Probabilistic Cloud Cover Forecasting from an Ensemble

Travis M. Harty*, Scott McKinley*, William F. Holmgren†, and Antonio T. Lorenzo‡.

Program in Applied Mathematics*, University of Arizona
Department of Physics*, University of Arizona.
Department of Hydrology & Atmospheric Sciences†, University of Arizona

Summary

Background: Ensemble forecasts of cloud index (CI) are created through the advection of a satellite image with a wind field derived from NWP and optical flow combined through data assimilation [1].

Motivation: A probabilistic forecast can be used to create a probabilistic forecast, but must be calibrated to produce a reliable forecast.


Results: Both methods of calibration result in a probabilistic forecast with better reliability and modestly lower Brier score.

Introduction

• The ensemble (20 members) is generated by advecting CI fields derived from GOES-16 satellite images using an ensemble of cloud motion fields.
• Five months of data (March-July 2019) are used in this study. Only days which contain clouds over the Tucson region (40 km x 56 km area) are used.
• A schematic of the forecast system (showing 15 minute satellite resolution rather than 5 minute for simplicity) is shown below.
• When taking the ensemble mean, the forecast method presented here has a skill of 16% for 15 minute forecasts to 12% for 60 minute forecasts.

Analysis of ensemble

• In this study we generate a probabilistic forecast of the form \( P(X < b) \) for a single location (over the University of Arizona) with \( b \) equal to 0.2.
• We use two ways of generating a probability distribution directly from the ensemble:
  1) an empirical distribution in which \( P(X < b) \) is determined by the number of ensemble members less than \( b \).
  2) A Gaussian distribution with mean and standard deviation defined by the forecasted ensemble.
• We calculate the above quantities by convolving the forecasted fields with a truncated 2-D gaussian with a standard deviation of 2 km in order to avoid under dispersion and account for positional uncertainty.
• We assess our forecasts using the Brier score, the reliability, the resolution, and the uncertainty:

\[
BS = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 = \frac{1}{N} \sum_{i=1}^{N} (y_i - \hat{y}_i)^2 - \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2 - \frac{1}{N} \sum_{i=1}^{N} (\hat{y}_i - \bar{y})^2
\]

where \( y_i \) is the forecasted probability, \( \hat{y}_i \) is the observation, \( \bar{y} \) is the average observation, and \( \hat{y}_i \) is the average observation conditioned on \( y_i \).

Calibration function

• If the reliability of a forecast is perfect, then \( y_i = \hat{y}_i \) for all forecasts and reliability will be equal to zero. We therefore estimate a calibration function \( \hat{y}_i = \hat{d} y_i \) and apply this to our forecasted probabilities to calibrate our forecasts [2].
• The figures below show this calibration process. We use a training set of observations (20% of our observations) and fit a 3rd order polynomial from our probability forecasts to our observations. We then apply this to our testing data to determine the calibrated forecasts.
• We will apply this to the forecasts derived from the Empirical distribution.

Logistic regression

• Alternatively, we forecast calibrated probabilities directly from the mean and standard deviation of our ensemble using logistic regression [3].
• We will do this by fitting our test data to a logistic curve that is a function of the mean and standard deviation,

\[
P(X < b) = \frac{1}{1 + \exp(1 + a_0 + a_1 \bar{y} + a_2 \bar{d})}
\]

• The result of such a fitting can be seen to the right.

Results

• Below are the Brier score, reliability, and resolution normalized by uncertainty and the skill of the calibrated vs uncalibrated forecasts.
• Resolution is relatively unaffected by the calibration process, reliability is improved (decreased) for longer horizons and harmed (increased) for shorter horizons. This results in a corresponding change in the skill of the calibrated forecasts.

Conclusions

• All forecasts presented here show a significant improvement over the persistence ensemble.
• Calibration results in greater improvement over longer forecast horizons.
• The calibration function performed similarly to logistic regression.
• Separate calibration by month would likely result in improved forecasts if more data were available.
• Should increase resolution of forecast possibly by decreasing std. of convolved gaussian.

References