OR Spektrum (2001) 23: 117–136

OR Spektrum

OLAF — A general modeling system to evaluate and optimize the location of an air polluting facility*

OLAF – ein Modellierungssystem zur Standortevaluierung und -optimierung einer luftverschmutzenden Anlage

Jörg Fliege

Institut für Angewandte Mathematik, Fachbereich Mathematik, Universität Dortmund, 44221 Dortmund, Germany (e-mail: fliege@math.uni-dortmund.de)

Abstract. We propose a general modeling and optimization system for the effects of a low toxic air pollutant on the inhabitants of different land-based ecosystems. The pollution load can be traced through different ecotrophic levels, while health effects are quantified on the basis of detailed cytodynamic effects in different organs and tissues. The software package OLAF (*O*ptimal *L*ocating *A*ir Polluting *F*acilities) developed on the basis of this model does not only evaluate quantitatively the effects of a proposed decision, but uses an efficient optimization technique to optimize the location of a polluting facility with respect to the adverse effects of the total pollution load on the inhabitants of the region under consideration.

Zusammenfassung. In dieser Arbeit wird ein allgemeines Modellierungs- und Optimierungssystem für einen schwach toxischen Luftschadstoff und seine Wirkung auf die Bewohner verschiedener landgestützter Ökosysteme vorgeschlagen. Dabei wird der Schadstoffeintrag durch verschiedene ökotrophe Stufen hindurch simuliert und gesundheitliche Effekte der Bewohner der betrachteten Ökosysteme auf der Basis detaillierter zytodynamischer Effekte in verschiedenen Organen und Geweben quantifiziert. Das auf der Basis dieses Systems entwickelte Softwarepaket OLAF (*O*ptimal *L*ocating *A*ir Polluting *F*acilities) quantifiziert nicht nur den Einfluß einer bestimmten Standortentscheidung, sondern optimiert den Standort einer schadstoffausstoßenden Anlage bezüglich der Negativeffekte auf die Bevölkerung der betrachteten Region.

* This work was sponsored by the Region Bruxelles Capital under a "Research in Brussels" grant. The author would hereby like to express his gratitude to Prof. Dr. F. Plastria from the Vrije Universiteit Brussel for the many fruitful discussions he enjoyed with him during his stay at the Vakgroep Beleidsinformatica.

Key words: Continuous facility location – Environmental planning – Health care – Nonlinear programming
Schlüsselwörter: Standortplanung – Umweltschutz – Gesundheitsvorsorge – Nichtlineare Optimierung

1 Introduction

Finding a location for an air polluting facility is usually a source of heated discussions among local inhabitants and decision makers, since the effects of polluted air on human health and the environment can be quite severe.

We propose a general modeling and optimization system for the effects of a low toxic air pollutant on the inhabitants of different land-based ecosystems, including health effects on humans or other beings. The pollution load can be traced through different ecotrophic levels, while health effects are quantified on the basis of detailed cytodynamic effects in different organs and tissues. The software package OLAF (*O*ptimal *L*ocating *A*ir Polluting *F*acilities) developed on the basis of this model does not only evaluate quantitatively the effects of a proposed decision, but uses an efficient optimization technique to optimize the location of the polluting facility with respect to the adverse effects of the pollution on the inhabitants of the region under consideration.

Up to now, the decision process with respect to environmental issues and health concerns was mainly based on simulations of a small number of different decisions. However, the need for an automated optimization process incorporating detailed estimates of effects of the pollution has recently been acknowledged by the European Environmental Agency [18,4]. Unfortunately, up to now only few attempts have been made to integrate the decision process into a general optimization procedure, thereby automatically minimizing a measure of risk. While some of the models developed with respect to this are rather simplistic [15,14,1,20], others use stochastic optimization techniques (see [16], [17] and the references therein), methods which are known for their notoriously slow behavior.

The general strategy of the task is clear: after a simulation module has been developed for simulating the environmental effects and the health effects for a given choice of parameters which characterize the polluting facility, an optimization module to be developed can handle the task of finding optimal choices for the describing parameter values.

However, it is quite obvious that a more detailed consideration of the problem at hand is necessary. To this end, let us first focus on two main questions.

1.1 Two main questions

1. What are the parameters of the facility that can be controlled by the decision makers?
2. How can the effects on the environment and on the health of certain individuals be quantified?

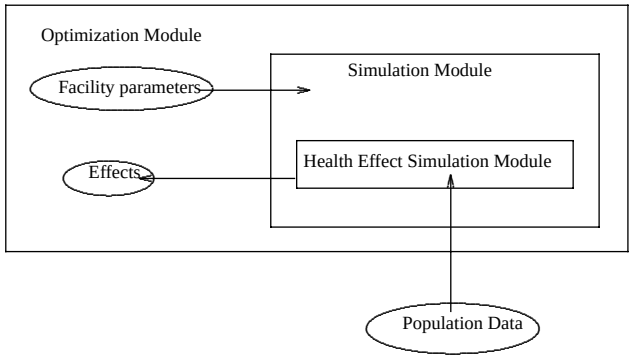

Fig. 1. General structure of the software developed including health effects

While there are no definite answers to these questions, we can give some recommendations, thereby providing at least partial answers.

1. Since we are concerned with locating a *new* facility, we can safely assume that the question ”Must the facility be built anyway?” has already been answered affirmatively. It remains to find an acceptable location for the facility. Therefore, the actual location is represented by two (real-valued) unknowns or parameters which have to be optimized. Moreover, since we are assuming that the main environmental effect of the facility is airborne pollution, we consider the height of the point of entry of the pollutant into the atmosphere (e. g. the height of the smokestack) as an additional unknown.
   We assume that production-dependent considerations have already fixed the amount of pollutant released in a given time period, and we assume that the amount released is constant in time. As a consequence, we are naturally led to the conclusion that all the physical and toxicological characteristics of the pollutant are prespecified. This leaves us with the width of the opening of the ”smokestack” (or, alternatively, with the speed of the exhaust gas) as one additional decision variable. For more details, we refer to Section 6.
2. On an individual level, a well-established measure for carcinogenic effects within a certain organ or tissue is the number of malignant cells appearing within a certain time period. Likewise, a measure for cytotoxic effects is the number of cells which have died during a fixed time period. Since we are dealing with low-toxic doses, actual outbreaks of symptoms are not an issue here. Taking now the position of a postulated central authority (i. e. a governmental health care official), *the total sum of all health effects over all individuals has to be minimized*. The same analysis holds with respect to environmental issues.
   More details with respect to the evaluation of the health effects can be found in Section 6.

The structure of the software system is depicted in Figure 1.

1.2 The submodules

Having given a more detailed account on the decision variables and the quantification of the effects to be minimized, we can now turn our attention to the simulation of the fate of a pollutant through the biosphere. It was chosen to develop a general modeling system to simulate the following subsystems of the environment and the fate of the pollutant in them:

1. atmosphere,
2. land-based ecosystems,
3. individuals, and
4. individual organs.

Of course, this list of subsystems is not exhaustive. We deliberately skipped the geosphere, the aquatic distribution of pollutants, the corresponding water-based ecosystems, the simulation of the effects of a pollutant on cells, etc.

Note that the meteorological conditions of the atmosphere are independent of the fate of the pollutant and of the location of the facility, but not vice versa. It makes therefore sense to calculate the meteorological variables needed over the time period of interest beforehand and use this data in subsequent simulation runs inside the optimization module.

The discussion above leads in a natural way to a breakdown of the software system into the following submodels and modules.

1. The meteorological preprocessor, which uses observational meteorological data to compute all the meteorological parameters needed for the time period of interest in the spatial domain considered,
2. the air dispersion model, in which the distribution of the pollutant in the atmosphere is simulated,
3. the ecosystem dispersion model, which simulates the fate of the pollutant in the ecosystems considered, through different trophic levels, through food chains, etc.,
4. the chemokinetic model, in which the concentration of the pollutant in different organs of different individuals, living in the abovementioned ecosystems, is traced,
5. the cytodynamic multistage model, which simulates the effects of the pollution on different cell populations in the organs of the individual life forms considered,
6. the effect quantification module, agglomerating the cytodynamic effects on the individuals to an effect at the population level, and finally
7. the optimization module, which minimizes this effect by optimizing the decision variables specified in Section 1.1.

The corresponding data flow between the submodules is depicted in Figure 2. In the following sections, each of these parts is shortly described. For more details, the reader is referred to the freely available project report [8].

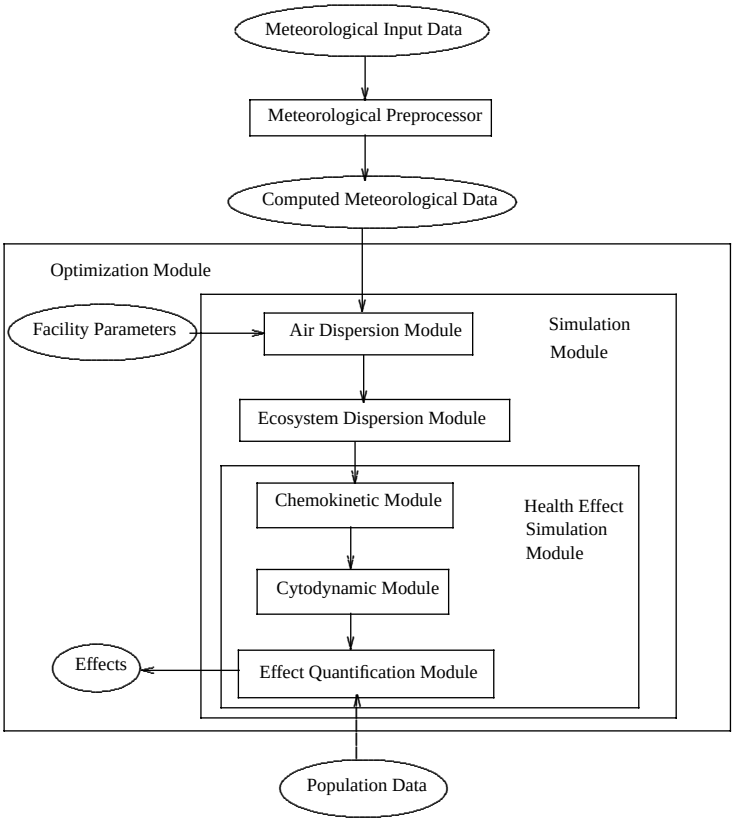

Fig. 2. Submodules (rectangles), main data (ellipses) and data flow (arrows)

2 The meteorological preprocessor

A meteorological preprocessor takes observational meteorological data from a given time period and a given spatial domain and computes out of this sparse data set all the meteorological information necessary to calculate pollutant concentrations in the air later on.

At the present moment, the preprocessor in use is a modified version of the MESOPAC II meteorological preprocessor [22,27], developed for the Environmental Protection Agency of the U.S.A. This preprocessor interpolates given meteorological data in time and space and produces a gridded field of data in two vertical layers of the atmosphere. This grid is the *computational grid*, whose grid points represent all the receptor locations of the pollutants. Within OLAF, each of these receptor locations contains an ecosystem and some individuals whose response on the pollution is modeled in other modules later on. More details with respect to the preprocessor can be found in the user's manual [27].

Since the meteorological data set provided by the preprocessor is independent of the parameters of the pollutant and its source, the preprocessor needs to be called only once during an optimization run. This strategy has been followed for

the design of the code. Subsequent computational tests showed significant savings in computation time.

3 The air dispersion module

The air dispersion module takes as input the data computed by the meteorological preprocessor, the locational characteristics of the polluting facility, and the physical and chemical characteristics of the pollutant, and computes out of this given data the concentration of the pollutant in the air at ground level within the time period considered. The amount of literature concerning solution methods to this problem is monumental, see, e. g. the review article [21], which lists more than 100 different algorithms.

The code used in OLAF to solve this problem is a modified version of the MESOPUFF II puff superposition code [22, 27, 6]. The MESOPUFF II code is able to simulate the dispersion of several different pollutants on a scale of ca. 1000km in west-east direction and ca. 600km in north-south direction. This makes the package suitable for medium to long-range transportation simulation. Simple atmospheric chemistry as well as wet and dry deposition processes are handled, too. All deposition takes place on the predefined Cartesian grid. Moreover, several pollutant sources including polluting areas can be present in the computational domain. In this way, a background pollution of the atmosphere can be modeled.

As it is the case for the preprocessor, the air dispersion code can be replaced by an arbitrary code capable of computing the desired values as long as the format of the input and the output stays the same. More details with respect to the dispersion model can be found in the user's manual [27] and in an application protocol [6].

4 Ecology and chemokinetics

Both the ecosystem dispersion model and the chemokinetic model are described in this section, since the modeling paradigm used for the two submodels is the same.

The ecosystem dispersion model takes as input the time- and space-dependent air concentration of the pollutant at the predefined grid points and follows the fate of the pollutant across the different trophic levels of the ecosystems considered. Each grid point of the air dispersion model represents one local food web, through which the trace of the pollutant is followed. The metabolism of a pollutant due to microbial degradation is modeled according to Monod kinetics.

The chemokinetic model takes as input the time- and space-dependent exposure of the individuals modeled to the pollutant under consideration and computes the uptake and the distribution of the chemical in the different body tissues and organs. Several different types of individuals can be simulated by specifying the corresponding parameters. Here, the metabolism of the pollutant under consideration is modeled by Michaelis-Menten kinetics, which is mathematically equivalent to Monod kinetics.

4.1 Compartment models

Both codes simulating the processes described above are based on a compartment model of the system to be simulated [13,26] (also called continuously stirred tank reactor (CSTR) model). Compartment models are well known in the literature and have been widely used to model pollutant transport and dispersion on an ecosystem level as well as on an individual level. These systems are of the following type, and one of these systems is set up for each grid point of the computational grid of the atmospheric dispersion model. The modeler divides the total system under consideration (i. e. the ecosphere and the geosphere) into $n \geq 1$ different *compartments*. Then, an exchange of pollution governed by first-order mass transfer equations takes place. Let $\Omega \subset \mathbb{R}^3$ be the computational domain. The total mass of pollutant over time in a compartment k, $k = 1, \ldots, n$, residing at a point $x \in \Omega$ is given by a function $m_k(x, \cdot) : [0, +\infty[\longrightarrow \mathbb{R}$, which has to be approximated.

For ecosystems and food webs, each compartment usually represents one species or one type of species. Species which feed on each other are connected by an interchange flow modeling their predating pattern, as described in the last section. See, e. g., [13,12,26] for a list of examples.

Of course, the more detailed the ecosystem description is, the more detailed and more realistic the compartment model will be. For example, it is possible to divide a species into several distinct classes, each living spatially separated or equipped with distinctive features (gender, age [24], etc). Each of these classes would then be represented by a separate compartment.

4.2 Individual based chemokinetic models

In chemokinetic models, the modeler has to divide the body of an individual under consideration into as many compartments as he or she wishes and define the pathways for the pollutant between the body compartments and the ecosystem compartments defined previously. Usually, each body compartment represents a specific organ or a specific tissue type. In toxicology, models of this type are also called physiologically based pharmacokinetic models (PBPK). See, e. g., [13, Section 7.7] for a model describing cadmium transport in mammals or [5] for a model describing the transport of the degreasing agent trichlorethylene and the monomer 1,1-dichlorethylene (widely used in the packaging industry) in rats. Several different types of individuals can also be modeled by specifying the defining compartments for each of them. For instance, it might make sense to use a group of compartments to model the adult members of a species and along with this another group of compartments as a more detailed model for the adolescents.

5 Toxicology: a cytodynamic multistage model

For each tissue and organ modeled in the chemokinetic model, the cytodynamic model calculates quantitatively the carcinogenic and/or cytotoxic effects of the pollutant in the corresponding tissue or organ. The approach adopted is an extension

of the Armitage-Doll model of carcinogenesis and follows the cytodynamic model as outlined, for example, in [7,3]. Note that this type of model is able to handle hormetic behavior of a tissue to an outside response, a feature of strong interests to toxicologists [2]. For each compartment k, the modeler has to specify the number $r(k)$ of different cell populations that reside in this compartment. The dynamic behavior of each population is then governed by the pollutant concentration inside the compartment (by a toxic or a mutagenic effect) and by the population numbers of the other cell types. The final output of this module consists of approximations to functions $p_{i,k}(x,\cdot) : [0,+\infty[\longrightarrow \mathbb{R}$ ($k = 1,\ldots,n$, $i = 1,\ldots,r(k)$, $x \in \Omega$), representing the time-dependent number of cells of type i in compartment k residing at point x, and calculated as a solution to system of ordinary differential equations.

6 Optimization

6.1 The effect function

The discussion in the last sections and the model derivation in Section 1 have shown that we are able to calculate not only the pollutant concentration in certain compartments, which can be anything from ecotrophic levels down to organs in individuals, but that we can also calculate the effect of a pollutant on certain cell populations in terms of toxicity and carcinogenicity. The minimization of *all* the quantified effects simultaneously is evidently a multicriteria problem for which the knowledge of some approximation of the efficient set would be of great use. However, the high computational demands to compute such an approximation as well as methodological difficulties force us at the present moment to globalize all the different effects by way of a utility function. The choice of such a utility is nontrivial, and it does not seem to be clear what effect a corresponding choice has on the solutions of the resulting single-criterion optimization problem. We prefer at the present moment to globalize all the different effects by way of a weighting function.

Let $G \subset \mathbb{R}^2$ be the ground surface of the modeled region, $n \geq 1$ the number of compartments per point in G and $m_k(x,t)$ be the pollutant mass in compartment k ($k = 1,\ldots,n$) at $x \in G$ and time $t \geq 0$. Furthermore, let $r(k) \geq 0$ be the number of cell populations in compartment k and $p_{i,k}(x,t)$ be the number of cells of type i in compartment k ($i = 1,\ldots,r(k)$) at the point $x \in G$ and time $t \geq 0$. Define $N := \sum_{k=1}^{n} r(k)$. The modeler has then to specify $n + N$ real-valued weights $\hat{\omega}_k, \omega_{i,k}$ ($k = 1,\ldots,n$, $i = 1,\ldots,r(k)$) (one for each compartment and one for each cell population). Let $\lambda = (y_1, y_2, w, h)^\top \in \mathbb{R}^4$ be the vector of decision variables. The vector $(y_1, y_2) \in G$ represents the location of the pollutant emitter, while $w \geq 0$ is the width of the "smokestack" and $h \geq 0$ is it's height. With the weights given as above, define the function

$$W(x,t,\lambda) := \sum_{k=1}^{n} \left(\hat{\omega}_k m_k(x,t) + \sum_{i=1}^{r(k)} \omega_{i,k} p_{i,k}(x,t) \right),$$

which measures the agglomerated effect of the decision λ at the point $x \in G$ at time $t \geq 0$. Here, the dependence of m_k and $p_{i,k}$ on λ has been skipped, as it has

been done above. The agglomeration of effects over a time interval $[0, T]$, $T > 0$, can be done in many different ways. The method employed in OLAF is the time integral over the effect

$$e(x, \lambda) := \int_0^T W(x, t, \lambda)\, \mathrm{d}t.$$

Note that this approach allows for hormetic values, i. e. $e(x, \lambda) < 0$ for some points $x \in G$ and decisions $\lambda \in \mathbb{R}^4$. In such a case, the pollutant effect has actually lead to a reaction of the cell population of an individual residing at x which can be interpreted as an increase in health [2,3].

6.2 The objective function

Let $e(x, \lambda)$ be an arbitrary agglomeration function for the pollutant effects on individuals residing at $x \in G$. If the cytodynamic model is used, this effect quantification measures the overall effect of the pollutant under consideration on a small number of individuals only. From the viewpoint of a health care decision maker, it is therefore important to calculate the pollutant effect on the whole population. Fortunately, this poses no problem. Suppose that we have given a population density function $p : G \longrightarrow \mathbb{R}$, $p(x) \geq 0$ for all $x \in G$. The objective function to be minimized is then given by

$$f(\lambda) := \int_G p(x) e(x, \lambda)\, \mathrm{d}x. \tag{1}$$

Obviously, to evaluate f we do not only need to evaluate $e(\cdot, \lambda)$ at several points, but we also need a numerical integration routine. Since in the present version of OLAF the pollutant concentration and the health effect function is computed on a fixed Cartesian grid of size $N \times M$ with grid points $x_i \in G$, the integral in (1) can be replaced by the weighted sum over all grid points considered, where the weights are $1/NM$.

6.3 The gradient of the objective function

An efficient optimization algorithm needs local information of higher order about the objective function. While there does not seem to be a simple way to compute $\partial f/\partial w$ and $\partial f/\partial h$, the gradient with respect to the locational decision variables y_1, y_2 (which are arguably the most important ones) can be approximated in the following way. Under suitable differentiability assumptions and integrability of $e(\cdot, \lambda)$, we have that

$$\frac{\partial f(\lambda)}{\partial y_1} = \frac{\partial}{\partial y_1} \int_G p(x) e(x, \lambda)\, \mathrm{d}x = \int_G p(x) \frac{\partial}{\partial y_1} e(x, \lambda)\, \mathrm{d}x. \tag{2}$$

Furthermore, a simple geometric argument shows that it is reasonable to make the assumption

$$e((x_1, x_2)^\top, (y_1 + \delta, y_2, w, h)^\top) = e((x_1 - \delta, x_2)^\top, (y_1, y_2, w, h)^\top) + o(|\delta|)$$

for $\delta \to 0$. As a consequence, for $e(\cdot,\lambda) \in C^1(\text{int}(G))$ we get

$$\begin{aligned}
&\frac{\partial}{\partial y_1} e((x_1,x_2)^\top,(y_1,y_2,w,h)^\top) \\
&= \lim_{\delta\to 0} \frac{1}{2\delta}\left(e((x_1,x_2)^\top,(y_1+\delta,y_2,w,h)^\top) - e((x_1,x_2)^\top,(y_1-\delta,y_2,w,h)^\top)\right) \\
&= \lim_{\delta\to 0} \frac{1}{2\delta}\left(e((x_1-\delta,x_2)^\top,(y_1,y_2,w,h)^\top) - e((x_1+\delta,x_2)^\top,(y_1,y_2,w,h)^\top)\right. \\
&\qquad\qquad \left. +o(|\delta|)\right) \\
&= -\frac{\partial}{\partial x_1} e((x_1,x_2)^\top,(y_1,y_2,w,h)^\top).
\end{aligned}$$

This is the approach chosen in OLAF. The differentials $\partial e/\partial x_1$ are approximated by finite differences of third order, where the discretization size is exactly the grid size of the computational grid in x_1-direction. In this way, the approximations to $e(\cdot,\lambda)$ on the grid discretizing G are used to calculate the differential of $e(\cdot,\lambda)$, and no further simulation run of the dispersion model or the cytodynamic model is necessary. Under analogous assumptions, the same reasoning holds for $\partial f/\partial y_2$. Additional objective function evaluations due to time-intensive numerical differentiation routines can therefore be avoided.

We may go even one step further. Using (2) and partial integration, we see that

$$\frac{\partial f(\lambda)}{\partial y_1} = -\int_{\partial G} p(z)e(z,\lambda)\nu_1(z)\,\mathrm{d}z + \int_G e(x,\lambda)\frac{\partial}{\partial x_1}p(x)\,\mathrm{d}y$$

holds whenever G has a smooth boundary and the function $p(\cdot)e(\cdot,\lambda)$ has compact support. Here, $\nu(z)$ is the outer normal of G at a boundary point $z \in \partial G$. By choosing G large enough (or choosing a population distribution p with compact support small enough), we can make sure that the boundary integral gets arbitrary small, and we can approximate the differential of f with respect to y_1 by

$$\frac{\partial f(\lambda)}{\partial y_1} \approx \int_G e(x,\lambda)\frac{\partial}{\partial x_1}p(x)\,\mathrm{d}y. \tag{3}$$

Since $\partial p(x)/\partial x_1$ is independent of λ, it can be computed or approximated in whatever way necessary before any optimization takes place. Approximating $\partial f(\lambda)/\partial y_1$ then involves only the evaluation of the integral in (3), which can be done in the same way as and in conjunction with the evaluation of f at λ.

Unfortunately, none of these schemes works for the derivatives with respect to stack height h and stack diameter w. As a consequence, numerical differentiation with central differences is used to approximate the derivative of f with respect to h and w. Another possibility would be to use automatic differentiation to compute these derivatives. Unfortunately, preliminary numerical experience [17, Chapter 1] suggests that this approach is computationally very expensive for parameterized air dispersion models. No effort has been made to use automatic differentiation within OLAF.

6.4 The optimization algorithm

At the present moment, an SQP-method [23] is used to solve the location problem at hand. The optimization routine considers the locational parameters of the polluting facility as parameters to be optimized and minimizes the function f as defined above. Usually, constraints on the decision vector $\lambda \in \mathbb{R}^4$ are natural results of the modeling process. They include representations of engineering restrictions on stack height resp. diameter as well as locational constraints representing coastlines, jurisdictional boundaries, etc. As a consequence, almost all optimization problems to be solved are constrained ones. We can, however, assume that all constraints are written in standard form $g_1(\lambda) \leq 0$ or $g_2(\lambda) = 0$ with functions $g_1 : \mathbb{R}^4 \longrightarrow \mathbb{R}^{n_1}$, $g_2 : \mathbb{R}^4 \longrightarrow \mathbb{R}^{n_2}$, $n_1, n_2 \geq 1$. Moreover, both functions can be assumed to be continuously differentiable. This is usually considered as a strong assumption in locational analysis, especially if one wants to model regions forbidden for placement of the smokestack. However, recent research [9] has shown that a corresponding problem formulation is not necessarily more complicated than other locational models, provided that the proper modeling tools are used.

7 A numerical study

The software package OLAF was used for a numerical case study concerning the location of an air polluting facility in the region of central Europe. A complete description of the study can be found in the technical report [8].

"Central Europe" is a rather vague term. We mean here a region roughly encompassing the Benelux countries (Belgium, The Netherlands, and Luxembourg), the western part of Germany, southeastern Britain, and northeastern France. More precisely, the region used in this case study is a spherical rectangle with the Greenwich meridian (0° longitude) as it's western bound and 8°E longitude as it's eastern boundary, while the latitude coordinate runs from 47°N to 53°N. In Figure 3 the region under consideration is depicted (the map has been produced with GMT (*G*eneric *M*apping *T*ools), see [28]). In total, ca. 58 million people live in the region which the case study considers.

7.1 Data

The origin of the population data is described in [25]. Figure 4 show the resulting discrete distribution function interpolated by bicubic splines. Land use data stems from [19], where the surface of the earth is divided into 44 different land ecosystem complexes. Meteorological data has been obtained from the FASTEX archive [11]. The data consists of, among others, all the information necessary for the preprocessor and covers a time period from January 6, 1997, to February 28, 1997. While this time period is shorter than actually necessary, we feel confident that we can at least demonstrate the feasibility of the approach taken by solving an optimization problem as realistic as possible. More details with respect to the meteorological data can be found in [8].

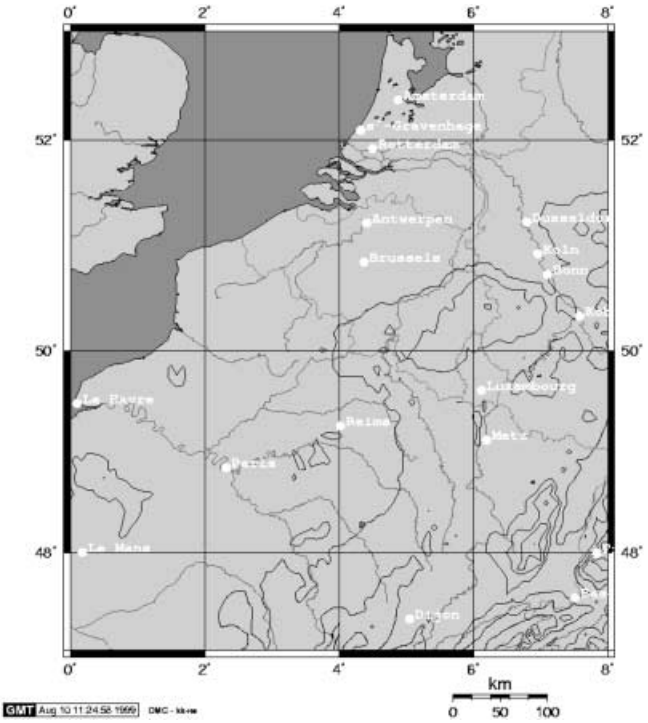

Fig. 3. The modeled region

In this case study, a hypothetical pollutant is used whose gas-phase characteristics are those of SO_2, a chemical often used in benchmarking atmospheric dispersion models [27, Tables 6-1,6-2]. However, for the chemokinetic submodel, it was assumed that the pollutant behaves similar to perchlorethylene (PERC), a typical and widely used carcinogen [10].

7.2 Modeling Issues

7.2.1 Meteorologics. With the data available, we decided to use a computational grid of size 80×60. Given the sparsity and the resolution of the population data, land use data, and meteorological data available, this seemed to be a comfortable compromise between computational accuracy and data accuracy. All settings in the air dispersion model used are standard ones as recommended in [6].

7.2.2 Chemokinetics. The chemokinetic model in use is an extension of the one found in [10] (cmp. Figure 5, p. 130). Note that the lung compartment has been split up into lung air and lung blood, since the main uptake route of a pollutant into the body should be modeled as detailed as possible. The only metabolism in this model takes place in the liver compartment, see [10].

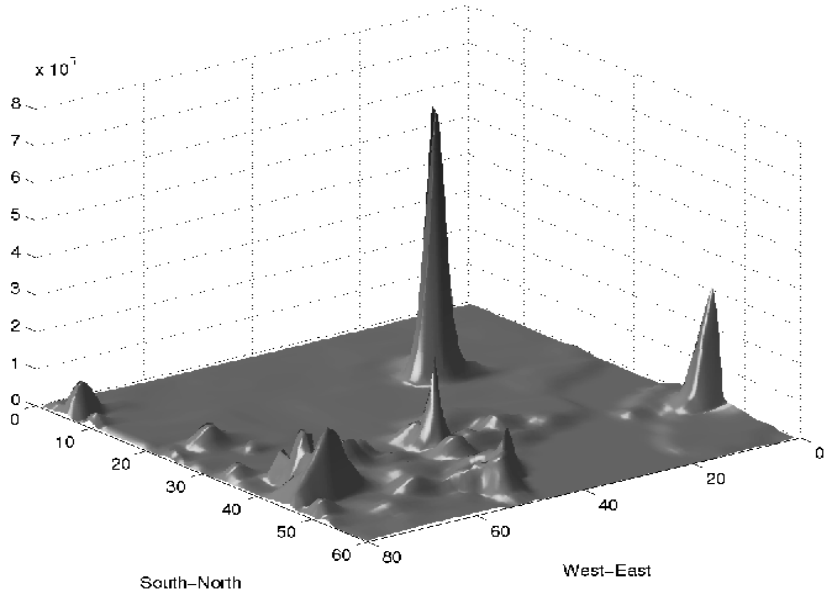

Fig. 4. Population distribution on the computational grid as seen from the north east. The highest peak represents Paris, the small peak just in front of Paris is Brussels. the "hills" near the closer corner are the cities in the Rhine-Ruhr area. On the right side, greater London area is just cut off

7.2.3 Cytodynamics. The cytodynamic model follows the one outlined in [3] and is depicted in Figure 6. This model is used only for the cell population in the liver compartment. In the model, stem cells (S) may give rise to premalignant cells (P), who, in turn, can mutate to malignant cells. But due to toxic effects, stem cells may also die or lose their ability to divide. In this case they are placed in a pool of 'dead' cells (D). Furthermore, cells from a reservoir (R) can replace stem cells, but also mutate into premalignant cells of a different type (Q).

7.2.4 The effect function. Given that a complete cytodynamic model is at our disposal and that in this study ecological consequences are, hypothetically, of no concern, we used weights $\hat{\omega}_k = 0$ for all compartments k, and choose $\omega_{1,9} = \omega_{2,9} = \omega_{3,9} = \omega_{4,9} = 0$, as well as $\omega_{5,9} = 10^8$ and $\omega_{6,9} = 1$. As a consequence, we value the carcinogenic effects as much more important than the cytotoxic effects (the integral over time of the number of dead cells, cell type no. 6). All other effects, especially the number of premalignant cells, are neglected.

7.2.5 Constraints. OLAF uses grid units for the locational decision variables (y_1, y_2) and meter as the physical dimension of the stack height resp. diameter. It is therefore comparatively simple to add constraints to the optimization process. The following constraints with respect to the decision variables (y_1, y_2, w, h) have been added:

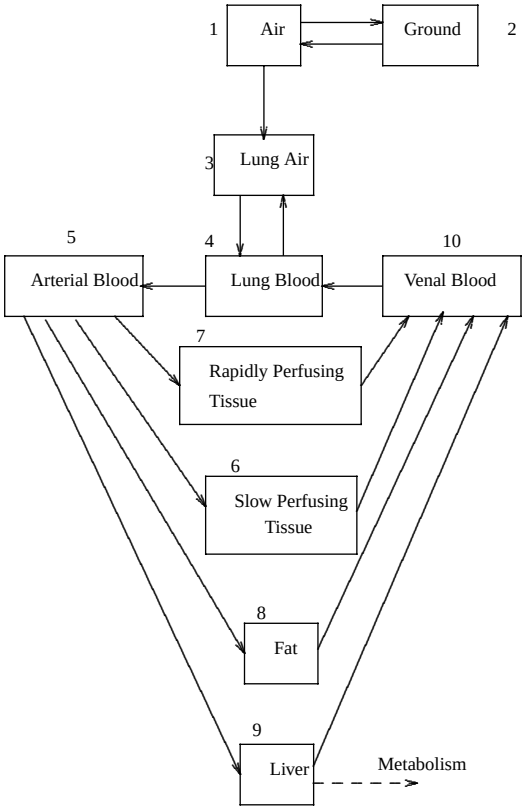

Fig. 5. Chemokinetic model used in the case study. Numbers represent the numbering scheme for the compartments

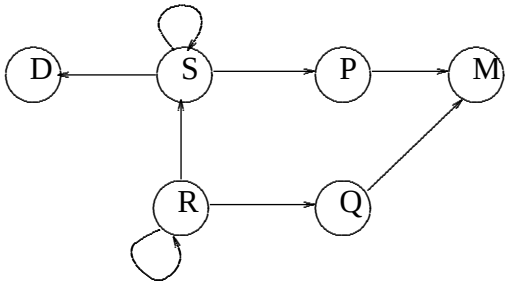

Fig. 6. Cytodynamic model used in the case study. Circles denote cell types, arrows denote procreation and mutagenic paths

1. Locational constraints:
   (a) The facility has to lie inside a circle with center at $4°$W longitude and $50°$N latitude and radius 180.738km (20 grid units). This translates readily into the convex quadratic constraint
   $(y_1 - 40)^2 + (y_2 - 30)^2 - 400 \leq 0$.
   (b) For hypothetical political reasons, the facility is not allowed to lie within a circle of radius 100km around Paris and within circles of radius 60km around Brussels and Reims. With a grid spacing of 9.0869km, these constraints can be expressed as concave quadratic constraints:
   $(60/9.0869)^2 - (y_1 - 45.33)^2 - (y_2 - 39)^2 \leq 0$
   for Brussels,
   $(100/9.0869)^2 - (y_1 - 24)^2 - (y_2 - 17.33)^2 \leq 0$
   for Paris, and
   $(60/9.0869)^2 - (y_1 - 40.33)^2 - (y_2 - 23)^2 \leq 0$
   for Reims (cmp. Figure 8).
2. Engineering constraints:
   (a) The stack height has to be between 20m and 90m, i. e.
   $20 \leq h \leq 90$.
   (b) The stack diameter at the top has to be between 2m and 7m, i. e.
   $2 \leq w \leq 7$.

   Note that these simple box constraints for the physical characteristics of the stack are rather unrealistic. More realistic constraints have to include a nonlinear coupling between h and w.

7.3 Results

The code was developed in Fortran 77 and tested on an Intel Pentium 330MHz with 128MB main memory, 512KB cache and Linux operating system (Kernel 2.0.35).

Preparing the input files by some filter and interpolation routines took 13 minutes on the Intel platform. After that, the MESOPAC preprocessor needed 30 minutes user time to process the data described above. Starting from the point

$$\lambda^{(0)} := \left(y_1^{(0)}, y_2^{(0)}, w^{(0)}, h^{(0)}\right)^\top = (40, 30, 4, 80)^\top$$

with an objective function value of 70.193, the code needed 23 iterations and ca. 260 hours to calculate the local minimum

$$\lambda^* = (58.291, 22.018, 4.000, 79.687)^\top$$

with objective function value 29.018. (All values are given with an accuracy of ca. 1%.) Figure 7 shows the effect function $e(\cdot, \lambda^{(0)})$ overlayed over a map of the region considered, while Figure 8 shows the changes the locational decision variables underwent during the optimization process. Note that this picture is partially incomplete, since the whole decision space is four-dimensional. As it can be seen, the code seems to encounter some convergence problems close to the local minimum found. This might be attributed to the inevitable noise in the objective function or to inevitable errors in the gradient approximation scheme used.

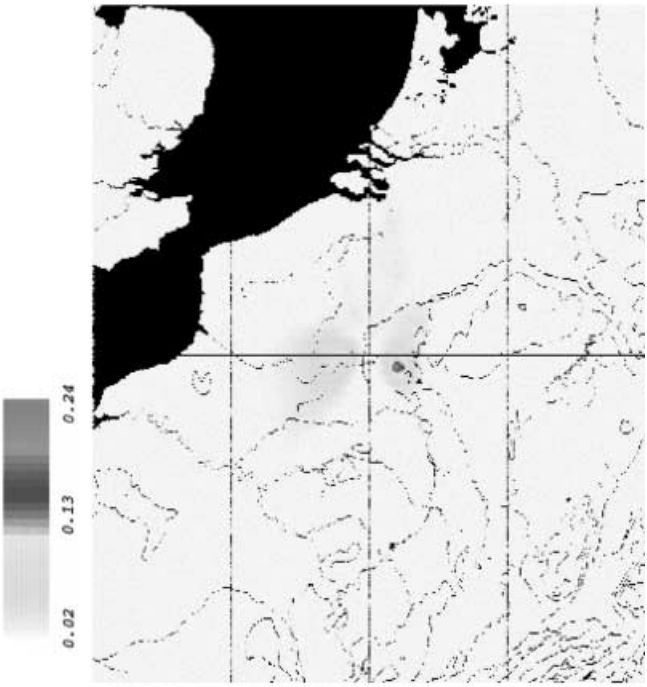

Fig. 7. The effect function $e(\cdot, \lambda^{(0)})$ for the starting configuration

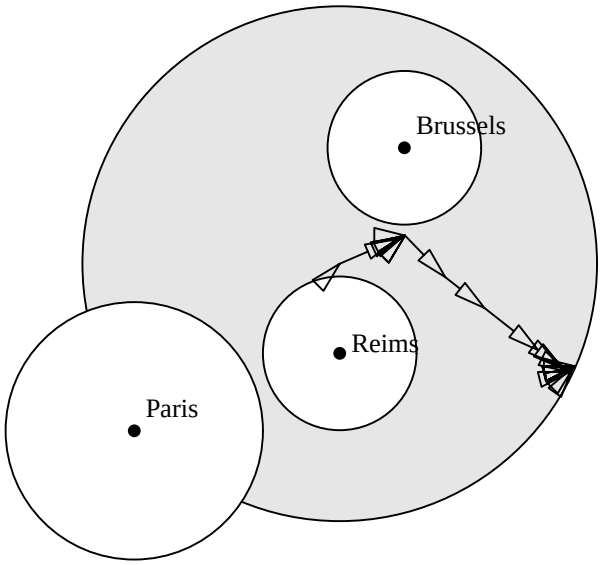

Fig. 8. Iterations in the locational domain. Line searches and iterations in the stack decision variables are not depicted

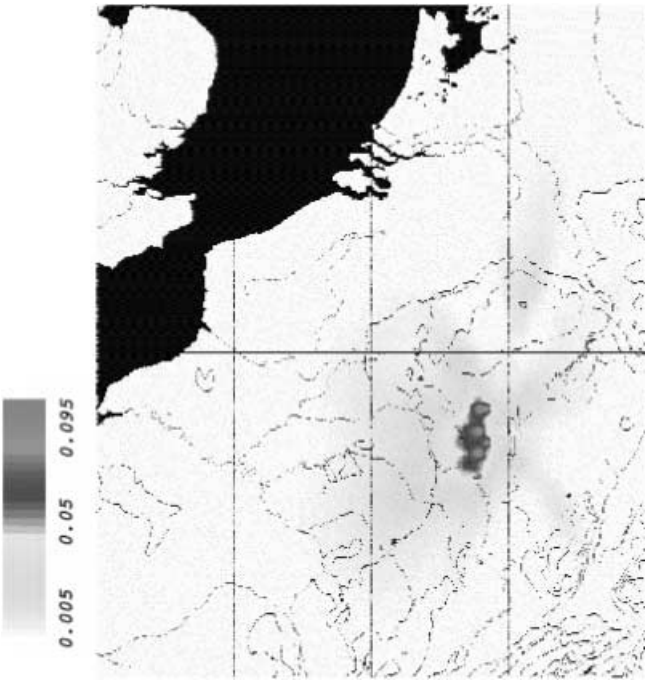

Fig. 9. The effect function $e(\cdot, \lambda^*)$ for the final configuration

Figure 9 shows the effect function $e(\cdot, \lambda^*)$ for the final point λ^* overlayed over a map of the region.

Note that the stack height and width remain virtually unchanged. This might indicate that the pollution load is insensitive to small scale changes of these variables on the data used. On the other hand, the local optimum found lies on the boundary of the feasible region with respect to the locational variables, far away from the starting point, and especially farther away from the main population centers (Paris and the larger cities of the Netherlands) than $(y_1^{(0)}, y_2^{(0)})$. The meteorological conditions simulated over the time period of interest seem to indicate a pollutant transport from $(y_1^{(0)}, y_2^{(0)})$ in the direction of these population centers, while these effects seem to be much smaller when the pollutant is emitted from λ^*. Note also that the total effect is much more evenly distributed from λ^*, when compared with the distribution emitted from $\lambda^{(0)}$.

More details with respect to the performance of the code can be found in the project report [8].

8 Conclusions

The developed prototype code OLAF is capable of solving complex, highly realistic air pollution facility location problems. The code is able to follow the trace of the pollutant under consideration over different ecotrophic levels and into different

individual organs of a modeled individual. There, detailed cytodynamic effects are simulated in order to evaluate the health effect of the pollutant on humans or other beings. A numerical study has shown that highly realistic models of pollutant fate and effect can be used together with realistic atmospheric dispersion models and in conjunction with an optimization process as a decision aiding tool. As it had to be expected, the computational workload to find optimal locations is far from trivial, and at the present moment the code can probably not be used in an interactive mode, except for extremely simple problems.

While we have shown that the overall strategy for estimating and minimizing ecological effects and health effects is feasible, several open questions remain. For example, at the present moment the code is able to compute local minima only. Global minima, are, of course, of high interest to the decision maker. It remains to be seen if highly efficient global optimization techniques can be applied to the pollution model developed or if it will be necessary to resort to stochastic techniques of inherently low efficiency. Moreover, several modeling issues, especially with respect to the agglomeration function, remain open. A robust, realistic and computationally effective formulation for the multicriteria problem at hand still needs to be found. Other atmospheric dispersion models which include nonlinear chemistry of the pollution should be tested, too, as it is the case with other numerical integration techniques for the ordinary differential equations occurring in the model. Another possibility would be to investigate the effect of time-dependent emission rates, a particular interesting subject since these emission rates might be considered as decision variables, albeit heavily constrained ones. Moreover, further work is necessary with respect to the evaluation of synergistic effects caused by several pollutants instead of only one. This, together with other open questions, is a subject of further research.

References

1. Amouzega, MA, Moshirvaziri, K (1999) Determining optimal pollution control policies: An application of bilevel programming. European Journal of Operational Research 119(1): 100–120
2. Dose-response studies of genotoxic rodent carcinogens: thresholds, hockey sticks, hormesis or straight lines? BELLE Newsletter 3(3), February 1995. Electronically distributed via http://www.belleonline.com/
3. Bogen, KT (1998) Mechanistic model predicts a U-shaped relation of radon exposure to lung cancer risk reflected in combined occupational and U.S. residential data. BELLE Newsletter 7(2). Electronically distributed via http://www.belleonline.com/n3v72.html
4. de Leeuw F, Berge E, Grønskei K, Tombrou M (1996) Review on requirements for models and model application. Electronically available via http://www.etcaq.rivm.nl/
5. El-Masri HA, Tessari JD, Yang, RS (1996) Exploration of an interaction threshold for the joint toxicity of trichlorethylene and 1,1-dichlorethylene: utilization of a PBPK model. Archive of Toxicology 70: 527–539
6. A modeling protocol for applying MESOPUFF II to long range transport problems. Technical Report EPA-454/R-92-021, U. S. Environmental Protection Agency, Office of Air Quality Planning and Standards, Technical Support Division, Research Triangle Park, North Carolina 27711, October 1992. Electronically distributed via http://www.epa.gov/ttn/scram/guidance/guide/mesoprot.zip

7. Chapter 8. Dose-Response Modeling for 2,3,7,8-TCDD. Workshop Review Draft (1997). Electronically distributed via http://www.epa.gov/ORD/WebPubs/dioxin/dioxin8.pdf
8. Fliege J (1999) Project report: OLAF – A General modeling system to optimize the location of an air polluting facility. Technical Report BEIF/116, Vakgroep Beleidsinformatica, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium Electronically available via http://www.mathematik.uni-dortmund.de/lsx/fliege/olaf/
9. Fliege J, Nickel S (2000) An interior point mehod for multifacility location problems with forbidden regions. Studies in Locational Analysis 14: 23–46
10. Isukapalli SS (1999) Uncertainty analysis of transport-transformation models. PhD thesis, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA. Electronically availaible at http://www.ccl.rutgers.edu/∼ssi/thesis/
11. Jaubert G, Piriou C, Loehrer SM, Petitpa A, Moore JA (1999) The FASTEX experiment data archive. Technical Report 57, Meteo-France/CNRM/GMME, Météo-France, 1 Quai Branly, 75340 Paris Cedex 07, France. Electronically available via http://www.cnrm.meteo.fr:8000/dbfastex/index.html
12. Jørgensen SE (1979) Modelling the distribution and effec of heavy metals in an aquatic ecosystem. Journal of Ecologic Modelling 6: 199–223
13. Jørgensen SE (1986) Fundamentals of ecological modelling. In: Developments in environmental modelling, number 9. Elsevier, Amsterdam Oxford New York Tokyo
14. Karkazis J (1998) Long range environmental impact assessment for activities causing air borne pollution – Applications in Belgium. Technical Report BEIF/104, Vakgroep Beleidsinformatica, Vrije Universiteit Brussel, Pleinlaan 2, 1050 Brussel, Belgium
15. Karkazis J, Boffey B (1994) Modelling pollution spread. Studies in Locational Analysis, Vol 7. EWGLA 7 Proceedings
16. Loughlin DH, Ranjithan S (1997) The neighborhood constraint method: a genetic algorithm-based multiobjective optimization technique. In: Proceedings of the 7th International Conference on Genetic Algorithms. Morgan Kaufman, San Francisco
17. Loughlin DH (1998) Genetic algorithm-based optimization in the development of tropospheric ozone control strategies: least cost, multiobjective, alternate generation, and chance-constrained applications. PhD thesis, Civil Engineering Department, North Carolina State University, Raleigh, North Carolina, USA. Electronically availaible at http://www4.ncsu.edu/∼dhloughl/Documents/Dissertation/
18. Moussiopoulos N, Berge E, Bøhler T, de Leeuw F, Grønskei K-E, Mylona S, Tombrou M (1996) Ambient air quality, pollutant dispersion and transport models. Electronically available via http://www.etcaq.rivm.nl/
19. Olson JS, Watts JA, Allison LJ (1983) Carbon in live vegetation of major world ecosystems. Technical Report Environmental Sciences Division Publication No. 1997, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA
20. Plastria F, Carrizosa E (1999) Undesirable facility location with minimal covering objectives. European Journal of Operational Research 119: 158–180
21. Rood RB (1987) Numerical advection algorithms and their role in atmospheric transport and chemistry models. Review of Geophysics 25(1): 75–100
22. Scire JS, Lurmann F, Bass A, Hanna SR (1984) User's guide to the MESOPUFF II model and related processor programs. U. S. Environmental Protection Agency, Research Triangle Park, North Carolina, EPA-600/8-84-013 edition
23. Spellucci P DONLP2 code and accompanying documentation. Electronically available via http://plato.la.asu.edu/donlp2.html or http://www.mathematik.tu-darmstadt.de/ags/ag8/spellucci/
24. Thomann RV (1989) Deterministic and statistical models of chemical fate in aquatic systems, pp 323–351. Springer, Berlin Heidelberg New York

25. Tobler W, Deichmann U, Gottsegen J, Maloy K (1995) The global demography project. Technical Report TR-95-6, National Center for Geographic Information and Analysis, Department of Geography, University of California, Santa Barbara, USA. Electronically available via http://www.ciesin.org/datasets/gpw/globldem.doc.html
26. Trapp S, Matthies M (1998) Chemodynamics and environmental modelling. Springer, Berlin Heidelberg New York
27. U.S. Environmental Protection Agency (1994) Research Triangle Park, North Carolina. A revised user's guide to MESOPUFF II (V5.1), EPA-454/b-94-025 edition. Electronically distributed via http://www.epa.gov/ttn/scram/userg/other/
28. Wessel P, Smith WHF (1995) New version of the generic mapping tools released. Technical report, American Geophysical Union. Electronically available as http://www.agu.org/eos_elec