MODEL EVALUATION AND TRACER EXPERIMENT UNCERTAINTIES

Silvia Trini Castelli and Domenico Anfossi

CNR, Institute of Atmospheric Sciences and Climate (ISAC) Turin, Italy

Abstract: The typical errors in the real field measurements are considered and their effects on model evaluation are investigated. We analyzed some tracer releases from the KATREX data set in which contemporary emissions of different gases at two heights were performed and parallel ground level sampling and analysis were done by two different teams. A median error of 14% among the various concentration estimations was found.

Key words: *Model evaluation, tracer experiments, measurement uncertainties*

INTRODUCTION

Air quality models are generally evaluated against tracer experiments before being accepted as reliable tools in regulatory applications. In such a way one can understand how well they are able to simulate different meteo-diffusion and terrain conditions. The evaluation of a model is usually based on statistical analyses based on the calculations of standard measures such as the fractional bias, normalized mean square error and correlation coefficient, from observed and predicted concentrations at the sampler locations. Implicit in this analysis is that the observations are affected by a negligible experimental error, which usually the modeler does not know. Thus, it might be interesting to investigate which is the typical error in the real field measurements. To do that, we examined some tracer data from the KATREX data set (Thomas et al., 1983). During one exercise a tracer was emitted at the 160 m of the Karlsruhe meteorological tower (Germany) and sampled at ground by two teams with their own devices for two consecutive 30-min periods. Two other tracers were simultaneously emitted at the 195 m of the tower and again sampled by the two teams. In particular, at one point each team positioned 2 samplers. Thus, four estimates, in principle coincident, were done there. In our investigation we firstly compared the different concentration estimates at the various sampler points, in order to assess the average error in the measurements. Then, we simulated these experiments with a simple model, the analytical Gaussian plume model, using in input the sigma curves evaluated by Thomas et al. (1983) in their original work based on this database. In this second phase of the work we could assess how much the above statistical measures would be different using the same model, by comparing it to the different set of simultaneous tracer observations. In this comparison it is not the model accuracy to be of interest but it is interesting how different (but in principle equal) tracer data may give different answers on the quality of the same model.

THE KATREX EXPERIMENT

The KATREX data set (Thomas et al., 1983; Thomas and Nester, 1984) concerns a series of tracer exercises performed at the Karlsruhe Nuclear Research Center (KNRC), Karlsruhe (Rhine Valley, Germany). Two non-buoyant tracers were simultaneously released from two heights, 160 m and 195 m, of the KNRC meteorological tower (200 m high): difluorodibromomethane (CF_2Br_2) at 195 m and Freon-ll ($CFCL_3$) at 160 m. Out of the comprehensive meteorological information recorded as 10-min averages at five tower levels, we utilized in this work the wind speed and direction at 160 m and 200 m only. The KATREX data set includes: the meteorological information and the estimation of the prevailing stability class, the emission data and ground level concentrations (g.l.c) at various sampling points located on five arcs (about 500 m, 1000 m, 2000 m, 4000 m, 8000 m) around the tower. Tracer samples were collected and analyzed in two subsequent 30-min periods starting about 1 h after the beginning of the emission.

During one exercise (exp. 72), in which neutral stability conditions prevailed, the tracer emitted at the 160 m in the two consecutive 30-min periods was independently sampled at ground and analyzed by two teams (Schuttelkopf et al., 1981), with their own devices: the KNRC team (K_team hereafter) and a second team from Ispra Joint Research Center (I_team hereafter). During the same exercise two other tracers were simultaneously emitted at the 195 m of the tower (CF_2Br_2 by K_team and SF_6 by I_team) and independently sampled and analyzed by the two teams. We mention that co-located sampler were about 2 m apart and that, in particular, at one point each team positioned 2 samplers.

Table 1 presents the eight contemporary cases (tracer experiments) considered in our analysis, whereas Table 2 reports the wind measurements. Figure 1 shows the sampler location during exp. 72.

Table 1. Presentation of the release experiments

Case	Emission height (m)	tracer	team	period
1	160	$CFCL_3$	K	1st
2	160	$CFCL_3$	K	2nd
3	160	$CFCL_3$	I	1st
4	160	$CFCL_3$	I	2nd
5	195	CF_2BR_2	K	1st
6	195	CF_2BR_2	K	2nd
7	195	SF_6	I	1st
8	195	SF_6	I	2nd

Table 2. Wind during the releases

Emission height (m)	period	wind speed (ms^{-1})	wind direction (deg)
40	1st	7.4	67
40	2nd	6.5	66
60	1st	8.5	64
60	2nd	8.2	63
100	1st	10.1	63
100	2nd	9.9	62
160	1st	12.0	67
160	2nd	12.0	65
200	1st	13.0	71
200	2nd	13.1	70

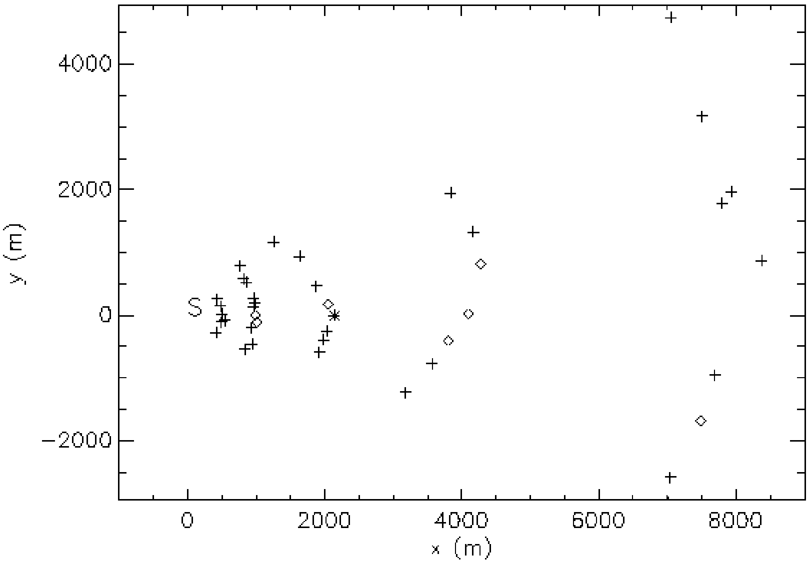

Figure 1 – Source (S) and samplers (crosses, diamonds and asterisk) locations: in the plot, the abscissa is aligned along the 160-m wind direction. The diamonds represents the samplers used in the comparison between the different observed datasets, the asterisk corresponds to the point at which four simultaneous measurement were carried out.

COMPARISON OF THE SIMULTANEOUS CONCENTRATION ESTIMATES AT THE VARIOUS SAMPLER POINTS

In order to be able to compare the estimations in the height exercises (two periods and three emitted gases) concentrations were made nondimensional as follows:

$$C = \frac{\chi}{Q}\overline{u}H^2 \qquad (1)$$

where C is the dimensionless concentration, χ the g.l.c. (μg m^{-3}), Q the emission rate (μg s^{-1}), H the emission height (m) and $\bar{u}$ the wind speed (m s^{-1}). Table 3 shows the nondimensional concentrations at the sampler locations at which the two teams performed simultaneous determinations. "R" is the downwind distance from the source (m) and "dir" its angle (degrees) from the North.

Table 3. Non dimensional tracer concentration, *C*, estimated at the co-located samplers by the two teams.

Case	**1**								
R (m)	**dir (deg)**	**Case 1**	**Case 2**	**Case 3**	**Case 4**	**Case 5**	**Case 6**	**Case 7**	**Case 8**
985	67	0.116	0.030	0.048	0.081	-	-	-	-
1015	73	0.084	0.064	0.054	0.054	-	-	-	-
2050	62	0.069	0.006	0.107	0.009	0.028	0.038	0.053	0.069
2140	67	0.244	0.186	0.914	-	0.205	0.176	0.108	-
2140	67	0.172	0.175	0.069	0.059	0.157	0.166	0.089	0.115
3820	73	0.085	0.058	0.018	0.028	0.109	0.095	0.050	0.052
4100	65	-	-	0.241	0.831	-	-	0.262	0.062
4350	56	-	-	0.011	0.002	-	-	0.003	0.025
7675	68	-	-	0.018	0.028	-	-	-	-
7800	68	-	-	0.140	0.937	-	-	-	-

A statistical analysis was done on the data reported in Table 3 joining together all estimations. In particular, the following indices were computed: Correlation Coefficient (COR), Normalized Mean Square Error (NMSE), Fractional Bias (FB), Factor of 2 (FA2) and the median of the relative errors ε_i (RE_median). We add that for each single pair, the *relative error* ε_i (i = 1, N of pairs) was computed as:

$$\varepsilon_i = \frac{0.5\left|C_{Ki} - C_{Ii}\right|}{0.5\left(C_{Ki} + C_{Ii}\right)} = \frac{\left|C_{Ki} - C_{Ii}\right|}{\left(C_{Ki} + C_{Ii}\right)} \qquad (2)$$

in which C_{Ki} and C_{Ii} represent the i_th nondimensional concentration evaluated by the K_team and I_team, respectively.

The results of the statistical analysis are reported in Table 4.

Table 4. Result of the statistical analysis on the co-located concentration estimation

Samplers	**Cor**	**NMSE**	**FA2 /%)**	**RE_median**
25	0.75	0.38	80	13.9

INFLUENCE OF THE TRACER CONCENTRATION ERROR ON THE MODEL EVALUATION

This part of the work aims at assessing how much the tracer concentration estimation error influences the evaluation of model accuracy. Thus, in this analysis it is not the model accuracy to be of interest but it is interesting how different (but in principle equal) tracer data may give different answers on the quality of the same model. For sake of simplicity the standard analytical Gaussian plume model was used and the σ_y and σ_z curves evaluated by Thomas et al. (1983) in their original work, based on this database, were used.

In Figures 2 and 3 the scatter plots of the model predictions versus the observations are plotted respectively for the two different height emissions. In general the disagreement between predictions and observations, related to an underestimation of the measurements by the model simulations, is analogous for both team datasets in both sampling periods.

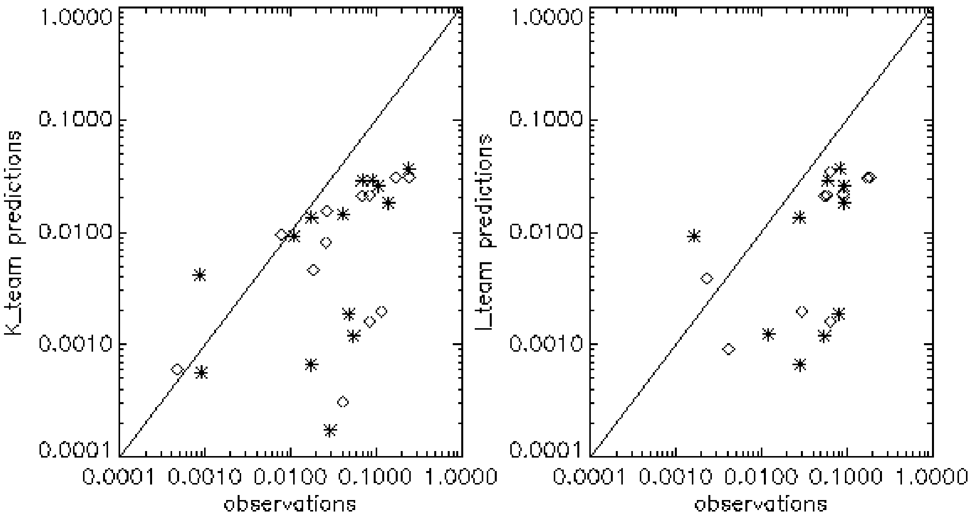

Figure 2. Scatter diagram between predicted and observed adimensional g.l.c. concentrations of the Katrex release 72, emission of CFCL3 at 160 m. K_team (left) and I_team (right) data, first (diamonds) and second (asterisks) periods.

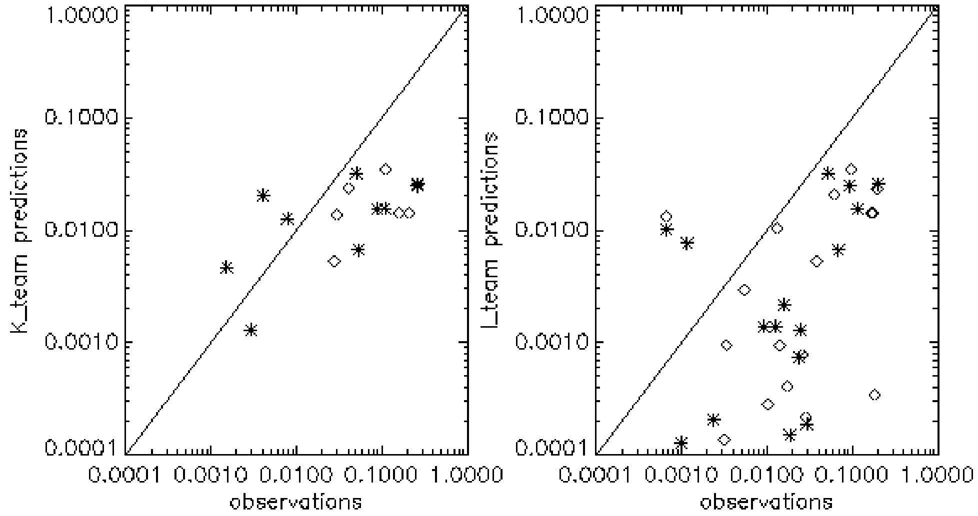

Figure 3. Scatter diagram between predicted and observed adimensional g.l.c. concentrations of the Katrex release 72, emission at 195 m. Tracer CF2BR2 for K_team (left) and tracer SF6 for I_team (right) data, first (diamonds) and second (asterisks) periods.

Next Table 5 shows the results of the comparison among the eight exercises and the corresponding model simulations, in terms of the above statistical indices.

Table 5 indicates that rather different values of the indices are found when using the different observed datasets from K_team and I_team: COR ranges from 0.61 to 0.88, FB from 0.78 to 1.25, NMSE from 1.48 to 9.41and FA2 from 3% to 20%. These results put in evidence that the uncertainties in the observations significantly affect the evaluation of the model performance (see, in particular the NMSE variability). In particular, it is worthwhile to point out that NMSE experienced a difference of a factor of about 6.

This means that the same model, run in the same meteorological conditions, may result to be a good or bad model according to which tracer experiment, that should have in principle all the same accuracy, is compared. This suggests that in any model evaluation the tracer experiment error should be accounted for. Since, in general, the latter is not known, we suggest considering a typical error of about 14% that was found in the present analysis.

Table 5. Result of the statistical analysis on the observed and prescribed concentration in the various cases

Case	Samplers	COR	FB	NMSE	FA2 /%)
1	17	0.87	1.25	5.96	6
2	14	0.79	0.94	2.56	14
3	13	0.88	0.82	1.85	15
4	11	0.61	0.78	1.48	18
5	6	0.87	1.19	3.31	17
6	10	0.81	0.98	3.54	20
7	33	0.66	1.28	9.41	6
8	31	0.67	1.17	7.45	3
all	135	0.76	1.07	4.45	10

CONCLUSIONS

The results of the present work suggest that the concentration determination error in a tracer exercise performed in the real terrain might be not negligible and, consequently, should be taken into account in the model evaluation. A model exercise, in which the concentration of eight tracer experiments, all in principle equivalent, simultaneously evaluated by two teams, were simulated by the same model gave rather different values of four standard measures (COR, NMSE, FB and FA2). In particular, NMSE experienced a difference of a factor of about 6 from one simulation to the other for the same model. This also highlights that the uncertainties in the observations significantly affect the evaluation of the model performance.

REFERENCES

Schuttelkopf, H., P. Thomas , S. Vogt, M. De Bortoli and P. Gaglione, 1981: Experimental Determination of the Atmospheric Dispersion Parameters: Experiment Performed July 2nd 1980 at the Karlsruhe Nuclear Research Centre. EUR-75577 EN Report

Thomas, P, W. Hubschmann, H. Schuttelkopf and S. Vogt., 1983: Experimental determination of the atmospheric dispersion parameters at the Karlsruhe Nuclear Research Center for 160 m and 195 m emission heights. Part 1: measured data. KfK Report 3456

Thomas, P. and K. Nester, 1984:) Experimental determination of the atmospheric dispersion parameters at the Karlsruhe Nuclear Research Center for 160 m and 195 m emission heights. Part 2: evaluation of measurements. KfK Report 3457