Evaluation of air quality forecasts CCA-WG1
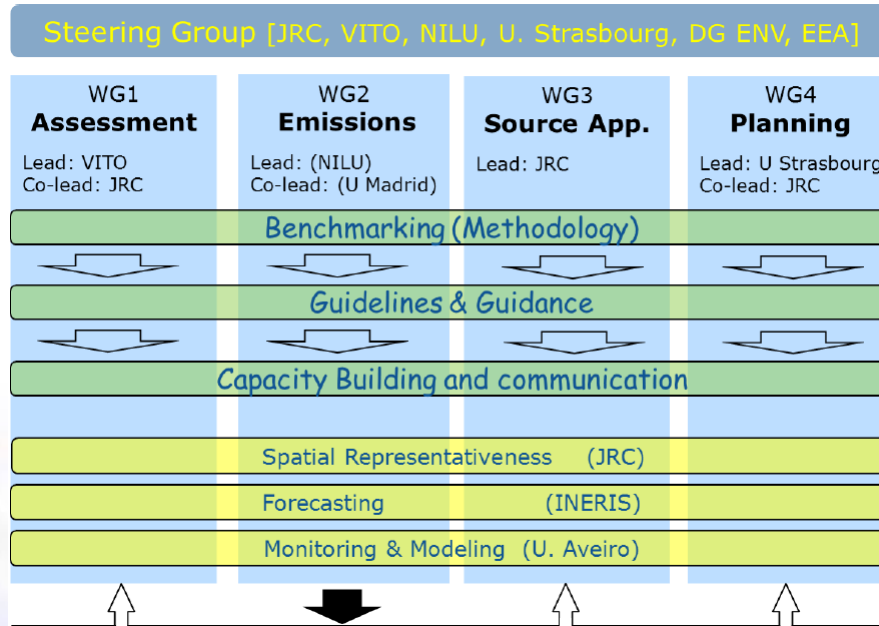
Frédérik Meleux

Institut National de l'Environnement Industriel et des Risques

CCA WG1: Forecasting

Steering Group [JRC, VITO, NILU, U. Strasbourg, DG ENV, EEA]

WG1 **Assessment**	WG2 **Emissions**	WG3 **Source App.**	WG4 **Planning**
Lead: VITO Co-lead: JRC	Lead: (NILU) Co-lead: (U Madrid)	Lead: JRC	Lead: U Strasbourg Co-lead: JRC

Benchmarking (Methodology)

Guidelines & Guidance

Capacity Building and communication

Spatial Representativeness (JRC)

Forecasting (INERIS)

Monitoring & Modeling (U. Aveiro)

14:00-15:00	WG1 – CCA: Forecasting		WG1
14:00-15:15	Introduction & Evaluation of MACC forecast	F. Meleux	
14:15-14:30	CALIOPE forecasts evaluated by DELTA	M.T. Pay	
14:30-14-45	Evaluation of DELTA forecast functionality	J. Stocker	
14:45-15:00	Discussion	All	

- Workplan 2014-2016
  - Q1: Can we use DELTA and its reporting template adapted for forecasting as in the current version? Or do we need to add other indicators/diagrams? If so which ones?
  - Q2: Do we need to add new MQO addressing the detection of threshold exceedances ?

INERIS

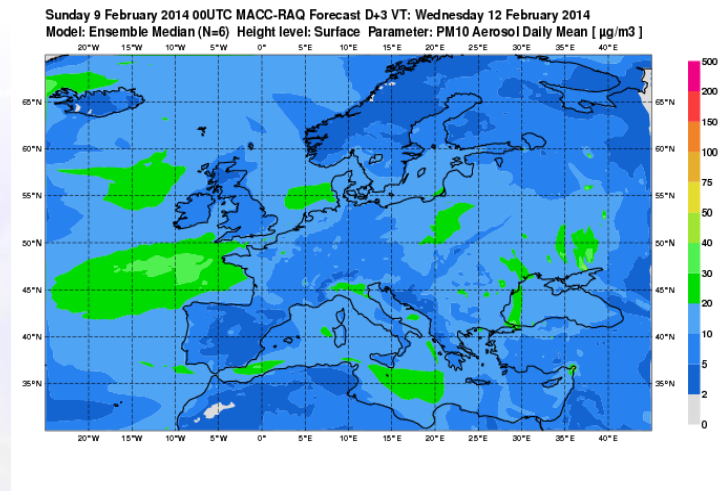maîtriser le risque
pour un développement durable

AQ forecasting system

- European and national air quality platforms produce daily forecasts for the D+0, D+1 and D+2.

  - provide every day information related to the air quality levels

  - targetted pollutants: O3,NO2,PM10,PM2.5

  - In case of pollution episode:
    - Support to policy users
      - Provide recommendations to the public
      - to identify the likely causes
      - to assess population exposure
      - to set-up the efficient measures (short term action plans)

INE-RIS
maîtriser le risque
pour un développement durable

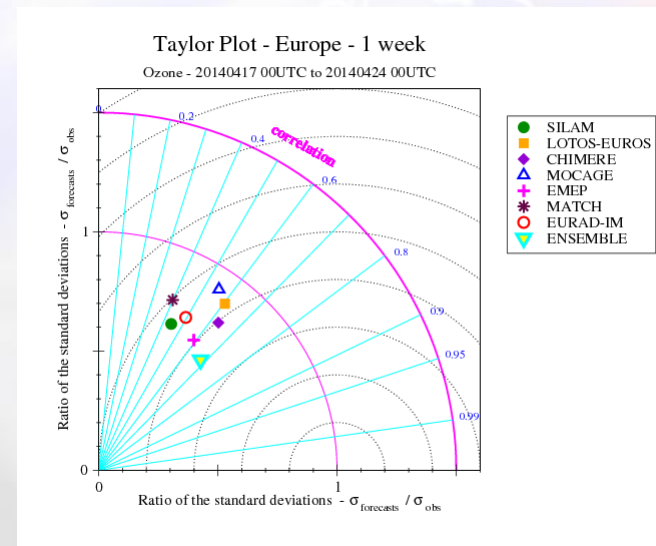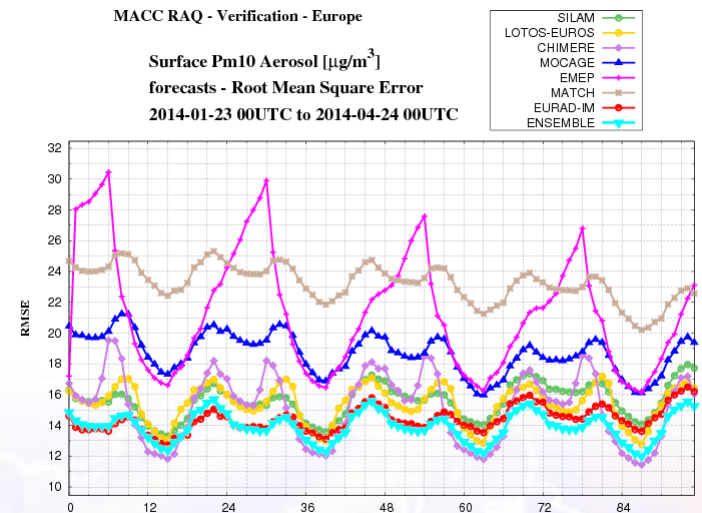# MACCII regional forecasts over Europe

- Based on an ensemble of 7 European models using the same input data (met, emissions, boundary conditions)

	Current geometry	*Assimilation method*
CHIMERE INERIS, CNRS	0.1°, L8, top : 500hpa	Optimal Interpolation
EMEP met.no	0.25°x0.125°, L20, top : 100hpa	Variational 3d-var
EURAD FRIUUK	15km, L23, top : 100hpa	Variational, 3d-var
L-EUROS TNO, KNMI	0.25°x0.125°, L4, top : 3.5km	Ensemble Kalman Filter
MATCH SMHI	0.2°, L40, top : 100hpa	Variational, 3d-var
MOCAGE MF, CERFACS	0.2°, L47, top : 5hpa	Variational, 3d-var
SILAM FMI	0.2°, L46/8, top : 100hpa	Variational, 4d-var

Sunday 9 February 2014 00UTC MACC-RAQ Forecast D+3 VT: Wednesday 12 February 2014
Model: Ensemble Median (N=6) Height level: Surface Parameter: PM10 Aerosol Daily Mean [µg/m3]

INERIS

maîtriser le risque
pour un développement durable

MACC AQ forecast evaluations on a daily basis

- ## MACCII operational evaluation relies on NRT obs data for computing:
  - Bias, RMSE and correlation
  - Calculated for the 96 hours
    - Timeseries of scores from D+0 to D+3 averaged over the last week and the last three months
    - Taylor diagrams

MACC RAQ - Verification - Europe

Surface Pm10 Aerosol [µg/m³] forecasts - Root Mean Square Error 2014-01-23 00UTC to 2014-04-24 00UTC

Legend: SILAM, LOTOS-EUROS, CHIMERE, MOCAGE, EMEP, MATCH, EURAD-IM, ENSEMBLE

Taylor Plot - Europe - 1 week
Ozone - 20140417 00UTC to 20140424 00UTC

Legend: SILAM, LOTOS-EUROS, CHIMERE, MOCAGE, EMEP, MATCH, EURAD-IM, ENSEMBLE

- Evaluations focus on RMSE, bias, correlation averaged over quarters
  - Analyses of the model scores compared to the ones computed one year before and compared to the ensemble performances
  - Teams provide explanations regarding the changes on their model behaviours
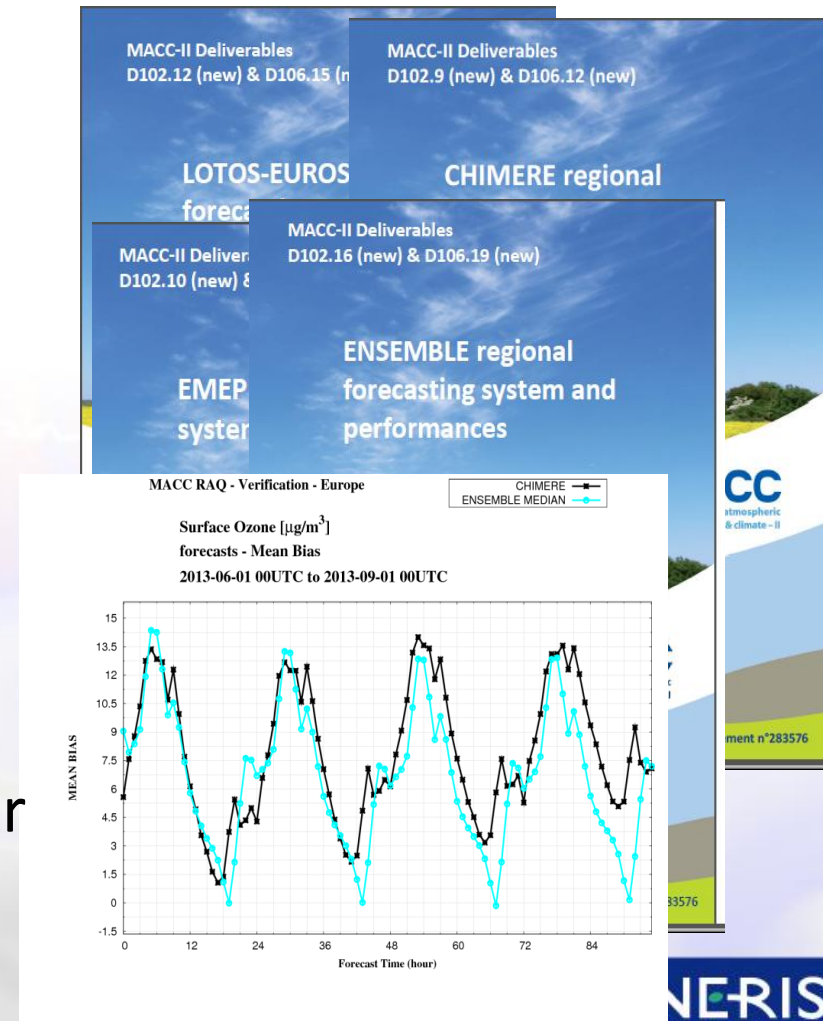
MACCII policy

- There are no scores referring to daily values and to regulatory threshold
  - To provide useful products for national or local applications dealing with air quality management
  - scientific evaluation of the model forecasts and of the ensemble multi-model.
- This approach might change with the operational set-up of the Copernicus atmospheric service (in 2015)
  - ENSEMBLE can test the FAIRMODE procedures for forecasting evaluation

INERIS

maîtriser le risque
pour un développement durable

FAIRMODE FORECAST TARGET

• A target has been designed in the previous FAIRMODE phase for forecasting applications :

$$\text{target} = \frac{\sqrt{\dfrac{1}{N}\sum_{1}^{N}(M_i - O_i)^2}}{\sqrt{\dfrac{1}{N}\sum_{1}^{N}(O_{i-1} - O_i)^2}}$$

Stating that the worst acceptable model is the persistent model, so at a given station the forecast (D+0) provides the observation of the eve (D-1).

FAIRMODE FORECAST TARGET

- What are the policy objectives for using AQ forecasts:
  - Predict the development of a pollution episode
    - How able are the model to reproduce the transition between non polluted regime to polluted regime?
    - How able the model are to detect / anticipate threshold exceedances ?
      - The triggering of measures is based on threshold values
      - Use of the group function in Delta tool to select an adapted range of station for such evaluation (including observation uncertainty)
    - How stable are the forecast scores from D+0 to D+n (usually n = 2 or 3)?
      - Policy measures are more efficient when they are taken earliest – so the goal is really to provide confident forecasts at least at D+1 (D+2 would be even better)

INERIS

maîtriser le risque |
pour un développement durable |

FAIRMODE FORECAST TARGET

- Focus on daily mean and daily maximum

- Various assessments should considered :
  - a global approach (whole domain and whole year)
  - sub-geographical groups (in linkages to the heterogeneity of the geographical distribution of the emission sources)
  - sub-temporal groups (seasonal evaluations…)
    - For instance the model ability to reproduce PM10 levels may vary from one season to another

INERIS
maîtriser le risque
pour un développement durable

French evaluation for episodes

- The evaluation of the French PREV'AIR system relies on the classical skill scores (RMSE…)
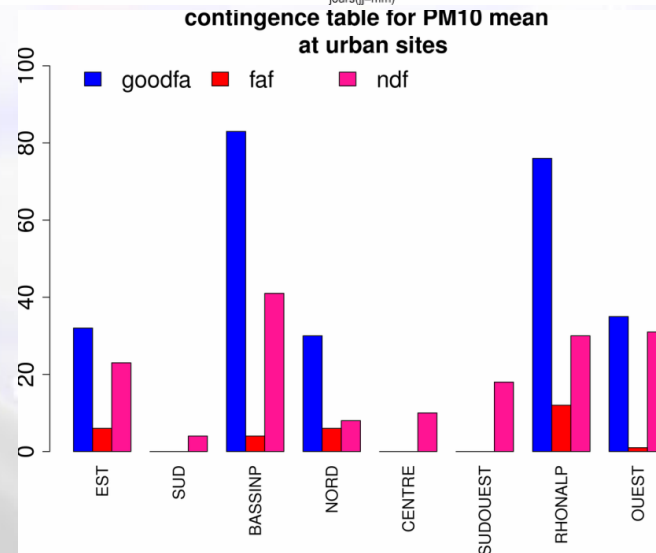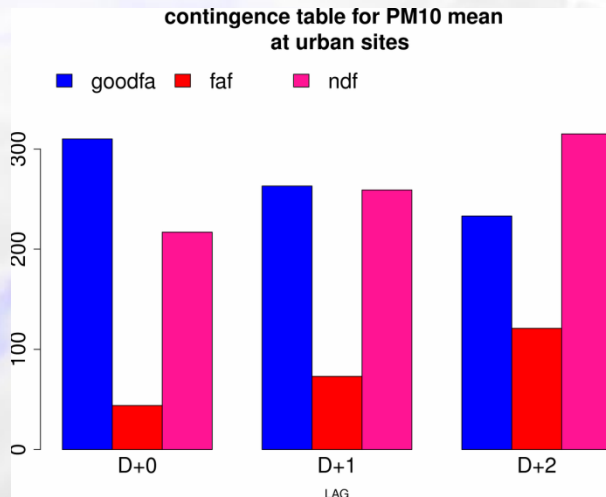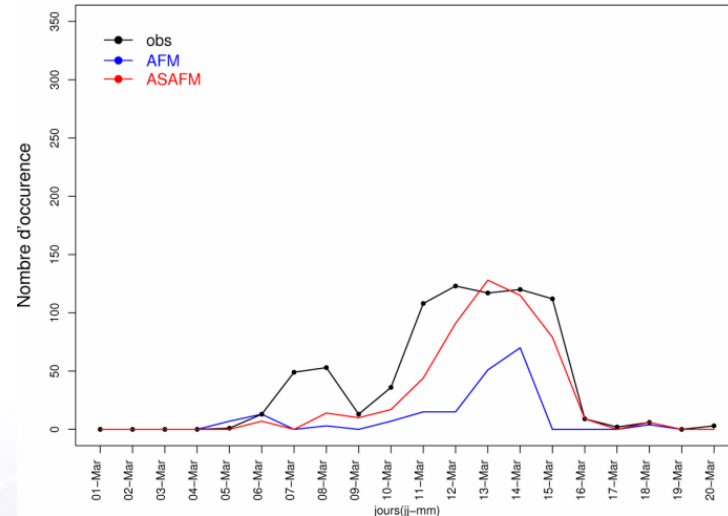- additional assessment are produced focusing on episodes
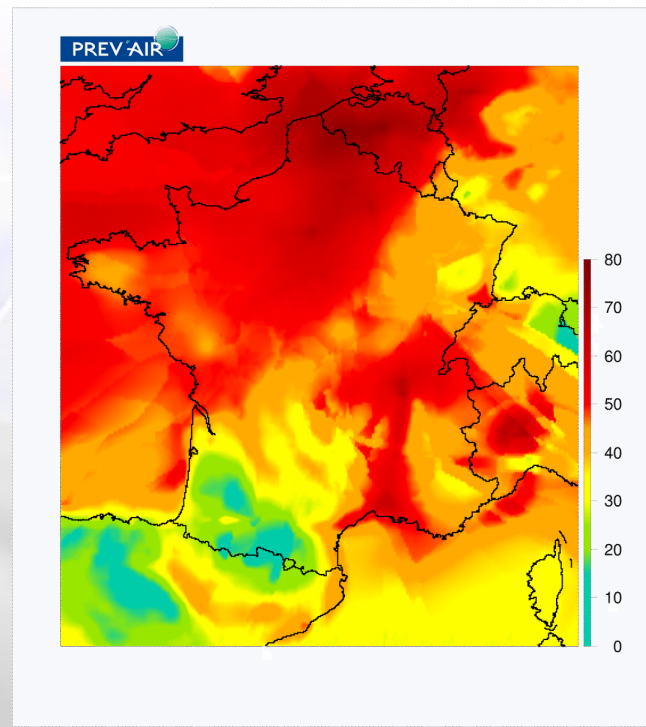
PM10 daily mean concentrations ($\mu g/m^3$); March 2014

French evaluation for episodes

- Timeseries of the number of threshold exceedances
- Contingency tables
  - D+0 to D+2
  - Regional representation

PM10 dépassements du seuil de 50 µg/m3 D+0 Année: 2014

contingence table for PM10 mean at urban sites

contingence table for PM10 mean at urban sites

INERIS

maîtriser le risque
pour un développement durable

Conclusions

- Forecast model should be assessed using classical scores and target defined in the delta tool -> but the best models doesn't mean ability to detect threshold exceedances

- Using forecast for AQ managements, additional evaluations should be requested
  - Is there possible adaptation of the Delta target for forecasting threshold exceedances ?
  - How can we connect this forecasting target to indicators dedicated to threshold detection ability (contingency table, odds ratio skill scores …) ?

- Do we need strict evaluation of the co-located threshold exceedances in obs & mod ?