



## Supplement of

## Seasonal streamflow forecasts for Europe – Part 2: Sources of skill

Wouter Greuell et al.

Correspondence to: Wouter Greuell (wouter.greuell@wur.nl)

The copyright of individual parts of the supplement might differ from the CC BY 3.0 License.



Figure S1 This figure supplements Fig. 4 by comparing the skill in runoff for the InitSH with that for the FullSH for verification with real observations and at three different lead times (months 0, 2 and 4). Figure S1a and b present actual skill for large and small basins, respectively. Large and small basins were defined on the basis of the observations of discharge, which were acquired from the Global Runoff Data Centre, 56068 Koblenz, Germany (GRDC) and gridded onto the 0.5° x 0.5° model grid. Large basins are catchments upwards from the monitoring station larger than 9900 km<sup>2</sup> and small basins are catchments upwards from the station with an area smaller than that of the corresponding grid cell.



Figure S2: This figure supplements Fig. S1a by comparing actual skill of discharge at three different lead times (months 0, 2 and 4) for a selection of the large catchments with relatively little human impact.



Figure S3 This figure supplements Fig. 4 by comparing the skill of the InitSH with that of the FullSH in terms of the percentage of cells with a significant Ranked Probability Skill Score (RPSS).



Figure S4 This figure supplements Fig. 4 by comparing the skill of the InitSH with that of the FullSH in terms of the percentage of cells with a significant ROC area for the BN tercile.



Figure S5 This figure supplements Fig. 4 by comparing the skill of the InitSH with that of the FullSH in terms of the percentage of cells with a significant ROC area for the AN tercile.



Figure S6 This figure supplements Fig. 4 by comparing the skill of the InitSH with that of the FullSH in terms of the domain-mean of R.



Figure S7 Like Fig. 5 this figure compares the annual cycles of the skill in the runoff hindcasts of the four experimental hindcasts for lead months 0 and 2. In Fig. S7a skill is measured in terms of the fraction of grid cells with a significant R. The other panels are similar but for the RPSS (Fig. S7b), for the ROC area of the AN tercile (Fig. S7c) and the ROC area of the BN tercile (Fig. S7d).



Figure S8 Comparison of the annual cycles of skill of the FullSH (red), the InitSH (blue) and the ESP (green) for three different lead times. Where blue or green symbols seem to be missing, they coincide.



Figure S9 This figure complements Fig. A1 but here, instead of pseudo-observations, real observations for large basins were used for verification. See Figure A1 for more explanation.