Specification of Mathematical Model for Cost-effective Control of Acidification and Ground Level Ozone

Marek Makowski Christopher Heyes Wolfgang Schöpp

July 1, 1998

This technical note provides a specification of the mathematical programming problem that is used for examination of various cost-effective policies aimed at reduction of acidification, eutrophication and ground-level ozone concentration in Europe.

1 Model definition

We should distinguish first 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.

1.1 Notation

The model definition requires the use of the following indices:

- Index $i \in I$ corresponds to emitters. The numbers of elements in I corresponds to a number of countries (about 40).
- 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 the i-th country. Sets S_i may have up to 5 elements.
- Index $j \in J$ corresponds to receptors. The numbers of elements in J corresponds to numbers of grids (about 600).
- Index $l \in L$ corresponds to a combination of ozone thresholds and a year. The set L may have up to 6 elements.
- Index $m \in M$ corresponds to a set of receptors for which balancing of violations and surplusses of targets is defined.

1.1.1 Emission sectors

Emissions are analysed for sets of emitters that are located in a certain area, which is typically a country. However, for some types of analysis sets of NO_x and VOC emitters may be further subdivided into subsets called *sectors*, in order to account for measures that can be applied for emitters that belong to a sector. In such a case emitters that belong to a particular sector emit either NO_x or VOC or a linear combination of them. A sector is defined by a quadruple:

$$\{is, i, \lambda, \mu\} \tag{1}$$

where is is an integer number that uniquely identifies a sector, an index i corresponds to a country in which emitters belonging to this sector are located, and the following convention is used for the λ member of this quadruple:

$$\lambda = \begin{cases} 0. & \text{if only NO}_x \text{ is emitted} \\ \text{a negative number} & \text{if only VOC is emitted} \\ \text{a positive number} & \text{if both NO}_x \text{ and VOC are emitted} \end{cases} \tag{2}$$

Moreover, a (positive) value of λ defines the following relation between the amount of VOC emission corresponding to a unit emission of NO_x:

$$v_{is} = \lambda n_{is} + \mu \tag{3}$$

1.2 Decision variables

The main decision variables are the annual emissions of the following four types of primary air pollutants from either sectors or countries:
n_{is} - annual emission of NO_x (nitrogen oxides)
v_{is} - annual emission of VOCs (Volatile Organic Compounds)
a_i - annual emission of NH_3 (ammonia)
s_i - annual emission of SO_2 (sulphur dioxide)

Additionally, optional decision variables are considered for scenarios which allow limited violations of air quality targets. For such scenarios variables corresponding to each type of the considered air quality targets are defined for each receptor. Each variable represents a violation of a given environmental standard. Optionally, violations of targets can be balanced with surpluses (understood as a difference between a target and a corresponding actual concentration/deposition). 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).

Therefore, the following decision variables are optionally defined:
y_{lj} - violation of ozone exposure targets (surplus if $y_{lj} < 0$);
ya_j - violation of acidification targets (surplus, if $ya_j < 0$);
ye_j - violation of eutrophication targets (surplus, if $ye_j < 0$).

1.3 Outcome variables

The consequences of applications of computed 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.

1.3.1 Auxiliary variables

n_i - annual emission of NO_x (nitrogen oxides) defined by:

$$n_i = \sum_{is \in S_i} n_{is} \tag{4}$$

v_i - annual emission of VOCs (Volatile Organic Compounds) defined by:

$$v_i = \sum_{is \in S_i} v_{is} \tag{5}$$

en_{lj} - the mean effective emissions of NO_x experienced at j-th receptor is given by:

$$en_{lj} = \sum_{i \in I} e_{lij} n_i + enn_{lj} \tag{6}$$

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

nlv_{lj} - the representation of another non-linear term defining ozone exposure at j-th receptor is defined by:

$$nlv_{lj} = \sum_{i \in I} d_{lij} v_i \tag{7}$$

1.3.2 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.

Annual costs related to the reduction of a corresponding emission to a certain level are given by a piece-wise linear (PWL) function for each type of emission and for each emitter.

A piece-wise linear (PWL) function is defined for each member of sets of NO_x and/or VOC emitters declared by (1) and for NH_3 and SO_2 emitters (the latter for each country). PWL functions are defined by the following three triples (one file for NO_x and VOC jointly, one for SO_2 and one for NH_3):

$$\{id, em\_level, cost\} \tag{8}$$

where:

- *id* is the sector id for NO_x and VOC emissions (denoted by *is* in (1)), and emitter (country) id (i.e. i index) for SO_2 and for NH_3 cost functions,
- *em_level* is emission level of either NO_x (for sectors that emit only NO_x or NO_x and VOC) or VOC (for sectors that emit only VOC), or of NH_3 or of SO_2.
- *cost* is the cost of emission reduction to this level.

The sets of points $\{em\_level, cost\}$ for each *id* are such that a convex PWL cost function is defined by them. Formally, the following PWL functions define the annual cost related to reducing the level of emission to a level given by argument(s) of the function:

$ca_i(a_i)$ - for a_i

$cs_i(s_i)$ - for s_i

$c_i(n_i, v_i)$ - for n_i and v_i

The term $c_i(n_i, v_i)$ is defined by:

$$c_i(n_i, v_i) = \sum_{s \in S_i} c_s(\cdot) \tag{9}$$

where $c_s(\cdot)$ is a cost function for NO_x or for VOC or for joint NO_x and VOC reduction, depending on $type\_def$ defined by (2).

Each cost function will define its domain by specifying lower and upper bounds for its argument(s). This implicitly defines lower and upper bounds for all emissions that are used as bounds defined in Section 1.4.1. Those bounds may be made tighter by an optional specification of bounds for emissions from countries or sectors.

For the sake of brevity, the sum of costs is defined by:

$$cost = \sum_{i \in I} (ca_i(a_i) + cs_i(s_i) + c_i(n_i, v_i)) \tag{10}$$

Such a function is continuous and convex but it is not smooth. Therefore, it has to be represented by another function that meets typical requirements of non-linear solvers.

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:

$$aot_{lj} = \sum_{i \in I} (a_{lij} v_i + b_{lij} n_i + \gamma_{lij} n_i^2) + \alpha_{lj} en_{lj}^2 + \beta_{lj} en_{lj} nlv_{lj} + k_{lj} \tag{11}$$

$ac1_j$ - acidification of type 1:

$$ac1_j = tns_j (\sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + kn_j) + \sum_{i \in I} ts_{ij} s_i + ks_j \tag{12}$$

$ac2_j$ - acidification of type 2:

$$ac2_j = \sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + tss_j (\sum_{i \in I} ts_{ij} s_i + ks_j) + kn_j \tag{13}$$

eu_j - eutrophication:

$$eu_j = \sum_{i \in I} tn_{ij} n_i + \sum_{i \in I} ta_{ij} a_i + kn_j \tag{14}$$

where tn_{ij}, ta_{ij}, ts_{ij} are transfer coefficients for NO_x, NH_3 and SO_2, respectively; kn_j and ks_j are constants for nitrogen and sulphur background depositions; tns_{ij}, tss_{ij} are scaling coefficients that convert acidification coefficients of one type into acidification coefficients of another type, for NO_x and NH_3, and for SO_2, respectively.

Environmental effects caused by the two types of acidification and by eutrophication are evaluated at each receptor by a Piece-Wise Linear (PWL) function which represents an accumulated excess of each type of the air quality index:
$aac1_j$ - accumulated excess of acidification of type 1:

$$aac1_j = PWL_j^{ac1}(ac1_j) \tag{15}$$

$aac2_j$ - accumulated excess of acidification of type 2:

$$aac2_j = PWL_j^{ac2}(ac2_j) \tag{16}$$

aeu_j - accumulated excess of europhication:

$$aeu_j = PWL_j^{eu}(eu_j) \tag{17}$$

1.4 Constraints

1.4.1 Bounds

Each of the decision variables declared in Section 1.2 for $i \in I$ or for $is \in S_i$ is implicitly bounded by a corresponding definition of the domain of the corresponding cost function, which represents costs associated with the reduction of emission (see Section 1.3.2). This domain may be restricted by a specification of optional bounds (see below).

Emissions of NO_x and VOC from each sector are bounded:

$$n_{is}^{min} \leq n_{is} \leq n_{is}^{max} \tag{18}$$

$$v_{is}^{min} \leq v_{is} \leq v_{is}^{max} \tag{19}$$

The total emissions of NO_x and VOC from i-th country (defined as sums of emissions from sectors located in this country by eq. (4, 5)) are bounded:

$$n_i^{min} \leq n_i \leq n_i^{max} \tag{20}$$

$$v_i^{min} \leq v_i \leq v_i^{max} \tag{21}$$

The total emissions of NH_3 and SO_2 from i-th country are also bounded:

$$a_i^{min} \leq a_i \leq a_i^{max} \tag{22}$$

$$s_i^{min} \leq s_i \leq s_i^{max} \tag{23}$$

If a total emission of NO_x or VOC is fixed (which can also be implied by a small difference between corresponding lower and upper bounds), then variables corresponding to sectoral emissions defined by (4) and/or by (5) are not generated. In other words, a sectoral split of a fixed emission is not subject to optimization.

In particular, a re-definition of bounds may be used for fixing a certain emission at a given level.

Violations of targets are constrained at each receptor by corresponding lower and upper limits specified for each target type and for each grid:

$$y_{lj}^{min} \leq y_{lj} \leq y_{lj}^{max} \tag{24}$$

$$ya_j^{min} \leq ya_j \leq ya_j^{max} \tag{25}$$

$$ye_j^{min} \leq ye_j \leq ye_j^{max} \tag{26}$$

1.4.2 Complex constraints

The acumulative excess of long-term ozone exposure of l-th type is also constrained at each receptor by:

$$aot_{lj} - y_{lj} \leq aot_{lj}^{max} \tag{27}$$

where aot_{lj} is defined by (11) and aot_{lj}^{max} is a given maximum ozone exposure for l-th threshold at j-th receptor.

Constraint (27) without the term $-y_{lj}$ represents a so-called *hard constraint* for accumulated excess of ozone exposure and such a formulation is typically used in a traditional formulation of optimization problems. It can be used also in the presented model by selecting an option that does not allow for generation of variables y_{lj}. However, an assumption of hard constraints for air quality targets results in forcing more expensive solutions which are 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. 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 violation of targets and corresponding costs of reducing emissions.

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

$$aac1_j - ya_j \leq aac_j^{max} \tag{28}$$

$$aac2_j - ya_j \leq aac_j^{max} \tag{29}$$

$$aeu_j - ye_j \leq aeu_j^{max} \tag{30}$$

Optionally, violations of targets can be balanced with surplusses of targets over sets of receptors denoted in the following constraints by J_m, where $m \in M$ is the index of a set of receptors. Obviously, lower bounds in conditions (24), (25), (26) have to be negative in such a case. The balances are represented by the following constraints:

$$\sum_{j \in J_m} wo_{lmj} y_{lj} \leq tbo_{lm} \qquad l = 0 \tag{31}$$

$$\sum_{l=1}^{L} \sum_{j \in J_m} wo_{lmj} y_{lj} \leq \sum_{l=1}^{L} tbo_{lm} \tag{32}$$

$$\sum_{j \in J_m} wa_{mj} ya_j \leq tba_m \tag{33}$$

$$\sum_{j \in J_m} we_{mj} ye_j \leq tbe_m \tag{34}$$

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 are target balances for m-th set of receptors for l-th type of ozone exposure, for acidification, and for eutrophication, respectively. The sets J_m are defined implicitly by non-zero elements of sparse matrices wo_l, wa and we, respectively.

1.5 Goal function

A composite goal function is used for single criterion optimization of the non-linear ozone model in order to support scenario analysis of trade-offs between the following three goals:

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

The first two components have been already discussed therefore only the last one requires a 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) for 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 and solutions of slightly perturbed problems may differ substantially. However, from an application point of view an equally important indication of solution robustness is some measure of closeness of solutions of perturbed problems. Consider, for the sake of illustration, two instances of the ozone 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 be hardly acceptable. In order to avoid this problem, a technique called regularization is applied.

The goal function is defined by:

$$goal\_function = cost + penalty + reg \tag{35}$$

where the cost term is defined by (10), the penalty term is defined by:

$$penalty = \rho \sum_{l \in L} \sum_{j \in J} y_{lj}^{\xi} + \rho_a \sum_{j \in J} ya_j^{\xi} + \rho_e \sum_{j \in J} ye_j^{\xi} \tag{36}$$

and the regularization term is defined by:

$$reg = \epsilon \| z - \bar{z} \| \tag{37}$$

where ρ, ρ_a, ρ_e are given positive (not necessarily large) numbers, and ϵ is a given positive (not necessarily small) number. The penalty term exponent ξ is equal to 1, if the corresponding lower bound in conditions (24), (25), (26), is equal to 0; this exponent is equal to 2, if such a lower bound is negative. In the latter case the balances between violations of environmental targets and surpluses (given by eq. (31, 32, 33, 34)) are accounted for.

The interpretation of each of the terms is as follows:

- The first term corresponds to the sum of costs of emission 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 (27, 28, 29, 30).
- The third term is $\epsilon \| z - \bar{z} \|$, where z denotes a vector composed of all decision variables (except of y_j that are implied decision variables for which the reference point is implied to be 0. by virtue of the penalty term of the goal function). This is a regularizing term introduced in order to avoid large variations of solutions (with almost the same value of the original goal function) for problems that differ very little. This term assures that the optimal solution (for a problem that does not have a unique local optimal solution) will be the optimal solution closest to the point defined by given reference vector $\bar{z}$.