SKILL AND UNCERTAINTY OF THE REGIONAL AIR QUALITY FORECAST SYSTEM FOR THE APULIA REGION (ITALY)

Ilenia Schipa[1], Angela Morabito[1], Annalisa Tanzarella[1], Francesca Intini[1], Alessio D'Allura[2], Camillo Silibello[2], Matteo Paolo Costa[2], Roberto Giua[1]

[1]Regional Environment Protection Agency (ARPA) Puglia, Bari, 70126, Italy.
[2]ARIANET Srl, via Gilino 9, Milan, 20128, Italy

THE AIR QUALITY FORECASTING SYSTEM (AQFS)

http://cloud.arpa.puglia.it/previsioniqualitadellaria/index.html

According to the Air Quality Directive 2008/50/EC, Apulia Regional Environmental Protection Agency (ARPA) has implemented, during 2016, an AQFS to provide information on the expected changes in pollution over the region (Southern Italy). This study reports the evaluation of the AQFS during this first operational year using statistical and category indices based on predicted and observed air quality data collected by the regional network. The AQFS (see Figure 1) is based on FARM chemical transport model (CTM) that implements different gas-phase chemical mechanisms and includes chemical and physical processes involving particulates. The AQFS consists of the following modules:

- GAP interpolates WRF meteorological fields on the simulation grid;
- SURF*Pro* computes additional fields used by the CTM (turbulent dispersion scale parameters, pollutants' deposition velocities, biogenic emissions, etc.);
- EMMA performs the spatial disaggregation, the time modulation and the VOC/PM speciation on the anthropogenic emissions derived from the regional inventory (INEMAR) and the Apulia Territorial Emission Register;
- QualeAria, national scale AQFS (*http://www.aria-net.it/qualearia/en/*), provides boundary conditions;

Finally, post-processing modules compute air quality indicators, verify possible air quality standards exceedances and disseminate results to stakeholders and to general public.

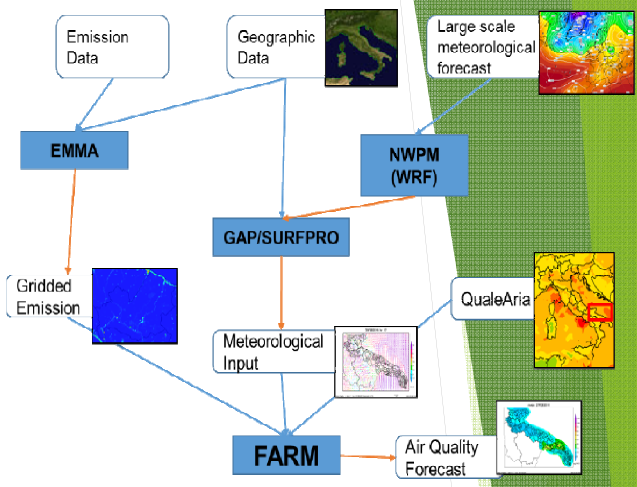

Figure 1. AQFS Schematic representation.

Table 1. Model evaluation statistics and their definition

Root mean square error	RMSE	$\sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i - O_i)^2}$
Correlation coefficient	r	$\frac{\frac{1}{N}\sum_{i=1}^{N}(O_i-\bar{O})(P_i-\bar{P})}{\sqrt{\frac{1}{N}\sum_{i=1}^{N}(O_i-\bar{O})^2}\sqrt{\frac{1}{N}\sum_{i=1}^{N}(P_i-\bar{P})^2}}$
Index of Agreement	IOA	$1-\frac{\sum_{i=1}^{N}(P_i-O_i)^2}{\sum_{i=1}^{N}[\lvert P_i-\bar{O}\rvert+\lvert O_i-\bar{O}\rvert]^2}$
Factor of two	FAC2	Fraction of data for which $0.5 \le \frac{P_i}{O_i} \le 2$

AQFS PERFORMANCE EVALUATION: METHODOLOGY AND RESULTS

To evaluate the AQFS performance, NO_2, O_3, PM_{10} and $PM_{2.5}$ predictions (+24h) have been compared with the observations collected by the regional air-monitoring network managed by Apulia ARPA. The network includes 61 stations; the AQFS evaluation has been performed considering the stations having a spatial representativeness equal or greater than the model horizontal resolution (4 km).

The following analysis was performed to evaluate the model's forecasting skills:

1) using four commonly used scores (see Table 1): root-mean-square error (RMSE), correlation coefficient (r), index of agreement (IOA) and the fraction within a factor of two (FAC2);
2) using four indices (see Table 2) to quantify forecast performance: the accuracy (A), the bias (BIAS), the probability of detection (POD) and the false alarm rate (FAR). The 75th percentile of the observed concentrations for each pollutants has been used as threshold value. These indices are based on the so-called "Contingency table" (Figure 2), that reports:
    - the number of occurrences in which observed data and model output were both above the selected threshold (hits, a) or both below (correct-negative, d);
    - the number of alarms missed by the model (misses, c);
    - the number of false alarms (b).

Event forecast	Event observed		
	Yes	No	Marginal total
Yes	a	b	a + b
No	c	d	c + d
Marginal total	a + c	b + d	a + b + c + d =n

Figure 2. Contingency table

Table 2. Categorical statistical indices

Index name	Formula	Range	Ideal value
Accuracy [%]	$A = \frac{a+d}{n}100$	0 to 100	100
Bias [%]	$BIAS = \frac{a+b}{a+c}100$	0 to 100	100
Probability of Detection [%]	$POD = \frac{a}{a+c}100$	0 to 100	100
False Alarm Ratio [%]	$FAR = \frac{b}{a+b}100$	0 to 100	0

SIMULATION RESULTS AND CONCLUSIONS

Model statistic results for the year 2016 are summarized in Table 3 evidencing a good agreement between predicted and observed mean annual values for all the species (FAC > 50%). The best performance is obtained for $PM_{2.5}$ (IOA = 0.7, r = 0.6 and FAC2 = 90.8 %). More in detail, the correlation coefficient *r* is in the range 0.4-0.7 and IOA shows the best agreement for ozone (0.8). As for the categorical indices, BIAS values show a slight tendency to underforecast for PM_{10}; this tendency increases for NO_2 and $PM_{2.5}$, while O_3 tends to be overpredicted. As of FAR, it can be seen that the AQFS performs well, maintaining a FAR value always smaller than 50%. The analysis of skill scores shows the capability of the AQFS to forecast O_3 exceedances, as indicated by the high POD values. The model skills are within accepted criteria for the considered pollutants, evidencing the good capability of the modelling system to forecast the pollutants levels across the region.

Table 3. Results of forecast evaluation for NO_2, PM_{10}, $PM_{2.5}$ and O_3

	NO_2	PM_{10}	$PM_{2.5}$	O_3
Number of stations	22	21	7	19
Mean obs. [μg m^{-3}]	15.1	18.9	11.9	63.7
Mean pred. [μg m^{-3}]	11.3	13.6	10.4	68.8
RMSE [μg m^{-3}]	13.2	9.4	4.9	23.9
r	0.5	0.4	0.6	0.7
IOA	0.7	0.6	0.7	0.8
FAC2 [%]	54.5	80.5	90.8	86.3
75° percentile [μg m^{-3}]	19.9	23.4	15.1	82.4
BIAS [%]	62.3	20.7	61.7	133
POD [%]	36.6	13.4	40.9	77.3
FAR [%]	41.2	35.2	33.7	42
ACC [%]	77.7	76.5	80.1	80.2

18th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, 9-12 October 2017, Bologna, Italy