ABOUT THIS ISSUE

What If

Years after taking a course in fluid dynamics, it occurred to me that the wind and temperature of the daily weather are variables in a fluid system that I happen to live inside of. The dynamics of this fluid system, governed by Euler, Navier-Stokes, and thermodynamic equations, are yet more complicated than textbook fluid dynamics due to rain, snow, and hail, all happening globally and at a wide range of altitudes. The gentle or strong wind I feel at any given moment is just a sampling of the velocity field at that location and instant; in other words, when you experience the weather, you're collecting data.

Digital Object Identifier 10.1109/MCS.2009. 932393

Contributors

S. Lakshmivarahan at the Lone Sailor Memorial, Vista Point, at the northern end of the Golden Gate Bridge in San Francisco, California.

Geir Evensen atop Austabotntind in Norway.

David Stensrud returning to the boat after parasailing over the ocean near Moreton Island, Australia.

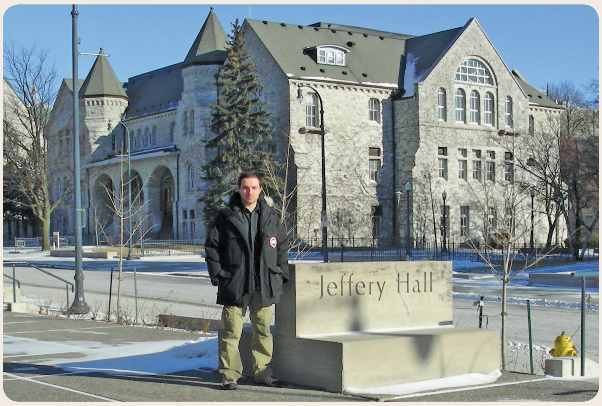
Jan Mandel.

Predicting the weather well is useful if you might need an umbrella but indispensable when lives are at stake. Although it's hard to predict the future, it helps to have a good initial condition. That's a basic rule in weather forecasting. Measurements—of wind speed and direction, temperature, and other quantities—help to determine that initial condition, but measurements in a distributed system are always more sparse than we would like. An estimator based on a model can fill in the gaps.

You might not have realized that variants of the Kalman filter (which, as pointed out to me by Bob Bass, should really be called the Kalman *estimator*) are used to predict the weather, or at least determine an initial condition to predict the weather. But what's interesting is that the "Kalman filter" used for weather prediction is not the Riccati-equation-based algorithm you learned in school and might have used for a missile or two. Since a weather model can have millions of states, the matrices in the Riccati equation would be enormous—too large for any modern computer to conveniently work with. If you were absolutely determined to do so, you could linearize the weather model and obtain these matrices, but weather scientists don't do that. So how do they do it?

The idea is both simple and surprising: Make multiple copies of the dynamics of the system (called an ensemble), give each copy in the ensemble a random initial condition and random driver (sample of the disturbance), and then propagate (simulate) each copy forward in time. When new measurements are available, stop the propagation and combine the values of the states from the ensemble members to update the error covariance. This matrix is used for the next step in which the most recent measurements are injected into the model. This two-step process appears in traditional Kalman filtering but has


Minjeong Kim.



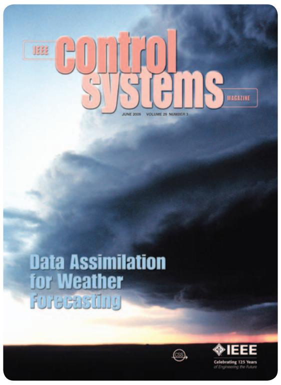
Pantelis Isaiah on the campus of Queens University.


Jonathan Beezley and his niece Dianna.


Jeffrey Anderson with his daughters Jennifer and Michelle.

a name specific to large-scale applications: data assimilation.

Note that the forecast step in data assimilation uses the ensemble of simulations to estimate the error covariance and thus avoids the need for linearization. Without having to linearize the dynamics, the implementation process is fairly easy. In fact, if the dynamics of a system were given by a computer program that wasn't based on any equations at all, that would be just as easy to handle: Just run copies of the program. For data assimilation, the measured data must be injected into the system, and this step is accomplished by mimicking the form of the Kalman filter gain.



What is surprising about ensemble methods is that you would expect that a Monte Carlo method of this type would require an enormous number of ensemble members to capture the statistics of the state. For large-scale systems, such as weather forecasts—which is only one of many applications—the number of ensemble members is typically 50–100. This range may reflect the available computational resources more than the ideal number, but the fact that good results are obtained with this modest number is no less than astounding. Determining the minimal ensemble size for the desired accuracy is one of the principal theoretical challenges of this approach.

Another remarkable feature of ensemble methods is the obvious fact that the ensemble members can be propagated on different processors; in other words, the algorithm is trivially parallelizable. That is not to say that you could not parallelize a Riccati equation with matrices of size 1e6 by 1e6, but I would rather not write such a program, much less derive the linearized matrices.

There are numerous varieties of ensemble Kalman filters, including the unscented Kalman filter (UKF), which is a deterministic variant involving the ideal number, namely, $2n+1$, of ensemble members, where n is the order of the system. Ensemble methods can be traced back to "particle filtering," which seems to have been discussed as early as the 1950s. Interestingly, UKF applied to linear systems gives exactly the same estimates as the classical result. It would be fun to speculate on how systems and control theory would have evolved if UKF had been discovered before the Riccati-based variety. In fact, matrix equations are becoming increasingly quaint and dated as estimation theory becomes increasingly Riccati free.

In this issue of *IEEE Control Systems Magazine*, we bring you four articles on data assimilation, that is, ensemble-type estimation methods for large-scale systems. The first two articles are applications oriented. The first article, by S. "Varahan" Lakshmivarahan and David Stensrud, provides an overview of the challenges of data assimilation for meteorology, tracing the development of numerical weather forecasting including data assimilation.

Next, Jan Mandel, Jonathan Beezley, Janice Coen, and Minjeong Kim describe the application of data assimilation to predicting the evolution of forest fires. This application is challenging due to uncertainties in the fuel content of brush, topographical features of the landscape, and feedback interaction with the weather.

The next two articles focus on data assimilation algorithms. In their respective articles, Jeffrey Anderson and Geir Evensen analyze the accuracy of the estimates provided by data assimilation algorithms in relation to ensemble size, model nonlinearity, and model error.

The image used to illustrate each article is a photo composite courtesy of David Stensrud. It is a severe thunderstorm moving across northern Oklahoma with a low-level, tube-shaped roll cloud seen stretching horizonally ahead of the storm along the boundary between the cold thunderstorm downdraft air and the warmer environmental air ahead of the storm.

This issue also brings you an "Ask the Experts" column in which Pantelis Isaiah provides a tutorial discussion of the control-theoretic principles behind parallel parking. The "Technical Activities" column in this issue reviews the history of the Technical Committee on Computer-Aided Control Systems Design. This is the first column on technical committee activities in 2009 courtesy of Shuzhi Sam Ge, CSS vice-president for Technical Activities.

For "People in Control," we have interviews with Rob Stengel and Martha (Gallivan) Grover. In addition, we present the 2008 IEEE Fellows elected by the IEEE Control Systems Society (CSS). Finally, we introduce the new associate editors for the magazine.

This issue includes as well two book reviews, two conference reports, and a new department that we hope to make a yearly event, a preview of the upcoming CDC. In 2009 our flagship conference will be held—for the first time—in China. This article is aimed at encouraging you attend this special CDC whether or not you submitted a paper. Finally, we end this issue with a special call for artistic renditions of the CSS logo. I look forward to your creative contribution.

Dennis S. Bernstein