



Seasonal Forcasting Verification

Supervised by Ouafae BADI, Idriss BARI, and Nicholas SAVAGE

Nohayla BERRAHMOUCH & Mohamed EL-BADRI

EHTP

3 décembre 2024

Outline

1 DATA

2 2M-TEMPERATURE

- DETERMINISTIC EVALUATIONS
- Probabilistic Evaluation Metrics

3 PRECIPITATION

- DETERMINISTIC EVALUATIONS
- Probabilistic Evaluation Metrics

4 CONCLUSION

LES VARIABLES ETUDES

ERA5-HINDCAST(1993-2016)

- 2-M TEMPERATURE
- PRECIPITATIONS

LES CENTRES DE HINDCASTS

- ukmo : UK Met Office.
- meteo_france : Modèles français de Météo-France.
- ecmwf : Modèles du Centre Européen pour les Prévisions Météorologiques à Moyen Terme.
- eccc : Environnement Canada (ECCC).
- dwd : Service météorologique allemand (Deutscher Wetterdienst).
- cmcc : Modèles du Centre Euro-Méditerranéen sur les Changements climatiques.

DETERMINISTIC EVALUATIONS

Spearman's Rank Correlation

$$r_s = \frac{\text{cov}(R[H], R[O])}{\sigma_{R[H]} \cdot \sigma_{R[O]}}$$

where :

- r_s : Spearman's rank correlation coefficient.
- H : Hindcast.
- O : Observation.
- $R[x]$: Rank of the variable x .
- σ_x : Standard deviation of the variable x .

DETERMINISTIC EVALUATIONS

Spearman's Rank Correlation

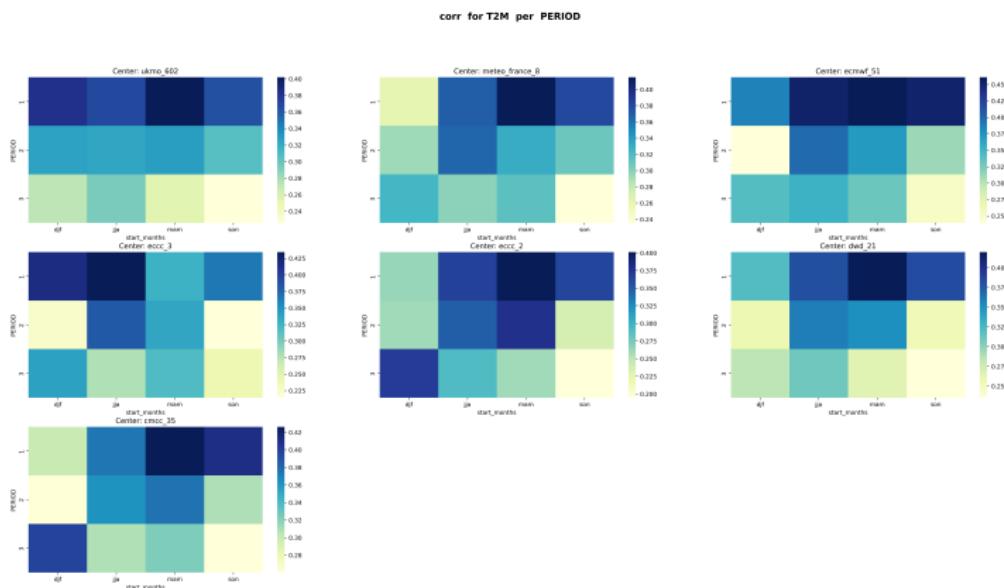


Figure – Heatmap for Spearman's Rank Correlation (2M Temperature)

DETERMINISTIC EVALUATIONS

RMSE : Root Mean Square Error

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_i - O_i)^2}$$

where :

- H : Hindcast.
- O : Observation.
- i : Index of valid time.
- n : Total number of observations.

DETERMINISTIC EVALUATIONS

RMSE : Root Mean Square Error

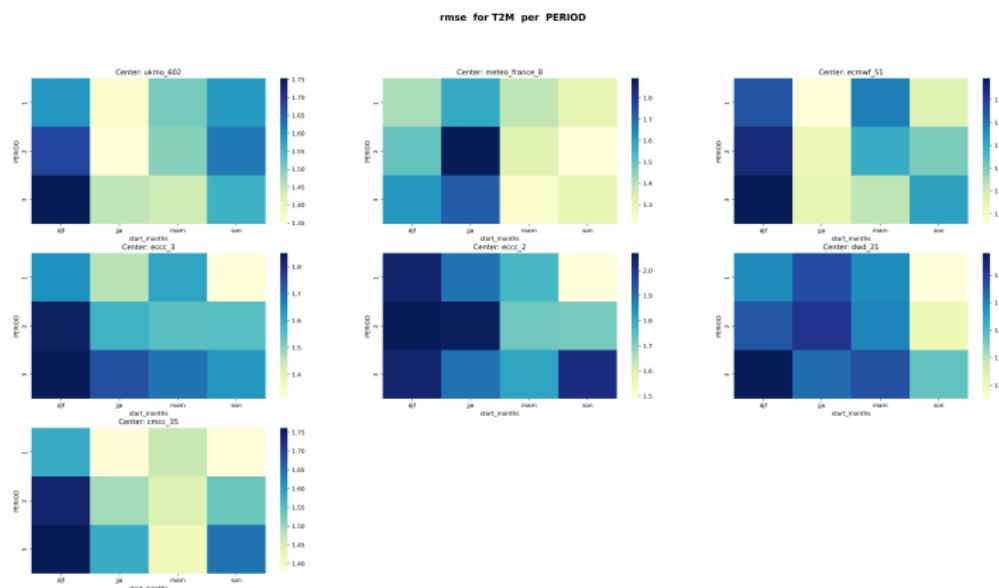


Figure – Heatmap for RMSE (2M Temperature)

DETERMINISTIC EVALUATIONS

R-squared (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^n (O_i - H_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}$$

where :

- R^2 : Coefficient of determination.
- H_i : Predicted value (Hindcast).
- O_i : Observed value (Observation).
- \bar{O} : Mean of observed values.
- $\sum_{i=1}^n (O_i - H_i)^2$: Residual sum of squares (unexplained variance).
- $\sum_{i=1}^n (O_i - \bar{O})^2$: Total sum of squares (total variance).

DETERMINISTIC EVALUATIONS

R-squared (R^2)

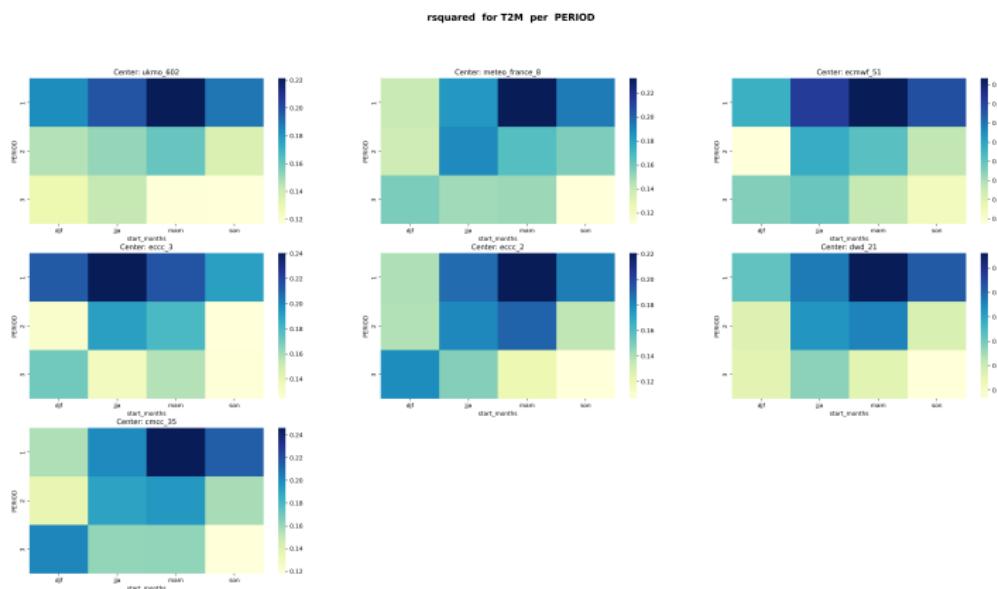


Figure – Heatmap for R² (2M Temperature)

Probabilistic Evaluation Metrics

The Brier Score (BS)

$$BS_j = \frac{1}{N} \sum_i^N (y_{j,i} - p_{j,i})^2$$

where :

- n is the number of forecasts
- $y_{j,i}$ is 1 if the i^{th} observation was in category j , and is 0 otherwise.
- $p_{j,i}$ is the i^{th} forecast probability for category j .

Probabilistic Evaluation Metrics

The Brier Score (BS)

bs T2M / lead_time

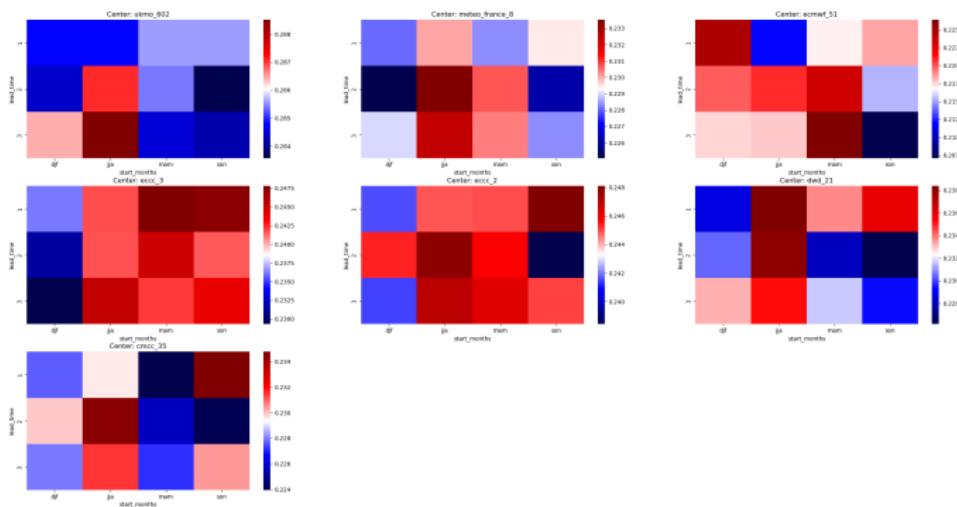


Figure – The Brier Score for each category . (0 represents perfect BS)

Probabilistic Evaluation Metrics

RELIABILITY

$$\text{Reliability} = \frac{1}{n} \sum_{k=1}^d n_k (\bar{p}_k - \bar{y}_k)^2$$

where :

- n_k is the number of forecasts for the $k_t h$ probability value (\bar{p}_k)
- (\bar{y}_k) is the observed relative frequency for that value.

Probabilistic Evaluation Metrics

ranked probability score (RPS)

$$RPS = \frac{1}{n(m-1)} \sum_{i=1}^n \sum_{k=1}^{m-1} \left(\sum_{j=1}^k (y_{j,i} - p_{j,i}) \right)^2$$

where :

- n is the number of forecasts.
- m is the number of categories.
- $y_{j,i}$ is 1 if the i^{th} observation was in category j, and is 0 otherwise.
- $p_{j,i}$ is the i^{th} forecast probability for category j

Probabilistic Evaluation Metrics

ranked probability score (RPS)

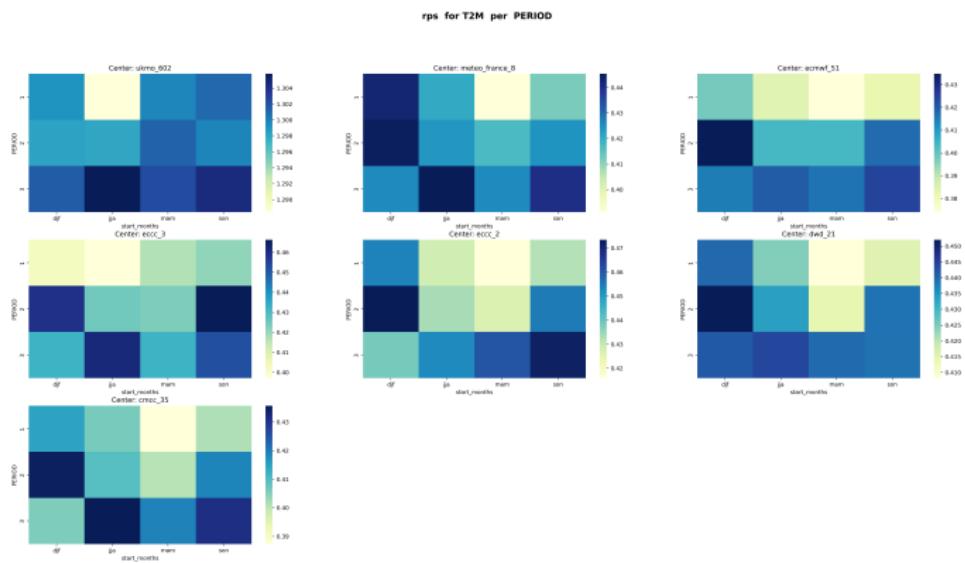


Figure – The average of RPS Score on all categories . (0 means perfect RPS)

Probabilistic Evaluation Metrics

Relative operating characteristics(ROC)

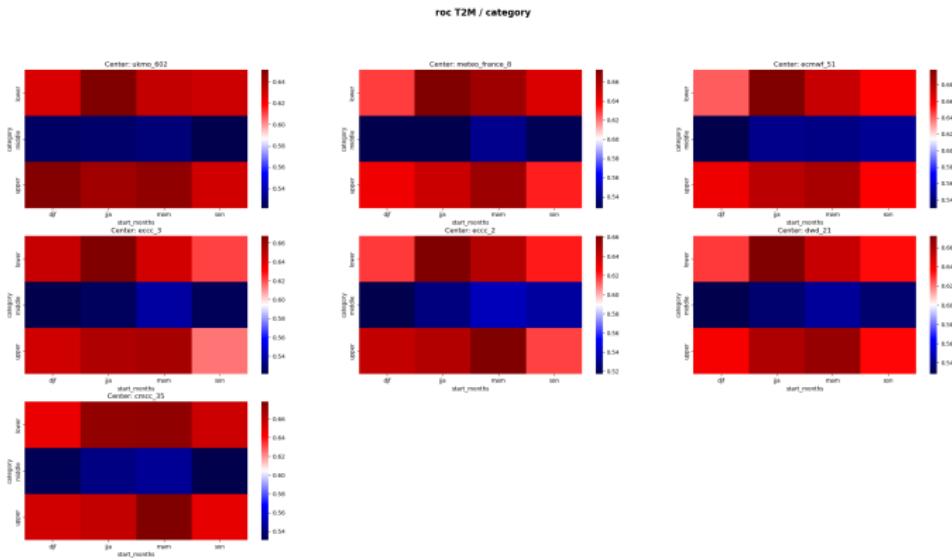


Figure – The ROC Score for each category . (1 means perfect ROC)

Probabilistic Evaluation Metrics

Relative operating characteristics(ROC)

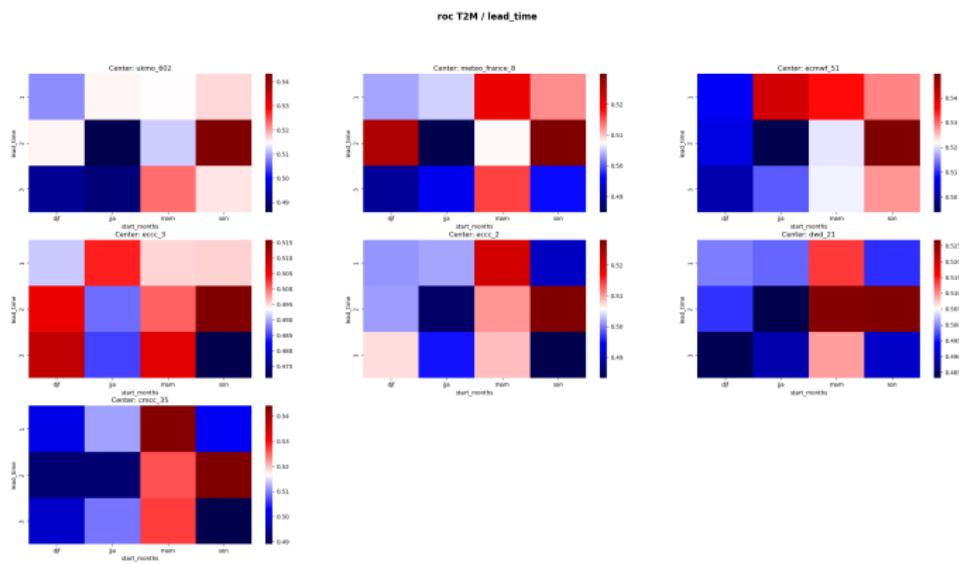


Figure – The ROC Score for each category . (1 means perfect ROC)

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

$$ROCSS = \frac{AUC - AUC_{no-skill}}{1 - AUC_{no-skill}}$$

where :

- AUC : Area Under the ROC Curve for the forecast being evaluated.
- $AUC_{no-skill}$: Area Under the Curve for a no-skill forecast 0.5 for our case.

Interpretation of ROCSS :

- 1 : Perfect discrimination ability.
- 0 : No skill (forecast performs no better than random guessing).
- Negative values : Forecast performs worse than random guessing.

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

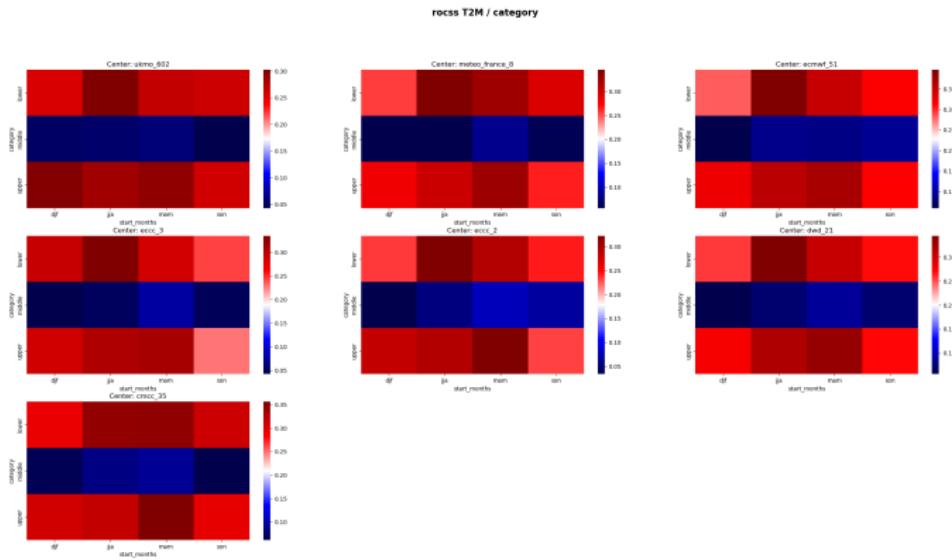


Figure – The ROCSS Score for each category . (1 means perfect ROCSS)

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

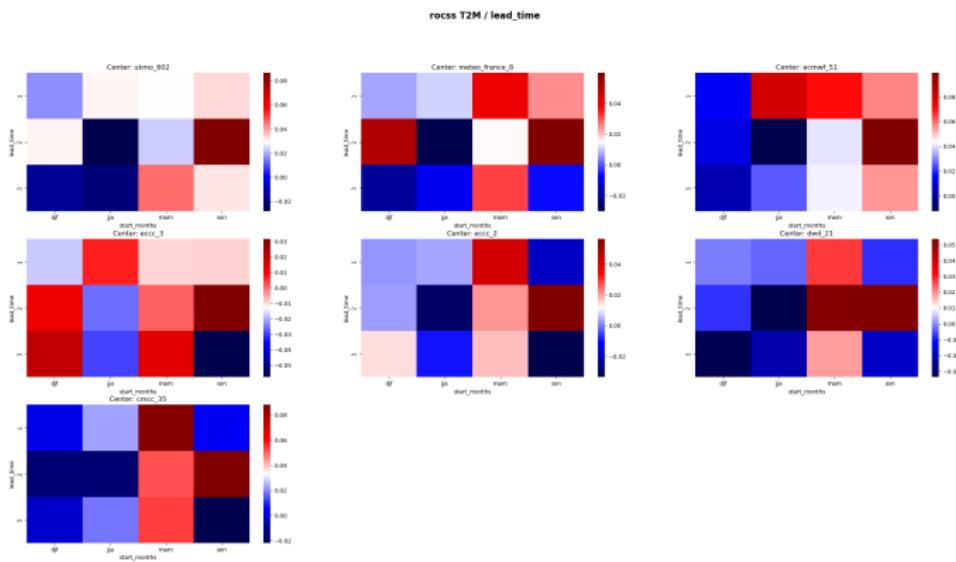


Figure – The ROCSS Score for each lead time . (1 means perfect ROCSS)

DETERMINISTIC EVALUATIONS

Spearman's Rank Correlation

$$r_s = \frac{\text{cov}(R[H], R[O])}{\sigma_{R[H]} \cdot \sigma_{R[O]}}$$

where :

- r_s : Spearman's rank correlation coefficient.
- H : Hindcast.
- O : Observation.
- $R[x]$: Rank of the variable x .
- σ_x : Standard deviation of the variable x .

DETERMINISTIC EVALUATIONS

Spearman's Rank Correlation

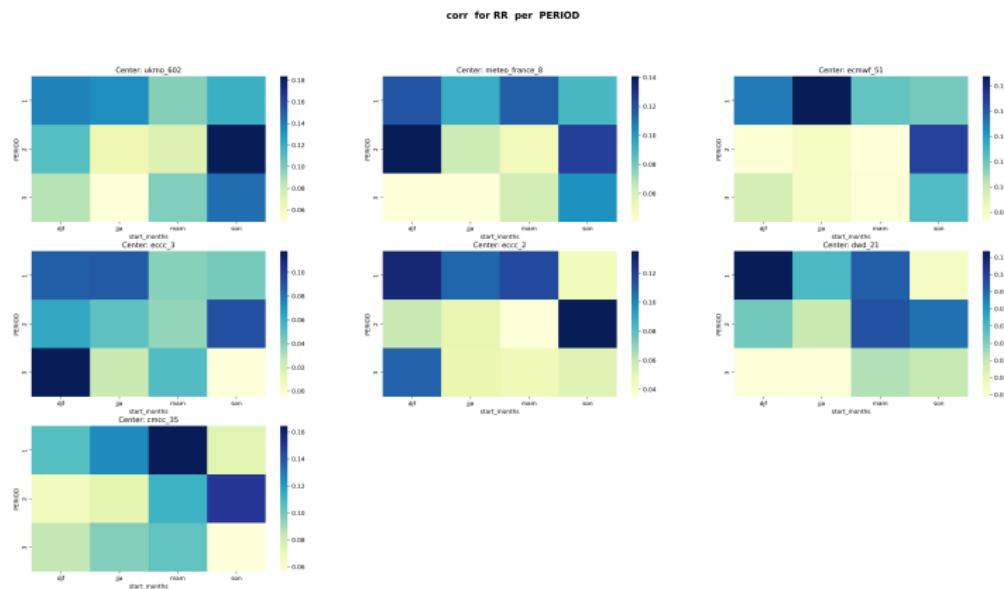


Figure – Heatmap for Spearman's Rank Correlation (PRECIPITATION)

DETERMINISTIC EVALUATIONS

RMSE : Root Mean Square Error

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (H_i - O_i)^2}$$

where :

- H : Hindcast.
- O : Observation.
- i : Index of valid time.
- n : Total number of observations.

DETERMINISTIC EVALUATIONS

RMSE : Root Mean Square Error

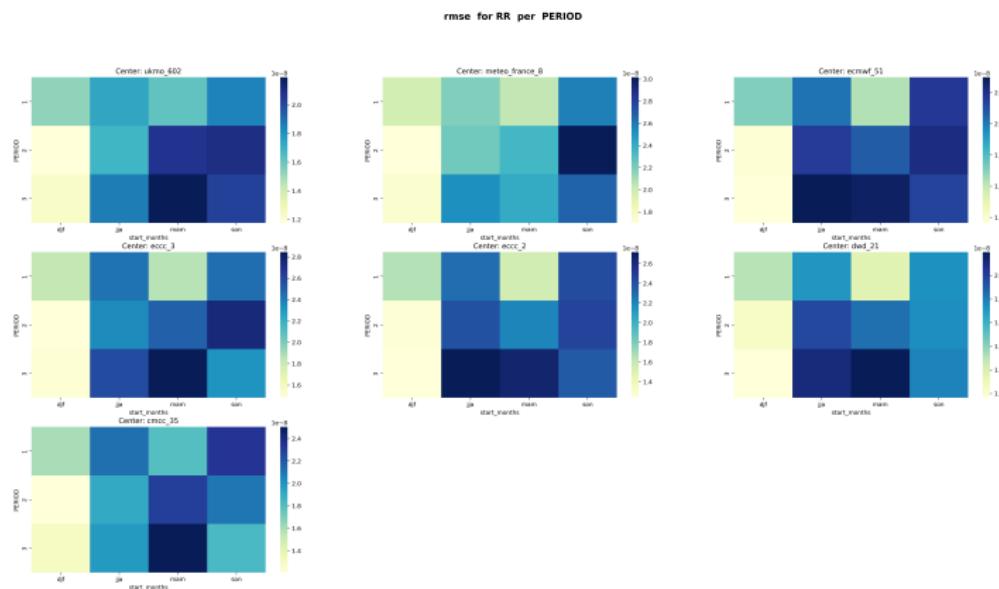


Figure – Heatmap for RMSE (PRECIPITATION)

DETERMINISTIC EVALUATIONS

R-squared (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^n (O_i - H_i)^2}{\sum_{i=1}^n (O_i - \bar{O})^2}$$

where :

- R^2 : Coefficient of determination.
- H_i : Predicted value (Hindcast).
- O_i : Observed value (Observation).
- \bar{O} : Mean of observed values.
- $\sum_{i=1}^n (O_i - H_i)^2$: Residual sum of squares (unexplained variance).
- $\sum_{i=1}^n (O_i - \bar{O})^2$: Total sum of squares (total variance).

DETERMINISTIC EVALUATIONS

R-squared (R^2)

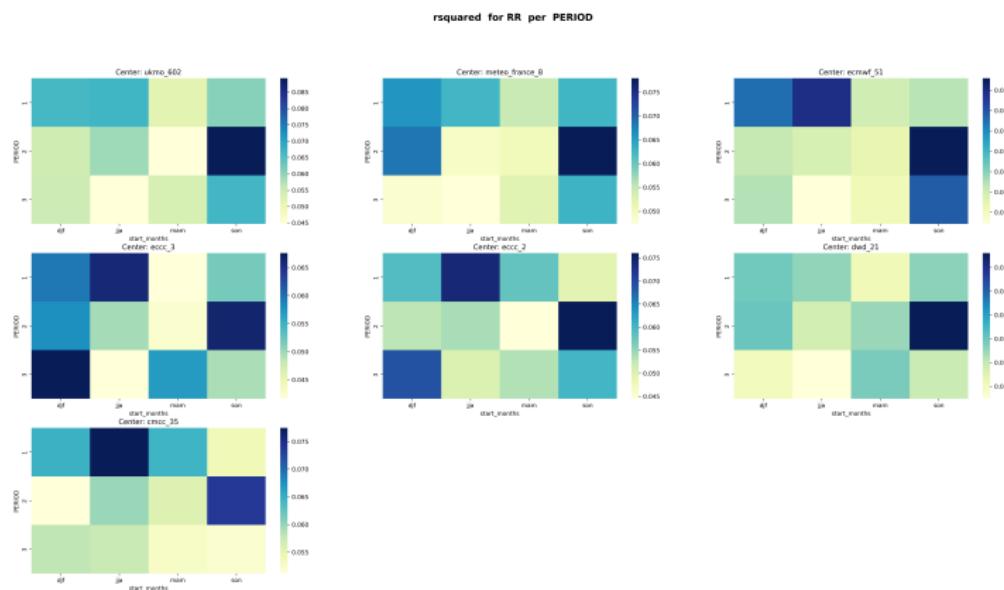


Figure – Heatmap for R^2 (PRECIPITATION)

Probabilistic Evaluation Metrics

RELIABILITY

$$\text{Reliability} = \frac{1}{n} \sum_{k=1}^d n_k (\bar{p}_k - \bar{y}_k)^2$$

where :

- n_k is the number of forecasts for the $k_t h$ probability value (\bar{p}_k)
- (\bar{y}_k) is the observed relative frequency for that value.

Probabilistic Evaluation Metrics

ranked probability score (RPS)

$$RPS = \frac{1}{n(m-1)} \sum_{i=1}^n \sum_{k=1}^{m-1} \left(\sum_{j=1}^k (y_{j,i} - p_{j,i}) \right)^2$$

where :

- n is the number of forecasts.
- m is the number of categories.
- $y_{j,i}$ is 1 if the i^{th} observation was in category j, and is 0 otherwise.
- $p_{j,i}$ is the i^{th} forecast probability for category j

Probabilistic Evaluation Metrics

ranked probability score (RPS)

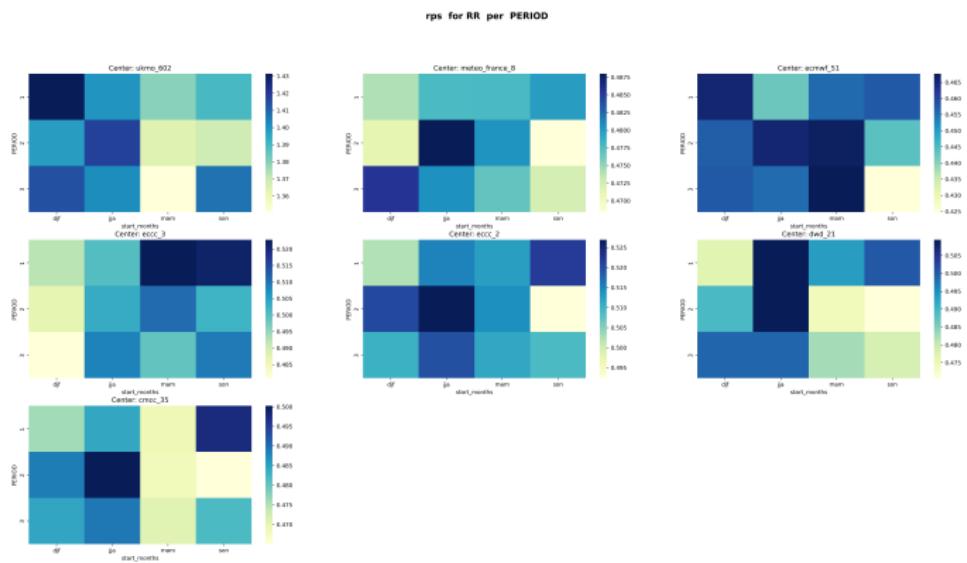


Figure – The average of RPS Score on all categories . (0 means perfect RPS)

Probabilistic Evaluation Metrics

The Brier Score (BS)

$$BS_j = \frac{1}{N} \sum_i^N (y_{j,i} - p_{j,i})^2$$

where :

- n is the number of forecasts
- $y_{j,i}$ is 1 if the i^{th} observation was in category j , and is 0 otherwise.
- $p_{j,i}$ is the i^{th} forecast probability for category j .

Probabilistic Evaluation Metrics

The Brier Score (BS)

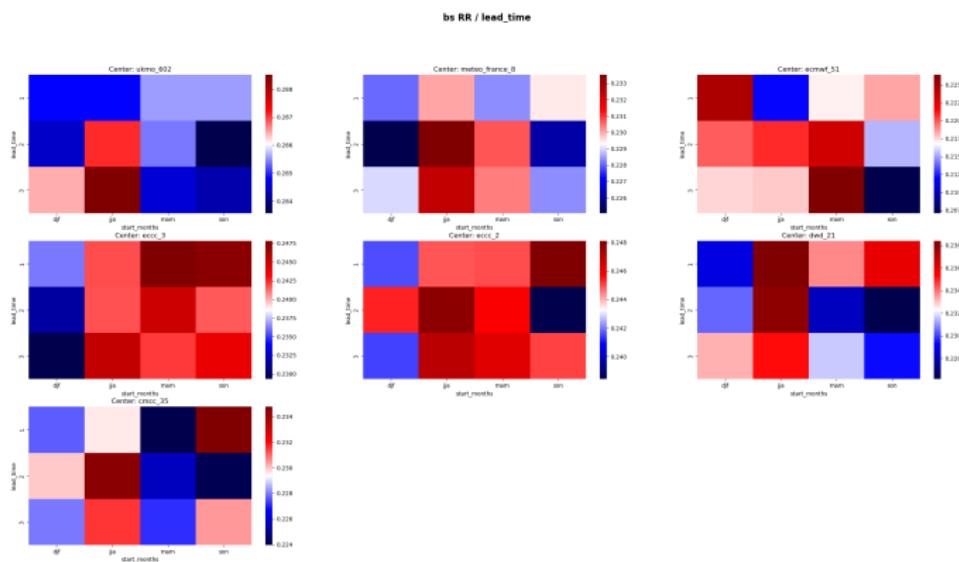


Figure – The Brier Score for each lead time . (0 represents perfect BS)

Probabilistic Evaluation Metrics

The Brier Score (BS)

bs RR / category

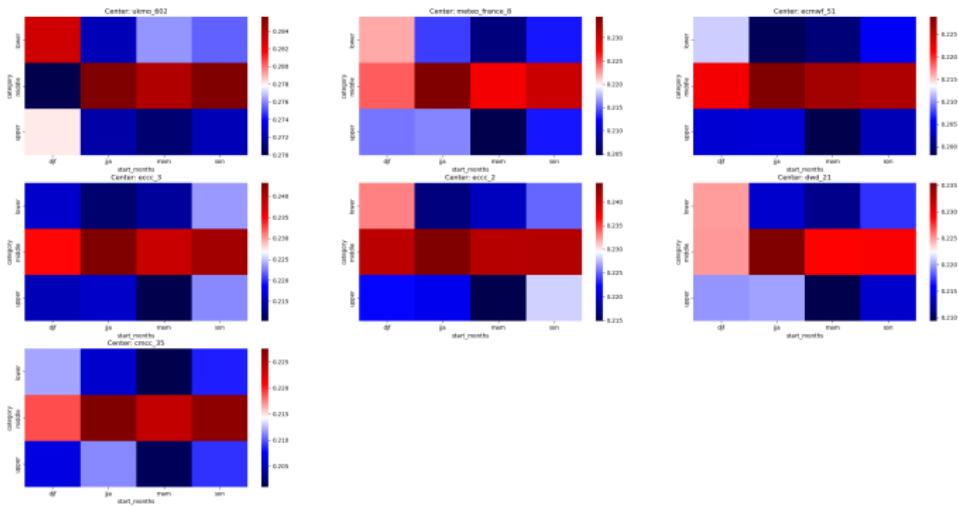


Figure – The Brier Score for each category . (0 represents perfect BS)

Probabilistic Evaluation Metrics

Relative operating characteristics(ROC)

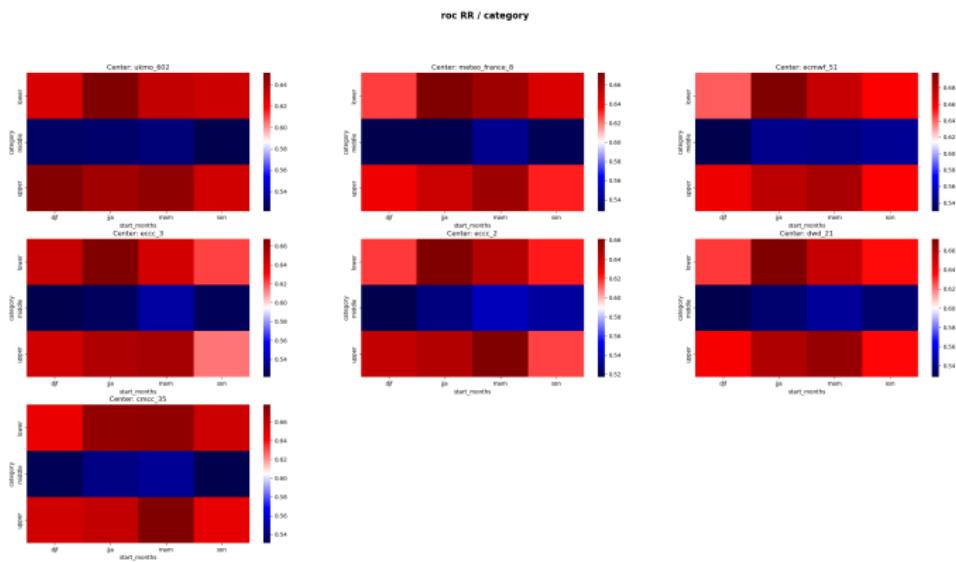


Figure – The ROC Score for each category . (1 means perfect ROC)

Probabilistic Evaluation Metrics

Relative operating characteristics(ROC)

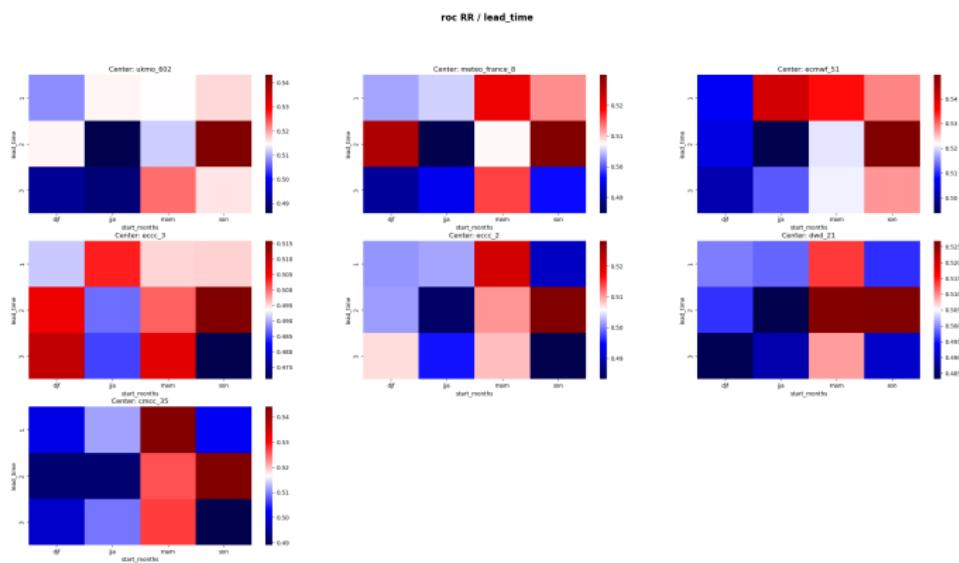


Figure – The ROC Score for each LEAD TIME. (1 means perfect ROC)

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

$$\text{ROCSS} = \frac{AUC - AUC_{no-skill}}{1 - AUC_{no-skill}}$$

where :

- AUC : Area Under the ROC Curve for the forecast being evaluated.
- $AUC_{no-skill}$: Area Under the Curve for a no-skill forecast 0.5 for our case.

Interpretation of ROCSS :

- 1 : Perfect discrimination ability.
- 0 : No skill (forecast performs no better than random guessing).
- Negative values : Forecast performs worse than random guessing.

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

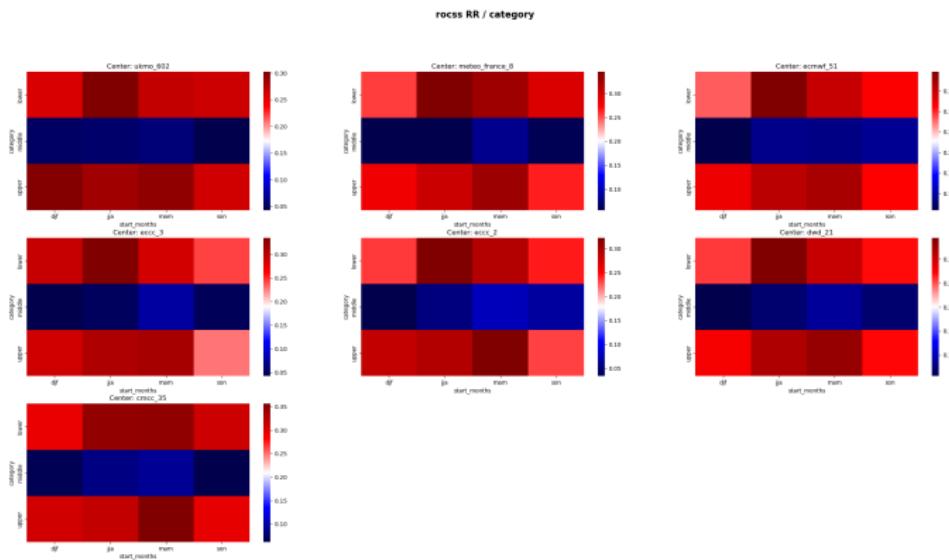


Figure – The ROCSS Score for each category . (1 means perfect ROCSS)

Probabilistic Evaluation Metrics

Relative operating characteristics Skill Score(ROCSS)

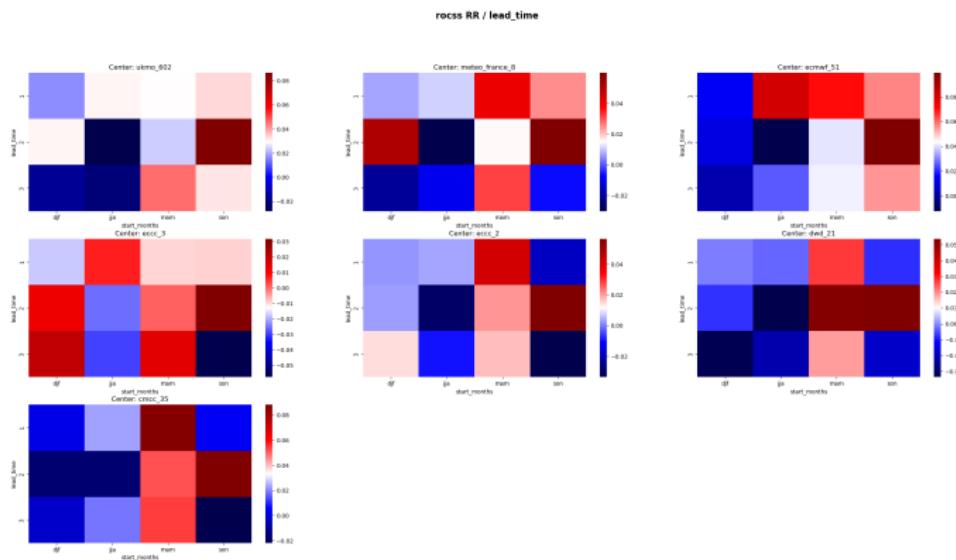


Figure – The ROCSS Score for each lead time . (1 means perfect ROCSS)

LES PAROLES S'ENVOIENT, LE CODE QUI RESTE

<https://github.com/Mohamed-01-git/MODELIZATION>
<https://www.kaggle.com/mohamedelbadri01>