Comparing four radar rainfall nowcasting algorithms for 1533 events

RUBEN IMHOFF[1,2], CLAUDIA BRAUER[1], AART OVEREEM[1,3], ALBRECHT WEERTS[1,2] & REMKO UIJLENHOET[1]

[1]WAGENINGEN UNIVERSITY, THE NETHERLANDS

[2]DELTARES, THE NETHERLANDS

[3]ROYAL NETHERLANDS METEOROLOGICAL INSTITUTE, THE NETHERLANDS

WAGENINGEN UNIVERSITY & RESEARCH

Deltares Enabling Delta Life

The need for very-short-term rainfall predictions

Early warning systems are key to ensure water safety in a changing climate [Pappenberger et al. (2015), Env. Sci. & Policy]

These systems highly depend on accurate and timely Quantitative Precipitation Forecasts (QPF)

For this purpose, NWP is often used, but this is not sufficient on the short term:

- The update frequency of issued forecasts is too low, especially for convective events
- Events are forecast, but not at the right location

Is nowcasting an option?

The New York Times, July 8, 2019

Washington Area Hit by Heavy Rains and Flash Floods

Severe Weather Europe, Aug. 4, 2018

Intense flash floods hit various parts of Europe this week

BBC News, July 10, 2017

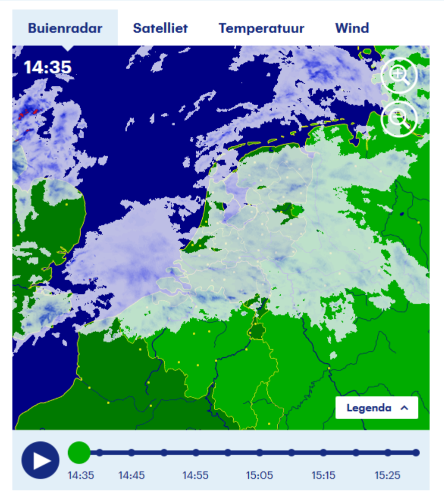
Paris flooding: Record rainfall hits French capital

Alex Brandon / New York Times

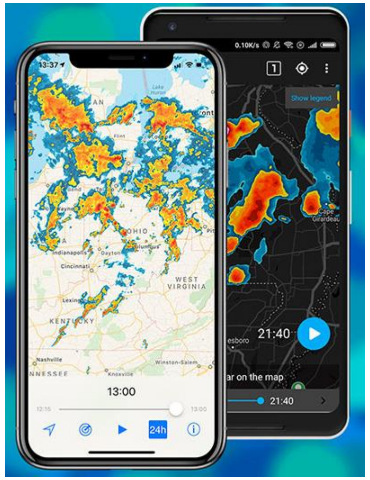
A short introduction to rainfall nowcasting

You probably use it every day!

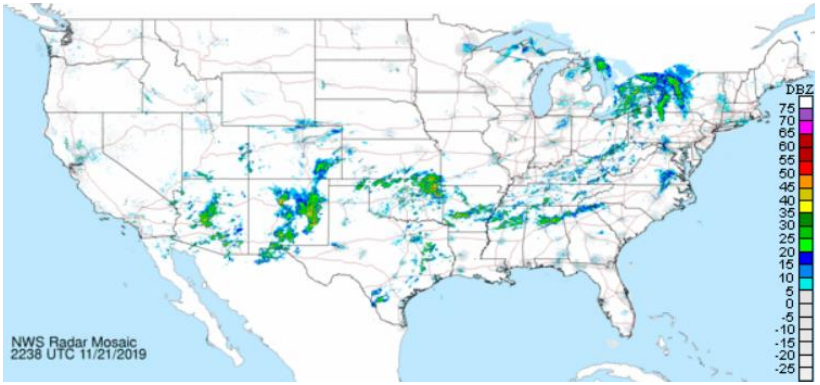
Nowcasts are short-term rainfall forecasts, available up to several hours ahead, based on extrapolation of (radar) rainfall fields and sometimes with the initiation, growth and dissipation of convective cells incorporated.



Source: www.buienradar.nl



Source: www.rainviewer.com



Source: www.weather.gov

Note that in this study, we only focus on cross-correlation based (extrapolation) nowcasting algorithms

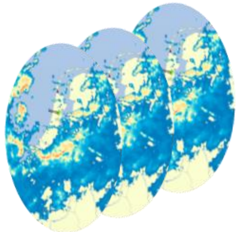
A short introduction to rainfall (extrapolation) nowcasting

Vector field from 2-24 previous radar images

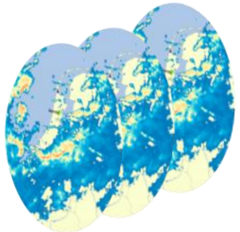
Optional: take the field development into account

Extrapolate rainfall fields to future

Interpolate the results to the original grid structure



[t-n] until [t=0]



[t+1] until [t+n]

Statistical extrapolation and processes

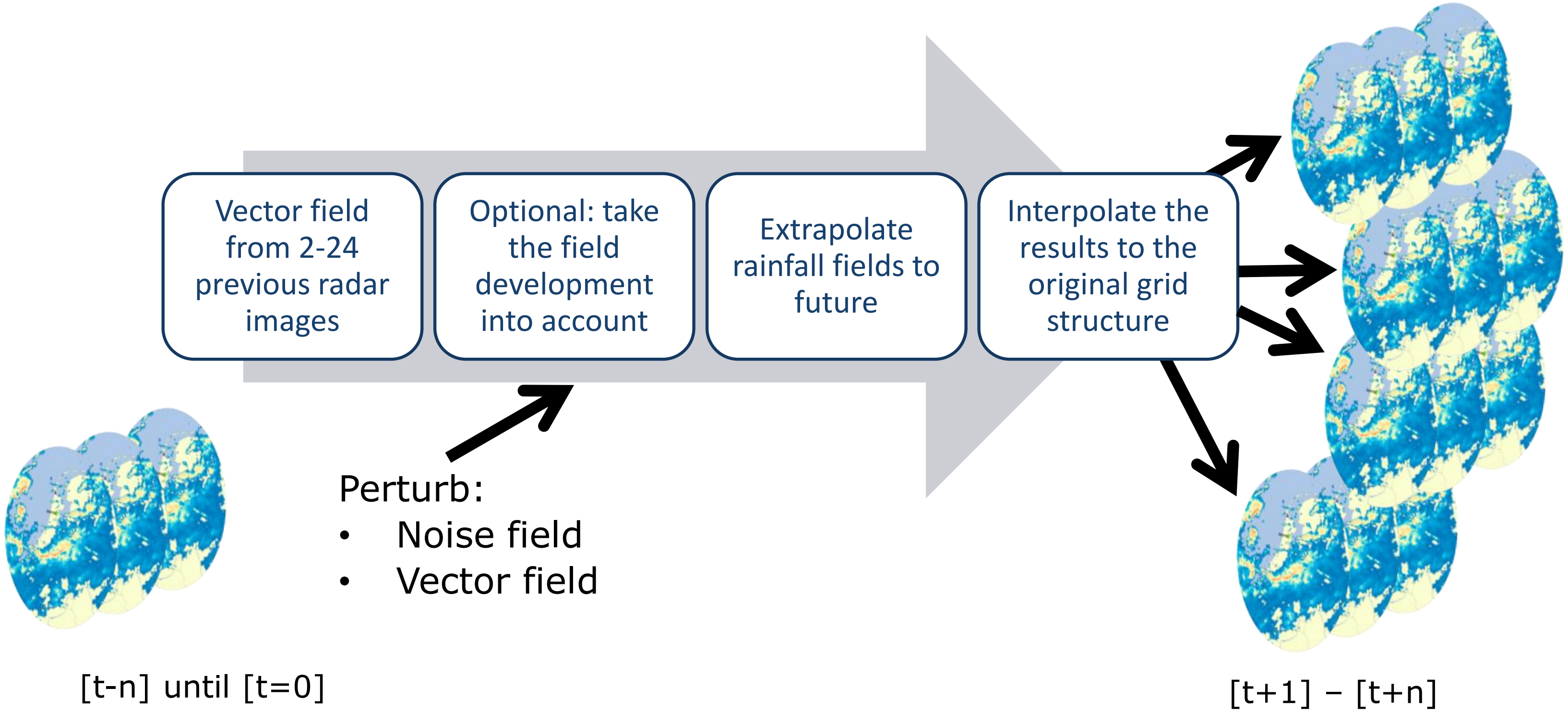
Rapid forecasting (e.g. new nowcasts available every 5 min.)

Absence of physical growth/dissipation processes

Deterministic nowcasting

[e.g. Seed, J. Appl. Meteorol., 2003; Ayzel et al., Geosci. Model. Dev., 2019]

A short introduction to rainfall (extrapolation) nowcasting



Statistical extrapolation and processes

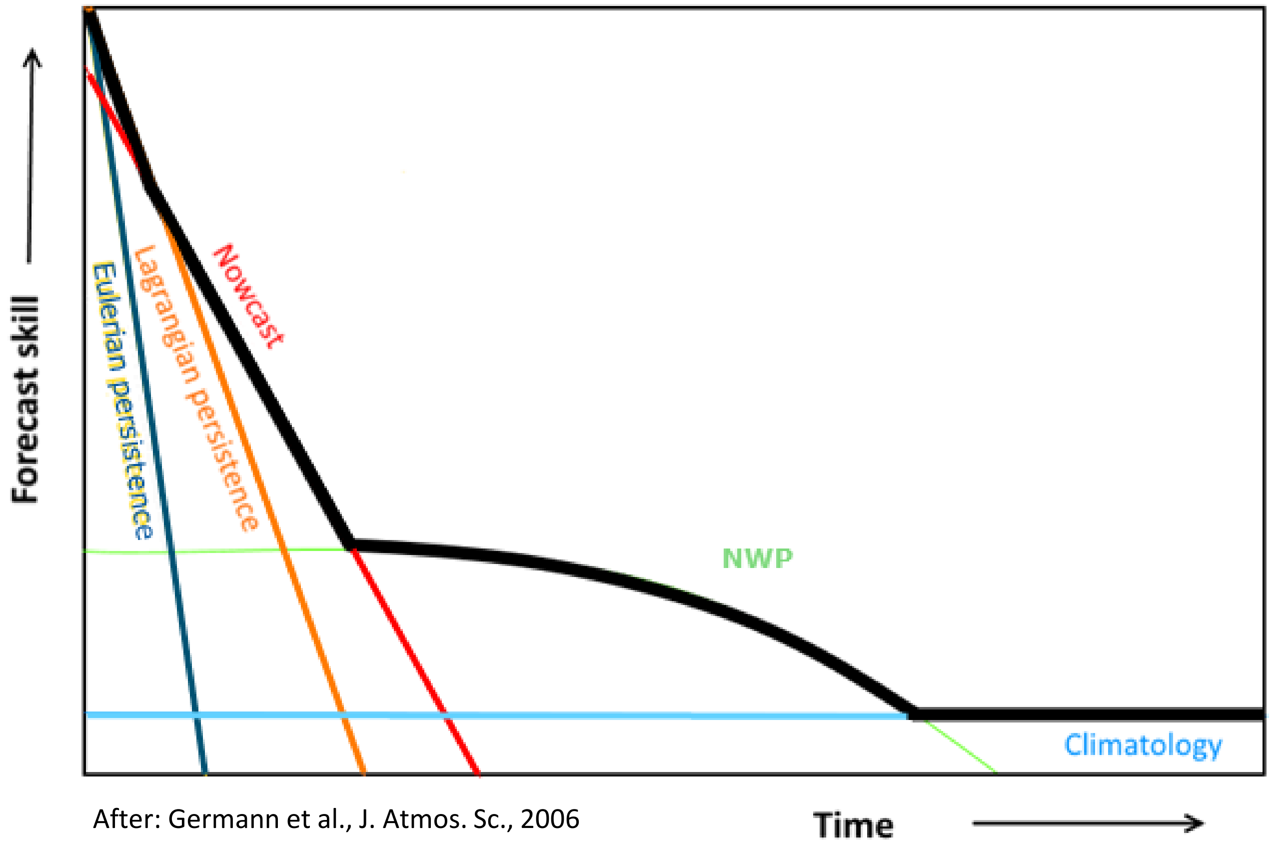
Rapid forecasting (e.g. new nowcasts available every 5 min.)

Absence of physical growth/dissipation processes

Probabilistic nowcasting

[e.g. Bowler et al., Q. J. Roy. Meteor. Soc., 2006; Berenguer et al., J. Hydrol., 2011; Seed et al., Water Resour. Res., 2013; Pulkkinen et al., Geosci. Model. Dev., 2019]

Can radar rainfall nowcasting fill the very-short-term predictions gap?



After: Germann et al., J. Atmos. Sc., 2006

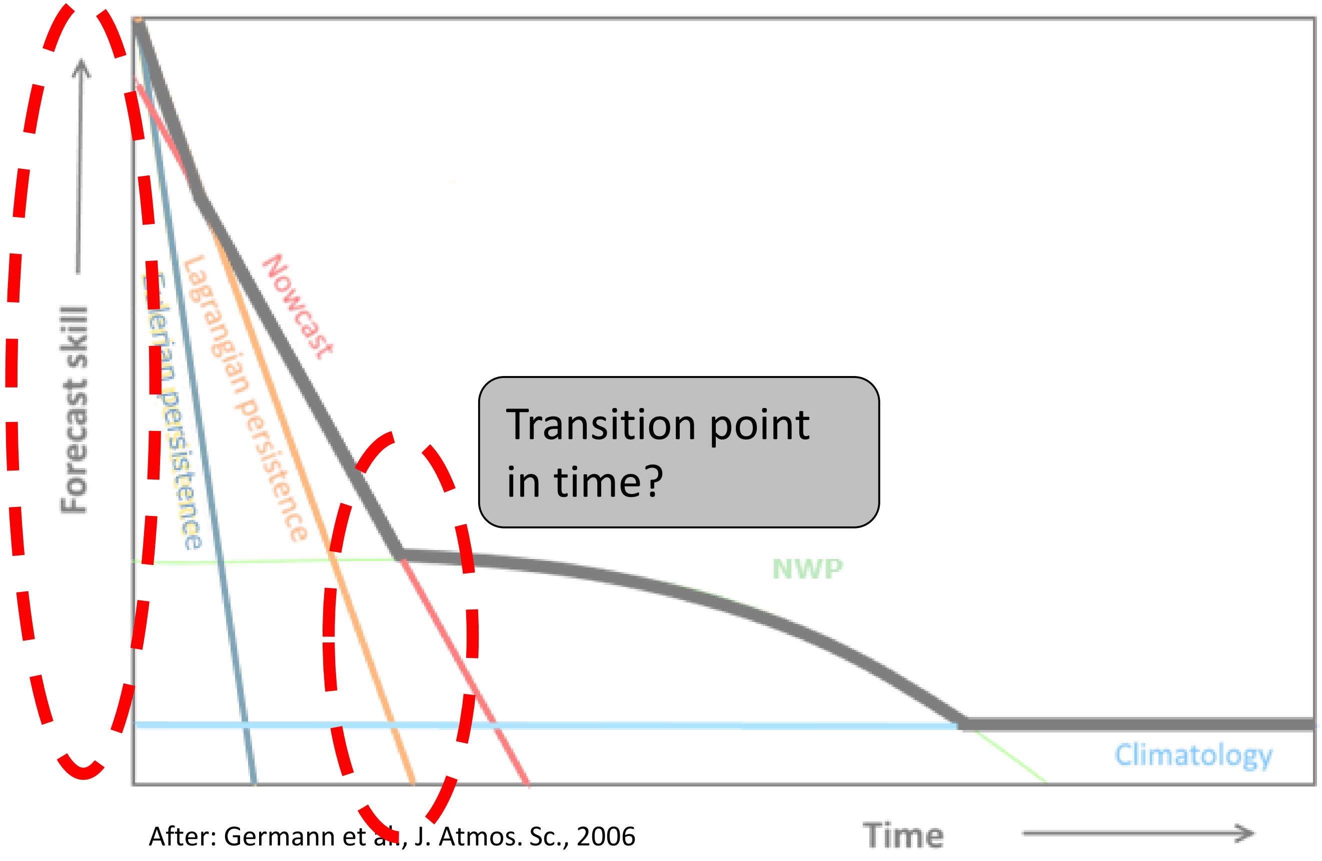
The state-of-the-art

- Nowcasting is applicable up to 3 - 6 h ahead [e.g. Lin et al., Geophys. Res. Lett., 2005; Germann et al., J. Atmos. Sc., 2006]
- 30 min for convective cells [e.g. Liguori and Rico-Ramirez, Atmos. Res., 2012; Foresti et al., Hydrol. Earth Syst. Sc., 2016]
- The focus is dominantly on model development, with currently a big interest in: ensemble predictions, deep Learning Models and model uncertainties.

Model skill is, however, often determined with relatively small analyses (2 – 15 events).

Can radar rainfall nowcasting fill the very-short-term predictions gap?

What skill can we expect?

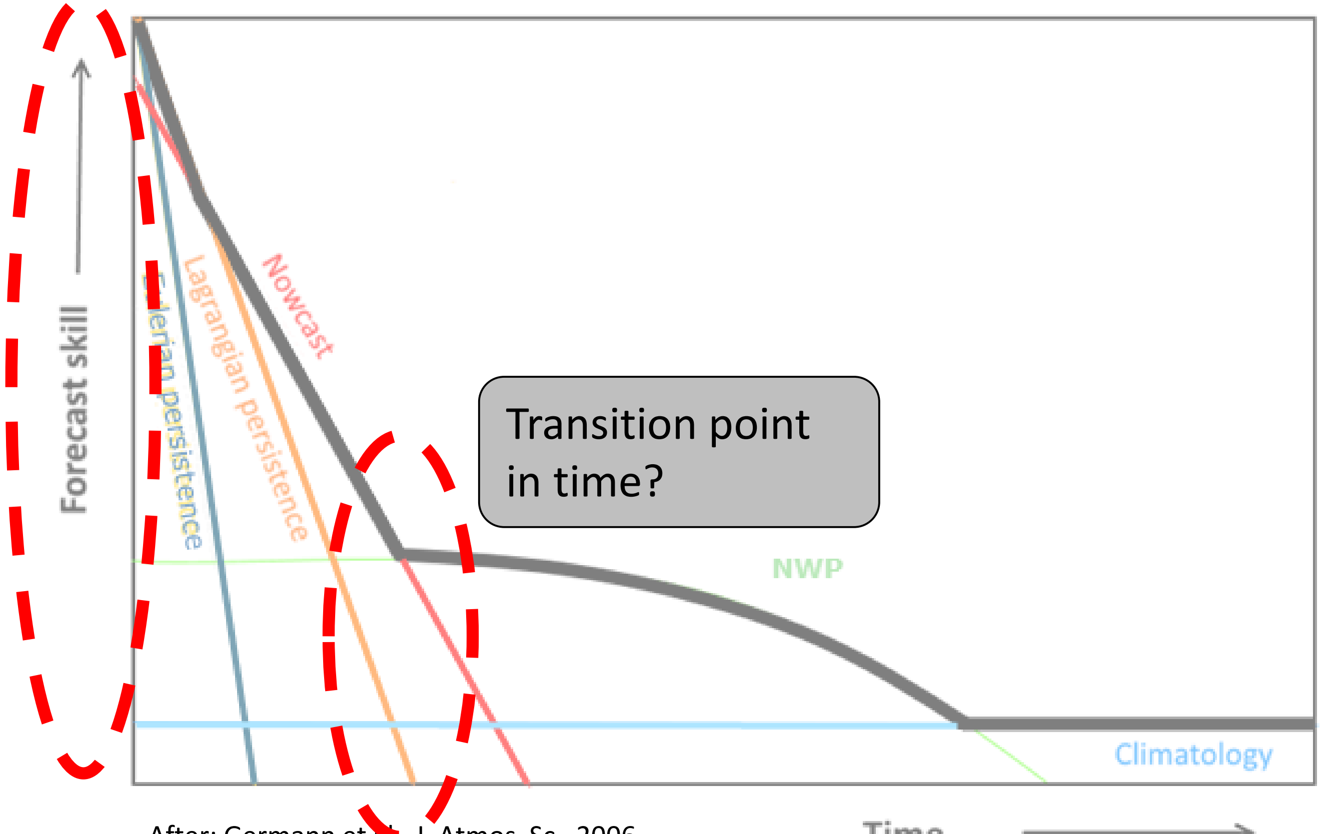


Transition point in time?

After: Germann et al., J. Atmos. Sc., 2006

Can radar rainfall nowcasting fill the very-short-term predictions gap?

What skill can we expect?



Transition point in time?

After: Germann et al., J. Atmos. Sc., 2006

And.. What is the dependence on climatic and environmental characteristics?

Our objective

To quantify the skill of radar rainfall nowcasting algorithms for the short-term predictability of rainfall for different regions in the Netherlands

Study area

Twelve catchments and polder areas in the Netherlands.

Areas range from 6.5 to 957 km^2

Employed nowcasting algorithms

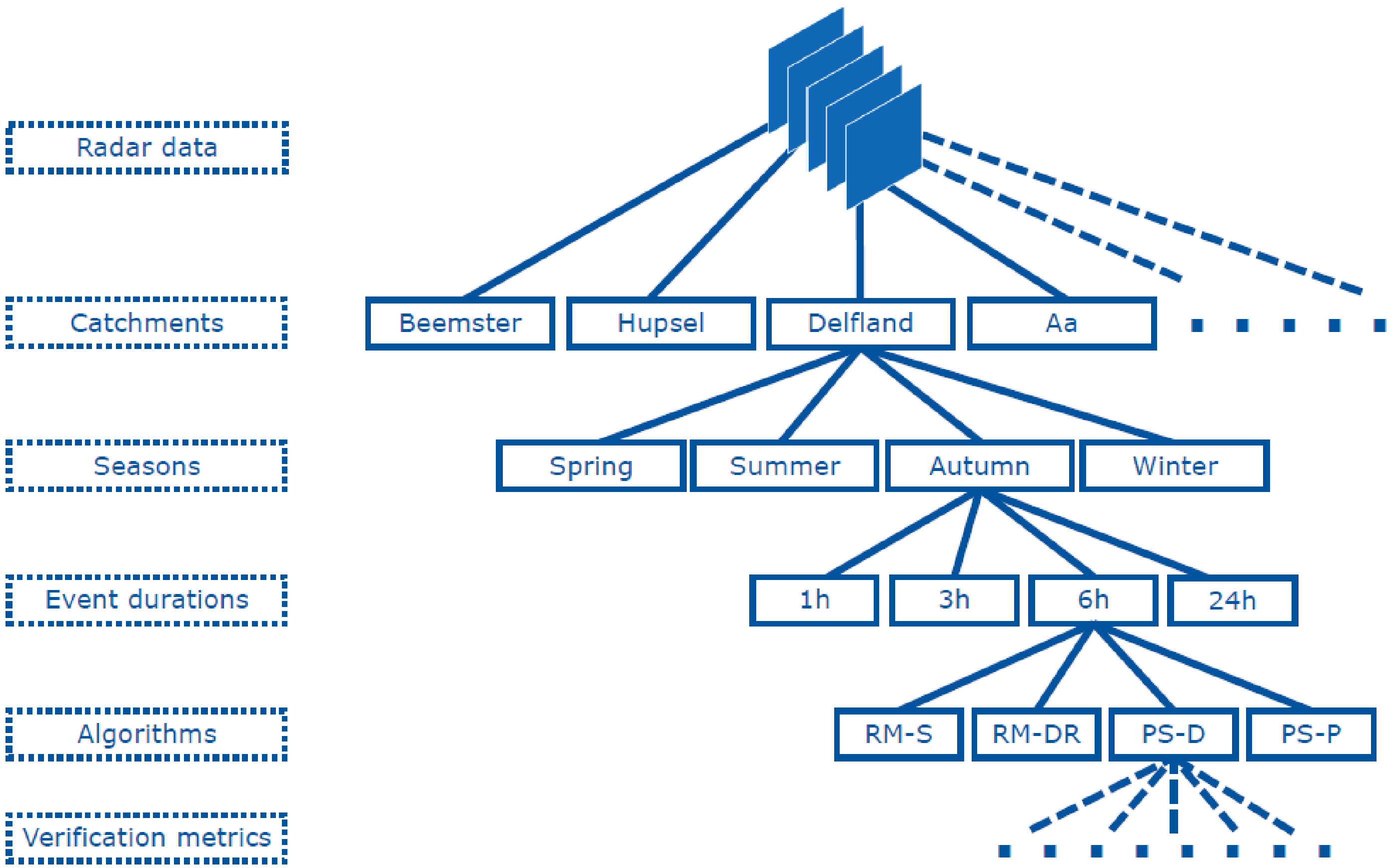
Name	Abbreviation	Type of system	Source
Eulerian Persistence	EP	-	-
rainymotion Sparse	RM-S	Advection / benchmark	Ayzel et al., GMD, 2019
rainymotion DenseRotation	RM-DR	Rotational advection / benchmark	Ayzel et al., GMD, 2019
pySTEPS deterministic (S-PROG)	PS-D	Deterministic nowcasting	Seed, J. Appl. Meteorol., 2003; Pulkkinen et al., GMD, 2019
pySTEPS probabilistic *20 ensemble members*	PS-P	Probabilistic nowcasting	Bowler et al., QJR. Meteor. Soc., 2006; Seed et al., WRR, 2013, Pulkkinen et al., GMD, 2019

Four nowcasting algorithms are used and compared to Eulerian Persistence.

All methods are cross-correlation based.

Rainymotion originally introduced as an alternative benchmark for the evaluation of other nowcasting methods.

Systematic event selection procedure



Radar data: 11 years, 5 min, 1 km^2

12 Dutch catchments

Per season and event duration, 8 events selected with highest rainfall amounts.

1536 events in total *(of which 1533 successfully run)*

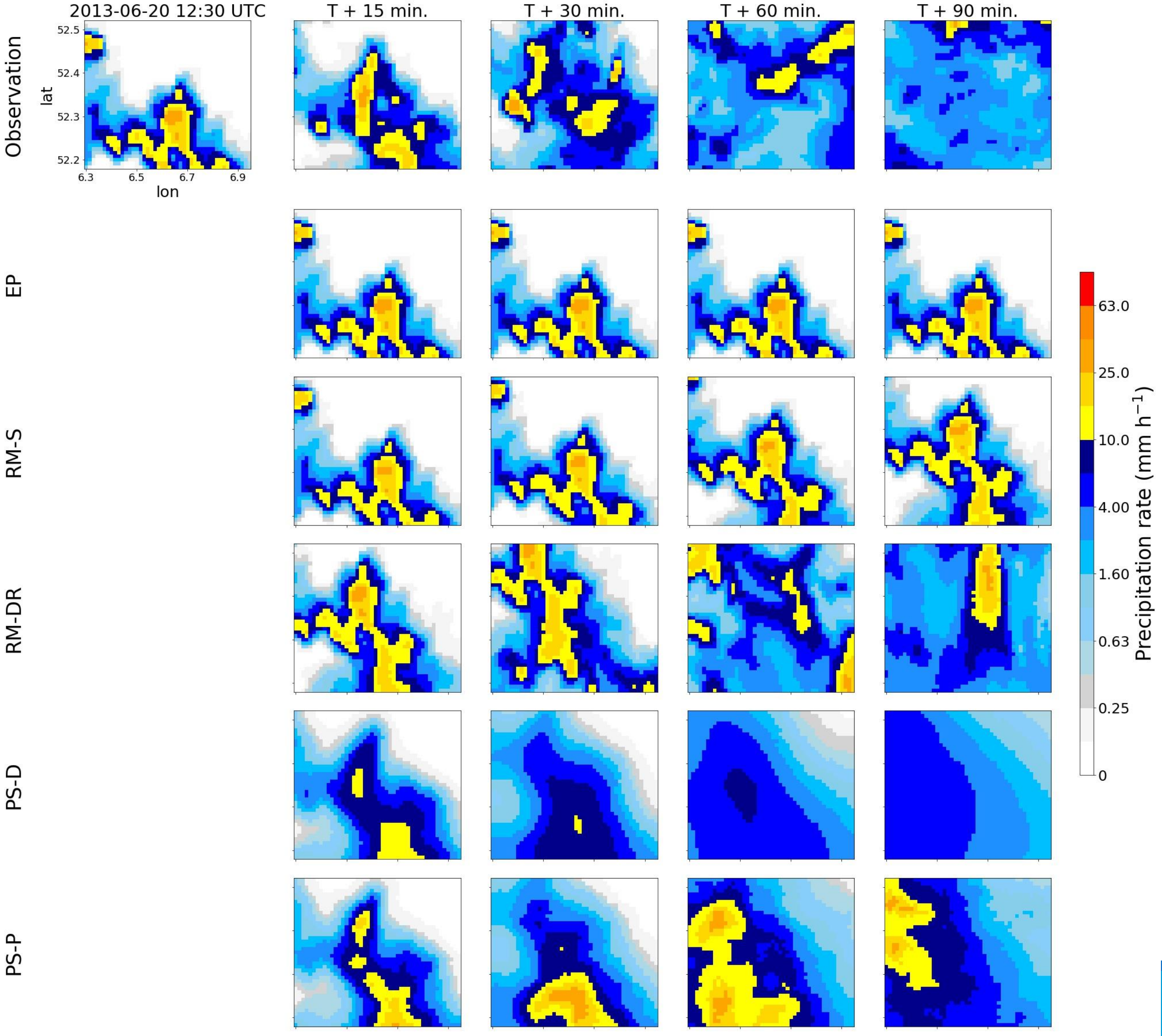
Results – Example nowcast

Example event for the Regge catchment (957 km^2; eastern Netherlands) on 20 June 2013.

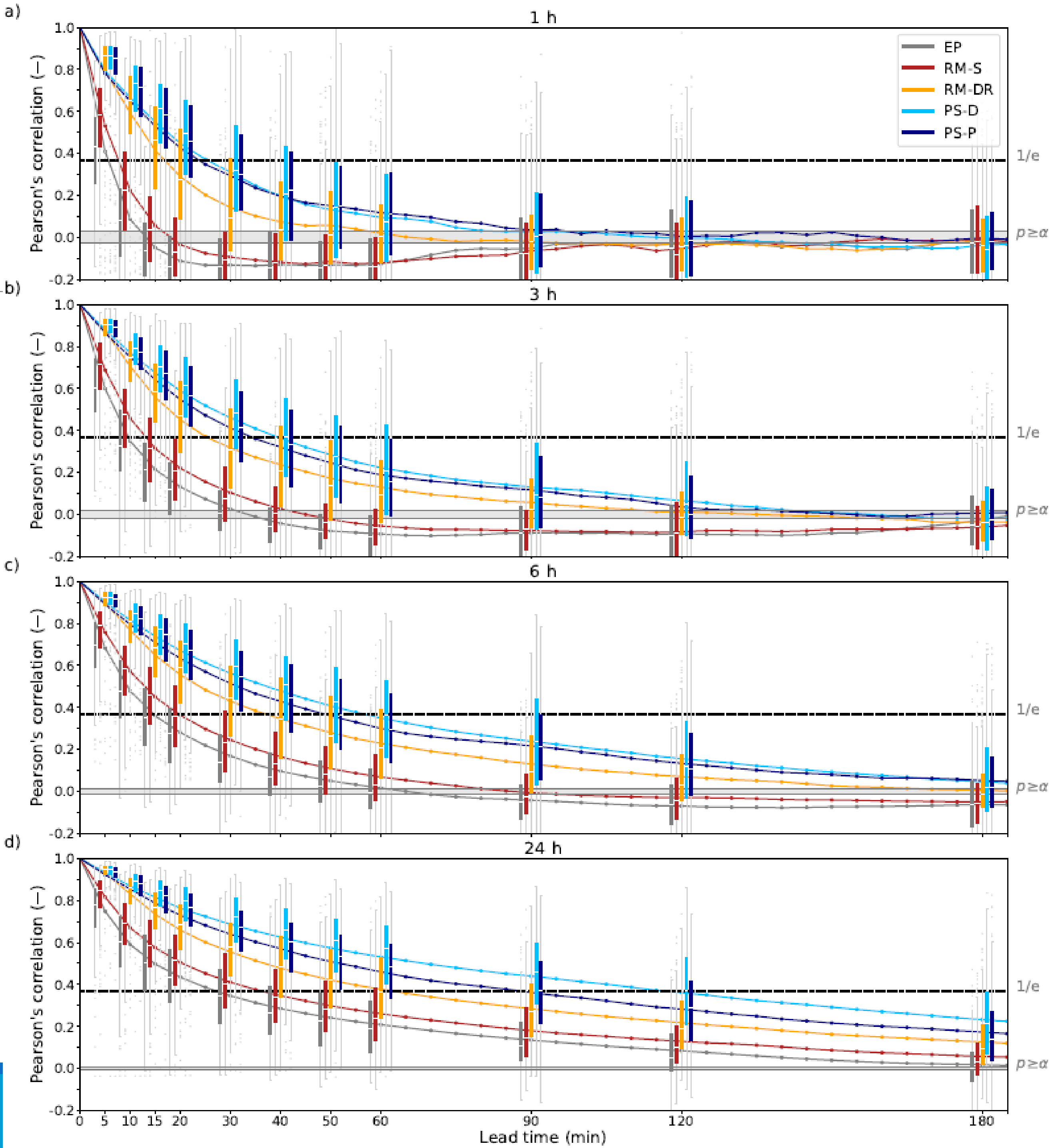
On a national scale, nowcasts often seem to capture the location of rainfall quite well. On a local scale, this is different:

None of the algorithms are able to fully capture the rainfall field evolution, but RM-DR, PS-D and PS-P are clearly outperforming RM-S and EP here. Until t + 15 min, all methods are performing relatively well, after that only the three aforementioned algorithms. Note the smoothing behavior of PS-D, which is more often found for S-PROG.

Only ensemble member 10 (out of 20) is shown for PS-P.



Results



Max. skillful lead time (PS-D):

25 min.

40 min.

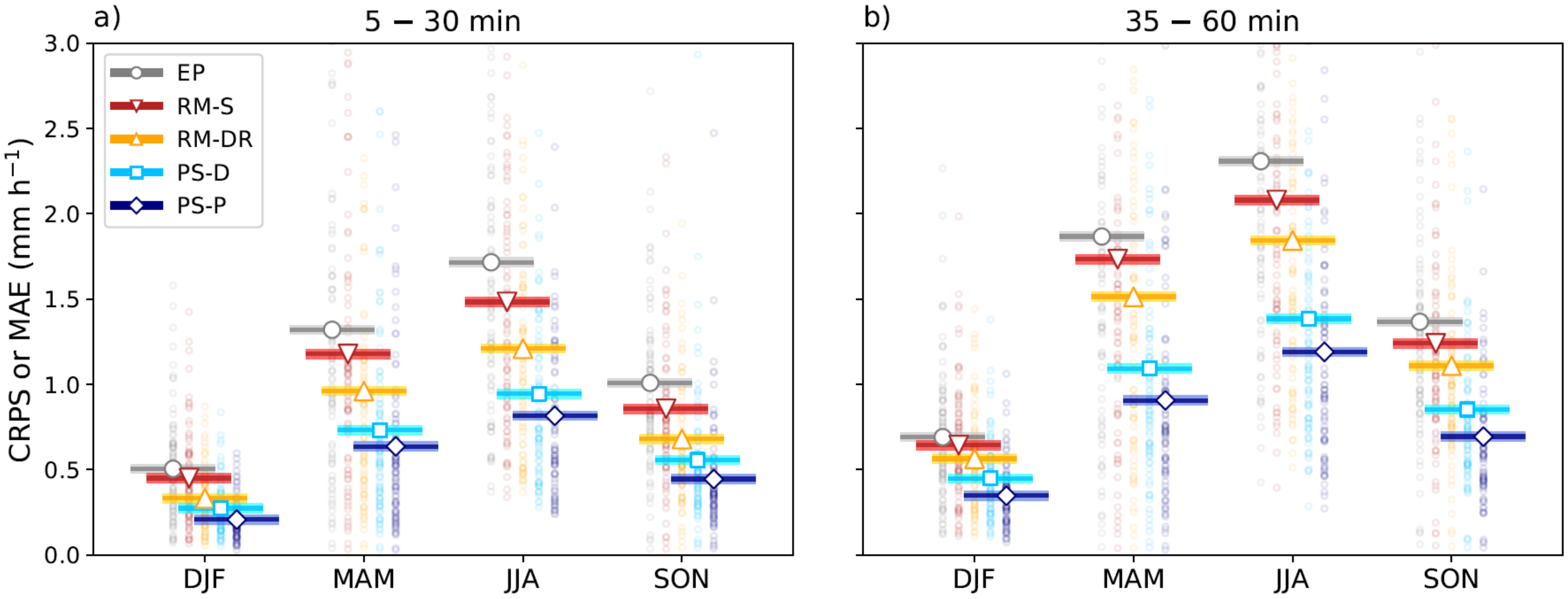
56 min.

116 min.

Shown: Event-averaged Pearson's correlation as a function of lead time for the four durations. The dotted line indicates a correlation of 1/e (min. correlation for a skillful nowcast). Elongated boxes indicate the event variability in the results, with: the median in white, the interquartile range (IQR) in colored boxes, 1.5 * IQR outside the boxes in grey bars and outliers in grey dots.

Around a correlation of 0.0, horizontal grey bands indicate insignificant correlations from 0.0, based on a two-tailed T-test with α = 5%.

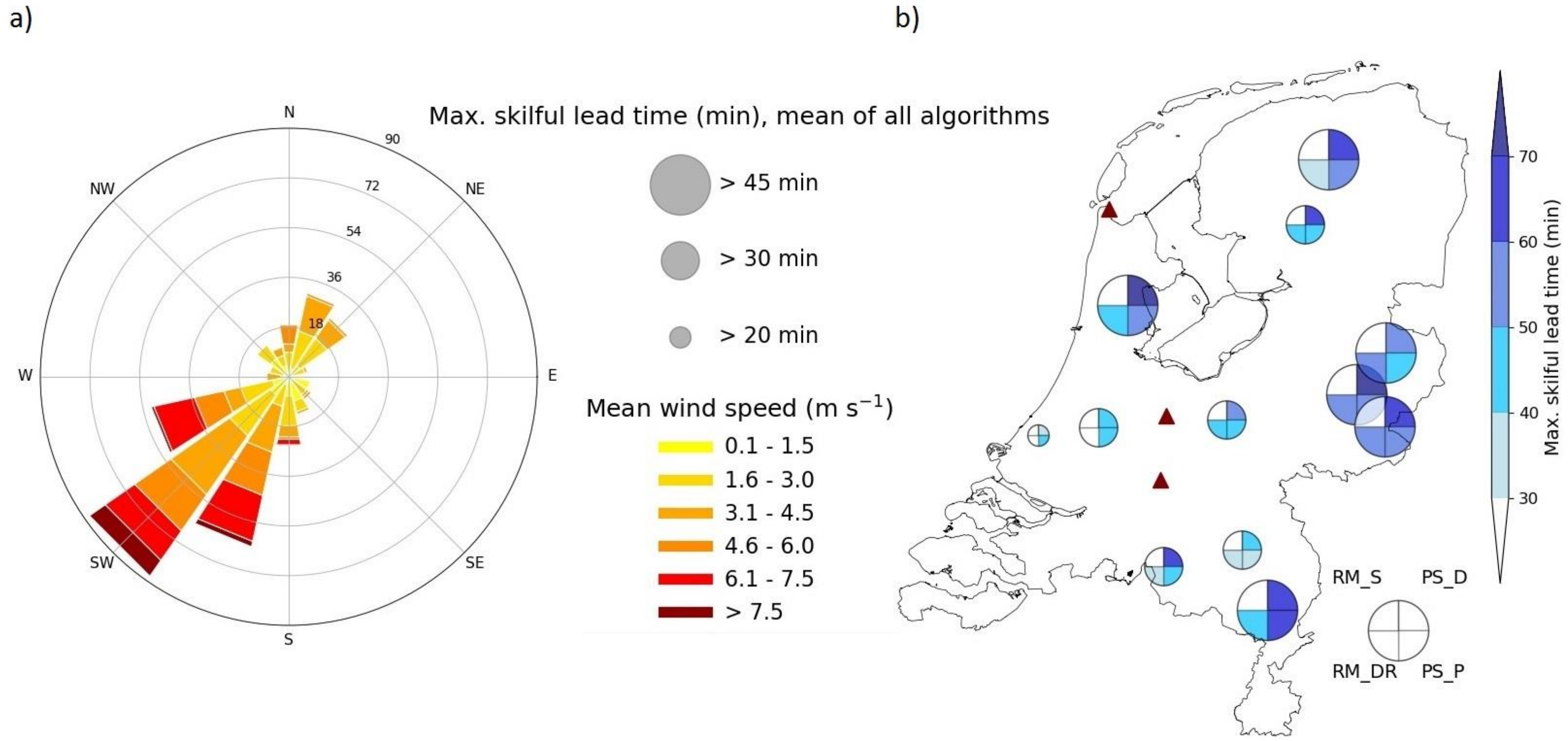
Results – Seasonal dependency



- Better forecasts during winter than during summer, with MAE/CRPS values that are approx. three times lower during winter than during summer. Difference remains the same for longer lead times, only the MAE/CRPS increases for all seasons.
- Consistent performance difference with from high to low: PS-P, PS-D, RM-DR, RM-S and EP.

Shown: Event-averaged CRPS (for PS-P) and MAE (for all other methods) per season for all catchments for the 6-h event duration. The scores are averaged over lead times of 5 - 30 min (a) and 35 - 60 min (b). Individual points indicate the mean CRPS or MAE per event.

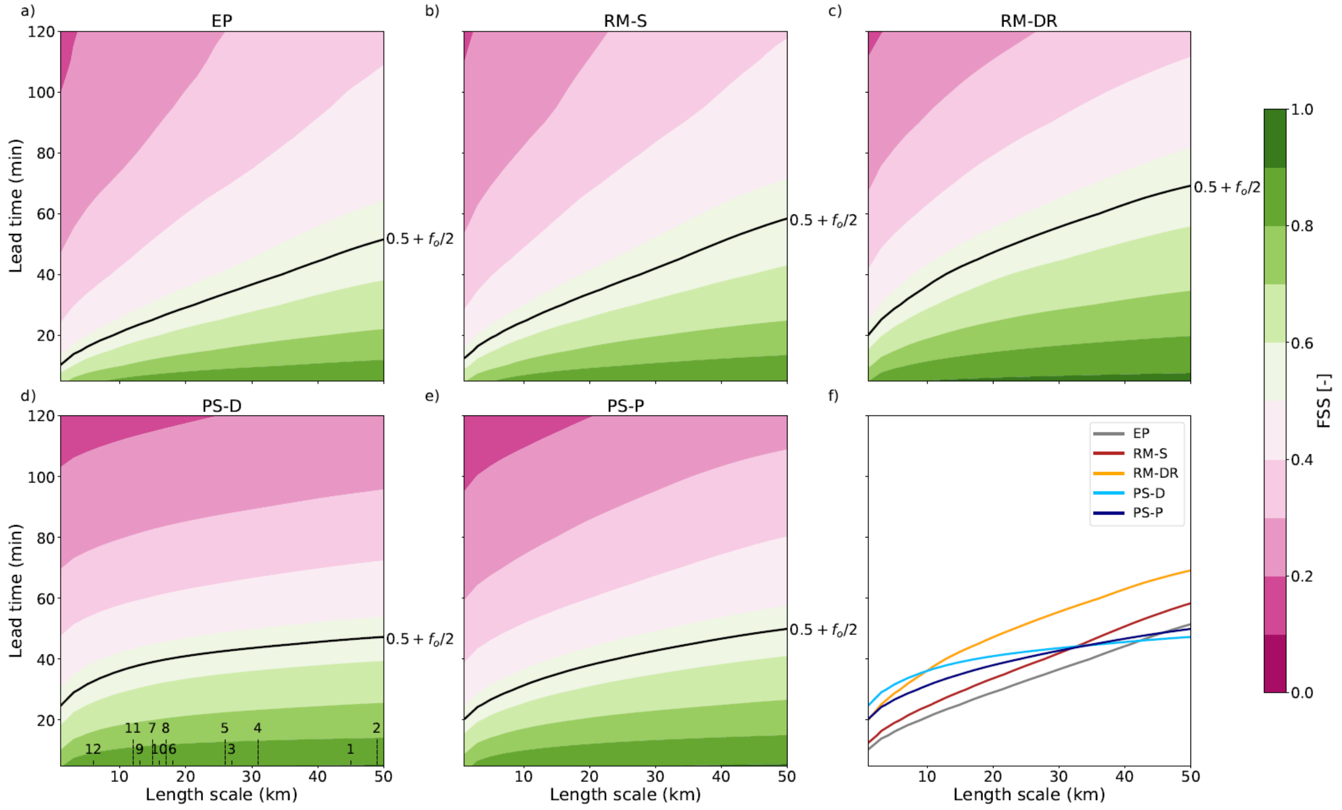
Results – Location dependency



- Catchment location matters for nowcast skill.
- Prevailing southwesterlies during the selected events with 6-h duration.
- Average max. skillful lead times generally increase in downwind direction.

Shown: a) Prevailing wind directions at KNMI station De Bilt for the events with the 6-h duration. The length of the bars indicates the number of events with that wind direction. b) Mean max. skillful lead time of all 6-h events (based on Pearson's correlation) for the 5x5 center cells per catchment (to make it catchment size independent). Circle sizes indicate the average of the four algorithms, whereas the hue per quarter indicates the max. skillful lead time per algorithm. Radar locations are indicated with red triangles.

Results – Catchment size dependency



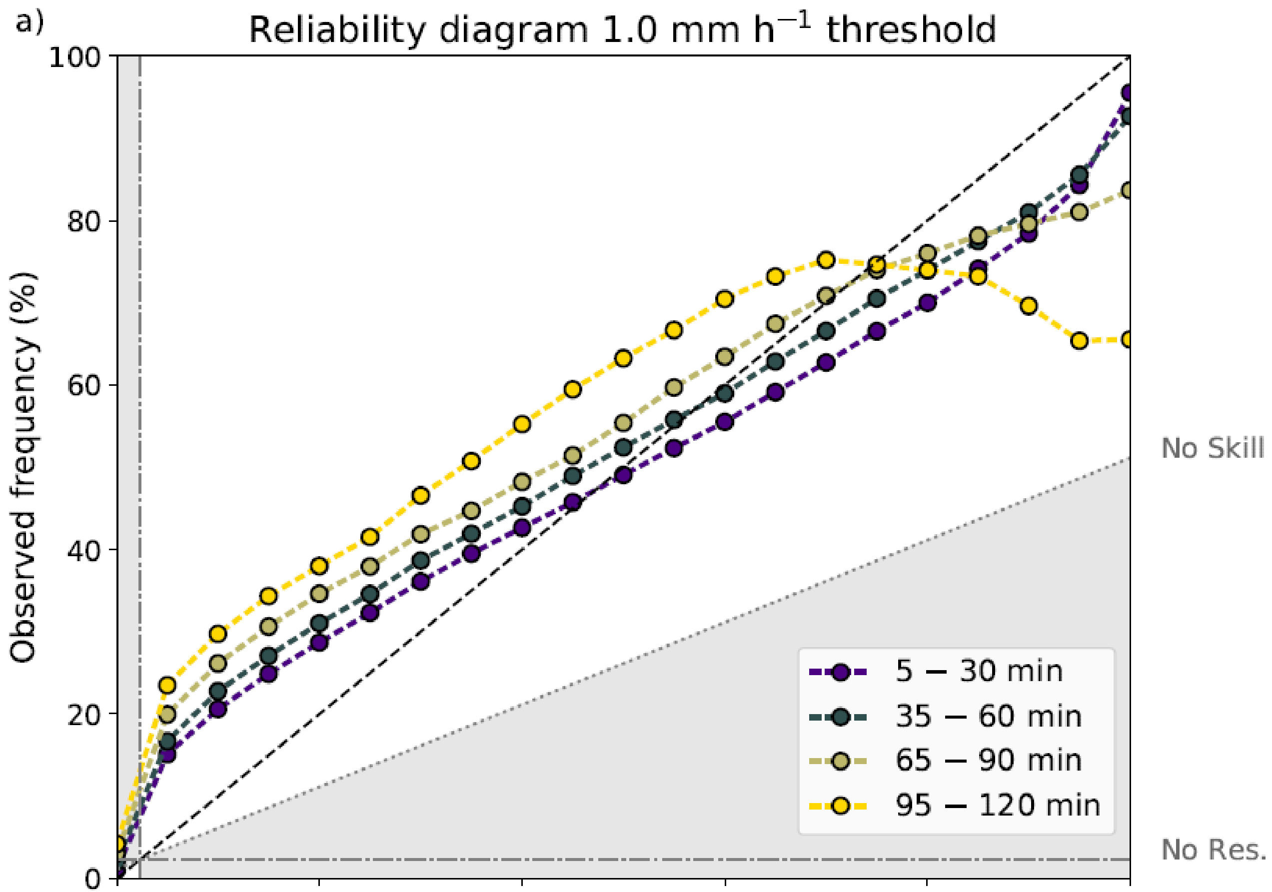
Shown: Event-averaged Fractions Skill Score (FSS) as a function of catchment length scale and lead time (for h-6 duration and 1.0 mm h^{-1} threshold, based on catchments Aa and Regge). FSS = 0.5 + f_0 / 2 (black line) indicates the min. FSS for a skillful nowcast. All contour lines are combined in (f).

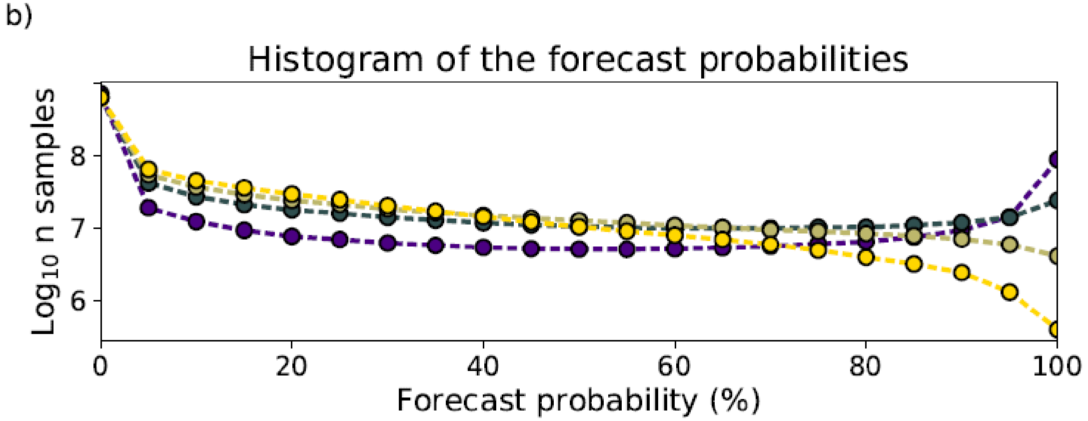
- FSS increases with increasing length scale (thus catchment size) and decreases with increasing lead time.
- FSS is sensitive to a bias in the forecast, which is present for all algorithms, but stronger for the pysteps algorithms.
- This explains the steeper increase in FSS with increasing length scale for RM-DR than the two pysteps algorithms.

WAGENINGEN UNIVERSITY & RESEARCH

Deltares Enabling Delta Life

Results – Ensemble verification (PS-P)

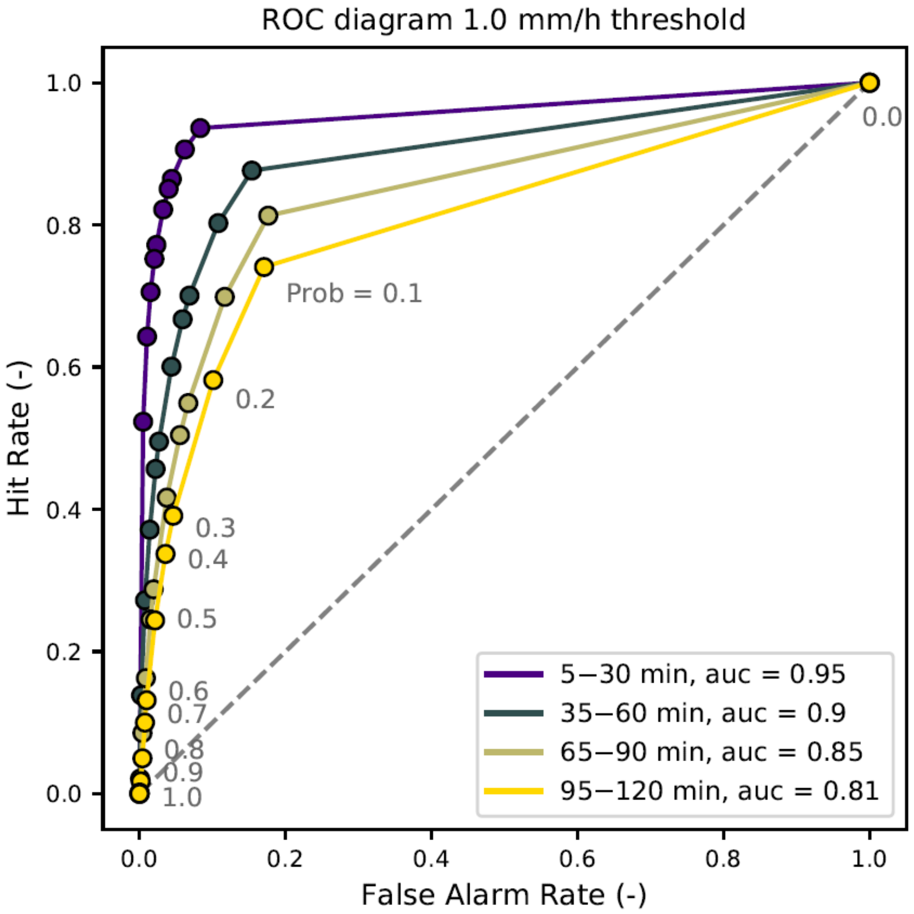




Shown for PS-P: (a) Reliability diagram of exceeding a threshold of 1.0 mm h^{-1} (6-h duration). The climatological frequency of exceeding the threshold in the events is used as reference. (b) Histogram of the frequency a probability (fraction of the ensemble members) was forecast for exceeding the threshold.

WAGENINGEN UNIVERSITY & RESEARCH
Deltares Enabling Delta Life

Results – Ensemble verification (PS-P)



ROC curve of exceeding a 1.0 mm h^{-1} threshold for the events with a 6-h event duration. auc (area under the curve) indicates the skill of the probabilistic forecast with 1.0 for a perfect forecast.

Conclusions

Skillfulness of radar rainfall nowcasting tested

1536 events

12 Dutch catchments

5 nowcasting methods

Conclusions

Skillfulness of radar rainfall nowcasting tested

1536 events

12 Dutch catchments

5 nowcasting methods

Skill was found to depend on:

1) **Event type and duration:** Increasing for longer events, max. skilfull lead times range from 25 min (1-h events) to 116 min (24-h)

2) **Season**: Decreasing skill towards summer

3) **Location**: Increasing in the downwind direction

4) **Catchment size**: Increasing with larger catchment size

Conclusions

Skillfulness of radar rainfall nowcasting tested

1536 events

12 Dutch catchments

5 nowcasting methods

Skill was found to depend on:

1) **Event type and duration**

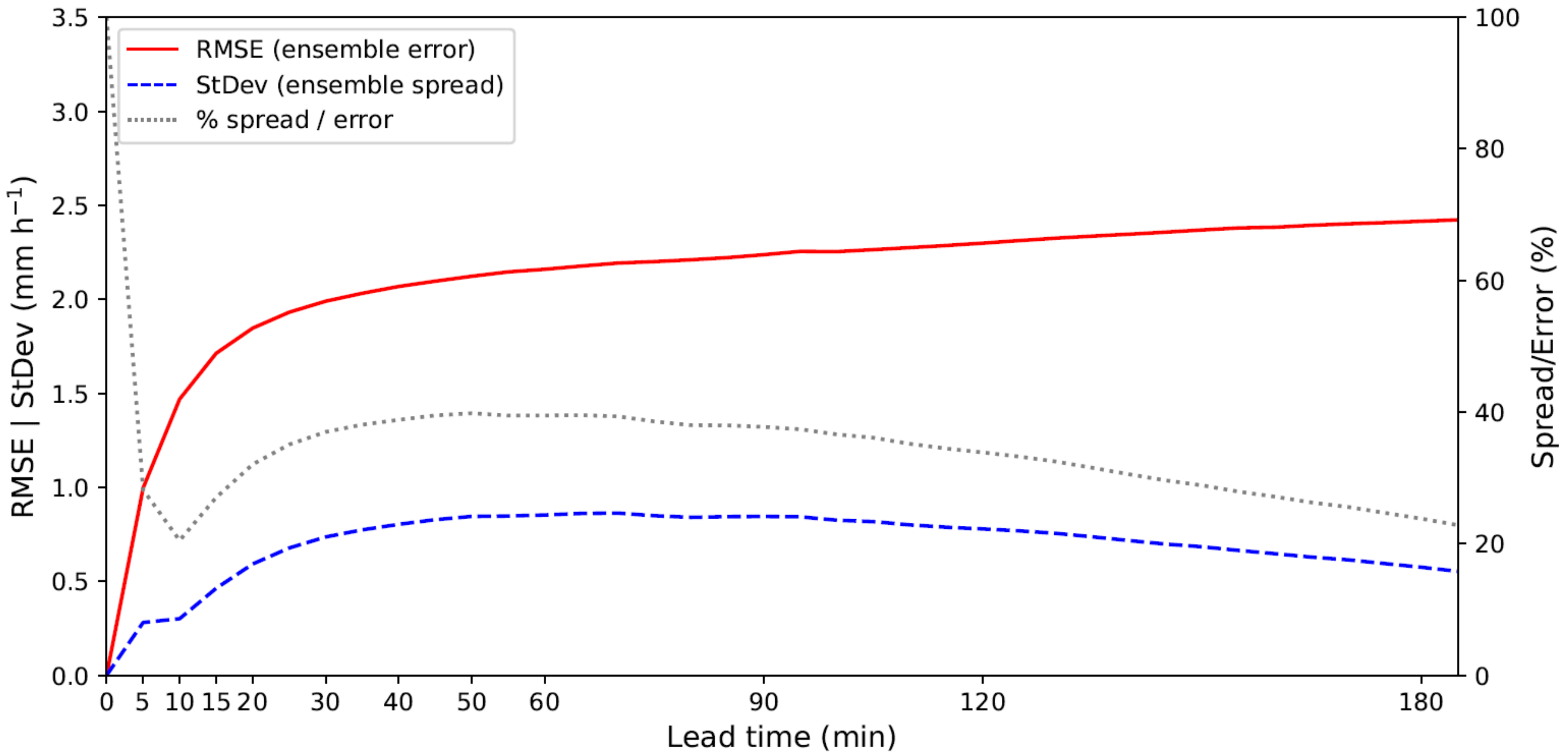
2) **Season**

3) **Location**

4) **Catchment size**

Despite the performance differences between the tested methods, none of them captures growth/dissipation processes well. This is currently a major shortcoming of cross-correlation based nowcasting methods.

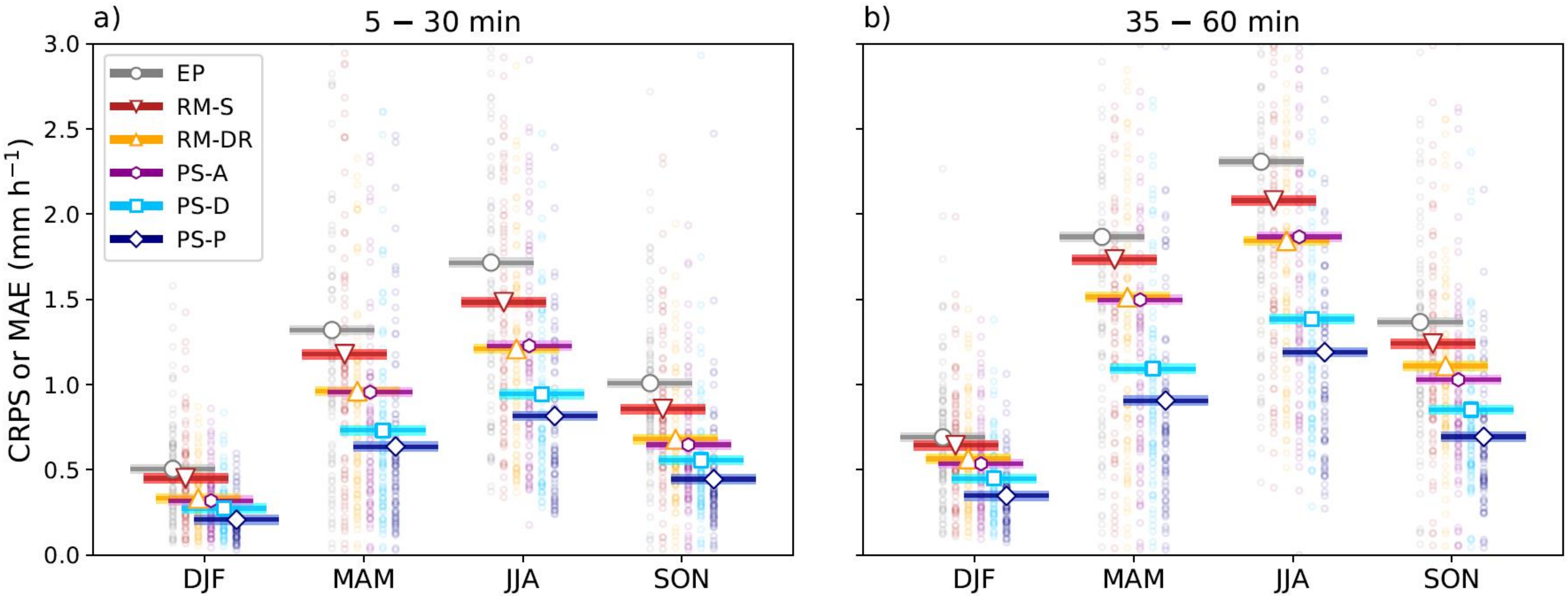
Extra – Ensemble spread vs error



Imhoff et al. (2020), WRR (in review)

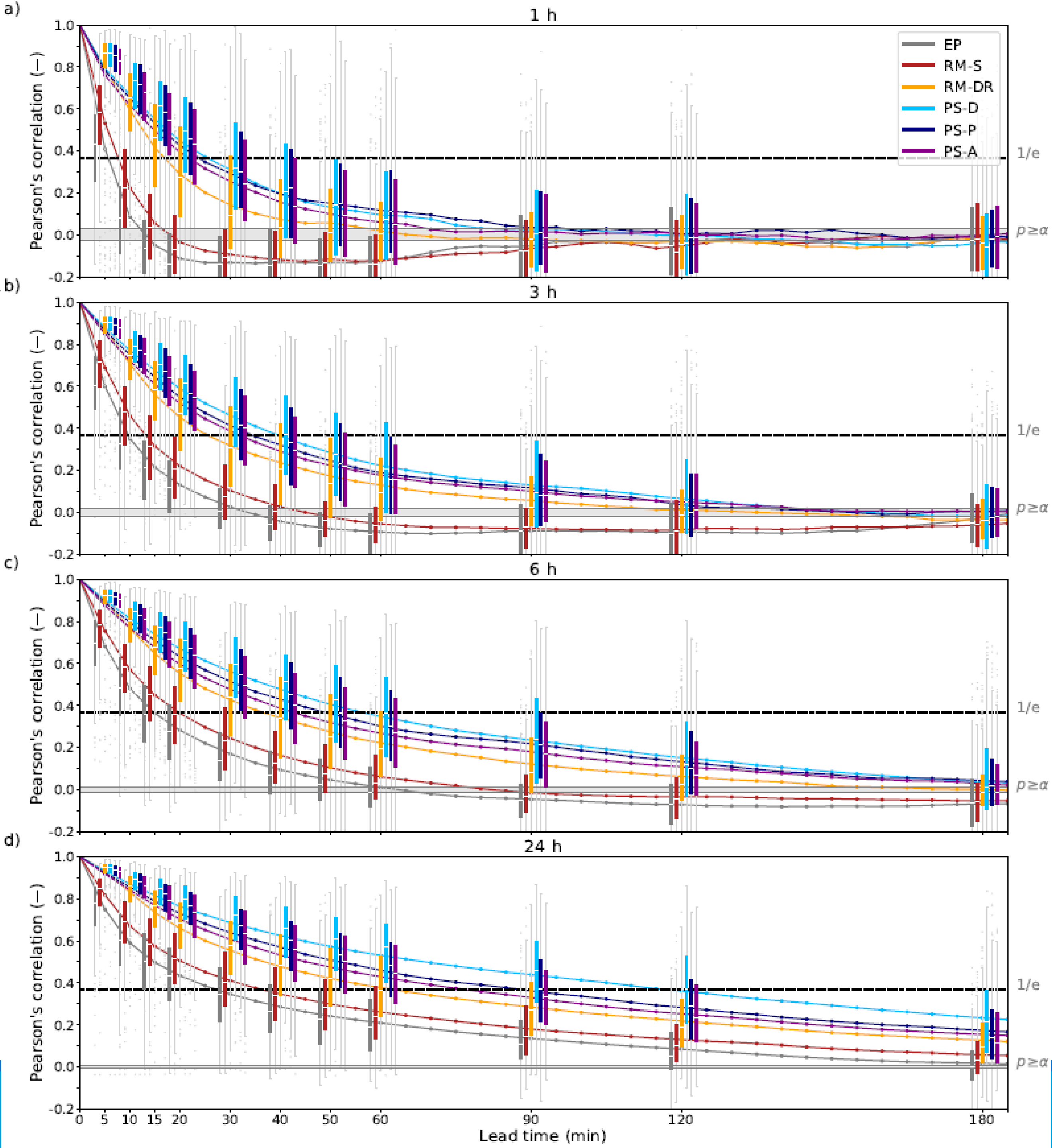
Extra – Seasonal dependency
pysteps advection

Pysteps run with only advection (PS-A)



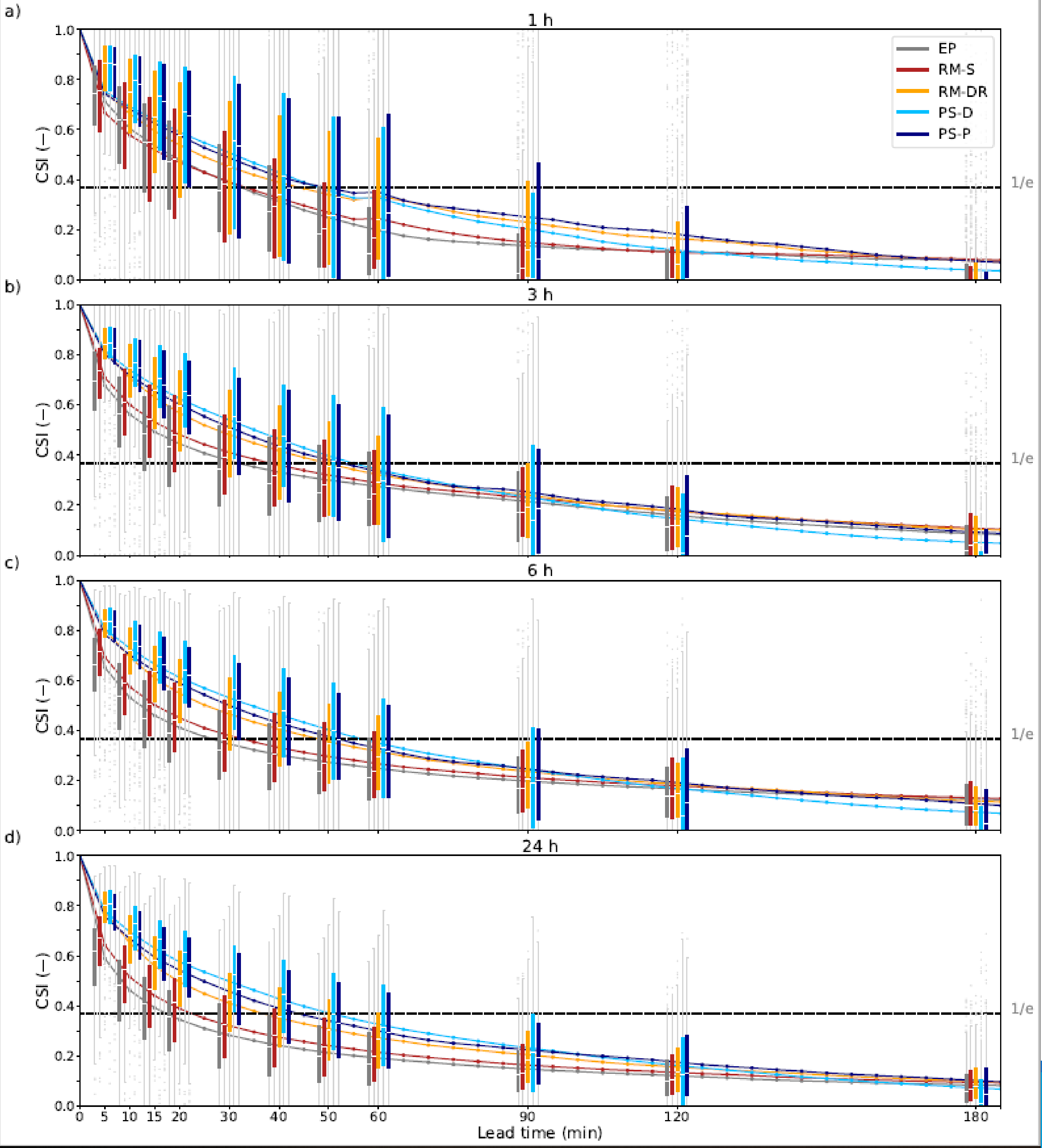
Extra – Event duration dependency *pysteps advection*

Pysteps run with only advection (PS-A)



Extra – Event duration dependency

Critical Success Index (1.0 mm h^{-1} threshold)



Imhoff et al. (2020), WRR (in review)