Earth Surf. Dynam. Discuss., 3, C582–C591, 2016
www.earth-surf-dynam-discuss.net/3/C582/2016/

Interactive comment on "Designing a suite of measurements to understand the critical zone" *by* S. L. Brantley et al.

S. L. Brantley et al.

dka12@psu.edu

Received and published: 7 January 2016

We appreciate this opportunity to discuss our approach toward understanding the critical zone (CZ). The CZ was originally defined as the zone that spans from the top of vegetation canopy to the bottom of ground water (1 - See PDF supplement for references). Our paper was written to provoke conversation about the worldwide efforts to understand the CZ. In particular, we focus on the question, how can we best investigate the CZ as an integral object of study? We target the three major questions and a few subsidiary questions posed by the editor and reviewers below.

1. Why did we expand the Critical Zone Observatory (CZO)?

In our CZO, we recently expanded from the original focus catchment of Shale Hills (0.08

km2) to the larger Shavers Creek catchment (164 km2). Shavers creek is a tributary in Pennsylvania (U.S.A.) nested within one of the largest watersheds of the eastern U.S.A., the Susquehanna River basin. Why did we expand? The answer lies in our need to address the two fundamental objectives of environmental science: i) to answer questions out of curiosity, ii) to answer questions of applied importance. As described below, both types of questions are addressed in upscaling the CZO.

What distinguishes CZ science from other disciplinary foci is that it invites practitioners to cross timescales from that of the geologist to that of the meteorologist. Each of the disciplines has tended to focus on questions boxed in by ranges in timescale: for example, the meteorologist or ecologist seldom ponders questions related to millennial timescales. CZ science invites the practitioner to build understanding that crosses from the past to the future. This is inherently curiosity-driven: how do the details of specific CZ processes integrate over geologic timescales? What aspects of the deep past have structured ecosystems that are still affecting life today? What aspects of the geological past have been coded in the genomes of today's organisms to allow adaptation to environmental change? What rates or extents of change can break through that resilience?

As we originally designed our CZO in 2006, we acknowledged the limits of what we could do by focusing efforts on a small catchment on one lithology experiencing one type of land use. Given that no humans live in Shale Hills, most measurements were driven by curiosity. We argued our observations would eventually be extrapolated to larger human-nature systems. Simply put, however, the world is not entirely shale. The CZO was therefore expanded both as a test of what we have learned at the smaller scale and as an exercise to learn about new phenomena and conditions within a somewhat larger region that includes humans.

The expansion of the CZO within Central Pennsylvania now provides us the opportunity to study stark contrasts in rock type (shale, carbonate, sandstone) and land use (forested vs. farmed) in what is still a compact, accessible area. Given that the use of the Penn State Integrated Hydrologic Model (PIHM) at the small Shale Hills catchment highlighted the linkage between land surface and subsurface across different topographic features, soil textures, land covers, and weather conditions (2), we explicitly incorporated land that is farmed into the expanded Observatory. Specifically, we will soon be focusing part of our efforts on a small, farmed catchment on calcareous shale. In this way, our efforts are growing to address more applied questions related to the CZ. Part of our focus is to understand the key processes that produce the ecosystem-supporting soil substrate within this region – processes that depend on rock properties, tectonic shaping of the landscape, water availability, temperature, and biota – including human activity.

In fact, however, although a molecular site on a mineral surface that is weathering within a soil may look mostly the same to the chemist, biologist, geochemist, or physicist regardless of where it is studied, a region as large as Shavers creek catchment looks different to each discipline. Therefore, whereas the reductionist approach to science can narrow the focus to a molecular level where disciplines can agree, projecting the future of the earth surface (earthcasting) requires both larger spatial scales and the conceptual and mathematical models that bridge scientific disciplines. Obviously, one cannot study the structure and function of landscapes by studying one ecosystem or one lithology. To expand the CZO scale while paring down the measurements to match the limited resources, therefore, we developed the approach described in the paper that uses systems-level models with targetted field monitoring and sample collection across the expanded CZO.

2. How can we use the knowledge gleaned from studying catchments at different spatial scales?

The second question, which follows directly from the first, is the question of what exactly can be learned from studying catchments at different scales. In large studies of the CZ, two experimental designs have been used. First, some investigators study entities at different spatial scales from different regions that share some environmental conditions

(e.g. 3). The second approach, a nested catchment study, is the approach used at the Susquehanna Shale Hills CZO. In this approach (e.g. 4), investigators study one or more small watersheds embedded in a larger catchment.

We first address the question of what can be learned at different spatial scales and then address the utility of nested watersheds in the context of understanding WEGSS fluxes – fluxes of water, energy, gas, solutes and sediments – in sedimentary catchments across scales of space and time. With the nested watershed approach, we can measure individual processes with great specificity. For example, we can measure attributes describing the history of a given parcel of rock at a point on a hillslope; we can infer its weathering history based on its chemistry and mineralogy; and we can measure its exposure history and residence time in the weathering zone over long timescales using cosmogenic isotopes.

In such an approach, however, we cannot integrate over space. This therefore limits the scope of insights gained. In contrast, taking a grab sample of water or sediment from a large river allows integration of insights from across the catchment – at the same time limiting interpretations to shorter temporal scales. On the other hand, the integrative aspect of the sample means that inverse modelling is necessary to understand processes controlling sample character. It is well known that such inverse modelling – guessing the processes or components that control a large system based on inputs and outputs – is always non-unique. By working with both forward models and inverse models – i.e. working from process-based models and integrative models – greater knowledge of the system is possible.

Nested catchments provide the data to pursue both forward and inverse modelling. Within our nested CZO design, we chose the largest catchments that are small enough to drain a single rock unit or land use type: the Shale Hills (forested shale) and Garner Run (forested sandstone) catchments. At the other spatial scale of the entire CZO, we chose the smallest catchment that integrates the full complexity of regional behavior: the 164 km2 Shavers Creek watershed. Despite its relatively small size, Shavers

Creek contains much of the variability in CZ parameter space contained within the Susquehanna River Basin. Geologically, it also contains almost all the variability of the Appalachian Valley and Ridge province in general.

With this field design, we can explore both small-scale processes that may be important at the large scale as well as the large-scale processes of uplift or ecosystem migration that may in turn affect or be encoded at the molecular or gene scale, respectively. As one simple cross-cutting example of the first endeavor, recent experiments have shown that water interactions with interlayer cations in clays may affect the oxygen isotopes of water in rocks (5): in turn, this molecular-scale effect may explain some of the so-called "two-water world" of plants and streams identified in some large catchments (6). The opposite effect – where a large spatial-scale process affects small-scale phenomena – also occurs. For example, the effects of mixing of deep ground water flow beneath the mouth of the Shale Hills catchment with subsurface hill runoff results in oxidative dissolution of pyrite at 8 m depth that has been imaged under scanning electron microscopy at the micron scale (7).

Of course, it is notoriously difficult to downscale or upscale between point observations and entire regions. Some of this difficulty can be attributed to the non-uniform character of the environment. Specifically, many models posit that the system of interest is a continuum while environmental systems are instead marked by heterogeneities. Heterogeneities in the rocks include point defects, lattice defects, grain boundaries, cleavage, lithologic boundaries, and faults. Heterogeneities in ecosystems likewise vary across spatial scales from the RNA and DNA within a cell to a single tissue to the interface between clusters of organisms to the interface between ecosystems. Scaling across heterogeneities is one of the biggest challenges facing the process-based CZ scientist. As we increase the spatial scale of our measurements and models, we move from models which treat individual heterogeneities (the cell wall, the organism boundary, the ecosystem boundary) to models of the whole where the heterogeneities are treated as a continuum. Intriguingly, the study of a larger system may in some

cases be easier than a smaller system: for example, the hydrologic study of a large watershed may allow treatment of the system as a continuum whereas treatment of a small catchment may force the hydrologist to measure the exact position of fractures and faults.

However, the presence of heterogeneities in the system of interest is not the only issue in building deep understanding of the environment. The other major problem is the problem of feedbacks that operate across spatial scales or thresholds that trigger system behavior changes across timescales. Investigations that target long timescales will best be able to tease out the effects of such feedbacks and thresholds in complex systems. Thus, developing a predictive and generalized understanding of CZ processes requires quantitative models covering a vast range in both spatial and temporal scales. Such models are integral not simply for predicting landscape and ecosystem response, but also to building a heuristic understanding of CZ processes that may not be apparent from 1st-order observations. Systems-level models are especially needed for understanding feedbacks between climate, biota, and Earth surface and near-surface processes.

3. How do we optimize coupling between data collection and model development in a CZO?

This is one of the big questions motivating our paper. We start with the belief that models are needed at all temporal and spatial scales. A measurement in most cases can be recorded as a number: the understanding that derives from that number requires a model. To the extent that models can be used to infer predictions about landscape behavior, field observations and measurements provide the data needed for calibration and testing. Thus, because our understanding of the complicated suite of CZ processes is still in its infancy, baseline measurements and curiosity-driven sample collection are still vital to determine the important processes. Thus, a balance of field measurements and model development must be maintained to enable the two-way exchange of insights needed to maximize the efficiency of CZ science.

We are agnostic on the type of model that can be useful: any type of model may lead to deeper understanding. Models can be conceptual or numerical. They can be statistical or deterministic. Given the focus of our paper on designing a CZO, we emphasize only the broader scales of models. Specifically, we emphasize the PIHM family of models as a tool to explore and understand the CZ across time scales. Our current conceptual understanding and our current computers do not allow us to produce one model that simulates the CZ at all timescales. Instead we propose a cascade of models that are built on one core model, PIHM. By building this suite of models we provide a way for different disciplines to converse through the use of a shared model. In turn these models can be used to propose and test hypotheses. Throughout the paper, however, we cite publications that describe the many smaller scales or disciplinary-specific models that have been invoked to learn about individual CZ systems.

To the extent possible, we parameterize the PIHM models with datasets and then evaluate the models with different datasets. The phrase "data assimilation" gets at the idea, however, that with more and more complex models, the data and the model output become harder to distinguish. For example, it is well known that a regression line may provide a researcher with a more accurate prediction of an experimental quantity than any individual observation itself. In the same manner, the output calculated for a given observable from a complex model may be more accurate than any individual measurement of that observable. As model output is used to parameterize other models, such data assimilation obscures the difference between model and data. Considered in a different way, data assimilation provides a means to combine the strengths of both in situ observations and numerical models. Neither observations nor models are perfect. Data assimilation can provide optimal estimates of observable variables and parameters, taking into account both the uncertainties of model predictions and observations. Assimilation thus provides four-dimensional dynamically consistent predictions for the study of CZ processes.

In our work with PIHM models, we usually use one time period to calibrate and another

to evaluate the model. For example, we calibrated the PIHM model using June 2009 discharge data and then tested the model using the whole 2009 discharge dataset as well as discharge from other years; thus, the calibration and evaluation data are of the same type but are independent. As new types of observations are provided, we may first evaluate model output against the new observations prior to calibrations to see if the current calibration predicts the new data. If the prediction is poor, we gain quantitative insight regarding the robustness of our model under new conditions. In some cases we discover that even with a new calibration we cannot successfully predict the new observations and that we instead must incorporate a new module that describes a new phenomenon in PIHM. By tracking which parameters must be tuned and which processes must be added, we gain insights into both the model and system dynamics, and we learn which parameters must be observed if we want to apply our model to a new site or a new time period.

An example where we discovered we needed a new module for Flux-PIHM is our effort to understand differences in aqueous fluxes from the sunny and shaded sides of the Shale Hills catchment (8). A module was developed to test the hypothesis that aspect controls these differences in water and energy fluxes. In turn, the new version of Flux-PIHM was used to drive the WITCH weathering model (9) to understand how aspect affects porewater chemistry. Although this work has not yet been published, we found that the effect of aspect and slope on solar radiation – not included in the earlier version of Flux-PIHM – was helpful in modelling weathering on the opposing hillslopes. The interplay between model and data as well as between models (Flux-PIHM and WITCH) yielded insights about the effect of solar radiation and evapotranspiration on water chemistry. Of course, such improvements are not confined just to PIHM models. During the Flux-PIHM-WITCH modelling, it was also discovered that the porewater chemistries of elements such as Ca and K were only well described by WITCH by building in a module describing uptake by vegetation.

We can also summarize the opposite type of model-observation linkage: a measurement whose importance became apparent to us through use of the PIHM model, and that was not part of the original observation array at Shale Hills. Specifically, in a synthetic data assimilation study, Shi et al. (10) showed that land surface temperature is an important observation that is needed to provide constraints to Flux-PIHM for accurate land surface process predictions. A four-component radiometer was installed on the eddy covariance tower at Shale Hills to provide land surface temperature measurements for the catchment.

4. What is the point of data collection?

At a CZO, the point of data collection is to understand the CZ both at the scale of interest of the individual investigator and at the full spatial and temporal scale needed to earthcast the CZ. For example, the biogeochemist may collect data to understand the controls on nitrogen cycling in the catchment. Very specific questions may be posed and answered through the data collection. As part of the CZO, however, some of the data collection must contribute to the larger goals of building understanding across scales of space and time. For the individual N project example, models might be used that are inherently disciplinary in nature and specific to N. At some point, however, the new observations will be incorporated into the suite of PIHM models to interconnect the researcher with the broader CZO endeavor and the overall goal of understanding the evolution of the CZ over time.

5. How will we know if we have been successful?

Ultimately, success means that we gain deeper understanding of the system. This in turn would mean that our models are successful predictors at other times or other places. Such tests are being made. For example, we incorporated what we knew from Shale Hills to develop a model for Garner Run using only published datasets. We are currently comparing that model prediction to the new Garner Run observations.

Another specific indicator of success could be the use of our datasets by other researchers who want to test their own models. In other words, as models are advanced

by others (models for the entire CZ or for subsystems of the CZ), success will mean they are tested against our datasets. A successful data strategy should attract other modellers. If other models provide better simulations of the catchment, model-model comparisons would in turn drive development of better models.

Another indicator of success could be the adoption by others of the approaches developed to study the CZ. These approaches could include the design of a CZO, design of a network of CZOs, development of a suite of models, or an approach toward data assimilation. This latter topic is a good example where learning within one discipline – specifically meteorology where the available datasets are huge – is driving more data-poor sciences such as ecology or geology to explore how learning can proceed. In fact, the sensitivity analysis and data assimilation methods described in our paper can be applied to any numerical watershed model to identify the key variables and to guide data collection. Overall, although the scientific community may not yet agree on which model to use, an indicator of success might be when the community adopts similar data assimilation approaches.

Two final indicators of success, both of which can already be documented, are growth in use of the PIHM suite of models in other places and growth in use of the CZO concept worldwide. For example, PIHM is now being used to simulate processes in the Southern Sierra CZO among other localities (11-16). Growth of CZOs worldwide is an implicit documentation of success of the idea of CZ science itself (17).

Please also note the supplement to this comment:
http://www.earth-surf-dynam-discuss.net/3/C582/2016/esurfd-3-C582-2016-supplement.pdf

---

Interactive comment on Earth Surf. Dynam. Discuss., 3, 1005, 2015.