8.7 Assessing ensemble forecast performance for select members available in the Community Leveraged Unified Ensemble (CLUE) during 2017 Hazardous Weather Testbed Spring Forecasting Experiment (HWT-SFE).

Wolff, Jamie K**.,** Michelle Harrold, **Lindsay Blank**, *National Center for Atmospheric research and Developmental Testbed Center (DTC),* and Jeff Beck, *Colorado State University/Cooperative Institute for Research in the Atmosphere/Affiliated with National Oceanic and Atmospheric Administration/Earth System Research Laboratory/ Global Systems Division, and Developmental Testbed Center (DTC)*

For the last several years, the Hazardous Weather Testbed (HWT) Spring Forecasting Experiment (SFE) has strived to coordinate the contributed model output from participating groups around a unified setup (e.g., WRF versions, domain size, vertical levels and spacing, etc.) to create a super-ensemble of over 60 members called the Community Leveraged Unified Ensemble (CLUE). The careful coordination and construction of CLUE allows for meaningful comparisons among a variety of members to be performed. The large datasets produced during the HWT-SFE provide an excellent opportunity to help identify and begin to answer the most pressing scientific questions that need to be addressed. In particular, many questions remain regarding the best approach to constructing a convection-allowing model (CAM) ensemble system. For example, should model uncertainty be addressed through multiple dynamic cores, multiple physics parameterizations, stochastic physics, or some combination of these? The careful coordination and construction of CLUE will provide the datasets necessary to begin to explore this question.

The forecast methods targeted for this presentation will include examining single physics vs. multi-physics approaches using CLUE 2017 data. Ultimately, the probabilistic forecast performance of each targeted ensemble subset will be examined. Individual deterministic forecasts from select members will also be to assessed to understand their contribution to the overall ensemble spread. The objective evaluation will be conducted using the Model Evaluation Tools (MET) software system. The metrics used for probabilistic and deterministic evaluation will range from traditional metrics widely used in the community (spread, skill, error, reliability, etc.) to newer methods that provide additional diagnostic information such as the Method for Object-based Diagnostic Evaluation (MODE) and neighborhood methods applied to deterministic and probabilistic output (e.g., Fractions Skill Score).