APR 3 1996

ORNL/M-4994

NATIONAL ACID PRECIPITATION ASSESSMENT PROGRAM

NAPAP

# Peer Review of the Tracking and Analysis Framework (TAF) for Use in the 1996 NAPAP Integrated Assessment

Hyatt Regency Hotel
Bethesda, Maryland
December 18-20, 1995

ORNL/M-4994

National Acid Precipitation Assessment Program

# PEER REVIEW OF THE TRACKING AND ANALYSIS FRAMEWORK (TAF) FOR USE IN THE 1996 NAPAP INTEGRATED ASSESSMENT

**Hyatt Regency Hotel**
**Bethesda, Maryland**
**December 18–20, 1995**

Organized by
OAK RIDGE NATIONAL LABORATORY
ORNL Assessment Services Office
Oak Ridge, Tennessee 37831
managed by
LOCKHEED MARTIN ENERGY RESEARCH CORP.
for the
U.S. DEPARTMENT OF ENERGY
under contract DE-AC05-96OR22464

## AGENDA

### Monday, December 18

| | |
|---|---|
| 8:30a–9:00a | *Welcome—NAPAP, Goals for Review*<br>Mike Uhart |
| 9:00a–9:50a | *Introduction and Objectives of TAF*<br>Cary Bloyd, Argonne National Laboratory<br>Max Henrion, Lumina Decision Systems |
| 9:50a–10:20a | *Peer Review Methodology*<br>John Malanchuk, International Technology Corporation |
| 10:30a–12:30p | *Presentation: Atmospheric Pathways Module*<br>Jack Shannon, Argonne National Laboratory<br>Ronald Marnicio, Foster Wheeler Environmental Corporation |
| 12:30p–1:30p | Lunch |
| 1:30p–3:10p | *Presentation: MAGIC Improvements Since 1990 NAPAP Assessment*<br>Tim Sullivan, E&S Environmental Chemistry, Inc.<br>Jack Cosby, University of Virginia |
| 3:10p–5:10p | *Presentation: Aquatics Effects Module*<br>Mitchell Small, Carnegie Mellon University<br><br>*Soils Effects Module*<br>Patrick Ryan, SAIC/Oak Ridge National Laboratory |
| 5:10p–5:30p | *Administrative Discussion—Write-ups*<br>John Malanchuk, International Technology Corporation |
| 8:00p–10:00p | Informal computer demonstration available |

### Tuesday, December 19

| | |
|---|---|
| 8:30a–10:10a | *Presentation: Visibility Module*<br>Jack Shannon, Argonne National Laboratory |
| 10:10a–12:10p | *Presentation: Emissions and Cost Modules*<br>John Molburg, Argonne National Laboratory<br>Jayant Kalagnanam, Carnegie Mellon University |
| 12:10p–1:10p | Lunch |

## Tuesday, December 19 (continued)

| | |
|---|---|
| 1:10p–3:10p | *Presentation: Health Effects Module*<br>Alan Krupnick, Resources for the Future<br><br>*Benefits Valuation Module*<br>Dallas Burtraw, Alan Krupnick, and David Austin, Resources for the Future |
| 3:10p–5:10p | *Presentation: Tracking and Analysis Framework (Collaborative Development of a Tool for Integrated Assessment)*<br>Max Henrion and Richard Sonnenblick, Lumina Decision Systems |
| 8:00p–10:00p | Informal computer demonstration available |

## Wednesday, December 20

| | |
|---|---|
| 8:00a–2:00p | (Closed meeting for reviewers only)<br>Panel discussion and preparation of presentations to Interagency Committee (45 min. each) (includes working lunch) |
| 2:00p–5:00p | Briefing to Interagency Committee (10 min.); Discussion of TAF Presentations (10 min. each) |

Time schedule for presentations:

| | |
|---|---|
| Presenter | 40 min. [except Pathways and Aquatics/Soils (50 min.)] |
| Questions | 40 min. |
| Administrative | 20 min. |

# EXECUTIVE SUMMARY

John L. Malanchuk, Ph.D.
International Technology Corp.

## INTRODUCTION

A peer review was convened December 18–20, 1995, to evaluate the NAPAP Tracking and Analysis Framework (TAF). The list of peer review members with their respective affiliations and responsibilities is given in Attachment 1. Substantial support for TAF was expressed. In fact, given the totality of constraints, time, availability of substitutes, etc., it is difficult to imagine how the NAPAP 1996 Integrated Assessment could be completed without reliance upon TAF and the TAF development team. Over the past eighteen months, substantial progress has been made in developing TAF and bringing the modules and framework to a usable point. In fact, reviewers commented that TAF represented a major advancement in our ability to perform integrated assessments. Significant work remains to be completed. This was part of the subject of this review. However, given the progress to date, the prognosis is excellent that TAF will be an important part of the 1996 assessment.

Logically, there were differences among reviewers and the degree of reviewer support for TAF and its various modules. However, there is no doubt that TAF provides a vital organizing framework for scenario evaluation and that the TAF development team should play an important role in the preparation, execution, and interpretation of the scenarios chosen for evaluation.

Recommendations also were made to improve TAF in the near term for use in the 1996 assessment and beyond, and these enhancements will serve only to strengthen TAF's credibility and flexibility as an assessment tool. Reviewers cautioned, however, that TAF not be applied or extrapolated to regions for which it has not been developed or intended. Finally, reviewers voiced strong sentiment that there can be no substitute for proper scientific and policy analysis in any assessment. Neither TAF nor any other assessment tool can be employed in a stand-alone fashion. Neither science nor public policy are well served if decision makers are provided information based upon poorly conceived or poorly executed technical analyses.

In this vein, adequate funding is imperative if NAPAP is to execute the mission given to it by Congress to perform the 1996 and 2000 integrated assessments. Major funding is not required; *adequate* funding *is* required, and it is the unanimous recommendation of these peer reviewers that NAPAP member agencies provide sufficient funding to conduct these assessments to evaluate the multibillion dollar investment already made on behalf of the American public.

## ORGANIZATION OF PEER REVIEW

Primary reviewers were designated for each of the areas to be reviewed. Originally there were eight categories to be reviewed: the seven modules and TAF itself. However, at the beginning of the review, it was decided that the Health Effects and the Benefits Valuation modules would be combined into one area and presented by Resources for the Future. Two primary reviewers were designated for each module, and three primary reviewers were designated for TAF overall. Remaining reviewers provided input to the primary reviewers, who employed this input as they deemed appropriate. Each primary reviewer was requested to produce a separate, independent report.

No reviewers for the health section were available during the review. The presentation was videotaped and sent to two reviewers. A subsequent conference call was held among presenters, reviewers, and NAPAP representatives, after which reviewers submitted their written reports.

Reviewers were given a series of standardized questions for each module and two standardized questions for TAF (Attachment 2). Module questions pertained to

- parent model credibility,
- reduced-form model credibility,
- suitability for use in NAPAP assessments,
- readiness for use in the 1996 NAPAP assessment, and
- specific recommendations for short-term improvement.

The two TAF questions concerned

- achievement of concept design goals and
- overall model credibility.

## EVALUATION AND RECOMMENDATIONS

This executive summary of reviewers' evaluations and recommendations, consisting of a compilation of primary reviewers' reports presented in the order of the agenda, attempts to capture associated key discussion. Table 1 quickly draws attention to areas in need of discussion and executive action and has been reviewed by the peer review team. Columns correspond to the five module questions and the two TAF questions discussed above; the footnotes are integral to proper understanding of the table. The table is not meant to substitute for the full write-ups prepared by the individual reviewers. Short-term recommendations are summarized after the table and are an integral part of the discussion.

**Table 1. Summary of reviewer responses to questions asked by NAPAP for each module and the Tracking and Analysis Framework (TAF)**

| Module/model | Parent model credibility | Reduced-form credibility | NAPAP 1996 | NAPAP 2000 |
|---|---|---|---|---|
| Atmospheric Pathways—$SO_x/NO_x$ | R[a] | R | R | R |
| Atmospheric—base cations | NA | U | U | U |
| MAGIC improvements | R | NA | R | R |
| MAGIC modules | | | | |
| Aquatics | R | R | R | R |
| Soils | NR | NR | NR | NR |
| Visibility | NA | R | R | U |
| Emissions/Cost | R | R | R | R |
| Health/Benefits | NA | R | R | R |
| TAF | R | NA | R | R |

[a] R = presently ready or can be made ready for use in the 1996 assessment.
NR = not ready for use in 1996 assessment.
U = unsatisfactory for use in assessment; see complete text for explanation.
NA = not applicable.

Long-term recommendations can be found in the individual reviewer reports. Short-term recommendations to be implemented prior to use in the 1996 assessment are as follows:

**Atmospheric Pathways—$SO_x/NO_x$**

- Perform additional module evaluation.
- Incorporate source-receptor matrix uncertainty into TAF.
- Address problem of location-dependent biases if shown to be significant.

**Atmospheric—Base Cations**

- Abandon modeling approach for 1996 assessment; use observed data and hypothesize trends for sensitivity studies concerning effects.

**MAGIC Improved Parent Structure**

- Improve Al-vs-pH relationship.
- Extrapolate to other regions only with great caution.

**MAGIC Modules—Aquatics/Soils**

Aquatics Module

- Make explicit in TAF the relationship between the Aquatics and Soils modules.
- Improve Al-vs-pH relationship (Aquatics Module).
- Make explicit the uncertainty characteristics in the ASI (acidic stress index) in the Aquatics Module.

Soils Module

- Incorporate lag-time dependency between effective deposition and soil base saturation (Soils Module).
- Develop relationships based upon "base saturation" instead of "changes to base saturation" in Soils Module.
- Examine significance of organic matter and carbon dioxide to base saturation predictions in Soils Module.

**Visibility Module**

- Examine significance of covariance between relative humidity and aerosol species.
- Use additional aerosol data to calibrate Visibility Module.

**Emissions/Cost Modules**

- Include in TAF, if desired.
- Evaluate significance of simplifications in cost model.
- Improve Canadian, Mexican, and $NO_x$ scenarios.
- Include scenarios with different geographic distribution of emissions.

**Health/Benefits Modules**

- Integrate how people actually view changes: thresholds, peaks, functional form.
- Quantitative coverage could be expanded to include forests, passive N fertilization, damage to car finish.
- Add treatment in the assessment of cultural materials and general existence/nonuse values.
- Provide clearer treatment of health and incorporate into uncertainty analysis the application of unit values to exposed and susceptible populations.
- Framework should be adapted to reflect important endpoints not quantifiable in monetary terms.

- Assessment should include discussion of limitations, including whether or not endpoints can be valued properly and whether or not concentration-response functions are sufficiently complete.
- Lack of formal treatment of ozone is a serious drawback in the 1996 assessment. Should NAPAP policy prevent inclusion of ozone, then its omission must be accomplished in a technically credible manner, fully caveated, and thoroughly discussed.
- Ambient temperature is an important consideration especially in evaluating the impact of PM10. To the extent possible, temperature effects should be included in the assessment.
- Whether or not formally included in TAF, the information contained in the Health/Benefits section is integral to the assessment. Estimates should be accompanied by uncertainty analyses, and methods of determining exposed and susceptible populations must be critically reviewed. Total health benefits should be distributed over different health impacts, e.g. chronic respiration, coughs, etc.

**Tracking and Analysis Framework**

- Stakeholders must be integrated into the TAF and assessment process.
- There should be an overall master plan, and someone must have lead and authority to make necessary assessment decisions.
- In the TAF causal chain, receptor modules (especially cost/benefits and effects) need to express input data requirements more specifically.
- TAF must be fully in the public domain, preferably DOS or Windows based.
- TAF must be made more transparent relative to underlying assumptions.

## ATMOSPHERIC PATHWAYS: $SO_x$ AND $NO_x$ SUBMODULE MODULE

**Presenter: Jack Shannon (Argonne National Laboratory)**

**Reviewer: Michael Moran**

### INTRODUCTION

The TAF Atmospheric Pathways Module is a key component of the overall TAF modeling system. It provides the link between anthropogenic emissions of $SO_x$ and $NO_x$ and their delivery to sensitive receptors. All of the TAF effects modules depend upon inputs from the Atmospheric Pathways Module.

This module is divided into two very different reduced-form submodules. The first, the $SO_x$ and $NO_x$ Submodule, consists of four seasonal sets of four atmospheric concentration and four deposition linear source-receptor matrices produced off-line by running the ASTRAP Lagrangian-statistical atmospheric long-range transport model. The second, the Additional Species Submodule, consists of a set of empirical, receptor-specific linear regression equations for 19 other quantities obtained from a limited set of precipitation-chemistry and dry-deposition station observations by means of stepwise multiple linear regression.

Given these very different formulations, these two submodules are reviewed separately.

### CREDIBILITY

The full-form "parent" model of the $SO_x$ and $NO_x$ Submodule is ASTRAP, the Advanced Statistical Trajectory Regional Air Pollution model, which was originally developed in the early 1980s by Jack Shannon of Argonne National Laboratory. As a single-level Lagrangian LRTAP model with highly parameterized representations of chemistry and deposition, ASTRAP was designed to calculate atmospheric average concentrations and wet and dry deposition of sulfur and nitrogen species over longer time periods on the order of seasons or years.

ASTRAP is much simpler in form than such comprehensive three-dimensional Eulerian episodic acid-deposition models as RADM, ADOM, and STEM-II. However, these latter models either require heroic efforts to be exercised directly for the long time periods that ASTRAP is routinely used for or else are subject to the additional approximations involved in using climatological and precipitation-chemistry statistics to estimate long-term patterns via the aggregation of a number of short-term episodes. Thus, a Lagrangian LRTAP model is a reasonable tool to use for predicting seasonal or annual atmospheric concentration or deposition patterns. It is worth noting that another Lagrangian model, the EMEP model, is used extensively in Europe in support of source-attribution and emissions-reduction planning work.

ASTRAP has been evaluated in a number of LRTAP model intercomparisons over the past decade and has performed respectably compared to other LRTAP models. These intercomparisons include the International Sulfur Deposition Model Evaluation (Clark et al. 1989) and several NAPAP intercomparisons (Binkowski et al. 1991, Shannon and Sisterson 1992). ASTRAP is also one of only a few Lagrangian models to have undergone continued development and improvement since the major NAPAP research program of the 1980s. ASTRAP is thus a credible choice from the available Lagrangian LRTAP models to serve as the TAF Atmospheric Pathways Module parent model.

It should be noted, however, that ASTRAP, unlike other Lagrangian LRTAP models, cannot be used to predict daily concentration and deposition fields. This restriction is due to its statistical representation of horizontal transport and diffusion and will present a problem for calculations requiring information about the variability of daily fields such as may be required in the Visibility Module.

The choice of source-receptor matrices as the form of the reduced-form model in TAF to represent atmospheric transport, diffusion, transformation, and deposition of sulfur species is also acceptable and scientifically credible. All other acid-deposition integrated assessment models developed to date, including the European RAINS and ASAM models (Alcamo et al. 1987, ApSimon et al. 1994) and the Canadian Integrated Assessment Model, have adopted the same approach.

In the case of nitrogen species, however, the use of linear source-receptor matrices is less good given the nonlinearity of the nitrogen system. For example, Olson et al. (1992) used a Lagrangian LRTAP model with nonlinear nitrogen chemistry and found nonlinear effects to be less than 10% for nitrogen wet deposition and less than 5% for nitrogen total deposition. The uncertainty introduced into the representation of the nitrogen cycle by linearization should be considered in the TAF uncertainty estimates. Note too that ammonia emissions are not considered by ASTRAP at the present time.

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

Most of these goals have been achieved. The use of source-receptor matrices permits easy substitution and upgrading to a new set of source-receptor matrices produced either by ASTRAP or by another LRTAP model. The source-receptor matrices also do not require a major simplification of the ASTRAP full-form model predictions.

One assumption that is required when using these source-receptor matrices to predict the concentration and deposition fields resulting from a change in emissions is that the horizontal and vertical distribution of emissions sources will remain constant when the magnitude of emissions within the source region is changed. The ASTRAP emissions inventory considers emissions over nine vertical layers and 127 × 127 km horizontal grid squares. For example, if emissions are reduced in one scenario by reducing emissions only for elevated sources, the use of the ASTRAP-derived source-receptor matrices assumes implicitly that these reductions are distributed over both ground-level and elevated sources in the same proportion as found in the original ASTRAP emissions inventory. Similarly, reductions in emissions obtained by shutting down a particular major point source such as a power plant will be modeled in the source-receptor calculation by distributing the reduction over all emissions grid squares contained in the source region of interest in proportion to the emissions associated with these grid squares.

It should be recognized too that the focus on only 15 irregularly spaced receptors in the present TAF source-receptor matrices does not permit spatial analyses of future-year concentration and deposition patterns to be undertaken because of the limited spatial coverage of these receptors. This may be a limitation in some policy analyses.

## READINESS

The $SO_x$ and $NO_x$ Submodule appears to be complete and ready to be used, with the one caveat that source-receptor matrix uncertainty does not appear to be specified or represented in the current version of TAF. Such uncertainty will include systematic ASTRAP biases as discussed in the next section.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENT

Only a limited comparison of the predictions of state-wide total wet deposition and dry deposition has been carried out against 1985–87 observations and RADM predictions for the new ASTRAP source-receptor matrices. However, a very useful additional performance evaluation could be carried out for little extra effort in the short term. For wet S and N deposition, 14 years of network precipitation-chemistry data are available for the years 1980–93. North American $SO_2$ emissions have decreased by about 23% during this period. Sulfur wet deposition could be estimated for the periods 1980–83 and 1990–93, and then the predicted change in this field could be compared against the observed change. Such additional evaluations would allow systematic biases in the source-receptor predictions to be identified and the module's performance in predicting changes in concentration and deposition fields due to changes in emissions fields to be evaluated.

The issue of source-receptor-relationship uncertainty still needs to be addressed. Interannual variations in meteorology will constitute one source of uncertainty. Quantification of this interannual variability could be used to specify one component of the intrinsic uncertainty in the source-receptor matrices in the short term. Additional uncertainty also arises because of random model prediction errors and, likely, systematic errors. No mention was made of how any source-receptor-relationship biases will be handled in TAF in the uncertainty estimation framework or otherwise.

Specification of the uncertainty arising from both random and systematic errors in the description of the source-receptor relationship must be addressed in the longer term.

One gap in the Atmospheric Pathways Module at present that may limit its usefulness in policy applications is the lack of a capability to calculate emissions reductions (most likely optimized by source region according to a specified cost function) given the deposition or concentration reductions required at a receptor site to meet some environmental objective or criterion. Such a capability has been very valuable in applications of the RAINS model and should be considered for implementation in TAF in the longer term.

Finally, in the longer term it may be necessary to consider the use of another parent model if ASTRAP's inability to predict daily fields becomes a significant restriction.

## REFERENCES

Alcamo, J., et al. 1987, "Acidification in Europe: A Simulation Model for Evaluating Control Strategies," *Ambio* 16, 232–45.

ApSimon, H. M., R. F. Warren, and J. J. N. Wilson 1994, "The Abatement Strategies Assessment Model—ASAM: Applications to Reductions of Sulphur Dioxide Emissions Across Europe," *Atmos. Environ.* 28, 649–63.

Binkowski, F. S., et al. 1991, *NAPAP SOS/T Report 3: Regional Acid Deposition Modeling*, National Acid Precipitation Assessment Program, Washington, D.C.

Clark, T. L., et al. 1989, "The Evaluation of Long-Term Sulfur Deposition Models," *Atmos. Environ.* 23, 2267–88.

Olson, M. P., J. W. Bottenheim, and K. K. Oikawa 1992, "Nitrogen Source-Receptor Matrices and Model Results for Eastern Canada," *Atmos. Environ.* 26A, 2323–40.

Shannon, J. D., and D. L. Sisterson 1992, "Estimation of S and $NO_x$-N Deposition Budgets for the United States and Canada," *Water, Air and Soil Pollu.* 63, 211–35.

<u>**Reviewer: Gary Stensland**</u>

## INTRODUCTION

The TAF Atmospheric Pathways Module predicts the seasonal and annual average atmospheric concentrations and cumulative wet and dry deposition of sulfur and nitrogen species. ASTRAP, a well-established model, is exercised off-line to produce linear source-receptor matrices that serve as the reduced-form model (RFM) for TAF. The RFM is designed to achieve the efficiency necessary for use as a component of an on-line integrated assessment.

## CREDIBILITY

The full-form model is ASTRAP, which has been in use and undergoing continuous development during the last 15 years and has been the subject of many peer-reviewed publications. Thus, the model has strong scientific credibility when used for appropriate time and space scales. It has been well tested in a variety of ways and is well regarded. An output of ASTRAP is a source-receptor matrix (SRM), which is the key component of the RFM for TAF. The TAF module should have strong credibility, as does ASTRAP itself. ASTRAP has been validated against both observations (dry and wet deposition) and against a very comprehensive process-oriented full-form module, namely RADM. Such validations should continue. Dr. Shannon points out that RADM simulations of long-term wet S deposition patterns indicated that the large time-scale relationship was approximately linear. This fact lends considerable credibility to using a linear model such as ASTRAP as the basis for the RFM in TAF. The 1992 Shannon and Sisterson paper summarizes another useful type of validation for ASTRAP, comparing horizontal net mass fluxes across the Atlantic coast of North America with estimates from field studies.

## ACHIEVEMENT OF DESIGN GOALS

The RFM has been implemented. Results for wet deposition of S in the Adirondacks were presented by Dr. Shannon and show the expected decrease in deposition over time. For this simulation the source-receptor matrix (generated off-line by exercising ASTRAP) had the sources aggregated by state and province, and fifteen receptors were considered. SRMs have been widely applied as RFMs, so the approach has credibility and provides considerable flexibility because it is straightforward to generate other SRMs (from ASTRAP runs) as new policy-relevant questions are formulated. Furthermore, as ASTRAP is modified to reflect new information or additional validation studies, then new sets of SRMs can be readily prepared for application in the RFM in TAF. Such ASTRAP improvements and the corresponding improvements in the RFM in TAF can be implemented without a long time lag, and this is an important advantage.

## READINESS

The RFM appears to be developed and running. However, it is not clear to this reviewer how much effort will be required to develop appropriate indicators of uncertainty that appear not presently to be in TAF.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENTS

The issue of incorporating uncertainty into TAF was not covered adequately in the presentation and materials by Dr. Shannon (considerable effort may have occurred that was not discussed). This issue needs to be thoroughly addressed in the short term and continued into the long term. The fact that the Aquatics Module for TAF is currently developed only for the single receptor, the Adirondacks, should be considered by Dr. Shannon and might suggest some short-term focused improvements in his Atmospheric Pathways Module to address relevant policy questions for this single region.

Dr. Shannon discussed several areas of future work that should be prioritized and addressed in the long term. Two areas are high on this reviewer's list:

- Dr. Shannon provided a rather lengthy list of changes in ASTRAP since about 1990 to improve accuracy and widen application. The RFM in TAF for atmospheric pathways will be given proper recognition and be thoroughly understood only if the out-of-date 1985 user's guide is updated. This new manual should also document the approaches used to address the issue of developing uncertainty values for TAF.

- Dr. Shannon should develop SRMs defined for specific major sources and then for other sources aggregated by state or province. This method will allow for explicit consideration of the impact sources on the Adirondack region emphasized in the Aquatics RFM.

## SUMMARY RECOMMENDATIONS

ASTRAP and its RFM in TAF should continue to be supported. The SRMs generated from ASTRAP should be intuitively attractive to the nonspecialist analyzing policy options with TAF.

In both the long and short term, TAF appears to be a very worthwhile endeavor to allow a relatively wide audience access to a tool that appears to be a reasonable balance between scientific rigor and ease of operation on widely available computer platforms. However, I would recommend that a version of TAF be developed to run on something other than Analytica™, a programming language that is not in widespread use.

The widespread educational use of TAF could become very substantial, and this possibility is itself an important justification for continuing development of TAF.

## ATMOSPHERIC PATHWAYS: ADDITIONAL SPECIES SUBMODULE

**Presenter: Ronald Marnicio (Foster Wheeler Environmental Corporation)**

**Reviewer: Michael Moran**

### INTRODUCTION

The Aquatics Effects Module requires information about depositions and concentrations of a number of species in addition to those predicted by ASTRAP. A good-faith effort has been made in this submodule to "piggyback" prediction of these quantities onto the ASTRAP predictions by carrying out regression analyses of station data to search for empirical relationships between the additional fields needed and the fields predicted by ASTRAP.

### CREDIBILITY

Unlike the $SO_x$ and $NO_x$ Submodule, there is no existing parent full-form model for this reduced-form model and hence no immediate credibility from previous peer reviews. The approach used is ad hoc, the results are poor overall, and the significant predictors vary from one empirical relation to another. There is also no reason to expect a priori that major anthropogenic $SO_x$ and $NO_x$ sources will be well correlated with base cation species. Overall, this approach is not very credible.

One additional issue is that this submodule adopts what is termed a "perfect prog" approach in statistical meteorology. That is, it assumes that the ASTRAP source-receptor matrices will produce perfect predictions. Any errors in the source-receptor predictions will propagate directly through the regression equations and add to the already considerable uncertainties in the regression relationships themselves.

### RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENT

A better approach for the 1996 assessment might be to use average observed values for the various additional species required, carry out bounding estimates in the Aquatics Effects Module, and plan to address this missing linkage before the year 2000 assessment. The inclusion of this submodule in TAF as currently proposed implies scientific credibility that is in fact lacking.

### SUMMARY RECOMMENDATIONS

The $SO_x$ and $NO_x$ Submodule appears to be sufficiently complete and credible to be used in TAF for the 1996 NAPAP assessment. However, module performance and uncertainty need to be better characterized. The Additional Species Submodule is not scientifically credible, and its use should be avoided if possible.

**Reviewer: Gary Stensland**

### INTRODUCTION

Dr. Ronald Marnicio of Foster Wheeler Environmental Corporation presented the approach and results for the TAF submodule that provides airborne concentrations and deposition of the other species: Ca, Mg, K, Na, Cl, $NH_4$, and H. The approach for this submodule is to use as input the results from the ASTRAP module (which predicts $SO_4$, $SO_2$, and

$NO_3$ concentrations and depositions) together with regression relationships between the other ions in precipitation and $SO_4$ and $NO_3$ in precipitation. When the regression relationships are found to be inadequate, measured air concentrations at one NDDN site in the Adirondacks are combined with the settling velocities and an empirical term describing the fraction of total deposition that is due to long-range transport to predict the deposition of the other ions. As implemented, air concentration data were reported for only a single Adirondack site, so only in this region is there an alternative to the regression approach.

## SCIENTIFIC CREDIBILITY

Several reviewers expressed concern about the regression approach as implemented. The concern is rooted in the fact that the source of other species, such as dust constituents, is likely not the same as the source of $SO_x$ and $NO_x$. Where regressions with current deposition data may suggest that for the last 10–15 years the dust sources may have been tracking smokestack sources of $SO_x$, this cannot in general be expected to hold for years or several decades into the future. Therefore, the credibility of a module assuming that $SO_x$ emission scenarios are useful as a surrogate for dust source trends is very questionable. The same concern holds for ammonium ion. Only for hydrogen ion does it seem credible to use ASTRAP output to evaluate effects of various emission scenarios for $SO_x$ and $NO_x$.

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

The TAF design goals with respect to other species cannot all be achieved because of the lack of a credible model to predict changes over time in dust and ammonia sources.

## READINESS

A default approach for TAF for the March 1996 assessment is simply to assign the other ion values based on observations and to do sensitivity tests based on hypothetical changes in the observed levels. The investigator for this project has organized TAF to accept such assigned values. However, the investigator does not appear to have experience in making dry-deposition measurements for the other species, nor is he currently familiar with the characteristics of various data sets and modeling techniques for inferential methods or extrapolations. Consequently, some time will be required to travel the learning curve and thus become skilled in critically evaluating available data sets to select averages and standard deviations for concentrations of other species.

## RECOMMENDATION FOR IMPROVEMENT

Implement an approach for TAF for the March 1996 assessment wherein other ion values are assigned values based on observations and not determined from a model. Considerable effort would seem to be required for the current principal investigator to become knowledgeable with air-concentration/dry-deposition/dustfall measurements that should be considered for incorporation into TAF. Such experience is necessary to choose reasonable values and uncertainties and to discuss potential biases between data sets. This change from a combination of modeling and the direct use of observations to an approach using only observations should be ready for implementation by March 1996, especially if someone familiar with the various dry-deposition/dryfall data sets is utilized.

## SUMMARY RECOMMENDATIONS

Considering the lack of knowledge about sources and transport for the other species, it seems best to assign other species concentrations and not attempt to model them with the regression approach now in TAF. It will probably be necessary to state clearly to the TAF user that the measurements used to assign mean values in TAF for dry deposition of the dust type species are controversial, because large biases between the data sets are to be expected based on the lack of widely accepted measurement approaches.

From the Aquatics and Soils Effects presentations, it does seem that the simple approach of providing air concentrations means, uncertainties, and hypothetical trends for the other species based on a careful review of ambient air literature for dust species will be useful for sensitivity studies of the effects modules. It might also be useful to do sensitivity studies on Aquatics Effects due to the very rare but extreme episodes of high dustfall that might be hypothesized. Someone familiar with the geological/soils science literature might be able to provide input on extreme dry-deposition values that have been observed or inferred for desert dust storms, volcanic dust fallout, etc.

## MAGIC IMPROVEMENTS SINCE 1990 NAPAP ASSESSMENT

**Presenters: Tim Sullivan (E&S Environmental Chemistry, Inc.) and Jack Cosby (University of Virginia)**

**Reviewer: David Lam**

### INTRODUCTION

The MAGIC (Model of Acidification of Groundwater In Catchments) Model was used in the NAPAP Integrated Assessment in 1990. It has been successfully applied to answer scientific and policy questions on changes in aquatics chemistry due to acidic precipitation in many watersheds in the United States. As true in many models, there is always room for improvement. MAGIC is no exception. Its developers, a dedicated team of world-class researchers, were aware of its shortcomings at the time of the 1990 assessment on the formulations of natural occurring acids (organic acids), the dynamics of different aluminum compounds, and the considerations for including nitrogen oxides as part of the acid-deposition input. The development of TAF has provided an excellent opportunity for making these improvements in MAGIC.

### CREDIBILITY OF MODEL CALCULATIONS

#### Organic Acid Formulations

For lakes in the Adirondack areas in northeastern United States, model credibility was enhanced, showing favorable comparison of observed data, including estimated data that were used to extrapolate back to preindustrial times, with model results using organic acids.

#### Aluminum Dynamics

Again mainly for lakes in the Adirondack area, the need for improved formulations for aluminum was demonstrated. Further scientific credibility was gained by successfully applying the new model formulation to so-called manipulated watersheds, i.e., watersheds that were artificially controlled by experiments on their acid-deposition inputs, in both the United States and Norway.

#### Nitrogen Oxides

Several assumptions, such as constant nutrient uptake by algae and separate conditions for increasing and decreasing nitrogen-oxide-loading scenarios, were required to make the model sufficiently credible in its first attempt to include the nitrogen component. The major finding was that nitrogen oxides were, for the most part, five to ten times less in their acidifying power than equivalent amounts of sulfur dioxide, based on model predictions for Adirondack lakes.

### ACHIEVEMENT OF MODEL DESIGN GOALS

#### Open Architecture with Accessibility for Substitution and Upgrade

In the context of improvements made to the original or parent model, fairly effective use of computer programming was carried out in modification of the three new model components.

### Utility of Reduced-Form Structure with Acceptable Additional Uncertainty

While not explicitly discussed in the presentation for this module in TAF, subsequent presentations on the Soils and Aquatics Effects modules showed very convincingly the benefits of using reduced-form models derived from the parent model. The reduced-form models were simpler than the parent model but were shown to produce similar results. The uncertainties were included in the predictions of both soils and aquatics effects through the use of mean conditions and probable extreme bounds of model results.

### Transparency of Assumptions

The presenters made a conscientious effort to describe the detailed assumptions used in the new model formulations, particularly for the nitrogen component. It seemed that, while the nitrogen component is probably the most inconsequential in the prediction of soils and aquatics effects among the three new additions to MAGIC, the assumptions for the nitrogen component were made very clear, although they turned out to be the most awkward to explain. For example, in the new MAGIC model, the ammonium deposition was assumed to be constant, whereas the Atmospheric Pathways Module reported a robust predictive relation, showing that ammonium concentration varied directly with sulfate concentration in the deposition data collected in the same area.

### Analysis and Presentation of Sensitivity

A good discussion was made by the presenter, for example, on the sensitivity of the aluminum-vs-pH relationships for the manipulated watersheds in Norway and the United States. Depending on the ratio chosen for model coefficients relating these two parameters, one can make the pH predictions better at the expense of the aluminum results and vice versa.

### Robustness for Addressing Alternative Issues

The introduction of nitrogen component is a first step to expand MAGIC to other forms of acidifying agents. However, model comparison with more receptor sites is required to establish its robustness.

## READINESS

On the one hand, the MAGIC model was used successfully for scientific input to integrated assessment in 1990. Therefore, it is ready for another round of policy input such as for the Title IV questions for the 1996 NAPAP assessment. The preliminary results using data from Adirondack lakes were very convincing. On the other hand, many reviewers including this one are of the opinion that caution needs to be exercised if and when MAGIC, with the newly added components—particularly the nitrogen module, is used for extrapolation to other parts of the United States. Applications to other receptor sites beyond the Adirondack area are required to establish the readiness and to remove this caution, which was raised by many reviewers during this meeting.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENT

While there seemed to be excellent dialogues between the parent model developers and the reduced-form model developers, better communication between these modules and the Atmospheric Pathways Module may lead to a quicker resolution on the consistency of emission-deposition scenario data used in TAF. Over the short term, it is worthwhile to explain how uncertainties could propagate from one part of TAF to another. If successful, these uncertainty analysis results could make the use of TAF for the 1996 NAPAP assessment even more convincing. Over the long term, more receptor sites are required to establish further the credibility of the improved parent model and also to confirm the reduced-form models. Also, one should carefully examine whether the reduced-form information from the Atmospheric Pathways Module may help sharpen the predictions. For example, one should re-examine the consistency on deposition data from the Atmospheric Pathways Module, such as calcium and ammonium deposition and their effects on soil and water chemistry. The investigation of these and other linkages in TAF should be made an iterative and long-term goal.

## SUMMARY RECOMMENDATIONS

On the whole, the MAGIC model with its recent improvement is a big step toward demonstrating its usefulness in TAF. It produced by far the best reduced-form models. The emulation of the parent model results by their reduced-form equivalents in predicting changes in soil and water chemistry is most remarkable. The use of estimated prehistorical data and data from the manipulated watersheds produced excellent examples for demonstrating how effects due to control of sulfur deposition can be predicted. Likewise, the use of the Adirondack data can be used to demonstrate potential changes for various loading scenarios in one of the most sensitive regions in the United States under the influence of acid rain. If used properly, these results will make it useful for the 1996 NAPAP assessment. One limitation is that the conclusions are no better than those made for the Adirondack areas, for the most part, and therefore caution must be exercised not to extrapolate the results to other parts of the United States, pending further confirmations with other receptor sites. In the long term, TAF is a worthwhile exercise that will see iterative improvement in all its modules.

## AQUATICS EFFECTS MODULE
## SOILS EFFECTS MODULE

**Presenters: Mitchell Small (Carnegie Mellon University) and Patrick Ryan (SAIC/ORNL)**

**Reviewer: Nikos P. Nikolaidis**

### INTRODUCTION

The Aquatics and Soils modules of TAF are "reduced-form" models based on the results derived from the MAGIC model. MAGIC is a spatially lumped parameter model that uses the principles of chemical equilibrium and mass balance equations to simulate the effects of acidic deposition to lake and watershed chemistry. MAGIC operates on a temporal scale of years to decades, and it was not designed to reproduce episodic event changes in chemistry. The model has been tested extensively during the past decade, and it has undergone comparisons with other models that have more detailed parameterization. MAGIC has already been used successfully to address policy-related questions as part of the 1990 NAPAP Integrated Assessment. The MAGIC model is the best choice in the selection of full-scale watershed models for the development of the Aquatics and Soils modules. Its input data requirements and computer time resources are significantly less than other freshwater acidification models, yet the model can simulate the major biogeochemical processes that control the water quality in watersheds.

### CREDIBILITY OF MODEL CALCULATIONS

The Aquatics and Soils modules of TAF represent a simplified representation of the MAGIC model. The modules can predict the concentrations of pH, calcium, aluminum, and acid-neutralizing capacity in the lake and the base saturation of soils. These parameters are further used to evaluate the effects of the predicted lake chemistry on the biological community through the calculation of indices such as the Acid Stress Index (ASI). The ASI index has been field collaborated. The Aquatics Module was shown to have excellent fidelity to the parent model results. It was shown that the uncertainty introduced by the reduced-form models is small compared to the uncertainty introduced by the parent model.

This is a credible approach that is consistent with the questions that relate to the 1996 NAPAP assessment. However, it is very important to stress that the modules have been tested only for the Adirondack subregion using data from 33 lakes. The modules should be tested using data for other regions in order to verify and validate their conceptual framework. The modules will not be robust unless the concept is shown to work in other regions/receptors.

### ACHIEVEMENT OF MODEL DESIGN GOALS

The implementation of the conceptual framework of the Aquatics and Soils modules in the TAF model is completed for the Adirondack subregion. Suggested changes and improvements in the modules will not delay the implementation of the 1996 NAPAP assessment. The modules have utilized an open architecture that makes upgrades easy to implement. This architectural style will aid in applying the TAF model to other acid-sensitive regions of the United States.

The first phase of uncertainty analysis of the modules has produced acceptable results. The modules produce a comparatively small additional uncertainty to the results compared with the inherent uncertainty produced by the MAGIC model. The planned activities that focus in better quantifying the additional uncertainty are expected to address this issue fully. Overall, the modules are appropriately detailed for addressing alternative scenarios policy questions consistent with the 1996 NAPAP assessment. To evaluate the robustness of the modules, it is necessary to show that the concept is working well for other regions.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENTS

There are several potential changes and improvements that could improve the results for the 1996 NAPAP assessment:

1. Re-examine the aluminum-vs-pH relationship to account for lake-to-lake variation rather than a regional basis. This could be done either by defining an aluminum adjustment factor (ALAF) where the regional relationship prediction is adjusted by ALAF on a lake-by-lake basis (similar to the watershed-specific deposition adjustment factors) or by defining an aluminum-vs-pH slope that would best fit each lake. Either approach should improve the aluminum predictions. This improvement is necessary because aluminum concentrations play a significant role in the ASI prediction.

2. The uncertainty characteristics of the ASI should be explicit and carried through. The index is a regression model with uncertainties that can be explicitly incorporated in TAF.

3. The relationship between the Aquatics and Soils modules needs to be explicit in the TAF model. As it presently stands, the two modules appear to be disconnected. After a close examination, it becomes apparent that the two modules are closely related because their relationships are produced from the same MAGIC simulations. In addition, the Aquatics Module has explicit parameters that relate to soil water chemistry and the geochemical processes that take place in the terrestrial portion of the watershed.

4. Evaluate the importance of calcium to the prediction of ASI. This evaluation should be conducted in light of the great uncertainties involved in the calcium-deposition predictions. For example, the fraction of calcium from deposition as opposed to the one generated by the watershed should be calculated.

5. With respect to the Soils Module, the following improvements could be made to reduce the uncertainty in the predicted results:

   a. Incorporate a lag-time dependency (similar to ones used in the Aquatics Module) between effective deposition and soil base saturation.

   b. Attempt to relate base saturation rather than changes to base saturation to effective deposition.

   c. Conduct a sensitivity analysis to examine the importance of soil organic matter and carbon dioxide concentrations to the predicted base saturation. If they are important, change the formulation of the module appropriately.

   d. Include a forest effect index model. Such an index exists in both Canada and Norway. They typically relate the calcium-to-aluminum ratio and other factors to the percent of canopy damaged.

The long-term improvement in the TAF Aquatics and Soils modules should include:

1. Application to other regions. From the perspective of robustness, the modules can be claimed robust only if they are applied successfully to other regions. In addition, from the economic perspective and fisheries, the Adirondack region is not the most important or sensitive region. This shortcoming necessitates the application of the framework to other regions.

2. All the modification and improvements conducted to the MAGIC model should also be carried through to the reduced-form modules.

**Reviewer: Arthur Bulger**

Both modules appear to be going in the right direction. The Aquatics Module is ready or nearly ready now; the Soils Module needs more work. Both are geographically limited at this time in their application to the Adirondacks region. The reviewers were optimistic that the reduced-form models can be successfully developed for other regions, but the reduced-form models for the Adirondacks are probably not exportable as is.

## INTRODUCTION

The Aquatics Module presentation by Small followed the MAGIC and Soils Module presentations by Sullivan and Ryan. These three topics were well-integrated, as is appropriate.

The full-form model (FFM) for the Soils and Aquatics modules is MAGIC, an intermediate-complexity, lumped-parameter model developed by Jack Cosby. MAGIC predicts a number of variables, but only those most strongly linked to aquatic effects have been developed for the Aquatics Module. A different variable has been developed for prediction in the Soils Module.

Improvements added to MAGIC since 1990 will make it more useful for these applications. MAGIC uses precipitation chemistry and soil chemistry to predict lake or stream chemistry, and its output chemistry can be used to predict biotic response. The modeled biotic response in the Aquatics Module is the ASI. It is a characteristic of a water body reflecting how toxic it is expected to be as a result of acidification. The ASI formulae used are the same for both the FFM and the RFM, and are discussed in SOS/T 13, NAPAP 1990. The formulae include the variables pH, calcium, and aluminum, which are the most important in understanding fish response. Low pH is toxic to fish and also mobilizes toxic aluminum from soils; calcium ameliorates the toxic effects of low pH and aluminum.

The Soils Module predicts percent base saturation, basically a sensitivity character.

The Soils and Aquatics modules have been developed solely for the Adirondacks subregion.

## CREDIBILITY OF MODEL/MODULE DESIGN GOALS

MAGIC has been well tested in a variety of acidification contexts and is highly regarded. The comparisons provided show excellent agreement between FFM and the Aquatics RFM outputs for ANC, pH, and calcium, as well as the ASI. The ASI ranges from 0 (no acid stress) to 100 (total mortality) and represents the percent mortality expected from a particular combination of pH, aluminum, and calcium. Much lab toxicity work goes into the development of each ASI formula. Because it is impractical to develop toxicity functions for all fish species, three fish species are taken to represent three sensitivity classes. ASIs are calculated for tolerant, sensitive, and intermediate fish species (brook trout, smallmouth bass, and rainbow trout, respectively). Both MAGIC and the ASI passed peer review tests for the 1990 document.

It is very important to emphasize that the ASI is an acute toxicity index developed in constant-condition lab studies, and it appears to underestimate negative effects (percent mortality) when related to base-flow chemistry conditions in nature, probably because baseflow chemistry conditions do not reflect acid episodes to which fish in nature are exposed. Nevertheless, while there is not a one-to-one relationship between the mortality in nature and the ASI, an upward trend in ASI over time denotes increasing acid stress in nature. This relationship is well corroborated with field studies.

The comparisons provided to us showed good agreement between FFM and the Soils RFM outputs for change in percent base saturation. One of the reviewers was concerned about the one-parameter model to replace structured variation in the response of watersheds to deposition and the biases that might be thereby introduced. The Soils RFM appears acceptable for the Adirondacks but would have to be recast for other regions.

Others thought that the uncertainty considerations were too "fancy" and that the supporting data were absent. One reviewer suggested more emphasis on uncertainty analysis. The RFM appears to introduce only a small fraction of the uncertainty inherent in the FFM.

In the Soils Module, the approach to uncertainty analysis appears to be good; it is a bit unclear how the RFM is derived, but it shows good agreement with the FFM. The bounds on deposition (+75% and ramps) seem appropriate.

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

The modules have open architecture, are apparently easy to modify, and have well-defined assumptions.

The Aquatics objectives are to model the changes in lake chemistry as a result of changes in acid deposition and to model the effects of lake chemistry on fish populations. The Aquatics Module can accomplish its objectives now for the Adirondacks. MAGIC produces a number of modeled parameters, but the Aquatics RFM has been developed only to use pH, calcium, and aluminum, which are used in the fish effects component, the ASI. If the goal is solely the production of input variables for the ASI, this method seems adequate. One reviewer suggested that it would be valuable to develop the Aquatics RFM for other MAGIC variables for completeness so that policy makers could select a desired outcome (such as a sulfate concentration) and work "backwards" to identify the deposition that would produce that outcome. Cosby pointed out that MAGIC could not be used in this way. He described a process of altering the depositions until the desired outcome was achieved. Lam also asked if the RFM output was constrained by mass/charge balance considerations. Apparently it is not; rather it is simply a statistical relationship with MAGIC's output. It was suggested that the charge/mass balance issue be considered for scientific credibility.

The Soils objectives are to assess changes in soil chemistry (as base saturation) that may be related to forest productivity or ecosystem function. The Soils Module can predict base saturation now, but it is not linked to effects. MAGIC produces a number of modeled parameters, but it is my understanding that the RFM has been developed only to percent base saturation. It is also my understanding that links have not yet been made to forest effects or ecosystem function.

The robustness of both Soils and Aquatics modules will depend on how well they can be made to work for other regions; the reviewers were optimistic about applications to other regions.

## READINESS

The Aquatics Module may be ready; whether it will be fully functional by the deadline requires clarification. It is probably in the best condition of all the modules.

The Soils Module does not appear to be ready in the sense that linkages to effects have not been made. However, percent of watersheds with low base saturation is available for model exercise.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENT

1. There is a real or perceived disjunction between the Soils and Aquatics modules. Because both modules use MAGIC, it would be helpful to have a brief text description of how they are related conceptually and how percent base saturation and the soil parameters of the Aquatics Module are related. It is suggested that the authors (in concert with the Aquatics Module authors) offer a rationalization for the lack of correlation between base saturation and lake ANC.

2. Clarification of the weighting of effects in the Aquatics Module was suggested because overall damage might be underestimated if much of the damage occurs in a small percentage of lakes.

3. It would be very valuable to develop the Aquatics Module for other regions; all the reviewers who gave their comments suggested this step. The mid-Appalachians and Southern Blue Ridge areas in particular contain many streams of intermediate pH and low ANC and are vulnerable to future acidification. It is possible that the Aquatics Module would identify a greater change in pH and the ASI over time in these areas versus the Adirondacks. One reviewer pointed out that this possibility would be more likely if it could be shown that the pH distribution of unacidified areas is not bimodal, as it often is in acidified areas: acidification moves mid-ANC lakes to the low ANC category.

4. In the Aquatics Module, it may be possible to improve the overall aluminum and pH prediction relationship for a region by including lake-specific aluminum-vs-pH relationships.

5. A comparison of region-to-region variability versus lake-to-lake calibration results in the expansion of the Aquatics Module to other regions was suggested.

6. Continue planned reformulation of the Soils RFM to incorporate exponential approach to equilibrium for base saturation.

7. It would be valuable, given resources, to develop the Soils Module for other regions.

## SUMMARY RECOMMENDATIONS

MAGIC continues to be a good choice for effects modeling, and the RFMs appears to mimic MAGIC's output well, at least for the region so far developed (Adirondacks); consequently, they should be developed for other areas to generalize their utility. The RFMs should be developed from the FFM for each region because it is likely that the colinearity structure among the input variables differs among regions, and this could affect the fidelity of the RFMs to the FFM. Continue planned work on incorporation of nitrogen. The ASI coupled to the chemistry models appears to be a good procedure for modeling an index of aquatic biotic response.

## VISIBILITY MODULE

**Presenter: Jack Shannon (Argonne National Laboratory)**

**Reviewer: Rudolf Husar**

### INTRODUCTION

The design goal of the Visibility Module is to estimate the daily and seasonal visibility changes due to sulfate and nitrate concentrations. Shannon's approach uses as input seasonal average sulfate and nitrate concentrations provided by the ASTRAP long-range transport models. Within the Visibility Module, the seasonal average aerosol concentrations are transformed into daily extinction coefficients using the Visibility Assessment Scoping Model (VASM). Within VASM the short-term (daily) aerosol concentration is computed from the long-term (seasonal) values using measured distribution functions at (IMPROVE) monitoring sites.

The randomized sulfate concentrations are converted into an extinction coefficient using climatological relative humidity values. The daily relative humidity (RH) is calculated by randomizing the climatological RH values. Finally, the randomized aerosol species are converted to extinction coefficient using an RH-dependent correction factor. The result of these computations is a distribution function of extinction coefficients that is expressed in deciview units.

### CREDIBILITY OF THE VISIBILITY MODULE

The Visibility Module is based explicitly on empirically derived distribution functions. Its strength is derived from the fact that the existing observations of aerosol species were used in the derivation of the distribution functions. Hence, it approximates reality reasonably well. The weaknesses of the model are also attributable to its empirical character. The model results are applicable only to the geographic locations and time epochs for which detailed aerosol chemical data are available. An additional weakness is transmitted from the ASTRAP Atmospheric Pathways Module, which provides only seasonal average aerosol concentrations. In fact, the need of stochastic perturbations in VASM along with the statistical assumptions could be eliminated entirely if the ambient concentrations were provided daily or hourly.

### ACHIEVEMENT OF THE MODULE DESIGN GOALS

In general, the model has achieved its stated goal of estimating changes in the extinction coefficient.

1. Open architecture: in principle, the model is well suited for open architecture. However, the distribution functions of aerosol species and for relative humidity were not accessible to this reviewer.

2. Utility of reduced-form model: the reduced-form model is virtually identical to the full VASM model; hence, there is no significant loss from the "parent."

3. Transparency of the assumptions: conceptually, the model is simple and transparent. However, the actual distribution functions were not accessible.

4. Sensitivity analysis: evidently a sensitivity analysis to different model assumptions has not been completed.

5. Robustness and flexibility for alternative approaches: the empirical Visibility Module is not easily extensible to alternative approaches. In fact, it appears to be an expedient but also a dead-end approach.

## READINESS

The Visibility Module is in a good state of readiness. Additional calibration would be desirable to improve the empirical fit of the measured distribution functions.

## RECOMMENDATIONS

Short-term improvements can be achieved by using additional aerosol chemistry data sets, more detailed examination, and statistical consideration of covariance between relative humidity and aerosol species.

In the long-term (e.g., the year 2000 assessment), this methodology should be replaced by dynamic daily model simulations as stated in the recommendations below.

## SUMMARY RECOMMENDATIONS

The Visibility Module provided by Shannon is a suitable short-term approach for visibility effects assessment. Being largely empirical, it adequately represents the reality where and when monitoring data are available. However, the statistical-empirical approach has inherent limitations that make it unsuitable for evolutionary development. Rather, the recommendation is that, for the long term, the approach be replaced by a dynamic and physically based relationship between aerosol chemical species, relative humidity, particle size, and extinction coefficient.

Reviewer: Daniel McNaughton

## INTRODUCTION

The Visibility Module of TAF serves as a link to tie atmospheric modeling of sulfate and nitrate concentrations to a valuation of changes in visibility impairment due to regional haze. Inputs used in the model are seasonal average concentrations of sulfate and nitrate aerosols from which light attenuation or extinction is estimated. This results in output expressed as visual range at each of seven selected receptors (two national parks and five urban areas). The model selected for TAF is a modified version of the Visibility Assessment Scoping Model (VASM) developed by the Department of Energy (Trexler and Shannon 1995).

Visibility models range in complexity from simple extinction relationships such as those used in regulatory modeling to complex aerosol dynamics models formulated as components of Eulerian grid models. VASM is designed as a stand-alone model to be applied using outputs of the ASTRAP air quality model. The Visibility Module is applied to the reduced-form (transfer matrix) Air Pathway Module in TAF, but the module can be used with an appropriate alternative air pathway model.

In its current configuration, the Visibility Module serves multiple functions including, most importantly, (a) extinction of light due to sulfates and nitrates in the context of all major light-scattering and -absorbing pollutants and (b) the estimation of short-term visual range values from the seasonal values input. The former requires parameterizations of the extinction efficiencies, particle size distributions, the effects of relative humidity on aerosol growth, nitrate aerosol speciation, and specification of concentration distributions for elemental carbon, fine and coarse dust, and organic carbon. The latter involves a Monte Carlo technique used to develop short-term distributions. Much of the effort in the Visibility Module, and many of the limitations, result from this attempt to derive short-term results from climatological averages.

The VASM was selected as an available model that is sufficiently complete and efficient to fit in TAF.

## CREDIBILITY OF THE MODEL/MODULE CALCULATIONS

Review panel comments on the Visibility Module ranged from acceptance to rejection, with the general finding being that the model is credible but only as an intermediate and approximate technique. Several elements of the model design were considered to be not fully complete or adequate for use in a final integrated assessment. Specific comments include:

1. Modeling for this module and all TAF effects modules should be enhanced to provide wider geographic coverage to resources at risk in the NAPAP region.

2. Confidence in TAF will result from careful quantification of module uncertainties through model evaluations of the reduced-form modules. This quantification is required for ASTRAP transfer matrices and Visibility Module estimates relative to observations. Evaluations should be performed with consistent summarization of observations, internal model parameters, and model outputs; that is, evaluation simulations and observations must have consistent assumptions as to scale, limitations, phenomena simulated, and scenarios evaluated. Initial evaluations of the model results suggest an overprediction of visibility improvement.

3. The ASTRAP/VASM pairing does not adequately address urban visibility receptors because of its regional scale and the local concentration variability characteristic of urban areas. The current model may be adequate to assess the changes due to contributions to regional air quality, but this may not be adequate for use in the valuation module.

4. Local-scale effects are also important in the use of Atlantic City as an urban receptor and the possible use of near-coastal sites as a data source for parameter development. At Atlantic City, the maritime influence introduces a significant local variability and could potentially result in the decoupling of the local and regional components to visual range.

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

The Visibility Module was successfully implemented as a highly parameterized model amenable to change or substitution. Explicit introduction of uncertainty estimates for the TAF uncertainty analysis should be pursued. There is no true parent model for the Visibility Module, as it differs little from VASM, but the potential exists for a range of relatively simple model improvements or replacement for later, more complete assessments. Replacement of the module or a change from ASTRAP as the input source to the Visibility Module may require modifications to accommodate inputs of shorter term than seasonal. Some reviewers recommended the explicit treatment of peak short-term concentrations to understand the tails of the visual range distributions. The module lacks some generality in its application because numerous visibility-degrading pollutants are specified rather than modeled.

## READINESS

The ASTRAP/VASM pairing represents a significant body of work that resulted in a simplified approach to regional visibility assessment. The review panel considered the module to be approximate and intermediate. It represents a complete modeling module that provides useful results for first approximations of visibility effects and valuation and a means of assessing the relative importance of the processes relative to other TAF components. Prior to its use in an interim assessment, further testing of this model is required to quantify module uncertainty relative to observed data.

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENTS

In the short term, the Visibility Module in TAF should undergo enhanced model evaluation to better quantify uncertainties prior to its use in the 1996 assessment. For the long term, an alternative model is probably indicated to introduce more complete parameterizations aerosols other than sulfate and nitrate and to replace the highly parameterized relationships currently used to estimate short-term concentration distributions from seasonal values. Some guidance in refining relationships should be available from complementary studies such as the Grand Canyon Visibility Transport Commission Study.

## SUMMARY RECOMMENDATIONS

The current Visibility Module is adequate for use in the 1996 NAPAP assessment as an approximate technique to visibility assessment. This use requires additional evaluation and uncertainty quantification.

## REFERENCE

Trexler, E. C., and J. D. Shannon 1995, "Assessing the Potential Visibility Benefits of Clean Air Act Title IV Emission Reductions," presented at Acid Rain and Electric Utilities: Permits, Allowances, Monitoring, and Meteorology, a meeting of the Air and Waste Management Association held in Tempe, Arizona.

## EMISSIONS AND COST MODULES

**Presenters: John Molburg (Argonne National Laboratory) and Jayant Kalagnanam (Carnegie Mellon University)**

Reviewers: Ben Hobbs and Michael Hanemann

### SUMMARY

In general, the models represent a not unreasonable compromise between TAF's need for compactness and rapidity and NAPAP's need for credible cost estimates based on the full range of options available to utilities. The model reflects our current knowledge based on a few years of experience with an incentive-based regulatory system; extrapolation to the future is difficult when the system is so complex and dynamic.

If one or more of the models are directly included in TAF, then TAF users will have the very desirable capability of exploring the implications for emissions of numerous alternative fuel price, demand, retirement, technology, and other scenarios. However, interpretations of cost differences should be cautious because (1) the models' assumptions of constant capacity factors and simplification of the post-2000 decision process may underestimate the allocative efficiency benefits of allowance trading and (2) X-efficiency benefits of trading will not be quantified. On the other hand, (3) the discouragement of trading resulting from state government actions and utility conservatism may instead mean that the models might overestimate the cost savings resulting from trading under Title IV.

### INTRODUCTION

The presentation consisted of two parts. The first part, presented by Dr. Molburg, summarized the Argonne National Laboratory methodology. Its purpose is to project $SO_2$ compliance costs for utilities under alternative policies and to provide scenarios of total $SO_2$ and $NO_x$ emissions for the United States, Canada, and Mexico. The ANL model's heart is a method to choose $SO_2$ emissions-reduction measures on a unit-by-unit basis starting in the year 2000, based on minimizing the total present worth of revenue requirements. The assumptions concerning Phase I and nonutility emissions were also described. Generators could choose from a set of fuel-switching and -scrubbing options. The ANL methodology has been used to create 17 emissions/cost scenarios; these model outputs are available in TAF to be used as inputs to the other modules. These scenarios differ in terms of policy (no Title IV, Title IV limits applied to units but no trading allowed, Title IV with trading, and a "Beyond Title IV" scenario of mandated scrubbers); different retirement dates for generators (40 and 60 years); and alternative load growth rates (1% and 3% per year).

Dr. Kalagnanam made the second part of the presentation, summarizing a methodology under development at Carnegie Mellon University (CMU) that represents $SO_2$ control costs for each unit as a continuous function rather than as a set of discrete alternatives. The model's purpose is to explore the cost implications of allowances trading. An important difference compared to the ANL model is that the CMU model can simulate Title IV compliance strategies in which no interutility trading takes place. Each utility can choose options for its units so as to minimize the cost of meeting its $SO_2$ cap.

### CREDIBILITY OF MODEL CALCULATIONS

#### Credibility of Cost Calculations

Title IV lowers the cost of emissions reduction compared to traditional command-and-control regulation because utilities are required only to meet an annual tonnage cap and can use any of a wide range of options (including allowance purchases from overcomplying utilities or utilization of banked allowances) to meet it. Cost reductions can be viewed as arising from two sources:

1. Allocative efficiency improvements, choosing to use less costly means of reduction than would have otherwise not been allowed. Some of these options, such as energy efficiency (conservation), emissions dispatch,[1] and paying other utilities to overcomply (the effect of allowance purchases) were not considered by traditional command-and-control rules.

2. X-efficiency improvements, in which sharpened competition results in decreases in costs of particular options.[2] For example, Burtraw and others argue that having the option of allowance purchases has given utilities leverage in their negotiations with suppliers of low-sulfur coal and scrubbers, with the result that those emission options have turned out to be much less costly than envisioned in 1990.

Optimization models can be used to assess the cost reductions due to the Title IV allowance system by simulating the cost-minimizing decisions by utilities under two sets of options: the limited options allowed by command-and-control legislation and the expanded options available under the allowance system. If the same cost assumptions are made for both cases, then only allocative efficiency improvements can be quantified. If decreases in the costs of individual $SO_2$ control measures due to trading are hypothesized, then an estimate of X-efficiency improvements can be obtained.[3] However, the magnitude of such hypothesized decreases would be speculative, even if the experience of the last six years indicates that they may be very large.

Therefore, it would be unfair to criticize the ANL and CMU models for their omission of X-efficiency impacts of legislation. No available model includes those. The question is then how adequately can they characterize allocative efficiency gains? In order to obtain a compact and quickly executing model that can yield state-level emissions scenarios in TAF, compromises were necessary that could lead to some distortions:

- Not all options available under Title IV are modeled. In particular, "systemic" options that are associated with the entire generation system rather than individual units are not considered. Fuel switching and scrubbers are the options available to an individual unit and are considered by the ANL and CMU systems.[4] Emissions dispatch, energy efficiency, and early retirement affect multiple units simultaneously and are not captured by a unit-by-unit cost function approach that assumes constant capacity factors. In this reviewer's opinion, the amount of additional energy efficiency induced by Title IV is likely to be relatively small, but emissions dispatch is used now by most or all Phase I utilities and will be ubiquitous in Phase II. The resulting emissions reductions may be as large as 10%, depending on the fuel mix of the system.

- If this reviewer's understanding of the models is correct, the cost minimization for the period 2000–30 under Title IV is accomplished by either a heuristic (ANL) or a static cost function approach (CMU) that may overestimate costs, underestimate emissions (in later years), and have a bias towards capital-intensive controls. For the ANL model, the rationale for this conclusion is as follows. The ANL model, as this reviewer understands it, attempts to find the set of controls for all plants that minimizes the present worth of revenue requirements associated with decreasing emissions over the entire period 2000–30. However, the model results indicate that the price of allowances will fall to zero sometime in the 2010–15 time frame because total emissions are projected to fall below 8.9 million tons by then. The model assumes that the emission-reduction measures adopted until that time will remain in place afterwards. However, if allowances become valueless, many or most of those utilities that switched to more expensive low-sulfur coal would switch back, and emissions dispatch would cease. The effect would be higher emissions and lower costs after 2015 than the ANL model would anticipate. Further, if utilities anticipated that allowance prices would crash by, say, 2015, they would be less likely to adopt capital-intensive measures (such as scrubbers) than if they thought that emission reductions would be required through the year 2030.

- In the interutility scenarios, it is assumed that all profitable trades take place. However, utility conservatism and state meddling in the market have prevented this from taking place. Simplifications 1 and 2 mean that the models may overestimate compliance costs under Title IV, while the third simplification instead implies a possible underestimation. The net effect is not known; however, some idea of the relative importance of the first two effects might be obtained by analysis of the output of the ICF coal-electricity market model because that model allows for optimal modification of capacity factors and timing of investments. The possible impact of the third simplification might be gauged by the CMU model by starting with the intrautility model and then allowing a limited number of trades between utilities with the greatest divergence of marginal costs.

Another reviewer has pointed out that an additional simplification, compared to the ICF model, is that fuel prices are parameterized rather than calculated endogenously. Effects of the legislation and alternative assumptions concerning, say, utility deregulation upon equilibrium fuel prices are not easily assessed. A possible fix might be to examine outputs from the National Energy Modeling System or the ICF model regarding the effects of changes in such assumptions upon fuel prices, which could be used to establish likely ranges of fuel prices for input to the ANL/CMU models.

However, the strength of the ANL and CMU approaches is their nimbleness. If they can be included as modules in the TAF system, then users can take advantage of them to explore a wide range of scenarios, not just the few predefined scenarios now available.[5] No single scenario is credible because of, for instance, utility deregulation, demand and fuel price uncertainties, and the possibility of greenhouse gas legislation. Users need to be able to create a range of possible scenarios and to explore the robustness of the assessment's conclusions.

### Credibility of Emissions Calculations

Both the timing of emissions and their spatial distribution are important to TAF's environmental modules. Because no single emissions scenario can be assigned a high likelihood, it is important to be able to explore the implications of alternative scenarios. TAF has the potential to be able to do this through the use of user-defined emissions scenarios, which can be created within TAF or imported from another model.

As noted above, emissions in the years 2015–30 might be underestimated by the ANL model. Differences between the ANL and ICF emissions levels for earlier years have also been noted by the modelers, and explanations for them are being sought. Likewise, differences in geographic distributions have also been found. The benefit of TAF is that alternative emissions scenarios can be quickly analyzed.[6]

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

The existing 18 scenarios (including the ICF/EPA scenario) accomplish the goal of allowing exploration of the effects of a limited range of alternative assumptions regarding legislation, demand growth rate, and unit retirement age. Incorporation of one or more of the emissions/cost models within TAF would improve TAF's robustness for addressing alternative issues and its ability to analyze and present sensitivity to assumptions underlying the emissions and cost projections.

The capability to produce multiple scenarios within TAF is especially desirable because of the great uncertainty concerning the future evolution of allowance markets, future technical change (both exogenous and induced by economic incentives), and future opportunities for additional improvements in X-efficiency. The model reflects our current knowledge based on a few years of experience with an incentive-based regulatory system; extrapolation to the future is difficult when the system is so complex and dynamic. A scenario analysis may be the best way to accommodate the substantial uncertainty about future market developments within the model structure.

The goal of open architecture seems fully accomplished, given the ability to import any state-level scenarios of emissions.

## READINESS

The ANL model scenarios are incorporated in TAF, and the model itself seems ready to be included in time for the March 1996 assessment. The CMU model is evidently undergoing testing, and it is unclear whether it could be incorporated in time. The comparisons of the temporal and spatial distribution of ANL's emissions with the ICF results are a very useful validation exercise. Costs should also be compared.

## RECOMMENDATIONS

### Short-Term Improvement (for the 1996 assessment)

- Consider incorporating either or both of the emissions/cost models as modules in TAF, rather than just the 17 scenarios. However, before doing so, assess whether potential users would value the resulting capability to generate their own emissions/cost scenarios. Workshops in which users test and evaluate preliminary versions of TAF would be an effective means of accomplishing this goal.

- Continue to compare the ANL and ICF results, including an analysis of the role of emissions dispatch and changing capacity factors in the differences between the models. Conduct a sensitivity analysis of the ANL model to see if a 15- or 20-year planning horizon results in significantly lower costs and different emission-control options than the present 30-year horizon. If so, consider using the shorter horizon. Include improved Canadian and Mexican emissions scenarios.[7] Assess impact of Title I $NO_x$ restrictions (especially in the Ozone Transport Region) on utility and nonutility emissions. Where there is uncertainty, it should be noted.[8] Although these emissions are not altered by Title IV, they might affect the net impacts of Title IV to the extent that the impact models are nonlinear.

- Assess, using USEPA data, the effect of Title IV–induced demand-side management on growth rates. Incorporate this effect, if significant.

- Consider and include emissions scenarios that are very different in the geographic distribution of emissions to assess whether geographic distribution is an important issue.

### Longer-Term Improvement

- Assess whether the assumption of constant capacity factors results in a significant upward bias in compliance costs. If so and if model users consider relatively accurate cost savings to be a priority, then consider how emissions dispatch, seasonal fuel switching, and economic (rather than fixed-year) retirement decisions could be modeled. Consider including multiple ICF model outputs as optional emissions/cost scenarios.

- Consider including restricted interutility trading in the CMU model to simulate the situation in which utility conservatism and state regulation prevents some efficient trades from being consummated. Consider modeling intrautility trading in the ANL model.

- Explore how X-efficiency improvements resulting from trading could be estimated and included in assessments.

## NOTES

1. Emissions dispatch is the alteration of generation system operation so that cleaner but more expensive generating units generate more power and dirtier yet cheaper units generate less. This shifting of generation results in a net decrease of emissions at some increase in cost.

2. The concept of X-inefficiency is perhaps misused here, as strictly speaking it refers to the failure of a firm to wring the maximum possible output from a given set of inputs. The concept has somewhat fallen out of favor with economists, but it is nevertheless useful in this context to distinguish competition-induced reductions in the cost of particular options from cost reductions that result from choosing a different mix of control options.

3. As was suggested by one reviewer that, if cost reductions result from adoption or development of innovative technologies (e.g., omitting redundant scrubbers because very high reliability is no longer required), then identification of these technologies could be the basis of an assessment of X-efficiency improvements.

4. Although even their adoption affects other units, deratings will require more output from other plants, and changes in marginal operating costs will result in dispatch changes. Furthermore, the entire range of possible fuel blends, including natural gas substitution during summers when prices are low, are not considered in the models.

5. If only predefined scenarios were to be used, as is now the case in TAF, more credible results might be obtained by running the ICF model for the 17 cases and using those results. However, the ability for users to test scenarios of their own would be lost.

6. One of the reviewers noted that it is possible to use optimization methods to define the worst and best cases, from an environmental point of view, of the spatial distribution of emissions. If the net environmental effects of the two cases are fairly similar, then it could be argued that the spatial rearrangement of emissions resulting from emissions trading is unlikely to be of national environmental importance. Such an analysis would require the capability to assess impacts throughout the United States rather than at just 15 receptors.

7. One reviewer pointed out that Canadian emissions projections for 5-year periods up to at least 2020 are available from the Environmental Protection Service in Ottawa (Libby Greenwood) and also from EPA, Research Triangle Park, North Carolina (Sharon Nizich).

8. More minor suggestions:

   - The present calculation of average cost of control in 2000–30 in TAF divides present worth of revenue requirements by total (undiscounted) emissions. This method results in a much smaller value than if revenue requirements are divided by discounted emissions, which is the proper procedure for calculating levelized cost.

   - The fourth edition of AP-42 is being used for $NO_x$ emissions and other parameters; however, a fifth edition was issued by EPA in 1995 and should be checked for significant changes.

   - The cost of $NO_x$ emission reductions has fallen dramatically in the past year or so; the implications of this trend for the cost estimates should be examined.

   - It is unclear why it is necessary in the CMU model to fit quadratic or exponential curves to the cost data. It increases computational effort (by requiring nonlinear programming) and introduces inaccuracies (due to the fitting process). The ANL approach defines a piecewise linear convex cost function for each unit. Such functions can be directly input to a linear program, permitting the use of linear programming rather than nonlinear programming for the CMU model and also yielding more accurate cost estimates (in that the nonlinear fitted functions often had R2 < 0.7).

## HEALTH EFFECTS MODULE

**Presenter: Alan Krupnick (Resources for the Future)**

**Reviewer: Jane Hall**

These comments are focused on the written questions provided to reviewers. Those that relate most directly to the Health Effects Module are (1) Q1F: Are endpoints well identified and related to human welfare? (2) Q1H: Are both monetary and nonmonetary benefits included? and (3) Q2D2: What about interaction of health with other related changes, including climate change?

1. The broad issue here (and for all similar large-scale complex modeling efforts) is the extent to which the necessarily reductionist approach leads to loss of information that might be important in forming policy. The objective of estimating benefits (and costs) in monetary terms means that endpoints that are recognized as "real" but subject to one or more difficulties in quantification tend to disappear from the analysis.

   In this context, the Health Effects Module does a very good job of meeting the objective of identifying endpoints that clearly affect human welfare. It is currently less well adapted to supporting the inclusion of both monetary and nonmonetary benefits. This is not a failing of the approach used, because the Resources for the Future (RFF) framework can be adapted to more completely reflect endpoints that are not quantifiable in monetary terms.

   I recommend that more consideration be given to how to deal with endpoints that cannot be quantified in monetary terms either because we don't know how to value them or because the concentration-response functions are not supported by robust concentration estimates or other necessary variables. Such data need to be portrayed in a way that provides perspective about the importance of the "missing" endpoints. Interpretation is necessary to convey some sense of what we really do and don't know in the nonquantified as well as the quantified group. For many health researchers, a central conundrum is how to deal with lung function changes. Several pollutants (including PM10 and ozone) are associated with decrements, and a number of health scientists, notably Lippman and Bates, conclude that this is a nontrivial effect. Economists have no idea how to place a value on this. We also do not have robust means to estimate frequencies of chronic conditions related to long-term exposures, yet people attach high values to avoiding such endpoints.

   The analysis required to begin to fill in the gaps might be beyond the scope of the 1996 assessment, but the issue should at least be clearly noted and discussed as a factor in interpreting the benefit estimates.

2. While the modeling difficulties for ozone explain why it is not included in the analysis, this is another potentially important loss of information about the overall costs and benefits of Title IV. Some discussion could be included about the potential for ozone benefits, to the extent that acid precursor reductions simultaneously reduce ozone concentrations to a greater degree (or sooner) than would be the case under NAAQS attainment strategies. This effect may fall outside of TAF, but it should fit into the assessment somewhere.

3. Regarding nitrate, there is not much work on its impact independent of PM10, but given its size, it is being treated in this analysis as sulfate, for which more data are available. Bart Ostro has done a study of two Southern California counties that shows mortality effects consistent with the results obtained from using existing PM10 concentration-response functions. Because PM10 in that region is relatively sulfate poor and nitrate rich, it may be an indicator of nitrates's role, either because of chemistry or particle size. At least it supports treating nitrate as a factor.

4. A longer-term research issue is the interaction of temperature and concentration. To the extent that Title IV affects fossil fuel use (mix and amount), there is a relationship to climate change. If temperatures rise, concentrations of

PM10 being constant, the combined effect will increase, and the reverse for temperature moderation. This effect is beyond the scope of the current work and may be too complex to address, but is directly related to Q2D2.

5. Regarding the question raised during the conference call about which endpoints have a consensus and which are more "fringe," I think the researchers have done a good job of sorting that out.

   Probably more problematic than agreeing on the association between concentration and effect is agreeing on what some endpoints are "worth." The most difficult issue here is how we value premature death when we do not have clear distributions for who is most at risk or how values vary among the risk groups. The written report addresses this issue, and it is an area for future research.

Reviewer: Richard Schlesinger

## INTRODUCTION

This model is designed to estimate the human health impacts from changes in air pollution levels (namely PM10, $NO_x$, $SO_2$, $SO_4$ and $NO_3$), expressed as the number of days of acute morbidity effects, the number of chronic disease cases, and the extent of premature mortality. The model inputs consist of ambient concentrations of selected pollutants, demographic information, and other information as needed. Both threshold and nonthreshold approaches can be used with the model.

## CREDIBILITY OF MODEL

In general, the model is credible for the inputs used and the outputs obtained. However, the issue of actual credibility in terms of predicting real benefits in health from pollutant reductions must take into account other factors that are not currently considered, as indicated below in the recommendations.

## ACHIEVEMENT OF GOALS

The model has a fairly open architecture, with ample accessibility for substitution of other factors that may be important as issues evolve. However, the actual sensitivity of the model to changes in input is not clear.

## READINESS

The model is considered to be in a preliminary stage because the status of verification and calibration is not clear. However, as a start, it is acceptable, again given the caveats indicated in the recommendations below.

## RECOMMENDATIONS FOR IMPROVEMENT

There are a number of items that were not considered in the current model but which are of concern in evaluating health impacts of ambient air pollution:

- A number of recent studies have strongly suggested that ambient temperature is an important factor to consider when evaluating the health impact of PM10. Thus, the interaction of temperature conditions with pollutant levels needs to be taken in account in the model.

- While the model indicates that nitrogen oxides are considered because they are a precursor of ozone (rather than an important contributor to health impact themselves), the model does not include any consideration of health impact of ozone itself. This is a significant omission because ozone is probably the one major gaseous pollutant having the strongest association with health endpoints. Furthermore, a number of epidemiological and controlled experimental studies strongly indicate that there could be significant interaction between ozone and particulate pollutants (such as acidic aerosols) affecting ultimate biological responses. The lack of any consideration of ozone itself and especially of such interactions of ozone with other pollutants is a critical omission in the current model.

- Given the lack of consideration of ozone in the model, it is surprising that so much emphasis has been placed on how to deal with nitrates. While it is true that there is a lack of epidemiological evidence for any health effects of nitrate other than perhaps those associated with particulates in general, controlled studies strongly suggest that nitrates are much less potent than acidic sulfates. Thus, a separate category for nitrates may not be necessary, and it clearly is inappropriate to consider that nitrates have the same potency as sulfates (option 4 in the diagram of mortality impacts). It is best to merely include nitrates as part of generic PM10.

- The basis for assuming that adult chronic bronchitis is the only health impact that needs a threshold seems to be that this would be the only way in which it would work in the model.

- The health outcomes used are not always appropriate. For example, asthmatic attacks should be included for $SO_2$ because there is some evidence of increased susceptibility of asthmatics to this pollutant.

## SUMMARY RECOMMENDATIONS

As currently constituted, the model is not suitable for use in the 1996 NAPAP assessment without consideration of ozone. This problem needs to be addressed in the short term, while other issues indicated could be incorporated into future iterations of the model.

## BENEFITS VALUATION MODULE

**Presenters: Alan Krupnick, Dallas Burtraw, and David Austin (Resources for the Future)**

**Reviewers: Michael Hanemann and Mac Callaway**

### INTRODUCTION

The benefits presentation was integrated with the health effects presentation, but this review covers only the Benefits Valuation Module. The Benefits Valuation Module is based largely on economic estimates for specific effects categories that have been derived from other studies. There is no parent model. The calculations and methodologies that have been borrowed from other studies are easily integrated within TAF.

The Benefits Valuation Module covers the following effects categories:

- Health
- Visibility
- Recreational fishing

Categories that were not included, but for which some estimates of benefits/damages may be available include:

- Forests—1990 Integrated Assessment Sensitivity Analysis
- Passive nitrogen fertilization—1990 Integrated Assessment, Integrated Analysis
- Damages to car finishes—EPA contractor report (ICF)

The coverage of these benefit categories could be expanded to show how these values compare with those that are estimated, based on the 1990 Integrated Assessment results and the ICF report on car finishes.

Three potentially large effects categories were excluded because there is currently no credible information to conduct a benefits assessment:

- Building materials
- Cultural materials
- General non-use/existence

The NAPAP assessment needs to wrestle with the issue of what to do about these categories because the potential benefits are quite large. To say nothing about these categories has the effect of placing zero benefit values on them. To try to value them with the existing information is also inappropriate. There is enough literature on these topics to perform an independent assessment of the potential impact of including these valuation categories and then to follow the lead of Hagler Bailly in the N.Y. State Externality Study and devote a section or sections to unquantified limitations, biases, errors, etc., and then, to the extent possible, speculate on the impact of the direction and magnitude of the total benefits estimate(s).

### CREDIBILITY OF MODEL/MODULE CALCULATIONS

The calculations for the three modeled categories are based on the concept of benefits transfer. This involves four basic steps:

1. reviewing published studies and evaluating their applicability for benefits transfer,

2. using information in one or more studies to estimate "unit" benefits values on a per person or per unit output basis as a function of an important driver variable of policy interest,

3. determining the appropriate segment(s) of the population or market to which to apply these estimates, and

4. combining steps 2 and 3 with changes in the driver variable to produce an estimate of the change in benefits associated with the change in the driver variable.

NAPAP should be commended for developing benefits transfer methods/modules that systematically include all these steps.

Benefits transfer estimates usually cannot be validated against observed data, except in studies specifically designed for this purpose. Where these studies have been performed—in the recreation demand area—benefits transfer methods have not proven to be particularly accurate. The accuracy of the benefits transfer calculation methods in this application, likewise, cannot be empirically validated. Therefore, different criteria need to be adopted for evaluating the applicability of the benefits transfer methodology.

One of these criteria involves the relationship between the locations/populations/groups/etc. that are included in the "parent" study and those in the transfer application. Related to this is the question of the magnitude of the impacts in regions for which no direct model studies are available. The reader would benefit from a discussion of these topics because it would help explain why certain types of studies have to be "stretched" to fit other regions.

Comments on individual sections follow.

**Health**

The approach here appears to be quite thorough and relies on the best available estimates of dose response functions and unit values for a wide variety of health impacts. We have two concerns:

- The documentation we reviewed did not present the methodology used to determine (a) what segments of the population are susceptible to certain health risks and (b) what segments of these groups are exposed to higher or lower health risks as a result of the Clean Air Act Amendments.

- The problem of valuing multiple endpoints was raised in the RFF presentation but was not presented.

Both of these concerns are especially important in view of the large magnitude of the benefits estimates in this area and the potential impact that assumptions made in the methodology might have on the size of the estimates. For these reasons, we recommend that the uncertainties associated with defining the exposed and susceptible populations and avoiding double counting of multiple endpoints be included in the uncertainty analysis. This change would be very useful in helping us to understand how robust the benefits assessment methodology is to assumptions in these areas.

**Visibility**

The visibility studies on which the valuation model is based for all user categories is the best available, or almost. Mitchell and Carson did a study of Cincinnati for EPA (which was never released for a variety of reasons) that might have some valuable information in it. The unit values used for residential demand in the Benefits Valuation Module are probably a little lower than the literature would support. Something in the neighborhood of $10, $15, $30–35 would better reflect more recent studies.

The Visibility Module employs a methodology based on valuation functions that is commonly found in the literature and has previously been used by NAPAP and EPA. We are concerned, however, that this methodology is too reductionist and may not adequately capture how people place a value on visibility. In particular, this methodology employs what may be an excessively simple functional form, and it may not adequately allow for variation in the psychophysics of visibility perception and valuation across different groups of individuals and different types of visibility (e.g., urban vs rural, east coast vs west coast, episodic vs continuous). It is true that the restrictive specification of the current visibility valuation function reflects the limited data available in the economics literature. However, we believe that a better approach would be to spend more time on basic aspects of visibility perception, casting a wider net that would encompass noneconomic and qualitative information about the human response to changes in visibility. We would recommend that the visibility valuation team be asked to review this broader literature and then develop a new modeling framework for visibility valuation functions. Also, uncertainty can be introduced into the analysis of functional form, for example, by fitting a linear model to the average "price"-quantity combination and then varying the "price" elasticity over a reasonable range.

**Recreational Fishing**

The NAPAP study by Englin et al. was used in this analysis because it is the only study that allows one to track the effects of changes in deposition through changes in ASI, catch per unit effort, and willingness to pay. However, this study has a number of limitations that need to be discussed and weighed more carefully:

- Parameter estimates from the recreation demand model were derived from a different sample of lakes than the ASI model parameters.
- The large ASI changes that "drive" the Englin results occurred in a small number of highly sensitive lakes, and data for these lakes were not used to estimate the parameters of the recreation demand model.
- The model included only trout anglers.

The NAPAP 1990 Integrated Assessment identifies a number of other limitations that should be reviewed with an eye toward evaluating the impacts on the benefits transfer estimates.

One possible alternative to investigate would be to use the avoided cost of liming a lake as a benefit measure for the "average" lake. A methodology for developing supply curves for lake liming in the Adirondacks is presented in both SOST-27 and the NAPAP 1990 Integrated Assessment. There, a number of supply curves are presented for liming Adirondack lakes with long residence times via surface liming methods. Total and average/incremental cost curves are presented. The incremental cost of liming an "average" Adirondack lake was quite low. Developing a methodology to accommodate the avoided cost of liming would help to bound the estimates of benefits on the lower end. The avoided-cost approach would not be applicable to small, higher-elevation lakes with shorter residence times for two reasons: (1) surface liming of these waters is not very effective, and (2) these are waters for which existence values may be more important than use values and also use values may be much higher because of the "pristine" site characteristics of these lakes. On the other hand, if these pristine characteristics were not altered by acid deposition and fishing were improved by stream-liming methods (which are more expensive than surface-liming methods), then perhaps the avoided-cost estimates would not be inappropriate if suitable estimates of stream liming costs could be found.

**Policy Issues**

The TAF methodology makes it possible to calculate incremental benefits and costs associated with different levels of emissions reductions. In past assessments, NAPAP was heavily criticized for not showing the relationship between incremental benefits and costs of Title IV reductions. NAPAP now has that capability, and it would be of interest to economists and policy makers to see whether the incremental benefits of Title IV are much greater, or smaller, than the incremental costs. Because point estimates may be misleading, this topic should be included in the uncertainty analysis to determine the relationship between the respective ranges, or confidence bands, of the incremental costs and benefits of various levels of emissions reductions.

## ACHIEVEMENT OF MODEL/MODULE DESIGN GOALS

The calculations in the different modules are computationally trivial to implement. However, there are a couple of other issues worth raising related to the achievement of design goals. First, the status of actual work completed and work remaining on this TAF module are not very clear. Second, the nature of the uncertainty analysis to be included in the TAF Benefits Valuation Module has not been made explicit.

## RECOMMENDATIONS

- Expand the treatment of areas not included in the assessment so that the reader can put the potential benefits in these categories into perspective with the included categories.
- Include discussion of the rationale for including or excluding certain geographic areas and the applicability of the benefits transfer results outside the context of the original study.
- In health effects analysis, incorporate the uncertainty associated with determining the susceptible and exposed population and multiple endpoints into the benefits estimates.
- In aquatics/recreation demand, include more thorough discussions of the limitations of the Englin et al. study and the potential impacts of these limitations on the results.
- Consider using an avoided-cost estimate based on lake liming to estimate the benefits of improved water quality in the Adirondacks. (Defensive expenditures/control cost has a theoretical justification if MC is less than MB of improved water quality.)
- In visibility, some of the values used could be adjusted to better reflect the literature (residential visibility). Also, the Carson and Mitchell study of Cincinnati may be useful.
- The visibility methodology could be improved by trying to incorporate (a) valuation of changes in "peaks," (b) consideration of perception thresholds, and (c) uncertainty due to functional form. These changes would require additional medium-term research.
- Visibility valuation has long been plagued by the paucity of national studies for benefits assessments. This problem can be rectified only through long-term research by developing primary research topics specifically to back up benefits assessments. EPA, DOE, and the park service have a joint interest in this type of research.
- Calculate and compare the ranges/confidence bands for incremental benefits and costs of different levels of emissions reductions.

## TRACKING AND ANALYSIS FRAMEWORK
## (COLLABORATIVE DEVELOPMENT OF A TOOL FOR INTEGRATED ASSESSMENT)

**Presenters: Max Henrion and Richard Sonnenblick (Lumina Decision Systems)**

Reviewer: Markus Amann

### INTRODUCTION

Dr. Max Henrion presented the collaborative development of the integrated assessment framework of the TAF model. The primary target of TAF is to link the various aspects relevant for the integrated assessment of acid rain emission-reduction strategies into a consistent framework, i.e., providing consistent and well-defined linkages between the various modules and offering a systematic tool for scenario and uncertainty analysis. Using the software package Analytica™, each module of the integrated assessment framework is defined by the mathematical relations between its input and output variable and the definition of the relevant databases. The model has been developed in a cooperative effort among geographically dispersed collaborators using advanced forms of telecommunication.

### CREDIBILITY OF MODEL CALCULATIONS

The credibility of TAF model calculations is strongly dependent on the credibility of its individual modules and thereby, as outlined by other reviewers, different for the various TAF modules. Most important for the integrated framework, however, is the complete and balanced description of all relevant processes and aspects of the entire chain from the sources of emissions, the possibilities and costs for controlling them, and the atmospheric dispersion, up to the environmental impacts and, if required, their monetary valuation.

A weak element of the current implementation of TAF is the treatment of other pollutants from nonenergy sources (e.g., $NH_3$ and base cations). A credible implementation should revise the presented approach, for example, by following the recommendations of the peer review panel.

A strength of the presented approach is its modularity and open architecture, which enable the exploration of a variety of aspects with great flexibility and eventually a replacement of individual modules by alternative formulations.

Although uncertainties of parameters are specified in the context of the individual submodules, the overall uncertainty analysis and propagation is to be accomplished in the TAF integration module. A thorough analysis and thoughtful presentation of the results (beyond the currently implemented features) will be essential for the overall credibility of the entire framework. Unfortunately, only limited documentation has been provided on the detailed approaches and methods currently implemented in TAF.

### ACHIEVEMENT OF THE MODEL DESIGN GOALS

The team is to be congratulated for the flexible structure and open architecture of TAF, which can keep TAF up to date for quite some time. This goal has been clearly achieved, but perhaps at the cost of reduced user-friendliness. To what extent the lower user-friendliness has actual relevance, however, depends crucially on the type of envisaged model users. If TAF is to be used by a limited group of experts only, the selected compromise is definitely justified.

The currently implemented approach for uncertainty analysis is a good starting point. It can be expected, however, that further refinements will be necessary to identify customized methods of analysis and presentation of results to provide practical and useful answers to the actual questions of model users.

## READINESS

Although the basic software package is ready, some refinements could facilitate its practical application for the 1996 NAPAP Integrated Assessment. In particular, the implementation of the Emissions and Cost modules in TAF would enable (1) the tracking of the implications of control measures at particular emission source groups and (2) the propagation of uncertainties of these modules throughout the entire chain to the benefits assessment. Both aspects could be important elements of a comprehensive analysis.

Furthermore, every effort should be undertaken to extend the geographical scope of the impact analysis (starting from deposition fields to monetary benefits) from the current 15 receptor sites to the entire area of the United States and possibly to North America, so that a full picture of the benefits could be obtained.

Despite these possible improvements, not all modules can be expected to be available in 1996 at comparable levels of quality and reliability. The use of TAF in an integrated assessment exercise in 1996 must consider this fact by not relying mechanically on model results.

## SUMMARY RECOMMENDATIONS

The development of TAF is a crucial and necessary step towards a systematic and comprehensive assessment of costs and benefits of strategies to reduce acid deposition. By employing reduced-form models covering the wide range of relevant aspects, additional and important insights can be gained that would not be possible without such an integrated assessment tool.

Significant progress was made in short time. Despite some shortcomings of the first version of TAF, some improvements will make it possible to use the model for the 1996 NAPAP Integrated Assessment as one tool among others. Because of the preliminary nature of many of its elements, careful interpretation of the results with active involvement of the key scientists responsible for the development of the various modules will be essential.

Reviewer: Hugh Ellis

## INTRODUCTION

This narrative takes as a basic premise the need for TAF to be reviewed and evaluated in a specific context, that is, a context defined in part by two concerns: (1) for whom is the integrated assessment being performed and (2) what are their needs and purposes relevant to the integrated assessment. These concerns, much more so than for the various scientific modules that compose TAF, will ultimately determine the project's usefulness and potential policy relevance.

Given that premise, a few general observations and a disclaimer are in order. Written evaluations provided to this reviewer ranged in tone and content from strongly supportive and complementary to somewhat skeptical (relevant quotes follow, as appropriate). This fact is important to lay out because the review and evaluation that follow are generally more critical, and in some cases much more critical, than the aforementioned evaluations.

## PRESENTATION

As far as it went, the presentation was useful, albeit a marginal enhancement over the original overview given on the first day. Essentially, the presentation consisted of a restatement of the philosophy and goals of integrated assessment as seen by Lumina Decision Systems and Cary Bloyd (more on his contribution later). Two attributes of the presentation—one important, one less so—made a distinct impression. The first and more important attribute was my perception of a disconnect between the decisions made in the design and implementation of TAF and the needs and purposes of NAPAP (and, ostensibly, the Interagency Committee). Repeated questioning failed to produce any substantive evidence that the Interagency Committee, for example, had any (let alone substantive) input into the original design of TAF (generally cast) and its development. If this assertion is correct, then it represents a potentially serious shortcoming, at least in process, and renders the favorable evaluation of TAF, in context, highly problematic. The apparent disconnect to which I refer seemed clear to me in Cary Bloyd's opening remarks in the presentation. Specifically, he noted that TAF was created "for NAPAP" then proceeded to explain that he hand-picked the TAF team, making no mention of collaboration or solicitation of advice from other presumably interested parties—the stakeholders as I understand them to be.

The second, and admittedly much less significant, attribute of the presentation that made a distinct impression on me was the altogether curious (and unsolicited) question/challenge posed by Dr. Henrion regarding size and complexity of the integrated assessment. The question (or challenge, depending on your viewpoint) went something like this: "If anyone has done an integrated assessment that's larger or more complex, I'd like to know about it." That is a curious statement for at least one reason. Specifically, a member of the review team (Amann) played and plays a central role in what is already the largest and, in several important respects, most complex integrated assessment performed anywhere (RAINS and now RAINS-ASIA as well). RAINS involved participation from over 60 countries and has resulted in the successful negotiation and implementation of first "Sulfur Protocol 1," then "Sulfur Protocol 2," and now a nitrogen protocol as part of a multipollutant agreement involving acid rain and tropospheric ozone. Slightly bruised pride prompts me to note as well that on a smaller scale (in terms of model size and complexity—not in terms of impact) another review member (Ellis) conducted an integrated assessment, the results of which became Canadian federal and provincial acid rain–control law, which has been implemented. The point here is that Dr. Henrion seems unaware of these developments and presumably of the lessons learned therein, and it is difficult to explain why that is the case.

On a less negative note, TAF (with some caveats noted below) looks promising with respect to the goal of producing a useful didactic tool. I saw occasional evidence in the scientific module presentations that TAF could be having a synergistic or otherwise catalytic effect, which of course is or should be one of its most important objectives. It is in this respect that many of the laudatory written responses were made ("a valuable and worthwhile tool," "TAF has been able to integrate multidisciplinary research into a great cause-and-effect tool," and "great concept").

## CREDIBILITY

Written responses ranged from "highly credible" to "lots of caveats but I think [that the] bottom line is yes, TAF has credibility." My stance is that TAF is credible but that it is much more credible as a didactic tool than as a mechanism to support and inform acid rain policy development. I have another concern that straddles credibility and achievement of design goals, which I reserve until the next section.

## ACHIEVEMENT OF DESIGN GOALS

One written response noted quite correctly that "they have pulled together a lot in a short time, which speaks well of the open architecture [goal (a)]," to which I would add my general agreement. I do, however, have some reservations regarding item (a) "open architecture," which reflects as well a written comment: "should be totally public domain." Specifically, I question the choice to base TAF on Analytica™. My difficulties in this regard lie not so much with Analytica per se (I am unfamiliar with it) but rather with the potential long-term consequences of adopting an integrated analysis framework based on a programming language that is not in widespread use. As far as I can tell, much of the expertise with Analytica lies with Lumina Decision Systems. Could not TAF be based on, for example, C++, which is ubiquitous and for which widespread expertise is readily available? Is Analytica truly so superior as to make it the only and obvious choice? No explanation of the rationale to use Analytica was ever given other than several vague references to its being "better" than Fortran and its lineage to DEMOS. A very fundamental issue—why TAF was designed to be run by (presumably) many users, giving rise to RAM budgets and their attendant consequences—was, to me at least, treated as intuitively obvious. I have not seen or read any strong or otherwise compelling reason to make me believe that numerous users want to run TAF for purposes related to policy support.

The adoption of reduced-form models looks, thus far, to be a success, and I take as particularly relevant, the oft-repeated assertion that something valuable is learned in building a reduced-form model. Of course, all decisions involve tradeoffs, and effort expended in making reduced-form models comes at the expense of, say, effort put into enhancing the scientific credibility of a full-form model.

Taking transparency of assumptions to refer to TAF itself and not to its modules, my assessment is that TAF is not yet acceptably transparent (to wit, numerous questions trying to probe how TAF does what it does, especially with regard to the propagation of uncertainty). One comment rated transparency as "poor." The treatment of uncertainty in TAF—as far as it goes—is generally acceptable, if occasionally grossly overemphasized. Adroit treatment of uncertainty can enhance the potential policy relevance of a scientifically credible assessment model, but it most assuredly cannot make a module, in any sense, more scientifically credible. This issue was pursued at some length, but I remain unsure as to whether Lumina Decision Systems recognizes or appreciates the distinction being made here. Additional issues concerning uncertainty speak to the need for (1) paying more attention to the communication of problematic results (i.e., presenting distribution functions to end users is not good enough), (2) considering both parameter (Type II) and structural model (Type I) uncertainty, and (3) requesting that module developers put, in some sense, a "value" on information, as well as a probabilistic characterization.

## READINESS

What few written comments were provided, either explicit or implicit, suggested that TAF is not ready for use in policy support. It is, almost by definition, ready and functioning in its didactic role.

## IMPROVEMENTS

As far as I can tell, TAF has not, for example, been exposed to the Interagency Committee, presumably its principal stakeholders. That, if true, is a serious omission and must be corrected to connect the disconnect stated in my opening remarks. Particular improvements that would arise from such exposure are unpredictable. Additional specific improvements provided in the written evaluations include the following: (1) better summarize results, (2) explicitly state key assumptions, (3) develop a reduced-form representation for uncertainty analysis and results, and (4) significantly expand geographic coverage.

## SUMMARY RECOMMENDATIONS

Lacking a serious and substantive effort to place TAF in a realistic, relevant context (e.g., involving Interagency Committee feedback) probably renders it unsuitable for use in the March 1996 NAPAP assessment. One of the written evaluations is particularly relevant in this regard. It referred to Mike Uhart's opening statements on Monday regarding the decision to have TAF run (for the 1996 NAPAP assessment) by "policy people" chosen "soon." The written comment went on to note that "there will be very little involvement of TAF integrators/developers/scientists. It sounds as if there is considerable potential for misuse of this tool." I concur with this concern. For all of the appropriate and well-intentioned concerns regarding "integration," TAF is remarkable in its lack of integration/collaboration with its intended near-term audience.

<u>Reviewer: David Lam</u>

## INTRODUCTION

For many years, there have been discussions about the concept of developing TAF as an integrated assessment tool that combines predictive models for air, soil, water, ecology, socioeconomics, and human health to answer policy questions on acidification problems. While the NAPAP 1990 Integrated Assessment was useful in gathering large volumes of data and knowledge, including models, there was no serious effort in actually combining these data and models for systematic overall investigation. Therefore, the TAF module presented here is most timely and necessary. The general objective is to support coordination and communication between policy makers and scientists within a tracking and analysis system that is comprehensive and technically credible, yet fast, responsive, and flexible. The specific objective is to provide an assessment framework on a PC to assist NAPAP in understanding the costs, benefits, and effectiveness of emissions-control strategies under Title IV.

## CREDIBILITY OF MODEL CALCULATIONS

Because TAF is a framework, not a model, the calculations pertain to those models implemented within it. It was also recognized that progressive refinement is important for integrated assessment. At this early stage of development, the results presented were preliminary and subject to further improvement. The credibility of TAF is best judged by its conceptual framework and by the quality of output as examples of more elegant results that may evolve as the system is being progressively refined.

### Conceptual Framework

The concept and system design were excellent, and the prototype system was proven to work for many modules and their linkages. It has brought together data and models for air, water, soil, socioeconomics, and human health as far as system management and computer memory (18 MB of RAM on a Power Macintosh) would allow. The integrity of data and model were shown to be liable and made available and accessible via World Wide Web (WWW) and other modern communication methods.

### Reduced-Form Models

One salient feature of TAF was the idea that full, comprehensive models could be emulated with much simpler forms of mathematical or statistical equations. These reduced-form models, when appropriately derived, could produce results very similar to their parent models with only small perturbations or uncertainties that were within the tolerance of observed data or the uncertainties of the parent models. The best example of the use of reduced-form models is in the Aquatics and Soils modules, derived from results of the MAGIC models.

## ACHIEVEMENT OF MODEL DESIGN GOALS

### Open Architecture with Accessibility for Substitution and Upgrade

With the use of WWW and other easily accessible means of communication, the system was open to all participants. That so much data and models were assembled in a short time speaks well of the open architecture concept. It is foreseeable that totally public domain software will be made available as the system evolves into more mature status.

### Utility of Reduced-Form Structure with Acceptable Additional Uncertainty

As noted before, the Aquatics and Soils modules provided the best example of reduced-form structure with acceptable uncertainty. The source-receptor matrix in the Atmospheric Pathways Module was another good example. During the presentation and from the submitted review comments, many reviewers held different opinions than the presenters on the notion that uncertainty analysis could lead to a better model. There was no doubt that uncertainty analyses on model input, output, and parameters were required to establish the validity of the reduced-form model relative to its parent model. As long as better models implied model structural change, however, such analyses alone might not be able to detect structural alteration requirements without in-depth investigation of models with different structures and, more importantly, without new knowledge input from process research.

### Transparency of Assumptions

The Public Index Library was a good approach to document model sources, linkages, and assumptions. At the moment, key assumptions may be buried in the references of parent models. In the future, it should include those assumptions that were relevant during the development or modification of parent models as well as reduced-form models.

### Analysis and Presentation of Sensitivity

There were mechanisms in TAF for sensitivity analysis of data and model results. Some graphical presentation modules also evidently helped visualize statistical variability and normalized distributions.

### Robustness for Addressing Alternative Issues

TAF is sufficiently generic for addressing other issues. Its robustness has been fairly well established. Given more time and successful examples, it will be further improved and progressively refined.

## READINESS

It was emphasized during the review meeting that TAF served a dual role: (a) as a framework for coordinating research efforts and thereby promoting creativity in combining research results from different disciplines and (b) as a framework for policy input. Because it is only a framework, not a model, its readiness depends on how each individual model, be it parent or reduced form, is ready or not. At the very least, this determination requires a judgment call on the part of the developers of the individual models as to whether they are comfortable with the scientific credibility and policy readiness of their models. Then there is the coherence or consistency requirement: are the individual models really linkable to each other with acceptable uncertainty? Chances are that some models are more ready than others. In the end, it is the user (i.e., both the scientist and the policy maker) who will decide its readiness. Inasmuch as TAF is an iterative process, no matter how many cautions and caveats one may impose on it, it needs to start somewhere. I think it has already started. (It may have been long overdue.)

## RECOMMENDATIONS FOR SHORT-TERM AND LONG-TERM IMPROVEMENT

While there seemed to be excellent dialogues between some parent model developers and the reduced-form model developers, better communication among all modules is required. Over the short term, it is worthwhile to explain how uncertainties could propagate from one part of TAF to another. If successful, these uncertainty analysis results could make the use of TAF for the 1996 NAPAP assessment even more convincingly. Over the long term, more receptor sites are required to further establish the credibility and to progressively refine the improved parent model and reconfirm the reduced-form models. So far, much work has been devoted to the system and software engineering design. To promote its use, porting the system over to the Windows environment for IBM-compatible PCs is a desirable long-term goal.

More attention, however, should be made to the knowledge-engineering and team-research efforts to acquire new levels of creativity. This may compete with the demand from policy advisors to make TAF a more user-friendly system, which too is a complex exercise in itself.

## SUMMARY RECOMMENDATIONS

On the whole, TAF is ready to be used as a framework for integrated assessment. There are supports from both the scientific community and policy/public consultation groups. It should be recognized that this is an iterative and progressively refined approach, with open architecture to accept necessary changes and improvement, of which there are plenty, as discussed in the review of all its modules.

- Clarification of the weighting of effects in the Aquatics Module was suggested because overall damage might be underestimated if much of the damage occurs in a small percentage of lakes.

- It would be very valuable to develop the Aquatics Module for other regions; all the reviewers who gave their comments suggested this. The mid-Appalachians and Southern Blue Ridge areas in particular contain many streams of intermediate pH and low ANC and are vulnerable to future acidification; it is possible that the Aquatics Module would identify a greater change in pH and the ASI over time in these areas versus the Adirondacks. One reviewer pointed out that this possibility would be more likely if it could be shown that the pH distribution of unacidified areas is not bimodal, as it often is in acidified areas: acidification moves mid-ANC lakes to the low ANC category.

- In the Aquatics Module, it may be possible to improve the overall aluminum and pH prediction relationship for a region by including lake-specific aluminum-pH relationships.

- It was suggested that a comparison be made of region-to-region variability versus lake-to-lake calibration results in the expansion of the Aquatics Module to other regions.

- Continue planned reformulation of the Soils RFM to incorporate exponential approach to equilibrium for base saturation.

- It would be valuable, given resources, to develop the Soils Module for other regions.

MAGIC continues to be a good choice for effects modeling, and the RFMs appears to mimic MAGIC's output well, at least for the region so far developed (Adirondacks). Consequently, they should be developed for other areas to generalize their utility. The RFMs should be developed from the FFM for each region because it is likely that the colinearity structure among the input variables differs among regions, and this could affect the fidelity of the RFMs to the FFM. Continue planned work on the incorporation of nitrogen. The ASI coupled to the chemistry models appears to be a good procedure for modeling an index of aquatic biotic response.

## ATTACHMENT 1. PARTICIPANTS LIST

### I. NAPAP Program Managers, Representatives, and Interagency Committee

| | |
|---|---|
| Rick Artz | DOC/NOAA |
| Jack Barnes | USDA |
| Beth Campbell | DOE |
| Noreen Clancy | NAPAP |
| Bruce Hicks | DOC/NOAA |
| Paul Kapinos | USGS |
| Joe Kruger | EPA |
| Karen Malkin | NPS |
| Robert Rosenthal | Office of Program Analysis |
| Mike Uhart | NAPAP |
| Walt Warnick | DOE/OPA |
| Derek Winstanley | DOC/NOAA |

### II. Oak Ridge National Laboratory Support Staff

| | |
|---|---|
| Angela Beach | OER Support |
| Pat Honeycutt | OER Support |

### III. Review Panel

**Executive Secretary**

| | |
|---|---|
| John Malanchuk | International Technology Corporation |

**Aquatics/Soils Effects Modules**

| | |
|---|---|
| Arthur Bulger | University of Virginia |
| Nikos Nikolaidis | University of Connecticut |

**Atmospheric Pathways Module ($SO_x$ and $NO_x$ Submodule and Additional Species Submodule)**

| | |
|---|---|
| Michael Moran | Atmospheric Environment Service |
| Gary Stensland | Illinois State Water Survey |

**Benefits Valuation Module**

| | |
|---|---|
| Mac Callaway | Hagler Bailly Consulting, Inc. |
| Michael Hanemann | University of California, Berkeley |

**Emissions and Cost Modules**

| | |
|---|---|
| Michael Hanemann | University of California, Berkeley |
| Ben Hobbs | Johns Hopkins University |

**Health Effects Module**

| | |
|---|---|
| Jane Hall | California State University |
| Richard Schlesinger | New York University Medical Center |

**MAGIC**

| | |
|---|---|
| David Lam | Water Research Institute |
| Nikos Nikolaidis | University of Connecticut |

III. **Review Panel (continued)**

**TAF**

| | |
|---|---|
| Markus Amann | IIASA, Austria |
| Hugh Ellis | Johns Hopkins University |
| David Lam | National Water Research Institute |

**Visibility Module**

| | |
|---|---|
| Rudy Husar | Washington University, St. Louis |
| Daniel McNaughton | Consultant |

IV. **Principal Investigators/Co-Investigators/Program Managers**

| | |
|---|---|
| David Austin | Resources for the Future |
| Bryan Bloomer | USEPA |
| Cary Bloyd | Argonne National Laboratory |
| Dallas Burtraw | Resources for the Future |
| Jack Cosby | University of Virginia |
| Deirdre Farrell | Resources for the Future |
| John Formento | Argonne National Laboratory |
| Chris Frey | North Carolina State University |
| Donald Hanson | Argonne National Laboratory |
| Max Henrion | Lumina Decision Systems |
| Jayant Kalagnanam | Carnegie Mellon University |
| Alan Krupnick | Resources for the Future |
| Erin Mansur | Resources for the Future |
| Ronald Marnicio | Foster Wheeler Environmental Corp. |
| John Molburg | Argonne National Laboratory |
| Gar Ragland | Resources for the Future |
| Patrick Ryan | SAIC/Oak Ridge National Laboratory |
| Jack Shannon | Argonne National Laboratory |
| Rajarishi Sinha | Carnegie Mellon University |
| Mitchell Small | Carnegie Mellon University |
| Richard Sonnenblick | Lumina Decision Systems |
| Kevin Soo Hoo | Lumina Decision Systems |
| Tim Sullivan | E&S Environmental Chemistry, Inc. |
| Robb Turner | Oak Ridge National Laboratory |

ATTACHMENT 2

Reviewer________________ Presenter(s) ____________

## TAF PEER REVIEW QUESTIONNAIRE

Q1. Parent Model Credibility (for cases of parent/reduced-form modules)

Scientific credibility of the enhancements or additions to the full-form (parent) model used as the basis for this module made since the 1990 NAPAP Assessment, including recalibration and reverification of the parent model with currently available empirical data.

Comments:

Q2. Credibility of TAF Module Derived from a Full-Form Parent Model or from an Empirical Data Set

Scientific credibility of TAF module (approach and implementation) including calibration and verification using results from a parent model within defined limits of applicability -- or from empirical data, both with explicit characterization of assumptions and unknowns.

Comments:

Q3. Suitability of Module for use in Policy-related NAPAP Assessments

Utility of module output measures for directly addressing the scoping, synthesis or analysis of the policy issues to be addressed by NAPAP in the 1996 Assessment (see Attachment A).

Comments:

Reviewer ____________________

## TAF PEER REVIEW QUESTIONNAIRE (CONT)

Q4. Readiness of Module for use in 1996 NAPAP Assessment in March, 1996

Completeness of module software, availability of appropriate data sets, need for additional validation.

Comments:

Q5. Recommendations for Improvement

Specific recommendations, if any, for possible short-term improvements that would enhance the credibility or availability of this module for the March, 1996 NAPAP Assessment. Recommendations, if any, for module development to support later assessments.

Reviewer______________ Presenter(s)_Henrion/Sonnenblick_

## TAF PEER REVIEW QUESTIONNAIRE (FRAMEWORK SUPPLEMENT)

Q1. Achievement of Concept Design Goals

Extent to which model implementation has achieved the goals of a) open architecture with accessibility for module substitution and upgrade, b) utility and credibility of reduced-form module concept with acceptable incremental uncertainty, c) transparency of assumptions, d) presentation of sensitivity, and e) robustness for addressing alternative issues.

Comments:

Q2. Overall Model Credibility

Scientific credibility of overall TAF model. Ability to support scientifically defensible integrated assessments of acidic deposition including effects, the benefits of emission reductions, and their associated costs.

Comments: