SEDIMENTS, SEC 3 • HILLSLOPE AND RIVER BASIN SEDIMENT DYNAMICS • RESEARCH ARTICLE

Factors controlling sediment yield at the catchment scale in NW Mediterranean geoecosystems

Joris de Vente • Rubia Verduyn • Gert Verstraeten • Matthias Vanmaercke • Jean Poesen

Received: 26 October 2010 / Accepted: 25 February 2011 / Published online: 23 March 2011 © Springer-Verlag 2011

Abstract

Purpose This study aimed to (1) increase understanding of the relation between sediment yield and environmental variables at the catchment scale; (2) test and validate existing and newly developed regression equations for prediction of sediment yield; and (3) identify how better predictions may be obtained. Materials and methods A correlation and regression analysis was performed between sediment yield and over 40 environmental variables for 61 Spanish catchments. Variables were selected based on availability and expected relation with diverse soil erosion and sediment transport processes. For comparison, the Area Relief Temperature (ART) sediment delivery model was applied to the same catchments. Sediment yield estimates obtained from reservoir surveys were used for model calibration and validation. Results and discussion Catchment area, catchment perimeter, stream length, relief ratio, Modified Fournier Index, the RUSLE's R factor, and catchments percentage with poor vegetation cover showed highest correlations with sediment

 Responsible editor: Marcel van der Perk

 J. de Vente (⊠)

 Estación Experimental de Zonas Áridas, EEZA-CSIC,

 Department of Desertification and Geoecology,

 Carretera de Sacramento s/n, La Cañada de San Urbano,

 04120 Almeria, Spain

 e-mail: Joris@sustainable-ecosystems.org

 J. de Vente

 School of Geosciences, University of Aberdeen,

 Aberdeen, UK

 R. Verduyn · G. Verstraeten · M. Vanmaercke · J. Poesen

 Geography Division Katholieke Universiteit Leuven

 GEO-INSTITUTE,

 Department Earth and Environmental Sciences,

Department Earth and Environmental Sciences Celestijnenlaan 200 E, 3001 Heverlee, Belgium vield. Stepwise linear regression revealed that variables representing topography, climate, vegetation, lithology, and soil characteristics are required for the best prediction equation. Although calibration results were relatively good, validation showed that the models were unstable and not suitable for extrapolation to other catchments. Reasons for this unstable model performance include (1) lack of detail and quality of the data sources; (2) large variation in catchment characteristics; (3) insufficient representation of all relevant erosion and sediment transport processes; and (4) the presence of nonlinear relations between sediment yield and environmental variables. The nonlinear ART model performed relatively well but systematically overpredicted sediment yield. A model reflecting human impacts, including dams and conservation measures, is expected to provide better results. This, however, requires significantly more input data. Conclusions Although important insight is obtained into the relation between sediment yield and environmental factors, prediction of sediment yield at the catchment scale requires alternative approaches. More detailed information is required on land cover (change), and the effect of soil conservation measures. Validation of regression equations is a necessity, and better predictions are obtained by nonlinear models.

Keywords ART model · Mediterranean · Multiple regression · Prediction · Sediment yield · Soil erosion · Spanish reservoirs

1 Introduction

Understanding and predicting sediment yield from catchments form an essential part of geomorphologic and ecosystem research, and are indispensable to support policy decisions dealing with off-site effects of soil erosion such as sediment deposition within river channels and reservoirs. flooding, coastal development, and contamination of floodplains and water bodies with agrochemicals and other pollutants associated with the eroded sediments (Clark 1985; Owens et al. 2005; Ramos and Martinez-Casasnovas 2004; Steegen et al. 2001; Syvitski and Milliman 2007; Verstraeten and Poesen 1999; WCD 2000; Woodward 1995). Moreover, river sediment fluxes play an important role in various natural geochemical cycles, such as the carbon cycle (Ludwig and Probst 1998; Ludwig et al. 1996). Before predictions of sediment yield can be made, understanding the integrated effect of diverse erosion and sediment transport processes in relation to land use and climate change is needed. In recent decades, significant progress has been made in the development of spatially distributed, process-based soil erosion models such as LISEM (de Roo 1998; de Roo et al. 1996), EUROSEM (Morgan et al. 1998), WEPP (Flanagan et al. 1995; Foster et al. 1995), LAPSUS (Schoorl and Veldkamp 2001) and PESERA (Kirkby et al. 2004). However, mainly because of the large data requirements and difficulty of describing the effects of all relevant erosion and sediment transport processes, operational application of these models at the catchment scale is often problematic. A first step towards understanding and prediction of sediment yield at the catchment scale, therefore, is the use of lumped regression equations (e.g. Achite and Touaibia 2000; Ali and de Boer 2008; Dendy and Bolton 1976; Grauso et al. 2008; Hadley et al. 1985; Langbein and Schumm 1958; Lixian et al. 1996; Ludwig and Probst 1998; Ludwig et al. 1996; Restrepo et al. 2006; Syvitski et al. 2005; Tamene et al. 2006; Verstraeten and Poesen 2001). In these equations, sediment yield is often related to catchment characteristics in terms of morphology, topography, lithology, climate, discharge and soil properties. Obviously, these regression relations not only serve prediction purposes but also provide insight into relations between catchment characteristics and sediment yield that can be used for the further development of more complex models.

Table 1 provides some examples of regression equations as reported in the literature for the prediction of sediment yield for various regions worldwide. An extensive overview and discussion of regression equations can be found in Ludwig and Probst (1998) and in Jansson (1982), who grouped equations according to drainage basin size and climatic conditions. The equations in Table 1 illustrate the range of variables used in regression studies. Some equations are based on only one variable, but most use several variables to predict sediment yield. The variables most often reported to be related to sediment yield can be grouped as *topography* (e.g. relief ratio, slope), *climate* (e.g. mean annual rainfall, erosivity index), *soil and lithology* (e.g. % of soil texture class or erodible lithology), *hydrology* (e.g. mean or maximum runoff discharge), *vegetation cover* or *land use* (e.g. % forest or orchards), *drainage network* (e.g. drainage density, stream length), and *catchment morphology* (e.g. basin area, form factor). Beside the variety in variables, it is interesting to observe the large differences in optimal prediction equations for sediment yield in different parts of the world. These differences highlight that (1) erosion and sediment transport processes depend on many interacting environmental factors; (2) a factor that is important in one area can be insignificant in another; and so (3) no straightforward globally effective relations are apparent, and hence region-specific relations are needed for predicting sediment yield.

The objective of this study is to assess the relation between sediment yield at the catchment scale and data on topography, climate, land use, catchment morphology, drainage network, lithology and soil properties, for NW Mediterranean geoecosystems, using 61 Spanish study catchments. Furthermore, the aim is to test to what extent these regional data can be used in a multiple linear regression model for the prediction of sediment yield. This study follows on from a previous study where the sediment yields of 22 Spanish catchments were related to environmental factors of climate, topography and land use (Verstraeten et al. 2003). In the latter study, it was observed that, although a relatively good regression equation was obtained ($R^2=0.80$), the model was very unstable since exclusion of outliers changed the model and greatly reduced the model performance ($R^2=0.29$). As explanation for the model instability, it was concluded that land cover information lacked detail, and that the process of gully erosion was not sufficiently represented by included variables. Therefore, beside a higher resolution for most data, in this present study, more factors that potentially can be a proxy for the occurrence of gully and bank erosion (e.g., lithology, drainage density) were included in the analysis, and the dataset was extended to 61 catchments for which measured sediment yield data were available from the literature. In addition to the first two objectives-for comparison and to test the potential of a global prediction equation as an alternative to regionally calibrated equationsthe Area Relief Temperature (ART) sediment delivery model that was developed for assessment of sediment load of global rivers (Syvitski et al. 2003; Syvitski et al. 2005) was also applied to and validated for the 61 Spanish catchments.

2 Materials and methods

2.1 Sediment yield data of Spain

Sediment yield was calculated from published sedimentation rates in 61 reservoirs in Spain (Avendaño Salas et al.

Table 1 Exam	ples of published regression relations for the prediction of sediment yield			
Location	Regression model	R^2	N	Reference
Algeria	SSY = 26.62*LI + 5.07*PI + 9.77*Tor - 593.56	0.92	30	Demmak (1984)
Belgium	SSY=25*A ^{-0.4} (with area in ha) SY=25*A ^{0.6} (with area in ha)	$0.64 \\ 0.80$	26 26	Verstraeten and Poesen (2001)
	$\ln SSY = 3.72 - 0.72* \ln A - 0.84* \ln HI + 0.11* \ln SL (with area in ha)$	0.76	26	
	$\ln SY = 0.21^{*}L + 22.2^{*}HD - 988^{*}HI$ (with area in ha and L in m)	0.92	26	
China	$\ln SSY = -2.65 + 0.96*S + 0.002*L^2 + 8.41*L^{0.5} - 4.16*L^{2/3} + 3.22*W^{0.5} - 1.45*W^{2/3} - 2.23*T + 2.46*T^2 - 1.36*For^2 - 1.26*For^2 - 1.2$	0.63	200	Lixian et al. (1996)
Colombia	$\log SSY = -0.88 + 0.81*\log Q - 0.39* \log Q_{max}$	0.58	32	Restrepo et al. (2006)
Ethiopia	SY = 690^{*} TDL - 0.58^{*} SWC SSY = 0.86^{*} S - 0.269^{*} SWC + 10	0.96 0.80	11 11	Haregeweyn et al. (2005)
Ethiopia	$\log SSY = 2.33 + 0.007*GulB + 0.003*EL + 0.002*RG - 0.007*Bush$	0.96	11	Tamene et al. (2006)
Italy	$SSY = 497.61 - 1.74^*A + 129.42^*E - 12.61^*S - 0.29^*MAP + 121.18^*DD$	0.82	42	Bazzoffi et al. (1997)
Maghreb	$\ln SSY = 4.79 + (54^*10^{-3*}K) + (4^*10^{-3*}R) - (5.6^*10^{-5*}A)$	0.71	36	Probst and Amiotte-Suchet (1992)
Romania	$\log SY = 4.50 - 0.18* \log O + 0.75* \log A + 0.04* \log FF + 0.10* \log DD + 0.33* \log RR + 0.54* \log MAP$	0.74	63	Ichim (1990)
Romania	$\log SY = 7.99 + 0.81* \log A - 0.30* \log BR + 0.15* \log DD - 0.15* \log RR + 0.09* \log For - 1.57* \log MAP$	0.92	36	Ichim (1990)
South-Africa	SSY = 1.45*P - 0.78	0.78	15	Roberts (1973)
Spain	SSY = 20.6*RR + 18*D - 14*For - 33*Foma - 31*Orch - 173*PV	0.80	22	Verstraeten et al. (2003)
USA	SSY = $1280^{*}0^{0.46} * (1.43 - 0.26^{*} \log A)$; for Q<2 in./year SSY = $1958e^{-0.055} * (1.43 - 0.26^{*} \log A)$; for Q>2 in./year	0.75	505	Dendy and Bolton (1976)
	(with SSY in t/mi^2 , Q in inches, and A in mi^2)			
USA	SSY = $0.39*DD + 1.85$; with DD based on permanent and ephemeral streams SSY = $2.81*DD + 1.81$; with DD based on permanent streams	0.02 0.08	17 17	Allen (1986)
	SSY = 0.75 + DD - 0.16; with DD based on aerial photos	0.83	17	
USA	$\log(SSY + 100) = 6.21 - 2.19^* \log(PT + 100) + 0.06^* \log(S + 100) - 0.02^* \log(CS + 100) + 0.04^* \log(AI + 100) + 0.04^* $	0.92	39	Flaxman (1972)
Global	(with SSY in acre feet per square mile) ln SSY = $-0.416^* \ln A + 4.26^* 10^{-4*} H_{max} + 0.15^* T_m + 0.095^* T_{range} + 0.0015^* R + 3.585$	0.49	86	Hovius (1998)
Global	$Q_{S} = 2^{*}10^{-5*}A^{0.50*}HD^{1.50*}e^{0.1*Tm}$ (for Polar regions, $T_{m} < 0^{\circ}C$) $Q_{C} = 6.1*10^{-5*}A^{0.55*}HD^{1.12*}e^{007*Tm}$ (for tourscore units) being a point of the	0.76	48	ART model: Syvitski
	$Q_{5} = 0.1 \text{ I}_{5} = 0.1 \text{ I}_{1} = 0.400 \text{ M}^{-1} = 0.1 \text{ M}^{-1} = 0.1 \text{ M}^{-1} = 0.1 \text{ M}^{-1} = 0.131 \text{ M}^{-1} = 0.400 \text{ M}^{-1} = 0.131 \text{ M}^{-1} = 0.400 \text{ M}^{-1} = 0.131 \text{ M}^{-1} = 0.400 \text{ M}^{-1} = 0.131 \text{ M}^{-1$	0.00 0 5 0	107 67	(CUUZ) .13
	$Q_{S} = 0.57^{*}A^{0.50*}HD^{0.37*}e^{-0.1^{*}Tm}$ (for tropical regions, 0.30°S)	0.67	42	
	$Qs = 1.3*10^{-3}*A^{0.43}*HD^{0.96}$ (for temperate regions with latitude>30°S)	0.54	26	
Global	Qs = $0.02^{*}B^{*}Q^{0.31*}A^{0.5*}R^{*}T_{m}$, for $T_{m} > 2^{\circ}C$ Qs = $0.04^{*}B^{*}Q^{0.31*}A^{0.5*}R$, for $T_{m} < 2^{\circ}C$	0.96	488	BQART model: Syvitski and Milliman (2007)
Global	SSY = 0.020*(R*Sr*Four)	0.91	58	Ludwig and Probst (1998)

1

ned	
atin	
<u>5</u>	
_	
le	
ac	

Location	Regression model	R^2	z	Reference
Global	$\begin{split} & SSY = 6.14^* p^2 / MAP - 49.78; (low relief, p^2 / MAP \geq 20) \\ & SSY = 27.12^* p^2 / MAP - 475.4; (low relief, p^2 / MAP > 20) \\ & SSY = 52.49^* p^2 / MAP - 513.2; (high relief, MAP > 600) \\ & SSY = 91.78^* p^2 / MAP - 737.62; (high relief, 200 < MAP < 600) \end{split}$		78	Fournier (1960)

precipitation temperature ratio (mm/°C), PV% poorly vegetated, Q s^{-1}), Qs long-term sediment load (kg s^{-1}), R specific runoff (mm a^{-1}), RG index of nuggedness ratio, Bush % bush/shrub cover, CS % soil particles coarser than 1 mm, D distance between valley outlet and highest point at the divide (km), DD drainage density (km km⁻²), E erodible surface area (km²), H_{max} maximum P LR/BR, MAP mean annual Tor R^2 correlation coefficient, N number of observations, A catchment area (km²), AI aggregation index, B factor representing glacial erosion, lithology and reservoir or lake trapping efficiency, BR bifurcation (km⁻¹), RR relief ratio (m km⁻²), S average slope gradient (%), Sr average slope gradient (radian), SL total stream length (m), SSY area-specific sediment yield (t km⁻² a⁻¹), SWC % of catchment treated with T_n mean annual temperature (°C), T_{range} annual temperature range (°C), TDL total drainage length (km), Fournier Index (mm), GulB index for gully and bank erosion (-), length ratio, R Index, LI Lithology length (km), precipitation (mm), PI precipitation erosivity index, PT factor, L valley forest cover, Four K soil erodibility mean maximum river discharge (m³ transition forest Matorral, For % terraces, point (m), SY absolute sediment yield (t a^{-1}), T % area with Orch % orchards, p maximum mean monthly and lowest \mathcal{Q}^{max} 1 *EL* % with erodible lithology, *FF* catchment form factor (-), *Foma* % HD difference between highest a⁻¹), s^{-1}), Qa mean annual discharge (km³ Forrentiality Index, W average catchment width (km) elevation (m), HI hypsometric integral, precipitation (mm), O catchment order, soil and water conservation measures, mean annual runoff (m³

1997a, b). Volumes of sediments retained in the reservoirs were measured by the Centro de Estudios Hidrográficos (CEH-CEDEX) of the Spanish Ministry of the Environment using bathymetric reservoir surveys. The methodology was extensively described in various reports (Avendaño Salas and Cobo Rayán 1997; Avendaño Salas et al. 1995; CEDEX 1992). At the moment of the survey, the reservoirs were operative between 20 and just over 100 years. Surveys were performed between 1977 and 1994. The reservoirs are located in diverse climatic, geologic and geomorphologic regions of the country (Fig. 1), but without representation of the relatively humid northwestern area. Almost half of all catchments are semiarid with an aridity index (UNEP 1997) below 0.5. Around 82% of the catchments have an aridity index between 0.29 and 0.65 and are thus considered to be semiarid to dry subhumid. The remaining catchments are humid-subhumid to humid. The measured total sediment yield (SY) varies between ~3.5 and $5,300 \times 10^3$ t a⁻¹, and the area-specific sediment yield (SSY) varies between 10 and 2,600 tkm⁻² a⁻¹. The average surface area draining to the reservoirs is about $1,251 \text{ km}^2$.

2.2 Quantitative catchment properties and correlations

An extensive database of quantitative variables was prepared for each of the 61 catchments (Table 2). The selection of variables included in the analysis was based on the availability of the data for all 61 catchments. Furthermore, it was intended to include variables that are most widely reported to be related to soil erosion and sediment transport processes characterising topography, climate, land use, catchment morphology, drainage network, lithology and soil properties. The variables related to catchment morphology, drainage network and topography were extracted from the Shuttle Radar Topography Mission (SRTM) 3 arcsecond Digital Elevation Model (DEM) distributed by the USGS EROS data centre (http://seamless.usgs.gov/; Rabus et al. 2003), using IDRISI software (Eastman 2003). This DEM was mosaiced and resampled to 100-m spatial resolution by bilinear interpolation. Beside the calculation of catchment area, slope gradient, minimum and maximum elevation, the DEM was used to extract a river network for all catchments by assuming a channel when the upslope contributing area is larger than 0.2 km². This threshold roughly corresponds with the river network and drainage density indicated on topographical maps at a scale of 1:50,000. One of the reasons for the use of drainage density as an explanatory factor is that it may be a proxy for gully presence. Since no numbers or maps were available on gully density for all catchments, it was decided to extract the river network as detailed as possible based on the SRTM DEM, in the hope that this drainage density would also be representative for gully density. The threshold of 0.2 km² provided the most detailed





Fig. 1 Location of the 61 reservoirs in Spain

reasonable drainage network without deformations of the drainage network. For calculation of the climate variables, gridded (1 km) mean monthly rainfall and air temperature data, based on station data of the Spanish National Meteorological Institute (INM) for the period 1971–2000, were used. Mean annual precipitation for the catchments varies between 350 and 1,400 mm, and the mean annual temperature varies between 8°C and 18°C. The Precipitation Concentration Index (PCI; Oliver 1980), the Modified Fournier Index (MF; Arnoldus 1977), and the rainfall erosivity R factor (Rf; Renard and Freimund 1994) from the Revised Universal Soil Loss Equation were calculated as:

$$PCI = \frac{\sum_{i=1}^{12} p_i^2}{P^2} *100$$
(1)

$$MF = \frac{\sum_{i=1}^{12} p_i^2}{P} \tag{2}$$

$$Rf = 0.7397*MF^{1.847}$$
, for MF < 55 mm (3)

$$Rf = 95.77 - 6.081*MF + 0.4770*MF^2$$
, for MF > 55 mm (4)

Here, p_i is the average monthly precipitation (millimetre), and P is the average annual precipitation (millimetre). Land

use information was extracted from the CORINE Land Cover map (CLC2000) of the European Environment Agency (EEA 2000). The CORINE map has a resolution of 100 m, and contains 44 land cover classes of which the classes that were considered most relevant for erosion prediction were selected. The variables related to lithology, soil texture and soil type were extracted from the European Soil Geographical Database at 1-km resolution (ESB 2004). This database (ESDBv2 Raster Archive) is freely available from the website of the European Soil Bureau (http://eusoils.jrc.it/).

2.3 Statistical analysis and model validation

A correlation analysis was performed to obtain insight into the explanatory value of individual variables for prediction of sediment yield and to identify the relation between the explanatory variables. Then, in order to find a predictive model of sediment yield, a stepwise linear regression analysis was performed with SAS statistical software (SAS 1999), using only those variables with a relatively high correlation coefficient from the correlation matrix. The stepwise regression combines forward and backward selection using a significance level threshold of 0.15. Since model development for absolute sediment yield (SY; t a^{-1}) or specific sediment yield (SSY; t km⁻² a^{-1}) provides different results (de Vente et al. 2005, 2008), model equations were calibrated for both SY and SSY.

To compare the obtained linear regression equations with other previously developed prediction equations, we also

Table 2 Summary and source of variables used in the correlation analysis

	Abbreviation	Description	Range	Mean	Source/Reference
Catchment	A (km ²)	Catchment area	29–13,248	1,251	SRTM DEM
morphology	Peri (km)	Catchment perimeter	31-1,188	249	SRTM DEM
	C (-)	Circle ratio is calculated as $4\pi A/Peri^2$	0.10-0.41	0.22	SRTM DEM
	D (km)	Distance between the reservoir and the highest point on the catchment divide.	7–126	44	SRTM DEM
	FF (-)	Form Factor, calculated as A/D^2	0.19-1.00	0.52	SRTM DEM
Drainage network	SL (km)	Total stream length automatically extracted	41–20,224	1,837	SRTM DEM
	DD (km km ^{-2})	Drainage Density, calculated as SL/A	1.20-2.07	1.44	SRTM DEM
Topography	S (%)	Mean slope gradient	7–37	18	SRTM DEM
	H _{min} (m)	Minimum elevation	13–935	458	SRTM DEM
	H _{max} (m)	Maximum elevation	516-3,135	1,707	SRTM DEM
	HD (m)	Height Difference as: H_{max} - H_{min}	503-2,812	1,249	SRTM DEM
	H _{mean} (m)	Mean elevation	147–1,613	905	SRTM DEM
	HI (-)	Hypsometric Integral, as: (H _{mean} -H _{min})/(H _{max} -H _{min})	0.18-0.55	0.37	SRTM DEM
	$RR (m km^{-2})$	Relief Ratio, calculated as: (H _{max} -H _{min})/A	0.12-20.08	2.97	SRTM DEM
Climate	MAP (mm)	Mean Annual Precipitation	362-1,362	661	INM data 1971-2000
	PCI (%)	Precipitation Concentration Index (see Eq. 1)	8.60-11.92	9.94	Michiels et al. (1992)
	MF (mm)	Modified Fournier index (see Eq. 2)	34–147	66	Arnoldus (1977)
	${ m Rf} ({ m MJ.mm.m}^{-2}.{ m h}^{-1}. { m yr}^{-1})$	RUSLE erosivity factor R (see Eqs. 3–4)	495–10,100	2,097	Renard and Freimund (1994)
	T _m (°C)	Mean annual temperature	8-18	13	INM data 1971-2000
Land use	Urban (%)	Percentage Urban area	0–4	1	CORINE (CLC2000)
	Ar (%)	Percentage Arable land	0-85	21	CORINE (CLC2000)
	OlAlm (%)	Percentage Olive and Almond orchards	0-37	7	CORINE (CLC2000)
	Past (%)	Percentage Pasture	0-31	7	CORINE (CLC2000)
	Compl (%)	Percentage Complex cultivation patterns	0–40	7	CORINE (CLC2000)
	For (%)	Percentage Forest cover	1–72	23	CORINE (CLC2000)
	Foma (%)	Percentage transition Forest Matorral	0–24	10	CORINE (CLC2000)
	PV (%)	Percentage Poorly Vegetated surfaces	0–24	3	CORINE (CLC2000)
	Mat (%)	Percentage Matorral	1-58	17	CORINE (CLC2000)
Lithology	Sand (%)	Percentage clastic sedimentary rock, sandstone lithology.	0-84	26	ESB (2004)
	Lime (%)	Percentage limestone lithology	0–100	31	ESB (2004)
	Marls (%)	Percentage Marls lithology	0–100	41	ESB (2004)
	Plut (%)	Percentage plutonic rock lithology	0–57	22	ESB (2004)
	Argil (%)	Percentage argillaceous rocks/shale	0-100	34	ESB (2004)
	Fll (%)	Percentage fluvial clays, silts, and loam	0-41	11	ESB (2004)
	Meta (%)	Percentage acid metamorphic rocks	0-77	23	ESB (2004)
Soil texture	Coarse (%)	Percentage coarse soil texture	0-86	25	ESB (2004)
and type	Medium (%)	Percentage medium soil texture	0-100	55	ESB (2004)
	Fine (%)	Percentage fine soil texture	2-100	34	ESB (2004)
	Fluv (%)	Percentage of Fluvisols	0-21	7	ESB (2004)
	Camb (%)	Percentage of Cambisols	1-100	54	ESB (2004)
	Lept (%)	Percentage of Leptosols	0-100	41	ESB (2004)
	Rego (%)	Percentage of Regosols	0–99	45	ESB (2004)
	Luvi (%)	Percentage of Luvisols	0–49	13	ESB (2004)

Also indicated are the range and mean of the measured variables

applied the Area Relief Temperature sediment delivery model (ART) that was developed for assessment of prehuman sediment load of rivers worldwide (Syvitski et al. 2003, 2005). The ART model is a nonlinear regression model that was trained on a global database of 340 rivers and consists of 5 regression equations for 5 climatic zones (see also Table 1). The ART model equations predict longterm sediment load (Qs; kg s⁻¹) based on catchment area (A), large-scale relief (height difference; HD), and mean annual air temperature (T_m). The ART regression equation for temperate regions at northern latitudes (>30°) was defined by Syvitski et al. (2003) as:

$$Qs = 6.1*10^{-5*} A^{0.55*} HD^{1.12*} e^{0.07^* Tm} (R^2 = 0.63; n = 162)$$
(5)

Although ART was calibrated mostly for large catchments, the model is assumed to be applicable at a global scale and was also applied to smaller catchments (Syvitski and Kettner 2007). Therefore, Eq. 5 was applied to the 61 Spanish catchments, and the predicted sediment load (Qs; kg s⁻¹) was converted to absolute sediment yield (SY; t a⁻¹) by multiplication of Qs with 31,536, and to areaspecific sediment yield (SSY; t km⁻² a⁻¹) by subsequent division with catchment area.

The prediction accuracy of all regression equations was evaluated by the proportion of explained variance (R^2) , the adjusted R^2 , the Model Efficiency (ME) and the Relative Root Mean Square Error (RRMSE). The ME and RRMSE are calculated as (Nash and Sutcliffe 1970):

$$ME = 1 - \frac{\sum_{i=1}^{n} (O_i - Pd_i)^2}{\sum_{i=1}^{n} (O_i - O_{mean})^2}$$
(6)

$$RRMSE = \frac{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (O_i - Pd_i)^2}}{\frac{1}{n} \sum_{i=1}^{n} O_i}$$
(7)

Here, *n* stands for the number of observations, O_i is the observed value, Pd_i is the predicted value and O_{mean} is the mean observed value. The ME can range from $-\alpha$ to 1 and represents the part of the initial variance accounted for by the model. So, the closer the ME approaches 1, the more efficient the model is. The RRMSE is independent of the units, and the smaller the RRMSE value, the more accurate the model prediction.

Validation is necessary to test the robustness of the linear regression equations and explore the potential for extrapolation to other datasets. This is mostly done by splitting the dataset into two parts, one for model development and one for model validation. In this case, the models were validated using the "Jackknife" procedure (Shao and Tu 1995). In this procedure, n-1 catchments are used for calibration of the model. This calibrated model is then used for prediction of sediment yield in the remaining catchment. This is repeated *n* times in order to obtain independently predicted sediment yield for each catchment. This means that 61 test models were constructed, with the same independent variables, each of which excluding one reservoir at a time, and after which the model was applied for predicting SY and SSY for the remaining reservoir. The observed sediment yield of the excluded catchment was compared to the predicted value of the corresponding test model. Model performance was evaluated by the R^2 , ME and the RRMSE.

3 Results

3.1 Correlation between catchment properties and sediment yield

Table 3 shows a correlation matrix between catchment properties and area-specific (SSY) and absolute (SY) sediment yield. The nonparametric Spearman correlation coefficient was used since practically none of the variables are normally distributed. Some of the variables have a lognormal distribution (e.g. A, SL, DD, S, RR, MAP, MF), while most of the land use, lithology and soil-related variables show a skewed distribution because of many zero values for catchments where this class does not occur. Only variables with a relatively high correlation coefficient (>0.20) and those for which a relevant relation with other variables was found are included in Table 3. Only a few variables are significantly (α =5%) correlated with sediment yield. The most important explanatory variables for both SY and SSY are catchment area (A), stream length (SL), relief ratio (RR), Modified Fournier Index (MF), RUSLE R factor (Rf), catchment perimeter (Peri) and the percentage of poorly vegetated areas (PV). The high correlation coefficients between some of the variables point towards (multi-)collinearity between some variables. For example, stream length (SL), catchment area (A) and perimeter (Peri) are strongly correlated, and also, the percentage of soils with a fine texture and the percentage of argillaceous rocks show a strong correlation. For comparison, correlations were also assessed based on the Pearson correlation coefficient. Although correlations were slightly different, this provided the same selection of main explanatory variables as based on the Spearman correlation coefficient.

Figure 2 illustrates some of the important correlations in scatter plots. Catchment area shows a strong negative

		Catchm morpho	lent logy		Drainage network	F	opograp	hy			ū	imate			Lano	luse				Lithc	logy			Soil te	exture a	nd type
	SSY	V	Peri	C	SL		111	lmin H	lmax F	R	M N	AP PC	I MF	Rf	Ar	OIAli	n For	ΡV	Mat	Lime	Marl	Argil	Meta	Med	Fine	Fluv
SY	0.58	0.45	0.47	-0.28	0.45	0.04	0.01 -	0.28	0.22 -	0.28 -	9.37 0	.23 0.	20 0.	26 0.	28 0.0	5 0.2	5 -0.0	5 0.3	5 -0.2	0 -0.1	1 -0.05	5 0.21	-0.04	-0.07	-0.07	0.40
SSY		-0.38	-0.36	0.22	-0.38	-0.23	0.37 -	0.09	0.02 -	0.10	0.41 (.38 0.	32 0.	44 0.	44 -0.2	3 0.0	1 0.0	8 0.2	9 -0.1	5 -0.10	30.0	3 -0.05	-0.22	-0.22	0.12	0.06
А			0.97	-0.49	0.99	0.18	-0.31 -	0.14	0.24 -	0.18 -	9.90 –0	.07 -0.	18 -0.	12 –0.	10 0.2	5 0.2	3 -0.0	6 0.0	6 -0.0	5 0.0	4 -0.12	2 0.30	0.21	0.18	-0.15	0.36
Peri				-0.64	0.97	0.23 -	- 0.37 -	0.21	0.21 -	0.20 -	9.89 -0	.14 –0.	13 -0.	17 -0.	16 0.3	3 0.2	<u>5</u> -0.1	0 0.0	5 -0.0	8 0.0	8 -0.05	5 0.25	0.17	0.14	-0.11	0.37
С					-0.50	-0.15	0.33	0.32	0.01 -	0.01	0.51 (.34 -0.	05 Ø.	33 0.	33 -0.4	<i>i</i> -0.3	0.0	9 0.1	4 0.1	0 -0.2	2 -0.13	3 0.01	0.05	-0.01	-0.06	-0.06
SL						0.27 -	- 0.36 -	0.14	0.22 -	0.22 -)- 16.	.0- 60.	16 -0.	13 –0.	11 0.2	8 0.2	5 -0.1	0 0.0	6 -0.0	6 0.0	3 -0.10	0.30	0.18	0.19	-0.14	0.38
DD						I	- 0.64 -	- 0.16	-0.11 -	0.37).22 -6	.50 -0.	05 – 0 .	47 -0.	46 0.5	0 0.2	1 -0.5	4 -0.0	1 0.0	1 -0.13	8 0.32	2 -0.02	2 -0.17	0.05	0.19	0.17
S								0.15	0.36	0.33	9.54 (47 -0.	03 0.	43 0.	43 -0.6	9 -0.2	9 0.5	7 0.3	7 0.0	9 0.20	30.0- 0	8 -0.27	0.21	-0.08	0.07	-0.02
Hmin									0.51	0.06	0.19 0	.13 –0.	49 -0.	03 -0.	0.0 -0.0	9.0- 6	1 0.4	2 0.0	6 -0.0	2 0.3	1 -0.13	3 -0.06	0.12	0.04	-0.12	-0.13
Hmax										0.29	0.10	.26 -0.	28 0.	14 0.	18 0.0	1 -0.3	5 0.3	2 0.4	3 -0.0	8 0.4	4 0.05	9 -0.15	0.19	-0.12	0.13	0.25
IH										-	0.09	.00 -0.	19 -0.	00 -0.	08 - 0. 4	8 -0.1	8 0.2	9 -0.2	2 0.3	6 0.1	1 -0.23	3 0.02	0.15	0.21	-0.18	-0.32
RR											0	.13 0.	04 0.	13 0.	12 -0.3	0 -0.3	0.1	9 0.1	5 0.0	2 0.0′	7 0.21	1 -0.45	-0.13	-0.23	0.26	-0.25
MAP												0.	22 0 .	94 0.	94 -0.4	7 -0.1	5 0.2	8 0.1	7 0.0	7 0.0	5 -0.35	3 0.31	0.02	0.05	-0.24	0.11
PCI													0.	49 0.	47 0.0	2 0.5	6 -0.3	5 0.0	2 -0.0	1 -0.1	0.02	2 0.22	2 -0.43	-0.24	0.01	-0.07
MF														0.	99 -0.4	2 0.0	4 0.1	2 0.1	7 0.0	8 0.0	3 -0.27	0.33	9 -0.10	-0.06	-0.17	0.13
Rf															-0.4	2 0.0	3 0.1	4 0.2	0.0	7 0.0	4 -0.28	8 0.33	9 -0.11	-0.05	-0.18	0.15
Ar																0.2	-0.5	0 -0.0	8 -0.1	6 0.2	2 0.44	4 -0.17	-0.12	-0.17	0.31	0.06
OlAlm																	-0.5	5 -0.0	5 0.0	5 -0.0	9 0.20	0.02	2 -0.31	-0.25	0.18	0.09
For																		0.1	2 -0.1	9 0.27	2 -0.34	4 -0.08	0.22	0.18	-0.25	-0.07
ΡV																			-0.1	6 0.27	2 0.02	2 -0.08	0.21	-0.11	-0.01	0.16
Mat																				-0.0	3 0.01	0.15	-0.09	0.19	0.01	-0.09
Lime																					0.13	3 -0.25	0.13	-0.12	0.17	-0.05
Marl																						-0.54	t -0.15	-0.64	0.91	0.09
Argil																							-0.01	0.43	-0.52	0.00
Meta																								0.16	-0.06	0.10
Med																									-0.59	-0.15
Fine																										0.18
SY abso mean s	olute s lope g	edimen	t yield (%), H	(t a ⁻¹),	<i>SSY</i> area imum a	1-specific	: sedime m), H _m	ant yield _{ax} maxi	l (t km ⁻ mum a	² a ⁻¹), . ltitude (4 catchi m), HI	nent are hypsorr	a (km ²) tetric in	, <i>Peri</i> p tegral (erimeter –), <i>RR</i>	(km), <i>C</i> relief ra	circle 1 tio (m	atio (-) km ⁻²),	, SL sti MAP	eam len mean a	.gth (km nnual p	ı), <i>DD</i> d	lrainage tion (m	density m), <i>PC</i>	(km k I Preci	m^{-2}), S pitation
Concen PV perc	tration entage	i Index (e poorly	%), <i>MF</i> vegetat	⁷ Modifi ed areas	ed Fourn , <i>Mat</i> pe	nier Inder rcentage	x (mm), Matorr	, <i>Rf</i> rain al, <i>Lime</i>	fall eros percen	iivity fa tage lim	ctor (MJ estone, .	mm m <i>Marl</i> per	² h ⁻¹ a centage	⁻¹), <i>Ar</i>] Marls,	percenta Argil pe	ge arable rcentage	e land, (Argilli	<i>JIAIm</i> p e rock,	ercenta Meta p	ge olive ercentag	e and alr ge of me	nond or etamorp	chards, hic rock	For per, Med p	centago	e forest, ge with
mediun	n soil t	texture,	Fine pe	rcentage	with fi	ne soil te	exture, J	Fluv pei	rcentage	Eluvise	slo															

Correlation statistics in *italic and bold* are significant at 1%; those underlined at 5%



Fig. 2 Scatter plots to illustrate some of the relations found between environmental variables and sediment yield

correlation with SSY and with slope gradient. In other words, large catchments have lower mean slope gradients and lower SSY than small catchments. All climate variables show a positive correlation with SSY, stressing the importance of rainfall amount and seasonality. Perhaps, counterintuitively, a negative correlation was observed between mean annual precipitation and drainage density. Less surprising is that most arable land appears to be on less steep slopes, while on steep slopes, high percentages of forest are found. Although the correlation matrix shows relatively high correlations, the scatter plots of Fig. 2 reveal ambiguous relations between SSY and most land cover and lithology classes, in part due to the many zero observations for these classes. The clearest positive correlations were found between the percentage of poorly vegetated areas (PV) and both SSY and SY, and the percentage olive and almond orchards (OlAlm) and SSY (see Fig. 2 and Table 3).

3.2 Multiple linear regression equations

Because of the lognormal distribution of SY and SSY, the natural logarithm of area-specific sediment yield (lnSSY) and absolute sediment yield (lnSY) were used as dependent variables in multiple linear regression analysis using stepwise forward and backward selection. The models with the highest explanatory value for lnSSY and lnSY respectively are:

$$\ln SSY = 1.353 + 0.065*S + 0.108*RR + 0.337*PCI - 0.026*Mat - 0.013*Lime - 0.041*Meta + 0.038*Fluv$$
(8)

 $(R^2=0.55; Adj-R^2=0.49; ME=0.55, RRMSE=0.14; p$ value<0.0001; n=61)

$$\ln SY = 10.18 + 0.049^{*}S + 0.011^{*}MF$$

$$-0.038$$
*Meta $+0.041$ *Fluv $+0.003$ *Peri (9)

 $(R^2=0.57; Adj-R^2=0.50; ME=0.57, RRMSE=0.07; p$ value<0.0001; n=61)

Although there is a relatively high correlation between slope gradient (S) and the Relief Ratio (RR; see Table 3), there is no problem of collinearity in Eq. 8 since the tolerance and variance inflation factor are within acceptable ranges (SAS 1999). The relatively good calibration results are, however, strongly influenced by a few observations as is indicated by the studentised residual of the predictions (i.e. studentised residual >2 or $\langle -2 \rangle$). Based on this criterion, the El Gergal catchment was identified as an outlier for SSY, and the catchments of Embarcaderos, Riudecañas and Oliana were considered as outliers for SY (Fig. 3). For the Oliana and El Gergal catchments, no explanation was found for the high residual of the predictions, whereas Embarcaderos and Riudecañas represent by far the largest and smallest catchments of the database, respectively, which may be an explanation as to why they appear as outliers. The models were calibrated again without the outliers using a stepwise regression with the same variables, resulting in the following equations for lnSSY and lnSY respectively:

$$\ln SSY = 1.891 + 0.067*S + 0.110*RR + 0.278*PCI - 0.029*Mat - 0.012*Lime - 0.042*Meta + 0.039*Fluv$$
(10)

 $(R^2=0.58; Adj-R^2=0.52; ME=0.58; RRMSE=0.13; p \text{ value} < 0.0001; n=60)$

$$\ln SY = 6.768 + 0.055 * S + 0.382 * PCI$$

- 0.025 * Mat - 0.012 * Lime
+ 0.059 * Fluv + 0.005 * Peri (11)

 $(R^2=0.58; Adj-R^2=0.53; ME=0.58; RRMSE=0.06; p \text{ value} < 0.0001; n=58)$

Again, the tolerance and variance inflation factor indicate that there is no collinearity problem in these equations, so all parameters were maintained.

The main part of variance in SSY is explained by the Relief Ratio (RR), both in the equation with and without outliers (Table 4). All other parameters explain much less of the variance in SSY, with lowest contribution from the percentage of Fluvisols (Fluv), which is also the least significant model parameter, with and without outliers. For SY, the main part of variance is explained by the catchment perimeter (Peri), followed by the Modified Fournier Index (MF). Least variance in SY is explained by the percentage of olive and almond orchards (OlAlm), which together with the percentage of Fluvisols (Fluv) are the least significant model parameters. Without outliers, Peri is the single most significant parameter explaining by far the main part of variance. In the model without outliers, the MF is substituted by the PCI, but with a much lower contribution to the explained variance in SY.

Calibration results of both models were only slightly better without outliers than with all observations (Table 5 and Fig. 3). Therefore, it was decided to validate the models both with and without the outliers using the "Jackknife" procedure as described above (Shao and Tu 1995). When validated for all 61 observations, the R^2 and the ME decrease dramatically compared to the calibration for both SY and SSY (Table 5 and Fig. 4). There is a large scatter around the 1:1 line, and the ME for both models is



Fig. 3 Calibration results of the multiple linear regression models. Presented are the natural logarithm of predicted and observed area-specific sediment yield (SSY) and absolute sediment yield (SY) with all observations (*upper*) and after removal of outliers (*lower*)

negative, meaning that the model produces more variation than is present in the measured values. Without the outliers, the validation results are better but still far worse than the calibration results. Validation results for SY are slightly better than for SSY, both with and without outliers. For comparison, it was also tested if a log transformation of those variables with a lognormal distribution included in the analysis would provide different results for the regression analysis. Although prediction equations were slightly different, validation results were very similar to those described above. Since normality of explanatory variables is not required for regression analysis, and to maintain the focus of this paper, we do not present all details of this exercise here.

3.3 The ART sediment delivery model

The ART model (Syvitski et al. 2003, 2005) was applied directly for prediction of sediment yield in the 61 Spanish

catchments. Table 6 and Fig. 5 show an evaluation of the quality of the predictions by comparison of predicted and observed sediment yield. Based on the studentised residual of the predictions with the ART model, several outliers were identified, and so, the model was also applied without these outliers (see Table 6 and Fig. 5). For SY, the catchments of Embarcaderos, Cijara and Doña Aldonza appeared as outliers, while for SSY, the outliers were Riudecañas, La Cierva, Conde de Guadalhorce, and El Gergal.

Both with and without outliers, the R^2 of the ART model predictions is relatively high. However, for all ART model predictions, the ME is very low, and the RRMSE is more than 1, meaning that predictions strongly deviate from the line of unity. The model generally overestimates sediment yield (SSY and SY), especially for the lower sediment yield values. Validation results for SSY are better than for SY, both with and without outliers. For SSY, removal of outliers results in a higher R^2 and ME, whereas for SY, removal of outliers results in a higher R^2 but much lower ME.

Table 4 Summary of statistics per variable for the regression equations for area-specific sediment yield (SSY) and absolute sediment yield (SY)

Variable	Partial R^2	Standard error	p value	Variable	Partial R^2	Standard error	p value
SSY	Full model (n	=61)		No outliers (<i>n</i> =60)		
Intercept		1.291	0.299	Intercept		1.264	0.141
S	0.051	0.017	< 0.001	S	0.032	0.017	< 0.001
RR	0.154	0.032	0.001	RR	0.165	0.031	< 0.001
PCI	0.031	0.123	0.008	PCI	0.037	0.121	0.026
Mat	0.061	0.009	0.005	Mat	0.077	0.009	0.002
Lime	0.038	0.005	0.010	Lime	0.065	0.005	0.011
Meta	0.080	0.010	< 0.001	Meta	0.081	0.010	< 0.001
Fluv	0.025	0.022	0.091	Fluv	0.028	0.021	0.071
SY	Full model (n	=61)		No outliers (n=58)		
Intercept		0.505	< 0.001	Intercept		1.283	< 0.001
S	0.04	0.020	0.016	S	0.023	0.016	0.001
MF	0.13	0.006	0.066	PCI	0.022	0.116	0.002
OlAlm	0.02	0.016	0.119	Mat	0.045	0.008	0.005
Mat	0.05	0.009	0.005	Lime	0.024	0.005	0.014
Lime	0.03	0.005	0.016	Fluv	0.061	0.021	0.007
Meta	0.06	0.011	< 0.001	Peri	0.258	0.001	< 0.001
Fluv	0.04	0.024	0.091				
Peri	0.21	0.001	< 0.001				

Indicated are the results with all observations (n=61) and after removal of outliers (n=60 or 58)

4 Discussion

Based on the calibration of the linear regression equations, it seems that predictions of absolute (SY) and area-specific sediment yield (SSY) are reasonably accurate. There is strong agreement in the environmental factors that appear in the equation for prediction of SY and SSY, and model performances are also comparable. Both regression equations represent at least one variable related to topography, climate, vegetation, lithology and soil characteristics, as such representing a range of environmental factors. Most significant model parameters were related to topographic relief (RR for SSY), catchment size and annual rainfall distribution (Peri and MF for SY). On the other hand, the poor validation results imply that the models are very unstable and are not useful for extrapolation to other catchments, although after removal of the outliers the validation results became somewhat better. This stresses the importance of performing a model validation and not trusting only calibration results, as is often done in regression studies. This was also confirmed by Verstraeten and Poesen (2001) who found that good regressions in calibration do not always imply good model validation results. Likewise, Grauso et al. (2008) reported that application of regression equations developed for a range of Italian catchments did not yield accurate results for

Table 5 Summary of calibration and validation results of the regression equations for area-specific sediment yield (SSY; t km⁻² a⁻¹) and absolute sediment yield (SY; t a⁻¹)

	Calibration	1			Jackknife v	validation		
	SSY	SY	SSY	SY	SSY	SY	SSY	SY
Equation nr.	8	9	10	11	_	_	_	_
R^2 pred-obs	0.55	0.57	0.58	0.58	0.03	0.07	0.17	0.25
ME	0.55	0.57	0.58	0.58	-0.63	-0.22	-0.10	0.17
RRMSE	0.14	0.07	0.13	0.06	0.27	0.11	0.22	0.09
n	61	61	60	58	61	61	60	58

 R^2 pred-obs R^2 between predicted and observed values, ME model efficiency, RRMSE relative root mean square error, n number of observations included in the analysis

Indicated are the results with all observations (n=61) and after removal of outliers (n=60 or 58)



Fig. 4 Jackknife validation results of the multiple linear regression models. Presented are the natural logarithm of predicted and observed areaspecific sediment yield (SSY) and absolute sediment yield (SY) with all observations (*upper*) and after removal of outliers (*lower*)

prediction of sediment yield in 16 Sicilian catchments. Yet, still most studies presenting a regression equation include a correlation coefficient for calibration but do not apply a proper validation procedure with independent data (e.g., Delmas et al. 2009; Grauso et al. 2008; Restrepo et al. 2006; Syvitski et al. 2003; Tamene et al. 2006; Verstraeten et al. 2003).

Notwithstanding the use of more variables representing a wider range of environmental characteristics for 61 catchments, the regression in this study was not more stable and did not perform better than the regression equation calibrated for a subset of 22 catchments by Verstraeten et al. (2003). Although fewer variables were used in their analysis, some very similar variables were selected in the stepwise regression of both studies (i.e. Relief Ratio, olive and almond orchards, % Matorral Forest, poorly vegetated areas). Yet, both the regression equations for the 22 and for the present 61 catchments performed worse than the semiquantitative Factorial Scoring Model (FSM) that was

applied to the same catchments and showed better calibration and validation results (i.e. $R^2=0.72$; ME=0.72; RRMSE=0.65; de Vente et al. 2005). FSM is an expert model where five factors (i.e. topography, lithology, vegetation cover, gullies, catchment shape) are used to characterise a watershed in the vicinity (~5 km) of the reservoir and the main tributaries by providing a score between 1 and 3. An index is calculated by multiplying the five scores, and together with catchment area, this index is used to obtain a regression equation for the prediction of SSY (de Vente et al. 2005, 2006; Verstraeten et al. 2003).

The ART model with only three variables that was calibrated for global rivers (Syvitski et al. 2003) provides much higher R^2 values between predicted and observed sediment yield (SSY and SY) of the Spanish catchments than the linear regression equations (Eqs. 8–11). However, the low ME and high RRMSE of the predictions with the ART model suggest that these are not necessarily better. For most catchments, the ART model strongly overpredicts

Table 6 Validation of the ART model (Syvitski et al. 2003) for a reaspecific sediment yield (SSY; t km⁻² a⁻¹) and absolute sediment yield (SY; t a⁻¹) as observed in the Spanish catchments

	Validation	ART model		
	SSY	SY	SSY	SY
R^2 pred-obs	0.38	0.17	0.54	0.31
ME	-0.22	-0.32	0.04	-2.62
RRMSE	1.29	2.30	1.18	1.84
n	61	61	57	58

 R^2 pred-obs R^2 between predicted and observed values, *ME* model efficiency, *RRMSE* relative root mean square error, *n* number of observations included in the analysis

Indicated are the results with all observations (n=61) and after removal of outliers (n=57 or 58)

sediment yield. According to Syvitski et al. (2003, 2005), the overestimation of sediment yield can be explained by the fact that reservoirs trap substantial volumes of sediments and therefore reduce the expected sediment yield for undisturbed conditions. For the Spanish reservoirs, this might indeed explain part of the overestimated sediment yield by the ART model, because there are reservoirs upstream of most of the reservoirs used for the present study. On the other hand, important errors can also be due to factors that are not considered by the model such as human impact (including conservation measures), lithology (de Vente et al. 2005; Woodward 1995) tectonics and the rate of rock uplift (Hovius 1998). Part of these issues may be solved by the BQART model, a successor of the ART model that integrates factors on river discharge (Q), lithology and human impacts, including dams and erosion control measures (B; Syvitski and Milliman 2007). Yet, reliable and detailed information on these factors is often difficult to obtain. It should also be remembered that the ART model was calibrated mostly for large global river basins worldwide. Calibration of the ART model for the Spanish data, similar to that done for the other regression equations, may give much better results.

Although some additional variables may be required, as well as a more regionalised calibration of model parame-



Fig. 5 Validation results of the ART model (Syvitski et al. 2003). Presented are the predicted and observed area-specific sediment yield (SSY) and absolute sediment yield (SY) with all observations (*upper*) and after removal of outliers (*lower*)

ters, it is certainly promising to observe that the ART model with only three variables has the potential for reasonable predictions of sediment yield at the global scale. Moreover, the relevance of these three variables is confirmed by the most significant model parameters in the multiple regression equations of this study. As in the ART model, factors related to topography, catchment size, and climate (RR, Peri, MFI) explained the main part of variation in sediment yield. The RR is a combination of the A and R factor of ART, whereas catchment perimeter is highly correlated with A (see Table 3). On the other hand, the ART model uses temperature instead of a rainfall-related factor. Syvitski et al. (2003) considered temperature as a more complete climate factor than rainfall alone, since temperature relates to the occurrence of convective, high-intensity rainstorms, but also to soil formation, soil erosion, runoff from melting snow, and freeze-thaw cycles (Harrison 2000; Syvitski et al. 2003). The role of mean annual temperature (Tm) was, however, not confirmed by our regression analysis, maybe because the differences in temperature between the Spanish study catchments were not large enough.

There are various possible explanations for the unstable model performance of the linear regression equations. First of all, land use information does not necessarily represent the percentage soil cover by vegetation, whereas the latter is especially important for erosion predictions (Casermeiro et al. 2004; Gyssels et al. 2005; Molina et al. 2007). Secondly, due to land use changes, land use as indicated on the CORINE Land Cover map (CLC2000) may not be representative for land cover conditions during the period for which the measured reservoir sedimentation rates are valid. Large-scale reforestations and extensions of olive and almond orchards that have taken place in the last decades (Cammeraat and Imeson 1999; García-Ruiz 2010; Poesen et al. 1997; Rojo Serrano 2003; van Wesemael et al. 2003) may provide a false view of the soil protection by vegetation during the lifetime of the reservoir. Furthermore, due to the low resolution of the soil and lithology maps, important detail is probably missed, and so, a higher resolution may increase the explanatory value of these variables. Next, drainage density is a typical example of a scale-dependent variable (Abrahams 1984; Vogt et al. 2003). In this study, we automatically extracted a river network from a DEM, using a threshold of 0.2 km², and used this river network to calculate the drainage density. Surprisingly, a (nonsignificant, p value of 0.08) negative correlation was found between drainage density and SSY, and no correlation was observed with SY, which contrasts with various studies that considered drainage density to be strongly related to both sediment yield and mean annual precipitation (Abrahams 1984; Langbein and Schumm 1958). In agreement with our results, some studies reported a decreasing or complex correlation between drainage density and mean annual precipitation (Abrahams 1984), whereas others reported an increasing drainage density with increasing precipitation (Daniel 1981). Most studies explained the relation between mean annual precipitation and drainage density by differences in vegetation cover and runoff generation under diverse climatic conditions. In our analysis, however, drainage density did not appear as a significant variable in any of the multiple regression models (alpha 5%). A better result may be obtained when additional information on lithology, vegetation cover and climate is used to determine the threshold value in a stratified manner (Colombo et al. 2007; Vogt et al. 2003), or when drainage density is extracted from aerial photographs (Allen 1986).

Errors in the measured sediment yield and in the source data of the explanatory variables may also explain part of the poor model validation results. Furthermore, the variation in environmental conditions and in the dominant erosion and sediment transport processes between the 61 catchments may be too large to be explained by one single regression equation, or it may be that important explanatory variables were missing in the analysis. For example, mean annual runoff-related variables might explain an important part of mean annual sediment yield (Restrepo et al. 2006; Syvitski et al. 2003). Comparison of soil loss rates obtained from field plot studies with reservoir sedimentation rates in Mediterranean environments also suggests that gully erosion, mass movements and bank erosion are mainly responsible for sediment export at the catchment scale (de Vente et al. 2007, 2008; Poesen and Hooke 1997; VanMaercke et al. 2011). These processes operate at distinct locations in the landscape, and factors controlling these processes are difficult to parameterise by one single value for the entire catchment.

Many of the possible improvements to deal with the abovementioned problems require a major amount of additional processing and input of data that are often not easily available for large watersheds. Yet, at least equally important as data quality is the model concept used. Apart from different variables used for prediction, some models multiply factors (ART and FSM), while other models use the sum of variables (regression). Based on the current information, it seems that those models accounting for the interaction between environmental factors and for nonlinear relations through multiplying factors and by using power functions, such as the FSM and ART models, potentially perform better than linear models such as the linear regression models (de Vente 2009).

Although the linear regression model was not stable, the correlation analysis and regression equations do provide more insight into relevant relations between environmental factors and sediment yield, and increase our understanding of sediment dynamics at the catchment scale. This helps to identify important factors that constrain erosion and sediment transporting processes, and to select factors to be considered in more sophisticated model approaches. For example, the negative relation between catchment area and SSY that was found in this study is often reported in the literature, and is explained by the fact that during transport of sediments through a catchment, increasingly more sediment is trapped on footslopes, alluvial plains and other sinks, while erosion rates decrease due to decreasing hillslope gradients downstream (de Vente et al. 2007; Walling 1983). This theory is supported by the negative correlation between catchment area and mean hillslope gradient and the positive correlation between slope gradient and SSY observed in this study. Furthermore, our results suggest a positive correlation between mean annual precipitation and sediment yield, although the increase in sediment yield with annual rainfall is not as strong as was reported for Moroccan catchments (Woodward 1995). This uninterrupted, positive relation deviates from the relations suggested in earlier studies where SSY often first increases and then decreases with increasing mean annual rainfall (Langbein and Schumm 1958; Walling and Kleo 1979; Wilson 1973). This trend is still frequently referred to and was originally explained by low runoff rates and the absence of a protective vegetation cover in areas with very low precipitation amounts, and the presence of a vegetation cover protecting the soil beyond a certain threshold of annual precipitation (Langbein and Schumm 1958). The deviation from these trends in Spain and Morocco may be due to widespread human disturbances that impede a direct relation between rainfall amounts and vegetation cover that is expected in areas with a more natural vegetation cover.

For practical applications and to support policy decisions in a given region, regression equations form an attractive option because they are relatively easily applicable and it is tempting to apply an existing equation. However, the results of this study confirm that the use of regression equations beyond the region and conditions for which the model was developed is not a guarantee of good results. Furthermore, the regression analysis is based on lumped variables, which allows prediction of sediment yield but does not identify where the sediments originate, where erosion problems are most severe, or which erosion and sediment transport processes are dominant. To answer these questions, other spatially distributed approaches are required.

5 Conclusions

Correlation analysis reveals that, for 61 Spanish catchments, drainage area, catchment perimeter, stream length, relief ratio, the Modified Fournier Index and the percentage poorly vegetated areas have the highest correlation coefficients with area-specific (SSY) and absolute (SY) sediment yield. However, a combination of topographic, climatic, lithologic, vegetation and soil variables is required to provide the best linear regression equation to predict sediment yield. Although calibration results after a stepwise regression are reasonably good, validation of these equations shows that they are very unstable and cannot be used for extrapolation. On the other hand, application of the ART sediment delivery model (Syvitski et al. 2003) for prediction of sediment yield in the Spanish catchments shows a relatively high correlation between predicted and observed values, although predictions systematically overestimate observed sediment yield. Beside measurement errors in sediment yield and in the explanatory variables, there are various explanations for the poor model validation results in general: (1) not all controlling erosion and sediment transport processes are well represented by the available data; (2) land use is not representative of vegetation cover; (3) the land use, lithology and soil data are of a too low resolution; (4) land use is not a static property and will have changed during the lifetime of the reservoirs; and (5) nonlinear relations between sediment yield and environmental variables, and nonlinear interactions between those variables are not (sufficiently) accounted for by the models.

Acknowledgements The research described in this paper was conducted in the framework of the EC-DG RTD 6th Framework Research Programme (sub-priority 1.1.6.3)—Research on Desertification—project DESIRE (037046): Desertification Mitigation and Remediation of Land—a global approach for local solutions. We thank Carles Balasch and an anonymous reviewer for their constructive and helpful comments that helped us to improve this paper significantly.

References

- Abrahams AD (1984) Channel networks: a geomorphological perspective. Water Resour Res 20:161–188
- Achite M, Touaibia B (2000) Analyse multivariée de la variable 'érosion spécifique': cas du bassin versant de l'Oued Mina (Wilaya de Relizane -Algérie). In: Séminaire International Montpellier 2000 Hydrologie des régions méditerranéennes. Programme Hydrologique International, Documents Techniques en Hydrologie 51, pp 119–127
- Ali FK, de Boer DH (2008) Factors controlling specific sediment yield in the upper Indus River basin, northern Pakistan. Hydrol Process 22:3102–3114
- Allen PB (1986) Drainage density versus runoff and sediment yield. In: Fourth Federal Interagency Sedimentation Conference, Las Vegas. USGS, pp. 3–38, 33–44
- Arnoldus HMJ (1977) Methodology used to determine the maximum potential average annual soil loss due to sheet and rill erosion in Morocco. FAO Soils Bulletin
- Avendaño Salas C, Cobo Rayán R (1997) Metodología para estimar la erosion de cuencas fluviales a partir de la batimetría de embalses.

In: Ibáñez JJ, Valero Garcés BL, Machado C (eds) El paisaje mediterráneo a través del espacio y del tiempo. Implicaciones en la desertificación. Logroño. Geoforma Ediciones, pp 239–257

- Avendaño Salas C, Cobo Rayán R, Gómez Montaña JL, Sanz Montero E (1995) Procedimiento para evaluar la degradación específica (erosión) de cuencas de embalses a partir de los sedimentos acumulados en los mismos. Aplicación al estudio de embalses Españoles. Ingeniería Civil 99:51–58
- Avendaño Salas C, Cobo Rayán R, Sanz Montero E, Gómez Montaña JL (1997a) Capacity situation in Spanish reservoirs. In: Dixneuvième Congrès des Grands Barrages, Florence. Commission Internationale Des Grands Barrages, pp 849–862
- Avendaño Salas C, Sanz Montero E, Gómez Montaña JL (1997b) Sediment yield at Spanish reservoirs and its relationship with the drainage basin area. In: Dix-neuvième Congrès des Grands Barrages, Florence. Commission Internationale De Grands Barrages, pp 863–874
- Bazzoffi P, Baldassarre G, Pellegrini S, Bassignana A (1997) Models for prediction of water storage decrease in Italian Reservoirs. In: Shady AM, Kassem A, Delise CE, Bouchard MA (eds) Proceedings of the IX World Water Congress IWRA, 1997. pp 249–252
- Cammeraat LH, Imeson AC (1999) The evolution and significance of soil-vegetation patterns following land abandonment and fire in Spain. Catena 37:107–127
- Casermeiro MA, Molina JA, de la Cruz Caravaca MT, Hernando Costa J, Hernando Massanet MI, Moreno PS (2004) Influence of scrubs on runoff and sediment loss in soils of Mediterranean climate. Catena 57:91–107
- CEDEX (1992) Informe sobre la situación de los estudios de aforo de solidos, batimetria y sedimentologia de embalses hasta 1992. Centro de Estudios Hidrográficos del CEDEX, Madrid
- Clark EH (1985) The off-site costs of soil erosion. J Soil Water Conserv 40:19–22
- Colombo R, Vogt JV, Soille P, Paracchini ML, de Jager A (2007) Deriving river networks and catchments at the European scale from medium resolution digital elevation data. Catena 70:296– 305
- Daniel JRK (1981) Drainage density as an index of climatic geomorphology. J Hydrol 50:147–154
- de Roo APJ (1998) Modelling runoff and sediment transport in catchments using GIS. Hydrol Process 12:905–922
- de Roo APJ, Wesseling CG, Ritsema CJ (1996) LISEM: a single event physically-based hydrologic and soil erosion model for drainage basins: I Theory, input and output. Hydrol Process 10:1107–1117
- de Vente J (2009) Soil erosion and sediment yield in Mediterranean geoecosystems. Scale issues, modelling and understanding. PhD thesis, K.U.Leuven, Leuven, Belgium
- de Vente J, Poesen J, Verstraeten G (2005) The application of semiquantitative methods and reservoir sedimentation rates for the prediction of basin sediment yield in Spain. J Hydrol 305:63–86
- de Vente J, Poesen J, Bazzoffi P, Van Rompaey A, Verstraeten G (2006) Predicting catchment sediment yield in Mediterranean environments: the importance of sediment sources and connectivity in Italian drainage basins. Earth Surf Process Landf 31:1017–1034
- de Vente J, Poesen J, Arabkhedri M, Verstraeten G (2007) The sediment delivery problem revisited. Prog Phys Geog 31:155– 178
- de Vente J, Poesen J, Verstraeten G, Van Rompaey A, Govers G (2008) Spatially distributed modelling of soil erosion and sediment yield at regional scales in Spain. Global Planet Change 60:393–415
- Delmas M, Cerdan O, Mouchel J-M, Garcin M (2009) A method for developing a large-scale sediment yield index for European river basins. J Soils Sediments 9:613–626

- Demmak A (1984) Recherche d'une relation empirique entre apports solides spécifiques et paramètres physico-climatiques des bassins: application au cas algérien. In: Walling DE, Foster SSD, Wurzel P (eds) Challenges in African hydrology and water resources, Harare. IAHS publication No. 144, pp 403–414
- Dendy FE, Bolton GC (1976) Sediment yield-runoff-drainage area relationships in the United States. J Soil Water Conserv 31:264– 266
- Eastman JR (2003) IDRISI Kilimanjaro guide to GIS and image processing; Manual Version, 14th edn. Clark Labs, Worcester
- EEA (2000) CORINE land cover 2000. European Environment Agency http://image2000.jrc.it
- ESB (2004) European Soil Database. European Soil Bureau Network and the European Commission, EUR 19945 EN, http://eusoils. jrc.it
- Flanagan DC, Ascough JC II, Nicks AD, Nearing MA, Lafle JM (1995) Overview of the WEPP erosion prediction model. In: Flanagan DC, Nearing MA (eds) USDA-Water Erosion Prediction Project. Hillslope profile and watershed model documentation, vol NSERL Report 10. USDA, Indiana, pp 1.1–1.12
- Flaxman EM (1972) Predicting sediment yield in western United States. J Hydraul Div ASCE 98(12):2073–2085
- Foster GR, Flanagan DC, Nearing MA, Lane LJ, Risse LM, Finker SC (1995) Hillslope erosion component. In: Flanagan DC, Nearing MA (eds) USDA-water erosion prediction project. Hillslope profile and watershed model documentation, vol NSERL Report 10. USDA, Indiana, pp 11.11–11.12
- Fournier F (1960) Climat et érosion: la relation entre l'érosion du sol par l'eau et les précipitations atmosphériques. Presses Universitaires de France, Paris
- García-Ruiz JM (2010) The effects of land uses on soil erosion in Spain: a review. Catena 81:1–11
- Grauso S, Pagano A, Fattoruso G, De Bonis P, Onori F, Regina P, Tebano C (2008) Relations between climatic–geomorphological parameters and sediment yield in a mediterranean semi-arid area (Sicily, southern Italy). Environ Geol 54:219–234
- Gyssels G, Poesen J, Bochet E, Li Y (2005) Impact of plant roots on the resistance of soils to erosion by water: a review. Prog Phys Geog 29:189–217
- Hadley RF, Lal R, Onstad CA, Walling DE, Yair A (1985) Recent developments in erosion and sediment yield studies. International Hydrological Programme (IHP). UNESCO, Paris
- Haregeweyn N, Poesen J, Nyssen J, Verstraeten G, de Vente J, Govers G, Deckers S, Moeyersons J (2005) Specific sediment yield in Tigray-Northern Ethiopia: assessment and semi-quantitative modelling. Geomorphology 69:315–331
- Harrison CGA (2000) What factors control mechanical erosion rates? Int J Earth Sci 88:752–763
- Hovius N (1998) Control of sediment supply by large rivers. In: Relative role of Eustasy, climate and tectonism in continental rocks. SEPM Special Publication, pp 3–16
- Ichim I (1990) The relationship between sediment delivery ratio and stream order: a Romanian case study. In: Walling DE, Yair A, Berkovicz S (eds) Erosion, transport and deposition processes, Jerusalem, 1990. IAHS publication 189, pp 79–86
- Jansson MB (1982) Land erosion by water in different climates. Ph.D thesis, UNGI report No. 57, Uppsala University, Sweden
- Kirkby MJ, Jones RJA, Irvine B, Gobin A, Govers G, Cerdan O, Van Rompaey AJJ, Le Bissonnais Y, Daroussin J, King D, Montanarella L, Grimm M, Vieillefont V, Puigdefabregas J, Boer M, Kosmas C, Yassoglou N, Tsara M, Mantel S, Van Lynden G (2004) Pan-European soil erosion risk assessment: the PESERA Map, Version 1 October 2003. Explanation of Special Publication. No.73 (S.P.I.04.73). European Soil Bureau Research Report No.16, EUR 21176, Office for Official Publications of the European Communities, Luxembourg

- Langbein WB, Schumm SA (1958) Yield of sediment in relation to mean annual precipitation. Trans Amer Geophys Union 39:1076– 1084
- Lixian W, Lida S, Xiying H, Boging Z, Qingyu W, Guiyun L (1996) Application of regression analysis to the prediction of average annual yield of silt. In: Lixian W (ed) Combating desertification in China. China Forestry Publishing House, pp 210–221
- Ludwig W, Probst J-L (1998) River sediment discharge to the oceans; present-day controls and global budgets. Am J Sci 298:265–295
- Ludwig W, Probst J-L, Kempe S (1996) Predicting the oceanic input of organic carbon by continental erosion. Global Biogeochem Cycles 10:23–41
- Michiels P, Gabriels D, Hartmann R (1992) Using the seasonal and temporal precipitation concentration index for characterizing the monthly rainfall distribution in Spain. Catena 19:43–58
- Molina A, Govers G, Vanacker V, Poesen J, Zeelmaekers E, Cisneros F (2007) Runoff generation in a degraded Andean ecosystem: interaction of vegetation cover and land use. Catena 71:357–370
- Morgan RPC, Quinton JN, Smith RE, Govers G, Poesen J, Auerswald K, Chisci G, Torri D, Styczen ME (1998) The European Soil Erosion Model (EUROSEM): a dynamic approach for predicting sediment transport from fields and small catchments. Earth Surf Process Land 23:527–544
- Nash JE, Sutcliffe JV (1970) River flow forecasting through conceptual models part I—a discussion of principles. J Hydrol 10:282–290
- Oliver JE (1980) Monthly precipitation distribution: a comparative index. Profess Geographer 32:300–309
- Owens PN, Batalla RJ, Collins AJ, Gomez B, Hicks DM, Horowitz AJ, Kondolf GM, Marden M, Page MJ, Peacock DH, Petticrew EL, Salomons W, Trustrum NA (2005) Fine-grained sediment in river systems: environmental significance and management issues. River Res Appl 21:693–717
- Poesen J, Hooke JM (1997) Erosion, flooding and channel management in Mediterranean environments of Southern Europe. Prog Phys Geog 21:157–199
- Poesen J, van Wesemael B, Govers G, Martinez-Fernandez J, Desmet P, Vandaele K, Quine T, Degraer G (1997) Patterns of rock fragment cover generated by tillage erosion. Geomorphology 18:183–197
- Probst JL, Amiotte-Suchet P (1992) Fluvial suspended sediment transport and mechanical erosion in the Maghreb (North Africa). Hydrol Sci J 37:621–637
- Rabus B, Eineder M, Roth A, Bamler R (2003) The Shuttle Radar Topography Mission—a new class of digital elevation models acquired by spaceborne radar. ISPRS J Photogramm 57:241–262
- Ramos MC, Martinez-Casasnovas JA (2004) Nutrient losses from a vineyard soil in northeastern Spain caused by an extraordinary rainfall event. Catena 55:79–90
- Renard KG, Freimund JR (1994) Using monthly precipitation data to estimate the R-factor in the revised USLE. J Hydrol 157:287–306
- Restrepo JD, Kjerfve B, Hermelin M, Restrepo JC (2006) Factors controlling sediment yield in a major South American drainage basin: the Magdalena River, Colombia. J Hydrol 316:213–232
- Roberts PJT (1973) A method of estimating mean annual sediment yields in ungauged catchments. Technical Note No. 44 Department of Water Affairs South Africa, Division of Hydrology
- Rojo Serrano L (2003) Description of the land use changes in relation to type of crops, landscapes and physiographic units involved in the Guadalentin basin. Deliverable 3.3a. DESERTLINKS, EVK2-CT-2001-00109 www.kcl.ac.uk/projects/desertlinks

- SAS (1999) SAS Online Doc version 8. SAS Institute, Cary, NC, USA. http://v8doc.sas.com/sashtml/
- Schoorl JM, Veldkamp A (2001) Linking land use and landscape process modelling: a case study for the Alora region (south Spain). Agr Ecosyst Environ 85:281–292
- Shao JX, Tu D (1995) The Jackknife and Bootstrap. Springer, New York
- Steegen A, Govers G, Takken I, Nachtergaele J, Poesen J, Merckx R (2001) Factors controlling sediment and phosphorus export from two Belgian agricultural catchments. J Environ Qual 30:1249– 1258
- Syvitski JPM, Kettner AJ (2007) On the flux of water and sediment into the Northern Adriatic Sea. Cont Shelf Res 27:296–308
- Syvitski JPM, Milliman JD (2007) Geology, geography, and humans battle for dominance over the delivery of fluvial sediment to the coastal ocean. J Geol 115:1–19
- Syvitski JPM, Peckham SD, Hilberman R, Mulder T (2003) Predicting the terrestrial flux of sediment to the global ocean: a planetary perspective. Sediment Geol 162:5–24
- Syvitski JPM, Vörösmarty CJ, Kettner AJ, Green P (2005) Impact of humans on the flux of terrestrial sediment to the global coastal ocean. Science 308:376–380
- Tamene L, Park SJ, Dikau R, Vlek PLG (2006) Analysis of factors determining sediment yield variability in the highlands of northern Ethiopia. Geomorphology 76:76–91
- UNEP (1997) World atlas of desertification. United Nations Environment Programme
- Vanmaercke M, Poesen J, Maetens W, de Vente J, Verstraeten G (2011) Sediment yield as a desertification risk indicator. Science of the Total Environment 409 (9):1715–1725
- van Wesemael B, Cammeraat E, Mulligan M, Burke S (2003) The impact of soil properties and topography on drought vulnerability of rainfed cropping systems in southern Spain. Agr Ecosys Environ 94:1–15
- Verstraeten G, Poesen J (1999) The nature of small-scale flooding, muddy floods and retention pond sedimentation in central Belgium. Geomorphology 29:275–292
- Verstraeten G, Poesen J (2001) Factors controlling sediment yield from small intensively cultivated catchments in a temperate humid climate. Geomorphology 40:123–144
- Verstraeten G, Poesen J, de Vente J, Koninckx X (2003) Sediment yield variability in Spain: a quantitative and semiqualitative analysis using reservoir sedimentation rates. Geomorphology 50:327–348
- Vogt JV, Colombo R, Bertolo F (2003) Deriving drainage networks and catchment boundaries: a new methodology combining digital elevation data and environmental characteristics. Geomorphology 53:281–298
- Walling DE (1983) The sediment delivery problem. J Hydrol 65:209– 237
- Walling DE, Kleo AHA (1979) Sediment yields of rivers in areas of low precipitation: a global view. In: IAHS (ed) The hydrology of areas of low precipitation, Canberra, 1979. IAHS publication 128, pp 479– 493
- WCD (2000) Dams and development. A new framework for decisionmaking. Earthscan Publications Ltd., London
- Wilson L (1973) Variations in mean annual sediment yield as a function of mean annual precipitation. Am J Sci 273:335–349
- Woodward JC (1995) Patterns of erosion and suspended sediment yield in Mediterranean river basins. In: Foster IDL, Gurnell AM, Webb BW (eds) Sediment and water quality in river catchments. Wiley, Chichester, pp 365–389