LARGE SCALE SYSTEMS ISSUES IN MODELING COST-EFFECTIVE POLICIES FOR IMPROVING THE EUROPEAN AIR QUALITY

Marek Makowski

International Institute for Applied Systems Analysis,
A-2361 Laxenburg, Austria
`http://www.iiasa.ac.at/~marek`

Abstract: The paper presents a large scale nonlinear model which is used for supporting international negotiations aimed at improving air quality in Europe. The model helps to identify cost-effective measures for reducing air pollution emissions that will result in meeting environmental standards for tropospheric ozone, acidification and eutrophication. Several methodological issues related to the specification, generation and optimization-based analysis of large nonlinear models for decision support that are of a more general interest are presented.

Keywords: decision support systems, air pollution, nonlinear models, object-oriented programming, preprocessing, robustness, multiple-criterion optimization, model management, criterion functions, constraint satisfaction problems.

1. INTRODUCTION

[1] In many parts of Europe the critical levels of air pollution indicators are exceeded and measures to improve air quality in these areas are needed to protect the relevant ecosystems. Cost effective measures aimed at the reduction of ground level ozone concentrations at several hundreds of receptors over Europe can be calculated by a minimization of a cost function that corresponds to the costs related to reductions of NO_x and VOC emissions subject to constraints on the resulting ozone concentrations. The Ozone model, cf e.g. (Heyes *et al.*, 1997), has been developed for analysis of various policy options that lead to improvement of the air quality by reductions of such emissions. However, the emissions of NO_x should also conform to the standards set at each receptor for acidification and for eutrophication. The latter problem is handled by the RAINS model (Alcamo *et al.*, 1990). An analysis of two separate models is cumbersome, therefore the RAINS model has been included in the Ozone model. This in turn requires a joint consideration of not only emissions of NO_x and VOC but also of NH_3 (ammonia) and SO_x (sulphur oxides). The resulting model is large and nonlinear with a large linear part.

There is a number of methodological and technical issues related to the specification, generation and optimization-based analysis of such a large model that are of a more general interest and therefore several of them are presented in this paper:

- The resulting model is a nonlinear one, therefore a problem specific generator has been developed and coupled with three nonlinear solvers. The generation of the model requires processing of a large amount of data coming from various sources. Object-oriented programming approach to the model generation and analysis has been applied.
- A representation of environmental targets by hard constraints would result in recommendations of expensive solutions, hence soft constraints (with compensations for violations of the original targets) are specified.
- The resulting optimization problem has typically non-unique solutions, therefore a technique called regularization was applied in order to

[1] This paper will appear in: Large Scale Systems: Theory & Applications, P. Groumpos (Ed.), Elsevier, 1998.

provide a suboptimal solution having additional properties that are specified by a user.
- A minimization of costs related to measures needed for improvement of air quality is a main goal; however, other objectives (such as robustness of a solution, trade-offs between costs and violations of environmental standards) are also important. Therefore, a multicriteria model analysis has been applied to this case study.

2. MODEL DEFINITION

One should first distinguish between a set I of sources of various types of air pollution, and a set J of areas for which various quality indicators are assessed. Conventionally, the names *emitter* and *receptor* are used for elements of such sets. In order to account for measures that can be applied to a group of emitters, sets of NO_x and VOC emitters are composed of subsets called *sectors*. Emitters that belong to a particular sector emit either NO_x or VOC or a linear combination of them.

The model definition requires the following indices:
- Index $i \in I$ corresponds to emitters. The number of elements in I is equal to the number of countries (about 50).
- Index $is \in S_i$ corresponds to a sector that emits either NO_x or VOC or a linear combination of them; S_i is a set of sectors in i-th country. A set S_i has typically about 5 elements.
- Index $j \in J$ corresponds to receptors. There are 598 receptors, each representing one 150x150 km grid.
- Index $l \in L$ corresponds to a combination of ozone thresholds and a year.
- Index $m \in M$ corresponds to a set of receptors for which balancing of violations and surpluses of targets is defined.

2.1 *Decision variables*

The main decision variables are the annual emissions of the following four types of primary air pollution emitted by either a sector or by a country:
n_{is} - emission of NO_x
v_{is} - emission of VOCs
a_i - emission of NH_3
s_i - emission of SO_x

Additionally, optional decision variables are considered for scenarios which allow controlled violations of air quality targets. For such scenarios variables corresponding to each type of the considered air quality targets are defined for each receptor. Optionally, violations of targets can be balanced with surpluses (understood as a difference between a target and a corresponding actual concentration). For efficiency reasons one variable is used for both violations of targets and surpluses (positive values represent violations while negative values correspond to a part of a surplus that is used to balance violations of targets with surpluses).

Therefore, the following decision variables are optionally defined for violations (surplus if a variable is negative) of the corresponding targets:
y_{lj} - for ozone exposure,
ya_j - for acidification,
ye_j - for eutrophication.

2.2 *Outcome variables*

The consequences of applications of computed (or provided) values of the decision variables are evaluated by values of outcome variables. However, several auxiliary variables needed for the definitions of outcome variables have to be specified first.

2.2.1. *Auxiliary variables*

n_i - emission of NO_x:

$$n_i = \sum_{is \in S_i} n_{is} \tag{1}$$

v_i - emission of VOCs:

$$v_i = \sum_{is \in S_i} v_{is} \tag{2}$$

en_{lj} - the mean effective emissions of l-th type of NO_x experienced at j-th receptor:

$$en_{lj} = \sum_{i \in I} e_{lij} n_i + enn_{lj} \tag{3}$$

where enn_{lj} are given effective natural emissions of NO_x.

nlv_{lj} - the representation of another nonlinear term defining the l-th type of ozone exposure at j-th receptor:

$$nlv_{lj} = \sum_{i \in I} d_{lij} v_i \tag{4}$$

2.2.2. *Definition of outcome variables*

One outcome variable represents the sum of costs of reductions of emissions; four sets of other outcome variables correspond to various indices of air quality.

The sum of annual costs related to the reduction of emissions is defined by:

$$cost = \sum_{i \in I} (ca_i(a_i) + cs_i(s_i) + c_i(n_i, v_i)) \tag{5}$$

where $ca_i(\cdot)$ and $cs_i(\cdot)$ are cost functions for reductions of NH_3 and SO_x, respectively, and $c_i(\cdot)$ are defined by:

$$c_i(n_i, v_i) = \sum_{is \in S_i} c_{is}(\cdot) \qquad (6)$$

where $c_{is}(\cdot)$ are cost functions for NO_x or for VOC or for joint NO_x and VOC reduction.

All cost functions are PWL (piece-wise linear), convex and monotonically decreasing.

For each receptor, the following four outcome variables correspond to various indices of air quality:
aot_{lj} - the long term ozone exposure of l-th type:

$$aot_{lj} = \sum_{i \in I}(a_{lij} v_i + b_{lij} n_i + \gamma_{lij} n_i^2) + \alpha_{lj} en_{lj}^2 + \beta_{lj} en_{lj} nlv_{lj} + k_{lj} \qquad (7)$$

$ac1_j$ - acidification of type 1, i.e. the sum of depositions of NO_x, NH_3 and SO_x:

$$ac1_j = tns_j(\sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + kn_j) + \sum_{i \in I} ts_{ij} s_i + ks_j \qquad (8)$$

$ac2_j$ - acidification of type 2:

$$ac2_j = \sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + tss_j(\sum_{i \in I} ts_{ij} s_i + ks_j) + kn_j \qquad (9)$$

eu_j - eutrophication, i.e. the sum of depositions of NO_x and NH_3:

$$eu_j = \sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + kn_j \qquad (10)$$

where tn_{ij}, ta_{ij}, ts_{ij} are transfer coefficients for NO_x, NH_3 and SO_x, respectively; kn_j and ks_j are constants for nitrogen and sulphur background depositions; tns_{ij}, tss_{ij} are scaling coefficients.

Environmental effects caused by the two types of acidification and by eutrophication are evaluated at each receptor by a PWL function which represents an accumulative excess of each type of the air quality index:
$aac1_j$ - accumulative excess of $ac1_j$:

$$aac1_j = PWL_j^{ac1}(ac1_j) \qquad (11)$$

$aac2_j$ - accumulative excess of $ac2_j$:

$$aac2_j = PWL_j^{ac2}(ac2_j) \qquad (12)$$

aeu_j - accumulative excess of eu_j:

$$aeu_j = PWL_j^{eu}(eu_j) \qquad (13)$$

2.3 *Constraints*

The accumulative excess of long-term ozone exposure is constrained by:

$$aot_{lj} - y_{lj} \leq aot_{lj}^{max} \qquad (14)$$

where aot_{lj} is defined by (7) and aot_{lj}^{max} is a given maximum ozone exposure for l-th threshold at j-th receptor.

Constraint (14) without the term $-y_{lj}$ would be a so-called *hard constraint* for the accumulative excess of ozone exposure. Such a formulation is typically used in a traditional formulation of optimization problems. It can also be used in the presented model by selecting an option that does not allow for generation of variables y_{lj}. However, an implementation of hard constraints for air quality targets would result in forcing much more expensive solutions caused by constraints that are active in only one or two receptors. Introduction of the term $-y_{lj}$ converts a *hard* constraint into a so-called *soft constraint*. This allows a violation of a target air quality. However, such a violation is:

- constrained by upper bounds on variables y_{lj},
- compensated by surpluses (i.e. differences between actual exposure and the corresponding target) in other receptors belonging to the same set of receptors (e.g. located in the same country or region),
- controllable by a trade-off between violations of targets and corresponding costs of reducing emissions.

The constraints for the accumulated excess of the two types of acidification and of eutrophication are defined in a similar way:

$$aac1_j - ya_j \leq aac_j^{max} \qquad (15)$$

$$aac2_j - ya_j \leq aac_j^{max} \qquad (16)$$

$$aeu_j - ye_j \leq aeu_j^{max} \qquad (17)$$

Optionally, violations of targets can be balanced with surpluses of targets over sets of receptors:

$$\sum_{j \in J_m} wo_{lmj} y_{lj} \leq tbo_{lm} \qquad l = 0 \qquad (18)$$

$$\sum_{l=1}^{L} \sum_{j \in J_m} wo_{lmj} y_{lj} \leq \sum_{l=1}^{L} tbo_{lm} \qquad (19)$$

$$\sum_{j \in J_m} wa_{mj} ya_j \leq tba_m \qquad (20)$$

$$\sum_{j \in J_m} we_{mj} ye_j \leq tbe_m \qquad (21)$$

where $wo_{lmj}, wa_{mj}, we_{mj}$ are given weighting coefficients, $J_m, m \in M$ are sets of receptors, and

$tbo_{lm}, tba_m, tbe_m, tbs_m$ are target balances for m-th set of receptors for l-th type of ozone exposure, two types of acidification, and eutrophication, respectively.

3. MODEL ANALYSIS

3.1 *Multiple-criterion optimization*

A composite criterion function (22) is applied in order to support analysis of trade-offs between the three criteria:

- minimization of total costs of emissions reduction,
- minimization of violations of environmental standards,
- robustness of solutions.

The first two components have already been discussed, therefore only the last one requires justification.

A typical problem with applications of optimization techniques for decision support is caused by very different solutions (with almost the same value of the original goal function) of various instances of a mathematical programming problem that differ very little. A quality of a solution is assessed from the optimization point of view primarily through the value of a goal function; therefore solutions of slightly perturbed problems may differ substantially. However, from an application point of view an equally important indication of a solution robustness is some measure of closeness of solutions of perturbed problems. Consider, for the sake of illustration, two instances of the model that differ very little. The values of goal functions for such solutions will be almost the same. However, it often happens that the optimal solution of the first instance has high reduction of emission in country A and low reduction in country B, while the optimal solution for the second instance has low reduction in country A and high reduction in country B. Such solutions would hardly be acceptable. In order to deal with this problem, a technique called regularization, cf. e.g. (Makowski, 1991) for a more detailed discussion, is applied.

The criterion function is defined by:

$$goal\_function = cost + \Theta + \Phi \qquad (22)$$

where the cost term is defined by (5), the penalty term Θ is defined by:

$$\Theta = \sum_{j \in J} \left(\sum_{l \in L} \rho_o y_{lj}^2 + \rho_a ya_j^2 + \rho_e ye_j^2 \right) \qquad (23)$$

and the regularization term Φ is defined by:

$$\Phi = \epsilon ||z - \bar{z}|| \qquad (24)$$

where ρ_o, ρ_a, ρ_e are given penalty coefficients (not necessarily large) and ϵ is a given (not necessarily small) positive number.

The interpretation of each of the terms is as follows:

- The first term corresponds to the sum of costs of emission's reduction of all types of pollution and at all emitters.
- The second term is the penalty term introduced to deal with the soft constraints defined by introduction of variables y_{lj}, ya_j, ye_j into constraints (14, 15, 16, 17).
- The third term is $\epsilon||z-\bar{z}||$, where z denotes a vector composed of all decision variables (except of the decision variables y_{lj}, ya_j, ye_j, for which the reference point is implied to be 0 by the virtue of the penalty term of the criterion function). This is a regularizing term introduced in order to avoid large variations of solutions having similar values of the original criterion function.

Note that the formulation of the optimization problem is single-objective – because such were the requirements of the modeler. However, the specifics of this model – in particular the penalty terms for soft constraint violations, the regularizing term – make it very similar to a multiobjective formulation, as applied *e.g.* to softly constrained inverse scenario analysis.

4. MODEL MANAGEMENT

Generation and management of the model under consideration is a challenging task from the operations research point of view. Several methodological and technical issues that are of a broader interest are discussed in subsequent subsections.

4.1 *Generation and solution of the model*

A commonly accepted rule of thumb for optimization of large nonlinear models is to try various solvers. Therefore three solvers, namely CFSQP (Lawrence *et al.*, 1996), CONOPT (Drud, 1996) and MINOS (Murtagh and Saunders, 1987) are used for solving the resulting optimization problem. For the reasons that are discussed in detail by Makowski (1998*a*) a problem specific model generator has been implemented in C++ for this model.

The task of implementation of software that uses several solvers is interesting from the software engineering point of view. Each solver has a different interface (the way of specification of an optimization problem). However, most of the software components are common to all the solvers. Therefore, object-oriented programming approach was a natural choice because it greatly simplifies the software development by handling common parts in base classes and by providing solver-specific interfaces through inherited classes. The approach is conceptually very simple. Each of the above mentioned solvers is available as a library of Fortran subroutines. The generator has C++ classes that are specific for each solver. These classes are inherited from base classes that handle a common part of the generator. A problem specific report writer processes the results into a form that eases their interpretations. Another class supports a portable interface between C++ and Fortran. Hence, three versions of executables can easily be produced, each is composed of the generator, report writer (postprocessor) and one of the solvers.

A nonlinear solver requires routines that compute values and Jacobian of the constraints and of the goal function. A remarkable part of total computation time is used for execution of these functions, therefore efficiency of their implementation is important. The code for the Jacobian has been generated by *Mathematica* (Wolfram, 1996) with a prior use of the *FullSimplify* operator that substantially simplifies the formulas. This is an easy way to generate a bug free and efficient code.

Finally, one should notice that the dimensions of the model are not fixed. For some scenarios a part of the constraints and/or variables does not need to be generated. Moreover, the dimensions of matrices and vectors used in the model definition vary substantially for various types of analysis. Fortunately, constructors of C++ classes handle such problems in a natural and efficient way.

4.2 *Data handling*

The model has a large number of parameters, but this itself would not be a problem. The challenge comes from the fact that various parts of the parameters are provided as a result of data processing that is performed on various computers. Data handling for the model has to meet the following requirements:

- efficient handling of a large amount of data,
- binary compatibility, at least for Unix and NT,
- easy handling of basic data structures (sparse and dense matrices having elements of basic types),
- no royalty fees.

The HDF (Hierarchical Data Format) public domain software by Koziol and Matzke (1998) is used for handling data in the model. The basic data structures are handled by a collection of well tested C++ classes that are also used for the LP-DIT. A C++ interface class has been implemented for an easy and efficient handling of the used data structures by the HDF library.

4.3 *Conversion of PWL functions*

Costs of emission reductions are given as PWL functions of the emission level. PWL functions are not smooth. Therefore, in order to be able to use efficient nonlinear solvers (which require smooth functions), the PWL cost functions are represented by corresponding smooth functions. However, the PWL functions (11, 12, 13) are replaced by sets of inequalities. Due to the space limitations these conversions are not presented here.

4.4 *Preprocessing of the optimization problem*

Preprocessing of an optimization problem is aimed at generating another problem that has the same goal function value as the original problem and fulfills its constraints, but which is easier to solve. It is a commonly known fact that a preprocessing of a large optimization problem can dramatically reduce computation time and memory requirements. Preprocessing is a standard feature of any good LP solver. However, preprocessing of nonlinear models is a much more difficult task, see e.g. (Drud, 1997). Generally, preprocessing of an optimization problem in a problem generator is much more efficient than an attempt to preprocess a nonlinear problem by a solver. Some instances of the model presented in this paper contain over 10,000 variables and contraints, therefore its preprocessing is essential.

Preprocessing in the generator is composed of the following elements:

- Outcome variables defined by equations (6) through (13) are not generated. The affected constraints are reformulated to equivalent forms without using these outcome variables (auxiliary functions are implemented to provide values of outcome variables for the report writer).
- The variables en_j and nlv_j and equations (3, 4) are eliminated and eq. (7) is modified accordingly.
- All linear constraints are combined into the LP-DIT format by Makowski (1998*b*), and the preprocessing implemented in LP-DIT, which is similar to that implemented by Gondzio (1997), is applied to these constraints. Only preprocessing methods based on the analysis of the primal problem can be applied. Nevertheless, for many types of scenarios even a majority of linear constraints can be removed from the optimization problem.

4.5 *Scaling*

Scaling of nonlinear models is an important element of a model specification. The experiences from the early stages of the model development show that a badly scaled model created numerical problems to all the solvers that are used [2]. A detailed discussion of scaling implemented in the model is far beyond the scope of this paper. Therefore we only mention that each instance of the optimization problem is scaled in the generator in such a way that:

- absolute values of the elements of the Jacobian and of the Hessian are smaller than 10^5,
- an attempt is also made to achieve a smallest (non-zero) absolute value of Jacobian to be *"not too small"*.

5. CONCLUSIONS

The paper illustrates methods and techniques applied to the generation and analysis of the optimization-based nonlinear model which is applied to the examination of various policy options aimed at improving the air quality in Europe. Extensions of traditional OR methods that enhance usefulness of model-based decision support for policy analysis have been presented. Software engineering issues pertinent to generation and analysis of complex and large nonlinear models were discussed.

ACKNOWLEDGEMENTS

The RAINS and Ozone models have been developed for several years by the TAP (Transboundary Air Pollution) Project [3] at IIASA in collaboration with several European institutions. The author has been collaborating with the TAP Project since several years by taking part in various activities related to the development of methodology and software for generation and optimization based analysis of RAINS and Ozone models. The author thanks all members of the TAP Project led by Dr Markus Amann for this collaboration, but especially Dr Chris Heyes and Dr Wolfgang Schöpp for many fruitful discussions.

The author also thanks Dr Arne Drud of ARKI Consulting and Development A/S, Denmark, for continuously providing the latest versions of the Conopt libraries (Drud, 1996), and for his consultations on advanced topics related to formulation and solving nonlinear optimization problems.

6. REFERENCES

Alcamo, J., Shaw, R. and Hordijk, L., (Eds.) (1990). *The RAINS Model of Acidification*. Kluwer Academic Publishers. Dordrecht, Boston, London.

Drud, A.S. (1996). CONOPT: A system for large scale nonlinear optimization. Reference manual for CONOPT subroutine library. ARKI Consulting and Development A/S. Bagsvaerd, Denmark.

Drud, A.S. (1997). Interactions between nonlinear programming and modeling systems. *Mathematical Programming* **79**(1-3), 99–123.

Gondzio, J. (1997). Presolve analysis of linear programs prior to applying an interior point method. *INFORMS Journal on Computing* **9**(1), 73–91.

Heyes, C., W. Schöpp, M. Amann, I. Bertok, J. Cofała, F. Gyarfas, Z. Klimont, M. Makowski and S. Shibayev (1997). A model for optimizing strategies for controlling ground-level ozone in Europe. Interim Report IR-97-002. International Institute for Applied Systems Analysis. Laxenburg, Austria. Available on-line from `http://www.iiasa.ac.at/cgi-bin/pubsrch?IR97002`.

Koziol, Q. and R. Matzke (1998). HDF5 – a new generation of HDF. Reference manual and user guide. National Center for Supercomputing Applications. Champaign, Illinois, USA. `http://hdf.ncsa.uiuc.edu/nra/HDF5/`.

Lawrence, C.T., J.L. Zhou and A. Tits (1996). User's guide for CFSQP version 2.4: A C code for solving (large scale) constrained nonlinear (minimax) optimization problems, generating iterates satisfying all inequality constraints. Technical report. Institute for Systems Research, University of Maryland. College Park, Maryland.

Makowski, M. (1991). Selected issues of design and implementation of decision support systems. Working Paper WP-91-16. International Institute for Applied Systems Analysis. Laxenburg, Austria.

Makowski, M. (1998*a*). Generation and analysis of a non-linear optimization problem: European ozone model case study. In: *Proceedings of the 18th IFIP TC7 Conference* (M. Polis, Ed.). Addison Wesley Longman. (in press).

Makowski, M. (1998*b*). LP-DIT++, C++ class library for Data Interchange Tool for Linear Programming Problems, (version 2.06). Interim Report IR-98-xxx. International Institute for Applied Systems Analysis. Laxenburg, Austria. (to appear).

Murtagh, B.A. and M.A. Saunders (1987). MINOS 5.1 user's guide. Technical Report SOL 83-20R. Stanford University. Stanford, CA 94305-4022, USA.

Wolfram, S. (1996). *The Mathematica Book, Third Edition, Mathematica Version 3*. Cambridge University Press. Cambridge.

[2] CONOPT provides diagnostics (that trace also wrong scaling) based on analysis of Jacobian and Hessian.

[3] `http://www.iiasa.ac.at/Research/TAP/`.