrach-Sagi, Felix Ament, and Clemens Simmer

Abstract The aim of this study is to develop a realistic model of water transport for the Sieg river catchment, a tributary of the Rhine river, by considering topography, land use, and the distribution of soil moisture. Realistic modelling of discharge in catchment areas has been provided for different models before. However, this approach was applied here for the first time for the COSMO model based on its own SVAT-module. The generation of runoff in a catchment area depends on multiple parameters in the atmosphere and in the ground, particularly on their temporal and spatial covariabilities. The vertical and horizontal movement of water at the land surface and in the soil depends on soil texture, soil moisture, topography, plant cover, and atmospheric conditions. We applied the SVAT-module Terra of the COSMO model in a stand alone version in order to calculate the water- and energy fluxes at the land surface. Terra was coupled to a river routing system to transform the runoff generated at each grid cell into discharge and to study Terras dependence on the spatial and temporal characteristics of the input data and parameters. This allowed the results generated by the model to be compared to actual measurements of discharge recorded at river gauges. These results were then used to produce forecasts. In addition, the multi-layer-version of Terra was used to allow for changes in spatial resolution with low computational costs. A grid-resolution of one kilometre was applied to calculate the discharge for the river catchment of the Sieg. To allow for the use of available data from Geologisches Landesamt Nordrhein-Westfalen and Rheinland Pfalz, the operational soil texture distribution of Terra was replaced. In addition, a new one kilometer resolution topography was implemented based on NASA-SRTM. Currently, the meteorological forcing for all driving parameters is provided by COSMO, except precipitation which is obtained from the REGNIE data set (one kilometer gridded daily rain accumulation derived operationally by DWD).

R. Graßelt (✉)
Meteorological Institute, University of Bonn, Auf dem Huegel 20, 53121 Bonn, Germany
e-mail: grasselt@uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_7,

Keywords SVAT-module · Discharge · Precipitation · Topography · Soil characteristics

1 Introduction

1.1 Introduction

Interactions between the atmosphere and the land surface strongly impact weather and climate. Currently, in numerical weather prediction (NWP) the interface between the atmosphere, soil and vegetation is described by land surface parametrizations (LSP). The primary purpose of these models is to separate the downward solar and longwave radiation into sensible, latent and ground heat fluxes and upward longwave radiation, rather than to predict discharge. In addition, hydrological models are principally able to advance numerical weather prediction, e.g. considering by mass conservation based on runoff observations. Coupling with hydrology can be realized by extensions to the LSP contained in SVAT-modules (soil-vegetation-atmosphere-transfer-schemes). Alternatively, existing hydrological models can be coupled with NWP models. Hydrological models describe processes in the soil which are rarely or only rudimentarily considered in NWP. One example is the insufficient consideration of horizontal transport processes in NWP models. On the other hand, SVATs usually focus only on determining the energy fluxes and conservations correctly (Warrach et al. 2002).

Lohmann (1996) stressed the importance of developing more realistic LSPs, especially for more accurate flood forecasting. The weather forecast model COSMO (Consortium for Small-Scale Modelling) originally developed by the German Weather Service (DWD) and now used in the international COSMO community does not allow flood forecasting for river systems. Lohmann (1996) also noted that LSP schemes do not consider the temporal movement of water in a horizontal direction. Soil, vegetation and topography usually change on small spatial scales. NWP models use a maximium operational spatial resolution of 2.8 km (e.g. COSMO-DE model). Thus an additional challenge is the realistic modelling of hydrological processes as well as the downscaling of atmosphere variables to the scales required by hydrology.

2 Model and Data

2.1 Model

The Consortium for Small-Scale Modelling (COSMO) represents an international user group which aims to develop and advance an operational meso-γ to convective scale weather prediction system. The COSMO model (Doms and Schättler

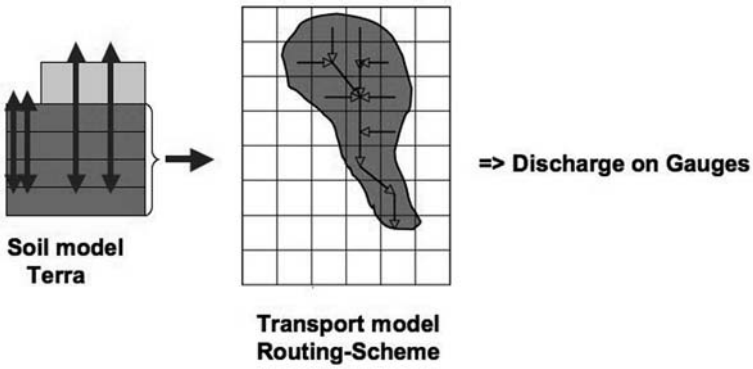

Fig. 1 Model configuration

2002), use by for various European weather services (e.g. German Weather Service, MeteoSwiss, Institute of Meteorology and Water Management Poland) encompasses regional scale weather models with different versions and spatial resolutions. For this study we applied the soil model Terra of the LM (Lokal-Modell) Version 3.16, an older version of the COSMO model, which describes exchange processes in the soil and between the land-surface and the atmosphere. The multi-layer soil model Terra- Stand-Alone is a decoupled version of the original Soil-Vegetation-Atmosphere-Transfer-Scheme (SVAT) of the LM (Schrodin and Heise 2001). A decoupled soil model enables the SVAT-module to be upgraded without additional computation costs. Due to the atmosphere part of the simulation. Terra uses meteorological forcing data to calculate, the surface energy and water fluxes, the soil moisture, soil temperature and runoff for each timestep based on soil- and plant-fields information. Since Terra only calculates runoff for each gridcell, the version of Terra coupled to the routing scheme by Lohmann et al. (1996) is applied to calculate river discharge at the gauging-stations. Both models, Terra and the routing scheme, work at the same spatial resolution (Fig.1), of 1 km.

2.2 Data

We present the results for the Sieg river, a tributary of the Rhine river. The Sieg watershed covers 2.832 km^2 and is located in western Germany. The landscape is characterized by a low mountain range in the middle and eastern part of the catchment and lowlands in the western part near the Rhine (Fig. 2). This watershed has been the focus of past research (e.g. Seuffert et al. 2002 coupled an older LM-version to the hydrological model TOPLATS in a 2-way mode). The coupled Terra - routing model was applied to this catchment requiring a set of input parameters and data described in this section. We distinguished between the soil fields and the meteorological forcing. The temporally constant soil fields contain information

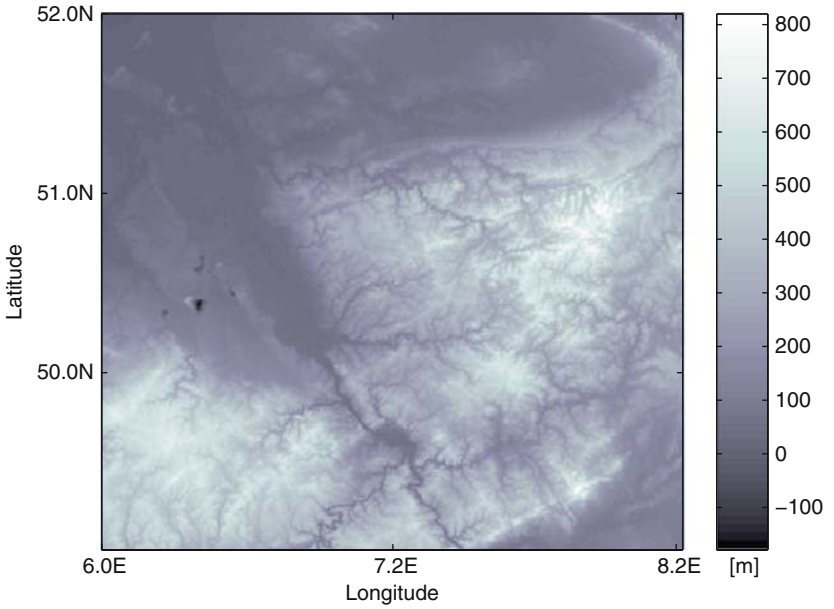

Fig. 2 Nasa-SRTM Dataset (Shuttle Radar Topography Mission, 2000) with a original resolution of 5 arc min

about vegetation, soil type and topography. These parameters are used as constant information in our calculations. For Terra-Stand-Alone the meteorological forcing variates in a high temporal resolution, thus atmosphere forcing was updated for each timestep of 1 min.

2.2.1 Constant Soil Fields

Our model configuration has a 1 × 1 km resolution and covers a model area of 142 × 101 km. This resolution requires some changes of the default constant soil fields provided with Terra. The COSMO model is currently able to run at resolutions up to 2.8 × 2.8 km (COSMO-DE). However, the LM 3.16 only runs with a resolution of 7 × 7 km. For the constant soil fields (Table. 1) information about topography, vegetation and soil texture is required. We adopted the surface elevation from the NASA-SRTM (Shuttle Radar Topography Mission, Farr et al. 2007) dataset. Figure 2 shows (in its original spatial resolution) the dataset. A 1 km spatial resolution is applied for the operational model area (Fig. 3).

The estimation of different vegetation parameters (Table.1) is based on the CORINE Land Cover (CLC) dataset (EEA, 2000) in combination with a reference table (Table. 2). The dataset has been build from satellite observations with a spatial resolution of approximately 100 m. For more information about the vegetation aspects see Chapter ‘The Impact of Landform Structure on the Formation of Fog’ by Alberts et al.

Table 1 Constant soil fields for Terra-Stand-Alone

Variable	Dataset	Annotation
Topography	NASASRTM	Elevation
Soil texture	BK 50	Ten classes
Plant cover min/max	CORINE	Minimum of plant cover in winter, maximum plant Cover during the vegetation period
Leaf area index min/max	CORINE	
Root depth	CORINE	In meter
Roughness length	CORINE	In meter

The operational soil type input used in the COSMO model was provided by the FAO (Food and Agricultural Organisation of UNO) from the global DSMW dataset (Digital Soil MAP of the World). Due to the coarse spatial resolution of 5 arc min (approximately 10 km) only 5 classes of soil type are distinguished. A high

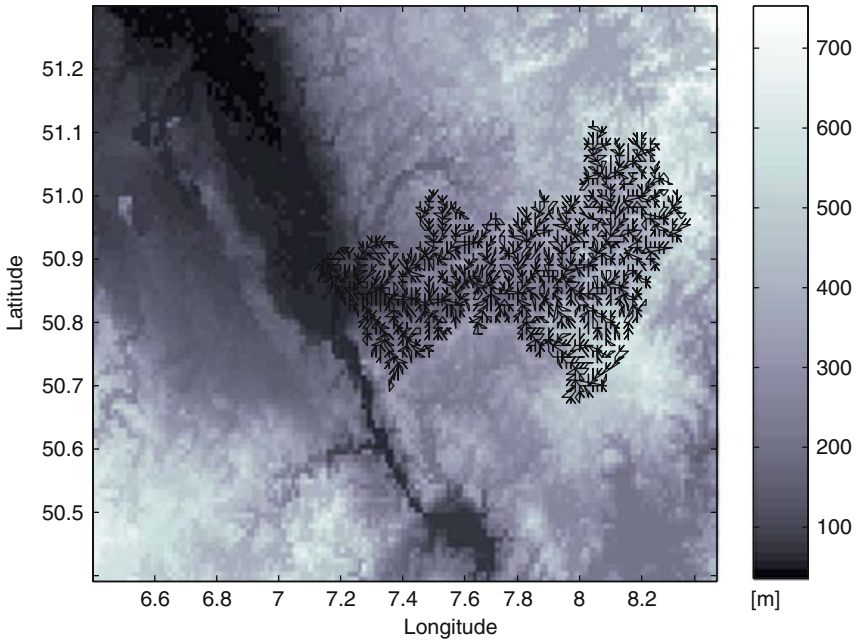

Fig. 3 Model domain topography in 1 km spatial resolution (grid cells in the Sieg catchment are descripted with the directions for horizontal transport processes)

Table 2 Look-up table of land-use dependent vegetation parameter adopted from the operational table of DWD

Variable	Town	Grass	Crop	Shrub	Decid. forest	Confi. forest	Mixed forest
Roughness length (m)z_0	1.0	0.03	0.1	0.1	1.0	1.0	1.0
Plant cover f_{veg} min/max	0.05/0.1	1/1	0.5/1	0.1/0.5	1/1	0.5/1	0.5/1
LAI f_{LAI} min/max	0.1/4	0.5/4	0.2/4	0.1/3	0/6	8/8	2.25/7
Root depth (m) z_{root}	0.3	0.15	0.3	0.4	0.8	0.8	0.8

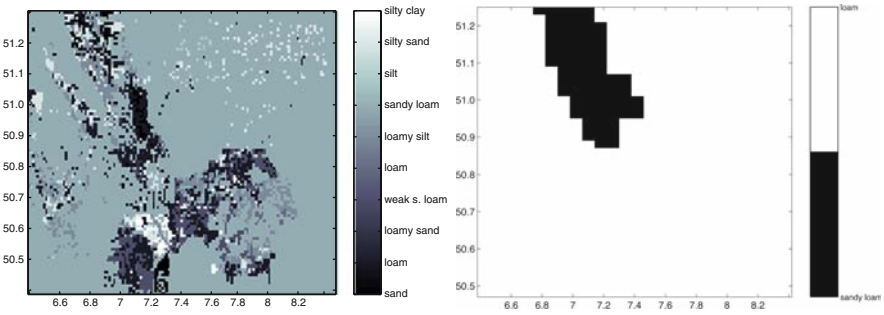

Fig. 4 New soil characteristics (BK 50) for the model domain (*left*) in comparison with the information of the FAO (Food and Agricultural Organisation of UNO) dataset (*right*)

spatial resolution dataset similar to CORINE for vegetation does not exist for Germany or Europe. However, the local BK 50 dataset from Geologisches Landesamt Nordrhein-Westfalen and Rheinland Pfalz was available for the Sieg catchment and its surroundings. Ten classes of soil types could be considered in the soil model. Comparing the DSMW 10 km dataset with the high resolution 1 km BK 50 dataset, it becomes obvious that the high resolution maps have significantly more variability (Fig. 4).

The routing scheme requires a flow direction for each grid cell within the Sieg catchment as input. This information is computed from the digital elevation model NASA-SRTM with the help of a geographic information system. The direction of the maximum hill slope of each grid box determines the flow direction for drainage.

2.2.2 Meteorological Forcing

Precipitation is the dominant factor for calculating runoff. The spatial variability of precipitation within the catchment area influences all runoff characteristics including timing and amplitude; thus convective rain will lead to different results compared to stratiform rain. Thus precipitation measurements should improve in high spatial and temporal resolution discharge estimations. We studied the effects of two different precipitation datasets on the modelled discharge. The first

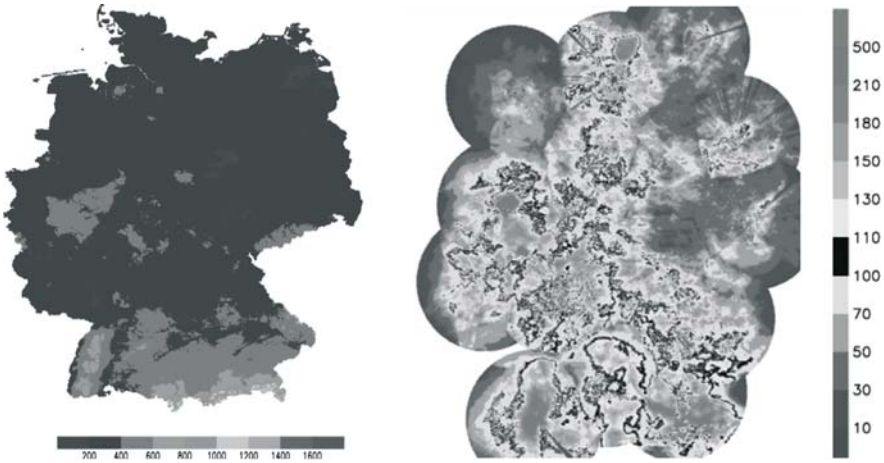

Fig. 5 Accumulated precipitation (mm) 03/05 – 09/05 by REGNIE (*left*) and accumulated precipitation (mm) 04.04.05 – 12.05.05 by RADOLAN RW (*right*)

dataset is REGNIE, a product of the DWD (Fig. 5). Its temporal resolution is 24 h which its spatial resolution is 1 km. REGNIE is a gridded dataset which is produced from original station observations by interpolation to a regular 1 km grid. The interpolation considers local reference values from the years 1961 to 1990. Local reference values constitute for example, the height of the measuring station, longitude and latitude, etc.

The second precipitation dataset in our study is RADOLON (Radar Online Calibration) (Bartels et al. 2004). This dataset has been operational since June 2005 and is also a product of DWD. We use the product RADOLAN RW which is a combination of station measurements and radar data designed for the use in high quality short range precipitation forecasting. In comparison with REGNIE, RADOLOAN RW is available with a temporal resolution of 1 h. Thus convective precipitation with their spatial and temporal distribution can influence the runoff calculation. Nevertheless, studies have shown that RADOLAN RW has some problems related to quality (Simmer et al. 2005). Due to combinations of radar errors with the interpolation scheme RADOLAN RW can largely overestimate or underestimate the amount of precipitation by strong rainfall events (Fig. 5). A reason for this problem is probably the creation of the radar composite. The interference band of two radars with their maximum reflectivity is always selected. So that the radar with a local low elevation angle will be choosen.

Terra-Stand-Alone also needs input information by the short- and longwave radiation, temperature, specific humidity, wind and pressure. These variables are provided by the LM-analyses data with a temporal resolution of 1 h. The LM-analyses data have a spatial resolution of 7 km. The meteorological forcing data are linearly interpolated to the 1 km grid.

3 Soil Moisture Dynamics and Runoff Generation

The water budget of the soil model layers depends on the conditions at the upper and the lower boundaries of the soil model, on the water extraction by evapotranspiration, on gravitational and capillary transports, and on the runoff formation.

The flux of soil water in unsaturated soil is the sum of drainage and diffusion. Horizontal transports are neglected due to the coarse horizontal resolution and therefore the flux of soil water F_η can be written as a one-dimensional Darcy equation. (e.g. Ament 2006 or Dingman, 2002)

$$F_\eta = K(\eta) + D(\eta)\frac{\partial \eta}{\partial z} \tag{1}$$

The hydraulic conductivity K and hydraulic diffusivity D depend both on the soil moisture η and the soil characteristics pore volume η_{PV} and air dryness point η_{ADP}. Both functions are parameterized in Terra by the exponential laws of Rijtema (1969):

$$K(\eta) = K_{0\exp}\left(K_1 \frac{\eta_{PV} - \eta}{\eta_{PV} - \eta_{ADP}}\right) \tag{2}$$

$$D(\eta) = D_{0\exp}\left(D_1 \frac{\eta_{PV} - \eta}{\eta_{PV} - \eta_{ADP}}\right) \tag{3}$$

The coefficients K_0, K_1, D_0 and D_1 as well the pore volume η_{PV} and the soil moisture at air dryness point η_{ADP} depend on the soil type (for the reference table see Schrodin 1995). These approaches are not the conventional parametrizations used in other meteorological models. Different authors developed various approaches, that follow a linear relationship (e.g. Van Genuchten 1980, Brooks and Corey 1988, Rawls et al. 1993). In Terra-Stand-Alone a linear formulation by Campbell (1974) was for the first time applied to test its influence on runoff and soil moisture calculation. Hydraulic conductivity and diffusivity dependent on soil moisture and soil characteristics are parameterized as follows:

$$K(\eta) = K_1 . \left(\frac{\eta}{\eta_{PV}}\right)^c, \tag{4}$$

$$C = 2 \cdot b + 3 \tag{5}$$

$$D(\eta) = -b \cdot \eta_{ADP} \cdot K_1 \cdot^{-b-3} \cdot \eta^{b+2} \tag{6}$$

The Brooks and Corey parameter b is an additional soil characteristic parameter for the calculation of hydraulic conductivity and diffusivity which did not occur in the approach cited above. The soil parameter b can be estimated by empirical and measured soil properties of various authors (e.g. Clapp and Hornberger 1978 or Pielke 1984).

The runoff from the individual soil layers determines the base flow component in Terra, which is an important component for the formation of runoff. In addition hydraulic conductivity and diffusivity described above influence runoff generation. Runoff from a soil layer k occurs if the total water content η_k of the layer exceeds field capacity η_{FC} and if the divergence of the fluxes (1) in this layer is negative. In this case, Terra parametrizes the runoff from soil layer according to:

$$R_k = -\frac{\eta_k - \eta_{FC}}{\eta_{PV} - \eta_{FC}} \left(\frac{\partial F}{\partial z}\right)_k \Delta z_k \tag{7}$$

The lower boundary condition at the bottom of the deep layer in TERRA-ML is formulated by a free drainage boundary condition. Soil water can drain from the lowest layer, but the flux due to diffusion is neglected. This means, ground water cannot wetten the soil by capillary rise from below.

4 Results

This chapter shows sensitivity studies using the different soil water transport parameterization and precipitation forcings. In order to test how the generation of discharge is influenced by a new vertical soil water parameterization. We focus on two gauging-stations of the Sieg catchment; Siegburg-Kaldauen and Menden. Both stations are located on the underflow of the river Sieg. Firstly, the modelled discharge is presented for the gauging-station Menden. We started a model run from April to September 2005 and a second run from September 2004 to Febuary 2005. Precipitation input from the REGNIE dataset was selected. For the spring and summer months the use of the Rijtema equations results in a stronger correlation between measured and modelled discharge compared to the use of Campbell equations (Fig. 6). Although the Rijtema formulas reproduce better peak runoff comparly to Campbell by precipitation events. Nevertheless the base flow in both simulations represent an unrealistic evolution.

The second model run (Fig. 7) presents the discharge for the same station during the autumn/winter season. Although the Rijtema formulas simulate a better peak runoff comparly to the Campbell equations for strong precipitation events, the base flow shows a more realistic result for the Campbell formulas. During both model runs, the differences in accumulated mass between measured and the modelled Campbell approach are less than with the Rijtema equations (Fig. 8).

Figure 9 presents the temporal evolution of simulated soil moisture in Terra-Stand-Alone in different layers by the impact of precipitation, in order to better understand the different results of both schemes. Overall, the amount of soil water content is lower when using the exponential soil water equations of Rijtema. Using this method, a strong precipitation event leads to a stronger reaction of the soil moisture in the upper four layers, wich have a total thickness of 0.81 m. In the lower soil layers the water content is higher using the Campbell equation. We can explain the lower soil water content and the higher reaction of the Rijtema approach

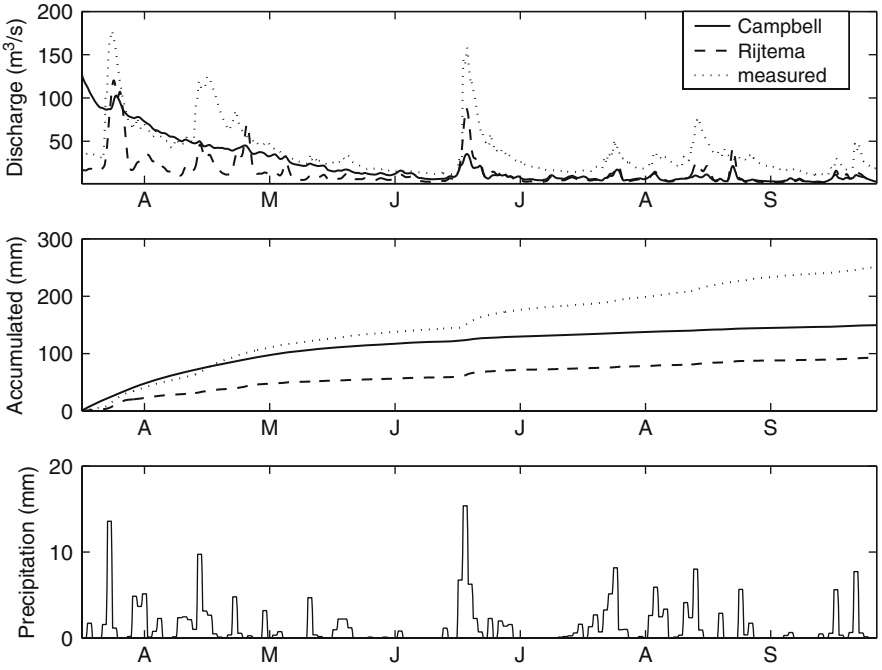

Fig. 6 Discharge for gauging-station Menden (*top*) and the accumulated mass (*middle*), *black* curve Campbell, *dashed* the parameterization by Rijtema, *doted* measured and precipitation input by REGNIE (*bottom*)

by strong precipitation events with the rapid saturation in the first layers and the subsequent reaction of the runoff from soil layers (Eq. 7). This approach produces more runoff in the first layers and a small vertical soil water transport to the lower layers. Furthermore, the drying out of Terra multilayer may explain the reduced base flow during the summer months. Studies have shown that drying of the lower layers increases during the summer months (Simmer et al. 2005). The main reason is the lower boundary condition at the bottom of the deep layer.

The comparison of the calculated water balance in Terra-Stand-Alone supports this result (Fig. 10). Strong differences show also the base flow (run_g). The base flow is higher using the Campbell equations and also a further argument for the more continuous base flow discharge in the winter season. Smaller differences occur between evapotranspiration (evapo.), surface runoff (run-s) and soil moisture (SM) compared to components of the water balance.

In order to evaluate the significance of precipitation information on runoff simulation we compare the results obtained with the use of RADOLAN RW. The discharge simulations (Fig. 11) show a higher correlation compared to the precipitation input by REGNIE. Both model approaches reproduce better peak runoff. Rijtema especially simulates by precipitation events a higher discharge. The use of RADOLAN RW data does, however, also lead to overestimation from July until

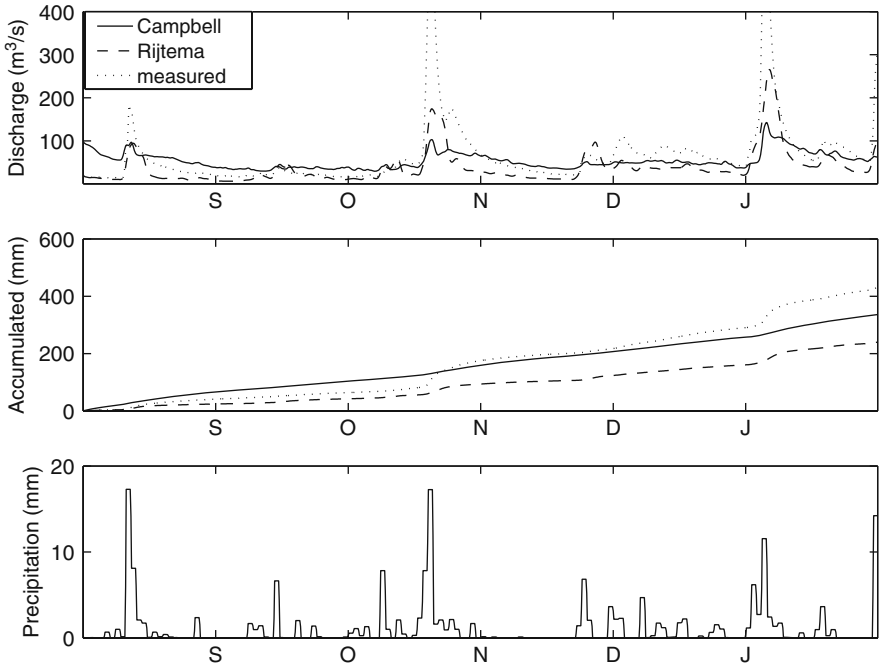

Fig. 7 Discharge for gauging-station Menden (*top*) and the accumulated mass (*middle*), *black* curve Campbell, *dashed* the parameterization by Rijtema, *doted* measured and precipitation input by REGNIE (*bottom*)

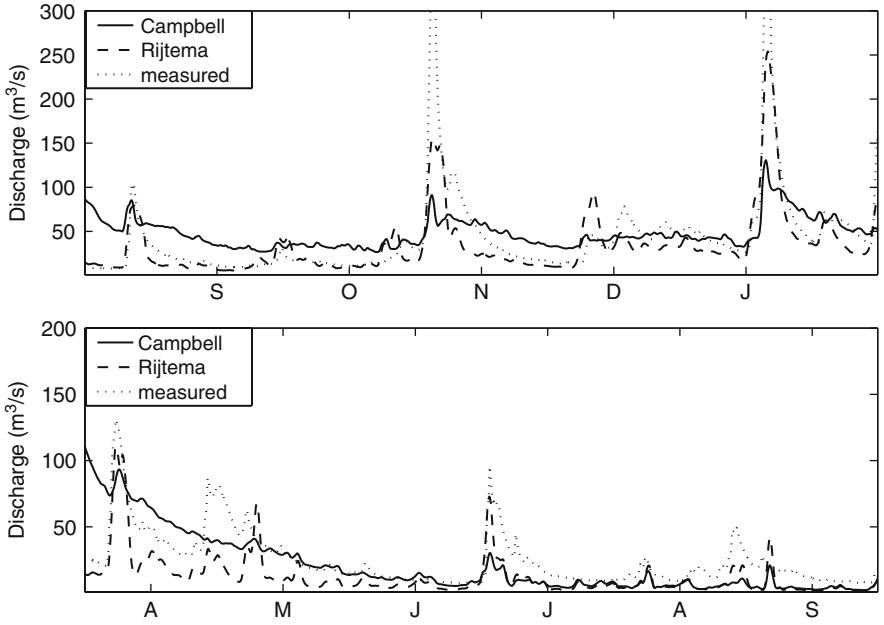

Fig. 8 Discharge for gauging-station Siegburg black curve Campbell, *dashed* the parametrization by Retijma, *doted* measured

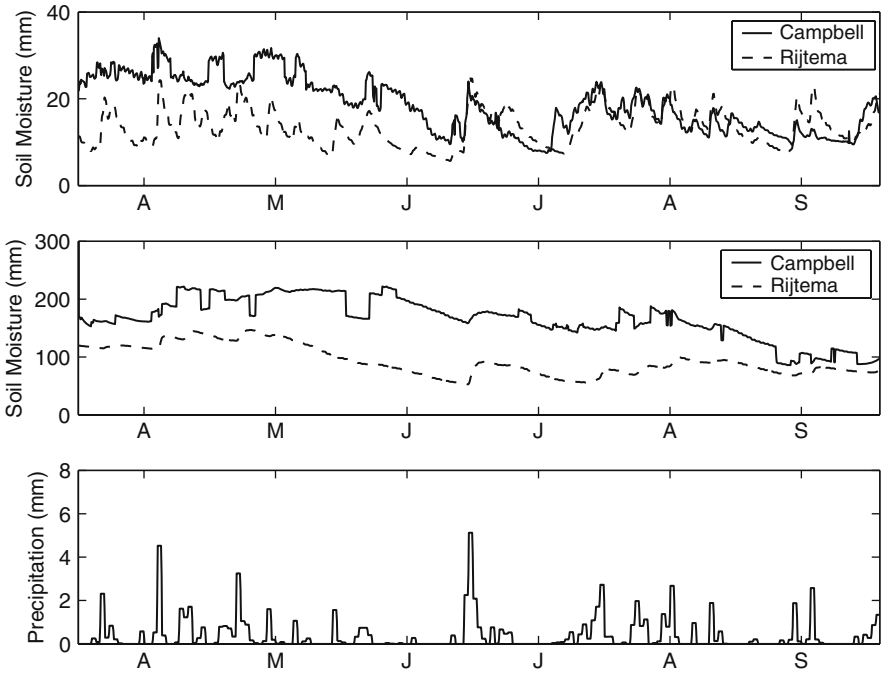

Fig. 9 Soil moisture for layer 1 – 4 thickness 0.27 m (*top*), layer 5 – 6 thickness 2.16 m (*middle*) and the precipitaion input by REGNIE (*bottom*)

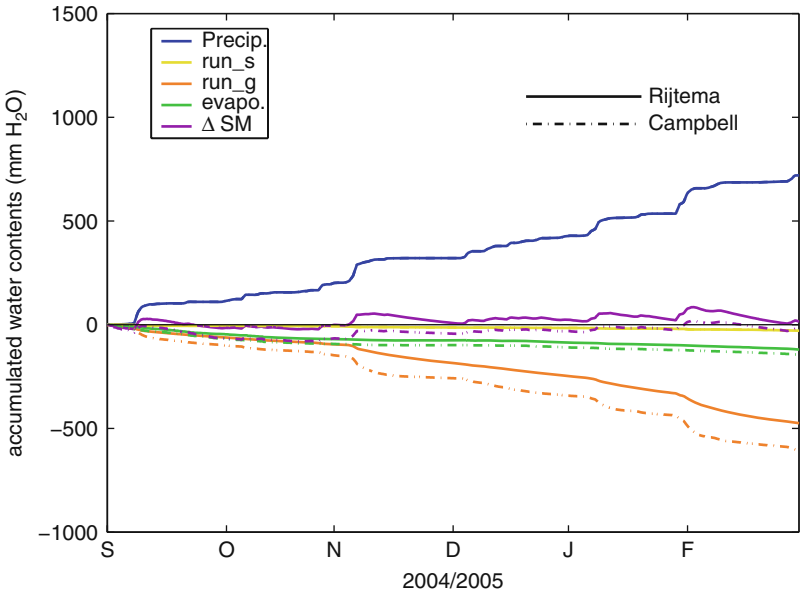

Fig. 10 Water balance for gauging-station Menden

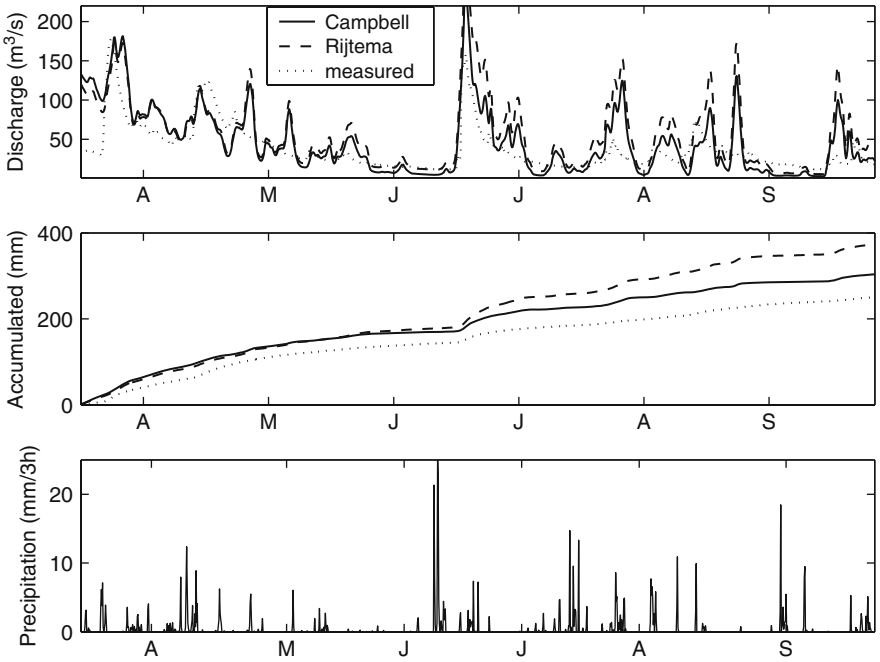

Fig. 11 Discharge for gauging-station Menden (*top*) and the accumulated mass (*middle*), *black* curve Campbell, *dashed* the parameterization by Rijtema, *doted* measured and precipitation input by RADOLAN RW (*bottom*)

September by strong precipitation events. Earlier studies have already shown that RADOLAN RW has some quality related problems (Simmer et al. 2005) (see also Sect. 2.2.2). RADOLAN RW shows a more realistic discharge for the Sieg river catchment. The accumulated mass of discharge shows, similary to the REGNIE run, a better result for the Campbell soil water parameterization. Although compared to REGNIE and the discharge measurements the accumulated mass with precipitation input by RADOLOAN RW simulates more discharge.

The daily temporal resolution of REGNIE and RADOLAN RW explains the differences in the discharge simulations. An uniform precipitation input with REGNIE in Terra-Stand-Alone simulates a high evapotranspiration and a too strong base flow. Especially the high soil water content modelled with the Campbell formulations lead to a high evapotranspiration. That result supports the drying out problem in Terra-Stand-Alone additionally. An uniform precipitation masks also the runoff peaks.

5 Conclusion

The coupling of Terra-Stand-Alone and a routing scheme allows horizontal transport processes in river systems to be considered. The results for the Sieg cachtment suggest that the applied changes in the vertical soil water parameterization

show small improvements of the modelled discharge. The Campbell approach shows a better correlation between accumulated measured mass and modelled mass. Due to its higher spatial resolution of rainfall data, RADOLAN RW gives more realistic results compared to the REGNIE dataset and it is concluded that a radar-based rainfall product is able to simulate the discharge for the catchment under consideration. For the future radar-based rainfall products calibrated by different ground measurements might improve the results for the introduced model approach. To further substantiate these results, we will analyze another catchment. We also plan to verify the new model setup with measurements of soil moisture produced by the DWD. An additional parameterization of a groundwater table for Terra-Stand-Alone should also help to improve runoff calculation for a river system. Up to now, the water table is kept constant in Terra. We are in the process of implementing a various ground water table based on the methodology by Stieglitz et al. (1997). We expect further improvements in runoff simulation and additon ways to validate the model.

Acknowledgements This project was founded by the Deutsche Forschungsgemeinschaft (DFG) within the Research Training Group 437: Landform a structured and variable boundary layer. We would like to thank all the people who have improved this work with their ideas and experience. We would also like to thank the DWD (German Weather Service) for their assistance and for provision of model (TERRA) and data.

References

Ament, F.: Energy and moisture exchange processes over heterogeneous land-surfaces in a weather prediction model. Dissertation University of Bonn, Bonn (2006)

Bartels, H., Weigl, E., Reich, T., Lang, P., Wagner, A., Kohler, O., Gerlach, N.: Routineverfahren zur Online-Aneichung der Radarniederschlagsdaten mit Hilfe von automatischen Bodenniederschlagsstationen. Abschlussbericht. Offenbach (2004)

Brooks, R. H., Corey, A. T.: Hydraulic properties of porous media. Hydrol. Paper. Colorado State University, (1988)

Campbell, G.S.: A simple method of determining unsaturated conductivity form moisture rentention data. Soil Science. 117, 311–314, (1974)

Clapp, R. B., Hornberger, G. M.: Empirical equations for some soil hydraulic properties. Water Resources Research. 14(4), 601–604, (1978)

Doms, G., Schättler, U.: A description of the nonhydrostatic regional model LM In: Part 1: Dynamics and Numerics, Offenbach (2002)

Dingman, S. L.: Physical hydrology, Second Edition. Prentice Hall, Upper Saddle River, NJ, (2002)

EEA: Corine Land Cover (CLC 90). European Environment Agency a http://dataservice.eea.eu.int/dataservice/, Copenhagen DK, (2000)

Van Genuchten, M. T.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soil. Soil Sience. 44, 892–898, (1980)

Farr, T. G., Rosen, P. A., Caro, E., Crippen, R., Duren, R., Hensley, S., Kobrick, M., Paller, M., Rodriguez, E., Roth, L., Seal, D., Shaffer, S., Shimada, J., Umland, J., Werner, M., Oskin, M., Burbank D., Alsdorf, D.: The shuttle radar topography mission. Reviews of Geophysics. 45 (2007) doi: 10.1029/2005RG000183

Lohmann, D., Nolte-Holube, R., Raschke E.: A large-scale horizontal model to be coupled to land surface parametrization schemes. Tellus. 48A, 708–721, (1996)

Lohmann, D.: Hydrologische Modellierung auf regionaler Skala. Dissertation University of Hamburg, Hamburg (1996)

Pielke, R. A.: Mesoscale meteorological modelling. Academic Press, Orlando (1984)

Rawls, W. J., Ahuja, L. R., Brakensiek, D. L., Shirmohammadi, A.: Infiltration and soil water movement. Handbook of hydrology, Chapter 5, D. R. Maidment Ed, McGraw Hill, New york, (1993)

Rijtema, P. E.: Soil moisture forecasting. Technical Report Nota 513. Instituut voor Cultuurtechniek en Waterhuishouding, Wageningen (1969)

Simmer, C., Ament, F., Graßelt, R., Girmes, R.: Abschlussbericht zum Forschungsprojekt 'Von Messdaten angetriebenes Wasserbilanzmodell. University of Bonn, Bonn (2005)

Schrodin, R., Heise, E.: The multi-layer version of the DWD soil model TERRA-LM. COSMO Technical Report No. 2. Offenbach (2001)

Schrodin, R.: Dokumentation des EM/DM Systems. DWD, available at Deutscher Wetterdienst, Postfach 100465, D-63004 Offenbach (1995)

Seuffert, G., Gross, P., Simmer, C., Wood, E. F.: The influence of hydrologic modeling on the predicted local weather: two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. Journal of Hydrometeorology. 3, 505–523, (2002)

Stieglitz, M., Rind, D., Famigliette J., Rosenzweig C.: An efficient approach to modeling topographic control of surface hydrology for regional and global climate modeling. Journal of Climate. 10, 118–137, (1997)

Warrach, K., Stieglitz, M., Mengelkamp, T., Raschke, E.: Adavantages of a topographically controlled runoff simulation in a soil-vegetation atmosphere transfer model. Journal of Hydrometeorology. 3, 131–147, (2002)

Landform – Hydrology Feedbacks

Günter Blöschl and Ralf Merz

Abstract The paper discusses the role of hydrological fluxes in landscape evolution and, conversely, the role of the landscape in modulating hydrological fluxes. Three examples from Austria are presented that highlight the nature of the feedbacks. The first example examines feedbacks between different time scales – storm runoff at the event scale and landform evolution at the long term scale. The second example examines more complex processes of interactions between macroscale terrain roughness, occurrence of convective storms, flash flooding and drainage density. The third example examines feedbacks between different space scales – precipitation, the regional water balance, local runoff dynamics and flooding, and illustrates that the feedbacks can be counter intuitive and inconsistent with what one would obtain by a mechanistic description that does not include feedback effects across scales.

Keywords Floods · Geomorphology · Runoff generation · Vegetation · Runoff coefficient

1 Introduction

The way in which the landscape influences the flow of water and how the hydrology in turn influences the form of the landscape has been a long standing research issue of geomorphology and, to a lesser degree, of hydrology. The landform – hydrology interactions may occur over a range of time scales. Most often, the interest resides in long-term landscape processes and short term hydrological processes. Kirkby (2005, p. 58), for example, notes: "... it is clear that the form of the landscape is vitally important in affecting the response of the landscape to a storm event. The form of

G. Blöschl (✉)
Institut für Wasserbau und Lngenieurhydrologie, Technische Universität Wien, Karlsplatz 13/222, A-1040 Vienna, Austria
e-mail: bloeschl@hydro.tuwien.ac.at

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_8,

the landscape, however, is already a product of the hydrology over a long period, determining the soil, vegetation, hillslope forms and channel network morphology." Landscape – hydrology interactions may also occur over a range of space scales from small scale erosion processes to continental scale feedbacks between rainfall and the sculpting of mountain ranges, for example.

The purpose of this symposium on landforms was to highlight the crucial role played by the landform as a boundary surface between the atmosphere, the hydrosphere, the biosphere, the pedosphere and the lithosphere. This paper, in particular, focuses on the interactions between geomorphological processes and hydrological processes with an emphasis on the hydrological part. Specifically, three examples will be presented to illustrate the nature of the interactions. Key questions concerning these interactions are: At what space-time scales do they occur? How effective are any feedback effects and are there positive or negative feedback loops? What is the role of these interactions *vis a vis* other controls such as vegetation and geology?

The three examples have been selected to address different aspects of the feedbacks (Table 1). The first example examines feedbacks between different time scales – storm runoff at the event scale and landform evolution at the long term scale. This is the classical landform – hydrology feedback related to by Kirkby (2005). The second example examines more complex processes of interactions between macroscale terrain roughness, occurrence of convective storms, flash flooding and drainage density. The third example examines feedbacks between different space scales – precipitation, the regional water balance, local runoff dynamics and flooding and analyses the role of the catchment water balance versus local soil and vegetation characteristics.

Table 1 Landform-hydrology feedback examples discussed in this paper

(1) *Precipitation* → incision ↔ *runoff dynamics*
(2) Macroscale terrain roughness → *flash flooding* ↔ drainage density
(3) *Precipitation* → *water balance* ↔ *runoff dynamics* → *flooding*

Plain font = Landform process
Italics = *Hydrological* process
↔ denotes feedback

All examples have been taken from Austria (Fig. 1). Runoff data from a total of about 500 stream gauges were available to this study, as well as rainfall data from about 1000 stations and air temperature data from about 300 stations.

2 Landscape Incision and Runoff Dynamics

Figure 2 shows a topographic contour line for each of two example catchments. The left panel shows the contour line of the Thal catchment in western Austria, the right panel that of the Oberlainsitz catchment in north-eastern Austria. The striking difference between the two contour lines is the degree of incision of the channels – deeply

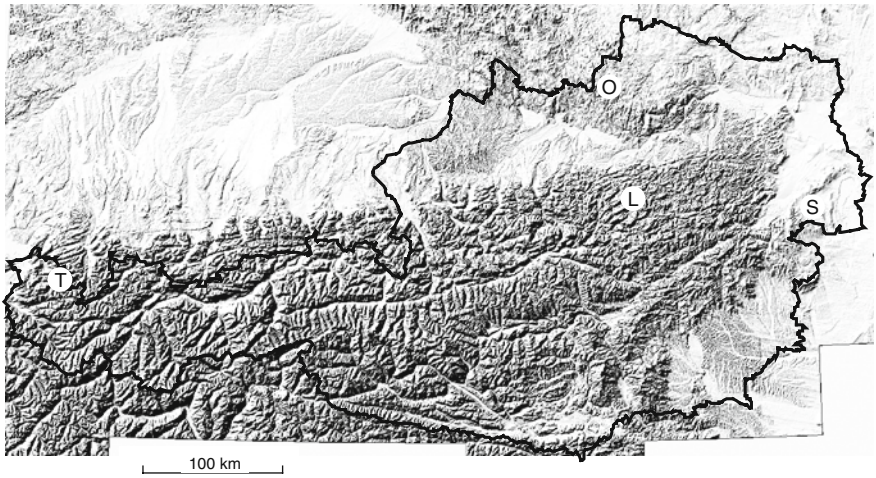

Fig. 1 Topography of Austria. Letters T, O, L, S relate to the example catchments in Figs. 2 and 6 (Thal, Oberlainsitz, Lunz, Schützen)

incised channels in the case of the Thal catchment as indicated by the breaks in the topographic contour lines and hardly any incision in the case of the Oberlainsitz catchment as indicated by smooth contour lines. Incised channels are, apparently, a result of erosive forces due to regular large floods, while the smooth landform points to low or moderate floods. Conversely, it is likely that the more incised landform will exacerbate runoff production and routing, thereby increasing flood flows in the Thal catchment. This is hence an example of landform-hydrology feedbacks. This hydrological assessment of flood behaviour of the two catchments based on the analysis of channel incision is clearly reflected in the observed flood data. The floods in the Thal catchment are much larger than those in the Oberlainsitz (Fig. 2 bottom) even though the catchment sizes are similar (90 and 81 km^2, respectively). The 10 year flood of the Thal catchment is about 140 m^3/s while that of the Lainsitz is only 10 m^3/s. Also, the shape of the flood frequency curve differs. The Thal catchment has the characteristics of a wet catchment with frequent large floods, i.e., the smallest floods are relatively large and the flood frequency curve continues as a straight line in the semi-logarithmic plot. Conversely, the Oberlainsitz catchment has the characteristics of a dry catchment where most floods are small and large floods are rare, i.e., the smallest floods are small and the flood frequency curve indicates an upward curvature. The main driver of the differences in the landscape and hydrological characteristics appears to be rainfall. Mean annual precipitation is 1800 mm and 830 mm for the Thal and Oberlainsitz catchments, respectively. Clearly, mean annual precipitation is an indicator of landform-hydrology feedbacks at the scale of centuries rather than at the event scale. There are also differences in the geology of these two catchments that likely affect the erosive processes. The Thal catchment consists of clay, marl, sand and sandstone. Due to the dominance

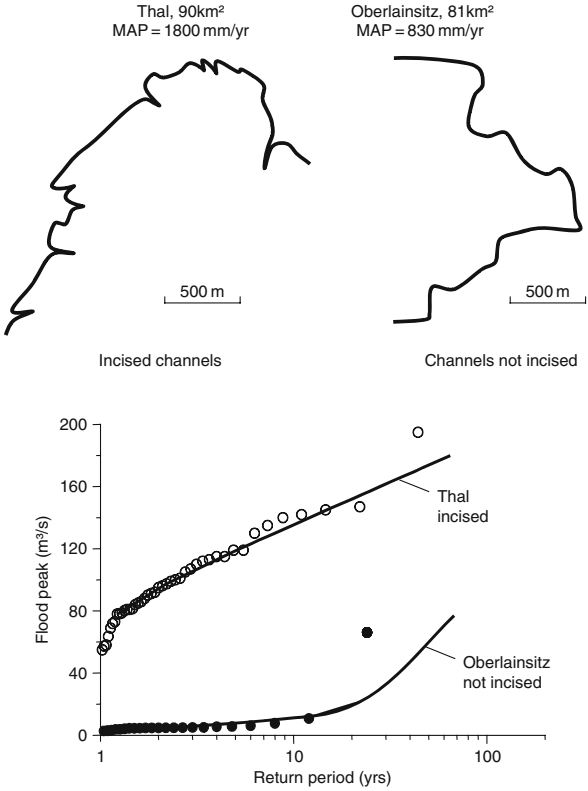

Fig. 2 *Top*: Topographic contour lines of two catchments. *Left*: Thal catchment in the West of Austria, 90 km^2 catchment area, mean annual precipitation (MAP) of about 1800 mm/yr. *Right*: Oberlainsitz catchment in the North-east of Austria, 81 km^2 catchment area, MAP of about 830 mm/yr. The topographic contours are at 800 m and 900 m a.s.l., respectively. *Bottom*: Flood frequency curves of the two catchments

of clay and marl, only a small part of rainfall infiltrates to recharge groundwater. The geology of the Lainsitz catchment is mainly granite and gneiss. Weathering has produced sandy soils with a large infiltration capacity. This example illustrates the presence of a positive feedback loop between landform and hydrology at different time scales where the main link is through erosion processes. A similar paired example, the Mitterbach and Ötscherbach catchments, is given in Blöschl (2005). In that example, the main drivers of differences between landscape-hydrology feedbacks, however, are differences in the depth of bedrock weathering as a result of differences in the rock chemistry (dolomite versus limestone).

3 Flash Flooding and Channel Evolution

The second example discusses a more complex case. Figure 3(a) shows the drainage density as estimated from the rivers displayed in the 1:50,000 topographic map of Austria. The drainage density is an indicator of a spectrum of hydrological processes. High densities typically indicate large runoff generation and fast runoff

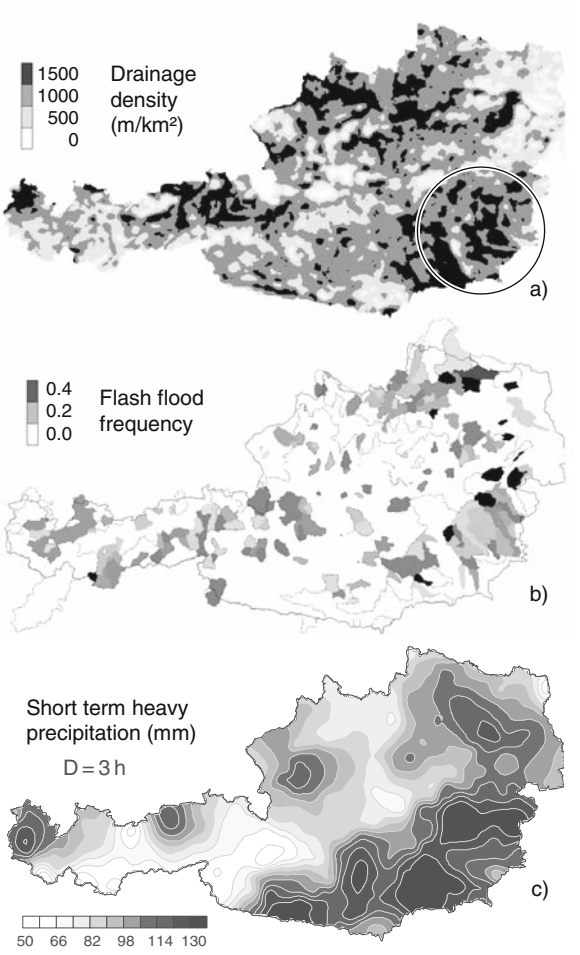

Fig. 3 (**a**) Drainage density as estimated from the rivers of the 1:50,000 topographic map of Austria. Circle shows area of high drainage density (redrawn after Merz and Blöschl, 2003). (**b**) Flash flood frequency as estimated by analysing stream gauge data and flood process indicators (redrawn after Merz and Blöschl, 2003). (**c**) Envelope curves of maximum observed 3 h precipitation in Austria (redrawn after Blöschl et al., 2005)

routing. Small densities typically indicate large infiltration potential, e.g., in Karst. Drainage density is hence frequently used as a summary index in hydrology for estimating floods (Merz and Blöschl, 2005), low flows (Laaha and Blöschl, 2006) and model parameters (Parajka et al., 2005b) if no local runoff data are available. The map in Fig. 3(a) shows high drainage densities in the South-east of Austria (indicated by a circle). Figure 3(b) indicates that, in that area, the frequency of flash floods is high. A flash flood is defined as a localised flood event, i.e., an event that covers a small area (e.g., 10 km^2), with short durations (of an hour or less), and is usually produced by short convective storms. The frequency of flash floods was estimated by Merz and Blöschl (2003) by analysing stream gauge data and flood process indicators such as storm size, storm duration and antecedent soil moisture. The maximum frequency is about 40% which indicates that 40% of the maximum annual floods are of a flash flood type. It should be noted that this frequency relates

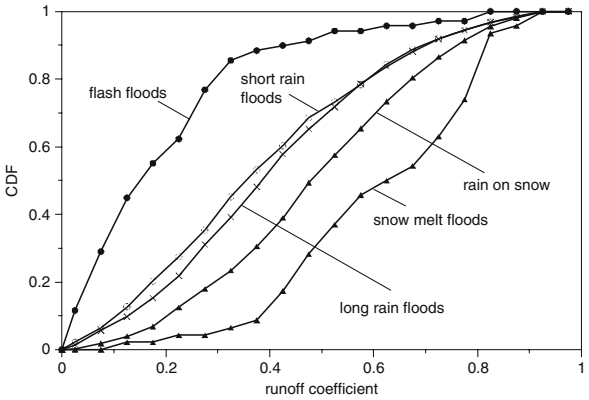

Fig. 4 Distribution function of the event runoff coefficients for maximum annual floods in Austria stratified by the event type (redrawn after Merz et al., 2006a). Runoff coefficients of 1 indicate that all the rainfall becomes runoff during an event while smaller values indicate less runoff

to gauged catchments only (median catchment size of 100 km^2). The flash flood frequencies are probably larger for smaller catchments but it is difficult to estimate these at the regional scale. A comparison of Fig. 3(a) and (b) suggests that the two patterns are likely related, i.e., the above-average flash flood occurrence is related to the high drainage density and *vice versa*. This is supported by the pattern of short term heavy precipitation presented in Fig. 3(c). Clearly, the largest rainfall intensities occur in this part of Austria. The high intensities are related to the landscape form as the undulating terrain in this region tends to trigger convective storms. There are other controls on flood occurrence also. Specifically, the runoff coefficient controls the amount of rainfall that actually becomes runoff. A comprehensive analysis of runoff coefficients of Merz et al. (2006a) has been used here, where the events have been stratified by event type. The events of the flash flood type are associated with the lowest runoff coefficients (Fig. 4). This is due to flood occurrence in summer, where the soils tend to be dry and due to the limited spatial extent of such storms. Specifically, the median runoff coefficient is 0.15 while the other event types (such as long rain synoptic floods, rain on snow floods) have median runoff coefficients of 0.36 or more over all of Austria. This implies that, while the undulating landform type does enhance floods through triggering storms, relatively little of that rainfall becomes runoff due to limited storm size and dry antecedent conditions, again due to the convective nature of the storms. One can hence think of these interactions as negative, or stabilising, feedbacks.

4 Water Balance and Runoff Dynamics

The last example, again, deals with runoff generation, but in a different context. Figure 5(a) shows event runoff coefficients as estimated from soil and vegetation data. The runoff coefficients have been expressed by the SCS curve number (CN) which is 100% if all the rainfall becomes runoff during an event and 0% if no runoff

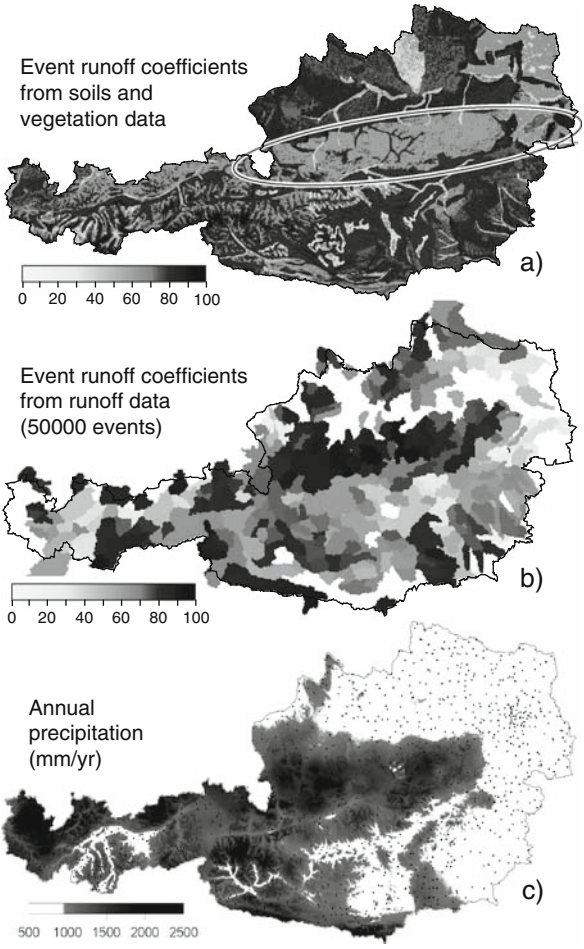

Fig. 5 (**a, b**) Event runoff coefficients as indexed by the SCS-Curve number. 100% indicates that all the rainfall becomes runoff during an event while lower values indicate less runoff (Redrawn after Merz et al., 2006b). (**a**) has been estimated from soil and vegetation data without using runoff data while (**b**) has been estimated directly from runoff data. (**c**) Long term mean annual precipitation. The dots indicate the stations used in the analysis (redrawn after Parajka et al., 2005a)

occurs. The SCS-CN method is widely used in many countries to estimate runoff coefficients in the absence of runoff data (i.e., in ungauged catchments) and the procedure (see, e.g., Dingman, 1994) has been mimicked in Fig. 5(a). Obvious in Fig. 5(a) is an elongated area (marked by an ellipse in the Figure) with below average runoff coefficients. This is a forested area. For forest soils the SCS-CN method predicts small runoff coefficients. This is because, usually, runoff coefficients in forests are small at small spatial scales (plots, hillslopes). However, at the catchment scale this is not the case due to feedback effects as illustrated in Fig. 5(b) which shows runoff coefficients (again indexed by the SCS-CN) estimated from runoff data. The runoff coefficients in Fig. 5(b) are certainly more reliable than those in Fig. 5(a) as actual runoff data have been used for estimating them. In the elongated area, the runoff coefficients are in fact much above the average in Austria in spite of the forest soils. The reason for this is the catchment water balance which is

controlled by rainfall at longer time scales (longer than an event). This is illustrated by the mean annual precipitation in Fig. 5(c) which shows a remarkable similar pattern to that of the event runoff coefficients. The feedbacks are hence across scales. Precipitation tends to increase the climatic soil moisture state which controls the runoff dynamics at the event scale, including extremes (floods). It is indeed surprising that the soil moisture effect is more important than the vegetation effect. Here, again, the important point is the space scale. The SCS-CN method has been designed for the hillslope and small catchment scale and is consistent with many local scale experiments in forests. However, at the catchment scale the soil water balance tends to dominate as runoff is not produced from all parts of the catchment but preferably from those that are close to the stream, hence the scale effect. Also lateral subsurface redistribution seems to be important. These interactions also have a bearing on the magnitude and frequency of floods. This is illustrated in Fig. 6 in terms of the flood frequency curves and associated runoff coefficients for two example catchments. The Lunz catchment (Fig. 6a) is a wet catchment (mean annual runoff of 1200 mm).

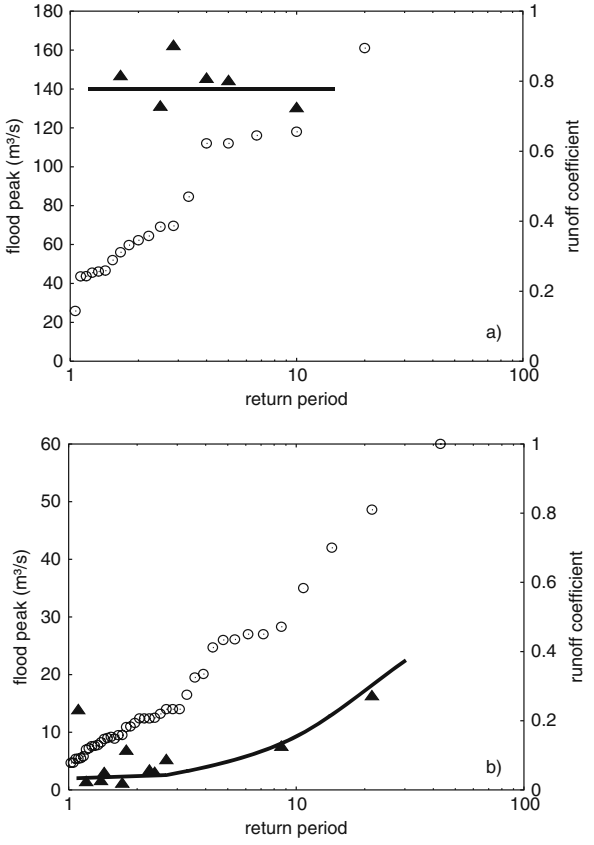

Fig. 6 Flood frequency plots (*circles*) and event runoff coefficients of the associated flood events (*triangles*). (**a**) Lunz catchment (118 km^2 catchment area). (**b**) Schützen catchment (383 km^2 catchment area)

The runoff coefficients (plotted as triangles) are always large (around 0.8) and do not change much with the magnitude of the event. In contrast, the Schützen catchment (Fig. 6b) is a dry catchment (mean annual runoff of 100 mm). The runoff coefficients are very small (around 0.05) for small events and increase significantly with the magnitude of the event. This produces a strong non-linearity in the flood frequency response, i.e., the floods increase more strongly with rainfall magnitude than does the rainfall. These are interesting non-linearities as they are controlled by feedback effects across scales.

5 Conclusions

This paper illustrates landform – hydrology feedbacks by three examples. The first example indicates that positive feedbacks exist between storm runoff at the event scale and landform evolution at the long term scale. As rainfall increases, runoff increases more than does rainfall. This is mainly due to erosion processes which are apparent in the landform. The second example indicates that undulating terrain may enhance convective storm activity which may increase flash floods and drainage density. Conversely, drainage density may increase flash flooding. However, there is a negative feedback loop involved as convective storms are usually of small spatial scale, hence runoff coefficients tend to be small which may decrease the magnitude of flash floods. The third example examines feedbacks between different space scales – precipitation, the regional water balance, local runoff dynamics and flooding. Runoff generation is usually thought of as a local phenomenon (with high infiltration in forest soils), but the data indicate that runoff generation at the catchment scale is largest in the forested catchments. This is because of the wet soil moisture states as a result of (long term) climate. This suggests that feedbacks can be counter intuitive and inconsistent with what one would obtain by a mechanistic description that does not include feedback effects across scales. The seasonal water balance and the hydrological regimes that have evolved in conjunction with landform processes are hence an excellent indicator for explaining the spatial variability of both low flow and flood flow fluxes. This is because of the crucial role of catchment soil moisture that is controlled by the water balance at the seasonal scale and, conversely, controls low flows and flood flows as well as landform processes.

Acknowledgments We would like to thank the FWF (project no. P18993-N10) and the ÖAW (APART programme) for financial support.

References

Blöschl, G. (2005) Rainfall-runoff modelling of ungauged catchments. Article 133 in: *Encyclopedia of Hydrological Sciences*, M. G. Anderson (Managing Editor), J. Wiley & Sons, Chichester, pp. 2061–2080.

Blöschl, G., R. Merz, G. Laaha, J. Parajka and J. Skøien (2005) Regionale hydrologische Untersuchungen in Österreich. Mitteilungsblatt des Hydrographischen Dienstes in Österreich, Nr. 83, Wien, pp. 83–98.

Dingman, S.L. (1994) Physical Hydrology. Macmillan New York, 575 pp.
Kirkby, M. (2005) Organisation and process. Article 4 in: *Encyclopedia of Hydrological Sciences*, M. G. Anderson (Managing Editor), J. Wiley & Sons, Chichester, pp. 41–85.
Laaha, G. and G. Blöschl (2006) A comparison of low flow regionalisation methods – catchment grouping. *Journal of Hydrology*, 323, pp. 193–214
Merz R. and G. Blöschl (2003) A process typology of regional floods. *Water Resources Research*, 39 (12), 1340, doi:10.1029/2002WR001952.
Merz, R. and G. Blöschl (2005) Flood frequency regionalisation – spatial proximity vs. catchment attributes. *Journal of Hydrology*, 302 (1–4), pp. 283–306.
Merz, R., G. Blöschl and J. Parajka (2006a) Spatio-temporal variability of event runoff coefficients in Austria. *Journal of Hydrology*, 331, pp. 591–604.
Merz, R., G. Blöschl and J. Parajka, D. (2006b) Raum-zeitliche Variabilität von Ereignisabfluss-beiwerten in Österreich. *Hydrologie und Wasserbewirtschaftung*, 50 (1), pp. 2–11.
Parajka J., R. Merz and G. Blöschl (2005a) Regionale Wasserbilanzkomponenten für Österreich auf Tagesbasis. *Österreichische Wasser- und Abfallwirtschaft*, 57, H 3/4. pp. 43–56.
Parajka J., R. Merz and G. Blöschl (2005b) A comparison of regionalisation methods for catchment model parameters. *Hydrology and Earth Systems Sciences*, 9, 157–171.

Hydrological Analyses as a Prerequisite for Soil Erosion Modeling – Landscape Related Studies in a Mesoscale Hydrological Catchment

Herwig Hölzel and Bernd Diekkrüger

Abstract This study focuses on the necessity of coupling hydrological analyses with event-based erosion models. Continuous runoff simulations for 2 years were conducted with the WaSiM-ETH model. The model was calibrated for 1999 and successfully validated for 2000 using observed runoff data (Pearson's $r = 0.9$ for both years). The validated model was used to compute spatial and temporal differentiated soil moisture as an essential prerequisite for event-based soil erosion modeling.

Commonly, soil properties are assumed to be independent of land use; however, this is rarely true. To be able to consider effects of different usages of grassland on runoff-generating processes, a database was generated based on field measurements from cut pasture and a neighboring pasture. On average, these measurements (305 samples) show significant differences in bulk density of the top soil layer (neighboring pasture = 1.26 g/cm^3, cut pasture = 1.57 g/cm^3). Mean measurements of saturated hydraulic conductivity (54 samples) were five times smaller from the cut pasture (96 cm/d) than they were from the neighboring pasture (488 cm/d). Agricultural practices were identified as the main cause for these differences. Mechanical pressure applied by heavy farming machines on the cut pasture led to higher bulk densities and lower saturated hydraulic conductivities. These effects surpassed the compacting effect of trampling by grazing animals on the pasture.

The land use effects were represented in the hydrological model. Modeling exercises that use the extended model may lead to differences in calculated infiltration rates and related surface runoff rates. For the pasture, 44 mm/a of surface runoff was calculated compared to 68 mm/a of runoff on the cut pasture for a small headwater catchment. An extrapolation to the whole catchment (54 km^2) leads to an improvement of model quality without recalibration. Pearson's r and the Model Efficiency were increased by 0.02. For the sub-catchment (0.28 km^2), where the investigated pasture and the cut pasture are located, Pearson's r was increased by 0.04 and the

H. Hölzel (✉)
Department of Geography, University of Bonn, Meckenheimer Allee 166, 53115, Bonn, Germany
e-mail: herwighoelzel@giub.uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_9,

Model Efficiency was increased by 0.06. Considering land use-dependent runoff processes may lead to an improved simulation of erosion risk for adjacent fields.

Keywords Hydrological modeling · WaSiM-ETH · Saturated hydraulic conductivity · Infiltration · Runoff building · Runoff · Pastures

1 Introduction

There is an increasing demand for models that can explain the influence of humans on the environment, especially when climate change effects need to be considered (IPPC, 2007). Process-based distributed hydrological models are useful for estimating changes in water and matter turnover through climate and land use changes (Refsgaard, 1996). Although conceptual models are often used successfully, the application and development of Richard's equation as the basis for process-based models of the unsaturated soil zone presents the appealing possibility of dealing with current and future model demands (Durner and Fühler, 2005). Current developments focus on scaling capabilities and are based on the assumption of the invariance of process algorithms for scale transfer (Wigmosta and Prasad, 2005). In a discussion with Beven, Refsgaard et al. (1996) determined the necessity of developing process-based distributed models and concluded that there is no alternative to other model concepts (stochastic, empirical). The distributed concept allows for a spatially explicit representation of landscape properties (Rosbjerg and Madsen, 2005). In contrast with the HRU concept (Flügel, 1997), the application of raster cells (grids) as regular spatial simulation entities allows for easier coupling with other models. Examples include those described by Warrach et al. (2002) and Ament and Simmer (2006); these authors coupled hydrological models with atmospheric models.

Šimunek (2005) defined important application fields for process-based models:

- Multiple coupling of physical, biological and geochemical processes;
- Interpretation of measurements and explanation of trends;
- Evaluation of the sensitivity of driving forces.

Incorporation of existing knowledge and improvement of process understanding. Nearing et al. (2005) emphasized the possibility of process-based erosion models to consider sedimentation processes and related off-site damages. As disadvantages of these models, they mentioned the considerable parameterization effort required and the necessary empirical components. Further, predictions with process-based distributed erosion models are still considered to be highly uncertain (Vigiak et al., 2006).

However, the advantage of these models is that the consideration of spatially heterogeneous landscape properties becomes possible (Rosbjerg and Madsen, 2005). Lørup and Styczen (1996) pointed out that process-based erosion models should be more sensitive to land use effects. They focused on the saturated hydraulic conductivity (K_{sat}) because this parameter plays a key role in modeling the water dynamics

of the unsaturated soil zone (Tietje and Hennings, 1996). The estimation of K_{sat} is often done using empirical models, the so-called pedotransfer functions (Rawls and Brakensiek, 1985; Vereecken et al., 1989). These functions are easy to apply and require no cost-intensive field investigations, however they are mainly based only on physical-mineralogy characterics (Schaap, 2005). The importance of K_{sat} for calculating infiltration and runoff is well accepted. However, this effect is superimposed by the land use effect as shown by Endreny (2005) for urban areas, by Burt and Slattery (2005) for farm land, and by Bruijnzeel (2005) for forested land.

The aim of this study was to take into account land use effects when estimating K_{sat} using pedotransfer-functions. It is assumed that the consideration of land use-dependent K_{sat} will improve the quality of runoff and erosion simulations. For this study, a well investigated mesoscale catchment (Wahnbach) was chosen in western Germany. Since extensive observations and simulations of runoff and erosion were conducted by Bogena (2001), the catchment is suitable for examining new methods in hydrological modeling.

Bogena (2001) applied the model system OPUS to simulate water dynamics, nitrogen dynamics and erosion in the Wahnbach catchment. The OPUS model is a hillslope model linked to the river system. For the simulation the entire catchment was discretized into 1000 hillslopes. Although reasonable results were obtained, some limitations have been pointed out (Bogena et al., 2003):

- No process-based description for erosion and transport on hillslopes;
- Assumption of homogeneity of the hillslopes without internal differentiation;
- No consideration of river or stream processes.

To deal with these disadvantages the LImburg Soil Erosion Model (LISEM) developed by de Roo et al. (1996) was chosen for further studies, which had the following features:

- Process-based descriptions of erosion, transport and sedimentation;
- Internal differentiation of spatial properties by raster cells;
- Consideration of river stream processes and therefore validation with gauge data.

One main problem with event-based erosion models like EROSION3D (Schmidt and von Werner, 1999) or LISEM is the definition of initial conditions (Bogena and Diekkrüger, 2001). The highest sensitivity for event-based erosion models consist initial soil moisture. Schmidt and von Werner (1999) determined that runoff increased by more than 200% when there was a 10% increase in soil moisture. Using LISEM, an increase in soil moisture by 10% often led to saturated conditions and therefore an overestimate of runoff peaks by several orders of magnitude (de Roo and Jetten, 1999). Thereby, Lørup and Styczen (1996) suggested using initial soil moisture patterns computed by a continuous hydrological model for event-based soil erosion analyses. This study will follow that suggestion.

In this study, we show the necessity of hydrological analyses as a prerequisite for soil erosion modeling. Moreover, land use effects on K_{sat} were investigated. By implementing these effects in our model, we expected to see an improvement in model certainty.

2 Research Area und Temporal Focus

The Wahnbach catchment is situated in the low mountain range of Bergisches Land on the border of the Rhenish Massiv 25 km north-east of Bonn, Germany (Fig. 1). There is a gradient in mean annual precipitation from 850 to 1,130 mm due to the elevation increasing from SW to NE from 130 to 380 m (Bogena, 2001). The bedrock is composed of moderate to low permeability clay, silt and sandstone, some of which is heavily weathered and overlain by a loess cover. The catchment is characterized by steep valley cuts due to high rainfall intensities.

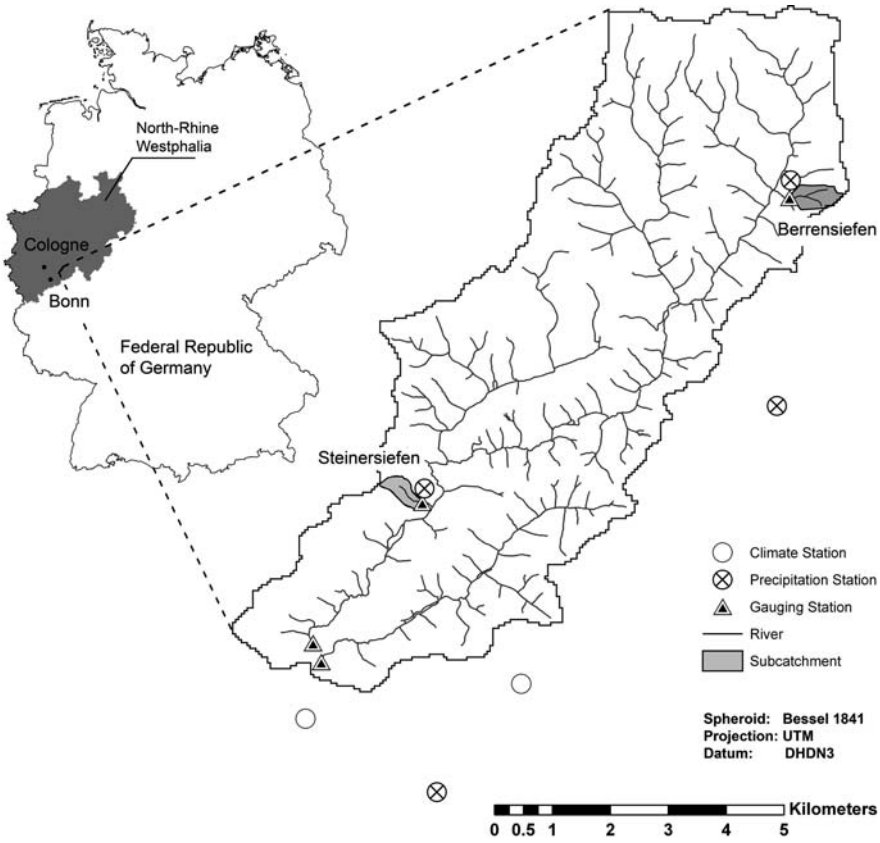

Fig. 1 The hydrological catchment of the Wahnbach river (54 km^2) and the investigated subcatchments (Berrensiefen 0.28 km^2; Steinersiefen 0.21 km^2)

Dominant soil types are fertile Cambisols and Luvisols on the hillslopes as well as Gleysols and Planosols in the valleys. The thickness of soil on the hillslopes is often less then one meter. A dense net of macropores is commonly observed. The thin soils and low permeability of the bedrock are responsible for a quick runoff response after rainfall (Fig. 2). Due to the number of rivers and the steep valley cuts, the discharge leaves the area in less than one day.

The area is intensively used for agriculture. The largest part of the region (> 50%) is used as grassland, followed by forest (20%). Despite the fertile soils, farm land occupies only 8% of the area; this activity is regulated because there are eutrophication problems in the water reservoir. An additional 10% of the land is used for settlements. The remaining areas are occupied by water, streets and other features. The grassland is subdivided into pastures (78.7%) and cut pastures (21.3%). The Wahnbach river drains into a reservoir. This reservoir stores on average 34×10^6 m^3 and provides drinking water for over 700,000 people of the surrounding area. Land use for both agriculture and the drinking water supply leads to conflicts between agriculture and the water economy. Because of this, investigations dealing with water and matter turnover were conducted from 1998 to 2001 and extensive data sets were generated by the Hydrology Research Group of the Geographical Institute, University of Bonn (Bogena, 2001). Because of good data availability, the study focused on 1999 and 2000.

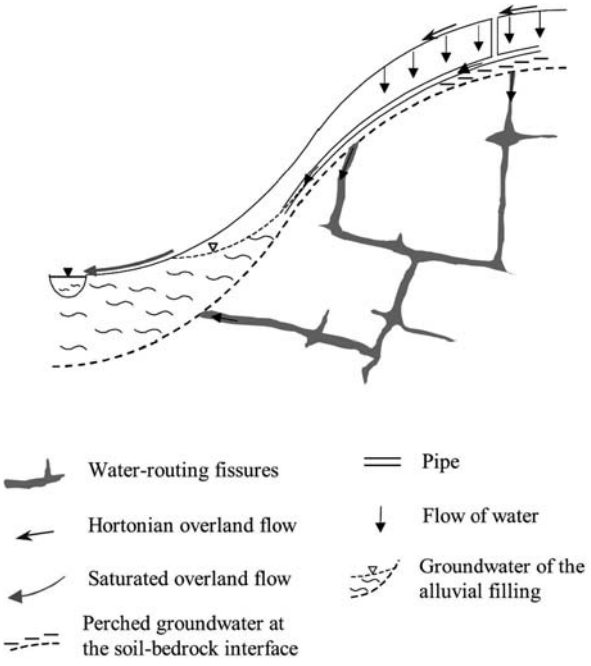

Fig. 2 Runoff generation processes in the Wahnbach catchment with special consideration of the interflow (Bogena, 2001)

Additional data were obtained by field investigations focusing on land use effects on runoff building and runoff. Two sub-catchments, Berrensiefen (0.28 km^2) and Steinersiefen (0.21 km^2) were chosen for field investigations (Fig. 1).

3 Methods

3.1 Field Investigations

Because operationally available data sets were not sufficient to cover all aspects, field investigations were conducted. Two neighboring sites that were used for different purposes (pasture, cut pasture) were investigated in the Berrensiefen sub-catchment (0.28 km^2) (Fig. 3).

Both sites have been used for the same purpose since 1968. The cut pasture is generally influenced more by farming machines and it is cut three times, once each in May, July and September yearly. During the spring, a heavy barrel (750 kp/m) for maintenance purposes is dragged over the ground. For this procedure, a tractor is used (unloaded weight of approximately 4 tons). Further, tumbrils increase the soil compactness when bunkering the harvested grass.

The pasture is only traversed for fertilization and maintenance purposes by a tractor with an unloaded weight of approximately 2 tons. Cutting is conducted only once in the autumn for maintenance. From the beginning of April to the end of November (depending on the weather), the pasture is used as a shared pasture. This means that the pasture is used as such for one week and then allowed a regrowth period of three weeks. In August, the breaks are increased to 4–5 weeks. The average number of cattle stocked is 17–20 (dairy cows) per hectare. At each site, a test area of 100 m^2 was chosen. The two test areas were directly adjacent, so the properties (relief, soil) would be similar for the two.

Undisturbed soil samples were taken directly from below the root zone (10–15 cm) to determine bulk density and soil texture. Nine point measurements were conducted at each test site with three replicates to increase statistical certainty. To

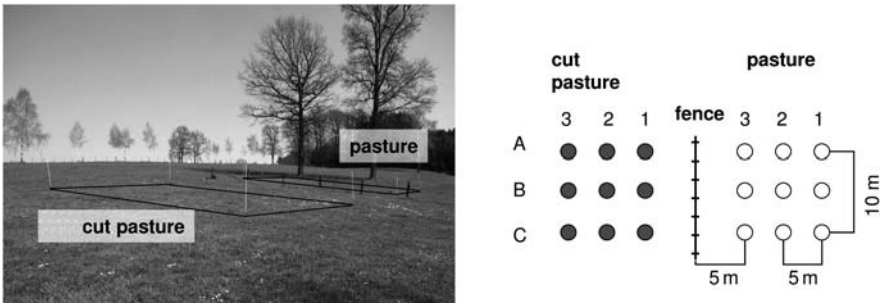

Fig. 3 Experimental setup in the Berrensiefen sub-catchment (0.28 km^2)

consider the temporal dynamics during the course of the year, measurements were conducted in the spring and in autumn. Bulk density was gravimetrically measured in the laboratory after drying the soil for 24 h at 105 °C. Samples were mixed from the three replicates to determine soil texture according to Köhn (1928) and organic carbon content was also determined. Infiltration rate and therefore K_{sat} was measured in situ at the surface using a double-ring infiltrometer.

In addition to these data, measurements were taken along transects at other sites in other pastures and cut pastures in the Wahnbach catchment to illustrate the transferability of the methods to other locations (Fig. 4).

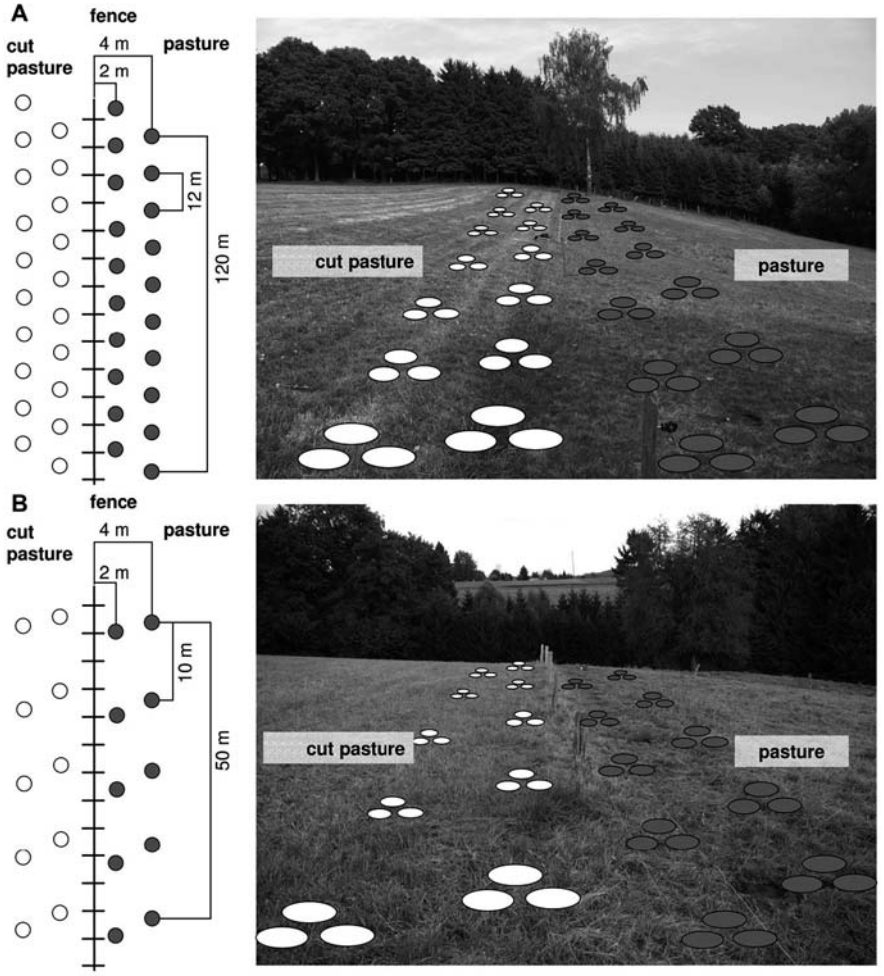

Fig. 4 Experimental setup for the test of transferability in the (**A**) Berrensiefen sub-catchment (0.28 km^2) and (**B**) Steinersiefen sub-catchment (0.21 km^2)

3.2 Runoff Analyses

For the hydrological simulation, the Water balance Simulation Model (WaSiM-ETH), version 7.2.7 (Schulla and Jasper, 2006) was chosen. WaSiM-ETH is a deterministic hydrological model and can be classified as process-based and distributive, as it is based on raster cells (Refsgaard, 1996).

Because WaSiM-ETH computes soil moisture dynamics continuously, the determination of initial conditions for event-based erosion models is possible. Moreover, the separation of fast flow components (surface runoff and interflow) from base flow is feasible, which enables the calibration and validation of the erosion models.

The model may be configured according to data availability and research aim. Table 1 lists the chosen model configuration. Excluding the groundwater model, all algorithms are characterized as process-based. Groundwater was described using a single linear storage approach, since a physical description of cleft-aquifers is missing. A summary of required input parameters and their origins are given in Table 2. Observed runoff for the sub-catchments and the entire catchment area was used for model calibration and validation. Through simultaneous measurements of electrical conductivity and runoff at the gauging stations of the sub-catchments, the direct runoff (surface and fast interflow) could be separated from the slow runoff components.

Table 1 Selected sub models and algorithms of WaSiM-ETH

Sub model	Algorithm
Meteorological data extrapolation	Thiessen-polygons
Meteorological data correction	Shadowing and exposition dependent adjustment for radiation and temperature (Oke, 1987)
Potential evapotranspiration	Penman-Monteith equation (Monteith and Unsworth, 1990)
Actual evapotranspiration	Suction depended reduction according to Feddes et al. (1978)
Interception	Leaf-area-index depended bucket approach
Snow accumulation and snowmelt	Degree-day method
Infiltration	Modified Green-Ampt approach according to Peschke (1987)
Unsaturated soil zone	Richard's equation parameterized based on van Genuchten (1980), macropores considered (Jansson and Karlberg, 2001)
Groundwater	Single linear reservoir
Discharge routing	Cinematic-wave using manning-strickler equation

Table 2 Input parameters with spatial-temporal resolution and data source

Input parameter	Data source	Spatial resolution of data	Time of measurement or temporal resolution
Relief information	State survey office of North Rhine-Westphalia	5 m	1998
Soil information	Geological service of North Rhine-Westphalia	1: 50,000	2003
Land use information	Wahnbach reservoir association	1: 5000	Yearly
Precipitation	Bogena (2001), State agency of North Rhine-Westphalia	2 + 2 stations	5–10 min
Global radiation	Wahnbach reservoir association	1 station	15 min
Air temperature	German weather service	1 station	Daily minimum and maximum values
Wind velocity	German weather service	1 station	Daily
Relative air humidity	German weather service	1 station	Daily
Relative sun duration	German weather service	1 station	Daily

For the simulation, a raster-cell size of 50 m and hourly time steps were chosen as a compromise of calculation effort, detail of input data and certainty of model results.

Due to the process-based model character, the number of effective parameters that need to be calibrated is limited (Jasper, 2005). However, some parameters still must be calibrated; this is performed by comparing simulated and measured discharge at the microscale level (Table 3). These parameters were taken to be constant for the entire catchment; this helps in accounting for the well known scale problem (Šimunek, 2005).

Different measures of quality were used to evaluate model fit. The Pearson's product-correlation-coefficient (r), the Coefficient of Model Efficiency (CME) (Nash and Sutcliffe, 1970) and the Index of Agreement (IoA) (Willmott, 1981) were used.

In addition to Pearson's r being used as a criterion for linearity dependence of both simulated and measured values (Legates, 1999), the CME is commonly used as a measure for the degree of fit, despite its sensitivity to extreme events. In addition, the IoA is used because it relates better to the temporal agreement (Krause et al., 2005).

Table 3 Calibrated parameters for the Wahnbach catchment (54 km^2) and the Steinersiefen sub-catchment (0.21 km^2)

Parameter	Effect	Wahnbach catchment (54 km^2)	Steinersiefen sub-catchment (0.21 km^2)	Unit
Recession constant for surface runoff	Runoff concentration	72	0.5	h
Recession constant for interflow	Runoff concentration	96	0.5	h
Drainage density	Runoff generation	20	20	1/m
Inverse recession constant for base flow	Runoff concentration	2.5	2.5	m
Scaling factor for base flow or maximum baseflow	Runoff generation	0.7	0.7	mm/h
Recession factor for decrease of K_{sat} with soil depth*	Runoff generation	0.9	0.9	[–]

*Solely used for soils with one horizon, otherwise deactivated.

3.3 Extended Runoff Analyses

The land use map of the Wahnbach catchment shows the differences between the pastures and the cut pastures.

By using this map, the extrapolation of local scale measurements to the scale of the Wahnbach catchment becomes possible. The soil map is used for the parameterization of K_{sat} in the model, although no information about land-use-dependent variations of K_{sat} is provided by this map. To consider land-use-dependent K_{sat} in the model, the soil map must be categorized by land use.

To account for this, a GIS-analysis composed of two steps was applied: First, pastures and cut pastures were separated using the land use map. Second, both land uses were cross-referenced to the soil map. As a result, an extended soil map was created dividing soil types by where pastures and cut pastures were located.

K_{sat} values provided by the soil map did not consider effects of macropores and were solely representative of the soil matrix. Other measurements of K_{sat} were obtained using the double-ring infiltrometer and considering macropores. Thus, instead of using the measurements directly, the ratio of the measured K_{sat} between pastures and cut pastures was applied to the parameterization. These relationships were used to modify the K_{sat} values given in the soil map.

A ratio-factor was composed for each area based on the geometric means of the measured K_{sat} values. The geometric mean is more appropriate for skewed distributed data than the arithmetic mean. The ratio-factor was calculated by taking the

geometric mean of all samples divided by the geometric mean of the specific area according to the following equation:

$$\mathrm{Ksat}_{x_i}^{\mathrm{mod}} = \mathrm{Ksat}^{\mathrm{map}} * \frac{\overline{\mathrm{Ksat}^{\mathrm{meas}}}_{x_{i\,\mathrm{geom}}}}{\overline{\mathrm{Ksat}^{\mathrm{meas}}}_{x\,\mathrm{geom}}}$$

where x = all samples, xi = samples of the specific area, mod = modified values, map = values given by the soil map, meas = measured values and geom = geometric mean. The reference basis was provided by the K_{sat} -values from the soil map.

The calculated ratio-factors were applied for all soil types where pastures and cut pastures were located. Runoff analyses were conducted without model recalibrations using the modified soil map. The results were compared with the runoff analyses based on the standard soil map.

4 Results

4.1 Field Investigations

Measurements in both areas showed on average a higher soil compaction factor and lower values of K_{sat} for the cut pasture (Fig. 5). To evaluate the statistical relationship between the two parameters, the Pearson's correlation coefficient was determined (Fig. 6). Soil water retention curves were estimated for both areas based on the measured soil textures, organic carbon content and temporally differentiated bulk densities using the pedotransfer-function of Vereecken et al. (1989) (Fig. 7). A measurement of bulk densities along transects in the Berrensiefen and the Steinersiefen sub-catchments validated the measured differences between pastures and cut pastures (Fig. 8).

4.2 Runoff Analyses

The model was calibrated for 1999 and validated with data from 2000 (Fig. 9). Model quality was evaluated by different measures (Table 4).

The validated model was applied to compute several runoff components and to derive soil moisture values for the erosion models. The total runoff for 1999 was subdivided into surface flow (6%), interflow (64%) and base flow (30%) (Fig. 10). According to the measurements of the electrical conductivity in the sub-catchment, direct runoff was estimated to be 6–8% for that year. As an example, Fig. 11 shows a typical rainfall-runoff response for the microscale Steinersiefen sub-catchment (0.21 km^2). Another important output was the spatially distributed soil moisture regime as shown in Fig. 12.

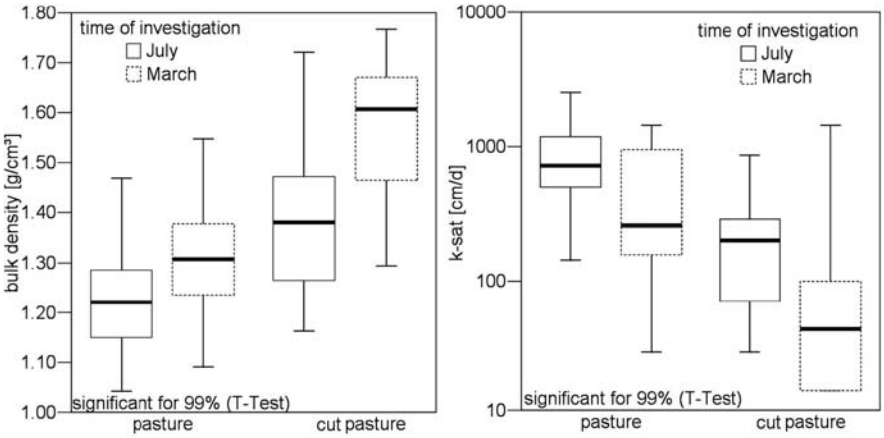

Fig. 5 Results of field investigations at the test sites in the Berrensiefen sub-catchment (0.28 km^2, 27 samples per area and time)

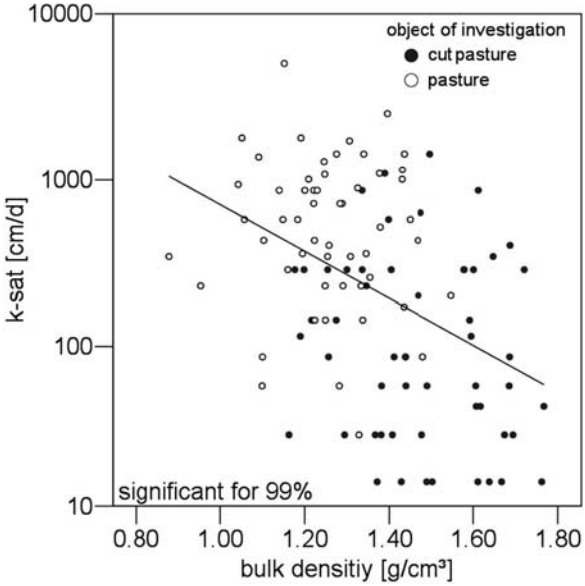

Fig. 6 Correlation of field measurements at the test sites in the Berrensiefen sub-catchment (0.28 km^2, 54 samples, Pearson's $r = -0.43$)

4.3 Extended Runoff Analyses

The ratio-factors for parameterization of K_{sat} differences in the extended soil map were calculated as described previously. The ratio-factors were determined on the basis of the geometric means of the measured K_{sat} for the pasture (487.5 cm/d), the cut pasture (95.7 cm/d) and in total (216 cm/d). For pasture, a ratio-factor of 2.26

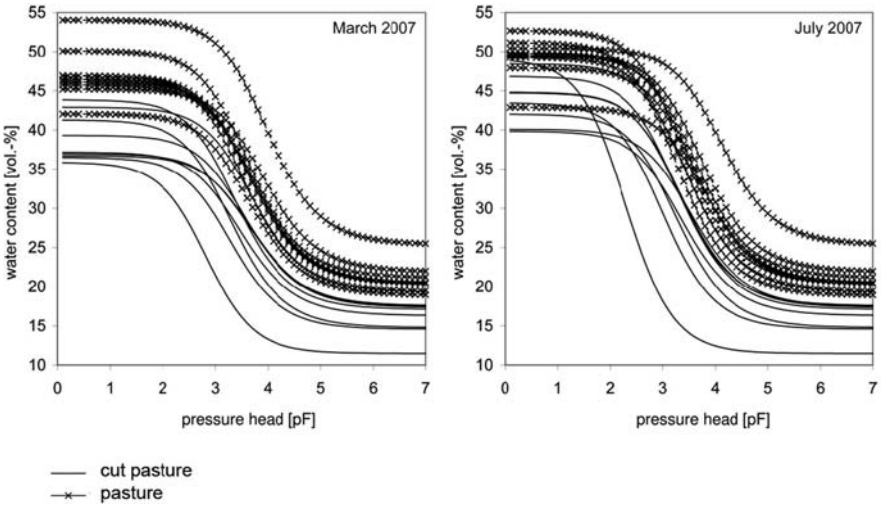

Fig. 7 Calculated pF-curves based on measured soil texture, soil compactness and organic carbon content at the test sites in the Berrensiefen sub-catchment according to Vereecken et al. (1989)

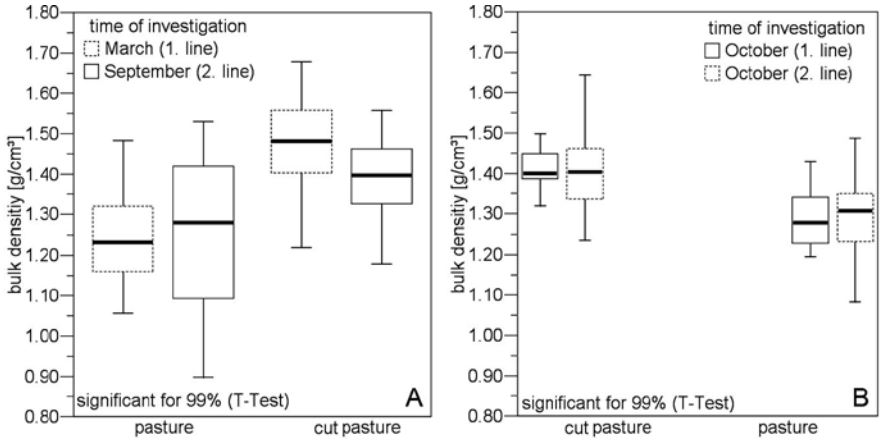

Fig. 8 Spatial-temporal variability of bulk density along transects in (**A**) the Berrensiefen sub-catchment (0.28 km^2, 30 samples per area and time of investigation) and (**B**) the Steinersiefen sub-catchment (0.21 km^2, 15 samples per area and time of investigation)

(216/487.5) was calculated; the ratio-factor for cut pasture was 0.44 (216/95.7). Therefore, values of K_{sat} in the extended soil map were increased for pasture and reduced for cut pastures using the calculated ratio-factor.

Model application of the extended soil map led to differences in calculated runoff building for both areas (Fig. 13).

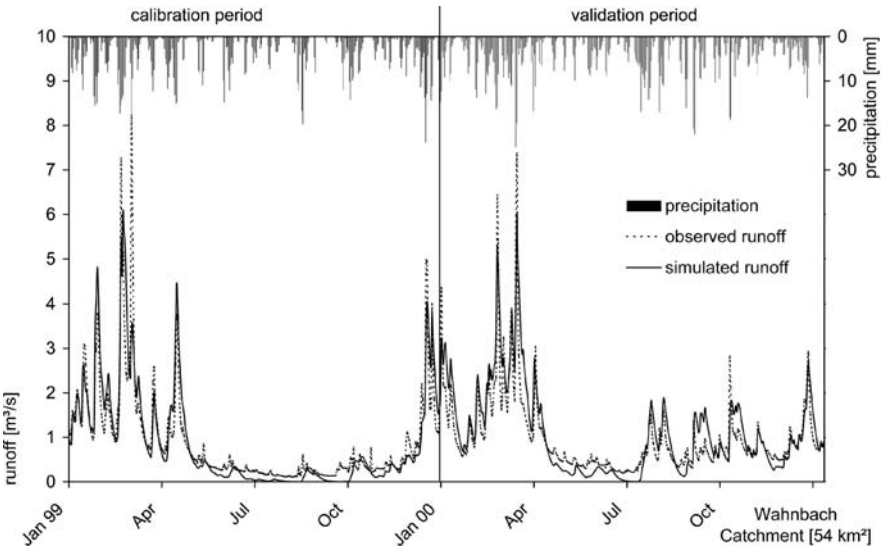

Fig. 9 Observed and simulated runoff of the Wahnbach catchment (54 km^2)

Table 4 Quality criteria for simulated runoff (hourly time steps) of the Wahnbach catchment (54 km^2)

Quality measure	1999 (calibration)	2000 (validation)
Pearson r	0.9	0.9
Model efficiency	0.79	0.74
Index of agreement	0.95	0.94
Observed runoff sum [mm]	550	573
Simulated runoff sum [mm]	533	628
Mass balance error [%]	3.02	8.8

The amount of annual surface runoff for 1999 was 50% higher for cut pasture (68 mm) compared to pasture (44 mm). Without considering these effects, the same surface runoff of 21 mm/a was calculated for both pastures. This is discussed later. The application of the extended soil map increased the quality of the runoff analyses in comparison with the application of the standard soil map without model recalibration.

For the Berrensiefen sub-catchment (0.28 km^2) and the Wahnbach catchment (54 km^2), an improvement of model quality was achieved (Table 5).

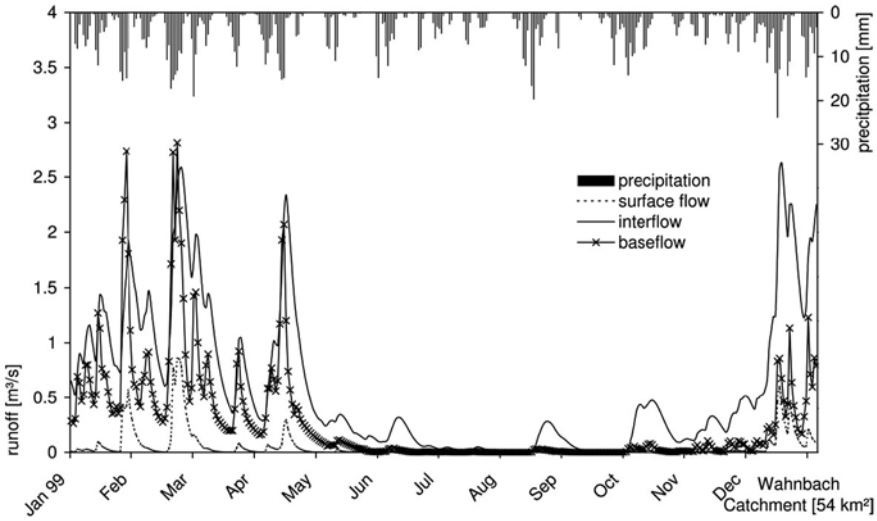

Fig. 10 Simulated runoff components of the Wahnbach catchment (54 km^2)

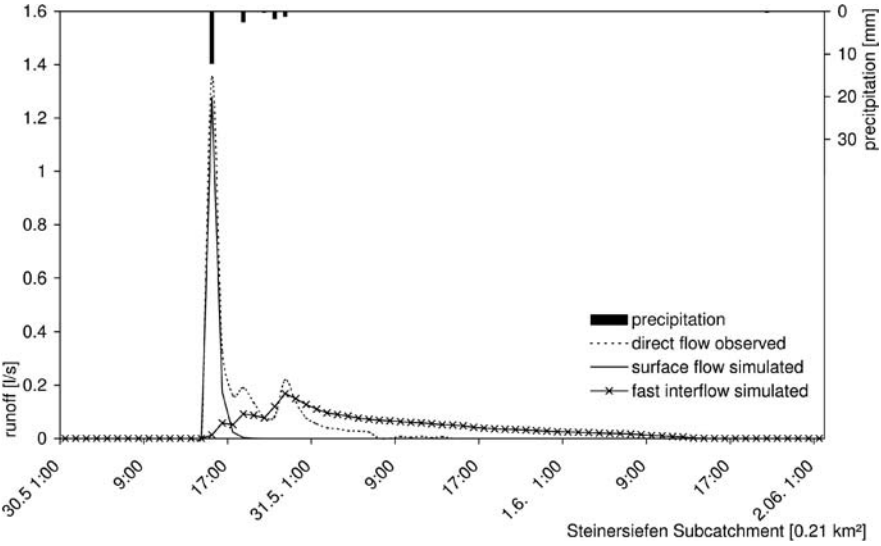

Fig. 11 Observed and simulated runoff response in the Steinersiefen sub-catchment (0.21 km^2)

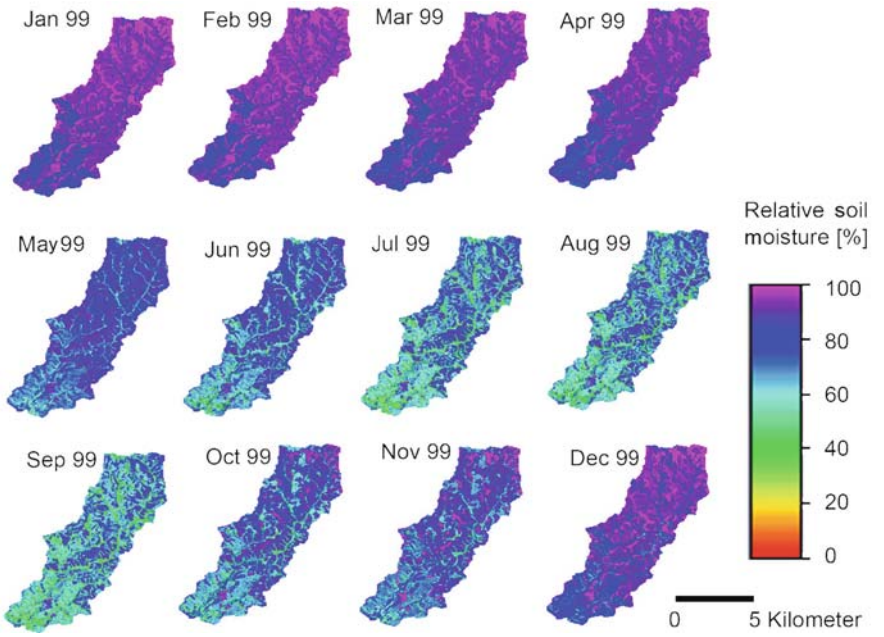

Fig. 12 Simulated mean monthly soil moistures for the Wahnbach catchment (54 km^2)

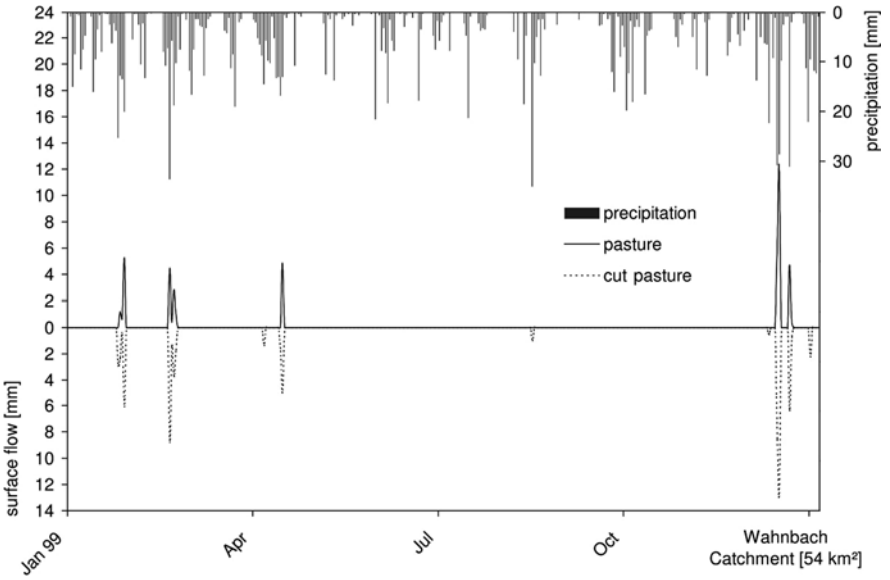

Fig. 13 Simulated surface runoff for the investigated areas in the Berrensiefen sub-catchment (0.28 km^2) with and without the consideration of land use effects

Table 5 Model quality calculated using the standard and the modified soil map for 1999

Quality measure	Berrensiefen sub-catchment (0.21 km^2) with standard soil map	Berrensiefen sub-catchment (0.21 km^2) with modified soil map	Wahnbach catchment (54 km^2) with standard soil map	Wahnbach catchment (54 km^2) with modified soil map
Pearson's r	0.85	0.89	0.9	0.92
Model Efficiency	0.72	0.78	0.79	0.81
Index of Agreement	0.9	0.93	0.95	0.95
Mass balance error [%]	5.69	5.93	3.02	2.97

5 Discussion

5.1 Field Investigations

Investigations have shown that mechanical pressures by heavy farm machinery leads to a compaction of topsoil and an increase in surface runoff (Burt and Slattery, 2005). As well, the trampling effect of grazing animals leads to a higher compaction of the topsoil in pastures (Greenwood and McKenzie, 2001). Kooistra and Boersma (1994) measured higher infiltration rates on permanently used pastures compared to farm land. They found that the macropore effects were responsible for this influence. Previous studies have focused on comparisons between farm land and grassland and a review of the literature reveals a gap in knowledge about the behavior of differently-used grassland (pasture, cut pasture). Measurements of both areas showed clear differences in K_{sat} and bulk density values. The reason for higher soil compaction and lower K_{sat} values is probably the mechanical pressures on the field, which exceed the trampling effects of grazing animals on the pasture. The measurements differ throughout the year, but the trend remains the same. There is high variability within both areas due to the effects of macropores. Therefore, no clear statistical relationship between bulk density and K_{sat} exists. The Pearson's correlation coefficient was significant, but with $r = -0.43$, which indicates only a low dependency. However, two clusters can clearly be seen in the bivariate plots (Fig. 6). Moreover, this clustering is also apparent in the calculated soil water retention curves (Fig. 7). Therefore, in further analysis, the bulk density was taken as the key parameter to explain the differences in K_{sat} values. Measurements of bulk density are quicker to obtain and less complicated than measurements of K_{sat} using the double-ring infiltrometer; this simplifies field work substantially. Therefore, the bulk density was used for the spatial validation of measured differences at other locations in the Wahnbach catchment.

5.2 Runoff Analyses

The calculated measures of model quality indicate good agreement between simulated and observed runoff for the Wahnbach catchment (Fig. 9 and Table 4). As Fig. 9 shows, the quality of low flow simulation during summer is poor. During this period, the runoff is mainly fed by base flow. WaSiM-ETH considers only one aquifer, which is not sufficient in this situation. The effect of the fractured aquifer could not be considered; this results in limited simulation quality during dry periods. Nevertheless, the simulation results are acceptable because the analyses focus on soil water content and surface runoff.

To reduce the effort of parameterization, both interflow and base flow are often neglected by erosion models. Therefore, runoff calculated with erosion models cannot be directly calibrated or validated using measured discharges. However, a good runoff analysis is an essential prerequisite for analyzing erosion processes. WaSiM-ETH is suitable for differentiating surface runoff from other runoff components. To check the validity of modeled surface runoff, simultaneous measurements of electrical conductivity and runoff at the gauging stations of the sub-catchments were used. Thus, the separation of direct runoff (surface and fast interflow) became possible. Between 6 and 8% of direct flow from total flow was determined for the subcatchments in 1999 (Bogena, 2001). With these results, the calculated surface runoff for the entire catchment of approximately 6% seems plausible.

Of note is the dominance of interflow, with this process accounting for 65% of total runoff; this deserves special consideration. Apart from shallow soils and dense bedrock, macropores are responsible for this effect (Bogena and Diekkrüger, 2001). Generation of macropores may result from permanent usage of land as pasture (Kooistra and Boersma, 1994). Macropores drastically influence runoff dynamics because of their high absorption capacity and quick drainage of water (Bronstert and Plate, 1996). The runoff response to rainfall events at small scales is typically characterized by two peaks (Bogena, 2001). The first quick high peak fed by surface flow is followed by a secondary lower peak fed by interflow through the macropores. This system behavior is reflected by the model (Fig. 11). The fitting of some conceptual parameters to the small-scale model is necessary (Table 3).

The limited consideration of runoff processes by erosion models for the purpose of reducing parameterization effort leads to an inaccurate description of the water balance. The remaining gaps in the amount of moisture accounted for are filled by initial soil moisture, which underlines the importance of this parameter in erosion models. As shown in Fig. 12, WaSiM-ETH allowed for the variation of such system conditions depending on spatial-temporal heterogeneity in connection with other landscape properties (relief, soil, land use). Moreover, the meteorological boundary conditions were considered. This is a clear advantage of WaSiM-ETH over to the application of constants or topographic indices (Beven and Kirkby, 1979) as the latter are commonly used to derive soil moisture in a qualitative manner. Topographic indices are based solely on relief properties and are therefore not suitable for adequately characterizing soil moisture values. Consideration of the spatial variation is limited due to soil and land use influences being neglected in this metric. The tem-

poral dynamic is ignored because meteorological boundary conditions and land use changes are not defined.

The coupling of WaSiM-ETH with other process-based models often allows for the direct transfer of input parameters. This guarantees consistent modeling of the hydrological processes in dynamic systems. Moreover, the parameterization effort is drastically reduced.

The output as raster files guarantees the simple transfer to other raster-based models.

5.3 Extended Runoff Analyses

The application of the modified soil map allowed land use effects on soil hydraulic parameters to be considered. This leads to differences in calculated runoff components (Fig. 13). The model quality is measurably improved without recalibration (Table 5). The improvement is more significant for the Berrensiefen sub-catchment than for the Wahnbach catchment (Table 5). This is probably because there is a smoothing effect of within-stream processes for the entire catchment compared to the dominant effect of surface runoff processes for the sub-catchment.

Surprisingly, if the extended soil map is applied, the surface runoff is doubled for the pasture despite an increase in K_{sat}.

The reason for the increase in K_{sat} is the relief parameters of the pasture. The pasture is located below the cut pasture within the flow path. Due to the increased runoff from the cut pasture, more lateral input of surface runoff reaches the pasture. This finding is important for further research and it should be used to increase the certainty of erosion analyses.

However, because of the small sample size, the application of the extended model should be taken solely as a demonstration. Due to the intensive effort required for field investigations, only two areas could be studied in detail. Therefore, the relationship that was identified had to be applied to all units in the soil map.

Additional uncertainties resulted from using the measurement of differences in K_{sat} instead of using the raw measurements themselves. This was necessary because of the infiltration model we used from Green-Ampt (1911). We also used Richard's equation, which does not consider macropores. WaSiM-ETH considers such influences by using a conceptual bypass model (Jansson and Karlberg, 2001).

It can be concluded that land use effects on soil hydraulic parameters were successfully implemented in this hydrological model. Thus, an extended process description was provided. This application leads to an improvement in model result certainty. With this, a methodical advancement in hydrological modeling was achieved.

5.4 Conclusion and Outlook

This study discusses two different factors related to hydrological modeling. First, the necessity of hydrological analyses for event-based erosion models was

demonstrated. Then, an improvement in the model result certainty was achieved by considering land use effects on soil hydrological properties.

The first component of this research showed that with a continuously working process-based distributed model (WaSiM-ETH), the determination of initial conditions of event-based erosion models is possible. Compared to constants or relief-based indices to represent the initial conditions, this method considers spatial variability and temporal dynamics depending on landscape properties and meteorological conditions. In addition to the derivation of initial conditions, the model was used to calibrate parameters for event-based erosion models. The often-mentioned disadvantage of increasing parameterization effort through the application of a hydrological model was diminished by using a rational model coupling strategy.

In the second part of our work, a newly developed method for considering land use effects on runoff building and runoff was demonstrated. Therefore, the parameterization of the model was improved. Field investigations were conducted to determine necessary input parameters. To validate the transfer of data to the model scale, measurements at different locations as well as temporal replications were conducted. Application of the extended model leads to an improvement in the certainty of the results.

Further studies will answer the question of whether the application of this model will help to improve the certainty of erosion analyses. It is expected that areas that are in a position below the cut pastures will have a higher erosion risk due to increased runoff (Fig. 14).

Moreover, land use patterns, runoff and erosion are caused by rural infrastructure (Duke et al., 2006). Therefore, further studies should investigate the effects of linear structures such as roads or field borders on runoff concentration, runoff and erosion.

Apart from methodological improvements upon this work, the development of risk analyses for management purposes should be the practical goal of further research.

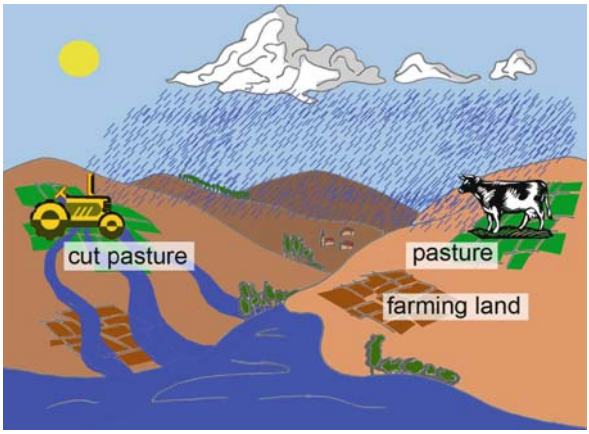

Fig. 14 Hypotheses for future research. Spatial arrangement of pasture and cut pasture is most important concerning erosion risk for adjacent fields

Acknowledgments This research was funded by the German Research Foundation (DFG). We would like to thank O. Rössler, A. Klose, M. Krautblatter and H. Busche for their constructive comments, which helped improved the manuscript considerably. We also would like to thank H. Neumeister and J.-C. Otto for fruitful discussions and H. Bogena for providing fundamental data for this catchment study.

References

Ament, F. and Simmer, C., 2006. Improved representation of land-surface heterogeneity in a non-hydrostatic numerical weather prediction model. Boundary-Layer Meteorology, 121: 153–174.

Beven, K.J. and Kirkby, M.J., 1979. A physically based variable contributing area model of basin hydrology. Bulletin of Hydrological Sciences, 24(1): 43–69.

Bogena, H., 2001. Analyzing and Modelling Solute and Sediment Transport at Different Spatial and Temporal Scales. Dissertation Thesis, Rheinische Friedrich-Wilhelm Universität, Bonn, 182 pp.

Bogena, H. and Diekkrüger, B., 2001. Modelling solute and sediment transport at different spatial and temporal scales. Earth Surface Processes and Landform, 27: 1475–1489.

Bogena, H., Diekkrüger, B., Klingel, K., Jantos, K. and Thein, J., 2003. Analysing and modelling solute and sediment transport in the catchment of the Wahnbach River. Physics and Chemistry of the Earth, 28: 227–237.

Bronstert, A. and Plate, E.J., 1996. Modelling runoff generation and soil moisture dynamics for hillslopes and micro-catchments. Journal of Hydrology, 198: 177–195.

Bruijnzeel, S.L., 2005. Land Use and Land Cover Effects on Runoff Processes: Forest Harvesting and Road Construction. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1813–1827.

Burt, T.P. and Slattery, M.C., 2005. Land Use and Land Cover Effects on Runoff Processes: Agricultural Effects. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1805–1812.

de Roo, A.P.J. and Jetten, V.G., 1999. Calibrating and validating the LISEM model for two data sets from the Netherlands and South Africa. Catena 37: 477–493.

de Roo, A.P.J., Wesseling, C.G., Jetten, V.G. and Ritsema, C.J., 1996. LISEM: A Physically-Based Hydrological and Soil Erosion Model Incorporated in a GIS. In: H.P. Nachtnebel and K. Kovar (Editors), IAHS Proceedings and Reports: Application of Geographic Information systems in Hydrology and Water Resource Management. IHAS Press, Wallingfort, pp. 395–403.

Duke, G.D., Kienzle, S.W., Johnson, D.L. and Byrne, J.M., 2006. Incorporating ancillary data to refine anthropogenically modified overland flow paths. Hydrological Processes, 20(8): 1827–1843.

Durner, W. and Fühler, H., 2005. Soil Hydraulic Properties. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1103–1119.

Endreny, T.A., 2005. Land Use and Land Cover Effects on Runoff Processes: Urban and Suburban Development. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1775–1803.

Feddes, R.A., Kowalik, P.J. and Zaradny, H., 1978. Simulation of Feld water use and crop yield. Simulation Monograph. PUDOC, Wageningen (NL), 189 pp.

Flügel, W.A., 1997. Combining GIS with regional hydrological modelling using hydrological response units (HRUs): An application from Germany. Mathematics and Computers in Simulation, 43: 297–304.

Greenwood, K.L. and McKenzie, B.M., 2001. Grazing effects on soil physical properties and the consequences for pastures: a review. Australian Journal of Experimental Agriculture, 41: 1231–1250.

IPPC, 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment.- Report of the Intergovernmental Panel on Climate, University Press, Cambridge, United Kingdom and New York.

Jansson, P.E. and Karlberg, L., 2001. Coupled Heat and Mass Transfer Model for Soil-Plant-Atmosphere System, Department of Civil and Environment Engineering, Royal Institute of Technology, Stockholm.

Jasper, K., 2005. Hydrological Modelling of Alpine River Catchments using Output Variables from Atmospheric Models, ETH, Zurich, 138 pp.

Kooistra, M.J. and Boersma, O.H., 1994. Subsoil compaction in Dutch marine sandy loams: loosening practices and effects. Soil and Tillage Research, 29: 237–247.

Krause, P., Boyle, D.P. and Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Advanced in Geosciences, 5: 89–97.

Legates, D.R., 1999. Evaluating the use of "goodness of fit" measures in hydrologic and hydroclimatic model validation Water Resources Research, 35: 233–241.

Lørup, J.K. and Styczen, M., 1996. Soil Erosion Modelling. In: M.B. Abott and J.C. Refsgaard (Editors), Distributed Hydrological Modelling. Kluwer Academic Publishers, Dortrecht, pp. 93–120.

Monteith, J.L. and Unsworth, M.H., 1990. Principles of Environmental Physics. Second Edward Arnold, London.

Nash, J.E. and Sutcliffe, J.V., 1970. River flow forecasting through conceptual models – part 1: a discussion of principles. Journal of Hydrology, 10: 282–290.

Nearing, M., Renard, K. and Nichols, M., 2005. Erosion Prediction and Modeling. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1221–1228.

Oke, T.R., 1987. Boundary Layer Climates. Routledge, London and New York, 435 pp.

Peschke, G., 1987. Soil Moisture and Runoff Components from a Physically Founded Approach. Acta hydrophysica, 31(3–4): 191–205.

Rawls, W.J. and Brakensiek, D.L., 1985. Prediction of soil water properties for hydrologic modelling. American Society of Civil Engineers: 293–299.

Refsgaard, J.C., 1996. Terminology, Modelling Protocol and Classification of Hydrological Model Codes. In: M.B. Abott and J.C. Refsgaard (Editors), Distributed Hydrological Modelling. Water Sinces and Technology Library. Kluwer Academic Publishers, Dortrecht, pp. 17–39.

Refsgaard, J.C., Storm, B. and Abbott, M.B., 1996. Comment on ′A Discussion of Distributed Hydrological Modelling by K.J. Beven. In: M.B. Abbott and J.C. Refsgaard (Editors), Distributed Hydrological Modelling. Water Sciences and Technology Library Kluwer Academic Publishers, Dortrecht, pp. 279–295.

Rosbjerg, D. and Madsen, H., 2005. Concepts of Hydrologic Modeling. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 155–163.

Schaap, M.G., 2005. Models for Indirect Estimation of Soil Hydraulic Properties. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1145–1150.

Schmidt, J. and von Werner, M., 1999. Application of the EROSION 3D Model to the CATSOP Watershed, The Netherlands. Catena, 37: 449–456.

Schulla, J. and Jasper, K., 2006. Model Description WaSiM-ETH, ETH, Zurich.

Šimunek, J., 2005. Models of Water Flow and Solute Transport in the Unsaturated Zone. In: M.G. Anderson and J.J. McDonnell (Eds.), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 1171–1180.

Tietje, O. and Hennings, V., 1996. Accuracy of the saturated hydraulic conductivity prediction by pedo-transfer functions compared to the variability within FAO textural classes. Geoderma, 69(1–2): 71–84.

van Genuchten, M.T., 1980. A closed-form equation for predicting the hydraulic conductivity of unsaturated soils. Soil Sciences Society of America, 44(5): 892–898.

Vereecken, H., Maes, H., Feyen, J. and Darius, P., 1989. Estimating the soil moisture retention characteristic from texture, bulk density, and carbon content. Soil Sciences, 148: 389–403.
Vigiak, O., Sterk, G., Romanowicz, R.J. and Beven, K.J., 2006. A semi-empirical model to assess uncertainty of spatial patterns of erosion. Catena, 66: 198–210.
Warrach, K., Stieglitz, M., Mengelkamp, H.T. and Raschke, E., 2002. Advantages of a topographically controlled runoff simulationin a soil-vegetation-atmosphere transfer model. Journal of Hydrometeorology, 3(2): 131–148.
Wigmosta, M. and Prasad, R., 2005. Upscaling and Downscaling: Dynamic Models. In: M.G. Anderson and J.J. McDonnell (Editors), Encyclopaedia of Hydrological Sciences. John Willey & Sons Lds., Chichester, pp. 165–176.
Willmott, C.J., 1981. On the validation of models. Physical Geography, 1: 184–194.

Snow Cover Duration in Relation to Topography in the Loetschental, Switzerland

Susanne Schmidt

1 Introduction

In high mountain environments the spatial distribution of seasonal snow cover depends decisively on topography. Yearly repeated snow cover patterns during snow melt clearly indicate the impact of topography and influence significantly ecological and geomorphological patterns, as well as hydrological and climatic processes. Therefore, an in-depth understanding of the relationship between topography and snow cover is required for an improved knowledge of these systems. Indeed, topographical effects on the spatial distribution of snow cover have been described in many studies, but until now the amount of studies which quantify this effect is poor (Blöschl & Kirnbauer 1992; Kölbel 1984; König & Sturm 1998; Tappeiner et al. 2001). The main problem to improve the uncertainties as well as to quantify the topographical impact is the availability of data with both high spatial and high temporal resolution. In fact, on large scale the observation of snow cover duration can be resolved with satellite data (Goodison et al. 1981; Hall & Martinec 1985; Seidel & Martinec 2004). On this scale satellite images provide a good database to analyse the duration and spatial distribution of snow cover (Winiger et al. 2005). On regional scale, where the snow cover distribution is mainly influenced by topography (McKay & Gray 1981) and varies rapidly over time and space, the usage of satellite images is limited by either coarse spatial and high temporal resolution or vice versa. Moreover, the temporal resolution of satellite images that depends on the repeat cycle of the sensor is additionally reduced by cloud cover. Even in high structured terrain with scattered snow cover patterns high spatial and temporal resolution data are required (Aschenwald et al. 2001; Corripio 2003: S. 46; Seidel & Martinec 2004: S. 34; Tappeiner et al. 2001). Thus, satellite images cannot be used for monitoring and analysing rapid changes of snow cover patterns in high mountain areas.

S. Schmidt (✉)
Südasien-Institut, University of Heidelberg, Im Neuenheimer Feld 330, D-69120 Heidelberg, Germany
e-mail: s.schmidt@sai.uni-heidelberg.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_10,

2 Aims

The aim of this study is to present a new high spatial and temporal resolution technique for monitoring snow dynamics in high mountain areas. This new technique is based on terrestrial images, which were used to analyse the impact of topography on the snow cover duration. For this purpose, a simple statistical regression analysis was derived. Furthermore, the regression equations were applied to model the snow cover duration in the whole catchment. In order to investigate the validity of extrapolation the modelled snow cover was compared with ASTER-satellite data.

3 Study Area

The study area is situated in the Lötschental (Swiss Alps), a northern tributary of the Rhone-Valley (Fig. 1). The valley extends over 160 km^2 and has a main direction from NE to SW. The elevation ranges from 1375 m up to 3200 m a.s.l. at the northern and up to 3400 m a.s.l. at the southern ridge. The highest mountain is the Bietschhorn (3953 m a.s.l.). Due to geological structure the valley has an asymmetrical cross section resulting in an average inclination of 30° on the SW-facing slope and 37° on the NE-facing slope. Thus, on the latter the critical limit for avalanche release is exceeded.

Both slopes are forested up to 2200 m a.s.l. with larch (Larix decidua) and spruce (Picea abies) which are interrupted by unforested tracks of avalanches. Above the timberline alpine meadows exist whereas in higher altitudes exposed

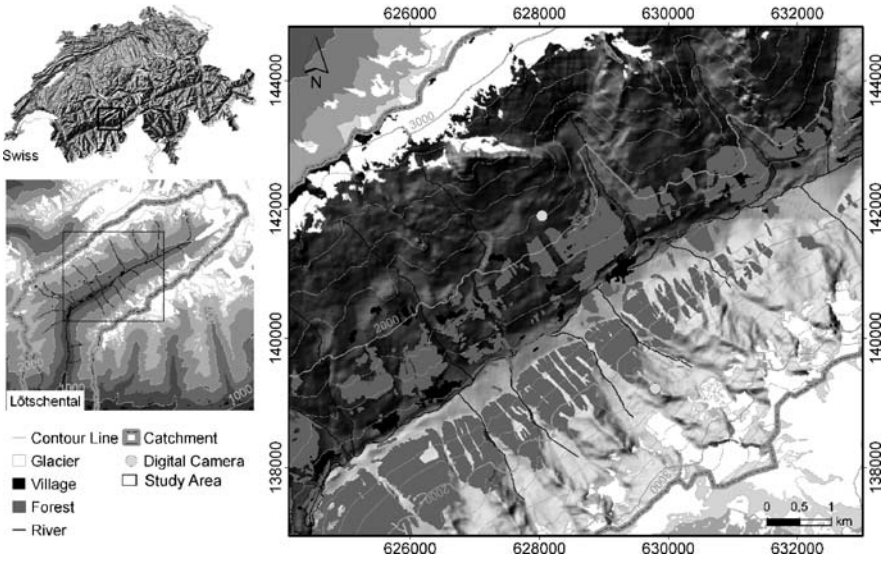

Fig. 1 The study area is situated in the Lötschental (Swiss Alps)

rocks and glaciers dominate. The proportion of the latter amounts 11% for the whole Lötschental (Hörsch 2003).

Based on data from the meteorological station Ried (1500 m a.s.l.), which provided climatic parameters from 1970 up to 1999, a long-term measurement series is available. The long-term average annual temperature is 4.7°C with an annual amplitude of –3.4°C (January) upto 13.8°C (July). From December until February the mean monthly temperature is below zero. Moreover, 50% of the total precipitation of 1,113 mm falls as snow in the valley floor and the proportion increases with altitude.

4 Methods and Data

In order to monitor the snow cover duration with a high spatial and temporal resolution two digital cameras were used. The terrestrial images were orthorectified and classified into snow covered and snow free areas to carry out quantitative analyses of snow cover duration in relation to topography (Fig. 2). A digital elevation model (DEM) was the basis for processes of orthorectification as well as for deriving of topographical parameters. The DEM with a resolution of 10×10 m^2 was generated by the means of geodetic points (SWISSTOPO 2004). To estimate the impact of topographical parameters (elevation, slope, potential solar radiation, curvature and topographical position) on the snow cover duration a simple semi-empirical statistical analysis was deployed. On the basis of this multivariate regression analysis the snow cover duration was modelled for the catchment area. Finally, the semi-empirical statistical model was validated based on the generated snow cover duration maps and selected ASTER-satellite images. The latter enables the investigation of the possibility of model extrapolation.

4.1 Terrestrial Image Processing and Analysis

Two digital camera systems (MetSupport) were mounted on opposite slopes above the timberline to monitor the variations of snow cover distribution on NW- and SE-facing slopes on a daily basis. The terrestrial images representing the SE-facing slope covered an area of 12 km^2 within an elevation range from 1600 m up to 3200 m a.s.l. On the NW-facing slope the covered area is reduced to 9 km^2, which elevation ranges from 1600 m up to 3800 m a.s.l. Using an automatic control system the terrestrial images were taken at predefined times. In order to improve classification results, it was necessary to prevent back light and shadow. Therefore, the image time was set to 12 a.m. for the images monitoring the SE-facing slope and to 8 a.m., respectively 2 p.m., for the images monitoring the NW-facing slope. Due to bad weather conditions, e.g., clouds, snowfall and fog, the temporal resolution of images which could be used for further analyses was restricted to approximately once every 5 days.

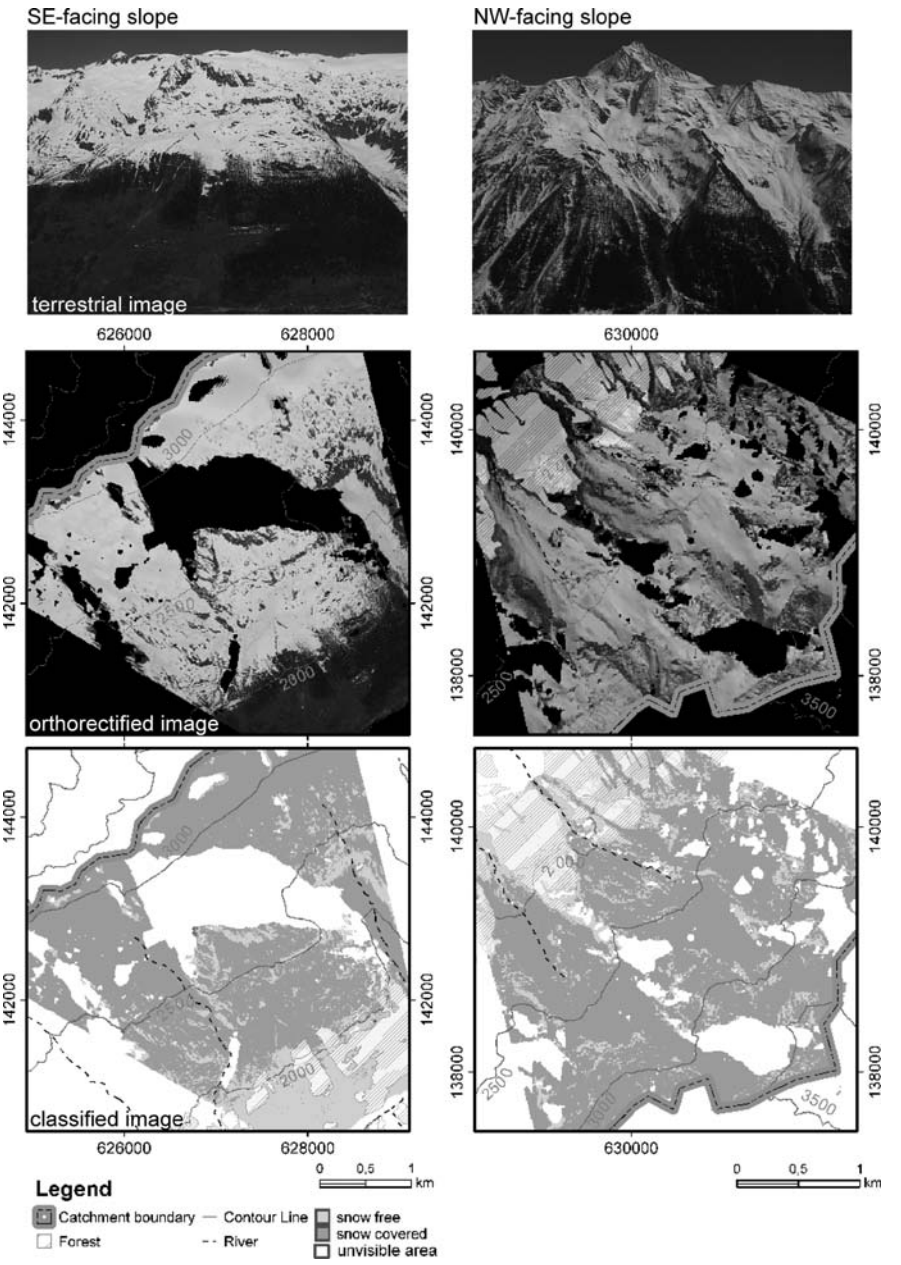

Fig. 2 Terrestrial image processing: The terrestrial images (*top*) were orthorectified (*middle*) and classified (*bottom*)

For further spatial analyses the terrestrial images have to be orthorectified using, Corripio's program "Georeferencing Terrestrial Photography" (Corripio 2004). Input parameters for the computation of the orthorectified images were:

– a DEM with a resolution of 10 m × 10 m (SWISSTOPO 2004),
– the orientation data of images, and
– the size of the CCD chip (KODAK camera manufacturer).

To achieve a high accuracy of the topographical dependent analysis a good consistency between the DEM and the orthorectified images is necessary. Orientation parameters were computed for each image individually, because the parameters may change from image to image due to climatic influences such as wind, snow and temperature variation. The orientation data for every image consisted of:

– three rotation parameters for spatial rotations,
– coordinates of the camera position in a ground control system (three parameters), and
– principle distance (approximately the focal length of the lens) (Schmidt & Weber 2008).

To compute the orientation data for each image a semi-automated orientation strategy was developed by B. Weber (Institute for Photogrammetry). The advantage of this developed strategy is that parts of the time-consuming manual and error-prone process are done automatically. Therefore, manual measurements are required only in few selected reference images which were to be matched with the other images. By means of image matching and on the basis of the reference images the orientation parameters could be calculated for each image (Schmidt & Weber 2008).

Finally, the orthorectified images were classified into snow free and snow covered areas. Because the CCD-Chip of the camera is only sensitive to the visible light, none of standardized classification methods, which additionally use spectral bands of the near-infrared, could be applied (Hinkler et al. 2002). Due to higher reflection properties of snow in the visible radiation spectrum compared to snow free objects a process of thresholds can be used to identify snow covered and snow free areas. To improve the classification results masks of shadow, which were modelled separately for each image by means of the DEM and the sun position, and different ground features such as geology and vegetation were applied. Therefore, different thresholds were applied to classify the images into snow free and snow covered areas. For the classification the Red-Green-Blue colour (RGB-Colour) was transformed to Hue, Lightness and Saturation colour (HLS-Colour), so that the spectral properties of objects are represented in a single band. The classification was carried out in ENVI, the resulting binary snow cover maps were export to ArcGIS and summarized to snow cover duration maps which were used for further statistical analyses (Schmidt 2007).

4.1.1 Statistical Analyses and Modelling

The impact of topography on the snow cover duration was quantified by multiple linear regression analyses (Chang & Li 2000). The aim of these statistical analyses was to develop a semi-empirical model, which describes the snow cover patterns (dependent variable *Y*) in relation to topographical parameters such as elevation, slope, curvature (independent variables X_i) in a simple form (s. Eq. 1)

$$Y = \alpha + \beta_1 X_1 + \beta_2 X_2 + ... + \beta_m X_m + \varepsilon \quad (1)$$

In equation 1, α means the intercept value and the RMS-Error. The partial regression coefficients β_i can be regarded as a measure of the weight of a single parameter X_i on the dependent variable *Y*. Due to different scales all topographic features were standardized to enable comparability (Bahrenberg et al. 2003: S. 33).

All used topographic parameters were derived from the DEM. 20% of terrestrial image pixels were selected randomly for the regression analyses that the model can be subsequently validated. The random selection was carried out by the application HAWTH's ANALYIS TOOLS V. 3.21. In the analyses the following areas were ignored: areas free of snow, forests, and areas which were snow covered on the last image. Moreover, the statistical analyses were conducted for both slopes separately due to different snow cover performances. Based on these multiple linear regression equation the snow cover was modelled for the Loetschental.

4.2 Model Validation

For a quantitative validation the modelled snow cover duration was compared with snow cover maps derived from the terrestrial images. For this purpose maps were calculated, by substraction of both binary maps from each other. The resulting deviation map contains information on accuracy predicted snow cover distribution and its spatial inaccuracies. Values close to zero indicate a good accordance; values greater than zero represent an overestimation and smaller than zero an underestimation of snow cover duration.

Furthermore, in order to show the potential use of terrestrial images and the possibility of extrapolation the modelled snow cover duration maps were compared with available ASTER-satellite images. For this purpose only cloudless ASTER-data could be used, so that their number was restricted to four images which were taken on 26th March, 19th May, 27th July and 8th September 2004. In this paper, the visual model validation is exemplified for the 26th March and 19th May.

5 Results

The snow cover maps (Fig. 3) which were derived from the orthorectified and classified terrestrial images indicate rapid changes of depletion patterns and highlight, the necessity of the application of data sets with high temporal and spatial res-

olution. The patchy snow cover on both slopes shows that the elevation dependant temporary snow line is strongly alternated by slope angle, slope aspect and curvature. Thus, very steep slopes and ridges are mainly free of snow during the whole winter and release zones of avalanches are snow-free very early in the year. Contrary to these parts, concave and flat topographical positions are characterized by comparatively long snow cover duration. Within both slopes comparable relief positions reveal much longer snow cover duration on the NW-than on the SE-facing slope. Moreover, snow re-distribution by avalanches turns out to be an important factor on the steep NW-facing slope, which slope angle averages more than 37°.

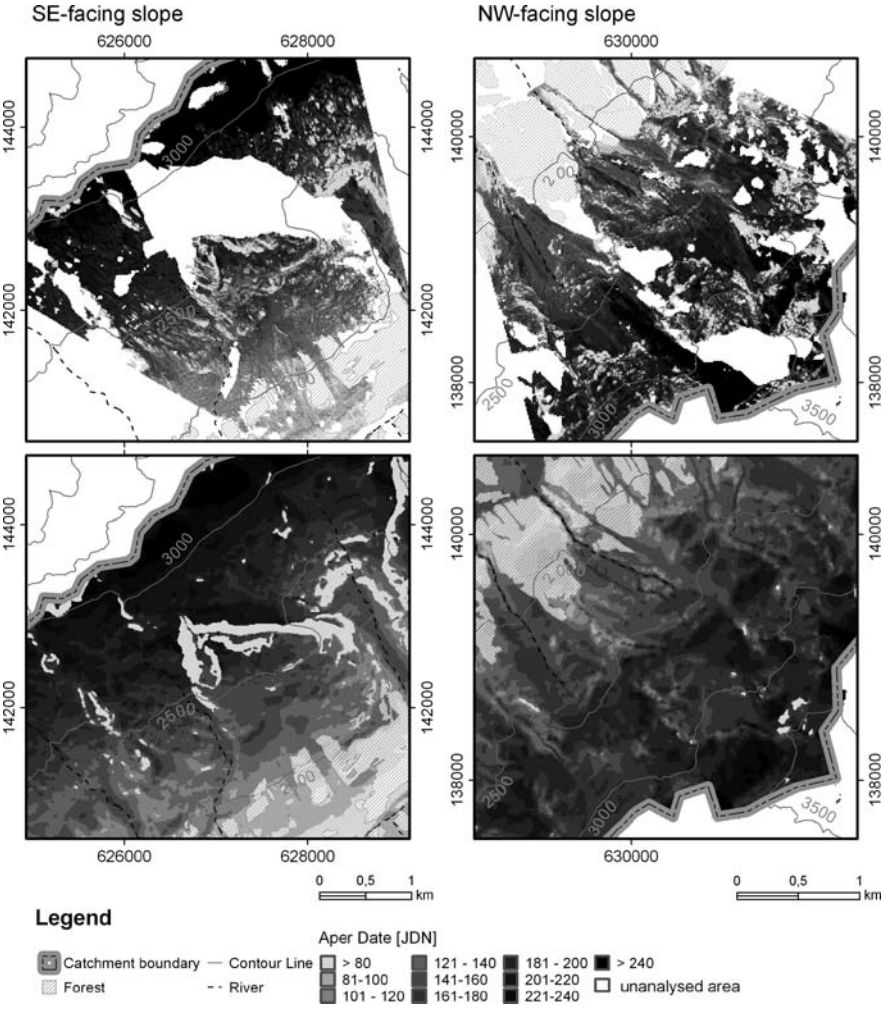

Fig. 3 Map of snow cover duration derived from terrestrial images (*top*) and modelled snow cover duration (*bottom*)

Due to the different snow cover characteristics on both investigated slopes, statistical analyses were separately calculated for each slope. The analyses show that snow cover patterns can effectual be described by the used topographic parameters (elevation, slope angle, potential radiation, topographical position, curvature). On the SE-facing slope 74.87% of the snow cover pattern is explainable by topographic features, whereas the snow cover pattern on the NW-facing slope is explainable to a proportion of only 36.02%. The standardized regression coefficients identify the elevation and slope angle as keyfactores on the NW-facing slope, while on the SE-facing slope elevation is the only superior factor controlling the snow cover patterns. All other topographical parameters are of minor effect on both slopes (Table 1).

The quantitative model validation (Fig. 4) shows that the snow duration was modelled within a temporal bias of ± 20 days correctly in 60% of the total area, whereas, the proportion is greater on the SE-facing slope (65%) than on the NW-facing slope (50%). By contrast, the snow cover duration is underestimated only in 26% of the total area. These areas are located predominantly in accumulation zones of avalanches, so that the proportion amounts to 20% on the NE-facing slope compared to 12% on the SE-facing slope (Table 2).

The visual comparison of the modelled snow cover map to the observed snow cover distribution by ASTER-images indicates similar characteristics. These similar patterns were exemplified by the means of visual comparison of two selected ASTER-images from 26th March 2004 and 19th May 2004 with the modelled snow cover.

The ASTER-image from 26th March 2004 represents snow patterns at the beginning of snow melt. The snow cover is nearly closed; only very steep slopes and some lower parts of the catchment are snow free. Furthermore, in the lower parts the snow cover is very thin so that partly vegetation is visible. The visual comparison between

Table 1 Regression analyses of snow cover duration in relation to topographical parameters

Intercept	Elevation [m]	Slope [°]	Pot. Rad. [W m^{-2}]	Topo scale	Min. Curv.	Max. Curv.	RMS	Adj. R^2
NW-facing slope 2003/04								
96.08	0.0627[4)]	–2.067[4)]	–0.0004[4)]	–0.083[4)]	927.9[4)]	–758.1[4)]	32.75	36.02*
	0.6381	–0.4938	–0.1204	–0.2097	0.1376	–0.1335	0.79	
SE-facing slope 2003/04								
–55.95	0.1115[4)]	–1.246[4)]	–0.0003[4)]	–0.130[4)]	697.3[4)]	–454.1[4)]	24.51	74.87*
	0.8655	–0.2186	–0.0449	–0.189	0.0635	–0.0542	0.50	
Both slopes 2003/04								
–10.23	0.0933[4)]	–1.550[4)]	–0.0001[4)]	–0.120[4)]	67.5[4)]	–47.8[4)]	28.65	60.84*
	0.7949	–0.3314	–0.0875	–0.0218		–0.0674	0.63	

Upper number: regression coefficient
Lower number: standardized regression coefficient
Level of significance of partial regression coefficient [1)]10%,[2)]5%, [3)]1%,[4)]0.1%
*Significant on 5%-level (F-distribution)

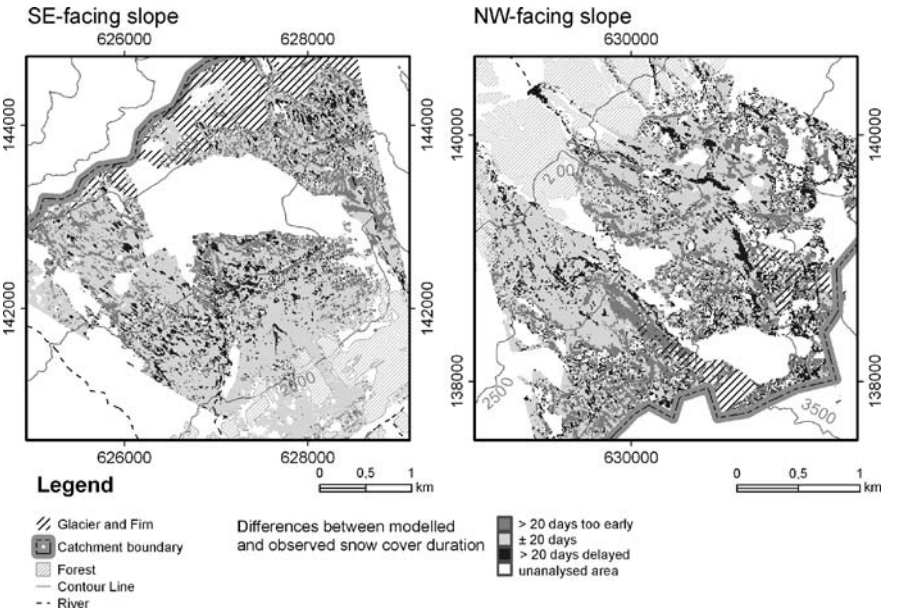

Fig. 4 Difference snow cover maps between modelled snow cover duration and by means of terrestrial images derived snow cover duration

Table 2 Proportion of underestimated, correctly and overestimated modelled snow cover duration

Area [%]	> 20 days earlier	± 20 days	> 20 days later
SE-facing slope	12	65	23
NW-facing slope	20	50	30
Total	14	60	26
Characteristics of topographical position	Accumulation zones of avalanches		Steep rock faces Track of avalanches

the modelled and observed snow cover also shows a nearly closed snow cover. The steep snow free slopes were modelled correctly. In the lower regions, where the snow melt has already started the snow duration was partly underestimated by the model (Fig. 5).

On 19th May 2004 the scattered temporary snow line was located above 2000 m a.s.l. On SW-facing slopes the snow line reaches an altitude up to 2500 m a.s.l. In contrast, slopes with flat angle and accumulation zones of avalanches are characterized by longer snow cover duration compared to neighbouring areas. Therefore, in accumulation zones of avalanches on the NW-facing slope the snow cover exists down to 1500 m a.s.l. These scattered shapes of snow line were also modelled correctly within a bias of ± 20 days. Only on flat slopes of the NW-facing slope a significant overestimation of snow cover duration is observable (Fig. 6).

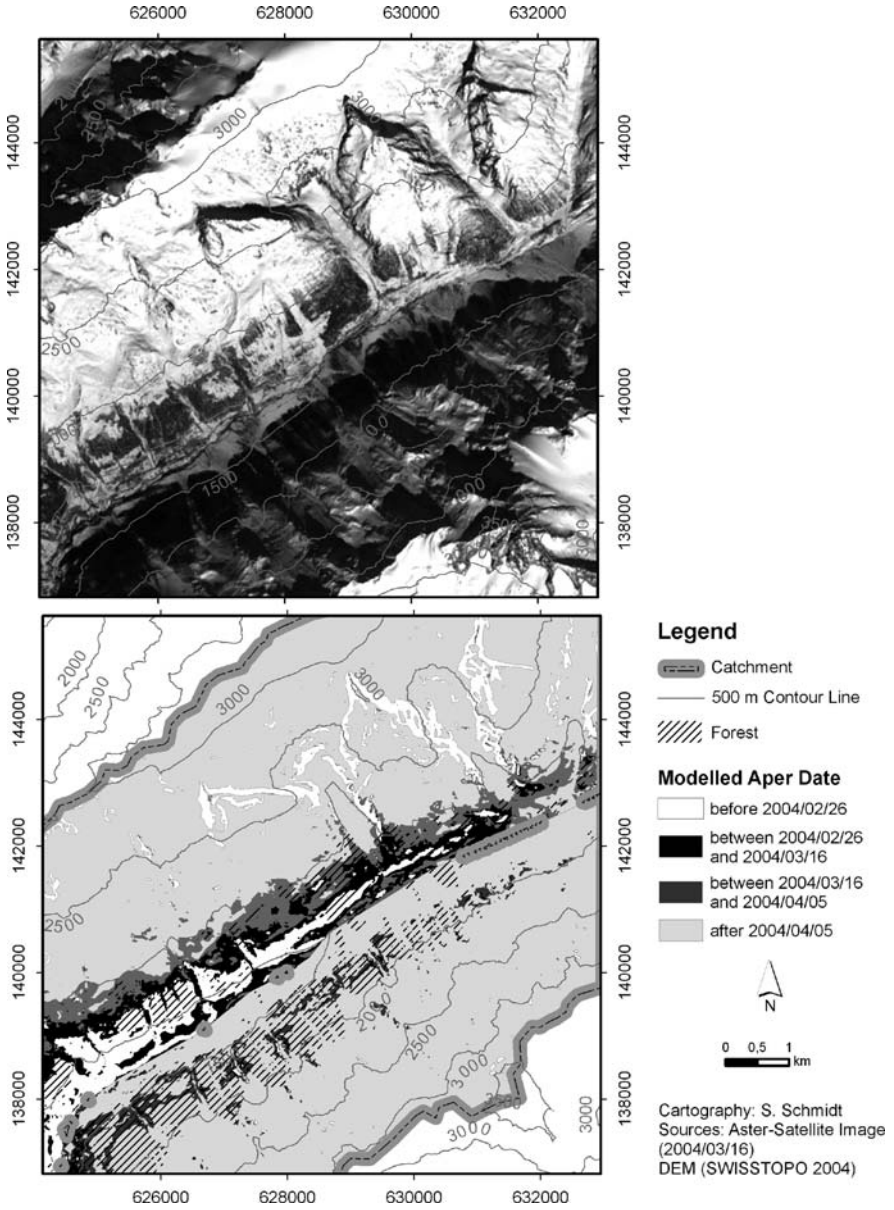

Fig. 5 Model validation by means of ASTER-Satellite image (2004/03/16)

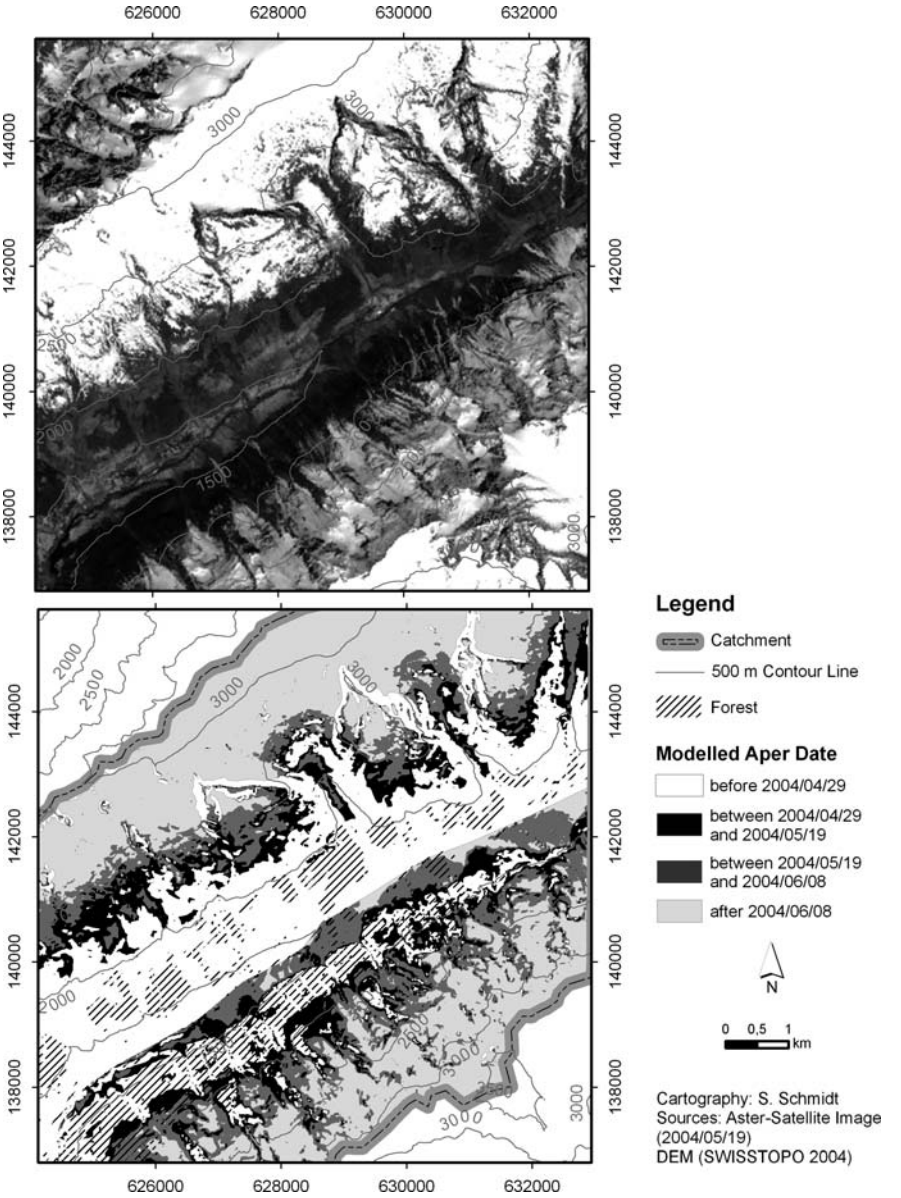

Fig. 6 Visual model validation by means of ASTER-Satellite image (2004/05/19)

6 Discussion

The study shows that snow cover patterns cannot be described by topographic features alone. The main structures of the pattern are expressed, but the temporal and spatial variations are not covered correctly. The terrestrial images indicate that avalanches modify the main structure considerably, even at the beginning of the snow melting period. Consequently, avalanches effect hydrological, geomorphological and ecological processes (Brugger 2004; de Jong et al. 2005; Phillips 2000; Rychetnik 1984). According to Blöschl & Kirnbauer (1991) and Tappeiner et al. (2001), these gravitative processes can neither be described by first order topographic features like slope aspect, angle and elevation nor by secondary features like curvature and topographic position. Hence, topographic features are to be developed that are capable of coping with these processes. First methodological approaches were done by Maggioni et al. (2005) who used slope length besides primary parameters.

Moreover, snow cover variations are influenced by turbulent and advective heat fluxes which originated from snow free areas (Stähli & Jansson 1998). These snow free areas offer a lower albedo compared to snow covered areas resulting in a higher energy budget for snowmelt (Marsh 1999). Depletion maps generated from terrestrial images show that melting propagates from the snow free areas, i.e., rocks, pikes, and release areas of avalanches. Additionally, snow cover patterns are modified by other parameters such as vegetation cover, glaciers and permafrost. Hence, variations of snow cover patterns are mainly controlled by snow height, which cannot be referred to by topographical features alone (Barry 1992; Bernhard & Weibel 1999; Holtmeier & Broll 1992; Löffler 2005; Luce et al. 1999; Phillips 2000).

7 Conclusion

The application of terrestrial images enables the monitoring of snow cover distribution with a high spatial and temporal resolution. This resolution cannot be achieved by the usage of conventional satellite images. Therefore, the application of terrestrial images allows the gap-filling between the resolution of observation data and the required resolution of data to analyse the snow cover distribution. The developed semi-automatic process facilitates a high spatial precision of the terrestrial images that is a crucial requirement for topographical analyses. Moreover, the orthorectification of each image allows using classification processes which are additionally using mask of shadow to improve the classification results. Furthermore, the possibility of model extrapolation within the bias highlights the potential of the application of terrestrial images.

The study shows that due to “topography-independent” variation, the snow cover pattern can be explained by topographical parameters only by approximation. Thus, on the NW-facing slope 36% of the snow cover duration can be explained by topographical parameters, whereas the proportion increased to 74% on the SE-facing

slope. Hence, the exact prediction of snow cover distribution will remain limited (Blöschl et al. 1991). But, with the use of spatial and temporal high resolution data these uncertainties can be recorded. Further topographical parameters, as well as algorithms to describe neighbouring relationships, are required to predict the snow cover distribution.

Acknowledgments This study was financed by a PhD fellowship of the German Research Group (Deutsche Forschungsgemeinschaft DFG). My special thanks goes to Bernhard Weber, who developed the semi-empirical process to orthorectify the images and who always supported me in the photogrammetic methods. Furthermore, I would like to thank Matthias Winiger, Richard Dikau and Carmen de Jong for their fruitful discussions.

References

Aschenwald, J., K. Leichter, E. Tasser & U. Tappeiner (2001): Spatio-temporal landscape analysis in mountainous terrain by means of small format photography: a methodological approach. *IEEE Transactions on Geoscience and Remote Sensing*, 39, S. 885–893.

Bahrenberg, G., E. Giese & J. Nipper (2003): Statistische Methoden in der Geographie 2. Multivariate Statistik. Stuttgart. B.G. Teubner, 415 S.

Barry, R. G. (1992): Mountain weather and climate. London. Routledge, 313 S.

Bernhard, L. & R. Weibel (1999): Modelling snowmelt using a digital terrain model and GIS-based techniques, In: R. Dikau & H. Saurer, [Eds.]: GIS for earth surface systems: analysis and modelling of the natural environment. Stuttgart, S. 25–47.

Blöschl, G. & R. Kirnbauer (1991): Point snowmelt models with different degrees of complexity – internal processes. *Journal of Hydrology*, 129, S. 127–147.

Blöschl, G. & R. Kirnbauer (1992): An analysis of snow cover patterns in a small alpine catchment. *Hydrological Processes*, 6, S. 99–109.

Blöschl, G., R. Kirnbauer & D. Gutknecht (1991): A spatially distributed snowmelt model for application in alpine terrain, In: H. Bergmann, H. Lang, W. Frey, D. Issler & B. Salm, [Eds.]: Snow, hydrology and forests in high alpine areas (Proceedings of the Vienna Symposium). IAHS 205, S. 51–60.

Brugger, S. (2004): Lawinen schaffen Lebensraum. Größere Artenvielfalt in Lawinenzügen. *Die Alpen*, 1, S. 29–31.

Chang, K.-T. & Z. LI (2000): Modelling snow accumulation with a geographic information system. *International Journal of Geographical Information Science*, 14, S. 693–707.

Corripio, J. (2004): Snow surface albedo estimation using terrestrial photography. *International Journal of Remote Sensing*, 25, S. 5705–5729.

Corripio, J. G. (2003): Modelling the energy balance of high altitude glacierised basins in the Central Andes. PhD-Thesis, University Edinburgh. 151. Unpublished.

De Jong, C., P. Ergenzinger, M. Borufka, A. Köcher & M. Dresen (2005): Geomorphological Zoning: An improvement to coupling alpine hydrology and meteorology?, In: C. De Jong, D. Collins & R. Ranzi, [Eds.]: Climate and hydrology in mountain areas. London, S. 247–260.

Goodison, B. E., H. L. Ferguson & G. A. Kay (1981): Measurement and data analysis, In: D. M. Gray & D. H. Male, [Eds.]: Handbook of snow. Principles, Processes, Management & Use., S. 194–274.

Hall, D. K. & J. Martinec (1985): Remote sensing of ice and snow. London, 189 S.

Hinkler, J., S. B. Pedersen, M. Rasch & B. U. Hansen (2002): Automatic snow cover monitoring at high temporal and spatial resolution, using images taken by a standard digital camera. *International Journal of Remote Sensing*, 23, S. 4669–4682.

Holtmeier, F.-K. & G. Broll (1992): The influence of tree islands and microtopography on pedoecological conditions in the Forest-Alpine Tundra Ecotone on Niwot Ridge, Colorado Front Range, U.S.A. *Arctic and Alpine Research*, 24, S. 216–228.

Hörsch, B. (2003): Zusammenhang zwischen Vegetation und Relief in alpinen Einzugsgebieten des Wallis (Schweiz). Ein multiskaliger GIS- und Fernerkundungsansatz. *Bonner Geographische Abhandlungen*, 110, Bonn, S. 256 S.

Kölbel, H. (1984): Die Schnee-Ausaperung im Gurgler Tal (Ötztal/Tirol): Salzburger Geographische Arbeiten, 12. Salzburg, 214 S.

König, M. & M. Sturm (1998): Mapping snow distribution in the Alaskan Arctic using aerial photography and topographic relationships. *Water Resources Research,*34, S. 3471–3483.

Löffler, J. (2005): Snow cover dynamics, soil moisture variability and vegetation ecology in high mountain catchments of central Norway. *Hydrological Processes*, 19, S. 2385–2405.

Luce, C. H., D. G. Tarboton & K. R. Cooley (1999): Subgrid parameterization of snow distribution for an energy and mass balance snow cover mode. *Hydrological Processes,*13, S. 1921–1933.

Maggioni, M., U. Gruber & A. Stoffel (2005): Definition and characterisation of potential avalanche release areas. http://gis.esri.con/library/userconf/proc02/pa1161/p11(Letzter Aufruf: 05.08.2006).

Marsh, P. (1999): Snowcover formation and melt: recent advances and future prospects. *Hydrological Processes,*13, S. 2117–2134.

McKay, G.A. & D.M. Gray (1981): The distribution of snowcover. In: D.M. Gray & D.H. Male (Eds.): Handbook of snow. Principles, Processes, Management & Use, 153–190.

Phillips, M. (2000): Influence of snow supporting structures on the thermal regime of the ground in alpine permafrost terrain: Eidgenössisches Institut für Schnee- und Lawinenforschung. Davos, 146 S.

Rychetnik, J. (1984): Methoden der Erfassung und Auswertung der Ausaperung und Lawinenaktivität an einem Lawinenhang, In: H. M. Brechtel, [Ed.]: Snow hydrologic research in Central Europe, 162, IAHS, S. 153–165.

Schmidt, S. (2007): Die Reliefabhängige Schneedeckenverteilung im Hochgebirge – ein multiskaliger Methodenverbund am Beispiel des Lötschentals (Schweiz). Dissertation. Universität Bonn. http://hss.ulb.uni-bonn.de/diss_online/math_nat_fak/2007/schmidt_susanne, 187 S.

Schmidt, S. & B. Weber (2008): The dilemma of resolution and seasonality of snow cover in alpine environments, In: U. Strasser, [Ed.]: Research report: Alpine snow workshop: National park research report, S. 103–110.

Seidel, K. & J. Martinec (2004): Remote sensing in snow hydrology- runoff modelling, effect of climate change. Heidelberg, 150 S.

Stähli, M. & P.-E. Jansson (1998): Test of two SVAT snow submodels during different winter conditions. *Agricultural and Forest Meteorology* 92, S. 29–41.

SWISSTOPO (2004): DHM 25 Das digitale Höhenmodell der Schweiz, Zürich.

Tappeiner, U., G. Tappeiner, J. Aschenwald, E. Tasser & B. Ostendorf (2001): GIS-based modelling of spatial pattern of snow cover duration in an alpine area. *Ecological Modelling*, 138, S. 265–275.

Winiger, M., M. Gumpert & H. Yamout (2005): Karakorum – Hindukush-western Himalaya: assessing high-altitude water resources. *Hydrological Processes*, 19, S. 2329–2338.

Sediment Transfer in Steep Upland Catchments (Northern England, UK): Landform and Sediment Source Coupling

Jeff Warburton

Abstract Headwater catchments have been viewed as active geomorphic environments which generate high sediment yields, however, direct measurements in such settings are rarely undertaken. The dynamics of sediment supply and transfer in steep, UK upland catchments are assessed in this chapter using a series of detailed short-term sediment budget studies. It is demonstrated that sediment transfer in these catchments is relatively inefficient and high rates of erosion and sediment transport only manifest themselves at the local scale. Sediment yields are episodic and often associated with extreme rainfall events. Low sediment fluxes are usually the result of poor coupling between slopes and channels; and headwater and lowland sediment systems. It is concluded that the understanding of upland sediment systems requires full-component sediment budgets carried out at nested scales which link headwater and valley sediment systems.

Keywords Catchment sediment budgets · Slope-channel coupling · Sediment yield · Landslides · Bank erosion

1 Introduction

Upland and mountain headwater catchments have traditionally been viewed as active geomorphic environments with some of the highest global specific sediment yields (Dietrich and Dunne, 1978; Church and Slaymaker, 1989; DeBoer and Crosby, 1996). Steep slopes, high runoff, widespread cryospheric activity, thin vegetation covers and active geomorphic processes all contribute to high rates of sediment production and transfer, particularly during extreme events (Johnson and Warburton, 2002a, b). These processes result in a suite of characteristic landform

J. Warburton (✉)
Department of Geography, Durham University, Science Laboratories,
South Road, Durham DH1 3LE, UK
e-mail: jeff.warburton@durham.ac.uk

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_11,

assemblages (e.g., landslide scars, debris flow tracks, incised river channels) that are often interpreted as indicative of high rates of sediment flux (Wells and Harvey, 1987). Such evidence of headwater erosion is often used to infer potential downstream sedimentation problems (Lawler et al., 1999; Orr et al., 2004). This assumption is implicit in the coupling of landforms in sediment cascade models (Caine, 1974; Barsch and Caine, 1984, Fig. 1). In such models, it is important to critically assess whether landform occurrence (frequency) provides reliable evidence for estimating process activity and determine whether there is a direct link between sediment sources and sinks.

However, without direct measurement of rates of sediment flux and estimates of coupling between hillslope and channel processes such relationships cannot be easily determined (Walling and Collins, 2000; Lawler et al., 2006). Figure 1 provides a conceptual model of the geomorphic processes that characterise UK upland sediment budgets. The general framework shown in Fig. 1 is common to many sediment budget studies (e.g., Dietrich and Dunne, 1978; Evans and Warburton, 2005) and the overall structure is divided between hillslope and channel processes with the steambank dividing these two landscape domains. In the UK uplands, a variety of hillslope erosion processes operate but in general these are dominated by water-driven erosion (Russell et al., 2001; Evans and Warburton, 2007). Stream channel

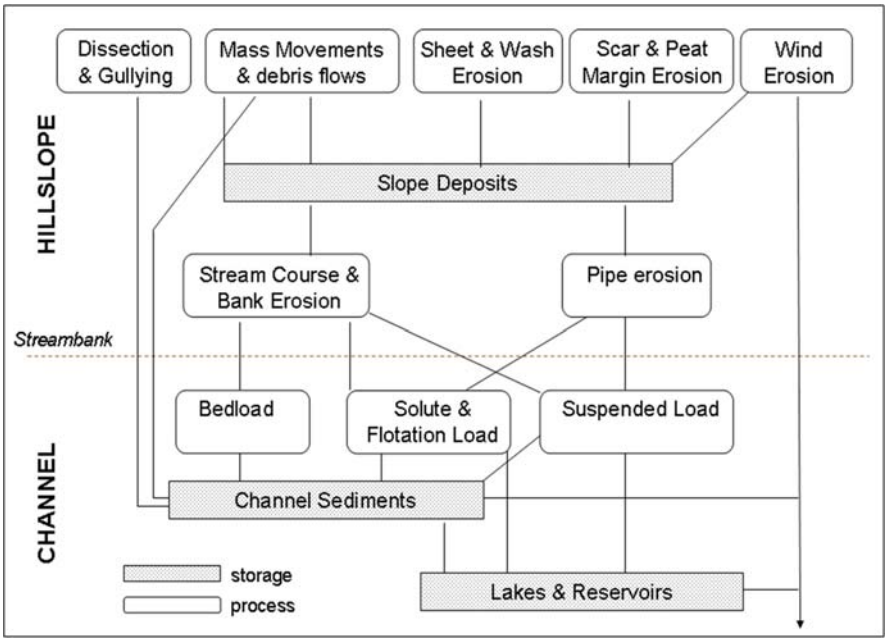

Fig. 1 The upland catchment sediment system. Sediment budget framework showing linkages between main landscape domains, storage components and geomorphic processes operating in steep UK upland catchments. Flotation load refers to debris (peat, vegetation) with a density close to 1.0

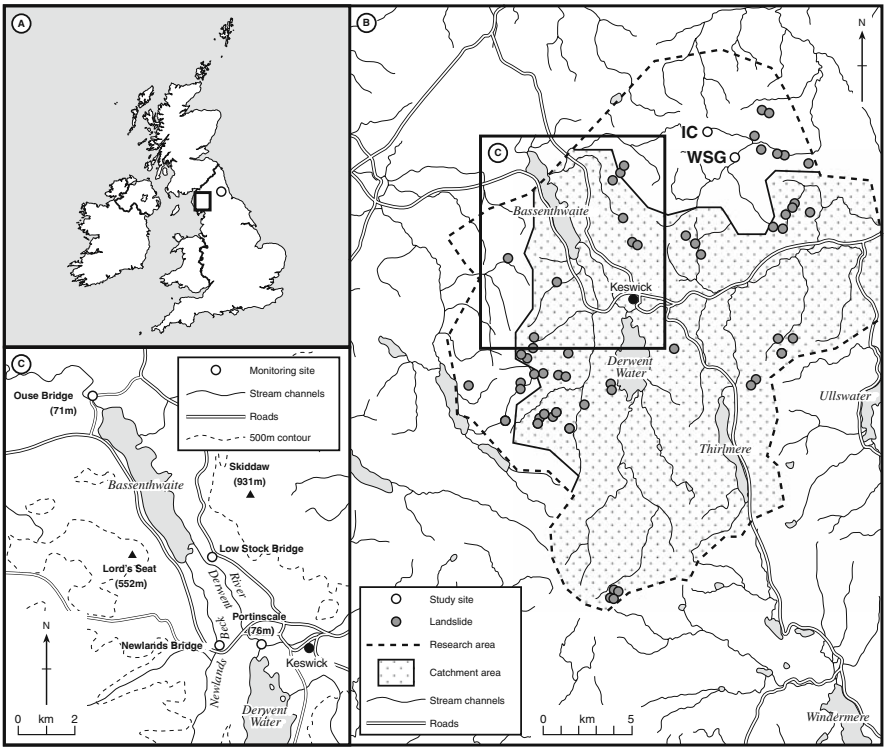

Fig. 2 Study area location map. (**a**) Location of Bassenthwaite catchment and Catcleugh Reservoir sites in Northern England. (**b**) Detail showing Bassenthwaite catchment, location of significant landslides (January 2005) and sites of detailed sediment budget studies referred to in the text (IC, Iron Crag and WSG, Wet Swine Gill). (**c**) Fine sediment flux monitoring sites in the Bassenthwaite catchment lowlands

sediment fluxes are dominated by suspended sediment discharge but during extreme events, especially in headwater locations, bedload may contribute significantly to the total sediment load (Ashbridge, 1995). In peatland areas large blocks of low density peat and turf may be transported as a "floatation load" (Evans and Warburton, 2007). This paper examines the importance of landform and sediment source coupling in controlling the sediment flux from steep upland catchments in Northern England (UK). Two main issues are examined: coupling of hillslope and channel processes; and the efficiency of the delivery of sediment from the headwaters to the lower catchment system. This is illustrated using a number of contemporary sediment budget studies which demonstrate sediment transfer is often an inefficient and poorly coupled process. However, it is important to recognise at the outset, that from a global perspective UK upland and mountain catchments have relatively low rates of geomorphic activity and small sediment yields (typically 10–50 t km^{-2} yr^{-1} (Ledger et al., 1974; Holliday et al., 2008)). Nevertheless, from a local perspective

erosion is still a significant problem and needs to be effectively managed. The specific objectives of this study are to:

1. Quantify the geomorphological contribution of the landslide events in terms of hillslope sediment delivery to the Bassenthwaite catchment fluvial system.
2. Determine the significance of channel and river bank erosion in contributing fine sediment load to the Bassenthwaite lowland fluvial system.
3. Assess the importance of headwater sediment dynamics and connectivity within UK (Northern England) upland catchments.

An important goal is to integrate several studies, carried out at a number of different temporal and spatial scales, to provide an overview on catchment sediment delivery. The case studies considered in this paper focus on the large Bassenthwaite upland catchment (347 km^2) in the Northern Lake District and three small headwater catchment sediment budget studies at Iron Crag, Wet Swine Gill (both in Cumbria) and Chattlehope Burn (Northumbria) (Fig. 2). Figure 1 provides a useful general framework in which process measurements can be undertaken and sediment fluxes compared. Detail of these case studies are given in the description of the study area characteristics and previously published work (Johnson and Warburton, 2002a, 2006a, b; Johnson et al., 2008).

2 Study Area Characteristics

2.1 Bassenthwaite Lake Catchment

Bassenthwaite Lake is the 4th largest lake (by area) in the English Lake District. It has the largest catchment area (347 km^2) of any of the Lake District lakes, a large relative relief and high annual precipitation (> 2000 mm yr^{-1}). The lake has an area of 5.28 km^2 but is the shallowest of the large Lake District lakes (mean depth 5.3 m, maximum depth 20 m) with two thirds of the lake being less than 5 m deep (Thackeray et al., 2006). Hence the catchment area is large compared to the lake volume. The lake is only weakly stratified (warm, monomictic) and susceptible to wind mixing (Bloesch, 1995; Hall et al., 2001). This often results in variable turbidity which is thought to be due to resuspension of lake bed sediments (Parker et al., 1999) in the shallow water environment. The mean hydraulic residence time of the lake is only 19 days although there is a marked seasonal pattern to discharge with a minimum in July and maximum in winter.

The catchment is underlain by Skiddaw Slates to the north and Borrowdale Volcanics (andesitic lava and tuff) to the south. The landscape has been heavily affected by Quaternary glaciation resulting in steep slopes and glacial debris slopes which have been subsequently modified by periglacial activity (Boardman, 1992). These Ordovician rocks and deposits are overlain mainly by nutrient and base-poor mountain soils. The land use of the catchment is dominated by upland moor (53%) and

improved pasture (21%) (Thackeray et al., 2006). Upstream of Bassenthwaite Lake are two other large lakes (Derwent Water and Thirlmere, Fig. 2) which both act as significant sediment traps for sediment transported from the upper catchment. Despite this there has been growing speculation that recent increases in lake sedimentation (primarily related to the flux of inorganic material; Cranwell et al., 1995, Bennion et al., 2000) has lead to a decline in the ecology of the lake (Orr et al., 2004).

The organic content of the lake sediments varies between 9–15% dry mass suggesting a predominantly inorganic input of material from the catchment (Bennion et al., 2000). Thackeray et al. (2006) summarise the available evidence on sediment sources within the Bassenthwaite catchment. They rely heavily on the analysis of Orr et al. (2004) and Nisbet et al. (2004) who carried out a geomorphological assessment of sediment delivery in the Bassenthwaite catchment. This included a fluvial audit of 110 km of stream channel and an aerial photography assessment of catchment erosion features. Orr et al. concluded that patches of eroded ground on the high mountains represented the largest potential sediment source and approximately 7% of the channels surveyed showed a high potential for fine sediment delivery; particularly in the Derwent floodplain area between Derwent Water and Bassenthwaite Lake (Fig. 2, Portinscale to Low Stock Bridge). Both Thackeray et al. (2006) and Orr et al. (2004) conclude that further research is necessary to quantify the significance of these potential sediment sources and particular attention should be paid to the streams entering the lake directly. Furthermore, Thackeray et al. (2006) also suggest that the highly episodic inflows of suspended sediment into the lake may be caused by mass movements in the catchment that deliver large quantity of new sediment into the river system (Hall et al., 2001). These processes are particularly problematic in the Bassenthwaite catchment where excess fine suspended sediment concentrations have been related to decline in the vendace (*Coregonus albula*), an endangered fish species only found in Bassenthwaite Lake and Derwent Water (Atkinson et al., 1989; Winfield and Durie, 2004; Winfield et al., 2004), and also have wider implications for riverine fish (Grieg et al., 2005). Although there is a considerable body of past research documentation sedimentation and environmental changes in Cumbrian lakes (Pennington, 1981; Oldfield and Appleby, 1984; van der Post et al., 1997; Anderton et al., 1998; Shen et al., 2008) surprisingly studies on Bassenthwaite have been rather limited (Parker et al., 1999; Bennion et al., 2000).

2.2 Headwater Catchments

Three small study catchments, Iron Crag, Wet Swine Gill (both in Cumbria) and Chattlehope Burn (Northumbria, Northeast England) (Fig. 2) provide information on headwater sediment dynamics. The three catchments have been sites of detailed sediment budget studies and are used here to illustrate catchment sediment fluxes over different timescales and for different types of event: annual sediment budget (Iron Crag, Johnson and Warburton, 2002a); impact of a large discrete slope failure

Table 1 Catchment characteristics of the summary sediment budget sites (Fig. 8)

	Iron Crag, Cumbria	Wet Swine Gill, Cumbria	Chattlehope, Northumbria
Type of study	Annual sediment budget	Landslide event	Flood event
Period	December 1998–December 1999	1 February 2002	2 August 2002
Catchment area (km^2)	0.024	0.65	11.5
Local geology	Microgranite	Siltstone & mudstone, dolerite	Sandstone & shale
Annual precipitation ($mm\ yr^{-1}$)	2200	2200	1500
Mean altitude (masl)	500	500	550
Channel slope	0.273	0.18	0.075
Specific sediment yield ($t\ km^{-2}$)	1916	3.1	2.5*(25)

*Takes account of reservoir sedimentation. Bracketed value is the sediment flux to the reservoir. See text for further explanation.

(Wet Swine Gill, Johnson et al., 2008); and the effects of a large localised flood event (Chattlehope). Table 1 summarises their key characteristics. Although the three sites are broadly representative of upland headwater catchments in the UK and have similar precipitation and altitude characteristics they are significantly different in size and steepness (Table 1). This point is revisited later in the paper.

3 The Significance of Sediment Supply from Upland (Headwater) Landslides

Following exceptional regional flooding in North West England on 7–8 January 2005 an unprecedented number of shallow landslides occurred throughout the Northern Lake District mountains, particularly in the Bassenthwaite catchment (Environment Agency, 2006; Fig. 2). Major disruption was caused by landslides and local flooding (daily rainfall >150 mm). The impacts in the region were widespread involving disruption and damage to roads, bridges and culverts; destruction of agricultural infrastructure; inundation of forestry plantations and farmland; loss of livestock; diversion of stream courses; and contamination of upland water courses with sediment. Fortunately because most of this occurred in the early hours of 8 January there was no direct loss of life.

In the aftermath of the flood it was assumed that the landslides had been responsible for much of the sediment that polluted the upland water courses. Therefore, a rapid aerial assessment of the full extent of the slope failures was undertaken. This provides immediate photographic evidence of the fresh failure features and was used to plan a programme of field-based assessments. At each landslide site a differential GPS survey was undertaken and a Slope Failure Reconnaissance Sheet was used to

document the morphometry, morphological characteristics, material characteristics, drainage setting, post failure development and degree of slope channel coupling. A total of 51 significant shallow landslides occurred over an area 457 km^2 which encompassed most of the Bassenthwaite upland catchment (Fig. 2B). The landslides varied in size from small stream side scars (6 m^2) to major hillslope failures (1120 m^2). Landslide source areas were widely distributed over a large altitudinal range from 200–700 masl. Most failures occurred on slopes greater than 20° and the depth of failure was shallow, rarely exceeding 1 m in depth. The majority of failures (65%) involve sediment volumes of < 100 m^3.

During the flood, fluvial sediment flux monitoring was undertaken at four sites in the lower Bassenthwaite catchment (Fig. 2C). At each monitoring site river stage (m) and turbidity (measured using a Analite 395 probe in nephlometric turbidity units (NTU)) were recorded continuously at 15 minute intervals and logged on a Campbell data logger. Where available discharge was calculated from established stage-discharge rating relationships collected by the Environment Agency.

The detailed landslide inventory and fluvial sediment flux monitoring provide an excellent opportunity to quantitatively assess the geomorphological impacts of the landslide events in terms of hillslope sediment production and delivery to the lowland stream system. These calculations were carried out for the Newlands Beck site and upstream catchment (Fig. 2B and C, Table 2). Based on the landslide inventory 16 landslides occurred upstream of the gauging site which collectively mobilised in total 1440 t of sediment. This mass was dominated by two large landslides which accounted for 91% of this total (Keskadale 52%, Rigg Beck West 39%). Given the steep nature of the catchment it can be hypothesised that these large events may have had a significant impact on the downstream fine sediment flux. Examining the sediment flux record for the downstream site clearly shows two large turbidity pulses on the falling limb of the flood hydrograph (Fig. 3). The total suspended sediment load for the flood event (shown in the rectangular box) is 980 t. The two distinct falling limb sediment spikes (shown by the ellipse) contribute approximately 54 t to

Table 2 Comparison between landslide sediment delivery estimates and sediment load calculations at the Newlands Beck monitoring site during the January 2005 flood event

Sediment load component	Mass (t)		
Fluvial suspended sediment – flood event*			980
Fluvial suspended sediment – spike events			54
Landslides	*Keskadale*	*Rigg Beck*	*Combined*
Total erosion	748	561	
Sediment delivered to the channel**	344	258	
Fluvial suspended sediment – from landslides***	34	26	60

* Load estimates (t) are derived from the turbidity record using a calibration relationship between turbidity (NTU) and directly sampled suspended sediment concentration (mg l^{-1}).
** Sediment delivery is calculated by subtracting the total volume of sediment mobilised by the landslide minus that sediment left in storage on the hillslope.
*** This is assumed to be the grain-size component less than 2 mm (~ 10% based on analysis of sediment delivered to the channel).

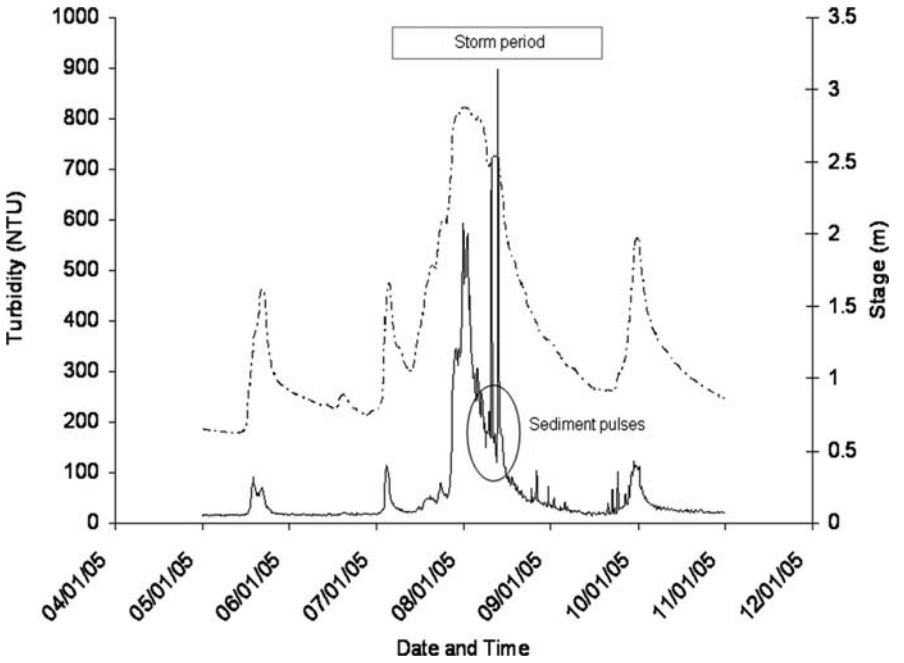

Fig. 3 Stage record (m – *dashed line*) and turbidity (NTU – *solid line*) at the Newlands Bridge monitoring site during the January 2005 flood event. Values recorded at 15 minute intervals. The period of the flood is highlighted by the rectangle and the ellipse shows the main sediment pulses observed on the falling limb of the event

the overall load. To test whether the two large landslides in the catchment are potentially responsible for this component of the load we must first calculate the mass of the landslides and then determine what proportion of this sediment entered the river channel (Table 2). This can be done by using the data in the landslide inventory. Based on this information we can determine that approximately 600 t of sediment entered the stream channel. However, because there are considerable transmission losses in the immediate reach downstream of the landslide only a small proportion (c.10%) contributes to the suspended fraction of the stream load. Based on these measurements it is estimated the fine load contributed from the landslides is approximately 60 t (Table 2). This load compares very favourably with the load (54 t) measured in the sediment spikes at the downstream monitoring site. Furthermore, local observations suggest the landslides occurred after the main flood peak which is consistent with the timing of the sediment pulses in the hydrological record and the fact that the Keskadale and Rigg Beck landslides are 10.1 and 5.6 km upstream respectively of the monitoring site.

Based on this simple analysis it can be seen that the shallow landslides provide a relatively small contribution to the overall suspended sediment load and given that they occur only in extreme events are unlikely to be of major long term significance to the catchment sediment flux. This raises the question as to where the fine sediment

is coming from? Consideration of Fig. 1 suggests several possibilities, namely: gullying, wash processes or bank erosion. Wind erosion and pipe erosion are insignificant in this event and stream side bluff erosion was mapped as part of the landslide inventory and shown to be a small component of the overall sediment flux from slopes (Johnson and Warburton, 2002b). We also note that in events of this type surface wash and gullying, although occurring locally, are of minor importance because the mass of sediment eroded is small and the mobilised sediment is generally very poorly connected to the river channels (Evans and Warburton, 2005; Johnson and Warburton, 2002a). Hence stream course and bank erosion must be considered to be a significant component of the overall load (Johnson and Warburton, 2002a).

4 The Significance of Upland Channel and Bank Erosion to the Fine Sediment Load in the Bassenthwaite Upland Catchment

In order to evaluate the importance of bank erosion in contributing fine sediment to the river load in the Bassenthwaite catchment direct monitoring of suspended sediment load has been undertaken, since the beginning of 2005, on the two main rivers entering the lake (two sites on the River Derwent and one site on Newlands Beck) (Fig. 2C). Results presented here are from an eight month period of the time series when simultaneous records of flow and turbidity are available for the three sites; these include estimates of suspended sediment loads (calculated from the turbidity record) for the River Derwent and Newlands Beck from 1st April to 30 November 2006.

Figure 4 show the cumulative suspended sediment loads measured at the three sites for the eight month period. It is immediately obvious that the main source of fine fluvial sediment to Bassenthwaite Lake is contributed from the River Derwent. Over the period of monitoring, the Derwent (upstream catchment area 238 km^2) contributes 5.2 times more sediment than Newlands Beck (upstream catchment area 34 km^2). If it is assumed these are the two major fluvial sediment sources to the southern end of the lake then the Derwent contributes over 80% of the fine sediment load to the lake. These estimates differ significantly from Parker et al. (1999) who calculated the load entering Bassenthwaite from fortnightly samples collected on the River Derwent. Back-calculating a comparable 8 month load based on their data suggests a total load for the period of approximately 1900 t which is less than half the estimate reported here. This is expected given the infrequent fortnightly sampling strategy used by Parker et al. (1999) and the inevitable underestimation of significant sediment events that will occur.

The results also demonstrate there is a large apparent increase in sediment load (approximately 738 t) between the upper Derwent site and the lower Derwent site at Low Stock Bridge (Fig. 2C). Figure 4 shows that the difference in load between the two sites is largely accounted for by a series of large step changes in sediment flux at the lower site. The interpretation of this is the load difference is greater during large

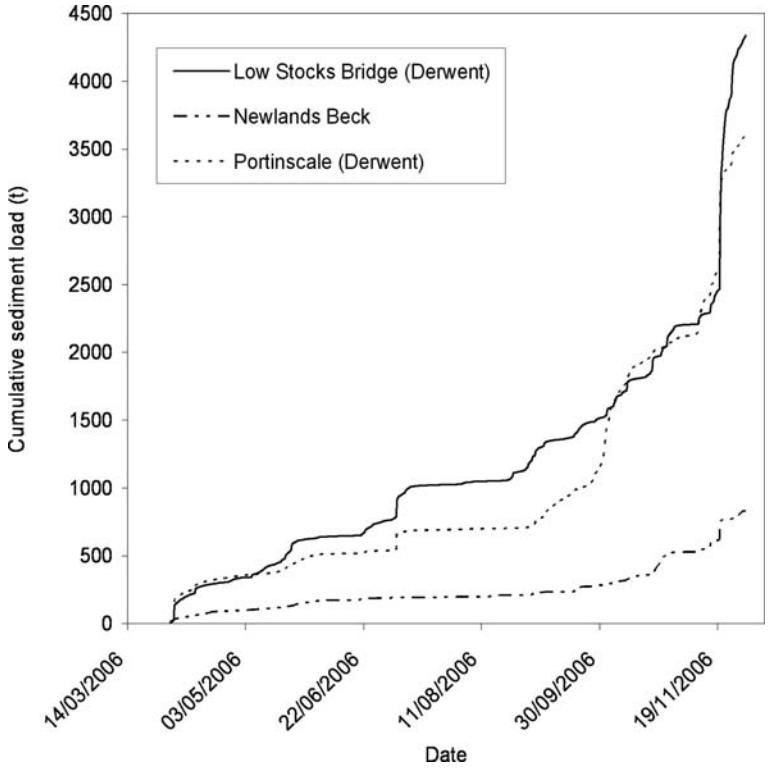

Fig. 4 Estimated cumulative suspended sediment loads (t) for the River Derwent and Newlands Beck, 1st April 2006–30 November 2006. The Low Stocks Bridge site is located approximately 3.7 km downstream of Portinscale (Fig. 2)

events (steps) and field observations together with the inspection of the suspended sediment traces suggest this is associated with bank failure and channel erosion producing pulses in fine sediment. Considering the potential sediment sources for the observed downstream increase in suspended sediment load, bank erosion is thought to be the dominant source but minor amounts of sediment may also be derived from in-channel storage of bed sediments and from two small headwater tributaries which join the main channel. If we assume that bank erosion contributes all the additional 738 t of sediment then this would equate to approximately 0.046 m of uniform bank erosion along the 3.7 km reach between the two sites (Fig. 2C). This calculation assumes a bulk density of 1.8 t m^{-3} (Dapporto et al., 2003), uniform erosion over the full length of the reach (3700 m) and along both channel banks (average bank height assumed to be 1.2 m). Field evidence and stream reconnaissance studies (Orr et al., 2004) suggest that this is a channel with active bank erosion along much of the reach although this is not evenly distributed. These figures must be treated cautiously because part of the load may be derived from temporary bed storage of fine sediment which is re-entrained during high flow and; the tributaries (mentioned above) could potentially provide additional load as these drain the south-western

slopes of Skiddaw (e.g., Applethwaite Gill) which has been previously identified as a potential significant sediment source (Orr et al., 2004). Both these sources would offset the inferred rates of bank erosion as calculated in the sediment balance equation.

These initial fine sediment load estimates should be regarded as preliminary for several reasons. Firstly, in order to calculate the load, turbidity values are converted to suspended sediment concentration (mg l^{-1}) using a conversion of 1.1NTU = 1 mg l^{-1}. Further work is underway to refine this calibration but experience shows from similar upland river basins that this relationship is fairly consistent. This does not affect the conclusions given here because relative values are unaffected by this as all turbidity meters are calibrated to the same NTU standards. Secondly, it is assumed that the discharge at Low Stocks Bridge (currently ungauged) is the same as at Portinscale. This is a reasonable assumption because only minor tributaries join the river between the two sites and the timing and form of flood peaks (based on a comparison of discharge and stage records at the two sites) is approximately coincident. Thirdly, the results represent a relatively short time series which essentially covers the main summer/autumn period, with the winter flows not fully represented. An analysis including a longer times series is required in order to assess whether the balance of loads is stable over this extended period.

It can therefore be concluded that the Derwent is the dominant source of fine sediment delivery to Bassenthwaite Lake and a significant proportion (c. 20%) of this sediment is derived from channel processes acting in the lower reaches between Portinscale and Low Stock Bridge. Considering that bank erosion is widespread in the catchment (Orr et al., 2004) it is hypothesised that this is the dominant sediment source in the fluvial system and is consistent with many upland sediment budget studies (Foster and Lees, 1999; Johnson and Warburton, 2002a, b; Evans and Warburton, 2007).

Direct coupling between the stream banks (sediment source) and river channel means that sediment transfer is an efficient process during high discharge events (Hooke, 1980; Thorne, 1982; Green et al., 1999). Figure 5 shows a 10 day period of high discharge (measured upstream at Portinscale) in November 2006. The flow was relatively high and experienced multiple discharge peaks with a particularly notable event on the 20 November rising to over 80 m^3 s^{-1}. From 15–20 November the two turbidity records are broadly similar with a slight downstream lag in suspended sediment concentration. However, following the large flood on 20/21 November the two records are markedly different. At the lower site the record is considerably noisier with peak turbidity values exceeding upstream readings. The multiple spikes and elevated background levels of turbidity are associated with local bank failures and the reworking and redistribution of cohesive sediment blocks in the channel (Grissinger, 1982) (Fig. 6). Much of this is complicated by local failures which deliver large blocks of cohesive sediment to the stream channel which act as local ephemeral sediment sources (Fig. 6). This sediment can remain in local storage for many successive discharge events until it is eventually dispersed (Dapporto et al., 2003). There is extensive evidence of this type of erosion on the lower Derwent immediately upstream of the lower sediment monitoring site (Fig. 6A, C).

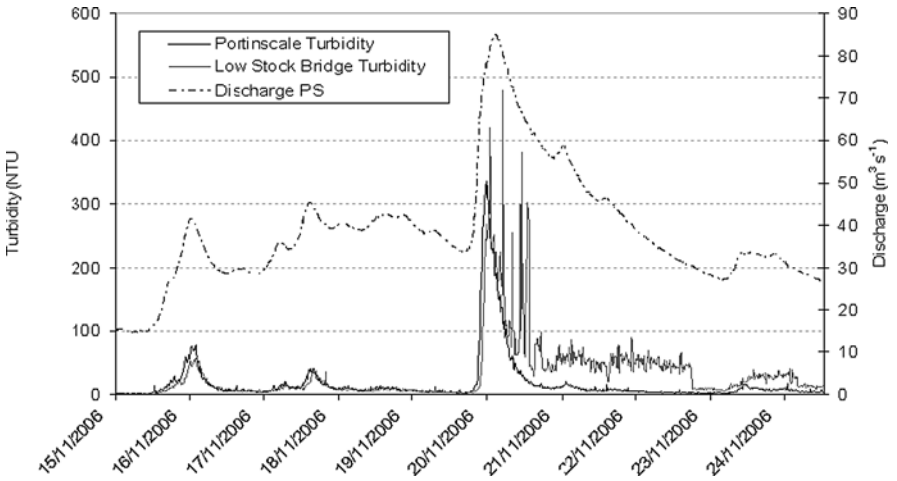

Fig. 5 Differences in suspended sediment flux between Portinscale and Low Stocks Bridge during a flood event (20 November 2006). Discharge at Portinscale is shown

Fig. 6 Examples of bank failures on the lower Derwent. (**a**) Recent slumps involving fine upper bank sediments and gravel (*lower bank*). (**b**) Block failures of fine cohesive material. (**c**) Extensive bank erosion upstream of the Low Stocks Bridge monitoring site

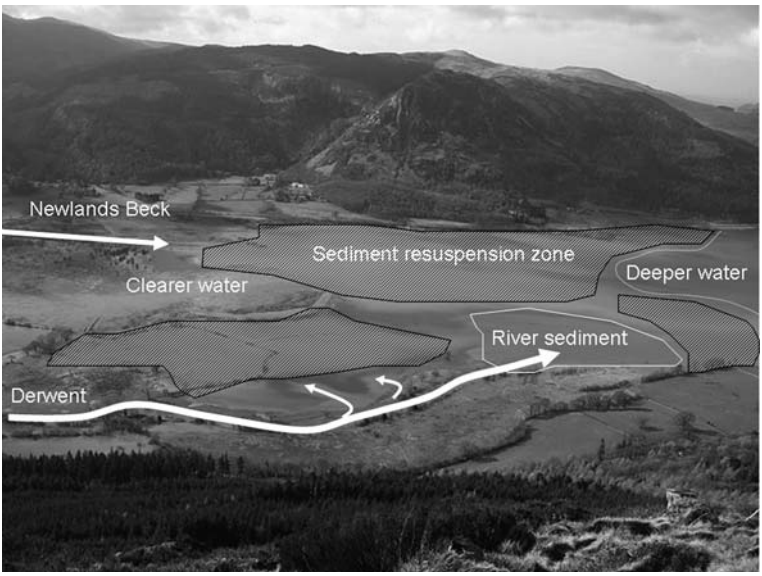

Fig. 7 Sediment delivery to the southern part of Bassenthwaite Lake on 6 March 2007. The image is annotated to show the relative significance of fluvial fine sediment delivery and lake resuspension

Eventually this river sediment is delivered to the head of the Lake (Fig. 7). However, there is a complex pattern of sediment dynamics in the shallow lake that has partly led to misconceptions about the relative importance of fine sediment delivery from the main catchment rivers (Parker et al., 1999). Figure 7 photographed on 6 March 2007 clearly show very high turbidity values in the shallow proximal lake basin. Four key zones can be identified: (1) delivery of turbid water from the river Derwent, (2) inflow of relatively clear water from Newlands Beck; (3) resuspension of lake bed sediment in the nearshore shallows; and (4) clear water in the deeper lake basin (Fig. 7). Hence the high turbidity in the lake is not derived solely from the delivery of fine sediment from the inflowing rivers (Hakanson and Jansson, 2002). Resuspension of fine sediment by high winds has been identified as a particular problem in Bassenthwaite (Parker et al., 1999; Hall et al., 2001) due to its shallow bathymetry and extensive near shore shallows. Modelling by Parker et al. (1999) has demonstrated that northerly winds of low to moderate magnitude disturb more of the lake bed than southerly winds of similar force.

5 Headwater Sediment Dynamics and Connectivity with the Wider Catchment

In the previous two sections we have established that: (1) the impact of catchment landslides is detectable, in the sediment flux record, if they are of sufficient size; however their overall significance is relatively minor and ; (2) bank erosion appears

to be the dominant source of fine sediment supply to the fluvial system (particularly in lowland reaches).

However it is important that we look more closely at the role of headwater catchments in sediment supply to determine whether this general conclusion is valid over a longer timescale and for diffident types of event. In this section we examine the annual sediment budget for the Iron Crag torrent system (Johnson and Warburton, 2002a); the impact of a large discrete slope failure in Wet Swine Gill (Johnson et al., 2008); and the significance of a large localised flood event on Chattlehope Burn in Northumbria (Table 1, Fig. 2A). Although these are not directly linked to the Bassenthwaite catchment they represent similar behaviour to those headwater areas in the catchment.

The Iron Crag headwater catchment is an upland torrent (Eisbacher, 1982). It has a steep, convex-concave catchment of 0.024 km^2 (Table 2).

Field investigations have established the link between the hydro-meteorology of the catchment and sediment dynamics over range of timescales. Monitoring began in 1998 and results have been published in a number of articles (Johnson and Warburton 2002a, b, 2006a, b). The annual sediment budget for the Iron Crag system, measured in 1998–1999, is summarised schematically in Fig. 8. The relative contributions from slope erosion, gullying, bank erosion and channel bed erosion is shown alongside floodplain and fan storage. The catchment sediment yield for this period (1916 t km^{-2}) is extremely high and only represents 20% of the sediment eroded in the catchment. This is not surprising given the locally weak bedrock, high rainfall and extremely steep nature of the torrent system. Most of the sediment is eroded from the channel system and redeposit in a basal fan. Sediment dynamics

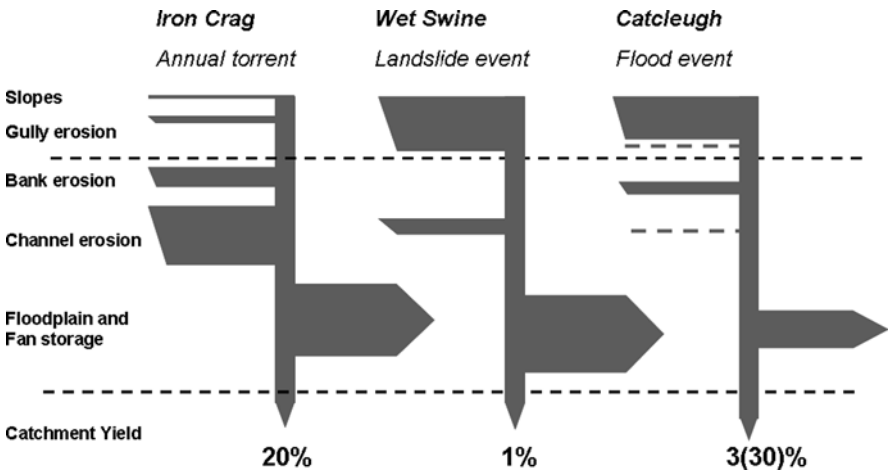

Fig. 8 Comparison of three summary upland sediment budgets: Annual sediment budget of the Iron Crag torrent system (December 1998–December 1999); Landslide event sediment budget at Wet Swine Gill (1 February 2002); and Flood event sediment budget on Chattlehope Burn, Catcleugh (24 August 2002). All values (*horizontal bars*) are proportional to the total catchment sediment yield (Table 1)

are characterised by a process of slow accumulation until a trigger event (usually heavy rainfall) results in catastrophic flushing of the sediment sink (c. 1 major event pear year), although not all events are equivalent due to a combination of supply – limited & transport-limited sediment transfer controls (Seeger et al., 2004).

Annual sediment budgets do not always capture extreme events such as landslides which often need to be evaluated after their initial occurrence. At Wet Swine Gill in February 2002 a large hillslope failure (181 m^3) occurred on a 30° slope, supplying a large sediment flux to the channel system (Table 2). Given the direct coupling of the event between the hillslope and channel it was hypothesised that this would result in enhance sediment transfer rates downstream, however, a careful sediment budget analysis of the event (Johnson, et al., 2008) demonstrated that only a tiny fraction (1%) of the eroded sediment escaped from the small catchment (Fig. 8). The majority of sediment was deposited on footslopes and along the valley axes. To some extent this might be understandable given that the slope instability can occur independently of valley stream flows competent enough to efficiently transfer the accumulated sediment (at least in the short-term) (Parsons et al., 2006). It might therefore be anticipated that a headwater flood event would evacuate some of this available sediment in the future. Headwater floods are noteworthy for the destruction they cause to hillslopes and channels (Macklin and Rumsby, 2007) but their efficiency in transporting sediment is not frequently evaluated as estimates of sediment fluxes are not known.

The final example given in Fig. 8 shows a summary sediment budget for an upland flood event that occurred on 2nd August 2002 in the Chattlehope Burn reservoir catchment (11.5 km^2) in Northumbria. Following severe rainfall the peak inflow into Catcleugh Reservoir was estimated at 50 $m^3\ s^{-1}$ which over a 12 hour flood period increased storage in the reservoir by approximately 10% of the volume. The geomorphic impact of the flood produced numerous streamside slope failures in the headwater streams resulting in an event sediment yield to the reservoir which was approximately equivalent to the average annual stream sediment load (established from long-term records of sedimentation). The summary sediment budget shows considerable erosion of the toe slopes and streambanks but nearly two thirds of this sediment is redeposited on the intervening floodplain. The net effect was that only about 305 t of the eroded material was transferred downstream and once entering the reservoir only 3% passed the dam wall as fine suspended material (Fig. 8). Although this is a small proportion of the eroded sediment it was still sufficient to temporarily suspend reservoir operations until suspendered solids returned to manageable levels.

These three examples when viewed collectively demonstrate that sediment transfer is a relatively inefficient process in upland catchments and sediment sinks (tributary valley floors, fans and upland floodplains) are a very important, and hitherto poorly recognised, component of the upland sediment system (Moore and Newson, 1986). It is therefore of great importance to fully recognise and quantify all components in a headwater sediment budget studies, particularly because the sediment balance will differ depending on the dominant geomorphic processes operating (Prestegaard, 1988; Lawler, 1995). Figure 8 clearly demonstrates the value of contemporary monitoring in understanding sediment transfer rates but care should

be taken in extrapolating measured sediment transfer rates over longer timescales (Dearing, 1992; Marsh and Sanderson, 1997; van de Post et al., 1997). Under current conditions the impact of these "extreme" events is probably equal to "1x the annual sediment load in a day" (~ 50–200 t km^{-2}) and although devastating in the short-term are not very significant at the larger catchment scale or over longer time periods. Event magnitude, erosion thresholds and event frequency would need to change significantly to generate large-scale change in the sediment transfer regime on the British Upland landscape (Macklin et al. 1992; Wilby et al., 1997).

6 Conclusions

This paper has used a variety of data sources to attempt to understand the dynamics of sediment supply and transfer in steep, UK upland catchment sediment systems and in particular the Bassenthwaite catchment in Northern England. Although these individual studies are not of equivalent temporal and spatial scales they are effectively combined in a general sediment budget framework (Fig. 1) and provide an excellent initial assessment of the catchment sediment dynamics. The primary conclusions of this work are:

1. Sediment transfer is an inefficient process in upland catchment systems and high rates of erosion and sediment transport often only manifest themselves at the local scale.
2. Low sediment fluxes are often the result of inefficient coupling between slope-channel linkages and headwater-mainstream channels.
3. Upland sediment yields are frequently episodic, typically associated with a large "annual" event often following a major rainstorm. These sediment flux events often manifest themselves as upland flash floods which although devastating in the short-term may not be very significant at the larger scale or over longer time periods.
4. Upland sediment systems are complex systems and to understand them fully we require full-component (Fig. 1) sediment budgets carried out at nested scales which link headwater and lowland sediment transfer regimes.

Although the synthesis presented in the paper represents probably one of the most complete assessments of upland sediment transfer for a large UK catchment the information given is still fragmentary in both its spatial and temporal coverage. This does not diminish the findings presented in this paper but it does highlight the need for further investigations to better understand the links between upland landforms and the continuity in sediment flux at various temporal and spatial scales. Future investigations should aim to integrate existing terrestrial sediment flux studies (as reported here) with detailed investigations of lake sedimentation in order to provide information for modelling the full sediment cascade.

Acknowledgments I am grateful to my collaborators Richard Johnson, Dave Milledge, Andrew Mills and Jon Hopkins who have contributed to various aspects of this work. Funding has been generously provided by NERC (Grant NE/D521481/1), Durham University, EA, and Northumbria Water. Ann Bowker kindly provided permission to reproduce the photograph used in Fig. 7.

References

Anderton, J, Howarth, EY, Horne, DJ and Wray, DS (1998) Environmental impacts of lead mining in the Ullswater catchment (English Lake District): dam failures and flooding. In Issues in Environmental Geology: a British Perspective. The Geological Society, London, pp 226–242.

Ashbridge D (1995) Processes of river bank erosion and their contribution to the suspended sediment load of the River Culm, Devon. In: Foster IDL, Gurnell AM, Webb, BW (eds) Sediment and water quality in river catchments. John Wiley and Sons: Chichester, pp 229–245.

Atkinson KM, Heaney SI, Elliot JM, Mills CA (1989) Bassenthwaite Lake: a general assessment of environmental and biological features and their susceptibility to change. Freshwater Biological Association, Ambleside.

Barsch D, Caine N (1984) The nature of mountain geomorphology. Mountain Research and Development 4(4): 287–298.

Bennion H, Montieth D, Appleby P (2000) Temporal and geographical variation in lake trophic status in the English lake district: evidence from (sub) fossil diatoms and aquatic macrophytes. Freshwater Biology 45: 394–412.

Bloesch J (1995) Mechanisms, measurement and importance of sediment resuspension in lakes. Marine and Freshwater Research 46: 295–304.

Boardman J (1992) Quaternary landscape evolution in the lake district – a discussion. Proceedings, Cumberland Geological Society 5: 285–315.

Caine, N (1974) The geomorphic processes of the alpine environment. In: Ives JD, Barry RG (eds) Arctic and Alpine Environments. Methuen, pp 721–748.

Church M, Slaymaker O (1989) Disequilibrium of Holocene sediment yield in glaciated British Columbia. Nature 337: 452–454.

Cranwell PA, Haworth EY, Davies PS, Lawler AJ, Lishman JP (1995) Bassenthwaite Catchment Studies Phase 3b. Assessment of historical changes in Bassenthwaite Lake by analysis of a deep sediment profile. IFE Report T1106011/1 to the National Rivers Authority (North West Region).

Dapporto S, Rinaldi M, Casagli N, Vannocci P (2003) Mechanisms of riverbank failure along the Arno River, central Italy. Earth Surface Processes and Landforms 28: 1303–1323.

Dearing JA (1992) Sediment yields and sources in a Welsh upland lake-catchment during the past 800 years. Earth Surface Processes and Landforms 17: 1–22.

De Boer DH, Crosby G (1996) Specific sediment yield and drainage basin scale. In: Walling DE, Webb BW (eds) Erosion and Sediment Yield: Global and Regional Perspectives, IAHS Publication No. 236, pp 333–338.

Dietrich WE, Dunne T (1978) Sediment budget for a small catchment in mountainous terrain. Zeitschrift fur Geomorphologie 29: 191–206.

Eisbacher GH (1982) Mountain torrents and debris flows. Episodes 4: 12–17.

Environment Agency (2006) Cumbria floods technical report: factual report on meteorology, hydrology and impacts of January 2005 flooding in Cumbria. Environment Agency.

Evans MG, Warburton J (2005) Sediment budget for an eroding peat moorland catchment in northern England. Earth Surface Processes and Landforms 30: 557–577.

Evans MG, Warburton J (2007) Upland Peat Erosion – Form, Processes and Landscape Change. Blackwells, Oxford.

Foster IDL, Lees JA (1999) Changing headwater suspended sediment yields in the LOIS catchments over the last century: a paleolimnological approach. Hydrological Processes 13: 1137–1153.

Green TR, Beavis SG, Dietrich CR, Jakeman AJ (1999) Relating stream-bank erosion to in-stream transport of suspended sediment. Hydrological Processes 13: 777–787.
Greig SM, Sear DA, Carling PA (2005) The impact of fine sediment accumulation on the survival of incubating salmon progeny: Implications for sediment management. Science of the Total Environment 344: 241–258.
Grissinger EH (1982) Bank erosion of cohesive materials. In: Hey RD, Bathurst JC, Thorne CR (eds) Gravel Bed Rivers. John Wiley and Sons, Chichester pp 273–287.
Hakanson L, Jansson M (2002) Principles of Lake Sedimentology. Blackburn Press: Caldwell.
Hall GH, Haworth EY, Lawlor AJ, Vincent C, Tipping EW (2001) The origin of the frequently resuspended sediment material in Bassenthwaite Lake. Centre for Ecology and Hydrology, Ambleside.
Holliday VJ, Warburton J, Higgitt DL (2008) Historic and contemporary sediment transfer in an upland Pennine catchment, UK. Earth Surface Processes and Landforms. DOI 10.1002/esp.1600, 33, 14, 2139–2155.
Hooke JM (1980) Magnitude and distribution of rates of river bank erosion. Earth Surface Processes 5: 143–157.
Johnson RM, Warburton J (2002a) Annual sediment budget of a U.K. Mountain Torrent. Geografiska Annaler 84A: 73–88.
Johnson RM, Warburton J (2002b) Flooding and geomorphic impacts in a mountain torrents: Raise Beck, Central Lake District, England. Earth Surface Processes and Landforms 27: 945–969.
Johnson RM, Warburton J (2006a) Variability in sediment supply, transfer and deposition in an upland torrent system: Iron Crag, Lake District, Northern England. Earth Surface Processes and Landforms 31: 844–861.
Johnson RM, Warburton J (2006b) Episodic discharge of coarse sediment in a mountain torrent. In: Rowan JS, Duck RW, Werritty A (eds) Sediment Dynamics and the Hydromorphology of Fluvial Systems. IAHS Publication, 306, pp 64–71.
Johnson, R.M., Warburton, J. and Mills, A. (2008) Hillslope-channel sediment transfer in a slope failure event: Wet Swine Gill, Lake District, Northern England. Earth Surface Processes and Landforms. DOI 10.1002/esp.1563, 33, 394–413.
Lawler DM (1995) The impact of scale on the processes of channel-side sediment supply: a conceptual model. Effects of Scale on Interpretation and Management of Sediment and Water Quality (Proceedings of a Boulder Symposium, July 1995). IAHS Publication 226, pp 175–184.
Lawler DM, Grove JR, Couperthwaite JS, Leeks GJL (1999) Downstream change in river bank erosion rates in the Swale-Ouse system, northern England. Hydrological Processes 13: 977–992.
Lawler DM, Petts GE, Foster IDL, Harper S (2006) Turbidity dynamics during spring storm events in an urban headwater river system: The Upper Tame, West Midlands, UK. Science of the Total Environment 360: 109–126.
Ledger DC, Lovell JPB, McDonald AT (1974) Sediment yield studies in upland catchment areas in south-east Scotland. Journal of Applied Ecology 11: 201–206.
Macklin MG, Rumsby BT, Heap T (1992) Flood alluviation and entrenchment: Holocene valley-floor development and transformation in the British uplands. Geological Society of America Bulletin 104: 631–643.
Macklin MG, Rumsby BT (2007) Changing climate and extreme floods in the British uplands. Transactions of the Institute of British Geographers NS 32: 168–186.
Marsh TJ, Sanderson FJ (1997). A review of hydrological conditions throughout the period of the LOIS monitoring programme – considered within the context of the recent UK climatic volatility. The Science of the Total Environment 194/195: 59–69.
Moore RJ, Newson MD (1986) Production, storage and output of coarse upland sediments: natural and artificial influences as revealed by research catchment studies. Journal of the Geological Society 143: 921–926.
Nisbet TR, Orr HG, Broadmeadow S (2004) Evaluating the role of woodlands in managing soil erosion and sedimentation within river catchments: Bassenthwaite Lake study. Final report to Forestry Commission (England).

Oldfield, F and Appleby, PG (1984) A Empirical testing of 210Pb-dating models for lake sediments. In: Haworth EY and Lund JWG (eds) Lake Sediments and Environmental History. Leicester University Press, Leicester, pp 93–124.

Orr H, Davies G, Quinton J, Newson M (2004) Bassenthwaite Lake Geomorphological Assessment Phase 2 Final Report, Unpublished Lancaster University report to the Environment Agency.

Parker JE, Lyle AA, Dent MM, James JB, Lawlor AJ, Simon BM, Smith E (1999) Final report: investigation into the nature of the material resuspended in Bassenthwaite Lake during mixing episodes. Unpublished CEH report to the Environment Agency. Report ref: WI/T11067/Q7/2.

Parsons AJ, Wainwright J, Brazier RE, Powell DM (2006) Is sediment delivery a fallacy? Earth Surface Processes and Landforms 31:1325–1328.

Pennington, W (1981) Records of a lakes life in time: the sediments. Hydrobiologia 79: 197–219.

Prestegaard KL (1988) Morphological controls on sediment delivery path-ways. In: Bordas MP, Walling DE (eds) Sediment Budgets. Proceedings, Symposium 11–15th December 1988, Porto Alegre, Brazil. IAHS Publication 174, pp 533–540.

Russell MA, Walling DE, Hodgkinson RA (2001) Suspended sediment sources in two small lowland agricultural catchments in the UK. Journal of Hydrology 252: 1–24.

Seeger M, Errea M-P, Begueria S, Arnaez J, Marti C, Garcia-Ruiz JM (2004) Catchment soil moisture and rainfall characteristics as determinant factors for discharge/suspended sediment hysteretic loops in a small headwater catchment in the Spanish Pyrenees. Journal of Hydrology 288: 299–311.

Shen, Z, Bloemendal, J, Mauz, B, Chiverrell, RC, Dearing, JA, Lang, A and Liu, Q (2008) Holocene environmental reconstruction of sediment-source linkages at Crummock Water, English Lake District, based on magnetic measurements. The Holocene, 18(1): 129–140.

Thackeray SJ, Maberly S, Winfield I (2006) The ecology of Bassenthwaite Lake (English Lake District). Freshwater Forum 25: 1–80.

Thorne CR (1982) Processes and mechanisms of river bank erosion. In: Hey RD, Bathurst JC, Thorne CR (eds) Gravel Bed Rivers. John Wiley and Sons, Chichester pp 227–271.

van de Post KD, Oldfield F, Haworth EY, Crooks PRJ, Appleby PG (1997) A record of accelerated erosion in the recent sediments of Blelham Tarn in the English Lake District. Journal of Paleolimnology 18: 103–120.

Walling DE, Collins AL (2000) Integrated assessment of catchment sediment budgets: A technical manual. University of Exeter, UK, pp 1–168.

Wells SG, Harvey AM (1987) Sedimentologic and geomorphic variations in storm-generated alluvial fans, Howgill Fells, northwest England. Geological Society of America Bulletin 98: 182–198.

Wilby RL, Dalgleish HY, Foster IDL (1997) The impact of weather patterns on historic and contemporary catchment sediment yields. Earth Surface Processes and Landforms 22: 353–363.

Winfield IJ, Durie NC (2004) Fish introductions and their management in the English Lake District. Fisheries Management and Ecology 11: 1–7.

Winfield IJ, Fletcher JM, James JB (2004) Conservation ecology of the vendace (Coregonus albula) in Bassenthwaite Lake and Derwent Water, U.K. Annales Zoologici Fennici 41: 155–164.

Volume Estimation, Kinematics and Sediment Transfer Rates of Active Rockglaciers in the Turtmann Valley, Switzerland

Isabelle Gärtner-Roer and Michael Nyenhuis

Abstract Rockglaciers represent major sediment storages and transport components of the periglacial system. A multi-method approach combining geomorphological mapping, DTM analyses, digital photogrammetry and geodetic survey, was applied to quantify volumes and kinematics for the calculation of sediment transfer rates of a rockglacier in the Turtmann Valley, Switzerland. Compared to other studies, the calculated thicknesses and also the volumes and masses of the landforms appeared to be rather small. Calculated with a mass of 0.53 million tons and with different annual velocities, the annual transfer rate of the rockglacier HuHH3 ranges between 0.13 M t/a for the period 1975–1993 and a maximum value of 0.9 M t/a for the year 2003/2004. Accounting for the simplicity of thickness models and due to the limited knowledge on internal deformation, the amount of moving sediment can only be estimated and conclusions have to be drawn with care.

Keywords Sediment transfer rates · Sediment budget · Periglacial · Rockglacier · European Alps

1 Introduction

Investigations of sediment budgets in high mountains improve the understanding of geomorphological processes shaping this landscape and allow for an explanation of Holocene landscape evolution. In addition, they serve as an important basis for the development of new process based models (Church & Slaymaker, 1989; Jordan & Slaymaker, 1991). Up to now, deficits exist in the knowledge of temporal and spatial coupling of nested geomorphological processes and complex interactions of different sediment storages (Dietrich & Dunne, 1978; Caine & Swanson, 1989; Schrott et al., 2003).

I. Gärtner-Roer (✉)
Department of Geography, University of Zürich, Winterthurerstrasse 190, 8057 Zürich, Switzerland
e-mail: iroer@geo.uzh.ch

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_12,

According to Reid & Dunne (1996, p. 3), a sediment budget is defined as "an accounting of the sources and disposition of sediment as it travels from its point of origin to its eventual exit from the drainage basin". This definition comprises geomorphic processes and process rates as well as sources and sinks (storages) as main components of the sediment transfer within geomorphic systems. The work of Jäckli and Rapp represented important first approaches to quantify geomorphic processes as a basis for sediment budgets in high mountain regions (Jäckli, 1957; Rapp, 1960). Meanwhile, the quantification of geomorphic processes and landforms and the analysis of process rates play an important role in the framework of sediment budget studies. Recent studies focus especially on sediment storages, on their interactions with each other and with associated geomorphic processes (e.g., Schrott & Adams, 2002; Schrott et al., 2003; Nyenhuis, 2006; Otto, 2006).

In order to analyse sediment transfer in high mountain environments in a systematic context, Caine (1974) developed his conceptual model of alpine sediment cascades. This model describes the logical, cascade-like sequence of processes, based on an idealized slope profile (Fig. 1). Later, the model was improved by the differentiation of four subsystems: glacial system, coarse debris system,fine sediment system, geochemical system (Barsch & Caine, 1984). Each of the subsystems has its own cascading system and is characterised by the intra-system flow of mass and energy. Interactions as well as movements of matter from one system to the other are possible. Application of this model facilitates comparative investigations at different study sites.

Regarding the significance of rockglaciers within the alpine geosystem, they represent one of the main storage components of the coarse debris system, following Barsch & Caine (1984). According to Barsch (1992, p. 176) "active rockglaciers are lobate or tongue-shaped bodies of perennially frozen unconsolidated material supersaturated with interstitial ice and ice lenses that move downslope or downvalley by creep as a consequence of the deformation of ice contained in them". Different methods to distinct active, inactive and relict rockglaciers are discussed in Roer & Nyenhuis (2007), highlighting associated methodological capabilities and limitations. Since a reactivation of an inactive rockglacier – regardless of whether inactivity was climatically or dynamically induced (Barsch 1996) – has never been observed, these landforms are considered to be sediment traps and are therefore closed coarse debris systems. Remobilisation of the sediment might take place only in a glacier advance. Hence, rockglaciers as well as debris cones are well suited for the calculation of denudation rates typical for their environment. The assessment of sediment volumes of these landforms and the calculation of the area providing the sediment (rockwall), represent important sources of error in the calculation of denudation rates. In addition, little is known on the dynamics of rockglacier creep as well as the age of the landforms.

The geomorphic significance of rockglaciers in periglacial geosystems, especially regarding the quantification of headwall recession rates, is supported by a series of studies (Messerli & Zurbuchen, 1968; Gray, 1970; White, 1971; Barsch, 1977, 1996; Höllermann, 1983; Gorbunov, 1983; Barsch & Jakob, 1998; Humlum,

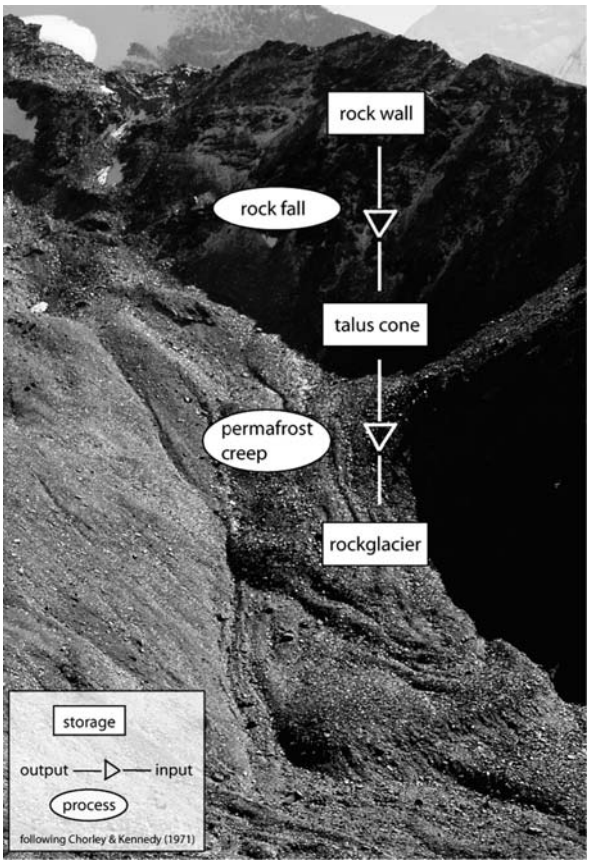

Fig. 1 The process-chain connecting rockwall, talus cone and rockglacier is a typical sediment cascade of the high mountain geosystem

2000). A detailed overview of these investigations is given by Barsch (1996) and Humlum (2000) and is summarised by Nyenhuis (2006).

It is assumed that rockglaciers represent a major sediment storage and transport component of the periglacial system. However, only few attempts have been undertaken to quantify the transfer of sediment by creeping rockglaciers and quantitative regional scale studies to determine the contribution of rockglaciers to the sediment budget of entire geosystems do not exist. The study presented here combines the results of two PhD theses conducted within the Research Training Group; on the one hand the calculated volumes and masses of the rockglaciers (Nyenhuis 2006), on the other hand the kinematics quantified for these landforms (Roer 2007). The purpose of the study is the combination of rockglacier masses and kinematics, in order to describe sediment transfer rates and to prepare further studies on regional scale relevance of rockglaciers for sediment budgets.

2 Study Site

The Turtmann Valley is a tributary of the River Rhone and is located in southern Switzerland between the Matter Valley and the Anniviers Valley (at about 46° 13′ N/7° 38′ E). Due to its inner-alpine location, the area is characterized by an intramontane climate. The valley's lithology mainly consists of Palaeozoic gneisses and schists. Its geomorphology is dominated by two large glaciers in the valley head and several hanging valleys above the shoulders of the trough. Here, periglacial landforms such as gelifluction lobes, ploughing boulders and rockglaciers are numerous (Otto & Dikau, 2004). Rockglaciers are the most striking landforms (covering 4.2% of the total area of 110 km^2) and occur in different states of activity. The high density of rockglaciers is one characteristic of the study area and facilitates intense investigation of the phenomenon.

3 Methods

Within the study area, two independent approaches are combined: first, the assessment of rockglacier thicknesses and masses by geomorphological mapping and DTM (Digital Terrain Model) analyses, and second, the quantification of rockglacier kinematics by geodetic and photogrammetric techniques.

3.1 Rockglacier Inventory and DTM Analyses

The rockglaciers in the Turtmann Valley have been mapped by geomorphological mapping in the field as well as by interpretation of high resolution aerial images and a corresponding DTM. The high resolution data have been created by a special campaign with an Airborne High Resolution Stereo Camera (HRSC-A) for multipurpose applications (Otto et al., 2007). The DTM served mainly as a basis for geomorphometric analyses to determine rockglacier characteristics like spatial dimensions, slope, aspect and curvature. Three dimensional visualizations provided a better overview of the landscape and enabled the mapping of landforms even in inaccessible terrain. All geomorphological and geomorphometric characteristics of the rockglaciers were compiled in an inventory, which is presented in more detail by Nyenhuis et al. (2005) and Nyenhuis (2006).

Direct information on rockglacier thicknesses, e.g., by borehole measurements or geophysical soundings, are not available for the study area. Therefore, rockglacier thicknesses were assessed by the application of a semi-quantitative approach using the high resolution DTM. Two methods were proposed:

The first method is based on the approach used by Calkin et al. (1987) and Humlum (2000), assessing the rockglacier thickness based on the thickness at a representative position at the front. The value is determined in a Geographical Information System (GIS) by a manual pixel enquiry. The resulting data are interpreted as maximum thicknesses, contrary to the assumption made by Barsch (1996).

For the second method, rockglacier thickness has been determined in different parts of the front and at the lateral margins and an average value has been calculated for each rockglacier.

To simplify the determination of thicknesses, no further landform characteristics, such as collapse structures often observed in the central part of relict rockglaciers, have been considered in the analysis. The error of the elevation in the DTM is specified with 50 cm and the accuracy of the thickness determination is assumed to be in the range of one to two meters.

The calculation of sediment volumes is based on a simplified three-layer-model of the internal structure of rockglaciers (Haeberli, 1985; Wagner, 1992; Barsch, 1996; Humlum, 2000): the uppermost layer is up to three meters thick and consists of very coarse blocks. The second layer represents the ice-rich, deforming permafrost core which is moving downslope and therefore covers the third layer, which results from blocks that fell from the front and are overridden by the rockglacier. The amount of sediment stored within rockglaciers is estimated to vary between 30 and 50% (Barsch, 1996; Humlum, 2000). Therefore, the rockglacier thicknesses determined according to the two methods described before were used to calculate two different sediment scenarios. Each scenario incorporates up to three different sediment ratios. For active rockglaciers, the masses were calculated with ratios of 30, 40 and 50% and for inactive landforms it was calculated with 40 and 50%. In relict rockglaciers the ice is melted out, the sediment is compressed and a lot of pore volumes are filled with fine material. Hence, sediment ratios of 50–60% are assumed. For the calculation of sediment masses, a density of the sediment of 2.6 g cm^3 was utilised.

3.2 Digital Photogrammetry and Geodetic Survey

Rockglacier kinematics is quantified by means of digital photogrammetry and by geodetic survey in order to describe rockglacier activity reliably. Analogue aerial stereo photographs for the years 1975 and 1993 were selected and analysed in combination with digital data of the HRSC-A survey in 2001. Horizontal displacements of single blocks at the rockglacier surfaces were quantified by the application of special image correlation software developed by Kääb and Vollmer (2000). Hence, permafrost creep rates were quantified for several landforms over a 26-year period, divided into two periods (1975–1993, 1993–2001). Further details on raw data, processing and accuracies are given in the work of Kääb & Vollmer (2000), Roer et al. (2005) and Roer (2007). In addition to the remote sensing approach, a geodetic network was established on two rockglaciers in order to get high precision data on spatio-temporal changes in rockglacier geometry. The survey was conducted annually between 2002 and 2006. More information on monitoring design and accuracies is described by Roer (2007). Assuming the life time of rockglaciers to be some thousand years (Frauenfelder & Kääb, 2000), this monitoring does not cover a long time period, but in comparison to other studies this is quite a long data series.

For the combination of movement rates and sediment volumes in order to describe sediment transfer rates, one representative annual velocity value has to be

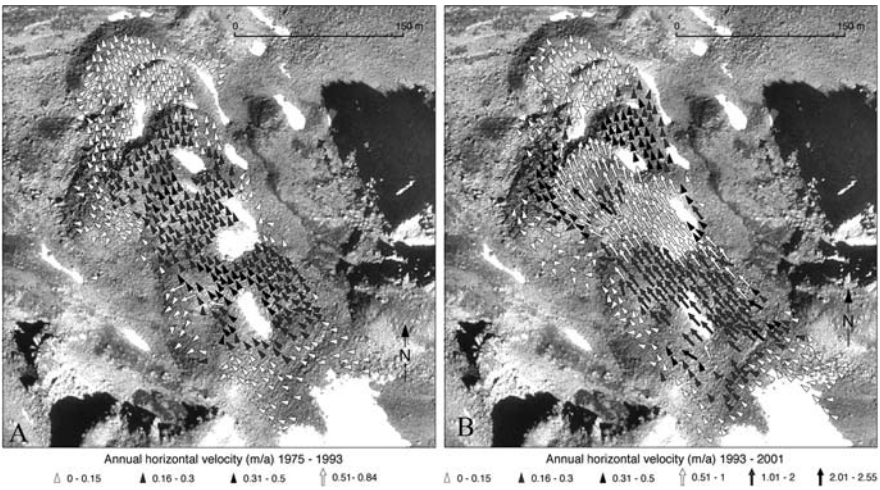

Fig. 2 Annual horizontal velocities during the periods 1975–1993 and 1993–2001 on rockglacier HuHH3, as derived from photogrammetric analyses (see also Roer, 2007; Otto et al., 2007). Underlying orthoimage of 20.08.1975 (aerial photographs taken by Swisstopo)

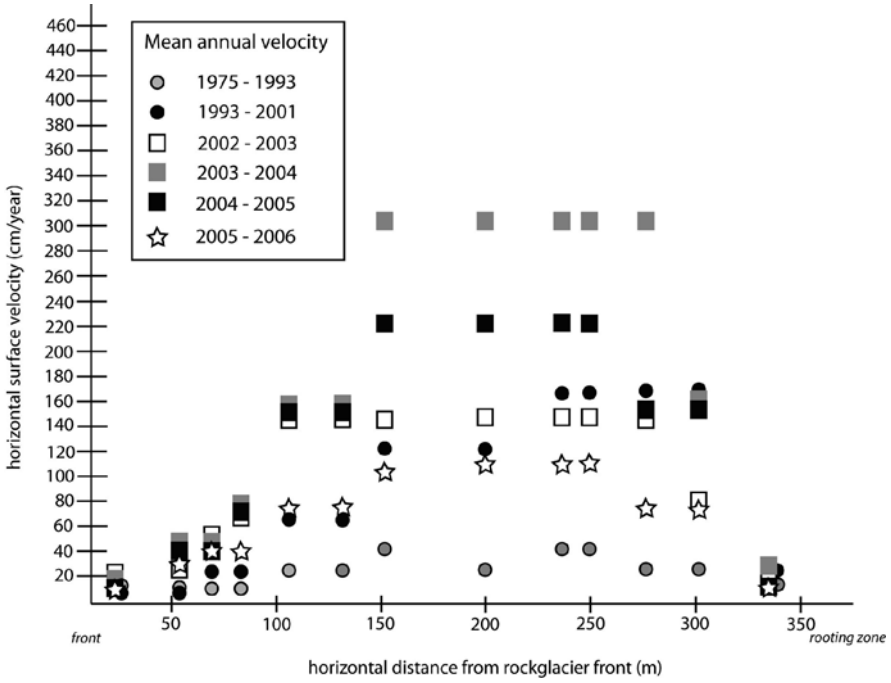

Fig. 3 Mean annual velocities of selected blocks (in a transect from the front to the rooting zone) at the surface of rockglacier HuHH3 for different periods. The values for 1975–1993 and 1993–2001 are derived from digital photogrammetry (compare Fig. 2), while the other values result from geodetic survey

determined for the entire landform. In general, movement rates are small at the rockglacier front, in the rooting zone as well as at the margins, while the central flow field shows the highest rates (see Fig. 2). Therefore, the mean annual velocities representative for the entire rockglacier are calculated for a longitudinal transect (compare Fig. 3).

4 Results and Discussion

4.1 Rockglacier Thicknesses and Volumes

The analysis of rockglacier thicknesses indicates that apparent differences exist between the two methods applied. The thicknesses calculated by Method 1 generally exceed those calculated by Method 2. Based on these methods, two scenarios for sediment volume have been calculated. The results show that Scenario 1 represents a maximum value for the estimation of rockglacier volumes, probably exceeding the true volumes; Scenario 2 is therefore regarded to represent a more realistic estimation for the rockglaciers in the Turtmann Valley (Nyenhuis, 2006). Regarding the different states of rockglacier activity, only small differences emerge in the statistical characteristics (Table 1). At least for the relict landforms, smaller thicknesses were expected. The reasons for the general similarity of the calculated thicknesses seem to be manifold. One reason may be the method itself; since the thickness is measured at the margin of the landform, the characteristic collapse structures, which are often observed at the surface of relict rockglaciers, are not reflected in the measurements. Therefore, the calculated thicknesses of relict rockglaciers may have been overestimated to a certain extent. In addition, the high value of the standard deviation complicates a general conclusion on characteristic thicknesses of active, inactive and relict rockglaciers. Furthermore, also geomorphological causes may account for the resulting values (Nyenhuis, 2006).

The calculated volumes of the rockglaciers, as well as the deduced sediment volumes and – masses, are summarised in the Tables 2a and 2b. The total volume of the rockglaciers in the Turtmann Valley amounts to 0.07 km^3 for Scenario 1 and to 0.05 km^3 for Scenario 2. Hence, the volume per rockglacier averages out at $0.79 \cdot 10^6 m^3$ (Scenario 1) and $0.55 \cdot 10^6 m^3$ (Scenario 2), with a mean real area of 0.06 km^2. For active landforms, the average sediment volume amounts to

Table 1 Statistical characteristics of the calculated rockglacier thicknesses (in metres) using Scenario 1 and 2 (Nyenhuis, 2006)

Count		Scenario 1			Scenario 2		
		Mean	Median	Standard deviation	Mean	Median	Standard deviation
Active rockglaciers	38	12.4	11	6.9	8.8	8.8	2.7
Inactive rockglaciers	24	11.5	10	5.2	7.6	6.3	3.7
Relict rockglaciers	21	11.1	10	5.5	8.7	8	3.6

Table 2a 2a + 2b: Calculated rockglacier volumes, as well as sediment volumes and – masses of active, inactive and relict rockglaciers in the Turtmann Valley by the application of Scenario 1 (Table 2a) and Scenario 2 (Table 2b) (Nyenhuis, 2006)

				Sediment content 30 %		Sediment content 40 %		Sediment content 50 %		Sediment content 60 %	
	Count	Real area [km^2]	Volume [$10^6 m^3$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]
Active rockglaciers	38	2.54	39.0	11.71	30.45	15.6	40.56	19.5	50.7	–	–
Inactive rockglaciers	24	0.68	8.73	–	–	3.49	9.07	4.37	11.35	5.24	13.62
Relict rockglaciers	21	1.73	17.7	–	–	–	–	8.85	23.01	10.62	27.61
Sum	83	4.95	65.5	–	–	–	–	32.72	85.06	–	–

Table 2b

				Sediment content 30 %		Sediment content 40 %		Sediment content 50 %		Sediment content 60 %	
	Count	Real area [km^2]	Volume [$10^6 m^3$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]
Active rockglaciers	38	2.54	24.86	7.48	19.4	9.94	25.94	12.43	32.32	–	–
Inactive rockglaciers	24	0.68	5.79	–	–	2.31	9.03	2.9	7.54	3.74	9.72
Relict rockglaciers	21	1.73	15.1	–	–	–	–	7.55	19.63	9.06	23.56
Sum	83	4.95	45.7	–	–	–	–	22.88	59.49	–	–

Table 3 Minimum and maximum estimations of sediment volumes and masses for the rockglaciers in the Turtmann Valley (Nyenhuis, 2006)

	Szenario 1		Szenario 2	
	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]	Sediment volume [$10^6 m^3$]	Sediment mass [$10^6 t$]
Minimum	24.05	62.53	17.33	45.06
Maximum	35.36	91.33	25.23	65.6

$0.41 \cdot 10^6 m^3$ (Scenario 1) and $0.26 \cdot 10^6 m^3$ (Scenario 2), with a mean real area of 0.07 km^2 per rockglacier. When calculating sediment volumes of rockglaciers in the Alps, Barsch (1977) refers to an average sediment volume of $0.66 \cdot 10^6 m^3$ with an area of 0.03 km^2. The considerably lower values for the Turtmann Valley relates to the smaller thicknesses determined for the rockglaciers. The thicknesses typically range between 3 and 38 m (Scenario 1) and 3 and 19 m (Scenario 2), with mean values of 12 m (Scenario 1) and 8.4 m (Scenario 2) respectively. According to Barsch (1996), characteristic thicknesses of alpine rockglaciers vary between 30 and 50 m, whereas for West Greenland Humlum (2000) gave a wider range of 15–80 m. The thicknesses determined here are significantly lower.

The minimum and maximum estimations of sediment volumes and masses for the rockglaciers in the Turtmann Valley are summarised in Table 3. They result from the smallest and highest possible values of each rockglacier class (active, inactive, relict). For example, the minimum estimation for Scenario 2 is composed of the sum of sediment volumes and sediment masses of the following percentage of sediment: active rockglaciers 30%, inactive rockglaciers 40% and relict rockglaciers 50%. The maximum estimation takes these sediment percentages into account: active rockglaciers 50%, inactive and relict rockglaciers 60% (Nyenhuis, 2006).

4.2 Rockglacier Kinematics

Rockglacier kinematics were quantified and analysed on various spatial and temporal scales by the combination of digital photogrammetric and geodetic survey. Maximum horizontal displacements in the range of 3–5 m/a were observed on several rockglaciers (Roer, 2007). Thus, most rockglaciers in the Turtmann Valley indicate above-average horizontal velocities compared to other rockglaciers in the Alps, which show rates in the range of centimetres to one meter per year (Haeberli, 1985). On the individual landforms, spatial variations in horizontal velocities mostly emphasize the rockglacier topography or rather vice versa. In general, velocities are highest in the central flowline of the individual lobes and depict a distinct decrease toward the margins, due to increased friction (see Figs. 2 and 3). Regarding temporal variations, an increase in horizontal velocities within 26 years of photogrammetric monitoring was ascertained for all investigated rockglaciers (Roer, 2007). Using geodetic survey on two landforms, continued high velocities are confirmed for the years 2002–2004 (with maximum values in the year 2003/2004), followed

by a distinct decrease in the years 2004–2006. Probable controls on the observed speed-up of rockglaciers are described and discussed in Roer et al. (2005), Roer (2007) and Kääb et al. (2007).

4.3 Sediment Transfer Rates

For the estimation of sediment transfer rates, data on thickness, volume and average horizontal velocity were combined and exemplified for one rockglacier (HuHH3). For the determination of representative average values of horizontal velocities and in order to compare the results of the two methods applied, the mean values were calculated on the basis of a longitudinal transect with 13 measurements for each period (Fig. 3). From this result the following mean annual values: 0.24 m/a (1975–1993), 0.9 m/a (1993–2001), 0.97 m/a (2002–2003), 1.7 m/a (2003–2004), 1.12 m/a (2004–2005) and 0.67 m/a (2005–2006). Hence, great differences in transfer rates result from the different average values. Table 4 gives a comparison of characteristics for rockglacier HuHH3 and Murtèl rockglacier (Barsch, 1977).

Based on the data given in Table 4, sediment transfer rates were calculated. With a mass of 1.696 million tons and an average speed of 0.06 m/a, the annual transfer rate of Murtèl rockglacier amounted to 0.098 M t/a (Barsch, 1977). For the estimation of sediment volume and mass for the rockglacier HuHH3, Scenario 2 was chosen. Calculated with the much smaller mass of 0.53 million tons but with different speeds, the annual transfer rate of rockglacier HuHH3 ranges between 0.13 M t/a for the period 1975–1993 and a maximum value of 0.9 M t/a for the year 2003/2004.

When calculating sediment transfer rates from the given data, it has to be considered that information on internal deformation is lacking due to the absence of boreholes. Horizontal velocities represent the cumulative deformation of the entire rockglacier, without giving information which part of the landform is moving.

Table 4 Comparison of characteristics from rockglacier Murtèl (from Barsch 1977) and HuHH3 (based on Scenario 2)

	Length	Width	Thickness (max)	Volume (M m^3)	Scenario 2, vol 40% (M m^3)	Scenario 2, mass 40% (M t)*	Average speed [measurement period]
Murtèl	370	150	44	1.6	0.64	1.696	0.06 m/a [unknown]
HuHH3	310	140	9	0.5	0.2	0.53	0.24 m/a [1975–1993]
HuHH3	310	140	9	0.5	0.2	0.53	1.7 m/a [2003/2004]

*Calculated with a specific weight of 2.65 t/m^3 (Barsch, 1977).

Arenson et al. (2002) analysed horizontal borehole deformation for several rockglaciers in the Alps and showed that on some landforms the whole mass is moving, while on others only the upper layer moved. Hence, uncertainties and errors may be included in the calculation of sediment transfer rates.

5 Conclusion

A combination of geomorphological mapping, DTM analyses, digital photogrammetry and geodetic survey has been applied in order to quantify sediment transfer rates of rockglaciers. Compared to other studies (Barsch, 1996; Humlum, 2000), the calculated rockglacier thicknesses are very small with 12 m (Method 1) and 8.4 m (Method 2), respectively. According to that and in combination with the given areal extent of the landforms, also the calculated volumes and masses are rather small (Scenarios 1 and 2). All active rockglaciers showed high horizontal velocities and indicated strong temporal variations in their kinematics during the last three decades. Based on the high horizontal velocities, the calculated sediment transfer rates were quite high with a maximum value of 0.9 Mt/a (2003/2004), even if the masses are small.

Accounting for the simplicity of thickness models and due to the limited knowledge on internal deformation as well as the strong spatial variations in velocity, the amount of moving sediment can only be estimated and conclusions have to be drawn with care. Hence the question arises, in which way the usage of rockglaciers as archives for the reconstruction of sediment transfer rates can be improved and existing uncertainties regarding the flow behaviour of rockglaciers can be reduced. Future studies will therefore aim at analysing sediment transfer rates of rockglaciers on a regional scale including the calculation of denudation rates, which will be compared with typical Alpine denudation rates determined within other studies in the Turtmann Valley and other nearby study areas. For further analyses, supplementary data like borehole information and geophysical soundings have to be included and correction factors have to be considered. In addition, the measurement of rockglacier kinematics needs to be continued, in order to further monitor the dynamic behaviour of the landforms.

In view of the given findings and due to the high density of rockglaciers in the study area, it can be concluded that rockglaciers contribute significantly to the sediment budget of the Turtmann Valley. The extraordinary temporal dynamics of the landforms, the resulting variability of sediment transport rates as well as the consequences regarding possible instabilities, need to be analysed in more detail.

Acknowledgments The study was funded by the Research Training Group 437 (Landform – a structured and variable boundary layer) of the Deutsche Forschungsgemeinschaft (DFG). Thanks are dedicated to the Swiss Federal Office of Topography (Swisstopo) for access to the aerial photographs. We gratefully acknowledge most valuable comments by Ian S. Evans.

References

Arenson, L.U., Hoelzle, M. & Springman, S., 2002. Borehole deformation measurements and internal structure of some rock glaciers in Switzerland. Permafrost and Periglacial Processes, 13: 117–135.

Barsch, D., 1977. Nature and importance of mass-wasting by rock glaciers in alpine permafrost environment. Earth Surface Processes, 2: 231–245.

Barsch, D., 1992. Permafrost creep and rockglaciers. Permafrost and Periglacial Processes, 3:175–188.

Barsch, D., 1996. Rockglaciers: Indicators for the present and former geoecology in high mountain environments. Springer, Berlin, 331 pp.

Barsch, D. & Caine, N., 1984. The nature of mountain geomorphology. Mountain Research and Development, 4 (4): 287–298.

Barsch, D. & Jakob, M., 1998. Mass transport by active rock glaciers in the Khumbu Himalaya. Geomorphology, 26: 215–222.

Caine, N., 1974. The geomorphic processes of the Alpine environment. In: Ives, J.D. & Barry, R.G. (eds.): Arctic and Alpine Environments. London: 721–748.

Caine, N. & Swanson, F.J., 1989. Geomorphic coupling of hillslope and channel systems in small mountain basins. Zeitschrift für Geomorphologie, 33: 189–203.

Calkin, P.E., Haworth, L.A. & Ellis, J.M., 1987. Rock glaciers of central Brooks Range, Alaska, USA. In: Giardino, J.R., Shroder, J.F. & Vitek, J.D. (eds.): Rock glaciers. London: 65–82.

Chorley, R.J. & Kennedy, B.A., 1971. Physical geography – a systems approach. Prentice Hall, London.

Church, M. & Slaymaker, O., 1989. Disequilibrium of Holocene sediment yield in glaciated British Columbia. Nature, 337: 452–454.

Dietrich, W.E. & Dunne, T., 1978. Sediment budget for a small catchment in mountainous terrain. Zeitschrift für Geomorphologie, Suppl.-Bd., 29: 191–206.

Frauenfelder, R. & Kääb, A., 2000. Towards a palaeoclimatic model of rock-glacier formation in the Swiss Alps. Annals of Glaciology, 31: 281–286.

Gorbunov, A.P., 1983. Rock glaciers of the mountains of Middle Asia. In: Pewe, T.L. & Brown, J. (eds.), 4th International Conference on Permafrost. Fairbanks, Alaska. National Academic Press, Washington: 359–362.

Gray, J.T., 1970. Mass wasting studies in the Oglivie and Wernecke Mountains, Central Yukon Territory. Geological Survey of Canada Papers, 70-1 (A): 192–195.

Haeberli, W., 1985. Creep of mountain permafrost: Internal structure and flow of Alpine rock glaciers. Mitteilungen der VAW-ETH Zürich, 17: 139 pp.

Höllermann, P., 1983. Blockgletscher als Mesoformen der Periglazialstufe. Studien aus europäischen und nordamerikanischen Hochgebirgen. Bonner Geographische Abhandlungen, 67: 73 pp.

Humlum, O., 2000. The geomorphic significance of rock glaciers: estimates of rock glacier debris volumes and headwall recession rates in West Greenland. Geomorphology, 35: 41–67.

Jäckli, H., 1957. Gegenwartsgeologie des bündnerischen Rheingebietes – ein Beitrag zur exogenen Dynamik alpiner Gebirgslandschaften. Beiträge zur Geologie der Schweiz. Geotechnische Serie, 36: 136 pp.

Jordan, P. & Slaymaker, O., 1991. Holocene sediment production in Lillooet River basin, British Columbia, a sediment budget approach. Géographie Physique et Quaternaire, 45: 45–57.

Kääb, A. & Vollmer, M., 2000. Surface geometry, thickness changes and flow fields on creeping mountain permafrost: automatic extraction by digital image analysis. Permafrost and Periglacial Processes, 11 (4): 315–326.

Kääb, A., Frauenfelder, R. & Roer, I., 2007. On the response of rockglacier creep to surface temperature increase. Global and Planetary Change, 56: 172–187.

Messerli, B. & Zurbuchen, M., 1968. Blockgletscher im Weissmies und Aletsch und ihre photogrammetrische Kartierung. Die Alpen, 3: 1–13.

Nyenhuis, M., 2006. Permafrost und Sedimenthaushalt in einem alpinen Geosystem. Bonner Geographische Abhandlungen, 116: 142 pp.

Nyenhuis, M., Hoelzle, M. & Dikau, R., 2005. Rock glacier mapping and permafrost distribution modelling in the Turtmanntal, Valais, Switzerland. Zeitschrift für Geomorphologie, 49 (3): 275–292.

Otto, J.-C., 2006. Paraglacial sediment storage quantification in the Turtmann Valley, Swiss Alps. URL: http://hss.ulb.uni-bonn.de/diss_online/math_nat_fak/2006/otto_jan-christoph.

Otto, J.-C. & Dikau, R., 2004. Geomorphologic system analysis of a high mountain valley in the Swiss Alps. Zeitschrift für Geomorphologie N.F., 48 (3): 323–341.

Otto, J.-C., Kleinod, K., König, O., Krautblatter, M., Nyenhuis, M., Roer, I., Schneider, M., Schreiner, B. & Dikau, R., 2007. HRSC-A data: a new high-resolution data set with multipurpose applications in physical geography. Progress in Physical Geography, 31 (2): 179–197.

Rapp, A., 1960. Recent development of mountain slopes in Kärkevagge and surroundings. Geografiska Annaler, 42A: 65–200.

Reid, L.M. & Dunne, T. (eds.), 1996. Rapid evaluation of sediment budgets. GeoEcology paperback. Catena Verlag, GeoScience Publisher, Reiskirchen: 164 pp.

Roer, I., 2007. Rockglacier kinematics in a high mountain geosystem. Bonner Geographische Abhandlungen, 117: 217 pp.

Roer, I. & Nyenhuis, M., 2007. Rockglacier activity studies on a regional scale: comparison of geomorphological mapping and photogrammetric monitoring. Earth Surface Processes and Landforms, 32: 1747–1758.

Roer, I., Kääb, A. & Dikau, R., 2005. Rockglacier acceleration in the Turtmann valley (Swiss Alps) – probable controls. Norwegian Journal of Geography, 59 (2): 157–163.

Schrott, L. & Adams, T., 2002. Quantifying sediment storage and Holocene denudation in an Alpine basin, Dolomites, Italy. Zeitschrift für Geomorphologie N.F., Suppl.-Bd., 128: 129–145.

Schrott, L., Hufschmidt, G., Hankammer, M., Hoffmann, T. & Dikau, R., 2003. Spatial distribution of sediment storage types and quantification of valley fill deposits in an alpine basin, Reintal, Bavarian Alps, Germany. Geomorphology, 55: 45–63.

Wagner, S., 1992. Creep of alpine permafrost, investigated on the Murtel Rock Glacier. Permafrost and Periglacial Processes, 3: 157–162.

White, S.E., 1971. Rock glacier studies in the Colorado Front Range, 1961–1968. Arctic and Alpine Research, 3: 43–64.

Patterns of Multiannual Aggradation of Permafrost in Rock Walls with and Without Hydraulic Interconnectivity (Steintälli, Valley of Zermatt, Swiss Alps)

Michael Krautblatter

Abstract This article shows monthly, annual and multiannual response of two rock permafrost systems with and without hydraulic interconnectivity. It is hypothesized, that interconnective systems with cleft water percolation from glacierets close to 0°C are high-entropy systems that balance rock masses at 0°C and prevent cooling but are not effective in excessive melting, as thermal conduction away from water pathways into compact rock is a function of sensitive temperature gradients. This was tested using static (permafrost distribution in 2005) and dynamic (permafrost aggradation 2005–2007) performance of two adjacent north-exposed transects with similar geometries, geology and discontinuity patterns. Transect NW is only affected by heat transfer from the rock surface. Transect NE conducts hydraulic heat transfer with glacierets by meltwater seepage. Electrical resistivity tomography (ERT) time-sections and mean apparent resistivity – median depth of investigation (AR/DOI) gradients of steep sections (> 60°) were analysed from 2005–2007. (i) In 2005, a body in a transitory (0°C) resistivity range (< 20 kΩm) was developed in Transect NE. Transect NW indicated a deeply frozen body (> 40 kΩm) of several meters diameter. (ii) Negative AR/DOI surface gradients indicate a pronounced short-term response of surface resistivities in Transect NW to surface chilling (August 13, 2007: –3.3 kΩm/m) and cool pulse propagation whereas Transect NE is well buffered (August 13, 2007: –0.1 kΩm/m). (iii) Cool mid-summer conditions in 2005 and 2006 initiated permafrost aggradation in both transects. In Transect NW, ERT displays resistivity increases by more than 70%, a spatially aggrading permafrost body and the formation of a new perennially frozen rock body. Resistivity in Transect NE increases 10–30% in the transitory body. (iv) In Transect NW, the AR/DOI gradient increased from 5 kΩm/m in August 2005 to 11 kΩm/m in August 2007, indicating significant permafrost aggradation and cooling. In Transect NE, AR/DOI increased from 0.6 kΩm/m in August 2005 to 1.0 kΩm/m in August 2007 but resistivities still do not exceed the initial freezing range significantly at any depth

M. Krautblatter (✉)
Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany
e-mail: michael.krautblatter@giub.uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*,
Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_13,

of investigation. Data reliability of both transects is assessed in terms of uncertainty and relative sensitivity plots.

Keywords Rock permafrost · Response time · Model uncertainty · ERT · Hydraulic interconnectivity · Entropy · DOI

1 Introduction

Recently, much attention is spent to bedrock permafrost systems as rockwall instabilities and permafrost-related rockfalls appear to be an increasing threat to human lives and anthropogenic structures in high-mountain areas (Gruber and Haeberli, 2007; Haeberli, 1992; Harris et al., 2001). The Dzhimarai-khokh rock/ice avalanche in 2002 (Russian Caucasus) detached approximately 4 million m^3 from a permafrost-affected starting zone and caused more than 140 casualties (Haeberli, 2005; Haeberli et al., 2003; Kääb et al., 2003). Even smaller permafrost rockfalls, such as the 2003 Matterhorn rockfall are considered as major hazards in densely populated high mountain areas (Gruber et al., 2004). Inventories show that the frequency of these rockfalls has considerably increased in the warm 1990 s and was boosted by the hot summer of 2003 (Schoeneich et al., 2004). Moreover, slow rock creep in permafrost rocks causes significant damage especially to tourist infrastructure in high mountain areas.

Besides temperature logger data, borehole information and rock temperature modeling approaches, geophysical applications provide a new perspective for the spatial and temporal analysis of rock permafrost (Krautblatter and Hauck, 2007). Permafrost geophysics provide a new insight into the spatial and temporal internal behavior of rock permafrost systems in response to external climatic forces on a monthly, annual and multiannual scale. A review of external factors influencing the rock permafrost system has been given by Gruber and Haeberli (2007) and these factors will not be discussed here. This article rather concentrates on internal system behaviour and tries to answer three questions:

(i) What factors govern response in hydraulically connective and non-connective rock permafrost systems in terms of thermodynamics? (ii) Do resistivity signals reveal spatial and temporal patterns of rock permafrost aggradation over time? (iii) How reliable is such data in terms of data uncertainty and sensitivity? (iv) Do time-lapse geophysics reveal systemic differences of permafrost aggradation in hydraulically connective and non-connective rock masses?

2 System Considerations

Permafrost is a thermally defined phenomenon as ground that remains below 0°C for at least 2 years irrespective of the presence of water in the system. Besides rock matrix, bedrock permafrost systems involve ice, air and often a residuum of unfrozen water in pores and rock discontinuities. Thus, the rock permafrost system

basically consists of rock, ice, air and water. The thermal energy ($U_{thermal}$) of this system is added up by both,

$$U_{thermal} = U_s + U_l \tag{1}$$

sensible energy (U_s), that is related to the molecular kinetic energy (translation, rotation, vibration and spin) and latent energy (U_l) that is associated with phase transition. To simplify the equations, less important factors in compact rock masses such as sensible heat transfer to air and water in pores/discontinuities will not be mentioned in detail.

2.1 Basic Sensitive System

Let us first consider a closed system that exclusively consists of rock without pores and discontinuities, has theoretically no spatial extension and no heat exchanges with the lithosphere below or the atmosphere above. Any heat (Q) put into or taken out of the system (ΔU_s=Q) will result in a change of temperature (ΔT),

$$\Delta T = \Delta U_s/(c^*m) \tag{2}$$

where m is the rock mass and c is the specific heat capacity, that typically approaches values of 0.8–1.3 (kJ/kg*K) according to rock type.

2.2 Adding the Spatial Dimension

As soon as the closed rock system has a certain spatial extent (and is observed for a restricted time) the entropy concept becomes important. According to the second law of thermodynamics, heat energy added to one part of the system will not be fully dispersed in the system but part of it will rather serve to increase the entropy of a system. This has two fundamental consequences. Heat will always be transferred from the warmer to the colder part and the speed of this transfer will decrease with lower temperature differences corresponding to an increase in entropy. The quantity of heat Q

$$(Q/A) = (k/l)^* \Delta T^* t \tag{3}$$

is transmitted during time t perpendicular to rock surface A through a rock wall thickness l for a given temperature gradient ΔT. The thermal conductivity k approaches values from 1 to 3.5 W/(mK) for most rock types.

2.3 Adding Latent Phase Transitions

While the presence of non-circulating air with its low heat capacity in pores is not very important in the system, water has a significant influence, as the phase transition from ice to water requires latent energy (U_l),

$$\Delta U_l = V_r{}^* p_w{}^* 334(MJ/m^3) \quad (4)$$

where V_r is the volume of rock (in m^3) and p_w is the porosity. E.g. to increase the temperature of a m^3 of granite or limestone (2700 kg) with water-filled porosity of 5% (50 kg) from –0.5 to 0.5°C, ca. 2.2 MJ of sensible energy and 16.7 MJ of latent energy are required. Thus, frozen and thawed rocks are separated by a phase transition with enormous energy consumption as soon as water is present in the system.

2.4 Adding Discontinuous Heat Flow

Heat transmission at constant temperature gradients as presented in Fig. 1 is not realistic due to two reasons. Firstly, rock surface temperatures change steadily according to differences in radiation, air temperatures and other factors such as snow cover or ice-coating. Secondly, during heat transmission, the temperature gradient changes as the temperature in the rock interior alters. This has major implications: (i) Heat transmission to greater depths adjusts to the integral temperature gradients over a longer time. (ii) Heat transmission by short rock surface anomalies initiates warm/cool waves that propagate through the rock according to the laws of thermal conduction. (iii) These waves are often absorbed at depths where latent phase transitions occur. (iv) Heat transfer is a multidirectional phenomenon whose most effective direction is adjusted to thermal gradients.

2.5 Adding Heat Transfer in Clefts

Anisotropic behaviour of clefts becomes important in the case of rock discontinuity systems with significant cross-sectional and longitudinal extension to allow fluids and air to circulate or ice to seal clefts. Air filled-discontinuities are most common at the rock surface, while below the internal rock water level most discontinuities are filled with either ice or water. In permafrost rocks, the blocking of discontinuities against seepage by ice generally causes high water levels. Anisotropic behaviour of ice-filled clefts is a relatively straightforward concept explained by the thermal conduction of ice. Thoroughly frozen rock along ice-filled discontinuities can be explained by the good thermal conduction ($k = 2.2\ W/(m^*K)$) of ice without latent buffers in the readily frozen ice in clefts.

In the case of mass transfer in water-filled clefts between the rock permafrost system and surroundings, the concept of entropy becomes increasingly important. This is best explained by an example. The system setting is frozen rock at 0°C,

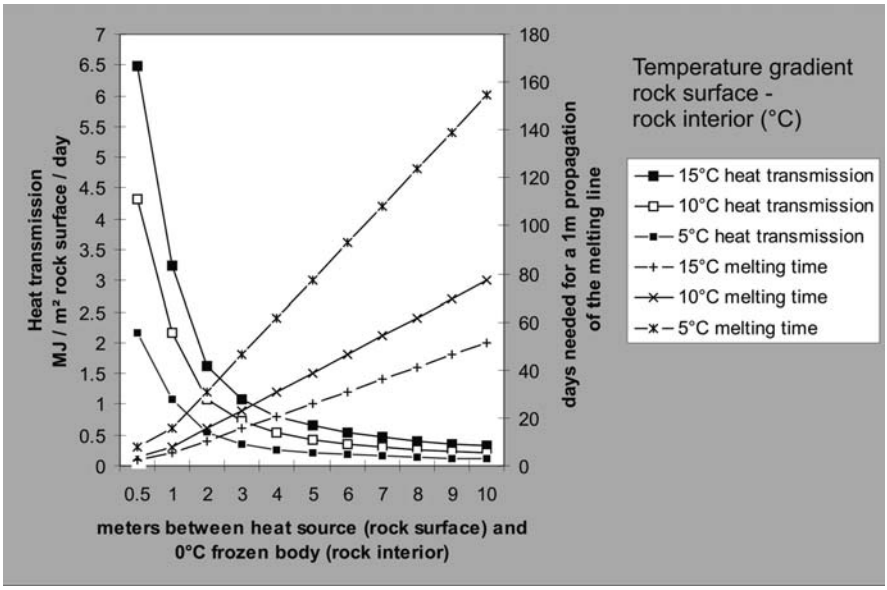

Fig. 1 Theoretical heat transmission and melting time at different depths. The left axis describes how much heat (GJ/day) is transmitted if a temperature difference of 5, 10 or 15°C exists between rock surface and the rock interior at a certain depth. The right axis shows how many days this specific heat flow must theoretically continue to let the melting line propagate one meter further. A steady state system with unidirectional heat flow, a *k*-value of 2.5, and a latent energy consumption of 16.7 MJ for melting a m^3 of rock with 5% water-filled porosity, are assumed

but still with ice-filled pores. Fluid water from the surrounding intrudes into permafrost rocks via cleft systems. If water temperature is above 0°C, water will be cooled as the heat flow away from the cleft system is proportional to the temperature gradient (Eq. 3). At 0°C, cleft water still contains 334 MJ/m^3 of latent energy, but this latent energy is no longer available for melting apart from the cleft, as heat transmission away from the cleft is determined by sensible temperature differences. During the described cooling process, entropy, the unavailability of the system to do work, increases to a maximum. While the heat uptake of sensible heat alongside the cleft follows Eq. (3), the entropy (S) [J/K] change in the system is theoretically approached by

$$\frac{dS}{dt} = \sum_{k=1}^{K} \dot{M}_k \hat{S}_k + \sum_{j=1}^{J} \frac{\dot{Q}_j}{T_j} + \dot{S}_{\text{int}} \tag{5}$$

where $\sum_{k=1}^{K} \dot{M}_k \hat{S}_k$ is the net rate of entropy flow due to mass exchange with the surroundings, $\sum_{j=1}^{J} \frac{\dot{Q}_j}{T_j}$ is the net rate of entropy due to heat exchange with the

surroundings and $\dot{S}_{int}$ is the internal entropy generation. This means that in the case of cleft water close to 0°C, the thermal effect is determined by hydraulic flow patterns in cracked rocks as heat transmission away from the clefts is not very effective.

2.6 Adding Feedbacks

The ice-water transition in the rock permafrost systems causes some positive/negative feedbacks. E.g. a rock mass with ice-filled porosity has a slightly higher *k*-value than thawed rock with water-filled porosity. This results in a slightly more effective cooling during the freezing period in comparison to warming in the summer period. The same is true for small clefts in which no turbulent mixing processes take place. However, these feedbacks are possibly not relevant on a larger scale.

2.7 Response Times

According to the mentioned thermodynamic rules, one can outline some general findings for rock permafrost response times: (i) The temporal dimension of response times is governed by the speed of heat propagation. (ii) Dynamics and direction of heat propagation are dominated by sensitive thermal gradients. (iii) Discontinuous heat flux at the rock surface results in steadily changing temperature gradients and results in daily warm/cold pulses in the upper meter, a propagation of several meters for pulses generated during several day-lasting unusual cold/warm periods and integral seasonal pulses that usually reach a penetration depth in the range of one decameter. (iv) Latent heat transfer from frozen to thawed rock requires high amounts of energy, often stops heat pulses and increases response times at greater depths enormously. (v) Permafrost rocks perform more heat transfer with the surrounding than non-permafrost rocks as latent phase transitions keep temperature gradients constant and as frozen rock has a higher thermal conductivity. (vi) Anisotropic behaviour of well-jointed rock masses significantly influences heat transfer and response times. (vii) In the case of hydraulic interconnectivity with water temperatures close to 0°C, spatial patterns of melting are orientated directly along hydraulic pathways as heat propagation away from clefts is not effective.

Fig. 2 (b) (continued) line is indicated by white points. The discussed part (–30 to –6 m) is the steep rock face above the (**c**) Transect NE is situated below a glacieret that feeds meltwater into bedrock cleft systems. The approximate position of the 60 m long electrode line is indicated by white points. The discussed part (–30 to +3 m) is the steep right part of the transect glacieret.

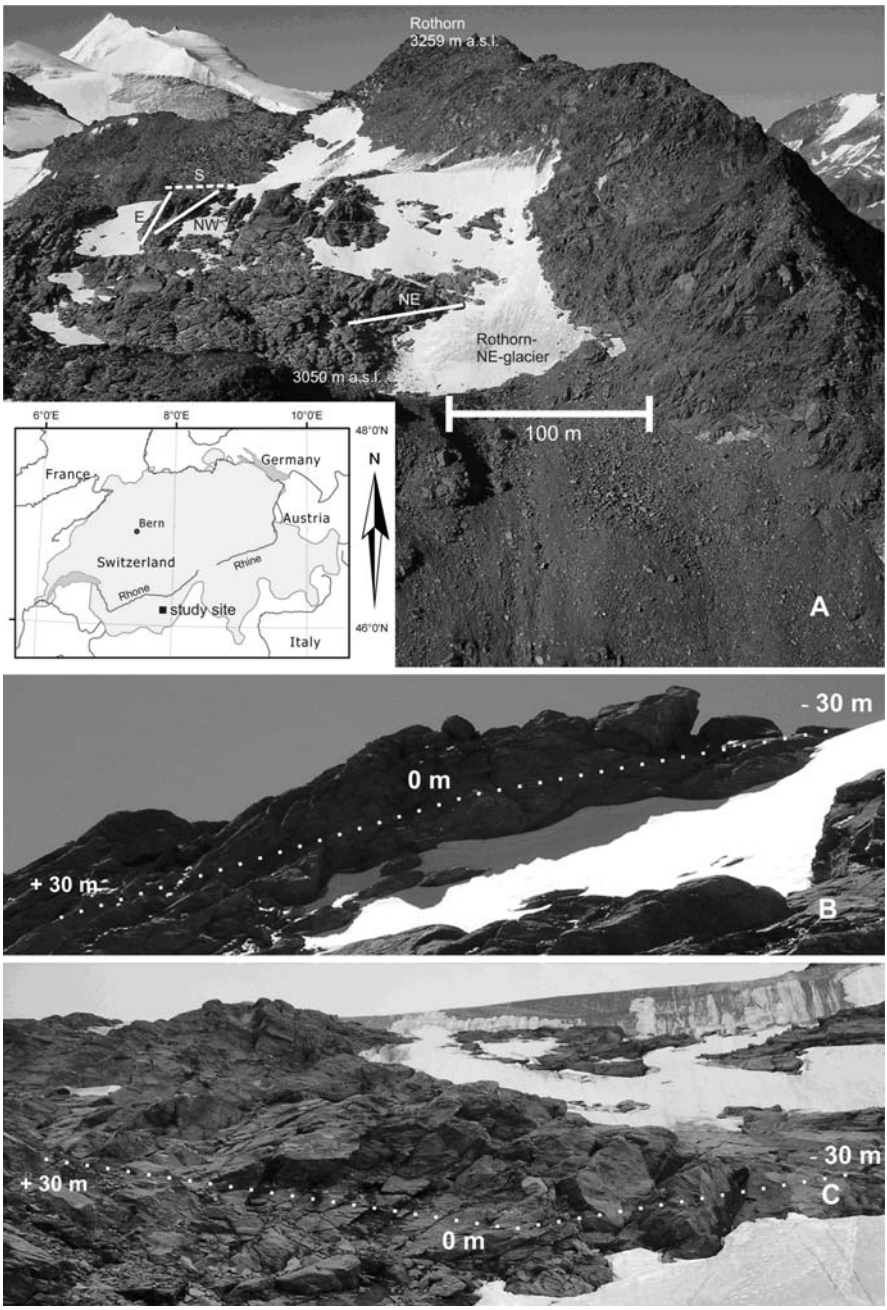

Fig. 2 (a) Overview of the study site "Steintälli". The Turtmann Valley is situated right (*West*) of the Rothorn, the Matter Valley is visible on the left hand side (*East*). Positions of ERT-transects are indicated. (**b**) Transect NW is affected by marginal onfreezing of a lateral glacieret but no hydraulic exchange in the level of the transect. The approximate position of the 60 m long electrode

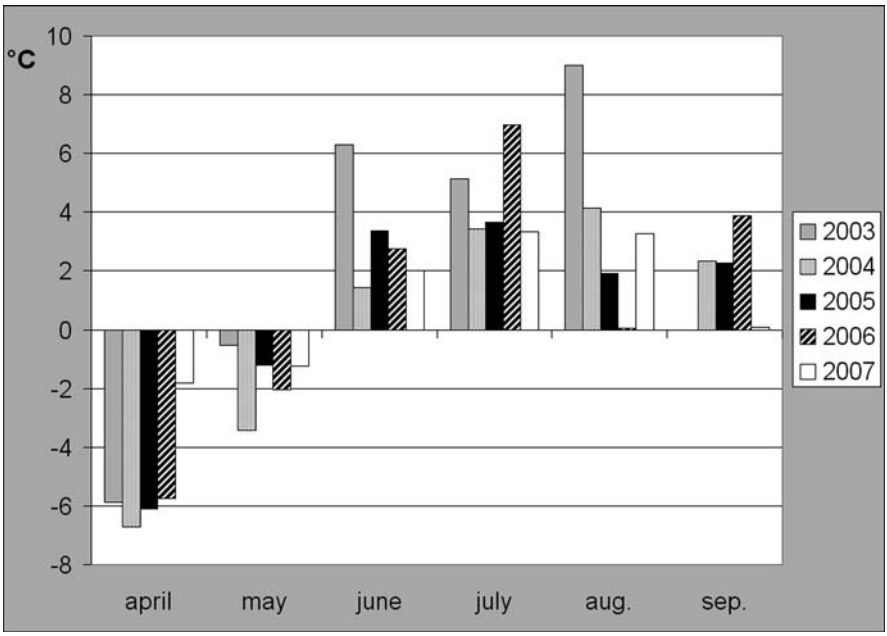

Fig. 3 Average monthly temperatures from 2003 to 2007 measured at a meteorological station located at a horizontal distance of 900 m and at an altitudinal difference of 330 m from the study site (2770 m a.s.l.). Temperature data were corrected by –1.98°C (0.6°C per 100 m) for the altitudinal difference to the Steintälli 3100 m a.s.l. The average annual temperature of 2004 at the study site was –3.7°C, which is in accordance with the long-term average value. Data from September 2003 are incomplete and, therefore, not displayed

3 Field Site

The geophysical measurements described in this article were conducted in the "Steintälli", a N-S exposed crestline (Matter-/Turtmann Valleys, Switzerland) at about 3070–3150 masl. that consists of slaty paragneiss (Fig. 2A). The annual average air temperature of the Steintälli approached values of –3.5°C according to measurements taken in the adjacent Matter Valley 1962–1990 (see Fig. 3). A local permafrost study conducted by Gruber and Hoelzle (2001) in the Matter Valley ranked the whole Steintälli in the highest possible permafrost category as "likely permafrost area". As all rock bars are surrounded by ice, the local air temperatures can often be significantly lower than the temperature calculated from the meteorological station. Transect NE receives meltwater from a glacieret above that performs intensive advective heat transfer in cleft systems, while Transect NW has only lateral contact to glacieret ice, with marginal onfreezing (Fig. 2B, C).

4 Geophysical Evidence

ERT is a key method in permafrost research as freezing and thawing of most materials are associated with a resistivity change that spans one or more orders of magnitude, which is, in turn, easily detectable. The fist approach to derive spatial information from steep rock faces by ERT was applied by Sass (2003). In subsequent studies, he provided further evidence that ERT measurements are capable of measuring the degree of rock moisture (Sass, 2005) and temporal and spatial variations of freeze and thaw limits (Sass, 2004) in rock faces. These ERT measurements were confined to the monitoring of the upper weathering crust (centimeter- to decimeter-scale) of non-permafrost rock faces. Krautblatter and Hauck (2007) extended this method to a decameter-scale and applied it to the investigation of active-layer processes in permafrost-affected rock walls. Due to the large electrode spacing and the resulting spatial resolution, the ER tomographies presented here are prone to investigate monthly to multiannual responses but are less applicable to short time series (see Sass, 2004).

4.1 Applied Methods and Error Sources

A systematic discussion of error sources associated with ERT measurements in rocks is given by Krautblatter and Hauck (2007). These include electrokinetic potentials created by percolating water, high resistivity gradients along the suspected permafrost bodies that can result in an over- or underestimation of the depth of the transition layer, high resistivities of frozen rock surface that may disturb electrode contact as well as topographic distortions when measuring on a structured rock face. Errors associated with different ERT-arrays were assessed quantitatively alongside with the impact of topography and other geometric error sources by Krautblatter and Verleysdonk (2008).

A number of hardware and software adaptations were made to cope with these error sources. Arrays with centimeter-long steel screws, lubricated with battery grease, were drilled into solid rock as electrodes to gain stable electric contact. These were left in place for identical monitoring conditions and were measured repeatedly with high voltages (mostly 10^2–10^3 V) to improve the signal to noise ratio. We applied a SAS 300C 41-electrode system that was sometimes supplemented with a booster to yield higher voltages. Resistivities were repeatedly measured with reversed currents to exclude the influence of self potentials. A detailed survey of hardware and software adaptations is provided by Krautblatter and Hauck (2007). The RES2DINV software was chosen as it is capable of topographic correction and time-lapse inversion of subsequent measurements. To cope with high resistivity gradients, inversion models with mesh size smaller than the electrode distance and robust inversion routines provide better results. Resistivity values that correspond to the transition between frozen and thawed rock were measured repeatedly

at the rock surface along different arrays and yielded evidence that the transition occurs between 13 and 20 kΩm for the slaty paragneiss present at the Steintälli (Krautblatter and Hauck, 2007) and is in the same range as the one established for carbonate rocks at the Zugspitze, German Alps, by Sass (2004).

The relation between measured resistivity and rock temperature changes at the freezing point. For temperatures below the freezing point, resistivity (p) depends mainly on unfrozen water content until most of the pore water is frozen. In Alpine environments, resistivity can be calculated based on a reference value p_o as an exponential response to the temperature below the freezing point (T_f) according to McGinnis et al. (1973):

$$\rho = \rho_0 e^{b(T_f - T)} \tag{6}$$

The factor b in Eq. (6) determines the rate of resistivity increase and can be derived empirically (Hauck, 2001, 2002). Short-term changes in resistivity can be attributed to changes in pore water content and temperature, while changes in porosity and water chemistry can be neglected over daily to monthly measurement intervals in low-porosity rocks. Due to the exponential response of resistivity to temperatures below 0°C, freeze-thaw transitions correspond to an increase in resistivity by one order of magnitude and are thus, a very sensitive method for detecting the state of rock permafrost close to 0°C.

4.2 Data Quality

Even if the RMS error provides some information on outliers, i.e. measurements that were take at a transitorily deeply frozen rock surface with decreased electrode contact, it is not a straightforward approach to judge data quality in terms of RMS errors. This is due to two factors. Firstly, rock permafrost systems have high resistivity gradients and even small distortions between apparent and modelled resistivity sections result in high RMS errors that are not comparable to systems with lower gradients (Blaschek et al., 2008). Secondly, RMS errors are possibly a systematic proxy for fractured rock masses as these show significant anisotropic behaviour in the subscale, that is not resolved by ERT measurements with meter-scale spacings (Linde, 2005)

To restrict overfitting of results, no more than five iterations were applied. Typical RMS errors usually range between 5 and 15% for 1.5 m electrode spacing at the Steintälli when measured under convenient conditions with only slightly frozen rock surfaces. Hardware errors of subsequent measurement with identical electrodes are mostly below 1%. Measurements during intensely frozen and ice-coated rock surface conditions inherit worse hardware measurement errors (> 1%) and result in RMS errors ranging from 20 to 30% with visible distortions in the inversion results. These were excluded from time-lapse interpretations.

Model uncertainties were calculated in RES2DInv according to the model covariance matrix method described in Alumbaugh and Newman (2000) (Fig. 4).

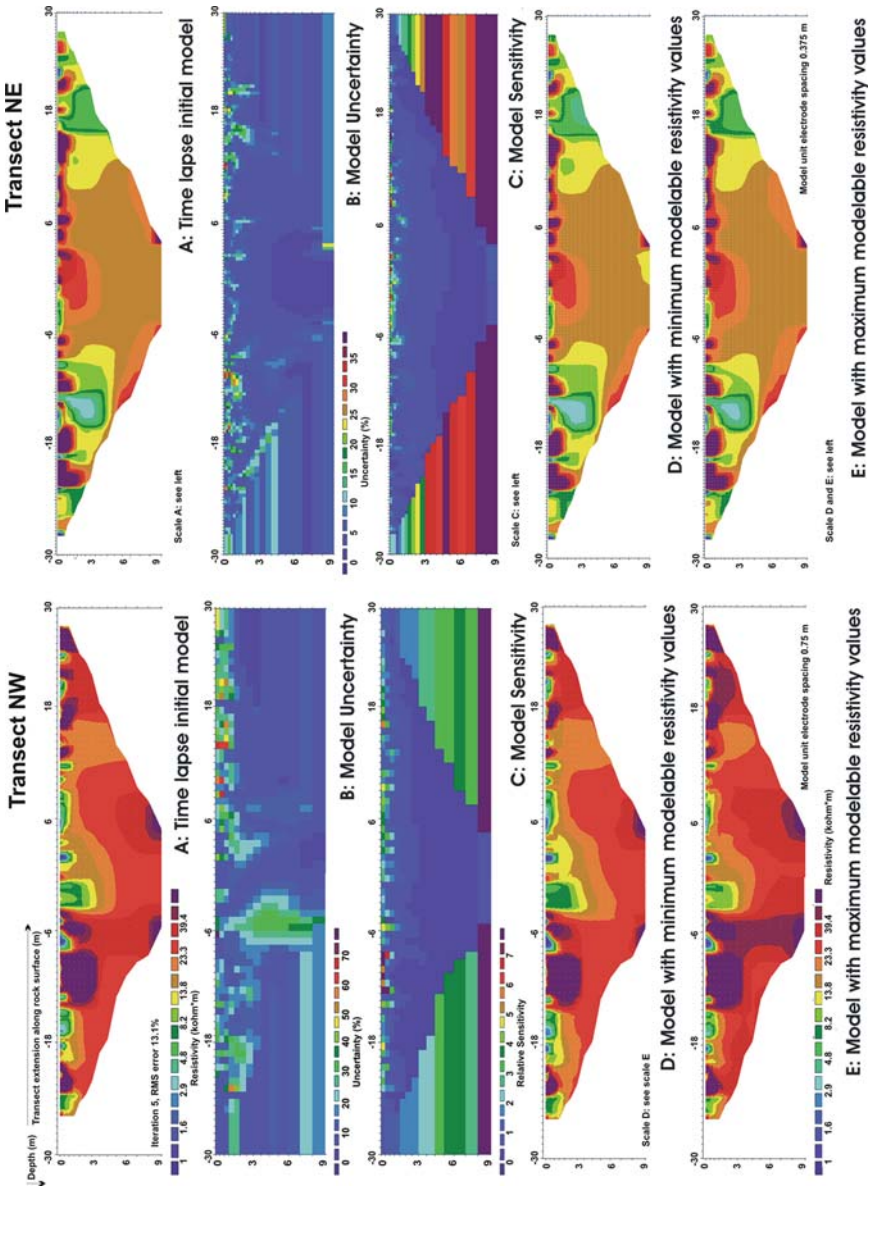

Fig. 4 Model uncertainty and relative sensitivity of the time lapse initial models of Transect NW and NE

Minimum and maximum models display average models plus or minus model uncertainty. Relative sensitivity is displayed in Fig. 4 and is divided by average sensitivity. Mention that high sensitivity of side blocks is an artificial effect due to their large size.

Uncertainties range between 0 and 5% in sections of *Transect NE* that are primarily interpreted in the article (–6 to +6; 3–8 m depth). Minimum and maximum modelable resistivities show only small alterations in the interpreted section. Interpreted monthly resistivity changes of 20–30% (–6 to +6; 3–8 m depth) in Fig. 6B, D exceed model uncertainty by at least 4 times in the largest part of the section. Interpreted annual changes of 10–30% (–6 to +6; 3–8 m depth) in Fig. 6C, E exceed model uncertainty by several times. The distribution of relative sensitivity indicates that the depth of investigation is higher in Transect NE and thus, to depth resistivity information is more reliable than in Transect NW (Fig. 4C).

Frozen surface rock (–15 to –3 m) in *Transect NW* causes model distortions centred at –6 m/2.5–8 m depth with model resistivity uncertainties ranging from 15 to 35%. Interpreted monthly changes of mean thaw line propagation (23 kΩm line) in Fig. 8A, B could vary up to 1–1.5 m according to minimum and maximum models (Fig. 4D, E). Interpreted annual changes (>70%) in the left part (–20 to –4.5 m/3–9 m depth) in Fig. 8C, D are located in a part of the pseudosection with 0–20% model

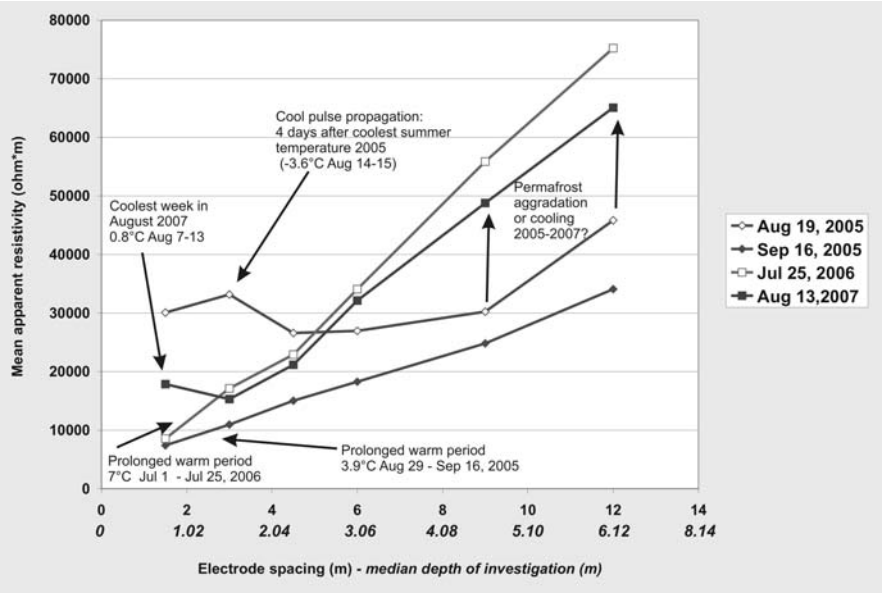

Fig. 5 Apparent resistivities of the permafrost-affected left side (electrode array midpoints –30 to –6 m) of Transect NW with increasing electrode spacing. The median depth of investigation (MDOI) is written in italics. It shows approximately at what depth one-half of the total resistivity signal originates from above and one-half from below (Barker, 1977)

resistivity uncertainty. Interpreted annual changes in the right part (4.5–20 m/2–9 m depth) in Fig. 8C, D have model resistivity uncertainties ranging from 5 to 15% with interpreted changes being in the range of 50–70%.

4.3 Monthly, Annual and Multi-annual Response in a Rock Wall Without Hydrologic Interconnectivity

4.3.1 Raw Data

Figure 5 shows mean apparent resistivities for different electrode spacings corresponding to different depths (Barker, 1977; Edwards, 1977; Loke and Barker, 1996) for Transect NW. Measurements taken after prolonged warm periods show constant resistivity gradients with increasing depth of investigation. This is in line with slow linear heat propagation in summer from the warm rock surface to a permafrost rock interior. Cool pulses cause significant distortion of this general behaviour at the rock surface. The plunge in temperature of the August 14–15, 2005 appears to propagate as a high resistivity pulse. After 4 days, at August 19, is no longer apparent at a MDOI of 0.8 m, is most pronounced at a MDOI of 1.5 m and has not reached a MDOI of 2.3 m. This corresponds to a mean speed of cool pulse propagation ranging from 30–50 cm/day to a MDOI of 1.5 m. Mean apparent resistivity data of August 2005 and August 2007 indicate an increase by 20 kΩm for a median depth of investigation of 4–6 m. This could be due to significant cooling or/and due to aggradation of frozen rock.

4.3.2 ERT Results

Figure 6 presents four time sections of a permafrost-affected rock face exposed to NW (see Fig. 2B). The most pronounced appearance of high resistivity bodies close to the rock surface is apparent in 6A following unusually cool Mid-August conditions in 2005. The warm September 2005 (3.9°C) coincides with the most pronounced advance of the 23 kΩm line from 5 m depth to almost 9 m in the centre of the transect (0 m), where the most reliable depth information is attained (Fig. 6B). The high resistivity body on the left side (–20 to –4.5 m) in 3–9 m depth, increases significantly in size and resistivity in the following 2 years (6C and 6D). On the right side (4.5–20 m) in 2–9 m depth a new high resistivity body evolves in 2006 and 2007 (6C and 6D).

4.3.3 ERT Interpretation

Figure 6 shows two time-scales of response times. On a monthly scale (6A and 6B) the impact of a warm month following a cool summer period can be observed. On

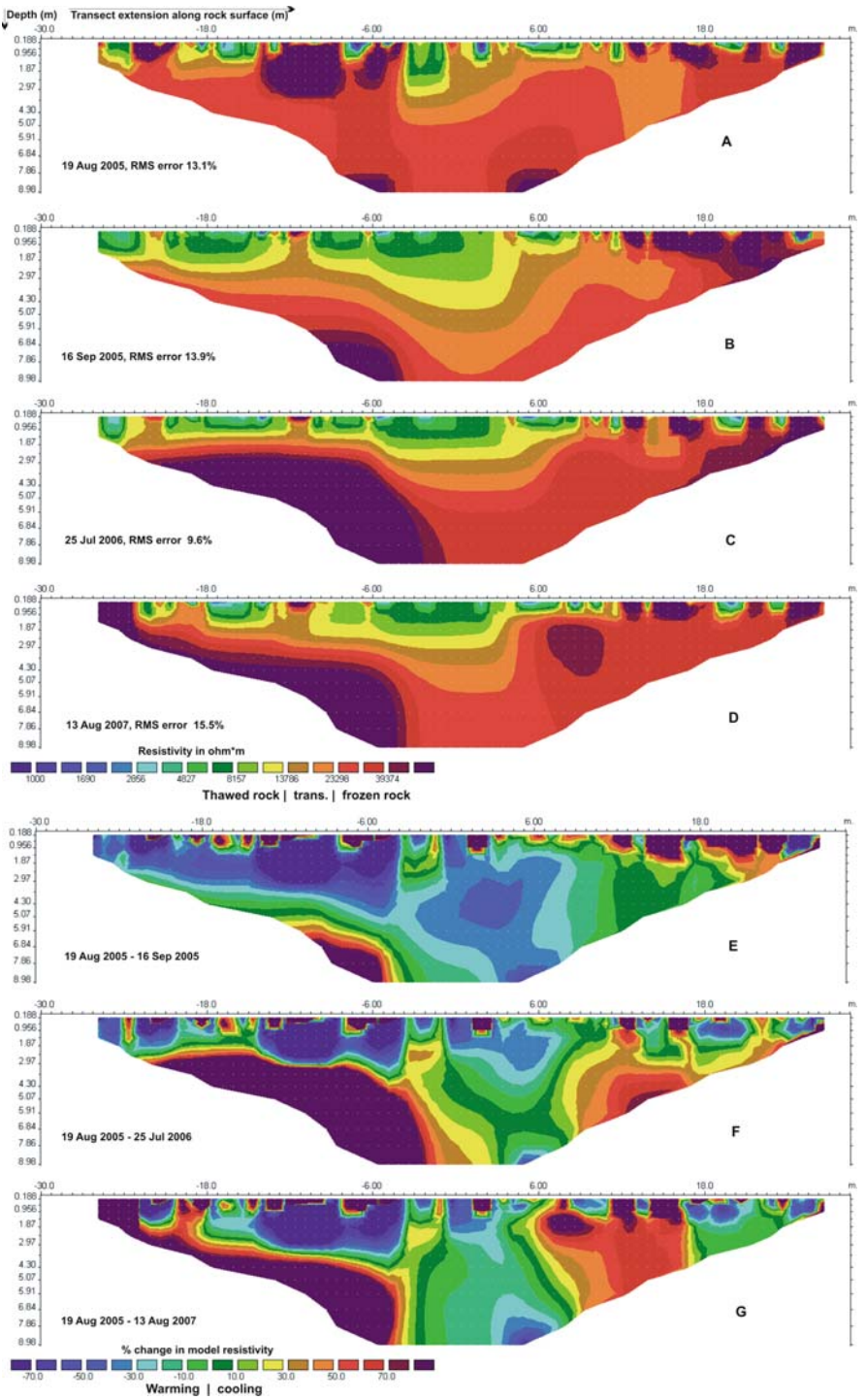

Fig. 6 (continued)

annual and multiannual scale permafrost aggradation in response to cool summer conditions 2005 and 2006 can be observed. For a better description, Transect NW will be divided in a (i) left section (–30 to –6 m), a (ii) middle section (–6 to 6 m) and a (iii) right section (6–30 m).

The surface layer of the *right section* is of stepped appearance and more exposed to radiation. Simultaneous ERT and P-wave velocity measurements (Krautblatter et al., 2007) yielded evidence that high ERT values in this part in the upper 1–2 m match with low P-wave velocities in the range of 3000 m/s, that correspond to a dry, air-filled rock mass, while in the middle and left part, high ERT-values coincide with high P-wave velocities in the range of 4000 m/s and above and are, thus attributed to frozen, ice-filled rock mass. This affects the interpretability of the monthly response. On the annual and multiannual timescale a permafrost body builds up between 6–18 m with deeply frozen resistivity values that exceed 40 kΩm in 2007. Figure 6F, G indicate a 20–60% increase in model resistivity until 2006 and a 30–80% increase when comparing 2007 and 2005. This data shows significant aggradation and cooling of rock permafrost from 2005 to 2007.

The *middle section* (–6 to 6 m) highlights a monthly response of the melting line in warm late summer conditions (6A and 6B). Resistivity losses of 20–50% illustrate enhanced heat transmission from the rock surface to the rock interior at all depths. The 14 kΩm line, which approaches thawed conditions slightly above the freezing point, advances as much as 2–3 m and the 23 kΩm line, which corresponds to frozen conditions slightly below the freezing point, advances as much as 4 m in 1 month. The transition zone between 14 kΩm and 23 kΩm appears to increase in thickness from roughly 1.5–2.5 m to 2–5 m. In the annual and interannual scale permafrost aggradation is visible in the proximity to the left section but is less pronounced than in the two other sections.

The *left section* also indicates strong surface heat transmission in late summer 2005 and resistivity values decrease more than 50% in the upper 3 m. This results in a 0–3 m advance of the 14 kΩm line and a 0–4 m advance of the 23 kΩm line gapped by a transition zone mostly smaller than 2 m. The impact of the warm summer is less pronounced than in the middle section and is possibly compensated by the deeply frozen permafrost lens (>50 kΩm) below. On an annual and multiannual scale, permafrost aggrades spatially only 1–2 m but resistivity values strongly increase by more than 70%. Below a relatively constant 3 m thick active layer the permafrost lens appears to cool down.

Two features visible in Fig. 6 represent additional factors that interfere with rock permafrost response and cause additional complexity. A totally shaded overhang position between –10.5 and –12 m, causes near surface rock permafrost that is also evidenced by permanently ice-filled crevasses in the field. Intense snow accumulation in 2006 caused an increase in thickness of the glacieret close

Fig. 6 Four ERT time-sections of Transect NW that are mostly determined by the rock surface temperature regime. 4A–4D show absolute resistivity values while 4E–4G show relative changes in resistivity. Raw data is displayed in Fig. 5

to the left part of the transect up to two meters and intense heat transmission from ice frozen to the rock up to the electrode level is visible from 24 to 27 m.

4.4 Monthly, Annual and Multi-annual Response in a Rock Wall with Hydraulic Interconnectivity

4.4.1 Raw Data

Melting in Transect NE in summer 2005 causes a decrease in apparent resisitivities at depth while mean surface resistivities (electrode spacings 1.5 and 3m) remain stable with a change of –0.0 (±0.5) kΩm (Fig. 7). Mean subsurface apparent resistivities, covering median depths of investigation greater than 2.5 m, decrease uniformly for all electrode spacings (4.5, 6, 9, 12 and 18 m) by 2.1 (±0.1) kΩm. In accordance to Transect NW mean apparent resisitivity values with a MDOI greater than 7 m possibly indicate aggradation of frozen rock and/or cooling. Table 1 shows that the AR/DOI gradient of Transect NE is an order of magnitude lower than the AR/DOI of Transect NW. This indicates that resistivity values do not signif-

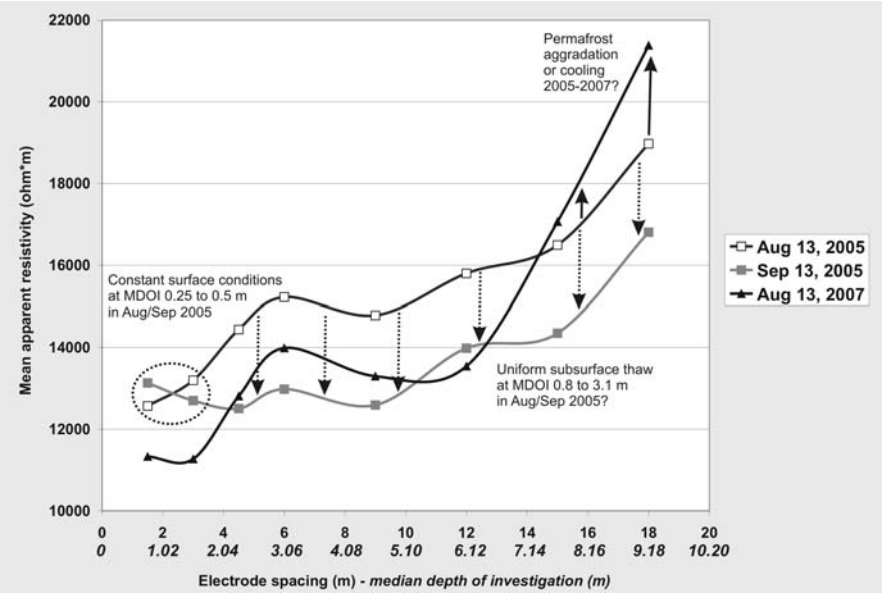

Fig. 7 Mean apparent resistivities of the left side of Transect NE (electrode array midpoints –30 to + 3 m) with increasing electrode spacing. The median depth of investigation, where half of the at which one-half of the total resistivity signal originates from above and one-half is given in italics (Barker 1977)

Table 1 Resistivity increase with penetration depth of resistivity measurements in the rock face. Data sets are identical with those displayed in Figs. 5 and 7. High AR/DOI gradients and high resistivities at maximum recorded depths correspond to deeply frozen permafrost conditions. Increases in AR/DOI gradients point towards permafrost aggradation. Surface distortions (1.5 and 3 m electrode spacing) due to cold surface condition in August 2005 and 2007 in Transect NW were excluded from the linear regression. As the analysed steep (> 60°) section of Transect NW only extends from –30 to –6 m, the maximum recorded MDOI is restricted to 6.1 m

Date	Gradient of a linear regression of median depth of investigation mean apparent resistivity (AR/DOI)	Apparent resistivity (at maximum recorded MDOI)
Transect NW – not hydraulically connective (steep > 60° section, electrode array midpoints –30 to –6 m)		
August 19, 2005	4.9 kΩm/m ($R^2 = 0.83$)	46 kΩm (6.1 m)
September 16, 2005	4.9 kΩm/m ($R^2 = 0.99$)	34 kΩm (6.1 m)
July 25, 2006	12.7 kΩm/m ($R^2 = 0.99$)	75 kΩm (6.1 m)
August 13, 2007	11.1 kΩm/m ($R^2 = 0.99$)	65 kΩm (6.1 m)
Transect NE- hydraulically connective (steep > 60° section, electrode array midpoints –30 m to + 3 m)		
August 13, 2005	0.6 kΩm/m ($R^2 = 0.90$)	19 kΩm (9.2 m)
September 13, 2005	0.4 kΩm/m ($R^2 = 0.70$)	17 kΩm (9.2 m)
August 13, 2007	1.0 kΩm/m ($R^2 = 0.81$)	21 kΩm (9.2 m)

icantly exceed the transitory range of freezing (13–20 kΩm) around 0°C at any depth.

4.4.2 ERT Results

Figure 8 displays a central high-resistivity body (–7 to 11 m) that is in the range of the transition from frozen to thawed rock and is restricted by two pronounced wet rock low-resistivity bodies at –22 to –8 m and 12 to 22 m. The central high-resistivity body shows a decrease in resistivity in the range of 10–40% that is more pronounced at greater depths in September 2005. Two years later resistivity values in the central transitory body increase by 0–40%.

4.4.3 ERT Interpretation

Figure 8 shows meta-stable rock permafrost (–6 to 12 m) close to the freezing point. Its spatial extension is delimited by two cleft water systems that create significant cleft water outflow at the rock surface with water temperatures close to 0°C.

On a monthly scale (Fig. 8B, D), the 14 kΩm line at position –9 to 3 m at 3.5 to 9 m depth retreats several meters in September 2005 which is also apparent as a 2.1 (±0.1) kΩm shift for 4.5, 6, 9, 12 and 18 m electrode spacing in the raw data (Fig. 7). There are some indications that this melting is due to cleft water propagation as heat transmission in the ERT and the raw data does not originate from the

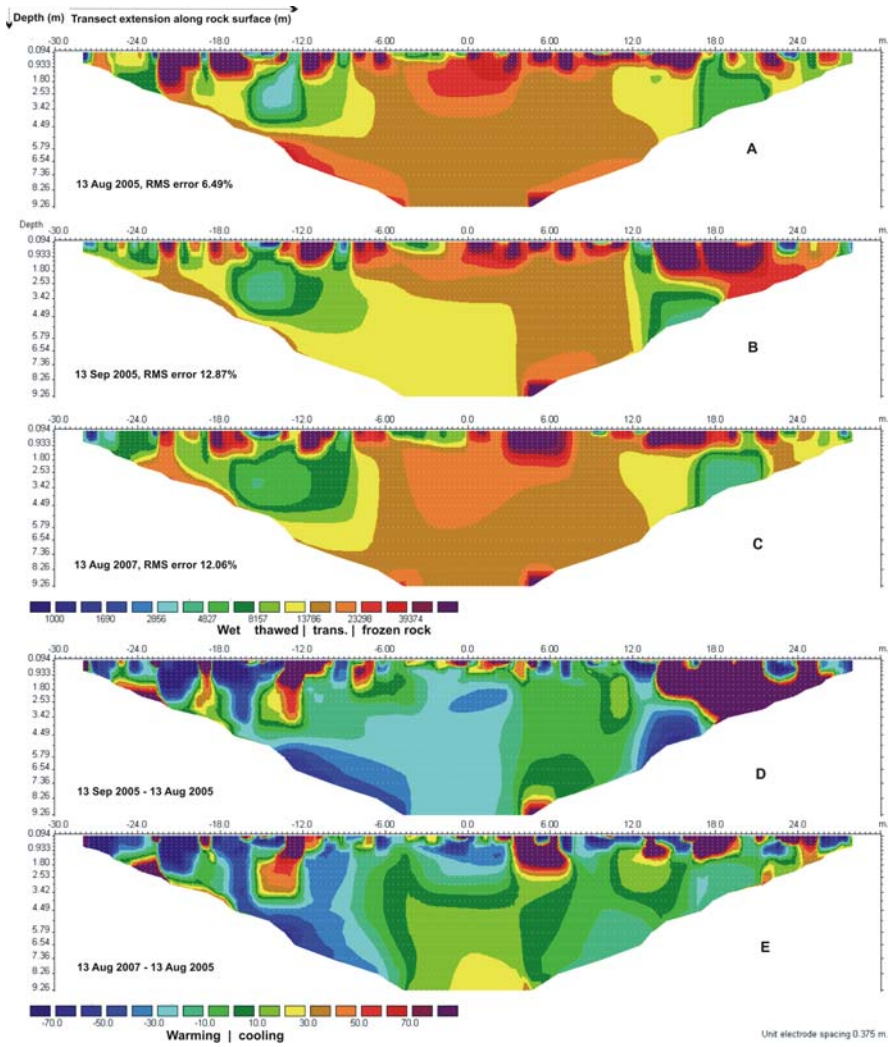

Fig. 8 Four ERT time-sections of Transect NE determined by both, the rock surface temperature regime and heat transfer by meltwater in water-filled cleft systems. 6A–C show absolute resistivity values while 6D–E show relative changes in resistivity

rock surface but starts in 2.5 to 3 m depth and the core zone of the cleft water body at –12 to –7 m with wet rock resistivity values indicated by the 8.2 and 10.6 kΩm (transition green – yellow) line extends simultaneously.

On a multiannual scale (Fig. 8C, E) the transitory permafrost body shows an increase in resistivity by 10–30% in the core area, that points towards a shift from a metastable permafrost body towards a more deeply frozen body with twice the extension of September 2005, but still close to 0° C. This cooling clearly derives

from the rock surface. However, ERT and AR/DOI indicate (i) that cleft water systems are still effective restricting the size of permafrost bodies (no significant aggradation) and (ii) even if weather conditions yielded significant cooling in the adjacent Transect NW, the transitory Transect NE did cool significantly below 0°C at any observed depth (no significant cooling).

5 Conclusion

In homogeneous rock wall sections, mean apparent resistivity – median depth of investigation (AR/DOI) diagrams appear to reveal information on the state of permafrost. Measurements after prolonged warm periods in Transect NW, not disturbed by cleft water, show constant resistivity gradients with increasing DOI corresponding to linear heat propagation from the warm rock surface to a permafrost rock interior. Cool pulse signals from transitorily chilling of the rock surface propagate through these linear functions as local maxima with speed 3–5 decimetres per day to a depth of 1.5 m. AR/DOI gradients are steep in the range of 5–13 kΩm/m. Apparent resistivity values reach values of 30–80 kΩm at depth indicating deeply frozen conditions. In Transect NE, disturbed by cleft water percolation, the AR/DOI curve deviates from a linear function with local maxima and minima that remain constant over time. Cool pulses are already buffered at the surface, AR/DOI gradients are low reaching values from 0.4 to 1 kΩm/m and apparent resistivities at depth do not significantly exceed 20 kΩm. This indicates that resistivity gradients with depth have persistent minima, possibly due to cleft water, and that no deeply frozen rock body can develop in this system.

In Transect NW, as a response to a monthly warm late summer period in 2005, the melting line was observed to propagate up to 4 m per month and significant melting evidenced by wide-spread resistivity changes above 30% were observed up to 9 m depth. As an annual response to cool summer conditions an existing permafrost body (–18 to –6 m) extends its deeply frozen core (> 40 kΩm) from more than 7 to 3–4 m from the rock surface which results in an resistivity increase greater than 70% in most of its extension. Transect NE is heavily affected by the heat transfer of two cleft water systems which only allow for the development of a meta-stable permafrost body, whose resistivity values indicate temperatures close to 0°C. There are some indications that the impact of cleft water-induced melting is visible on a monthly scale in September 2005.

From 2005 to 2007, both rock transects provide indications of permafrost aggradation and/or cooling. Anomalously cool August temperatures in 2005 (1.9°C) and 2006 (0.1°C) with fresh snow and ice-coating of rock faces reduced active layer propagation during the short summer period significantly. *Transect NW* indicates both, the consolidation of existing permafrost bodies and the development of new perennial frozen rock bodies. Above the existing permafrost body in the steep section (–30 to –6 m), ERT indicates a reduction of the active layer thickness and an increase in resistivity by more than 70%. For this body mean apparent resistivity

data of August 2005 and August 2007 indicate an increase by 20 kΩm for a MDOI of 4.6 m and 6.1 m which is interpreted as permafrost aggradation and/or cooling. A new high-resistivity body builds up at 6–18 m mostly with resistivity changes of 30–70% from 2005 to 2007. In *Transect NE*, ERT indicates a moderate cooling in the high-resistivity body with a resistivity increase of 10–30% from 2005 to 2007 paralleled by apparent resistivity raw data. In spite of similar surface climate conditions from 2005 and 2007, Transect NW shows both, permafrost aggradation and cooling while the hydraulically-interconnective Transect NE shows nor pronounced aggradation nor pronounced cooling.

Aggradation in permafrost rocks with and without active cleft water systems evokes significantly different patterns. While heat propagation in cleft-water affected systems occurs locally and is necessarily delimited by hydrological conductivity, systems without cleft-water influence indicate a fast and deep-reaching transmission of heat from all parts of the rock surface. While permafrost aggradation in the latter is merely an integral signal from cool pulses coming from the rock surfaces, sealing of subordinate cleft systems with ice may constitute a major influence on lateral permafrost aggradation in cleft water-disturbed systems. 3D thermal effects could play an vital role in this context.

References

Alumbaugh, D.L. and Newman, G.A., 2000. Image appraisal for 2-D and 3-D electromagnetic inversion. Geophysics, 65(5): 1455–1467.

Barker, R.D., 1977. Depth of investigation of collinear symmetrical four-electrode arrays. Geophysics, 54(8):1031–1037.

Blaschek, R., Hördt, A. and Kemna, A., 2008. A new sensitivity-controlled focusing regularization scheme for the inversion of induced polarization data based on the minimum gradient support. Geophysics, 73(2): F45–F54.

Edwards, L.S., 1977. A modified pseudosection for resistivity and IP. Geophysics, 42(5): 1020–1036.

Gruber, S. and Haeberli, W., 2007. Permafrost in steep bedrock slopes and its temperature-related destabilization following climate change. J. Geophys. Res.Earth Surf, 112(F2): F02S13.

Gruber, S. and Hoelzle, M., 2001. Statistical modelling of mountain permafrost distribution: Local calibration and incorporation of remotely sensed data. Permafrost Periglac. Process., 12(1): 69–77.

Gruber, S., Hoelzle, M. and Haeberli, W., 2004. Permafrost thaw and destabilization of Alpine rock walls in the hot summer of 2003. Geophys. Res. Lett., 31(13): L15054.

Haeberli, W., 1992. Construction, environmental problems and natural hazards in periglacial mountain belts. Permafrost Periglacial Processes, 3: 111–124.

Haeberli, W., 2005. Investigating glacier-permafrost relationships in high-mountain area: historical background, selected examples and research needs. In: C. Harris and J.B. Murton (Editors), Cryospheric Systems: Glaciers and Permafrost. Geological Society Special Publication, London, pp. 29–37.

Haeberli, W. et al., 2003. Permafrost conditions in the starting zone of the Kolka-Kamadon rock/ice-slide of the 20th September 2002 in North Osetia (Russian Caucasus). . In: W. Haeberli and D. Brandova (Editors), ICOP 2003 Permafrost: Extended Abstracts, Zürich, pp. 49–50.

Harris, C., Davies, M.C.R. and Etzelmüller, B., 2001. The assessment of potential geotechnical hazards associated with mountain permafrost in a warming global climate. Permafrost Periglacial Processes, 12(1): 145–156.

Hauck, C., 2001. Geophysical methods for detecting permafrost in high mountains. Mitt.Versuchsanst. Wasserbau, Hydrologie und Glaziologie, PhD-thesis ETH Zürich, 171: 1–204.

Hauck, C., 2002. Frozen ground monitoring using DC resistivity tomography. Geophys. Res. Lett., 29, 2016, doi: 10.1029/2002GL014995: 12-1.

Kääb, A. et al., 2003. Rapid aster imaging facilitates timely assessments of glacier hazards and disasters. EOS, 13(84): 117, 121.

Krautblatter, M. and Hauck, C., 2007. Electrical resitivity tomography monitoring of permafrost in solid rock walls. J. Geophys. Res. Earth Surf, 112. doi:10.1029/2006JF000546.

Krautblatter, M., Hauck, C. and Wolf, S., 2007. Geophysical 2D and 3D-monitoring of permafrost in rock walls. Geophys. Res. Abstr., 9: A-09884.

Krautblatter, M. and Verleysdonk, S., 2008. Rock wall permafrost monitoring with high-resolution 2D-ERT: lessons learnt from error estimates and a comparison of Wenner, Schlumberger, Gradient and Dipole-type arrays. Geophys. Res. Abstr., 10: A-10383.

Linde, N., 2005. Characterization of Hydrogeological Media using electromagnetic geophysics. Digital Summaries of the Uppsala Dissertations from the Faculty if Science and Technology, 86: 65.

Loke, M.H. and Barker, R.D., 1996. Rapid least-squares inversion of apparent resistivity pseudosections by a quasi-Newton method. Geophys. Prospecting, 44(1): 131–152.

McGinnis, L.D., Nakao, K. and Clark, C.C., 1973. Geophysical identification of frozen and unfrozen ground, Antarctica, 2nd Int. Conf. on Permafrost, Yakutsk, Russia, pp. 136–146.

Sass, O., 2003. Moisture distribution in rockwalls derived from 2D-resistivity measurements. Z. Geomorph. N.F., Suppl.-Bd. 132(51–69).

Sass, O., 2004. Rock moisture fluctuations during freeze-thaw cycles: Preliminary results from electrical resistivity measurements. Polar Geography, 28(1): 13–31.

Sass, O., 2005. Rock moisture measurements: Techniques, results, and implications for weathering. Earth Surf. Process. Landforms, 30(3): 359–374.

Schoeneich, P. et al., 2004. A new Alpine rockfall inventory, Swiss Geoscience Meeting, Lausanne.

Resilience, Integrity and Ecosystem Dynamics: Bridging Ecosystem Theory and Management

Felix Müller, Benjamin Burkhard and Franziska Kroll

Abstract In this paper different approaches to elucidate ecosystem dynamics are described, illustrated and interrelated. Ecosystem development is distinguished into two separate sequences, a complexifying phase which is characterized by orientor optimization and a destruction based phase which follows disturbances. The two developmental pathways are integrated in a modified illustration of the "adaptive cycle". Based on these fundamentals, the recent definitions of resilience, adaptability and vulnerability are discussed and a modified comprehension is proposed. Thereafter, two case studies about wetland dynamics are presented to demonstrate both, the consequences of disturbance and the potential of ecosystem recovery. In both examples ecosystem integrity is used as a key indicator variable. Based on the presented results the relativity and the normative loading of resilience quantification is worked out. The paper ends with the suggestion that the features of adaptability could be used as an integrative guideline for the analysis of ecosystem dynamics and as a well-suited concept for ecosystem management.

Keywords Orientor · Resilience · Adaptability · Ecosystem integrity · Ecosystem indicators · Disturbance

1 Ecosystems are Dynamic Systems, and So are their Scientific Perceptions

Although the term "ecosystem" has an age of only 70 years since Tansley's grounding in 1935, the fundamental understanding of ecosystem dynamics has been based upon a multitude of consecutive ideas. If we take a look back to the second half of the last century, the idea that increased complexity automatically leads to increased *stability* was a "conventional wisdom" in ecology (Begon et al. 1990). Elton's

F. Müller (✉)
Department of Ecosystem Research, Ecology Centre, University of Kiel, Olshausenstr. 75, 24118 Kiel, Germany
e-mail: fmueller@ecology.uni-kiel.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_14,

respective arguments from 1958 supported the development of the long-term dominating stability paradigm: Ecosystems have to be managed in a way that guarantees their *stability*, *resistance* and *robustness*.[1] These ideas were promoted accessory by the application of cybernetic approaches in ecology. They are founded on the concept of *equilibrium*, and thus the inherent target of an object was comprehended as its ability to return to the initial state after any perturbation (Grimm and Wissel 1997; Joergensen and Straskraba 2000).

In the 1970's these convictions became *vulnerable* and *fragile*. Initially, succession analyses and theories demonstrated that there are constructive pathways of (instable) dynamics in ecosystem history (Odum 1969), valuing mature states as the final developmental targets. Nearly simultaneously, the model experiments of May (1972) had shown declines of local stability with complexity. Besides the correlated doubts in the stabilizing functions of biodiversity, thereafter the whole stability concept became open to attack. Thermodynamic investigations demonstrated that in nature *irreversible* (and therefore instable) reactions are dominant (Prigogine 1980), and chaos and catastrophe theory emphasized the unpredictability of ecosystem states, thereby undermining important stability statements and opinions (Joergensen and Müller 2000). *Disequilibrium* theories were developed (Schneider and Kay 1994), and the new *resilience* paradigm started its development after Holling's paper (1986) about local surprise and global change. Destruction became an accepted agent of evolutionary development, and the perception of change has been altered totally, temporary ending in the panarchy theory of Gundersson and Holling (2002). Typically enough, this theory has been nominated with reference to the Greek god Pan, who feels responsible for the power of nature as well as panics and destabilisation. Thus, finally, we have returned to the knowledge of the Greek philosopher Heraclitus who stated that *nothing is permanent but change* already 500 years BP. Today this recognition is being claimed by the United Nations, postulating that "ecosystem change is inevitable" in their CBD ecosystem approach (see Müller and Burkhard 2007) which functions as a guideline for nature protection.

Remembering these historical trails in science, the appreciation of ecosystem dynamics provides a metaphoric example for the dynamics of the natural systems themselves: We will find no stability at all, "everything is flowing", and the question is whether the concept of resilience is able to provide better answers.[2] Referring to the recent situation we want to illuminate some components of the described "adaptive knowledge cycle" and discuss the following questions:

- Is there a general tendency in undisturbed ecosystem development?
- How can this trend be indicated?
- Which is the role of disturbance throughout ecosystem dynamics?

---

[1] several technical terms from systems theory and resilience research are defined briefly in Box 1.

[2] At least it is providing many answers, moving from 50 topical publications in 1992 to more than 250 in 2006 (Janssen 2007) and increasing the scientific and communicative complexity enormously. Thus the resilience concept is developing in a scientific version of orientor dynamics.

- What are the recent comprehensions within the resilience and adaptability concept?
- Is there a potential to quantify these attributes?

2 Ecosystems Follow Orientor Dynamics and Increase Complexification

Now – being aware of the uncertainty and vulnerability of scientific statements – the first of these questions can be discussed. In fact even throughout the dynamics of paradigms as they are described above, typical developmental patterns can be found. In ecosystems, different kinds of successions (autogenous vs. exogenous, primary vs. secondary, progressive vs. retrogressive successions) take place as ordered processes of community development at varying temporal and spatial scales (Walker and del Moral 2003; Walker et al. 2007). Changes in community structures cause changes in the physical ecosystem functions (e.g., energy budget, water and matter cycling), and these modifications vice versa cause changes of the system's processors composition, its content of information and the degree of complexity. In the end, usually a number of similarities appear in successions, leading to the typical sequences of pioneer, middle aged and mature ecosystems, which Odum has compiled in 1969. Due to individual exterior inputs, to the dynamics of constraints,

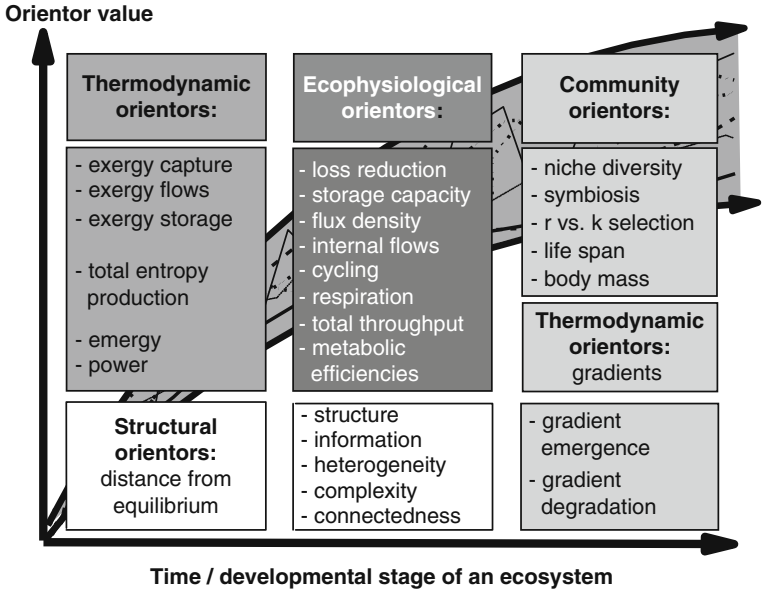

Fig. 1 Some ecosystem orientors: If an ecosystem develops in an undisturbed manner the values of the orientor variables will be optimized until they reach a site specific maximum. For further information see Joergensen et al. (2007)

Box 1: Glossary; more detailed definitions can be found in the text

Adaptability	A system has a high adaptability if the sum of all disturbances and changes in the attractor domains do not reduce the system's degree of self-organization.
Adaptive cycle	The adaptive cycle is a metaphor and a conceptual model of cyclic developmental phases of ecosystems and socio-ecological systems. The identified phases are: growth or exploitation, conservation, collapse or release, and reorganization (Holling et al. 2002).
Disturbance	Any relatively discrete event in space and time that disrupts ecosystem, community, or population structure and changes resources, substrates, or the physical environment is called disturbance (Picket and White 1985).
Ecosystem integrity	Integrity is a management target to support and preserve those processes and structures which are essential prerequisites of the ecological ability for self-organization (Barkmann et al. 2001).
Orientor	Orientors are indicating ecosystem components which are generally optimized throughout undisturbed successions.
Panarchy	Panarchy is the temporal structure in which systems are interlinked in continual adaptive cycles of growth, accumulation, restructuring, and renewal (Gundersson and Holling 2002).
Resilience	Resilience refers to the ability of a system to reorganize after a disturbance and remain in the previous basin of attraction.
Resistance	Resistance characterizes a system which is not influenced by a disturbance.
Robustness	Robustness refers to the capacity of a system to maintain the structure despite perturbations (Gallopin 2006).
Self-organization	Self-organization is the spontaneous creation of macroscopic order from microscopic disorder, whereby structural and functional gradients are created.
Stability	Stability is the ability of a system to return to an equilibrium state after a temporary disturbance (Holling 1973).
Vulnerability	Vulnerability is the exposure to contingencies and stress, and the difficulty in coping with them as a function of sensitivity and resilience (Millennium Ecosystem Assessment 2005).

and to the high degree of developmental uncertainty, maturity must be characterized as an attractor function. The concrete values of development-based indicators cannot be foreseen, and some communities even never reach the climax stage (see Fig. 1). But the general direction of that development is a clear consequence of self-organization.

In order to describe and evaluate ecosystems' developmental stages and their varying properties, *ecosystem orientors* have been proposed as indicating criteria of ecosystem states, realizing the inherent uncertainties (Müller and Leupelt 1998). Depending on the investigated dynamics and processes, different types of orientors have been defined, such as thermodynamic orientors, network orientors, ecophysiological orientors, eco-dynamic orientors and community orientors (see Fig. 1). Based on non-equilibrium thermodynamics (Joergensen 2000; Schneider and Kay 1994; Kay 2000), network development (Fath and Pattten 1998) and succession theory (e.g., Odum 1969), the *orientor approach* was developed as a systems-based outcome from the theory of self-organized ecosystem dynamics (Bossel 1998, 2000; Müller and Leupelt 1998).

The succeeding processes are based on the fact that self-organised entities like ecosystems are able to generate structures and gradients if exergy passes through the system. Exergy is the energy fraction of a system which can be converted into other forms of energy or mechanical work (Joergensen 2000). During the development, gradients and structures are built up, maintained and operated, transforming the incoming exergy (e.g., high frequency solar radiation) within metabolic reactions into non-convertible energy fractions (entropy production; e.g., CO_2 or heat from respiration). Thereby, parts of the captured exergy are stored as biomass, detritus and information. These processes of ecosystem development, including phases of exploitation, growth, and conservation, characterize a development towards an attractor state of high complexity and connectivity. Developments towards attractor states and related orientor dynamics are – in general – associated with increase of stored biomass (captured exergy, see above), more complex food webs, more diversity in species and habitats, more matter cycling and reduced nutrient loss. These developments take place in long-term trajectories on varying spatial scales (see Fig. 1 and Table 2).

What does this mean for environmental indication and management? In the following chapter we will try to build a bridge and make those theoretical conceptions applicable.

3 Ecosystem Integrity Indicates Orientor Development

To describe orientor dynamics towards maturity, comprehensive sets of long-term data from different sites at varying conditions are needed. Alternatively, or in addition, model simulations can be used for respective quantifications. As the number of elements and processes that can be measured or modelled is however limited, applicable assessments have to be based on a selected set of manageable indicators, representative for the system in focus. Nevertheless, respective indicators have to denote ecosystem processes (such as energy, matter and water budgets and cycling) and structures (abiotic and biotic components) as a whole in order to assess and evaluate the system's stage of self-organization (Müller 2004, 2005). An overview of such indicators is given in Table 1. Exemplary applications in case studies will be described later on.

The resulting indicators are representatives for orientor dynamics in ecosystems. Thus, they provide information about the developmental stage and the degree of self-organization of ecosystems and landscapes. If needed, a comparison of the actual developmental stage with an observer-defined (and thus, normative) mature or target stage can be made. By doing so, estimations of potentials for future self-organization can be suggested.

The developmental contexts and variables described above have also been used in environmental management. The respective concepts of ecosystem health (e.g., Rapport and Moll 2000) and integrity have been discussed intensively in the last years. Ecosystem health and integrity are comparable concepts, representing the

Table 1 Selected indicators of ecosystem integrity in the wetland case study

Ecosystem component	Indicator
Biotic structures	Number of plant species
Energy budgets, exergy capture	Net primary production (NPP)
Energy budgets, entropy production	Microbial soil respiration (MSR)
Energy budgets, metabolic efficiency	NPP/soil respiration
Hydrological budgets, biotic water flows	NPP/transpiration
Chemical budgets, nutrient loss	Net nitrogen mineralization (NNM)
	Nitrate leaching
	Denitrification
Chemical budgets, storage capacity	Nitrogen balance
	Carbon balance

same concern: a system based description of environmental entities to support ecosystem functioning (Rapport 1989; Haskell et al. 1993). The term "integrity" has been introduced by Leopold (1944) to characterize requirements for the stability of biotic communities. During the last decades the concept has been further developed, e.g., by Woodley et al. (1993), Westra and Lemons (1995), Crabbé et al. (2000) and Barkmann (2002). In some of these interpretations integrity is strongly related to the idea of wilderness, other authors refer to a normative, social perspective, and in a third group of interpretations integrity represents a complex systems approach, which is mainly based upon variables of energy flows, matter budgets and structural features of whole ecosystems (Barkmann et al. 2001).

Taking into account the contexts described above, it becomes clear that the ability for future self-organizing processes within the respective system has to be preserved (Kay 1993). Ecosystem integrity thus describes ecosystem functionality as a result of self-organized dynamics. Applying this viewpoint, Barkmann et al. (2001) have defined the management target of ecological integrity as "a support and preservation of those processes and structures which are essential prerequisites of the ecological ability for self-organisation".

4 Disturbance Causes Stress and Innovation

In the preceding chapter, the focus has been put on growth and development processes in ecosystems. In fact, these are important features of ecosystem dynamics. Most of its life time the system follows these traits, and they provide the origins of various emergent ecosystem properties. But the picture remains totally incomplete if disturbance and decay are not taken into account (Joergensen et al. 2007). These putatively "destructive" processes can be observed on all relevant scales: For example, death and decay of organisms and their components are integral elements of natural dynamics, populations have limited life spans at certain places, and also ecosystems themselves exist for a limited period of time only. The disturbing processes can be observed on many different scales, as results of climatic changes,

shifts of biomes, or as continuous invasions of new species. On the contrary, abrupt processes often modify ecosystems very efficiently within rather short periods of time. Besides many different human interventions, also "natural" disturbances can be found, such as volcanic eruptions, droughts, soil erosion events, avalanches, landslides, fires, windstorms, pests, or pathogen outbreaks. The consequences of such rare events can be enormous (see Picket and White 1985; Scheffer and Carpenter 2003; White and Jentsch 2001; Joergensen et al. 2007).

If we apply these general points to the conceptions mentioned before and if we restrict our argumentation to natural phenomena, it can be stated that two general processes are governing the dynamics of ecosystems: On the one hand, there are long phases of *complexification*: starting with the pioneer stage, orientor dynamics bring about slow mutual adaptation processes with long durations, if there is a dominance of biological processes. A system of interacting structural gradients is created, which provoke very intensive internal flows and regulated exchanges with the environment (Müller 1998). The processes are linked hierarchically (Müller 1992), and the domain of the governing attractor remains rather constant, whereupon optimization reactions provoke a long-term increase of orientors, efficiencies and information dynamics.

The highest state of internal mutual adaptation is attained at the *maturity* domain. But the further the system has been moved away from thermodynamic equilibrium, the higher seems to be the risk for the system of getting moved back (Schneider and Kay 1994). The more time has been used for complexification, the higher is the risk of being seriously hit by disturbance, and the longer the elements of the system have increased their mutual connectedness, the stronger is the mutual interdependency and the total system's brittleness (Holling 1986). In general it can be concluded that the ability to adapt after changes of the constraints may be decreased when a high degree of maturity is attained. The respective system features have been arranged in Table 2.

In Fig. 2, the sequence of ecosystem states has been illustrated as a function of the system's internal connectedness and the stored exergy, in parallel to the adaptive cycle from the resilience alliance (Gundersson and Holling 2002). Starting with the exploitation function, there is a slow development. The trajectory demonstrates a steady increase in mutual interactions as well as an increase in exergy stored. As

Table 2 Compilation of basic phases in ecosystem development and some summarized ecosystem features, after Holling (1986) and Joergensen et al. (2007)

Developmental stage	Productivity	Connectivity and self-regulation	Nutrient loss	Flexibility
Start	Low	Low	High	High
Fast growth	High	Low	High	High
Fast development	Decreasing	Increasing	Decreasing	Decreasing
Maturity	Low	High	Low	Low
Breakdown	Low	Low	High	High
Reorganization	Low	Low	High	High

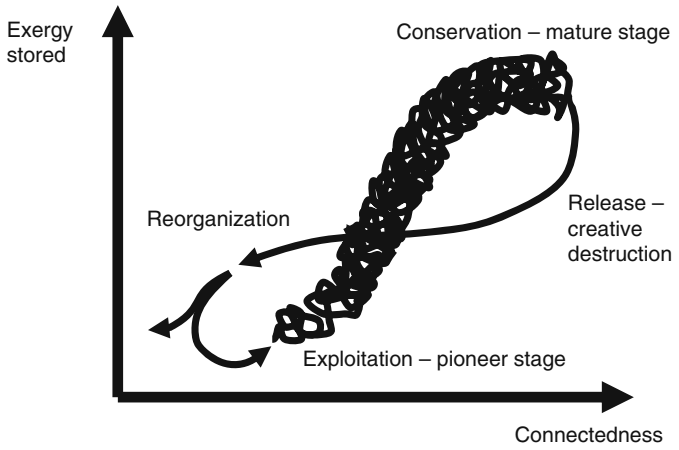

Fig. 2 Ecosystem trajectories of connectedness and stored exergy, after Holling (1986) and Joergensen et al. (2007)

has been described above, this energetic fraction can be distinguished into a material fraction and the specific exergy which refers to a complexification of the system's structure. In spite of several variabilities (e.g., annual cycles), the long term development shows a steady increase up to the mature state. Here the maximum connectivity can be found, which on the one hand is a product of the system's orientation, but which also is correlated with the risk of missing adaptability that has been nominated as overconnectedness by some authors. After the fast releasing event, the short-term conditions determine the further trajectory of the system. It might turn into a similar trajectory or find a very different pathway.

Throughout the introduction of disturbance it was mentioned that destruction and decay take place on multiple scales. Now we can extrapolate this fact to the "adaptive cycle" of Fig. 2: All ecosystem components potentially develop with the described phases, and many ecological processes support these dynamics. Observing this context from a hierarchical viewpoint, four different elements with typical frequencies and amplitudes may be distinguished:

- The dynamics of constraints
- The inherent dynamics of the focal variable
- The dynamics of the underlying fast processes
- The dynamics of the disturbance itself

The performance of the ecosystem arises from the complex interactions of these components. And furthermore, the long-term development of a system itself can be viewed from a hierarchical viewpoint: Within certain thresholds the ecosystem (as a focal level of observation) can follow the changes of the constraints. Due to the internal "brittleness" there will be a breakdown, if the constraints change or if a disturbance appears. In this situation the "creative destruction" may enable the system to start a reset under the new prevailing conditions. If the respective state

variables are able to return to orientor dynamics, in the end disturbance can be understood as a part of ecosystem growth and development on a higher temporal scale; disturbance may even be extremely necessary to enable a conituation of the complexifying trajectory of the overall system. Disturbance thus can cause stress, it can support adaptation and it may create innovative structures. Of course, on the other hand many disturbances, i.e., the consequences of human interventions, can set the system back to a less complex state, with minor mutual adaptations and a reduced diversity, thereby reducing the options of future emergence of complexity. Therefore, it might be helpful to take a look at the prevailing definitions of the respective variables of ecosystem dynamics and to introduce the concepts of resilience and adaptability, which have been elaborated on the base of the "adaptive cycle" concept.

5 Elements of Ecosystem Change Can be Characterized

The recognition that permanent stability is an unattainable state for ecosystems has caused a long lasting debate about how the diverse changes that occur following a disturbance can be characterized. This chapter gives an overview about the numerous concepts that have been developed in order to characterize changes and responses to disturbances, focusing on the concepts of resilience, adaptability and vulnerability.

An ecosystem can respond to a disturbance in two opposite ways. Either it stays essentially the same without structural changes, or it collapses into a qualitatively different state. The first reaction is commonly assigned to a resilient system with adaptive capacity, the second one to a vulnerable system. But there is a variety of definitions that concretize these concepts, without an overall consensus on their meanings (Brand and Jax 2007). However, there seems to be an agreement that "vulnerability is not always a negative property" (Gallopin 2006) and "resilience is not always a good thing" (Walker et al. 2004).

The term resilience has been used in various disciplines (see Table 3). It was introduced to ecology by Holling (1973) who illustrated in his paper "resilience and

Table 3 Utilization of the term "resilience" in different scientific and applied disciplines

Physics	Material resilience: capacity of a material to absorb energy when it is deformed elastically (Joergensen et al. 2007)
Network theory	Network resilience: ability to provide an acceptable level of service in the face of various faults (Najjar and Gaudiot 1990)
Industry	Engineering resilience : create processes that are robust and flexible (Hollnagel et al. 2006)
Psychology	Resilience: capacity of people to cope with stress and catastrophe (Luthar et al. 2000)
Economics	Resilience: ability of a local economy to retain function, employment and prosperity in the face of perturbations (Farber 1995; Briguglio 2004)

stability of ecological systems" the existence of multi-stable states of ecosystems and thereby prepared an end of the perception of single equilibria and global stability. He states that "resilience determines the resistance of relationships within a system and is a measure of the ability of these systems to absorb changes of state variables, driving variables, and parameters, and still persist" whereas stability is "the ability of a system to return to an equilibrium state after a temporary disturbance". In his definition, resilience refers to a dynamic system far from equilibrium whilst stability refers to a system close to an equilibrium state (Gunderson, 2000). There are, however, other definitions of resilience that are similar to the stability concept and assume only one stable steady state, e.g., the definition of Pimm (1984) which calculates resilience as the return times of variables "towards their equilibrium following a perturbation". This conception has later been termed "engineering resilience" by Holling (1996).

More recent definitions of resilience presume the existence of multiple stability domains, but extend Holling's (1973) definition by including hierarchy, cross-scale interactions and the theory of complex adaptive systems, thereby intermingling the concepts of resilience and adaptability (e.g., in Gunderson and Holling 2002; Walker et al. 2004). These definitions are called "ecological resilience" in contrast to the "engineering resilience". An often cited definition of ecological resilience is that of Walker et al. (2004) who characterize resilience as "the capacity of a system to absorb disturbance and reorganize while undergoing change so as to still retain essentially the same function, structure, identity, and feedbacks". Hence, the focus in this definition is on the maintenance of functions and structure of an ecosystem. Walker et al. (2004) specify four aspects of resilience which they also use to characterize the "basin of attraction", that is the sum of all attractors that induce the system to tend toward an equilibrium state. Accordingly, the first aspect of resilience is latitude, or the maximum amount a system can be changed before losing its ability to recover. Assigned to the basin of attraction, this aspect refers to the width of the basin. Resistance, the second aspect, is defined as the difficulty of changing the system, or the depth of the basin of attraction. The third aspect, precariousness, is the proximity of the system to a threshold, and the last one, panarchy, refers to influences of systems of another hierarchical level.

The resilience-perception of Carpenter et al. (2001) builds a bridge to the concept of adaptability. They define resilience as being composed of the following three properties:

(1) the amount of change the system can undergo and still remain within the same domain of attraction,
(2) the degree to which the system is capable of self-organization, and
(3) the degree to which the system can build the capacity to learn and adapt.

The adaptive capacity that Gallopin (2006) equates with the term adaptability, therefore is a component of resilience sensu Carpenter et al. (2001), but the relation between adaptive capacity and resilience, the question of which one is an element of which, remains unclear "because of the diversity of views" (Gallopin 2006). For instance, the perception of Norberg and Cumming (2006) that adaptive processes

relate to the capacity to tolerate and deal with change and those of Smit and Wandel (2006) that adaptive capacity allows for continuous development, resemble the above mentioned definitions of resilience. Thus, it doesn't become clear if resilience includes adaptability or is a component of the latter.

Both concepts are by now used in various disciplines and frequently applied to social-ecological systems (e.g., in the journal Ecology and Society 11, 2006). Here, the distinction between resilience and adaptability is clearer, due to the definition of Walker et al. (2006) that determines adaptability to be the capacity of the actors in a system to influence resilience. Social-ecological systems are, however, not the subject of this article and in spite of the inherent commendable distinctions, the applicability as well as the precision of this definition may be observed with some doubts.

Now, how does the previously mentioned term "vulnerability" of an ecosystem fit into this mess of definitions? Just as the relationship between resilience and adaptability, the linkage of both concepts with vulnerability remains unclear. Smit and Wandel (2006) identify three elements of vulnerability, which are

(1) sensitivity,
(2) adaptive capacity, and
(3) exposure.

Hence, adaptability is seen as an element of vulnerability. Sensitivity, the first element of vulnerability, has been defined by Gallopin (2003) as the "amount of transformation of the system per unit of change in the disturbance" and exposure – the last element – is defined by Adger (2006) as an "attribute of the relationship between the system and the perturbation", therefore the exposure is not an internal property of the system, but a characteristic of the relationship between the system and the disturbance (Gallopin 2003). This is the reason why Gallopin (2003), unlike Smit and Wandel (2006) does not consider the exposure to be an element of vulnerability. Also the relationship between resilience and vulnerability is being discussed by several authors (Gallopin 2003, 2006; Walker et al. 2004; Young et al. 2006; Smit and Wandel 2006). There is an agreement, that both attributes are ambivalent (Walker et al. 2004) insofar as a resilient system is less vulnerable than a non-resilient one (Gallopin 2006). In other words, a system that reacts vulnerably to a disturbance changes its structure and functions, whereas a resilient system does not. The question emerges from this statement, if vulnerability is the opposite of resilience. Gallopin (2006) concludes that the ambivalent relationship between both does not imply symmetry, since resilience applies to the system's persistence in the considered domain of attraction while vulnerability refers to transformations that change the system fundamentally and may go beyond a single domain of attraction. According to Gallopin (2006) the opposite of vulnerability would rather be robustness, which refers to the capacity to maintain the structure despite perturbations.

It becomes obvious that the terms of resilience, adaptability and vulnerability are not clearly distinguished in existing definitions but are often used interchangeably (Gallopin 2006). They are used in a very wide extension and are affected by a conceptual vagueness (Brand and Jax 2007) that hinders their practical application. Brand and Jax (2007) also criticize the increasing degree of normativity in current

definition and group the existing definitions in descriptive, hybrid (descriptive-normative), and normative definitions.

Another problem of application is the missing definition of the terms identity, structure and functions, and basin of attraction, which should all be maintained in resilient systems. Carpenter et al. (2001) therefore ask "resilience of what to what"?, and Brand and Jax (2007) question if there are "any possibilities to estimate or measure the resilience of an ecosystem". But can a definition of resilience really be purely descriptive as Brand and Jax (2007) demand in order to be a suitable concept for application within ecological science? The missing specifications of "identity", "structure and functions" and "basin of attraction" are inevitable as they strongly vary from ecosystem to ecosystem, depending on the spatial and temporal scale (Walker et al. 2002; Redman and Kinzig 2003), the system's boundaries, the hierarchical level, the variables of interest and many other factors. In other words, resilience, adaptability and vulnerability are always observer-dependent and therefore are always normative.

However, there is a necessity to clearly distinguish the ideas of resilience and adaptability in order to enhance conceptual clarity. We try to do this by including the concept of ecological integrity into the debate: *A system has a high adaptability if the sum of all disturbances and changes in the attractor domains do not reduce the system's degree of self-organization.*

In this definition, adaptability of an ecosystem refers to long-term changes and thus to a higher hierarchical level than resilience. Hence, the important property of a system with adaptive capacity is a trajectory that follows orientor dynamics, which is a general feature of ecosystem dynamics, but also a human target and therefore can become a normative property. In contrast, *resilience refers to the ability of a system to reorganize after a disturbance and remain in the previous basin of attraction*. Resilience does not correspond to a long-term trajectory of the attractor, but to the reaction after a single disturbance, excluding the dynamics on higher hierarchical levels.

6 Case Studies About Ecosystem Integrity and Resilience

To demonstrate the proposed interrelationship between integrity and the measures for ecosystem dynamics, two case studies will be sketched. Both cases are related to the development of wetland ecosystems. To summarize their basic features, the stability checklist from Gigon and Grimm (1997) has been applied in Table 4.

6.1 Case Study Wetland Retrogression

The first case study demonstrates disturbance dynamics in the wetlands of the Bornhöved Lake District in Northern Germany. Here a holistic indicator system, which has been developed on the base of orientor theory (see Table 1) has been used

Table 4 An application of the stability check list after Gigon and Grimm (1997). Following the authors, the methodological questions asked in this list have to be taken into account referring to any characterization of ecosystem dynamics. The data have been originally presented in Müller et al. (2006) and Schrautzer et al. (2007)

	Wet grassland retrogression	Wet grassland resilience
Which level of organization?	Ecosystem	Ecosystem
Which spatial scale?	Bornhöved lake district	Northern Germany
Which temporal scale?	~30 years development 30 years simulations	~100 years development 60 years simulations
Which disturbance?	Eutrophication Drainage	Deforestation Eutrophication Drainage Mowing Grazing
Which indicator in the ecosystem?	Ecosystem integrity indicator set	Ecosystem integrity indicator set (Table 1)
Which stability feature?	Change of single indicator values	Change of single indicator values
Which reference points or dynamics?	Mostly disturbed system	Maximum performance of single indicators
Which method of quantification?	Measurements and model runs	Measurements and model runs
Which normative aspects are taken into account?	Ecosystem approach	Ecosystem approach of the CBD

to demonstrate some steps of wetland retrogression as provoked by eutrophication and drainage. A comprehensive description of the study can be found in Müller et al. (2006).

On the base of field measurement, mappings and ecosystem classifications different wetland types have been analysed with the computer based digital landscape analysis system DILAMO (Reiche 1996) and the modelling system WASMOD-STOMOD (Reiche 1996) which was used to simulate the dynamics of water budgets, nutrient and carbon fluxes based on a 30 years series of daily data for meteorological and hydrological forcing functions. The model outputs were validated by measured data in representatives of these systems. The model results were extended to include data sets concerning ecosystem features by the integrity variables from Table 1.

The wet grasslands of the Bornhöved Lake District are exposed to the following management measures: drainage, fertilisation, grazing, and mowing in a steep gradient of ecosystem disturbances. The systems have been classified due to these external input regimes, and in Fig. 3 the sequential consequences of these disturbances can be seen in a synoptic manner: As the farmer's targets, the improvement of production and yield, are successfully fulfilled, the net primary production (NPP) is increasing by a factor of 10. Simultaneously the structural indicator (no. of plant species) is decreasing enormously throughout the retrogression. Also the efficiency

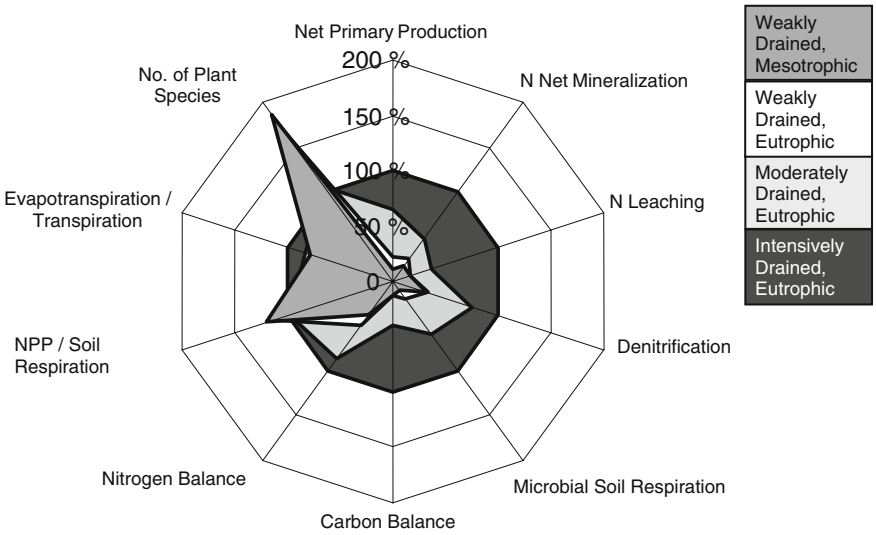

Fig. 3 Amoeba diagram representing indicators of ecosystem integrity during a retrogressive development in different wet grasslands of the Bornhöved Lake District. The values of the most degraded state have been set as 100%. The degradation has been mainly caused by different degrees of eutrophication and drainage. Starting with the initial state ("weakly drained, mesotrophic"), drainage and eutrophication of the wet grassland ecosystems affect irreversible changes up to the reference state "intensively drained, eutrophic". During that development ecosystem structures are reduced, energy and matter efficiencies decrease, and the originally sink function of the ecosystem turns into a source for nitrogen and carbon compounds. After Müller et al. (2006)

measures (NPP/soil respiration) are going down, and the biotic water flows (transpiration/evapotranspiration) get smaller. On the other hand, the development of the nitrogen and carbon balances demonstrates that the system is turning from a sink function into a source, the storage capacity is being reduced, and the loss of carbon and nitrogen compounds (all indicators on the right side of the figure) is rising enormously. Due to these dynamics we can state that there has been an enormous decrease of ecosystem integrity. As many of the processes are irreversible, the capacity for future self-organization is reduced to a very small degree.

If we now take a critical look at the resilience definitions and apply them on a small scale, it turns out that this variable behaves opposite to integrity: If this index is low, the potential for recovery is high, and also the necessary return time would be small. Thus, the higher the degradation of an ecosystem, the higher will be its resilience. Of course this context raises a further question for the resilience of the whole successional series: Is there a way back to a more integer ecosystem state after such a sequence of degradations?

6.2 Case Study Restoration Potential

In this second case study the scope of successional stages has been enhanced. The series leads from nature-near alder breaks to degraded wet pastures in 5 steps. The

depiction in Fig. 4 is based on an intensive data sampling in wetlands all over Northern Germany, the interpretation of several ecological and botanical time series and on the thorough interpretation of successional studies in Northern German wetlands (see Schrautzer 2004). By model applications it was possible to quantify the indicator set of the retrogressional study. Therefore, the results which are illustrated in Fig. 4 allow for a wide range of interpretations, which will be restricted here to the basic elements of ecosystem integrity.

The amoeba diagrams in Fig. 4, which were used to summarize the outcomes of the investigation, once more represent the selected indicators that have been shown in Table 1. Their position can be found on the lower left side of the Figure. The reference values (100%) were chosen from the whole data set. They represent the highest value found for the respective variables. Negative values can be found with reference to the budgets of carbon and nitrogen, which in that cases function as landscape sinks. On the left hand sides of the amoeba diagrams those indicators are arranged which represent a high level of integrity with high values, while the parameters on the right sides of the amoeba demonstrate a loss of nutrients and gradients if their values are high. Consequently a qualitative estimation of the systems' integrity can be derived from the form of the amoeba value areas.

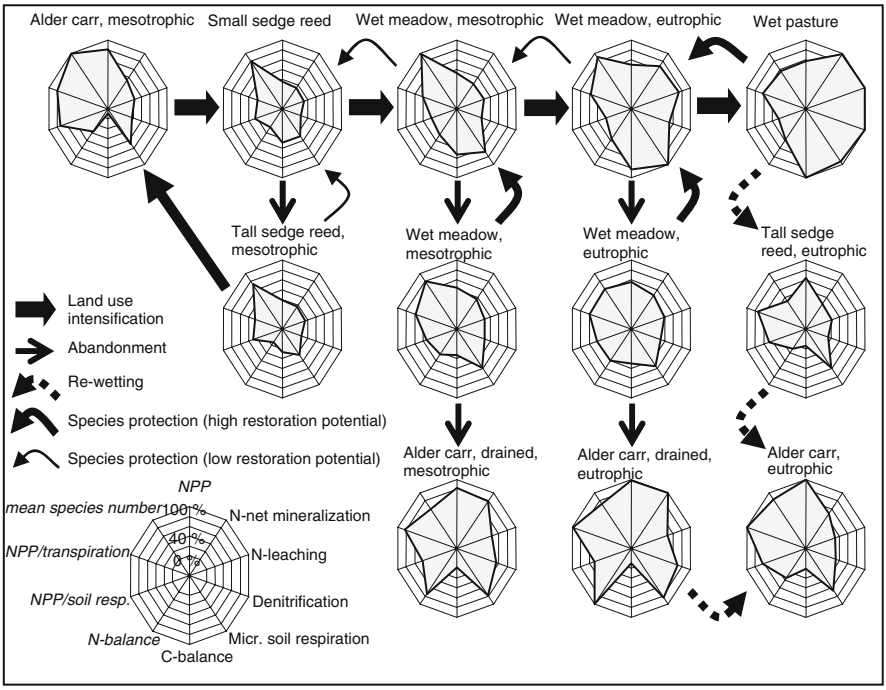

Fig. 4 Amoeba diagrams of sequential indicator values for ecosystem integrity during a retrogressive succession following land use intensification, and progressive successions following abandonment. The highest parameter values for all successional stages were set as 100%. Negative values concerning C- and N-balances represent a sink function of the systems. After Schrautzer et al. (2007)

The sequence of retrogressional states starts in the upper row of the Figure with the mesotrophic alder carr. Moving to the right (symbolized by the bold arrows), land use is intensified by deforestation, drainage, eutrophication, grazing and mowing, and in the end the state of a wet pasture is reached, which can be characterized by a low diversity, decreased ecological efficiencies and by a loss of nutrients, e.g., by high mineralization rates, high nitrate outputs by leaching, and high nitrogen loss through denitrification. Consequently these systems turn the landscape function of a nutrient and carbon sink into a source.

In contrast to the preceding case study, here also the question whether the changes are reversible or not has been posed to the data, and four different landscape management measures have been distinguished; abandonment, species protection measures with two different restoration potentials, and re-wetting. In Fig. 4 these modifications are symbolized by different arrows. The small straight arrows represent the consequences of abandonment, which could be perceived as a natural recovery measure. Thus, if a system is able to return to the starting formation, it would behave in a resilient manner. Studying the successional series, it becomes obvious that only in one case (alder carr → small sedge reed → tall sedge reed → alder carr) such dynamics are possible. Only in this case the original ecosystem can be restored. All other series of abandoned ecosystems lead to different ecosystem types: drained and mostly eutrophicated alder forests, which are depicted in the lower row of the figure. At a first glance, these systems look very similar to the initial state, but concerning the diversity, the metabolic efficiency, the N balances and the nitrogen loss, the indicator values indeed differ a lot from the starting state. As the systems are defined functionally, there is a decreased resilience in comparison with the only returning pathway. This context may be underlined if we use Pimm's measure of resilience: The successions which lead to the lower line of the figure need much longer times than the reversible one.

Additionally two results can be derived from the study: The species protection measures mostly enable a one-step-return only and they do not re-develop the initial stage. And even if extreme management activities like rewetting are carried out, the result will not be identical with the stating point, although the process takes a long time.

With reference to the discussion of ecosystem dynamics and resilience, this case study documents the high normative loading of all potential interpretations: The evaluation is a result of the observer's objectives, her or his indicators and the applied thresholds of reversibility, which are also defined by the observer:

If we return to the idea of attractor states to characterize resilience, the question is which width of the attractor basin is defined before a change of the system's regime is accepted. Are wetlands with alders healthy alder carrs? Or do we have to take into account their functionality as well?

The next normative point is the scale of observation: If we concentrate on the recovery potential of single states, the results might be satisfactory, because some of the species protection measures in fact are successful. If we choose the total retrogressional series, the result will be frustrating. Furthermore, the recovery potential of course also is a function of the selected time scales. And additionally also the

potential investment into the treated wetlands is a parameter with important consequences for the evaluation.

The observer might also concentrate on a smaller amount of variables. The targets, results and the restorer's satisfaction might be very different if he or she wants to improve biodiversity or climate protection or ecosystem integrity.

Regrettably, one further consequence can be drawn: Degraded and "simple" ecosystems provide a much higher resilience than healthy ones (e.g., it might not be hard to restore a simple pioneer system but much effort and patience will be necessary for the restoration of a complex ecological entity), they have a much higher buffer capacity than complex systems, but the latter represent a high degree of ecosystem integrity.

In spite of the multitude of consequences and interpretations, the two wetland case studies hopefully could show that it is possible to indicate ecosystem integrity, thus to characterize the state of an ecosystem based on a concept of complexifying dynamics. And additionally it is also possible to depict resilience or adaptability on the base of ecosystem data and indicator sets.

7 Concluding Remarks

In this paper we have tried to discuss and illustrate some items of the actual debate about ecosystem dynamics. Returning to the initial objectives and questions of the article, some assertions can be formulated in the end. They are stated under the impression that stability of ecosystems is an accepted illusion, while dynamic development is a fact. Within this final discussion some aspects of this idea can be summarized, taking the initial questions of the paper as a guideline:

- *Is there a general tendency in undisturbed ecosystem development?*

In undisturbed ecosystem development there are different tendencies if we choose different scales of observation. Therefore, the consideration of constructive or destructive developmental phases is a function of the observer's objectives. For example, evolutionary investigations have to cope with all phases of the adaptive cycle, while the analysis of successions might lead to a preference of complexifying phases, and disturbance ecology will often concentrate on the consequences of the "release stage". Ecosystem theory has put emphasis on the phase of complexification, because nature-near ecosystems operate in a complexifying developmental trail for the longest time of their existence. Furthermore, a break down on a lower level of observation might be part of an adaptation process on a higher level that optimizes orientor values at a long-term scale.

Within these boundary conditions, general developmental trends of ecosystem features can be detected, understood and forecasted. When an ecosystem approaches the bifurcation point of Holling's release phase, its future fate can not be foreseen.

- *How can this trend be indicated?*

During the complexifying phase, orientors are optimized within the limits of the respective site conditions. These variables can be used as indicators to illuminate the state of the system. To describe an ecosystem in a holistic way, structural and functional attributes should be included, water, matter and energy budgets should be characterized and inputs, outputs, internal flows and efficiencies should be taken into account. These functional requirements are preferable demands for environmental applications: If an index with a higher degree of aggregation is selected (e.g., exergy storage, ascendancy or emergy), the overall results will also be visible, but it is hard to understand the functionality of disturbances and to derive respective management measures.

• *Which is the role of disturbance throughout ecosystem dynamics?*

In disturbed phases of development the orientor system and the ecosystem's hierarchical structure is broken. Functional modifications will lead to structural changes, and the system will approach a new state. If this state is situated within the old domain of attraction the system behaves resilient. If its position is outside the latitude (the width of the domain of attraction), a new steady state will be approached. The magnitude of the disturbance will influence the longer-term consequences and resilience as well as adaptability. Destruction and decay can be understood as "normal" components of natural development, and consequently – in an attitude towards nature conservation which still lacks a theory-based discussion – they could also be seen as basic processes supporting long-term innovation and adaptation.

• *What are the recent comprehensions within the resilience and adaptability concept?*

In the literature several features characterizing ecosystem dynamics can be found: e.g., stability, resistance, resilience, buffer capacity, elasticity, adaptability on the "stable" side and vulnerability, fragility, transformability to describe the conditions of the receptors. As a result of the preceding discussion of these terms, two focal comprehensions are proposed as guiding indicators of ecosystem dynamics: A system has a high adaptability if the sum of all disturbances and changes in the attractor domains do not reduce the system's degree of self-organization.

Resilience refers to the ability of a system to reorganize after a disturbance and remain in the previous basin of attraction. The difference between these concepts arises from the scale of observation: While adaptability considers longer developmental durations and attractor dynamics, resilience should be used for short time investigations, when the attractor composition does not change noticeably.

Resilience and integrity can not be linked directly. The respective investigations demonstrate that the higher an ecosystem's complexity and integrity is, the smaller is its resilience. Therefore the two conceptions seem to be counter moving; simple ecosystems provide a high resilience while systems with a high integrity have a low resilience. Consequently, they are more vulnerable with reference to disturbances. In contrast, adaptability in this comprehension can be directly related to integrity because orientor dynamics are taken into account by the inclusion of self-organized processes.

- *Is there a potential to quantify these attributes?*

The case studies have shown that an indication of the dynamic variables is possible. In the presented cases the resilience of the integrity variables has been characterized. These studies can also be used to demonstrate the demands and challenges for future work.

A focal problem arises from the *normative loading* of the resilience and adaptability concepts. This problem is not new. For example Picket and White (1985) have used a structural approach to define disturbance as *any relatively discrete event in space and time that disrupts ecosystem, community, or population structure and changes resources, substrates, or the physical environment*. The arising question referring to this definition is: how to denote the "normal state" of an ecosystem as a reference state with regard to disturbed states (Jentsch and White 2001)? Applying this problem to the resilience concept, similar questions would be:

- What is a regime of system states?
- Which are the thresholds of such a regime?
- How to define a regime shift?
- What is the function of the ecosystem?
- Which are the thresholds of the system's identity?

Actually there are no objective criteria to answer those queries. The observer defines his or her system, and the change of a system's resilience is one of the points which are defined case by case, lacking a scientific generalisation.

Finally, the question arises, which of the two above defined concepts will have a higher significance in the future. The difference is that resilience still considers the return to a former state, may be with a higher tolerance, than the stability concept, referring to a greater state space. Adaptability, following the proposed definition describes the ability of an ecosystem to return to an orientor trajectory. Thus, this concept is much more related to long-term developments on the one hand, and to the dynamic nature of the system's constraints on the other. Therefore, it might be very helpful in the context of the multiple changes we are facing. For example global climate change will modify the ecosystemic constraints drastically, making it impossible to return to the former situation in several instances. Thus, a search for the optimal conditions for adaptability might be much more helpful than resisting on a stability related resilient development, which might support systems that could become extremely dependent on external, protecting inputs of energy and work. The respective discussions and the development of adapted targets for ecosystem dynamics is an important task. It will be helpful to consider some ecosystem theoretical arguments within this process to find a balanced societal consensus.

Acknowledgments The data presented in the two case studies have been measured and worked out by numerous colleagues from the Bornhöved-Project. We want to thank them for their collaboration.

References

Adger WN (2006) Vulnerability. Global Environmental Change 16: 268–281

Barkmann J (2002) Modellierung und Indikation nachhaltiger Landschaftsentwicklung – Beiträge zu den Grundlagen angewandter Ökosystemforschung. Ph.D. thesis, University of Kiel

Barkmann J, Baumann R, Meyer U, Müller F, Windhorst W (2001) Ökologische Integrität: Risikovorsorge im Nachhaltigen Landschaftsmanagement. Gaia 10(2): 97–108

Begon M, Harper JL, Townsend CR (1990) Ecology: individuals, populations and communities. Blackwell Publ., Boston, Oxford, London

Bossel H (1998) Ecological orientors: emergence of basic orientors in evolutionary self-organization. In: Müller F and Leupelt M (1998) Eco targets, goal functions and orientors. Springer, Berlin Heidelberg New York, pp 19–33

Bossel H (2000) Sustainability: application of systems theoretical aspects to societal development. In: Jørgensen SE, Müller F (eds.) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton London New York Washington DC, pp 519–536

Brand FS, Jax K (2007) Focusing the meaning(s) of resilience: resilience as a descriptive concept and a boundary object. Ecology and Society 12(1): 23

Briguglio L (2004) Economic vulnerability and resilience: concepts and measurements. In: Briguglio L, Kisanga EJ (eds.) Economic vulnerability and resilience of small states, Formatek Ltd, Malta

Carpenter S, Walker B, Anderies JM, Abel N (2001) From methaphor to measurement: resilience of what to what? Ecosystems 4: 765–781

Crabbé P, Holland A, Ryszkowski L, Westra L (2000) Implementing ecological integrity. Kluwer Academic Publishers, Dordrecht

Elton CS (1958) The ecology of invasion by animals and plants. Methuen Press, London

Farber S (1995) Economic resilience and economic policy. Ecological Economics 15(2): 105–107

Fath B, Patten BC (1998) Network orientors: a utility goal function based on network synergism. In: Müller F, Leupelt M (1998) Eco targets, goal functions and orientors. Springer, Berlin, Heidelberg, New York, pp 161–176

Gallopin GC (2003) A systematic synthesis of the relations between vulnerability, hazard, exposure and impact, aimed at policy identification. In: Economic Commission for Latin American and the Caribbean: Handbook for Estimating the Socio-Economic and Environmental Effects of Disasters. ECLAC, LC/MEX/G.S., Mexico, D.F., pp 2–5

Gallopin GC (2006) Linkages between vulnerability, resilience, and adaptive capacity. Global Environmental Change 16: 293–303

Gigon A, Grimm V (1997) Stabilitätskonzepte in der Ökologie: Typologie und Checkliste für die Anwendung. In: Fränzle O, Müller F, Schröder W (eds.) Handbuch der Umweltwissenschaften. Ecomed, Landsberg, III-2.3 pp 1–19

Grimm V, Wissel C (1997) Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. Oecologia 109: 323–334

Gunderson LH (2000) Ecological resilience – in theory and application. Annual Review of Ecology and Systematics 31: 425–439

Gundersson LH, Holling CS (eds.) (2002) Panarchy. Island Press, Washington Covelo London

Haskell BD, Norton BG, Costanza R (1993) Introduction: what is ecosystem health and why should we worry about it? In: Costanza R, Norton BG, Haskell BD (eds.) Ecosystem Health. Island Press, Washington, Covelo London, pp 3–22

Holling CS (1973) Resilience and stability of ecological systems. Annual Review of Ecological Systems 4: 1–23

Holling CS (1986) The resilience of terrestrial ecosystems: local surprise and global change. In: Clark WC and Munn RE (eds.) Sustainable development of the biosphere. Cambridge University Press, Cambridge, pp 292–320

Holling CS (1996) Engineering resilience vs. ecological resilience. In: PC Schulze (eds.) Engineering within ecological constraints. National Academies Press, Washington, pp 31–43

Holling CS, Gunderson L, Ludwig D (2002) In quest of a theory of adaptive change. In: Gunderson LH and CS Holling (eds.) Panarchy: understanding transformations in human and natural systems. Island Press, Washington, D.C., pp 3–24

Hollnagel E, Woods, DD, Leveson NG (2006) Resilience engineering: Concepts and precepts. Ashgate Publishing Co., Aldershot

Janssen MA (2007) An update on the scholarly networks on resilience, vulnerability, and adaptation within the human dimensions of global environmental change. Ecology and Society 12(2):9

Joergensen SE (2000) The tentative fourths law of thermodynamics. In: Joergensen SE, Müller F (eds.) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton, London, New York, Washington, DC, pp 161–176

Joergensen SE, Müller F (eds.) (2000) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton

Joergensen SE, Straskraba M (2000) Ecosystems as cybernetic systems. In: Joergensen SE, Müller F (eds.) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton, London, New York, Washington, DC, pp 249–264

Joergensen SE, Fath B, Bastianoni S, Marques J, Müller F, Nielsen SN, Patten B, Tiezzi E, Ulanowicz R (2007) A new ecology – the systems perspective. Elsevier, Amsterdam

Kay JJ (1993) On the nature of ecological integrity: some closing comments. In: Woodley S, Kay J, Francis G (eds.) Ecological integrity and the management of ecosystems. University of Waterloo and Canadian Park Service, Ottawa

Kay JJ (2000) Ecosystems as self-organised holarchic open systems: Narratives and the second law of thermodynamics. In: Joergensen SE, Müller F (eds.) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton, London, New York, Washington, DC, pp 135–160

Leopold A (1944, 1991) Conservation: In whole or in part? In: Flader S, Callicott JB (1991) The river of the mother of God and other essays by Aldo Leopold. University of Wisconsin Press, Madison, pp 310–319

Luthar SS, Cicchetti D, Becker B (2000) The construct of resilience: A critical evaluation and guidelines for future work. Child Development 71(3): 543–562

May, RM (1972) Will large complex systems be stable? Nature 238: 413–414

Millennium Ecosystem Assessment (2005) Ecosystems and human well-being. A framework for assessment. p.109, Island Press, Washington.

Müller F (1992) Hierarchical approaches to ecosystem theory. In: Ecological Modelling 63: 215–242

Müller F (1998) Gradients, potentials and flows in ecological systems. Ecological Modelling 108: 3–21

Müller F (2004) Ecosystem indicators for the integrated management of landscape health and integrity. In: Joergensen SE, Costanza R, Fu-Liu X (eds.) Ecological indicators for assessment of ecosystem health. Taylor & Francis Group, Boca Raton, pp 277–303

Müller F (2005) Indicating ecosystem and landscape organisation. Ecological Indicators 5(4): 280–294

Müller F, Burkhard B (2007) An ecosystem based framework to link landscape structures, functions and services. In: Mander Ü, Wiggering H, Helming K (eds.) Multifunctional land use – meeting future demands for landscape goods and services. Springer, Berlin, pp 37–64

Müller F, Leupelt M (1998) Eco targets, goal functions and orientors. Springer, Berlin Heidelberg New York

Müller F, Schrautzer J, Reiche EW, Rinker A (2006) Ecosystem based indicators in retrogressive successions of an agricultural landscape. Ecological Indicators 6(1): 63–82

Najjar W, Gaudiot JL (1990) Network resilience – a measure of network fault tolreance. IEEE Transactions on Computers 39(2): 174–181

Norberg J, Cumming GS (2006) Complexity theory for a sustainable future. Columbia University Press, New York

Odum EP (1969) The strategy of ecosystem development. Science 164: 262–270

Picket STA, White PS (1985) Natural disturbance and patch dynamics: An introduction. In: Picket STA, White PS (eds.) The ecology of disturbance and patch dynamics. Academic Press, Orlando, pp 3–13

Pimm SL (1984) The complexity and stability of ecosystems. Nature 307: 321–326

Prigogine I (1980) From being to becoming: Time and complexity in the physical sciences. Freeman, San Francisco

Rapport DJ (1989) What constitutes ecosystem health? Perspectives in Biology and Medicine 33(1): 120–132

Rapport DJ, Moll R (2000) Applications of ecosystem theory and modelling to assess ecosystem health. In: Joergensen SE, Müller F (eds.) Handbook of ecosystem theories and management. Lewis Publishers, Boca Raton, London, New York, Washington, DC, pp 487–496

Redman CL, Kinzig AP (2003) Resilience of past landscapes: resilience theory, society, and the Longue Durée. Conservation Ecology 7(1): 14

Reiche EW (1996) WASMOD: Ein modellsystem zur gebietsbezogenen simulation von Wasser- und Stoffflüssen. EcoSys 4: 143–163

Scheffer M and SR Carpenter (2003) Catastrophic regime shifts in ecosystems: linking theory to observation. Trends in Ecology and Evolution 12: 648–656

Schneider ED, Kay JJ (1994) Life as a manifestation of the second law of thermodynamics. Mathematical and Computer Modelling 19: 25–48

Schrautzer J (2004) Niedermoore Schleswig-Holsteins: Charakterisierung und Beurteilung ihrer Funktion im Landschaftshaushalt. Mitteilungen der Arbeitsgemeinschaft Geobotanik in Schleswig-Holstein und Hamburg 63

Schrautzer J, Rinker A, Jensen K, Müller F, Schwartze P, Dierßen K (2007) Succession and restoration of drained fens: perspectives from Northwestern Europe. In: Walker L, Hobbs RJ, Walker J (eds.) Linking restoration and ecological succession. Springer, Berlin, Heidelberg, New York

Smit B, Wandel J (2006) Adaptation, adaptive capacity and vulnerability. Global Environmental Change 16: 282–292

Tansley AG (1935) The use and misuse of vegetational terms and concepts. Ecology 16: 284–307

Walker B, Carpenter S, Anderies J (2002) Resilience management in social-ecological systems: a working hypothesis for a participatory approach. Conservation Ecology 6(1): 14

Walker B, Holling CS, Carpenter SR, Kinzig A (2004) Resilience, adaptability and transformability in social-ecological systems. Ecology and Society 9(2): 5

Walker B, Gunderson L, Kinzig A, Folke C, Carpenter S, Schulz L (2006) A handful of heuristics and some propositions for understanding resilience in Social-Ecological Systems. Ecology and Society 11(1): 13

Walker LR, del Moral R. (2003) Primary succession and ecosystem rehabilitation. Cambridge University Press, Cambridge

Walker LR, Walker J, Hobbs R. (eds.) (2007) Linking restoration and ecological succession. Springer.

Westra L, Lemons J (eds.) (1995) Perspectives on ecological integrity. Kluwer Academic Publishers, Dordrecht

White PS, Jentsch A (2001) The search for generality in studies of disturbance and ecosystem dynamics. Progress in Botany 62: 95–113

Woodley S, Kay JJ, Francis G (1993) Ecological integrity and the management of ecosystems. St. Lucie Press, Ottawa

Young OR, Berkhout F, Gallopin G, Janssen MA, Ostrom E, van der Leeuw S (2006) The globalization of socio-ecological systems: an agenda for scientific research. Global Environmental Change 16: 304–316

Analyzing Spatio-Temporal Hydrological Processes and Related Gradients to Improve Hydrological Modeling in High Mountains

Ole Rößler and Jörg Löffler

Abstract Mountain hydrology suffers from insufficient data availability and partly coarse process understanding. But the improvement of our process knowledge is the key to manage the mountain water resources in present and future. Hydrological models like WaSiM-ETH are used to simulate the water balance in areas were less data are available. The validation and estimation of (un-)certainties of the model are essential to assess the accuracy and applicability of model results. In this project extensive hydrological data monitoring serve (a) to improve process understanding, (b) to enable multi-validation of the model, and (c) to estimate sensitive hydrological parameter for hydrological modeling. The latter can support the efficient monitoring of unmeasured catchments in future case studies. In this paper, we present the methodological concept, the monitoring program and preliminary results. Soil moisture was analyzed for 2007 in reliance on elevation and exposition and found to be highly variable at all investigated plots. But, a superior dynamic was found that can be characterized by a drying period after snow melt, a rise due to high precipitation amounts, and a following second drying period in autumn. The superior dynamic of soil moisture has to be interpreted as a direct function of temporal precipitation distribution. Measurements show that low soil moistures (9% vol) during dry periods occur at lower elevations (up to 1800 m) and lead to drought stress for plants. In contrast, soil moisture at higher altitudes (1800–2700 m) never attains a critical level for plants. Moreover, a slight increase of soil moisture with altitude can be derived. The dynamic of storages like soil moisture, snow and glaciers play a key role to characterize the alpine water balance. The validation of these storages at the same time is therefore very important to assess the accuracy and certainty of the model results and has not yet been conducted for WaSiM-ETH. Monitored microscale processes and mesoscale gradients are to be applied to a macroscalic catchment using the multi-validated model to ensure precise estimation of mountain water balance.

O. Rößler (✉)
Department of Geography, University of Bonn, Meckenheimer Allee 166, D-53115 Bonn, Germany
e-mail: o.roessler@giub.uni-bonn.de

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0_15,

Finally, the most crucial parameter and processes for hydrological modeling with WaSiM-ETH can be derived to improve the efficiency of future monitoring.

Keywords Mountain hydrology · Monitoring · Gradients · Soil moisture · Modeling · WaSiM-ETH

1 Introduction

Mountains are of particular importance within in the global water cycle and are regarded as essential water towers for the lowlands (Mountain Agenda 1998). This superior relevance is contrasted with the relative low level of knowledge about mountain hydrology (Viviroli et al. 2003). One problem is the lack of data in these remote areas. Furthermore, in comparison with lowlands more data are needed because mountain areas are known to "have the most intimate and complex interactions and variability on short space and time scales" (Roots & Glen 1982). Moreover, snow and glaciers occur as temporal storages of water and their distribution and dynamic. This has to be taken into consideration when analyzing the water balance of alpine catchments (Gurtz et al. 1999, Zappa et al. 2003).

Not least due to this complexity and the lack of data, hydrological processes and their gradients are not fully understood in high mountains (de Jong et al. 2005). But the improvement of our process knowledge is key to manage the water resources of the mountains in future (Winiger et al. 2005), especially regarding the challenges of a changing climate. To face these challenges more detailed studies providing more precise and accurate data are needed to improve our understanding of processes and to implement the gained knowledge in hydrological models.

One very prominent state-of-the-art model for mountain areas is the distributive, physically based hydrological model WaSiM-ETH that has already been successfully applied to mountain areas (Verbunt et al. 2003). The model was used to forecast floods (Jasper et al. 2002), and effects of land use change (Bronstert et al. 2007) as well as future hydrological scenarios under climate change (Jasper et al. 2004). The validation and the estimation of the (un-)certainties of the model results are crucial. Most models are validated using observed discharge data, but to assess the accuracy of the model in terms of other processes like evapotranspiration as well as the spatio-temporal dynamic of storages like snow, glaciers, and soil moisture, additional data are required to validate the model results.

Hydrological parameters like soil moisture and snow cover are keys for the characterization and the formation of mountain ecosystems (Löffler 2007). The determination of the spatial distribution of hydrological parameters are therefore crucial to characterize present and future ecosystem functioning. Not least to this, the validation of the hydrological model should not only focus on the runoff but requires multi-validation of further parameters (Verbunt et al. 2003). Hydrological models are used to simulate the water balance in areas were less data are available. The monitoring of all parameters needed is often laborious and cost intensive. Especially in mountain areas, the input parameters are uncertain but need careful consideration

in hydrological models (Verbunt et al. 2003). To design an effective monitoring program it is essential to detect the influence of single parameters, which have to be measured in future studies to receive the optimal model results. The analysis of a single but coupled process within a complex system like the hydrological system requires the use of process based models like WaSiM-ETH (Šimůnek 2005). Furthermore, using process based models sensitivity analyses can be conducted to elaborate the influence of processes and parameters (Šimůnek 2005).

Due to these current challenges of mountain hydrology the aims of this project are derived:

The aims of this project are:

- the characterization of spatio-temporal dynamics of hydrological processes and related gradients in high mountain areas,
- the estimation of the accuracy of WaSiM-ETH in terms of different hydrological parameters, and
- the estimation of sensitive processes and gradients to be analyzed for characterizing and modeling mountain hydrology.

In this paper, we present the methodological concept and the used semi-empirical methods. Preliminary results focusing on the soil moisture dynamics are exposed before an outlook is given.

2 Methodological Concept

To characterize the spatio-temporal dynamics of hydrological processes and related gradients in high mountain areas as well as to improve hydrological models, a methodological concept was developed. Thereby, gradients are regarded as the continuous spatial change of the hydrological processes in reliance to underlying environmental parameter. The methodological concept enables the characterization of the hydrological processes at three different spatial scales, the derivation of hydrological gradients, the estimation of the accuracy of the used model, and the extraction of the sensitive processes and related gradients for mountain hydrological modeling.

At the microscale, hydrological parameters are measured at single points focusing on the temporal gradients of processes. The determination of the magnitude of processes and storages are keys to this scale. At the mesoscale, the gradients of the hydrological processes are studied as to their spatio-temporal change with environmental parameters like altitude and land use that are assumed to have the most profound effect on the processes. At the macroscale the spatial characterization of the processes and derived gradients are studied within a catchment. These processes and related gradients cannot be directly measured at the macroscale, but need to be simulated using a hydrological model. Therefore, two different levels exist: an empirical level where hydrological processes are measured and a model level where these processes are transferred to the broad scale of the entire catchment.

The two levels are connected in a double manner: the microscale and the mesoscale serve as a resource for input parameters and as the validation reference to estimate the validity and accuracy of the model quality. On the one hand, this procedure enables the exact and manifold validation of the model results. On the other hand, investigated processes and related gradients can be evaluated as to their ability to improve the model significantly using sensitivity analyses. The outcome of this procedure is the identification of the most sensitive processes and an estimation of the certainty and accuracy of the model itself. Figure 1 illustrates this concept. This iterative process provides an optimized model configuration to characterize the hydrological processes. The spatio-temporal distribution of the processes enables the derivation of hydrological gradients in reliance on environmental parameters like topography and land use.

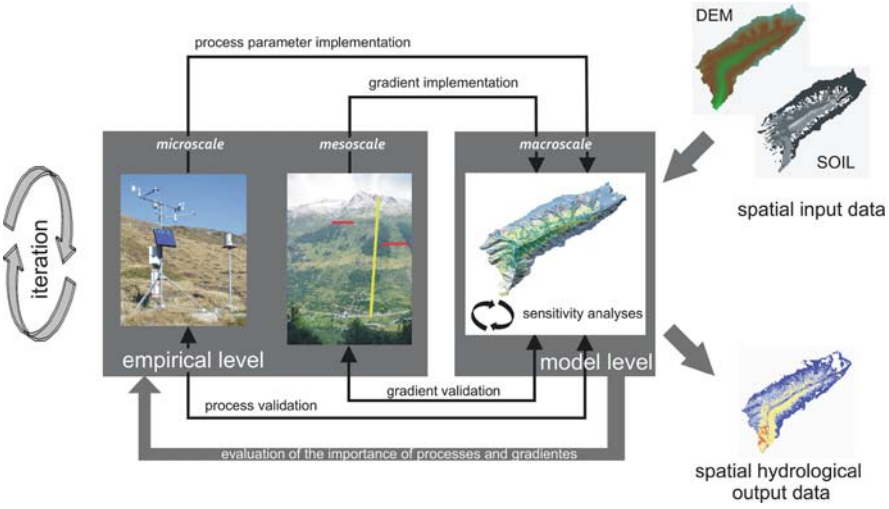

Fig. 1 The methodological research concept. Empirical studies at the microscale and at the mesoscale are implemented in the model, and serve as a resource for model validation. Iterations are done to improve the model and sensitivity analyses are calculated to assess the importance of the used processes and gradients form microscale and mesoscale, respectively. With the use of spatial data sets, hydrological processes are simulated for the entire catchment

3 Study Area and Monitoring Design

3.1 Empirical Level

To achieve the aims of the study, comprehensive data sets are needed. The alpine valley Lötschental (Switzerland, Fig. 2) is predestinated due to the extensive research work that has been conducted in this valley for years. Besides climatologic measurements, spatial data sets like vegetation (Hörsch 2001), relief (digital elevation

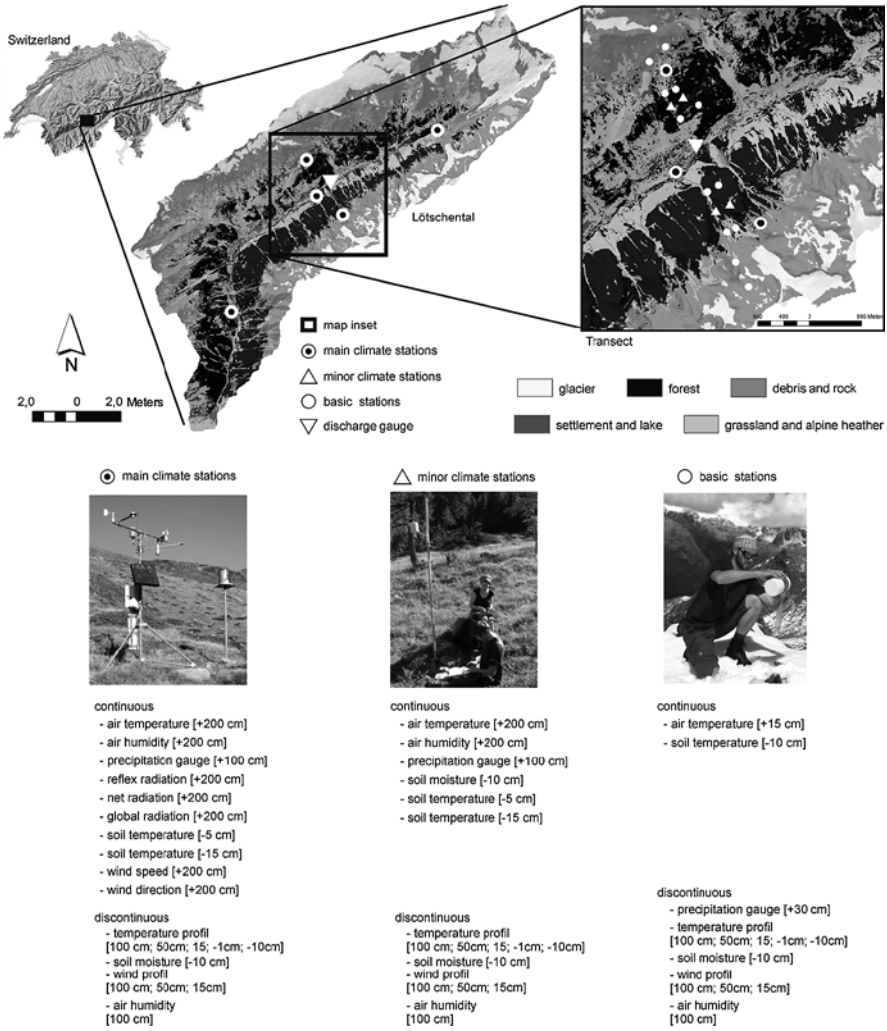

Fig. 2 Location of the alpine valley Lötschental in the Bernese Alps, its land cover and the research design. The major meteorological stations are distributed along the valley floor and on the two opposite slopes covering the total area. The focus of the measurements is laid on the transect in the middle of the valley stretching across all elevational belts from valley at 1400 m to the nival zone at 2700 m

model, 10 m resolution), and snow cover distribution of 2003 and 2004 (Schmidt 2007) are available. The Lötschental is a northern tributary valley of the Rhone Valley in the Swiss Alps. In this study the catchment of the main river in the valley (Lonza) is investigated (164 km^2) covering altitudes from 600 m up to ~4000 m. The catchment is partly glaciated (13%), and half of the total annual precipitation (1200–3000 mm, Kirchhofer & Sevruk 2001) falls as snow leading to a dominant

glacial discharge regime. Vegetation consists of meadows, coniferous forests and alpine grassland extensively used as pasture. The melt water of snow and glaciers were used irregularly in the valley to irrigate pastures in dry seasons. Soils are mostly shallow (~ 50 cm) and the texture except for the alluvial fans and the outwash plains in front of the glaciers consist of sandy silt. Bedrock material consists of amphibolites and gneiss (Börst 2006).

Five major meteorological stations were mounted in the year of 2000 representing two crossing transects: one transect stretches along the valley floor, the second covers the north- and the south-facing slopes. In addition, since autumn 2005 four additional minor stations have been installed to refine the elevational resolution. The minor stations are situated at half elevation between the stations covering the north- and south-facing slopes, two stations on each slope. All four stations were installed at 1900 m, on each slope one within the deciduous forest and one within an avalanche slope, respectively (cp. Fig. 2).

Moreover, the main investigation transect in the middle of the valley was extended both by elevation and elevational resolution. 13 basic stations were installed, six on the north-facing slope and seven on the south-facing slope. This transect stretches from 1400 m up to 2700 m recording hydrological processes from subalpine grassland to the nival belt. Within the subalpine belt, basic stations were set up within the forest and on non-forest sites like avalanche slopes at the same elevation. Figure 2 presents the location of the indifferent stations and summarizes the measured data at each station. Soil moisture was measured at each station with three samples and three repetitions using an un-calibrated handheld TDR-probe (Imko TRIME-FM). Major-, minor- and basic stations were located on local ridges or concave structure to ensure similar micro-topographical conditions.

Figure 2 summarizes the measurements conducted. The major and minor stations automatically measured the hydrological and meteorological parameters with an hourly temporal resolution, while the basic stations besides temperature were run using handheld measurements. These discontinuous measurements were accomplished and attained a temporal resolution of a few days. To obtain comparable data, discontinuous handheld measurements were conducted at major and minor stations, too. The results of the handheld measurement were correlated with the major stations, and temporal interpolation of the data was intended. Microscale processes were obtained by analyzing the measured data of each station. Mesoscale gradients were identified and derived by analyzing the differences of the station data against environmental parameters.

3.2 Model Level

Besides empirical investigations, a hydrological model is needed to simulate the distribution of hydrological processes and related gradients in the entire catchment (macroscale). For this purpose the physically based, distributed raster based model WaSiM-ETH in its current version WaSiM-ETH 7.10.1 (Schulla & Jasper 2007)

was used. The model is able to simulate the water balance at different spatial and temporal scales including the simulation of single raster cell dynamics. Discharge, evapotranspiration, soil moisture, snow distribution, as well as glacier melt are generated. The minimum input parameter needed are temperature, precipitation, topography, soil attributes, glacier distribution, and vegetation cover. For a detailed model description see Schulla & Jasper (2007).

Spatial data are needed to model the water balance of the valley. Vegetation data and digital elevation model (DEM) are available with high spatial resolutions of 5 and 10 m, respectively. Soil data are not available at a sufficient quality and hence a soil map has to be elaborated. Therefore, a random data sampling of 250 soil profiles was conducted in the field campaign. Based on these profiles a soil map is compiled using supervised classification. Discharge data for the validation of the model is provided by the state agency of Switzerland, which measures the runoff in the middle of the valley with a temporal resolution of 1 hour.

The superior importance of glacier and snow melt to the total discharge requires a validation of these hydrological storages, too. Snow cover distribution, and snow water equivalents were validated using the results of Schmidt (2007). Moreover, glacier mass balance was estimated using glacier ablation gauges that were installed in 2006.

4 Preliminary Results

4.1 Temporal Soil Moisture Dynamics

Figure 3 illustrates the interannual dynamics of soil moisture in reliance on climatic parameters. During that time the soil moisture varied between 67% vol and 9% vol, reaching its maximum in spring 2006 and the lowest value in November 2006. The investigated years show different dynamics in principle. In 2006 the soil moisture showed much higher amplitudes than 2007. In 2006 the temperature was higher (11.5°C in snow free times) and the amount of summerly precipitation was lower (585 mm) than in 2007 (9.3°C, 956 mm) leading to dryer conditions and a desiccation of the top soil (9% vol). During this time the permanent wilting point was nearly reached. Moreover, due to an earlier avalanche accumulation at the investigated plot the snow cover was thicker in spring 2006 resulting in a higher snow melt water contribution. This explains the high soil moisture content in May 2006. In 2007 the soil moisture amplitudes were more frequent as a result of recurring rainfall infiltration. These moist conditions together with cooler temperatures caused relatively moderate, balanced soil moisture contents.

The mean soil moisture was similar in both years with approx. 35% and never undershot the wilting point (WP, Fig. 3). But in July and August 2006 the soil underwent a period of very dry climatic conditions resulting in values between 15% vol and 9.4% vol. The low soil moisture values in November 2006 have to be interpreted as a result of freezing ground. The air and soil temperature indicates that in

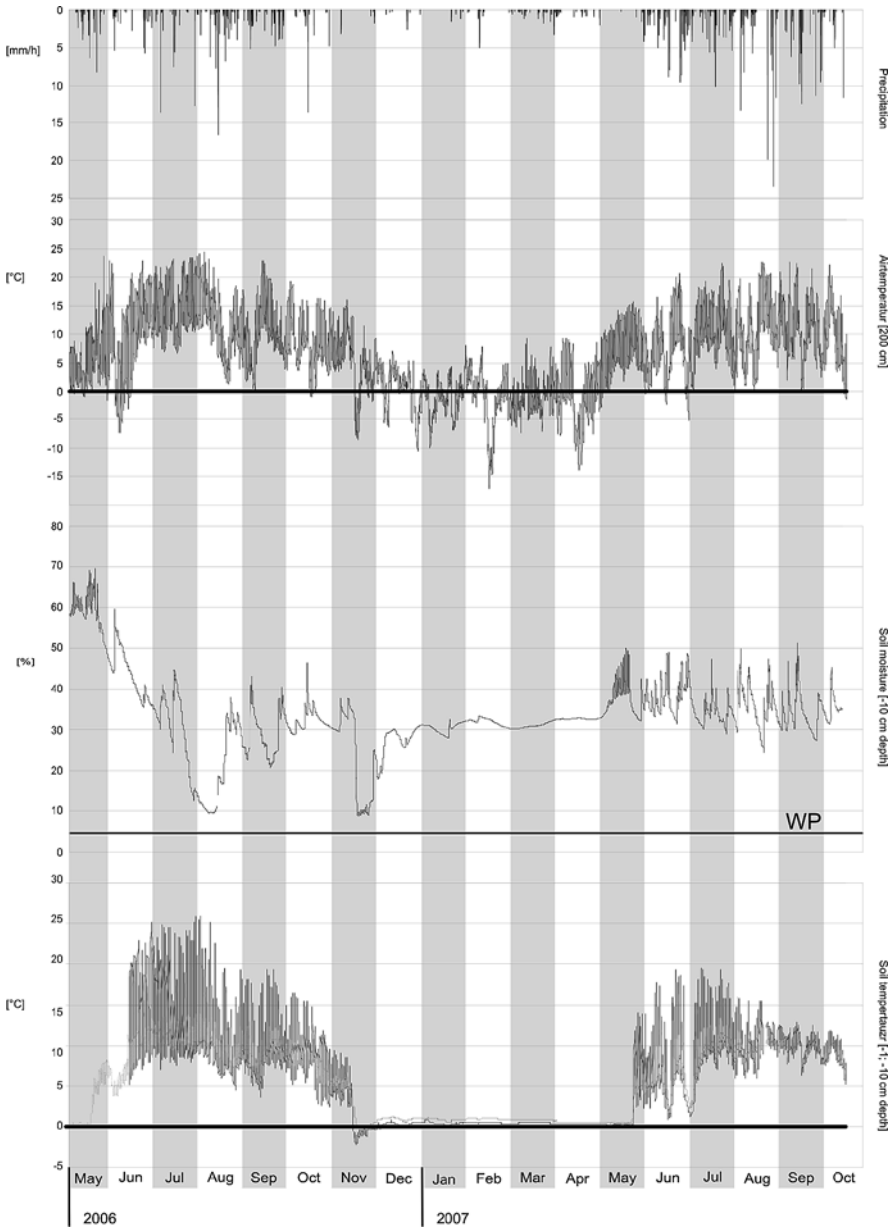

Fig. 3 The dynamics of soil moisture and effecting parameters like precipitation (*above*), air temperature (*second from above*), and soil temperature at two different depths

the mentioned time period the temperature declined rapidly down to –8.5°C and this cold period lasts four days making the freezing of the soil very likely. Rising temperatures after this period cause a thawing of the soil.

During the winter of 2006/2007 the soil moisture remained almost constant at approx. 30% vol and thereby slightly below the mean value. In according with the soil temperature this shows that the ground had never been frozen in winter, except for described four days in November 2006. Nearly constant soil temperatures with no daily amplitude indicated a permanent snow-cover from November 2006 till end of May 2007. The time of snowmelt can easily be assessed as a period of soil moisture fluctuations that corresponds with positive air temperatures and coeval constant soil temperature. The fluctuations of soil moisture fluctuations result from the infiltration of snow melt water. Figure 4 displays the snow melt period in detail for 2007. Soil moisture and air temperature oscillate parallelly. The amplitudes of soil moisture content increase in time and respond to the air temperature magnitude with a short delay. After complete snow melt as indicated by daily soil temperature fluctuations, soil moisture is primarily responding to liquid precipitation that had no influence on snow moisture as long as the snow persisted. Soil moisture responses from now on directly to liquid precipitation and these responses are more direct than snow melt infiltration.

Regarding the contribution of snow melt water to the annual water budget, the monitoring indicates the soil water surplus to be a function of snow cover thickness. In 2006 the high soil water content during snow melt remains above the average soil moisture over 6 weeks while in 2007 this surplus was limited to 2 weeks.

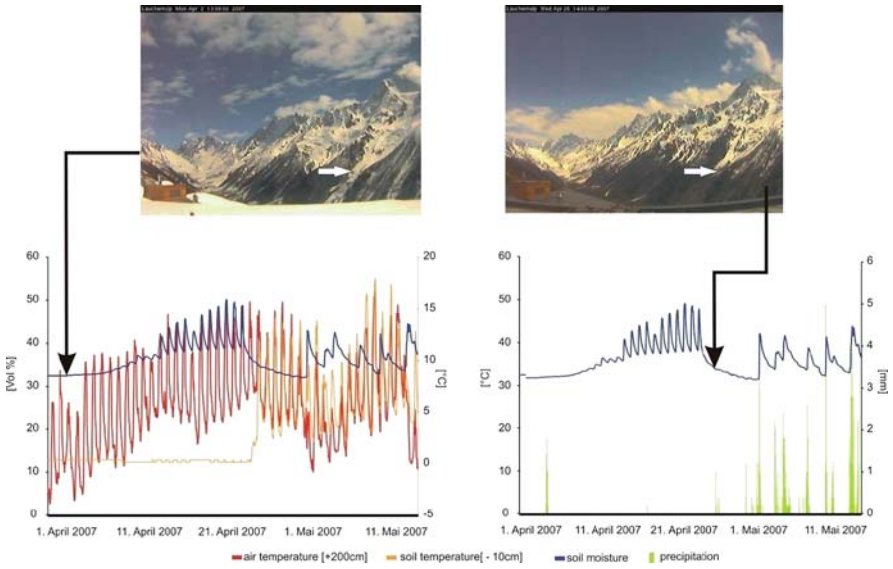

Fig. 4 Soil moisture dynamics during snow melt in April and May 2007 on the north-facing slope at 1900 m on an avalanche slope. Soil moisture dynamics is graphed against air- and soil-temperature (*left*) and precipitation (*right*). The pictures illustrate the snow cover distribution at the minor station (*white arrow*)

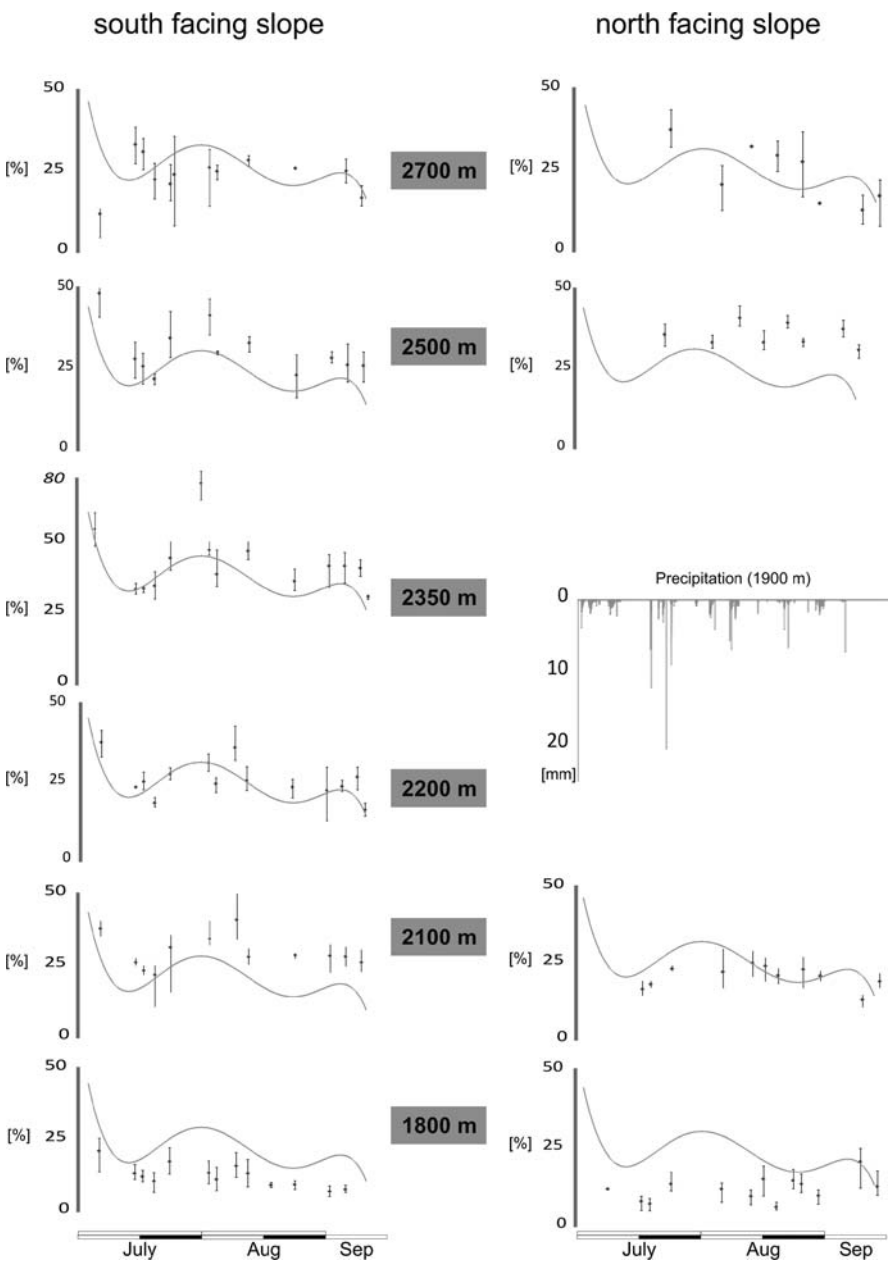

Fig. 5 Soil moisture in summer 2007 in reliance on elevation for the north- and the south-facing slopes. Error bars mark on-site variability, grey line on a polynomial function that was introduced to illustrate similar soil moisture dynamic at all sites. The temporal distribution of precipitation indicates the relation of rainfall and soil moisture dynamic

4.2 The Dynamics of Soil Moisture at Different Elevations and Exposition

The analysis of soil moisture shows the clear relation between soil moisture and precipitation. Figure 5 presents the soil moisture dynamics in summer 2007 in reliance on elevation and exposition. A superior dynamics is visible in most of the investigated plots that can be approximately described by a polynomial function (grey curve). This function represents an oscillation of soil moisture starting in July with relative high values as a result of earlier snow melt (cp. Fig. 3), followed by a drying period. High precipitation amounts in the beginning of August caused a rise of soil moisture at nearly all plots, but the amount of change differs. At 1800 mon the south-facing slope the increase was very small (+ 7% points) while at 2350 m on the same slope resulted in very high values (+ 35% points). Moreover, soil moisture at higher altitudes on the south-facing slope (at 2350 m and 2500 m) show strongest rise after the drying period. We assume that this is an effect of higher precipitation caused by rain clouds restricted by the northern rim of the valley during north-western flows. Despite further smaller rainfall events, a second drying period lasted till the end of August causing soil moistures at all sites to descend. Soil moisture rises again due to heavier rainfall events in mid of September.

On both slopes the soil moisture slightly increase with altitude, but this trend is not significant. Comparing north- and south-facing slopes, the investigated plots show similar values and similar dynamics indicating that exposition has no controlling effect on soil moisture. One exception marks the dynamics of highest situated plots on the north-facing slope with little or contrasting fluctuations. Low elevational sites (1800 and 2100 m) on both slopes show the lowest soil moisture most likely as a result of higher temperatures and hence higher evapotranspiration. These sites were the only ones that experienced drought stress for vegetation in summer 2007. Drought stress was detected by desiccated plant leafs of dwarf shrubs and grasses. All other sites although situated on ridges and concave positions maintained moderate soil moisture contents.

5 Discussion

We found the soil moisture dynamics to be highly variable at all plots. The superior dynamic of soil moisture were interpreted as a direct function of spatio-temporal precipitation distribution. Measurements showed that low soil moistures (9% vol) during dry periods occurred at lower elevations (up to 1800 m) and led to drought stress for plant. In contrast, soil moisture at higher elevational sites (1800–2700 m) never attained a critical level for plants. Moreover, a slight increase of soil moisture with altitude was derived. Löffler (2005) found similar annual dynamics in the Norwegian high mountains with high soil moisture values after snow melt, dryer conditions in summer but no drought stress and constant winterly soil moisture.

Our findings are in line with the early studies of Tranquilini (1964) who reported for the Alps that summer drought is seldom severe. Instead winter frost desiccation

has a major influence on the ecosystem. The latter can also be found in our data for some days in November 2006. Moreover, Löffler (2005) described frozen ground during winter leading to constant soil moisture contents in winter and surface runoff on frozen ground during the snow melt season. In contrast, we found that the soil at the north-facing slope was not frozen throughout the winter, at least up to this altitude. Therefore, snow melt water infiltrates into the soil resulting in interflow and base flow discharge and hence results in a delayed snow melt runoff. This process does not hold true for higher elevations since the glaciological map (Programmleitung NFP 31, 1998) declared permafrost to occur at app. 2500 m on the north-facing slopes. Here, conditions as described by Löffler (2005) must be assumed.

Gurtz et al. (2006) reported measured soil moisture dynamics in a neighboring valley but at lower elevation (700 m) with less snow fall. The measured annual dynamics of soil moisture showed the same characteristic with lowest values found in late summer and highest values occurred during the snow melt season in spring. Due to the higher snow pack and assumed higher percolation rates on steeper slopes, this characteristic shows a pronounced dynamics in alpine areas.

Assuming increasing precipitation and decreasing evapotranspiration, soil moisture in higher altitudes must consequently increase with altitude. This general assumption was proved, i.e., by Schawe (2005) for tropical mountain forests in Bolivia. We found only a slight, non-significant and non-linear trend of soil moisture with altitude and assume that this is mainly caused by the present precipitation patterns.

6 Research Perspectives

Microscale and mesoscale data revealed a detailed description of soil moisture dynamics and led to assumptions that can be proved by macroscale simulation. We found soil moisture to be highly variable with no significant dependency to elevation and exposition. These findings can now be verified at the macroscale using the hydrological model that was validated against the soil moisture dynamics at the measured stations. The assumption that the soil moisture dynamics and patterns are mainly caused by the present precipitation patterns have to be proved on the macroscale. The microscale temporal dynamics will be used to validate the simulated dynamics and will be used to characterize and to understand different principles of soil moisture dynamics at the macroscale. For example, the effect of frozen and non-frozen soils on soil moisture and runoff during snow melt will be spatially assessed and consequences for runoff building will be calculated.

The dynamic of storages like soil moisture, snow and glaciers play a key role to characterize the alpine water balance. The validation of these storages at the same time is therefore very important to assess the accuracy and certainty of the model results and has not yet been conducted for WaSiM-ETH. Moreover, the relevance of the hydrological processes and gradients as well as input parameters will be assessed by adding more detailed information to a basic model. This procedure will serve to evaluate the sensitive parameters for future simulation in similar alpine catchments.

References

Börst, U. 2006. Nachhaltige Entwicklung im Hochgebirge. Eine Systemanalyse von Mensch-Umwelt-Szenarien im Lötschental (Zentral-Alpen). Dissertation, Bonn

Bronstert, A., Bárdossy, A., Bismuth, C., Buiteveld, H., Disse, M., Engel, H., Fritsch, U., Hundecha, Y., Lammersen, R., Niehoff, D. & Ritter, N. 2007. Multi-scale modeling of land-use change and river training effects on floods in the Rhine basin. River Research and Applications 23, 1102–1125.

De Jong, C., Whelan, F. & Messerli, B. 2005. Preface: The importance of a hydrological research framework for water balance studies in mountain basins. Hydrological Processes 19, 2323–2328.

Gurtz, J., Baltensweiler, A. & Lang, H. 1999. Spatially distributed hydrotope-based modeling of evapotranspiration and runoff in mountains basins. Hydrological Processes 13, 2751–2768.

Gurtz, J. et al. 2006. Analysis of the time series of meteorological and hydrological measurements in the Rietholzbach research catchment between 1976 and 2005 with special consideration of the dry summer 2003. Institut für Atmosphäre und Klima der ETH Zürich.

Hörsch, B. 2001. Zusammenhang zwischen Vegetation und Relief in alpinen Einzugsgebieten des Wallis (Schweiz), Dissertation, Bonn.

Jasper, K., Gurtz, J. & Land, H. 2002. Advanced flood forecasting in Alpine watersheds by coupling meteorological observations and forecasts with a distributed hydrological model. Journal of Hydrology 267, 40–52.

Jasper, K., Calanca, P., Gyalistras, D. & Fuhrer, J. 2004. Differential impacts of climate change on the hydrology of two alpine river basins. Climate Research 26, 113–129.

Kirchhofer & Sevruk, B. 2001. Hydrological Atlas of Switzerland, Swiss Federal Office for Water and Geology, Berne.

Kirnbauer, W. & Sevruk, B. 2001. Mittlere Jährliche korrigierte Niederschlagshöhen 1951–1980. In: Landeshydrologie, Bundesamt für Wasser und Geologie: Hydrologischer Atlas der Schweiz, Bern.

Löffler, J. 2005. Snow cover dynamics, soil moisture variability and vegetation ecology in high mountain catchments of central Norway. Hydrological Processes 19, 2385–2405.

Löffler, J. 2007. The influence of micro-climate, snow cover, and soil moisture on ecosystem functioning in high mountains. Journal of Geographical Science 17, 3–19.

Mountain Agenda 1998. Mountains of the World – Water Towers for the 21st Century. Centre for Development and Environment (CDE), University of Berne, Switzerland.

Programmleitung NFP 31 1998. Glaziologische Karte Julier-Bernina (Oberengadin), Synthesekarte NFP 31, Maßstab 1:60,000.

Roots, E.F. & Glen, J.W. 1982. Hydrological aspects of Alpine and high mountain areas (Proceedings of the Exeter Symposium, July 1982). International Association of Hydrological Sciences 138, V–VI.

Schawe, M. 2005. Altitudinal change of climate and soils in Bolivian tropical montane rainforest ecosystems. Dissertation, Uni Göttingen.

Schmidt, S. 2007. Die reliefabhängige Schneedeckenverteilung im Hochgebirge. Ein multiskaliger Methodenverbund am Beispiel des Lötschentals (Schweiz). Dissertation. Bonn.

Schulla, J. & Jasper, K. 2007. Model description WaSiM-ETH.

Šimůnek, J. 2005. Models of water flow and solute transport in the unsaturated zone. Encyclopedia of Hydrological Sciences 1171–1180.

Tranquilini, W. 1964. The physiology of plants at high altitude. Annual Review of Plant Physiology 15, 345–362.

Verbunt, M., Gurtz, J., Jasper, K., Lang, H., Warmerdam, P. & Zappa, M. 2003. The hydrological role of snow and glaciers in alpine river basins and their distributed modeling. Journal of Hydrology 282, 36–55.

Viviroli, D., Weingartner, R. & Messerli, B. 2003. Assessing the hydrological significance of the world's mountains. Mountain Research and Development 23(1):32–40.

Welpmann, M. 1997. Geomorphologie und Naturgefahren im Lötschental (Schweiz). Diploma thesis, Bonn.

Winiger, M., Gumpert, M. & Yamout, H. 2005. Karakorum-Hindukush-Western himalaya: assessing high altitude water resources. Hydrological Processes 19: 2329–2338.

Zappa, M., pos, F., Strasser, U., Warmerdam, P. & Gurtz, J. 2003. Seasonal water balance of an alpine catchment as evaluated by different methods for spatially distributed snowmelt modeling. Nordic Hydrology 34, 179–202.

Index

J.-C. Otto, R. Dikau (eds.), *Landform – Structure, Evolution, Process Control*, Lecture Notes in Earth Sciences 115, DOI 10.1007/978-3-540-75761-0,