Dear Niels:

Thank you for considering publication of our manuscript, "Designing a suite of suite of measurements to understand the critical zone", in Earth Surface Dynamics. We appreciated the careful reviews we received and your comments. We have tried to address all the comments in the attached revision. I also enclose a Response to Editor text that goes through all the comments one by one and summarizes what we did.

We have enjoyed working on this manuscript in the context of this unusual journal format. Thank you for giving us this opportunity. We hope you will now consider the paper ready for publication.

Regards,
Susan Brantley

## Response to review:

We attempt to reply to reviewer comments in a color coded fashion below.
*Italicized text:* These are comments from editor/reviewers that do not bring up specific questions
*Italicized text*: These are specific comments from editor/reviewers with important parts ***bolded***.
For the three big questions indicated by the editor: We already posted our answers to these big picture questions publicly online. We also incorporate text throughout our manuscript that addresses these points. We kept track changes on and include one version of our manuscript that tracks all changes (there are many).
Text highlighted below in green: These are targeted topics that we respond to specifically here.
Normal text below: assorted quick comments and easy fixes from the reviewers.
Yellow text below in italics: Our answers.

*Dear Sue and colleagues,*

*First, thank you for submitting this unconventional manuscript. It is good to see that ESurf is used to broaden scientific dialogues beyond the traditional confines. The careful thought and strategy invested in the design of new experiments and observatories often remain unreported in ways that help others to shape their rationale. ESurf warmly welcomes your initiative to change this.*

*We have obtained two thorough and constructive reviews of your manuscript. Both reviewers agree that the manuscript makes a significant and novel contribution, in the unconventional way stressed above, and they advise that it should be considered for publication in ESurf. I am also of this opinion. However, I ask you to carefully address the major comments made by the reviewers.*

*They are:*
***1) Explain why it was decided to expand the SSHCZO.*** *This is now explained in Section 1, and was explained in our posted comment.*
***2) Explain the benefits of using experimental catchments across scales.*** *This is now explained in Section 1, and was explained in our posted comment. See also below.*
***3) Address the links and feedbacks between CZO design and the state and development of models.*** *This is now explained in Sections 1 and 4, and was explained in our posted comment. We have emphasized this answer throughout the paper.*

*Reviewer 2 also urges you to analyse the new geochemical data in the manuscript. I would advise to include only those data needed in the explanation of your CZO expansion strategy, and to focus the writing entirely on this explanation and an exploration of the consequences of your choices. If this requires new data, then a brief analysis of their meaning and significance should be added, but this should not distract from the main aim of your contribution.*

*Otherwise, I ask you to carefully consider the detailed comments of both reviewers in preparing a revised manuscript, which I hope you will do at the earliest opportunity.*

*Best wishes,*
*Niels Hovius (Stand-in AE)*

## Reviewer 1

*The manuscript entitled, "Designing a suite or measurements to understand the critical zone" by Brantley et al. describes the on-going expansion of the Susquehanna Shale Hills Critical Zone observatory (SSHCZO) from the original 0.08 $km^2$ catchment to a much larger 164 $km^2$ catchment. To continue to make progress towards understanding the fluxes of water, energy, gas, solute, and sediment fluxes, the expansion of the SSHCZO demands that some measurements are prioritized above others. In their manuscript, the authors outline the rationale for their new sampling strategy and explain how discrete measurements of many properties will be used together to extrapolate across the entire SSHCZO study area.*

*This manuscript is not a typical research paper. Instead of discussing results and drawing new conclusions, this paper explains, in detail, an experimental design. Consequently, it is not clear to me if the manuscript is appropriate for publication in Earth Surface Dynamics given that it does not fit the typical model of a research or review article.*

*That said, I do think that manuscript by Brantley et al. makes a valuable scientific contribution for a couple of reasons. First, the Earth science community has implicitly invested in the success of the Critical Zone Observatory (CZO) program. The CZO program has and promises to continue to provide new datasets and models that are useful to a range of sub-disciplines within the Earth sciences. By providing a glimpse into the design of the new SSHCZO, the authors are, in a way, helping the Earth science community to better understand the utility of the CZO program. Similarly, future researchers will likely also struggle with need to prioritize sampling strategies. While the effectiveness of the sampling strategy proposed by Brantley et al. is not yet known, it is likely to be useful as a comparison for future iterations.*

1. *Given that this manuscript is not a typical paper, I have only a few comments about the underlying science. In particular, I got the impression that the main purpose of the new instrumentation that the authors are installing at the SSHCZO is to provide calibration data for a suite of models that the authors are developing. Importantly, this suite contains only a single model per processes (i.e. one hydrologic model, one reactive transport model, one land-surface model, ETC).* ***If, as the authors state, the scientific community does not yet agree on which model or which parameters we need to measure in order to understand the critical zone, how will the proposed sampling strategy make progress towards this knowledge gap if it is focused on specific models with a particular set of calibration requirements?***

   A final paragraph was added at the end of the paper in Conclusions that discusses this explicitly. One attribute of success in our approach would be if other models are tested against our dataset. In addition, we already use multiple models (for example WITCH and RT-Flux-PIHM as shown in Table 1, and we have now emphasized such model comparison in the paper as well. However, overall, we run our models and test them against new data over and over and find out whether the model works. We keep adding modules (aspect, bio-uptake, etc.) when the model does not work. We try to explain this throughout. We also mention that we over-constrain some measurements (redundancy in data collection) to enable testing. We also collect some data (e.g. sapflux) which is not yet used in any of our models: we anticipate future use.

2. *Similarly, the authors mention that our understanding of the key processes that govern the critical zone is far from complete and that there are likely many important and unknown thresholds and feedbacks (page 1007-line 22).* ***How can/will the new sampling network at the SSHCZO be used to elucidate these thresholds and feedbacks? How can/will the calibration of a suite of models that do not contain these unknowns help us better understand critical zone processes? Are any of the proposed measurements useful for distinguishing between competing models with different assumptions/parameterizations?***

   We have indicated now in several places that numerical models at all timescales can elucidate thresholds and feedbacks. In fact, such features will emerge in the models if they are correctly set up. If we suspect a feedback based on our observations, we can test it with the numerical model. The reviewer is correct that we may not measure or model certain parameters or processes and so we may miss thresholds and feedbacks. However, we added the explanation that small targeted investigations are also planned with respect to individual controlling variables to enable new conceptual models to emerge as needed (Section 1, 4, 5) and these can drive new numerical modules to be incorporated in the suite of models.

*I understand that the questions I have included above are difficult to answer and I do not necessarily expect the final version of this paper to solve all of them. It is also likely that many of these issues have already been carefully considered by the authors. I just think that this manuscript would be improved by elaborating on how the expansion of the SSHCZO can/will lead to new insights into critical zone processes independent of any specific numerical model.*

*Detailed Comments*

3. *1007 - Line 16: I'd be interested to know if there are any specific disagreements that the authors have in mind. This is relevant since it seems as if the proposed sampling strategy is focused on calibrating a*

*specific set of models. If the community does not yet agree on which measurement to make, will the proposed sampling strategy for the expanded SSHZO only be useful for one modeling group? Or, are enough additional measurements being made in order for competing models to be tested?*

Each coauthor probably has a disagreement in mind based on their own discipline. We added the explanation that small targeted investigations are also planned with respect to individual controlling variables to enable new conceptual models to emerge as needed (Section 1, 4, 5). We now emphasize that disciplinary investigations are ongoing that target unknown processes or new specific hypotheses. We mention in section 5 that we hope other modelers will target our datasets.

4. *1008 - Line 16: I am not sure what "This" is referring to at the beginning of the paragraph.*

   We rephrased this text.

5. *1010 - Line 10:* ***This sentence points out that there needs to be a feedback between model collection and development, but this is not major focus of the manuscript.*** *To me, expanding the discussion about plans for model-data feedback at the SSHCZO would greatly improve the manuscript. The authors have a lot of complex models to calibrate, but will they use any independent measurements (i.e. measurements not involved in model calibration) to assess the performance of their models? The only mention of this that I saw was on page 1020 - line 28, which stated that independent measures of evaporation/transpiration could be compared with FLUX-PIHM. Is this the only plan for model validation?*

   We added discussion in Section 1 and 2 and Section 4 about model-data feedbacks. The reviewers were correct that this idea of model data interaction was not well described in our original manuscript.

6. *Potentially, the authors could also use a water isotope mass balance as an additional check on estimates of evaporation and transpiration from field measurements and FLUX-PIHM simulations. Can FLUX-PIHM be used to track fluid transit times? If so, is there a plan to compare these estimates with tracer based estimates? If they disagree, will the design of the sensor network, which was based on the FLUX-PIHM simulations, still be useful for other research questions? Can RT-FLUX-PIHM be used to predict concentration-discharge relationships? I suppose that the reaction kinetics will always be an adjustable parameter that can be used to fit the data. However, the optimal reaction kinetics would make predictions about long-term regolith development that could be tested with elemental profiles in the soil pits. I guess that this sort of model testing is what the authors are planning, but it would be nice to get more specific details of how different models will be tested with independent data.*

   Many new datasets could be collected and tested against the PIHM models and we hope this will happen. This is mentioned in Section 5. One of the problems with CZ science is money and time are tight: isotopic measurements are being made but probably not to the extent desired by the reviewer. The PIHM family of models can be used in the ways that the reviewer suggests.

7. *1011 - Line 17: At least to me, it is never made exactly clear why the distinction between planar hillslopes and swales is so important. Furthermore, the authors state later in the manuscript (1024-line 20) that swales are absent in portions of the catchments underlain by sandstone. If this is the case, is the geomorphological knowledge we learnt from the original SSHCZO actually translatable to the expanded SSHZO? Or, when we incorporate different lithologies, do we need to resume the random sampling strategy?*

   We added explanations about swales/planar slopes throughout the manuscript beginning in Section 1. At Shale Hills the basic hydrologic unit is either a planar hillslope or a swale and this is fundamental to understanding the catchment. It is of interest that swales have not developed in the sandstone catchment: we hope to develop greater understanding as to why this is true. If we had a good reason for why we do not see swales in sandstone, then we would have predicted this as we extrapolated from Shale Hills to Garner Run: we are working on this.

8. *1013 - line 5: I guess this is where I got the impression that the focus of the sampling strategy is to parameterize a specific model. While I understand the desire to find a robust way of extrapolating discrete measurements, are the authors doing anything to independently verify model results? Also, can these parameters be used to calibrate other hydrologic models?*

   We added discussion throughout the manuscript, especially in Section 2 and 4 to address this. The idea of data assimilation is also explained more fully now. In fact, we are constantly testing model output against data, but then we assimilate the new data and sometimes build new modules (aspect, vegetation) to constantly improve the model. This idea of data assimilation is an example of a new approach that has crossed from one field (meteorology) to another (geoscience) through CZ efforts.

9. *1025 - line 16: The authors write four rock samples here. Later on (1026 - line 14), the authors write 5 rock samples. Which is it?*

   It should have been 5. We had several mistakes about these rock samples. We fixed this.

10. *1032 - line 19: The authors mention characterizing groundwater residence times, but never mention how this will be determined. I know that there are a variety of approaches and each has its own caveats. So, it might be interesting to know which will be tested at the SSHCZO.*

    We have used $SF_6$ and CFCs at Shale Hills. This is the plan for Garner Run if we have enough money as well and we now mention it in Section 3.

11. *1036 - line 12: The authors state that, "we are testing the hypothesis that fewer soil pits are needed because we are using a regolith formation model and geological knowledge to site the few pits that we dig". I agree that this is what they are doing. But, it is not clear to me how they will evaluate the results of this test. I'd appreciate more detail about how the authors will determine whether or not digging a few pits is all that is really necessary.*

    This is a good question. In fact, the reviewer could ask this about any set of our measurements. In fact, this question can be asked about any study of a field system: is the subset of measurements (pit, samples, etc.) enough to derive conclusions for the system? The only answer we have is that we use our measurements to parameterize models that we then use to make predictions in new settings (i.e. the Shale Hills model is now being tested without tweaking at Garner Run) and when it works (i.e. we can answer a questions adequately given the data available), we conclude that the data for parameterization was adequate to the question asked – we infer that both the dataset and the model are adequate.

    We added this sentence right after the sentence quoted by the reviewer: If we find that our limited digging of soil pits is not successful in characterizing the regolith adequately – if our models of regolith formation do not match observations or our models of water flow through regolith do not simulate observations – more pits can be dug or new approaches toward geophysical measurements can be refined.

12. *Figure 1: I got a bit confused by this figure. So, I am going to write down my impressions so that the authors can see whether or not I understood it correctly. Why are some blue arrows pointing up and others pointing down? I get the impression that this is because the arrows pointing up refer to processes above the land-surface and vice versa, but this is not explicitly stated in the captions. It this is the case, shouldn't "other gas fluxes" be changed to "gas effluxes"? The left-to-right order of the blue arrows and brown boxes refers to relative time for each to reach steady-state, right? Why does surface runoff have sediments while subsurface runoff has particulates? Also, I am supposed to assume that the gradients on the top of the diagram (distribution of above ground biota and land surface elevation) arise from the above ground fluxes and not the below ground ones? I'd assume that the distribution of above ground biota is also sensitive to the gradients in regolith composition. These two properties seem disconnected in the figure, but I do not know if this is intentional or not.*

We have now indicated in the caption that the up and down arrows are related to above and below ground processes. We changed gas fluxes to gas effluxes. The left-to-right order of the brown boxes refers to relative time for each to reach steady-state. We now use “water, solutes, particulates” in both flux arrows. In the caption we now emphasize that there are many many feedbacks that could be drawn, for example from below ground to above ground, and we have eliminated those to make the figure clearer. This is now emphasized repeatedly in the caption. Putting in the feedbacks is prohibitively complex in such a schematic. The schematic has many disconnected properties that are, in actuality, connected. This is, in fact, why numerical models are needed so desperately, i.e. to understand such a complex system.

*I assume that the inset describing the tower is located where the black box is on the map. An arrow or something could be helpful here.*

This sentence by the reviewer refers to a different figure. We have now fixed this on the figure.

13. *Figure 8: Are the uncertainties associated with the analytical measurements and the range of parent material composition smaller than the point size in this figure? If not, it might be worth plotting them as well.*

    This is an excellent point on the part of the reviewer and we are chagrinned that we had not explored this before. Once we explored it (see the new figure 8) we realized that some of our claims about additions of elements such as Mg and Al were not necessarily robust because of the large error bars. The point of our discussion, i.e., that it looks like there are significant additions to the soil other than from protolith, is still true but we can only assert it for K, given the large error bars. We now present the discussion differently.

14. *Figure 9: are there any measurements of soil depths that could be plotted for comparison?*

    We provide a new version of Figure 9 which shows the pit depths superimposed.

15. *Figure 2: The caption says black dots = trees. But, the dots are actually green.*

    We have fixed this in caption.

## Reviewer 2

*This manuscript is an interesting contribution from a diverse team who are seeking to understand processes and fluxes in the Critical Zone (CZ). It is not a ‘standard’ paper, in that it considers aspects of experimental design, as the title suggests, which the authors hope will allow CZ research priorities to be achieved. I quite liked this take, we don’t really discuss these themes much (usually while writing proposals and designing/doing the research, rather than putting these decisions through peer review) - perhaps we should do more of this. In that regard, this is a relatively novel and quite refreshing perspective. While not having a large new dataset present (although there is new data which could be used more), this novelty clearly makes it relevant and interesting to the readership at Earth Surface Dynamics. It fits the journal remit well (linking atmosphere, pedosphere/lithosphere and hydrosphere) and in particular seeks to enhance theory–observation links. In my opinion it should eventually be published here, and should form an important paper for discussing CZO design and implementation, while also very relevant to others working on similar questions at non-CZO sites.*

*However, I have three important comments which I feel need to be addressed before it is ready for publication:*

1. ***Making it clear why the expansion being undertaken**: The paper discusses the expansion from Shale Hills CZO (0.08 km2) to the Shavers Creek (160km2). It outlines how Garners Cr was selected as a site to help inform how the increased size of the sample network might work. However, it was never really made clear what the scientific objectives for Shavers Creek were. Later in the article, nitrate export from farm land was mentioned briefly. I think some addition information is needed to explain why, in this case, the expansion in being undertaken. This is important because, from reading this article one may conclude that it is simply the ‘next step’ to install a sampling network downstream of a small (<1km2) CZO. A nested catchment approach can be very powerful to examine solute fluxes (for example see recent work by Torres et al., 2015, in the Andes). However, **it may instead be preferable to install monitoring at a completely***

***different larger catchment****, which is not necessarily draining the same geographic area as a CZO. For example, the SHCZO has shown us that these things are crucial (x,y,z) in terms of sedimentary rock weathering, now let's test this in a larger, sedimentary-rock dominated catchment system (where existing work may also be available). Or, one could imagine the case opposite to that posed here, where we have good data and information at a catchment ~1000km2, and we wish to install measurements at a 1st order site. I think some reflection on that could be helpful, and help provide a broader context for the discussions and decision making steps being explained in the paper.*

We have now discussed reasons for the expansion in Section 1 and in our Conclusions and in our posted comment. We also mention in Conclusions that we have some outside-of-the-CZO investigations ongoing for precisely the reasons cited by the reviewer.  Briefly, we point out in Conclusions that we agree with the reviewer that the use of PIHM models (or any of our models) in other settings is a useful way to proceed, but we also argue that our expansion in our own CZO allows us a firm foundation for strengthening our understanding and our modelling efforts.

2. ***Making it clear how we can use experimental catchments across scales****: Linked to the previous comment, I think the paper also needs to summarise what complementary information can be gained from monitoring at small scales (0.08km2), to medium (100km2) and larger (>1000km2) catchments. To me, it seems that small scale (most of the current CZOs), intensively monitored catchments are very helpful for informing us about processes (what controls them, their rates, variability etc.,). However, larger catchments allow us to identify the dominant water/solute/sediment fluxes, and target the major players (perhaps in terms of major processes, major sources etc.). This is key for upscaling and thinking about larger scale river geochemical fluxes (and what they tell us about the changing dynamics of erosion and weathering). A lot of work has been done on geochemical fluxes at Earth surface at the larger catchment scale (>100 km2) and at* ***I think some discussion of that literature could be very useful here. This could be linked to point 1 above, to a revised introduction.*** *Perhaps deal with these general themes first, before explaining the Shale Hills – Shavers Creek case (and its associated scientific priorities and objectives).*

   We have now discussed this to some extent in Section 1 but to a greater extent in our posted comment. We actually intended to incorporate more discussion of this topic into the current version of the manuscript, including citations to big river papers, but at the end of the day, our paper became so long that we left this extended discussion out. Specifically, what we did do here was to give a detailed explanation for expansion of our CZO but we did not give a detailed explanation in the paper of the rationale for nested watersheds versus paired watersheds. We decided that such an extended discussion, which was somewhat included in our posted reply to reviewers, is more of discussion of a CZO network, than a discussion of a CZO. But mostly we left it out because our paper became long.

3. ***More analysis of the geochemical data****: The article includes some new data (soil/solute data) from the Garner run catchment and other nested sample locations. Not much is made of this data, which is a pity. In addition to discussing experimental design, I think this work would be useful to the community with an expanded discussion of these datasets (and much clearer links to the objectives of the paper, i.e. how this data helps inform the future sampling efforts) in place of some of the 'will be done' parts of the discussion. See more detailed line-by-line comments below.*

   The editor suggested we not add more data analysis so we did not amplify this in response to the reviewer and actually made the discussion of soil data more brief. We did try to make it more clear in this section why this sort of data is important to measure early on during the development of a conceptual model of the CZ.

*Having completed my review, I also looked over comments by Reviewer #1s which were published earlier in the month. I see quite a few parallels with my main comments above, and I hope together they can be used to improve the manuscript further. I've also identified other comments as I worked through the paper, and list them here as they appear in the manuscript (with page and line, Px-Ly):*

4. *P06-L9: 'much' is true, but not in the context of other work on geochemical transfers at earth surface, 164 km2 is still a small catchment. Perhaps remove 'much'*

Removed

5. *P06-L16: It would be useful to have a little bit more about what these measurements seek to tell us*

   We have now tried to do this in Sections 2 and 3 and 4, to a small extent.

6. *P06-L25: The abstract didn't really specify what the scoping measurements were and what they show in terms of CZ science. I think this would strengthen the abstract to include these findings and conclusions*

   We were not sure what is meant by "scoping" so we only changed the abstract in minor ways.

7. *P07-L1: Given the potentially broad readership, it would be useful to perhaps have a little more about the CZ, what it is, the overarching questions which the community is seeking to address.*

   Again, we intended to do this but did not add significantly to answer this question to not make the paper too long. One of the problems with CZ science is that each discipline has their own question. The answer to this question, succinctly, is in the opening sentences: *The critical zone (CZ) is changing due to human impacts over large regions of the globe at rates that are geologically significant (Crutzen, 2002; Vitousek et al., 1997a; Vitousek et al., 1997b; Wilkinson and McElroy, 2007). To maintain a sustainable environment requires that we learn to project the future of the CZ. Models are therefore needed that accurately describe CZ processes and that can be used to project, or "earthcast," the future using scenarios of human behavior.*

8. *P07: somewhere in the intro, could it be useful to spell out what the point of data collection is? Is it to build empirical models for forecasting, or to test physically numerical models? It's both I imagine depending on the question. But could be useful to explain somewhere in the intro. In fact, my comments 1 and 2 above suggest the introduction can be refocused somewhat, which may help this.*

   We have tried to address this in the Intro and throughout the paper by emphasizing the model-data feedbacks as well as the different types of models we use: 1) CZ-level models such as the PIHM family in Table 1; 2) disciplinary specific models. We explicitly state now that measurements are numbers and to understand the numbers requires models and that the point of data + modelling is to create deeper understanding.

9. *P08-L11: perhaps change to "Different disciplines may focus on"*

   Revised.

10. *P09-L26: In fact, this is not a new research design, and has been hugely successful for understanding geochemical fluxes and revealing the dominant processes in too many studies to mention. However, what is now key, which is a point not made clearly here, is a sample design which allows you to answer what you want to know now (what you might have funding for) and parallel projects (or projects you may not have even thought of). For instance, if one is interested in carbon (inorganic and organic) fluxes, it makes sense to try and collect samples for BOTH suspended load and dissolved load, and archive/store these in a way that they can be used by others (in parallel projects or at the same time). Or, say if I wanted to use Li isotopes to look at weathering processes, collecting a sample set which also allows the samples to be used to examine organic matter fluxes and origins. There has been much progress in the 'large river' community with this in mind, which we've seen in the last few years (e.g. Ganges, Amazon, Mackenzie). To me, this is a critical need which CZOs really contribute to, and it doesn't come across clearly on reading the manuscript (also see main comments above)*

We think the reviewers are addressing this sentence: *By necessity, to understand the interaction of WEGSS 25 fluxes in Shavers Creek, we must move beyond the paradigm of measuring* *"everything, everywhere" (Fig. 2) to an approach of measuring "only what is needed."* We are a little unclear as to what the reviewer wants here because we think we addressed this in our entire paper. So we did not change anything to address this point.

11. *P10-L9: could this heading be better as 'Existing measurement network" or something like that?*

    We revised headings but did not use this phrasing.

12. *P10-L20+L24: 'extremely' and 'exceedingly' don't seem to be needed here.*

    "Exceedingly" was removed.

13. *P11-L24: Do you propose this to be a first step everywhere there is a CZO which may wish to upscale? Could be worth discussing briefly. See main comments above.*

    We are not really arguing that everyone has to do what we are doing. We thought that putting out a description of what we are doing would get others to improve on it!

14. *P12: when you mention 'small number's, it would be useful to clarify what is meant in terms of spatial or temporal frequency (+ at other points in the discussion)*

    This is described explicitly in the rest of the paper so we made no change.

15. *P13-L5: 'Data Assimilation' – this wasn't a very clear title. I found it a little hard to navigate the paper overall, and I wonder if these subheadings could be more clear.*

    See point **(11)** above. We also describe data assimilation more thoroughly now: data assimilation means something explicit and we try to define it explicitly now.

16. *P15-L15-20: I think the first thing to discuss is why this sub-catchment was picked –comes a bit later, but in the context of what the paper is trying to do, this would be better to come straight up.*

    We included this rationale as requested.

17. *P16-L19: Some of the 'will' be done parts of the paper are interesting (relating to my overall feeling about the manuscript, and this occurs in a few other places too). I think more can be made on what has been measured and found so far, and how that might inform these future campaigns.*

    The editor suggested we not add more data analysis so we did not amplify this in response to the reviewer and actually made the discussion of soil data more brief. We did try to make it more clear in this section why this sort of data is important to measure early on during the development of a conceptual model of the CZ.

18. *P17-L1: Could be nice to see the slope probability distributions for the two sub catchments and then Shavers' Creek, to get an idea how they sample the geomorphic variability.*

    We did this but did not include it because it was not that elucidating at the level we can complete the analysis now.

19. *P18-L1-17: Seems like a link to a recent paper by Prasciek et al., (2015) could be a useful way to support these discussions – they argue the rate at which a glacially sculpted landscape is modified relates to the tectonic uplift (and river incision) rate.*

Actually, we argue that the Prasicek et al. paper is not particularly relevant to Shavers Creek, as it focuses on glacial processes and the switching between U-shaped and V-shaped valleys in rapidly uplifting landscapes. In Shavers Creek, erosion rates are so slow that there has been no possibility of Cenozoic resetting of relief, and the dominant features are periglacial. This is more thoroughly emphasized in the new manuscript.

20. *P20-L2: First mention of the goals of upscaling to Shavers Creek. This needs to come in the opening exchanges – see main comment above.*

    We now have addressed this in Introduction.

21. *P25-L21: because of the very high SiO2 proportion in these rocks, and low [Ti], does this lead to larger uncertainty when quantifying these normalised concentrations (i.e. if [Ti] is very low)?*

    This is an excellent point on the part of the reviewer and we are chagrinned that we had not explored this before. Once we explored it (see the new figure 8) we realized that some of our claims about additions of elements such as Mg and Al were not accurate because of the large error bars. The point of our discussion, i.e., that it looks like there are significant additions to the soil other than from protolith, is still true but we can only assert it for K, given the large error bars.

22. *P26-L3: More discussion on the preliminary results would be good. I note that the conclusion comments on the dust deposition for example, but this idea is not discussed in enough detail here. How does this compare to known dust inputs at Shale Hills?*

    The editor suggested we not add more data analysis so we did not amplify this in response to the reviewer and actually made the discussion of soil data more brief. We did try to make it more clear in this section why this sort of data is important to measure early on during the development of a conceptual model of the CZ.

23. *P27-L1-26: while I argue in my opening comments that this paper is an interesting take (discussing the rationale behind measurement set up), it seems that much of this section would be better in the paper which actually provides the results of these investigations.*

    The editor suggested we not add more data analysis so we did not amplify this in response to the reviewer and actually made the discussion of soil data more brief. We did try to make it more clear in this section why this sort of data is important to measure early on during the development of a conceptual model of the CZ.

24. *P30-L10: Does the reader really need this detail? A lot of this description can be summarised in a few lines which document the challenges and uncertainties in the methods.*

    This was condensed.

25. *P31-L20: How do these measurements compare to the geochemical measurements? seems these were shallower than these GPR depths?*

    We have changed the figure to show a comparison.

26. *P32-L1-20: This section highlights some really important issues which CZOs can help inform us off. However, I think it would be useful to comment and summarise on work which has sought to do this at a larger scale (100 to >1000km2 catchments), using runoff vs concentration trends, and runoff vs ion/ratio trends, to better understand processes and sources of ions and water (e.g. Tipper et al., 2006; Calmels et al., 2011; Maher, 2011; Torres et al., 2015)*

    We actually intended to incorporate more discussion of this topic into the current version of the manuscript, including citations to big river papers, but at the end of the day, our paper became so long that we left this extended discussion out. We decided that such an extended discussion, which was somewhat included in our posted reply to reviewers, is more of discussion of a CZO network, than a discussion of a CZO. But mostly we left it out because our paper became long.

27. *P33-L26: were these samples for dissolved chemistry filtered, would be useful to clarify*

    Caption was altered for figure.

28. *P34-L9+P35-L7: To me, this dataset analysis needs more discussion and warrants more space in the manuscript. The Shale Hills data could be added to the plots, and discussed in much more detail (see main comment above and comment on the figure). How do we then use these datasets to make decisions about new sample networks?*

    See above comment **(3)** from Reviewer 2 – not sure we need to go into more detail.

29. *P36-L16-28: these aspects need more discussion in the main text.*

    Not sure this is necessary as indicated by editor

30. *Fig. 3 – nice to see this analysis. However, not that much is made of this in the paper. Could you do a little quantitative comparison? For example, slope probability (and or elevation probability plots) for the whole Shavers creek, vs Shale Hills and Garner run?*

    We did not add a slope map because we felt it was beyond the scope.

31. *Fig. 5. Can you add more to the caption which explains what the legend describes and whether these have been installed?*

    We revised the caption.

32. *Fig. 6. Not much was made of this topographic analysis, but to me, it makes sense as a way to inform of sampling locations etc.,*

    We did not add more discussion because we felt it was beyond the scope

33. *Fig. 8. More can be discussed on this in the main text. It would be useful to compare to Shale Hills profiles.*

    We did not add more discussion because we felt it was beyond the scope

34. *Fig. 9. Could be useful to have a zoomed in inset which shows the transition of soil-bedrock boundary in more detail. I feel like the main text didn't explain quite how this line was decided on.*

    We changed Figure 9 to add the zoomed in parts.

35. *Fig. 10. Like fig. 8, much more can be discussed in the main text. Also, given the drainage area changes a lot, it would be useful to plot a water discharge normalised to drainage area (i.e an instantaneous runoff depth, say mm/hr). On panel A, my guess is that if you did that, you'd find the yellow points form part of the decrease seen in the other catchments?*

    We did not add more discussion because we felt it was beyond the scope

    .

*References mentioned here not cited in the manuscript:*

*Calmels et al., 2011, Contribution of deep groundwater to the weathering budget in a rapidly eroding mountain belt, Taiwan. Earth Planet. Sci. Lett. 303, 48–58.*

*Maher, 2011, The role of fluid residence time and topographic scales in determining chemical fluxes from landscapes. Earth Planet. Sci. Lett. 312, 48–58.*

*Prasicek et al., 2015, Tectonic control on the persistence of glacially sculpted topography, Nature Communications, DOI:10.1038/ncomms9028*

*Tipper et al., 2006, The short term climatic sensitivity of carbonate and silicate weathering fluxes: insight from seasonal variations in river chemistry. Geochim. Cosmochim. Acta 70, 2737–2754.*

*Torres et al., 2015, Geomorphic regime modulates hydrologic control of chemical weathering in the Andes–Amazon, Geochim. Cosmochim. Acta 166, 105–128.*

**Brantley_etal_CZOMeasure_dka-01-25-2016-1-slb-final.doc [Compatibility Mode]**

## Main document changes and comments

| Page 1: Deleted | sue | 1/26/2016 6:03:00 AM |
|---|---|---|

**A. L. Neal[4],**

| Page 1: Inserted | sue | 1/26/2016 6:03:00 AM |
|---|---|---|

**, A. L. Neal[4], Kristen M. Brubaker[5]**

| Page 1: Formatted | sue | 1/26/2016 6:03:00 AM |
|---|---|---|

Not Superscript/ Subscript

| Page 1: Formatted | sue | 1/26/2016 6:04:00 AM |
|---|---|---|

Superscript

| Page 1: Inserted | sue | 1/26/2016 6:04:00 AM |
|---|---|---|

[5] {Department of Environmental Studies, Hobart and William Smith Colleges, Geneva, NY, USA}

| Page 2: Inserted | sue | 1/19/2016 5:49:00 AM |
|---|---|---|

n

| Page 2: Inserted | sue | 1/19/2016 5:50:00 AM |
|---|---|---|

integral

| Page 2: Deleted | sue | 1/19/2016 5:49:00 AM |
|---|---|---|

worthy

| Page 2: Inserted | sue | 1/19/2016 5:49:00 AM |
|---|---|---|

worthy

| Page 2: Deleted | sue | 1/19/2016 5:50:00 AM |
|---|---|---|

study

| Page 2: Inserted | sue | 1/19/2016 5:50:00 AM |
|---|---|---|

focus

| Page 2: Deleted | sue | 1/19/2016 5:51:00 AM |
|---|---|---|

original

| Page 2: Inserted | sue | 1/19/2016 5:51:00 AM |
|---|---|---|

itself

| Page 2: Inserted | sue | 1/19/2016 5:52:00 AM |
|---|---|---|

within a geomorphological framework

| Page 3: Inserted | sue | 1/19/2016 5:56:00 AM |
|---|---|---|

using scenarios of human behavior

| Page 3: Deleted | Roman DiBiase | 1/16/2016 11:52:00 AM |
|---|---|---|

even

| Page 3: Deleted | sue | 1/19/2016 5:56:00 AM |
|---|---|---|

using scenarios of human behavior

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

run

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

s

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

to project

| Page 3: Inserted | sue | 1/19/2016 5:57:00 AM |
|---|---|---|

estimates

| Page 3: Deleted | sue | 1/19/2016 5:57:00 AM |
|---|---|---|

projections

| Page 3: Inserted | sue | 1/19/2016 5:57:00 AM |
|---|---|---|

today, let alone projections for tomorrow

| Page 3: Deleted | sue | 1/19/2016 5:57:00 AM |
|---|---|---|

predict

| Page 3: Inserted | sue | 1/19/2016 5:57:00 AM |
|---|---|---|

estimate

| Page 3: Deleted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

the

| Page 3: Deleted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

in a catchment

| Page 3: Inserted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

in a given catchment

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

we are often uncertain

| Page 3: Inserted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

of difficulties characterizing

| Page 3: Deleted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

predict

| Page 3: Inserted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

estimate

| Page 3: Deleted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

fully

| Page 3: Inserted | sue | 1/19/2016 5:58:00 AM |
|---|---|---|

adequately

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

means

| Page 3: Inserted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

reflect

| Page 3: Inserted | sue | 1/19/2016 5:59:00 AM |
|---|---|---|

s

| Page 3: Inserted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

the fact

| Page 3: Deleted | sue | 1/19/2016 5:59:00 AM |
|---|---|---|

can be

| Page 3: Inserted | sue | 1/19/2016 5:59:00 AM |
|---|---|---|

is

| Page 3: Deleted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

The

| Page 3: Inserted | Roman DiBiase | 1/17/2016 12:08:00 AM |
|---|---|---|

However, the

| Page 3: Deleted | Roman DiBiase | 1/16/2016 11:55:00 AM |
|---|---|---|

with respect to the

| Page 3: Inserted | Roman DiBiase | 1/16/2016 11:55:00 AM |
|---|---|---|

of

| Page 3: Deleted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

disparate

| Page 3: Inserted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

are repeated from site to site despite variations in environmental conditions.

| Page 3: Deleted | sue | 1/21/2016 5:36:00 AM |
|---|---|---|

reveal aspects of the underlying complex behavior

| Page 3: Inserted | sue | 1/21/2016 5:33:00 AM |
|---|---|---|

Such patterns

| Page 3: Deleted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

– properties

| Page 3: Inserted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

include

| Page 3: Deleted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

such as

| Page 3: Inserted | sue | 1/21/2016 5:34:00 AM |
|---|---|---|

s

| Page 3: Deleted | sue | 1/21/2016 5:34:00 AM |
|---|---|---|

of regolith

| Page 3: Inserted | sue | 1/19/2016 6:00:00 AM |
|---|---|---|

across landscapes or versus depth of such observables as regolith, fractures, bacterial species,

| Page 3: Deleted | sue | 1/21/2016 5:35:00 AM |
|---|---|---|

the distribution of soil gas

| Page 3: Inserted | sue | 1/21/2016 5:35:00 AM |
|---|---|---|

gas composition

| Page 4: Deleted | sue | 1/21/2016 5:35:00 AM |
|---|---|---|

versus depth

| Page 4: Deleted | sue | 1/19/2016 6:01:00 AM |
|---|---|---|

At the most general level, g

| Page 4: Inserted | sue | 1/19/2016 6:01:00 AM |
|---|---|---|

G

| Page 4: Deleted | sue | 1/21/2016 5:35:00 AM |
|---|---|---|

of these

| Page 4: Inserted | sue | 1/26/2016 6:14:00 AM |
|---|---|---|

observable

| Page 4: Deleted | sue | 1/26/2016 6:13:00 AM |
|---|---|---|

temperature,

| Page 4: Inserted | sue | 1/26/2016 6:14:00 AM |
|---|---|---|

and

| Page 4: Deleted | sue | 1/26/2016 6:14:00 AM |
|---|---|---|

,

| Page 4: Deleted | sue | 1/19/2016 6:01:00 AM |
|---|---|---|

drive

| Page 4: Inserted | sue | 1/26/2016 6:15:00 AM |
|---|---|---|

emerge as indicators of

| Page 4: Inserted | sue | 1/21/2016 5:36:00 AM |
|---|---|---|

and reveal aspects of the underlying complex behavior

| Page 4: Deleted | sue | 1/19/2016 6:01:00 AM |
|---|---|---|

characterized by

| Page 4: Inserted | sue | 1/19/2016 6:01:00 AM |
|---|---|---|

experiencing

| Page 4: Deleted | sue | 1/17/2016 5:38:00 AM |
|---|---|---|

forms determined by the operative feedbacks.

| Page 4: Inserted | sue | 1/17/2016 5:38:00 AM |
|---|---|---|

conditions, i.e., gradients that remain constant over some interval of time.

| Page 4: Inserted | sue | 1/19/2016 6:04:00 AM |
|---|---|---|

In Fig. 1, some of these important gradients are arrayed from left to right to indicate the increasing length of time it takes for each gradient in general to achieve such a steady state. In other words, a steady-state soil gas depth profile might develop more rapidly than a steady-state regolith chemistry depth profile.

| **Page 4: Inserted** | **sue** | **1/19/2016 6:06:00 AM** |
|---|---|---|

(different gradients)

| **Page 4: Inserted** | **sue** | **1/21/2016 5:37:00 AM** |
|---|---|---|

under transient and steady-state conditions

| **Page 4: Inserted** | **sue** | **1/19/2016 6:06:00 AM** |
|---|---|---|

and modelling should

| **Page 4: Deleted** | **sue** | **1/19/2016 6:06:00 AM** |
|---|---|---|

should

| **Page 4: Deleted** | **sue** | **1/26/2016 6:17:00 AM** |
|---|---|---|

emerge

| **Page 4: Inserted** | **sue** | **1/26/2016 6:17:00 AM** |
|---|---|---|

characterize

| **Page 4: Deleted** | **sue** | **1/26/2016 6:17:00 AM** |
|---|---|---|

in

| **Page 4: Inserted** | **sue** | **1/19/2016 6:07:00 AM** |
|---|---|---|

Given that the

| **Page 4: Deleted** | **sue** | **1/19/2016 6:07:00 AM** |
|---|---|---|

Building a model for the CZ is a challenging task given these vast timescales: driving

| **Page 4: Inserted** | **sue** | **1/19/2016 6:09:00 AM** |
|---|---|---|

driving CZ change

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 11:57:00 AM** |
|---|---|---|

,

| **Page 4: Inserted** | **Roman DiBiase** | **1/17/2016 12:10:00 AM** |
|---|---|---|

over thousands of years

| **Page 4: Inserted** | **sue** | **1/19/2016 6:08:00 AM** |
|---|---|---|

, building a model of the CZ is daunting

| **Page 4: Deleted** | **sue** | **1/21/2016 5:38:00 AM** |
|---|---|---|

.

| Page 4: Inserted | sue | 1/21/2016 5:38:00 AM |
|---|---|---|

and no single model has been developed. Instead,

| Page 4: Deleted | sue | 1/19/2016 6:09:00 AM |
|---|---|---|

One way to do this is to use a

| Page 4: Inserted | sue | 1/19/2016 6:10:00 AM |
|---|---|---|

s

| Page 4: Inserted | sue | 1/17/2016 5:39:00 AM |
|---|---|---|

or cascade

| Page 4: Inserted | sue | 1/19/2016 6:10:00 AM |
|---|---|---|

have been used to

| Page 4: Deleted | sue | 1/19/2016 6:10:00 AM |
|---|---|---|

that both

| Page 4: Inserted | Roman DiBiase | 1/16/2016 11:58:00 AM |
|---|---|---|

that both address specific

| Page 4: Deleted | sue | 1/17/2016 5:39:00 AM |
|---|---|---|

specific

| Page 4: Inserted | sue | 1/17/2016 5:39:00 AM |
|---|---|---|

important

| Page 4: Inserted | Roman DiBiase | 1/16/2016 11:58:00 AM |
|---|---|---|

processes

| Page 4: Inserted | sue | 1/17/2016 5:38:00 AM |
|---|---|---|

over different

| Page 4: Deleted | sue | 1/17/2016 5:38:00 AM |
|---|---|---|

and

| Page 4: Inserted | Roman DiBiase | 1/16/2016 11:58:00 AM |
|---|---|---|

and timescales (e.g., Godderis and Brantley, 2014)

| Page 4: Inserted | sue | 1/19/2016 6:11:00 AM |
|---|---|---|

(e.g., Godderis and Brantley, 2014).

| Page 4: Deleted | sue | 1/17/2016 5:39:00 AM |
|---|---|---|

and that can be interconnected to address broad overarching CZ problems as shown in Table 1 (Duffy et al., 2014; Table 1).

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 11:58:00 AM** |
|---|---|---|

and that can be interconnected to address broad overarching CZ problems

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 12:01:00 PM** |
|---|---|---|

; Table 1

| **Page 4: Inserted** | **sue** | **1/17/2016 5:40:00 AM** |
|---|---|---|

To enable treatment using such a suite of models,

| **Page 4: Deleted** | **Roman DiBiase** | **1/16/2016 12:01:00 PM** |
|---|---|---|

Then, f

| **Page 4: Deleted** | **sue** | **1/17/2016 5:40:00 AM** |
|---|---|---|

For

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 12:01:00 PM** |
|---|---|---|

F

| **Page 4: Inserted** | **sue** | **1/19/2016 6:12:00 AM** |
|---|---|---|

setting for

| **Page 4: Inserted** | **sue** | **1/19/2016 6:12:00 AM** |
|---|---|---|

research including CZ

| **Page 4: Inserted** | **sue** | **1/19/2016 6:12:00 AM** |
|---|---|---|

ies (CZOs)

| **Page 4: Deleted** | **sue** | **1/19/2016 6:12:00 AM** |
|---|---|---|

y

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 12:01:00 PM** |
|---|---|---|

| **Page 4: Inserted** | **sue** | **1/17/2016 5:40:00 AM** |
|---|---|---|

must

| **Page 4: Deleted** | **sue** | **1/17/2016 5:40:00 AM** |
|---|---|---|

then, it is necessary to

| **Page 4: Inserted** | **Roman DiBiase** | **1/16/2016 12:01:00 PM** |
|---|---|---|

then

**Page 4: Inserted** sue 1/17/2016 5:40:00 AM

with the necessity of measuring

**Page 4: Deleted** sue 1/17/2016 5:40:00 AM

with

**Page 4: Inserted** sue 1/17/2016 5:40:00 AM

the

**Page 4: Deleted** sue 1/19/2016 6:13:00 AM

Critical Zone Observatory

**Page 4: Inserted** sue 1/19/2016 6:13:00 AM

CZO

**Page 4: Inserted** sue 1/19/2016 6:12:00 AM

At the same time, we have been developing a suite of models that can be interconnected to address broad overarching CZ problems (Duffy et al., 2014; Table 1).

**Page 4: Inserted** sue 1/21/2016 5:38:00 AM

focus of the effort has been the

**Page 4: Deleted** sue 1/21/2016 5:38:00 AM

,

**Page 4: Inserted** sue 1/21/2016 5:39:00 AM

which was

**Page 4: Inserted** sue 1/19/2016 6:14:00 AM

hydrologic

**Page 4: Inserted** sue 1/19/2016 6:14:00 AM

was

**Page 4: Inserted** sue 1/19/2016 6:14:00 AM

with other disciplinary studies

**Page 4: Inserted** sue 1/26/2016 6:18:00 AM

as part of a network of CZOs in the U.S.A

**Page 4: Deleted** sue 1/19/2016 6:14:00 AM

, has been a successful location for CZ research

**Page 4: Inserted** sue 1/19/2016 6:14:00 AM

spatial

**Page 4: Deleted** sue 1/19/2016 6:14:00 AM

has

**Page 5: Inserted** Roman DiBiase 1/16/2016 12:03:00 PM

(up to XXX

**Page 5: Deleted** Dan K. Arthur 1/21/2016 2:57:00 PM

XXX

**Page 5: Formatted** sue 1/19/2016 6:15:00 AM

Highlight

**Page 5: Inserted** Dan K. Arthur 1/21/2016 2:57:00 PM

50

**Page 5: Inserted** Roman DiBiase 1/16/2016 12:03:00 PM

m deep)

**Page 5: Deleted** Roman DiBiase 1/17/2016 12:10:00 AM

**Page 5: Inserted** Roman DiBiase 1/17/2016 12:10:00 AM

**Page 5: Deleted** sue 1/26/2016 6:20:00 AM

The Shale Hills catchment is also situated on a single lithology (shale), which reduced the complexity ofsimplified the boundary conditions for models with respect to the initial chemical and physical conditions.

**Page 5: Inserted** Roman DiBiase 1/16/2016 12:04:00 PM

simplified the

**Page 5: Inserted** Roman DiBiase 1/17/2016 12:10:00 AM

**Page 5: Deleted** Roman DiBiase 1/17/2016 12:10:00 AM

**Page 5: Inserted** sue 1/21/2016 5:39:00 AM

he approach at Shale Hills has been t

| Page 5: Inserted | sue | 1/19/2016 6:17:00 AM |
|---|---|---|

incrementally

| Page 5: Inserted | sue | 1/19/2016 6:17:00 AM |
|---|---|---|

CZ

| Page 5: Deleted | sue | 1/19/2016 6:17:00 AM |
|---|---|---|

varying

| Page 5: Inserted | sue | 1/19/2016 6:17:00 AM |
|---|---|---|

increasing

| Page 5: Inserted | sue | 1/26/2016 6:20:00 AM |
|---|---|---|

. The catchment itself is situated on a single lithology (shale), which simplified the boundary conditions for models with respect to initial chemical and physical conditions. W

| Page 5: Deleted | sue | 1/21/2016 5:40:00 AM |
|---|---|---|

, w

| Page 5: Inserted | sue | 1/19/2016 6:20:00 AM |
|---|---|---|

have

| Page 5: Deleted | sue | 1/19/2016 6:20:00 AM |
|---|---|---|

catenas

| Page 5: Inserted | sue | 1/19/2016 6:20:00 AM |
|---|---|---|

planar hillslopes

| Page 5: Inserted | sue | 1/19/2016 6:20:00 AM |
|---|---|---|

within swales and

| Page 5: Deleted | sue | 1/19/2016 6:20:00 AM |
|---|---|---|

integrated across

| Page 5: Deleted | sue | 1/26/2016 6:20:00 AM |
|---|---|---|

is

| Page 5: Inserted | sue | 1/26/2016 6:20:00 AM |
|---|---|---|

must be considered in full

| Page 5: Inserted | sue | 1/19/2016 6:18:00 AM |
|---|---|---|

Where possible, t

| Page 5: Deleted | sue | 1/19/2016 6:18:00 AM |
|---|---|---|

T

**Page 5: Deleted** sue 1/19/2016 6:18:00 AM

were

**Page 5: Inserted** sue 1/19/2016 6:18:00 AM

have

**Page 5: Inserted** sue 1/19/2016 6:18:00 AM

been

**Page 5: Deleted** sue 1/19/2016 6:31:00 AM

The characterization and modelling of settings as

**Page 5: Inserted** sue 1/19/2016 6:31:00 AM

Using the conceptualization of "

**Page 5: Deleted** sue 1/19/2016 6:31:00 AM

or

**Page 5: Inserted** sue 1/19/2016 6:31:00 AM

and

**Page 5: Inserted** sue 1/19/2016 6:31:00 AM

"

**Page 5: Inserted** sue 1/19/2016 6:31:00 AM

settings

**Page 5: Inserted** sue 1/19/2016 6:30:00 AM

has allowed measurements and modelling to proceed in a synergistic fashion: the reduction of

**Page 5: Deleted** sue 1/19/2016 6:18:00 AM

has allowed us to reduce some of the

**Page 5: Inserted** sue 1/19/2016 6:32:00 AM

in 1D and 2D sites

**Page 5: Deleted** sue 1/19/2016 6:19:00 AM

of

**Page 5: Inserted** sue 1/19/2016 6:29:00 AM

enabled

**Page 5: Deleted** sue 1/19/2016 6:29:00 AM

the

| Page 5: Inserted | sue | 1/19/2016 6:29:00 AM |
|---|---|---|

development of

| Page 5: Inserted | sue | 1/19/2016 6:29:00 AM |
|---|---|---|

s

| Page 5: Deleted | sue | 1/19/2016 6:29:00 AM |
|---|---|---|

systems,

| Page 5: Inserted | Roman DiBiase | 1/16/2016 12:06:00 PM |
|---|---|---|

, and

| Page 5: Deleted | sue | 1/19/2016 6:19:00 AM |
|---|---|---|

and

| Page 5: Inserted | sue | 1/19/2016 6:19:00 AM |
|---|---|---|

but also

| Page 5: Deleted | sue | 1/19/2016 6:32:00 AM |
|---|---|---|

to

| Page 5: Inserted | Roman DiBiase | 1/16/2016 12:06:00 PM |
|---|---|---|

to focus

| Page 5: Inserted | sue | 1/19/2016 6:32:00 AM |
|---|---|---|

ed

| Page 5: Inserted | Roman DiBiase | 1/16/2016 12:06:00 PM |
|---|---|---|

our sampling schemes

| Page 5: Deleted | Roman DiBiase | 1/16/2016 12:10:00 PM |
|---|---|---|

This formulation also led to recognition of the importance of two types of hillslopes: planar hillslopes that experience downslope but nonconvergent flow of water and soil (essentially 2D evolution), and swales that experience downslope convergent flow of water and soil (3D evolution). Much of our effort focused on understanding soils and waters in these two hydrologic units (Fig. 2).

| Page 5: Inserted | sue | 1/19/2016 6:21:00 AM |
|---|---|---|

conceptualizations

| Page 5: Deleted | sue | 1/19/2016 6:21:00 AM |
|---|---|---|

s

**Page 5: Deleted** sue 1/19/2016 6:32:00 AM

have been

**Page 5: Inserted** sue 1/19/2016 6:32:00 AM

were

**Page 5: Deleted** sue 1/19/2016 6:32:00 AM

strictly

**Page 5: Inserted** sue 1/19/2016 6:32:00 AM

first

**Page 5: Inserted** sue 1/19/2016 6:46:00 AM

s

**Page 5: Inserted** Roman DiBiase 1/16/2016 12:10:00 PM

(1D)

**Page 5: Deleted** sue 1/19/2016 6:32:00 AM

or

**Page 5: Inserted** sue 1/19/2016 6:32:00 AM

and then for

**Page 5: Inserted** Roman DiBiase 1/16/2016 12:10:00 PM

(2D)

**Page 5: Inserted** sue 1/19/2016 6:19:00 AM

and have been highly influenced by our soil chemistry measurements on ridgetops and planar hillslope catenas

**Page 5: Inserted** sue 1/21/2016 5:41:00 AM

; Lebedeva and Brantley, 2013; West et al., 2013; Ma et al. 2013

**Page 5: Formatted** sue 1/21/2016 5:42:00 AM

Highlight

**Page 5: Inserted** sue 1/19/2016 6:15:00 AM

. In some cases modelling and measurement proceed hand in hand while in others, the modelling lags. For example, soil measurements have been collected in hillslopes characterized by convergent water and soil flow regimes, i.e., swales (Jin and Brantley, 2011) and soil observations have been collected across much of the catchment, but soil formation models for swales or the entire catchment still remain to be developed.

In contrast to the soil formation models that have targeted the 1D and 2D sites

| Page 5: Formatted | sue | 1/19/2016 6:46:00 AM |
|---|---|---|

Line spacing: 1.5 lines

| Page 5: Deleted | sue | 1/19/2016 6:16:00 AM |
|---|---|---|

while

| Page 5: Inserted | sue | 1/19/2016 6:22:00 AM |
|---|---|---|

have been

| Page 5: Deleted | sue | 1/19/2016 6:43:00 AM |
|---|---|---|

are heavily influenced by the convergent swales

| Page 5: Inserted | Roman DiBiase | 1/16/2016 12:10:00 PM |
|---|---|---|

convergent

| Page 5: Inserted | sue | 1/21/2016 5:47:00 AM |
|---|---|---|

e.g.,

| Page 5: Inserted | sue | 1/19/2016 6:22:00 AM |
|---|---|---|

| Page 5: Inserted | sue | 1/21/2016 5:48:00 AM |
|---|---|---|

In fact, study of an entire catchment with a hydrologic model is sometimes more tractable than for smaller sub-systems because the large-scale study allows a continuum treatment whereas treatment of smaller scale sub-systems within the catchment might require measurements of the exact positions of heterogeneities such as fractures, faults, low-permeability zones, etc.

| Page 5: Deleted | sue | 1/19/2016 6:45:00 AM |
|---|---|---|

| Page 5: Deleted | sue | 1/19/2016 6:48:00 AM |
|---|---|---|

upscale

| Page 5: Inserted | sue | 1/19/2016 6:48:00 AM |
|---|---|---|

grapple with some of these down- and up-scaling issues by expanding the CZO

| Page 6: Deleted | sue | 1/19/2016 6:48:00 AM |
|---|---|---|

us to

**Page 6: Inserted** sue 1/19/2016 6:48:00 AM

ion of

**Page 6: Deleted** sue 1/19/2016 6:48:00 AM

e

**Page 6: Deleted** sue 1/19/2016 6:48:00 AM

our developing

**Page 6: Deleted** sue 1/19/2016 6:49:00 AM

For our approach

**Page 6: Inserted** sue 1/19/2016 6:49:00 AM

To enable understanding of the larger watershed

**Page 6: Inserted** sue 1/17/2016 5:43:00 AM

-

**Page 6: Deleted** sue 1/17/2016 5:43:00 AM

**Page 6: Inserted** Roman DiBiase 1/17/2016 12:11:00 AM

Thus, we can directly evaluate

**Page 6: Deleted** sue 1/17/2016 5:43:00 AM

us, we can directly evaluate

**Page 6: Inserted** sue 1/17/2016 5:43:00 AM

is allows evaluation of

**Page 6: Inserted** Roman DiBiase 1/17/2016 12:11:00 AM

how much of our understanding from the Shale Hills site is transferable to

**Page 6: Deleted** sue 1/17/2016 5:43:00 AM

the

**Page 6: Deleted** sue 1/17/2016 5:43:00 AM

site

**Page 6: Inserted** sue 1/17/2016 5:43:00 AM

other lithologies with

**Page 6: Inserted** Roman DiBiase 1/17/2016 12:11:00 AM

different boundary

| Page 6: Deleted | sue | 1/17/2016 5:43:00 AM |
|---|---|---|

boundary

| Page 6: Inserted | sue | 1/17/2016 5:43:00 AM |
|---|---|---|

initial

| Page 6: Inserted | Roman DiBiase | 1/17/2016 12:11:00 AM |
|---|---|---|

conditions

| Page 6: Inserted | sue | 1/19/2016 6:50:00 AM |
|---|---|---|

but with the same climate

| Page 6: Deleted | sue | 1/17/2016 5:43:00 AM |
|---|---|---|

in isolation

| Page 6: Inserted | Roman DiBiase | 1/17/2016 12:12:00 AM |
|---|---|---|

in isolation.

| Page 6: Inserted | sue | 1/19/2016 6:50:00 AM |
|---|---|---|

of the mainstem of the stream

| Page 6: Inserted | sue | 1/19/2016 6:50:00 AM |
|---|---|---|

within the larger watershed

| Page 6: Inserted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

| Page 6: Deleted | sue | 1/19/2016 6:51:00 AM |
|---|---|---|

,

| Page 6: Inserted | sue | 1/19/2016 6:51:00 AM |
|---|---|---|

s

| Page 6: Deleted | Roman DiBiase | 1/16/2016 12:15:00 PM |
|---|---|---|

for example, in

| Page 6: Deleted | Roman DiBiase | 1/16/2016 12:16:00 PM |
|---|---|---|

over

| Page 6: Inserted | Roman DiBiase | 1/16/2016 12:16:00 PM |
|---|---|---|

across

| Page 6: Deleted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

By necessity, t

| Page 6: Inserted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

T

| Page 6: Deleted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

, including the

| Page 6: Inserted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

and its

| Page 6: Inserted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

it is necessary to

| Page 6: Deleted | sue | 1/21/2016 5:51:00 AM |
|---|---|---|

we must

| Page 6: Inserted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

in the CZO;

| Page 6: Deleted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

and

| Page 6: Inserted | sue | 1/26/2016 6:24:00 AM |
|---|---|---|

(Fig. 1)

| Page 6: Deleted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

an

| Page 6: Inserted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

our

| Page 6: Deleted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

ideal

| Page 6: Inserted | sue | 1/21/2016 5:52:00 AM |
|---|---|---|

proposed

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

differ from such designs at other s

| Page 6: Deleted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

vary from s

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

s

| Page 6: Deleted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

to site

| Page 6: Deleted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

Nonetheless w

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

W

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

nonetheless

| Page 6: Deleted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

as a way to hypothesize an answer to the

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

to stimulate focus on the broad

| Page 6: Inserted | sue | 1/21/2016 5:53:00 AM |
|---|---|---|

to best learn about its evolution and function

| Page 6: Inserted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

To exemplify our design, w

| Page 6: Deleted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

W

| Page 6: Deleted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

then

| Page 6: Inserted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

also

| Page 6: Deleted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

present specific examples for

| Page 6: Inserted | sue | 1/21/2016 5:54:00 AM |
|---|---|---|

describe

| Page 7: Deleted | sue | 1/17/2016 6:00:00 AM |
|---|---|---|

**Building a co**

| Page 7: Inserted | Roman DiBiase | 1/17/2016 12:16:00 AM |
|---|---|---|

**Building a co**

| Page 7: Inserted | sue | 1/17/2016 6:00:00 AM |
|---|---|---|

**Co**

| Page 7: Inserted | Roman DiBiase | 1/17/2016 12:20:00 AM |
|---|---|---|

**nnection**

| Page 7: Inserted | sue | 1/17/2016 6:00:00 AM |
|---|---|---|

**s**

| Page 7: Inserted | Roman DiBiase | 1/17/2016 12:20:00 AM |
|---|---|---|

**between model development and field measurements**

| Page 7: Deleted | Roman DiBiase | 1/16/2016 1:18:00 PM |
|---|---|---|

**Rationale for the measurement plan**

| Page 7: Inserted | sue | 1/17/2016 5:46:00 AM |
|---|---|---|

The suite of models shown in Table 1 is designed to develop understanding over the entire CZ as an integral object of study, i.e., one system.

| Page 7: Deleted | sue | 1/21/2016 5:55:00 AM |
|---|---|---|

Ideally, the connection between model development and field measurements is two-directional

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:28:00 PM |
|---|---|---|

Ideally, the connection between model development and field measurements is two-directional

| Page 7: Inserted | sue | 1/26/2016 6:26:00 AM |
|---|---|---|

F

| Page 7: Deleted | sue | 1/26/2016 6:26:00 AM |
|---|---|---|

; f

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:28:00 PM |
|---|---|---|

; field measurements are prioritized and driven by data needs for developing models (e.g., Table 1), and at the same time model development is dictacted by observations in the field.

| Page 7: Deleted | sue | 1/21/2016 5:59:00 AM |
|---|---|---|

,

| Page 7: Deleted | sue | 1/21/2016 5:59:00 AM |
|---|---|---|

at the same time

| Page 7: Deleted | sue | 1/17/2016 5:46:00 AM |
|---|---|---|

c

**Page 7: Inserted** sue 1/21/2016 5:57:00 AM

Hand in hand with this system-level approach, researchers from different disciplines also bring discipline-specific hypotheses to their research that are related to disciplinary gaps in knowledge. Thus, disciplinary-level hypotheses also drive CZO research and sometimes these hypotheses feed directly into the overall CZ suite of models. Furthermore, because our understanding of the complicated suite of CZ processes is still in its infancy, both baseline measurements and curiosity-driven sample collection are still vital to determine the important processes. Throughout, models and observations are allowed to evolve to enable the two-way exchange of insights needed to maximize CZ science.

**Page 7: Deleted** Roman DiBiase 1/16/2016 12:32:00 PM

The choices of measurements to be made during the expansion are driven by data needs for the models under development (Table 1), at the same time that the models are driven by observations in the field.

**Page 7: Deleted** sue 1/17/2016 5:46:00 AM

The suite of models shown in Table 1 is one way tois aimed at understanding the entire CZ as an object of study, rather than as a set of disparate systems.

**Page 7: Inserted** Roman DiBiase 1/17/2016 12:15:00 AM

is aimed at

**Page 7: Inserted** Roman DiBiase 1/16/2016 12:33:00 PM

ing

**Page 7: Deleted** Roman DiBiase 1/16/2016 12:33:00 PM

In coordination with this modelling approach

**Page 7: Deleted** sue 1/17/2016 5:49:00 AM

Similarly

**Page 7: Inserted** Roman DiBiase 1/16/2016 12:33:00 PM

Similarly

**Page 7: Inserted** sue 1/17/2016 5:49:00 AM

Given all the needs for data

**Page 7: Deleted** Roman DiBiase 1/16/2016 12:33:00 PM a

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:33:00 PM |
|---|---|---|

the

| Page 7: Deleted | Roman DiBiase | 1/16/2016 12:33:00 PM |
|---|---|---|

stratified

| Page 7: Inserted | sue | 1/17/2016 5:49:00 AM |
|---|---|---|

which is

| Page 7: Deleted | Roman DiBiase | 1/16/2016 12:33:00 PM |
|---|---|---|

is

| Page 7: Deleted | sue | 1/17/2016 5:49:00 AM |
|---|---|---|

being

| Page 7: Deleted | sue | 1/17/2016 5:49:00 AM |
|---|---|---|

the

| Page 7: Inserted | sue | 1/17/2016 5:49:00 AM |
|---|---|---|

a

| Page 7: Deleted | sue | 1/26/2016 6:30:00 AM |
|---|---|---|

target

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:33:00 PM |
|---|---|---|

target

| Page 7: Inserted | sue | 1/26/2016 6:30:00 AM |
|---|---|---|

must provide both

| Page 7: Deleted | sue | 1/17/2016 5:50:00 AM |
|---|---|---|

s

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:33:00 PM |
|---|---|---|

s focused measurements for testing

| Page 7: Deleted | sue | 1/17/2016 5:50:00 AM |
|---|---|---|

focused

| Page 7: Deleted | sue | 1/17/2016 5:50:00 AM |
|---|---|---|

for testing

| Page 7: Inserted | sue | 1/17/2016 5:50:00 AM |
|---|---|---|

to test

**Page 7: Inserted** **Roman DiBiase** **1/16/2016 12:34:00 PM**

disciplinary hypotheses (e.g., eddy flux covariance, soil gas composition) and observations

**Page 7: Deleted** **sue** **1/21/2016 6:01:00 AM**

(e.g., eddy flux covariance, soil gas composition)

**Page 7: Inserted** **sue** **1/17/2016 5:50:00 AM**

necessary to bridge

**Page 7: Deleted** **sue** **1/17/2016 5:50:00 AM**

that are central

**Page 7: Inserted** **Roman DiBiase** **1/16/2016 12:36:00 PM**

that are central across disciplines (e.g., soil thickness, topography). Additionally,

**Page 7: Deleted** **sue** **1/21/2016 6:01:00 AM**

(e.g., soil thickness, topography)

**Page 7: Inserted** **sue** **1/17/2016 5:51:00 AM**

certain measurements such as

**Page 7: Inserted** **Roman DiBiase** **1/16/2016 12:40:00 PM**

geophysical and remote sensing surveys, catchment-integrating stream measurements,

**Page 7: Inserted** **sue** **1/17/2016 5:51:00 AM**

and

**Page 7: Inserted** **Roman DiBiase** **1/16/2016 12:42:00 PM**

time-integrating analysis of alluvial and colluvial sediments

**Page 7: Inserted** **sue** **1/17/2016 5:51:00 AM**

can be made along with

**Page 7: Deleted** **sue** **1/17/2016 5:52:00 AM**

, and

**Page 7: Inserted** **Roman DiBiase** **1/16/2016 12:42:00 PM**

, and model simulations will be used to

**Page 7: Deleted** **sue** **1/17/2016 5:52:00 AM**

will be used

**Page 7: Deleted** **Roman DiBiase** **1/16/2016 12:41:00 PM**

and paired with geophysical, catchment-scale stream and remote sensing measurements. The models will then be used to

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

across space (

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

)

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

time (

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

to

**Page 7: Deleted** sue 1/17/2016 5:52:00 AM

over longer

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

geological

**Page 7: Inserted** sue 1/17/2016 5:52:00 AM

)

**Page 7: Deleted** sue 1/21/2016 6:03:00 AM

including

**Page 7: Inserted** sue 1/21/2016 6:03:00 AM

ranging from the assessment of

**Page 7: Inserted** sue 1/21/2016 6:03:00 AM

to variations in land use

**Page 7: Inserted** Roman DiBiase 1/16/2016 12:48:00 PM

Because the mixing timescales of biota, regolith, and bedrock are relatively slow (compared to mixing of atmospheric and surface water reservoirs), the assessment of the spatial distribution of biota, regolith, and bedrock properties is both important and extremely challenging (Niu et al., 2014).

**Page 7: Deleted** Roman DiBiase 1/16/2016 12:45:00 PM

In other words, while assessment of atmospheric and surface water pools can be technically challenging, mixing of these pools is much faster than mixing of the biotic pool, the regolith and rock reservoir, or the pool of soil porewater, making assessment of the spatial distribution of these latter reservoirs exceedingly difficult (Niu et al., 2014).

| **Page 7: Deleted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

the

| **Page 7: Inserted** | **Roman DiBiase** | **1/16/2016 12:52:00 PM** |
|---|---|---|

rapid

| **Page 7: Deleted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

rates of

| **Page 7: Deleted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

in the land surface pools is generally much slower than the changing

| **Page 7: Inserted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

in the

| **Page 7: Deleted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

, making

| **Page 7: Inserted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

makes

| **Page 7: Inserted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

robust

| **Page 7: Deleted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

sensor requirements

| **Page 7: Inserted** | **Roman DiBiase** | **1/16/2016 12:53:00 PM** |
|---|---|---|

measurements

| **Page 7: Inserted** | **sue** | **1/24/2016 7:55:00 PM** |
|---|---|---|

The hydrologic state is intermediate, exhibiting large spatial and temporal variability.

| **Page 7: Formatted** | **sue** | **1/24/2016 7:56:00 PM** |
|---|---|---|

Highlight

| **Page 7: Deleted** | **sue** | **1/26/2016 6:33:00 AM** |
|---|---|---|

Thus, the largest difficulty in *temporally* characterizing today's fluxes in the CZ in any observatory may be measuring the fast-changing fluxes of thechanges in atmospheric pools and fluxes.

| Page 7: Inserted | Roman DiBiase | 1/16/2016 12:54:00 PM |
|---|---|---|

changes in

| Page 7: Inserted | sue | 1/16/2016 6:01:00 AM |
|---|---|---|

In recognition of these difficulties, the project started at Shale Hills precisely because it is a catchment almost 100% underlain by Rose Hill formation shale and strictly managed as forestland.

| Page 7: Deleted | sue | 1/16/2016 6:01:00 AM |
|---|---|---|

In recognition of these difficulties, the project started at Shale Hills precisely because it is a catchment almost 100% underlain by Rose Hill formation shale with land use strictly as managed forest.

| Page 7: Inserted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

differences in sedimentary layers,

| Page 8: Deleted | sue | 1/26/2016 6:34:00 AM |
|---|---|---|

1

| Page 8: Inserted | sue | 1/26/2016 6:34:00 AM |
|---|---|---|

3

| Page 8: Inserted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

the term

| Page 8: Deleted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

is used to

| Page 8: Inserted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

s

| Page 8: Deleted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

graded

| Page 8: Inserted | sue | 1/21/2016 6:04:00 AM |
|---|---|---|

eroded

| Page 8: Inserted | sue | 1/21/2016 6:05:00 AM |
|---|---|---|

not to be random but rather to be

| Page 8: Deleted | sue | 1/21/2016 6:05:00 AM |
|---|---|---|

using a stratified sampling plan

| Page 8: Inserted | sue | 1/21/2016 6:05:00 AM |
|---|---|---|

stratified

| Page 8: Inserted | sue | 1/21/2016 6:05:00 AM |
|---|---|---|

d

| Page 8: Deleted | sue | 1/21/2016 6:05:00 AM |
|---|---|---|

rather than random sampling

| Page 8: Deleted | sue | 1/21/2016 6:06:00 AM |
|---|---|---|

when it is designed as

| Page 8: Inserted | sue | 1/21/2016 6:06:00 AM |
|---|---|---|

for

| Page 8: Inserted | sue | 1/17/2016 5:54:00 AM |
|---|---|---|

especially

| Page 8: Inserted | sue | 1/17/2016 5:54:00 AM |
|---|---|---|

)

| Page 8: Deleted | sue | 1/17/2016 5:54:00 AM |
|---|---|---|

)

| Page 8: Inserted | sue | 1/17/2016 5:54:00 AM |
|---|---|---|

a first-order observation about hillslope morphology in Shale Hills based in long-standing observations from hillslope geomorphology is the delineation between planar slopes and swales: the former experience largely 2D nonconvergent flow while the latter experience 3D convergent flow of water and soil.

| Page 8: Deleted | sue | 1/17/2016 5:56:00 AM |
|---|---|---|

where

| Page 8: Inserted | sue | 1/17/2016 5:56:00 AM |
|---|---|---|

Where

| Page 8: Deleted | sue | 1/17/2016 5:54:00 AM |
|---|---|---|

many

| Page 8: Deleted | sue | 1/17/2016 5:56:00 AM |
|---|---|---|

not recognized

| Page 8: Inserted | sue | 1/17/2016 5:56:00 AM |
|---|---|---|

ignored

**Page 8: Deleted** sue 1/17/2016 5:56:00 AM

if

**Page 8: Inserted** sue 1/17/2016 5:56:00 AM

when

**Page 8: Deleted** sue 1/17/2016 5:56:00 AM

these two features are recognized and

**Page 8: Inserted** sue 1/17/2016 5:56:00 AM

separately

**Page 8: Deleted** sue 1/17/2016 5:57:00 AM

Furthermore, one of the models under development for the CZO is a regolith formation model: by using this model to understand regolith formation, the number of pits in regolith can similarly be minimized.

**Page 8: Inserted** sue 1/21/2016 6:09:00 AM

Another aspect of our stratified sampling plan is to complement m

**Page 8: Deleted** sue 1/21/2016 6:09:00 AM

M

**Page 8: Deleted** sue 1/21/2016 6:10:00 AM

will soon be supplemented with

**Page 8: Inserted** sue 1/21/2016 6:10:00 AM

by

**Page 8: Deleted** sue 1/21/2016 6:10:00 AM

instrumentation

**Page 8: Inserted** sue 1/21/2016 6:10:00 AM

measurements

**Page 8: Deleted** sue 1/21/2016 6:10:00 AM

the

**Page 8: Inserted** sue 1/21/2016 6:10:00 AM

**Page 8: Deleted** sue 1/21/2016 6:10:00 AM

watershed. The subcatchments were

| Page 8: Inserted | sue | 1/21/2016 6:11:00 AM |
|---|---|---|

. Once again the stratification of the sampling design is dictated by geological knowledge: b

| Page 8: Deleted | sue | 1/21/2016 6:11:00 AM |
|---|---|---|

, as b

| Page 8: Inserted | sue | 1/21/2016 6:11:00 AM |
|---|---|---|

| Page 8: Inserted | sue | 1/21/2016 6:12:00 AM |
|---|---|---|

such new subcatchment

| Page 8: Deleted | sue | 1/21/2016 6:12:00 AM |
|---|---|---|

a new

| Page 8: Inserted | sue | 1/21/2016 6:12:00 AM |
|---|---|---|

and

| Page 8: Deleted | sue | 1/21/2016 6:12:00 AM |
|---|---|---|

subcatchment

| Page 8: Inserted | sue | 1/21/2016 6:20:00 AM |
|---|---|---|

o upscale from subcatchments to Shavers Creek, t

| Page 8: Deleted | Roman DiBiase | 1/17/2016 12:30:00 AM |
|---|---|---|

t

| Page 8: Deleted | sue | 1/21/2016 6:20:00 AM |
|---|---|---|

to upscale from Shale Hills to Shavers Creek

| Page 8: Deleted | sue | 1/21/2016 6:21:00 AM |
|---|---|---|

only a

| Page 8: Inserted | sue | 1/21/2016 6:21:00 AM |
|---|---|---|

the

| Page 8: Inserted | sue | 1/21/2016 6:21:00 AM |
|---|---|---|

chosen

| Page 9: Inserted | sue | 1/21/2016 6:22:00 AM |
|---|---|---|

e.g.,

| Page 9: Inserted | sue | 1/21/2016 6:23:00 AM |
|---|---|---|

i) integrative measurements (geophysics, lidar, stream, atmosphere), and ii)

| **Page 9: Deleted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

through

| **Page 9: Inserted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

through

| **Page 9: Deleted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

se

| **Page 9: Inserted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

upland catchments under study in the CZO

| **Page 9: Deleted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

small catchments

| **Page 9: Deleted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

is necessary to

| **Page 9: Inserted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

will enable better

| **Page 9: Inserted** | **sue** | **1/21/2016 6:24:00 AM** |
|---|---|---|

ions of

| **Page 9: Deleted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

in the catchment

| **Page 9: Inserted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

Of course, t

| **Page 9: Deleted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

T

| **Page 9: Inserted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

s

| **Page 9: Inserted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

continually

| **Page 9: Deleted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

with

| **Page 9: Inserted** | **sue** | **1/21/2016 6:25:00 AM** |
|---|---|---|

against

**Page 9: Inserted** | **sue** | **1/15/2016 7:23:00 PM**

For clarity in describing the measurements i

**Page 9: Deleted** | **sue** | **1/15/2016 7:23:00 PM**

I

**Page 9: Inserted** | **sue** | **1/15/2016 7:26:00 PM**

that are needed for the models

**Page 9: Inserted** | **sue** | **1/26/2016 6:42:00 AM**

(Table S1)

**Page 9: Deleted** | **sue** | **1/15/2016 7:28:00 PM**

for clarity of description

**Page 9: Inserted** | **sue** | **1/21/2016 6:26:00 AM**

The Ground HOG deployments also are the locations for assessments of v

**Page 9: Deleted** | **sue** | **1/21/2016 6:26:00 AM**

V

**Page 9: Deleted** | **sue** | **1/21/2016 6:27:00 AM**

is being assessed

**Page 9: Deleted** | **sue** | **1/21/2016 6:27:00 AM**

t

**Page 9: Inserted** | **sue** | **1/21/2016 6:27:00 AM**

cross

**Page 9: Deleted** | **sue** | **1/21/2016 6:26:00 AM**

located coincident with the Ground HOG

**Page 9: Deleted** | **sue** | **1/21/2016 6:27:00 AM**

**Page 9: Deleted** | **sue** | **1/21/2016 6:27:00 AM**

being

**Page 9: Deleted** sue 1/21/2016 6:27:00 AM

being

**Page 9: Inserted** sue 1/21/2016 6:28:00 AM

As discussed above, t

**Page 9: Deleted** sue 1/21/2016 6:28:00 AM

T

**Page 9: Inserted** sue 1/21/2016 6:28:00 AM

provide

**Page 9: Deleted** sue 1/21/2016 6:28:00 AM

are

**Page 9: Deleted** sue 1/21/2016 6:28:00 AM

integrators

**Page 9: Inserted** sue 1/21/2016 6:28:00 AM

integrations

**Page 9: Inserted** sue 1/21/2016 6:33:00 AM

Data from Ground HOG and Tower HOG

**Page 9: Deleted** sue 1/21/2016 6:33:00 AM

Stream and ground water data

**Page 9: Inserted** sue 1/21/2016 6:29:00 AM

be used to

**Page 9: Deleted** sue 1/17/2016 6:00:00 AM

, as described below

**Page 9: Inserted** sue 1/21/2016 6:33:00 AM

**Page 9: Deleted** sue 1/15/2016 7:28:00 PM

**Page 9: Deleted** Roman DiBiase 1/16/2016 1:17:00 PM

## Data assimilation

| Page 9: Inserted | sue | 1/15/2016 6:52:00 PM |
|---|---|---|

In fact, t

| Page 9: Deleted | sue | 1/15/2016 6:55:00 PM |
|---|---|---|

T

| Page 9: Deleted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2014b

| Page 9: Inserted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2015b

| Page 10: Deleted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2014b

| Page 10: Inserted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2015b

| Page 10: Deleted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2014a

| Page 10: Inserted | Dan K. Arthur | 1/25/2016 9:35:00 AM |
|---|---|---|

2015a

| Page 10: Deleted | Dan K. Arthur | 1/25/2016 9:36:00 AM |
|---|---|---|

2014b

| Page 10: Inserted | Dan K. Arthur | 1/25/2016 9:36:00 AM |
|---|---|---|

2015b

| Page 11: Inserted | sue | 1/26/2016 6:40:00 AM |
|---|---|---|

(

| Page 11: Deleted | sue | 1/26/2016 6:40:00 AM |
|---|---|---|

. Notably,

| Page 11: Inserted | sue | 1/26/2016 6:40:00 AM |
|---|---|---|

)

| Page 11: Deleted | sue | 1/21/2016 6:36:00 AM |
|---|---|---|

ness of

| Page 11: Inserted | sue | 1/21/2016 6:36:00 AM |
|---|---|---|

spatial data available in

| Page 11: Deleted | sue | 1/26/2016 6:43:00 AM |
|---|---|---|

in the catchment

| Page 11: Inserted | sue | 1/21/2016 6:38:00 AM |
|---|---|---|

To the extent possible, we parameterize these 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, 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. Data assimilation can thus provide optimal estimates of observable variables and parameters, taking into account both the uncertainties of model predictions and observations.

As new types of observations are provided, we first evaluate PIHM model output against the new observations prior to calibrations to see if the current calibration predicts the new data. This comparison is ongoing for the Garner Run subcatchment. If the prediction is poor, this yields insight into the capabilities of our model under new conditions. If we discover that even with a new calibration we cannot successfully predict the new observations, we will incorporate a new module that describes a new phenomenon in PIHM. For example, discrepancies between model output and preliminary observations at Garner Run has led us to hypothesize that the distribution of boulders on the land surface – a phenomenon not observed in the Shale Hills catchment – must be incorporated into the PIHM models. 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.

| Page 12: Formatted | sue | 1/15/2016 7:18:00 PM |
|---|---|---|

Line spacing: 1.5 lines

| Page 12: Deleted | sue | 1/21/2016 6:37:00 AM |
|---|---|---|

| Page 12: Inserted | sue | 1/17/2016 6:02:00 AM |
|---|---|---|

These discussions about the design of a sampling strategy can best be explained through examples.

| Page 12: Deleted | sue | 1/22/2016 6:06:00 AM |
|---|---|---|

areas

| Page 12: Inserted | sue | 1/22/2016 6:06:00 AM |
|---|---|---|

subcatchments

| Page 12: Inserted | sue | 1/22/2016 6:06:00 AM |
|---|---|---|

To exemplify the approach, w

| Page 12: Deleted | sue | 1/17/2016 6:03:00 AM |
|---|---|---|

In addition to describing the geologic and geomorphologic setting, we detail the sampling strategy.

| Page 12: Inserted | sue | 1/17/2016 6:03:00 AM |
|---|---|---|

e describe the setting and some p

| Page 12: Deleted | sue | 1/17/2016 6:03:00 AM |
|---|---|---|

P

| Page 12: Deleted | sue | 1/26/2016 6:45:00 AM |
|---|---|---|

surface

| Page 12: Deleted | sue | 1/22/2016 6:06:00 AM |
|---|---|---|

are also presented

| Page 12: Deleted | sue | 1/22/2016 6:09:00 AM |
|---|---|---|

the

| Page 12: Inserted | sue | 1/22/2016 6:09:00 AM |
|---|---|---|

SSH

| Page 12: Deleted | Roman DiBiase | 1/17/2016 12:31:00 AM |
|---|---|---|

the

| Page 12: Inserted | sue | 1/22/2016 6:10:00 AM |
|---|---|---|

start by describing

**Page 12: Deleted** sue 1/22/2016 6:10:00 AM

include a description of

**Page 12: Inserted** sue 1/22/2016 6:10:00 AM

the

**Page 12: Deleted** sue 1/22/2016 6:10:00 AM

new

**Page 12: Inserted** sue 1/22/2016 6:10:00 AM

Garner Run

**Page 12: Inserted** sue 1/26/2016 5:51:00 AM

and land use

**Page 12: Deleted** sue 1/22/2016 6:10:00 AM

at Garner Run

**Page 12: Deleted** sue 1/26/2016 6:46:00 AM

quartzite

**Page 12: Inserted** sue 1/26/2016 6:46:00 AM

sandstone

**Page 13: Deleted** sue 1/22/2016 6:10:00 AM

the

**Page 13: Inserted** sue 1/22/2016 6:11:00 AM

roughly planar

**Page 13: Deleted** sue 1/22/2016 6:11:00 AM

are roughly planar and

**Page 13: Inserted** Roman DiBiase 1/17/2016 12:31:00 AM

are roughly planar and

**Page 13: Inserted** sue 1/22/2016 6:11:00 AM

the

**Page 13: Inserted** sue 1/26/2016 6:46:00 AM

4,

**Page 13: Deleted** Roman DiBiase 1/17/2016 12:32:00 AM

One question that must be addressed is whether the Rose Hill Shale of the Clinton Group, which underlies Shale Hills and lies stratigraphically above the Tuscarora, may be present in the Garner Run subcatchment. Down-valley of the Garner Run study area, the Rose Hill Shale has been mapped in a low-sloping bench at the foot of Tussey Mountain (Figs. 3,4, Flueckinger, 1969). Although the entirety of the Garner Run study area is mapped as the Tuscarora Formation, the continuation of a low-sloping bench along the entire valley (Fig. 7) could be consistent with the presence of Rose Hill Shale throughout the catchment. In general, bedrock exposure is poor in the Shavers Creek watershed, but lidar topographic analysis, field mapping, and targeted geophysical surveys will aid in resolving uncertainties in subsurface composition necessary for modeling water, solute, and sediment fluxes.

| Page 13: Inserted | sue | 1/22/2016 6:11:00 AM |
|---|---|---|

also contrasts strikingly in several ways from

| Page 13: Deleted | sue | 1/22/2016 6:11:00 AM |
|---|---|---|

shows a number of striking contrasts to

| Page 13: Deleted | sue | 1/22/2016 6:12:00 AM |
|---|---|---|

the

| Page 13: Deleted | sue | 1/22/2016 6:12:00 AM |
|---|---|---|

catchment

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

sandstone

| Page 13: Deleted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

(Fig. 6)

| Page 13: Deleted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

(including Shale Hills)

| Page 13: Deleted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

(Fig. 6)

| Page 13: Inserted | sue | 1/22/2016 6:12:00 AM |
|---|---|---|

significantly

| Page 13: Deleted | sue | 1/26/2016 6:48:00 AM |
|---|---|---|

(quartzite vs. shale)

| Page 13: Deleted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

This

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:16:00 PM |
|---|---|---|

The

| Page 13: Deleted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

their presumably

| Page 13: Inserted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

that both subcatchments are presumed to have experienced

| Page 13: Inserted | sue | 1/22/2016 6:16:00 AM |
|---|---|---|

local

| Page 13: Inserted | sue | 1/22/2016 6:16:00 AM |
|---|---|---|

the same regional

| Page 13: Inserted | sue | 1/22/2016 6:17:00 AM |
|---|---|---|

that

| Page 13: Deleted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

Garner Run

| Page 13: Inserted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

the sandstone

| Page 13: Deleted | sue | 1/22/2016 6:17:00 AM |
|---|---|---|

to

| Page 13: Inserted | sue | 1/22/2016 6:17:00 AM |
|---|---|---|

would

| Page 13: Deleted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

with time

| Page 13: Deleted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

have

| Page 13: Inserted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

generate

| Page 13: Inserted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

. Thus the shallower slopes on the resistant sandstone contradicts the general idea that

| Page 13: Deleted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

in order to erode

| Page 13: Inserted | sue | 1/22/2016 6:14:00 AM |
|---|---|---|

erosion

| Page 13: Deleted | sue | 1/22/2016 6:13:00 AM |
|---|---|---|

its

| Page 13: Inserted | sue | 1/22/2016 6:14:00 AM |
|---|---|---|

of a

| Page 13: Deleted | sue | 1/22/2016 6:17:00 AM |
|---|---|---|

and

| Page 13: Inserted | sue | 1/22/2016 6:17:00 AM |
|---|---|---|

that produces a larger grain-size

| Page 13: Deleted | sue | 1/22/2016 6:16:00 AM |
|---|---|---|

coarser

| Page 13: Inserted | sue | 1/22/2016 6:14:00 AM |
|---|---|---|

generally requires steeper hillslopes

| Page 13: Inserted | sue | 1/22/2016 6:18:00 AM |
|---|---|---|

| Page 13: Deleted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

However, t

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:17:00 PM |
|---|---|---|

However, t

| Page 13: Inserted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

T

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:17:00 PM |
|---|---|---|

wo issues

| Page 13: Inserted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

may explain this apparent contradiction

| Page 13: Deleted | sue | 1/22/2016 6:15:00 AM |
|---|---|---|

complicate such an interpretation

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:17:00 PM |
|---|---|---|

complicate such an interpretation. First, while the morphology of the Shale Hills catchment bears little resemblance to the underlying structure of steeply-dipping shale beds, the topography of Garner Run is nearly entirely controlled by

| Page 13: Inserted | sue | 1/22/2016 6:18:00 AM |
|---|---|---|

underlying Paleozoic structure (Fig. 4). Specifically,

| Page 13: Deleted | sue | 1/22/2016 6:18:00 AM |
|---|---|---|

Paleozoic structures, with

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:20:00 PM |
|---|---|---|

Paleozoic structures, with hillslope angles reflecting dip-slopes rather than a morphodynamic equilibrium. Second,

| Page 13: Deleted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

ing

| Page 13: Inserted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

as a headwater stream in

| Page 13: Deleted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

its location in the headwaters of

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:21:00 PM |
|---|---|---|

its location in the headwaters of

| Page 13: Inserted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

the

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:21:00 PM |
|---|---|---|

Shavers Creek

| Page 13: Inserted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

watershed, Garner Run i

| Page 13: Deleted | sue | 1/22/2016 6:19:00 AM |
|---|---|---|

has

| Page 13: Inserted | Roman DiBiase | 1/17/2016 1:21:00 PM |
|---|---|---|

has

**Page 13: Inserted** sue 1/22/2016 6:19:00 AM

s

**Page 13: Inserted** Roman DiBiase 1/17/2016 1:21:00 PM

isolated the hillslopes of Garner Run from the regional baselevel controls that influence the Shale Hills and other downstream catchments

**Page 13: Deleted** sue 1/22/2016 6:19:00 AM

the hillslopes of Garner Run

**Page 13: Deleted** sue 1/22/2016 6:20:00 AM

the Shale Hills and other

**Page 13: Inserted** sue 1/22/2016 6:20:00 AM

such as Shale Hills (Fig. 6)

**Page 13: Inserted** Roman DiBiase 1/17/2016 1:21:00 PM

.

**Page 13: Deleted** Roman DiBiase 1/17/2016 1:17:00 PM

This hillslope conundrum could be related to the role of local base level and transient landscape adjustment (Whipple et al., 2013).

**Page 13: Deleted** Roman DiBiase 1/17/2016 1:22:00 PM

of base level fall upstream and downstream of the knickpoint. Equivalently, the knickpoints could delineate different rates

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:23:00 PM

(e.g., Whipple et al., 2013)

**Page 14: Inserted** sue 1/22/2016 6:21:00 AM

(Ma et al., 2013; West et al., 2014; West et al., 2013)

**Page 14: Deleted** sue 1/22/2016 6:21:00 AM

(Ma et al., 2013; West et al., 2014; West et al., 2013)

**Page 14: Inserted** sue 1/22/2016 6:22:00 AM

e

**Page 14: Deleted** sue 1/22/2016 6:22:00 AM

e

**Page 14: Inserted** sue 1/22/2016 6:22:00 AM

**Page 14: Inserted** sue 1/22/2016 6:22:00 AM

**Page 14: Deleted** Roman DiBiase 1/17/2016 1:23:00 PM

. Such models

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:23:00 PM

that

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:23:00 PM

Garner Run and

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:23:00 PM

in general

**Page 14: Inserted** sue 1/22/2016 6:24:00 AM

limit of the

**Page 14: Deleted** sue 1/22/2016 6:24:00 AM

limit

**Page 14: Inserted** sue 1/26/2016 6:51:00 AM

, Fig. 5

**Page 14: Inserted** sue 1/26/2016 6:51:00 AM

abundant

**Page 14: Deleted** sue 1/26/2016 6:51:00 AM

observed to be common

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

, which appear to have accumulated as a large, valley-filling deposit

**Page 14: Deleted** sue 1/22/2016 6:26:00 AM

,

**Page 14: Deleted** sue 1/22/2016 6:26:00 AM

,

**Page 14: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

| Page 15: Inserted | sue | 1/22/2016 6:27:00 AM |
|---|---|---|

Many of these geomorphological features have controlled or been imprinted on CZ processes and human activities in Garner Run. For example, the modern flow pathways for surface and groundwater in Garner Run are significantly influenced by the forcing factors of tectonism, climate, and anthropogenic activity. Flow pathways are influenced i) by topography inherited from geologic events from $10^8$ years before present, ii) by variations in soil grain size as dictated by periglacial processes operating $10^4$ years ago, and iii) by modern land use over the last $10^2 – 10^3$ y.

| Page 15: Formatted | sue | 1/24/2016 7:59:00 PM |
|---|---|---|

Normal

| Page 15: Formatted | sue | 1/22/2016 6:32:00 AM |
|---|---|---|

Superscript

| Page 15: Formatted | sue | 1/22/2016 6:32:00 AM |
|---|---|---|

Superscript

| Page 15: Inserted | Dan K. Arthur | 1/26/2016 10:36:00 AM |
|---|---|---|

In terms of land use, the influence of anthropogenic activity in the catchment is relatively minor and consistent with the surounding region. Neither Shale Hills nor Garner Run subcatchments show signs of having been plowed or farmed in row-crop agriculture, although some grazing may have occurred. The top of one of the ridges in Shale Hills appears to define a field edge. Both subcatchments were forested for at least 100 years. Based on historic aerial photographs, both watersheds contain intact, closed canopy forests in 1938 and show no sign of obvious stand level disturbance since that year.. In the mid 1800s, significant quantities of charcoal were made in this region to run several nearby iron furnaces. Given that charcoal hearths have been identified in the subcatchments from lidar, the subcatchments were probably cleared in the mid to late 1800s as most available wood was used for charcoal making. This land use was also often associated with fires.

| Page 15: Deleted | Dan K. Arthur | 1/26/2016 10:36:00 AM |
|---|---|---|

In terms of land use, the influence of anthropogenic activity in the catchment is relatively minor. Neither Shale Hills nor Garner Run subcatchments show signs of having been plowed or farmed in row-crop agriculture, although some grazing may have occurred. The top of one of the ridges in Shale Hills appears to define a field edge. Both subcatchments were forested for at least 100 years. Based on historic aerial photographs, both watersheds contain intact, closed canopy forests in 1938 and show no sign of obvious disturbance since that year. Both watersheds experienced limited management activities (selected cuts, salvage harvests, etc.) as determined by the presence of stumps. In the mid 1800s, significant quantities of charcoal were made in this region to run several nearby iron furnaces. Given that charcoal hearths have been identified in the subcatchments from lidar, the subcatchments were probably cleared in the mid to late 1800s as most available wood was used for charcoal making. This land use was also often associated with fires.

| **Page 15: Inserted** | **sue** | **1/26/2016 6:53:00 AM** |
|---|---|---|

In terms of land use, the influence of anthropogenic activity in the catchment is relatively minor. Neither Shale Hills nor Garner Run subcatchments show signs of having been plowed or farmed in row-crop agriculture, although some grazing may have occurred. The top of one of the ridges in Shale Hills appears to define a field edge. Both subcatchments were forested for at least 100 years. Based on historic aerial photographs, both watersheds contain intact, closed canopy forests in 1938 and show no sign of obvious disturbance since that year. Both watersheds experienced limited management activities (selected cuts, salvage harvests, etc.) as determined by the presence of stumps. In the mid 1800s, significant quantities of charcoal were made in this region to run several nearby iron furnaces. Given that charcoal hearths have been identified in the subcatchments from lidar, the subcatchments were probably cleared in the mid to late 1800s as most available wood was used for charcoal making. This land use was also often associated with fires.

| **Page 15: Inserted** | **Roman DiBiase** | **1/17/2016 1:24:00 PM** |
|---|---|---|

The above

| **Page 15: Formatted** | **sue** | **1/24/2016 7:59:00 PM** |
|---|---|---|

Normal

| **Page 15: Deleted** | **sue** | **1/22/2016 6:29:00 AM** |
|---|---|---|

e above

| **Page 15: Inserted** | **sue** | **1/22/2016 6:29:00 AM** |
|---|---|---|

is short

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

analysis

**Page 15: Inserted** sue 1/22/2016 6:29:00 AM

of the geomorphology and land use

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

highlights the influence of

**Page 15: Inserted** sue 1/22/2016 6:29:00 AM

the

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

forcing mechanisms

**Page 15: Inserted** sue 1/22/2016 6:29:00 AM

(tectonism, climate, anthropogenic activity) that

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

operat

**Page 15: Inserted** sue 1/22/2016 6:29:00 AM

e

**Page 15: Deleted** sue 1/22/2016 6:29:00 AM

ing

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

ing over a wide range of timescales

**Page 15: Inserted** sue 1/22/2016 6:30:00 AM

and yet

**Page 15: Deleted** sue 1/22/2016 6:30:00 AM

to

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

to influence modern CZ processes,

**Page 15: Deleted** sue 1/22/2016 6:30:00 AM

,

**Page 15: Inserted** sue 1/22/2016 6:30:00 AM

. The CZO efforts document the

**Page 15: Deleted** sue 1/22/2016 6:30:00 AM

and emphasizes the

**Page 15: Inserted** Roman DiBiase 1/17/2016 1:24:00 PM

and emphasizes the importance of providing geologic and geomorphic context for any investigation of the CZ. For example, the modern flow pathways for surface and groundwater in Garner Run are significantly influenced by topography inherited from geologic events from $10^8$ years before present, heterogeneous soil properties (e.g., clay and boulder content) controlled by periglacial processes operating $10^4$ years ago, and modern land use.

**Page 15: Deleted** sue 1/22/2016 6:27:00 AM

any

**Page 15: Deleted** sue 1/22/2016 6:27:00 AM

For example, the modern flow pathways for surface and groundwater in Garner Run are significantly influenced by topography inherited from geologic events from $10^8$ years before present, heterogeneous soil properties (e.g., clay and boulder content) controlled by periglacial processes operating $10^4$ years ago, and modern land use.

Water and energy flux measurements at Garner Run: Tower HOG

**Page 15: Formatted** Roman DiBiase 1/17/2016 1:30:00 PM

Superscript

**Page 15: Formatted** Roman DiBiase 1/17/2016 1:30:00 PM

Superscript

**Page 15: Inserted** sue 1/24/2016 7:58:00 PM

## 0.0 Water and energy flux measurements at Garner Run: Tower HOG

Surface energy balance measurements (eddy covariance (EC) measurements of sensible and latent heat fluxes or upwelling terrestrial radiation/skin temperature) are needed to constrain Flux-PIHM (Shi et al., 2014). Measurements of precipitation, atmospheric state and incoming radiation are needed as inputs to the model. These measurements provide the data needed to simulate the catchment hydrology that is critical to understanding today's WEGSS fluxes. In addition, these fluxes are drivers for millennial-timescale landscape evolution (Fig. 1).

Instrumentation for measurements of water and energy flux measurements are designed as part of the "tower hydrological observation gear" – referred to here as Tower HOG (Table 2, Table S1). While the ideal plan would locate Tower HOG within the Garner run watershed itself, the remote, rocky, heavily wooded terrain makes this too challenging. Therefore, precipitation will be measured near Garner Run on a road crossing Tussey Mountain that is also the site of a pre-existing communications tower (see Fig. 3). A disdrometer (LPM, Theis Clima GmbH) and weighing rain gauge have been in use at Shale Hills since 2009 and 2006 respectively to measure precipitation. To measure precipitation amount at Garner Run, we are installing the simpler instrument (Pluvio$^2$, OTT Hydromet weighing rain gauge). Measurements will be compared to the National Atmospheric Deposition Program measurements and samples of rainwater. According to the nearest NADP site, Garner Run receives 1006 mm/y precipitation with an average pH of 5.0 (Thomas et al., 2013).

EC and radiation instrumentation (Table 2) will also be implemented on the pre-existing communications tower on the Tussey Mountain ridgeline (Fig. 3). Although located out of the subcatchment, the measurement footprint for the tower will be sensitive to fluxes from forests representative of those in Garner Run. The complex terrain at Shale Hills and Garner Run make EC measurements difficult to interpret in stable micrometeorological conditions. Since the primary energy partitioning happens during the day when the atmosphere is typically unstable, daytime sensible and latent heat flux measurements are sufficient to constrain the hydrologic modeling system. Daytime carbon dioxide flux measurements will inform the biogeochemical modeling system.

**Page 16: Deleted** **sue** **1/22/2016 6:55:00 AM**

One of our major focuses is measuring precipitation and evapotranspiration (ET). These fluxes are drivers for landscape evolution as they are manifested today (Fig. 1). These measurements also are needed for the land surface water balance to constrain today's WEGSS fluxes. First, we are installing a laser disdrometerweighing rain gauge (LPM, Theis Clima GmbHPluvio$^2$, OTT Hydromet) to measure precipitation amount and type in Garner Run. Another disdrometer (LPM, Theis Clima GmbH) weighing rain gaugehas have been in use at Shale Hills since 20082009 2006. The disdrometer will be the Jo Hays Vista on Tussey Mountain near Garner Rundeployed as part of the "tower hydrological observation gear" – referred to here as Tower HOG (Table 2)covariance. Tower HOG will be placed outside the watershed on Tussey Mountain ridgeline

(Fig. 3). The remote, rocky terrain in Garner Run made constructing a new tower in the center of the watershed challenging.  In contrast, a communications tower that is surrounded by representative forests already exists on the ridge top above the watershed, and we have therefore chosen this to host the eddy covariance flux instrumentation. Although the measurement footprint (i.e. fetch) for the tower measurements will include other areas, the tower instrumentation will be sensitive to fluxes from the forest in Garner Run. The tower measurements can also be compared to regional measurements such as the National Atmospheric Deposition Program measurements and samples of rainwater. For example, according to the nearest NADP site, Garner Run receives 1006 mm/y precipitation with an average pH of 5.0 (Thomas et al., 2013).

In addition to precipitation, sensible and latent heat fluxes (i.e. using eddy covariance), or skin temperature (upwelling terrestrial radiation) must also be measured to constrain Flux-PIHM (Shi et al., 20132014). A small clearing below the tower site on the Tussey ridgeline makes the site unsuitable for skin temperature measurements representative of the forest, so we are only collecting eddy covariance measurements at the Tussey ridgeline.  Of course, the complex terrain at both Shale Hills and Garner Run make eddy covariance measurements difficult to interpret in stable micrometeorological conditions. Since the primary energy partitioning happens during the day, however, daytime heat flux measurements  are sufficient to constrain the modeling.  For the Garner Run subcatchment, in addition, we also may be able to use upwelling infrared radiation measurements currently being made at the nearby Shale Hills. These radiative energy fluxes are measured using a four component radiometer, i.e., one that measures upwelling and downwelling terrestrial and solar radiation (Table 2). With both the EC measurements at Garner Run and radiative flux measurements at Shale Hills, Flux-PIHM should be well constrained.

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:16:00 AM |
|---|---|---|

**weighing rain gauge**

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:17:00 AM |
|---|---|---|

**Pluvio$^2$, OTT Hydromet**

| Page 16: Formatted | Dan K. Arthur | 1/21/2016 9:18:00 AM |
|---|---|---|

Superscript

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:19:00 AM |
|---|---|---|

**(LPM, Theis Clima GmbH) weighing rain gauge**

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:19:00 AM |
|---|---|---|

**have**

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:20:00 AM |
|---|---|---|

**2009 2006**

| Page 16: Inserted | Dan K. Arthur | 1/21/2016 9:21:00 AM |
|---|---|---|

**the Jo Hays Vista  on Tussey Mountain near Garner Run**

| Page 16: Formatted | sue | 1/22/2016 6:36:00 AM |
|---|---|---|

Not Highlight

| Page 16: Inserted | Yuning Shi | 1/21/2016 11:01:00 AM |
|---|---|---|

**covariance**

| Page 16: Formatted | sue | 1/22/2016 6:59:00 AM |
|---|---|---|

Highlight

| Page 16: Inserted | Yuning Shi | 1/21/2016 11:04:00 AM |
|---|---|---|

**2014**

| Page 16: Inserted | Yuning Shi | 1/21/2016 11:13:00 AM |
|---|---|---|

**heat**

| Page 16: Inserted | Yuning Shi | 1/21/2016 11:06:00 AM |
|---|---|---|

| Page 16: Deleted | sue | 1/17/2016 6:45:00 AM |
|---|---|---|

has important

| Page 16: Deleted | sue | 1/17/2016 6:45:00 AM |
|---|---|---|

on the

| Page 16: Inserted | sue | 1/17/2016 6:44:00 AM |
|---|---|---|

today's

| Page 16: Inserted | sue | 1/17/2016 6:44:00 AM |
|---|---|---|

is known to have

| Page 16: Deleted | sue | 1/17/2016 6:44:00 AM |
|---|---|---|

has important but poorly understood

| Page 16: Inserted | sue | 1/17/2016 6:45:00 AM |
|---|---|---|

influenced

**Page 16: Deleted** **sue** **1/17/2016 6:44:00 AM**

impacts on

**Page 16: Inserted** **sue** **1/17/2016 6:45:00 AM**

over geologic time

**Page 16: Deleted** **sue** **1/26/2016 5:39:00 AM**

individual

**Page 16: Deleted** **sue** **1/23/2016 5:44:00 AM**

WEGSS

**Page 16: Inserted** **sue** **1/26/2016 7:00:00 AM**

vegetation

**Page 16: Deleted** **sue** **1/26/2016 5:39:00 AM**

the fluxes

**Page 16: Inserted** **sue** **1/26/2016 7:00:00 AM**

WEGSS fluxes

**Page 16: Deleted** **sue** **1/23/2016 5:44:00 AM**

entire

**Page 16: Inserted** **sue** **1/17/2016 6:45:00 AM**

. As described below, we once again use the geomorphological framework to design the measurement strategy for vegetation. We also want

**Page 16: Deleted** **sue** **1/17/2016 6:45:00 AM**

: we therefore seek

**Page 16: Deleted** **sue** **1/17/2016 6:46:00 AM**

WEGSS

**Page 16: Inserted** **sue** **1/23/2016 5:44:00 AM**

of nutrients such as nitrate

**Page 16: Deleted** **sue** **1/23/2016 5:44:00 AM**

.

**Page 16: Inserted** **sue** **1/17/2016 6:47:00 AM**

out of Shavers creek. Ultimately, an OSSE will be run to compare measurements to model predictions as a way to determine the important parameters for predicting carbon and nitrogen fluxes. It

| Page 16: Deleted | sue | 1/17/2016 6:47:00 AM |
|---|---|---|

For example, one of our goals is to assess nitrate fluxes out of Shavers creek. To do this, it

| Page 16: Inserted | sue | 1/23/2016 6:06:00 AM |
|---|---|---|

also

| Page 16: Inserted | sue | 1/17/2016 6:47:00 AM |
|---|---|---|

the effect of individual

| Page 16: Inserted | sue | 1/17/2016 6:47:00 AM |
|---|---|---|

on N flux

| Page 16: Deleted | sue | 1/17/2016 6:46:00 AM |
|---|---|---|

Ultimately, we plan to run an OSSE to compare model predictions to measurements as a way to determine the important parameters for predicting carbon and nitrogen fluxes.

| Page 17: Inserted | sue | 1/23/2016 6:07:00 AM |
|---|---|---|

As part of the geomorphological measurement strategy, we

| Page 17: Deleted | sue | 1/23/2016 6:07:00 AM |
|---|---|---|

The vegetation has already been

| Page 17: Inserted | sue | 1/23/2016 6:07:00 AM |
|---|---|---|

the vegetation

| Page 17: Inserted | sue | 1/23/2016 5:45:00 AM |
|---|---|---|

across the Ground HOG catena (ridge top, midslope, and valley floor positions on one side of catchment and one midslope site on the other side, Fig. 5)

| Page 17: Deleted | sue | 1/23/2016 6:08:00 AM |
|---|---|---|

ground-based

| Page 17: Inserted | sue | 1/23/2016 6:08:00 AM |
|---|---|---|

catena-based stratified

| Page 17: Deleted | sue | 1/23/2016 5:46:00 AM |
|---|---|---|

vegetation

| Page 17: Deleted | sue | 1/26/2016 7:01:00 AM |
|---|---|---|

for the subcatchment

**Page 17: Inserted** sue 1/23/2016 6:08:00 AM

, i.e., under the assumption that landscape position was an important control on vegetation

**Page 17: Deleted** sue 1/23/2016 5:45:00 AM

across the catena (ridge top, midslope, and valley floor positions) defined for GroundHOG

**Page 17: Deleted** sue 1/23/2016 5:46:00 AM

later

**Page 17: Inserted** sue 1/23/2016 5:46:00 AM

planned

**Page 17: Deleted** sue 1/26/2016 5:41:00 AM

carbon

**Page 17: Inserted** sue 1/26/2016 5:41:00 AM

C

**Page 17: Deleted** sue 1/26/2016 5:41:00 AM

nitrogen

**Page 17: Inserted** sue 1/26/2016 5:41:00 AM

N

**Page 17: Inserted** sue 1/23/2016 5:46:00 AM

such

**Page 17: Deleted** sue 1/23/2016 5:46:00 AM

such

**Page 17: Deleted** sue 1/26/2016 5:41:00 AM

The v

**Page 17: Inserted** sue 1/26/2016 5:41:00 AM

V

**Page 17: Deleted** sue 1/26/2016 7:02:00 AM

being deployed at

**Page 17: Inserted** sue 1/26/2016 7:02:00 AM

planned for

**Page 17: Inserted** sue 1/23/2016 5:48:00 AM

LRRT, LRMS, LRVF, TMMS,

**Page 17: Deleted** sue 1/26/2016 7:09:00 AM

4

**Page 17: Inserted** sue 1/26/2016 7:09:00 AM

3

**Page 17: Deleted** sue 1/23/2016 5:47:00 AM

at each of the following pits:

**Page 17: Inserted** sue 1/23/2016 5:48:00 AM

at

**Page 17: Deleted** sue 1/23/2016 5:48:00 AM

(LRRT)

**Page 17: Deleted** sue 1/23/2016 5:48:00 AM

(LRMS)

**Page 17: Deleted** sue 1/23/2016 5:48:00 AM

(LRVF)

**Page 17: Inserted** sue 1/26/2016 7:03:00 AM

,

**Page 17: Deleted** sue 1/23/2016 5:48:00 AM

(TMMS)

**Page 17: Inserted** sue 1/23/2016 5:48:00 AM

respectively

**Page 18: Inserted** sue 1/23/2016 5:49:00 AM

transects

**Page 18: Deleted** sue 1/23/2016 5:49:00 AM

y

**Page 18: Inserted** sue 1/23/2016 5:49:00 AM

,

**Page 18: Deleted** sue 1/23/2016 5:56:00 AM

results from these linear

**Page 18: Inserted** sue 1/23/2016 5:56:00 AM

observations

| Page 18: Deleted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

s

| Page 18: Deleted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

across

| Page 18: Inserted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

along the

| Page 18: Deleted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

positions

| Page 18: Deleted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

within a

| Page 18: Inserted | sue | 1/23/2016 5:56:00 AM |
|---|---|---|

at each

| Page 18: Deleted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

tree

| Page 18: Deleted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

the total

| Page 18: Deleted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

; however, BA

| Page 18: Inserted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

with

| Page 18: Inserted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

ing

| Page 18: Deleted | sue | 1/23/2016 5:57:00 AM |
|---|---|---|

ed

| Page 18: Inserted | sue | 1/26/2016 5:43:00 AM |
|---|---|---|

The subcatchment contains a dry oak-heath community type (Fike 1999), primarily consisting of chestnut oak (*Q. montana*), red maple (*Acer rubrum*), black birch (*Betula lenta*), black gum (*Nyssa sylvatica*), and white pine (*Pinus strobus*) in the overstory, with a thick heath understory of mountain laurel (*Kalmia latifolia*), blueberry (*Vaccinium sp.*) and huckleberry (*Gaylussacia sp.*) species, and rhododendron (*Rhododendron maximum*) along Garner Run.

| **Page 18: Deleted** | **sue** | **1/23/2016 5:58:00 AM** |
|---|---|---|

Similarly

| **Page 18: Inserted** | **sue** | **1/26/2016 5:43:00 AM** |
|---|---|---|

T

| **Page 18: Deleted** | **sue** | **1/26/2016 5:43:00 AM** |
|---|---|---|

,

| **Page 18: Inserted** | **sue** | **1/23/2016 6:00:00 AM** |
|---|---|---|

he transect work also highlighted a type of measurement that we had not needed for Shale Hills but which our models and observations are showing is important in the new subcatchment: the fraction of land surface covered by boulders.

| **Page 18: Inserted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

At LRRT,

| **Page 18: Deleted** | **sue** | **1/23/2016 6:01:00 AM** |
|---|---|---|

in LRRT

| **Page 18: Inserted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

. Furthermore,

| **Page 18: Deleted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

, yet

| **Page 18: Inserted** | **sue** | **1/23/2016 5:59:00 AM** |
|---|---|---|

rock coverage

| **Page 18: Deleted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

certain

| **Page 18: Inserted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

some transect

| **Page 18: Deleted** | **sue** | **1/23/2016 6:02:00 AM** |
|---|---|---|

along the transect rock cover

| **Page 18: Deleted** | **sue** | **1/23/2016 6:03:00 AM** |
|---|---|---|

measurements will be combined with data on

| Page 18: Inserted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

and

| Page 18: Inserted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

data

| Page 18: Deleted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

(

| Page 18: Deleted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

)

| Page 18: Inserted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

will be combined with

| Page 18: Deleted | sue | 1/23/2016 6:03:00 AM |
|---|---|---|

and

| Page 18: Deleted | sue | 1/23/2016 6:04:00 AM |
|---|---|---|

resampling

| Page 18: Inserted | sue | 1/23/2016 6:04:00 AM |
|---|---|---|

re-measurements along

| Page 18: Deleted | sue | 1/23/2016 6:04:00 AM |
|---|---|---|

of linear

| Page 19: Deleted | sue | 1/26/2016 7:05:00 AM |
|---|---|---|

3

| Page 19: Inserted | sue | 1/26/2016 7:05:00 AM |
|---|---|---|

4

| Page 19: Inserted | sue | 1/26/2016 7:06:00 AM |
|---|---|---|

and Table S2

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

a combination of

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

accurately

| Page 19: Inserted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

allow estimation of

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

inform

| Page 19: Inserted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

. Depth of tree water use

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

which is

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

.

| Page 19: Inserted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

, is c

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

C

| Page 19: Inserted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

derived from

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

,

| Page 19: Deleted | sue | 1/23/2016 6:36:00 AM |
|---|---|---|

is used

| Page 19: Deleted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

in the PIHM suite of models

| Page 19: Inserted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

We will explore whether the use of

| Page 19: Deleted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

Using

| Page 19: Inserted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

-

| Page 19: Deleted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

may

| Page 19: Inserted | sue | 1/23/2016 6:37:00 AM |
|---|---|---|

s

**Page 19: Deleted** sue 1/23/2016 6:37:00 AM

are also being used to assess

**Page 19: Inserted** sue 1/23/2016 6:37:00 AM

will collect

**Page 20: Inserted** sue 1/23/2016 6:38:00 AM

for assessment

**Page 20: Deleted** sue 1/16/2016 7:03:00 AM

To first order, the

**Page 20: Inserted** sue 1/23/2016 6:40:00 AM

The uplands of the

**Page 20: Inserted** sue 1/23/2016 6:39:00 AM

The coarse blocks of the Tuscarora sandstone range in diameter from ~10-200 cm, making it challenging to excavate large soil pits (Table 3).

**Page 20: Inserted** sue 1/23/2016 6:41:00 AM

therefore

**Page 20: Inserted** sue 1/23/2016 6:39:00 AM

Three pits were dug by hand until deepening was impossible (LRRT, LRMS, and TMMS). The Leading Ridge Valley Floor (LRVF) pit was dug by hand and then deepened using a jackhammer until the inferred contact with intact bedrock was reached. The pits were excavated in the following soil series: TMMS, LRRT and LRMS (Hazleton-Dekalb association, very steep), and LRVF (Andover extremely stony loam, 0-8% slopes).

**Page 20: Inserted** sue 1/26/2016 7:16:00 AM

, S2

**Page 20: Inserted** sue 1/23/2016 6:41:00 AM

This design was informed by observations at Shale Hills and the new subcatchment and by modelling conceptualizations. As discussed earlier, the Shale Hills subcatchment upland land surface falls into one of two categories: hillslopes or swales. In contrast, we observed little evidence for swales in Garner Run.

**Page 20: Deleted** sue 1/23/2016 6:39:00 AM

In addition, the surface cover at Garner Run consists of coarse blocks of the Tuscarora sandstone ranging in diameter from ~10-200 cm, making it challenging to excavate large soil pits, limiting the number of such installations (Table 3). Three pits were dug entirely by hand (LRRT, LRMS, and TMMS). The Leading Ridge Valley Floor (LRVF) pit was dug by hand and was deepened using a jackhammer until the inferred contact with intact bedrock was reached.

| Page 20: Deleted | sue | 1/16/2016 7:04:00 AM |
|---|---|---|

s location

| Page 20: Inserted | sue | 1/23/2016 6:42:00 AM |
|---|---|---|

in the new subcatchment

| Page 20: Deleted | sue | 1/23/2016 6:42:00 AM |
|---|---|---|

selected

| Page 20: Inserted | sue | 1/23/2016 6:42:00 AM |
|---|---|---|

therefore located

| Page 20: Deleted | sue | 1/26/2016 7:12:00 AM |
|---|---|---|

slopes that were

| Page 20: Inserted | sue | 1/26/2016 7:12:00 AM |
|---|---|---|

roughly

| Page 20: Inserted | sue | 1/26/2016 7:12:00 AM |
|---|---|---|

or

| Page 20: Deleted | sue | 1/26/2016 7:12:00 AM |
|---|---|---|

in planview to avoid areas of convergent flow. The midslope pits were located on

| Page 20: Inserted | sue | 1/26/2016 7:12:00 AM |
|---|---|---|

somewhat

| Page 20: Deleted | sue | 1/16/2016 7:04:00 AM |
|---|---|---|

for reasons discussed

| Page 20: Inserted | sue | 1/16/2016 7:04:00 AM |
|---|---|---|

(see

| Page 20: Inserted | sue | 1/16/2016 7:04:00 AM |
|---|---|---|

)

| Page 20: Inserted | sue | 1/23/2016 6:43:00 AM |
|---|---|---|

| Page 20: Deleted | sue | 1/16/2016 7:04:00 AM |
|---|---|---|

Given our catena design, we excavated pits in the following soil series: TMMS, LRRT and LRMS (Hazleton-Dekalb association, very steep), and LRVF (Andover extremely stony loam, 0-8% slopes).

| Page 20: Deleted | sue | 1/23/2016 6:44:00 AM |
|---|---|---|

n Ground HOG are

| Page 20: Inserted | sue | 1/23/2016 6:44:00 AM |
|---|---|---|

s

| Page 20: Deleted | sue | 1/23/2016 6:46:00 AM |
|---|---|---|

water and earth materials is largely 1D: i.e., net

| Page 20: Deleted | sue | 1/23/2016 6:46:00 AM |
|---|---|---|

flux

| Page 20: Inserted | sue | 1/23/2016 6:46:00 AM |
|---|---|---|

largely

| Page 20: Inserted | sue | 1/23/2016 6:44:00 AM |
|---|---|---|

We are now developing

| Page 20: Deleted | sue | 1/23/2016 6:44:00 AM |
|---|---|---|

is a model under development to

| Page 20: Inserted | sue | 1/23/2016 6:44:00 AM |
|---|---|---|

to

| Page 20: Inserted | sue | 1/16/2016 7:05:00 AM |
|---|---|---|

The next level of complexity is a convex-upward but otherwise planar hillslope. The intent for

| Page 21: Deleted | sue | 1/16/2016 7:06:00 AM |
|---|---|---|

Second,

| Page 21: Deleted | sue | 1/16/2016 7:06:00 AM |
|---|---|---|

will

| Page 21: Inserted | sue | 1/16/2016 7:06:00 AM |
|---|---|---|

is that it will

| Page 21: Deleted | sue | 1/16/2016 7:06:00 AM |
|---|---|---|

also

| Page 21: Deleted | sue | 1/16/2016 7:07:00 AM |
|---|---|---|

convex-upward hillslopes by assessment of the

| Page 21: Inserted | sue | 1/16/2016 7:07:00 AM |
|---|---|---|

s

| Page 21: Deleted | sue | 1/16/2016 7:07:00 AM |
|---|---|---|

a

| Page 21: Inserted | sue | 1/16/2016 7:07:00 AM |
|---|---|---|

s

| Page 21: Deleted | sue | 1/23/2016 6:47:00 AM |
|---|---|---|

that incorporates downslope transport of water and soil

| Page 21: Deleted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

By analyzing soil

| Page 21: Inserted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

Soil

| Page 21: Inserted | sue | 1/26/2016 7:13:00 AM |
|---|---|---|

convex-upward but otherwise planar hillslope

| Page 21: Deleted | sue | 1/26/2016 7:13:00 AM |
|---|---|---|

planar hillslope

| Page 21: Deleted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

as we did

| Page 21: Inserted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

such as those described

| Page 21: Deleted | Dan K. Arthur | 1/25/2016 9:41:00 AM |
|---|---|---|

b

| Page 21: Inserted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

can be used to parameterize

| Page 21: Deleted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

,

| Page 21: Deleted | sue | 1/23/2016 6:45:00 AM |
|---|---|---|

will be enabled

| Page 21: Deleted | sue | 1/23/2016 6:48:00 AM |
|---|---|---|

With such conceptual and numerical models, we will extrapolate to other hillslopes within Shavers creek watershed.

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

at Shale Hills we discovered that

| **Page 21: Inserted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

while

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

were

| **Page 21: Inserted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

are

| **Page 21: Inserted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

at Shale Hills

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

, requiring measurements at both

| **Page 21: Inserted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

the lack of

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

. No such

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

have been observed

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

,

| **Page 21: Deleted** | **sue** | **1/23/2016 6:48:00 AM** |
|---|---|---|

ing

| **Page 21: Inserted** | **sue** | **1/15/2016 8:01:00 PM** |
|---|---|---|

(In fact the lack of swales in the sandstone catchment is one of the observations that we hope we can eventually explain).

| **Page 21: Deleted** | **sue** | **1/26/2016 7:14:00 AM** |
|---|---|---|

at Shale Hills

| **Page 21: Inserted** | **sue** | **1/26/2016 7:15:00 AM** |
|---|---|---|

o

**Page 21: Deleted** sue 1/26/2016 7:14:00 AM

(Graham and Lin, 2011; Graham and Lin, 2010 ; Ma et al., 2011; West et al., 2014) o

**Page 21: Inserted** sue 1/26/2016 7:14:00 AM

at Shale Hills (Graham and Lin, 2011; Graham and Lin, 2010 ; Ma et al., 2011; West et al., 2014)

**Page 21: Inserted** sue 1/26/2016 7:14:00 AM

in Pennsylvania o

**Page 21: Deleted** sue 1/26/2016 7:14:00 AM

in Pennsylvania

**Page 21: Inserted** sue 1/23/2016 6:49:00 AM

Ground HOG includes

**Page 21: Deleted** sue 1/23/2016 6:49:00 AM

additional

**Page 21: Deleted** sue 1/23/2016 6:49:00 AM

was sited

**Page 21: Deleted** sue 1/23/2016 6:49:00 AM

to make observations to constrain the effect of aspect

**Page 21: Inserted** sue 1/23/2016 6:49:00 AM

We will use numerical models to explore regolith formation and to extrapolate to other hillslopes within Shavers creek watershed.

**Page 21: Deleted** sue 1/23/2016 6:50:00 AM

**Page 21: Inserted** sue 1/23/2016 6:50:00 AM

This highlights the importance of understanding the soil to the CZ effort. Soil provides a record of both transport of rock-derived material as well as fluxes of water over the period of pedogenesis. For example

**Page 21: Deleted** sue 1/16/2016 7:10:00 AM

At each pit location

| Page 21: Deleted | sue | 1/23/2016 6:51:00 AM |
|---|---|---|

we described the soil profile

| Page 21: Inserted | sue | 1/23/2016 6:52:00 AM |
|---|---|---|

the pits at

| Page 21: Deleted | sue | 1/16/2016 7:10:00 AM |
|---|---|---|

,

| Page 21: Inserted | sue | 1/23/2016 6:51:00 AM |
|---|---|---|

Garner Run

| Page 21: Deleted | sue | 1/16/2016 7:10:00 AM |
|---|---|---|

which typically had the following structure:

| Page 21: Inserted | sue | 1/23/2016 6:51:00 AM |
|---|---|---|

are characterized from land surface downward by

| Page 21: Inserted | sue | 1/23/2016 7:33:00 AM |
|---|---|---|

a thin organic layer,

| Page 21: Deleted | sue | 1/23/2016 7:33:00 AM |
|---|---|---|

n upper

| Page 21: Deleted | sue | 1/23/2016 7:33:00 AM |
|---|---|---|

with a thin organic soil

| Page 21: Inserted | sue | 1/23/2016 7:33:00 AM |
|---|---|---|

characterized by sand-sized grains

| Page 21: Inserted | sue | 1/23/2016 6:52:00 AM |
|---|---|---|

few

| Page 21: Deleted | sue | 1/23/2016 6:52:00 AM |
|---|---|---|

mostly absent

| Page 21: Inserted | sue | 1/23/2016 6:53:00 AM |
|---|---|---|

interspersed

| Page 21: Deleted | sue | 1/23/2016 6:53:00 AM |
|---|---|---|

interspersed

| Page 21: Inserted | sue | 1/23/2016 6:53:00 AM |
|---|---|---|

Depth intervals of the soil

| Page 21: Deleted | sue | 1/23/2016 6:53:00 AM |
|---|---|---|

Additionally, for each pit we sampled soils at

| Page 21: Inserted | sue | 1/23/2016 6:53:00 AM |
|---|---|---|

every

| Page 21: Inserted | sue | 1/26/2016 7:16:00 AM |
|---|---|---|

and from basal rocks

| Page 21: Deleted | sue | 1/23/2016 6:54:00 AM |
|---|---|---|

intervals for

| Page 21: Inserted | sue | 1/23/2016 6:54:00 AM |
|---|---|---|

show variations in

| Page 21: Inserted | sue | 1/23/2016 6:54:00 AM |
|---|---|---|

(Supplement Tables S3, S4)

| Page 21: Inserted | sue | 1/25/2016 5:07:00 AM |
|---|---|---|

and are being analyzed for

| Page 21: Deleted | sue | 1/25/2016 5:07:00 AM |
|---|---|---|

composition analysis

| Page 21: Inserted | sue | 1/25/2016 5:07:00 AM |
|---|---|---|

mineralogy

| Page 21: Deleted | sue | 1/23/2016 6:54:00 AM |
|---|---|---|

(Supplement Table S4)

| Page 21: Inserted | sue | 1/23/2016 6:55:00 AM |
|---|---|---|

These soil observations yield further clues to the history of the landscape.

| Page 21: Deleted | sue | 1/25/2016 5:08:00 AM |
|---|---|---|

Most of t

| Page 21: Inserted | sue | 1/25/2016 5:08:00 AM |
|---|---|---|

T

| Page 21: Inserted | sue | 1/16/2016 7:11:00 AM |
|---|---|---|

Lower Silurian

| Page 21: Deleted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

This sandstone, deposited in the Lower Silurian, has been i

| Page 21: Inserted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

I

| Page 21: Deleted | sue | 1/16/2016 7:11:00 AM |
|---|---|---|

during original deposition

| Page 21: Inserted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

, this sandstone has been

| Page 21: Deleted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

.

| Page 21: Deleted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

The unit has been

| Page 21: Deleted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

| Page 21: Inserted | sue | 1/24/2016 7:23:00 AM |
|---|---|---|

to a highly indurated quartzite

| Page 21: Deleted | sue | 1/16/2016 7:12:00 AM |
|---|---|---|

so that pressure solution has cemented the fabric of the rock: as such, the unit is often referred to as a quartzite

| Page 21: Deleted | sue | 1/16/2016 7:14:00 AM |
|---|---|---|

Cotter reported the unit to be close to 98% $SiO_2$. Weathering of sandstone is largely controlled by the porosity, the fraction of non-quartz grains, the composition of the cement (Turkington and Paradise, 2005), and the pH of soil porewaters (Certini et al., 2003). The porosity is important because it dictates how much water enters the weathering rock; in addition, during seasonal drying, salts deposited inside a sandstone can crystallize and disintegrate the rock (Labus and Bochen, 2012). Thermal cycling can also crack sandstones (Turkington and Paradise, 2005) as can tree roots (Amundson, 2004).

The average of the b

| Page 21: Inserted | sue | 1/24/2016 7:24:00 AM |
|---|---|---|

B

| Page 21: Deleted | sue | 1/23/2016 7:00:00 AM |
|---|---|---|

four

| Page 21: Inserted | sue | 1/23/2016 7:00:00 AM |
|---|---|---|

five

| Page 21: Deleted | sue | 1/24/2016 7:25:00 AM |
|---|---|---|

| Page 21: Deleted | sue | 1/24/2016 7:25:00 AM |
|---|---|---|

amples

| Page 21: Deleted | sue | 1/23/2016 7:07:00 AM |
|---|---|---|

s

| Page 21: Inserted | sue | 1/25/2016 5:08:00 AM |
|---|---|---|

the five

| Page 21: Deleted | sue | 1/24/2016 6:55:00 AM |
|---|---|---|

the

| Page 21: Inserted | sue | 1/23/2016 7:00:00 AM |
|---|---|---|

| Page 21: Deleted | sue | 1/23/2016 7:22:00 AM |
|---|---|---|

soil

| Page 21: Deleted | sue | 1/16/2016 7:13:00 AM |
|---|---|---|

| Page 21: Deleted | sue | 1/23/2016 7:23:00 AM |
|---|---|---|

used

| Page 21: Inserted | sue | 1/23/2016 7:23:00 AM |
|---|---|---|

averaged

| Page 21: Deleted | sue | 1/23/2016 7:23:00 AM |
|---|---|---|

an average

| Page 21: Deleted | sue | 1/24/2016 7:25:00 AM |
|---|---|---|

quartzite for comparison to similar analyses of bulk regolith

| Page 21: Inserted | sue | 1/23/2016 7:08:00 AM |
|---|---|---|

protolith

| Page 21: Deleted | sue | 1/16/2016 7:13:00 AM |
|---|---|---|

samples

**Page 21: Deleted** sue 1/23/2016 7:11:00 AM

all measured using Li metaborate fusion followed by analysis by inductively coupled plasma atomic emission spectroscopy,

**Page 21: Deleted** sue 1/23/2016 7:11:00 AM

In Garner run samples, the

**Page 21: Inserted** sue 1/23/2016 7:11:00 AM

These

**Page 21: Deleted** sue 1/24/2016 7:25:00 AM

Tuscarora

**Page 21: Inserted** sue 1/23/2016 7:11:00 AM

samples

**Page 21: Deleted** sue 1/23/2016 7:11:00 AM

was observed to be close

**Page 21: Inserted** sue 1/23/2016 7:11:00 AM

contain

**Page 21: Deleted** sue 1/23/2016 7:11:00 AM

to

**Page 21: Inserted** sue 1/23/2016 7:23:00 AM

>

**Page 21: Deleted** sue 1/23/2016 7:23:00 AM

98

**Page 21: Inserted** sue 1/23/2016 7:23:00 AM

96 wt.

**Page 21: Deleted** sue 1/16/2016 7:14:00 AM

.

**Page 21: Inserted** sue 1/16/2016 7:14:00 AM

, i.e., very similar to published Tuscarora compositions (Cotter, 1982).

**Page 21: Deleted** sue 1/16/2016 7:14:00 AM

A small amount of

**Page 21: Inserted** sue 1/16/2016 7:14:00 AM

Minor

**Page 21: Deleted** sue 1/16/2016 7:14:00 AM

observed to be

**Page 21: Inserted** sue 1/24/2016 7:26:00 AM

in the parent

**Page 21: Deleted** sue 1/24/2016 7:26:00 AM

Supplement

**Page 21: Inserted** sue 1/24/2016 7:26:00 AM

and at even higher concentration

**Page 21: Deleted** sue 1/24/2016 7:26:00 AM

.

**Page 21: Inserted** sue 1/24/2016 6:58:00 AM

in soils (Table S4). This enrichment in soil could be due to several processes during weathering: for example, retention of Ti from the protolith, losses of elements other than Ti, or addition of Ti to the soil. If Ti in the soil was derived from protolith,

**Page 22: Deleted** sue 1/16/2016 7:15:00 AM

By calculating the normalized concentrations for elements assuming Ti is insoluble, we assessed the

**Page 22: Inserted** sue 1/24/2016 7:27:00 AM

other

**Page 22: Deleted** sue 1/24/2016 7:00:00 AM

from the regolith as compared to Ti

**Page 22: Deleted** sue 1/16/2016 7:15:00 AM

underlying Tuscarora

**Page 22: Inserted** sue 1/16/2016 7:15:00 AM

can be calculated from the

**Page 22: Deleted** sue 1/16/2016 7:15:00 AM

. These normalized concentrations are referred to as

**Page 22: Deleted** sue 1/24/2016 7:00:00 AM

s

**Page 22: Deleted** sue 1/24/2016 7:27:00 AM

the immobile element

**Page 22: Inserted** sue 1/16/2016 7:15:00 AM

Ti

**Page 22: Deleted** sue 1/16/2016 7:16:00 AM

the

**Page 22: Inserted** sue 1/16/2016 7:16:00 AM

an

**Page 22: Deleted** sue 1/16/2016 7:16:00 AM

mobile

**Page 22: Inserted** sue 1/16/2016 7:16:00 AM

that was lost or gained

**Page 22: Deleted** sue 1/23/2016 7:24:00 AM

From this assessment of regolith mass balance, it was observed that

**Page 22: Inserted** sue 1/23/2016 7:25:00 AM

Assuming Ti in soil was derived from the protolith, $\tau_{Ti,j}$ values = 0 within error for

**Page 22: Deleted** sue 1/23/2016 7:25:00 AM

Ca

**Page 22: Inserted** sue 1/23/2016 7:25:00 AM

Mg

**Page 22: Deleted** sue 1/23/2016 7:25:00 AM

Na, Si,

**Page 22: Deleted** sue 1/23/2016 7:25:00 AM

P

**Page 22: Inserted** sue 1/23/2016 7:25:00 AM

Fe, indicating they

**Page 22: Deleted** sue 1/23/2016 7:26:00 AM

either largely unchanged

**Page 22: Inserted** sue 1/24/2016 7:02:00 AM

neither added nor depleted

**Page 22: Deleted** sue 1/24/2016 7:28:00 AM

$(\tau \approx 0)$

**Page 22: Deleted** sue 1/23/2016 7:26:00 AM

or highly depleted $(\tau < 0)$

**Page 22: Inserted** sue 1/16/2016 7:16:00 AM

Ti.

**Page 22: Deleted** sue 1/24/2016 7:02:00 AM

the underlying rock.

**Page 22: Inserted** sue 1/24/2016 7:29:00 AM

$\tau_{Ti,K}$

**Page 22: Deleted** sue 1/23/2016 7:26:00 AM

Mg, K, and Fe were

**Page 22: Inserted** sue 1/23/2016 7:26:00 AM

**Page 22: Deleted** sue 1/23/2016 7:26:00 AM

all significantly enriched in the soils ($\tau$

**Page 22: Deleted** sue 1/24/2016 7:02:00 AM

)

**Page 22: Inserted** sue 1/24/2016 7:02:00 AM

, consistent with addition of K to the soil

**Page 22: Deleted** sue 1/16/2016 7:17:00 AM

compared to the protolith

**Page 22: Deleted** sue 1/23/2016 7:27:00 AM

Supplement

**Page 22: Deleted** sue 1/23/2016 7:27:00 AM

S3

**Page 22: Inserted** sue 1/23/2016 7:27:00 AM

8

**Page 22: Inserted** sue 1/25/2016 5:09:00 AM

Error bars on many of the elements are very large because of the variability in the low concentrations of all elements except Si and O.

| Page 22: Deleted | sue | 1/23/2016 7:27:00 AM |
|---|---|---|

On each plot, the star represents the parent composition ($\tau \approx 0$), plotted at an arbitrary depth.

| Page 22: Deleted | sue | 1/23/2016 7:28:00 AM |
|---|---|---|

These observations are c

| Page 22: Inserted | sue | 1/24/2016 7:02:00 AM |
|---|---|---|

According to

| Page 22: Deleted | sue | 1/24/2016 7:02:00 AM |
|---|---|---|

onsistent with

| Page 22: Inserted | sue | 1/16/2016 7:17:00 AM |
|---|---|---|

published

| Page 22: Deleted | sue | 1/16/2016 7:17:00 AM |
|---|---|---|

in the literature that

| Page 22: Inserted | sue | 1/16/2016 7:17:00 AM |
|---|---|---|

for

| Page 22: Deleted | sue | 1/16/2016 7:17:00 AM |
|---|---|---|

ridgetop soils

| Page 22: Inserted | sue | 1/16/2016 7:17:00 AM |
|---|---|---|

this formation in this region, the thin and poorly developed ridgetop soil

| Page 22: Deleted | sue | 1/24/2016 7:29:00 AM |
|---|---|---|

are

| Page 22: Inserted | sue | 1/24/2016 7:29:00 AM |
|---|---|---|

is likely

| Page 22: Deleted | sue | 1/23/2016 7:29:00 AM |
|---|---|---|

, poorly developed, and thin

| Page 22: Deleted | sue | 1/16/2016 7:18:00 AM |
|---|---|---|

downslope

| Page 22: Inserted | sue | 1/16/2016 7:18:00 AM |
|---|---|---|

on hillslopes likely

| Page 22: Deleted | sue | 1/16/2016 7:18:00 AM |
|---|---|---|

generally

| Page 22: Inserted | sue | 1/25/2016 5:10:00 AM |
|---|---|---|

previous researchers have pointed out that

| Page 22: Inserted | sue | 1/24/2016 7:10:00 AM |
|---|---|---|

central

| Page 22: Deleted | sue | 1/24/2016 7:35:00 AM |
|---|---|---|

color

| Page 22: Deleted | sue | 1/25/2016 5:10:00 AM |
|---|---|---|

has been attributed to

| Page 22: Inserted | sue | 1/25/2016 5:10:00 AM |
|---|---|---|

may indicate two generations of

| Page 22: Deleted | sue | 1/16/2016 7:19:00 AM |
|---|---|---|

exposure of earlier regolith to

| Page 22: Inserted | sue | 1/16/2016 7:19:00 AM |
|---|---|---|

, i.e., a previously weathered

| Page 22: Deleted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

(producing the

| Page 22: Deleted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

)

| Page 22: Deleted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

followed by emplacement

| Page 22: Inserted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

which was then covered by a

| Page 22: Deleted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

of

| Page 22: Inserted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

al layer

| Page 22: Deleted | sue | 1/16/2016 7:20:00 AM |
|---|---|---|

um

**Page 22: Deleted** sue 1/24/2016 7:30:00 AM

Such polygenetic histories will make regolith formation modelling more complex.

**Page 22: Inserted** sue 1/24/2016 7:11:00 AM

Although the soils here did not show a strong brown over red color signature (Supplement Fig. S3), clay-rich soil at depth may document soil formation before the LGM (Table S2).

**Page 22: Deleted** sue 1/24/2016 7:12:00 AM

Mg,

**Page 22: Deleted** sue 1/24/2016 7:12:00 AM

, and Fe

**Page 22: Inserted** sue 1/25/2016 5:11:00 AM

in the residual soils

**Page 22: Deleted** sue 1/25/2016 5:11:00 AM

where downslope transport is unlikely to have been significant

**Page 22: Inserted** sue 1/24/2016 7:36:00 AM

, is another complexity. K could have been added as exogenous dust inputs which were very important during and immediately after glacial periods (Ciolkosz et al., 1990). Alternately, K-containing clay particles could have percolated downward from weathering of the overlying units such as the Rose Hill shale before it was eroded away (Fig. 4). Such movement of fines downward from the Rose Hill have been observed at Shale Hills (Jin et al., 2010a): such particles could have been added to the underlying Tuscarora and then retained in the soil. In that case, the assumed protolith composition could be erroreous, especially if Ti was added from the downward infiltrating fines. K enrichment could also be explained by

**Page 22: Deleted** Dan K. Arthur 1/25/2016 9:41:00 AM

a

**Page 22: Deleted** sue 1/16/2016 7:21:00 AM

,

**Page 22: Deleted** sue 1/16/2016 7:21:00 AM

could either be explained by exogenous additions to the soil or by protolith compositional variation which was not assessed in the small set of 5 rock samples. For example, some interfingered

| Page 22: Deleted | sue | 1/24/2016 7:14:00 AM |
|---|---|---|

are known to occur

| Page 22: Inserted | sue | 1/25/2016 5:16:00 AM |
|---|---|---|

itself

| Page 22: Inserted | sue | 1/24/2016 7:21:00 AM |
|---|---|---|

. If these interfingered shales were the protolith of the observed soils

| Page 22: Deleted | sue | 1/24/2016 7:21:00 AM |
|---|---|---|

and could have provided the excess Mg, K

| Page 22: Inserted | sue | 1/24/2016 7:13:00 AM |
|---|---|---|

, this would mean that our estimated protolith composition was K-deficient

| Page 22: Deleted | sue | 1/24/2016 7:13:00 AM |
|---|---|---|

, and Fe

| Page 22: Deleted | sue | 1/24/2016 7:20:00 AM |
|---|---|---|

Alternately, addition of these elements could have been caused by i) dust inputs (Ciolkosz et al., 1990) which were likely to be important especially during the glacial period and just after, or ii) fines percolating downward from weathering of the overlying Rose Hill shale before it was eroded away (Fig. 4).

| Page 22: Inserted | sue | 1/24/2016 7:42:00 AM |
|---|---|---|

Thus, soil analysis (Fig. 8) leads to interesting hypotheses that will be investigated.

| Page 22: Deleted | sue | 1/16/2016 7:23:00 AM |
|---|---|---|

Movement of fines out of the Rose Hill shale is known to be happening today from our work at Shale Hills (Jin et al., 2010a).

| Page 23: Deleted | sue | 1/26/2016 7:22:00 AM |
|---|---|---|

taken by

| Page 23: Inserted | sue | 1/26/2016 7:22:00 AM |
|---|---|---|

at

| Page 23: Deleted | sue | 1/26/2016 7:21:00 AM |
|---|---|---|

instrumentation

| Page 23: Deleted | sue | 1/26/2016 7:22:00 AM |
|---|---|---|

the

| Page 23: Inserted | sue | 1/26/2016 7:22:00 AM |
|---|---|---|

Ground HOG

| Page 23: Deleted | sue | 1/25/2016 5:18:00 AM |
|---|---|---|

, which meant

| Page 23: Inserted | sue | 1/25/2016 5:18:00 AM |
|---|---|---|

. Thus,

| Page 23: Inserted | sue | 1/26/2016 7:23:00 AM |
|---|---|---|

(Fig. S2)

| Page 24: Inserted | sue | 1/25/2016 5:20:00 AM |
|---|---|---|

n

| Page 24: Deleted | sue | 1/25/2016 5:20:00 AM |
|---|---|---|

soil gas

| Page 24: Inserted | Dan K. Arthur | 1/25/2016 9:44:00 AM |
|---|---|---|

;

| Page 24: Deleted | Dan K. Arthur | 1/25/2016 9:43:00 AM |
|---|---|---|

)

| Page 24: Deleted | sue | 1/25/2016 5:22:00 AM |
|---|---|---|

can

| Page 24: Inserted | sue | 1/25/2016 5:22:00 AM |
|---|---|---|

s

| Page 25: Deleted | sue | 1/26/2016 7:25:00 AM |
|---|---|---|

valuable

| Page 25: Deleted | sue | 1/25/2016 5:23:00 AM |
|---|---|---|

will

| Page 25: Inserted | sue | 1/25/2016 5:23:00 AM |
|---|---|---|

can also

| Page 25: Deleted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

in the previous studies

| Page 25: Inserted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

between and compare with

| Page 25: Deleted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

between

| Page 25: Inserted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

For example, d

| Page 25: Deleted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

D

| Page 25: Deleted | sue | 1/25/2016 5:24:00 AM |
|---|---|---|

also

| Page 25: Inserted | sue | 1/25/2016 5:25:00 AM |
|---|---|---|

(Fig. 5)

| Page 25: Deleted | Dan K. Arthur | 1/25/2016 9:44:00 AM |
|---|---|---|

, in press

| Page 25: Deleted | sue | 1/25/2016 5:25:00 AM |
|---|---|---|

An example of

| Page 25: Inserted | sue | 1/25/2016 5:25:00 AM |
|---|---|---|

To exemplify

| Page 25: Inserted | sue | 1/25/2016 5:25:00 AM |
|---|---|---|

, results

| Page 25: Deleted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

is shown here

| Page 25: Inserted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

| Page 25: Deleted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

. The unit was

| Page 25: Inserted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

is shown in

| Page 25: Deleted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

(

| Page 25: Deleted | sue | 1/26/2016 7:27:00 AM |
|---|---|---|

5

| Page 25: Inserted | sue | 1/26/2016 7:27:00 AM |
|---|---|---|

9

| Page 25: Deleted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

)

| Page 26: Inserted | sue | 1/25/2016 5:26:00 AM |
|---|---|---|

errors

| Page 26: Deleted | sue | 1/26/2016 7:27:00 AM |
|---|---|---|

line lengths recorded by the survey wheel which were slightly different than the actual lengths

| Page 26: Deleted | sue | 1/25/2016 5:27:00 AM |
|---|---|---|

In order t

| Page 26: Inserted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

R

| Page 26: Deleted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

o surface normalize the radar records collected, r

| Page 26: Inserted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

as described below

| Page 26: Deleted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

at major slope breaks

| Page 26: Deleted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

with an engineering level and stadia rod

| Page 26: Inserted | sue | 1/25/2016 5:30:00 AM |
|---|---|---|

to surface normalize the data

| Page 26: Inserted | sue | 1/25/2016 5:28:00 AM |
|---|---|---|

from near Garner Run to the summit

| Page 26: Deleted | sue | 1/25/2016 5:27:00 AM |
|---|---|---|

essentially

| Page 26: Inserted | sue | 1/25/2016 5:27:00 AM |
|---|---|---|

nominally

| Page 26: Inserted | sue | 1/25/2016 5:27:00 AM |
|---|---|---|

~494 to ~588 masl

| Page 26: Deleted | sue | 1/25/2016 5:28:00 AM |
|---|---|---|

near Garner Run to the summit, running

| Page 26: Deleted | sue | 1/25/2016 5:27:00 AM |
|---|---|---|

from about 494 to 588 masl

| Page 26: Deleted | sue | 1/25/2016 5:35:00 AM |
|---|---|---|

5

| Page 26: Inserted | sue | 1/25/2016 5:35:00 AM |
|---|---|---|

9

| Page 26: Deleted | sue | 1/26/2016 7:27:00 AM |
|---|---|---|

:

| Page 26: Deleted | sue | 1/25/2016 5:29:00 AM |
|---|---|---|

. These obstacles

| Page 26: Inserted | sue | 1/25/2016 5:29:00 AM |
|---|---|---|

that

| Page 26: Deleted | sue | 1/25/2016 5:29:00 AM |
|---|---|---|

In this study, f

| Page 26: Inserted | sue | 1/25/2016 5:29:00 AM |
|---|---|---|

F

| Page 26: Deleted | sue | 1/25/2016 5:29:00 AM |
|---|---|---|

by

| Page 26: Deleted | sue | 1/25/2016 5:35:00 AM |
|---|---|---|

we describe the

| Page 26: Inserted | sue | 1/25/2016 5:35:00 AM |
|---|---|---|

are described

| Page 26: Deleted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

the

| Page 26: Deleted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

that were

| Page 26: Deleted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

down

| Page 26: Inserted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

from the summit of

| Page 26: Deleted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

from the summit area

| Page 26: Deleted | sue | 1/25/2016 5:36:00 AM |
|---|---|---|

the

| Page 26: Deleted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

(the stream)

| Page 26: Deleted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

In these plots, the d

| Page 26: Inserted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

D

| Page 26: Deleted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

scale

| Page 26: Deleted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

While d

| Page 26: Inserted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

D

| Page 26: Deleted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

on these images

| Page 26: Inserted | sue | 1/25/2016 5:37:00 AM |
|---|---|---|

but

| Page 26: Deleted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

However, w

| Page 26: Inserted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

W

| Page 26: Deleted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

has

| Page 26: Inserted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

| Page 26: Deleted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

-

| Page 27: Inserted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

ranged from 0.58 to 2.42 m and

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

(

| Page 27: Deleted | sue | 1/25/2016 5:38:00 AM |
|---|---|---|

, with a range of 0.58 to 2.42 m.

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

,

| Page 27: Deleted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

summarizes the depth to bedrock values estimated from the two radar traverses shown in

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

)

| Page 27: Deleted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

the t

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

T

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

4

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

,

| Page 27: Deleted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

that descended from the Leading Ridge.  Data are

| Page 27: Inserted | sue | 1/25/2016 5:39:00 AM |
|---|---|---|

(Fig. 9, Table S2)

| Page 27: Inserted | sue | 1/26/2016 5:38:00 AM |
|---|---|---|

## 0.0 Hydrology: Groundwater measurements

Several methods are needed to characterize physical and chemical interactions of water with regolith and rock in a catchment. First, physical inputs and outputs to a catchment, including precipitation, interception, ET, soil infiltration, and groundwater discharge, must be understood. Often, groundwater flows are omitted from comprehensive hydrology-meteorology-vegetation models such as the Variable Infiltration Capacity (VIC) hydrologic model, or the Noah Land-Surface Model (LSM). However at Shale Hills, we have estimated that rough 50% of incoming water is evapotranspired and 5% reaches the regional groundwater table and returns to the stream as baseflow (Sullivan et al., subm.). At Garner Run, we also expect groundwater to play a significant role in streamflow and geochemical dynamics. For example, some researchers have found that drainage and runoff on sandstone catchments is controlled to great extent by bedrock (Hattanji and Onda, 2004), and specifically by flow through fractures in the upper meters of sandstone beneath the soil (Williams et al., 2010). In this section and the next section we focus on quantifying fluid flow and transport of solutes into surface water and groundwater. We aim to measure the relative magnitudes, timing, and spatial variability of these fluxes. We emphasize methodologies for measuring and characterizing groundwater and streamwater to identify subsurface flow paths of groundwater, and the drivers and controls on water-rock interactions.

In the spirit of “measuring only what is needed,” well installation and solid earth sampling by coring will be reduced compared to Shale Hills. At Shale Hills, 28 wells were emplaced and then intermittently monitored (Fig. 2). In Garner run, deep samples (> 8m) have been extracted between Garner Run and Roaring Run from the Harrys Valley 1 well (HV1) drilled within the Garner Run catchment (see Fig. 3). Using a hand-held drill, three shallow wells will be installed and cores will be collected at the catena sites (Fig. 5) and additional monitoring wells will be installed along hillslopes and the valley floor. From these wells, we will also sample solid-phase chemistry and mineralogy.

All core samples will be analyzed for bulk chemistry and mineralogy to characterize the weathering reactions and protolith. Where possible, we will install groundwater monitoring wells in boreholes, with screened intervals spanning the water table. Monitoring at the wells will include hourly water level measurements using autonomous pressure loggers, hourly temperature measurements at two depths below the water table, and monthly water samples collected and analyzed for major ion chemistry. A pumping test will be conducted at the adjacent valley floor wells to measure aquifer storativity and hydraulic conductivity. Relative residence time of groundwater will be assessed from pathway analysis. If resources permit, $SF_6$ and chlorofluorocarbons (CFCs) will be measured in groundwater samples to assess residence time in the subsurface, as we have done for Shale Hills (Sullivan et al., in press).

Deep core samples and groundwater monitoring will provide a baseline understanding of the geologic and hydrologic system on the new sandstone lithology. Subsequent hypotheses about controls on weathering and hydrologic dynamics, as well as historical flow and solute fluxes, will be constrained by these observations at the catchment boundaries.

## 0.0 Hydrology: Streamwater flow and chemistry measurements

The Garner Run study reach is approximately 500 m long (Fig. 5) and consists of a rocky, often braided, channel. We have installed a flume at the downstream end of the reach to measure discharge. Stage is continuously monitored using a pressure transducer (Hobo U-20, Onset Computer Corp., Hyannis, MA). Surface water – groundwater (SW-GW) exchange characteristics have been measured using a short-term deployment of a fiber-optic distributed temperature sensor (FO-DTS), and two tracer injection tests. Stream chemistry, including dissolved oxygen (DO), pH, total dissolved solids (TDS), $NO_3^-$, $SO_4^{2-}$, Ca, K, Mg, Mn, Na, Fe, and Si, are measured biweekly or monthly in the field with handheld electrodes along the 500 m reach, or by grab sampling and laboratory analysis (inductively coupled plasma-atomic emission spectroscopy, organic carbon analyzer, and ion chromatography).

Stream chemistry is also explored using higher temporal resolution by using a s::can spectrometer and an autosampler during storm events (s::can, GmBH, Vienna, Austria). The s::can is an in-situ instrument capable of measuring such water quality parameters as pH, TDS, dissolved organic carbon (DOC), $NO_3^-$, DO, $NH_4^+$, $K^+$, and $F^-$. The chemistry and tracer test data will help quantify the flux of fluid and solutes through the subcatchment. The stream chemistry and discharge data will be combined with soil moisture, soil pore water chemistry, and groundwater data to estimate relative contributions to the stream, and underlying processes related to weathering in the near surface and aquifer.

Preliminary results from Garner Run indicate lower concentrations of Ca, Mg, and K compared to the stream discharging from Shale Hills. In addition, as expected, an initial constant injection tracer test at Garner Run revealed significant exchange with the subsurface during low-flow conditions (~0.004 $m^3 s^{-1}$). Tracer test and temperature results suggest that the stream is losing water along some sections of the 500 m experimental reach and gaining water in others. Both the FO-DTS and stream chemistry data indicate significant input of spring water at ~100 m downstream of the Ground HOG catena (Fig. 5), which is chemically distinct from the upstream surface water and local groundwater sampled from the deep HV1 well. The DTS time series data will be analyzed to identify locations and magnitudes of inputs to the stream, as well as characteristic responses to rainfall events. In combination with the tracer tests, DTS, and chemistry results, we will use well logs and lidar topography to explain the lithological and geomorphologic controls on the surface water – ground water (SW-GW) system.

To characterize the major controls and processes governing WEGSS fluxes through the entire Shavers Creek catchment, we are making strategic measurements across the watershed to represent variability: stream discharge, stream chemistry, lithology, and geomorphology. Specifically, stream discharge and chemistry are being monitored along the main stem of Shavers Creek (SCAL, SCBL, and SCO, Fig. 3). At each location we are monitoring stage continuously using pressure transducers (Hobo U-20, Onset Computer Corp., Hyannis, MA), and using periodic discharge measurements to construct stage-discharge rating curves. SW-GW exchange characteristics will be measured as the channel crosses varying lithologies using a series of tracer injection tests. Analyses of stream chemistry from the main stem of Shavers Creek provide a spatial integration of solute behavior from upstream lithologies and land-use types. Eventually, with data from the three subcatchments on shale, sandstone, and calcareous shale (Fig. 3), we will make estimates for non-monitored catchments and test up-scaled estimates of the processes observed in each small watershed.

Preliminary stream chemistry and discharge results indicate significant variability among the three monitoring locations along Shavers Creek (Fig. 10). We see declining concentrations with increasing discharge for Mg and Ca (not shown), and somewhat chemostatic behavior for Si, K, nitrate and others. In this context, chemostatic is used to refer to concentrations of a stream that vary little with discharge (Godsey et al., 2009). Concentrations of Si decrease downstream (a dilution trend), while concentrations increase for Mg and nitrate, presumably due to agricultural amendments in the lower half of Shavers Creek watershed where land use includes farmland. The variety of behaviors will be investigated with respect to land use and lithology changes through the catchment.

| Page 30: Deleted | sue | 1/26/2016 5:38:00 AM |
|---|---|---|

### Hydrology: Groundwater measurements

Several methods are needed in a catchment to characterize physical and chemical interactions of water with regolith and rock. First, physical inputs and outputs to a catchment, including precipitation, interception, ET, soil infiltration, and groundwater discharge, must be understood. In fact, however, groundwater flows are often omitted from comprehensive hydrology-meteorology-vegetation models (e.g. the Variable Infiltration Capacity (VIC) hydrologic model, or the Noah Land-Surface Model (LSM)); however, at Shale Hills, we have estimated that 5% of the nonevapotranspired water that enters the catchment reaches the regional groundwater table and flows to the stream as a deep flow component (Sullivan et al., subm.). At Garner Run, we also expect groundwater to play a significant role in streamflow and geochemical dynamics. For example, some researchers have found that drainage and runoff on sandstone catchments is controlled to great extent by bedrock (Hattanji and Onda, 2004), and specifically by flow through fractures in the upper meters of sandstone directly beneath the soil (Williams et al., 2010). In this section and the next section we focus on quantifying flows through and between surface water and groundwater. We aim to measure the relative magnitudes, timing, and spatial variability of these fluxes. We emphasize methodologies for measuring and characterizing groundwater and streamwater to characterize groundwater residence times, identify subsurface flow paths, and the drivers and controls on water-rock interactions.

Our plans for well installation and solid earth sampling by coring are reduced compared to sampling at Shale Hills. At Shale Hills, 28 wells were emplaced and then intermittently monitored (Fig. 2). In Garner run, two deep cores (> 50m) will be extracted at two locations near the Garner Run catchment, one (~ 100m) on Tussey ridge, i.e. the ridge that divides Shavers Creek from the watersheds to the northwest and one (50-75m) on the smaller divide within Shavers Creek between Garner Run and Roaring Run (see Fig. 3). Three shallow wells will be installed and cores (~10m) will be collected at the catena sites (Fig. 5). Two to four additional monitoring wells will be installed along the stream reach on the valley floor. In drilling boreholes for assessment of groundwater, we also sample borehole solid-phase chemistry and mineralogy.

All core samples will be analyzed for bulk chemistry and mineralogy to characterize the weathering reactions and protolith in the critical zone. All boreholes will have groundwater monitoring wells installed, with screened intervals spanning the water table and with instrumentation as shown in Fig. 5. Monitoring at the wells will include hourly water level measurements using autonomous pressure loggers, hourly temperature measurements at two depths below the water table, and monthly water samples collected and analyzed for major ion chemistry. A pumping test will be conducted at the adjacent valley floor wells to measure aquifer storativity and hydraulic conductivity.subm.

Deep core samples and groundwater monitoring will provide a baseline understanding of the geologic/pedologic and hydrologic system on the new sandstone lithology. Subsequent hypotheses about controls on weathering and hydrologic dynamics, as well as historical flow and solute fluxes, will be constrained by these observations at the catchment boundaries.

### Hydrology: Streamflow and chemistry measurements

The Garner Run study reach is approximately 500 m long within the catchment (Fig. 5) and consists of a rocky, often braided, channel. We have deployed a flume at the downstream end of the reach to measure discharge, and are monitoring stage continuously using a pressure transducer (Hobo U-20, Onset Computer Corp., Hyannis, MA). Surface water – groundwater (SW-GW) exchange characteristics have been measured using a short-term deployment of a distributed temperature sensor (DTS), and will be supplemented by a series of tracer injection tests to investigate hyporheic exchange characteristics over a wider range of stream discharges. Stream chemistry, including DO, pH, TDS, $NO_3^-$, $SO_4^{2-}$, Ca, K, Mg, Mn, Na, Fe, and Si, are being measured biweekly or monthly in the field with handheld electrodes along the 500 m reach, or by grab sampling and laboratory analysis (inductively coupled plasma atomic emission spectroscopy or ion chromatography).

Stream chemistry is also being monitored intermittently using higher temporal resolution by using a s::can spectrometer and an autosampler during storm events. The s::can is an in-situ measurement instrument for several water quality parameters (pH, TDS, DOC, $NO_3^-$,DO, $NH_4^+$, K, F (s::can, GmBH, Vienna, Austria). The chemistry and tracer test data will help quantify the flux of fluid and solutes through the subcatchment. The stream chemistry and discharge data will be combined with soil moisture, soil pore water chemistry, and groundwater data to estimate relative contributions to the stream, and underlying processes related to weathering in the near surface and aquifer.

Preliminary results from Garner Run indicate lower concentrations of Ca, Mg, and K, compared to Shale Hills. In addition, as expected, an initial constant injection tracer test at Garner Run revealed significant exchange with the subsurface during low-flow conditions (~0.004 $m^3$ $s^{-1}$). Tracer test and temperature results documented that the stream sometimes loses and sometimes gains water in different sections over the 500 m experimental reach. One point stood out along the reach: both DTS and stream chemistry measurements are consistent with a significant input of groundwater at ~100m downstream of the catena (Fig. 5). The DTS time series data will be analyzed to identify locations and magnitudes of groundwater inputs, as well as characteristic responses to rainfall events or changes in stream discharge. In combination with the tracer tests, DTS, and chemistry results, we will use well logs and lidar topography to explain the lithological and geomorphologic controls on the SW-GW system.

To characterize the major controls and processes governing WEGSS fluxes through the entire Shavers Creek catchment, we are making strategic measurements across the watershed to represent variability: stream discharge, stream chemistry, lithology, and geomorphology. Stream discharge and chemistry are being monitored along the main stem of Shavers Creek (SCAL, SCBL, and SCO) as shown in Fig. 3. At each location we are constructing a stage-discharge rating curve, and monitoring stage continuously using pressure transducers (Hobo U-20, Onset Computer Corp., Hyannis, MA). Streamwater-groundwater exchange characteristics will be measured as the channel crosses varying lithologies using a series of tracer injection tests. Stream chemistry will be measured monthly at each sampling site along Shavers creek. Analyses from the main stem of Shavers Creek provides a spatial integration of solute behavior from upstream lithologies and land use types. Eventually, with data from the three subcatchments on shale, sandstone, and calcareous shale, we will make estimates for nonmonitored catchments and test up-scaled estimates of the processes observed in each small watershed.

Preliminary stream chemistry and discharge results indicate significant variability among the three monitoring locations along Shavers Creek (Fig. 10). We see declining concentrations with increasing discharge for Mg and Ca (not shown), and somewhat chemostatic behavior for Si, K, nitrate and others. In this context, chemostatic is used to refer to concentrations of a stream that vary little with discharge (Godsey et al., 2009). Concentrations of Si decrease downstream (a dilution trend), while concentrations increase for Mg and nitrate, possibly due to agricultural amendments in the lower half of the watershed. The variety of behaviors will be investigated with respect to land use and lithology changes through the catchment.

| Page 30: Formatted | sue | 1/26/2016 7:37:00 AM |
|---|---|---|

Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.13" + Tab after: 0.43" + Indent at: 0.43"

| Page 30: Inserted | Dan K. Arthur | 1/25/2016 9:46:00 AM |
|---|---|---|

**subm.**

| Page 30: Inserted | sue | 1/15/2016 7:40:00 PM |
|---|---|---|

## 4 Model-data feedbacks

Throughout this paper we have described the two-way exchange of field-model insights needed to maximize the efficiency of CZ science. To understand the CZ requires models 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 are necessary to provide data for calibration and testing.

The CZ approach of using models to cross from short to long timescales has an important major benefit. Investigations that target long timescales can tease out the effects of feedbacks and thresholds in complex systems that are difficult to discern in short-timescale studies. We thus use quantitative models to explore a vast range in both spatial *and* temporal scales. In this paper we emphasized our approach toward designing a CZO as a tool to understand the CZ as one integral system. We therefore emphasized only one modelling tool, the PIHM family of models. This cascade of models provides a quantitative way for different disciplines to interact about the CZ through the use of a shared model. Our current conceptual understanding and our current computers do not allow us to produce one model that simulates the CZ at all timescales, hence the cascade of models (Table 1).

Such a suite of models is integral not simply for predicting landscape and ecosystem response, but also to building a heuristic understanding of individual CZ processes that may not be apparent from 1st-order observations. Systems-level models are especially needed for proposing and testing hypotheses about feedbacks between climate, biota, and Earth surface and near-surface processes. Although not emphasized here, we have also cited publications throughout this paper that describe the many smaller scales or disciplinary-specific hypotheses and models that have been invoked to learn about individual CZ systems. For example, we point to our earlier observation of Ti and K enrichment in Garner Run soils (Fig. 8). We suggested several processes that could interact to explain data in Fig. 8, including preferential retention of some elements in protolith compared to others, depth variations in protolith composition, accumulation of fines from weathering formations above the current protolith, and dust additions. While first-order mass balance model calculations such as those implicit to Fig. 8 can be used to propose or test such hypotheses, use of Regolith-RT-PIHM (Table 1) or WITCH (Godderis et al., 2006) to model regolith formation are necessary to quantitatively test feasibility of such ideas. Better understanding of regolith formation will in turn inform the permeability distributions needed for hydrologic flow models in the CZO.

| Page 30: Formatted | sue | 1/15/2016 7:44:00 PM |
|---|---|---|

Line spacing: 1.5 lines

| Page 31: Formatted | sue | 1/15/2016 7:40:00 PM |
|---|---|---|

Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.13" + Tab after: 0.43" + Indent at: 0.43"

| Page 31: Deleted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

| Page 31: Deleted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

f course, o

| Page 31: Deleted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

low or no

| Page 31: Inserted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

little

| Page 31: Inserted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

This results from the

| Page 31: Deleted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

Obviously, there is a

| Page 31: Inserted | sue | 1/25/2016 6:33:00 AM |
|---|---|---|

ground- and tower-based measurements (

| Page 31: Inserted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

,

| Page 31: Deleted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

and

| Page 31: Inserted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

)

| Page 31: Deleted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

,

| Page 31: Inserted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

and

| Page 31: Deleted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

and

| Page 31: Inserted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

as well as

| Page 31: Deleted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

as described in

| Page 31: Inserted | sue | 1/25/2016 6:34:00 AM |
|---|---|---|

(

| Page 31: Inserted | sue | 1/25/2016 6:35:00 AM |
|---|---|---|

)

| Page 32: Inserted | sue | 1/16/2016 6:40:00 AM |
|---|---|---|

If we find that our limited digging of soil pits is not successful in characterizing the regolith adequately – if our models of regolith formation do not match observations or our models of water flow through regolith do not simulate observations – more pits can be dug or new approaches toward geophysical measurements can be refined.

| **Page 32: Deleted** | **sue** | **1/25/2016 6:46:00 AM** |
|---|---|---|

As an example of this approach, we point to our earlier observation of loss of Al, Na, Si and P from the soils at the same time that we identified significant enrichment in Mg, K, and Fe (Fig. 8). Simple mass balance arguments can be used to show that the enrichments in these latter elements are not likely due to residual accumulation during weathering of the parent orthoquartzite: prohibitively large thickness of quartzite would have had to weather away without loss of any Mg, K or Fe to enrich the soils adequately. On the other hand, accumulation of dust during weathering over a significant time period could explain the enrichment. Alternately, downward mobilization of fine particles from weathering of the overlying Rose Hill shale or interfingered shaley units might adequately explain the enrichment in these elements. Use of Regolith-RT-PIHM (Table 1) or WITCH (Godderis et al., 2006) to model regolith formation should allow testing of the feasibility of these or other ideas. With

| **Page 32: Inserted** | **sue** | **1/25/2016 6:46:00 AM** |
|---|---|---|

As we build understanding,

| **Page 32: Inserted** | **sue** | **1/25/2016 6:43:00 AM** |
|---|---|---|

will be used to

| **Page 32: Deleted** | **sue** | **1/25/2016 6:43:00 AM** |
|---|---|---|

we can also

| **Page 32: Deleted** | **sue** | **1/16/2016 6:34:00 AM** |
|---|---|---|

Of course, we can also augment t

| **Page 32: Inserted** | **sue** | **1/16/2016 6:34:00 AM** |
|---|---|---|

T

| **Page 32: Inserted** | **sue** | **1/16/2016 6:34:00 AM** |
|---|---|---|

is also being augmented

| **Page 32: Deleted** | **sue** | **1/26/2016 7:44:00 AM** |
|---|---|---|

inside and

| **Page 32: Deleted** | **sue** | **1/26/2016 7:44:00 AM** |
|---|---|---|

or

**Page 32: Inserted** sue 1/26/2016 7:44:00 AM

and outside

**Page 32: Deleted** sue 1/16/2016 6:34:00 AM

can

**Page 32: Deleted** sue 1/25/2016 6:47:00 AM

,

**Page 32: Deleted** sue 1/16/2016 6:34:00 AM

may

**Page 32: Inserted** sue 1/16/2016 6:34:00 AM

are

**Page 32: Deleted** sue 1/16/2016 6:34:00 AM

characterize

**Page 32: Inserted** sue 1/16/2016 6:34:00 AM

characterizing

**Page 32: Deleted** sue 1/16/2016 6:34:00 AM

that we

**Page 32: Inserted** sue 1/16/2016 6:35:00 AM

This is a good example of targeted investigations that are not directly related to parameterization of the models in Table 1 for our CZO itself but are rather aimed at improving the process-based understanding that underlies models of CZ evolution.

**Page 32: Deleted** sue 1/25/2016 6:48:00 AM

**Page 32: Deleted** sue 1/16/2016 6:36:00 AM

in a catchment to the north of Shavers creek

**Page 32: Inserted** sue 1/16/2016 6:37:00 AM

in other catchments north of Shavers creek

**Page 32: Inserted** sue 1/16/2016 6:37:00 AM

s

**Page 32: Deleted** sue 1/16/2016 6:37:00 AM

is considerably faster

**Page 32: Inserted** sue 1/16/2016 6:37:00 AM

differ

**Page 32: Deleted** sue 1/16/2016 6:37:00 AM

At this site, we anticipate learning

**Page 32: Formatted** sue 1/25/2016 6:51:00 AM

Highlight

**Page 32: Inserted** sue 1/16/2016 6:37:00 AM

Such targeted investigations can also be compared to output from sensitivity tests where pertinent models are used to explore the effect of the targeted variables

**Page 32: Deleted** sue 1/16/2016 6:37:00 AM

how to parameterize or run models of regolith formation

**Page 32: Deleted** sue 1/16/2016 6:37:00 AM

by exploring the impact of the rate of erosion

**Page 32: Inserted** sue 1/25/2016 6:48:00 AM

Measurements outside the CZO may therefore highlight problems in our limited sampling scheme or modelling approaches that must be improved.

**Page 32: Formatted** sue 1/25/2016 6:48:00 AM

Highlight

**Page 32: Formatted** sue 1/25/2016 6:48:00 AM

Highlight

**Page 32: Formatted** sue 1/15/2016 7:04:00 PM

Normal

**Page 32: Deleted** sue 1/25/2016 6:51:00 AM

critical zone

**Page 32: Inserted** sue 1/25/2016 6:51:00 AM

CZ

**Page 33: Deleted** sue 1/25/2016 6:52:00 AM

critical zone

**Page 33: Inserted** sue 1/25/2016 6:52:00 AM

CZ

**Page 33: Inserted** sue 1/25/2016 6:56:00 AM

(CZOs)

**Page 33: Inserted** sue 1/25/2016 6:53:00 AM

Of great interest are

**Page 33: Deleted** sue 1/25/2016 6:53:00 AM

The most

**Page 33: Deleted** sue 1/25/2016 6:53:00 AM

models we have are

**Page 33: Inserted** sue 1/25/2016 6:53:00 AM

that aid in understanding the CZ

**Page 33: Inserted** sue 1/25/2016 6:54:00 AM

but such conceptual understanding must also be encoded within

**Page 33: Deleted** sue 1/25/2016 6:54:00 AM

and the most predictive are

**Page 33: Inserted** sue 1/25/2016 6:54:00 AM

that allow quantitative predictions for testing

**Page 33: Deleted** sue 1/25/2016 6:55:00 AM

However

**Page 33: Inserted** sue 1/25/2016 6:55:00 AM

Nonetheless

**Page 33: Inserted** sue 1/25/2016 6:55:00 AM

conceptual and numerical models still

**Page 33: Deleted** sue 1/25/2016 6:55:00 AM

typically

**Page 33: Deleted** sue 1/25/2016 6:55:00 AM

critical zone

**Page 33: Inserted** sue 1/25/2016 6:55:00 AM

CZ

**Page 33: Inserted** sue 1/25/2016 6:55:00 AM

To really understand WEGSS fluxes quantitatively requires

| Page 33: Deleted | sue | 1/25/2016 6:55:00 AM |
|---|---|---|

We seek

| Page 33: Inserted | sue | 1/25/2016 6:55:00 AM |
|---|---|---|

biota, atmospheric conditions,

| Page 33: Inserted | sue | 1/15/2016 7:04:00 PM |
|---|---|---|

In our efforts, new observations are tested against and incorporated into the PIHM models to explore the evolution of the CZ over time. In this endeavor, we can also ask, what does success look like? 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 project (earthcast) the CZ. Ultimately, success means that we gain deeper understanding of the system and can predict behavior in other places or with other datasets (e.g. tracers, water isotopes, etc.). Such testing is built in to our nested watershed approach (Fig. 3) and is also implicit to the design of the greater CZO network.

We can also imagine other indicators of success. For example, successful datasets will attract other researchers using other models. This in turn can lead to model-model inter-comparisons. If other models provide better simulations of the catchment, this will drive development of better models. One example of a model – model inter-comparison (RT-Flux-PIHM versus WITCH-Flux-PIHM, Table 1) has already driving new insights.

Another indicator of success is adoption by others of the strategies developed to study the CZ. Such strategies include design of a sampling paradigm for an individual CZO, design of a larger network of CZOs, development of suites of models, or approaches for data assimilation. While the CZO enterprise is still young, publications in the literature already attest to growth in use of the PIHM suite of models in other places (11-16) and growth in use of the CZO concept worldwide (17).

| Page 34: Inserted | sue | 1/22/2016 7:09:00 AM |
|---|---|---|

and for PIHM

| Page 34: Inserted | Dan K. Arthur | 1/25/2016 9:39:00 AM |
|---|---|---|

J. Doolittle from the USDA-NRCS is acknowledged for help with GPR data collection.

**Page 34: Inserted** **sue** **1/25/2016 7:04:00 AM**

**Page 34: Inserted** **Dan K. Arthur** **1/25/2016 9:26:00 AM**

Logistical support and data were provided by the NSF-supported Shale Hills Susquehanna Critical Zone Observatory. This research was conducted in Penn State's Stone Valley Forest, which is supported and managed by the Penn State's Forestland Management Office in the College of Agricultural Sciences.

**Page 34: Inserted** **sue** **1/25/2016 7:04:00 AM**

The Penn State Earth and Environmental Systems Institute and College of Agricultural Sciences is acknowledged for staff and funding support.

**Page 35: Deleted** **Dan K. Arthur** **1/25/2016 10:04:00 AM**

Certini, G., Hillier, S., McMurray, E. and Edwards, A. (2003) Weathering of sandstone clasts in a forest soil in Tuscany (Italy). Geoderma 116, 357-372.

**Page 38: Deleted** **Dan K. Arthur** **1/25/2016 10:11:00 AM**

Jin, L., Ravella, R., Ketchum, B., Bierman, P.R., Heaney, P., White, T. and Brantley, S.L. (2010a) Mineral weathering and elemental transport during hillslope evolution at the Susquehanna/Shale Hills Critical Zone Observatory. Geochimica et Cosmochimica Acta 74, 3669-3691.

**Page 38: Deleted** **Dan K. Arthur** **1/25/2016 10:11:00 AM**

b

**Page 38: Deleted** **Dan K. Arthur** **1/25/2016 10:13:00 AM**

Labus, M. and Bochen, J. (2012) Sandstone degradation: an experimental study of accelerated weathering. Journal of Soil Science 32, 453-459.

**Page 39: Deleted** **Dan K. Arthur** **1/25/2016 9:47:00 AM**

Seaber, P., Kapinos, F. P., and Knapp, G. L.: Hydrologic Unit Maps, US Geological Survey, Water-Supply Paper, 2294, Denver, CO, 1987.

**Page 40: Deleted** **Dan K. Arthur** **1/25/2016 10:22:00 AM**

Turkington, A. and Paradise, T.: Sandstone weathering: a century of research and innovation, Geomorphology, 67, 229–253, doi:10/1016/j.geomorph.2004.09.028, 2005.

Table 2. Measurements and instrumentation for Tower HOG system

| Measurement | Manufacturer | Model | Collection frequency |
|---|---|---|---|
| $[CO_2]$, $[H_2O]$ | Li-cor | LI-7500A $CO_2/H_2O$ analyzer | 10Hz‡ |
| 3-D wind velocity, virtual temperature | Campbell Scientific | CSAT3 sonic anemometer | 10Hz‡ |
| Precipitation | OTT Hydromet | $Pluvio^2$ Weighing Rain Gauge | Every 10 min |
| $T_{air}$ | Vaisala | HMP60 humidity and temperature probe | Every 30 min |
| Relative Humidity | Vaisala | HMP60 humidity and temperature probe | Every 30 min |
| Longwave Radiation* | Kipp & Zonen | CGR3 pyrgeometer | Every 30 min |
| Shortwave Radiation* | Kipp & Zonen | CMP3 pyranometer | Every 30 min |
| Snow depth† | Campbell Scientific | SR50A sonic ranging sensor | Every 30 min |
| Digital Imagery | Campbell Scientific | CC5MPX digital camera | Every 24 hr |

* All four components of radiation (upwelling and downwelling (longwave and shortwave)) will only be measured at Shale Hills Tower HOG due to the location of the Garner Run Tower HOG. To model Garner Run we will use the Shale Hills data.

† originally designed as part of tower system but will be deployed at LRVF Ground HOG location because the Garner Run tower will be located outside of the catchment.

‡The turbulent fluxes (sensible and latent heat) and the momentum flux are computed at 30 minute intervals via eddy covariance using these data collected at 10 Hz.

| Page 43: Deleted | sue | 1/24/2016 7:51:00 PM |
|---|---|---|

Table 2. Measurements and instrumentation for Tower HOG system

| Measurement | Manufacturer | Model | Collection frequency |
|---|---|---|---|
| [$CO_2$], [$H_2O$] | Li-cor | LI-7500A $CO_2/H_2O$ analyzer | 10Hz‡ |
| 3-D wind velocity, virtual temperature | Campbell Scientific | CSAT3 sonic anemometer | 10Hz‡ |
| Precipitation | Thies ClimaOTT Hydromet | LPM disdrometerPluvio[2] Weighing Rain Gauge | Every 30 10 min |
| $T_{air}$Precipitation type | VaisalaThies Clima | HMP60 humidity and temperature probeLPM disdrometer | Every 30 minEvery 30 min |
| Relative Humidity$T_{air}$ | VaisalaVaisala | HMP60 humidity and temperature probeHMP60 humidity and temperature probe | Every 30 minEvery 30 min |
| Longwave Radiation*Relative Humidity | Kipp & ZonenVaisala | CGR3 pyrgeometerHMP60 humidity and temperature probe | Every 30 minEvery 30 min |
| Shortwave Radiation*Longwave Radiation* | Kipp & ZonenKipp & Zonen | CMP3 pyranometerCGR3 pyrgeometer | Every 30 minEvery 30 min |
| Snow | Campbell | SR50A sonic ranging | Every 30 minEvery |

| | | | |
|---|---|---|---|
| depth†Shortwave Radiation* | ScientificKipp & Zonen | sensorCMP3 pyranometer | 30 min |
| Digital ImagerySnow depth† | Campbell ScientificCampbell Scientific | CC5MPX digital cameraSR50A sonic ranging sensor | Every 24 hrEvery 30 min |
| Digital Imagery | Campbell Scientific | CC5MPX digital camera | Every 24 hr |

* All four components of radiation (upwelling and downwelling (longwave and shortwave)) will only be measured at Shale Hills Tower HOG due to the location of the Garner Run Tower HOG. To model Garner Run we will use the Shale Hills data.

† originally designed as part of tower system but will be deployed at LRVF Ground HOG location because the Garner Run tower will be located outside of the catchment.

‡The turbulent fluxes (sensible and latent heat) and the momentum flux are computed at 30 minute intervals via eddy covariance using these data collected at 10 Hz.

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:30:00 AM |
|---|---|---|

OTT Hydromet

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:30:00 AM |
|---|---|---|

Pluvio$^2$ Weighing Rain Gauge

| Page 43: Formatted | Dan K. Arthur | 1/21/2016 9:30:00 AM |
|---|---|---|

Superscript

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:30:00 AM |
|---|---|---|

10

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

$T_{air}$

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Vaisala

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

HMP60 humidity and temperature probe

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Every 30 min

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Relative Humidity

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Vaisala

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

HMP60 humidity and temperature probe

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Every 30 min

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Longwave Radiation*

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Kipp & Zonen

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

CGR3 pyrgeometer

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Every 30 min

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Shortwave Radiation*

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Kipp & Zonen

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

CMP3 pyranometer

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Every 30 min

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Snow depth†

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

Campbell Scientific

**Page 43: Inserted** **Dan K. Arthur** **1/21/2016 9:29:00 AM**

SR50A sonic ranging sensor

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Every 30 min

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Digital Imagery

| Page 43: Formatted Table | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Formatted Table

| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Campbell Scientific

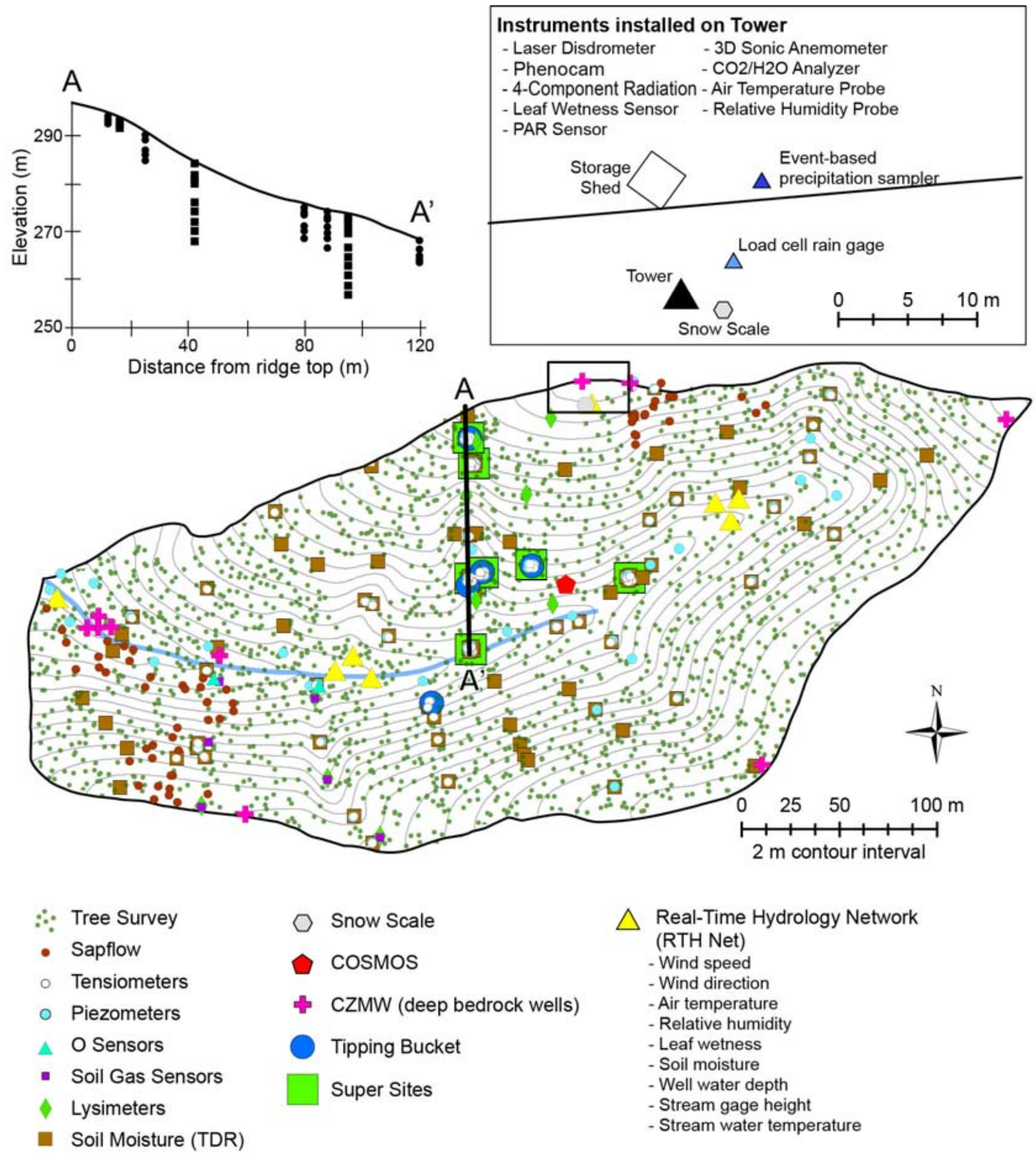
| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

CC5MPX digital camera

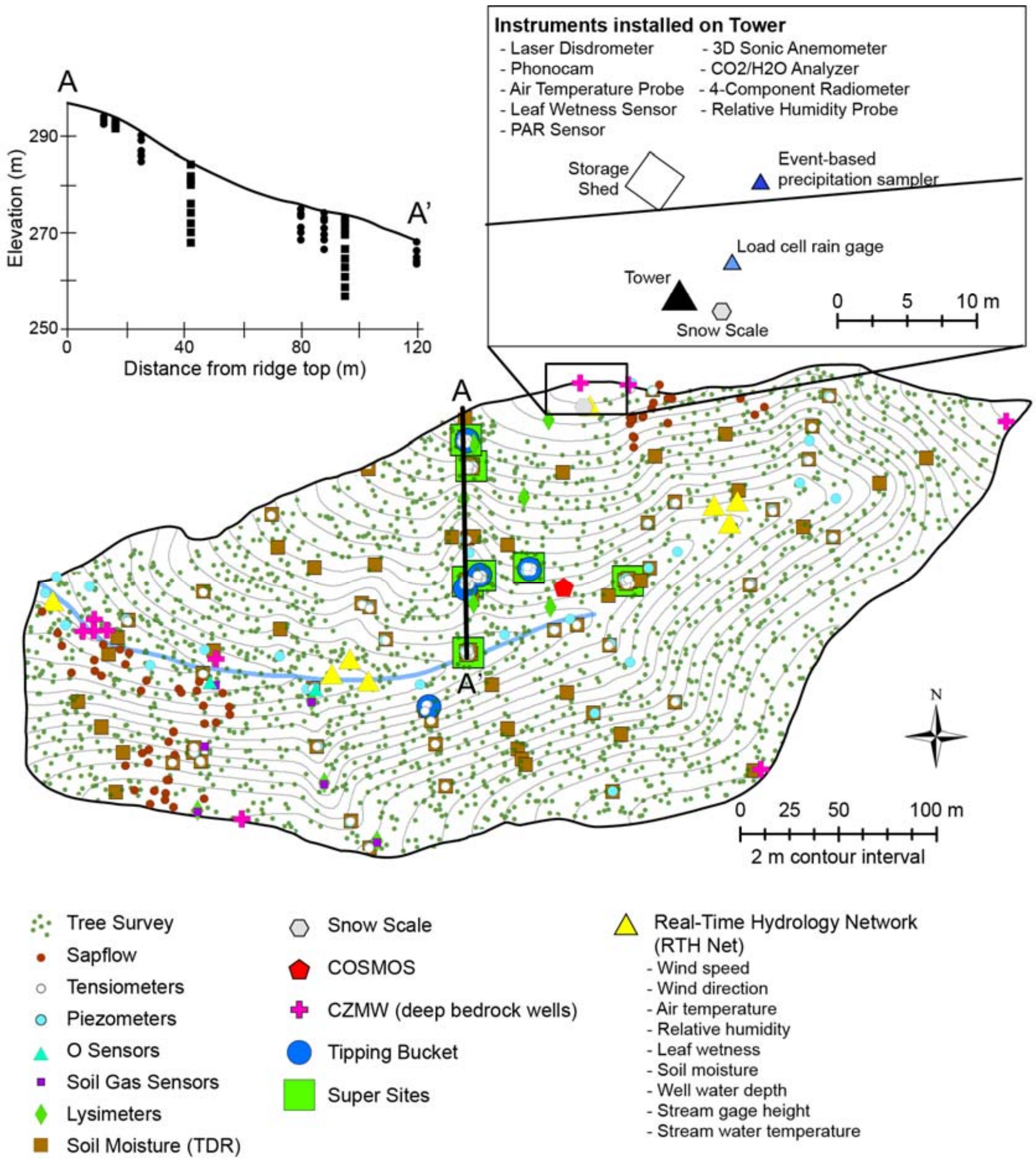
| Page 43: Inserted | Dan K. Arthur | 1/21/2016 9:29:00 AM |
|---|---|---|

Every 24 hr

| Page 47: Deleted | Dan K. Arthur | 1/21/2016 9:32:00 AM |
|---|---|---|




| Page 47: Inserted | Dan K. Arthur | 1/21/2016 9:32:00 AM |
|---|---|---|




| Page 47: Deleted | sue | 1/26/2016 5:45:00 AM |
|---|---|---|

black

| Page 47: Inserted | sue | 1/26/2016 5:45:00 AM |
|---|---|---|

green

| Page 47: Inserted | sue | 1/26/2016 5:46:00 AM |
|---|---|---|

:

| Page 47: Deleted | sue | 1/26/2016 5:46:00 AM |
|---|---|---|

.

| Page 47: Inserted | sue | 1/26/2016 5:46:00 AM |
|---|---|---|

the subcatchment contains a dry oak-mixed hardwood community type (Fike, 1999) with an extremely diverse mix of hardwood and softwood species, including white oak (*Quercus alba*), sugar maple (*Acer saccharum*), pignut hickory (*Carya glabra*), eastern hemlock (*Tsuga canadensis*, and chestnut oak (*Q. montana*). The sparse understory consists of American hop-hornbeam (*Ostrya virginiana*) and serviceberry (*Amelanchier sp.*).

| **Page 48: Inserted** | **sue** | **1/26/2016 5:45:00 AM** |
|---|---|---|

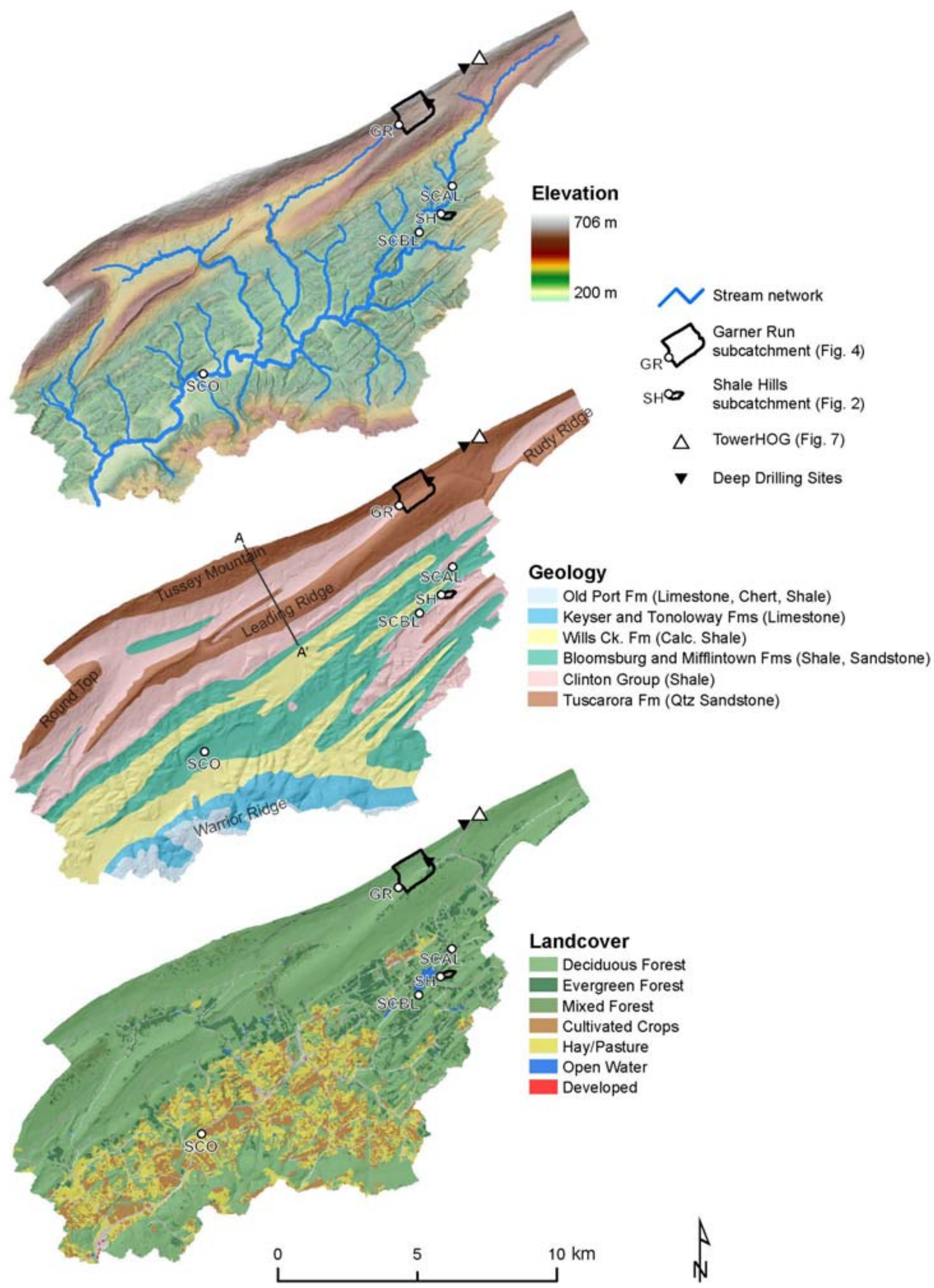
in the Shale Hills subcatchment

| **Page 48: Inserted** | **sue** | **1/26/2016 5:45:00 AM** |
|---|---|---|

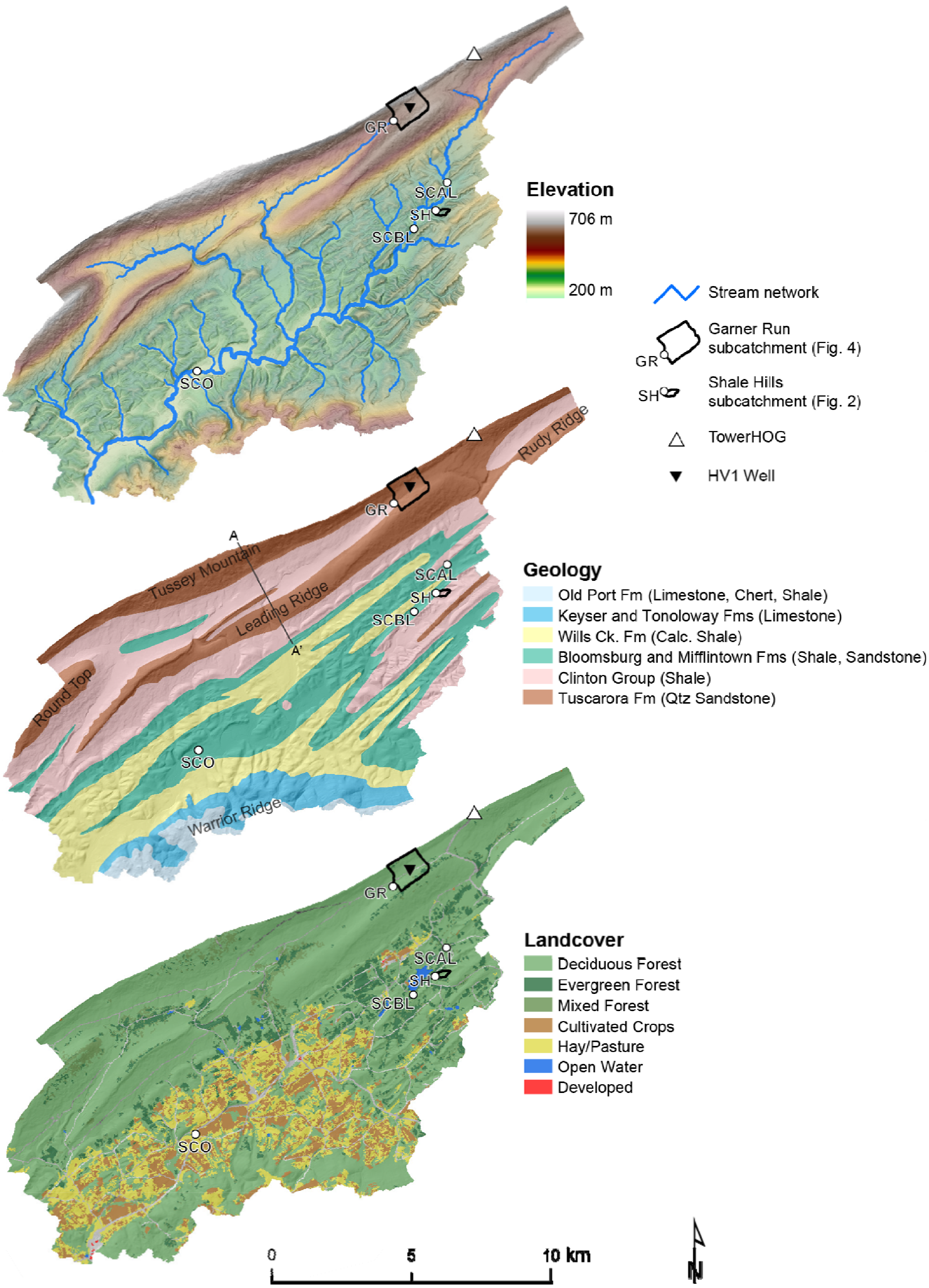
| **Page 48: Formatted** | **sue** | **1/26/2016 5:45:00 AM** |
|---|---|---|

German (Germany)

| **Page 49: Deleted** | **Dan K. Arthur** | **1/25/2016 1:34:00 PM** |
|---|---|---|




| Page 49: Inserted | Dan K. Arthur | 1/25/2016 1:34:00 PM |
|---|---|---|


GR
SCAL
SH
SCBL
SCO
Elevation
706 m
200 m
Stream network
Garner Run subcatchment (Fig. 4)
Shale Hills subcatchment (Fig. 2)
TowerHOG
HV1 Well
Rudy Ridge
Tussey Mountain
Leading Ridge
A
A'
Round Top
Warrior Ridge
Geology
Old Port Fm (Limestone, Chert, Shale)
Keyser and Tonoloway Fms (Limestone)
Wills Ck. Fm (Calc. Shale)
Bloomsburg and Mifflintown Fms (Shale, Sandstone)
Clinton Group (Shale)
Tuscarora Fm (Qtz Sandstone)
Landcover
Deciduous Forest
Evergreen Forest
Mixed Forest
Cultivated Crops
Hay/Pasture
Open Water
Developed
0
5
10 km
N

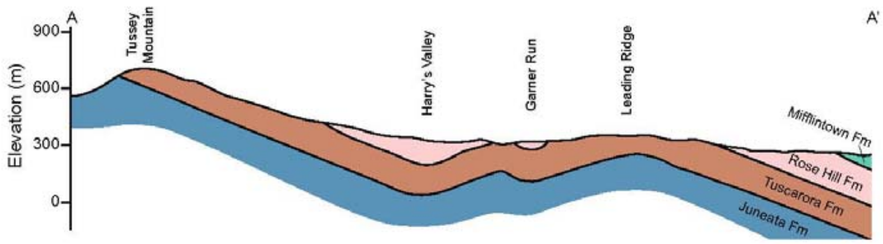
| Page 50: Deleted | Dan K. Arthur | 1/25/2016 9:47:00 AM |
|---|---|---|

NLCD

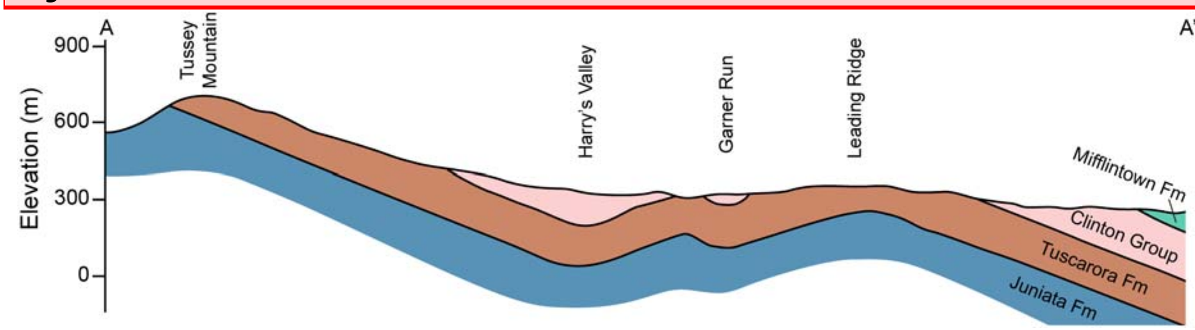
| Page 50: Inserted | Dan K. Arthur | 1/25/2016 9:47:00 AM |
|---|---|---|

Homer et al.

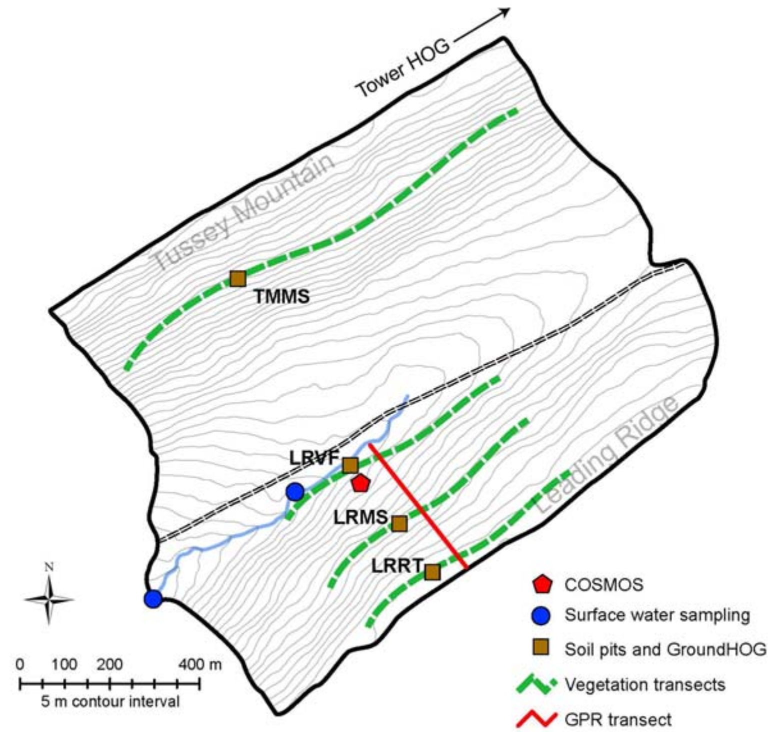
| Page 51: Deleted | Dan K. Arthur | 1/21/2016 9:40:00 AM |
|---|---|---|



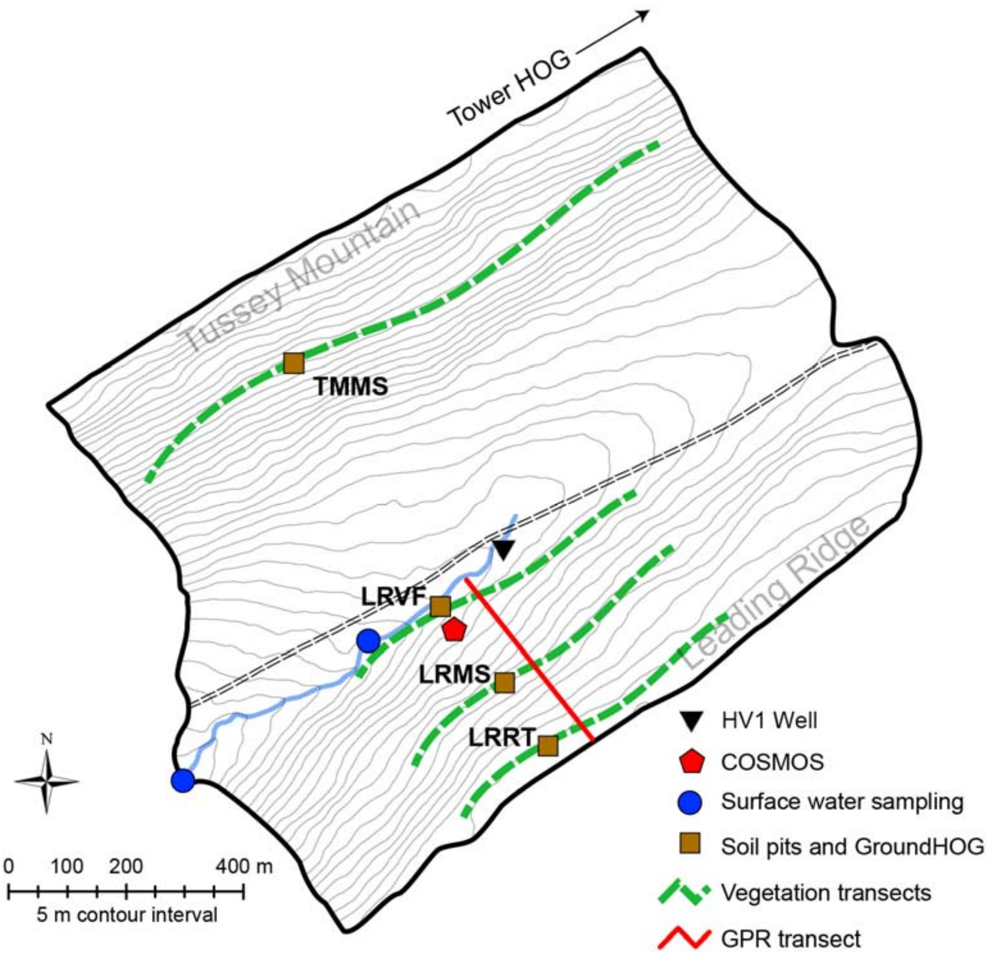

| Page 51: Inserted | Dan K. Arthur | 1/21/2016 9:40:00 AM |
|---|---|---|




| Page 52: Deleted | Dan K. Arthur | 1/25/2016 1:44:00 PM |
|---|---|---|



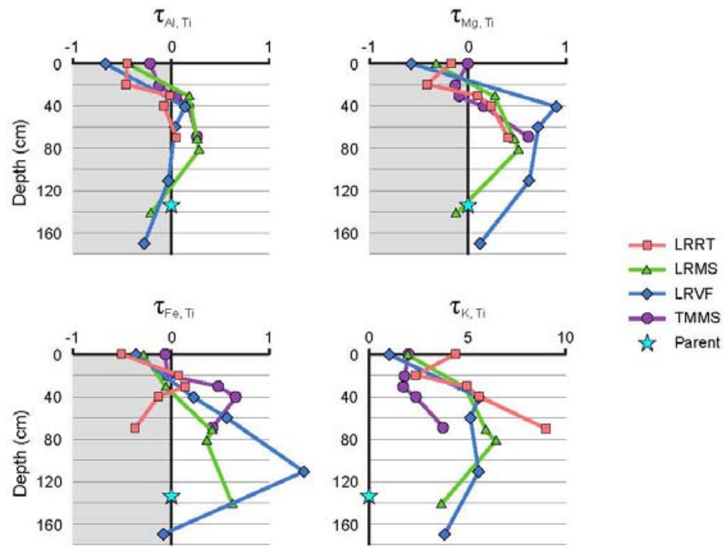

| Page 52: Inserted | Dan K. Arthur | 1/25/2016 1:44:00 PM |
|---|---|---|



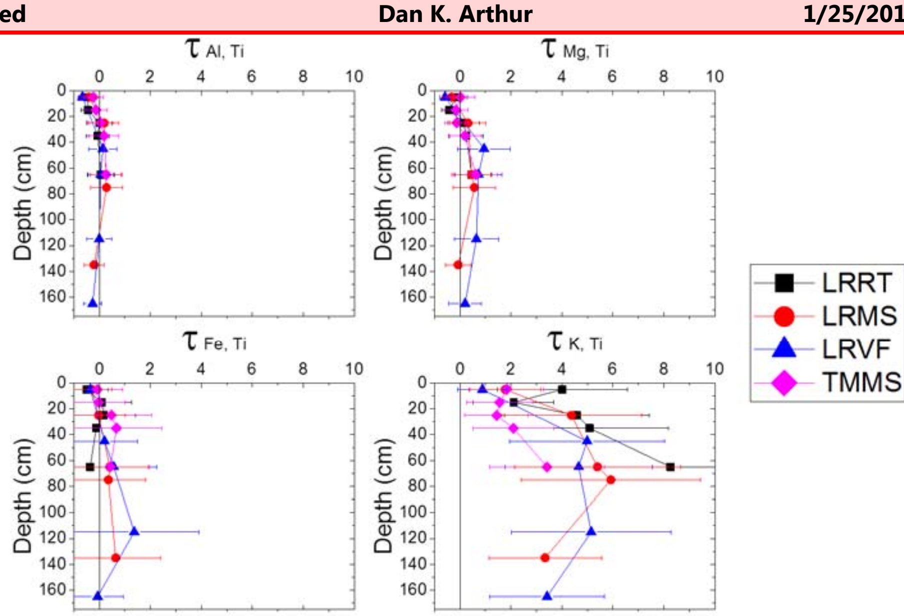

| Page 56: Formatted | Dan K. Arthur | 1/25/2016 11:26:00 AM |
|---|---|---|

Centered

| Page 57: Deleted | Dan K. Arthur | 1/25/2016 11:25:00 AM |
|---|---|---|




| Page 57: Inserted | Dan K. Arthur | 1/25/2016 11:25:00 AM |
|---|---|---|




| Page 57: Deleted | sue | 1/19/2016 5:43:00 AM |
|---|---|---|

Four plots

| Page 57: Inserted | sue | 1/19/2016 5:43:00 AM |
|---|---|---|

Plots

| Page 57: Inserted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

Garner Run sub-catchment

| Page 57: Inserted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

τ

| Page 57: Deleted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

The normalized concentration

| Page 57: Deleted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

average

| Page 57: Inserted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

estimated as the average of

| Page 57: Deleted | sue | 1/19/2016 5:47:00 AM |
|---|---|---|

from five

| Page 57: Inserted | sue | 1/18/2016 4:30:00 PM |
|---|---|---|

four

| Page 57: Inserted | sue | 1/18/2016 4:30:00 PM |
|---|---|---|

see data in

| Page 57: Deleted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

and

| Page 57: Inserted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

based on the assumption that

| Page 57: Inserted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

is

| Page 57: Deleted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

as the

| Page 57: Deleted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

element

| Page 57: Deleted | sue | 1/18/2016 4:31:00 PM |
|---|---|---|

One explanation for t

| Page 57: Inserted | sue | 1/18/2016 4:31:00 PM |
|---|---|---|

T

| Page 57: Inserted | sue | 1/18/2016 4:31:00 PM |
|---|---|---|

may indicate

| Page 57: Deleted | sue | 1/18/2016 4:31:00 PM |
|---|---|---|

is

| Page 57: Inserted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

If parent is correctly estimated,

| Page 57: Deleted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

In these plots,

| Page 57: Deleted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

completely

| Page 57: Inserted | sue | 1/19/2016 5:48:00 AM |
|---|---|---|

100%

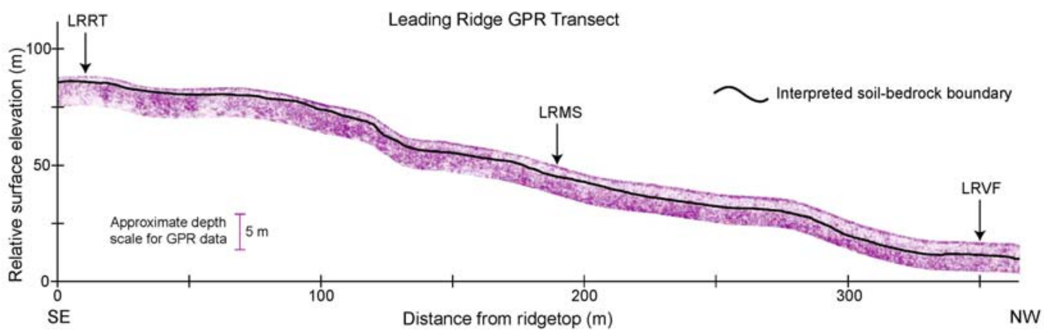
| Page 57: Deleted | sue | 1/19/2016 5:49:00 AM |
|---|---|---|

compared to Ti in the parent material

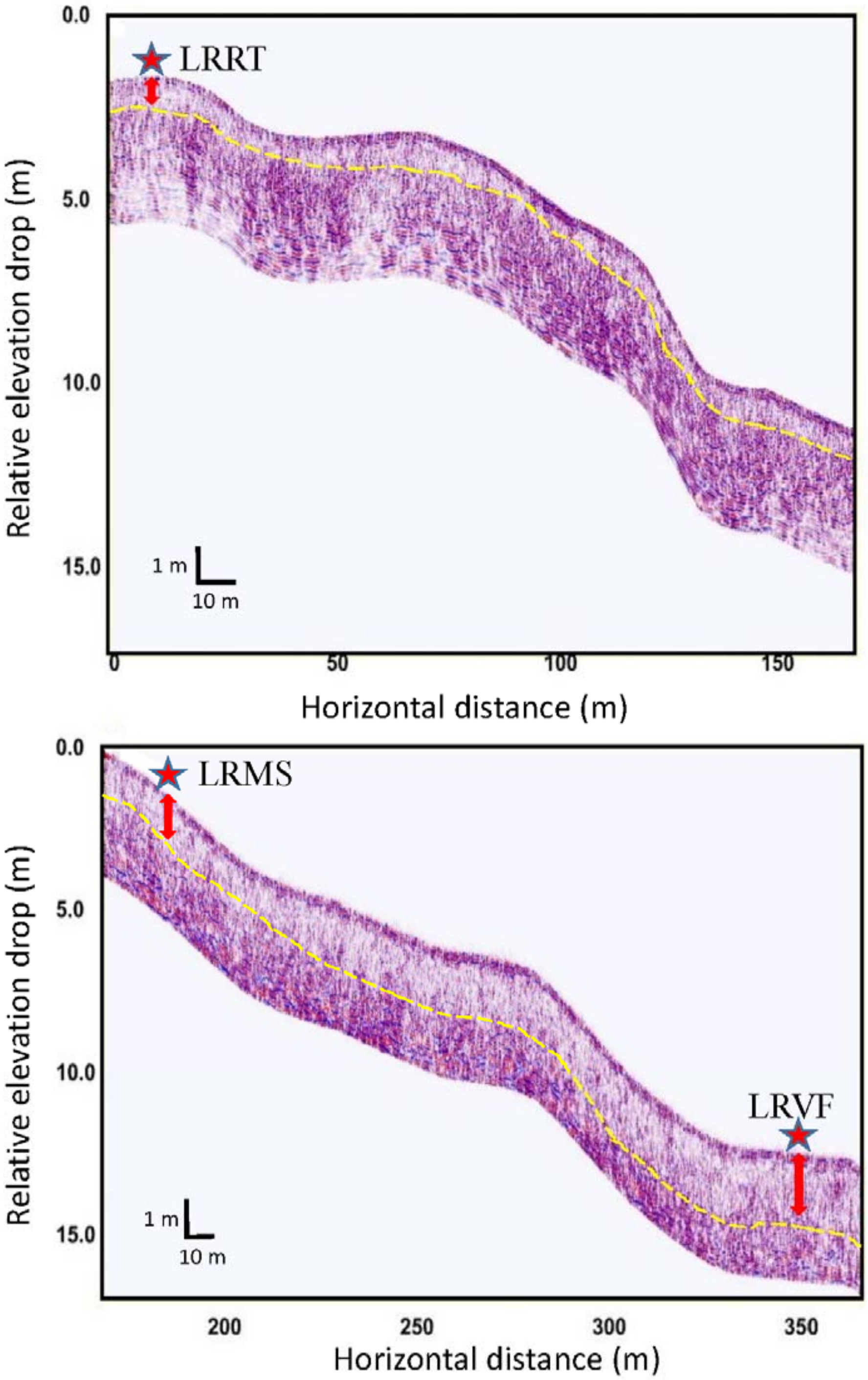
| Page 57: Inserted | sue | 1/19/2016 5:49:00 AM |
|---|---|---|

compared to Ti in the parent material

| Page 58: Deleted | Dan K. Arthur | 1/25/2016 11:11:00 AM |
|---|---|---|




| Page 58: Inserted | Dan K. Arthur | 1/25/2016 11:21:00 AM |
|---|---|---|




| Page 58: Inserted | Dan K. Arthur | 1/25/2016 11:21:00 AM |
|---|---|---|

Ground Penetrating Radar (GPR) transect of the Leading Ridge catena, showing inferred location of bedrock-soil interface (yellow dashed curve). The three soil pits (LRRT, LRMS, LRVF) are indicated by stars, with their observed depth to bedrock indicated by red arrow bar (note the observed depth was limited by excavation depth that was possible with the available digging tool). GPR data are exaggerated by 10x in vertical dimension as compared to surface topography. Summary values are tabulated in Table 4 from these GPR measurements.

**Page 59: Deleted** **Dan K. Arthur** **1/25/2016 11:21:00 AM**

Ground Penetrating Radar (GPR) transect of Leading Ridge Catena, showing inferred location of bedrock-soil interface. GPR data is exaggerated by 4x compared to surface topography. Summary values are tabulated in Table 4 from these GPR measurements.

**Header and footer changes**

**Text Box changes**

**Header and footer text box changes**

**Footnote changes**

**Endnote changes**

| Page 44: Deleted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

,

| Page 44: Inserted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

(

| Page 44: Deleted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

(

| Page 44: Deleted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

,

| Page 44: Inserted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

(

| Page 44: Deleted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

(

| Page 44: Deleted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

.,

| Page 44: Inserted | sue | 1/27/2016 5:41:00 AM |
|---|---|---|

. (

| Page 45: Deleted | sue | 1/27/2016 5:42:00 AM |
|---|---|---|

originally

| Page 45: Inserted | sue | 1/27/2016 5:42:00 AM |
|---|---|---|

Originally

| Page 46: Inserted | sue | 1/27/2016 5:42:00 AM |
|---|---|---|

(Fig. 5)

| Page 47: Inserted | sue | 1/27/2016 5:44:00 AM |
|---|---|---|

bedrock

| Page 47: Deleted | sue | 1/27/2016 5:44:00 AM |
|---|---|---|

to bedrock

| Page 47: Inserted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

measurements

| Page 47: Deleted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

the

| Page 47: Inserted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

GPR

| Page 47: Formatted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

Font: Not Bold

| Page 47: Formatted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

Font: Not Bold

| Page 47: Formatted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

Font: Not Bold

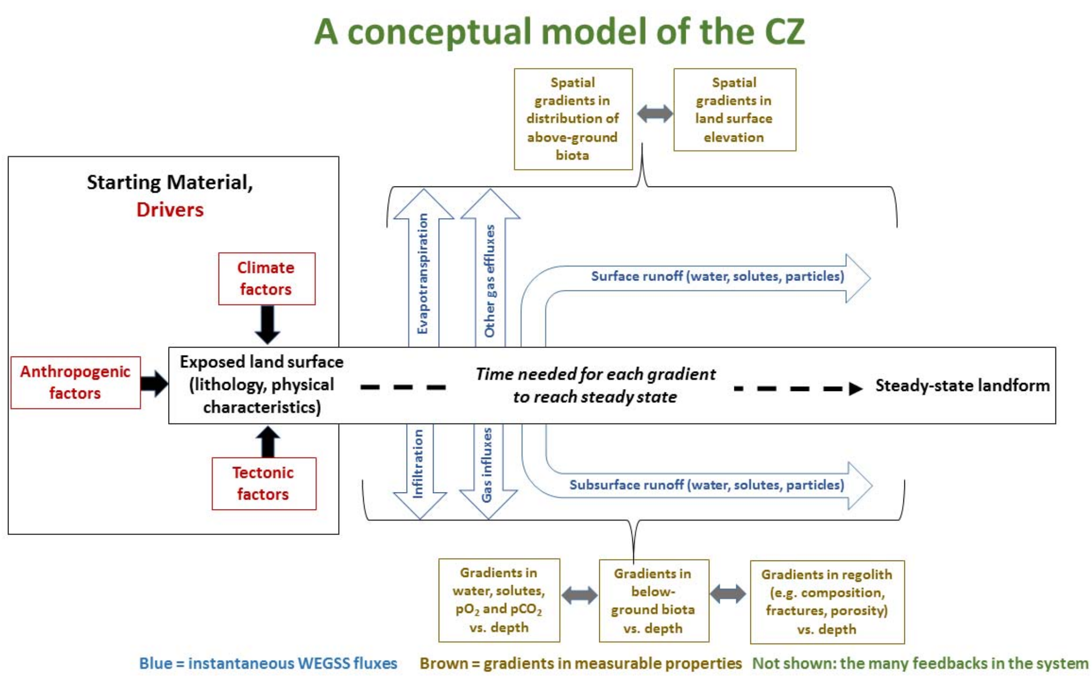
| Page 47: Formatted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

Font: Not Bold

| Page 47: Formatted | sue | 1/27/2016 5:43:00 AM |
|---|---|---|

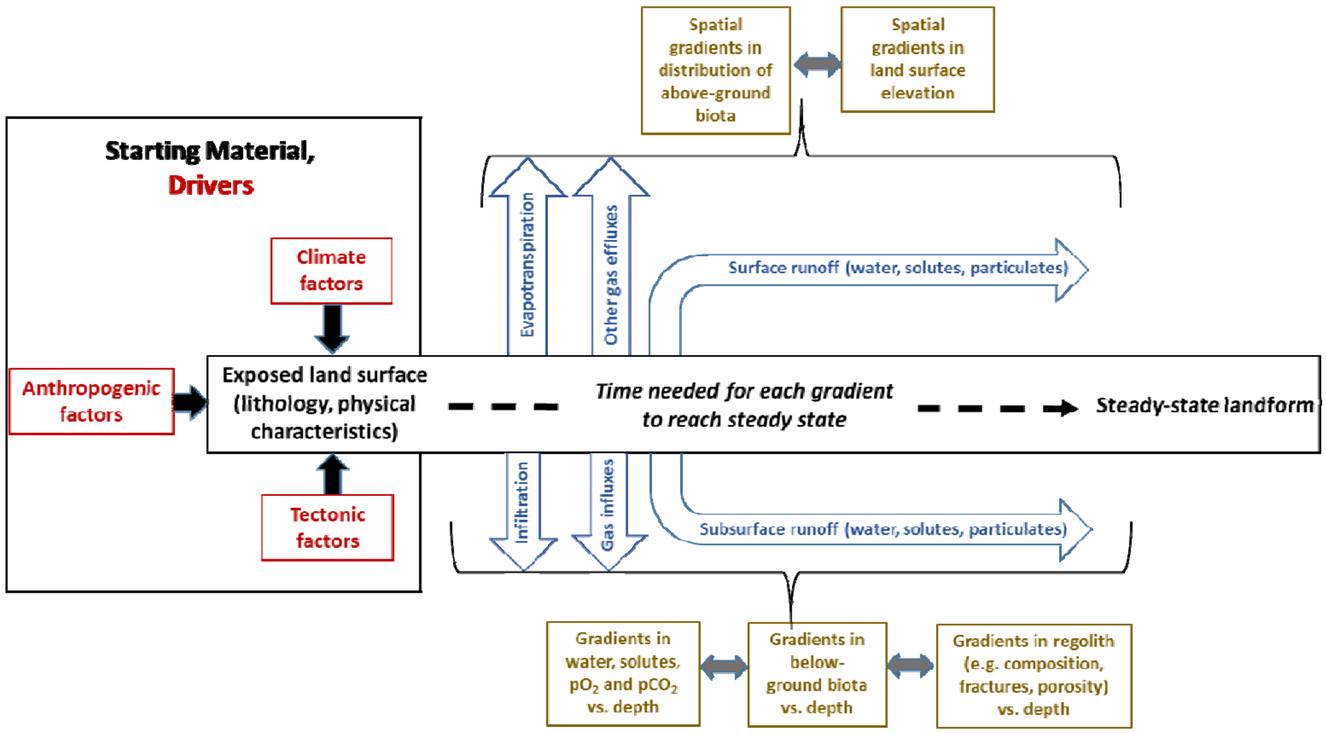
Font: Not Bold

| Page 48: Inserted | Dan K. Arthur | 1/27/2016 8:42:00 AM |
|---|---|---|




| Page 48: Deleted | Dan K. Arthur | 1/27/2016 8:42:00 AM |
|---|---|---|

# A conceptual model of the CZ

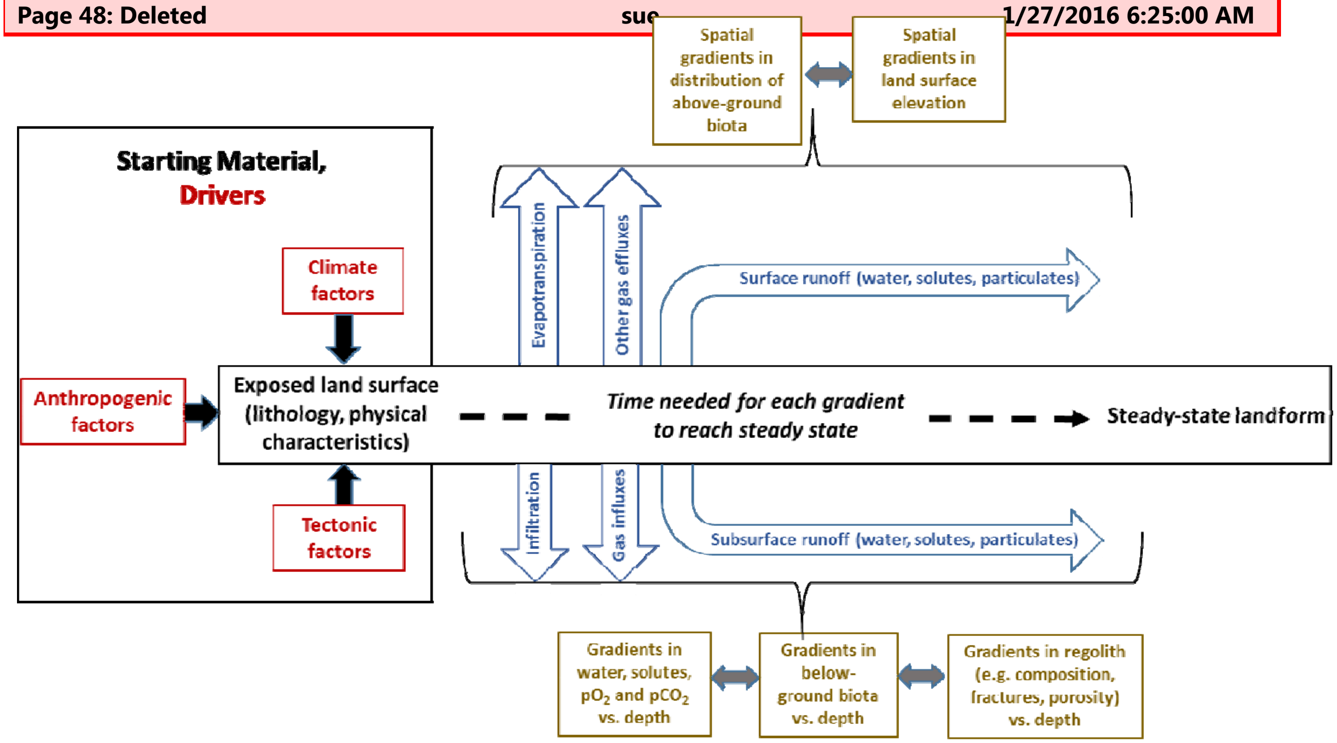



Blue = instantaneous WEGSS fluxes  Brown = gradients in measurable properties  Not shown: the many feedbacks in the system

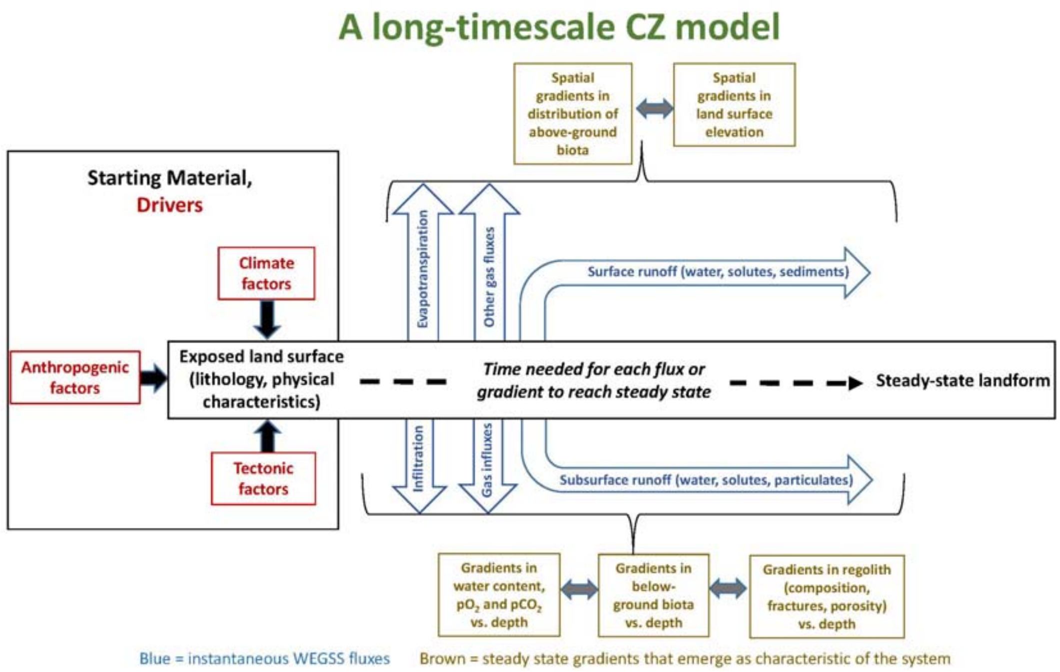
Page 48: Inserted  sue  1/27/2016 6:25:00 AM

A conceptual model of the CZ

Page 48: Deleted  sue  1/27/2016 6:25:00 AM




Blue = instantaneous WEGSS fluxes  Brown = gradients in measurable properties  Not shown: the many feedbacks in the system




| **Page 48: Deleted** | **sue** | **1/27/2016 5:45:00 AM** |
|---|---|---|

is aimed at understanding

| **Page 48: Inserted** | **sue** | **1/27/2016 5:55:00 AM** |
|---|---|---|

investigates

| **Page 48: Deleted** | **sue** | **1/27/2016 5:55:00 AM** |
|---|---|---|

| **Page 48: Deleted** | **sue** | **1/27/2016 6:05:00 AM** |
|---|---|---|

system

| **Page 48: Inserted** | **sue** | **1/27/2016 5:55:00 AM** |
|---|---|---|

from vegetation canopy to deep ground water

| **Page 48: Deleted** | **sue** | **1/27/2016 5:45:00 AM** |
|---|---|---|

at all different

| **Page 48: Inserted** | **sue** | **1/27/2016 5:56:00 AM** |
|---|---|---|

at all

| **Page 48: Deleted** | **sue** | **1/27/2016 5:56:00 AM** |
|---|---|---|

| **Page 48: Inserted** | **sue** | **1/27/2016 6:18:00 AM** |
|---|---|---|

As

**Page 48: Deleted** sue 1/27/2016 6:18:00 AM

As

**Page 48: Deleted** sue 1/27/2016 5:50:00 AM

set of

**Page 48: Deleted** sue 1/27/2016 5:51:00 AM

istics

**Page 48: Deleted** sue 1/27/2016 5:45:00 AM

rock

**Page 48: Inserted** sue 1/27/2016 5:45:00 AM

it

**Page 48: Inserted** sue 1/27/2016 5:57:00 AM

This transformation, shown in the black box, is catalyzed by biota (a feedback which is not shown explicitly).

**Page 48: Deleted** sue 1/27/2016 5:46:00 AM

All of the p

**Page 48: Inserted** sue 1/27/2016 5:46:00 AM

Gradients of properties describing the CZ are shown

**Page 48: Deleted** sue 1/27/2016 5:46:00 AM

roperties

**Page 48: Inserted** sue 1/27/2016 5:46:00 AM

brown

**Page 48: Deleted** sue 1/27/2016 6:23:00 AM

to the right

**Page 48: Inserted** sue 1/27/2016 5:57:00 AM

.

**Page 48: Inserted** sue 1/27/2016 5:58:00 AM

These gradients can

**Page 48: Deleted** sue 1/27/2016 5:46:00 AM

of the diagram can be considered properties which may sometimes

**Page 48: Inserted** sue 1/27/2016 5:47:00 AM

become time-independent (steady state)

**Page 48: Deleted** sue 1/27/2016 5:47:00 AM

reach a steady state

**Page 48: Inserted** sue 1/27/2016 5:47:00 AM

due to the many feedbacks which are not shown. Boxes are placed from left to right to note the increasing duration of exposure time needed to achieve such steady states

**Page 48: Deleted** sue 1/27/2016 5:48:00 AM

after increasingly long exposure times

**Page 48: Deleted** sue 1/27/2016 5:48:00 AM

In other words, after an initial transient period, these characteristics can reach dynamic equilibrium.

**Page 48: Deleted** sue 1/27/2016 6:07:00 AM

regolith thickness

**Page 48: Inserted** sue 1/27/2016 6:07:00 AM

depth profiles of regolith composition

**Page 48: Deleted** sue 1/27/2016 6:07:00 AM

equals the

**Page 48: Inserted** sue 1/27/2016 6:07:00 AM

=

**Page 48: Inserted** sue 1/27/2016 6:08:00 AM

in the presence of feedbacks related to pore water chemistry, soil gas composition, and grain size

**Page 48: Deleted** sue 1/27/2016 6:08:00 AM

Likewise, the nature and distribution of biota may become constant for some period. As emphasized by the

**Page 48: Inserted** sue 1/27/2016 6:02:00 AM

The

**Page 48: Deleted** sue 1/27/2016 6:02:00 AM

,

**Page 48: Inserted** sue 1/27/2016 6:02:00 AM

emphasizes that gradients to the left

**Page 48: Deleted** sue 1/27/2016 6:21:00 AM

ecosystems

**Page 48: Inserted** sue 1/27/2016 6:19:00 AM

can

**Page 48: Deleted** sue 1/27/2016 6:03:00 AM

are

**Page 48: Inserted** sue 1/27/2016 6:20:00 AM

achieve steady state

**Page 48: Deleted** sue 1/27/2016 6:20:00 AM

established

**Page 48: Deleted** sue 1/27/2016 6:20:00 AM

some geological changes

**Page 48: Inserted** sue 1/27/2016 6:08:00 AM

properties to the right. T

**Page 48: Deleted** sue 1/27/2016 6:20:00 AM

, and can t

**Page 48: Inserted** sue 1/27/2016 6:20:00 AM

properties to the left are

**Page 48: Deleted** sue 1/27/2016 6:20:00 AM

be

**Page 48: Deleted** sue 1/27/2016 6:03:00 AM

some of the other characteristics

**Page 48: Inserted** sue 1/27/2016 6:21:00 AM

properties in boxes to the right

**Page 48: Deleted** sue 1/27/2016 6:03:00 AM

in the diagram (e.g. regolith thickness and character, uplift rate, landscape curvature)

**Page 48: Inserted** sue 1/27/2016 6:24:00 AM

. The complexity of feedbacks (which are not shown for simplicity) can also create thresholds in system behavior

**Page 48: Deleted** sue 1/27/2016 6:04:00 AM

,

**Page 48: Inserted** sue 1/27/2016 6:04:00 AM

and

**Page 48: Deleted** sue 1/27/2016 6:04:00 AM

black indicates the system under study,

**Page 48: Inserted** sue 1/27/2016 6:22:00 AM

arrows are

**Page 48: Deleted** sue 1/27/2016 6:04:00 AM

indicates the

**Page 48: Inserted** sue 1/27/2016 7:08:00 AM

(up arrows for above-ground and down arrows for below-ground)

**Page 48: Deleted** sue 1/27/2016 6:04:00 AM

, and brown boxes indicate gradients

**Page 49: Deleted** sue 1/26/2016 5:45:00 AM

black

**Page 49: Inserted** sue 1/26/2016 5:45:00 AM

green

**Page 49: Inserted** sue 1/26/2016 5:46:00 AM

:

**Page 49: Deleted** sue 1/26/2016 5:46:00 AM

.

**Page 49: Inserted** sue 1/26/2016 5:46:00 AM

the subcatchment contains a dry oak-mixed hardwood community type (Fike, 1999) with an extremely diverse mix of hardwood and softwood species, including white oak (*Quercus alba*), sugar maple (*Acer saccharum*), pignut hickory (*Carya glabra*), eastern hemlock (*Tsuga canadensis*, and chestnut oak (*Q. montana*). The sparse understory consists of American hop-hornbeam (*Ostrya virginiana*) and serviceberry (*Amelanchier sp*.).

**Page 50: Inserted** sue 1/26/2016 5:45:00 AM

in the Shale Hills subcatchment

**Page 50: Inserted** sue 1/26/2016 5:45:00 AM

**Page 50: Formatted** sue 1/26/2016 5:45:00 AM

German (Germany)

**Page 52: Deleted** Dan K. Arthur 1/25/2016 9:47:00 AM

NLCD

**Page 52: Inserted** Dan K. Arthur 1/25/2016 9:47:00 AM

Homer et al.

**Page 53: Deleted** sue 1/27/2016 6:28:00 AM

are labelled from youngest to oldest: Smm (Rochester and McKenzie Members of the Mifflintown Formation), Smk (Keefer Member of the Mifflintown Formation), Srh (Rose Hill Formation), St (Tuscarora Formation), Oj (Juniata Formation).

**Page 53: Inserted** sue 1/27/2016 6:28:00 AM

include

**Page 53: Deleted** sue 1/27/2016 6:28:00 AM

is

**Page 53: Inserted** sue 1/27/2016 6:28:00 AM

(

**Page 53: Inserted** sue 1/27/2016 6:28:00 AM

)

**Page 53: Inserted** sue 1/27/2016 6:29:00 AM

Clinton group (including

**Page 53: Inserted** sue 1/27/2016 6:29:00 AM

formation)

**Page 53: Deleted** sue 1/27/2016 6:28:00 AM

are

**Page 53: Inserted** sue 1/27/2016 6:28:00 AM

(

**Page 53: Inserted** sue 1/27/2016 6:29:00 AM

)

| Page 53: Deleted | sue | 1/27/2016 6:29:00 AM |
|---|---|---|

is

| Page 53: Inserted | sue | 1/27/2016 6:29:00 AM |
|---|---|---|

(

| Page 53: Inserted | sue | 1/27/2016 6:29:00 AM |
|---|---|---|

)

| Page 53: Inserted | sue | 1/27/2016 6:30:00 AM |
|---|---|---|

(not shown)

| Page 53: Inserted | sue | 1/27/2016 6:29:00 AM |
|---|---|---|

in Garner Run subcatchment, i.e., the Tuscarora no longer outcrops upstream of this cross-section and Garner Run lies in the axis of Harry's valley

| Page 53: Deleted | sue | 1/27/2016 6:31:00 AM |
|---|---|---|

Nonetheless, t

| Page 53: Inserted | sue | 1/27/2016 6:31:00 AM |
|---|---|---|

T

| Page 54: Inserted | sue | 1/27/2016 8:02:00 AM |
|---|---|---|

The Harrys Valley well (HV1) is shown along with the location of the COSMOS unit and the outlet weir (blue dot to the southwest). The blue dot to the northeast indicates the approximate range of surface water sampling that is ongoing. Soil pits have been emplaced as shown, along with the Ground HOG deployment. Location of vegetation and GPR transects reported in this paper are also shown.

| Page 55: Deleted | sue | 1/27/2016 6:32:00 AM |
|---|---|---|

3

| Page 55: Inserted | sue | 1/27/2016 6:32:00 AM |
|---|---|---|

4

| Page 55: Deleted | sue | 1/27/2016 6:32:00 AM |
|---|---|---|

for bedrock geology map

| Page 58: Formatted | sue | 1/27/2016 6:33:00 AM |
|---|---|---|

Superscript

| Page 58: Formatted | Dan K. Arthur | 1/25/2016 11:26:00 AM |
|---|---|---|

Centered

**Page 59: Deleted** sue 1/19/2016 5:43:00 AM

Four plots

**Page 59: Inserted** sue 1/19/2016 5:43:00 AM

Plots

**Page 59: Inserted** sue 1/19/2016 5:47:00 AM

Garner Run sub-catchment

**Page 59: Inserted** sue 1/19/2016 5:47:00 AM

$\tau$

**Page 59: Deleted** sue 1/19/2016 5:47:00 AM

The normalized concentration

**Page 59: Deleted** sue 1/19/2016 5:47:00 AM

average

**Page 59: Inserted** sue 1/19/2016 5:47:00 AM

estimated as the average of

**Page 59: Deleted** sue 1/19/2016 5:47:00 AM

from five

**Page 59: Inserted** sue 1/27/2016 6:34:00 AM

5

**Page 59: Deleted** sue 1/19/2016 5:48:00 AM

and

**Page 59: Inserted** sue 1/19/2016 5:48:00 AM

based on the assumption that

**Page 59: Inserted** sue 1/27/2016 6:34:00 AM

derives from protolith and is

**Page 59: Deleted** sue 1/19/2016 5:48:00 AM

as the

**Page 59: Deleted** sue 1/19/2016 5:48:00 AM

element

**Page 59: Deleted** sue 1/18/2016 4:31:00 PM

One explanation for these plots is that Al has largely been removed or moved downward in the profile while Mg, K, and Fe have largely been added to the profile.

**Page 59: Inserted** sue 1/19/2016 5:48:00 AM

If parent is correctly estimated,

**Page 59: Deleted** sue 1/19/2016 5:48:00 AM

In these plots,

**Page 59: Deleted** sue 1/19/2016 5:48:00 AM

completely

**Page 59: Inserted** sue 1/19/2016 5:48:00 AM

100%

**Page 59: Deleted** sue 1/19/2016 5:49:00 AM

compared to Ti in the parent material

**Page 59: Inserted** sue 1/19/2016 5:49:00 AM

compared to Ti in the parent material

**Page 60: Inserted** Dan K. Arthur 1/25/2016 11:21:00 AM

Ground Penetrating Radar (GPR) transect of the Leading Ridge catena, showing inferred location of bedrock-soil interface (yellow dashed curve). The three soil pits (LRRT, LRMS, LRVF) are indicated by stars, with their observed depth to bedrock indicated by red arrow bar

**Page 61: Inserted** sue 1/27/2016 6:36:00 AM

. LRRT and LRMS were dug by hand until refusal and LRVF was dug by hand and deepened with a jack hammer

**Page 61: Deleted** sue 1/27/2016 6:36:00 AM

(note the observed depth was limited by excavation depth that was possible with the available digging tool)

**Page 61: Inserted** Dan K. Arthur 1/25/2016 11:21:00 AM

(note the observed depth was limited by excavation depth that was possible with the available digging tool). GPR data are exaggerated by 10x in vertical dimension as compared to surface topography. Summary values

**Page 61: Deleted** sue 1/27/2016 6:36:00 AM

values

| Page 61: Inserted | sue | 1/27/2016 6:36:00 AM |
|---|---|---|

bedrock depths

| Page 61: Inserted | Dan K. Arthur | 1/25/2016 11:21:00 AM |
|---|---|---|

are tabulated in Table 4

| Page 61: Inserted | sue | 1/27/2016 6:36:00 AM |
|---|---|---|

.

| Page 61: Inserted | Dan K. Arthur | 1/25/2016 11:21:00 AM |
|---|---|---|

from these GPR measurements.

| Page 61: Deleted | sue | 1/27/2016 6:36:00 AM |
|---|---|---|

from these GPR measurements.

| Page 61: Deleted | Dan K. Arthur | 1/25/2016 11:21:00 AM |
|---|---|---|

Ground Penetrating Radar (GPR) transect of Leading Ridge Catena, showing inferred location of bedrock-soil interface. GPR data is exaggerated by 4x compared to surface topography. Summary values are tabulated in Table 4 from these GPR measurements.

| Page 62: Deleted | sue | 1/27/2016 6:37:00 AM |
|---|---|---|

Nitrate

| Page 62: Inserted | sue | 1/27/2016 6:37:00 AM |
|---|---|---|

nitrate dissolved

| Page 62: Inserted | sue | 1/27/2016 7:58:00 AM |
|---|---|---|

(filtered at 0.45 µm)

| Page 62: Inserted | sue | 1/27/2016 6:37:00 AM |
|---|---|---|

as shown on Fig. 3

**Header and footer changes**

**Text Box changes**

**Header and footer text box changes**

**Footnote changes**

**Endnote changes**

# Designing a suite of measurements to understand the critical zone

**Susan L. Brantley[1], R. A. DiBiase[1], T. A. Russo[1], Y. Shi[2], H. Lin[2], Kenneth J. Davis[3], M. Kaye[2], L. Hill[2], J. Kaye[2], D. M. Eissenstat[2], B. Hoagland[1], A. L. Dere[1,a], A. L. Neal[4], Kristen M. Brubaker[5], Dan K. Arthur[1]**

[1]{Earth and Environmental Systems Institute, Department of Geosciences, Pennsylvania State University, PA, USA}

[2]{Department of Ecosystem Science and Management, Pennsylvania State University, PA, USA}

[3]{Earth and Environmental Systems Institute, Department of Meteorology, Pennsylvania State University, PA, USA}

[4]{Earth and Environmental Systems Institute, Pennsylvania State University, PA, USA}

[a]{now at: Department of Geography and Geology, University of Nebraska Omaha, Nebraska, USA}

[5] {Department of Environmental Studies, Hobart and William Smith Colleges, Geneva, NY, USA}

Correspondence to: Susan L. Brantley (sxb7@psu.edu)

## Abstract

Many scientists have begun to refer to the earth surface environment from the upper canopy to the depths of bedrock as the critical zone (CZ). Identification of the CZ as an integral ~~worthy~~ object worthy of study implicitly posits that the study of the whole earth surface will provide benefits that do not arise when studying the individual parts. To study the CZ, however, requires prioritizing among the measurements that can be made -- and we do not generally agree on the priorities. Currently, the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO) is expanding from a small original ~~study~~ focus area (0.08 $km^2$, Shale Hills catchment), to a ~~much~~ larger watershed (164 $km^2$, Shavers Creek watershed) and is grappling with the ~~necessity of~~ prioritization. This effort is an expansion from a monolithologic first-order forested catchment to a watershed that encompasses several lithologies (shale, sandstone, limestone) and land use types (forest, agriculture). The goal of the project remains the same: to understand water, energy, gas, solute and sediment (WEGSS) fluxes that are occurring today in the context of the record of those fluxes over geologic time as recorded in soil profiles, the sedimentary record, and landscape morphology.

Given the small size of the ~~original~~ Shale Hills catchment, the original ~~measurement~~ design ~~resulted in~~ incorporated measurement of as many parameters as possible at high temporal and spatial density. In the larger Shavers Creek watershed, however, we must focus the measurements. We describe a strategy of data collection and modelling based on a geomorphological and land use framework that builds on the hillslope as the basic unit. Interpolation and extrapolation beyond specific sites relies on geophysical surveying, remote sensing, geomorphic analysis, the study of natural integrators such as streams, ground waters or air, and application of a suite of CZ models. ~~In essence, w~~We are hypothesizing that ~~pinpointed~~ measurements of a few important variables at strategic locations within a geomorphological framework will allow development of predictive models of CZ behavior. In turn, the measurements and models will reveal how the larger watershed will respond to perturbations both now and into the future.

## 1 Introduction

The critical zone (CZ) is changing due to human impacts over large regions of the globe at rates that are geologically significant (~~Crutzen, 2002;~~ Vitousek et al., 1997a; Vitousek et al., 1997b; Crutzen, 2002; Wilkinson and McElroy, 2007). To maintain a sustainable environment requires that we learn to project the future of the CZ. Models are therefore needed that accurately describe CZ processes and that can be used to project, or "earthcast," the future using scenarios of human behavior. At present we cannot earthcast all the properties of the CZ ~~even using scenarios of human behavior~~, but rather must ~~run~~ model~~s~~ ~~to project~~ individual processes (Godderis and Brantley, 2014). Even so, many of our models are inadequate to make successful estimates ~~projections~~ of first-order CZ behavior today, let alone projections for tomorrow. For example, we cannot *a priori* ~~predict~~ estimate ~~the~~ streamflow ~~in a catchment~~ even if we know the climate conditions, soil properties, and vegetation in a given catchment, because ~~we are often uncertain~~of difficulties characterizing how much water is lost to evapotranspiration and to groundwater (Beven, 2011). Likewise, we cannot *a priori* ~~predict~~ estimate the depth or chemistry of regolith on a hillslope even if we know its lithology and tectonic and climatic history, because we do not ~~fully~~ adequately understand what controls the rates of regolith formation and transport (Amundson, 2004; Brantley and Lebedeva, 2011; Dietrich et al., 2003; Minasny et al., 2008). Perhaps even more unexpectedly, we often do not even agree upon which minimum measurements are needed to answer these questions at any location.

Such difficulties are largely due to two factors: i) we cannot adequately quantify spatial heterogeneities and temporal variations in the reservoirs and fluxes of water, energy, gas, solutes, and sediment (WEGSS); and ii) we do not adequately understand the interactions and feedbacks among chemical, physical, and biological processes in the CZ that control these fluxes. This latter problem ~~means~~ reflects the fact that the CZ (Fig. 1) is characterized by tight coupling between chemical, physical, and biological processes which exert both positive and negative feedbacks on surface processes. Modelling the CZ is fraught with problems precisely because of these feedbacks and because the presence of thresholds means that extrapolation from sparse measurements ~~can be~~ is challenging (Chadwick and Chorover, 2001; Ewing et al., 2006).

~~The~~ However, the result of these couplings and feedbacks is that patterns ~~with respect to the~~of measureable properties emerge during evolution of ~~disparate~~ Critical Zone systems

**Formatted:** Indent: First line: 0.5"

**Formatted:** Indent: First line: 0.49"

that are repeated from site to site despite variations in environmental conditions. ~~reveal aspects of the underlying complex behavior~~Such patterns~~— properties~~ include ~~such as~~ the distributions ~~of regolith~~ across landscapes or versus depth of such observables as regolith, fractures, bacterial species, or ~~the distribution of soil gas~~gas composition~~ versus depth~~. ~~At the most general level, g~~Gradients in some ~~of these~~ important observable properties (e.g., surface slope, ~~temperature,~~ chemistry of water and~~,~~ regolith) ~~drive~~ emerge as indicators of the evolution of the CZ and reveal aspects of the underlying complex behavior (brown boxes, Fig. 1). For systems ~~characterized by~~experiencing negative feedbacks, such gradients are thought to move toward steady-state ~~forms determined by the operative feedbacks.~~conditions, i.e., gradients that remain constant over some interval of time.

In Fig. 1, some of these important gradients are arrayed from left to right to indicate the increasing length of time it takes for each gradient in general to achieve such a steady state. In other words, a steady-state soil gas depth profile might develop more rapidly than a steady-state regolith chemistry depth profile. ~~Scientists from d~~Different disciplines tend to focus on different emergent properties (different gradients), and thus tend to ~~think about~~ emphasize processes operating at disparate timescales. However, CZ science is built upon the hypothesis that an investigation of the entire object – the CZ – across all timescales under transient and steady-state conditions (Fig. 1) will yield insights that disciplinary-specific investigations cannot. In turn, such integrative study and modelling should ~~should~~ allow deeper understanding of the patterns that ~~emerge~~ characterize ~~in~~ the CZ.

Given that the ~~Building a model for the CZ is a challenging task given these vast timescales: driving~~ mechanisms driving CZ change range from tectonic forcing over millions of years, to glacial-interglacial climate change over thousands of years, to the recent influence of humans on the landscape, building a model of the CZ is daunting~~.~~ and no single model has been developed. Instead, ~~One way to do this is to use a~~ suites or cascade of simulation models have been used to ~~that both~~ address ~~specific~~important processes over different ~~and~~ timescales (e.g., Godderis and Brantley, 2014). ~~and that can be interconnected to address broad overarching CZ problems as shown in Table 1 (Duffy et al., 2014; Table 1).~~ To enable treatment using such a suite of models, ~~Then, f~~F~~For~~ each setting for CZ research including CZ observatories (CZOs (White et al., 2015))~~y~~ must~~then, it is necessary to~~ grapple with the necessity of measuring ~~with~~ the processes at different timescales to understand the dynamics and evolution of the system.

At the Susquehanna Shale Hills ~~Critical Zone Observatory~~ CZO (SSHCZO), we have been investigating this challenge by studying the CZ in a 0.08 km$^2$ watershed located in central Pennsylvania (the Shale Hills catchment, Fig. 2).  At the same time, we have been developing a suite of models that can be interconnected to address broad overarching CZ problems (Duffy et al., 2014; Table 1). The focus of the effort has been the small Shale Hills catchment~~,~~ which was established for hydrologic research in the 1970s (Lynch, 1976) and was expanded with other disciplinary studies as a CZO in 2007 as part of a network of CZOs in the U.S.A~~, has been a successful location for CZ research~~. The small spatial scale of Shale Hills ~~has~~ allowed the development of a diverse but dense monitoring network that spans disciplines from meteorology to groundwater chemistry to landscape evolution (Fig. 2). Given the small size, we referred to our measurement paradigm as "measure everything, everywhere". For example, we inventoried all of the ~2000 trees with diameter greater than 20 cm at breast height, drilled 28 wells (up to ~~XXX~~50 m deep), sampled soil porewaters at 13 locations at multiple depths approximately every other week during the non-snow covered seasons for more than a year, and measured soil moisture at 105 locations (Fig. 2).

~~The Shale Hills catchment is also situated on a single lithology (shale), which reduced the complexity of~~~~simplified the boundary conditions for models with respect to the initial chemical and physical conditions.~~

~~–~~The approach at Shale Hills has been to develop understanding incrementally by studying CZ systems of ~~varying~~ increasing complexity. The catchment itself is situated on a single lithology (shale), which simplified the boundary conditions for models with respect to initial chemical and physical conditions.  W~~, w~~e have monitored at ridgetops (where water and soil transport is approximately one-dimensional (1D)), along ~~catenas~~ planar hillslopes (transects where such transport is essentially 2D), and within swales and ~~integrated across~~ the full catchment (where transport ~~is~~ must be considered in full 3D). Where possible, t~~T~~hese observations ~~were~~ have then been paired with 1D, 2D and 3D model simulations. ~~The characterization and modelling of settings as~~ Using the conceptualization of "1D, 2D, ~~or~~ and 3D" settings in the catchment has allowed measurements and modelling to proceed in a synergistic fashion: the reduction of ~~has allowed us to reduce some of the~~ complexity in 1D and 2D sites ~~of~~ enabled~~the~~ development of model~~s systems,~~ ~~and~~ but also ~~to~~ focused our sampling schemes. ~~This formulation also led to recognition of the importance of two types of hillslopes: planar hillslopes that experience downslope but nonconvergent flow of water and~~

**Formatted:** Indent: First line: 0.49"

soil (essentially 2D evolution), and swales that experience downslope convergent flow of water and soil (3D evolution). Much of our effort focused on understanding soils and waters in these two hydrologic units (Fig. 2). For example, our model conceptualizationss of soil formation have beenwere developed strictly first for ridgetops (1D) or and then for planar (2D) hillslope systems and have been highly influenced by our soil chemistry measurements on ridgetops and planar hillslope catenas (Jin et al., 2010; Lebedeva and Brantley, 2013; West et al., 2013; Ma et al. 2013). In some cases modelling and measurement proceed hand in hand while in others, the modelling lags. For example, soil measurements have been collected in hillslopes characterized by convergent water and soil flow regimes, i.e., swales (Jin and Brantley, 2011) and soil observations have been collected across much of the catchment, but soil formation models for swales or the entire catchment still remain to be developed.

In contrast to the soil formation models that have targeted the 1D and 2D sites, while our models of water flow have been developed for the entire catchment are heavily influenced by the convergent swales (e.g., Qu and Duffy, 2007). In fact, study of an entire catchment with a hydrologic model is sometimes more tractable than for smaller sub-systems because the large-scale study allows a continuum treatment whereas treatment of smaller scale sub-systems within the catchment might require measurements of the exact positions of heterogeneities such as fractures, faults, low-permeability zones, etc.

**Formatted:** Indent: First line: 0.49", Line spacing: 1.5 lines

The goal of the SSHCZO project now is to upscale grapple with some of these down- and up-scaling issues by expanding the CZO from Shale Hills to the encompassing 164 km$^2$ Shavers Creek watershed (Fig. 3). The expansion was designed to allow us to investigation ofe a broader range of lithologies (sandstone, calcareous shale, minor limestone) and land use (agriculture, managed forest, minor development), and to test our developing models at larger spatial scales. For our approachTo enable understanding of the larger watershed, we chose to analyze a suite of smaller subcatchments in detail, each of which were selected to be the largest that still drain a single rock unit or land--use type. Thus, we can directly evaluateis allows evaluation of how much of our understanding from the Shale Hills site is transferable to other lithologies with different boundary initial conditions but with the same climate in isolation. Additionally, we are making targeted measurements of the mainstem of the stream in nested catchments of differing size within the larger watershed, in order to upscale our site-specific models to a relatively complex watershed.

**Formatted:** Indent: First line: 0.49"

Despite its small size, Shavers Creek contains much of the variability in CZ parameter space found within the Susquehanna River Basin~~,~~ and the Appalachian Valley and Ridge province in general. By measuring in detail paired catchments of similar size but different underlying conditions, along with targeted measurements in nested catchments of differing size, we aim to test theories of CZ evolution, parameterize models (~~for example, in ~~Table 1) in different settings, and explore approaches toward upscaling ~~over ~~across different size watersheds.

~~By necessity, t~~To understand the interaction of WEGSS fluxes in Shavers Creek~~,~~ ~~including the~~ and its smaller subcatchments, it is necessary to ~~we must ~~move beyond the paradigm of measuring "everything, everywhere" (Fig. 2) to an approach of measuring "only what is needed". This phrasing, although simplistic, should resonate with any field scientist: the choice of measurement design is at the heart of any field project. But when we study the CZ as a whole, we are asking, how does one allocate resources to measure and model the dynamics and evolution of the entire CZ system? This paper describes our philosophy of measurement in the CZO; ~~and ~~our previous paper describes the modelling approach (Duffy et al., 2014). Obviously, due to the wide range of CZ processes across environmental gradients (Fig. 1), the specifics of ~~an ~~our ~~ideal ~~proposed sampling design will differ from such designs at other s~~vary from s~~ites~~ to site~~. ~~Nonetheless w~~We nonetheless describe the philosophy behind our approach ~~as a way to hypothesize an answer to the~~to stimulate focus on the broad question: how can we adequately and efficiently measure the entire CZ to best learn about its evolution and function? To exemplify our design, w~~W~~e ~~then ~~also ~~present specific examples for ~~describe the first part of our expansion from Shale Hills to a sandstone subcatchment within Shavers Creek.

Formatted: Indent: First line: 0.5"

## 2 ~~Building a co~~Connections between model development and field measurements ~~Rationale for the measurement plan~~

The suite of models shown in Table 1 is designed to develop understanding over the entire CZ as an integral object of study, i.e., one system. ~~Ideally, the connection between model development and field measurements is two-directional~~F~~; f~~ield measurements are prioritized and driven by data needs for developing models (e.g., Table 1)~~,~~ and ~~at the same time ~~model development is dicta~~e~~ted by observations in the field. Hand in hand with this system-level approach, researchers from different disciplines also bring discipline-specific

Formatted: Indent: First line: 0.5"

hypotheses to their research that are related to disciplinary gaps in knowledge. Thus, disciplinary-level hypotheses also drive CZO research and sometimes these hypotheses feed directly into the overall CZ suite of models. Furthermore, because our understanding of the complicated suite of CZ processes is still in its infancy, both baseline measurements and curiosity-driven sample collection are still vital to determine the important processes. Throughout, models and observations are allowed to evolve to enable the two-way exchange of insights needed to maximize CZ science.

~~The choices of measurements to be made during the expansion are driven by data needs for the models under development (Table 1), at the same time that the models are driven by observations in the field. The suite of models shown in Table 1 is one way to~~~~is aimed at~~ ~~understanding the entire CZ as an object of study, rather than as a set of disparate systems. In coordination with this modelling approach~~~~Similarly~~Given all the needs for data, ~~a~~ the ~~stratified~~ sampling plan which is ~~is being~~ implemented in ~~the~~ a CZO ~~target~~must provide boths ~~focused~~ measurements ~~for testing~~ to test disciplinary hypotheses ~~(e.g., eddy flux covariance, soil gas composition)~~ and observations necessary to bridge ~~that are central~~ across disciplines ~~(e.g., soil thickness, topography)~~. Additionally, certain measurements such as geophysical and remote sensing surveys, catchment-integrating stream measurements, and time-integrating analysis of alluvial and colluvial sediments can be made along with~~, and~~ model simulations ~~will be used~~ to ~~and paired with geophysical, catchment-scale stream and remote sensing measurements. The models will then be used to~~ upscale across space (from limited point or subregion measurements to the whole watershed) and time (from limited temporal measurements to ~~over longer~~geological timescales).

Perhaps the largest difficulty in *spatially* characterizing the CZ in any observatory is the assessment of the extremely heterogeneous land surface, ~~including~~ ranging from the assessment of regolith and pore fluids down to bedrock to variations in land use. Because the mixing timescales of biota, regolith, and bedrock are relatively slow (compared to mixing of atmospheric and surface water reservoirs), the assessment of the spatial distribution of biota, regolith, and bedrock properties is both important and extremely challenging (Niu et al., 2014). ~~In other words, while assessment of atmospheric and surface water pools can be technically challenging, mixing of these pools is much faster than mixing of the biotic pool, the regolith and rock reservoir, or the pool of soil porewater, making assessment of the spatial distribution of these latter reservoirs exceedingly difficult (Niu et al., 2014).~~ On the other

**Formatted:** Not Highlight

**Formatted:** Indent: First line: 0.49"

hand, ~~the~~ rapid ~~rates of~~ changes ~~in the land surface pools is generally much slower than the changing~~ in the atmospheric reservoir~~, making~~ makes robust atmospheric ~~sensor requirements~~measurements technically difficult. The hydrologic state is intermediate, exhibiting large spatial and temporal variability. ~~Thus, the largest difficulty in *temporally* characterizing today's fluxes in the CZ in any observatory may be measuring the fast-changing fluxes of the~~~~changes in~~ ~~atmospheric pools and fluxes.~~

In recognition of these difficulties, the project started at Shale Hills precisely because it is a catchment almost 100% underlain by Rose Hill formation shale and strictly managed as forestland. ~~In recognition of these difficulties, the project started at Shale Hills precisely because it is a catchment almost 100% underlain by Rose Hill formation shale with land use strictly as managed forest.~~ Surface heterogeneities at Shale Hills were largely related to hillslope position, colluvium related to the Last Glacial Maximum (LGM), fracturing, differences in sedimentary layers, and relatively limited spatial variations in vegetation. To understand the CZ at the Shavers creek watershed, on the other hand, we must grapple with a more complex set of variations related to differences in lithology, land use, climate change, and landscape adjustment to changes in base level due to tectonics, eustasy or stream capture (Fig. ~~1~~3). Here, the term base level ~~is used to~~ refers to the reference level or elevation down to which the watershed is currently being ~~graded~~eroded.

In recognition of the new complexities within Shavers creek, the sampling strategy was designed not to be random but rather to be ~~using a stratified sampling plan~~ stratified based on geological and geomorphological knowledge~~ rather than random sampling~~. An implicit hypothesis underlying this approach is the idea that sampling can be more limited ~~when it is designed as~~ for a stratified approach based on geological (especially geomorphological) knowledge~~)~~. For example, a first-order observation about hillslope morphology in Shale Hills based in long-standing observations from hillslope geomorphology is the delineation between planar slopes and swales: the former experience largely 2D nonconvergent flow while the latter experience 3D convergent flow of water and soil. ~~where~~ Where many ~~many~~ randomly chosen soil pits might be necessary if the delineation of swales versus planar hillslopes was ~~not recognized~~ignored, ~~if~~ when ~~these two features are recognized and~~ representative pits are dug to investigate these features separately, the number of pits can be minimized. ~~Furthermore, one of the models under development for the CZO is a regolith~~

~~formation model: by using this model to understand regolith formation, the number of pits in regolith can similarly be minimized.~~

Another aspect of our stratified sampling plan is to complement m~~M~~easurements at Shale Hills ~~will soon be supplemented with~~ by targeted ~~instrumentation~~ measurements in ~~the~~ two new subcatchments of Shavers creek ~~, watershed. The subcatchments were~~ chosen to represent two of the new lithologies in the watershed. Once again the stratification of the sampling design is dictated by geological knowledge: b~~, as b~~edrock geology is known to exert a first-order control on WEGSS fluxes in the CZ (e.g. Duvall et al., 2004; Williard et al., 2005). The first such new subcatchment is ~~a new~~ forested and ~~subcatchment~~ underlain only by sandstone. The second subcatchment for targetted measurements is currently being identified on calcareous shale. This second subcatchment will also host several farms and will allow assessment of the effects of this land use on WEGSS fluxes.

To upscale from subcatchments to Shavers Creek, the target~~t~~ed subcatchment data will be amplified by measurements of chemistry and streamflow along the mainstem of Shavers Creek as well as catchment-wide meteorological measurements ~~to upscale from Shale Hills to Shavers Creek~~ (Fig. 3). The upscaling will rely on ~~only a~~the small number of sites chosen for soil, vegetation, pore fluid, and soil gas measurements in each subcatchment. To extrapolate from and interpolate between these limited land surface measurements, models of landscape evolution (LE-PIHM), soil development (e.g., Regolith-RT-PIHM, WITCH), distribution of biota (BIOME4, CARAIB), C and N cycling (Flux-PIHM-BGC), sediment fluxes (PIHM-SED), solute fluxes (RT-Flux-PIHM, WITCH), soil gases (CARAIB), and energy and hydrologic fluxes (PIHM, Flux-PIHM) will be used. In effect, the plan is to substitute "everything everywhere" with measurements of "only what is needed" by using i) integrative measurements (geophysics, lidar, stream, atmosphere), and ii) models of the CZ. As a simple example, a regolith formation model is under development that will predict distributions of soil thickness on a given lithology under a set of boundary conditions. Since much of the water flowing ~~through~~ through the~~se~~ upland catchments under study in the CZO ~~small catchments~~ flows as interflow through the soil and upper fractured zone (Sullivan et al., subm.), use of the regolith formation model ~~is necessary to~~ will enable better predictions of the distribution of permeability ~~in the catchment~~. Of course, t~~T~~he models will be continually groundtruthed ~~with~~ against pinpointed field measurements. With this approach, water fluxes in the subcatchments and in Shavers creek watershed itself will eventually be estimated.

For clarity in describing the measurements i

~~I~~n each subcatchment that are needed for the models, we have given names to arrays of instruments (Table S1) ~~for clarity of description~~. The array of instruments in soil pits (1 m x 1 m x ~2 m deep) and in trees near the pits along a catena is referred to as "ground hydrological observation gear" (Ground HOG). The Ground HOG deployments also are the locations for assessments of v~~V~~egetation ~~is being assessed~~ a~~t~~cross transects ~~located coincident with the Ground HOG~~. Geophysical surveys and geomorphic analysis using lidar -are ~~being~~ conducted to interpolate between or extrapolate beyond the catenas.

In addition to Ground HOG, the energy, water, and carbon fluxes are ~~being~~ measured using "tower hydrologic observation gear" (Tower HOG). Ground and Tower HOGs are in turn accompanied by measurements of stream flow, chemistry and temperature, groundwater levels and chemistry. As discussed above, t~~T~~hese streams and ground waters provide ~~are~~ natural spatial and temporal ~~integrators~~ integrations over the watershed and therefore provide constraints on the 3D-upscaled models.

**Formatted:** Indent: First line: 0.49"

Data from Ground HOG and Tower HOG ~~Stream and ground water data~~ will be used to parameterize and constrain model-data comparison and data assimilation~~, as described below~~.

**Formatted:** Indent: First line: 0.5"

**Formatted:** Not Highlight

## ~~3 Data assimilation~~

In fact, t~~T~~he choice of targeted measurements are derived at least in part from an observational system simulation experiment (OSSE) completed for the Shale Hills catchment using the Flux-PIHM model (Table 1) (Shi et al., ~~2014b~~2015b). The OSSE evaluates how well a given observational array describes the state variables that are targeted by Flux-PIHM. Specifically, this OSSE (Shi et al, ~~2014b~~2015b) emphasized water and energy fluxes for the catchment.

**Formatted:** Indent: First line: 0.5"

Prior to the OSSE, a sensitivity analysis was performed (Shi et al. ~~2014a~~2015a) to determine the six most influential model parameters that were needed to constrain and produce a successful simulation. We defined "successful simulation" as one that reproduced the temporal variations of the four land surface-hydrologic fluxes (stream discharge, sensible heat flux, latent heat flux, and canopy transpiration), and the three state variables (soil

**Formatted:** Indent: First line: 0.49"

moisture, water table depth, and surface brightness temperature) (Table 1) with high correlation coefficients and small root mean square errors. Once the six most influential model parameters were determined -- porosity, van Genuchten alpha and beta, Zilitinkevich parameters, minimum stomatal resistance, and canopy water storage -- the OSSE was then performed.

The OSSE evaluated which of the fluxes and state variables were most important in constraining those model parameters. Shi et al. (~~2014b~~2015b) found that the calibration coefficients for the most important model parameters were most sensitive to observations of i) stream discharge, ii) soil moisture, and iii) surface brightness temperature. (Alternately, instead of brightness temperature, measurements could focus on sensible and latent heat fluxes.) The OSSE has also been validated with assimilation of field observations at Shale Hills (Shi et al., 2015b).

On the basis of this OSSE, we are targeting measurement of stream discharge, soil moisture, and surface brightness temperature for each of the SSHCZO subcatchments on shale, sandstone, and calcareous shale. These measurements should allow us to reproduce subcatchment-averaged land-atmosphere fluxes and subsurface hydrology adequately. Once the three subcatchments are parameterized, the models will then be upscaled to the entire Shavers Creek watershed using information from lidar, SSURGO, geological maps, geophysical surveying, and land use.

Currently, the OSSE has only been used for assimilation of water and energy data but is being expanded to include biogeochemical variables. In other words, our ultimate aim is to complete an OSSE for C and N fluxes in each subcatchment. In the long run, we could also extend the OSSE to assimilate data for other solutes and for sediments.

Modeling results from Shale Hills indicated that an accurate simulation of the sub-catchment spatial patterns in soil moisture were achieved using a relatively limited set of hydrologic measurements made at a few points (Shi et al., 2015a). Specifically, we had to measure i) stream discharge at the outlet, ii) soil moisture at a few locations, and iii) groundwater levels at a few locations. The soil moisture (ii) and groundwater (ii) data used to calibrate the model were from 3 nearly co-located sites in the valley floor. These sites (referred to as RTHnet on Fig. 2) were the only sites with continuous data at the time of model calibration (~~. Notably, ~~COSMOS data were not yet available). The measurements were averaged across the three RTHnet sites (see data posted at http://criticalzone.org/shale-hills/data/dataset/3615/) to provide one calibration point in the model. Extending from this calibration point to the entire catchment was attempted using data from the SSURGO database (http://www.nrcs.usda.gov/wps/portal/nrcs/main/soils/survey/). However, because of the coarseness of spatial data available in SSURGO, this was not successful for the very small Shale Hills catchment. Therefore, porosity, horizontal and vertical saturated hydraulic conductivity, and the van Genuchten parameters $\alpha$ and $\beta$ were separately measured for each soil series and then were averaged for the whole soil column for each soil series (Supplement Table S2). These soil core measurements for each soil series were used to constrain the shape of the soil water retention curve for each soil series in the catchment in the model.

The result of this effort was that for the monolithologic 0.08 $km^2$ catchment of Shale Hills, five soil series were identified and soil properties measured (Lin et al., 2006). As we proceed with work on the new subcatchments, one of two approaches will be used. First, it is possible that relatively few soil moisture measurement locations are required in any given catchment, as long as we can obtain soil hydraulic properties for each soil series. Using the SSURGO soils database, such measurements could be made to parameterize the model. Alternately, spatially extensive soil moisture measurements based on COSMOS may be adequate to infer the variations in soil hydraulic properties on a series-by-series basis or based on geomorphological criteria. The overall plan is to use i) SSURGO, ii) geomorphological constraints, iii) COSMOS, and iv) soil moisture measurements along the catenas to parameterize Flux-PIHM.

To the extent possible, we parameterize these 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, 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. Data assimilation can thus provide optimal estimates of observable variables and parameters, taking into account both the uncertainties of model predictions and observations.

**Formatted:** Not Highlight

As new types of observations are provided, we first evaluate PIHM model output against the new observations prior to calibrations to see if the current calibration predicts the new data.

**Formatted:** Not Highlight

**Formatted:** Line spacing: 1.5 lines

This comparison is ongoing for the Garner Run subcatchment. If the prediction is poor, this yields insight into the capabilities of our model under new conditions. If we discover that even with a new calibration we cannot successfully predict the new observations, we will incorporate a new module that describes a new phenomenon in PIHM. For example, discrepancies between model output and preliminary observations at Garner Run has led us to hypothesize that the distribution of boulders on the land surface – a phenomenon not observed in the Shale Hills catchment – must be incorporated into the PIHM models. 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.

## ~~4~~3 Implementation in the Garner Run subcatchment

These discussions about the design of a sampling strategy can best be explained through examples. In this section we introduce the Garner Run subcatchment, one of the two new focus ~~areas~~ subcatchments planned within the Shavers Creek watershed. To exemplify the approach, w~~In addition to describing the geologic and geomorphologic setting, we detail the sampling strategy.~~e describe the setting and some p~~P~~reliminary observations and measurements from soil pits, vegetation surveys, and ~~surface~~ water monitoring~~ are also presented~~.

Formatted: Indent: First line: 0.5"

### ~~4.1~~3.1 Geologic, ~~and~~ geomorphic, and land use context of Garner Run

A central underlying hypothesis of ~~the~~ SSHCZO work is that the use of geomorphological and land use analysis can inform ~~the~~ sampling strategy so that measurements can be limited in number. Therefore, we start by describing ~~include a description of~~ the current knowledge of the geomorphological setting of the ~~new~~ Garner Run subcatchment and land use~~ at Garner Run~~. The subcatchment drains a synclinal valley underlain by the Silurian Tuscarora Formation between the NW-SE trending ridges of Tussey Mountain and Leading Ridge (Figs. 3-5). The Tuscarora Formation, which locally consists of nearly pure ~~quartzite~~ sandstone with minor interbedded shales, is the ridge-forming unit that caps the highest topography in ~~the~~ Shavers Creek watershed. The hillslopes of both Tussey Mountain and Leading Ridge are nearly dip slopes, i.e, the roughly planar hillslopes ~~are~~

Formatted: Indent: First line: 0.4"

~~roughly planar and~~ parallel the bedding in the sandstone (Fig. 4, 5). Indeed, subtle bedding planes can be observed in lidar-derived elevation data (Fig. 6B). The strong lithologic control on landscape form is manifested clearly in the high-resolution (1 m) bare-earth lidar topography.

~~One question that must be addressed is whether the Rose Hill Shale of the Clinton Group, which underlies Shale Hills and lies stratigraphically above the Tuscarora, may be present in the Garner Run subcatchment. Down-valley of the Garner Run study area, the Rose Hill Shale has been mapped in a low-sloping bench at the foot of Tussey Mountain (Figs. 3,4, Flueckinger, 1969). Although the entirety of the Garner Run study area is mapped as the Tuscarora Formation, the continuation of a low-sloping bench along the entire valley (Fig. 7) could be consistent with the presence of Rose Hill Shale throughout the catchment. In general, bedrock exposure is poor in the Shavers Creek watershed, but lidar topographic analysis, field mapping, and targeted geophysical surveys will aid in resolving uncertainties in subsurface composition necessary for modeling water, solute, and sediment fluxes.~~

The hillslope morphology of the Garner Run subcatchment also contrasts strikingly in several ways from ~~shows a number of striking contrasts to~~ that of ~~the~~ Shale Hills ~~catchment~~. Most notably, the sandstone hillslopes of Tussey Mountain and Leading Ridge are nearly planar in map-view: they have not been dissected with the streams and swales common in the shale topography of much of Shavers Creek ~~–~~ (Fig. 6)~~(including Shale Hills)~~. Hillslopes underlain by the Tuscarora Formation are also nearly 10X longer (300-600 m) than those underlain by other geologic units within Shavers Creek ~~(Fig. 6)~~, including shales. In Shale Hills, for example, hillslopes are 50-100 m in length (Fig. 7). Furthermore, the hillslopes at Garner Run are less steep (mean slope = 12°-17°) compared to those at Shale Hills (mean slope = 14°-21°), despite having significantly stronger underlying bedrock ~~(quartzite vs. shale)~~.

~~This~~ The observation of steeper hillslopes in Shale Hills versus Garner Run is particularly curious given ~~their presumably~~that both subcatchments are presumed to have experienced similar histories of climate and tectonism. If the two landscapes were in a topographic steady-state with local erosion rate equal to the same regional rock uplift rate, we would expect that ~~Garner Run~~the sandstone ~~to~~ would have evolved ~~with time~~ to ~~have~~ generate steeper slopes. Thus the shallower slopes on the resistant sandstone contradicts the general

**Formatted:** Indent: First line: 0.49"

idea that ~~in order to erode~~erosion and transport ~~its~~ of a more resistant bedrock ~~and~~ that produces a larger grain-size ~~coarser~~ sediment generally requires steeper hillslopes.

~~However, t~~Two issues may explain this apparent contradiction~~complicate such an interpretation~~. First, while the morphology of the Shale Hills catchment bears little resemblance to the underlying structure of steeply-dipping shale beds, the topography of Garner Run is nearly entirely controlled by underlying Paleozoic structure (Fig. 4). Specifically, ~~Paleozoic structures, with~~ hillslope angles reflect~~ing~~ dip-slopes rather than a morphodynamic equilibrium. Second, as a headwater stream in ~~its location in the headwaters of~~ the Shavers Creek watershed, Garner Run i~~has~~s isolated ~~the hillslopes of Garner Run~~ from the regional baselevel controls that influence ~~the Shale Hills and other~~ downstream catchments such as Shale Hills (Fig. 6).~~This hillslope conundrum could be related to the role of local base level and transient landscape adjustment (Whipple et al., 2013).~~

Specifically, analysis of stream longitudinal profiles on Garner Run and the mainstem of Shavers Creek reveals prominent knickpoints at elevations of ~320 m and 380 m, respectively (Fig. 7). Such breaks in channel slope geomorphically insulate the upper stream reaches from the mainstem of Shavers Creek and could be consistent with different rates ~~of base level fall upstream and downstream of the knickpoint. Equivalently, the knickpoints could delineate different rates~~ of local river incision into bedrock in the upper and lower reaches (e.g., Whipple et al., 2013). Published cosmogenic nuclide-derived bedrock lowering rates ranging from 5-10 m/Myr from similar nearby watersheds (Miller et al., 2013; Portenga et al., 2013) may be a good estimate for rates in Garner Run upstream of the knickpoint (Fig. 7). These rates are indeed 3-4 times lower than bedrock lowering rates inferred for the Shale Hills catchment (20-40 m/Myr) (Ma et al., 2013; West et al., 2014; West et al., 2013), which lies downstream of the knickpoint on Shavers Creek ~~(Ma et al., 2013; West et al., 2014; West et al., 2013)~~.

The origin and genesis of these knickpoints is likely due to some combination of the following: regional baselevel adjustment on the Susquehanna River since the Neogene (3.5-15 Ma) due to epei~~e~~rogenic uplift (Miller et al., 2013), stream capture and drainage reorganization (e.g. Willett et al., 2014), or temporal and spatial variations in bedrock exposure at the surface (e.g. Cook et al., 2009). Testing these competing controls will require additional direct measurements of bedrock lowering rates with cosmogenic nuclides at Garner

Run, in addition to bedrock river incision models. Such models that can account for both variations in rock strength and temporal changes in relative base level.

In addition to variations in structure, lithology, and base level, Quaternary climate variations have left a strong imprint on the landscape of Garner Run and Shavers Creek in general. While the relict of the periglacial processes at Shale Hills are mostly observed in the subsurface colluvial stratigraphy (West et al., 2013), at Garner Run these processes have left behind boulder fields, solifluction lobes, and landslides observed at the land surface (Fig. 6). Such features are found throughout central Pennsylvania south of the limit of the LGM (last glacial maximum) limit (Gardner et al., 1991). These features document a major reorganization of the uppermost CZ by processes such as permafrost thaw. For example, the Leading Ridge hillslope (the southern hillslope defining the Garner Run subcatchment, Fig. 5) is characterized by a hummocky topography at the 5-10 m scale, with abundant partially vegetated boulder fields observed to be common. The other side of the catchment -- Tussey Mountain hillslope -- is steeper at the top, has greater relief, retains evidence of past translational slides, and contains open, unvegetated boulder fields. At the foot of the Tussey Mountain hillslope is a strong slope break that demarcates a low-sloping region characterized by abundant solifluction lobes, which appear to have accumulated as a large, valley-filling deposit (Figs. 6, 7). Such features were either not as active, or their evidence has been erased or buried, at the Shale Hills subcatchment.

Many of these geomorphological features have controlled or been imprinted on CZ processes and human activities in Garner Run. For example, the modern flow pathways for surface and groundwater in Garner Run are significantly influenced by the forcing factors of tectonism, climate, and anthropogenic activity. Flow pathways are influenced i) by topography inherited from geologic events from $10^8$ years before present, ii) by variations in soil grain size as dictated by periglacial processes operating $10^4$ years ago, and iii) by modern land use over the last $10^2$ – $10^3$ y.

In terms of land use, the influence of anthropogenic activity in the catchment is relatively minor and consistent with the surounding region. Neither Shale Hills nor Garner Run subcatchments show signs of having been plowed or farmed in row-crop agriculture, although some grazing may have occurred. The top of one of the ridges in Shale Hills appears to define a field edge. Both subcatchments were forested for at least 100 years. Based on historic aerial photographs, both watersheds contain intact, closed canopy forests in 1938 and

**Formatted:** Normal

**Formatted:** Superscript

**Formatted:** Superscript

**Formatted:** Indent: First line: 0.49"

show no sign of obvious stand level disturbance since that year. ~~Both watersheds experienced limited management activities (selected cuts, salvage harvests, etc.) as determined by the presence of stumps~~. In the mid 1800s, significant quantities of charcoal were made in this region to run several nearby iron furnaces. Given that charcoal hearths have been identified in the subcatchments from lidar, the subcatchments were probably cleared in the mid to late 1800s as most available wood was used for charcoal making. This land use was also often associated with fires.

The~~e above~~is short analysis of the geomorphology and land use highlights the influence of the forcing mechanisms (tectonism, climate, anthropogenic activity) that operat~~e~~ing over a wide range of timescales and yet ~~to~~ influence modern CZ processes~~,~~. The CZO efforts document the ~~and emphasizes the~~ importance of providing geologic and geomorphic context for ~~any~~ investigation of the CZ. ~~For example, the modern flow pathways for surface and groundwater in Garner Run are significantly influenced by topography inherited from geologic events from $10^8$ years before present, heterogeneous soil properties (e.g., clay and boulder content) controlled by periglacial processes operating $10^4$ years ago, and modern land use.~~

~~**4.2** Water and energy flux measurements at Garner Run: Tower HOG~~

## 3.2 Water and energy flux measurements at Garner Run: Tower HOG

Surface energy balance measurements (eddy covariance (EC) measurements of sensible and latent heat fluxes or upwelling terrestrial radiation/skin temperature) are needed to constrain Flux-PIHM (Shi et al., 2014). Measurements of precipitation, atmospheric state and incoming radiation are needed as inputs to the model. These measurements provide the data needed to simulate the catchment hydrology that is critical to understanding today's WEGSS fluxes. In addition, these fluxes are drivers for millennial-timescale landscape evolution (Fig. 1).

Instrumentation for measurements of water and energy flux measurements are designed as part of the "tower hydrological observation gear" – referred to here as Tower HOG (Table 2, Table S1). While the ideal plan would locate Tower HOG within the Garner run watershed itself, the remote, rocky, heavily wooded terrain makes this too challenging. Therefore, precipitation will be measured near Garner Run on a road crossing Tussey Mountain that is also the site of a pre-existing communications tower (see Fig. 3). A

**Formatted:** Superscript

**Formatted:** Superscript

**Formatted:** Normal

**Formatted:** Indent: First line: 0.5"

**Formatted:** Not Highlight

disdrometer (LPM, Theis Clima GmbH) and weighing rain gauge have been in use at Shale Hills since 2009 and 2006 respectively to measure precipitation. To measure precipitation amount at Garner Run, we are installing the simpler instrument (Pluvio[2], OTT Hydromet weighing rain gauge). Measurements will be compared to the National Atmospheric Deposition Program measurements and samples of rainwater. According to the nearest NADP site, Garner Run receives 1006 mm/y precipitation with an average pH of 5.0 (Thomas et al., 2013).

EC and radiation instrumentation (Table 2) will also be implemented on the pre-existing communications tower on the Tussey Mountain ridgeline (Fig. 3). Although located out of the subcatchment, the measurement footprint for the tower will be sensitive to fluxes from forests representative of those in Garner Run. The complex terrain at Shale Hills and Garner Run make EC measurements difficult to interpret in stable micrometeorological conditions. Since the primary energy partitioning happens during the day when the atmosphere is typically unstable, daytime sensible and latent heat flux measurements are sufficient to constrain the hydrologic modeling system. Daytime carbon dioxide flux measurements will inform the biogeochemical modeling system.

Formatted: Indent: First line: 0.49"

One of our major focuses is measuring precipitation and evapotranspiration (ET). These fluxes are drivers for landscape evolution as they are manifested today (Fig. 1). These measurements also are needed for the land surface water balance to constrain today's WEGSS fluxes. First, we are installing a laser disdrometerweighing rain gauge (LPM, Theis Clima GmbHPluvio[2], OTT Hydromet) to measure precipitation amount and type in Garner Run. Another disdrometer (LPM, Theis Clima GmbH) weighing rain gaugehas have been in use at Shale Hills since 20082009 2006. The disdrometer will be the Jo Hays Vista on Tussey Mountain near Garner Rundeployed as part of the "tower hydrological observation gear" referred to here as Tower HOG (Table 2)covariance. Tower HOG will be placed outside the watershed on Tussey Mountain ridgeline (Fig. 3). The remote, rocky terrain in Garner Run made constructing a new tower in the center of the watershed challenging. In contrast, a communications tower that is surrounded by representative forests already exists on the ridge top above the watershed, and we have therefore chosen this to host the eddy covariance flux instrumentation. Although the measurement footprint (i.e. fetch) for the tower measurements will include other areas, the tower instrumentation will be sensitive to fluxes from the forest in Garner Run. The tower measurements can also be compared to regional measurements such

Formatted: Superscript

Formatted: Not Highlight

as the National Atmospheric Deposition Program measurements and samples of rainwater. For example, according to the nearest NADP site, Garner Run receives 1006 mm/y precipitation with an average pH of 5.0 (Thomas et al., 2013).

In addition to precipitation, sensible and latent heat fluxes (i.e. using eddy covariance), or skin temperature (upwelling terrestrial radiation) must also be measured to constrain Flux-PIHM (Shi et al., 20132014). A small clearing below the tower site on the Tussey ridgeline makes the site unsuitable for skin temperature measurements representative of the forest, so we are only collecting eddy covariance measurements at the Tussey ridgeline. Of course, the complex terrain at both Shale Hills and Garner Run make eddy covariance measurements difficult to interpret in stable micrometeorological conditions. Since the primary energy partitioning happens during the day, however, daytime heat flux measurements, are sufficient to constrain the modeling. For the Garner Run subcatchment, in addition, we also may be able to use upwelling infrared radiation measurements currently being made at the nearby Shale Hills. These radiative energy fluxes are measured using a four component radiometer, i.e., one that measures upwelling and downwelling terrestrial and solar radiation (Table 2). With both the EC measurements at Garner Run and radiative flux measurements at Shale Hills, Flux-PIHM should be well constrained.

### 4.33.3 Vegetation mapping

Vegetation has important impacts on the today's WEGSS fluxes and is known to have has important but poorly understood influenced impacts on regolith formation and sediment transport over geologic time. As we study individual subcatchments to understand WEGSS budgets, we seek to learn enough about vegetation the fluxes to extrapolate WEGSS fluxes to the entire Shavers creek watershed. As described below, we once again use the geomorphological framework to design the measurement strategy for vegetation. We also want : we therefore seek to understand some of the biogeochemical controls on WEGSS fluxes of nutrients such as nitrate. out of Shavers creek. Ultimately, an OSSE will be run to compare measurements to model predictions as a way to determine the important parameters for predicting carbon and nitrogen fluxes. It For example, one of our goals is to assess nitrate fluxes out of Shavers creek. To do this, it may also be necessary to determine the effect of individual tree species on N flux (Williard et al., 2005). Ultimately, we plan to run an OSSE to compare model predictions to measurements as a way to determine the important parameters for predicting carbon and nitrogen fluxes.

**Field Code Changed**

**Formatted:** Not Highlight

**Formatted:** Indent: First line: 0.4"

As part of the geomorphological measurement strategy, we ~~The vegetation has already been~~ mapped the vegetation in Garner Run subcatchment across the Ground HOG catena (ridge top, midslope, and valley floor positions on one side of catchment and one midslope site on the other side, Fig. 5). The objective of the ~~ground-based~~ catena-based stratified ~~vegetation~~ sampling design ~~for the subcatchment~~ was to measure spatial variability in vegetation, i.e., under the assumption that landscape position was an important control on vegetation ~~across the catena (ridge top, midslope, and valley floor positions) defined for GroundHOG~~. These measurements set the stage for ~~later~~ planned re-measurements to understand temporal variability. For example, future assessments will quantify above-ground biomass, an important carbon pool. Variability in forest composition, standing biomass, and productivity across a watershed is generally related to gradients in biotic and abiotic resources such as soil chemistry or structure, water flux, and incoming solar energy. Therefore, the relatively restricted vegetation analysis design (Fig. 5) will be upscaled based on the team's developing knowledge of the distribution of soils across the watershed as well as lidar-based estimates of tree biomass and seasonal patterns of leaf area index and tree diameter growth. Given that we have not yet run an OSSE for ~~carbon~~ C or ~~nitrogen~~ N fluxes, our measurements of vegetation are relatively broad to enable such future ~~such~~ analysis.

**Formatted:** Indent: First line: 0.5"

~~The v~~Vegetation measurements are important not only for C and N fluxes, but also for water flux. At Shale Hills, seasonal variation in tree transpiration has been estimated using tree sap flux sensors (Meinzer et al., 2013). While we sampled many different tree species in multiple locations at Shale Hills (Fig. 2), a more restricted number will be sampled at Garner Run. For example, sapflux sensors are ~~being deployed at~~planned for only the midslope positions of Ground HOG (Fig. 5). While eddy flux and soil moisture dynamics provide estimates of total transpiration and evaporation, sap flux provides direct estimates of tree transpiration that can constrain model predictions of transpiration. Collectively, these measures will help evaluate Flux-PIHM model processes. In addition, all approaches to measuring water fluxes are imperfect; errors can best be constrained when multiple approaches are used.

**Formatted:** Indent: First line: 0.49"

In addition to these sapflux measurements limited to midslope pits, vegetation has been sampled in linear transects parallel to the slope contour at each of the four soil pits (LRRT, LRMS, LRVF, TMMS, Fig. 5, Section ~~4~~3.4), i.e., ~~at each of the following pits:~~ at Leading Ridge ridge top ~~(LRRT)~~, Leading Ridge midslope ~~(LRMS)~~, Leading Ridge valley floor ~~(LRVF)~~, Tussey Mountain midslope, ~~(TMMS)~~ respectively. Each vegetation transect was 10-m along the direction perpendicular to the valley axis and ~700-1400 m parallel to the valley axis.

Measurements along the transects yielded vegetation and forest floor cover data for 4.1 ha in the subcatchment (Table 3). The transects~~y~~ provide vegetation input data for land surface hydrologic models, and also evaluation data for a spatially-distributed biogeochemistry model (Flux-PIHM-BGC, Table 1). In the transected area, 2241 trees >10 cm diameter at breast height were measured, mapped, and permanently tagged. Understory vegetation composition was measured at 5 m intervals along transects and coarse woody debris was measured in 25 m planar transects parallel to the main transect, spaced every 100-m. Forest floor cover was classified as rock (typically boulder clasts from periglacial block fall), bare soil, or leaf litter every 1 m along each transect, and the dimensions (a, b, c axes) of the five largest exposed rocks was recorded every 25 m. Forest floor biomass was measured every 25 m along transects by removing the organic horizon from a 0.03 $m^2$ area for laboratory analysis: samples were dried, weighed and measured for carbon loss on ignition.

The ~~results from these linear~~ transect observations~~s~~ document variations in vegetation ~~across~~ along the catena ~~positions~~ (Table 3), as well as spatial variation in vegetation ~~within a~~at each position. For example, mean tree basal area (BA; the ratio of the total cross-sectional area of ~~tree~~ stems ratioed to ~~the total~~ land surface area) in the LRRT transect is 25.3 $m^2$ $ha^{-1}$~~;~~ ~~however, BA~~ with measurements ranging~~ed~~ from 0 to79 $m^2$ $ha^{-1}$. The subcatchment contains a dry oak-heath community type (Fike 1999), primarily consisting of chestnut oak (*Q. montana*), red maple (*Acer rubrum*), black birch (*Betula lenta*), black gum (*Nyssa sylvatica*), and white pine (*Pinus strobus*) in the overstory, with a thick heath understory of mountain laurel (*Kalmia latifolia*), blueberry (*Vaccinium sp.*) and huckleberry (*Gaylussacia sp.*) species, and rhododendron (*Rhododendron maximum*) along Garner Run.

~~Similarly~~T~~,~~he transect work also highlighted a type of measurement that we had not needed for Shale Hills but which our models and observations are showing is important in the new subcatchment: the fraction of land surface covered by boulders. At LRRT, 16% of points sampled every meter ~~in LRRT~~ fell on rock. Furthermore,~~, yet~~ rock coverage at ~~certain~~ some transect points ~~along the transect rock cover~~ was as high as 100% or as low as 0%. Vegetation ~~measurements will be combined with data on~~and surface rockiness data ~~(~~from transects~~)~~ will be combined with ~~and~~ a suite of ground and remotely sensed measurements from the watershed such as slope, curvature, aspect, solar radiation, and soil depth to model vegetation dynamics from environmental conditions and interpolate vegetation structure in areas of the watershed not directly sampled. Future ~~resampling~~ re-measurements along~~of linear~~ transects will allow assessment of carbon uptake in vegetation, as well as changes in forest composition and structure.

Additional key vegetation parameters will be assessed at the soil pits described in Section ~~3~~4.4 and Table S2. These additional measurements include root distributions, leaf area index (LAI, described in the next paragraph), litter fall, tree diameter growth and tree sap flux. Root distributions are being measured at all four soil pits in Garner Run using ~~a combination of~~ soil cores to ~~accurately~~ assess the high length densities near the surface. Root distributions, combined with soil water depletion patterns, can allow estimation of ~~inform~~ depth of tree water use over the season. Depth of tree water use, ~~which is~~ an input parameter in the PIHM suite of models~~.~~, is c~~C~~urrently derived from~~,~~ a look-up table (http://www.ral.ucar.edu/research/land/technology/lsm/parameters/VEGPARM.TBL) ~~is used~~ to determine the rooting depth of each landcover type~~ in the PIHM suite of models~~. We will explore whether the use of ~~Using~~ field- measured rooting depth as model input ~~may~~ improves the modeling of water uptake. In addition, profile wall mapping is being used to analyze the architecture, mycorrhizal colonization, and anatomy of deep roots. By characterizing and understanding the controls on root traits along a hillslope, we will eventually be able to use such observations to inform both models of water cycling (Flux-PIHM) and regolith formation (RT-Flux-PIHM, see Table 1).

At weekly intervals in the spring and fall and monthly during the summer, LAI will be assessed with a Li-2200 plant canopy analyzer (LI-COR Inc., Lincoln, Nebraska USA). The Moderate Resolution Imaging Spectroradiometer (MODIS) also provides remotely-sensed 8-day composite LAI (Knyazikhin et al., 1999; Myneni et al., 2002). The MODIS LAI product, however, has a spatial resolution of 1 $km^2$, which cannot resolve the spatial structure in LAI within small watersheds. The product also has a notable bias compared to field measurements (e.g. Shi et al., 2013). The LAI field measurements will be used for detailed information on leaf phenology, which is an important driver for the modeling of water and carbon fluxes for land surface and hydrologic models (e.g., PIHM, Flux-PIHM (Table 1)), and provides calibration or evaluation data for biogeochemistry models like Flux-PIHM-BGC (Naithani et al., 2013; Shi et al., 2013).

Another important value we must estimate is net primary productivity (NPP). With NPP it is possible to constrain carbon and nutrient fluxes in vegetation stocks, which can be large components of the overall budgets. To estimate aboveground NPP, we will measure annual variation in trunk growth with dendrobands emplaced on examples of each of the six dominant tree species near each soil pit site. In addition, traps at each soil pit are also being used to assess will collect litter fall for assessment. One of the key model outputs of Flux-PIHM-BGC is NPP, which can be evaluated using these measured data.

Formatted: Indent: First line: 0.5"

### 4.43.4 Soil pit measurements and Ground HOG instrumentation

#### 4.4.13.4.1 Soil observations

To first order, the The uplands of the Garner Run subcatchment land surface falls into one of three categories: i) fully soil mantled with few boulders emerging at the ground surface, ii) boulder-covered with tree canopy, and iii) boulder-covered without tree canopy. The coarse blocks of the Tuscarora sandstone range in diameter from ~10-200 cm, making it challenging to excavate large soil pits (Table 3). To assess the spatial heterogeneity of soils in the Garner Run subcatchment, we therefore focused efforts on four soil pits: three on the north-facing planar slope of Leading Ridge (LRRT, LRMS, LRVF) and one mid-slope pit on the south-facing slope of Tussey Mountain (TMMS) (Fig. 5). Three pits were dug by hand until deepening was impossible (LRRT, LRMS, and TMMS). The Leading Ridge Valley Floor (LRVF) pit was dug by hand and then deepened using a jackhammer until the inferred contact with intact bedrock was reached. The pits were excavated in the following soil series: TMMS, LRRT and LRMS (Hazleton-Dekalb association, very steep), and LRVF (Andover extremely stony loam, 0-8% slopes). This deployment of observations in soil pits along a catena, with an additional pit on the opposite valley wall, is here referred to as "Ground HOG" (ground hydrological observation gear) (Fig. 5, Supplement Fig. S1, S2) and is the result of our focus on a minimalist sampling design.

Formatted: Indent: First line: 0.49"

This design was informed by observations at Shale Hills and the new subcatchment and by modelling conceptualizations. As discussed earlier, the Shale Hills subcatchment upland land surface falls into one of two categories: hillslopes or swales. In contrast, we observed little evidence for swales in Garner Run. In addition, the surface cover at Garner Run consists of coarse blocks of the Tuscarora sandstone ranging in diameter from ~10-200 cm, making it challenging to excavate large soil pits, limiting the number of such installations

Formatted: Not Highlight

(Table 3). Three pits were dug entirely by hand (LRRT, LRMS, and TMMS). The Leading Ridge Valley Floor (LRVF) pit was dug by hand and was deepened using a jackhammer until the inferred contact with intact bedrock was reached. All four pits locationss in the new subcatchment were selected therefore located on slopes that were roughly planar or in planview to avoid areas of convergent flow. The midslope pits were located on somewhat convex-up hillslopes for reasons discussed(see below). Given our catena design, we excavated pits in the following soil series: TMMS, LRRT and LRMS (Hazleton-Dekalb association, very steep), and LRVF (Andover extremely stony loam, 0-8% slopes).

The rationale for the positions of the pits in Ground HOG ares as follows. First, regolith formation at a ridge top is the simplest to understand and model (see, for example, Lebedeva et al., 2007; Lebedeva et al., 2010) because net flux of water and earth materials is largely 1D: i.e., net water flux is largely downward and net earth material flux is upward over geological time. We are now developing Regolith-RT-PIHM is a model under development to to simulate regolith development quantitatively for such 1D systems, using constraints from cosmogenic isotope analysis (Table 1). The next level of complexity is a convex-upward but otherwise planar hillslope. The intent forSecond, Regolith-RT-PIHM will is that it will also be able to model convex-upward hillslopes by assessment of the hillslopes as a 2D systems that incorporates downslope transport of water and soil (e.g.Lebedeva and Brantley, 2013). By analyzing soilSoil pits along a convex-upward but otherwise planar hillslope planar hillslope as we did such as those described for Shale Hills (Jin et al., 2010b) can be used to parameterize , both 1D and 2D models of regolith formation will be enabled. With such conceptual and numerical models, we will extrapolate to other hillslopes within Shavers creek watershed. Third, at Shale Hills we discovered that while both planar hillslopes and swales were are important at Shale Hills, requiring measurements at both (Graham and Lin, 2010 ; Jin et al., 2011; Thomas et al., 2013) the lack of . No such swales have been observed at Garner Run, allowing focus on just one catena in the minimalist design. (In fact the lack of swales in the sandstone catchment is one of the observations that we hope we can eventually explain). Finally, the importance of aspect on soil development and WEGSS fluxes at Shale Hills has been noted o(Graham and Lin, 2011; Graham and Lin, 2010 ; Ma et al., 2011; West et al., 2014) on shale at Shale Hills (Graham and Lin, 2011; Graham and Lin, 2010 ; Ma et al., 2011; West et al., 2014), as well as on sandstones in Pennsylvania oin Pennsylvania (Carter and Ciolkosz, 1991). For that reason, Ground HOG includes one additional pit was sited on

Formatted: Not Highlight

the northern side of the catchment to make observations to constrain the effect of aspect (Fig. 5).

We will use numerical models to explore regolith formation and to extrapolate to other hillslopes within Shavers creek watershed.

This highlights the importance of understanding the soil to the CZ effort. Soil provides a record of both transport of rock-derived material as well as fluxes of water over the period of pedogenesis. For exampleAt each pit location, we described the soil profile the pits at, Garner Run which typically had the following structure:are characterized from land surface downward by a thin organic layer, an upper rocky layer with a thin organic soil, a leached layer characterized by sand-sized grains with few large clasts mostly absent, a sandy mineral soil with a thin layer of accumulated organic and sesquioxide material, and a deeper clay-rich layer with larger interspersed rock fragments interspersed (Supplement Fig. S3, Supplement Table S2). Depth intervals of the soil Additionally, for each pit we sampled soils atevery 10 cm and from basal rocks intervals for show variations in chemistry (Supplement Tables S3, S4), and are being analyzed for grain size, organic matter, and composition analysismineralogy (Supplement Table S4).

Formatted: Not Highlight

These soil observations yield further clues to the history of the landscape. Most of tThe Garner Run subcatchment has been mapped to lie on Lower Silurian Tuscarora sandstone (Flueckinger, 1969). This sandstone, deposited in the Lower Silurian, has been iInterpreted as reworked beach sediments during original deposition (Cotter, 1982), this sandstone has been. The unit has been mildly metamorphosed – to a highly indurated quartziteso that pressure solution has cemented the fabric of the rock: as such, the unit is often referred to as a quartzite. Cotter reported the unit to be close to 98% $SiO_2$. Weathering of sandstone is largely controlled by the porosity, the fraction of non-quartz grains, the composition of the cement (Turkington and Paradise, 2005), and the pH of soil porewaters (Certini et al., 2003). The porosity is important because it dictates how much water enters the weathering rock; in addition, during seasonal drying, salts deposited inside a sandstone can crystallize and disintegrate the rock (Labus and Bochen, 2012). Thermal cycling can also crack sandstones (Turkington and Paradise, 2005) as can tree roots (Amundson, 2004).

The average of the bBulk compositions of four five rock-samples collected from the bottoms of the five the Ground HOG soil pits –were used averaged to estimate an average composition of the quartzite for comparison to similar analyses of bulk regolith protolith samples (all measured using Li metaborate fusion followed by analysis by inductively coupled plasma atomic emission spectroscopy, Supplement Table S3). In Garner run samples, the These Tuscarora samples was observed to be close contain to >9896 wt. % $SiO_2$. , i.e., very similar to published Tuscarora compositions (Cotter, 1982). A small amount of Minor titanium (Ti), generally present in sandstones in highly insoluble minerals, was observed to be present in the parent (Supplement Table S3) and at even higher concentration. in soils (Table S4). This enrichment in soil could be due to several processes during weathering: for example, retention of Ti from the protolith, losses of elements other than Ti, or addition of Ti to the soil. If Ti in the soil was derived from protolith, By calculating the normalized concentrations for elements assuming Ti is insoluble, we assessed the loss or gain of other elements from the regolith as compared to Ti in the underlying Tuscarora sandstone can be calculated from the . These normalized concentrations are referred to as mass transfer coefficients, $\tau_{ij}$, where $i$ is the immobile element Ti and $j$ is the an mobile element that was lost or gained (Anderson et al., 2002; Brimhall and Dietrich, 1987). From this assessment of regolith mass balance, it was observed that Assuming Ti in soil was derived from the protolith, $\tau_{Ti,j}$ values = 0 within error for Al, CaMg, Na, Si, and P Fe, indicating they were either largely unchanged neither added nor depleted($\tau \approx 0$) or highly depleted ($\tau < 0$) compared to Ti. the underlying rock. In contrast, $\tau_{Ti,K}$ Mg, K, and Fe were all significantly enriched in the soils ($\tau > 0$) , consistent with addition of K to the soil compared to the protolith (Supplement Fig. S38). Error bars on many of the elements are very large because of the variability in the low concentrations of all elements except Si and O. On each plot, the star represents the parent composition ($\tau \approx 0$), plotted at an arbitrary depth.

These observations are eAccording to onsistent with published arguments in the literature that for ridgetop soils this formation in this region, the thin and poorly developed ridgetop soil are is likely residual, poorly developed, and thin (Ciolkosz et al., 1990). In contrast, downslope soils on hillslopes likely generally developed not only from rock in place but also from colluvium (Fig. 5). Furthermore, previous researchers have pointed out that soils in central Pennsylvania commonly show a brown over red color layering that has been attributed to may indicate two generations of exposure of earlier regolith to weathering, i.e., a previously weathered (producing the red layer) followed by emplacement which was then covered by a of colluvial layerum that experienced additional weathering (the brown layer) (Hoover and Ciolkosz, 1988). Such polygenetic histories will make regolith formation

~~modelling more complex.~~ Although the soils here did not show a strong brown over red color signature (Supplement Fig. S3), clay-rich soil at depth may document soil formation before the LGM (Table S2). The addition of ~~Mg,~~ K~~, and Fe~~ to the soils, even in the residual soils at the ridgetop ~~where downslope transport is unlikely to have been significant~~ (Supplement Fig. S3), is another complexity. K could have been added as exogenous dust inputs which were very important during and immediately after glacial periods (Ciolkosz et al., 1990). Alternately, K-containing clay particles could have percolated downward from weathering of the overlying units such as the Rose Hill shale before it was eroded away (Fig. 4). Such movement of fines downward from the Rose Hill have been observed at Shale Hills (Jin et al., 2010~~a~~): such particles could have been added to the underlying Tuscarora and then retained in the soil. In that case, the assumed protolith composition could be erroreous, especially if Ti was added from the downward infiltrating fines. K enrichment could also be explained by~~, could either be explained by exogenous additions to the soil or by protolith compositional variation which was not assessed in the small set of 5 rock samples. For example, some interfingered~~ shales ~~are known to occur~~ within the Tuscarora formation itself (Flueckinger, 1969). If these interfingered shales were the protolith of the observed soils~~ and could have provided the excess Mg, K~~, this would mean that our estimated protolith composition was K-deficient~~, and Fe~~. ~~Alternately, addition of these elements could have been caused by i) dust inputs (Ciolkosz et al., 1990) which were likely to be important especially during the glacial period and just after, or ii) fines percolating downward from weathering of the overlying Rose Hill shale before it was eroded away (Fig. 4).~~ Thus, soil analysis (Fig. 8) leads to interesting hypotheses that will be investigated. ~~Movement of fines out of the Rose Hill shale is known to be happening today from our work at Shale Hills (Jin et al., 2010a).~~

**Formatted:** Not Highlight

### ~~4.4.2~~3.4.2 Ground HOG

**Formatted:** Indent: First line: 0.49"

The Ground HOG instrumentation enables the *in situ* measurement of soil moisture and temperature, as well as gas and pore-fluid compositions, all at multiple depths (Fig. 5, Supplement Fig. S2). Ground HOG complements the atmospheric measurements ~~taken by~~ at Tower HOG ~~instrumentation~~ (Section 3.2). Because ~~the~~ Ground HOG sites are difficult to access, measurements were automated to the extent possible. However, the lack of access to electricity and the cost of automated sensors (for $CO_2$ for example) meant that a completely automated monitoring system was unfeasible as well. Therefore, our final approach (Supplement Fig. S2) included a few automated components recording a continuous time series of data, coupled with additional components to be monitored manually, but with lower temporal resolution.

In selecting depths for soil sampling we wanted to instrument the site so that results could be compared across all watersheds~~, which meant~~. Thus, we focused on a depth-based (as opposed to horizon-based) sampling scheme. In addition, we wanted to emphasize surface soils that have the highest water and biogeochemical flux rates. These layers also have the strongest influence on the atmospheric boundary layer. At the same time, we wanted to also document deep soil processes critical to understanding weathering and subsurface flowpaths. Thus, our final depth distribution included samples at 10, 20, and 40 cm from the top of the mineral soil (we used the top of the mineral soil as the depth reference because the O horizon depth varies greatly across the sites and among land-use types) and 20 cm above the bottom of the soil pit (coded “D-20”). At these four depths we installed from 1 to 4 component devices of the Ground HOG in each pit.

Automated soil moisture and temperature sensors (Hydra Probe, Stevens Water Monitoring Systems, Inc. Portland, OR) were emplaced to monitor at 10, 20 and 40 cm depths on the uphill face of each pit (Fig. S2). In addition, TDR waveguides (Jackson et al., 2000) for manual point estimates of soil moisture were installed at the same depths plus D-20 on the uphill pit face, and the left and right pit faces (facing uphill). Wave guides are paired metal rods on a single cable that conduct a signal for time-domain reflectometry. The rods are 20 cm long and hand-made (Hoekstra and Delaney 1974, Topp et al 1980; Topp and Ferre 2002). We placed 12 (4 depths x 3 pit faces) in each pit. The automated sensors were emplaced at depths expected to have the most dynamic soil moisture. In contrast, the waveguides measure deep soil moisture where temporal variability is expected to be low. The use of waveguides added spatial replication at all depths (Fig. 5, Supplement Fig. S2).

Co-located with every soil moisture waveguide is an ~~soil gas~~ access tube to sample soil gas for measurements of the depth distribution of $CO_2$ and $O_2$ at a low temporal frequency. At 20 cm below the soil surface and 20 cm above the bottom of the uphill face of the pit, sensors are continuously measuring soil $CO_2$ (GP001 $CO_2$ probe, Forerunner Research, Canada) and $O_2$ (SO-110 Sensor, Apogee Intruments, Utah, USA) at the two midslope catena positions. We selected the midslope catenas for these sensors because they provide the best locations for contrasting north and south aspects. We placed one sensor at the D-20 location to document controls on acid and oxidative weathering near the bedrock interface. The second sensor is near the surface to monitor a zone of high biological $CO_2$ and $O_2$ processing. We did not install the sensors at the shallowest depth (10 cm) because we found that high diffusion and advection at shallower depths causes the gas concentrations at 10 cm to reflect atmospheric conditions, providing less information on soil biology (Jin et al., 2014; ) (Hasenmueller et al., 2015).

Lysimeters (Super Quartz, Prenart Equipment ApS, Denmark) have been emplaced to allow periodic manual sampling of soil pore water for chemical analysis at 20 cm and D-20 cm depths in all catena locations. The rationale for these depths is the same as described above for the automated $CO_2$ and $O_2$ sensors (they are co-located in the midslope pits). Overall, these Ground HOG measurements will parameterize the regolith formation models (Table 1) and will be used to test hypotheses linking hydrology, biotic production/consumption of soil gases, and weathering rates.

**Formatted:** Indent: First line: 0.5"

### 4.53.5 Upscaling from the pits to the catena using geophysics

To supplement the Ground HOG observations, we use geophysical and large-footprint methods to interpolate between and extrapolate beyond soil pits. For example, a cosmic-ray neutron detector (CR-1000B, Hydroinnova Inc.) has been emplaced to measure large-scale (~0.5 km radius) average soil moisture every 30 minutes. This COSMOS unit, already used in a variety of ecosystems (Zreda et al., 2013), will measure spatially averaged (3D) soil moisture content within the watershed. Data processing methods have been developed that can accounts for various types of moisture storage (e.g. canopy storage, snow, water vapor (Franz et al., 2013; Zweck et al., 2013). The sensor has been installed near the LRVF (Leading Ridge valley floor) pit to provide spatially averaged moisture estimates across the valley.

**Formatted:** Indent: First line: 0.5"

The COSMOS fills in the gap between small-scale point measurements (Fig. 5) and large-scale satellite remote sensing. The footprint of COSMOS is optimal for hydrometeorological model calibration and validation at small watersheds. One sensor was installed at Shale Hills in 2011 and we are currently testing the COSMOS data with PIHM. We anticipate the results from both catchments will yield insights into the capabilities of cosmic-ray moisture sensing technology in steep terrain and will offer valuable insights into the problem of upscaling soil moisture measurements.

**Formatted:** Indent: First line: 0.49"

Ground HOG measurements will be further complemented by geophysical mapping along the catenas, including ground penetrating radar (GPR) transects of subsurface structure. Electromagnetic induction (EMI) mapping of soil electric conductivity will similarly be used to measure soil spatial variations between pits. We plan repeated GPR and EMI surveys, in combination with terrain analysis using lidar topography, to identify subsurface hydrological features and soil distribution using published procedures (Zhu et al., 2010a, b). We will also field check regolith depths using augers, drills, etc. With repeated geophysical surveys over time (e.g., different seasons and/or before and after storm events), we ~~will~~ can also explore temporal changes in heterogeneous soilscapes and subsurface hydrologic dynamics, as demonstrated ~~in the previous studies~~ at Shale Hills (Guo et al., 2014; Zhang et al., 2014).

Such geophysical mapping is necessary to link between and compare with ~~between~~ soil-pit point measurements. For example, d~~D~~epth to bedrock along the catenas will ~~also~~ be mapped using the geophysical surveys and compared to pit measurements (Fig. 5). These data can be used for upscaling biogeochemical patterns and processes. For example, we expect that soil depth and soil moisture exert the strongest controls on variation in soil gas concentrations, as observed in many places, including Shale Hills (Hasenmueller et al., 2015~~, in press~~; Jin et al., 2014). Empirical relationships among these variables developed at Ground HOG points can be coupled with catchment scale soil moisture (from COSMOS) and soil depth (from GPR) data to upscale soil gas characteristics to the whole catchment.

**Formatted:** Indent: First line: 0.49"

~~An example of~~ To exemplify the utility of this approach, results ~~is shown here~~ from an investigation completed using a ground penetrating radar unit (TerraSIRch Subsurface Interface Radar System-3000) ~~. The unit was~~ used to map the depth to bedrock in the Garner Run hillslope near the three major monitoring sites (LRVF, LRMS, LRRT) is shown in ~~(~~Fig. ~~5~~9~~)~~. Multiple GPR traverses were completed by pulling the antennae along the ground surface. A distance-calibrated survey wheel with encoder was bolted onto these antennae to provide greater control of signal pulse transmission and data collection. The survey wheel occasionally slipped in the challenging terrain, resulting in some errors~~line lengths recorded by the survey wheel which were slightly different than the actual lengths~~. ~~In order t~~R~~o surface normalize the radar records collected, r~~elative elevation data were collected as described below ~~at major slope breaks~~ along the traverse line ~~with an engineering level and stadia rod~~to surface normalize the data.

A traverse line from near Garner Run to the summit was established that ascends Leading Ridge in ~~essentially~~ a nominally west to east direction from ~494 to ~588 masl~~near Garner Run to the summit, running from about 494 to 588 masl~~ (Fig. ~~5~~9). The dominant soils mapped along this traverse line (Supplement Table S2) include~~:~~ Andover, Albrights, Hazleton, and Dekalb. The very deep, poorly drained Andover and moderately well to somewhat poorly drained Albrights soils have been reported in general to have formed in colluvium derived from acid sandstone and shale on upland toe-slope and foot-slope positions. The moderately deep, excessively drained Dekalb and the deep and very deep, well-drained Hazleton soils formed on higher-lying slope positions in residuum weathered from acid sandstone. These soils have moderate potential for penetration with GPR.

The traverse line was cleared of debris but the ground surface remained highly irregular with numerous rock fragments and exposed tree roots~~. These obstacles~~ that often halted the movement and caused poor coupling of the antennas with the ground. ~~In this study, f~~Flags were inserted in the ground at noticeable breaks in the topography along the traverse line. User marks were inserted on the radar records as the antenna passed ~~by~~ these survey flags. Later, the elevations of these points were determined using an engineering level and stadia rod. The elevation data were entered into the radar data files and used to “surface normalize” or “terrain correct” the radar records.

In this preliminary investigation, the soil-bedrock interface was not easy to identify. This was attributed to poor antenna coupling with the ground surface in the challenging rocky terrain, noise in the radar records caused by rock fragments in the overlying soil, irregular and fractured bedrock surfaces, and varying degrees of hardness in both rock fragments and the underlying bedrock. These factors weakened the amplitude, consistency and continuity of reflections from the soil-bedrock interface. Nevertheless, ~~we describe the~~ preliminary results are described below.

Figure 9 shows two surface-normalized plots of ~~the~~ data ~~that were~~ collected with the 400 MHz antenna as it was pulled ~~down~~ from the summit of Leading Ridge ~~from the summit area~~ to near ~~the~~ Garner Run ~~(the stream)~~. ~~In these plots, the d~~Distance ~~scale~~ is measured from the summit area to near Garner Run. ~~While d~~Differences in gross reflection patterns can be used to differentiate rock from soil, ~~on these images~~ but the soil-bedrock interface is diffuse. ~~However, w~~We collected four repeated GPR transects using both 400 and 270 MHz antenna. Compared with the 400 MHz antenna, the lower resolution of the 270 MHz antenna ~~has~~

**Formatted:** Indent: First line: 0.5"

smoothed ~~-~~out irregularities in the bedrock surface and reduced the noise from smaller, less extensive subsurface features, thus improving the interpretability of the soil-bedrock interface. Based on a total of 14,748 soil-depth measurements from ~400 m long GPR images along this traverse line, the interpreted depth to bedrock ranged from 0.58 to 2.42 m and averaged 1.37 m (~~, with a range of 0.58 to 2.42 m.~~ Table 4, ~~summarizes the depth to bedrock values estimated from the two radar traverses shown in~~ Fig. 9). Each entry in ~~the t~~Table 4 indicates the frequency of depth to bedrock data collected with the 400 MHz antenna along a traverse line, ~~that descended from the Leading Ridge. Data are~~ grouped into four soil depth classes. The GPR-derived soil depths are reasonable compared to the values we estimated in the soil pits (Fig. 9, Table S2).

## 3.6 Hydrology: Groundwater measurements

Several methods are needed to characterize physical and chemical interactions of water with regolith and rock in a catchment. First, physical inputs and outputs to a catchment, including precipitation, interception, ET, soil infiltration, and groundwater discharge, must be understood. Often, groundwater flows are omitted from comprehensive hydrology-meteorology-vegetation models such as the Variable Infiltration Capacity (VIC) hydrologic model, or the Noah Land-Surface Model (LSM). However at Shale Hills, we have estimated that rough 50% of incoming water is evapotranspired and 5% reaches the regional groundwater table and returns to the stream as baseflow (Sullivan et al., subm.). At Garner Run, we also expect groundwater to play a significant role in streamflow and geochemical dynamics. For example, some researchers have found that drainage and runoff on sandstone catchments is controlled to great extent by bedrock (Hattanji and Onda, 2004), and specifically by flow through fractures in the upper meters of sandstone beneath the soil (Williams et al., 2010). In this section and the next section we focus on quantifying fluid flow and transport of solutes into surface water and groundwater. We aim to measure the relative magnitudes, timing, and spatial variability of these fluxes. We emphasize methodologies for measuring and characterizing groundwater and streamwater to identify subsurface flow paths of groundwater, and the drivers and controls on water-rock interactions.

**Formatted:** Indent: First line: 0.5"

In the spirit of "measuring only what is needed," well installation and solid earth sampling by coring will be reduced compared to Shale Hills. At Shale Hills, 28 wells were emplaced and then intermittently monitored (Fig. 2). In Garner run, deep samples (> 8m) have been extracted between Garner Run and Roaring Run from the Harrys Valley 1 well (HV1)

**Formatted:** Indent: First line: 0.49"

drilled within the Garner Run catchment (see Fig. 3). Using a hand-held drill, three shallow wells will be installed and cores will be collected at the catena sites (Fig. 5) and additional monitoring wells will be installed along hillslopes and the valley floor. From these wells, we will also sample solid-phase chemistry and mineralogy.

All core samples will be analyzed for bulk chemistry and mineralogy to characterize the weathering reactions and protolith. Where possible, we will install groundwater monitoring wells in boreholes, with screened intervals spanning the water table. Monitoring at the wells will include hourly water level measurements using autonomous pressure loggers, hourly temperature measurements at two depths below the water table, and monthly water samples collected and analyzed for major ion chemistry. A pumping test will be conducted at the adjacent valley floor wells to measure aquifer storativity and hydraulic conductivity. Relative residence time of groundwater will be assessed from pathway analysis. If resources permit, $SF_6$ and chlorofluorocarbons (CFCs) will be measured in groundwater samples to assess residence time in the subsurface, as we have done for Shale Hills (Sullivan et al., in press).

**Formatted:** Not Highlight

Deep core samples and groundwater monitoring will provide a baseline understanding of the geologic and hydrologic system on the new sandstone lithology. Subsequent hypotheses about controls on weathering and hydrologic dynamics, as well as historical flow and solute fluxes, will be constrained by these observations at the catchment boundaries.

**Formatted:** Indent: First line: 0.5"

## 3.7 Hydrology: Streamwater flow and chemistry measurements

The Garner Run study reach is approximately 500 m long (Fig. 5) and consists of a rocky, often braided, channel. We have installed a flume at the downstream end of the reach to measure discharge. Stage is continuously monitored using a pressure transducer (Hobo U-20, Onset Computer Corp., Hyannis, MA). Surface water – groundwater (SW-GW) exchange characteristics have been measured using a short-term deployment of a fiber-optic distributed temperature sensor (FO-DTS), and two tracer injection tests. Stream chemistry, including dissolved oxygen (DO), pH, total dissolved solids (TDS), $NO_3^-$, $SO_4^{2-}$, Ca, K, Mg, Mn, Na, Fe, and Si, are measured biweekly or monthly in the field with handheld electrodes along the 500 m reach, or by grab sampling and laboratory analysis (inductively coupled plasma-atomic emission spectroscopy, organic carbon analyzer, and ion chromatography).

**Formatted:** Indent: First line: 0.5"

Stream chemistry is also explored using higher temporal resolution by using a s::can spectrometer and an autosampler during storm events (s::can, GmBH, Vienna, Austria). The s::can is an in-situ instrument capable of measuring such water quality parameters as pH, TDS, dissolved organic carbon (DOC), $NO_3^-$, DO, $NH_4^+$, $K^+$, and $F^-$. The chemistry and tracer test data will help quantify the flux of fluid and solutes through the subcatchment. The stream chemistry and discharge data will be combined with soil moisture, soil pore water chemistry, and groundwater data to estimate relative contributions to the stream, and underlying processes related to weathering in the near surface and aquifer.

Preliminary results from Garner Run indicate lower concentrations of Ca, Mg, and K compared to the stream discharging from Shale Hills. In addition, as expected, an initial constant injection tracer test at Garner Run revealed significant exchange with the subsurface during low-flow conditions (~0.004 $m^3\ s^{-1}$). Tracer test and temperature results suggest that the stream is losing water along some sections of the 500 m experimental reach and gaining water in others. Both the FO-DTS and stream chemistry data indicate significant input of spring water at ~100 m downstream of the Ground HOG catena (Fig. 5), which is chemically distinct from the upstream surface water and local groundwater sampled from the deep HV1 well. The DTS time series data will be analyzed to identify locations and magnitudes of inputs to the stream, as well as characteristic responses to rainfall events. In combination with the tracer tests, DTS, and chemistry results, we will use well logs and lidar topography to explain the lithological and geomorphologic controls on the surface water – ground water (SW-GW) system.

**Formatted:** Indent: First line: 0.49"

To characterize the major controls and processes governing WEGSS fluxes through the entire Shavers Creek catchment, we are making strategic measurements across the watershed to represent variability: stream discharge, stream chemistry, lithology, and geomorphology. Specifically, stream discharge and chemistry are being monitored along the main stem of Shavers Creek (SCAL, SCBL, and SCO, Fig. 3). At each location we are monitoring stage continuously using pressure transducers (Hobo U-20, Onset Computer Corp., Hyannis, MA), and using periodic discharge measurements to construct stage-discharge rating curves. SW-GW exchange characteristics will be measured as the channel crosses varying lithologies using a series of tracer injection tests. Analyses of stream chemistry from the main stem of Shavers Creek provide a spatial integration of solute behavior from upstream lithologies and land-use types. Eventually, with data from the three subcatchments on shale, sandstone, and calcareous shale (Fig. 3), we will make estimates for non-monitored catchments and test up-scaled estimates of the processes observed in each small watershed.

Preliminary stream chemistry and discharge results indicate significant variability among the three monitoring locations along Shavers Creek (Fig. 10). We see declining concentrations with increasing discharge for Mg and Ca (not shown), and somewhat chemostatic behavior for Si, K, nitrate and others. In this context, chemostatic is used to refer to concentrations of a stream that vary little with discharge (Godsey et al., 2009). Concentrations of Si decrease downstream (a dilution trend), while concentrations increase for Mg and nitrate, presumably due to agricultural amendments in the lower half of Shavers Creek watershed where land use includes farmland. The variety of behaviors will be investigated with respect to land use and lithology changes through the catchment.

**Formatted:** Indent: First line: 0.5"

## 4.6 Hydrology: Groundwater measurements

Several methods are needed in a catchment to characterize physical and chemical interactions of water with regolith and rock. First, physical inputs and outputs to a catchment, including precipitation, interception, ET, soil infiltration, and groundwater discharge, must be understood. In fact, however, groundwater flows are often omitted from comprehensive hydrology-meteorology-vegetation models (e.g. the Variable Infiltration Capacity (VIC) hydrologic model, or the Noah Land-Surface Model (LSM)); however, at Shale Hills, we have estimated that 5% of the nonevapotranspired water that enters the catchment reaches the regional groundwater table and flows to the stream as a deep flow component (Sullivan et al., subm.). At Garner Run, we also expect groundwater to play a significant role in streamflow and geochemical dynamics. For example, some researchers have found that drainage and runoff on sandstone catchments is controlled to great extent by bedrock (Hattanji and Onda, 2004), and specifically by flow through fractures in the upper meters of sandstone directly beneath the soil (Williams et al., 2010). In this section and the next section we focus on quantifying flows through and between surface water and groundwater. We aim to measure the relative magnitudes, timing, and spatial variability of these fluxes. We emphasize methodologies for measuring and characterizing groundwater and streamwater to characterize groundwater residence times, identify subsurface flow paths, and the drivers and controls on water-rock interactions.

Our plans for well installation and solid earth sampling by coring are reduced compared to sampling at Shale Hills. At Shale Hills, 28 wells were emplaced and then intermittently monitored (Fig. 2). In Garner run, two deep cores (> 50m) will be extracted at two locations near the Garner Run catchment, one (~ 100m) on Tussey ridge, i.e. the ridge that divides Shavers Creek from the watersheds to the northwest and one (50-75m) on the smaller divide within Shavers Creek between Garner Run and Roaring Run (see Fig. 3). Three shallow wells will be installed and cores (~10m) will be collected at the catena sites (Fig. 5). Two to four additional monitoring wells will be installed along the stream reach on the valley floor. In drilling boreholes for assessment of groundwater, we also sample borehole solid phase chemistry and mineralogy.

All core samples will be analyzed for bulk chemistry and mineralogy to characterize the weathering reactions and protolith in the critical zone. All boreholes will have groundwater monitoring wells installed, with screened intervals spanning the water table and with instrumentation as shown in Fig. 5. Monitoring at the wells will include hourly water level measurements using autonomous pressure loggers, hourly temperature measurements at two depths below the water table, and monthly water samples collected and analyzed for major ion chemistry. A pumping test will be conducted at the adjacent valley floor wells to measure aquifer storativity and hydraulic conductivity.subm.

**Formatted:** Not Highlight

Deep core samples and groundwater monitoring will provide a baseline understanding of the geologic/pedologic and hydrologic system on the new sandstone lithology. Subsequent hypotheses about controls on weathering and hydrologic dynamics, as well as historical flow and solute fluxes, will be constrained by these observations at the catchment boundaries.

## 4.7 Hydrology: Streamflow and chemistry measurements

The Garner Run study reach is approximately 500 m long within the catchment (Fig. 5) and consists of a rocky, often braided, channel. We have deployed a flume at the downstream end of the reach to measure discharge, and are monitoring stage continuously using a pressure transducer (Hobo U-20, Onset Computer Corp., Hyannis, MA). Surface water – groundwater (SW-GW) exchange characteristics have been measured using a short term deployment of a distributed temperature sensor (DTS), and will be supplemented by a series of tracer injection tests to investigate hyporheic exchange characteristics over a wider range of stream discharges. Stream chemistry, including DO, pH, TDS, $NO_3^-$, $SO_4^{2-}$, Ca, K, Mg, Mn, Na, Fe, and Si, are being measured biweekly or monthly in the field with handheld electrodes along the 500 m reach, or by grab sampling and laboratory analysis (inductively coupled plasma atomic emission spectroscopy or ion chromatography).

Stream chemistry is also being monitored intermittently using higher temporal resolution by using a s::can spectrometer and an autosampler during storm events. The s::can is an in-situ measurement instrument for several water quality parameters (pH, TDS, DOC, $NO_3^-$,DO, $NH_4^+$, K, F (s::can, GmBH, Vienna, Austria). The chemistry and tracer test data will help quantify the flux of fluid and solutes through the subcatchment. The stream chemistry and discharge data will be combined with soil moisture, soil pore water chemistry, and groundwater data to estimate relative contributions to the stream, and underlying processes related to weathering in the near surface and aquifer.

Preliminary results from Garner Run indicate lower concentrations of Ca, Mg, and K, compared to Shale Hills. In addition, as expected, an initial constant injection tracer test at Garner Run revealed significant exchange with the subsurface during low-flow conditions (~0.004 $m^3$ $s^{-1}$). Tracer test and temperature results documented that the stream sometimes loses and sometimes gains water in different sections over the 500 m experimental reach. One point stood out along the reach: both DTS and stream chemistry measurements are consistent with a significant input of groundwater at ~100m downstream of the catena (Fig. 5). The DTS time series data will be analyzed to identify locations and magnitudes of groundwater inputs, as well as characteristic responses to rainfall events or changes in stream discharge. In combination with the tracer tests, DTS, and chemistry results, we will use well logs and lidar topography to explain the lithological and geomorphologic controls on the SW-GW system.

To characterize the major controls and processes governing WEGSS fluxes through the entire Shavers Creek catchment, we are making strategic measurements across the watershed to represent variability: stream discharge, stream chemistry, lithology, and geomorphology. Stream discharge and chemistry are being monitored along the main stem of Shavers Creek (SCAL, SCBL, and SCO) as shown in Fig. 3. At each location we are constructing a stage-discharge rating curve, and monitoring stage continuously using pressure transducers (Hobo U-20, Onset Computer Corp., Hyannis, MA). Streamwater-groundwater exchange characteristics will be measured as the channel crosses varying lithologies using a series of tracer injection tests. Stream chemistry will be measured monthly at each sampling site along Shavers creek. Analyses from the main stem of Shavers Creek provides a spatial integration of solute behavior from upstream lithologies and land use types. Eventually, with data from the three subcatchments on shale, sandstone, and calcareous shale, we will make estimates for nonmonitored catchments and test up-scaled estimates of the processes observed in each small watershed.

Preliminary stream chemistry and discharge results indicate significant variability among the three monitoring locations along Shavers Creek (Fig. 10). We see declining concentrations with increasing discharge for Mg and Ca (not shown), and somewhat chemostatic behavior for Si, K, nitrate and others. In this context, chemostatic is used to refer to concentrations of a stream that vary little with discharge (Godsey et al., 2009). Concentrations of Si decrease downstream (a dilution trend), while concentrations increase for Mg and nitrate, possibly due to agricultural amendments in the lower half of the watershed. The variety of behaviors will be investigated with respect to land use and lithology changes through the catchment.

# 4 Model-data feedbacks

Throughout this paper we have described the two-way exchange of field-model insights needed to maximize the efficiency of CZ science. To understand the CZ requires models 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 are necessary to provide data for calibration and testing.

The CZ approach of using models to cross from short to long timescales has an important major benefit. Investigations that target long timescales can tease out the effects of feedbacks and thresholds in complex systems that are difficult to discern in short-timescale studies. We thus use quantitative models to explore a vast range in both spatial *and* temporal scales. In this paper we emphasized our approach toward designing a CZO as a tool to understand the CZ as one integral system. We therefore emphasized only one modelling tool, the PIHM family of models. This cascade of models provides a quantitative way for different disciplines to interact about the CZ through the use of a shared suite of models. Our current conceptual understanding and our current computers do not allow us to produce one model that simulates the CZ at all timescales, hence the cascade of models (Table 1).

Such a suite of models is integral not simply for predicting landscape and ecosystem response, but also to building a heuristic understanding of individual CZ processes that may not be apparent from 1st-order observations. Systems-level models are especially needed for

**Formatted:** Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.13" + Tab after: 0.43" + Indent at: 0.43"

**Formatted:** Not Highlight

**Formatted:** Indent: First line: 0.5", Line spacing: 1.5 lines

**Formatted:** Not Highlight

**Formatted:** Not Highlight

**Formatted:** Indent: First line: 0.49", Line spacing: 1.5 lines

**Formatted:** Indent: First line: 0.5", Line spacing: 1.5 lines proposing and testing hypotheses about feedbacks between climate, biota, and Earth surface and near-surface processes. Although not emphasized here, we have also cited publications throughout this paper that describe the many smaller scales or disciplinary-specific hypotheses and models that have been invoked to learn about individual CZ systems. For example, we point to our earlier observation of Ti and K enrichment in Garner Run soils (Fig. 8). We suggested several processes that could interact to explain data in Fig. 8, including preferential retention of some elements in protolith compared to others, depth variations in protolith composition, accumulation of fines from weathering formations above the current protolith, and dust additions. While first-order mass balance model calculations such as those implicit to Fig. 8 can be used to propose or test such hypotheses, use of Regolith-RT-PIHM (Table 1) or WITCH (Godderis et al., 2006) to model regolith formation are necessary to quantitatively test feasibility of such ideas. Better understanding of regolith formation will in turn inform the permeability distributions needed for hydrologic flow models in the CZO.

## 5 Conclusions: Measuring and modelling the CZ

Many environmental scientists worldwide are embracing the concept of the critical zone – the surface environment considered over all relevant timescales from the top of the vegetation canopy to the bottom of ground water. CZ science is built upon the hypothesis that an investigation of the entire object – the CZ – will yield insights that more disciplinary-specific investigations cannot. To understand the evolution and dynamics of the CZ, we are developing a suite of simulation models as shown in Table 1 (Duffy et al., 2014). These models are being parameterized based on measurements made at the Susquehanna Shale Hills Critical Zone Observatory (SSHCZO) which is currently expanding from less than 1 $km^2$ to 165 $km^2$.

In this paper we described an approach for assessing the CZ in the larger watershed. In effect, our measurement design is a hypothesis in answer to this question: if we want to understand the dynamics and evolution of the entire CZ, -what measurements are needed and where should they be made? Our approach emphasizes upscaling from 1D to 2D to 3D using a catena paradigm for ground measurements that are extended with geological, geophysical, lidar, stream and meteorological measurements. Of course, our dataset has very low or no little sampling replication within each catchment and we have only designed for one catchment per parent material. This results from theObviously, there is a tension between

Formatted: Outline numbered + Level: 1 + Numbering Style: 1, 2, 3, … + Start at: 1 + Alignment: Left + Aligned at: 0.13" + Tab after: 0.43" + Indent at: 0.43"

Formatted: Indent: First line: 0.5"

Formatted: Indent: First line: 0.49"

monitoring a core dataset over time (a geological or hydrological approach) versus the replication that is needed for spatial characterization (a soil science or ecological approach). Our spatial design was chosen based on the implicit assumption that implementation of ground- and tower-based measurements (Ground HOG, ~~and~~ Tower HOG) in each subcatchment could be upscaled to the entire watershed by interpolation~~,~~ and extrapolation, ~~and~~ as well as modelling ~~as described in~~ (Table 1). For example, we are testing the hypothesis that fewer soil pits are needed because we are using a regolith formation model and geological knowledge to site the few pits that we dig. If we find that our limited digging of soil pits is not successful in characterizing the regolith adequately – if our models of regolith formation do not match observations or our models of water flow through regolith do not simulate observations – more pits can be dug or new approaches toward geophysical measurements can be refined.

Formatted: Not Highlight

~~As an example of this approach, we point to our earlier observation of loss of Al, Na, Si and P from the soils at the same time that we identified significant enrichment in Mg, K, and Fe (Fig. 8). Simple mass balance arguments can be used to show that the enrichments in these latter elements are not likely due to residual accumulation during weathering of the parent orthoquartzite: prohibitively large thickness of quartzite would have had to weather away without loss of any Mg, K or Fe to enrich the soils adequately. On the other hand, accumulation of dust during weathering over a significant time period could explain the enrichment. Alternately, downward mobilization of fine particles from weathering of the overlying Rose Hill shale or interfingered shaley units might adequately explain the enrichment in these elements. Use of Regolith-RT-PIHM (Table 1) or WITCH (Godderis et al., 2006) to model regolith formation should allow testing of the feasibility of these or other ideas. With~~ As we build understanding, regolith formation models will be used to ~~we can also~~ extrapolate point measurements of soil thickness and porosity from catena observations to the broader Garner Run subcatchment and to other similar subcatchments in the Shavers creek watershed. In other words, the numerical models in Table 1 will be used to extend beyond the limited observations.

~~Of course, we can also augment t~~The sampling design described here is also being augmented with brief measurement campaigns ~~inside and~~ outside the subcatchments ~~or~~ and outside Shavers creek watershed as warranted. For example, while we will only monitor soil $CO_2$ continuously at a few catena positions and soil depths, we ~~can~~ augment these high frequency data with spatially extensive, but temporally limited measurements using manual soil gas samplers. Likewise, we may are characterize characterizing vegetation and surface soil properties at 3-5 additional catchments of each parent material type using the transect design that we initiated at Garner Run (Fig. 5). In general, these outside measurements will be discipline-specific excursions to understand a specific variable. This is a good example of targeted investigations that are not directly related to parameterization of the models in Table 1 for our CZO itself but are rather aimed at improving the process-based understanding that underlies models of CZ evolution. Another –example is a set of measurements that are ongoing in a catchment to the north of Shavers creek to investigate regolith formation and hillslope form in other catchments north of Shavers creek where the erosion rates is considerably fasterdiffer. At this site, we anticipate learning Such targeted investigations can also be compared to output from sensitivity tests where pertinent models are used to explore the effect of the targeted variableshow to parameterize or run models of regolith formation by exploring the impact of the rate of erosion (Table 1). Measurements outside the CZO may therefore highlight problems in our limited sampling scheme or modelling approaches that must be improved.

As we improve our understanding of the behavior of components of the critical zoneCZ, the point is to discover system-wide patterns and processes. Throughout, upscaling will remain a challenge. There is no comprehensive mathematical model of the critical zone, partly because it would be arduous to parameterize and perhaps more importantly because we do not yet understand all the interacting governing processes (Fig. 1). The research in Shavers Creek, and the work done at other critical zone CZ observatories (CZOs) around the world, is an attempt to develop a system-wide process model (or ensemble of models) and to identify the essential measurements required for parameterization. Of great interest are The most robust models we have are conceptual models that aid in understanding the CZ, but such conceptual understanding must also be encoded within and the most predictive are complex numerical simulations that allow quantitative predictions for testing. HoweverNonetheless, both conceptual and numerical models still typically include only a portion of the critical zoneCZ. To really understand WEGSS fluxes quantitatively requires We seek a model that successfully explains the dynamics between topography, groundwater levels, biota, atmospheric conditions, and regolith thickness – at present we are working mostly with conceptual relationships drawn between pairs of factors (Fig. 1).

**Formatted:** Not Highlight

**Formatted:** Not Highlight

**Formatted:** Not Highlight

In our efforts, new observations are tested against and incorporated into the PIHM models to explore the evolution of the CZ over time. In this endeavor, we can also ask, what does success look like? 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 project (earthcast) the CZ. Ultimately, success means that we gain deeper understanding of the system and can predict behavior in other places or with other datasets (e.g. tracers, water isotopes, etc.). Such testing is built in to our nested watershed approach (Fig. 3) and is also implicit to the design of the greater CZO network.

We can also imagine other indicators of success. For example, successful datasets will attract other researchers using other models. This in turn can lead to model-model inter-comparisons. If other models provide better simulations of the catchment, this will drive development of better models. One example of a model – model inter-comparison (RT-Flux-PIHM versus WITCH-Flux-PIHM, Table 1) has already driving new insights.

Another indicator of success is adoption by others of the strategies developed to study the CZ. Such strategies include design of a sampling paradigm for an individual CZO, design of a larger network of CZOs, development of suites of models, or approaches for data assimilation. While the CZO enterprise is still young, publications in the literature already attest to growth in use of the PIHM suite of models in other places (Kumar et al., 2013; Wang et al., 2013; Yu et al., 2015; Jepsen et al., 2015a; Jepsen et al., 2015b; Jepsen et al., in press.)~~(11-16)~~ and growth in use of the CZO concept worldwide (Banwart et al., 2012)~~(17)~~.

**Formatted:** English (U.S.), Not Highlight

**Formatted:** English (U.S.), Not Highlight

**Formatted:** English (U.S.), Not Highlight

## Author Contributions

S. L. Brantley, H. Lin, K. J. Davis, A. L. Neal, J. Kaye, and D. Eissenstat designed the study and wrote the sections on soil analysis, GPR, eddy correlation, GIS analysis, biogeochemical analysis, and tree root analysis respectively. R. DiBiase spearheaded the geomorphological treatment, and T. Russo and B. Hoagland led the ground and stream water research. Y. Shi performed the OSSE and worked on coupling models with PIHM. M. Kaye designed and led the vegetation surveys; L. Hill and J. Kaye designed and implemented the soil pit sensor research; A. Dere completed the soil descriptions; K. Brubaker contributed forest and land use descriptions. D. K. Arthur supervised overall data contributions.

## Acknowledgements

Field help was provided by B. Forsythe, D. Pederson, and R. Davis. D. Arthur and J. Williams are acknowledged for data organization; A. Orr for photographing the soil pits. C. Duffy is acknowledged for leadership at Shale Hills and for PIHM. S. Macdonald is acknowledged for Garner Run soil geochemistry analysis. L. Li is acknowledged for implementation of the COSMOS at Garner Run, and E. Kirby and P. Bierman for planning the cosmogenic measurements. J. Doolittle from the USDA-NRCS is acknowledged for help with GPR data collection. This work was funded by NSF Critical Zone Observatory program grants to C. Duffy (EAR 07-25019) and SLB (EAR 12-39285, EAR 13-31726). Logistical support and data were provided by the NSF-supported Shale Hills Susquehanna Critical Zone Observatory. This research was conducted in Penn State's Stone Valley Forest, which is supported and managed by the Penn State's Forestland Management Office in the College of Agricultural Sciences. The Penn State Earth and Environmental Systems Institute and College of Agricultural Sciences is acknowledged for staff and funding support.

## References


Amundson, R.: Soil formation, in: Treatise in Geochemistry: Surface and Ground Water, Weathering, and Soils, edited by: Drever, J. I., Elsevier Pergamon, Amsterdam, 1–35, 2004.

Anderson, S. P., Dietrich, W. E., and Brimhall, G. H.: Weathering profiles, mass balance analysis, and rates of solute loss: linkages between weathering and erosion in a small, steep catchment, Geol. Soc. Am. Bull., 114, 1143–1158, 2002.

Banwart, S., Menon, M., Bernasconi, S.M., Bloem, J., Blum, W.E.H., de Souza, D.M., Davidsdotir, B., Duffy, C., Lair, G.J., Kram, P., Lamacova, A., Lundin, L., Nikolaidis, N.P., Novak, M., Panagos, P., Ragnarsdottir, K.V., Reynolds, B., Robinson, D., Rousseva, S., de Ruiter, P., van Gaans, P., Weng, L., White, T., Zhang, B.: Soil processes and functions across an international network of Critical Zone Observatories: Introduction to experimental methods and initial results. Comptes Rendus Geoscience 344(11–12):758-772, 2012.

Bao, C., Li, L., Shi, Y., Sullivan, P. L., Duffy, C. J., and Brantley, S. L.: RT-Flux-PIHM: a hydrogeochemical model at the watershed scale, Water Resour. Res., submitted, 2015.

Berg, T. M., Edmunds,W. E., Geyer, A. R., and others: Geologic map of Pennsylvania, 2nd edn., Pennsylvania Geological Survey, 4th ser., Map 1, 3 sheets, scale 1 : 250,000, 1980.

Beven, K. J.: Rainfall–Runoff Modelling: The Primer, Wiley-Blackwell, 2011.

Brantley, S. L. and Lebedeva, M.: Learning to read the chemistry of regolith to understand the Critical Zone, Annu. Rev. Earth Pl. Sc., 39, 387–416, 2011.

Brimhall, G. and Dietrich, W. E.: Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenisis, Geochim. Cosmochim. Ac., 51, 567–587, 1987.

Carter, B. and Ciolkosz, E.: Slope gradient and aspect effects on soils developed form sandstone in Pennsylvania, Geoderma, 49, 199–213, 1991.

Chadwick, O. A. and Chorover, J.: The chemistry of pedogenic thresholds, Geoderma, 100, 321–353, 2001.

Ciolkosz, E. J., Carter, B., Hoover, M. T., Cronce, R., Waltman, W., and Dobos, R.: Genesis of soils and landscapes in the Ridge and Valley province of central Pennsylvania, Geomorphology, 3, 245–261, 1990.

Cook, K. L., Whipple, K. X., Heimsath, A. M., and Hanks, T. C.: Rapid incision of the Colorado River in Glen Canyon – insights from channel profiles, local incision rates, and modeling of lithologic controls, Earth Surf. Proc. Land., 34, 994–1010, 2009.

Cotter, E.: Tuscarora formation of Pennsylvania : guidebook, Society of Economic Paleontologists and Mineralogists, Eastern Section, 1982 field trip, Lewisburg, Pa., Dept. of Geology and Geophysics, Bucknell University, 1982.

Crutzen, P. J.: Geology of mankind, Nature, 415, 23–23, 2002.

Dietrich, W. E., Bellugi, D. G., Sklar, L. S., Stock, J. D., Heimsath, A. M., and Roering, J. J.: Geomorphic Transport Laws for Predicting Landscape Form and Dynamics, Dynamics, In Prediction, in: Geomorphology, Geophyscial Monograph, edited by: Wilcock, P. R., Iverson, R. M., 135, American Geophysical Union, 103–132, doi:10.1029/135GM09, 2003.

Duffy, C., Shi, Y., Davis, K., Slingerland, R., Li, L., Sullivan, P. L., Godderis, Y., and Brantley, S. L.: Designing a suite of models to explore critical zone function, Procedia Earth and Planetary Science, 10, 7–15, doi:10.1016/j.proeps.2014.08.003, 2014.

Duvall, A., Kirby, E., and Burbank, D.: Tectonic and lithologic controls on bedrock channel profiles and processes in coastal California. J. Geophys. Res.-Earth, 109, F03002 doi:10.1029/2003JF000086, 2004.

Ewing, S. A., Sutter, B., Owen, J., Nishiizumi, K., Sharp, W., Cliff, S. S., Perry, K., Dietrich, W., McKay, C. P., and Amundson, R.: A threshold in soil formation at Earth's arid-hyperarid transition, Geochim. Cosmochim. Ac., 70, 5293–5322, 2006.

Fike, J. Terrestrial and palustrine plant communities of Pennsylvania. Bureau of Forestry, PA. Department of Conservation and Natural Resources, 1999.

Flueckinger, L. A.: Report 176, Geology of a portion of the Allensville Quadrangle, Centre and Huntingdon Counties, Pennsylvania, Bureau of Topographic and Geologic Survey, Commonwealth of Pennsylvania State Planning Board, 1969.

Franz, T. E., Zreda, M., Rosolem, R., and Ferre, T. P. A.: A universal calibration function for determination of soil moisture with cosmic-ray neutrons, Hydrol. Earth Syst. Sci., 17, 453–460, doi:10.5194/hess-17-453-2013, 2013.

Gardner, T. W., Ritter, J. B., Shuman, C. A., Bell, J. C., Sasowsky, K. C., and Pinter, N.: A periglacial stratified slope deposit in the valley and ridge province of central Pennsylvania, USA: Sedimentology, stratigraphy, and geomorphic evolution, Permafrost Periglac., 2, 141–162, 1991.

Godderis, Y. and Brantley, S. L.: Earthcasting the future Critical Zone, Elementa, 1, 1–10, doi:10.12952/journal.elementa.000019, 2014.

Godderis, Y., Francois, L., Probst, A., Schott, J., Moncoulon, D., Labat, D., and Viville, D.: Modelling weathering processes at the catchment scale: the WITCH numerical model, Geochim. Cosmochim. Ac., 70, 1128–1147, 2006.

Godsey, S. E., Kirchner, J. W., and Clow, D. W.: Concentration-discharge relationships reflect chemostatic characteristics of US catchments, Hydrol. Process., 23, 1844–1864, doi:10.1002/hyp.7315, 2009.

Graham, C. and Lin, H. S.: Controls and frequency of preferential flow occurrence: a 175-event analysis, Vadose Zone J., 10, 816–831, 2011.

Graham, C. G. and Lin, H. S.: Spatial–temporal patterns of preferential flow occurrence in the Shale Hills CZO based on real-time soil moisture monitoring, Vadose Zone J., special issue, 10, 816–831, doi:10.2136/vzj2010.0119, 2010.

Guo, L., Chen, J., and Lin, H.: Subsurface lateral flow network on a hillslope revealed by timelapse ground penetrating radar, Water Resour. Res., 50, 21, 2014.

Hasenmueller, E., Jin, L., Stinchcomb, G., Lin, H., Brantley, S. L., and Kaye, J. P.: Topographic controls on the depth distribution of soil CO2 in a small temperate watershed, Appl. Geochem., 63, 58–69, doi:10.1016/j.apgeochem.2015.07.005, 2015.

Hattanji, T. and Onda, Y.: Coupling of runoff processes and sediment transport in mountainous watersheds underlain by different sedimentary rocks, Hydrol. Process., 18, 623–636, doi:10.1002/hyp.1262, 2004.

Hoekstra, P., and A. Delaney: Dielectric properties of soils at UHF and microwave frequencies, J. Geophys. Res., 79 (11), 1699–1708, doi:10.1029/JB079i011p01699, 1974.

Homer, C.G., Dewitz, J.A., Yang, L., Jin, S., Danielson, P., Xian, G., Coulston, J., Herold, N.D., Wickham, J.D., and Megown, K.: Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information. Photogrammetric Engineering and Remote Sensing, v. 81 (5), 345-354, 2015.

Hoover, M. T. and Ciolkosz, E. J.: Colluvial soil parent material relationships in the Ridge and Valley physiographic province of Pennsylvania, Soil Sci., 145, 163–172, 1988.

Jackson, R.B., Anderson, L.J., Pockman, W.T.: Measuring water availability and uptake in ecosystem studies, Methods in Ecosystem Science, Springer New York, 199-214, doi:10.1007/978-1-4612-1224-9_14, 2000.

Jepsen, S.M., Harmon, T., Meadows, M., and Hunsaker, C.: Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, USA. in American Geophysical Union Annual Meeting (American Geophysical Union, San Francisco), https://agu.confex.com/agu/fm15/meetingapp.cgi/Paper/70885, 2015a.

Jepsen, S.M., Harmon, T.C., and Shi, Y.: Assessment of a watershed model calibrated to the shape of baseflow recession with and without evapotranspiration effects. Water Resources Research, in revision, 2015b.

Jepsen, S.M., Harmon, T.C., Meadows, M.W., and Hunsaker, C.T.: Hydrogeologic influence on changes in snowmelt runoff with climate warming: Numerical experiments on a mid-elevation catchment in the Sierra Nevada, U.S.A. Journal of Hydrology, in press.

Jin, L. and Brantley, S. L.: Soil chemistry and shale weathering on a hillslope influenced by convergent hydrologic flow regime at the Susquehanna/Shale Hills Critical Zone Observatory. Applied Geochemistry 26:S51–S56, doi:10.1016/j.apgeochem.2011.03.027, 2011.

Jin, L., Ravella, R., Ketchum, B., Bierman, P. R., Heaney, P., White, T. S., and Brantley, S. L.: Mineral weathering and elemental transport during hillslope evolution at the Susquehanna/Shale Hills Critical Zone Observatory, Geochim. Cosmochim. Ac., 74, 3669–3691, 2010.

Jin, L., Andrews, D. M., Holmes, G. H., Lin, H., and Brantley, S. L.: Opening the “black box”: water chemistry reveals hydrological controls on weathering in the Susquehanna Shale Hills Critical Zone Observatory, Vadose Zone J., 10, 928–942, doi:10.2136/vzj2010.0133, 2011.

Jin, L., Ogrinc, N., Yesavage, T., Hasenmueller, E. A., Ma, L., Sullivan, P. L., Kaye, J., Duffy, C., and Brantley, S. L.: The CO2 consumption potential during gray shale weathering:

insights from the evolution of carbon isotopes in the Susquehanna Shale Hills critical zone observatory, Geochim. Cosmochim. Ac., 142, 260–280, doi:10.1016/j.gca.2014.07.006, 2014.

Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P., Christensen, T. R., Cramer, W., Matveyeva, N. V., McGuire, A. D., Murray, D. F., Razzhivin, V. Y., Smith, B., Walker, A. D., Anderson, P. M., Andreev, A. A., Brubaker, L. B., Edwards, M. E., and Lozhkin, A. V.: Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, amid future projections, J. Geophys. Res., 108, 8171, doi:10.1029/2002JD002559, 2003.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, P., Myneni, R. B., Nemani, R. R., and Running, S. W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm, Theoretical Basis Document, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA, 1999.

Kumar, M., Marks, D., Dozier, J., Reba, M., and Winstral, A.: Evaluation of distributed hydrologic impacts of temperature-index and energy-based snow models. Advances in Water Resources 56:77-89, 2013.

Lebedeva, M. and Brantley, S. L.: Exploring geochemical controls on weathering and erosion of convex hillslopes: beyond the empirical regolith production function, Earth Surf. Proc. Land., 38, 1793–1807, doi:10.1002/esp.3424, 2013.

Lebedeva, M. I., Fletcher, R. C., Balashov, V. N., and Brantley, S. L.: A reactive diffusion model describing transformation of bedrock to saprolite, Chem. Geol., 244, 624–645, 2007.

Lebedeva, M. I., Fletcher, R. C., and Brantley, S. L.: A mathematical model for steady-state regolith production at constant erosion rate, Earth Surf. Proc. Land., 35, 508–524, 2010.

Lin, H.S., Kogelmann, W., Walker, C., and Bruns, M.A.: Soil moisture patterns in a forested catchment: A hydropedological perspective. Geoderma 131:345-368, doi:10.1016/j.geoderma.2005.03.013, 2006.

Lynch, J. A.: Effects of Antecedent Soil Moisture on Storm Hydrographs, Doctor of Philosophy, Forestry, The Pennsylvania State University, University Park, PA, 1976.

Ma, L., Jin, L., and Brantley, S. L.: How mineralogy and slope aspect affect REE release and fractionation during shale weathering in the Susquehanna/Shale Hills Critical Zone Observatory, Chem. Geol., 290, 31–49, 2011.

Ma, L., Chabaux, F., West, N., Kirby, E., Jin, L., and Brantley, S. L.: Regolith production and transport in the Susquehanna Shale Hills Critical Zone Observatory, Part 1: Insights from U-series isotopes, J. Geophys. Res.-Earth, 118, 722–740, doi:10.1002/jgrf.20037, 2013.

Meinzer, F. C., Woodruff, D. R., Eissenstat, D. M., Lin, H. S., Adams, T. S., and McCulloh, K. A.: Above- and belowground controls on water use by trees of different wood types in an eastern US deciduous forest, Tree Physiol., 33, 345–356, 2013.

Merritts, D.J., Walter, R.C., Blair, A., Demitroff, M., Potter Jr., N., Alter, S., Markey, E., Guillorn, S., Gigliotti, S., and Studnicky, C.: LIDAR, orthoimagery, and field analysis of periglacial landforms and their cold climate signature, unglaciated Pennsylvania and Maryland: Abstract 329-1 presented at 2015 Annual Meeting, GSA, Baltimore, MD, 1–4 November, 47 (7), 831, 2015.

Miller, S. R., Sak, P. B., Kirby, E., and Bierman, P. R.: Neogene rejuvenation of central Appalachian topography: evidence for differential rock uplift from stream profiles and erosion rates, Earth Planet. Sc. Lett., 369–370, 13, 2013.

Minasny, B., McBratney, A. B., and Salvador-Blanes, S.: Quantitative models for pedogenesis – a review, Geoderma, 144, 140–157, 2008.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y.,Wang, Y., Song, X., Zhang, Y., Smith, G. R., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., and Running, S. W.: Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data, Remote Sens. Environ., 83, 214–231, 2002.

Naithani, K. J., Baldwin, D., Gaines, K., Lin, H. S., and Eissenstat, D. M.: Spatial distribution of tree species governs the spatio-temporal interaction of leaf area index and soil moisture across a landscape, PloS ONE, 8, 12, doi:10.1371/journal.pone.0058704, 2013.

Niu, X., Williams, J. Z., Miller, D., Lehnert, K. A., Bills, B., and Brantley, S. L.: An ontology driven relational geochemical database for the Earth's Critical Zone: CZchemDB, J. Environ. Inf., 23, 10–23, doi:10.3808/jei.201400266, 2014.

Portenga, E. W., Bierman, P. R., Rizzo, D. M., and Rood, D. H.: Low rates of bedrock outcrop erosion in the central Appalachian Mountains inferred from in situ Be-10, Geol. Soc. Am. Bull., 125, 201–215, 2013.

Qu, Y. and Duffy, C. J.: A semi-discrete finite volume formulation for multiprocess watershed simulation, Water Resour. Res., 43, 1–18, 2007.

Shi, Y., Davis, K. J., Duffy, C. J., and Yu, X.: Development of a coupled land surface hydrologic model and evaluation at a critical zone observatory, J. Hydrometeorol., 14, 1401–1420, 2013.

Shi, Y., K. J. Davis, F. Zhang, C. J. Duffy, and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using the ensemble Kalman Filter: A synthetic experiment. Water Resources Research, 50, 706-724, doi:10.1002/2013WR014070, 2014.

Shi, Y., Baldwin, B. C., Davis, K. J., Yu, X., Duffy, C. J., and Lin, H.: Simulating high resolution soil moisture patterns in the Shale Hills catchment using a land surface hydrologic model, Hydrol. Process., doi:10.1002/hyp.10593, 2015a.

Shi, Y., Davis, K. J., Zhang, F., Duffy, C., J., and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: a multivariate real data experiment, Adv. Water Resour., 83, 421–427, doi:10.1016/j.advwatres.2015.06.009, 2015b.

Sullivan, P. L., Hynek, S., Gu, X., Singha, K., White, T. S.,West, N., Kim, H., Clarke, B., Kirby, E., Duffy, C., and Brantley, S. L.: Oxidative dissolution under the channel leads geomorphological evolution at the Shale Hills catchment, Am. J. Sci., submitted, 2015.

Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resour. Res., 33, 309–319, doi:10.1029/96WR03137, 1997.

Thomas, E., Lin, H., Duffy, C., Sullivan, P., Holmes, G. H., Brantley, S. L., and Jin, L.: Spatiotemporal patterns of water stable isotope compositions at the Shale Hills Critical Zone: linkages to subsurface hydrologic processes, Vadose Zone J., 12, doi:10.2136/vzj2013.01.0029, 2013.

Topp, G.C., Davis, J.L., and Annan, A.P.: Electromagnetic determination of soil water content: Measurement in coaxial transmission lines, Water Resour. Res., 16, 574-582, 1980.

Topp, G. & Ferre, P.-A.: Water content, in Method of soil analysis, Part 4, Physical methods, SSSA Book, vol.5 ed. J.H. Dane and G.C. Topp, Soil Sci. Soc. of Am., Madison, Wis., 417-446, 2002.

Vitousek, P. M., Aber, J. D., Horwath, R. W., Lubchenco, J., and Melillo, J. M.: Human alteration of the global nitrogen cycle, Science, 277, 494–499, 1997a.

Vitousek, P. M., Mooney, H. A., Lubchenco, J., and Melillo, J. M.: Human domination of Earth's ecosystems, Science, 277, 494–499, 1997b.

Wang, R., Kumar, M., and Marks, D.: Anomalous trend in soil evaporation in a semi-arid, snow-dominated watershed. Advances in Water Resources 57:32-40, 2013.

Warnant, P., Francois, L., Strivay, D., and Gerard, J.-C.: CARAIB: a global model of terrestrial biological productivity, Global Biogeochem. Cy., 8, 255–270, 1994.

West, N., Kirby, E., Bierman, P. R., Slingerland, R., Ma, L., Rood, D., and Brantley, S. L.: Regolith production and transport at the Susquehanna Shale Hills Critical Zone Observatory: Part 2 – Insights from meteoric 10Be. J. Geophys. Res.-Earth, 118, 1–20, doi:10.1002/jgrf.20121, 2013.

West, N., Kirby, E., Bierman, P. R., and Clarke, B. A.: Aspect-dependent variations in regolith creep revealed by meteoric 10Be, Geology, 42, 507–510, doi:10.1130/G35357.1, 2014.

Whipple, K. X., Dibiase, R. A., and Crosby, B. T.: Bedrock rivers, in: Treatise in Geomorphology, Fluvial Geomorphology, edited by: Wohl, E., Academic Press, San Diego, CA, 550–573, 2013.

White T., Brantley S., Banwart S., Chorover J., Dietrich W., Derry L., Lohse K., Anderson S., Aufdendkampe A., Bales R., Kumar P., Richter D., McDowell B.: The Role of Critical Zone Observatories in Critical Zone Science, in Developments in Earth Surface Processes 19, Principles and Dynamics of the Critical Zone (eds. Giardino, J., and Houser, C.), Elsevier, 15–78, 2015.

Wilkinson, B. H. and McElroy, B. J.: The impact of humans on continental erosion and sedimentation, Geol. Soc. Am. Bull., 119, 140–156, 2007.

Willett, S. D., McCoy, S. W., Perron, J. T., Goren, L., and Chen, C. Y.: Dynamic reorganization of river basins, Science, 343, 1–9, doi:10.1126/science.1248765, 2014.

Williams, J., Reynolds, R., Franzi, D., Romanowics, E., and Paillet, F.: Hydrogeology of the Potsdam sandstone in northern New York, Can. Water Resour. J., 35, 399–415, 2010.

Williard, K., Dewalle, D., and Edwards, P.: Influence of bedrock geology and tree species composition on stream nitrate concentrations in mid-Appalachian forested watersheds, Water Air Soil Poll., 160, 55–76, 2005.

Yu, X., Lamačová, A., Duffy, C., Krám, P., Hruška, J., White, T., and Bhatt, G.: Modelling long-term water yield effects of forest management in a Norway spruce forest. Hydrol. Sci. J.-J. Sci. Hydrol. 60(2):174-191, doi: 10.1080/02626667.2014.897406, 2015.

Zhang, J., Lin, H. S., and Doolittle, J.: Soil layering and preferential flow impacts on seasonal changes of GPR signals in two contrasting soils, Geoderma, 213, 560–569, doi:10.1016/j.geoderma.2013.08.035, 2014.

Zhu, Q., Lin, H., and Doolittle, J.: Repeated electromagnetic induction surveys for determining subsurface hydrologic dynamics in an agricultural landscape, Soil Sci. Soc. Am. J., 74, 12, 2010a.

Zhu, Q., Lin, H., and Doolittle, J.: Repeated electromagnetic induction surveys for improved soil mapping in an agricultural landscape, Soil Sci. Soc. Am. J., 74, 1763–1774, doi:10.2136/sssaj2010.0056, 2010b.

Zreda, M., Shuttleworth, W. J., Zeng, X., Zweck, C., Desilets, D., Franz, T., and Rosolem, R.: Corrigendum to "COSMOS: the COsmic-ray Soil Moisture Observing System" published in Hydrol. Earth Syst. Sci., 16, 4079–4099, 2012, Hydrol. Earth Syst. Sci., 17, 1065–1066, doi:10.5194/hess-17-1065-2013, 2013.

Zweck, C., Zreda, M., and Desilets, D.: Snow shielding factors for cosmogenic nuclide dating inferred from Monte Carlo neutron transport simulations, Earth Planet. Sc. Lett., 379, 64–71, 2013.

~~Amundson, R. (2004) Soil Formation, in: Drever, J.I. (Ed.), Treatise in Geochemistry: Surface and Ground Water, Weathering, and Soils. Elsevier Pergamon, Amsterdam, pp. 1-35.~~

~~Anderson, S.P., Dietrich, W.E. and Brimhall, G.H. (2002) Weathering profiles, mass balance analysis, and rates of solute loss: Linkages between weathering and erosion in a small, steep catchment. Geological Society of America Bulletin 114(9), 1143-1158.~~

Bao, C., Li, L., Shi, Y., Sullivan, P.L., Duffy, C., J. and Brantley, S.L. (2015, subm.) RT-Flux-PIHM: A hydrogeochemical model at the watershed scale. Water Resour. Res., submitted, 2015.

Berg, T.M., Edmunds, W.E., Geyer, A.R. and others, a. (1980) Geologic map of Pennsylvania (2nd ed.). Pennsylvania Geological Survey, 4th ser., Map 1, 3 sheets, scale 1:250,000.

Beven, K.J. (2011) Rainfall-Runoff Modelling: The Primer. Wiley-Blackwell.

Brantley, S.L. and Lebedeva, M. (2011) Learning to read the chemistry of regolith to understand the Critical Zone. Annual Review Earth Planetary Science 39, 387-416.

Brimhall, G. and Dietrich, W.E. (1987) Constitutive mass balance relations between chemical composition, volume, density, porosity, and strain in metasomatic hydrochemical systems: results on weathering and pedogenisis. Geochimica et Cosmochimica Acta 51, 567-587.

Carter, B. and Ciolkosz, E. (1991) Slope gradient and aspect effects on soils developed form sandstone in Pennsylvania. Geoderma 49, 199-213.

Chadwick, O.A. and Chorover, J. (2001) The chemistry of pedogenic thresholds. Geoderma 100, 321-353.

Ciolkosz, E.J., Carter, B., Hoover, M.T., Cronce, R., Waltman, W. and Dobos, R. (1990) Genesis of soils and landscapes in the Ridge and Valley province of central Pennsylvania. Geomorphology 3, 245-261.

Cook, K.L., Whipple, K.X., Heimsath, A.M. and Hanks, T.C. (2009) Rapid incision of the Colorado River in Glen Canyon – insights from channel profiles, local incision rates, and modeling of lithologic controls. Earth Surface Processes and Landforms 34, 994-1010.

Cotter, E.: Tuscarora formation of Pennsylvania: guidebook, Society of Economic Paleontologists and Mineralogists, Eastern Section, 1982 field trip, Lewisburg, Pa., Dept. of Geology and Geophysics, Bucknell University, 1982.

Crutzen, P.J. (2002) Geology of mankind. Nature 415, 23-23.

Dietrich, W. E., Bellugi, D. G., Sklar, L. S., Stock, J. D., Heimsath, A. M., and Roering, J. J.: Geomorphic Transport Laws for Predicting Landscape Form and Dynamics, Dynamics, In Prediction, in: Geomorphology, Geophyscial Monograph, edited by: Wilcock, P. R., Iverson, R. M., 135, American Geophysical Union, 103–132, doi:10.1029/135GM09, 2003.

Duffy, C., Shi, Y., Davis, K., Slingerland, R., Li, L., Sullivan, P. L., Godderis, Y., and Brantley, S. L.: Designing a suite of models to explore critical zone function, Procedia Earth and Planetary Science, 10, 7–15, doi:10.1016/j.proeps.2014.08.003, 2014.

Duvall, A., Kirby, E., and Burbank, D.: Tectonic and lithologic controls on bedrock channel profiles and processes in coastal California. J. Geophys. Res.-Earth, 109, F03002, doi:10.1029/2003JF000086, 2004.

Ewing, S.A., Sutter, B., Owen, J., Nishiizumi, K., Sharp, W., Cliff, S.S., Perry, K., Dietrich, W., McKay, C.P. and Amundson, R. (2006) A threshold in soil formation at Earth's arid-hyperarid transition. Geochimica Cosmochimica Acta 70, 5293-5322.

Flueckinger, L.A. (1969) Report 176, Geology of a portion of the Allensville Quadrangle, Centre and Huntingdon Counties, Pennsylvania. Bureau of Topographic and Geologic Survey, Commonwealth of Pennsylvania State Planning Board

Franz, T.E., Zreda, M., Rosolem, R. and Ferre, T.P.A. (2013) A universal calibration function for determination of soil moisture with cosmic-ray neutrons. Hydrol. Earth Syst. Sci. 17, 453-460.

Gardner, T.W., Ritter, J.B., Shuman, C.A., Bell, J.C., Sasowsky, K.C. and Pinter, N. (1991) A periglacial stratified slope deposit in the valley and ridge province of central Pennsylvania, USA: Sedimentology, stratigraphy, and geomorphic evolution. Permafrost and Periglacial Processes 2, 141-162.

Godderis, Y. and Brantley, S. L.: Earthcasting the future Critical Zone, Elementa, 1, 1-10, doi:10.12952/journal.elementa.000019, 2014.

Godderis, Y., Francois, L., Probst, A., Schott, J., Moncoulon, D., Labat, D. and Viville, D. (2006) Modelling weathering processes at the catchment scale: The WITCH numerical model. Geochimica et Cosmochimica Acta 70, 1128-1147.

Godsey, S.E., Kirchner, J.W. and Clow, D.W. (2009) Concentration-discharge relationships reflect chemostatic characteristics of US catchments. Hydrological Processes 23, 1844-1864, doi:1810.1002/hyp.7315.

Graham, C. and Lin, H.S. (2011) Controls and frequency of preferential flow occurrence: A 175-event analysis. Vadose Zone Journal 10, 816-831.

Graham, C. G. and Lin, H. S.: Spatial–temporal patterns of preferential flow occurrence in the Shale Hills CZO based on real-time soil moisture monitoring, Vadose Zone J., special issue, 10, 816–831, doi:10.2136/vzj2010.0119, 2010.

Guo, L., Chen, J. and Lin, H. (2014) Subsurface Lateral Flow Network on a Hillslope Revealed by Time-lapse Ground Penetrating Radar. Water Resources Research 50, 21.

Hasenmueller, E., Jin, L., Stinchcomb, G., Lin, H., Brantley, S. L., and Kaye, J. P.: Topographic controls on the depth distribution of soil CO2 in a small temperate watershed, Appl. Geochem., 63, 58–69, doi:10.1016/j.apgeochem.2015.07.005, 2015.

Hattanji, T. and Onda, Y. (2004) Coupling of runoff processes and sediment transport in mountainous watersheds underlain by different sedimentary rocks. Hydrological Processes 18, 623-636, doi:610.1002/hyp.1262.

Hoover, M.T. and Ciolkosz, E.J. (1988) Colluvial soil parent material relationships in the Ridge and Valley physiographic province of Pennsylvania. Soil Sci. 145, 163-172.

Jin, L., Andrews, D.M., Holmes, G.H., Lin, H. and Brantley, S.L. (2011) Opening the "black box": water chemistry reveals hydrological controls on weathering in the Susquehanna Shale Hills Critical Zone Observatory. Vadose Zone Journal 10, 928-942, doi:910.2136/vzj2010.0133.

Jin, L., Ogrinc, N., Yesavage, T., Hasenmueller, E. A., Ma, L., Sullivan, P. L., Kaye, J., Duffy, C., and Brantley, S. L.: The $CO_2$ consumption potential during gray shale weathering: insights from the evolution of carbon isotopes in the Susquehanna Shale Hills critical zone observatory, Geochim. Cosmochim. Ac., 142, 260–280, doi:10.1016/j.gca.2014.07.006, 2014.

Jin, L., Ravella, R., Ketchum, B., Bierman, P.R., Heaney, P., White, T.S. and Brantley, S.L. (2010) Mineral weathering and elemental transport during hillslope evolution at the Susquehanna/Shale Hills Critical Zone Observatory. Geochimica et Cosmochimica Acta 74, 3669-3691.

Kaplan, J. O., Bigelow, N. H., Prentice, I. C., Harrison, S. P., Bartlein, P., Christensen, T. R., Cramer, W., Matveyeva, N. V., McGuire, A. D., Murray, D. F., Razzhivin, V. Y., Smith, B., Walker, A. D., Anderson, P. M., Andreev, A. A., Brubaker, L. B., Edwards, M. E., and Lozhkin, A. V.: Climate change and Arctic ecosystems: 2. Modeling, paleodata-model comparisons, amid future projections, J. Geophys. Res., 108, 8171, doi:10.1029/2002JD002559, 2003.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., Votava, P., Myneni, R. B., Nemani, R. R., and Running, S. W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm, Theoretical Basis Document, NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA, 1999.

Lebedeva, M. and Brantley, S. L.: Exploring geochemical controls on weathering and erosion of convex hillslopes: beyond the empirical regolith production function, Earth Surf. Proc. Land., 38, 1793–1807, doi:10.1002/esp.3424, 2013.

Lebedeva, M.I., Fletcher, R.C., Balashov, V.N. and Brantley, S.L. (2007) A reactive diffusion model describing transformation of bedrock to saprolite. Chemical Geology 244, 624-645.

Lebedeva, M.I., Fletcher, R.C. and Brantley, S.L. (2010) A mathematical model for steady-state regolith production at constant erosion rate. Earth Surface Processes and Landforms 35, 508-524.

Lynch, J. A.: Effects of Antecedent Soil Moisture on Storm Hydrographs, Doctor of Philosophy, Forestry, The Pennsylvania State University, University Park, PA, 1976.

Ma, L., Chabaux, F., West, N., Kirby, E., Jin, L., and Brantley, S. L.: Regolith production and transport in the Susquehanna Shale Hills Critical Zone Observatory, Part 1: Insights from U-series isotopes, J. Geophys. Res.-Earth, 118, 722–740, doi:10.1002/jgrf.20037, 2013.

Ma, L., Jin, L. and Brantley, S.L. (2011) How mineralogy and slope aspect affect REE release and fractionation during shale weathering in the Susquehanna/Shale Hills Critical Zone Observatory. Chemical Geology 290, 31-49.

Meinzer, F.C., Woodruff, D.R., Eissenstat, D.M., Lin, H.S., Adams, T.S. and McCulloh, K.A. (2013) Above- and belowground controls on water use by trees of different wood types in an eastern US deciduous forest. Tree Physiology 33, 345-356.

Miller, S.R., Sak, P.B., Kirby, E. and Bierman, P.R. (2013) Neogene rejuvenation of central Appalachian topography: Evidence for differential rock uplift from stream profiles and erosion rates. Earth and Planetary Science Letters 369-370, 13.

Minasny, B., McBratney, A.B. and Salvador-Blanes, S. (2008) Quantitative models for pedogenesis - A review. Geoderma 144, 140-157.

Myneni, R.B., Hoffman, S., Knyazikhin, Y., Privette, J.L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G.R., Lotsch, A., Friedl, M., Morisette, J.T., Votava, P., Nemani, R.R. and Running, S.W. (2002) Global products of vegetation leaf area and fraction absorbed PAR from year one of MODIS data. Remote Sens. Environ. 83, 214-231.

Naithani, K. J., Baldwin, D., Gaines, K., Lin, H. S., and Eissenstat, D. M.: Spatial distribution of tree species governs the spatio-temporal interaction of leaf area index and soil moisture across a landscape, PloS ONE, 8, 12, doi:10.1371/journal.pone.0058704, 2013.

Niu, X., Williams, J. Z., Miller, D., Lehnert, K. A., Bills, B., and Brantley, S. L.: An ontology driven relational geochemical database for the Earth's Critical Zone: CZchemDB, J. Environ. Inf., 23, 10–23, doi:10.3808/jei.201400266, 2014.

Portenga, E.W., Bierman, P.R., Rizzo, D.M. and Rood, D.H. (2013) Low rates of bedrock outcrop erosion in the central Appalachian Mountains inferred from in situ Be-10. Geological Society of America Bulletin 125, 201-215.

Qu, Y. and Duffy, C.J. (2007) A semi-discrete finite volume formulation for multiprocess watershed simulation. Water Resources Research 43, 1-18.

Shi, Y., Baldwin, B. C., Davis, K. J., Yu, X., Duffy, C. J., and Lin, H.: Simulating high resolution soil moisture patterns in the Shale Hills catchment using a land surface hydrologic model, Hydrol. Process., doi:10.1002/hyp.10593, 2015a.

Shi, Y., Davis, K.J., Duffy, C.J. and Yu, X. (2013) Development of a coupled land surface hydrologic model and evaluation at a critical zone observatory. Journal of Hydrometeorology 14, 1401-1420.

Shi, Y., Davis, K. J., Zhang, F., Duffy, C., J., and Yu, X.: Parameter estimation of a physically-based land surface hydrologic model using an ensemble Kalman filter: a multivariate real-data experiment, Adv. Water Resour., 83, 421–427, doi:10.1016/j.advwatres.2015.06.009, 2015b.

Sullivan, P. L., Hynek, S., Gu, X., Singha, K., White, T. S., West, N., Kim, H., Clarke, B., Kirby, E., Duffy, C., and Brantley, S. L.: Oxidative dissolution under the channel leads geomorphological evolution at the Shale Hills catchment, Am. J. Sci., submitted, 2015.

~~Tarboton, D. G.: A new method for the determination of flow directions and upslope areas in grid digital elevation models, Water Resour. Res., 33, 309–319, doi:10.1029/96WR03137, 1997.~~

~~Thomas, E., Lin, H., Duffy, C., Sullivan, P., Holmes, G. H., Brantley, S. L., and Jin, L.: Spatiotemporal patterns of water stable isotope compositions at the Shale Hills Critical Zone: linkages to subsurface hydrologic processes, Vadose Zone J., 12, doi:10.2136/vzj2013.01.0029, 2013.~~

~~Vitousek, P.M., Aber, J.D., Horwath, R.W., Lubchenco, J. and Melillo, J.M. (1997a) Human alteration of the global nitrogen cycle. Science 277, 494-499.~~

~~Vitousek, P.M., Mooney, H.A., Lubchenco, J. and Melillo, J.M. (1997b) Human domination of Earth's ecosystems. Science 277, 494-499.~~

~~Warnant, P., Francois, L., Strivay, D. and Gerard, J. C. (1994) CARAIB: A global model of terrestrial biological productivity. Global Biogeochemical Cycles 8, 255-270.~~

~~West, N., Kirby, E., Bierman, P. R., Slingerland, R., Ma, L., Rood, D., and Brantley, S. L.: Regolith production and transport at the Susquehanna Shale Hills Critical Zone Observatory: Part 2 – Insights from meteoric 10Be. J. Geophys. Res.-Earth, 118, 1–20, doi:10.1002/jgrf.20121, 2013.~~

~~West, N., Kirby, E., Bierman, P. R., and Clarke, B. A.: Aspect-dependent variations in regolith creep revealed by meteoric 10Be, Geology, 42, 507–510, doi:10.1130/G35357.1, 2014.~~

~~Whipple, K.X., Dibiase, R.A. and Crosby, B.T. (2013) Bedrock rivers, in: Wohl, E. (Ed.), Treatise in Geomorphology, Fluvial Geomorphology. Academic Press, San Diego, CA, pp. 550-573.~~

~~Wilkinson, B.H. and McElroy, B.J. (2007) The impact of humans on continental erosion and sedimentation. Geological Society of American Bulletin 119, 140-156.~~

~~Willett, S. D., McCoy, S. W., Perron, J. T., Goren, L., and Chen, C. Y.: Dynamic reorganization of river basins, Science, 343, 1–9, doi:10.1126/science.1248765, 2014.~~

~~Williams, J., Reynolds, R., Franzi, D., Romanowics, E. and Paillet, F. (2010) Hydrogeology of the Potsdam sandstone in northern New York. Canadian Water Resources Journal 35, 399-415.~~

Williard, K., Dewalle, D. and Edwards, P. (2005) Influence of bedrock geology and tree species composition on stream nitrate concentrations in mid-Appalachian forested watersheds. Water, Air, and Soil Pollution 160, 55-76.

Zhang, J., Lin, H. S., and Doolittle, J.: Soil layering and preferential flow impacts on seasonal changes of GPR signals in two contrasting soils, Geoderma, 213, 560-569, doi:10.1016/j.geoderma.2013.08.035, 2014.

Zhu, Q., Lin, H. and Doolittle, J. (2010a) Repeated Electromagnetic Induction Surveys for Determining Subsurface Hydrologic Dynamics in an Agricultural Landscape. Soil Science Society of America Journal 74, 12.

Zhu, Q., Lin, H., and Doolittle, J.: Repeated electromagnetic induction surveys for improved soil mapping in an agricultural landscape, Soil Sci. Soc. Am. J., 74, 1763-1774, doi:10.2136/sssaj2010.0056, 2010b.

Zreda, M., Shuttleworth, W.J., Zeng, X., Zweck, C., Desilets, D., Franz, T. and Rosolem, R. (2013) COSMOS: the COsmic-ray Soil Moisture Observing System (vol 16, pg 4079, 2012). Hydrol. Earth Syst. Sci. 17, 1065-1066.

Zweck, C., Zreda, M. and Desilets, D. (2013) Snow shielding factors for cosmogenic nuclide dating inferred from Monte Carlo neutron transport simulations. Earth and Planetary Science Letters 379, 64-71.

Table 1. Designing a suite of CZ models

| | Modeling purpose | Model | Timescale of interest |
|---|---|---|---|
| Numerical models in use at SSHCZO | Topography (land scape evolution) | LE-PIHM | Days—millions of years |
| | Regolith composition and structure | Regolith-RT-PIHM, WITCH[a] | Hours—millions of years |
| | Distribution of biota | BIOME4[b], CARAIB[c], ED2 | Days—centuries |
| | C and N pools and fluxes | Flux-PIHM-BGC | Days—decades |
| | Sediment fluxes | PIHM-SED | Hours—decades |
| | Solute chemistry and fluxes | RT-Flux-PIHM[d], WITCH | Hours—decades |
| | Soil $CO_2$ concentration and fluxes | CARAIB | Hours—decades |
| | Energy and hydrologic fluxes | PIHM[d], Flux-PIHM[f] | Hours—decades |
| Geological factor | Uplift rate, bedrock composition, bedrock physical properties, pre-existing geological factors such as glaciation | | |
| External driver | Energy inputs, chemistry of wet and dry deposition, atmospheric composition, climate conditions, anthropogenic activities | | |

a: ~~(~~Godderis et al~~.,~~. (2006)

b: ~~(~~Kaplan et al~~.,~~. (2003)

c: ~~(~~Warnant et al~~.,~~. (1994)

d: ~~(~~Bao et al., 2015~~,~~ (subm.)

e: ~~(~~Qu and Duffy~~,~~ (2007)

f: ~~(~~Shi et al~~.,~~. (2013)

Table 2. Measurements and instrumentation for Tower HOG system

| Measurement | Manufacturer | Model | Collection frequency |
|---|---|---|---|
| [$CO_2$], [$H_2O$] | Li-cor | LI-7500A $CO_2$/$H_2O$ analyzer | 10Hz‡ |
| 3-D wind velocity, virtual temperature | Campbell Scientific | CSAT3 sonic anemometer | 10Hz‡ |
| Precipitation | OTT Hydromet | Pluvio$^2$ Weighing Rain Gauge | Every 10 min |
| $T_{air}$ | Vaisala | HMP60 humidity and temperature probe | Every 30 min |
| Relative Humidity | Vaisala | HMP60 humidity and temperature probe | Every 30 min |
| Longwave Radiation* | Kipp & Zonen | CGR3 pyrgeometer | Every 30 min |
| Shortwave Radiation* | Kipp & Zonen | CMP3 pyranometer | Every 30 min |
| Snow depth† | Campbell Scientific | SR50A sonic ranging sensor | Every 30 min |
| Digital Imagery | Campbell Scientific | CC5MPX digital camera | Every 24 hr |

* All four components of radiation (upwelling and downwelling (longwave and shortwave)) will only be measured at Shale Hills Tower HOG due to the location of the Garner Run Tower HOG. To model Garner Run we will use the Shale Hills data.

† ~~originally~~ Originally designed as part of tower system but will be deployed at LRVF Ground HOG location because the Garner Run tower will be located outside of the catchment.

‡The turbulent fluxes (sensible and latent heat) and the momentum flux are computed at 30 minute intervals via eddy covariance using these data collected at 10 Hz.

Table 3. Vegetation sampling in the Garner Run subcatchment

| Site[1] | Sample area (ha) | Tree basal area ($m^2$ $ha^{-1}$) | Tree density (trees $ha^{-1}$) | Tree species richness (# species) | Dominant tree species (% basal area) | Forest floor cover (% rock) | Mean rock diameter (cm) | Organic horizon C (g $m^{-2}$) |
|---|---|---|---|---|---|---|---|---|
| LRRT | 1 | 25.3 | 607 | 9 | *Quercus prinus* (44%) *Acer rubrum* (19%) *Pinus strobus* (19%) *Nyssa sylvatica* (12%) | 16 | 29 | 1775 |
| LRMS | 1.4 | 25.1 | 610 | 12 | *Betula lenta* (37%) *Quercus prinus* (21%) *Nyssa sylvatica* (15%) *Quercus rubra* (10%) | 28 | 45 | 2208 |
| LRVF | 0.7 | 24.6 | 371 | 14 | *Quercus rubra* (26%) *Betula lenta* (23%) *Quercus prinus* (20%) *Acer rubrum* (14%) | 36 | 43 | 1122 |
| TMMS | 1 | 18.5 | 519 | 9 | *Acer rubrum* (32%) *Betula lenta* (29%) *Nyssa sylvatica* (25%) | 34 | 60 | n/a |

[1]LRRT: Leading Ridge ridge top, LRMS: Leading Ridge midslope, LRVF: Leading Ridge valley floor, TMMS: Tussey Mountain midslope. Measurements were made in linear belt transects 700 to 1400 m long and 10 m wide centered at each soil pit position (Fig. 5).

Table 4. Frequency distribution of bedrock depth ~~to bedrock~~ measurements along ~~the~~ GPR transect (Fig. 9)

| *Depth to bedrock* | *Upper section* | *Lower section* |
|---|---|---|
| Shallow (< 0.5 m) | 0.00 | 0.00 |
| Moderately Deep (0.5 to 1 m) | 0.26 | 0.04 |
| Deep (1 to 1.5 m) | 0.51 | 0.48 |
| Very Deep (> 1.5 m) | 0.24 | 0.48 |

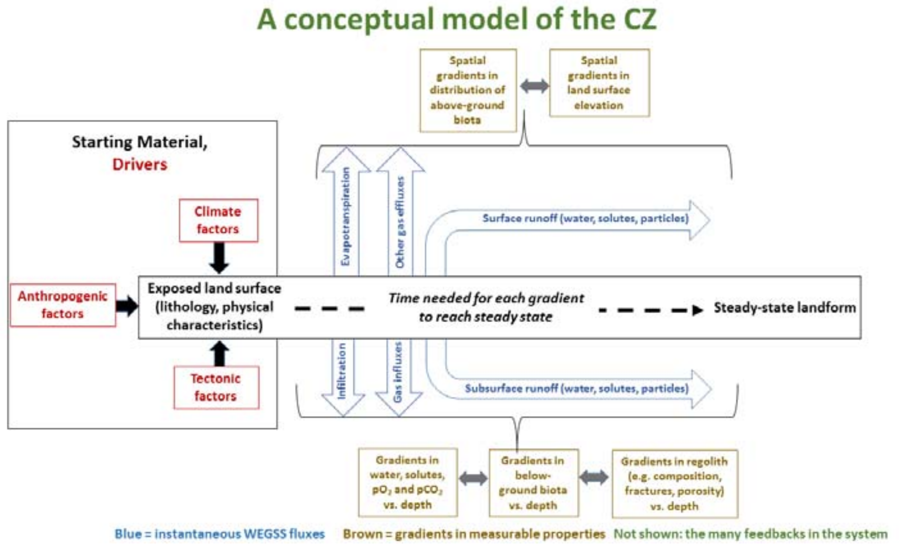
**Formatted:** Font: Not Bold

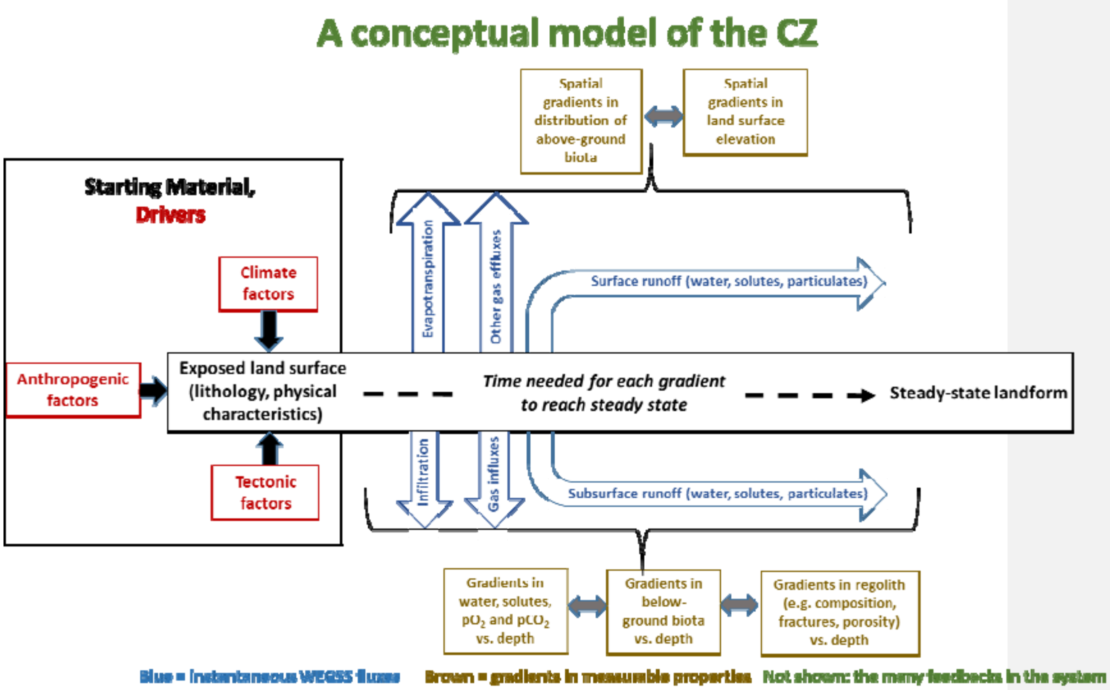
**Formatted:** Font: Not Bold

**Formatted:** Font: Not Bold

**Formatted:** Font: Not Bold

**Formatted:** Font: Not Bold


A conceptual model of the CZ
Spatial gradients in distribution of above-ground biota
Spatial gradients in land surface elevation
Starting Material, Drivers
Climate factors
Anthropogenic factors
Exposed land surface (lithology, physical characteristics)
Tectonic factors
Evapotranspiration
Other gas effluxes
Surface runoff (water, solutes, particles)
Time needed for each gradient to reach steady state
Steady-state landform
Infiltration
Gas influxes
Subsurface runoff (water, solutes, particles)
Gradients in water, solutes, $pO_2$ and $pCO_2$ vs. depth
Gradients in below-ground biota vs. depth
Gradients in regolith (e.g. composition, fractures, porosity) vs. depth
Blue = instantaneous WEGSS fluxes
Brown = gradients in measurable properties
Not shown: the many feedbacks in the system


A conceptual model of the CZ
Spatial gradients in distribution of above-ground biota
Spatial gradients in land surface elevation
Starting Material, Drivers
Climate factors
Anthropogenic factors
Exposed land surface (lithology, physical characteristics)
Tectonic factors
Evapotranspiration
Other gas effluxes
Surface runoff (water, solutes, particulates)
Time needed for each gradient to reach steady state
Steady-state landform
Infiltration
Gas influxes
Subsurface runoff (water, solutes, particulates)
Gradients in water, solutes, $pO_2$ and $pCO_2$ vs. depth
Gradients in below-ground biota vs. depth
Gradients in regolith (e.g. composition, fractures, porosity) vs. depth
Blue = instantaneous WEGSS fluxes
Brown = gradients in measurable properties
Not shown: the many feedbacks in the system

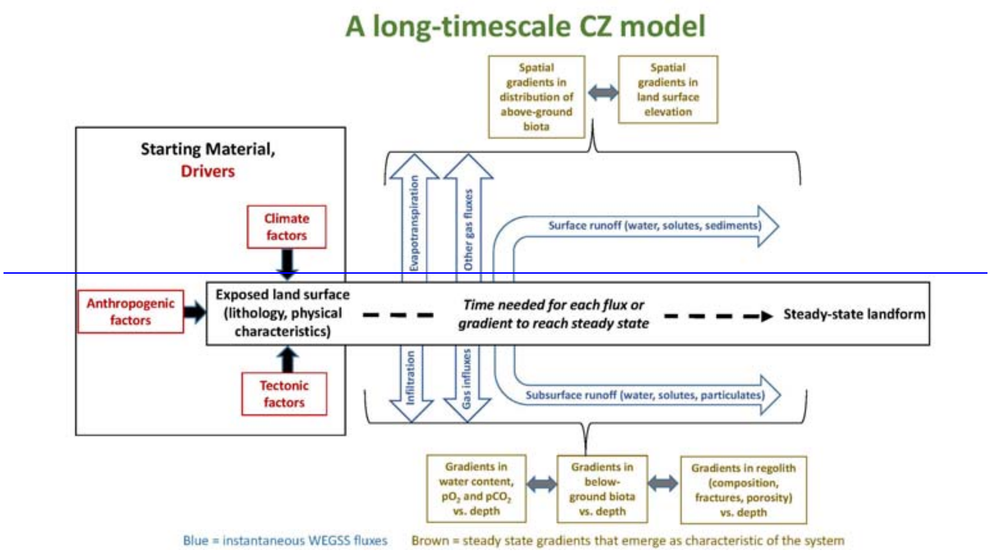



Figure 1. Critical zone science is aimed at understandinginvestigates –the architecture, character, and dynamics of the earth surface system from vegetation canopy to deep ground water at all differentat all -time scales. As As rock of a certain lithology and set of structural characteristics is exposed at earth's surface due to uplift or erosion, climate-driven inputs transform rock it to regolith. This transformation, shown in the black box, is catalyzed by biota (a feedback which is not shown explicitly). All of the pGradients of properties describing the CZ are shown roperties in brown boxes to the right. These gradients can of the diagram can be considered properties which may sometimes become time-independent (steady state) reach a steady statedue to the many feedbacks which are not shown. Boxes are placed from left to right to note the increasing duration of exposure time needed to achieve such steady states after increasingly long exposure times. In other words, after an initial transient period, these characteristics can reach dynamic equilibrium. For example, regolith thickness depth profiles of regolith composition can become constant when rate of erosion equals the = rate of weathering advance in the presence of feedbacks related to pore water chemistry, soil gas composition, and grain size. Likewise, the nature and distribution of biota may become constant for some period. As emphasized by theThe figure, emphasizes that gradients to the left ecosystems can are achieve steady state established quickly compared to some geological changesproperties to the right. T, and can therefore properties to the left are often be studied as if some of the other characteristics properties in boxes to the right in the

~~diagram (e.g. regolith thickness and character, uplift rate, landscape curvature)~~ are constant boundary conditions. However, over the longest timescales, all properties vary and can affect one another. The complexity of feedbacks (which are not shown for simplicity) can also create thresholds in system behavior. Red boxes indicate drivers~~,~~ and ~~black indicates the system under study,~~ blue arrows are ~~indicates the~~ WEGSS fluxes (up arrows for above-ground and down arrows for below-ground)~~, and brown boxes indicate gradients~~.

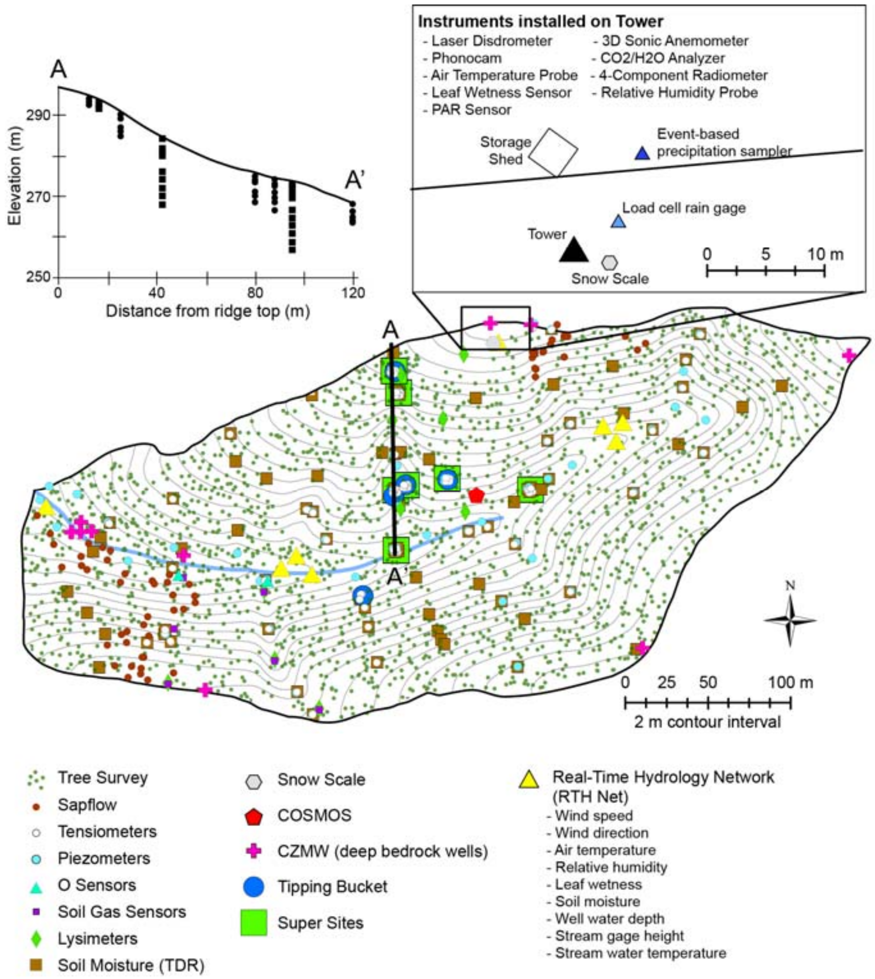
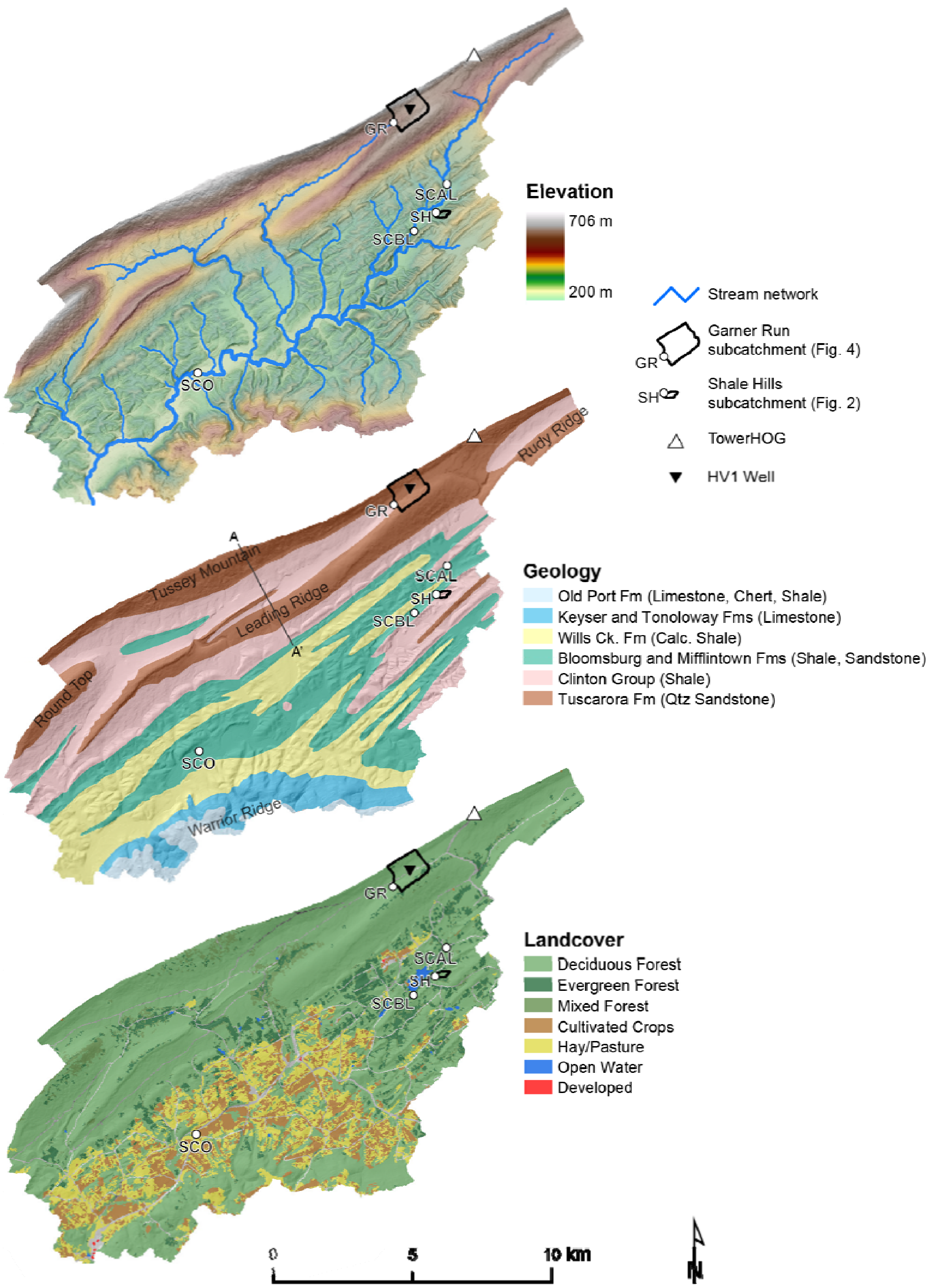



Figure 2. Mapped summary of the "everything, everywhere" sampling strategy at the Shale Hills subcatchment. Insets show soil moisture sensors (circles) and lysimeters (squares) along the transect shown on the map. Sensor and lysimeter depths are exaggerated five times compared to the land surface elevation. Second inset shows instrumentation deployed at the meteorological station on the northern ridge. Small ~~black~~ green dots on the map are the trees that were surveyed and numbered: ~~.~~ the subcatchment contains a dry oak-mixed hardwood community type (Fike, 1999) with an extremely diverse mix of hardwood and softwood species, including white oak (*Quercus alba*), sugar maple (*Acer saccharum*), pignut hickory (*Carya glabra*), eastern hemlock (*Tsuga canadensis*, and chestnut oak (*Q. montana*). The

sparse understory consists of American hop-hornbeam (*Ostrya virginiana*) and serviceberry (*Amelanchier sp*.). As we upscale the CZO to all of Shavers creek, many measurements will be eliminated in the Shale Hills subcatchment as we emphasize only a Ground HOG and Tower HOG deployment as described for the Garner Run subcatchment.

**Formatted:** German (Germany)



GR
SCAL
SH
SCBL
SCO
Elevation
706 m
200 m
Stream network
Garner Run
subcatchment (Fig. 4)
Shale Hills
subcatchment (Fig. 2)
TowerHOG
HV1 Well
Rudy Ridge
Tussey Mountain
Leading Ridge
Round Top
Warrior Ridge
A
A'
Geology
Old Port Fm (Limestone, Chert, Shale)
Keyser and Tonoloway Fms (Limestone)
Wills Ck. Fm (Calc. Shale)
Bloomsburg and Mifflintown Fms (Shale, Sandstone)
Clinton Group (Shale)
Tuscarora Fm (Qtz Sandstone)
Landcover
Deciduous Forest
Evergreen Forest
Mixed Forest
Cultivated Crops
Hay/Pasture
Open Water
Developed
0
5
10 km
N

Figure 3. Map of Shavers Creek Watershed, highlighting (a) topography derived from airborne lidar, (b) geology (Berg et al., 1980), and (c) landuse (~~NLCD~~Homer et al., 2011). In moving from measure-everything-everywhere (our paradigm in the 8 ha Shale Hills catchment (SH) to measure-only-what-is-needed in the Shavers Creek Watershed (164 $km^2$)), we chose to investigate two new first-order sub-catchments: a forested sandstone site (along Garner Run, marked GR) and an agricultural calcareous shale site (to be determined). In addition, three sites on Shavers Creek have been chosen as stream discharge and chemistry monitoring sites (marked SCAL – Shavers Creek Above Lake, SCBL – Shavers Creek Below Lake, and SCO – Shavers Creek Outlet).

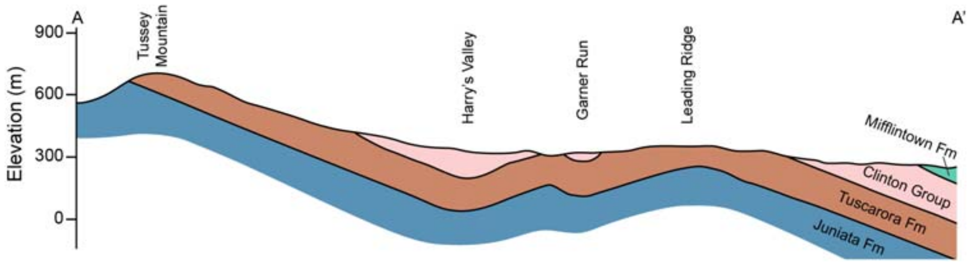



Figure 4. Geologic cross-section across Garner Run subcatchment reproduced from Flueckinger (1969). Map units ~~are labelled from youngest to oldest: Smm (Rochester and McKenzie Members of the Mifflintown Formation), Smk (Keefer Member of the Mifflintown Formation), Srh (Rose Hill Formation), St (Tuscarora Formation), Oj (Juniata Formation).~~include Mifflintown ~~is ~~(Middle Silurian), Clinton group (including Rose Hill formation) and Tuscarora ~~are ~~(Lower Silurian), and the Juniata ~~is ~~(Upper Ordovician). Cross section position is downstream from the targeted subcatchment (see Fig. 3). The published map (Flueckinger, 1969) of the actual sub-catchment (not shown) shows no remaining Rose Hill formation outcrop in Garner Run subcatchment, i.e., the Tuscarora no longer outcrops upstream of this cross-section and Garner Run lies in the axis of Harry's valley. ~~Nonetheless, t~~This cross-section from down valley of Garner Run subcatchment emphasizes that Rose Hill shale was originally present above the Tuscarora.

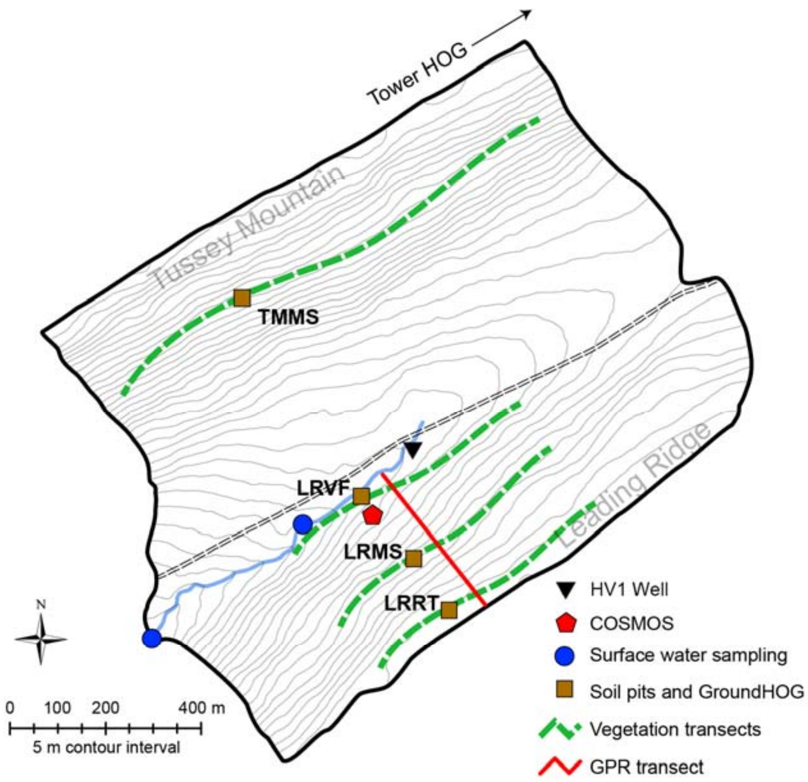



Figure 5. Map showing Garner Run subcatchment (blue line is the stream). Black dashed lines delineate Harry's Valley Road. The Harrys Valley well (HV1) is shown along with the location of the COSMOS unit and the outlet weir (blue dot to the southwest). The blue dot to the northeast indicates the approximate range of surface water sampling that is ongoing. Soil pits have been emplaced as shown, along with the Ground HOG deployment. Location of vegetation and GPR transects reported in this paper are also shown.

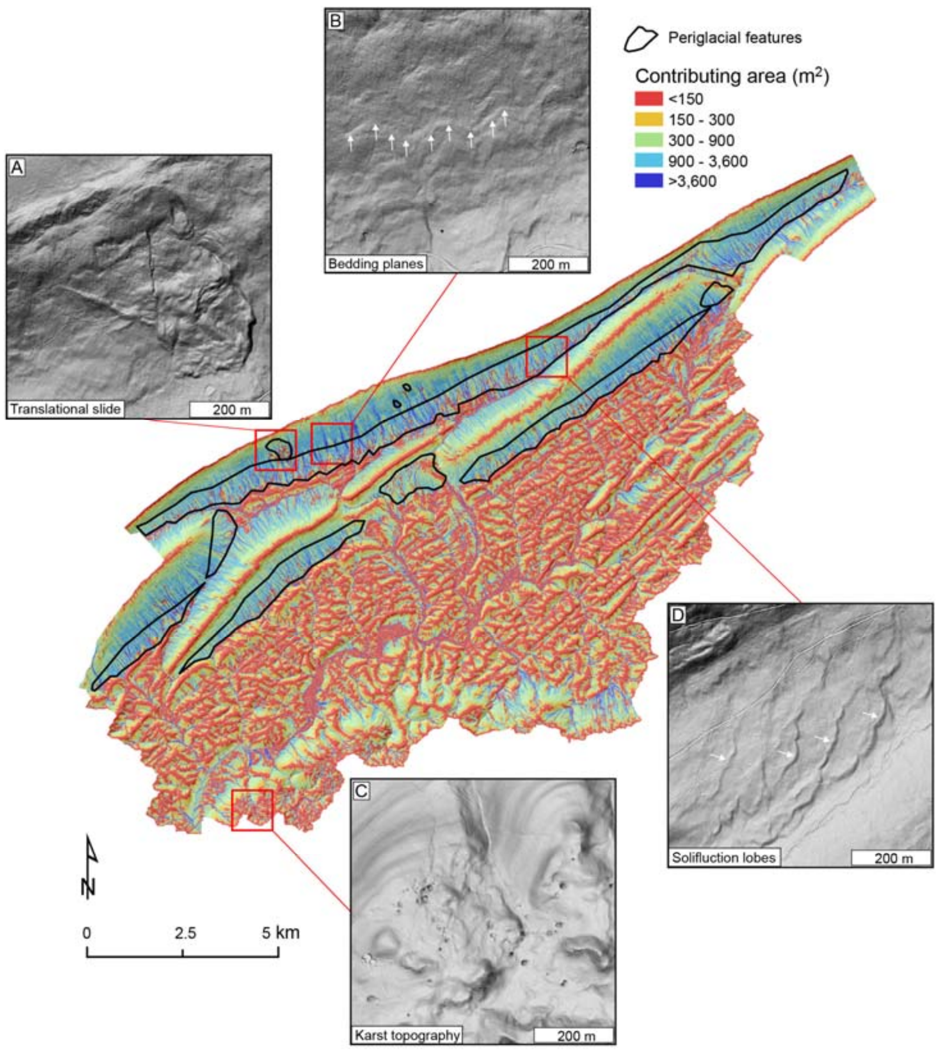



Figure 6. Map of bedrock and periglacial process controls on topography in Shavers Creek watershed. The contributing area was determined using the D-Infinity flow routing algorithm (Tarboton, 1997). The map highlights spatial variations in drainage density that correspond to sandstone (low drainage density and long hillslopes), shale (high drainage density and short hillslopes), and carbonate (intermediate drainage density and hillslope length) bedrock (see Figure ~~3~~ 4~~for bedrock geology map~~). Black outlines correspond to periglacial features expressed in the 1 m lidar topography, such as landslides (inset A) and solifluction lobes

(inset D). Sandstone bedding planes (inset B) and limestone karst topography (C) are also prominent.

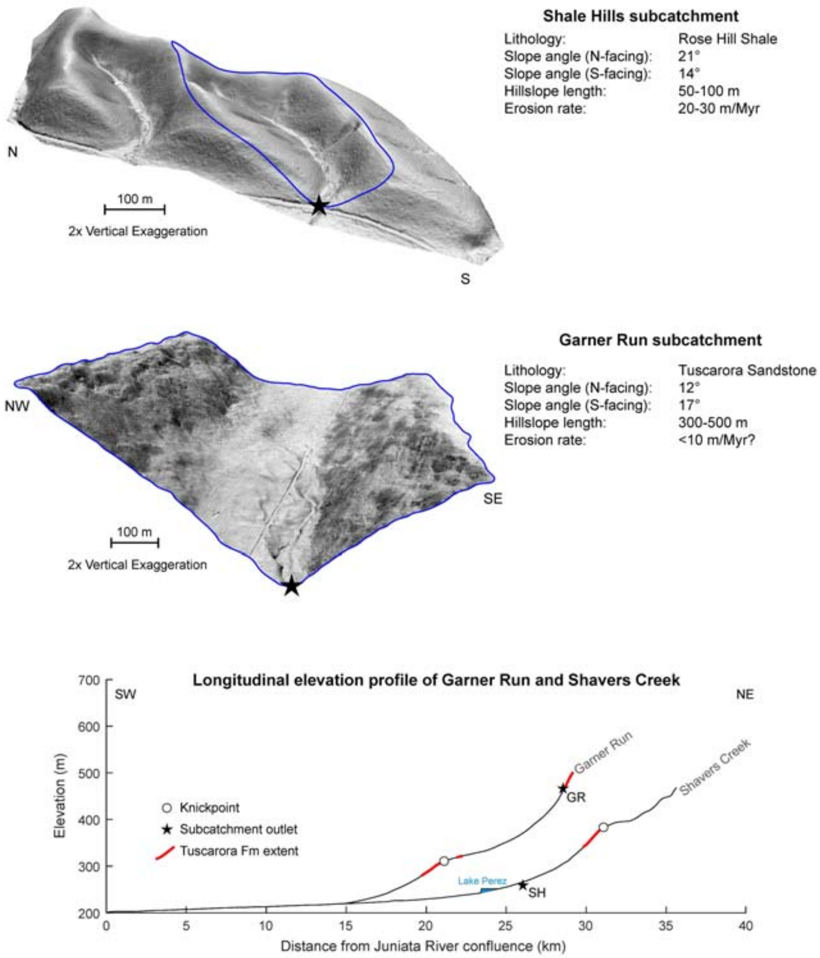



Figure 7. Perspective slopeshade maps (darker shades = steeper slopes) of Shale Hills (top panel) and Garner Run (middle panel) subcatchments, emphasizing differences in slope

asymmetry and hillslope length. Soil production and erosion rates for Shale Hills subcatchment were measured based on U-series isotopes and meteoric $^{10}$Be concentrations in regolith respectively (Ma et al., 2013; West et al., 2013; 2014). Erosion rate for Garner Run subcatchment is estimated based on detrital 10Be concentrations from nearby sandstone catchments with similar relief (Miller et al., 2013). Bottom panel shows stream longitudinal profiles, highlighting the lithologic control on knickpoint locations. Note the location of the Shale Hills subcatchment (SH) downstream of the knickpoint on Shavers Creek and the location of the Garner Run subcatchment (GR) upstream of the knickpoint on Garner Run.

**Formatted:** Superscript

**Formatted:** Centered

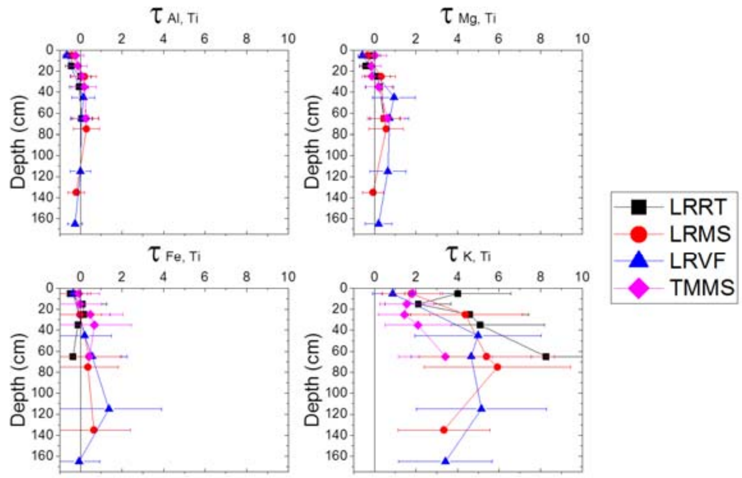



Figure 8. ~~Four plots~~Plots of normalized concentration (τ) versus depth for soils analyzed from the four Garner Run sub-catchment soil pits (LRVF, LRMS, LRRT, TMMS). Y axis indicates the depth below the organic – mineral horizon interface. τ~~The normalized concentration~~ is the mass transfer coefficient determined using ~~average~~ parent composition estimated as the average of ~~from five~~ 5 rocks (Supplement Table S3) from the bottom of several of the pits ~~and~~ based on the assumption that Ti derives from protolith and is ~~as the~~ immobile~~ element~~. ~~One explanation for these plots is that Al has largely been removed or moved downward in the profile while Mg, K, and Fe have largely been added to the profile.~~ If parent is correctly estimated, ~~In these plots,~~ τ = -1 when an element is ~~completely~~ 100% depleted~~ compared to Ti in the parent material~~, τ = 0 when no loss or gain has occurred, and is τ > 0 when the element has been added to the profile compared to Ti in the parent material.

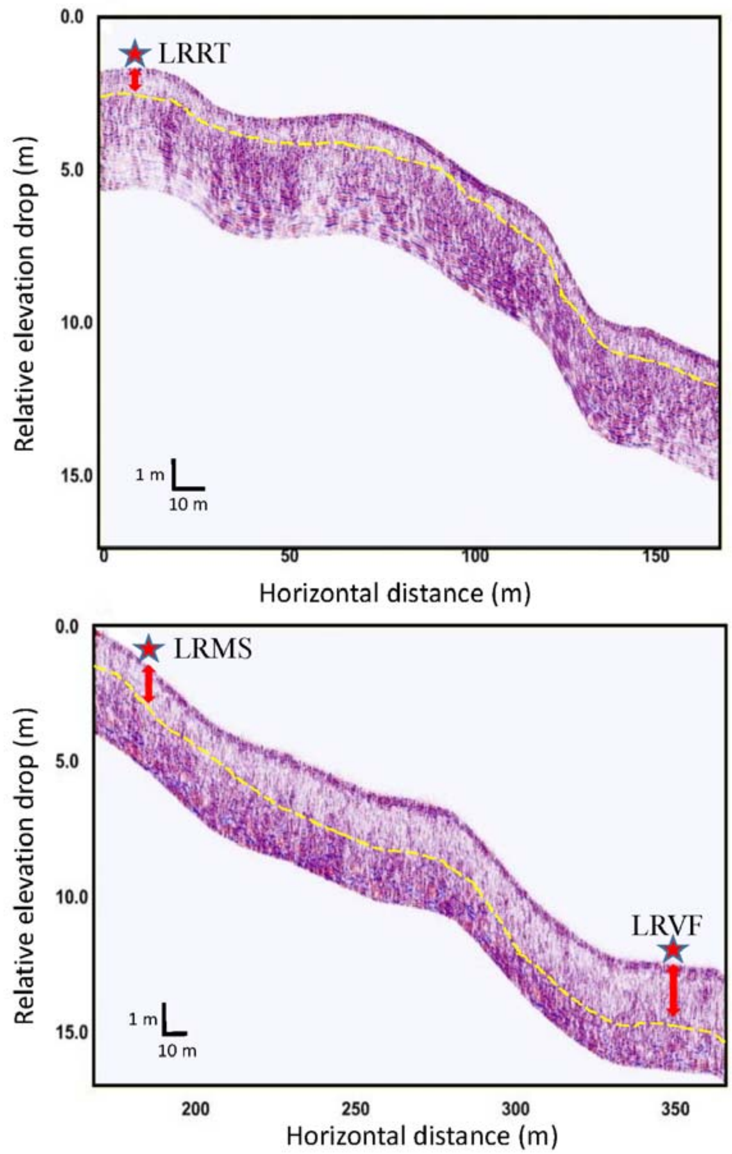



Figure 9. Ground Penetrating Radar (GPR) transect of the Leading Ridge catena, showing inferred location of bedrock-soil interface (yellow dashed curve). The three soil pits (LRRT,

LRMS, LRVF) are indicated by stars, with their observed depth to bedrock indicated by red arrow bar. LRRT and LRMS were dug by hand until refusal and LRVF was dug by hand and deepened with a jack hammer (note the observed depth was limited by excavation depth that was possible with the available digging tool). GPR data are exaggerated by 10x in vertical dimension as compared to surface topography. Summary values bedrock depths are tabulated in Table 4. from these GPR measurements.Ground Penetrating Radar (GPR) transect of Leading Ridge Catena, showing inferred location of bedrock-soil interface. GPR data is exaggerated by 4x compared to surface topography. Summary values are tabulated in Table 4 from these GPR measurements.

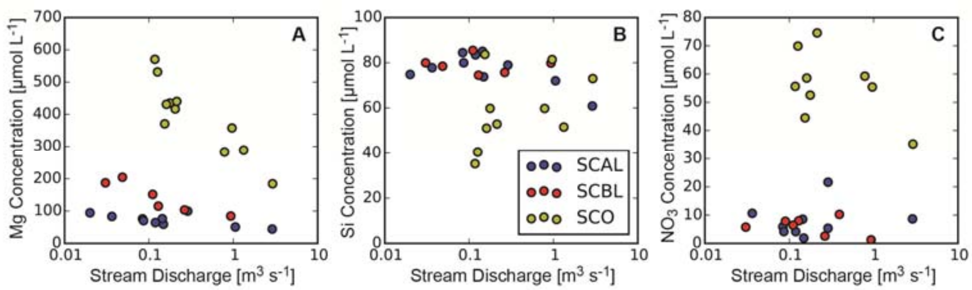



Figure 10. (A) Mg, (B) Si, and (C) ~~Nitrate~~ nitrate dissolved concentrations (filtered at 0.45 µm) and stream discharge measured at three locations on Shavers Creek: Above Lake (SCAL, blue), Below Lake (SCBL, red), and the Outlet (SCO, yellow) as shown on Fig. 3.