
Quality assurance and improvement of micro-scale meteorological models

Michael Schatzmann*

Meteorological Institute (ZMAW),
University of Hamburg,
Bundesstr. 55, D-20146 Hamburg, Germany
E-mail: schatzmann@zmaw.de

*Corresponding author

Rex Britter

Engineering Department,
University of Cambridge,
Trumpington Street, Cambridge, CB2 1PZ, UK
E-mail: rb11@eng.cam.ac.uk

Abstract: Models have begun to play an important role in environmental assessment and urban climate studies. Their increasing use, however, is paralleled by a growing awareness that the majority of these models have never been the subject of rigorous evaluation. Consequently there is a lack of confidence in the modelled results. The paper reports on a European initiative (COST 732) which was launched 2005 with the purpose of establishing a generally accepted procedure for the improvement and quality assurance of micro-scale meteorological models that are applied for predicting flow and transport processes in urban or industrial environments.

Keywords: model evaluation strategy; urban flow and dispersion; fitness for purpose; European harmonisation; COST action.

Reference to this paper should be made as follows: Schatzmann, M. and Britter, R. (xxxx) 'Quality assurance and improvement of micro-scale meteorological models', *Int. J. Environment and Pollution*, Vol. x, No. x, pp.xxx-xxx.

Biographical notes: Michael Schatzmann is Professor at the Meteorological Institute of the University of Hamburg. He is in charge of the Technical Meteorology Division which develops meso- and micro-scale meteorological models and operates a boundary layer wind tunnel laboratory. He chairs action COST 732.

Rex Britter is the Professor of Environmental Fluid Dynamics in the Department of Engineering at the University of Cambridge, UK. His research interests include fundamental studies into turbulent fluid dynamics, particularly those involving buoyancy. This is paralleled with operational interests in the flow and dispersion of hazardous materials, conventional pollutant dispersion problems in complex geometries such as cities, formalised model evaluation procedures, urban air quality, sustainable energy use in cities and security issues. He is vice-chairman of COST 732.

1 Introduction

The Framework Directive on Air Quality Assessment and Management together with its ‘daughter’ directives is a key element of present-day European environmental legislation. It addresses air quality within conurbations and near to major sources. It is the first time that a European directive explicitly requires the use of models as tools for the execution of air quality policy.

Though the directive does not require the harmonisation of models across Europe, the performance, representativeness and accuracy of results should be based on quality assured models that can be inter-compared across national borders in order to ensure sound, equitable and effective protection and/or mitigation measures.

For the short-range local problems (0–5 km) at stake here, simple Gaussian type models have generally been used. These models are applicable for pollutant emissions into uniform atmospheric flows (for example tall stack releases in flat, unobstructed terrain and averaged over a large number of atmospheric conditions). It is accepted that these models are not appropriate for predicting flow and concentration in complex structured urban or industrial areas, which is unfortunately where pollutants that are of major concern at present (NO_x, PM₁₀, VOC) are emitted. They result from traffic and low elevation domestic or industrial sources. Additionally, hazardous and critical situations generally occur under certain specific flow conditions embedding meteorological and dynamical disturbances that make the flow conditions far from uniform or stationary, for which these models have been developed. Other types of substances that deserve attention are toxic or flammable gases that escape during accidents or that are deliberately released in the context of terrorist attacks. It is common in all these cases that the dispersion takes place near to the ground within or slightly above the urban or industrial canopy layer. In this layer, the flow is significantly disturbed by buildings and other obstacles. The meteorological and concentration fields are very inhomogeneous and vary rapidly with time.

The emergence of increasingly powerful computers enabled the development of numerical tools that have the potential to meet the new demand for predictions from models. These new tools are micro-scale meteorological models of prognostic or diagnostic type. Prognostic models are based on the Reynolds-Averaged Navier-Stokes (RANS) equations, whereas diagnostic models are less sophisticated and only ensure the conservation of mass. These two model types are presently supplemented by even simpler engineering tools. It is to be expected, however, that the latter will sooner or later be replaced by RANS codes or the even more complex Large Eddy Simulation (LES) models.

The RANS codes belong to the family of Computational Fluid Dynamics (CFD) tools as they are used in various engineering contexts. 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 and they may contain modules that have the potential to simulate chemical transformations, aerosol formation or other important atmospheric physico-chemical processes. 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 parameterisation schemes.

Models have begun to play an important and often dominant role in environmental assessment and urban climate studies that are undertaken to investigate and to quantify the effects of human activity on air quality and the local climate. The increasing use of micro-scale meteorological models is paralleled by a growing awareness that the majority of these models have never been the subject of rigorous evaluation. Consequently, to a certain degree, there is a lack of confidence in the modelled results. To cast doubt on the results is perfectly justified, as was shown by systematic studies in which applications of the same model by different modellers to a given problem (Hall, 1997) and applications of different models by either the same or different modellers to the same problem (Ketzel et al., 2001) revealed significant differences. Nevertheless, these models are used in the preparation of decisions with profound economic and political consequences.

The reason that most of the models lack quality assurance is not due to insufficient efforts made by the model developers. It is mainly caused by

- the lack of a commonly accepted quality assurance procedures for such models
- the lack of data sets that are quality checked and commonly accepted as a standard for model validation purposes

2 Objective and methodology

COST is an intergovernmental European framework for international cooperation between nationally funded research activities. COST creates scientific networks and enables scientists to collaborate in a wide spectrum of activities in research and technology (see www.cost.esf.org). The objective of action COST 732 (see <http://www.mi.uni-hamburg.de/Home.484.0.html>) which comprises scientists from 22 European countries 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
- 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.

3 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 CEC’s Major Industrial Hazards Programme, the US-Environmental Protection Agency’s requirements for quality assurance of atmospheric dispersion models (Irwin, 1998, 1999) and the experience gathered within the initiative for harmonisation 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 UK’s “Royal Meteorological Society” (1995) and by Germany’s “VDI Commission on Clean Air” (VDI, 2002), were carefully considered. With respect to data the considerations outlined in Schatzmann and Leidl (2002) and Schatzmann et al. (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 (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 and Schatzmann, 2007a) and a much shorter
- model evaluation guidance and protocol document (Britter and Schatzmann, 2007b).

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 to that problem by drafting a third document, the

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

The recommendations given in the set of COST documents are presently tested by the action itself. The Mock-Up Urban Setting (MUST) data set (Biltoft, 2001; Yee and Biltoft, 2004) 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.

4 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.

Acknowledgements

The authors are grateful for financial support from the European Union through the COST office in Brussels. REB acknowledges support by the Singapore National Research Foundation through the Singapore – MIT Alliance for Research and Technology (SMART), Center for Environmental Sensing and Modelling.

References

- Biltoft, C.A. (2001) *Customer Report for Mock Urban Setting Test*, DPG Document No. WDTCFR-01-121, West Desert Test Center, US Army Dugway Proving Ground, Dugway, Utah, USA, p.58.
- 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, Brussels/Belgium, p.88.

- Britter, R. and Schatzmann, M. (Eds.) (2007b) *Model Evaluation Guidance and Protocol Document*, COST Office Brussels, ISBN 3-00-018312-4, Brussels/Belgium, p.28.
- Borrego, C. and Tchepel, O. (1999) 'General requirements for a quality assurance project plan', *Proceedings, 3rd SATURN Workshop*, Aveiro, Portugal, p.28.
- Franke, J., Hellsten, A., Schlünzen, H. and Carissimo, B. (Eds.) (2007) *Best Practice Guideline for the CFD Simulation of Flows in the Urban Environment*, COST Office Brussels, ISBN 3-00-018312-4, Brussels/Belgium, p.52.
- Hall, R.C. (Ed.) (1997) *Evaluation of Modelling Uncertainty – CFD Modelling of Near Field 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, p.13.
- 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, p 12.
- 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, pp.29–36.
- 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*, Vol. 36, pp.157–158.
- Model Evaluation Group (MEG) (1994) '*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, p.9.
- 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, p.25.
- Schatzmann, M., and Leidl, B. (2002) 'Validation and application of obstacle resolving urban dispersion models', *Atmospheric Environment*, Vol. 36, pp.4811–4821.
- Schatzmann, M., Grawe, D., Leidl, 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*, 26–30 May, Istanbul, Turkey, pp.529–537.
- 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, Brussels/Belgium, p.150.
- Verein Deutscher Ingenieure (VDI) (2002) *Richtlinie 3783, Blatt 9, VDI Guideline on Environmental Meteorology – Prognostic Microscale Windfield Models – Evaluation for Flow around Buildings and Obstacles*, Beuth Verlag, Berlin.
- Yee, E. and Biltoft, C.A. (2004) 'Concentration fluctuation measurements in a plume dispersing through a regular array of obstacles', *Boundary-Layer Meteorology*, Vol. 111, pp.363–415.