QUALITY ASSURANCE AND IMPROVEMENT OF MICRO-SCALE METEOROLOGICAL MODELS

Michael Schatzmann[1] and Rex Britter[2]
[1]University of Hamburg, Meteorological Institute (ZMAW), Hamburg, Germany
[2]University of Cambridge, Engineering Department, Cambridge, UK

INTRODUCTION

The emergence of increasingly powerful computers enabled the development of tools that have the potential to predict flow and transport processes within the urban canopy layer. These new tools are micro-scale meteorological models of prognostic or diagnostic type. Micro-scale meteorological models are special in so far as they are tailored to the needs of meteorologists. They are adjusted to domain sizes of the order of several decametres to a few kilometres (street canyons, city quarters). They usually use boundary conditions based on surface characteristics like land use, roughness and displacement thickness. Typically these models contain a substantial amount of empirical knowledge, not only in the turbulent closure schemes but also in the use of wall functions and in other parameterisations. Models play an important and often dominant role in environmental assessment and urban climate studies that are undertaken to investigate and quantify the effects of human activity on air quality and the local climate. Their increasing use is paralleled by a growing awareness that the most of these models have never been subject to rigorous evaluation. Consequently there is often a lack of confidence in the modelled results.

OBJECTIVE AND METHODOLOGY

The main objective of action COST 732 is to improve and assure the quality of micro-scale meteorological models that are applied for predicting flow and transport processes in urban or industrial environments. In particular it is intended

- to develop a coherent and structured quality assurance procedure for this type of model,
- to provide a systematically compiled set of appropriate and sufficiently detailed data for model validation (www data bank),
- to invite from all participating states model developers and users to apply the procedure and to prove its serviceability,
- to build a consensus within the community of micro-scale model developers and users regarding the usefulness of the procedure,
- to stimulate a widespread application of the procedure and the preparation of quality assurance protocols which prove the 'fitness for purpose' of the models,
- to contribute to the proper use of models by disseminating information on the range of applicability, the potential and the limitations of such models,
- to establish a consensus on 'best practises' in current model use and
- to give recommendations for focussed experimental programmes in order to improve the data base.

It is to be expected that the very existence of a widely accepted European standard for quality assurance in the field of micro-scale meteorological models in combination with the provision of suitable validation data will significantly improve "the culture" within which such models are developed and applied. European model developers shall find step-by-step guidance on how to demonstrate that their models are fit for a particular purpose. Data sets (both flow and concentration data) obtained from extensive experiments will be made accessible and more widely exploited. Relevant expertise available within the member states will be brought together and combined to develop a consensus for appropriate model use and model improvement.

RESULTS AND DISCUSSION

The action started in July 2005 with a joint ESF/COST 732 Exploratory Workshop on ‘Quality Assurance of Micro-Scale Meteorological Models’ in Hamburg. About 45 scientists from Europe and the US (the number of participants was limited in order to allow ample discussions) attended the workshop. The workshop proceedings (Schatzmann and Britter 2005) contain a state of the art report on former quality assurance initiatives in the field of micro-scale meteorological models. These initiatives comprise the 'General Requirements for a Quality Assurance Project Plan' by Borrego and Tchepel (1999), the 'Guidelines for Model Developers' and the 'Model Evaluation Protocol' which were worked out by the Model Evaluation Group (MEG, 1994) under the CECs Major Industrial Hazards Programme, the US-Environmental Protection Agency’s requirements for quality assurance of atmospheric dispersion models (Irwin, 1998 and 1999) and the experience gathered within the initiative for harmonization of atmospheric dispersion modelling for regulatory purposes (Olesen, 1999 and subsequent papers). Results from similar initiatives in related fields were also taken into account, for example from the investigations carried out within the ‘Podbi’-model inter-comparison exercise (Lohmeyer et al., 2002), from the FP5 project EMU (Hall 1997), the thematic network QNET-CFD or COST Action C14 which dealt with the industrial application of CFD codes for engineering applications. Finally, the recommendations given by national bodies, e.g., the Quality Assurance Guidelines released by a task force of UKs 'Royal Meteorological Society' (1995) and by Germanys 'VDI Commission on Clean Air' (2002), were carefully considered. With respect to data the considerations outlined in Schatzmann et al (2002, 2003) were taken into account and standards for validation data were defined which can generally only be met by data sets based on a combination of field and wind tunnel experiments.

Strategies for assuring the quality of a numerical model can only be based on very generic scientific principles such as the principle of falsification (K. R. Popper, 1959). The decision about which particular tests should be performed and which particular data sets should be used for comparisons between model results and observations can ultimately be only based on a consensus built up within and by the scientific community. The impact of COST 732 is dependent on whether the quality assurance procedures suggested by the Action are accepted by the community of model developers and users or not. Therefore, the next logical step was to draft a first version of the evaluation procedure and its underlying motivation in order to provide the basis for subsequent discussions within the scientific community. This was done in form of two related documents: A rather lengthy

- Background and justification document to support the model evaluation guidance and protocol document (Britter, R., and Schatzmann, M. 2007 a) and a much shorter
- Model evaluation guidance and protocol document (Britter, R., and Schatzmann, M. 2007 b).

The first document contains detailed explanations concerning the general model evaluation philosophy and the sequence of tasks that should be completed. These tasks are

- Model description: this should be a brief description of the characteristics of the model, the intended range of applicability, the theoretical background on which the model development was based, the software and hardware requirements, etc.

- Database description: a complete description of the database that is to be employed for the evaluation of the model, including the reasons why this specific database was chosen. An estimation of the data variability is required.
- Scientific Evaluation: this is a description of the equations employed to describe the physical and chemical processes that the model has been designed to include. If appropriate it should justify the choice of the numerical modelling procedures and it should clearly state the limits with respect to the intended applications.
- (Code) verification: this process is to verify that the model produces results that are in accordance with the actual physics and mathematics that have been employed. This is to identify, quantify and reduce errors in the transcription of the mathematical model into a computational model and the solution (analytical or numerical) of the model.
- Model validation: this is a structured comparison of model predictions with experimental data and is based on statistical analyses of selected variables. It seeks to identify and quantify the difference between the model predictions and the evaluation datasets; it provides evidence as to how well the model approximates to reality. A quantification of the uncertainty of the model predictions should be produced.
- User-oriented assessment: is there a readable, comprehensive documentation of the code including technical description, user manual and evaluation documentation? The range of applicability of the model, the computing requirements, installation procedures, and troubleshooting advice should be available.

Five of the steps of the evaluation procedure described above are relatively straightforward but the model validation is complex and requires more attention. Unfortunately this has led to the often-seen model evaluation study that is no more than the validation step. At the heart of the complexity of the model validation process is the stochastic nature of atmospheric flows, whether real or physically modelled. For example, and prior to any comparison between mathematical model and experimental results, the user or model evaluator needs to address issues such as:

- Which quantities should be compared?
- At which point within the area of interest should the comparison take place?
- Should the comparison take place on a point-to-point basis or on an area averaged basis?
- Should the compared quantities be averaged over a specific period of time and if so what is the time over which the averaging should take place?
- Should the quantities be compared at the same time or at different times?

The answers to these questions become clearer when the purpose of the model is precisely stated. The various metrics to be used need to be carefully selected and agreed upon. Experience has shown that there may be some generally expected values for these metrics for “state of the art/science” models when applied to particular data sets subject to a specified protocol.

A special section is devoted to validation data requirements. A suite of data sets with increasing geometrical complexity is needed that allows systematic testing of numerical codes. The data sets must be ‘complete’, i.e. they must contain sufficient information to set up a model run without further assumptions concerning the model input parameters and the uncertainty of the data must be known. It is explained that the uncertainty of field data cannot easily be quantified based on the results of field measurements alone. It is not just the accuracy of the instrumentation used for field measurements that defines the reliability of field data. In addition, the repeatability of field measurements for similar boundary conditions as well as the spatial representativeness of individual measurement locations with respect to a particular flow and dispersion problem must be evaluated and quantified with respect to the measured quantities before corresponding data can be used safely for model validation purposes. This is why COST 732 suggests validation data sets that always comprise combinations of field and laboratory experiments. The background document closes with a glossary of terms since words like ‘validation’, ‘verification’, ‘evaluation’, ‘quality assurance’ etc. are not unambiguously defined and used.

The second document (model evaluation guidance and protocol document) is a condensed version of the background document. It gives step-by-step guidance to model developers and users on how to assure the quality of a micro-scale meteorological model. The final guidance and protocol document will come along with recommendations for data sets that should be used during the validation work. These data sets will be made accessible in a unified format via a www data bank. In practise the quality of model output depends not only on the accuracy of the model itself and the model input. Likewise important is the qualification of the person running a model. Numerical simulation is a knowledge-based activity. Appropriate knowledge can be transferred to users by recommendations concerning the proper use of models. For obstacle resolving CFD codes such recommendations are not straightforward. COST 732 tried to respond that problem by drafting a third document, the

- Best practice guideline for the CFD simulation of flows in the urban environment (Franke et al., 2007a)

The recommendations given in the set of COST documents are presently tested by the action itself. The Mock-Up Urban Setting (MUST) data set which comprises field and wind tunnel experiments from flow and dispersion experiments carried out within and above an urban building array made up by 256 ship containers was selected and brought into a usable form. 11 groups of numerical modellers (9 CFD and 2 non-CFD) started to simulate the MUST case thereby following the evaluation guideline. At a first workshop that took place in Hamburg in January 2007 the results were presented and compared and the differences were discussed. Different evaluation metrics were tested and recommendations for fair comparisons were given. It followed another meeting in February in Brussels that was mainly used to draw conclusions from the MUST exercise, to discuss the next steps and to make appropriate changes in the three before-mentioned documents.

FUTURE WORK

Europe-wide discussion of the quality assurance procedure, the use of specific data sets and the recommendations given in the Best Practise Guideline will lead to a harmonised and accepted approach. A quality assurance activity will be launched and the community of model developers and users will be invited to apply the procedure to their models and to prepare model evaluation protocols based on selected data sets. This will be combined with a model inter-comparison exercise within which several model developers and users will simulate identical cases. The intent is not to pillory models that perform badly or to rank the models in one way or the other. That only blocks the flow of information and obstructs scientific exchange. The differences in model results should be discussed and the reasons for deviant model results should be investigated. The strengths and weaknesses of particular modules, parameterisations or closure schemes will be determined. It is expected that modellers will take this opportunity to test various modules, develop common views about the most appropriate set-up of micro-scale meteorological models and, thereby, the quality standard of micro-scale meteorological models and their application will significantly improve. This leads to the expectation expressed in the Introduction that the 'culture' within which urban air pollution models are developed and applied will be significantly improved.

REFERENCES

Britter, R., and Schatzmann, M. (Eds.) 2007a: Background and justification document to support the model evaluation guidance and protocol document. COST Office Brussels, ISBN 3-00-018312-4.

Britter, R., and Schatzmann, M. (Eds.) 2007b: Model evaluation guidance and protocol document. COST Office Brussels, ISBN 3-00-018312-4.

Borrego, C., and Tchepel, O. 1999: General Requirements for a Quality Assurance Project Plan. Proceedings, 3rd SATURN Workshop, Aveiro, Portugal.

Franke, J., Hellsten, A., Schlünzen, H., and Carissimo, B. (Eds.) 2007a: Best Practice Guideline for the CFD simulation of flows in the urban environment. COST Office Brussels, ISBN 3-00-018312-4.

Hall, R.C. (Ed.) 1997: Evaluation of modelling uncertainty - CFD modelling of nearfield atmospheric dispersion. EU Project EV5V-CT94-0531, Final Report. WS Atkins Consultants Ltd., Woodcote Grove, Ashley Road, Epsom, Surrey KT18 5BW, UK.

Irwin, J.S. 1998: Statistical Evaluation of Atmospheric Dispersion Models. Proceedings, 5th Int. Conf. on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Rhodes, Greece.

Irwin, J.S. 1999: Effects of Concentration Fluctuations on Statistical Evaluations of Centreline Concentration Estimates by Atmospheric Dispersion Models. Proceedings, 6th Int. Conf. on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Rouen, France.

Ketzel, M., Louka, P., Sahm, P., Guilloteau, E., Sini, J.F., and Moussiopoulos, N. 2001: Inter-Comparison of Numerical Urban Dispersion Models. Proceedings, 3rd Int. Conf. on Urban Air Quality, Loutraki, Greece.

Lohmeyer, A., Müller, W.J., and Bächlin, W. 2002: A comparison of street canyon concentration predictions by different modellers. Final results from the Podbiexercise. *Atmospheric Environment*, **36**, pp. 157-158.

MEG 1994: Model Evaluation Group: 'Guideline for Model Developers' and 'Model Evaluation Protocol'. European Community, DG XII, Major Technological Hazards Programme, Brussels, Belgium.

Olesen, H.R. 1999: Model Evaluation Kit – Recent Developments. Proceedings, 6th Int. Conf. on Harmonization within Atmospheric Dispersion Modelling for Regulatory Purposes, Rouen, France.

Popper, K. R. 1959: The Logic of Scientific Discovery. Basic Books, New York.

Royal Meteorological Society 1995: Atmospheric Dispersion Modelling. Guidelines on the Justification of Choice and Use of Models, and the Communication and Reporting of Results. RMS, 104 Oxford Road, Reading, RG1 7LJ, UK.

Schatzmann, M., and Leitl, B. 2002: Validation and application of obstacle resolving urban dispersion models. *Atmospheric Environment*, **36**, pp. 4811-4821.

Schatzmann, M., Grawe, D., Leitl, B., and Müller, W.J. 2003: Data from an urban street monitoring station and its application in model validation. Proceedings, 26th International Technical Meeting on Air Pollution Modelling and its Application. Istanbul, Turkey, May 26-30.

Schatzmann, M., and Britter, R. (Eds.) 2005: Proceedings from the International Workshop on 'Quality assurance of microscale meteorological models'. European Science Foundation, ISBN 3-00-018312-4.

Verein Deutscher Ingenieure 2002 Richtlinie 3783, Blatt 9, VDI Guideline on Environmental Meteorology - Prognostic microscale windfield models - Evaluation for flow around buildings and obstacles. Beuth Verlag Berlin.