# 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]Environmental Protection Agency, ARPA Puglia, Bari, Italy
[2]ARIANET Srl, via Gilino 9, Milan, I-20128, Italy

**Abstract**: During the last year, ARPA Puglia has developed a regional air quality forecasting system (AQFS) to inform the population, as established by the Italian Legislative Decree 155/2010, about the potential risks of limit exceedances. The AQFS, based on WRF prognostic meteorological model and FARM chemical transport model, simulates the fate and the chemical transformation of airborne pollutants over Puglia region. The regional AQFS provides daily 72-hour forecasts for a range of primary and secondary pollutants, including $NO_2$, CO, $O_3$ and airborne particles ($PM_{10}$ and $PM_{2.5}$). Forecasted air quality maps are freely accessible to the public through the ARPA Puglia web site (http://cloud.arpa.puglia.it/previsioniqualitadellaria/index.html) .
A full yearly assessment of the AQFS performances, based on classical statistical parameteres and skill scores, has been undertaken by considering the experimental data collected by the regional air quality monitoring network. The statistical evaluation evidences the good capability for the AQFS to reproduce the pollutants levels across the region.

***Key words:*** *Forecasting system, statistical indicators, skill scores.*

## INTRODUCTION

In order to improve air quality, the European Union introduced the New Air Quality Directive in 2008 and set its Member States strict targets on air pollution concentrations for the most harmful substances, such as $NO_2$, $O_3$ and fine particles. The same Directive requires the local authorities to inform the population not only on the air quality status, but also on the predictable trend for the following days by implementing an air quality forecasting system. Therefore, the performance evaluation of an air quality forecasting system is an important issue and a thorough assessment of forecast quality requires the computation of statistical parameters and skill scores (Bennet *et al.*, 2013; Zhang *et al.*, 2012).
In order to activate an adequate information and alert system for population, ARPA Puglia, following article 14 and article 18 of Legislative Decree 155/2010, has developed a modelling system for the Apulia Region. The goal of the system, which couples the meteorological model WRF with the photochemical model FARM, is to provide daily forecasts for the current day and the next two days through their daily web-site publication (http://cloud.arpa.puglia.it/previsioniqualitadellaria/index.html).
This work describes the full yearly assessment of the AQFS performances, based on classical statistical parameteres and skill scores, considering the experimental data collected by the regional air quality monitoring network.

## FORECASTING SYSTEM STRUCTURE

Forecasting system structure FARM is a 3-dimensional multiphase Eulerian atmospheric chemistry and transport model, able to work with different chemical schemes and to treat chemical and physical processes involving the particulates.
The meteorological fields, provided by WRF on a wider domain, are interpolated on the simulation grid through the application of the interface module GAP. After this step, the meteorological module SURFPro is used to calculate the turbulent dispersion scale parameters, the deposition velocitiesy of pollutants.
The emissions, gridded on the regional domain by EMMA processor, are derived from the regional INEMAR inventory (http://www.inemar.arpa.puglia.it/) and the Territorial Emission Register of the Apulia region (http://www.cet.arpa.puglia.it), appropriately integrated and updated with available information. The

initial and boundary conditions are provided by the QualeAria national air quality forecasting system (http://www.aria-net.it/qualearia/en/).
Post-processing modules compute air quality indicators, verify possible air quality standards exceedances and disseminate results to stakeholders and general public.

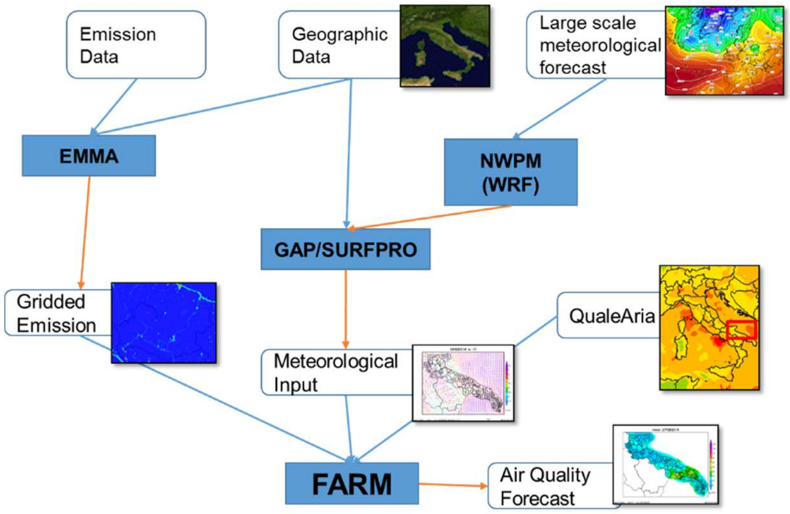


**Figure 1**. Schematic representation of AQFS.

## AQFS PERFORMANCE EVALUATION: METHODOLOGY AND RESULTS

The regional air-monitoring network, managed by the Regional Environmental Protection Agency (ARPA), is equipped with 61 stations of different type, all active in the year 2016. To evaluate the performances of the adopted modelling system, the $NO_2$, $O_3$, $PM_{10}$ and $PM_{2.5}$ predictions were compared with the observations with a spatial representativeness equal or greater than the model horizontal resolution.
Two types of estimation have been conducted to evaluate the model's forecasting skills. The first type has been made by 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). These scores assess different aspects of forecast quality. RMSE is one of the most basic and widely used methods of verification and assesses the average magnitude of forecast errors (Stanski *et al.*, 1989); smaller values indicate better agreement between measured and calculated values. The correlation coefficient reflects linear association between the forecasts and observations. Index of agreement (IOA) is a standardized measure of the degree of model prediction error; it is a nondimensional and bounded measure with values closer to 1 indicating better agreement. FAC2 is a measure of the proportion of predictions within a factor two of the observed concentration; it is recommended that an air quality model is considered acceptable if more than half of the model predictions lie within a factor of 2 of the observations and faulty if not.

**Table 1.** Model evaluation statistics and their definition

| Name | Symbol | 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 $0.5 \leq \frac{P_i}{O_i} \leq 2$ |

*Note.* $P_i$ is the *i*-th predicted value, $O_i$ is the *i*-th observed value, $N$ is the number of observed and predicted pairs, $\bar{P}$ is the mean predicted, $\bar{O}$ is the mean observed.

The second type of estimation has been made by using some categorical indices. These are based on the so-called ”Contingency table” (Figure 1), that reports the number of occurrences in which observed data and model output were both above the selected threshold (hits, a), the number of occurrences in which they were both below (correct-negative, d), the number of alarms missed by the model (misses, c) and the number of false alarms (b). The contingency table is a useful way to see what types of errors are being made. A perfect forecast system would produce only hits and correct negatives, and no misses or false alarms. We used five indices to quantify forecast performance: the accuracy (A), the bias (BIAS), the probability of detection (POD) and the false alarm rate (FAR). Table 2 shows some details of these categorical indices. The 75th percentile of the observed concentrations for each pollutants has been used as treshold value, according to Pay *et al.* (2014).

Observed Events

| Predicted Events | YES | NO | TOTAL |
|---|---|---|---|
| YES | HITS *a* | False alarms *b* | Predicetd Yes |
| NO | Misses *c* | Correct negatives *d* | Predicted No |
| TOTAL | Observed Yes | Observed No | Total *n* |

**Figure 2.** Structure of a contingency table

**Table 2.** Categorical statistical indices

| Index name | Formula | Range | Ideal value | Note |
|---|---|---|---|---|
| **Accuracy [%]** | $A=\frac{a+b}{n}100$ | 0 to 100 | 100 | *The level of agreement between the forecast and the truth (as represented by observations). It indicates the percentage of forecasts that correctly predicts an exceedance or a nonexceedance.* |
| **Bias [%]** | $BIAS=\frac{a+b}{a+c}100$ | 0 to 100 | 100 | *Measures the ratio of the frequency of forecast events to the frequency of observed events. Indicates whether the forecast system has a tendency to underforecast (BIAS<100) or overforecast (BIAS>100) events.* |
| **Probability of Detection [%]** | $POD=\frac{a}{a+c}100$ | 0 to 100 | 100 | *Is the fraction of observed exceedance conditions that are correctly predicted.* |
| **False Alarm Ratio [%]** | $FAR=\frac{b}{a+b}100$ | 0 to 100 | 0 | *Measures the percentage of times an exceedance was forecast when exceedance did not occur.* |

Model statistic results for the year 2016 are summarized on Table 3. The best performance (RMSE = 4.9 μgm$^{-3}$, IOA = 0.7, r = 0.6 and FAC2 = 90.8 %) is obtained for $PM_{2.5}$, but the comparison between the mean annual values is generally quite good and the model well reproduces observed data for all the species (FAC > 50%). In detail, the correlation coefficient r is in the range 0.4-0.7; the RMSE shows better results for particulate species (9.4 for $PM_{10}$ and 4.9 for $PM_{2.5}$) and IOA shows the best agreement for ozone (0.8).
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. In the case of FAR it can be seen that 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 accuracy values satisfy the performance goal, confirming the good capability of the modelling system to forecast pollutant species over the regional grid.

**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 station** | 22 | 21 | 7 | 19 |
| **Mean (O) [μg m$^{-3}$]** | 15.1 | 18.9 | 11.9 | 63.7 |
| **Mean (P) [μg m$^{-3}$]** | 11.3 | 13.6 | 10.4 | 68.8 |
| **Threshold [μg m$^{-3}$]** | 19.9 | 23.4 | 15.1 | 82.4 |
| **r** | 0.5 | 0.4 | 0.6 | 0.7 |
| **RMSE [μg m$^{-3}$]** | 13.2 | 9.4 | 4.9 | 23.9 |
| **IOA** | 0.7 | 0.6 | 0.7 | 0.8 |
| **FAC2 [%]** | 54.5 | 80.5 | 90.8 | 86.3 |
| **BIAS [%]** | 62.3 | 20.7 | 61.7 | 133 |
| **POD [%]** | 36.6 | 13.4 | 40.9 | 77.3 |
| **FAR [%]** | 41.2 | 35.2 | 33.7 | 42.0 |
| **A [%]** | 77.7 | 76.5 | 80.1 | 80.2 |

## CONCLUSIONS
This study has evaluated the performance of the air quality forecasting system AQFS over the Apulia region, Southern Italy. The evaluation has been made comparing the baseline simulation and the

experimental data, provided by the regional monitoring network, at a 4 km grid resolution, by using statistical indices. The model skills are within accepted criteria for the considered pollutants, evidencing the good capability of the modelling system to reproduce the pollutants levels across the region.
Future work is planned to assess the forecast variations at finer grid reasolution and time scales (diurnally) and at individual monitoring stations.

## REFERENCES


Bennett N.D., Croke B.F.W., Guariso G., Guillaume J.H.A., Hamilton S. H., Jakeman A.J., Marsili-Libelli S., Newham L.T.H., Norton J.P., Perrin C., Pierce S.A., Robson B., Seppelt R., Voinov A.A., Fath B.D., Andreassian V., 2013: Characterising performance of environmental models. *Environmental Modelling & Software*, **40**, , 1-20.

Pay M.T., Martinez, F., Guevara M. Baldasano J.M., 2014: Air quality forecasts on a kilometer-scale grid over complex Spanish terrains. *Geosci. Model Develop.*, **7**, 1979-1999.

Stanski H.R., Wilson L.J., Burrows W.R., 1989: Survey of common verification methods in meteorology. WMO World Weather Watch Technical Report No.8, WMO/TD No. 358.

Zhang Y., Bocquet M., Mallet V., Seigneur C., Baklanovet A., 2012: Real-time air quality forecasting, part I: History, techniques, and current status, *Atmospheric Environment*, **60**, 632-655.