SESSION: (B5) Hindcast and forecast quality assessment

(B5-03)

Signal and noise in regime systems: understanding NAO predictability

Strommen, Kristian (1), Palmer, Tim (2)

Oxford University

Over the last decade, NWP models have begun to exhibit significant skill at predicting the wintertime North Atlantic Oscillation (NAO) index. In particular, recent studies conducted by the UK Met Office reported significant skill over a hindcast period ranging from 1980 to 2015. At the same time, a very low signal-to-noise ratio was observed. This was captured using the `ratio of predictable components' (RPC) metric, designed to measure the extent to which reality is more or less predictable than the model itself. An RPC less than 1 indicates that the model is more predictable than reality, and greater than 1 that it is less predictable. A `signal-to-noise paradox' was described in Siegert et. al (2016) as the observation that, while the ensemble mean of their hindcasts correlated strongly with observations, the RPC was significantly greater than 1, suggesting the model is in some sense highly underconfident despite its notable level of skill.

Many studies have emphasized the role played by regimes in modulating North Atlantic weather. We will examine how actual skill, potential predictability and RPC are expected to behave in an idealized two-state regime system. Here the two states can be loosely be thought of as `positive NAO' (zonal jet structure) and `negative NAO' (more wavy jet structure), and regime transitions are assumed to take place on daily timescales. Predictability is assumed to be induced as a consequence of seasonal deviations of the persistence probabilities of the two states from their climatological means. Model skill is then assumed to be a consequence of the models ability to correctly capture these seasonal variations. We show that in this framework, the combination of high levels of skill and large RPC values can be expected for any model which captures the signal, but systematically reduces the overall level of regime persistence. In particular, a high RPC value on seasonal timescales may be expected even when the models level of internal noise on daily timescales is perfectly realistic. This lack of persistence is shown to be a common feature of several current NWP models.

Finally, we will compare the RPC metric with another commonly used method of evaluating over/under-confidence of models, namely the ratio of the RMS error to ensemble spread. It is shown that for models that are not statistically perfect, these metrics may not give the same answer, depending on if the underlying system is assumed to be highly linear, or e.g. a more non-linear regime-like system. This emphasizes the importance, when evaluating hindcast quality, of the choice of statistical model used to describe the system.