



# Space Analog's Searching to Improve Deterministic Forecasting Using Analog Ensemble Method Over Morocco

**Badreddine Alaoui [1]** , Driss Bari [1] and Yamna Ghabbar [2]

**1. CNRM/SMN, National Weather Department (DGM), Casablanca, Morocco**

**2. Department of Mathematics, Computer sciences and Geomatics, Hassania School for Public Works, Casablanca P.O. Box 8108, Morocco**

EWGLAM/SRNWP-2024, CHMU, Czech, Prague 30 September-03 October 2024

# Surface Weather Parameters & Ensemble Forecasting

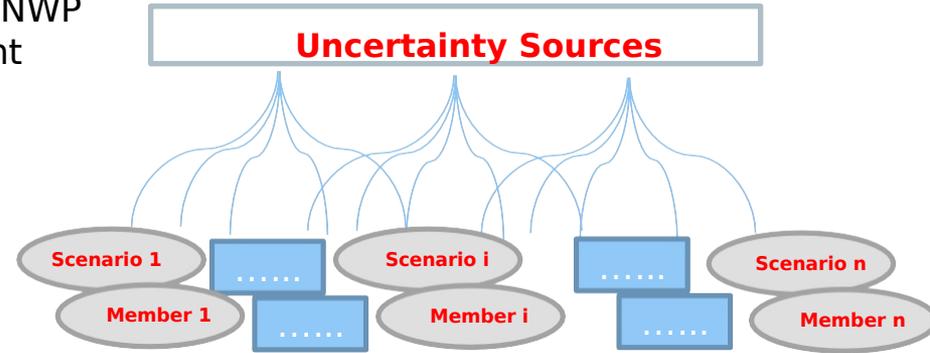
- High socio-economic impact + Elementary entry of NWP
- Wildfires, Wind Gusts, Heat /Cold waves and Drought



- **NWP models sources of uncertainty (IC, Model, Parametrization, LBCs)**
- **Users are seeking for probabilistic products !**

## The Analog Ensemble Method?

- EPS
- AI
- Low cost



- Perturbating IC, DA
- Multi-models
- Multi-Parametrizations
- Hybride techniques
- **Analogue Ensemble Method**

?

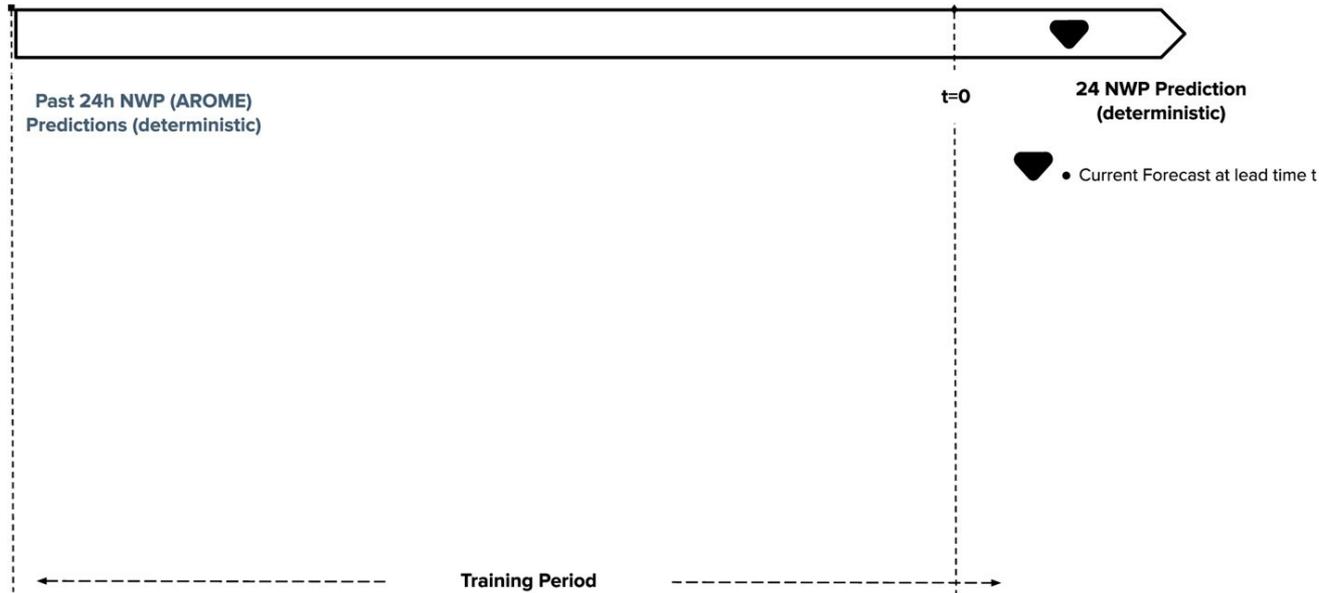
Ensemble Weather Prediction systems are computationally expensive !

(Delle Monache et al, 2013)

# Analog Ensemble Method

## [Step 0] : For a lead time $t$ (24h) and a station $s$ :

- An archive of past forecasts of AROME model at lead time  $t$ .
- An archive of past observations at station  $s$  and lead time  $t$ .
- Target forecast at lead time  $t$  (24h).

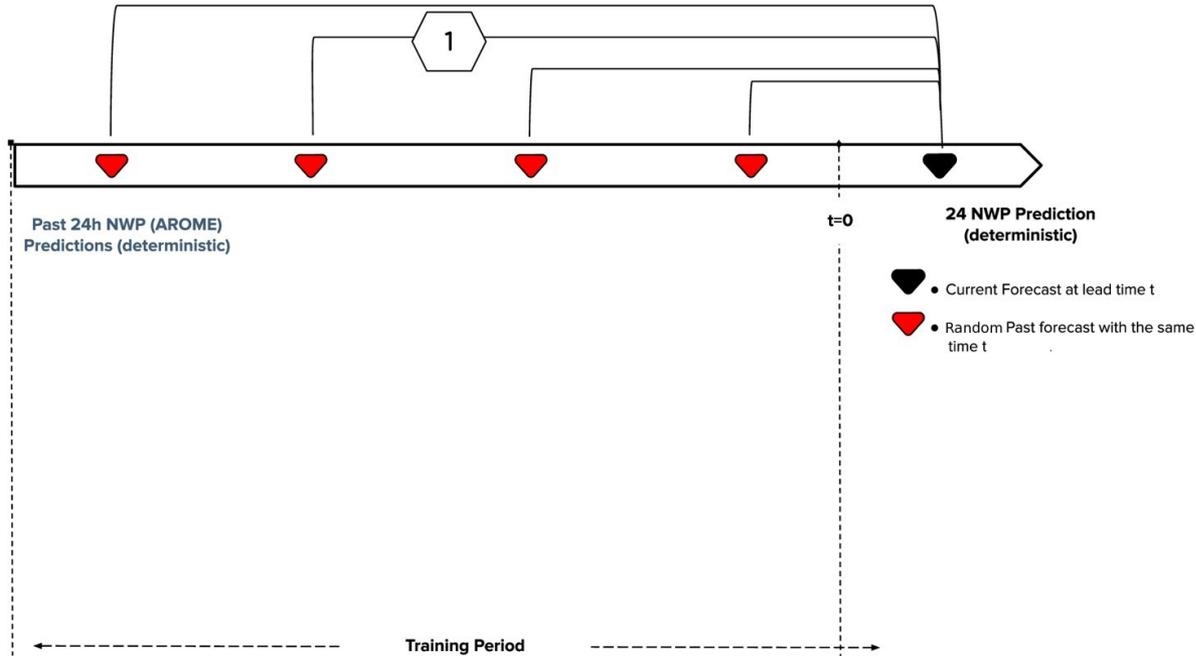


Goal : Build ensemble forecasts at leadtime  $t$  for station  $s$  using the past forecasts and their associated observations.

# Analog Ensemble Method

## [Step 1] : Searching for all the past forecasts similar to the actual forecast at time $t$

- Finding all the past forecasts (analogues) similar to the actual forecast (target) at lead time  $t$  (24h) .
- Distance between past and actual forecast is evaluated using a similarity

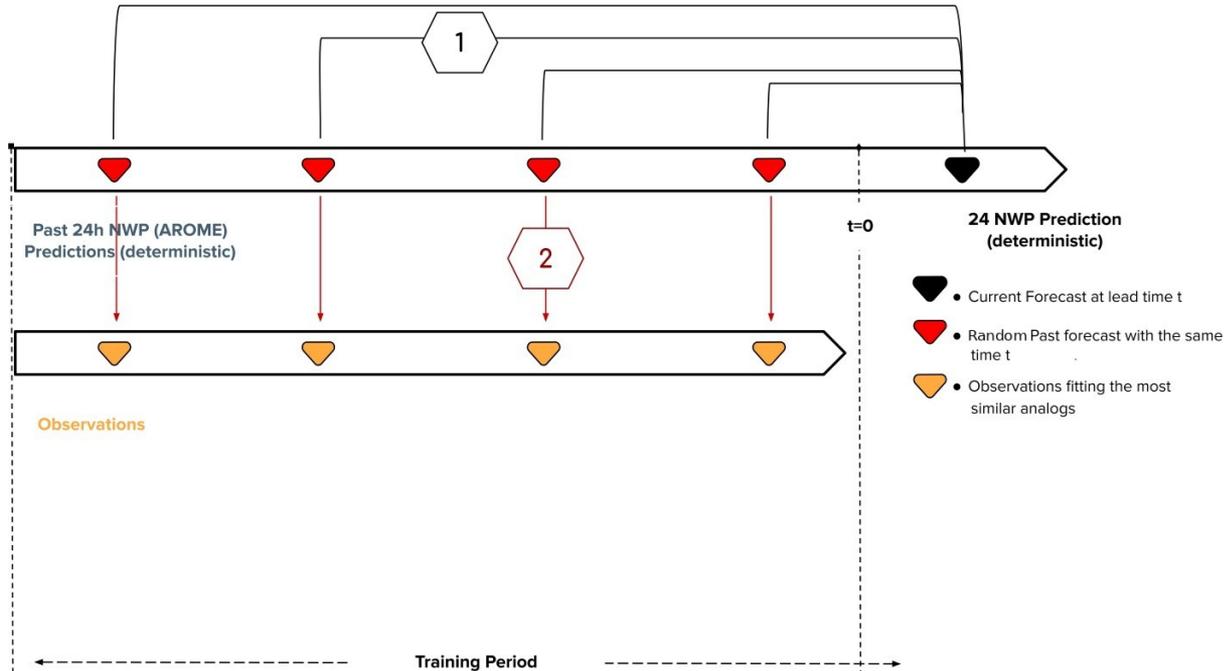


Objectif : Finding all the past forecasts similar to the actual one at time  $t$  and transferring their respective datetimes.

# Analog Ensemble Method

## [Step 2] : Finding the past observations associated with the best analogs

- Once the best analogues at leadtime  $t$  (24h) are found  $s$ .
- Start searching for the past observations at station  $t$  and lead time  $t$  associated with the best analogues datetimes.

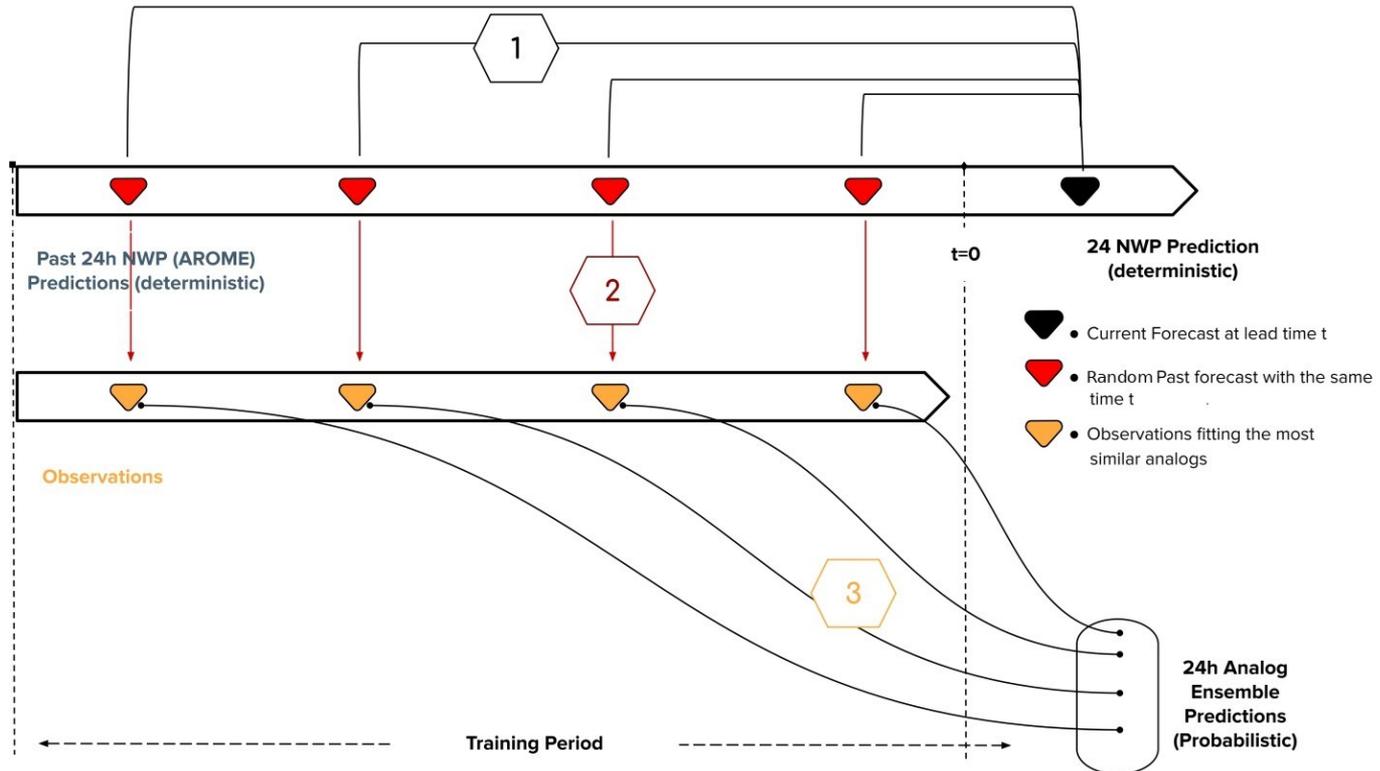


Goal : Finding observations at station  $s$  and leadtime  $t$  matching the datetimes of the best analogues

# Analog Ensemble Method

## [Step 3] : Constructing AnEn Ensemble using observations of the best analogues

- Observations associated with the best analogs are the members of the AnEn Ensemble.
- Ensemble scores may be evaluated upon those members (ensemble mean, spread, ....)



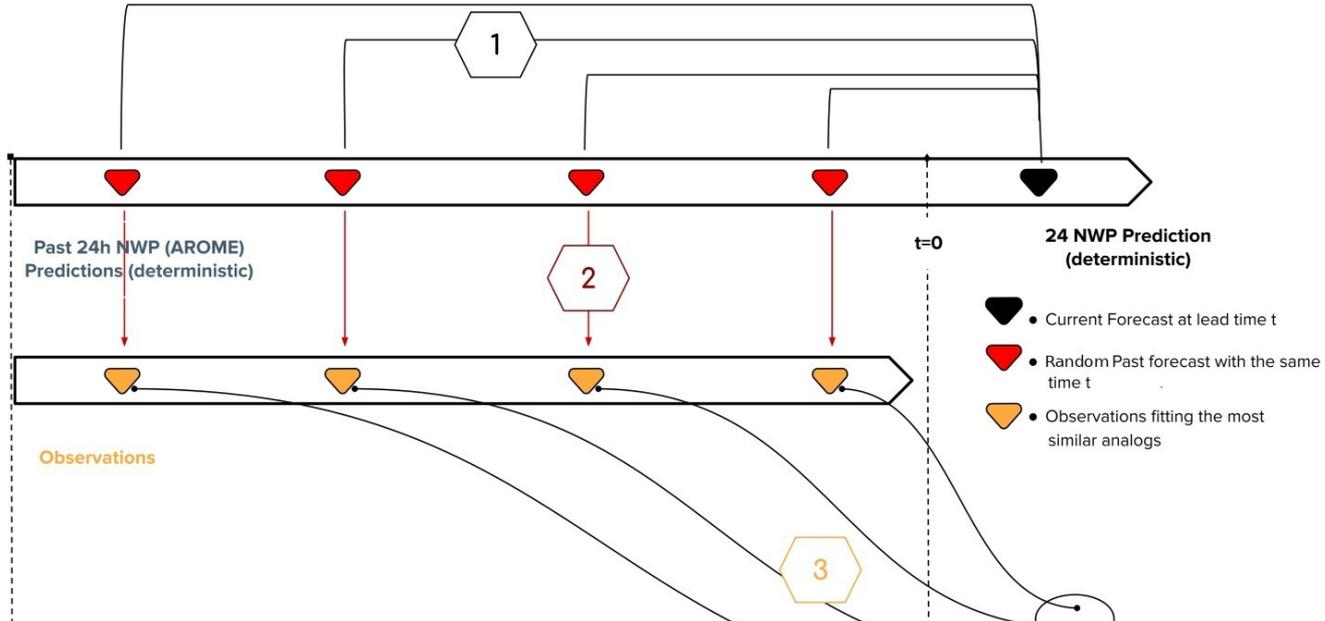
# Analog Ensemble Method

**[Step 0]** : Choosing a station  $s$  and a lead time  $t$  + Access to archives of past forecasts and observations

**[Step 1]** : Searching for all the past forecasts similar to the actual forecast at time  $t$

**[Step 2]** : Finding the past observations associated with the best analogs

**[Step 3]** : Constructing AnEn Ensemble using observations of the best analogues



?

How similarity between past and actual forecasts is evaluated?

# Similarity Metric



**Similarity Metric (Delle Monach et al 2013)**

$$\| \mathbf{F}_t, \mathbf{A}_{t'} \| = \sum_{i=1}^{N_V} \frac{w_i}{\sigma_i} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}$$

# Similarity Metric



Predictors number

Time neighboring window

**Similarity Metric**

Analog Forecast within the time neighboring window

$$\|F_t, A_{t'}\| = \sum_{i=1}^{N_v} \frac{w_i}{\sigma_i} \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}$$

Predictor weight

Actual Forecast within the time neighboring window

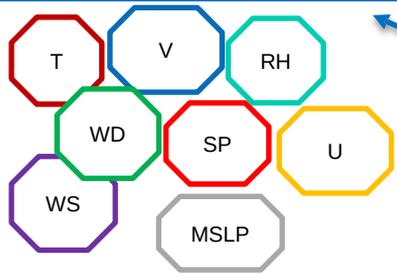
Predictor standard deviation

# Our Contributions



## 1 Predictors Set

8 surface weather parameters (in place of 3)



## Similarity Metric

$$\|F_t, A_{t'}\| = \sum_{i=1}^{N_V} w_i \sqrt{\sum_{j=-\tilde{t}}^{\tilde{t}} (F_{i,t+j} - A_{i,t'+j})^2}$$

## 3 Analogs Search Space

Spatial Extension of analogs search space

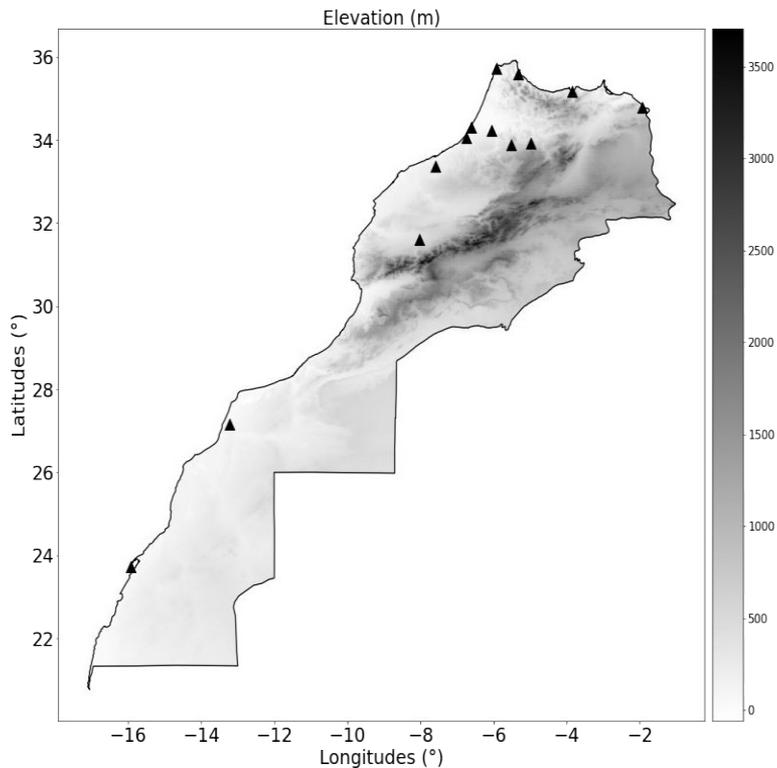
## 2 Predictor's weighting strategies

- Same weight for all predictors in literature / Expensive weighting strategies **Brute Force (Junk et al 2015)**
- New predictors weighting strategies based on AI
- Feature importance
- Linear Regression (**Hope 2020**), Random Forest (**Brieman 2001**) and Xgboost (**Chen and Guestrin 2016**)



Goal: Improving AnEn performance in forecasting Surface weather parameters using AI

# Study Domain & Experimental Design



Locations	13 airports of Morocco
Predictors	8 (T2m, RH2m, MSLP, SURFP, ZW10m, MW10m, WS10m, WD10m)
NWP Model	AROME, 2,5km, CYC41, 90 NV
Data total period	2016-2020
Training period	2016-2018
Test period	2019
Members number	25
Time window	1 hour
Weighting strategies	DM13, Xgboost, Random Forest and Linear Regression



Alaoui, B., Bari, D. & Ghabbar, Y. Surface Weather Parameters Forecasting Using Analog Ensemble Method over the Main Airports of Morocco. *J Meteorol Res* 36, 866–881 (2022). DOI: 10.1007/s13351-022-2019-0

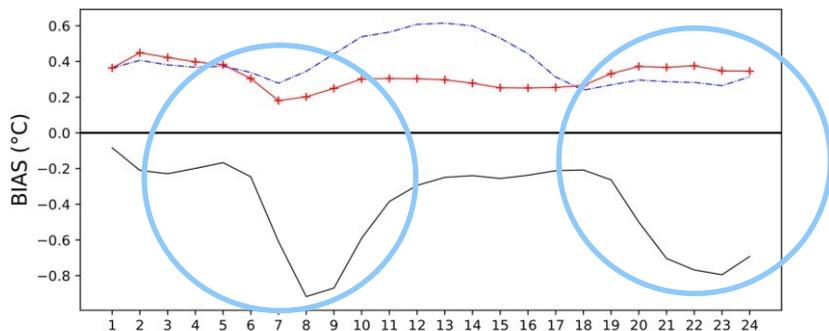
# Performances by lead times

**R: DM13 AnEn Basic;**

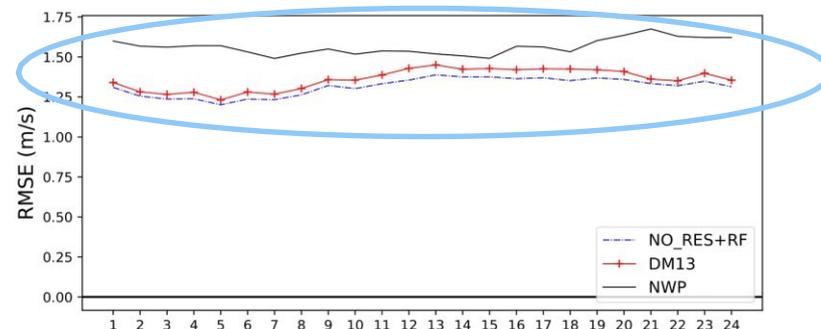
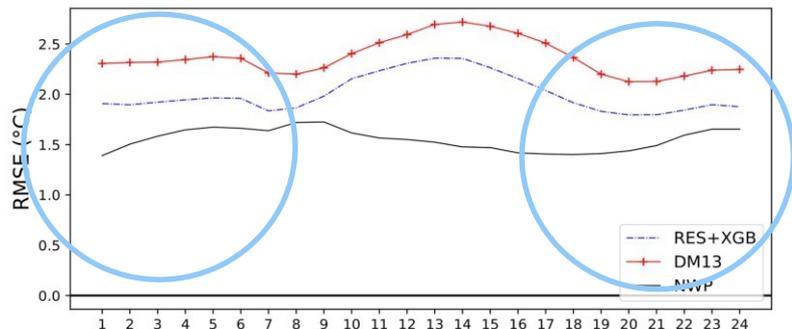
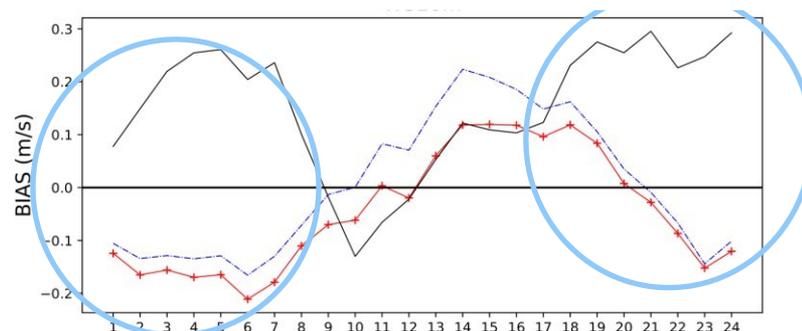
**B: AnEn + AI;**

**N: AROME;**

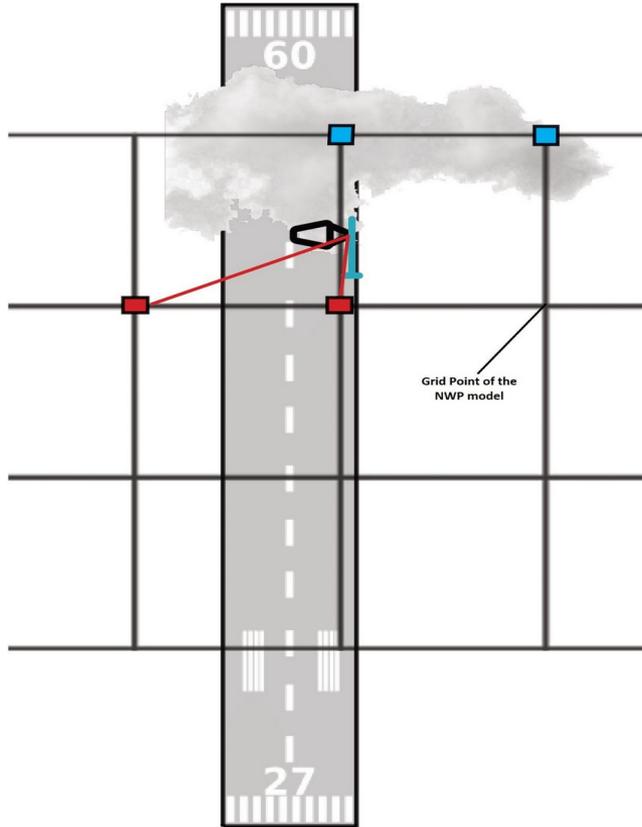
**T2m : Temperature at 2m**



**WS10m: Wind speed at 10m**



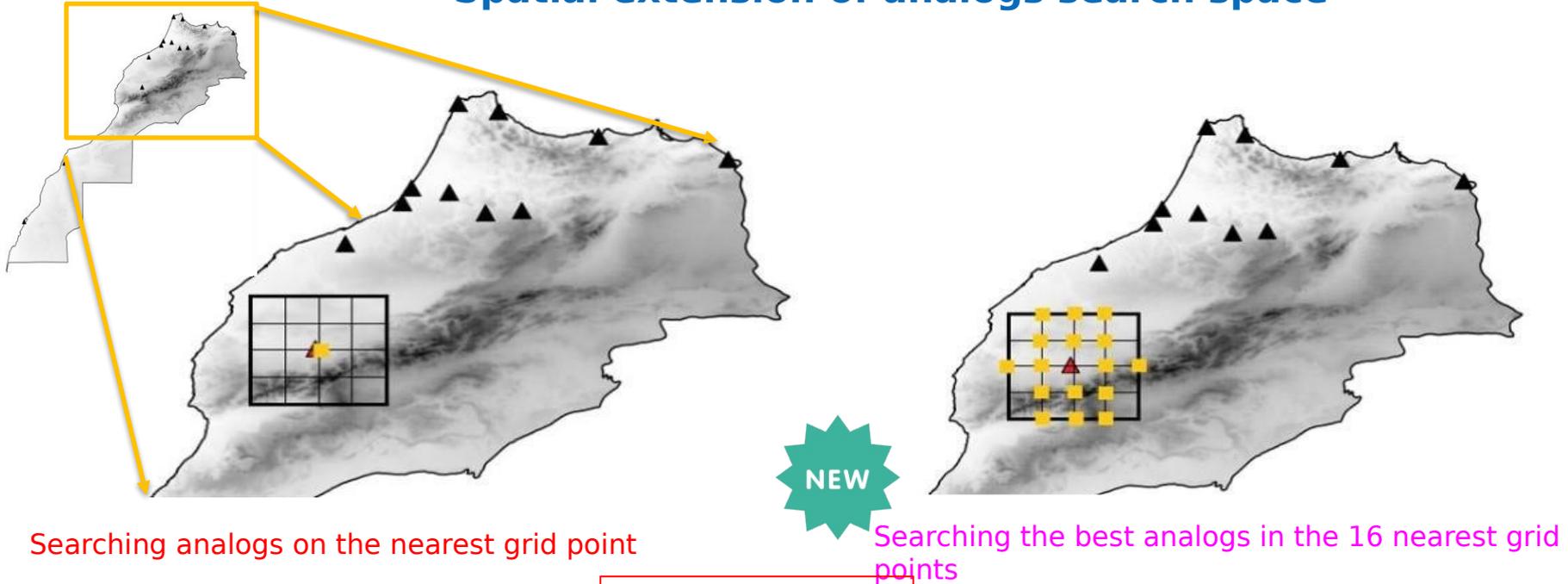
- **T2m : Best bias during night for AnEn + AI and better RMSE than DM13 AnEn Basic**
- **WS10m : Best bias during night for AnEn+AI, best RMSE for all lead times**
- **During the day lead times performances remain close.**



## ● Why spatial extension?

- AnEn method: “frozen system” = No changes on NWP model !
- Cycles, Archiving problem of same NWP version!
- For example if we are interested on low visibility phenomenon/ fog our transmissometer is seeing fog.
- **But its nearest grid points are not !**
- **Farest grid points are catching Fog !**

## Spatial extension of analogs search space



$$\min_{gp=1\dots 16} \|F_t^{gp}, A_t^{gp}\|$$

Where  $gp$  are the nearest grid points,  $F_t^{gp}$  the current forecast at  $gp$  and  $A_t^{gp}$  is the analog forecast at  $gp$

Alaoui, B., Bari, D. & Ghabbar, Y. Space Analog's Searching to Improve Deterministic Forecasting Using Analog Ensemble Method Over Morocco, Recent Advancements from Aquifers to Skies in Hydrogeology, Geocology, and Atmospheric Sciences 2024. DOI : 10.1007/978-3-031-47079-0

# Performances by lead times

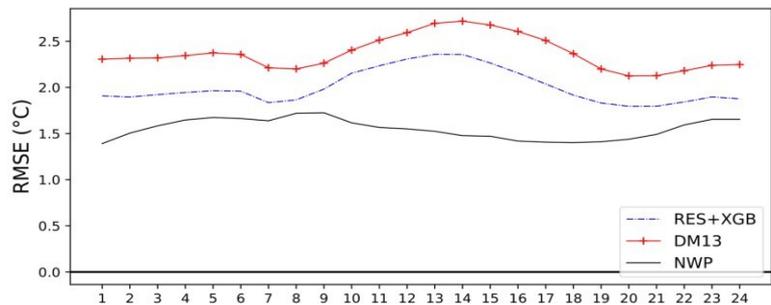
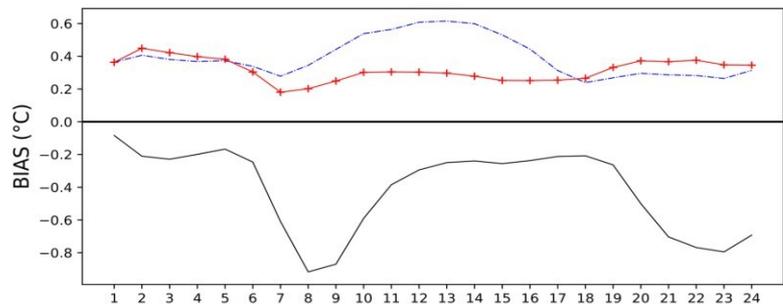
T2m : Temperature at 2m

R: DM13 AnEn Basic;

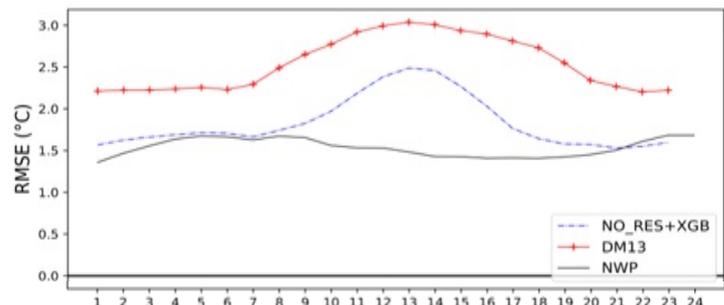
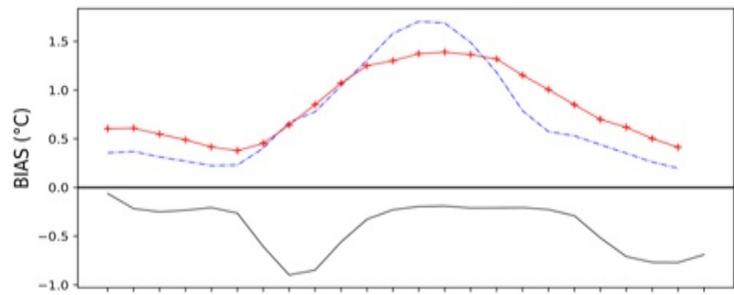
B: AnEn + AI;

N: AROME;

1 grid point



16 grid points



- T2m (16GP) : Best Bias always during the night with noticeable improvement of RMSE for 16 grid points version .

# Performances by lead times

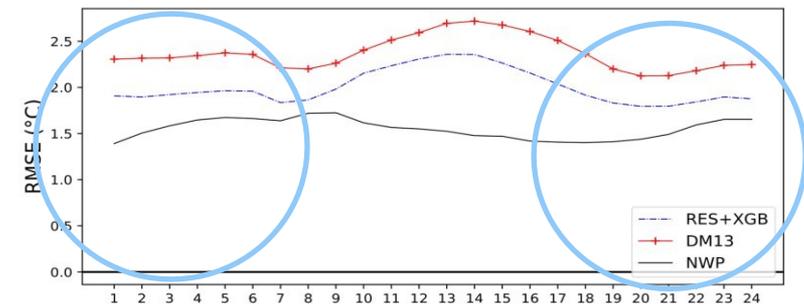
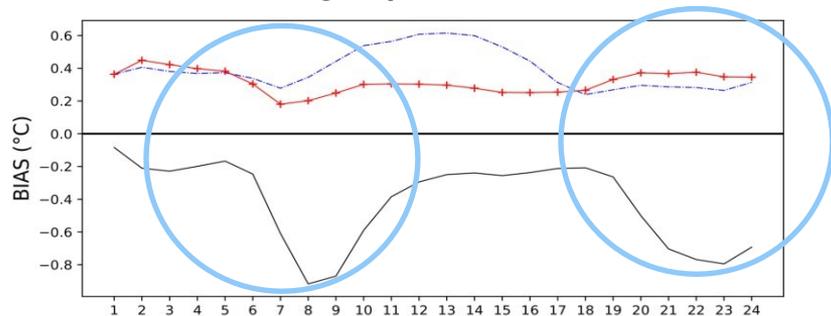
T2m : Temperature at 2m

R: DM13 AnEn Basic;

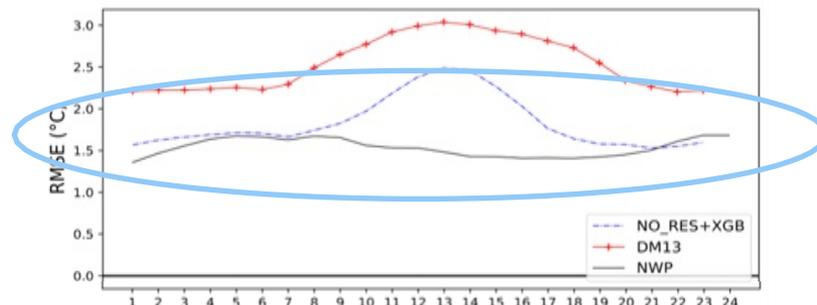
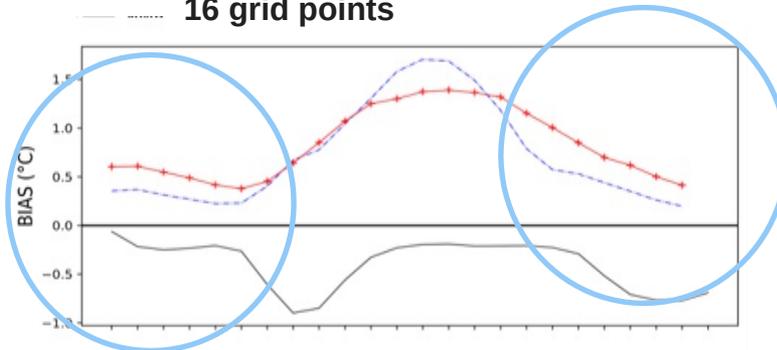
B: AnEn + AI;

N: AROME;

1 grid point



16 grid points



- T2m (16GP) : Best Bias always during the night with noticeable improvement of RMSE for 16 grid points version .

# Performances by lead times

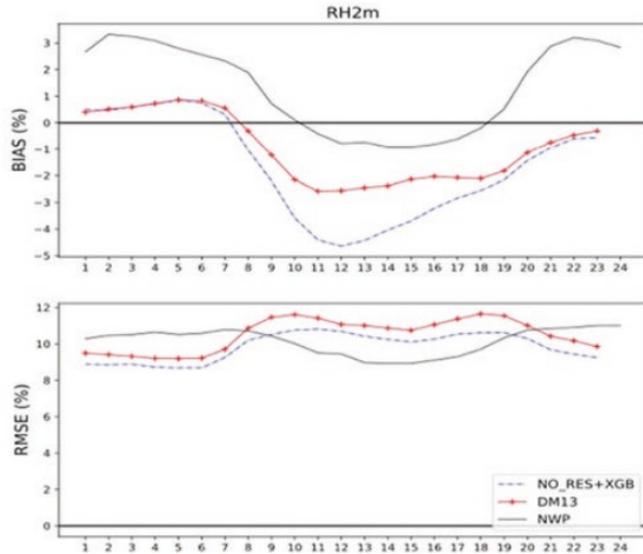
RH2m : Relative Humidity at 2m

R: DM13 AnEn Basic;

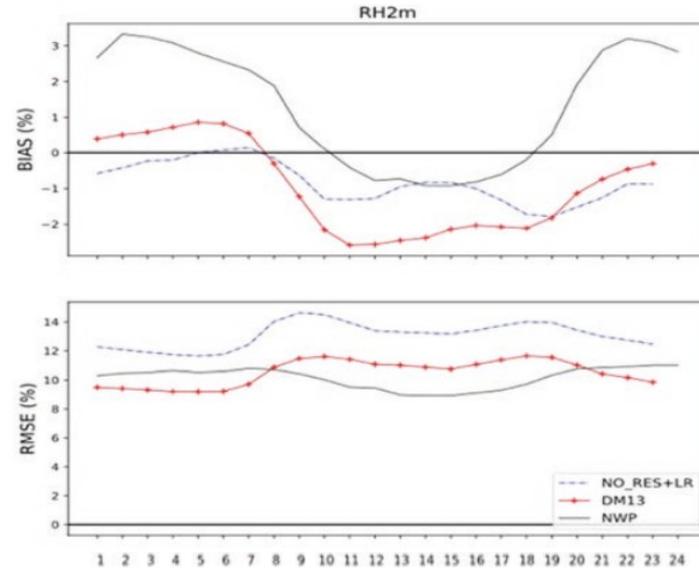
B: AnEn + AI;

N: AROME;

1 grid point



16 grid points



- RH2m (16GP) : Best Bias for most lead times worsening of RMSE compared to single grid point version .

# Performances by lead times

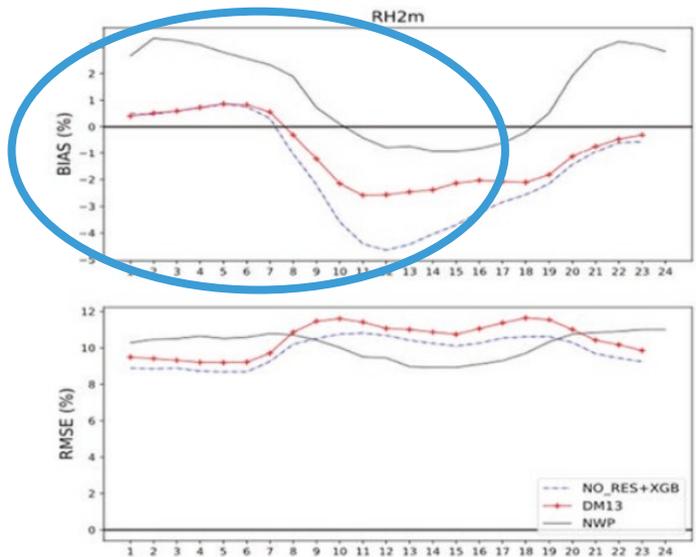
RH2m : Relative Humidity at 2m

R: DM13 AnEn Basic;

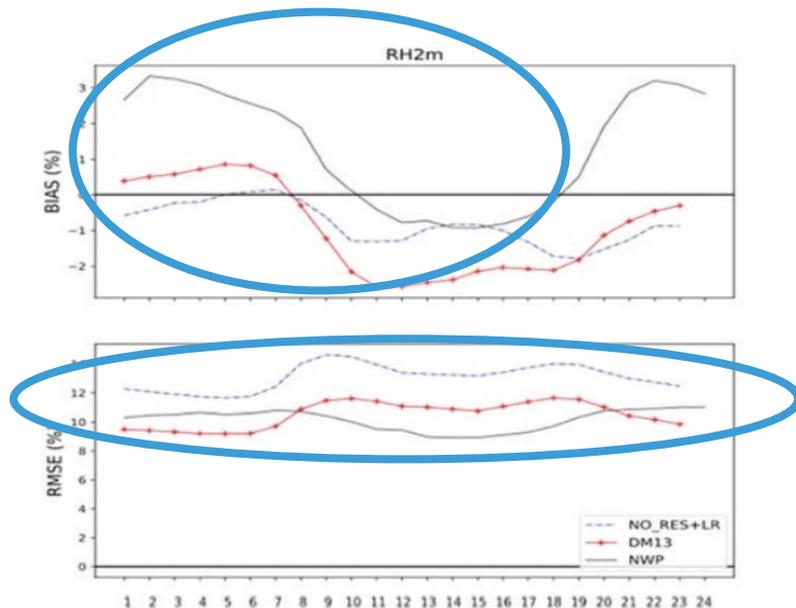
B: AnEn + AI;

N: AROME;

1 grid point



16 grid points



- RH2m (16GP) : Best Bias for most lead times worsening of RMSE compared to single grid point version .

# Performances by lead times

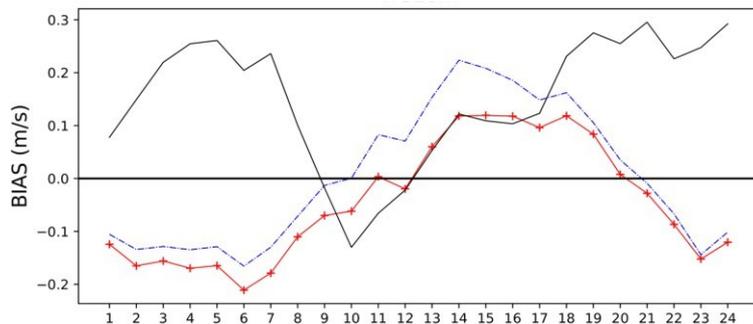
WS10m: Wind Speed at 10m

R: DM13 AnEn Basic;

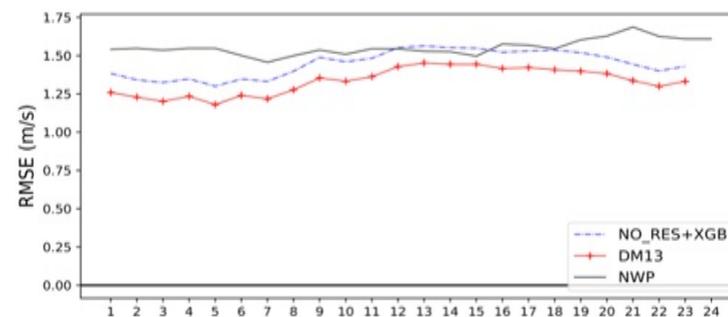
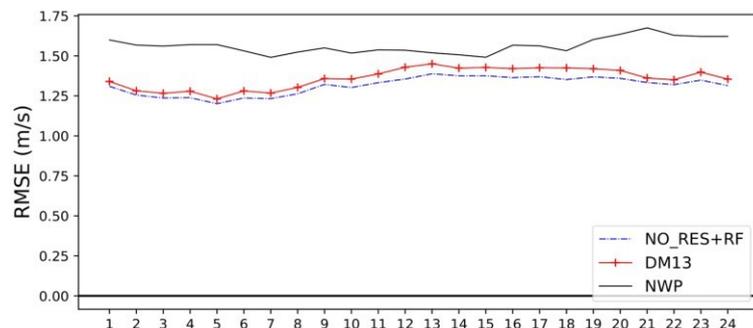
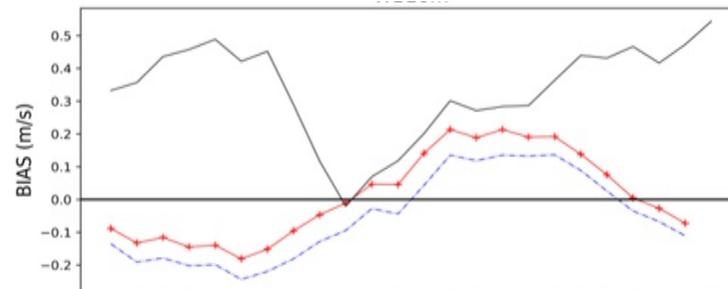
B: AnEn + AI;

N: AROME;

1 grid point



16 grid points



- WS10m (16GP): Better Bias during night lead times (close to 0m/s)
- WS10m (16GP) : Slight worsening on RMSE overall

# Performances by lead times

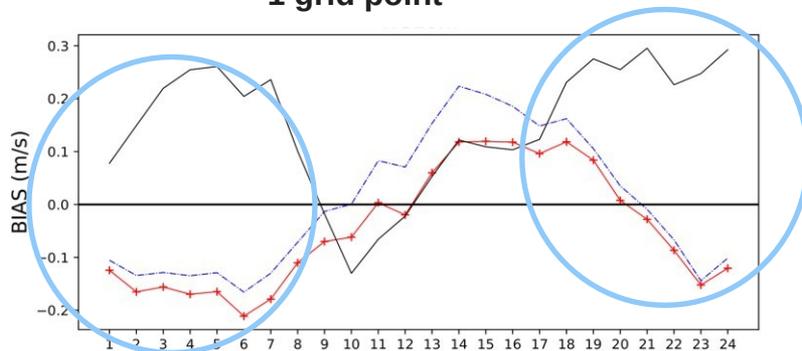
WS10m: Wind Speed at 10m

R: DM13 AnEn Basic;

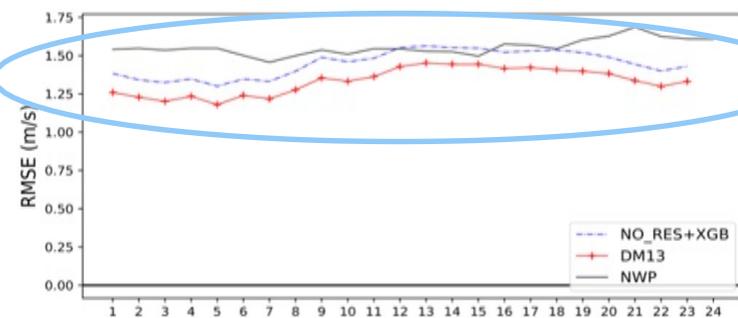
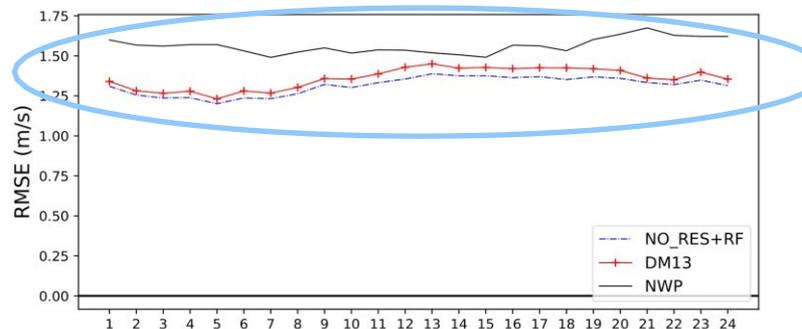
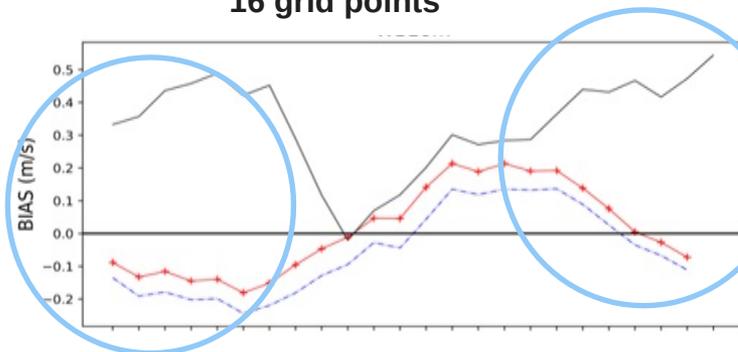
B: AnEn + AI;

N: AROME;

1 grid point

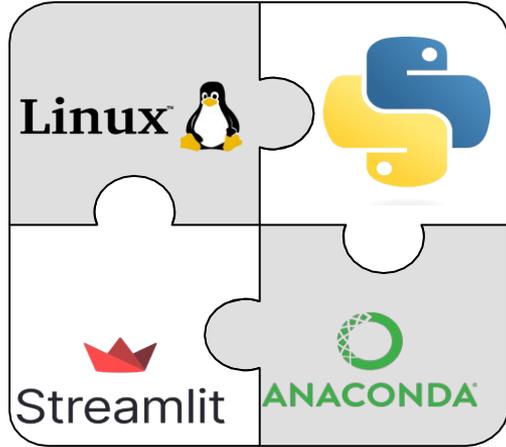


16 grid points



- WS10m (16GP): Better Bias during night lead times (close to 0m/s)
- WS10m (16GP) : Slight worsening on RMSE

## Toward an operational AnEn Ensemble System



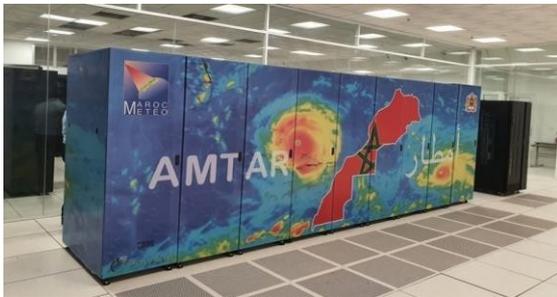
25 membres

Up to 72h

15 airports

2min on 1 node

Operational since July/2023



# Toward an operational AnEn Ensemble System

Drag and drop file here

Limit 200MB per file

Browse files



test\_2.csv  
1.6MB



## Navigation

Select what you want to display:

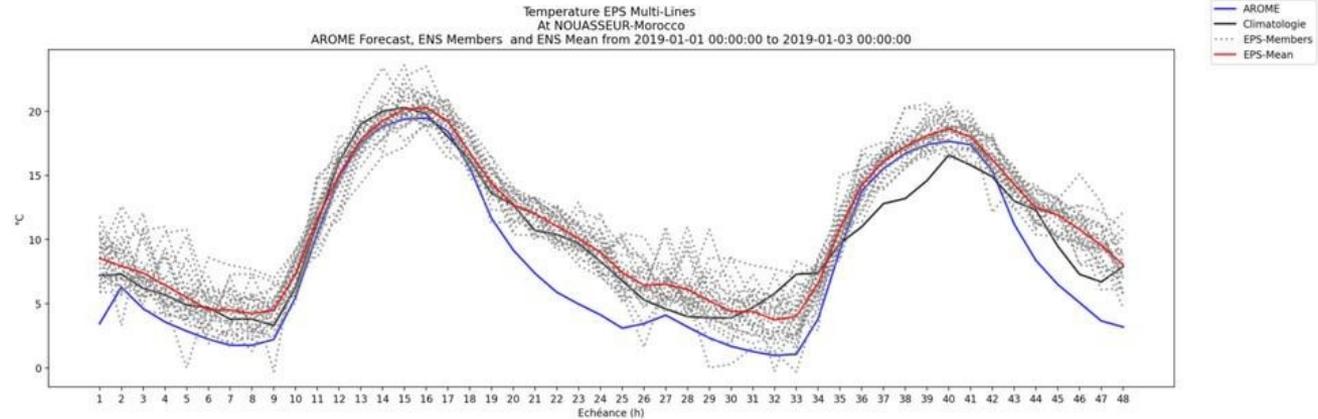
- Home
- EPS Summary
- Data Header
- EPS MeteoGram
- EPS Multi-Lines
- AnEn Spread & Mean
- Scenarios Clusters
- Probabilities Barplots
- Probabilities Maps
- Probabilities Maps By Steps

Which Leadtime?



You have chosen to see forecasts from 2019-01-01 00:00:00 to 2019-01-03 00:00:00

Available data is from 2019-01-01 00:00:00 to 2019-01-02 23:00:00



Download Figure



# Toward an operational AnEn Ensemble System



Drag and drop file here

Limit 200MB per file

Browse files



test\_2.csv  
1.6MB



## Navigation

Select what you want to display:

- Home
- EPS Summary
- Data Header
- EPS MeteoGram
- EPS Multi-Lines
- AnEn Spread & Mean
- Scenarios Clusters
- Probabilities Barplots
- Probabilities Maps
- Probabilities Maps By Steps

Choose City

KENITRA x

MEKNES x

LAAYOUNE x

NOUASSEUR x

RABAT\_SALE x

TETOUAN x

TANGER\_AERO x

OUJDA x

MARRAKECH x

SIDI\_SLIMANE x

Choose Parameter

temp x

Choose Day

2019/01/01

Choose Leadtime

00:00

You have chosen to see forecasts from 2019-01-01 00:00:00 to 2019-01-04 00:00:00

Which Leadtime?

6

48

72

# Toward an operational AnEn Ensemble System

Drag and drop file here  
Limit 200MB per file

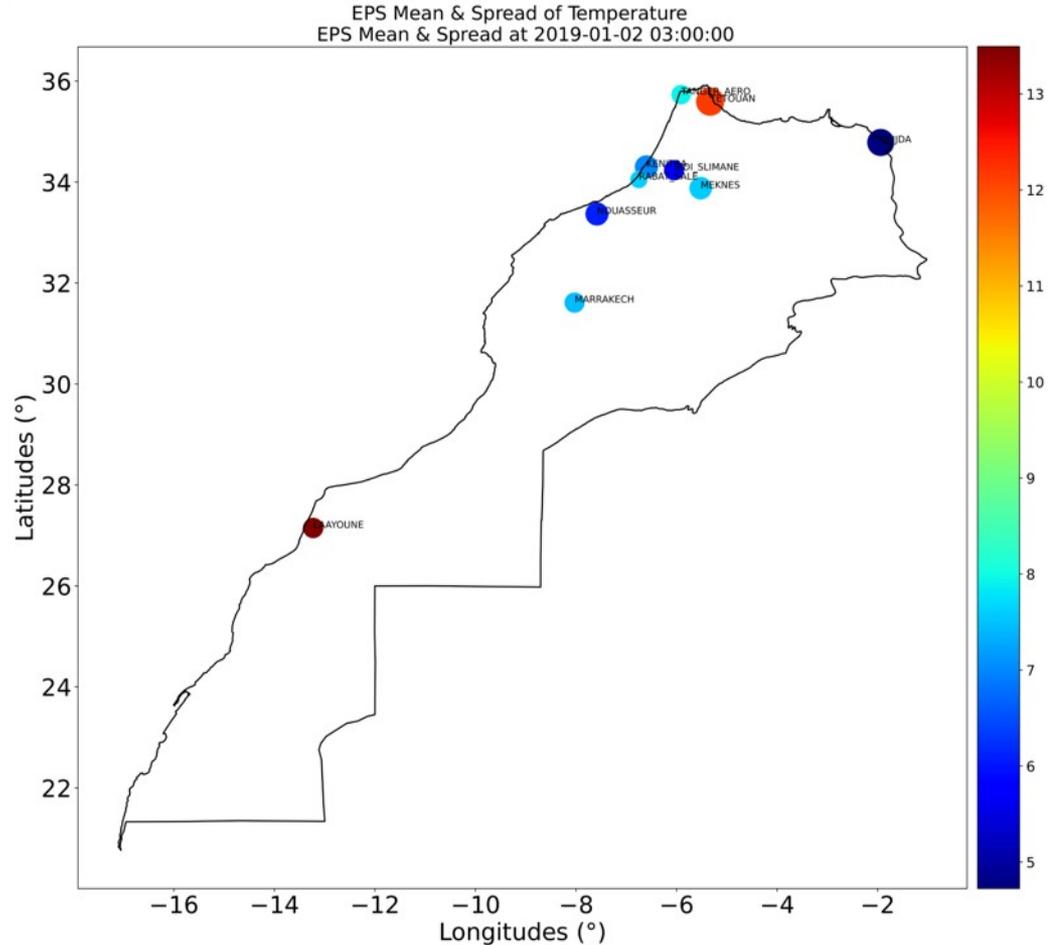
Browse files

test\_2.csv 1.6MB

### Navigation

Select what you want to display:

- Home
- EPS Summary
- Data Header
- EPS MeteoGram
- EPS Multi-Lines
- AnEn Spread & Mean
- Scenarios Clusters
- Probabilities Barplots
- Probabilities Maps
- Probabilities Maps By Steps



# Toward an operational AnEn Ensemble System

Drag and drop file here  
Limit 200MB per file

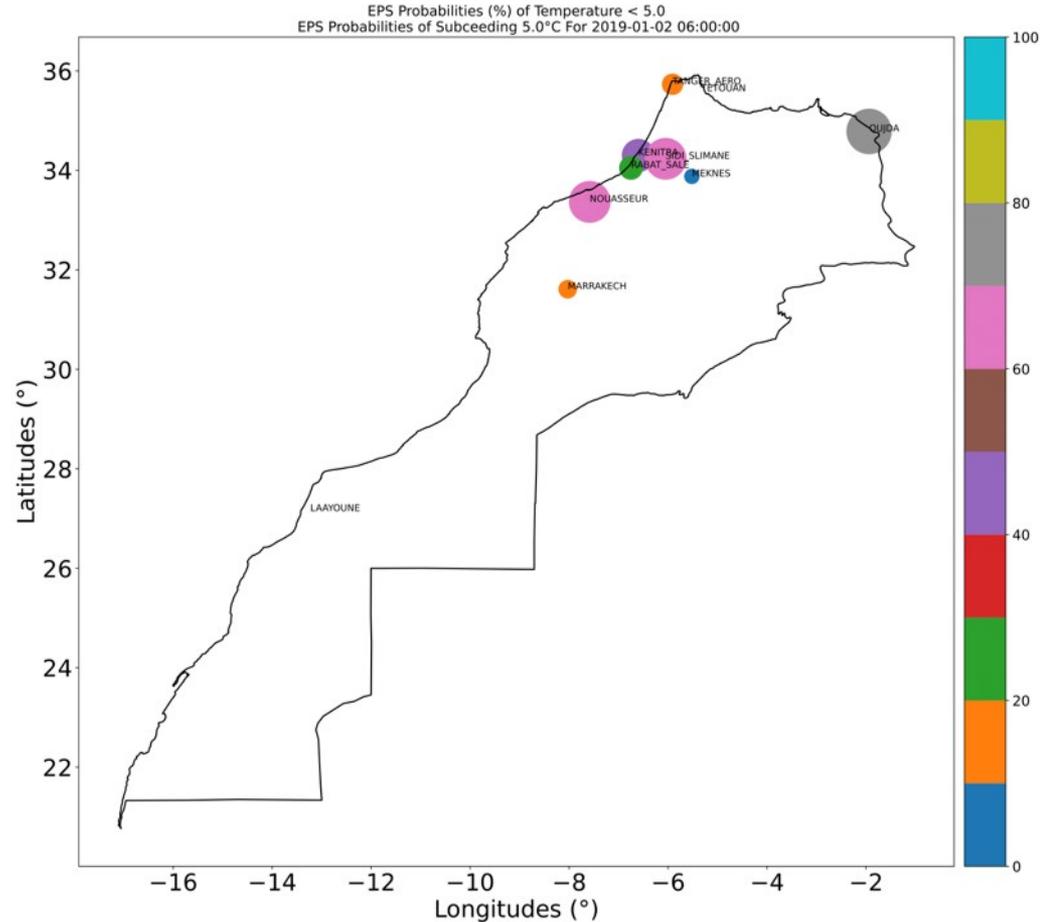
Browse files

test\_2.csv 1.6MB

### Navigation

Select what you want to display:

- Home
- EPS Summary
- Data Header
- EPS MeteoGram
- EPS Multi-Lines
- AnEn Spread & Mean
- Scenarios Clusters
- Probabilities Barplots
- Probabilities Maps
- Probabilities Maps By Steps



## Conclusion

Limitations of AnEn DM13	Our contributions	Gains
Limited set of predictors	Enlarging the set of predictors to cover 8 surface weather parameters	Targeting additional weather phenomena
Limited analogs search space	Spatial extension of analogs search space	Improvement of performances for some parameters (T2m, RH2m)
Constant predictors weight / Expensive weighting strategies	New low cost predictors weighting strategies based on AI	<ul style="list-style-type: none"><li>- Gain of 50% in Biases and 30% in RMSE for most airports</li><li>- Low cost</li><li>- Physics correlations friendly</li></ul>

## Perspectives

Our contributions	Future works
Enlarging the set of predictors to cover 8 surface weather parameters	1)Applying the analog ensemble method to ERA5 Reanalyses: <ul style="list-style-type: none"><li>• Deep Archive of data (up to 1940)</li><li>• Access to altitude weather parameters</li></ul> 2) Extending the application of AnEn method to cover all the ground observation network of Morocco (In progress)
Spatial extension of analogs search space	
New low cost predictors weighting strategies based on AI	Applying AnEn on satellite images and exploring new similarity metrics via Deep Learning (In progress)

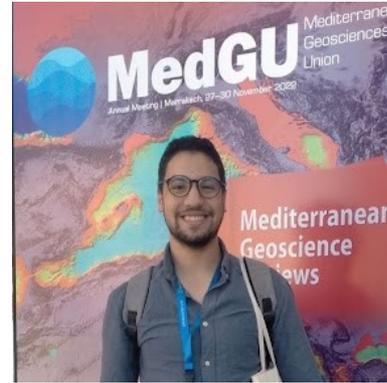


# Thank you!

**Badreddine Alaoui**

[alaoui.badreddine.abe@gmail.com](mailto:alaoui.badreddine.abe@gmail.com)

[alaoui.badreddine@marocmeteo.ma](mailto:alaoui.badreddine@marocmeteo.ma)



## Related Papers:

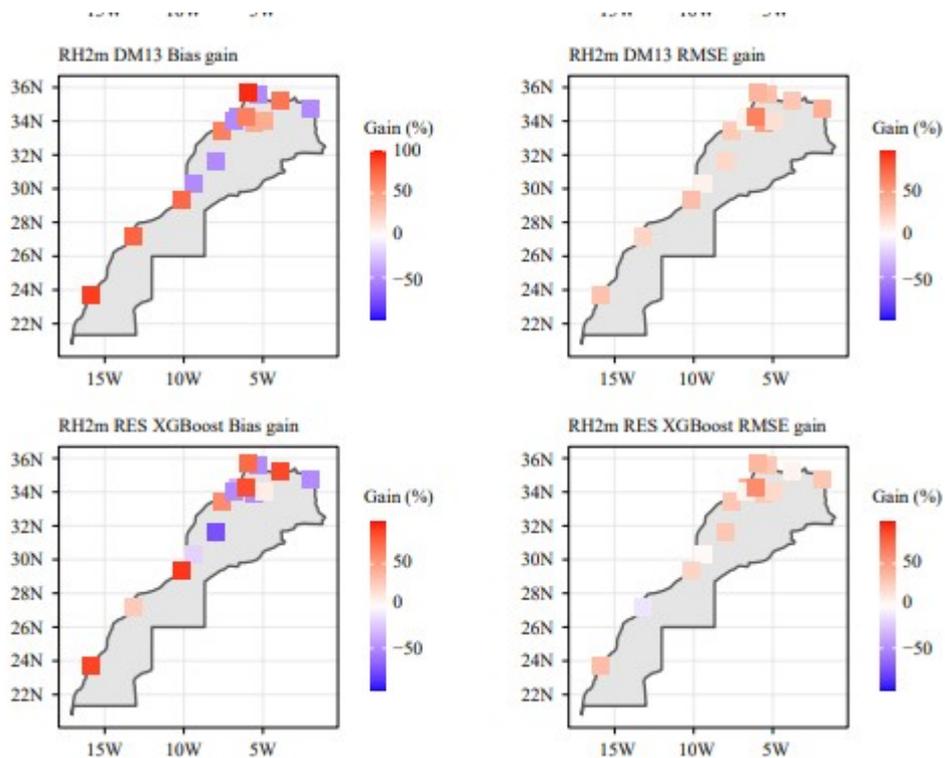
- Alaoui, B., Bari, D. & Ghabbar, Y. Surface Weather Parameters Forecasting Using Analog Ensemble Method over the Main Airports of Morocco. *J Meteorol Res* 36, 866–881 (2022). DOI: 10.1007/s13351-022-2019-0
- Alaoui, B.; Bari, D.; Bergot, T.; Ghabbar, Y. Analog Ensemble Forecasting System for Low-Visibility Conditions over the Main Airports of Morocco. *Atmosphere* 2022, 13, 1704. DOI: 10.3390/atmos13101704
- Alaoui, B., Bari, D. & Ghabbar, Y. Space Analog's Searching to Improve Deterministic Forecasting Using Analog Ensemble Method Over Morocco, Recent Advancements from Aquifers to Skies in Hydrogeology, Geoecology, and Atmospheric Sciences. DOI : 10.1007/978-3- 031-47079-0

# References

- Lorenz, E. Deterministic nonperiodic flow. *Journal Of Atmospheric Sciences*. 20, 130-141 (1963)
- Leith, C. Theoretical skill of Monte Carlo forecasts. *Monthly Weather Review*. 102, 409-418 (1974)
- Buizza, R., Leutbecher, M. & Isaksen, L. Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quarterly Journal Of The Royal Meteorological Society: A Journal Of The Atmospheric Sciences, Applied Meteorology And Physical Oceanography*. 134, 2051-2066 (2008)
- Lorenz, E. Atmospheric predictability as revealed by naturally occurring analogues. *Journal Of Atmospheric Sciences*. 26, 636-646 (1969)
- Seity, Y., Brousseau, P., Malardel, S., Hello, G., B' enard, P., Bouttier, F., Lac, C. & Masson, V. The AROME-France convective-scale operational model. *Monthly Weather Review*. 139, 976-991 (2011)
- Callado, A., Escrib' a, P., Garcia-Moya, J., Montero, J., Santos, C., Santos-Mu'noz, D. & Simarro, J. Ensemble forecasting. *Climate Change And Regional/Local Responses*. (2013)
- Eckel, F. & Mass, C. Aspects of effective mesoscale, short-range ensemble forecasting. *Weather And Forecasting*. 20, 328-350 (2005)
- Nurmi, P. Recommendations on the verification of local weather forecasts. (European Centre for Medium-Range Weather Forecasts, 2003)
- Wilks, D. Forecast verification. *International Geophysics*. 100 pp. 301-394 (2011)
- Toth, Z. & Kalnay, E. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin Of The American Meteorological Society*. 74, 2317-2330 (1993)
- Hamill, T. & Colucci, S. Verification of Eta/RSM short-range ensemble forecasts. *Monthly Weather Review*. 125, 1312-1327 (1997)
- Toth, Z. & Kalnay, E. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin Of The American Meteorological Society*. 74, 2317-2330 (1993)
- Toth, Z. & Kalnay, E. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*. 125, 3297-3319 (1997)
- Buizza, R. & Palmer, T. The singular-vector structure of the atmospheric global circulation. *Journal Of The Atmospheric Sciences*. 52, 1434-1456 (1995)
- Hamill, T., Mullen, S., Snyder, C., Toth, Z. & Baumhefner, D. Ensemble forecasting in the short to medium range: Report from a workshop. *Bulletin Of The American Meteorological Society*. 81, 2653-2664 (2000)
- Hamill, T. & Whitaker, J. Probabilistic quantitative precipitation forecasts based on reforecast analogs: Theory and application. *Monthly Weather Review*. 134, 3209-3229 (2006)
- Delle Monache, L., Eckel, F., Rife, D., Nagarajan, B. & Searight, K. Probabilistic weather prediction with an analog ensemble. *Monthly Weather Review*. 141, 3498-3516 (2013)
- Alessandrini, S., Delle Monache, L., Sperati, S. & Cervone, G. An analog ensemble for short-term probabilistic solar power forecast. *Applied Energy*. 157 pp. 95-110 (2015)
- Alessandrini, S., Delle Monache, L., Sperati, S. & Nissen, J. A novel application of an analog ensemble for short-term wind power forecasting. *Renewable Energy*. 76 pp. 768-781 (2015)
- Davo, F., Alessandrini, S., Sperati, S., Delle Monache, L., Airolidi, D. & Vespucci, M. Post-processing techniques and principal component analysis for regional wind power and solar irradiance forecasting. *Solar Energy*. 134 pp. 327-338 (2016)
- Alaoui, B., Bari, D. & Ghabbar, Y. Surface Weather Parameters Forecasting Using Analog Ensemble Method over the Main Airports of Morocco. *Journal Of Meteorological Research*. 36, 866-881 (2022)
- Alaoui, B., Bari, D., Bergot, T. & Ghabbar, Y. Analog Ensemble Forecasting System for Low-Visibility Conditions over the Main Airports of Morocco. *Atmosphere*. 13, 1704 (2022)
- PalMazumder, D. & M'olders, N. Theoretical assessment of uncertainty in regional averages due to network density and design. *Journal Of Applied Meteorology And Climatology*. 48, 1643-1666 (2009)
- Orrell, D., Smith, L., Barkmeijer, J. & Palmer, T. Model error in weather forecasting. *Nonlinear Processes In Geophysics*. 8, 357-371 (2001)
- Palmer, T. A nonlinear dynamical perspective on model error: A proposal for non-local stochastic-dynamic parametrization in weather and climate prediction models. *Quarterly Journal Of The Royal Meteorological Society*. 127, 279-304 (2001)
- Derksen, S. & Keselman, H. Backward, forward and stepwise automated subset selection algorithms: Frequency of obtaining authentic and noise variables. *British Journal Of Mathematical And Statistical Psychology*. 45, 265-282 (1992)
- Eckel, F. & Mass, C. Aspects of effective mesoscale, short-range ensemble forecasting. *Weather And Forecasting*. 20, 328-350 (2005)
- Zhu, Y. Ensemble forecast: A new approach to uncertainty and predictability. *Advances In Atmospheric Sciences*. 22, 781-788 (2005)
- Zhou, X., Zhu, Y., Hou, D., Luo, Y., Peng, J. & Wobus, R. Performance of the new NCEP Global Ensemble Forecast System in a parallel experiment. *Weather And Forecasting*. 32, 1989-2004 (2017)
- Toth, Z. Ensemble forecasting in WRF. *Bulletin Of The American Meteorological Society*. 82, 695-697 (2001)
- Toth, Z. & Kalnay, E. Ensemble forecasting at NMC: The generation of perturbations. *Bulletin Of The American Meteorological Society*. 74, 2317-2330 (1993)
- Toth, Z. & Kalnay, E. Ensemble forecasting at NCEP and the breeding method. *Monthly Weather Review*. 125, 3297-3319 (1997)
- Buizza, R., Leutbecher, M. & Isaksen, L. Potential use of an ensemble of analyses in the ECMWF Ensemble Prediction System. *Quarterly Journal Of The Royal Meteorological Society: A Journal Of The Atmospheric Sciences, Applied Meteorology And Physical Oceanography*. 134, 2051-2066 (2008)
- Buizza, R. & Richardson, D. 25 years of ensemble forecasting at ECMWF | ECMWF (No.153). *ECMWF Newsl.* (2017)
- Hope, T. Linear regression. *Machine Learning*. pp. 67-81 (2020)
- Chen, T. & Guestrin, C. Xgboost: A scalable tree boosting system. *Proceedings Of The 22nd Acm Sigkdd International Conference On Knowledge Discovery And Data Mining*. pp. 785-794 (2016)
- Breiman, L. Random forests. *Machine Learning*. 45, 5-32 (2001)

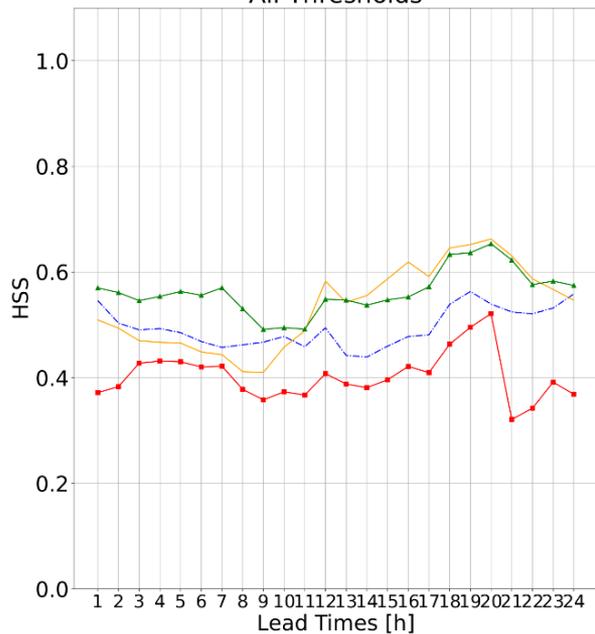
**Table 3.** Percentage of occurrence of possible weights values for each predictor for T2m as predictand

Variable	Weighting value										
	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
T2m (%)	19.4	9.7	–	6.5	–	3.2	12.9	22.6	25.8	–	–
RH2m (%)	58.1	41.9	–	–	–	–	–	–	–	–	–
WS10m (%)	83.9	16.1	–	–	–	–	–	–	–	–	–
MSLP (%)	51.6	19.4	3.2	–	9.7	16.1	–	–	–	–	–
SURFP (%)	61.3	6.5	–	6.5	9.7	16.1	–	–	–	–	–
WD10m (%)	83.9	16.1	–	–	–	–	–	–	–	–	–
MW10m (%)	74.2	22.6	3.2	–	–	–	–	–	–	–	–
ZW10m (%)	71	29	–	–	–	–	–	–	–	–	–

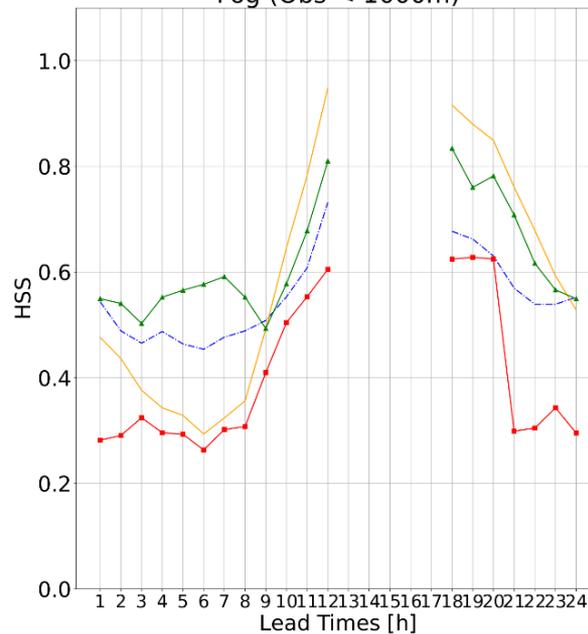


Alaoui, B., Bari, D. & Ghabbar, Y. Surface Weather Parameters Forecasting Using Analog Ensemble Method over the Main Airports of Morocco. *J Meteorol Res* 36, 866–881 (2022). DOI: 10.1007/s13351-022-2019-0

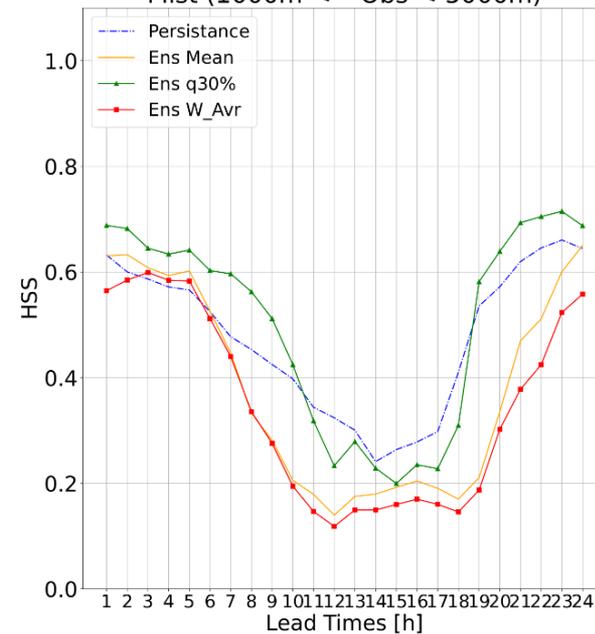
All Thresholds



Fog (Obs < 1000m)



Mist (1000m <= Obs < 5000m)



Alaoui, B.; Bari, D.; Bergot, T.; Ghabbar, Y. Analog Ensemble Forecasting System for Low-Visibility Conditions over the Main Airports of Morocco. *Atmosphere* 2022, 13, 1704. DOI: 10.3390/atmos13101704