Met Office

सत्यमेव जयते
Ministry of Earth Sciences

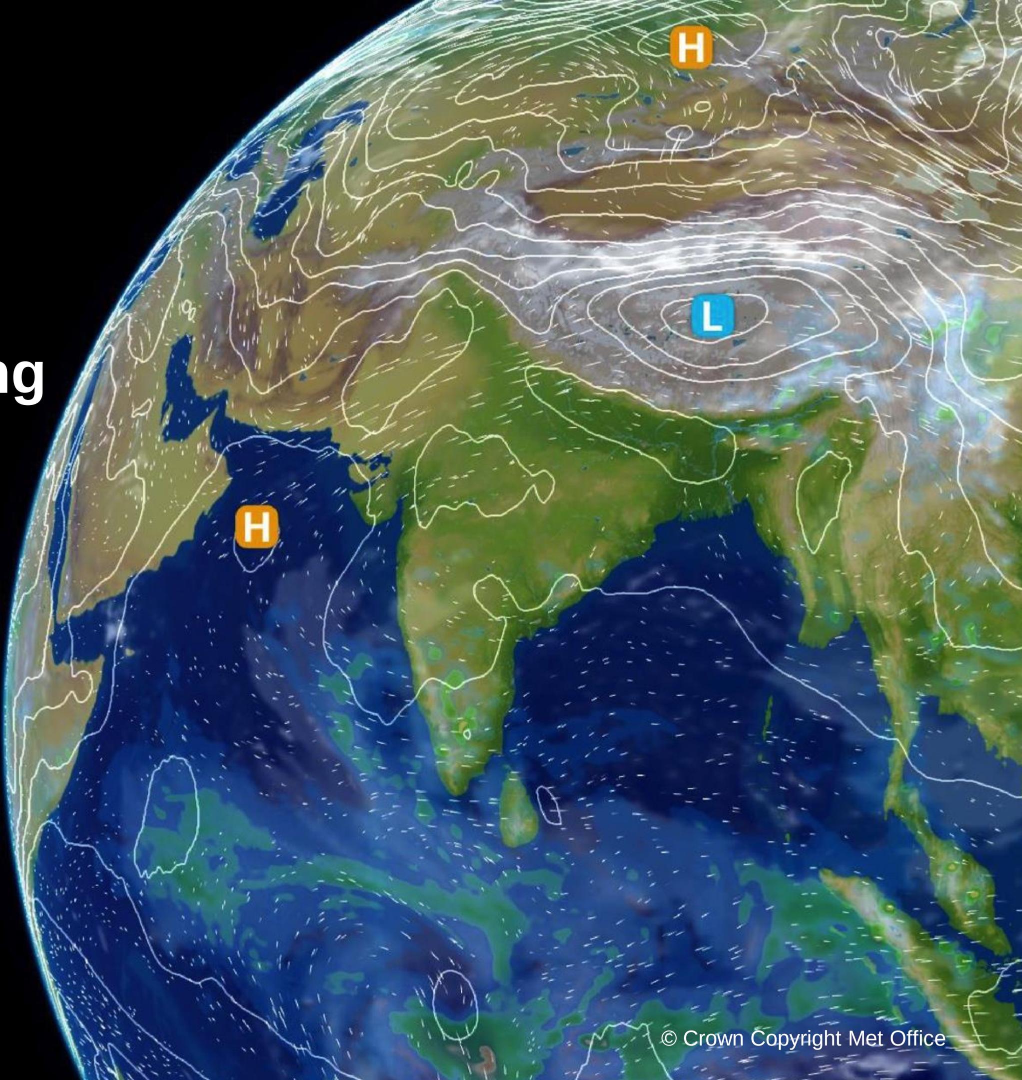
Quantifying observation representativeness errors using a spatial verification method: A lightning-based illustration

Marion Mittermaier and

Jonathan Wilkinson

9th International Verification Methods Workshop
Cape Town, May 2024

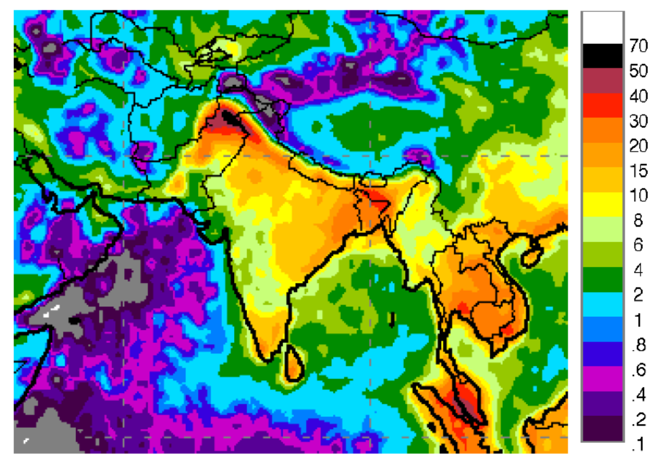
Mittermaier and Wilkinson, 2024, WAF, “accepted” subject to minor revisions.

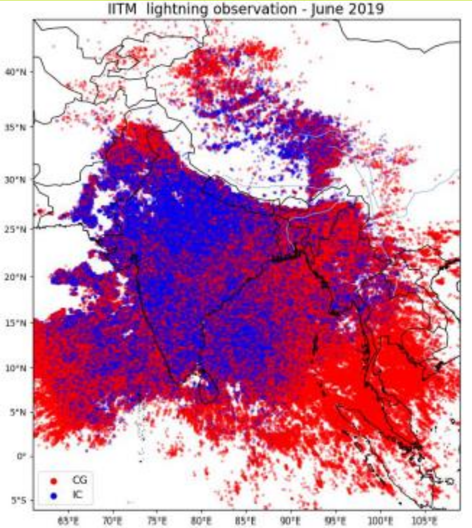


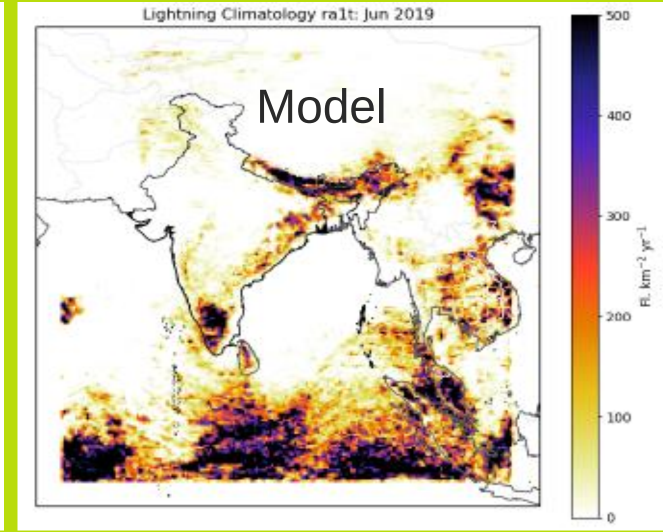
Lightning over India

Is a public safety issue. More people are killed by lightning than any other weather phenomenon.
Improving lightning forecasts is high up on the national forecast agenda.
Assessing lightning forecast performance is <u>non-trivial</u>

Annual lightning strikes per square km: 1995-2003 average around the India domain (Source: NASA)







ENGLN strikes for one month and what a model "sees"

Overview

Model

- 4.4 km forecasts from versions of the Met Office Unified Model over India for the monsoon. The model uses the **McCaul scheme**.

Observations

- **IITM lightning observations (ENGLN)** counts accumulated onto the model grid.

$$F = 0.95F1 + 0.05F2$$
$$F1 = k1 \times graupel\ flux$$
$$F2 = k2 \times total\ ice\ water\ path$$

Where F1/F2 are flash rates per 5 minutes, F is the weighted flash rate per 5 minutes and graupel flux is at the -15°C isotherm

We are intent on **improving the forecasts of lightning flash counts** in the model. India presents a uniquely active environment for development, tuning and testing of model parameterisations and subsequent forecast performance.

Observation representativeness mismatch

- **Lightning "flash" essentially treated as a point**, but a flash tends to consist of many strokes which can travel up to ~20 km in the horizontal. This spatial dimension is not reflected in the observations we use.
- This leads to a potentially **large representativeness mismatch** even between the 4.4 km model and aggregated lightning flashes onto the same grid.
- **Gaussian kernel smoothing** was added to categorical analysis of elements to mitigate against this impact this source of observation uncertainty has on the metrics.
- **Recently, the intensity has been re-imputed into the smoothed field but scaled such that total lightning over the domain is conserved**. This implies that the observed peaks are reduced in preference to spreading the lightning out in space.

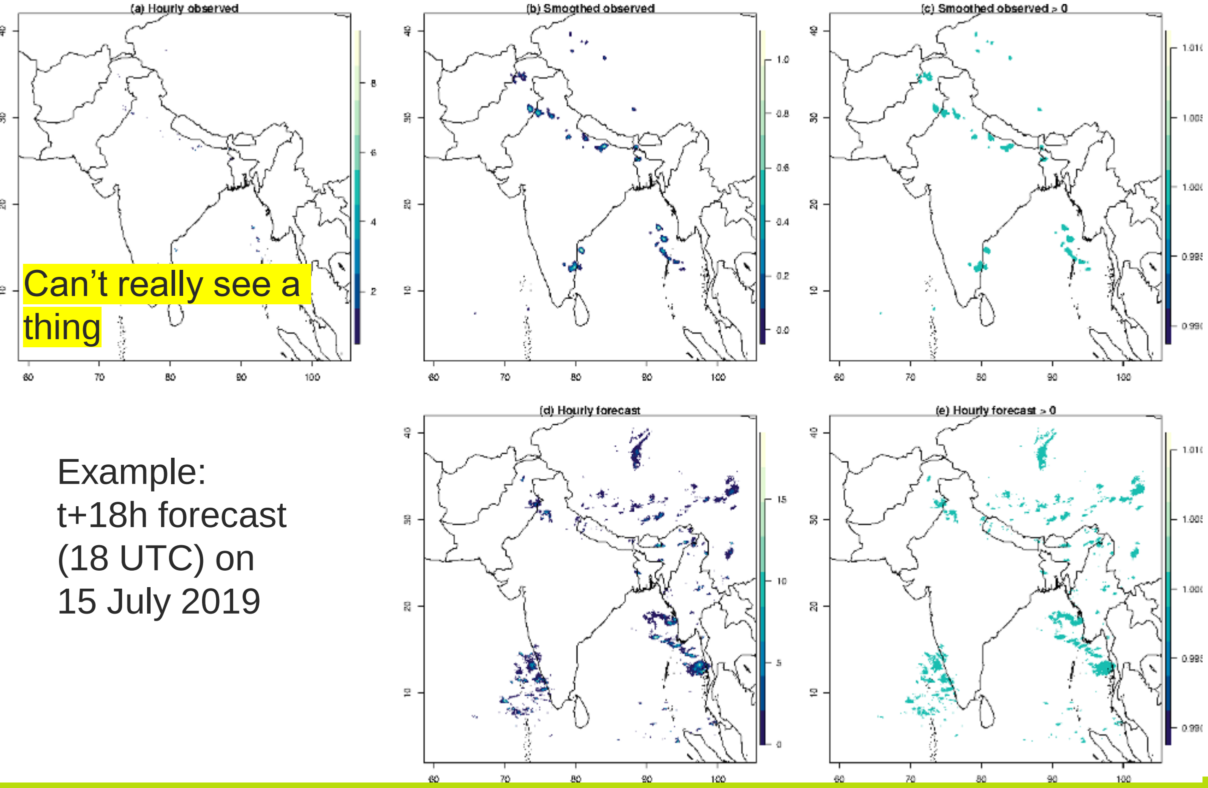
Addressing representativeness of the lightning observations

Met Office

Ministry of Earth Sciences

Department for Science, Innovation & Technology

Hourly



Example:
t+18h forecast (18 UTC) on 15 July 2019

Gaussian kernel dressing applied to hourly gridded lightning observations. (a) Gridded lightning flash counts; (b) kernel dressed and mass conserved flash counts; (c) binary field created from (b). (d) Hourly model flash counts and (e) binary field produced from (d).

Representativeness mismatch reduced

Used a ~12 km kernel to acknowledge that lightning can travel ~20 km in the horizontal, i.e. the location that the strike is registered to is not necessarily representative of the physical extent.

See also Mittermaier et al. (2022a, Meteorol. Apps.)

Spatial method required?

Yes. But why *not* the FSS?

- Just because you want to compare two gridded fields, the **FSS is not always appropriate.**
- For example, the **characteristics of the fields** are an important consideration.

In this instance:

- Lightning fields are **generally very sparse**, and therefore liable to large AFSS mismatches.
- **Without addressing intensity biases the mismatch** in what the forecast produces and what the observations provide is **too severe** to yield robust results.
- The FSS **can be improved by over-forecasting** (increasing the extent exceeding a threshold), thus increasing the likelihood of, or increasing the extent of the overlap in a neighbourhood sense. As a result, it is not a good score to use for tuning forecast performance. Without accounting for the spatial observation uncertainty, the temptation to tune the model to over-forecast extent even more is very real.

Met Office

सत्यमेव जयते
Ministry of Earth Sciences

Coverage-Distance-Intensity (CDI)

Department for Science, Innovation & Technology

- **Method specifically developed for evaluating lightning forecasts.**
- The method uses observed lightning flash counts put on the model grid.
- **Treats each grid point separately without doing precise matching.**
- Could be classed as a hybrid method with some commonalities with SAL combined with a distance metric.
- Here a modified coverage component is used (the paper used SEDS).

A Technique for Verification of Convection-Permitting NWP Model Deterministic Forecasts of Lightning Activity

JONATHAN M. WILKINSON
Met Office, Exeter, United Kingdom

Wilkinson, J. M. (2017). A Technique for Verification of Convection-Permitting NWP Model Deterministic Forecasts of Lightning Activity, *Weather and Forecasting*, *32*(1), 97-115.

CDI method II

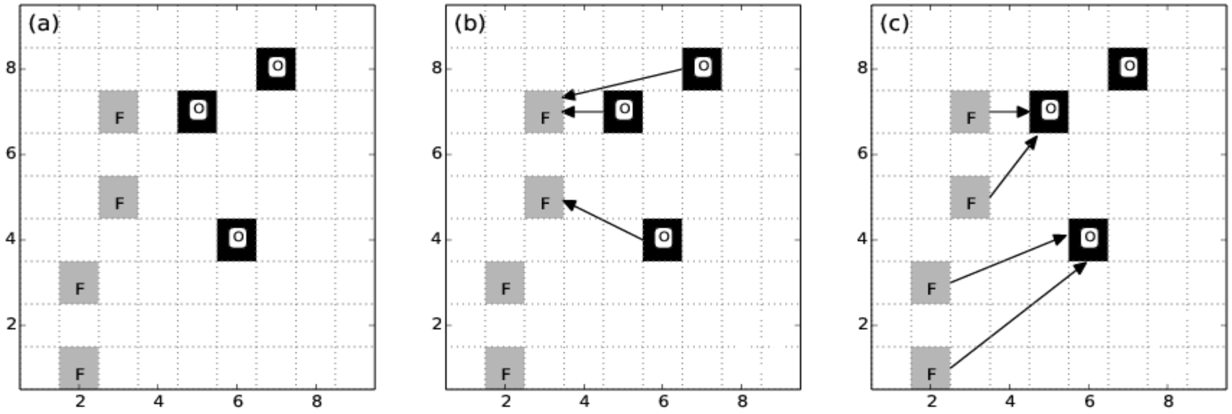
The revised CDI is based on three components:

- **Coverage:** Do the forecasts cover too large an area?
  Dominated by storm structure [-1,1]

$$C = \frac{P_m - P_o}{P_m + P_o}$$

> 0 over-forecast
= 0 unbiased
< 0 under-forecast

How is the location error computed?



Forecasts Grey, Observations Black

For each observation, look for nearest model point and define distance. Take mean distance as ***B (O2M)***

For each model, find nearest observation. Mean is ***C (M2O)***

Define a **Displacement Distance** **D_{dist}**(in km)

$$\frac{M2O + O2M}{2}$$

Analogous to a Modified Hausdorff Distance

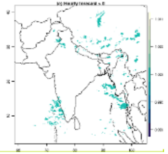
Met Office

सत्यमेव जयते
Ministry of Earth Sciences

Department for Science, Innovation & Technology

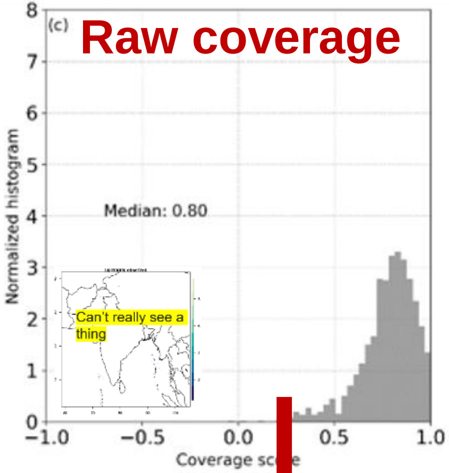
CDI results and sensitivity tests

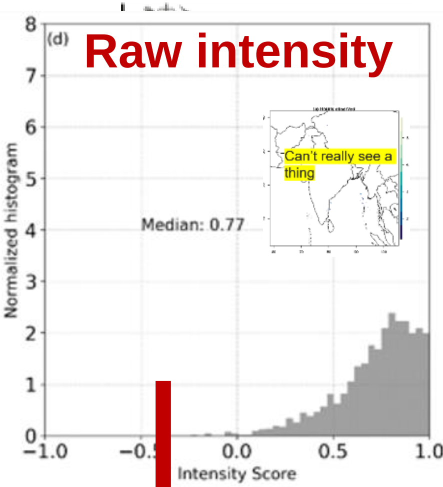
- Previously CDI was only computed with the “as is” observations.
- **Smoothed binary and mass-conserving gridded observations were used with the CDI method to investigate the sensitivity of representativeness mismatches** on the coverage score in particular (to see if the reduction in the representativeness mismatch is reflected there).
- As shall be seen, other impacts are also noted, which are interesting and worth noting for future exploration.

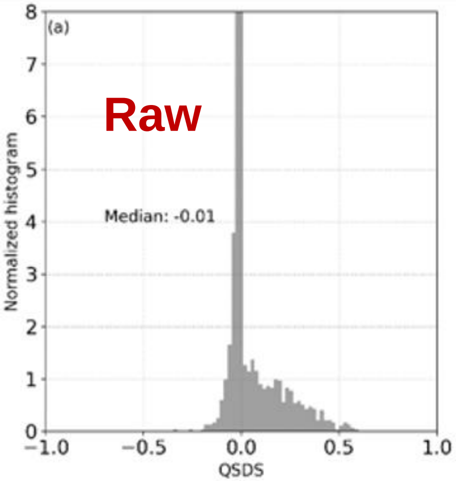


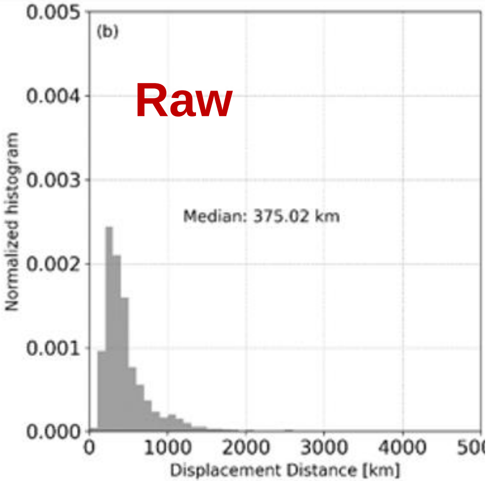
Distributional changes due to representativeness mismatches

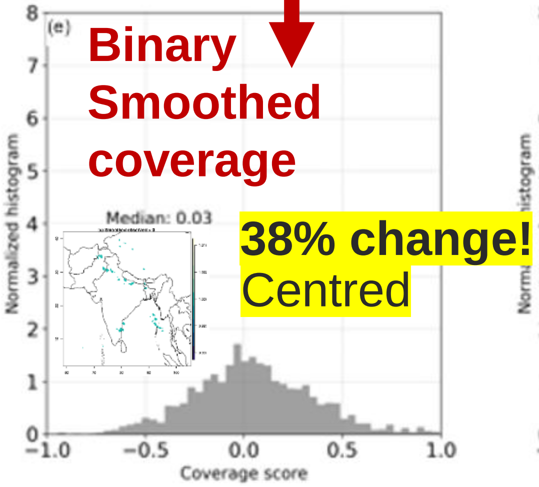
March to June 2019 and all lead times in the 24h forecast.

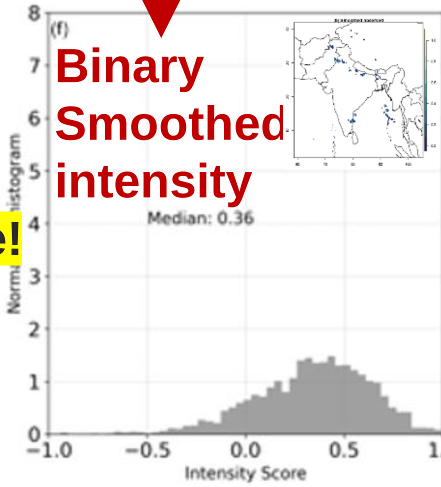












Diurnal cycle of domain average lightning

Met Office

Ministry of Earth Sciences

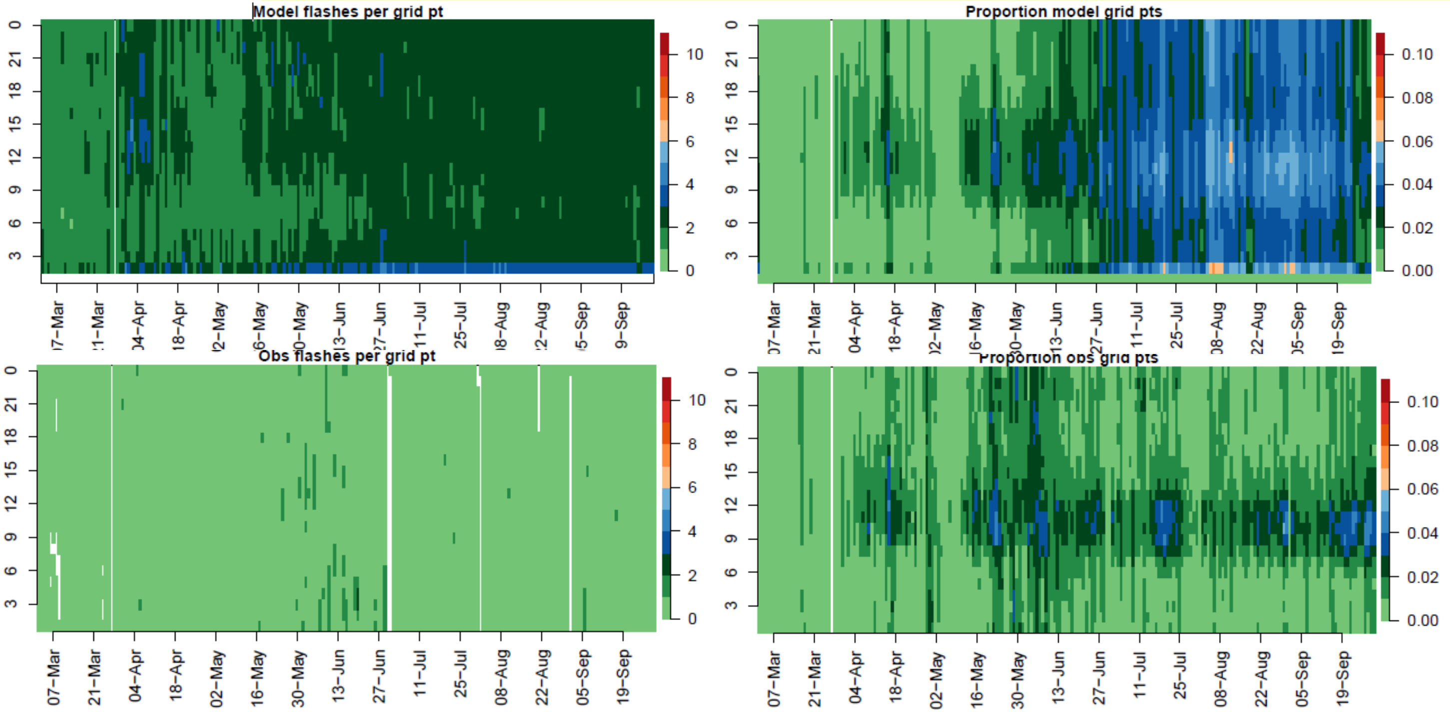
Department for Science, Innovation & Technology

Time series for the 2019 monsoon season.
Forecast initialisations on the x-axis with the 24h forecast scores plotted along the y-axis

Model and observed flashes per grid point aggregated over the domain along with the proportion of total grid points (900 x 900) that contain lightning for each.

Smoothed binary fields created using the Gaussian kernel smoothing increases the proportion of grid points with lightning by a factor of 4—5.

Mass-conserved smoothing reduces peak intensities.



Met Office

सत्यमेव जयते
Ministry of Earth Sciences

Department for Science, Innovation

CDI components

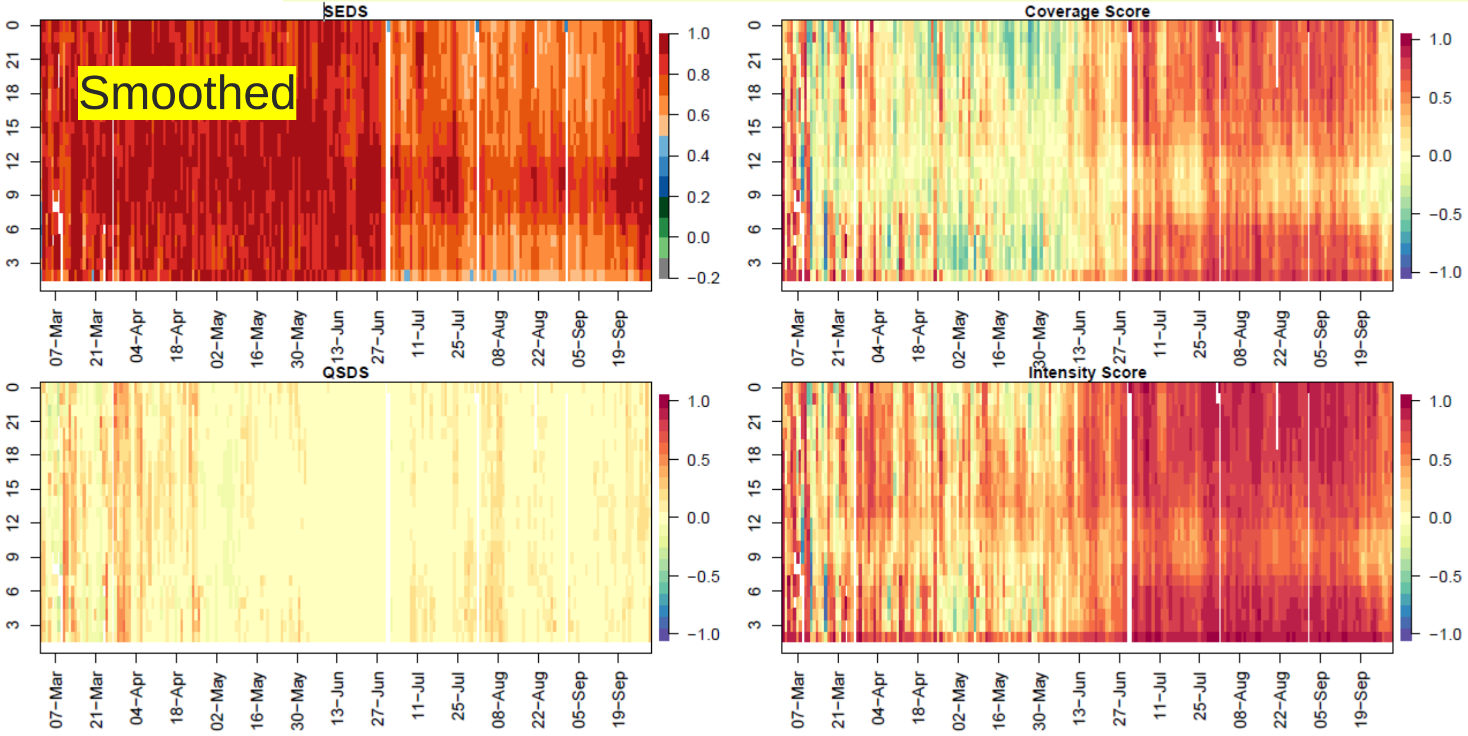
Time series for the 2019 monsoon season.
Forecast initialisations on the x-axis with the 24h forecast scores plotted along the y-axis

CDI components were calculated for the raw and smoothed binary gridded lightning observations

The onset of the monsoon leads to a marked change in the component scores related to location and intensity.

Raw observations show poor SEDS (bias dominating) but poor coverage intensity scores.

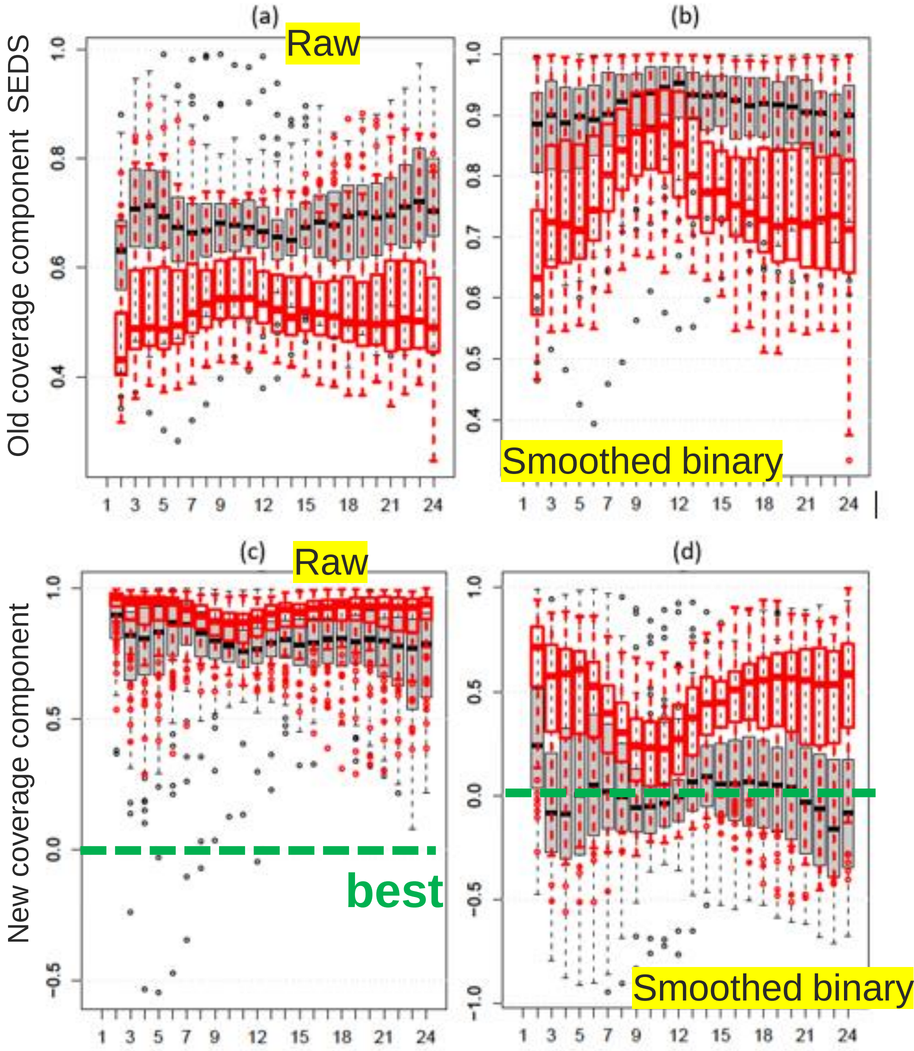
Smoothed binary observations show a marked improvement in SEDS, coverage and intensity scores (but improvements in intensity not for the right reasons).



Coverage stratification

Box plot representation of the **hourly SEDS and coverage component distributions** stratified by pre- (defined as MAM, **gray**) and main monsoon (JJAS, **red**) periods.
(a) Using the raw gridded observations and (b) Using the smoothed binary observations, (c) shows the coverage score with raw observations and
(d) the coverage scores with the smoothed binary observations.

Clear difference between the pre- and main monsoon seasons. The impact of the score is clear in (a) and (c) whereas the improvement in the score when taking representativeness into account is marked.



Met Office

सत्यमेव जयते
Ministry of Earth Sciences

Conclusions

Department for Science, Innovation & Technology

- CDI spatial verification method used to illustrate and quantify the impact of representativeness mismatches. There is a **38% change in the median coverage component for the period March-June 2019**.
- **Representativeness errors can skew the results** even for gridded methods affecting their interpretation. Representativeness mismatches should not be ignored.
- **Gaussian kernel dressing used in conjunction with a spatial verification method provides a means of dealing with extremely sparse observations by enlarging their footprint.** In doing so the representativeness mismatch between what the model can resolve and what the observation represents, is reduced and the method is able to correctly diagnose forecast behaviour.
- **CDI could be a suitable metric for quantifying the impact of representativeness errors** and for testing whether the right smoothing length has been applied before attempting to interpret any verification results.
- **Intensity biases are difficult to assess correctly because the lightning detection from a network is not 100%.** Lack of detection represents a source of observation uncertainty, which could be as low as 50%. Results presented here using the binary smoothed observed fields provide some insight into the impact of this aspect of observation uncertainty in an almost accidental or secondary way.

Mittermaier, M.P. and J. Wilkinson, 2024, WAF, accepted subject to minor revisions.

India partners