Applying tests of univariate Gaussianity on short range forecasts from an ensemble of variational assimilations.

Y. Michel[a] , R. Legrand[a]
[a] *CNRM-GAME, Météo-France and CNRS, Toulouse, France.*

Data assimilation schemes used in Numerical Weather Prediction have mainly relied on Gaussian Statistical Models. It is well known that non-linearities in the forecast model will produce non-Gaussian probability density functions, such that this assumption is questionable. Because the state and observation spaces have very large dimension, estimation of second order moments is already difficult, and estimating the full state probability density function or higher order moments is still an open area of research.

Measuring the deviation from Gaussianity produced by the non-linear model in a Gaussian-based assimilation cycle requires objectives measures or statistical tests [3]. The goal of this presentation is to describe the application of simple statistical tests on an ensemble of variational assimilations. In such an ensemble, each member consists in a cycled, four dimensional data assimilation (4D-Var), with assimilation of perturbed observations [2]. Such an ensemble is used to estimate background error covariances. Multiplicative inflation that insures that the short-range ensemble spread matches the forecast skill. The size of the ensemble has been extended to 90 members and the ensemble has been run over 10 days over a HyMeX period dedicated to the study of high impact weather over the Mediterranean [5].

Sample values of analysis and forecast errors for all variables and grid points are given to standard tests of Gaussianity, such as D'Agostino's chi2 [4] and Anderson-Darling [1] tests. Small samples almost always pass a normality test. Normality tests have little power to tell whether or not a small sample of data comes from a Gaussian distribution. With large samples, minor deviations from normality may be flagged as statistically significant (physical bounds on variables...). Thus, a hundred members is a typical size where those tests can indicate likely deviation from Gaussianity with enough power. We will describe those local deviations from Gaussianity as a function of the variables considered, the vertical levels, and the analysis-forecast cycle.

References

[1] Anderson, T.W. and Darling, D.A., 1954 : A Test of Goodness-of-Fit. Journal of the Am. Stat. Association 49 : 765-769.
[2] Berre, L. and G. Desroziers, 2010 : Filtering of Background Error Variances and Correlations by Local Spatial Averaging : A Review. Mon. Wea. Rev., 138, 3693-3720.
[3] Bocquet, M. et. al, 2010 : Beyond Gaussian Statistical Modeling in Geophysical Data Assimilation. Mon. Wea. Rev., 138, 2997-3023.
[4] D'Agostino, R.B., 1970 : Transformation to normality of the null distribution of g1. Biometrika 57 (3) : 679-681.
[5] Michel Y. et. al, 2012 : A large ensemble of data assimilations for studying forecast errors and predictability over the Mediterranean. 6th HyMeX Workshop, Croatia, May 2012.