

Challenges In Evaluating Lower-Dimensional Features

Gad Levy, Process Study & Model Improvement (PSMI) Panel co-Chair

Following common practice in data assimilation schemes, most diagnostic tools and metrics for intercomparison of reanalyses, correct a model forecast (hincast) or background field of continuous variables based on optimal minimization of the model variable with respect to observed values, summed over some or all grid points in a discretization. Evaluated properly, this procedure allows for effective utilization of innovation, increments, and residuals to improve parameterizations and physical understanding. The least squares difference is often used as a basic measure of accuracy that is then normalized to form an agreement index/metric and to quantify the correction needed. These metrics are most appropriate for continuous fields where the observed and model variables are commensurate (i.e., measured with the same units). They are, however, flawed when used in the presence of sharp gradients and discontinuities and when used to evaluate a model's success in predicting or reproducing smaller scale lower dimensional features contained within a bulk simulation. These features, occur frequently in geophysical climate applications and often represent discontinuities that are associated with important climate physical and dynamic processes.

Ideally, a validation and intercomparison scheme should maintain the physical principles embodied in the model and be able to evaluate and utilize lower dimensional information (i.e., information contained within a bulk simulation even when not directly observed or represented by model variables). Nonetheless, physical principles are often violated, and lower dimensional information usually ignored. Conversely, models that resolve such information and the associated physics well, yet imprecisely are penalized by traditional schemes. This can lead to (perceived or real) poor model performance and predictability and can become deleterious in model improvements when observations are sparse, fuzzy, or irregular. It also impedes our ability to evaluate how well the models represent the relevant processes at different space and time scales, and what resolution is required to adequately simulate key processes. This talk intends to start a discussion on how address these issues.