Propagating subsurface uncertainty to the atmosphere using fully-coupled, stochastic simulations

JOHN L. WILLIAMS, III[1] & REED M. MAXWELL[1,2]

1 Colorado School of Mines, Department of Geology and Geological Engineering, Hydrologic Science and Engineering Program, 1500 Illinois St, Golden, Colorado 80401, USA
johwilli@mines.edu

2 Integrated Groundwater Modeling Center, 1500 Illinois St, Golden, Colorado 80401, USA

Abstract Feedbacks between the land surface and the atmosphere, manifested as mass and energy fluxes, are strongly correlated with soil moisture under dry conditions, making soil moisture an important factor in land–atmosphere interactions. We show that uncertainty in subsurface properties propagate into atmospheric variables, and therefore reduction of uncertainty in hydraulic conductivity will propagate through land–atmosphere feedbacks to yield more accurate weather forecasts. Using ParFlow-WRF, a fully-coupled groundwater-to-atmosphere model, we demonstrate responses in land–atmosphere feedbacks and wind patterns due to subsurface heterogeneity. An idealized domain with heterogeneous subsurface properties is used in ensembles of coupled-simulations. These ensembles are generated by varying the spatial location of the subsurface properties, while honouring the global statistics and correlation structure, an approach common to the hydrologic sciences but never-before used in atmospheric simulations. We clearly show that different realizations of hydraulic conductivity produce variation in soil moisture, latent heat flux and wind for both point and domain-averaged quantities. A single random field is chosen as the "actual" case and varying amounts of hydraulic conductivity data are sampled from this realization. Using these conditional Monte Carlo simulations, we incorporate subsurface data into the ensemble of realizations. We also show that the difference between the ensemble mean prediction and the actual saturation, latent heat flux and wind speed are reduced significantly via conditioning of hydraulic conductivity. By reducing uncertainty associated with land–atmosphere feedback mechanisms, we also reduce uncertainty in both spatially distributed and synoptic wind speed magnitudes, thus improving our ability to make more accurate forecasts important for many applications such as wind energy.

Key words uncertainty; land–atmosphere feedbacks; wind; heterogeneity

INTRODUCTION

The direct effects of subsurface heterogeneity have not been included in atmospheric studies to date. While the land surface and groundwater have historically been treated as simplified systems in atmospheric forecast and prediction models (Golaz *et al*., 2001; Kumar *et al*., 2006), early work by Chen & Avissar (1994), among others, has shown that soil moisture has a profound effect on local and mesoscale atmospheric processes. It has also been shown in work by Betts *et al*. (1996), Beljaars *et al*. (1996), Seuffert *et al*. (2002) and Holt *et al*. (2006), for example, that more advanced land surface model formulations and initialization, which generate more realistic soil moisture fields, result in better skill in mesoscale, regional and local scale weather forecasts. The reliance of these land surface parameterizations on accurate representations of surface soil moisture fields can be problematic because soil moisture is a transient quantity that is variable and heterogeneous in space and time (Wendroth *et al*., 1999; Western *et al*., 2004; Famiglietti *et al*., 2008). Hydraulic conductivity, while highly variable and heterogeneous in space (several orders of magnitude), is static in time and has been shown to exhibit spatial correlation (Rubin, 2003). The uncertainty in the hydraulic conductivity correlated random field can be evaluated through multiple realizations in Monte Carlo ensemble simulations (Gelhar, 1986), and reduced by assimilating observational data (Rehfeldt *et al*., 1992).

Ensemble, or stochastic approaches are common in both the atmospheric and hydrologic/hydro-geologic sciences. However the approaches used in each of the communities differ significantly. Atmospheric ensembles, from numerical weather prediction to climate change simulations, are commonly generated through perturbations of initial conditions and choice of model parameterization (e.g. Leutbecher & Palmer, 2008), while ensembles in hydrogeology are motivated through uncertainty in input parameters, typically spatial variability in the hydraulic conductivity, K (e.g. Criminisi *et al*., 1997; Nowak *et al*., 2010). A common subsurface characterization approach in risk assessment, solute transport and aquifer remediation studies employs Monte Carlo simulation ensembles to back-calculate *K* using observations of solute concentration or arrival times to condition realizations of the subsurface (e.g. Graham & McLaughlin, 1989; Katul *et al*., 1993; James & Gorelick, 1994; Harvey & Gorelick, 1995; Yeh *et al*., 2005). Another approach directly conditions subsurface realizations to observed hydraulic conductivity data by assimilating point observations into a statistical representation of the subsurface (Maxwell *et al*., 1999).

We apply a conditioning method whereby the distribution of hydraulic conductivity values in a correlated stochastic random field is controlled by enforcing "observed" point values drawn from a control, or "truth," simulation using a linear regression technique through which the stochastic random field honours both the observational data and the specified global statistics (Goovaerts, 1997). Using a Monte Carlo simulation technique for both unconditioned and conditioned simulations, we show that hydraulic conductivity, saturation, latent heat flux and wind speed magnitudes more closely match hypothetical observed data, and that improvements in atmospheric ensembles can be achieved by assimilating subsurface data.

METHODS

We use PF.WRF to simulate subsurface, surface and atmospheric conditions in a hypothetical 15 × 15 km basin. PF.WRF is a combination of the mesoscale Weather Research and Forecasting (WRF) atmospheric model (Skamarock & Klemp, 2008) and ParFlow, a 3D variably-saturated subsurface model that simulates both subsurface and surface flow via an overland-flow boundary condition (Ashby & Falgout, 1996; Jones & Woodward, 2001; Kollet & Maxwell, 2006). The two models are coupled via mass and energy fluxes passed through the Noah Land Surface Model (Chen & Dudhia, 2001), resulting in a single model of the hydrologic cycle (Maxwell *et al*., 2010). Details of the coupling process, along with model equations, are presented by Maxwell *et al*. (2010).

Using PF.WRF, we ran four sets of Monte Carlo simulations. Each simulation comprised 10 realizations using different, yet statistically equivalent, heterogeneous *K* fields in the subsurface. One additional realization, with its own random seed not included in the unconditional ensemble, was used to represent the control (*CTRL*) conditions of the hypothetical domain. The conditioned sets of Monte Carlo simulations used the same statistical parameters and random seeds to generate the random fields as the unconditioned sets, but each was conditioned with an increasing number of data points drawn from the *CTRL* hydraulic conductivity field. We ran simulations with 60, 120 and 200 points of conditioning data (hereafter referred to as *CO60*, *CO120* and *CO200*, respectively) sampled from the *K* field of the *CTRL* case, in addition to the unconditioned (*NO*) case. Specifics of the methods used in this study are detailed by Williams & Maxwell (2011).

RESULTS AND DISCUSSION

We focus our analysis on saturation, latent heat flux—variables which provide an indicator of surface conditions as they relate to land–atmosphere feedbacks—and wind speed magnitude as the primary atmospheric variable of interest. We also focus on these variables as we expect the most direct and significant effect of conditional simulation on *K*, then saturation, then latent heat and finally wind speed magnitude, spanning the subsurface to the atmosphere. We show results only for the unconditioned (*NO*) and 200 conditioning points (*CO200*) endpoint cases. Results from the cases with 60 and 120 conditioning points fall within the bounds of these endpoints.

We first examine *K*, saturation and latent heat flux at the surface, and wind speeds at the pressure level closest to the surface in two dimensions at a time slice 8.0 h following the cessation of uniform rainfall, corresponding to the peak domain averaged wind speed magnitude. We examine the lowest elevation pressure level because, with a nominal vertical resolution of approximately 200 m, this is the level for which a wind forecast would be most relevant to wind energy applications. We compare the mean squared residual (γ) for the unconditioned and conditioned cases, calculated as:

$$\gamma = \frac{1}{n}\sum_{a=1}^{n}\left(X_a^{ijk} - CTRL\right)^2$$

where *X* is the individual measurement for a realization and *n* is the number of realizations. This measure was used to quantify the residual between simulated and *CTRL* and to capture the variance within the ensemble in a single metric.

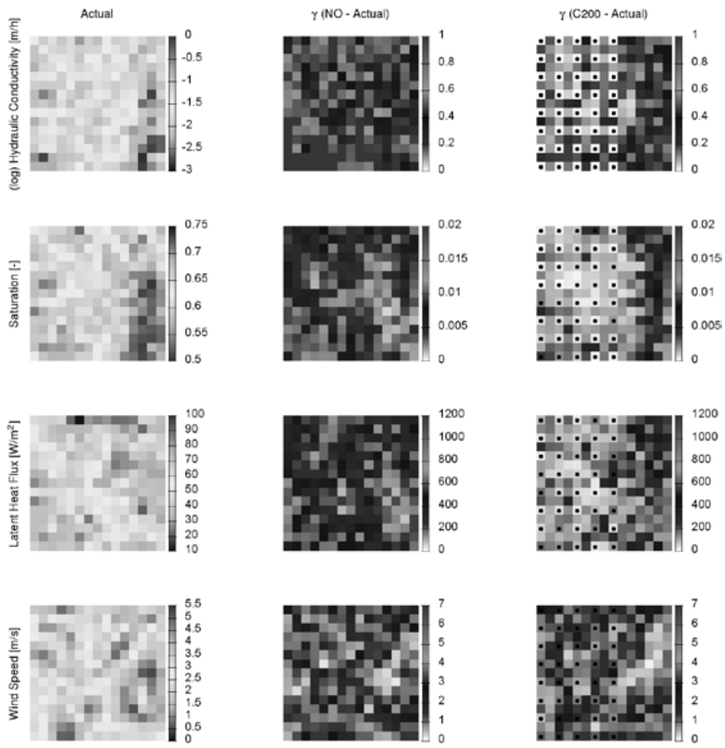

Fig. 1 Pointwise results at time t = 8.0 h after cessation of rainfall. The left column shows the CTRL fields of K, saturation, latent heat flux and wind speed (top to bottom). The remaining plots show mean squared residuals g between simulated realizations and CTRL conditions for the NO case (centre) and the *CO200* case (right).

Hydraulic conductivity, shown in the first row of Fig. 1, shows high mean squared residual (*g*) values at several points in the *NO* case. With conditioning, we see significant reduction in the *g* values throughout the conditioned area. As expected, *g* goes to zero at the conditioning points where the observed value of *K* is enforced. The spatial effect of the enforced *K* values can be clearly seen in the reductions of *g* values in the vicinity of the conditioning points. Similar behaviour is seen for saturation (second row of Fig. 1), owing to the strong correlation between saturation and hydraulic conductivity. Latent heat flux (third row of Fig. 1), heat transfer from the surface via evapotranspiration, is a process that is limited by water availability and strongly correlated with saturation. As can be expected with this strong correlation, the behaviour of latent heat flux closely resembles that of saturation and *K*. We also see small changes in *g* values for latent heat flux in the eastern part of the domain outside the conditioned area, indicating that the effects of conditioning the subsurface may influence land–atmosphere feedbacks and weather patterns, not only in the area where conditioning takes place but elsewhere as well.

Wind speed magnitude, shown in the bottom row of Fig. 1, exhibits the highest *g* values where the wind speed is highest for the *NO* case. The strongest winds in the *CTRL* case and in each ensemble average are predominantly from the west, defining clear up- and downwind directions.

While the differences in *g* values between the unconditioned and conditioned cases are not as dramatic on a domain-wide basis for wind as they are for land based variables, the influence of subsurface conditioning is clear, both in and out of the conditioned area, particularly in the high wind speed areas. Also noteworthy is the reduction in *g* values downwind of the strongest winds on the eastern part of the domain outside the conditioned area. The effects of subsurface conditioning are distributed across the domain and are not localized in the conditioning area, in contrast to the land-based variables.

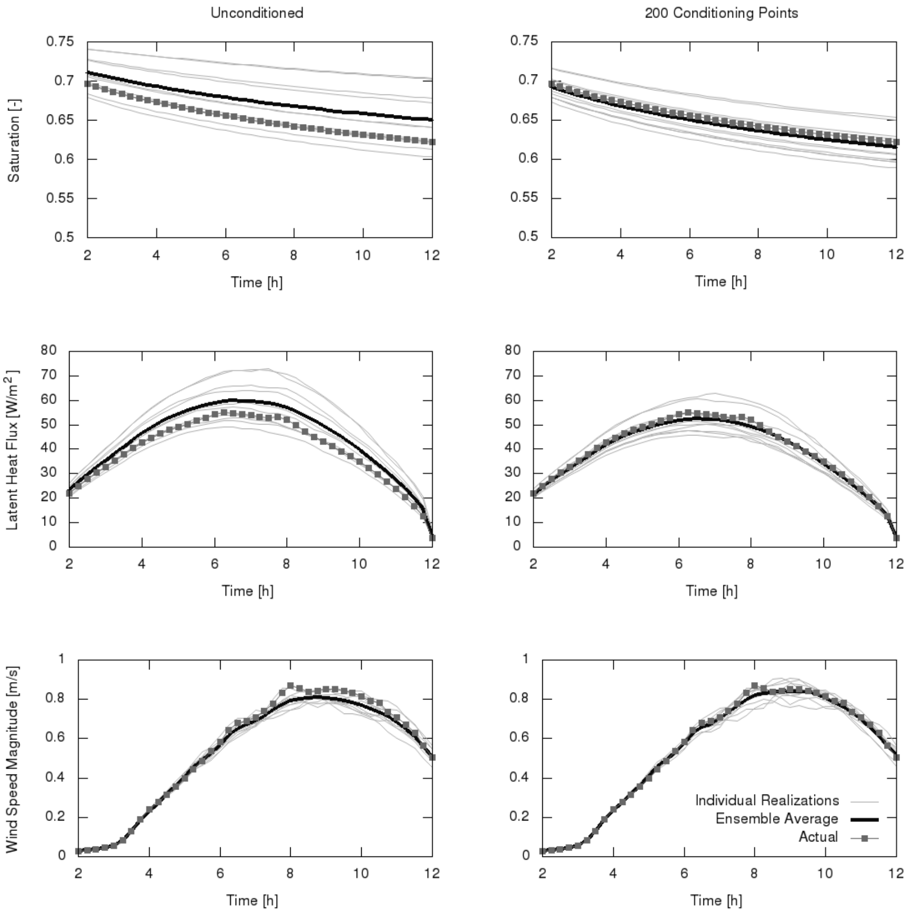

Fig. 2 Domain-averaged time series.

We then analyse domain-averaged time series for saturation and latent heat averaged over the land surface and wind speed magnitude averaged over the entire atmospheric domain (Fig. 2). For saturation, latent heat flux and wind speed, there is a clear improvement in forecast accuracy between the ensemble averages and the *CTRL* conditions from the unconditioned to the conditioned cases. Also notable is the reduced spread of the ensemble members (shown in light grey). For the wind case, the ensemble spread does not appear to reduce, however more ensemble members appear to concentrate around the *CTRL* conditions. While the maximum variance between realizations and the ensemble average does not decrease appreciably with conditioning (in contrast with saturation and latent heat flux), the maximum mean squared residual does, indicating a greater likelihood that the ensemble members fall near the *CTRL* values. This is evident at the peak *CTRL* wind shown in Fig. 2 for the *CO200* case where more than half of the realizations approach the peak at time 8.0 h—a significant improvement. Only one realization approaches the peak in the *NO* case.

CONCLUSION

Using a fully-coupled subsurface-to-atmosphere model, we demonstrate that an atmospheric simulation ensemble can be generated with different realizations of subsurface K: a new finding. We further demonstrate that by conditioning K with an increasing number of observations, it is possible to reduce uncertainties in not only subsurface variables like saturation, but also in atmospheric variables such as wind; also a new finding. It has previously been established that ABL conditions are tightly coupled to soil moisture and latent heat flux from the land surface. It has also been previously established that soil moisture is a function of, among other variables, K. Through conditioning of K fields, we bridge these previous findings and demonstrate reduced uncertainty in the predicted soil moisture field, in latent heat flux and in wind speed. The effects of conditioning the K field are evident in both spatially distributed cases and domain-averaged cases.

The reduction of uncertainty in K and the associated reduction of uncertainty in atmospheric variables is applicable to wind energy forecasts and to weather and atmospheric forecasts in general. Using a relatively small number of measurable observations, it is possible to reduce uncertainty in wind speed forecasts, which should prove useful in a wide range of atmospheric forecasting applications and climate change predictions.

REFERENCES

Ashby, S. F. & Falgout, R. D. (1996) A parallel multigrid preconditioned conjugate gradient algorithm for groundwater flow simulations. *Nucl. Sci. Eng*. 124, 145–159.

Beljaars, A. C. M., Viterbo, P. & Miller, M. J. (1996) The anomalous rainfall over the United States during July 1993: Sensitivity to land surface parameterization and soil moisture anomalies. *Mon. Weath. Rev*. 124, 362–383.

Betts, A. K., Ball, J. H., Beljaars, A. C. M., Miller, M. J. & Viterbo, P. A. (1996) The land surface–atmosphere interaction: A review based on observational and global modeling perspectives. *J. Geophys. Res*. 101(D3), 7209–7225.

Chen, F. & Dudhia, J. (2001) Coupling an advanced land surface–hydrology model with the Penn State-NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weath. Rev*. 129, 569–585.

Chen, F. & Avissar, R. (1994) Impact of land–surface moisture variability on local shallow convective cumulus and precipitation in large-scale models. *J. Appl. Meteorol*. 33, 1382–1401.

Criminisi, A., Tucciarelli, T. & Karatzas, G. P. (1997) A methodology to determine optimal transmissivity measurement locations in groundwater quality management models with scarce field information. *Water Resour. Res*. 33(6), 1265–1274.

Famiglietti, J. S., Ryu, D., Berg, A. A., Rodell, M. & Jackson, T. J. (2008) Field observations of soil moisture variability across scales. *Water Resour. Res*. 44, W01423.

Gelhar, L.W. (1986) Stochastic subsurface hydrology from theory to applications. *Water Resour. Res*. 22(9), 135S–145S.

Golaz, J.-C., Jiang, H. & Cotton, W. R. (2001) A large-eddy simulation study of cumulus clouds over land and sensitivity to soil moisture. *Atmospheric Research*, 59–60, 373–392.

Goovaerts, P. (1997) *Geostatistics for Natural Resources*. Oxford University Press, 483 pp.

Graham, W. & McLaughlin, D. (1989) Stochastic analysis of nonstationary subsurface solute transport, 1, Unconditional moments. *Water Resour. Res.* 25(2), 215–232.

Harvey, C. F. & Gorelick, S. M. (1995) Mapping hydraulic conductivity: Sequential conditioning with measurements of solute arrival time, hydraulic head, and local conductivity. *Water Resour. Res.* 31(7), 1615–1626.

Holt, T. R., Niyogi, D., Chen, F., Manning, K., LeMone, M. A. & Qureshi, A. (2006) Effect of land–atmosphere interactions on the IHOP 24–25 May 2002 convection case. *Mon. Weath. Rev*. 134, 113–133.

James, B. R. & Gorelick, S. M. (1994) When enough is enough: The worth of monitoring data in aquifer remediaiton design. *Water Resour. Res*. 30(12), 3499–3513.

Jones, J. E. & Woodward, C. S. (2001) Newton-Krylov multigrid solvers for large-scale, highly heterogeneous variably saturated flow problems. *Adv. Water Resour*. 24, 763–774.

Katul, G. G., Wendroth, O., Parlange, M. B., Puente, C. E., Folegatti, M. V. & Nielsen, D. R. (1993) Estimation of *in situ* hydraulic conductivity function from nonlinear filtering theory. *Water Resour. Res*. 29(4), 1063–1070.

Kollet, S. J. & Maxwell, R. M. (2006) Integrated surface–groundwater flow modeling: A free surface overland flow boundary condition in a parallel groundwater flow model. *Adv. Water Resour.* 29, 945–958.

Kumar, V., Kleissl, J., Meneveau, C. & Parlange, M. B. (2006) Large-eddy simulation of a diurnal cycle of the atmospheric boundary layer: Atmospheric stability and scaling issues. *Water Resour. Res*. 42, W06D09.

Leutbecher, M. & Palmer, T. N. (2008) Ensemble forecasting. *J. Computational Physics*, 227(7), 3515–3539.

Maxwell, R. M., Lundquist, J. K., Mirocha, J., Smith, S. G., Woodward, C. S. & Tompson, A. F. B. (2011) Development of a coupled groundwater–atmospheric model. *Mon. Weath. Rev*. 139(1), 96–116.

Maxwell, R. M., Kastenberg, W. E. & Rubin, Y. (1999) A methodology to integrate site characterization information into groundwater-driven health risk assessment. *Water Resour. Res*. 35(9), 2841–2855.

Nowak, W., de Barros, F. P. J. & Rubin, Y. (2010) Bayesian geostatistical design: Task-driven optimal site investigation when the geostatistical model is uncertain. *Water Resour. Res*. 46, W03535.

Rehfeldt, K. R., Boggs, J. M. & Gelhar, L.W. (1992) Field study of dispersion in a heterogeneous aquifer 3. Geostatistical analysis of hydraulic conductivity. *Water Resour. Res*. 28(12), 3309–3324.

Rubin, Y. (2003) *Applied Stochastic Hydrogeology*. Oxford University Press, 391 pp.

Seuffert, G., Gross, P. Simmer, C. & Wood, E. F. (2002) The influence of hydrologic modeling on the predicted local weather: Two-way coupling of a mesoscale weather prediction model and a land surface hydrologic model. *J. Hydrometeor.* 3, 505–523.

Skamarock, W. C. & Klemp, J. B. (2008) A time-split nonhydrostatic atmospheric model for weather research and forecasting applications. *Journal of Computational Physics* 227, 3465–3485.

Wendroth, O., Pohl, W., Koszinski, S., Rogasik, H., Ritsema, C. J. & Nelson, D. R. (1999) Spatio-temporal patterns and covariance structures of soil water status in two Northeast- German field sites. *J. Hydrol.* 215, 38–58.

Western, A. W., Zhou, S.-L., Grayson, R. B., McMahon, T. A., Blöschl, G. & Wilson, D. J. (2004) Spatial correlation of soil moisture in small catchments and its relationship to dominant spatial hydrological processes. *J. Hydrol.* 286, 113–134.

Williams, J. L. & Maxwell, R. M. (2011) Propagating subsurface uncertainty to the atmosphere using fully-coupled, stochastic simulations. *J. Hydrometeor.* 12(4), 690–701.

Yeh, T.-C.J., Ye, M. & Khaleel, R. (2005) Estimation of effective unsaturated hydraulic conductivity tensor using spatial moments of observed moisture plume. *Water Resour. Res.* 41, W03014.