

EFAS Evaluation Layers



Shaun Harriagn (shaun.harrigan@ecmwf.int)

with contributions from the CEMS-Flood Computational team @ECMWF

EFAS Annual Meeting | 27 October 2021



[About](#)[Products](#)[Data access](#)[Collaborate](#)[News and events](#)[Resources](#)[Training](#)[Notifications](#)[EFAS Wiki](#)

European Flood Awareness System

<https://www.efas.eu>

Latest events



CEMS Week 2021

Access the map viewer



Live map

Latest news

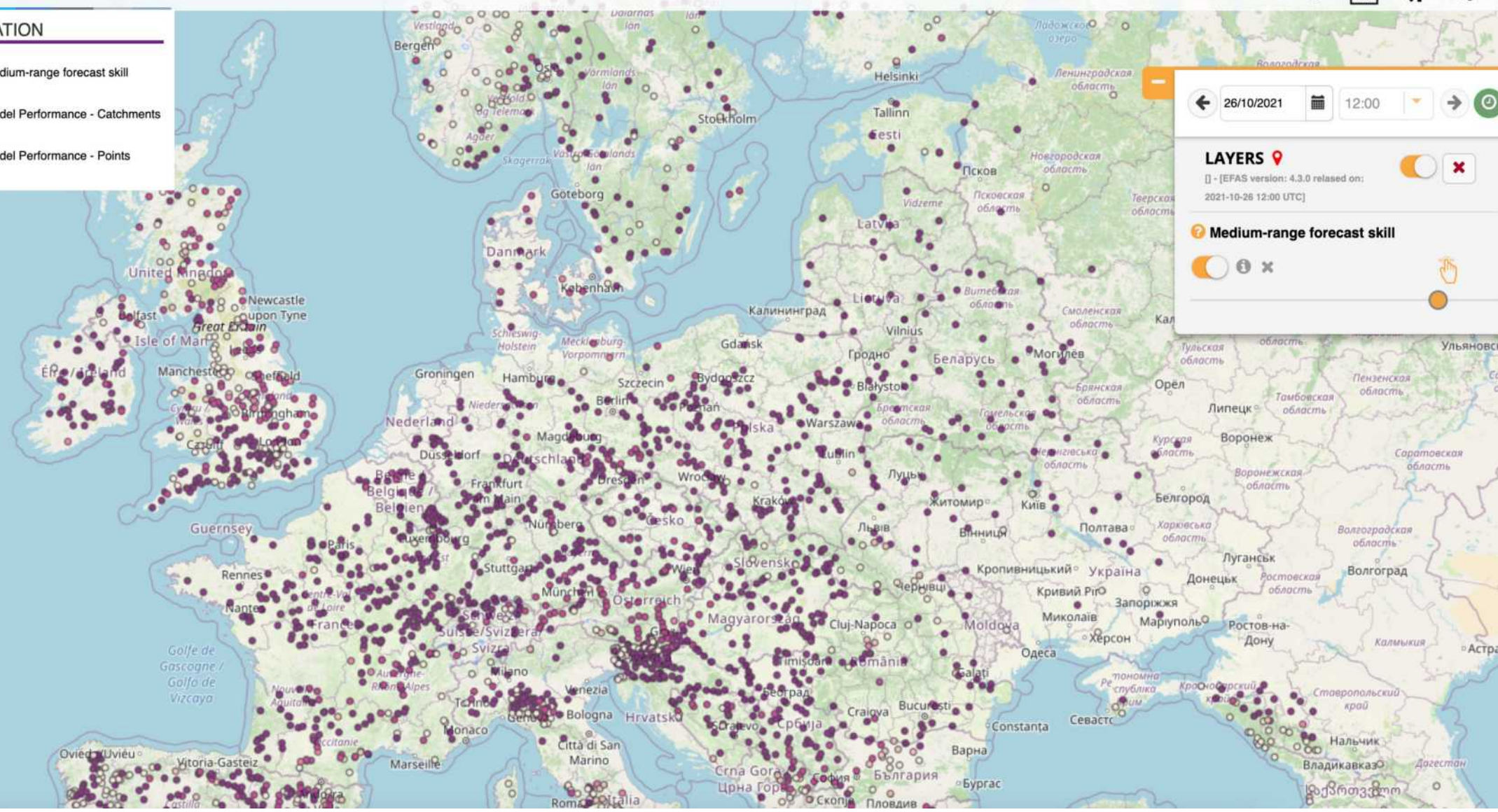


Release of New EFAS version 4.3



EVALUATION

- Medium-range forecast skill
- Model Performance - Catchments
- Model Performance - Points



26/10/2021 12:00

LAYERS

[] - [EFAS version: 4.3.0 released on: 2021-10-26 12:00 UTC]

Medium-range forecast skill

Search, zoom in (+), zoom out (-), and scale (0-300km) controls.



Feedback on layers via Padlet: http://tiny.cc/efas_padlet_1



padlet



Shaun Harrigan • 1m

EFAS Evaluation Layer Feedback

Session 1

Shaun Harrigan 3m
1a) Hydrological model performance - station points map



'Model performance - Points' layer on EFAS showing overall modified Kling Gupta Efficiency (KGE) per station (Darker blue is better, grey worse performance).

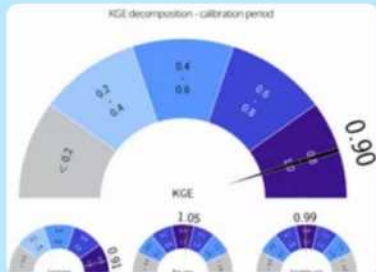
Please "Add comment" below to give your feedback

☆ Rate 1

Shaun Harrigan 17h
Please see EFAS Wiki for 'Hydrological model performance' documentation:
<https://confluence.ecmwf.int/display/COPSRV/EFAS+hydrological+model+performance>

Add comment

Shaun Harrigan 2m
1b) Hydrological model performance - KGE 'speedometer' plots



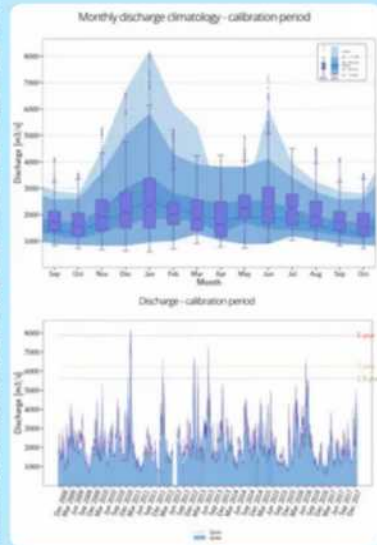
KGE' decomposition speedometer plot for the Rhine at Lobith (Darker blue is better, grey worse performance). Optimum value is 1 for each component.

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment

Shaun Harrigan 10h
1c) Hydrological model performance - Discharge time-series plots



Observed (Qobs) and simulated (Qsim) river discharge time-series for monthly average (top) and full available series (bottom) for the Rhine at Lobith

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment

Shaun Harrigan 1m
2a) Medium-range forecast skill - station points map



'Medium-range forecast skill' layer on EFAS showing the headline score (CRPSS, Continuous Ranked Probability Skill Score). Darker purple shows higher skill.

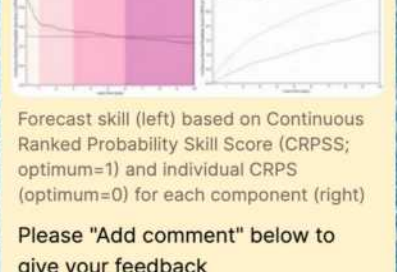
Please "Add comment" below to give your feedback

☆ Rate 1

Shaun Harrigan 10h
Please see EFAS Wiki for 'Medium-range forecast skill' documentation:
<https://confluence.ecmwf.int/display/COPSRV/EFAS+medium-range+forecast+skill>

Add comment

Shaun Harrigan 10h
2b) Medium-range forecast skill - CRPSS and CRPS plots



Forecast skill (left) based on Continuous Ranked Probability Skill Score (CRPSS; optimum=1) and individual CRPS (optimum=0) for each component (right)

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment

Shaun Harrigan 1m
General ideas for future EFAS evaluation layers



Please "Add comment" below to share your ideas

☆ Rate 0

Add comment





Feedback on layers via Padlet: http://tiny.cc/efas_padlet_2



padlet



Shaun Harrigan • 1m

EFAS Evaluation Layer Feedback 2

Session 2

Shaun Harrigan 1m
1a) Hydrological model performance - station points map



'Model performance - Points' layer on EFAS showing overall modified Kling Gupta Efficiency (KGE) per station (Darker blue is better, grey worse performance).

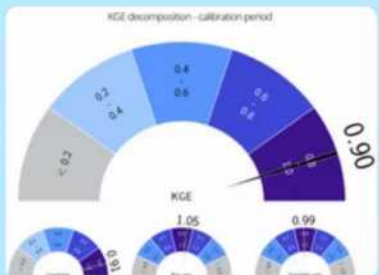
Please "Add comment" below to give your feedback

☆ Rate 1

Shaun Harrigan 1m
Please see EFAS Wiki for 'Hydrological model performance' documentation:
<https://confluence.ecmwf.int/display/COPSRV/EFAS+hydrological+model+performance>

Add comment

Shaun Harrigan 1m
1b) Hydrological model performance - KGE 'speedometer' plots



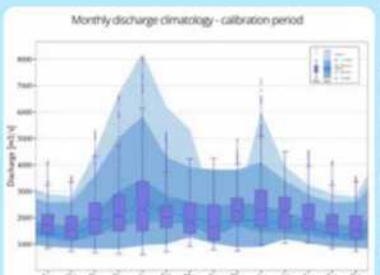
KGE' decomposition speedometer plot for the Rhine at Lobith (Darker blue is better, grey worse performance). Optimum value is 1 for each component.

Please "Add comment" below to give your feedback

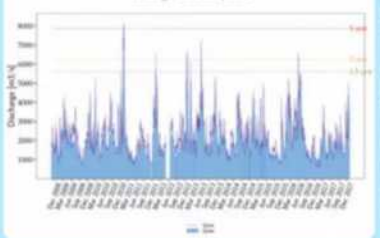
☆ Rate 0

Add comment

Shaun Harrigan 1m
1c) Hydrological model performance - Discharge time-series plots



Observed (Qobs) and simulated (Qsim) river discharge time-series for monthly average (top) and full available series (bottom) for the Rhine at Lobith



Observed (Qobs) and simulated (Qsim) river discharge time-series for monthly average (top) and full available series (bottom) for the Rhine at Lobith

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment

Shaun Harrigan 1m
2a) Medium-range forecast skill - station points map



'Medium-range forecast skill' layer on EFAS showing the headline score (CRPSS, Continuous Ranked Probability Skill Score). Darker purple shows higher skill.

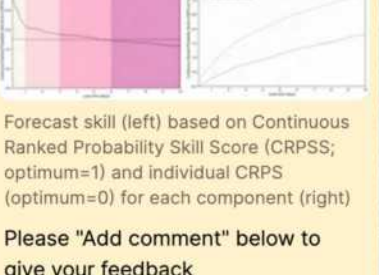
Please "Add comment" below to give your feedback

☆ Rate 1

Shaun Harrigan 1m
Please see EFAS Wiki for 'Medium-range forecast skill' documentation:
<https://confluence.ecmwf.int/display/COPSRV/EFAS+medium-range+forecast+skill>

Add comment

Shaun Harrigan 1m
2b) Medium-range forecast skill - CRPSS and CRPS plots



Forecast skill (left) based on Continuous Ranked Probability Skill Score (CRPSS; optimum=1) and individual CRPS (optimum=0) for each component (right)

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment

Shaun Harrigan 1m
General ideas for future EFAS evaluation layers



Please "Add comment" below to share your ideas

☆ Rate 0

Add comment



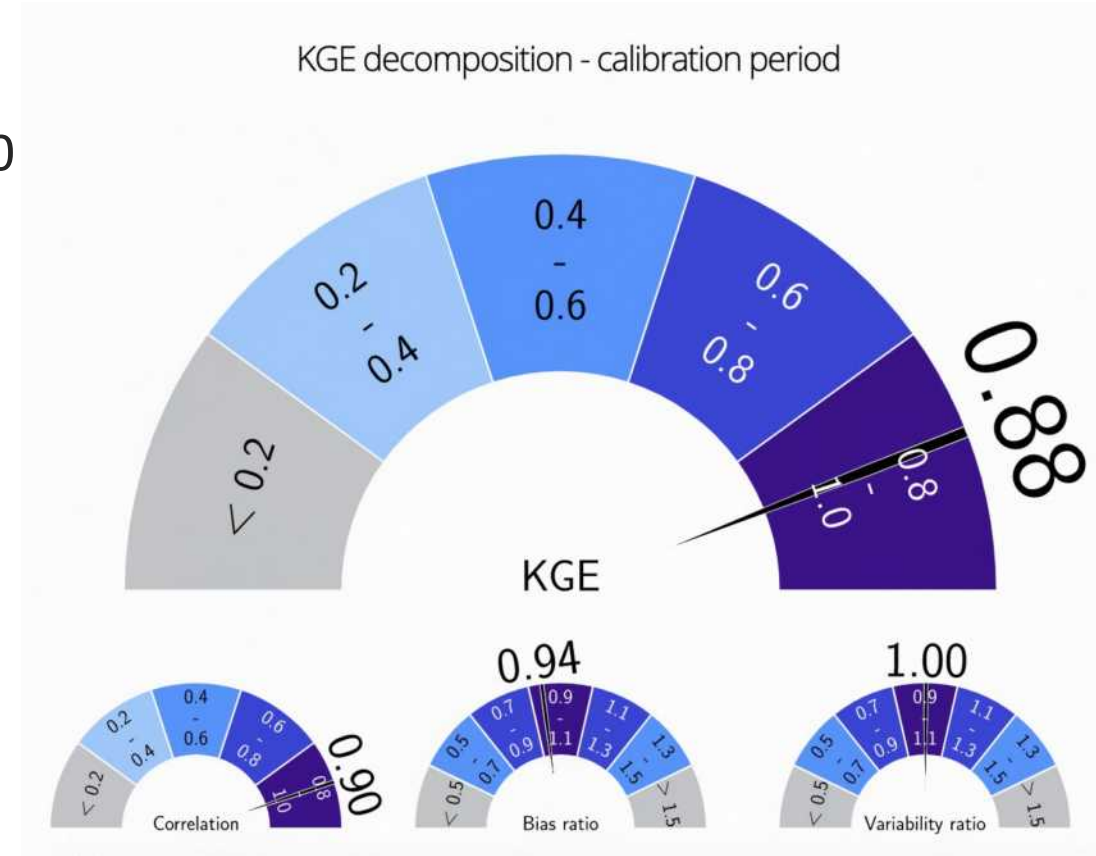


Part 1: Hydrological model performance



Method overview

- Evaluation for 1137 observation stations from 1990 to 2017 with at least 4 years of data
- **Metric:** Modified Kling-Gupta Efficiency (KGE'), decomposed into:
 - Pearson correlation
 - Bias ratio
 - Variability ratio
- Evaluated at 6-hourly time steps, but model series averaged to daily for stations with only daily data



KGE 'speedometer' plots



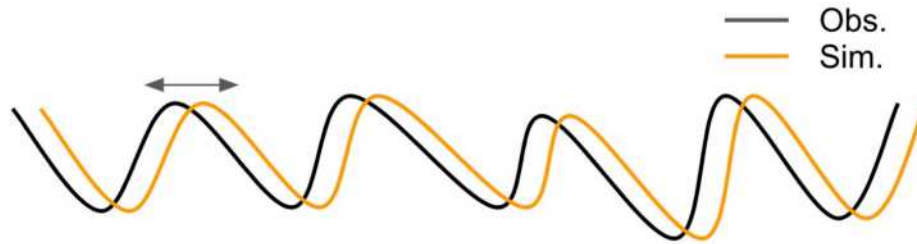
Model performance metric

modified Kling-Gupta Efficiency metric (KGE')
(Gupta et al., 2009; Kling et al., 2012)

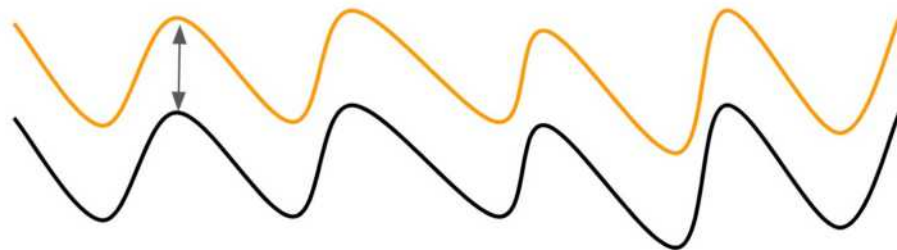
$$KGE' = 1 - \sqrt{(r - 1)^2 + (\beta - 1)^2 + (\gamma - 1)^2}$$
$$r = \frac{cov_{so}}{\sigma_s \cdot \sigma_o}, \quad \beta = \frac{\mu_s}{\mu_o}, \quad \gamma = \frac{\sigma_s / \mu_s}{\sigma_o / \mu_o}$$

Where, μ = mean & σ = standard deviation
KGE' Range: $-\infty$ to 1 (perfect)

Correlation
(timing, e.g. peaks too late)



Bias ratio
(e.g. positive bias)



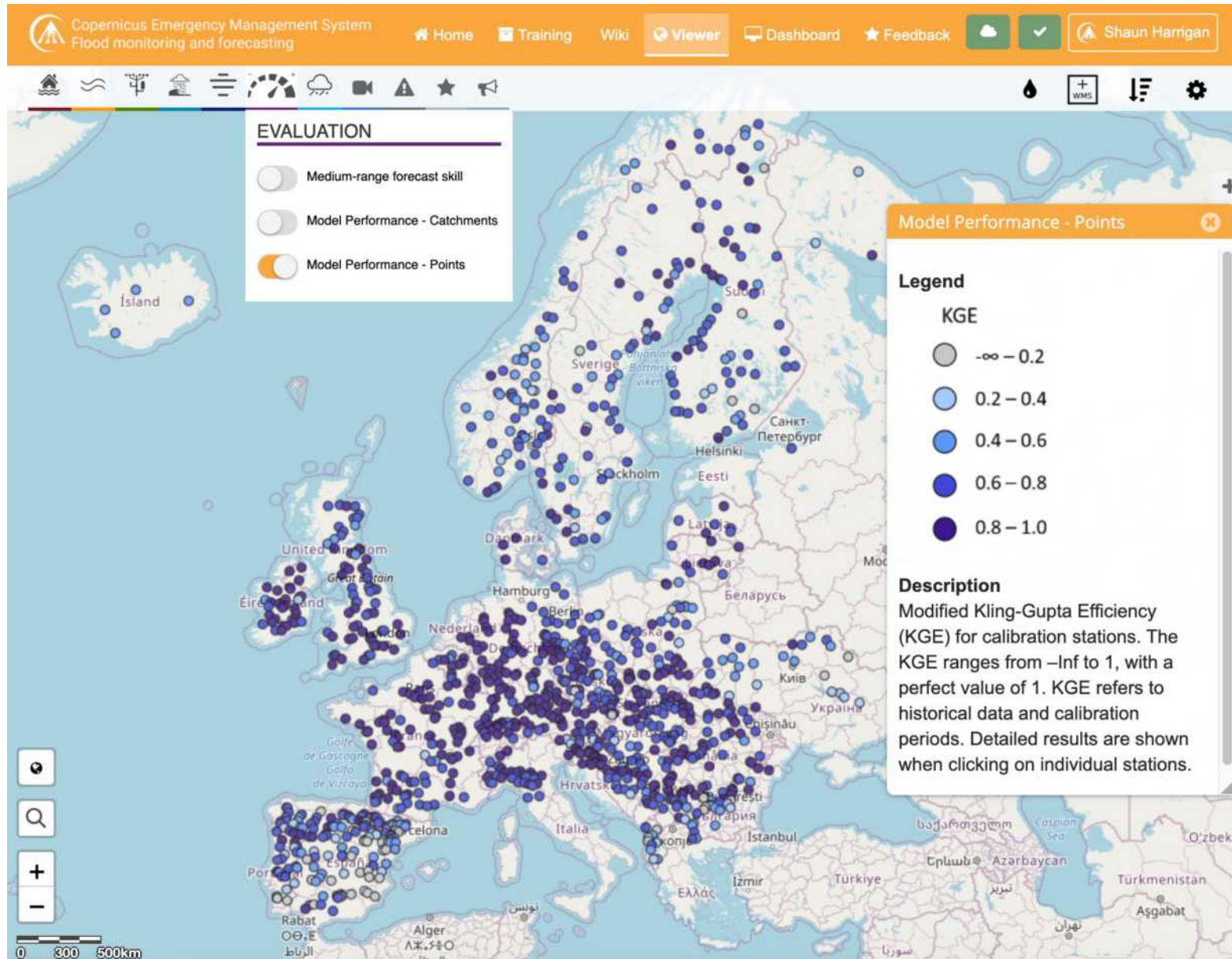
Variability ratio
(e.g. variability too low)





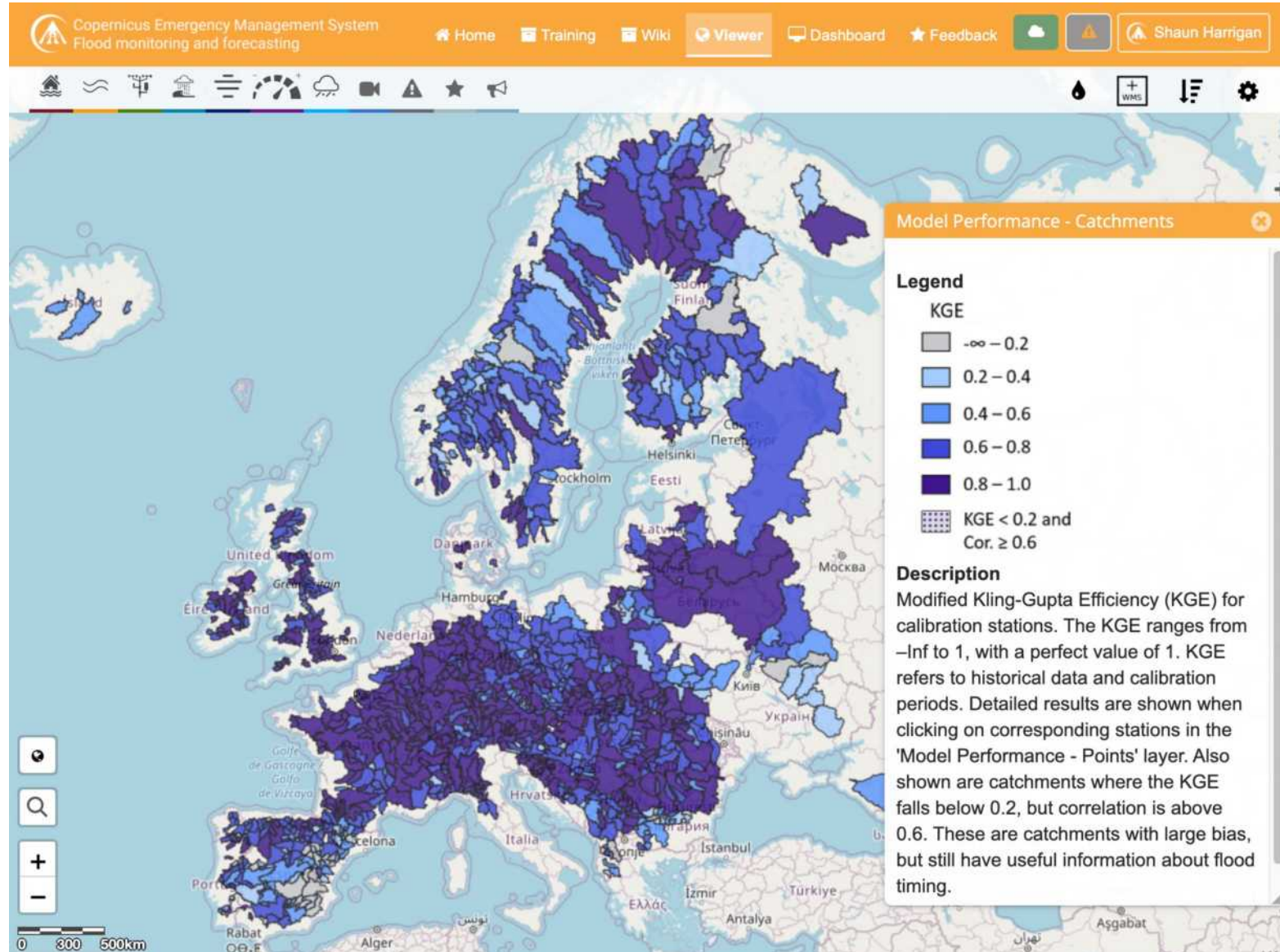
Model performance – Points

Emergency Management



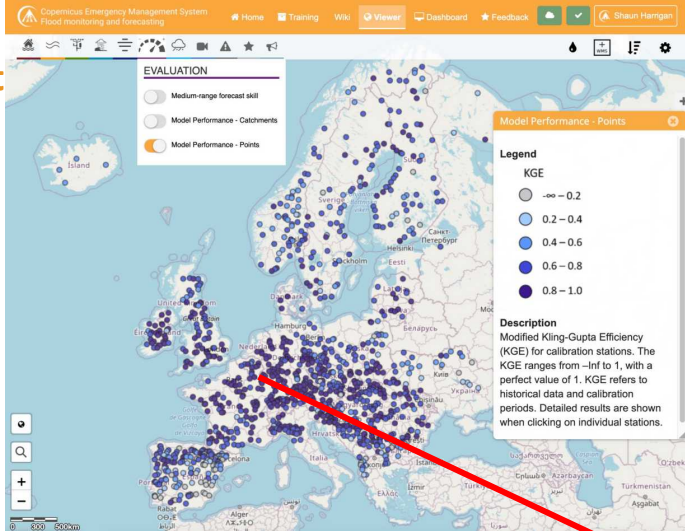


Model performance – Catchments

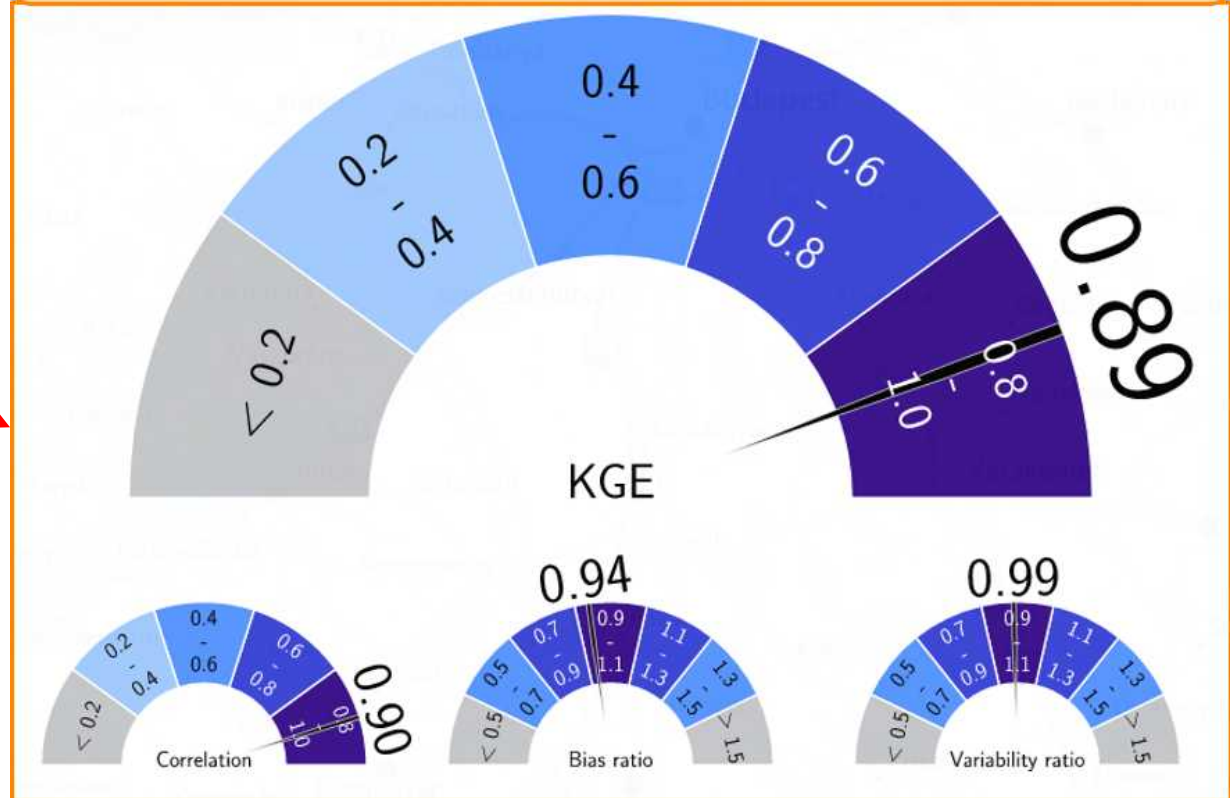




Station-based KGE 'speedometer' plots



Model performance score. Diagram showing the hydrological modelling performance score KGE' and its three components



KGE' score over the observational record length

Ideal value for KGE' and its components is 1

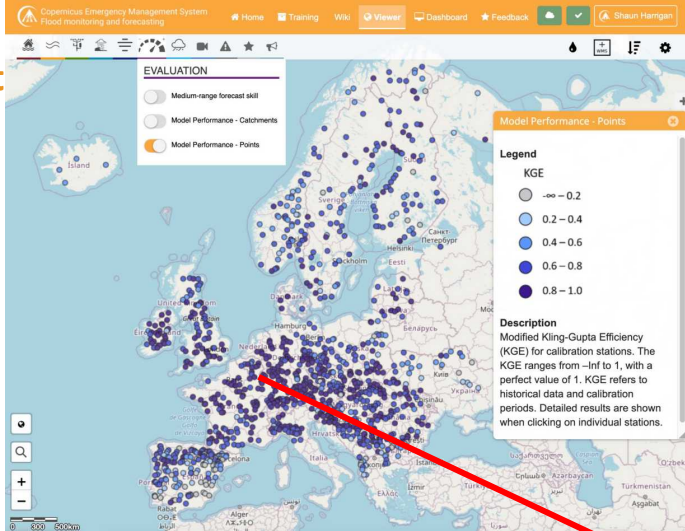
Correlation value informing on timing errors

Bias ratio informing on water balance errors

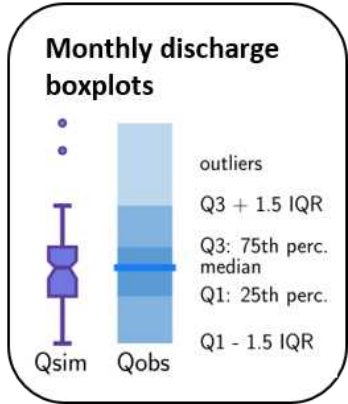
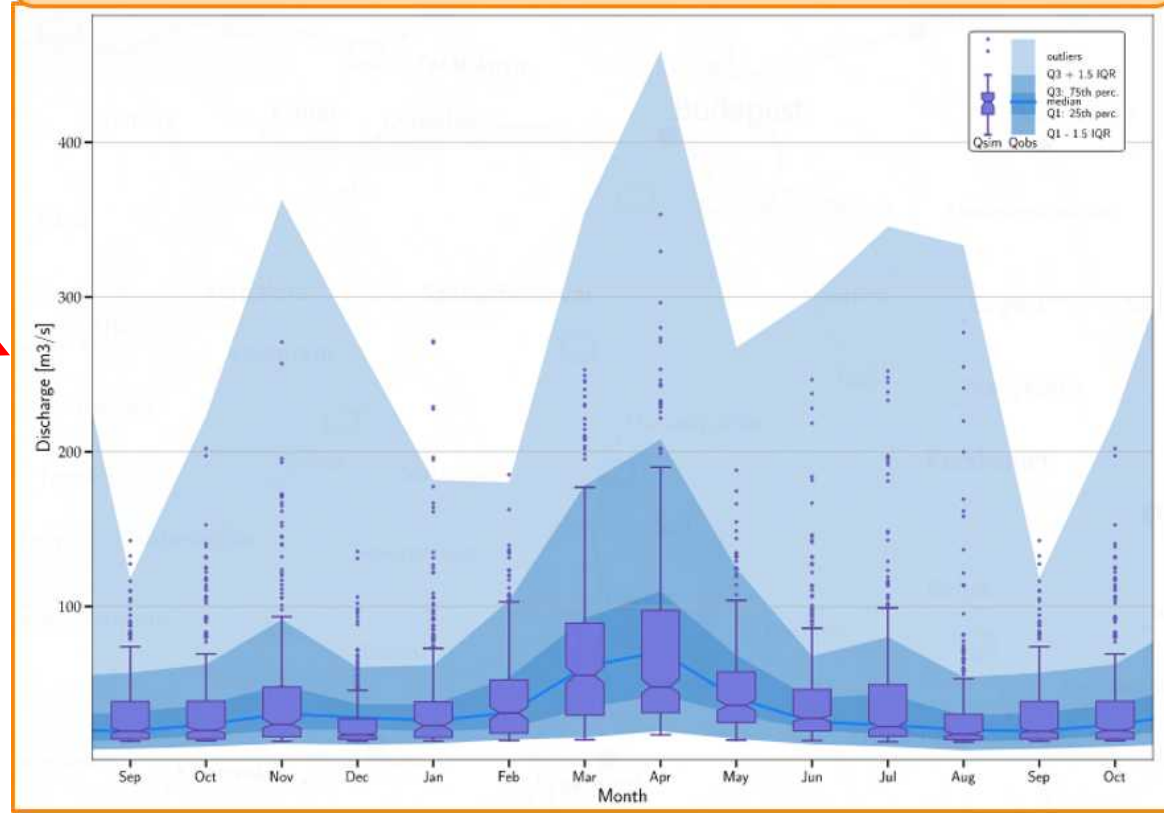
Variability ratio informing on variability errors



Station-based monthly discharge climatology plot



Monthly discharge climatology. Time series plot showing the observed and simulated discharge as monthly statistics calculated over the observational record length

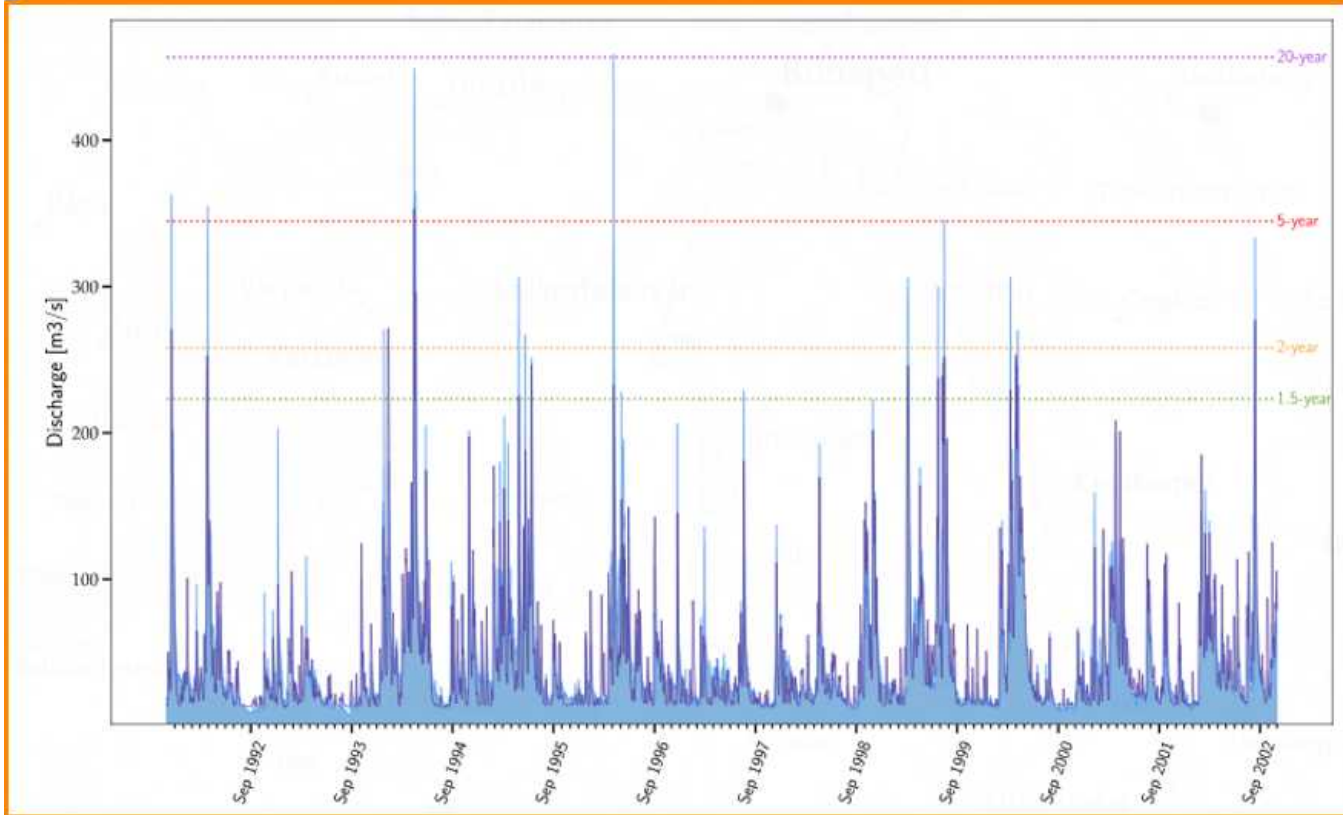


Hydrological year shown as 14-month starting in September



Station-based discharge time-series

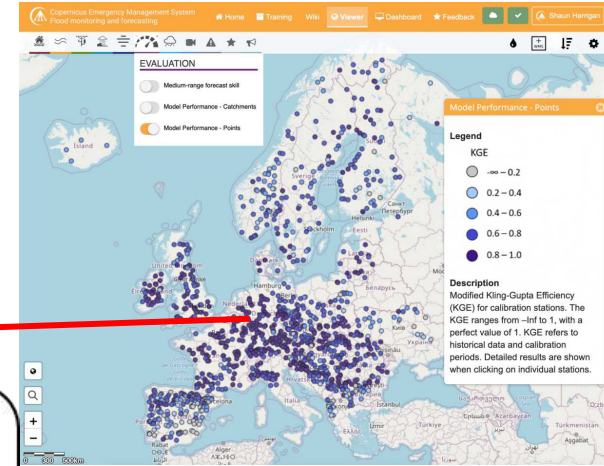
Simulated and modelled hydrograph. Time series plot showing the observed and simulated discharge over the calibration period



Hydrographs shown over the full calibration periods

Severity levels
Return periods (green to purple colours) derived from the model climatology

Hydrographs
— Qsim
■ Qobs





Feedback on Hydrological model performance products (1a, 1b, 1c)

http://tiny.cc/efas_padlet_1

The screenshot shows a Padlet board titled "EFAS Evaluation Layer Feedback Session 1" by Shaun Harrigan. It contains three posts:

- 1a) Hydrological model performance - station points map:** A map of Europe showing station points. Below it, text asks for feedback and includes a link to EFAS Wiki documentation: <https://confluence.ecmwf.int/display/COPSRV/EFAS+hydrological+model+performance>.
- 1b) Hydrological model performance - KGE 'speedometer' plots:** A speedometer plot for the KGE decomposition (bias, accuracy, reliability) for the Rhine at Lobith. The KGE value is 0.98.
- 1c) Hydrological model performance - Discharge time-series plots:** Two time-series plots for monthly average and full available series of river discharge for the Rhine at Lobith, comparing observed (Qobs) and simulated (Qsim) data.





Feedback on Hydrological model performance products (1a, 1b, 1c)

http://tiny.cc/efas_padlet_2

The screenshot shows a Padlet board titled "EFAS Evaluation Layer Feedback Session 1" by Shaun Harrigan. It contains three posts:

- 1a) Hydrological model performance - station points map:** A map of Europe showing station points. Below it, text asks for feedback and includes a link to EFAS Wiki documentation: <https://confluence.ecmwf.int/display/COPSRV/EFAS+hydrological+model+performance>.
- 1b) Hydrological model performance - KGE 'speedometer' plots:** A speedometer plot for the KGE decomposition (bias, accuracy, reliability) for the Rhine at Lobith. The KGE value is 0.98.
- 1c) Hydrological model performance - Discharge time-series plots:** Two time-series plots for the Rhine at Lobith: the top one shows monthly average discharge (simulated vs observed), and the bottom one shows full available series.





Part 2: Forecast skill



Forecast skill and benchmark forecast

Forecast *skill* refers to the relative accuracy of a set of forecasts, with respect to some set of *standard benchmark forecasts* (Wilks, 2011)

Common **benchmark forecasts** in hydrological forecasting (Pappenberger et al., 2015):

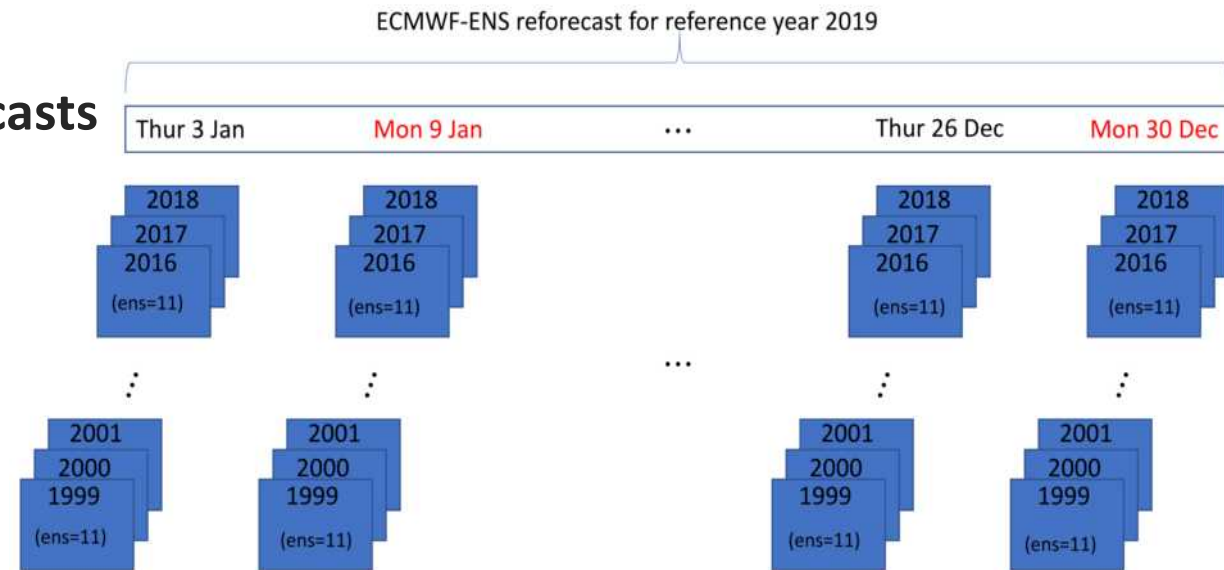
- **Persistence** (Last observed discharge persisted for each lead time):
 - Most common in short to medium-range forecasting
- **Climatology** (Seasonal average of observed based on historic data)
 - Most common in extended- and seasonal-range forecasting



Method overview

- Generation of long-term large-sample **reforecasts**

- ECMWF-ENS for reference year 2019
- 2 per week; 11 ensemble members
- 20-years (1999-2018), 6-hourly



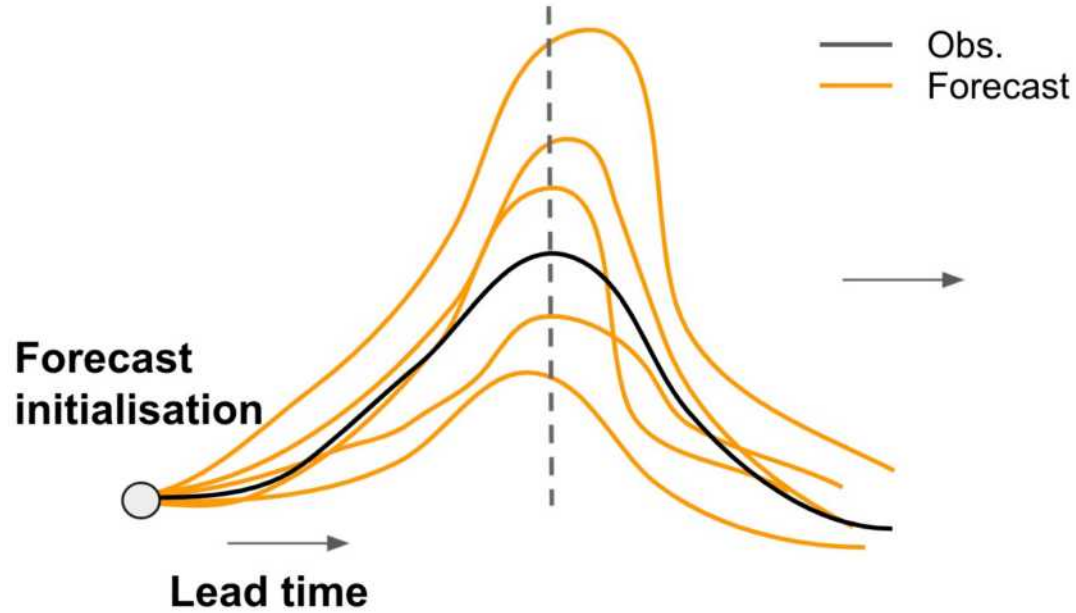
- **Benchmark forecast** = 6 hr river discharge persistence from previous time step

- Evaluated against **proxy observations (EFAS 4 forced simulation (sfo))** at n=2651 fixed reporting points

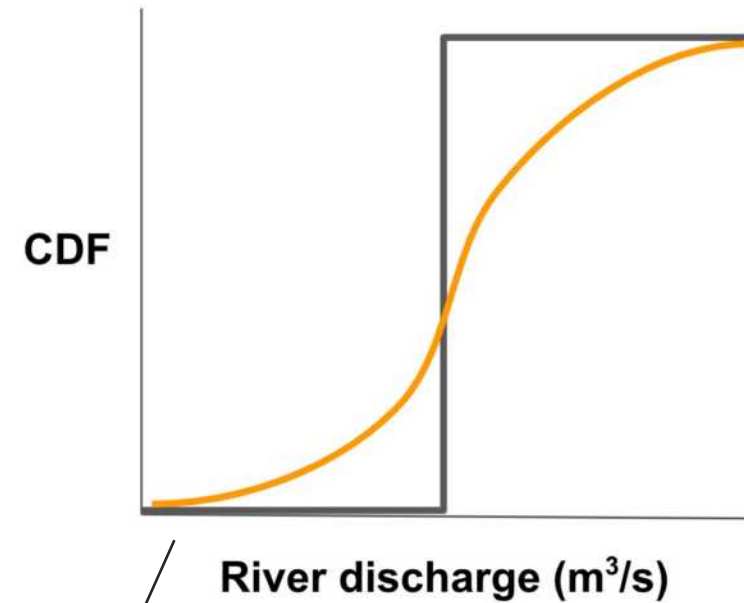


Ensemble forecast skill metric: CRPS & corresponding skill score

Ensemble forecast



Continuous Ranked Probability Score (CRPS)



$$\text{CRPSS} = 1 - \frac{\overline{\text{CRPS}_{fc}}}{\overline{\text{CRPS}_{bench}}}$$

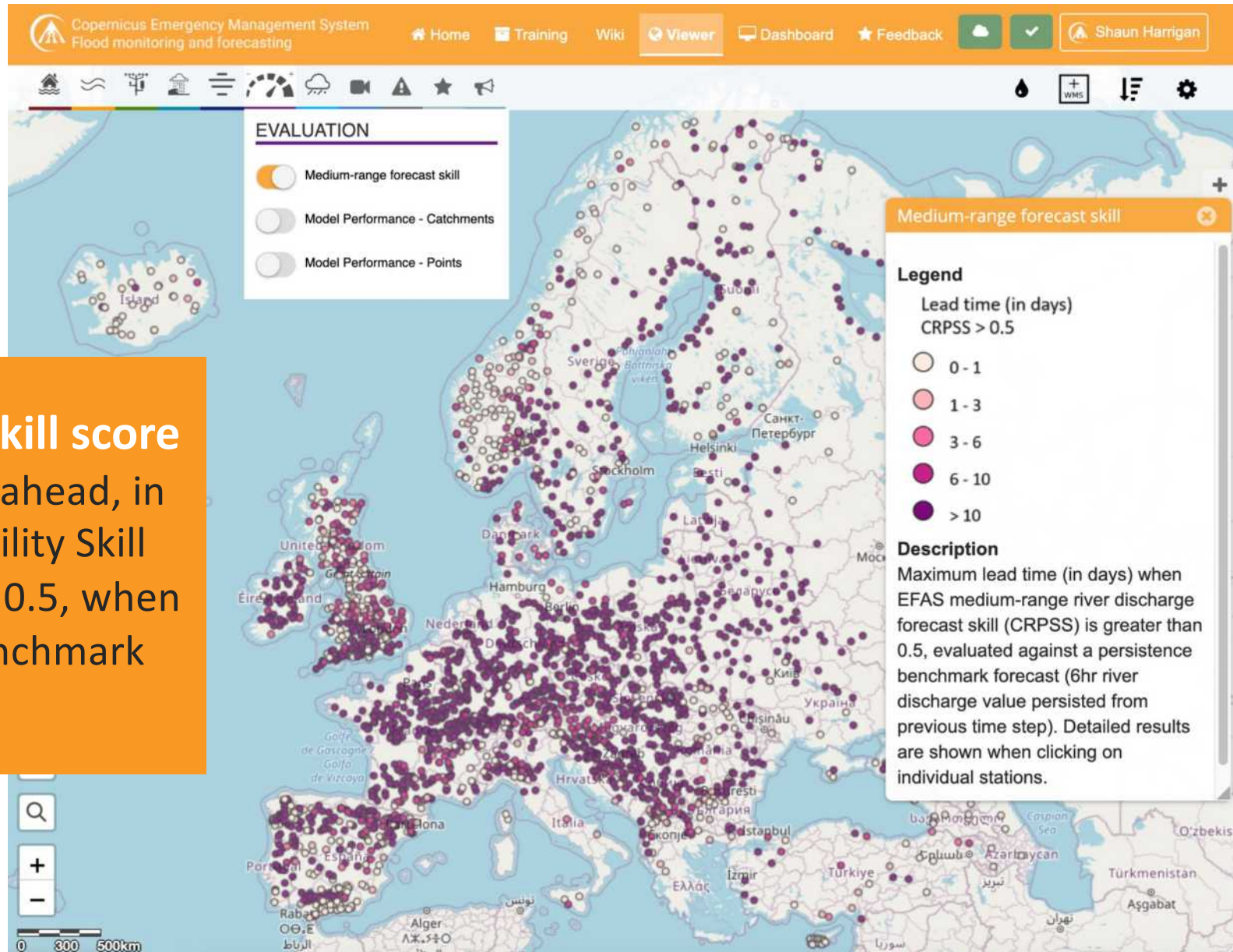
No Skill=0; Perfect skill=1



Medium-range forecast skill (Headline score)

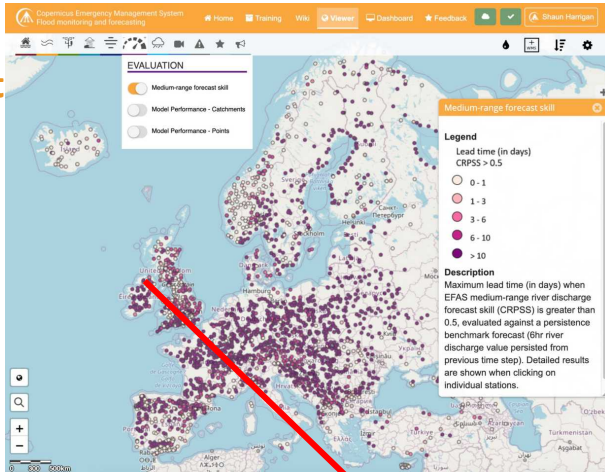
Headline medium-range forecast skill score

Max lead time (in days), up to 10-days ahead, in which the Continuous Ranked Probability Skill Score (CRPSS) is greater than a value of 0.5, when compared to a simple persistence benchmark forecast

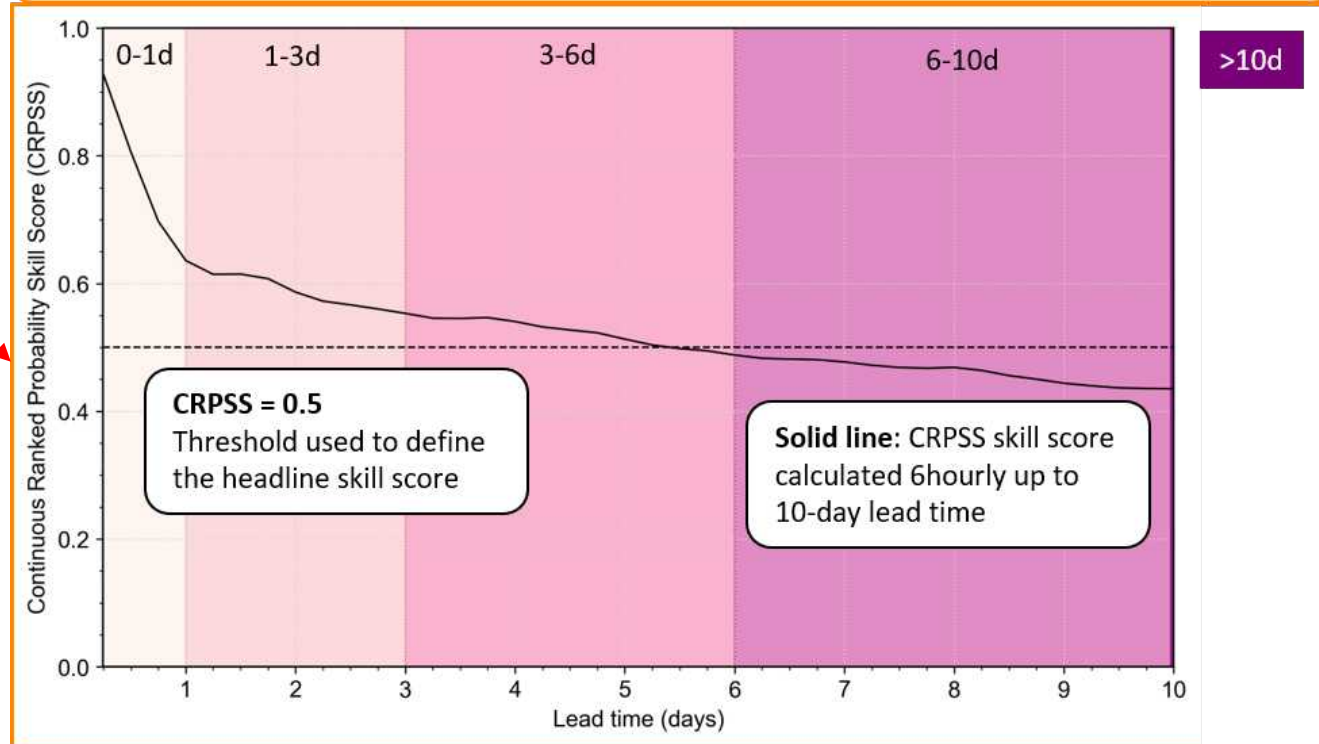




Station-based medium-range forecast skill (CRPSS)



Medium-range forecast skill. Time series plot of CRPSS as a function of the forecast lead time. CRPSS is calculated against a persistence benchmark



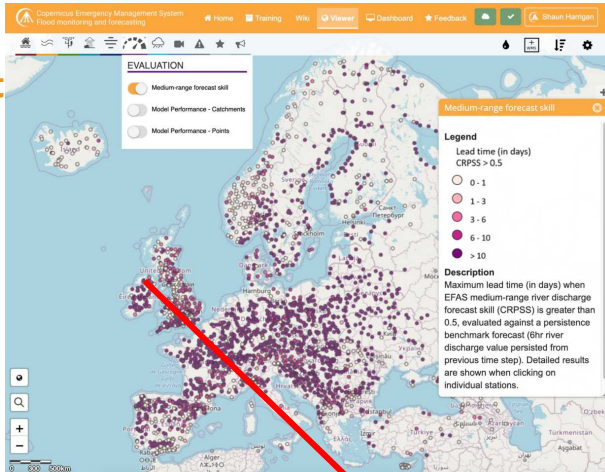
Forecast lead time (days).

Headline score thresholds

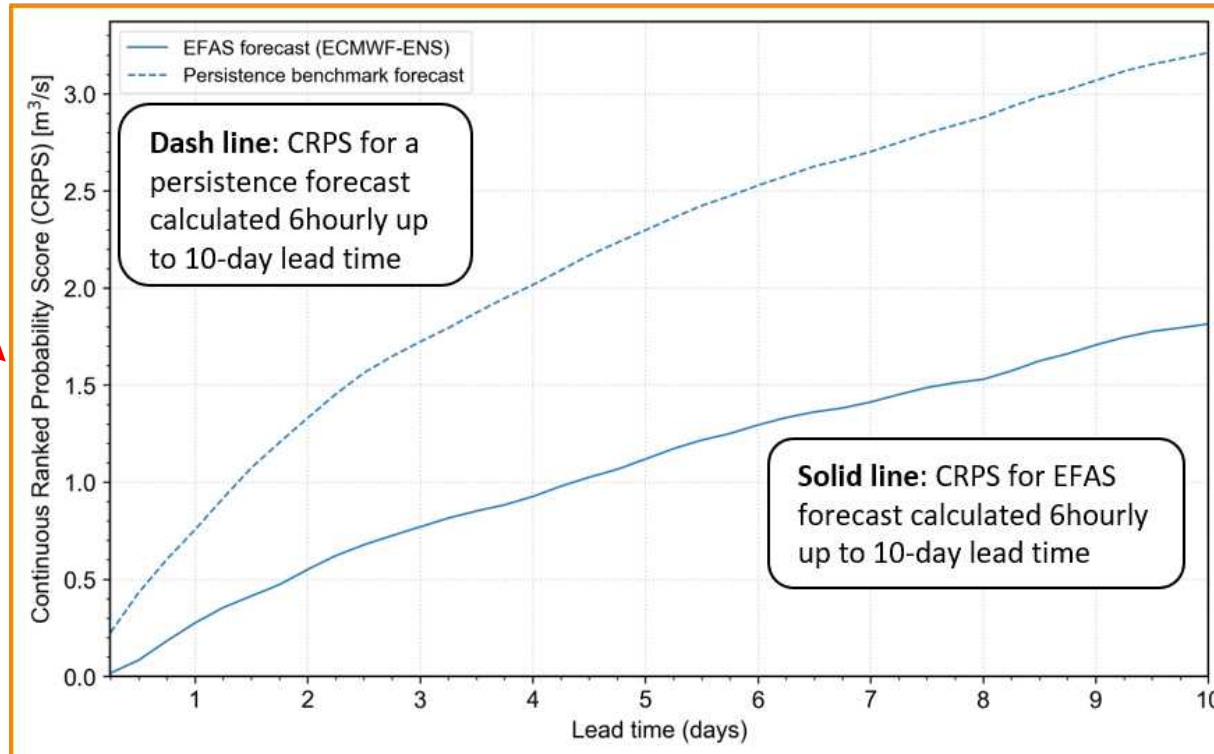
Threshold category (shaded background) defining the maximum lead time (in days) the CRPSS > 0.5 (when CRPSS crosses the 0.5 threshold line; here 3-6d). The higher the CRPSS, the less forecast error compared with the benchmark forecast



Station-based medium-range forecast skill (Individual CRPS)



Medium-range forecast skill. Time series plot of CRPS as a function of the forecast lead time. CRPS is calculated for EFAS forecast (ECMWF-ENS) and a persistence benchmark forecast



Forecast lead time (days).

CRPS
Measure of error between an ensemble forecast and (proxy) observations (here using EFAS forced simulation (sfo)). For deterministic forecasts, such as the persistence benchmark, the CRPS is equivalent to the Mean Absolute Error (MAE). A perfect forecast has a CRPS equal to zero.



Feedback on Forecast skill products (2a & 2b)

http://tiny.cc/efas_padlet_1

Shaun Harrigan 32m

2a) Medium-range forecast skill - station points map

'Medium-range forecast skill' layer on EFAS showing the headline forecast score

Please "Add comment" below to give your feedback

☆ Rate 0

Shaun Harrigan 1m

Please see EFAS Wiki for 'Medium-range forecast skill' documentation:
<https://confluence.ecmwf.int/display/COPSRV/EFAS+medium-range+forecast+skill>

Add comment

Shaun Harrigan 32m

2b) Medium-range forecast skill - CRPS and CRPS plots

Please "Add comment" below to give your feedback

☆ Rate 0

Add comment





Feedback on Forecast skill products (2a & 2b)

http://tiny.cc/efas_padlet_2

The image shows a screenshot of a Padlet board with a blue ocean background. It contains two posts by Shaun Harrigan, posted 32 minutes ago.

Post 1: 2a) Medium-range forecast skill - station points map
The post features a map of Europe with numerous purple dots representing station points. A legend on the right indicates lead times from 1 to 20 days. Below the map, the text reads: "'Medium-range forecast skill' layer on EFAS showing the headline forecast score". It includes a "Rate" button (0) and an "Add comment" button. A comment from Shaun Harrigan (1m) provides a link to EFAS Wiki documentation: <https://confluence.ecmwf.int/display/COPSRV/EFAS+medium-range+forecast+skill>.

Post 2: 2b) Medium-range forecast skill - CRPSS and CRPS plots
The post features two line graphs. The left graph is titled "Forecast skill (CRPSS)" and the right is "Individual CRPS (ECMWF, GFS, and GEM) plots". Below the graphs, the text reads: "Please 'Add comment' below to give your feedback". It includes a "Rate" button (0) and an "Add comment" button.





- Go-to for extended documentation
- Methods, products, operational system
- Information on version upgrades, evaluation etc...

Latest major release: EFAS v4.0

Created by Christel Prudhomme, last modified 14 minutes ago

This is a description of EFAS v4.0. For an overview of other EFAS releases, please see: [EFAS versioning system](#)

Summary

EFAS 4.0 follows a complete upgrade of EFAS hydrological modelling system, with effect on all EFAS flood forecast products except ERICHA nowcasting; it also includes a number of new products and some changes in the way the flood alerts are calculated, as well as some minor changes in the web interface and general bug fixes.

Here is a summary of the main changes:

- Hydrological modelling upgraded to 6hourly time step
- Upgrade of some static fields used in the hydrological modelling
- Recalibration of the hydrological model using 1137 river stations across the EFAS pan-European domain
- Upgrade of the statistical post-processing layer with new calibration and availability for over 2000 stations (more than 1000 daily, and more than 1000 six hourly)
- Forecasts weights for total probability updated (same weights for all forecasts)
- Three new static layers describing some hydrological modelling features (drainage network, location of location of fixed reporting points)
- Upgraded reporting points pop-up window graphics and tables:
 - Plots accommodated to 6-hourly time step and show the initial conditions
 - Modelled discharge hydrograph plots now available (with return period axis)
 - Persistence table showing most recent forecasts on top
- New forecast performance layers (medium-range, monthly and seasonal)
- Improved hydrological modelling skill layer
- Re-organisation of the layers menu for a more intuitive navigation
- One-click icon to access the EFAS wiki

Technical details

Release date	2020-xx-xx 12UTC
In test suite	2020-xx-xx 12

EFAS 4.0 ERIC flash flood forecast skill

Created by Christel Prudhomme, last modified by Calum Baugh on Sep 02, 2020

This page documents the performance of the ERIC forecast methodology obtained with EFAS 4.0.

Evaluation Time Period: 1st January 2019 - 31st December 2019

Observation Datasets:

FloodList.com

- A total of 176 observations where the flood type was recorded as 'flash flood' were extracted from the flood event databa which is populated with information from FloodList.com (see figure below)
 - This flood type was chosen in order to exclude riverine floods
- [/home/mo/moch/Flood_Events/Flash_Floods_2019](#)
- [/floodlist_events_2019-01-01_to_2019-12-31_FlashFloods_Europe.csv](#)



EFAS evaluation

Created by Christel Prudhomme, last modified on Aug 27, 2020

At each new system upgrade, a robust evaluation procedure is applied to all results affected by the changes to assess the performance of the system. When possible, comparison with the previous operational system is also conducted (from EFAS-4). This section summarised the main evaluation procedures applied to EFAS.

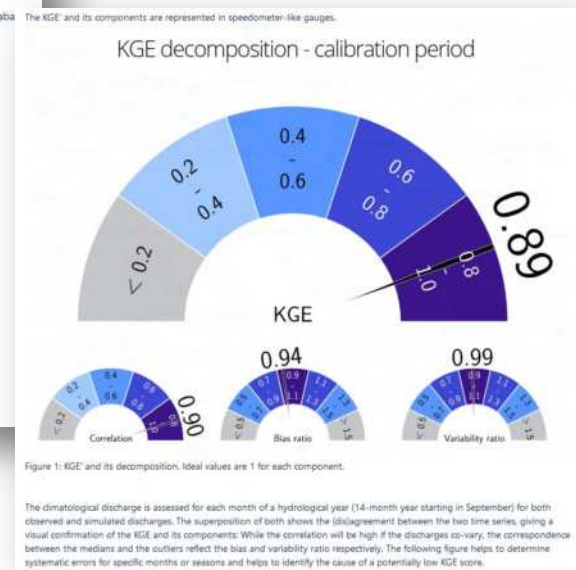
- EFAS hydrological model performance
- EFAS medium-range forecast performance
- ERIC flash flood forecast performance
- EFAS sub-seasonal and seasonal forecast performance
- EFAS seasonal forecast performance

Like Be the first to like this No labels

Write a comment...

Powered by a free Atlassian Confluence Open Source Project License granted to ECMWF. Evaluate Confluence today. This Confluence installation runs a Free Giffy License - Evaluate the Giffy Confluence Plugin for your Wiki!

Powered by Atlassian Confluence 7.3.1 - Report a bug - Atlassian News





Next steps

- Operational monitoring of forecast skill
- Scalability of evaluation products for EFAS version 5
- Padlet open for feedback for one week (3 November)
- Your input will form next iterations of Evaluation tab!

Verification September 2021



Figure 1. EFAS CRPSS at lead-time 1 day for September 2021, for all catchments. The reference score is persistence.



Figure 2. EFAS CRPSS at lead-time 3 days for September 2021, for all catchments.

Thank you!

shaun.harrigan@ecmwf.int

EFAS website:
<https://www.efas.eu/>

CEMS-Flood documentation
wiki:
[https://confluence.ecmwf.int/
display/COPSRV/CEMS-Floods](https://confluence.ecmwf.int/display/COPSRV/CEMS-Floods)

CEMS-Flood Data via CDS:
[https://cds.climate.copernicus.
eu/#!/home](https://cds.climate.copernicus.eu/#!/home)



References

- Gupta, H. V., Kling, H., Yilmaz, K. K., and Martinez, G. F.: Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling, *Journal of Hydrology*, 377, 80–91, <https://doi.org/10.1016/j.jhydrol.2009.08.003>, 2009.
- Kling, H., Fuchs, M., and Paulin, M.: Runoff conditions in the upper Danube basin under an ensemble of climate change scenarios, *Journal of Hydrology*, 424–425, 264–277, <https://doi.org/10.1016/j.jhydrol.2012.01.011>, 2012.
- Pappenberger, F., Ramos, M. H., Cloke, H. L., Wetterhall, F., Alfieri, L., Bogner, K., Mueller, A., and Salamon, P.: How do I know if my forecasts are better? Using benchmarks in hydrological ensemble prediction, *Journal of Hydrology*, 522, 697–713, <https://doi.org/10.1016/j.jhydrol.2015.01.024>, 2015.
- Wilks, D. S.: *Statistical methods in the atmospheric sciences*, Academic press, 2011.