COST
EUROPEAN COOPERATION
IN SCIENCE AND TECHNOLOGY

17th HARMO

9–12 May 2016 **Budapest, Hungary**

Conclusions and reflections from COST Action ES1006 activity: what do we miss for the application of models in local-scale emergency response in built environments?

Trini Castelli S., Leitl B., Baumann-Stanzer K.,Reisin T.G., Armand P., Andronopoulos S. and all COST ES1006 Members

COST Action ES1006

Outline

- COSTES1006 Results
- Lessons learnt
- Conclusions
- Some guidelines
- Open issues

http://www.elizas.eu

Motivation

… not just another model validation exercise!

- establishing a **platform for information exchange** to characterize, quantify and verbalize limitations and advantages of modelling approaches
- establishing **consensus on state-of-the-art in local-scale hazmat dispersion modeling**
- developing **methodologies, means, tools and data for rigorous testing** and task-specific evaluation of models
- providing **context specific guidance** for the reliable use of models
- drafting **concepts for improvements** in modelling and model application

The test cases

Wind tunnel experiments (EWTL, Inst. Met., Hamburg University)

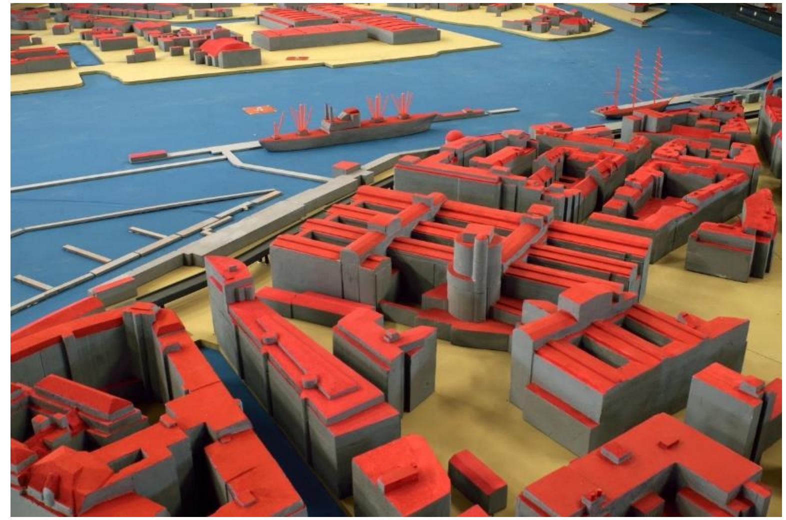
Michelstadt

A **typical** **European urban site** is reproduced.

Several **continuous** and **puff releases** from six different source locations: concentration measured at more than 30 points

Non-blind and **blind** tests

CUTE 3



A **real** **European city** is reproduced.

Several **continuous** and **puff releases** from three different source location: concentration measured at more than 30 points

Blind tests

The test cases

A field experiment: CUTE 1

Continuous 45-minutes release of SF6 with a flow rate of 2 g/s, from a b
the harbor area.

Concentration detected by 2
measurement stations locate
positions.

Each measurement station h
samplers. Each bag was fillec
minutes => 10-minute average values.

A real accident: AGREE

Vinyl Chloride Monomer accidentally
released inside a building in a liquid state
evaporated causing high
ns in the air outside the

nts gathered by the local VCM
onitoring network (more than
) installed around the plant

intervention of Firemen: the
managed and closed after
about 50 minutes

Blind tests!

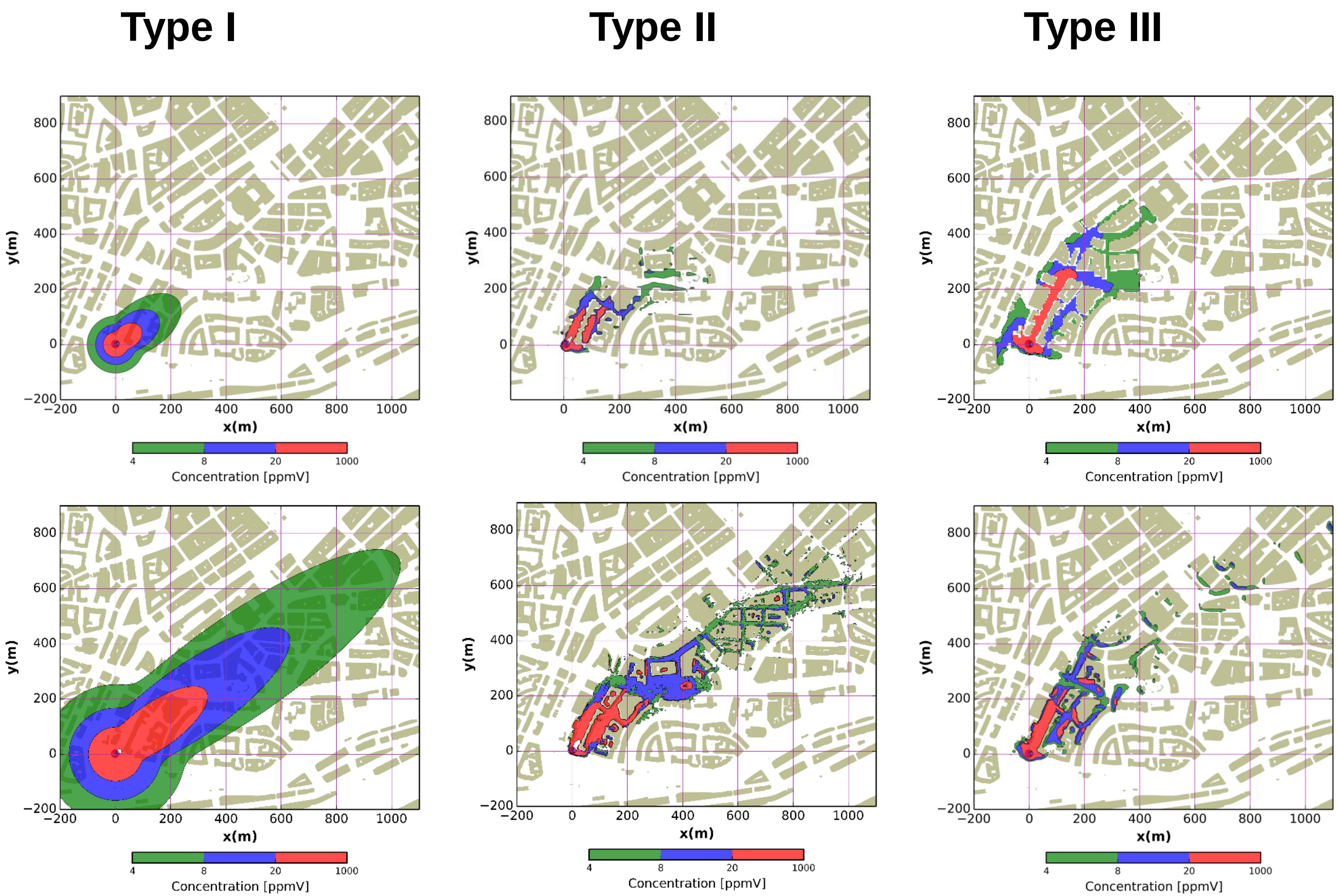
The models

Model type	Flow modelling approach	Dispersion modelling approach
Type I	models that do not resolve the flow between buildings	Gaussian
Type II	models for which the flow is resolved diagnostically or empirically, although not dynamically resolving the flow between buildings	Lagrangian
Type III	models that resolve the flow between buildings	Eulerian

Test case	# of ERT used			# of ADM used		
	Type I	Type II	Type III	Type I	Type II	Type III
Michelstadt	4	2	0	4	4	7
CUTE	2	0	0	2	5	5
AGREE	2	1	0	1	3	3

Test case	# modellers applying ERT			# modellers applying ADM		
	Type I	Type II	Type III	Type I	Type II	Type III
Michelstadt	5	2	0	5	4	12
CUTE	2	0	0	2	5	5
AGREE	2	1	0	2	5	4

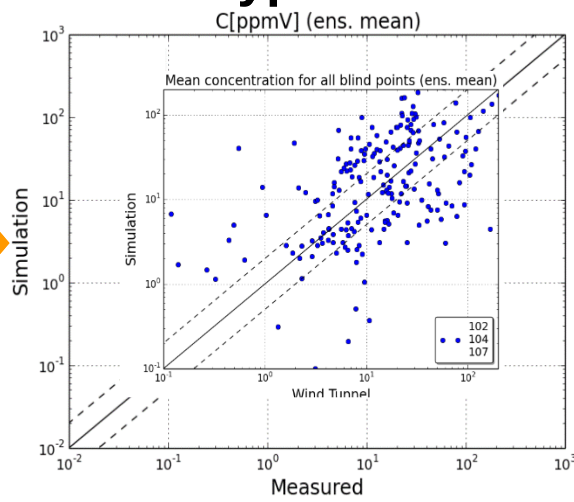
CUTE – blind test, continuous release



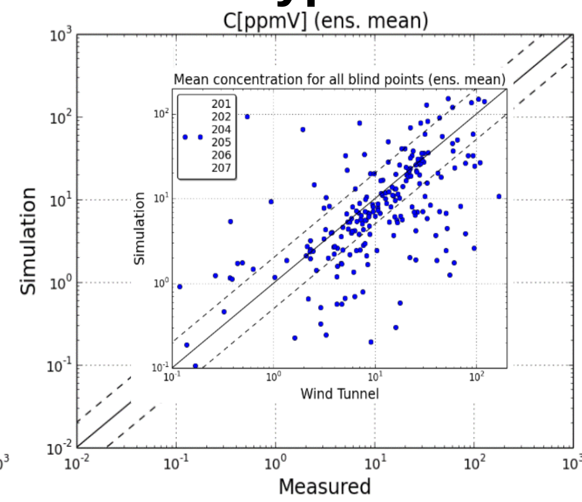
COST
EUROPEAN COOPERATION IN SCIENCE AND TECHNOLOGY

Michelstadt and CUTE - **blind** test, continuous release

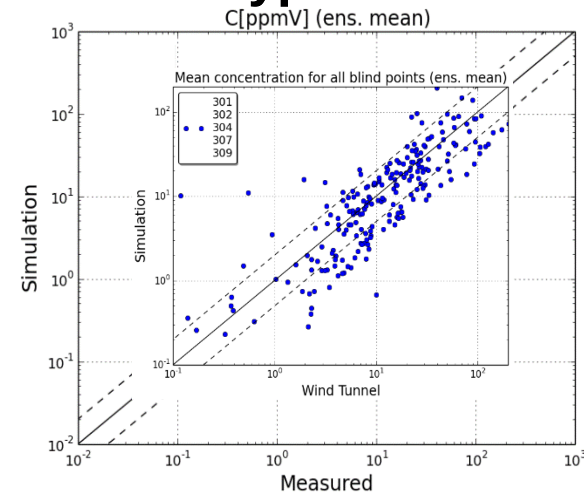
Type I

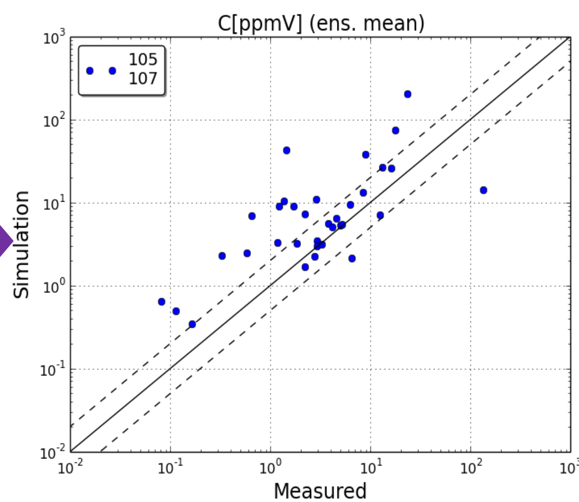


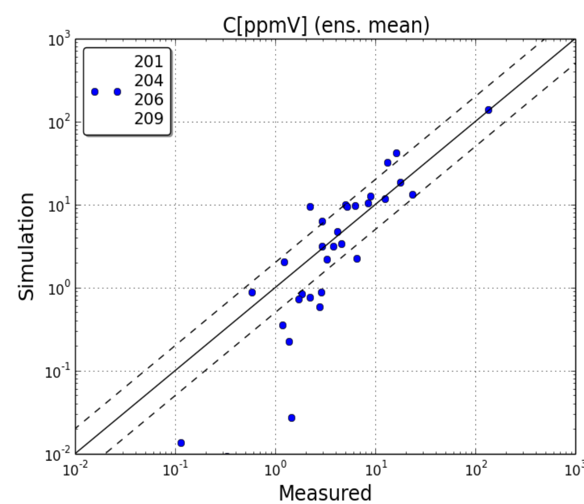
Type II

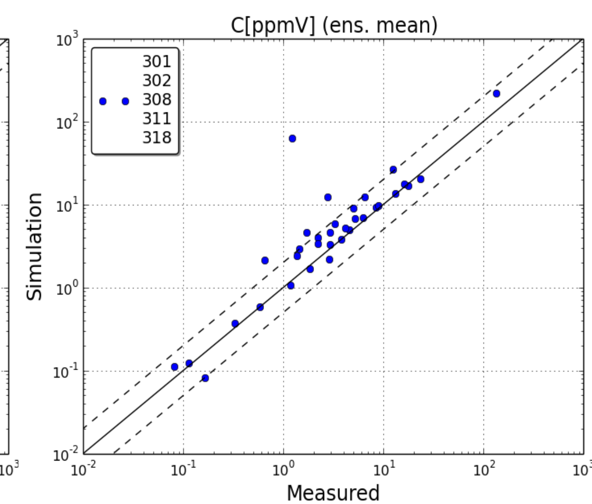


Type III





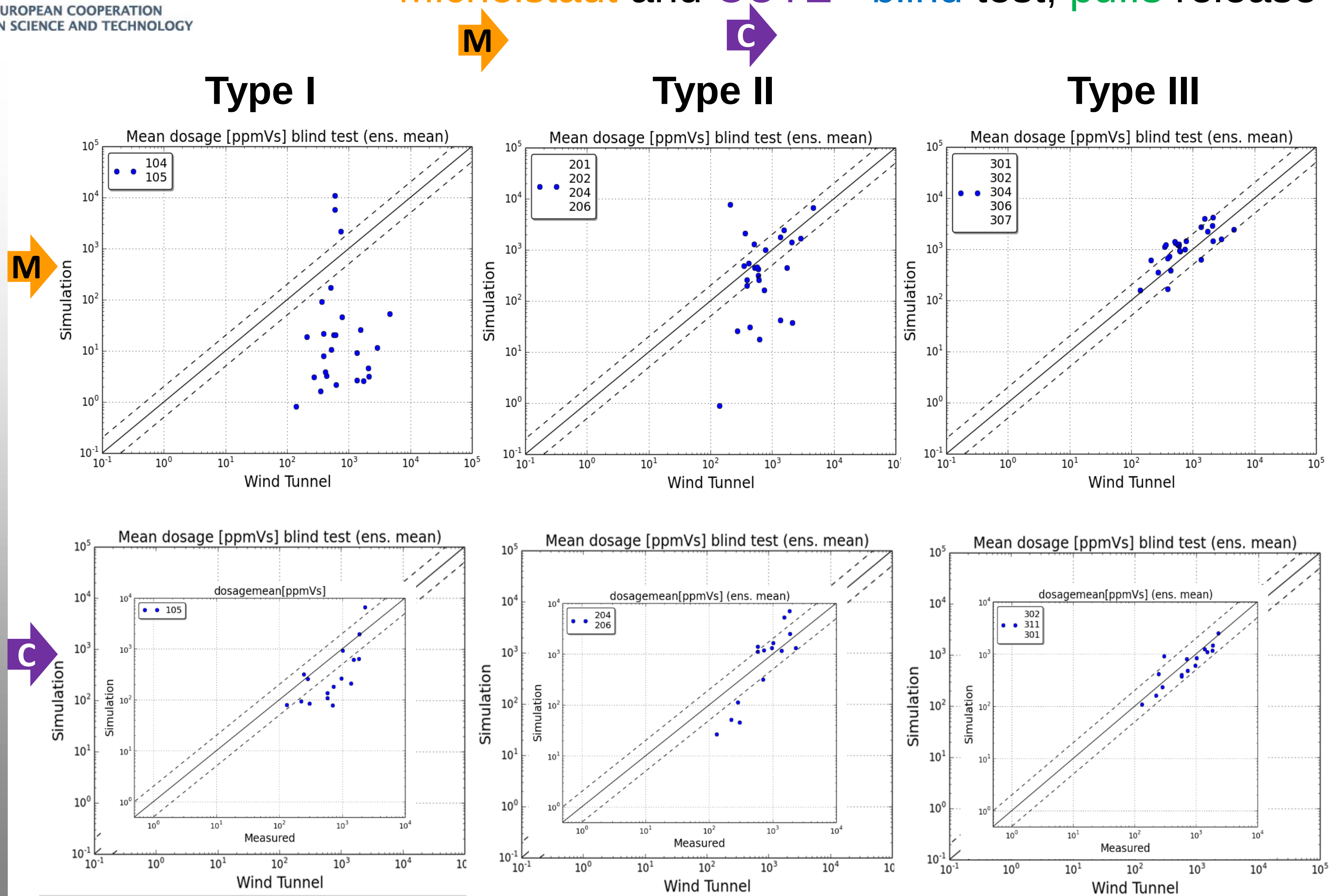




Mean concentration, ensemble average

COST Action ES1006

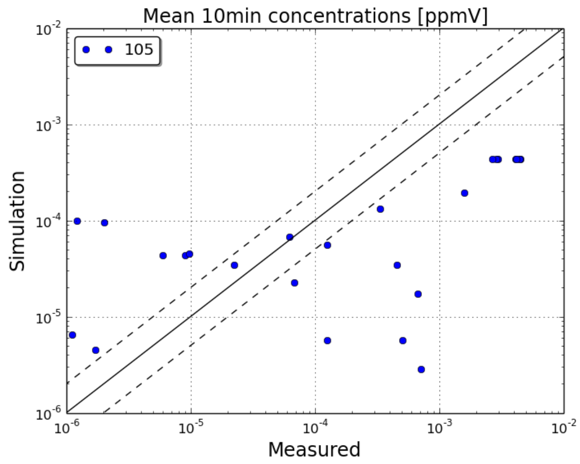
Michelstadt and CUTE - blind test, puffs release



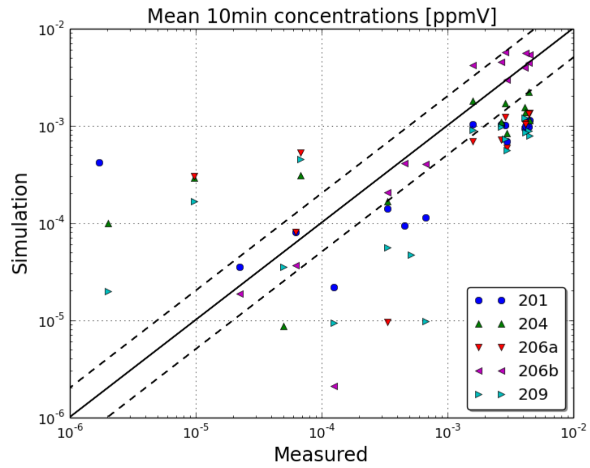
Mean dosage, ensemble average

10-minutes mean concentration (ppmV)

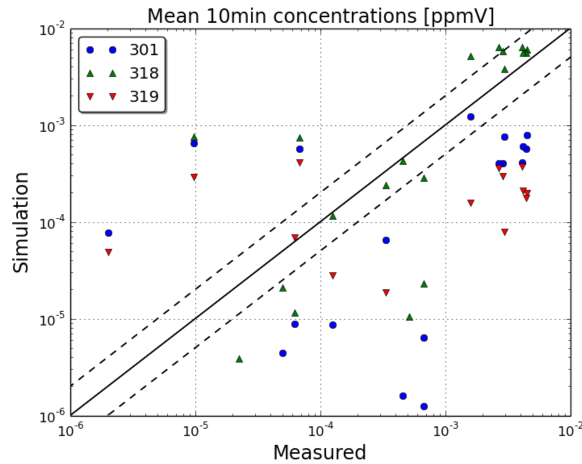
Type I



Type II



Type III



Observations and predictions mainly show rather small concentration values, in a range from 10^{-6} to 10^{-2} ppmV.

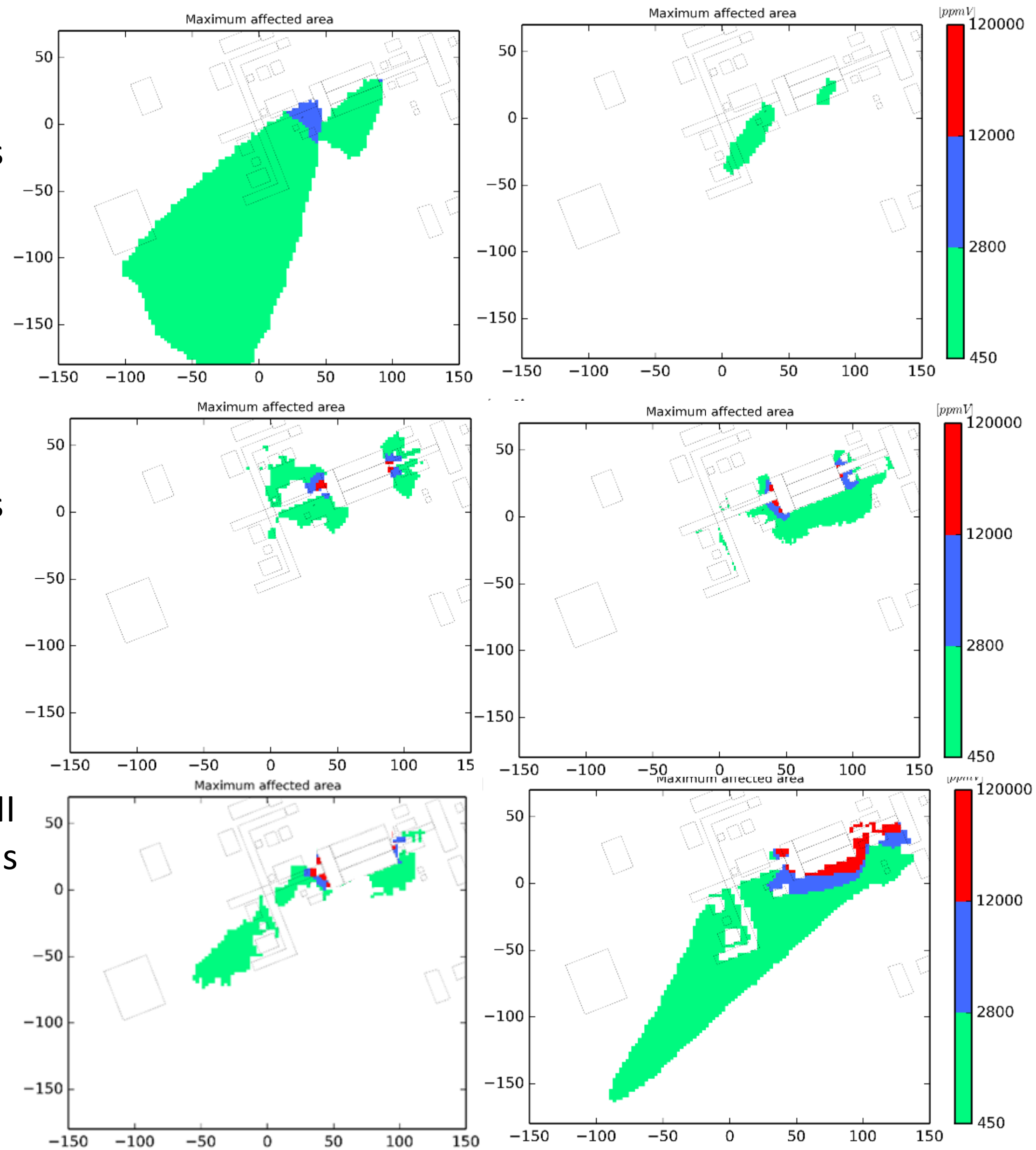
COST Action ES1006

AGREE real accident - blind test

Type I models

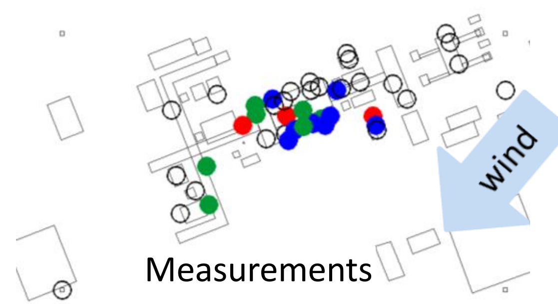
Type II models

Type III models



Comparison of size and position of predicted affected areas (Levels Of Concern, AEGL)

Threshold values:
AEGL values 450, 2800, 12000 Vppm VCM



Some statistics……….

CONTINUOUS RELEASE Mean concentration		model	|FB|		NMSE		FAC2	
			min	MAX	min	MAX	min	MAX
Michelstadt	NonBlind	Type I	0.06	0.37	2.6	9.7	0.15	0.44
		Type II	0.06	0.72	2.09	4.35	0.39	0.64
		Type III	0.02	0.27	0.39	9.44	0.25	0.71
	Blind	Type I	0.36	0.50	7.42	16.11	0.09	0.34
		Type II	0.36	0.67	9.01	13.98	0.34	0.45
		Type III	0.03	0.68	1.35	15.76	0.17	0.77
CUTE	Case 1	Type I	1.20		27.50		0.10	
		Type II	0.03	1.07	3.01	23.01	0.30	0.57
		Type III	0.10	1.41	4.82	39.85	0.18	0.42
	Case 3	Type I	0.56	1.05	12.69	16.21	0.24	0.30
		Type II	0.07	0.34	0.28	2.27	0.35	0.61
		Type III	0.02	1.02	0.30	24.08	0.29	0.77

PUFF RELEASE 15s-average peak concentration		model	|FB|		NMSE		FAC2	
			min	MAX	min	MAX	min	MAX
Michelstadt	NonBlind	Type II	0.40	1.25	1.55	3.39	0.13	0.67
		Type III	0.19	1.72	0.46	23.20	0.13	0.63
	Blind	Type II	0.35	1.69	2.81	17.82	0.08	0.31
		Type III	0.13	1.55	0.43	12.10	0.15	0.77
CUTE	Case 3	Type II	0.17	0.77	0.44	2.71	0.38	0.50
		Type III	0.33	0.64	0.38	1.95	0.44	0.63

Minimum and maximum values of the metrics of Michelstadt and CUTE experiments.

Acceptance criteria in built environments: |FB|<0.67, NMSE<6, FAC2>0.3
(Hanna and Chang, *Met. Atm. Phys.* 116, 2012)

What we have learnt – I – idealised scenarios

- the model performance is influenced significantly by the location of sources and receptor points, *due to the complexity of the geometry*
- metrics are within the acceptance values for most models
- the model performance increases with increasing model complexity
- difference was observed between the blind and non-blind tests, *but not systematic*
- consistency of results increases with model complexity, *for more advanced type, results of different models look more similar*
- the model evaluation for puff releases is by far more complex than for continuous release, *the puff-to-puff variation affects also the observed values and needs to be accounted for with a statistical approach*
- the performance of the models is mostly fair also for the puffs release, *higher bias are found with respect to the continuous case*
- type I models give generally the worst results for the puff releases

What we have learnt – II – real scenarios

- the intrinsic variability of atmospheric motions in real field makes the comparison more critical
- sensitivity to input data: non-representative reference measurements, especially in real field, may heavily affect the performance and response of the models
- …….. real life is turbulent and difficult …..
- real accident datasets, like AGREE, cannot be properly used for model evaluation: large input uncertainties e.g. emissions, unevenly placed on-site concentration measurements with unknown accuracy
- the buildings at this (typical) industrial site influence flow and dispersion
- all models tend to underestimate the measured concentrations in this case
- the simple Type I models fail in predicting the spread of the plume
- Type II and Type III models predicted a wide lateral spread of the plume: this was confirmed by the measurements

Conclusions

- despite uncertainties, models represent a valid tool to better support handling of emergency situations
- a higher level of physical description in a model is worthwhile to achieve higher simulation accuracy, also in difficult scenarios like emergency response
- choice of model/modelling approach is always a compromise between performance, reliability and response time
- different modelling approaches are used in the different phases of the response process
- more reliable results from complex models facilitate better emergency response

A model verified and validated for air quality is not automatically valid for local scale hazmat dispersion: transient phenomena !!

Some guidelines - I

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)

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 physic-chemical phenomena
(v) addressing the required computing resources (computing time and hardware)

Some guidelines - II

In addition to the standard statistical metrics, 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.*

The model 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).

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.

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.

Quino

Open Issues - I

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?

how to systematically take into account and communicate stochastic and epistemic uncertainties?

In emergency response, it is crucial not to under-estimate the actual consequences of a noxious dispersion event, yet providing realistic patterns of the pollutant distribution:

how to produce reasonably conservative results?

how to overcome different results obtained by different models or operators?

Open Issues - II

ADM (verified and validated) results do not correspond to the field measurements:

how to reconcile the modelling results and the field measurements, with ***consensus****?*

how to reconcile the needs and demands of the people involved, to support decision makers?

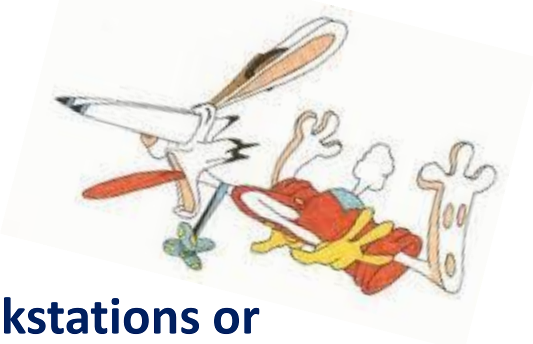
how to choose the appropriate ADM and ERT?

Consensus! … taking into account
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.

Comments?

.... and the computational time?

Test case	computational range times required for the model runs		
	Type I	Type II	Type III
Michelstadt	1 – 5 minutes	2 min – 5 hours	2 hours – 4 days
CUTE	1 – 5 minutes	1 - 3 hours	1 – 19 days



Hardware used here was ranging from laptops to workstations or supercomputers, clusters

Run configuration, scalar or parallel processing, also determines the needed run time

COST Action ES1006