# Increasing NWP thunderstorm predictability using ensemble data and machine learning

Kianusch Vahid Yousefnia   Christoph Metzl   Tobias Bölle

Deutsches Zentrum für Luft- und Raumfahrt, Institut für Physik der Atmosphäre, Oberpfaffenhofen, Germany

## Motivation

SALAMA 1D (see Poster P60) is a machine learning (ML) model determining the probability of thunderstorm occurrence, given an NWP forecast member. It is well-known that the skill of deterministic NWP forecasting systems can be extended by considering the ensemble mean:

NWP member #1 $\xrightarrow{\text{SALAMA 1D}}$ $p^{(1)}$

NWP member #2 $\xrightarrow{\text{SALAMA 1D}}$ $p^{(2)}$

...

NWP member #20 $\xrightarrow{\text{SALAMA 1D}}$ $p^{(20)}$

$$\rightarrow \langle p \rangle = \frac{1}{N_e} \sum_{k=1}^{N_e} p^{(k)}$$

- By how much does ensemble-averaging increase thunderstorm identification skill?
- What is the benefit of SALAMA 1D being an ML model?

## Ensemble-averaging indeed improves skill

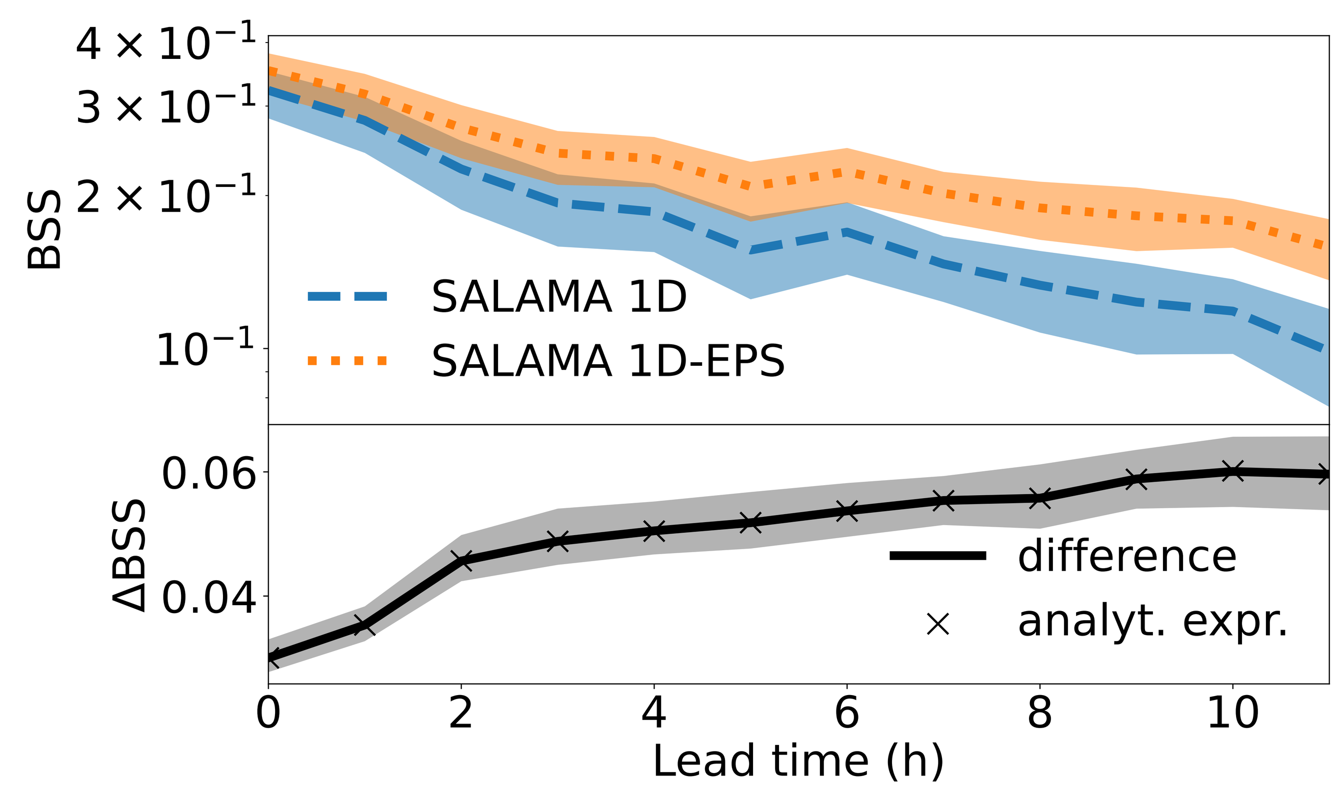


Figure 1. Lead-time dependence of skill, quantified by the Brier skill score (BSS), of single-member forecasts (SALAMA 1D) and ensemble forecasts (SALAMA 1D-EPS) of thunderstorm occurrence. Lower panel shows difference in skill, together with the prediction from the analytic expression (1). Shaded bands on this poster correspond to sampling uncertainty for a symmetric 90 % confidence interval.

## Exchange symmetry between NWP members

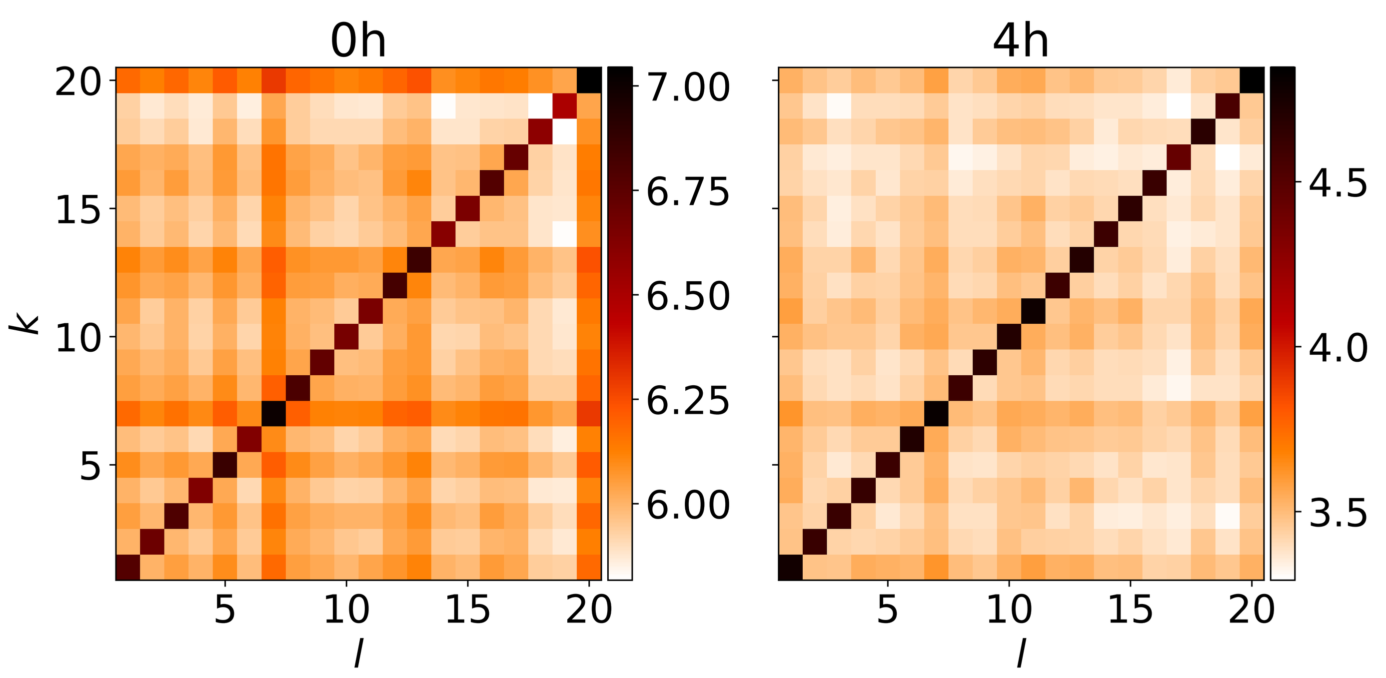


Figure 2. Sample covariance matrix $\mathrm{Cov}[p^{(k)}, p^{(l)}]/10^{-3}$, estimated for 0-hour lead time (left) and 4-hour lead time (right). If the members of the ensemble are exchangeable, the covariance matrix is fully determined by two numbers (one number for the diagonal entries of the matrix, one number for the off-diagonal entries), which is approximately the case.

$$\Rightarrow \mathrm{Cov}\left[p^{(k)}, p^{(l)}\right] = \begin{cases} \sigma^2 & \text{if } k = l \\ \gamma & \text{otherwise} \end{cases}$$

## Analytic formula for skill difference

Difference in Brier skill score (BSS) with climatology as reference:

$$\Delta \mathrm{BSS} = \frac{N_e - 1}{N_e g (1-g)} (\sigma^2 - \gamma) \tag{1}$$

with climatological probability $g$ of thunderstorm occurrence. This analytic formula is valid for all binary classification problems in severe weather forecasting!

## Machine learning extends thunderstorm predictability

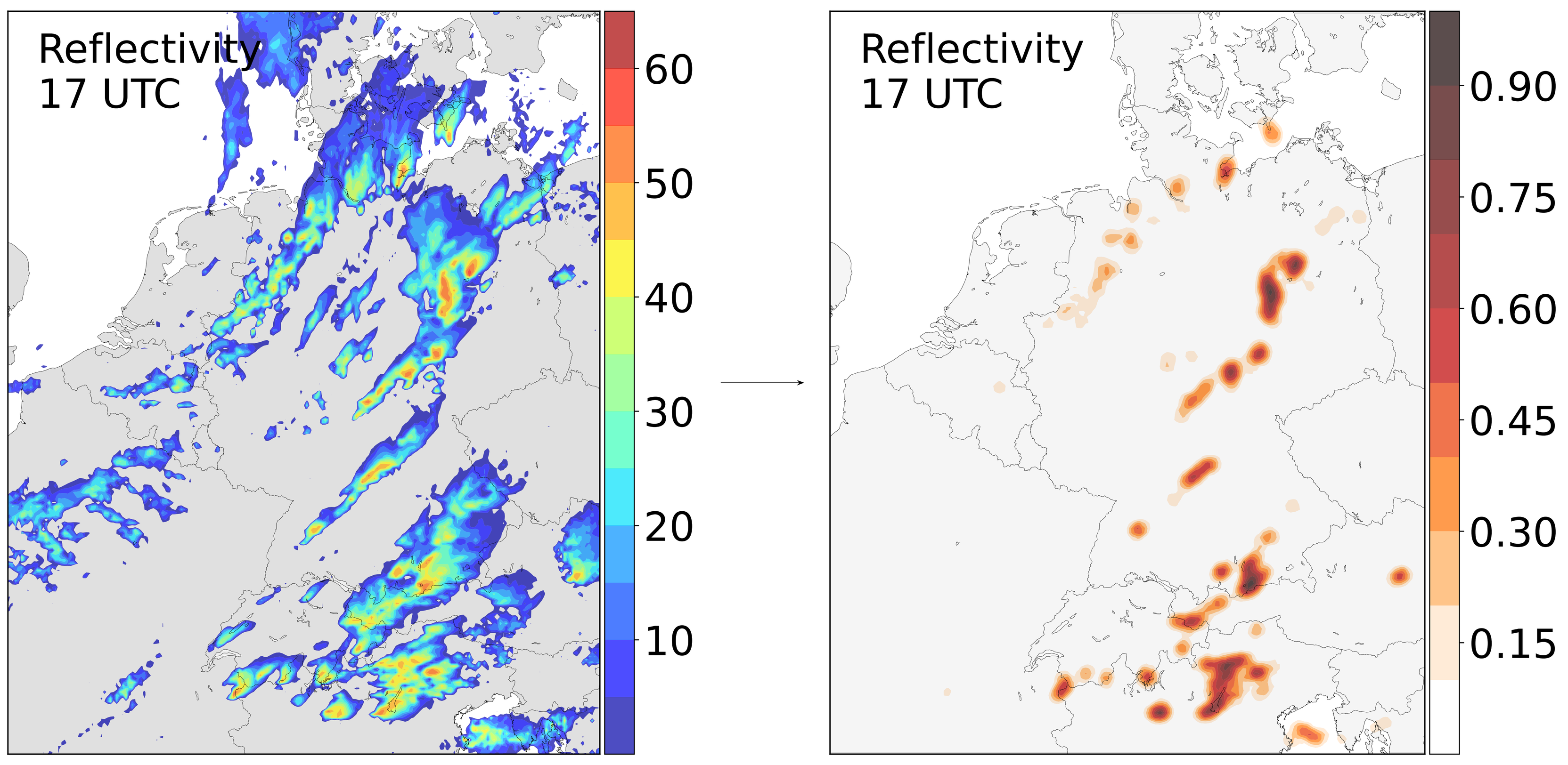


Figure 3. Proxy for convection in raw NWP forecasts: Fraction of grid points within $\Delta r = 15$ km for which 37 dBZ is exceeded.

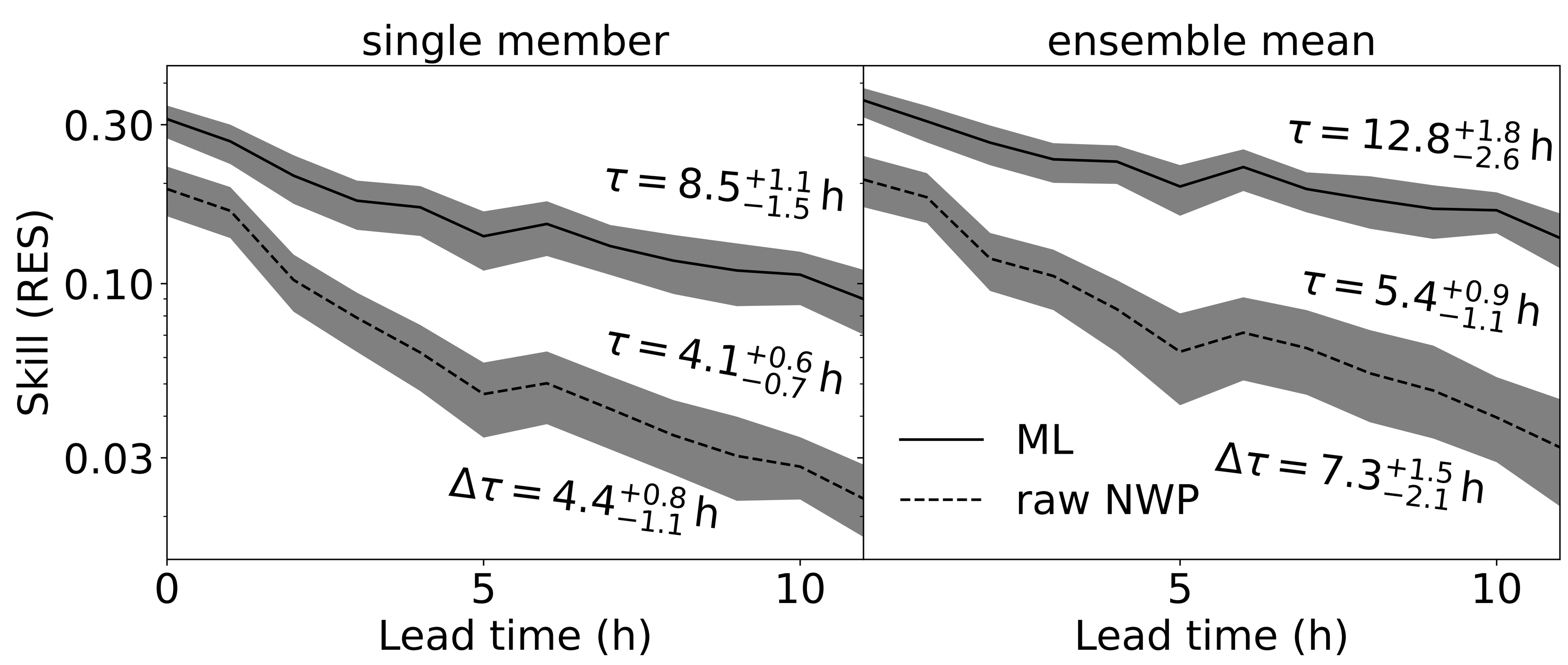


Figure 4. Lead-time dependence of skill for deterministic forecasts (left panel) and ensemble-averaged forecasts (right panel). Each panel displays the results for SALAMA 1D and a simple convection proxy based on raw NWP output without any ML corrections (Fig. 3). For each line, we fit an exponential function $\propto \exp(-t_{\mathrm{lead}}/\tau)$ to introduce a characteristic time scale $\tau$ of skill decay. Across all lines, the skill of ML-based forecasts decays more slowly than raw NWP forecasts.

## More information

K. Vahid Yousefnia et al. (2025): Increasing NWP Thunderstorm Predictability Using Ensemble Data and Machine Learning. https://arxiv.org/abs/2502.13316

K. Vahid Yousefnia et al (2025): Inferring Thunderstorm Occurrence from Vertical Profiles of Convection-Permitting Simulations: Physical Insights from a Physical Deep Learning Model. *Artif. Intell. Earth Syst.*, 4, 240096, doi: 10.1175/AIES-D-24-0096.1

The code for SALAMA 1D is available from https://github.com/kvahidyou/SALAMA

We present SALAMA 1D at Poster P60.

Deutsches Zentrum für Luft- und Raumfahrt
German Aerospace Center

*github.com/kvahidyou/SALAMA*
*kianusch.vahidyousefnia@dlr.de*