Improvement of Extreme Temperatures
Probabilistic Short-Term Forecasting

Adriana Gogonel $^{1,2}$  Jérôme Collet $^1$  Avner Bar-Hen$^2$

$^1$EDF R&D

$^2$Paris-Descartes University

May 29, 2012
1. Problem statement

2. A use of Best Member Method

3. Improvement using Quantile Regression

4. Other tentative improvements

5. Conclusion
Temperature-caused risk management

- Impact of temperature on electric system: heating and cooling of homes, cooling of thermal pants.
- Main input: Ensemble Prediction Systems (EPS), many meteorological scenarios, differing by its initial state.
- Lack of accuracy for tails, lack of smoothness.
- Best Member Method: modeling of the error of the EPS member closest to the realization.
Variables

- variables summarizing features of the current EPS: mean, spread of the ensemble;
- variables to take account of a smooth influence of the date;
- the rank of the member, considered as a discrete variable;
- since central ranks behave similarly; variable edge, equal to 0 when the rank is central, to the rank elsewhere.

with interactions between discrete and continuous variables.
Selection summary

- edge always preferred to the rank;
- ensemble mean significant for the modeling of the mean error, the spread for its variance;
- useless to take account of a smooth influence of the date,
- interactions between edge and ensemble mean or spread useful.
Results

Important improvement of the reliability of the temperature density forecast:

- reliability component of the CRPS
- plot of the distribution of the Probability Integral Transform (rank of the realization among the simulations)
The quantile regression method

Use of a different loss function: slope constant on $\mathbb{R}^+$ and $\mathbb{R}^-$

Intuition of the minimization:
- variation in the estimate generates
  - variation in the loss, depending only on:
    - 2 slopes, number of points on each side.
Use of quantile regression

Use to forecast 2 central quantiles: first and third quartiles. Same variables selected, significance proved using resampling.

**Measure of improvement**: for tail representation, take account of relative errors.

Use of a $\chi^2$ distance, with classes:

- $[0; 0.01]$
- $[0.01; 0.02]$
- $[0.02; 0.05]$

and symmetric classes for upper tail.
Results

Graph showing the relationship between real data and forecast values, with different lines representing various types of models.
Results

![Graph showing extreme temperatures probabilistic short-term forecasting results]
Error when best member is extreme

- more frequent,
- approximately 5 times bigger,
- dissymmetric distribution, fitting well with GEV type.

Little improvement using specific modeling.
Conditional modelings

We tried different distributions for some partitioning of data.

by horizon

by season
Conclusion

Summary of the results:

- statistical processing compulsory to improve reliability,
- tail representation needs specific modeling,
- quantile regression provides important improvement,
- specific modeling of extreme errors does not.

Further work: significant dependence between best member distribution and

- month,
- spread,
- forecasted temperature.

Use for further improvements?
## CRPS for raw EPS and BMM

<table>
<thead>
<tr>
<th>horizon</th>
<th>pot BMM</th>
<th>Reliab. BMM</th>
<th>pot MF</th>
<th>Reliab. MF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.25</td>
<td>3.47E-03</td>
<td>0.24</td>
<td>2.59E-02</td>
</tr>
<tr>
<td>3</td>
<td>0.39</td>
<td>3.60E-03</td>
<td>0.38</td>
<td>6.92E-03</td>
</tr>
<tr>
<td>5</td>
<td>0.64</td>
<td>3.21E-03</td>
<td>0.63</td>
<td>3.49E-03</td>
</tr>
<tr>
<td>7</td>
<td>0.97</td>
<td>2.43E-03</td>
<td>0.96</td>
<td>4.61E-03</td>
</tr>
<tr>
<td>9</td>
<td>1.26</td>
<td>2.47E-03</td>
<td>1.24</td>
<td>4.67E-03</td>
</tr>
<tr>
<td>11</td>
<td>1.47</td>
<td>2.97E-03</td>
<td>1.43</td>
<td>3.96E-03</td>
</tr>
<tr>
<td>13</td>
<td>1.57</td>
<td>2.93E-03</td>
<td>1.54</td>
<td>4.19E-03</td>
</tr>
</tbody>
</table>