17th International Conference on
Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes
9-12 May 2016, Budapest, Hungary

CONCLUSIONS AND REFLECTIONS FROM COST ACTION ES1006 ACTIVITY: WHAT DO WE MISS FOR THE APPLICATIONS OF MODELS IN LOCAL-SCALE EMERGENCY RESPONSE IN BUILT ENVIRONMENTS?

Trini Castelli S.[1], Leitl B.[2], Baumann-Stanzer K[3]., Reisin T.G.[4], Armand P.[5], Andronopoulos S.[6], and all COST ES1006 Members[7]

[1]Institute of Atmospheric Sciences and Climate, National Research Council, Torino, Italy
[2]Meteorological Institute, University of Hamburg, Hamburg, Germany
[3]Central Institute for Meteorology and Geodynamics, Vienna, Austria
[4]SOREQ NRC, Yavne, Israel
[5]CEA, DAM, DIF, Arpajon, France
[6]National Centre for Scientific Research "Demokritos", Institute of Nuclear and Radiological Sciences and Technology, Energy and Safety, Aghia Paraskevi, Greece
[7]http://www.cost.eu/domains_actions/essem/Actions/ES1006, http://www.elizas.eu/

Abstract: The main focus of the COST ES1006 Action has been the evaluation of the airborne pollutant dispersion models applied to accidental or intentional releases in complex built environments, urban or industrial, in the emergency response framework. The results of the three modelling exercises performed during the Action research activity are summarized and discussed, referring to the related model evaluation protocol and the connected best-practice guidelines elaborated during the Action.

Key words: *emergency response, accidental releases, model evaluation*

INTRODUCTION

A main research task of COST Action ES1006 was to evaluate the available local-scale models in built environments, by model inter-comparison, as well as by comparison against test data from qualified field and laboratory experiments. Assuming that a typical atmospheric dispersion model has already been validated with regard to local-scale dispersion simulation, the existing model evaluation and validation strategies were extended towards task- and application-specific measures for accidental release scenarios in emergency response conditions. Thus, additional quantities, such as extreme value prediction and exposure assessment, have been considered for the evaluation of the model performances. Both continuous and puff releases were taken into account in three modelling exercises carried out during the Action. The model evaluation focuses on the output obtained by the different model categories considered. On the basis of the results, and for supporting their analysis, a model evaluation protocol was established with an application-oriented approach and recommendations for its adoption in emergency response were proposed. The comprehensive analysis was then finalized issuing best-practice guidelines for consolidating the use of atmospheric dispersion models into emergency response tools, with a particular attention to the needs of model users and stakeholders.

The main results of the Action are summarized and the main issues still open are discussed. In the context of the conference, we like to propose an interactive discussion on what has been learnt, what are the limitations found in applying up-do-date air pollution models in emergency response, and, in particular, what the scientific community needs still to do in order to address and support the use of atmospheric dispersion models in such context

THE MODELLING EXERCISES

Three case studies were considered as benchmarks for the modelling evaluation exercises (Baumann-Stanzer et al., 2015): **(1)** the Michelstadt case, based on flow and dispersion data gathered in a wind-tunnel experiment carried out at the Environmental Wind Tunnel Laboratory (Hamburg University), where an idealized urban environment was modelled and both continuous and puff releases were reproduced; **(2)** a real-field campaign with continuous and puff releases conducted in a European harbour, named as CUTE 1 case, which was also reproduced in the wind tunnel, named as CUTE 3; **(3)** a real industrial accident occurred in a European Country, named as AGREE case.

In the experimental cases, the hazardous releases occurred in different locations in the built-up area: in open squares, small or wide streets, perpendicular or parallel to the prevailing large-scale flow or in court yards. After a preliminary non-blind test in one Michelstadt case, in all other cases the exercises were performed as ‘blind tests’, that is only the minimum information on the flow was provided to the modellers, while no information on the measured concentrations were available.

Different modelling tools were applied, from Gaussian type, to Lagrangian and advanced Eulerian CFD and LES. Given the varying airborne hazards flow and dispersion modelling approaches that were used, models were classified as three main types, according to their flow and dispersion approach characteristics, as reported in Table 1.

Table 1. Classification of the types of models applied in the COST ES1006 Action modelling exercises

Model type	Modelling approach: Flow	Dispersion
Type I	models that do not resolve the flow between buildings	Gaussian
Type II	models resolving the flow diagnostically or empirically	Lagrangian
Type III	models that resolve the flow between buildings	Eulerian

In case of continuous releases, steady-state concentrations, and the area affected by values above a relevant threshold are usually the information expected from an emergency response model. In case of puff releases, the dosage, the arrival time of the puff at given locations, the duration of the puff passage and the peak concentration values are of interest. In Figures 1 and 2 examples of the results based on scatter plots are shown.

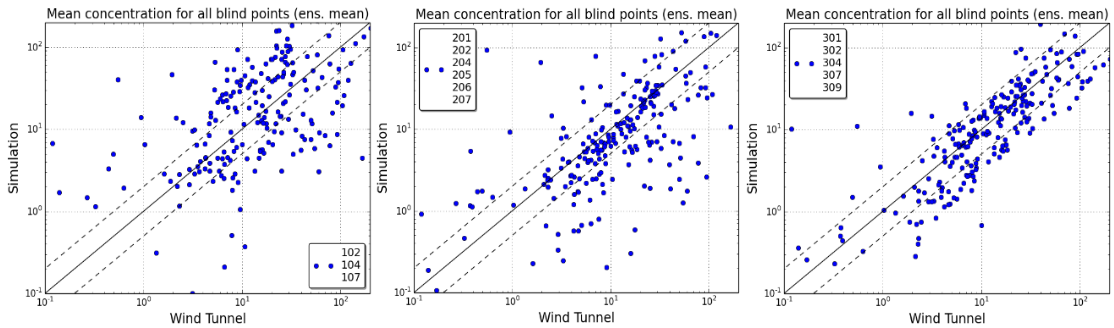

Figure 1. Michelstadt test-case. Scatter-plots of measured versus ensemble averaged modelled mean concentration values at all receptors for all blind continuous releases for model Type I (left), Type II (centre) and Type III (right).

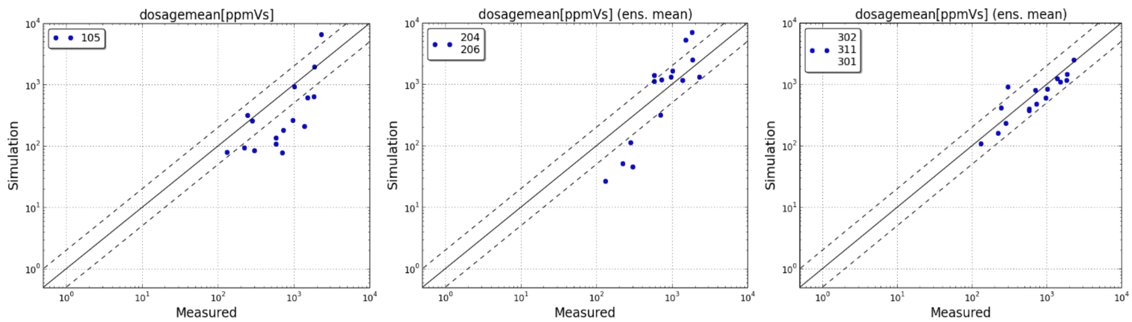

Figure 2. CUTE 3 test-case. Scatter-plots of measured versus ensemble averaged modelled mean dosage values at all receptors for all blind puff releases for model Type I (left), Type II (centre) and Type III (right).

To provide an overall view on the different test cases, a statistical analysis was applied to Michelstadt and CUTE experiments. Summarizing the results, it was found:

for continuous releases, (i) the metrics are generally better for the non-blind case, where more detailed meteorological information for the model input was available; (ii) the metrics show a better performance and agreement with the observations under controlled conditions (wind tunnel data); (iii) Type II and Type III models are generally superior to Type I ones, in particular for the FAC2 and the correlation coefficient R; (iv) some specific metrics, MG and VG, hardly get good values, since they are strongly influenced by extremely low values and are undefined for zero values: this type of problems are discussed in the Model Evaluation Protocol document (see next section);

for puff releases, (i) Type II and Type III statistics are often comparable, with some better performances for Type III models; (ii) the metrics do not differ substantially with respect to the continuous case, even if in some cases they are a bit worse: (iii) the metrics are not always better for the non-blind case with respect to the blind one.

The comparison of the statistics between the continuous and puff releases confirms the robustness of the models even in simulating short and transient events and suggests that, since the main difficulty here is the correct reproduction of transient events, the 'blindness' of the test play a secondary role.

In general, despite of the trend to improving performances with a higher complexity in modelling, the accuracy of the results produced by the most advanced models is still not guaranteed to be always and fully satisfying. The availability of proper inputs was proved once more to play a fundamental role for obtaining reliable results, based on sensitivity analyses. Yet, the models showed to be robust enough even when dealing with poor driving information, as it generally happens in case of accidental releases. Models showed to represent a valid tool to support handling emergency situations and can be applied with a certain confidence, all uncertainties considered when dealing with unexpected situations. It was established that to drive the choice of the kind of modelling approach, a balance between the model performances, thus its reliability, and the run-time effort, given that a fast answer is required, has to be considered. Different modelling approaches can be used in the different phases of the response process: preparatory, emergency and post-analysis. However, a fast but inaccurate model output can compromise the effectiveness of a response action, and this is another criterion to consider when taking decisions on what modelling tool to adopt.

MODEL EVALUATION PROTOCOL AND BEST PRACTICE GUIDELINES

The pertinence and applicability of the standard evaluation protocols and, specifically, of the commonly used statistical figures for the evaluation of the model performances in the emergency response context was a main subject matter for discussion. Recommendations based on the experiences gained through the course of COST Action ES1006 and, specifically, from the modelling exercises, are detailed in the Model Evaluation Protocol (MEP) document (Andronopoulos et al., 2015). The Best Practice Guidelines (BPG) document (Armand et al., 2015) integrates the results obtained and the analysis performed in the course of the Action, to organize information and guidelines based on state-of-the-art atmospheric dispersion models (ADM), at the same time tailoring them to the needs of the emergency responders, decision makers and/or stakeholders.

In applying ADM for accidental or deliberate hazardous releases in built areas, it is necessary to address both general and specific requirements for each of the three distinct phases of emergency response and preparedness: (1) pre-accidental analysis and planning (a priori predictions); (2) predictions during an actual emergency; (3) post-accidental analysis (a posteriori simulations). The final objective is to propose guidance for using ADM integrated (or not) in Emergency Response Tools (ERT, Figure 3), dedicated to decision-support in case of releases of hazardous materials into the atmosphere

The basic structure of the MEP adopted in the Action assumes to provide the following: (i) model description; (ii) description of the database used for the validation; (iii) scientific assessment of the model; (iv) user-oriented assessment of the model; (v) code verification (software errors); (vi) validation (or corroboration) of the model by comparing model predictions with observations; (vii) sensitivity and uncertainty analyses. It was then established that using ADM for emergency response, in application to releases of airborne hazardous materials in urban areas, needs specific important requirements: (i) computation of dispersion from transient-in-time releases, (ii) computation of flow and dispersion in built up (urban or industrial) environment, (iii) computation of affected areas based on a defined threshold of a quantity of interest, (iv) modelling of special relevant physico-chemical phenomena, (v) addressing the required computing resources (computing time and hardware).

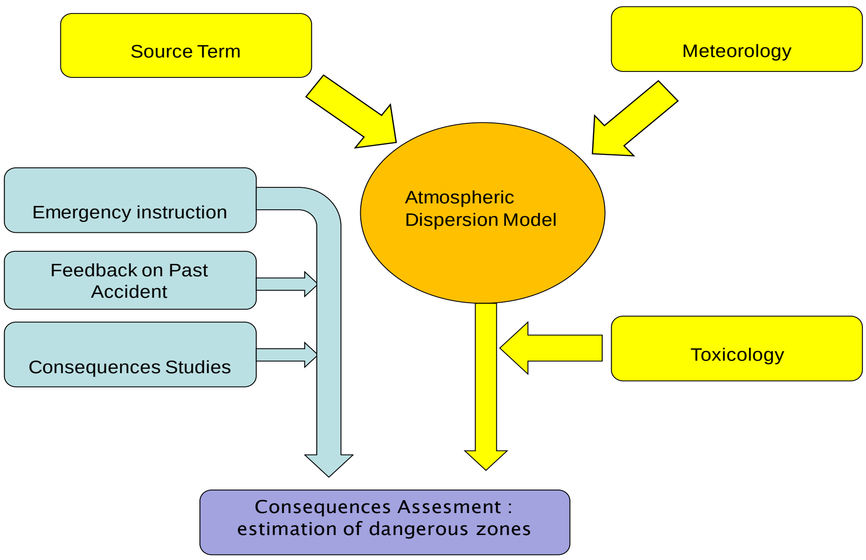

Figure 3. General sketch-up of an Emergency Response Tool (ERT).

The use of standard statistical metrics provides a usual platform for the validation of any ADM. In addition to these, within the context of emergency response and releases in urban areas, an important indicator of a model's fitness for purpose is the correct prediction of spatial and temporal extension of risk zones or affected areas. As stated in the BPG document, this evaluation should be based on exposure values depending on specific threat scenarios, which define the interval of concentrations that lie in the specific hazard zone. The affected areas can be defined through different quantities, but for emergency response cases, it is recommended to define them through Levels of Concern (LOC) values, such as AEGLs (Acute Exposure Guideline Levels) or IDLH (Immediate Danger to Life or Health).

A fundamental aspect is that the evaluation process should reflect a consensus among the various parties involved: the model developer, the model user and the stakeholder who undertakes the task of decision making. The improvement of the model has to be guided by the user and stakeholder requirements, and supported by the provision of guidance for its application within the context of emergency response: this topic is thoroughly addressed in the BPG document. The interaction between the different parties is thus highly recommended. Since the simulation results contain uncertainties due to the model formulation, to the input data and to the inherent variability of the physical system, it is also recommended to communicate to the stakeholders in a transparent and understandable manner the quantified uncertainties of the numerical outcomes. The BPG document is also devoted to highlight the supporting information which can be provided to first responders by ADM and ERT in an emergency.

CONCLUSIONS AND OPEN ISSUES

From the experience gained, in the context of emergency response, it is highly recommended to validate an ADM against several sets or test cases of observational data. These sets should include both dedicated boundary layer wind tunnel measurements as well as dedicated field trials. Experiments should be designed as application-oriented, so that the observed data reflect the difficulties associated with the specific type of the problem, which is accidental or deliberate release in real urban areas. Test cases should treat many different release scenarios: varying source release rates, different source locations and even simultaneous releases from different sources in a city. This implies the need for a close collaboration of experts in the fields of observational techniques and modelling as well as first responders and key policy and decision makers. Each test case dataset has to be easily accessible and well documented. When it gets difficult to locate experimental data sets descriptive of conditions of interest for validation of models for emergency response, it is suggested to consider also inter-comparisons among the results from different models in real cases. Model to model validation needs to be treated with care, and more systematic work is needed to come up with clear rules on this subject. In particular, each reference model data set that is used for the comparison needs to have its uncertainty adequately quantified. These requirements need to be more precisely defined with additional research work and sufficient application experience.

Several open issues need further investigation and a concerted and harmonized approach to address them. Among them, here we briefly list some main items to open a discussion in the scientific community.
In a real emergency situation, the required input information to ADM is not always available: *how to deal with incomplete information of source term and meteorological data and their uncertainties?* The source term is the most difficult information to retrieve. Meteorological input data should not be an issue in principle, but their access needs to be facilitated. *How to systematically take into account and communicate stochastic and epistemic uncertainties,* on model physics and input data, is a very complicated matter and still a topic for scientific research. In emergency response, it is crucial not to under-estimate the actual consequences of a noxious dispersion event: *how to produce reasonably conservative results?* Conservative results should be in a range that enables to adopt actions for which the benefits are greater than their drawbacks, with the need of being realistic in the dispersion computation and the distribution of the noxious agent. Related to this, it is important to establish *how to overcome different results obtained by different models or operators*: consistency has to be found between models in the same or in neighbouring categories. Moreover, even if and ADM is verified and validated, it cannot give results that correspond exactly to the field measurements: an effort and an agreed approach are needed to establish *how to reconcile the modelling results and the field measurements*. This is an important issue especially when supporting decision makers, responding at the same time to the question on *how to reconcile the needs and demands of the people involved*, whom in the end take care of the actions to face the emergency. The final question, on which consensus needs to be built, is: *how to choose the appropriate ADM and ERT*? The choice should take into consideration a series of conditions and parameters, such as the topography and morphological characteristics of the area, the climatology of the area, the scenario of the release, the expertise of operators, the computational and operative resources available, the emergency phase and the time restrictions.

The COST ES1006 Action succeeded in addressing these items and in proposing possible solutions to some open issues, based on an international debate in the frame of the Action activity. We like to promote further the discussion inside the ADM community, in order to establish a harmonized approach to the problem of applying models for emergency response scenarios in complex, built environments.

REFERENCES

Andronopoulos S., Barmpas F., Bartzis J.G., Baumann-Stanzer K., Berberkar E., Efthimiou G., Gariazzo C., Harms F., Hellsten A., Herring S., Jurcakova K., Leitl B., Trini Castelli S., 2015. COST ES1006 - Model evaluation protocol, Publisher: University of Hamburg, ISBN: 987-3-9817334-1-9 http://www.elizas.eu/images/Documents/Model%20Evaluation%20Protocol_web.pdf

Armand, P., J. Bartzis, K. Baumann-Stanzer, E. Bemporad, S. Evertz, C. Gariazzo, M. Gerbec, S. Herring, A. Karppinen, J.-M. Lacome, T. Reisin, R. Tavares, G. Tinarelli, and S. Trini Castelli, 2015 COST ES1006 Best Practice Guidelines. COST Action ES1006, April 2015. Publisher: University of Hamburg, 987-3-9817334-0-2 http://elizas.eu/images/Documents/Best%20Practice%20Guidelines_web.pdf

Baumann-Stanzer, K., S. Andronopoulos, P. Armand, E. Berbekar, G. Efthimiou, V. Fuka, C., Gariazzo, G. Gasparac, F. Harms, A. Hellsten, K. Jurcacova, A. Petrov, A. Rakai, S. Stenzel, R. Tavares, G. Tinarelli, S. Trini Castelli, 2015. COST ES1006 Model evaluation case studies: Approach and results. COST Action ES1006, April 2015. Publisher: University of Hamburg, 987-3-9817334-2-6 http://elizas.eu/images/Documents/Model%20Evaluation%20Case%20Studies_web.pdf