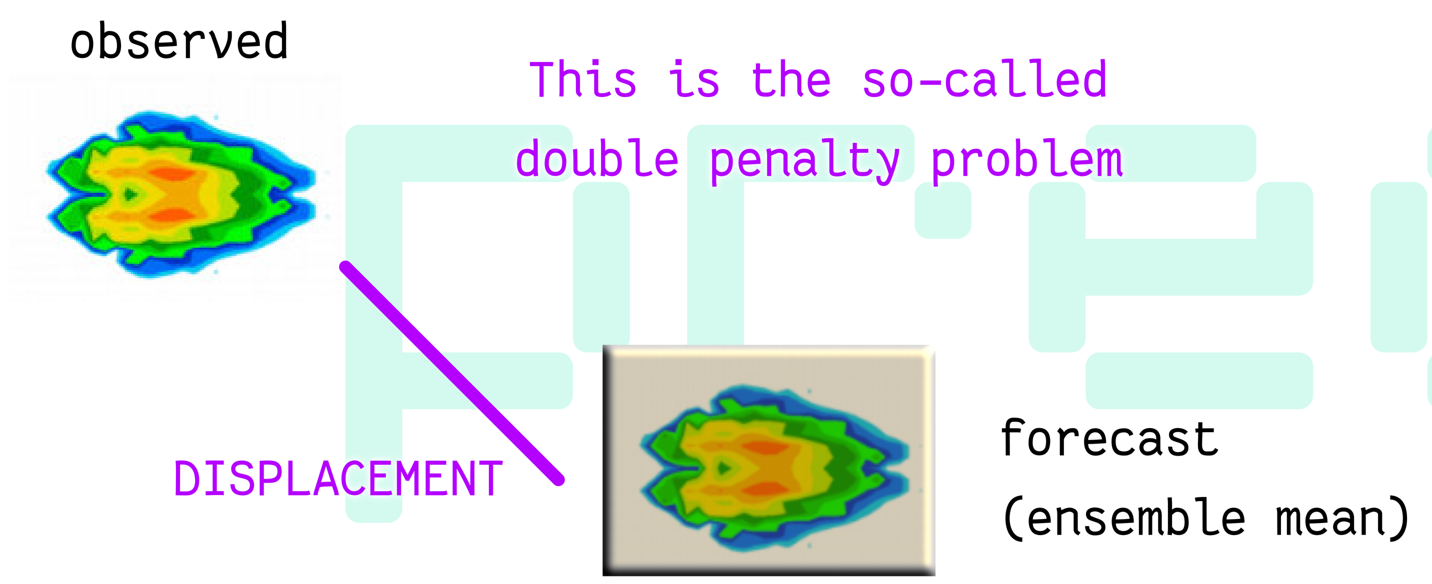
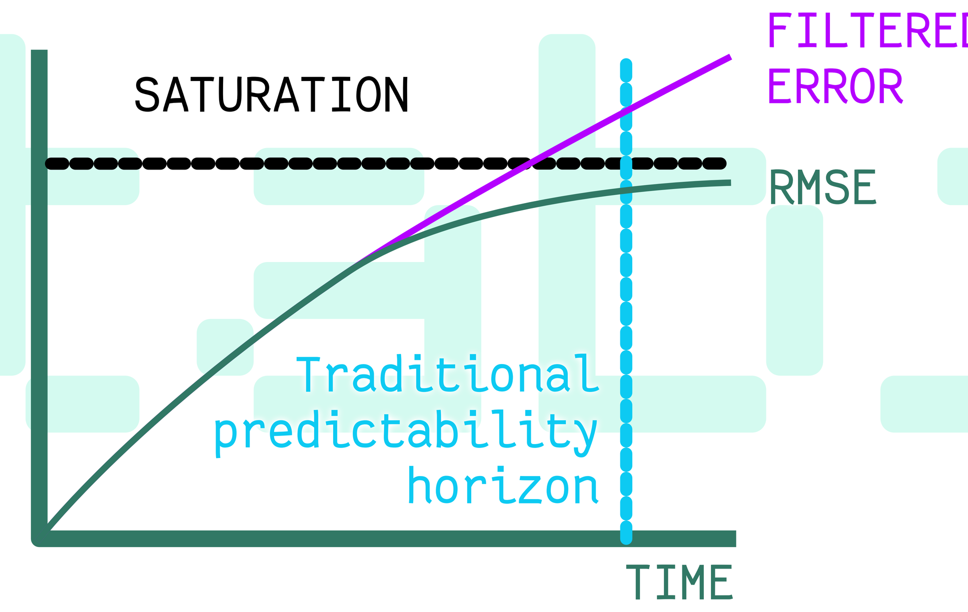


Consider a thunderstorm object identified in both observed and forecasted reflectivity. The forecasted storm has a locational (or timing) error equal to the diameter (or lifespan) of the storm itself.

Predictability: traditional versus filtered

Right: schematic showing "filtered" predictability based on object-based or scale-aware metrics doesn't saturate as quickly as defined by RMSE. The error value is not relevant here (different y-axes).



Traditional chaos theory defines a predictability horizon as reached when the forecast state is no more skillful than a state chosen randomly from the climatology (i.e., error growth slows, reaching saturation). But the forecast is useful: there *is* **information**. So how do we fix this?

Lorenzian predictability does not account for the filtering performed by the forecaster or user. The effective predictability horizon is in fact later in time.

We must reinterpret predictability on scales where sharp discontinuities yield unfairly severe penalties when performing evaluation, or where spatial and temporal tolerances are larger (e.g., issuing severe-weather polygons).

A long-lived supercell may be predictable from a phenomenological paradigm (i.e., a supercell is forecasted in most ensemble members but not in the same place) but not in the traditional sense (gridpoint-to-gridpoint evaluation).

For more, see Lawson (in review, JAS). Future work by JRL/CKP/MF will address the interpretation of predictability, and explore new evaluation metrics.

Information, Predictability, and Verification at the Thunderstorm Scale

John R. Lawson
CIMMS, Norman, OK, USA
NOAA/NSSL, Norman, OK, USA

Corey K. Potvin
CIMMS, Norman, OK, USA
NOAA/NSSL, Norman, OK, USA
OU School of Metr., OK, USA

Monte Flora
OU School of Metr., OK, USA
CIMMS, Norman, OK, USA
NOAA/NSSL, Norman, OK, USA

john.lawson@noaa.gov

Information Theory states that information, as a set of 1s and 0s (bits), measures the surprise that information carries; or, the minimum number of questions required to determine the state of a system. A large "surprise" (i.e. deviation from expectation) denotes a larger amount of Information.

Fractional Ignorance gives full reward to probabilistic forecasts in a way that Fractions Skill Score would not (i.e., an average of FSS over all members loses information regarding the ensemble pdf). FI, like FSS, is also scale-aware, reducing the double-penalty problem.

Fractional Ignorance: a scale-aware probabilistic metric

Fractional Ignorance was developed to bridge the gap between probabilistic and scale-aware verification schemes.

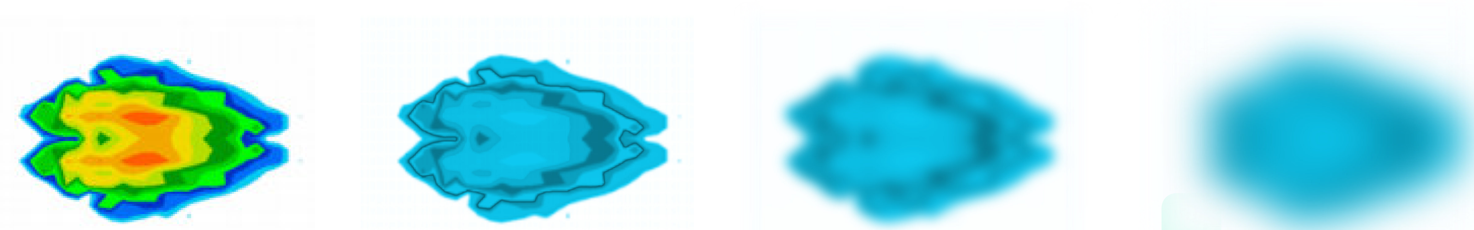
First, take an ensemble of forecasts and the observations for a given field. Let's say composite reflectivity. Threshold all fields (e.g., 30 dBZ) to create binary arrays.

Above: set point to 1 if forecast exceeds threshold, otherwise 0.

FI is proper, probabilistic, time-aware, scale-aware, has units of bits, is related to CRPS/FSS/cost-loss scores, does not require a climatology to measure skill.

For a given forecast time, choose spatial and temporal windows (e.g., 9 square kilometres; 15 minutes) and pixellate all ensemble members, and observations, at progressively larger windows.

This can be done with Fast Fourier Transforms, and accounts for tolerable errors in time and space.



Above: progressively pixellated supercells in the latitude-longitude dimension. Similarly for temporal dimension, if desired.

Apply the continuous/ranked form of Ignorance (CRIGN):

$$\text{CRIGN} = \text{REL} - \text{RES} + \text{UNC} \quad \text{(lower is better)}$$

to the resulting fractional fields. Repeat over multiple spatial and temporal windows for all times. (See ref: Tödter and Ahrens, 2012, MWR)

Ignorance Score measures information deficit in a probabilistic forecast, and can be decomposed into reliability, resolution, and uncertainty.

Right: difference in the skill score FISS (defined right) between a 3-km and 1-km ensemble forecast. Negative values indicate 3-km domain is more skillful than 1-km. This is due to poor reliability (REL) in this case.

$$\text{FISS} = (\text{RES}-\text{REL})/\text{UNC}$$