EMS Annual Meeting Abstracts
Vol. 15, EMS2018-510, 2018

Use of spatial predictors in clustered model output statistics (MOS) forecasting system

Jussi Ylhäisi (1), Laila Daniel (1), Matti Horttanainen (1), Marko Laine (1), Leila Hieta (1), and Reinhold Hess (2)
(1) Finnish Meteorological Institute, Helsinki, Finland (jussi.ylhaisi@fmi.fi), (2) Deutscher Wetterdienst, Offenbach, Germany

Weather forecast users are typically interested in point forecasts at their specific geographical location. However, forecasts interpolated to point-scale can have substantial errors caused mainly by two issues: Firstly, model values are not representative to those observed in the real world. Secondly, model forecasts themselves can have errors. Systematic error components of these sources can be tackled through bias correction, which essentially translates to downscaling grid-box scale model values to point-scale values. A traditional way of doing this is to derive station-based corrections through Model Output Statistics (MOS). Historical model data is used to derive some statistical relationship between model-observation pairs, which are then applied to real-time forecasts. Most typically, statistical model used in operative systems is multiple linear regression. These kind of postprocessing systems are currently in use at least in NOAA/NWS/MDL, DWD, FMI, and KMA.

To be able to realistically downscale model data to forecast users' location that is almost never near the surface station used in calibration, terrain characteristics data can be used. In this study, we present preliminary results on their use in downscaling. We use station-interpolated values of elevation and several other derived topography-related variables from a digital elevation map data with a 100 meter resolution. As the cause-effect relationship between measured values at the station site and the terrain characteristics can be highly complex, we derive the variables using several different resolutions of the underlying topography dataset. Finally, we feed our topography variables as predictors in DWD's MOS forecasting system which outputs ca. 150 variables. We train in-sample linear regressions for forecast hours 1...19 and three station clusters in Germany, each with different topography characteristics, using 5 years training data from COSMO-DE-EPS high-resolution LAM.

Over mountaneous station cluster, the topography variables receive the highest fractional weight whereas their contribution in the regression equations is, unsurprisingly, very small over the flat terrain station cluster. The performance of these variables varies with each forecast hour, but for individual forecast hours of several estimated variables the total contribution of all topography-related variables can sum up to $\sim$30% of accumulated RV (Reduction of Variance). The most important topography variables selected are the station-model grid box elevation difference on a 10 km resolution and the steepest slope elevation in the station vicinity on a 20 km resolution. The contribution of topography variables is clearly of secondary importance as compared to e.g. model data, but their strength is that they are available to the users' specific location on a high spatial resolution. This allows bias corrections to be made on a more high resolution than is available from the model data. Our results also encourage to experiment with other topography-derived predictors besides plain elevation difference.