

# Report on ACCORD VS at GeoSphere

## Implementation and Validation of Spatial Verification Methods for Extreme Events

Carlos Peralta, Polly Schmederer and Philipp Scheffknecht

2025-08-11

### Introduction

During a three-week scientific visit at GeoSphere, the main focus was the **implementation and validation of spatial verification methods** for high-impact meteorological forecast assessment, particularly for extreme weather events. This work contributed to the ACCORD MQA initiative, targeting operational model evaluation and development.

### General Review of Verification Methods

A comprehensive review of advanced spatial verification techniques set the context for evaluating high-resolution weather forecasts, especially regarding the challenge of location-dependent forecast errors and the “double penalty” effect. Classical scores were contrasted with neighborhood-based and feature-based metrics[<sup>1</sup>review]. <sup>1</sup>

- **Traditional metrics:** TS (Threat Score), ETS (Equitable Threat Score).
- **Spatial metrics:** Address limitations of grid-point comparison and emphasize the importance of spatial structure<sup>2</sup>.

---

<sup>1</sup>review.

<sup>2</sup>review.

# Implementation of SLX Score

## Overview

### SLX Score

The **SLX score** was implemented to quantify the similarity between forecast and observed fields, systematically reducing penalties for small spatial displacement errors.

- **Key properties:**
  - Focuses on matching patterns and features within spatial neighborhoods.
  - Sensitive to both the location and the structure of extreme events.
  - Effective for high-resolution forecasts where traditional point-by-point metrics fail<sup>3</sup>.

### Agreement scales

---

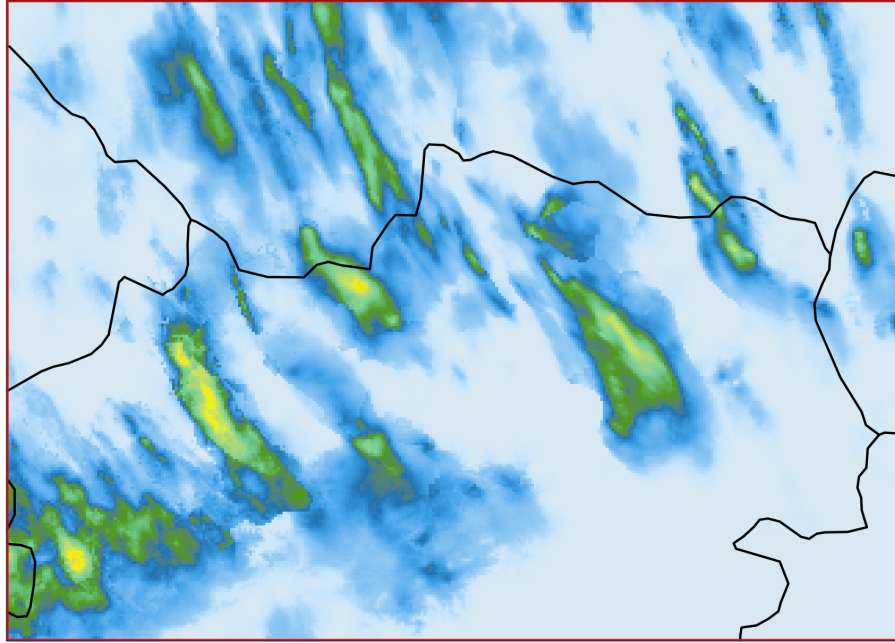
### Case studies

- Case 1: Convection driven event on July 28 2025 (Austria domain)

---

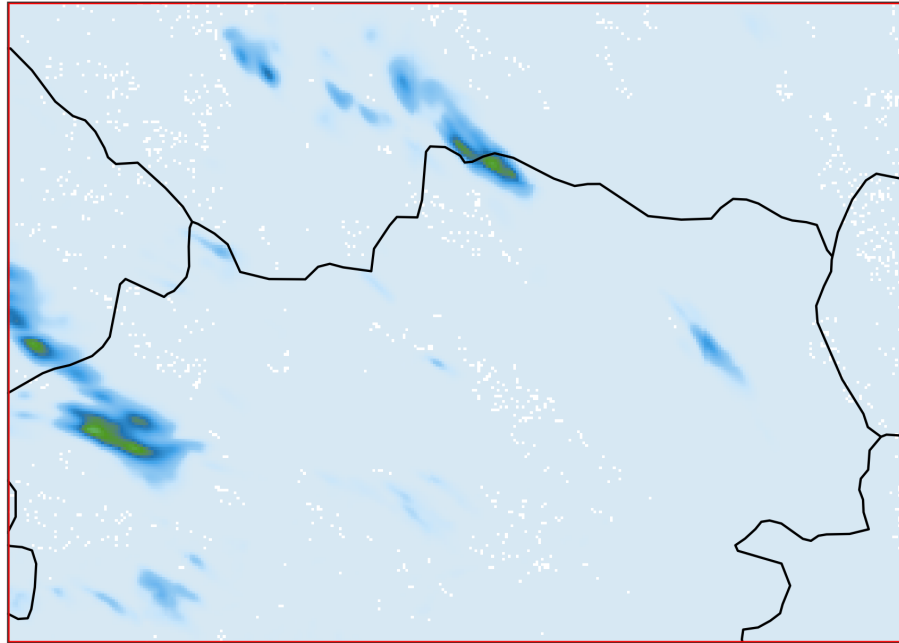
<sup>3</sup>See `review_sass/slx_presentation.qmd`; `slx_presentation.pdf`

Rain from INCAPlus



*Figure 1.1: INCAPlus gridded data*

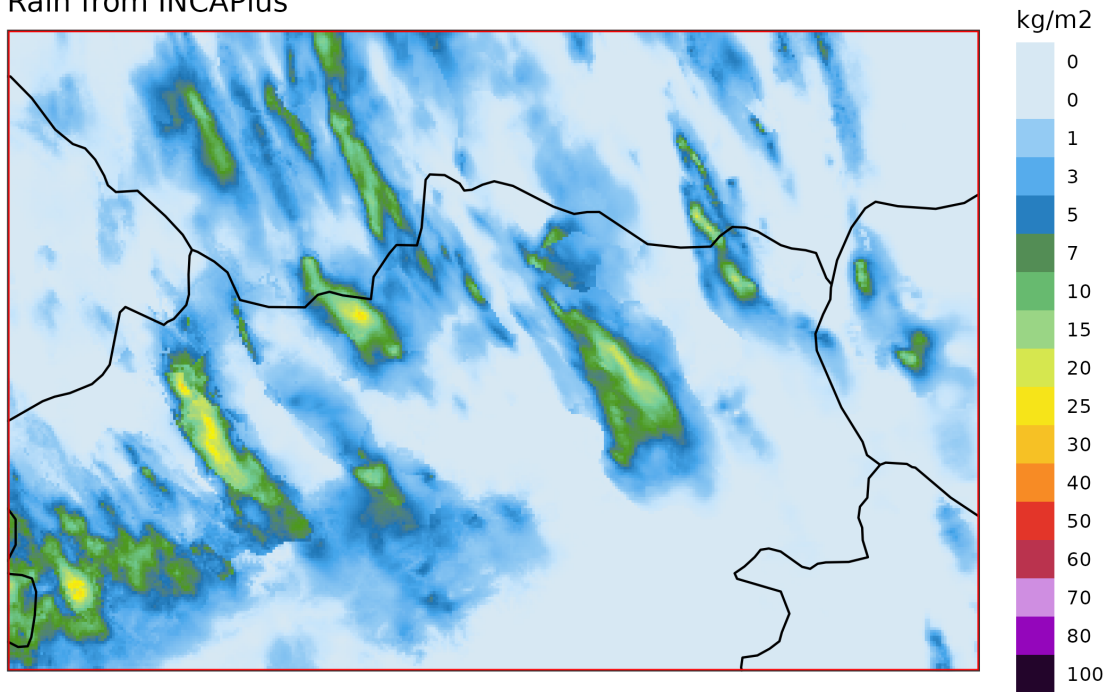
accrr1h 202507281800 2025-07-28 18:00:00 + kg/m2



*Figure 1.2: CLAEF 1km member 0*

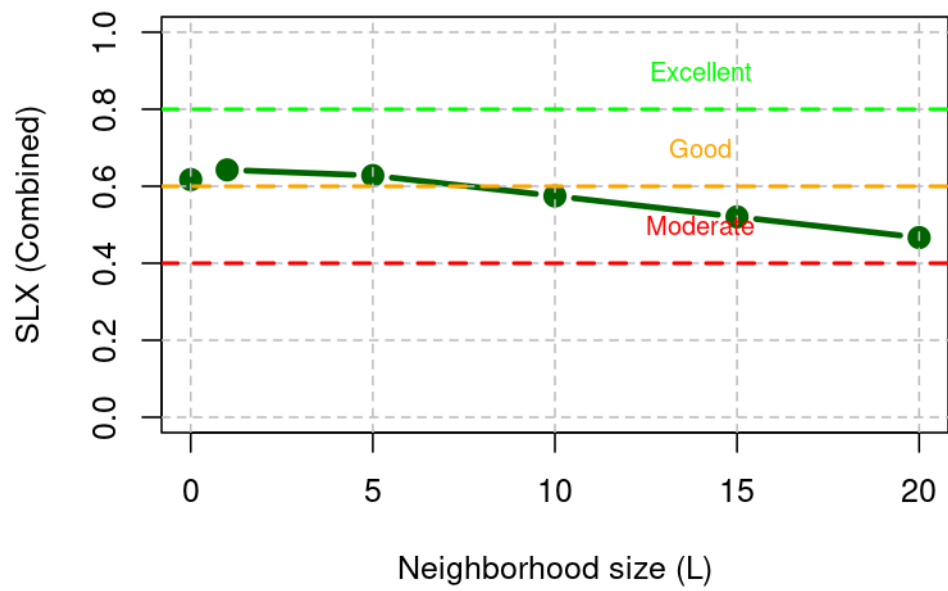
- **Case 2: Orographic driven precipitation event on July 8 2025 (Austria domain)**

Rain from INCAPlus



*Figure 1: OVerall SLX scores*

**Overall SLX Score vs Neighborhood Size**



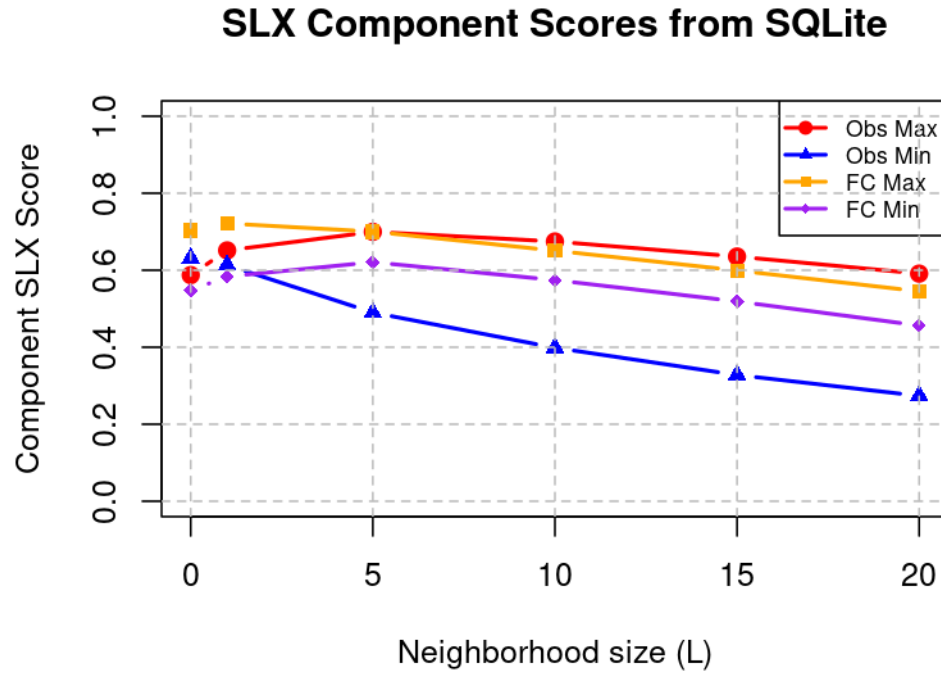


Figure 2: Individual SLX scores

- **Case 2: Field comparison for ensemble spread** Figure 2: Histogram of SLX scores across ensemble members

#### Findings:

SLX score provided enhanced discrimination of location errors and improved assessment of feature matching for extremes, outperforming classical methods in both sensitivity and operational utility<sup>4</sup>.

### Agreement Scales Score (Dey Method)

#### Method Summary

The **Agreement Scales Score** identifies the smallest neighborhood size at which forecast and observation fields “agree”, mapping scale-dependent skill and uncertainty.

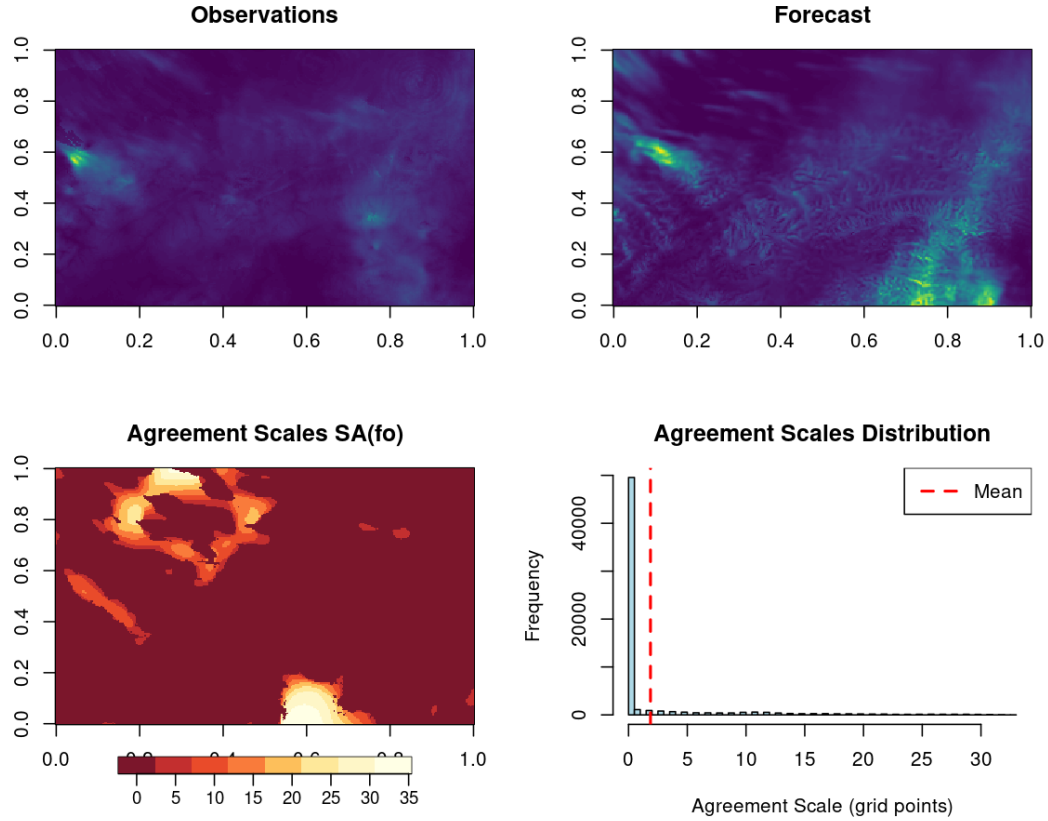
- **Key properties:**
  - Produces spatial maps of agreement scale.
  - Reveals local skill variations and ensemble spread.

<sup>4</sup>See [review\\_sass/slx\\_validation.qmd](#); [slx\\_validation.pdf](#)

- Enables targeted verification of prediction uncertainty for extremes<sup>5</sup>.

## Validation Example

- **Case: orographically driven test case on** *Figure 3: Agreement scale map for ensemble forecasts of a convective event*



### *Interpretation:*

The agreement scale identified regions of both high and low skill, illustrating the method's capacity to diagnose event-localized ensemble forecast confidence<sup>6</sup>.

Top Panels: Observations vs Forecast

Observations (top left): Very sparse, localized precipitation Forecast (top right): Much more extensive precipitation coverage

Bottom Left: Agreement Scales SA(fo) Map Looking at the color bar correctly (0-35 scale, darker = smaller values, lighter = larger values):

<sup>5</sup>See [review\\_dey/dey\\_agreement\\_scales\\_presentation.qmd](#); [dey\\_agreement\\_scales\\_presentation.pdf](#)

<sup>6</sup>See [review\\_dey/agreement\\_scales\\_validation.qmd](#); [agreement\\_scales\\_validation.pdf](#)



Dark red/purple areas (low values ~0-5 grid points): Good local agreement between forecast and observations - small neighborhoods needed  
 Light/white areas (high values ~30-35 grid points): Poor agreement - large neighborhoods needed to find any similarity

Bottom Right: Distribution The histogram shows:

Massive peak at small scales (0-5 grid points): Most grid points show good local agreement  
 Very few points at larger scales  
 Mean very low (~2-3 grid points): Overall good forecast-observation agreement

What This Actually Tells Us  
 1. Good Spatial Agreement Most of the domain (the dark areas) shows small agreement scales, meaning the forecast and observations agree well at the grid scale or small neighborhoods.  
 2. Limited Areas of Disagreement Only small portions of the domain (light areas) require large neighborhoods to find agreement between forecast and observations.  
 3. Successful Forecast Performance This appears to be a case where:

The forecast correctly predicted the dry areas (most of domain is dark = small agreement scales)  
 There are only localized disagreements where precipitation occurred in forecast vs observations  
 The SA(fo) method shows this is actually a reasonably good forecast when considering spatial uncertainty

## Results and Implications for ACCORD VS

The applied methods (SLX and Agreement Scales) successfully highlighted:

- Enhanced assessment of displacement and structure errors in extreme event forecasts.
- Improved operational diagnostics for ensemble spread and model skill.
- Practical, validated algorithms for routine use in operational forecast verification, supporting both feedback and model inter-comparison.

### Summary Table: Core Methods

Method	Purpose	Key Benefit
SLX Score	Spatial field pattern matching	Less penalization for small shifts, robust for extremes
Agreement Scales Score	Scale-dependent skill mapping	Local skill, uncertainty, ensemble spread visualization

## References

Include all references from the documents. For external scientific background, see: - Skok, G., et al. “Spatial verification of global precipitation forecasts,” arXiv (2025)[1]. - Hutchison, K., et al. “Climate Cloud Model Forecast Verification—an Engineering Perspective,” AEER (2020)[2].

Further internal references are embedded in the notebook via in-text citations to your documentation:

- ACCORD VS documentation:
    - `general_rev/review.qmd`
    - `review_sass/slx_presentation.qmd` & `slx_validation.qmd`
    - `review_dey/dey_agreement_scales_presentation.qmd` & `agreement_scales_validation.qmd`
    - `case_studies/results.md`
- 

## Appendix

### Plots and Code Snippets