Using statistical data to improve weather forecasting: a variance-constraining Kalman filter for data assimilation

Lewis Mitchell*

An important problem in numerical weather prediction is how to optimally estimate the full state of the atmosphere at any one time given noisy observations and an imperfect model. Data assimilation aims to solve the problem of finding this optimal atmospheric state, called the *analysis*. The process is complicated by the fact that the dimension of the system is generally orders of magnitude higher than the number of observations available, meaning that the problem is drastically underdetermined. Furthermore, the chaotic dynamics of both the atmosphere and the model can lead to a situation in which the analysis moves far away from the truth over time, or an ensemble of initial approximations of the state become unbounded in finite time. This problem is called *filter divergence*.

One of the standard tools for doing data assimilation is the ensemble Kalman filter (EnKF). Using this technique, one propagates an ensemble of initial conditions forward in time, and then uses the mean and variance of this ensemble to solve a least-squares minimisation problem to find the optimum state estimate which combines information from the forecast and the observations into the analysis [1].

We consider the case when the state space can be split into *observables*, for which observations are available, and *pseudo-observables*, for which only climatic information of a statistical nature is a priori known. Such a situation arises in practical applications when observations are only sparsely available, as well as in the mesosphere, where no direct observations are available. However, it is known that dynamically active gravity waves propagate [3]. These fast waves are spread over a large frequency range, and so the analysis may benefit from added information on the statistical behaviour of the gravity waves in the mesosphere.

In [2] we derive a variance-constraining Kalman filter (VCKF) which incorporates information from both observables and pseudo-observables to improve the forecast. We apply the filter to a pair of noisy coupled linear oscillators, which allows us to analytically study the VCKF and show under what conditions it outperforms a standard EnKF. The analysis shows that if the time scale of the pseudo-observables is much shorter than that of the observed variables, meaning if pseudo-observables have relaxed towards equilibrium and acquired the climatic covariance within a short enough time, then the variance constraint is beneficial. Thus, observing only

Lewis Mitchell was awarded the T.M. Cherry Prize for the best student presentation at ANZIAM 2010. This extended abstract is an invited contribution to the *Gazette*.

*School of Mathematics and Statistics, University of Sydney, NSW 2006.
E-mail: lewis.mitchell@sydney.edu.au

the 'slow' dynamics of the system and using averaged statistics only for the 'fast' dynamics may be sufficient for accurate assimilation. We then apply the VCKF to a more complicated nonlinear system, the Lorenz-96 model, a standard toy model for midlatitude atmospheric dynamics. Numerical simulations show that the VCKF analysis skill is better than that of the EnKF for both the observed and unobserved variables, and that the VCKF helps to stabilise the filter and avoids catastrophic filter divergence.

References

[1] Evensen, G. (2006). *Data assimilation: the ensemble Kalman filter*. Springer, New York.

[2] Gottwald, G., Mitchell, L. and Reich, S. (2010). A variance constraining Kalman filter. (Submitted.)

[3] Polavarapu, S., Shepherd, T.G., Rochon, Y. and Ren, S. (2002). Some challenges of middle atmosphere data assimilation. *Q. J. R. Meteorol. Soc.* **128,** 1–11.

Lewis studied mathematics and physics at the University of Wollongong, and graduated in 2007 with Honours in applied mathematics and the University Medal for informatics. He is currently studying for a PhD on dynamical systems techniques in data assimilation under the supervision of Georg Gottwald.