Royal Netherlands
Meteorological Institute
*Ministry of Infrastructure and the Environment*

# A comparative verification of high resolution precipitation forecasts using model output statistics

*Emiel van der Plas [1], Maurice Schmeits [1] (schmeits@knmi.nl), Nicolien Hooijman [1] and Kees Kok [1]*

**Despite the realism of precipitation forecasts from the non-hydrostatic Harmonie model (2.5 km resolution), they only perform similarly as precipitation forecasts from 2 lower resolution hydrostatic models (ECMWF with 16 km and Hirlam with 11 km resolution), at least in the Netherlands.**

## 1 Introduction

Verification of localized events such as precipitation has become even more challenging with the advent of high resolution mesoscale numerical weather prediction (NWP). The realism of a forecast suggests that it should compare well against precipitation radar imagery with similar resolution, both spatially and temporally. Spatial verification methods solve some of the representativity issues that point verification gives rise to. In this study a verification strategy based on model output statistics is applied that aims to address both double penalty and resolution effects that are inherent to comparisons of NWP models with different resolutions. This method, extensively described in [1], attempts to extract all relevant characteristics (predictors) from the available precipitation forecasts that have predictive potential for the probability of rain (exceeding some threshold) at a location or area during a time interval (predictand).

## 2 Data and method

### 2.1 Data

- Data availability: 2 years of 'summer' days (April - October) from August 2012 to July 2014
- The predictands are defined as follows:
  - 3-h precipitation sums at 7 SYNOP stations in The Netherlands (Fig. 1)
  - areal mean and maximum 3-h precipitation sums from calibrated radar data in circles (with 5 and 25 km radii) around 7 SYNOP stations in The Netherlands (Fig. 1)
- The potential predictor groups are defined as follows:
  - `q`: the exceedance threshold $q$,
  - `dmo`: precipitation as given by the model at the closest grid point to the station (direct model output),
  - `dist`: predictors based on the (weighted) distance from the station to the closest wet or dry grid point,

  and within a disk of radius $R$ around a station, with $R = 25$, 50, 75 and 100 km:

  - `max`: the (square root of the) maximum areal precipitation,
  - `coverage`: predictors based on the coverage of precipitation (i.e. the percentage of the area covered by precipitation),
  - `total`: the (square root of the) areal mean precipitation,
  - `weightmax`: predictors based on the maximum precipitation weighted by the distance to the station,
  - `weightint`: predictors based on the precipitation integrated over the area, weighted by an exponentially decaying function from the station.
- Because predictors within 1 group are highly correlated, only 1 predictor from each group can be selected.

### 2.2 Method: Extended Logistic regression (ELR)

As our statistical post-processing method we have used Extended Logistic Regression (ELR; [3]), where the threshold is also a predictor so that the regression equation is a function of the threshold, effectively yielding a complete probability distribution.

For a binary predictand $y_i$, here for an event with precipitation exceeding a threshold $q$, we try to find the probability $p_i$ as a function of the threshold $q$ and the other predictors $\mathbf{s} \equiv [s_{1,i}, s_{2,i}, ...]^T$, according to the nonlinear logistic equation

$$p_i(\mathbf{s}, q) = \frac{\exp(b_0 + b_1 s_{1,i} + b_2 s_{2,i} + ... + f(q))}{1 + \exp(b_0 + b_1 s_{1,i} + b_2 s_{2,i} + ... + f(q))}.$$

In this study, we have chosen a linear function of the threshold: $f(q) = aq$.

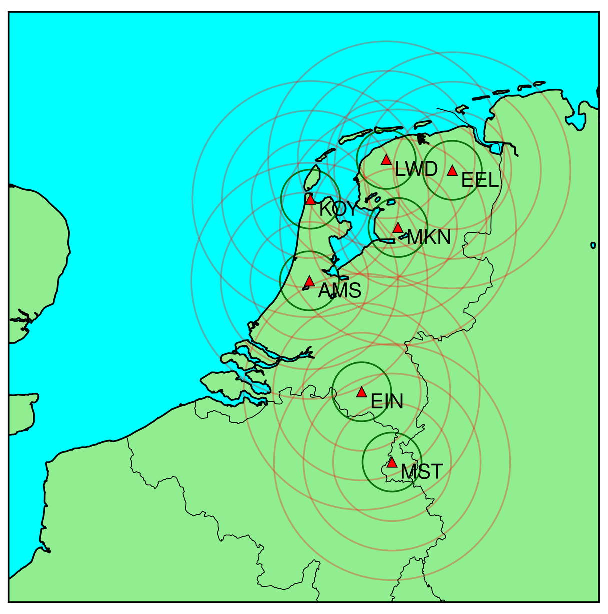


*Figure 1: Location of the stations on a map in the Netherlands: from north to south Leeuwarden (LWD), Eelde (EEL), De Kooy (KOY), Marknesse (MKN), Amsterdam Airport Schiphol (AMS), Eindhoven (EIN) and Maastricht (MST). Circles with a radius of 25 (green), 50, 75 and 100 km (grey) around each station are indicated (see text).*

## 3 Results

### 3.1 Selected predictors

In principle, each model may use very different predictors to achieve the best correlation or highest skill given the restrictions discussed in [2]. When we look at which predictor was selected first for the different models (see Fig. 2 for precipitation measured by SYNOP stations as predictand), we see that in general the large scale predictors, such as the square root of the mean areal precipitation from the `total` group and `coverage` (the fraction of the circular area where the precipitation exceeds 0.3 mm/3h) are preferred. The threshold is always selected as the second predictor. A predictor of the `dist` group is often picked third, indicating the (relative) importance of the distance of the predicted precipitation to the SYNOP station. Also, it is noteworthy that the predictor based on the direct model output (`dmo`) is never selected before the stopping criterion is met. Apparently, it contains no skilful information in addition to the selected predictors. This is in complete agreement with the fact that high resolution forecasts should not be taken at face value but should be interpreted probabilistically. For predictands that involve maxima, such as the maximum precipitation within a radius of 25 km (not shown), the predictors from the `max` group are selected more often, as third or fourth predictor.

In Fig. 3 the distribution of the scales of the selected predictors is given for the same four lead times and predictand as in Fig. 2. As expected, there is a general tendency to select predictors on larger spatial scales with increasing lead time. It is remarkable that even for SYNOP observations as a predictand source and especially for Harmonie, predictors defined on disks with a radius of 25 and 50 km are hardly selected.

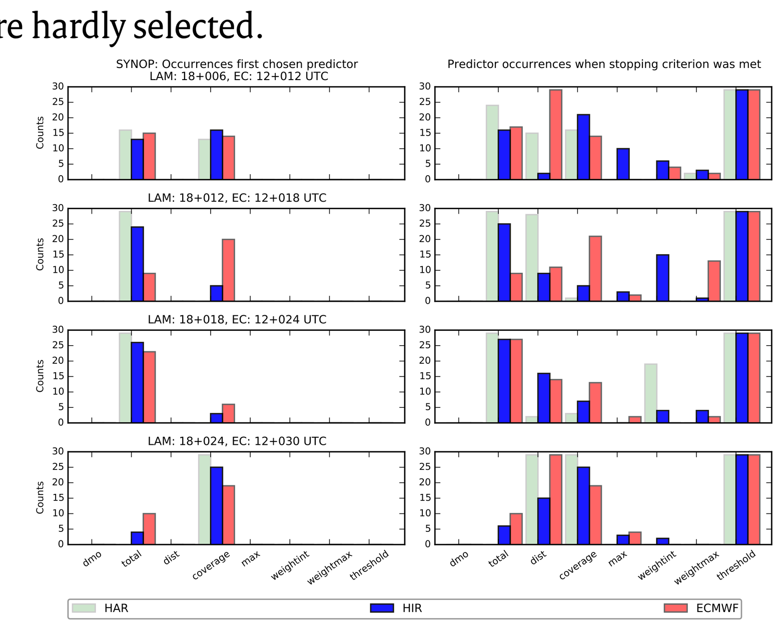

*Figure 2: The distribution of the predictors (over the 30 different training periods; [2]) that were selected first (left) and the predictors that were selected when the stopping criterion was met (right) for the SYNOP observations as predictand source, for the lead times 18+006 to 18+024 UTC (limited area models) and 12+012 to 12+030 UTC (ECMWF) from top to bottom.*

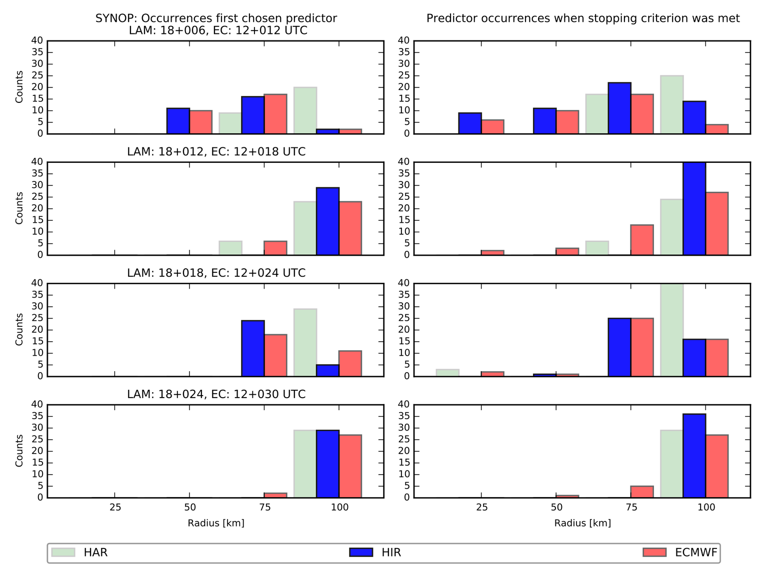

*Figure 3: As Figure 2 but for the distribution of the scales of the selected predictors.*

### 3.2 Verification

Overall the Brier skill scores (BSS) for the 3 post-processed precipitation forecasts are similar (Fig. 4), but larger differences are found for individual lead times [2].

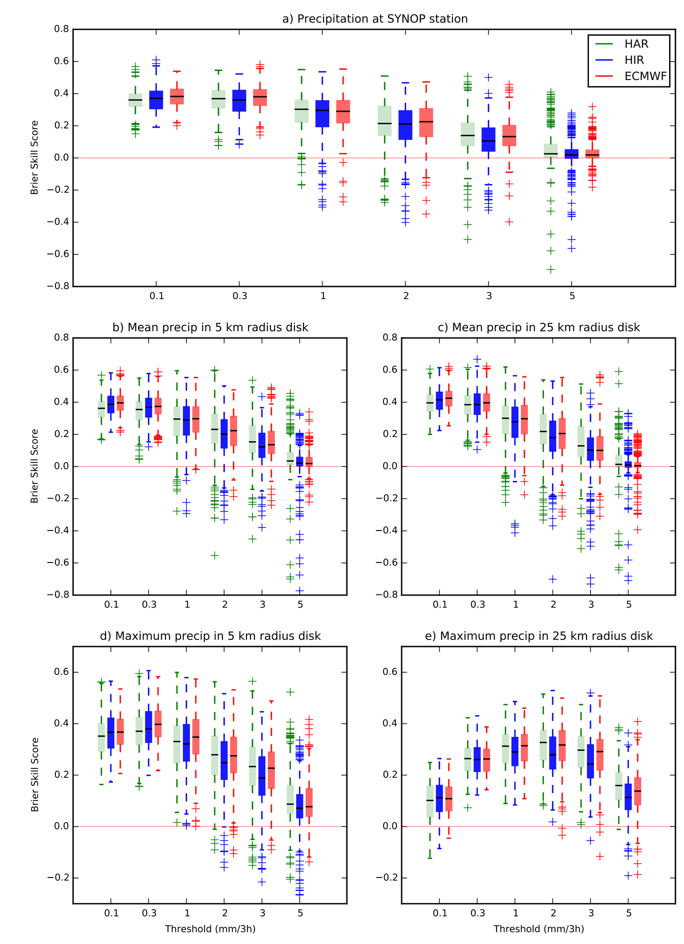

*Figure 4: BSS for the three post-processed NWP models for three predictand types: SYNOP observations (top), radar mean precipitation in circles with a radius of 5 and 25 km (middle) and radar maximum precipitation within circles with a radius of 5 and 25 km (bottom), averaged over all analysis and lead times, as a function of the threshold, and tested on independent data. Boxes represent the 25 to 75 percentile range or IQR with a line indicating the median. The whiskers (dashed lines) represent the 25th percentile minus 1.5 times the IQR and the 75th percentile plus 1.5 times the IQR, extending to the most extreme data point in this data range. The plus signs indicate data outside these ranges.*

## 4 Conclusion

When looking at the capability to estimate the probability of (the mean or maximum) precipitation exceeding some threshold in a point (an area), overall there appears to be only relatively small differences in the BSS between post-processed Harmonie, Hirlam and ECMWF model precipitation output using ELR, but the differences are larger for individual lead times.

## References

[1] C.J. Kok, B.G.J. Wichers Schreur, and D.H.P. Vogelezang. Valuing information from meso-scale forecasts. *Meteorological Applications*, 15(1):103–111, 2008.

[2] E.V. van der Plas, M.J. Schmeits, N. Hooijman, and C.J. Kok. A comparative verification of high resolution precipitation forecasts using model output statistics. *Monthly Weather Review*, submitted, 2017.

[3] D.S. Wilks. Extending logistic regression to provide full-probability-distribution MOS forecasts. *Meteorological Applications*, 16:361–368, 2009.

[1] Royal Netherlands Meteorological Institute (KNMI), PO Box 201, 3730 AE De Bilt, The Netherlands. Contact: schmeits@knmi.nl or plas@knmi.nl.