Forecasting Solar Radiation with Online Ensemble Learning

Jean Thorey\textsuperscript{1,2}, Vivien Mallet\textsuperscript{2}, Christophe Chaussin\textsuperscript{3}, Laurent Descamps\textsuperscript{4}, Philippe Blanc\textsuperscript{5}

\textsuperscript{1} CEREA, Joint Laboratory of École des Ponts ParisTech - EDF R&D\textsuperscript{2} INRIA\textsuperscript{3} EDF R&D\textsuperscript{4} Météo-France\textsuperscript{5} Center for Energy and Processes, MINES ParisTech

ICEM 2015
1. Ensemble of forecasts for solar radiation

2. Online ensemble learning
Typical days

The variable of interest is Global Horizontal Irradiance (GHI in $\text{W m}^{-2}$).

<table>
<thead>
<tr>
<th></th>
<th>Forecast</th>
<th>Verification (satellite)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example 1</td>
<td><img src="image" alt="Forecast" /></td>
<td><img src="image" alt="Verification" /></td>
</tr>
<tr>
<td>Example 2</td>
<td><img src="image" alt="Forecast" /></td>
<td><img src="image" alt="Verification" /></td>
</tr>
</tbody>
</table>

→ Need for multiple forecasts.
→ Spatial resolution.
Forecast horizon from several hours to a few days: we use Numerical Weather Predictions (NWP).

Many model runs form an ensemble of forecasts (different atmospheric models, parameterizations, perturbations).

Current Research

Study ensemble forecasting with statistics and online learning tools. Application: solar forecasting.
Ensemble of forecasts

Solar radiation data has a high variability and a low predictability.

The ensemble of forecasts is designed to catch the uncertainty of the observations.
The experiment

Goal: estimate the average GHI between 0600 and 1200 UTC, on the high resolution grid.

Horizon: 12 h, with possible extensions.

Verification data: real-time satellite-derived estimations, resolution of $\sim 10$ km, averaged from the database HelioClim 3.

Forecasts: TIGGE ensembles, resolution of $\sim 25$ km, available online.

<table>
<thead>
<tr>
<th>Center</th>
<th>Origin</th>
<th>Number of members</th>
<th>Run</th>
</tr>
</thead>
<tbody>
<tr>
<td>CMA</td>
<td>China</td>
<td>14</td>
<td>2400</td>
</tr>
<tr>
<td>ECMWF</td>
<td>UE</td>
<td>50</td>
<td>2400</td>
</tr>
<tr>
<td>UKMO</td>
<td>UK</td>
<td>23</td>
<td>2400</td>
</tr>
<tr>
<td>KMA</td>
<td>Korea</td>
<td>23</td>
<td>2400</td>
</tr>
<tr>
<td>CPTEC</td>
<td>Brazil</td>
<td>14</td>
<td>2400</td>
</tr>
<tr>
<td>Météo France</td>
<td>France</td>
<td>34</td>
<td>1800</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td></td>
<td>158</td>
</tr>
</tbody>
</table>

→ Combine TIGGE forecasts.
Forecasts from various meteorological centers

Annual average of the forecasts in $\text{W m}^{-2}$ for year 2012.
Rank diagrams. They show underdispersed ensembles.

Monthly scores.
- TIGGE ensemble means.
- – ECMWF deterministic forecast.
- · · Our forecast.
Online ensemble learning

Principle: create a linear combination $x^*$ of the members with best performance than any individual member.

$$x^*(t) = \sum_{m=1}^{M} u_m(t) \times x_m(t).$$

$x_m$ : value of $m$-th member.
$u_m$ : weight of $m$-th member.
y : observation (or verification) to be forecasted.

The members are sorted in each TIGGE ensemble.

The method is applied at each grid point on the high resolution grid.
Online ensemble learning

\[ x^*(t) = \sum_{m=1}^{M} u_m(t) \times x_m(t). \]

\( x_m \): value of \( m \)-th member.

\( u_m \): weight of \( m \)-th member.

\( y \): observation (or verification) to be forecasted.
Online ensemble learning

\[ x^*(t) = \sum_{m=1}^{M} u_m(t) \times x_m(t). \]

- \( x_m \): value of \( m \)-th member.
- \( u_m \): weight of \( m \)-th member.
- \( y \): observation (or verification) to be forecasted.
Weights computation

Online regularized regression.

Loss: \( \ell_t(u) = (u^\top x_t - y_t)^2 \).

Regularization term: \( r(u) = \lambda \|u\|^2 \).

The weight vector \( u_t = [u_1(t), u_2(t), ..., u_M(t)]^\top \) is chosen as:

\[
\arg\min_{w \in \mathbb{R}^M} \left[ r(w) + \sum_{t' = 1}^{t-1} \ell_t(w) \right].
\]
Weights computation

Online regularized regression.

Loss: \( \ell_t(u) = (u^\top x_t - y_t)^2. \)

Regularization term: \( r(u) = \lambda \|u\|^2. \)

The weight vector \( u_t = [u_1(t), u_2(t), ..., u_M(t)]^\top \) is chosen as:

\[
\arg\min_{w \in \mathbb{R}^M} \left[ r(w) + \sum_{t' = 1}^{t-1} \beta_{t,t'} \ell_t(w) \right].
\]

Focus on recent data with discount factors:

\[
\beta_{t,t'} = \frac{\gamma}{(t - t')^2}
\]

In practice: low sensitivity around optimal parameters \( \lambda, \gamma. \)
Theoretical guarantee

Case of online ridge regression (under essentially no assumptions):

- Tends to zero with increasing $T$.
- Cumulated loss of our forecasts.

$$\frac{1}{T} \left( \sum_{t=1}^{T} \ell_t(u_t) \right) - \min_{w \in \mathbb{R}^M} \sum_{t=1}^{T} \ell_t(w) \leq O \left( \frac{\ln T}{T} \right)$$

- Cumulated loss of the best combination with constant weights (more skillful than the ensemble mean and each individual member).

Robustness: the bound holds for any time series $y_t, x_{m,t}$. 
Typical forecast

Localized corrections $\rightarrow$ improved resolution.

Best member aggregated forecast

Verification

2012/05/16.
Annual average of the forecasts

Local biases are corrected.

reference forecast  aggregated forecast

verification

Irradiance online learning
Performance maps

RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (u_t \top x_t - y_t)^2}.

relative RMSE = \frac{\text{RMSE}}{\left( \frac{1}{T} \sum_{t=1}^{T} y_t \right)}.
Weights and Lead Time

Hourly forecasts: $x_{m,h,D,k}$ (member $m$, hour $h$, day $D$, and lead time $k>1$).

To estimate the observation $y_{h,D+k}$.

- Weights defined per hour and lead time, based on $x_{m,h,D,k}$.

  \[ W_{m,h,D,k} \rightarrow \hat{y}_{h,D+k} \]

  \[ \begin{array}{c}
  \text{D - 1} \\
  \text{D} \\
  \text{D + k}
  \end{array} \]

- Weights defined per hour, based on $x_{m,h,D,1}$ (for the shortest lead time), by lagging the weights: $w_{m,h,D,1} = w_{m,h,D,2} = \ldots = w_{m,h,D,k}$.

  \[ \begin{array}{c}
  W_{m,h,D,1} \rightarrow \hat{y}_{h,D+k} \\
  \text{D - 1} \\
  \text{D} \\
  \text{D + 1} \\
  \text{D + k}
  \end{array} \]
Weights and Lead time

With lagged weights computed for first next forecast of 1200 UTC.

Weights can be kept for longer lead times.

Thorey et al. (EDF R&D, INRIA)
Conclusion

- Ensemble forecasting of solar radiation.
- From low resolution forecasts to high resolution observations.
- Online ensemble learning brings theoretical guarantee.
- Robust and simple method, compatible with operational forecasts.

Perspectives

- Combine forecasts with finer resolution in time and space.
- Use of spatial information, aggregation of spatial structures.


Thank you for your attention.
Assumptions for the theoretical guarantee

\[
\frac{1}{T} \left( \sum_{t=1}^{T} \ell_t(u_t) - \min_{w \in \mathbb{R}^M} \sum_{t=1}^{T} \ell_t(w) \right) \leq O \left( \frac{\ln T}{T} \right)
\]

- Our losses are bounded: \( \ell_t(u_t) < C_1 \).
- The oracle weights are bounded for the 2-norm: \( w^\top w < C_2 \).

The oracle can be built on selected orthogonalized members to avoid overfitting.
Alternatives

- Pre-process time series: clear sky normalization.
- Chose few members in each sub-ensemble.
- Two-step aggregation (first within each sub-ensemble).
- Identify the members per rank or?
- Introduce new members:
  - lagged ensemble,
  - analogues,
  - near grid points forecasts.
Correction of the largest errors

- ECMWF deterministic forecast.
- ECMWF deterministic forecast with local corrections.
- Aggregated forecast.
Size effect for the aggregation

- No sorting.
- Sorting full ensemble.
- Sorting for each TIGGE center.
- Sorting for each TIGGE center + ECMWF deterministic forecast.
Performance maps

RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (u_t^\top x_t - y_t)^2}.

Relative RMSE = \frac{\text{RMSE}}{\left( \frac{1}{T} \sum_{t=1}^{T} y_t \right)}.
Performance maps

autumn winter

RMSE

relative RMSE

RMSE = \sqrt{\frac{1}{T} \sum_{t=1}^{T} (u_t^T x_t - y_t)^2}.

relative RMSE = \frac{\text{RMSE}}{\left(\frac{1}{T} \sum_{t=1}^{T} y_t\right)}.
Rank diagram of the whole ensemble (158 members).

The gray scale indexes the number of ensembles whose spread contains the observation.
### Conversion methods

<table>
<thead>
<tr>
<th>label</th>
<th>conversion formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>glob</td>
<td>$x_{tigge} \times 1.18$</td>
</tr>
<tr>
<td>mult</td>
<td>$x_{tigge} \times \frac{x_{ssrd}}{x_{ssr ECMWF}}$</td>
</tr>
<tr>
<td>add</td>
<td>$x_{tigge} - \left( x_{ssr ECMWF} - x_{ssrd ECMWF} \right)$</td>
</tr>
<tr>
<td>lin</td>
<td>$x_{tigge} \times a_{center} + b_{center}$</td>
</tr>
<tr>
<td>no conversion</td>
<td></td>
</tr>
</tbody>
</table>

---

![Graph showing RMSE for different models](image-url)