Best practices for collaboration: Working at the Susquehanna Shale Hills Critical Zone Observatory

Jason Kaye[1], Susan L. Brantley[2,3], and Jennifer Z. Williams[3] with the SSHCZO team*

[1]Department of Ecosystem Science and Management, The Pennsylvania State University, University Park, PA U.S.A. 16802

[2]Department of Geosciences, The Pennsylvania State University, University Park, PA U.S.A. 16802

[3]Earth and Environmental Systems Institute, The Pennsylvania State University, University Park, PA U.S.A. 16802

This living document is provided as a guide to best practices for collaborative science at the Susquehanna Shale Hills Critical Zone Observatory. We provide it to coordinate among scientists from different backgrounds and disciplines who follow different norms or practices. Our goal is to always produce the best science, education, and outreach that we can muster at our CZO.

I. Best practices for authorship on peer reviewed papers:

Based on the Geological Society of America and Ecological Society of America code of ethics and on the American Geophysical Union definition of authorship:

Authorship may be anticipated if researchers have made substantial contributions in one or more of the following areas:

1) creation of the conceptual ideas or experimental design;
2) management or execution of the study;
3) analysis or interpretation of data; or
4) writing of the manuscript.

We do not prescribe levels for substantial contribution, and so each manuscript will require an open discussion regarding authorship. However, to provide some guidance, substantial is taken here to mean a contribution that either involves planning and analysis beyond that available at a commercial laboratory, creative or long-term field work, development of models, or other similar contributions. In general, engagement in writing is often a key delineation of co-authorship. Thus, it is important that that scientists contributing to #2 (e.g. collecting field data) are sought out and afforded the opportunity to contribute in the analysis and writing stages of the manuscript. It must be recognized that different disciplines have different codes of authorship and so flexibility must be retained. Regardless, the discussion and agreement should be achieved early in the collaboration and the senior scientist should promote this discussion. In ambiguous cases, we are inclined to err on the side of being more generous with authorship. Once established, authorship and the order of authors shall not be changed without consulting all the authors on the manuscript. No author shall be included on a manuscript that has not agreed to

the content in the final version. This means that every author must be given a reasonable amount of time to read revisions of the manuscript.

Some questions that may arise are discussed below.

If I use someone's old, published data, should they be included as a co-author?

- No, prior publishing of data does not, in itself, constitute a significant contribution to new papers.

If I use someone's old, unpublished data should they be a co-author?

- If the data are unpublished but also not embargoed then we encourage the authors to engage the scientist who collected the data at a level that would constitute a substantial contribution. However, if a good faith effort is made to engage the scientist in charge of the original data and that scientist has not responded, then it would not be appropriate to include them as a coauthor (but it would be appropriate to acknowledge them). If the old data are embargoed (i.e., not yet public) then the authors must gain permission to use the data. At this time, the two parties (paper authors and data collector) should discuss authorship in the context of the criteria described above. In unusual cases, a researcher who collected embargoed data may not make appropriate progress in publishing a dataset. In that case, the CZO team may need to decide on a course of action with respect to publication of the embargoed data that, in the best case would involve discussion with the original researcher, but might have to proceed without such discussion. Such unusual circumstances should be well discussed among the steering committee for guidance. Ultimately, a researcher who makes a substantial contribution to a manuscript should be included as a coauthor on a publication.

If I use someone's code or model output that from a previously published paper, should they be included as a co-author?

- No, unless the code developer is intellectually engaged in the manuscript development. A couple of examples that can expect authorship: 1) The code developer provides new model outputs and is engaged in output analysis; 2) the code developer runs new model simulations for the manuscript (i.e., performs new calibration, collects new driver data), or adds new functionalities to the model.

If I use someone's code that has not been published in a paper, should they be included as a co-author?

- Similar to using someone's unpublished data, we encourage the authors to engage the code developer at a level that would constitute a substantial contribution.

If I collect field samples for someone should I expect to be a co-author on their paper?

- Field sampling is often an overlooked component of the creative scientific process where critical decisions are made that affect the quality and value of the data. However, field sampling alone is not a contribution that warrants co-authorship. We encourage discussions that enable people who have contributed substantially to field work to become engaged in analysis and writing at a level that warrants co-authorship.

If I test an idea from a CZO proposal, should the PIs be co-authors on the paper?

- This depends on how specific the idea is. If the authors of the proposal conceived of the idea and described an experimental design to test it, then they may have met criterion #1 for co-authorship, and they should be given the opportunity to meet other criteria for co-authorship. On the other hand, if the research is not tied to hypotheses that are described in the proposal then the proposal PIs should not be included as authors simply because they were a PI on the proposal.

If an undergraduate REU collected some of the data, should they be a co-author?

- Undergraduate REUs should be considered for authorship under the same criteria as other scientists. We should promote co-authorship in this regard by giving REUs opportunities to contribute to data analysis and writing if the student is ready for such efforts. However, in some cases, a worker may only do "what is told" and not participate in planning or thinking about the results in any substantial way: in these cases, inclusion as a coauthor may not be warranted.

How long should co-authors have to review a manuscript?

- Coauthors should discuss timelines for each manuscript. However, a reasonable expectation is that coauthors will read a draft within one month of receiving it, assuming that the author has established some sort of reasonable timeline with respect to vacations, trips, etc. Shorter turnaround times may be appropriate for revisions, but co-authors are still expected to read the final (revised) version.

What if I try repeatedly but I cannot get a coauthor to read the manuscript?

- An appropriate approach is the following. When the author finishes a version of the manuscript, he or she discusses with the possible coauthors a timeline or sequence of review (in other words, the authors must have some ability to frame up the timeline – it is not just at the discretion of the first author). If a potential coauthor does not read or comment appropriately on a manuscript, the author should propose a reasonable deadline and write in an email, "we will submit this paper without your name unless you read it and comment on it by such and such date: we prefer to retain you as coauthor but we must move forward". In case the potential coauthor still does not respond, it is appropriate to remove the potential coauthor from the authorship list. If possible it would be appropriate for the other authors to contact the coauthor by two means (e.g. email and phone) and make it clear that they will be removed from the authorship list if they do not respond in a specified amount of time. One possibility is also to submit a paper without a

coauthor (because the coauthor cannot participate in paper writing at the time) and then add the coauthor back in later if they re-engage appropriately.

Who will decide the final author list in cases of contention?

- We expect coauthors to handle this problem in a collegial way. Best practice will always dictate that the discussion of coauthorship be initiated early in the process and be continued throughout the process. The senior scientist on each project should guide this process along. Guidance can also always be sought from the CZO Steering Committee.

Who will decide the order of authors?

- Best practice would be for all of the coauthors to decide this in a collegial way; in most cases, the senior author will decide the order of authors.

II. Best practices for installing infrastructure or experiments:

Best practices for new infrastructure require a careful consideration of impacts on existing infrastructure, permitted use of sites, and usage fees. Scientists (including CZO PIs) that would like to initiate new work that is co-located within the bounds of the CZO must discuss this idea with the CZO Steering Committee, the Program Coordinator, and the Watershed Specialist. This includes discussion at the time of proposal writing if new money is sought (even for PIs inside of Penn State). The Steering Committee will help identify key CZO scientists that should be immediately consulted regarding the new project. At some point (perhaps even at the proposal stage), an email that describes the activity should be sent to all CZO scientists (including students) to determine who may be affected by the work. PIs are encouraged to share the email with everyone in the lab group. A second email should be sent prior to installing new equipment. If the new research includes destructive sampling that could affect many projects then the Steering Committee should present the proposed work in an all-hands meeting to discuss the viability of the new project. If there are conflicting deployments, then the Steering Committee has the responsibility to determine whether new installations should go forward.

The CZO Watershed Specialist should be included in both preliminary and developing conversations regarding new equipment. The final placement of all new field infrastructure (e.g. sensors, pvc, etc) must be approved by the CZO Watershed Specialist. In addition, materials that will stay in the field must be marked with a PI-specific color. Even non Penn State personnel will be assigned a specific color and will be expected to maintain their color coding while working in the project. Color coding should be managed directly with the Watershed Specialist and the Program Coordinator. Immediately after installing new equipment it is a best practice to take a photograph of the installation and share it, by email, with the entire team. In addition, the location of the instrument must be communicated to the Data Manager.

The current usage agreement with Penn State Forest Lands Office allows CZO top-tier priority research within the Shale Hills catchment. Any outside funded project will need approval from the CZO Steering Committee, followed by approval by the Penn State Forester. The Forester may impose a separate research permit and usage contract and usage fee. No projects are to default under the CZO umbrella (though they may eventually be placed there), with the exception of seed grant projects funded by the SSHCZO.

As the CZO expands outside of Penn State lands, new rules are being developed. Specifically, the CZO has an agreement with the PA Department of Conservation of Natural Resources (DCNR) for specific activities in Garner Run watershed. Every person who works in that area as part of the CZO (student or faculty, inside or outside of Penn State) and every advisor for a student working at the CZO on the specific activities described must sign the agreement with the PA DCNR and this must be kept on record by the CZO Program Coordinator. If a PI initiates new work in the area that is not listed in our permit, a new permit must be requested and negotiated directly between the PI and the DCNR, and a record of this documentation must be kept on hand by the Program Coordinator. It can take up to 3 months for the permit process with the DCNR. If work is pursued in these lands without signing the form, or if work is pursued which is not described on the agreements, the CZO will rescind permission to work on the project and will work with DCNR to rectify the situation. As we expand to private lands additional guidelines will be developed and extra care will need to be taken to respect the wishes of the land owners.

New research that is not co-located with existing CZO infrastructure may require a revision of the CZO permit and will need to be discussed with the Forest Lands Management Office or the DCNR. The CZO Project Coordinator and Watershed Specialist should be included in these discussions. In general, best practice will initiate discussions with the CZO Steering Committee, followed by discussions with the Penn State Forester or DCNR. When new funding is garnered for new research at the CZO, a new fee will generally be paid to the Penn State Forester for this work. This fee will be negotiated directly with the Forester

When a PI receives new funding for new instrumentation, the CZO itself will not become responsible for the new infrastructure that is emplaced in the CZO catchments. Likewise, the new PI will be encouraged to use the CZO's data infrastructure for publication of data; however, the CZO will not become responsible for the data from the new project nor will the CZO police publication of the new data. Ultimately however it is recognized that the PI is generally co-locating the experiment at a CZO catchment due to the pre-existing research and infrastructure. Given this "value added" by the CZO, the CZO Steering Committee and Watershed Specialist can ultimately decide whether certain activities are pursued in the CZO catchments. For example, a proposal might be funded to do geophysical monitoring in Shale Hills and might involve a new fee to the Forester. After initiation of the work, the PI of the new proposal might decide he/she wants to do trenching up the middle of the catchment. If the CZO Steering Committee decides this is inappropriate, then the new PI will not be enabled to complete the

trenching. In this regard, the Steering Committee will work closely with the Forester or DCNR to maintain appropriate activity.

III. Best practices for using, maintaining, and sharing existing field infrastructure:

All infrastructure at the CZO is linked to a PI via color coding. This PI is responsible for maintaining and promoting collaborative use of the equipment. While the color codes denote the PI in charge, they do not denote ownership of equipment. All CZO field infrastructure and data are shared. However, no field equipment should be used without first notifying the PI in charge and establishing the terms of use and collaboration. Shared use and collaboration is expected and in some cases, this may mean developing a plan of collaboration that could lead to co-authorship if criteria in Section I are met. If PIs cannot agree on a terms of shared use then they should bring the issue to the Steering Committee.

The PI in charge may decide that it is best not to maintain equipment in working order, even though the equipment can remain in the field for future activities. For example, lysimeters can stay in place for years without being sampled. In these cases, the PI in charge should notify the Watershed Specialist and any co-PIs that have used the equipment in the past. A new PI may want to initiate the use of that instrumentation. In that case, the new PI and the original PI will be considered in charge of the equipment and its use. Any time infrastructure is moved or removed, the person in charge should contact the Data Manager to report the equipment, PI, geolocation data, and the date of change.

While shared use is the overarching goal, there may be some equipment for which shared use is not appropriate. For example, some cases might involve equipment which is very sensitive or difficult to maintain or expensive or rented or borrowed. These can be handled on a case by case basis.

Questions that may arise:

What if I can't maintain the equipment myself?

- There are cases in which the CZO support staff or collaboration among co-PIs may be required to maintain field infrastructure. These will need to be handled on a case by case basis with consideration of the availability of support staff time. In general, when a PI begins a sub-project that will require time from support staff, that requirement must be vetted through the Steering Committee. The Watershed Specialist will generally be the person to help in maintenance.

What if the PI in charge is not maintaining critical equipment in a way that promotes shared use?

- In these cases a broader discussion may be needed in which the team may decide to transfer maintenance responsibilities to a different PI or to allocate more project resources (support staff time or funds for maintenance) to the equipment.

IV. Best practices for removing field infrastructure:

If field infrastructure has reached the end of its useful life it should be removed by the PI in charge, as denoted by the color coding, and the landscape returned to original form. There may also be cases in which the equipment is still functional but the PI wants to remove the equipment to reduce the maintenance burden. Before removing equipment for any reason the PI should work with the Watershed Specialist to email the CZO team (all co-PIs plus support staff) to determine whether the removal will affect other users.

When the CZO ceases to be a continuing research project, each PI has the responsibility to remove equipment with their color code, or negotiate a new use agreement with the PSU Forester or DCNR. Our use agreements stipulate that we will restore the landscape to a pristine condition when we are finished with the project.

V. Best practices for collecting, sharing and archiving samples:

Before going to the field to collect samples, scientists should make the Watershed Specialist and/or Program Coordinator aware of their sampling schedule. This is typically done via quarterly planning that is solicited by email. Sampling protocols should be posted on the CZO PSUBox and CZEN.org so that all future users can use the same sampling protocol or deviate intentionally. CZO workers should attempt to share samples so that multiple analyses are conducted on the same sample. The scientists sharing the samples should agree on the terms of the collaboration, including the potential for co-authorship.

Every solid and liquid sample collected from the field should become archived if there is sufficient sample and if it is likely or possible that future users might want to access this sample. The Program Coordinator is responsible for sample archiving. PIs and their students and postdocs should consult with the Program Coordinator prior to collecting any samples so that the archive protocol can be established. The CZO has dry storage for solid and water samples in Research Unit C. No archive is available for frozen samples. All samples must be registered with IGSNs prior to archival. CZO personnel should attempt to share archived samples with one another and with the broader scientific community. Scientists who want to use archived samples are required to contact the scientist that collected the sample and describe how the sample will be used. The Program Coordinator is responsible for facilitating this communication and sharing. The collector of the archived sample should agree on the terms of collaboration before the archived sample is released. However, in cases when the collector cannot be consulted or doesn't consent to the release, the case can go the Steering Committee. If archive sample retrieval becomes overly time-consuming, arrangements may need to be made to pay someone to find samples.

Questions that may arise:

What if I want to deviate from the established CZO sampling protocol?

- We expect this to happen. A rationale should be provided for the change and methodologies should be noted with respect to the CZO sampling protocol at both the ANGEL and CZEN.org sites so that others will know how and why the change was made. The Program Coordinator will facilitate and oversee modifications to the protocols.

What if there is only a little bit of an archived sample left and someone wants to use it up?

- If the collector and user of the archived sample and PI of the CZO agree that this is a good use of the sample, then it can be used. In general, however, samples should not be used up. If there is disagreement then the Steering Committee can be consulted.

VI. Best practices for sharing data:

Guidelines for sharing CZO data are outlined here: http://criticalzone.org/national/data/access-czo-data-1national/#DataUseAgreement. Where possible, a PI should get a doi for datasets for future citation. In general, we consider that data storage in the CZO data infrastructure is advisable, even for data funded by entities outside of Penn State NSF CZO funds. However, the CZO does not become responsible for archiving these data.

It is a best practice not to directly share your copies of data with third parties. For example, if you have an excel spreadsheet of data that another student or PI has shared with you, you should not share those data with a third scientist. Instead, it is best to have that scientist access the data by going directly to the CZO web page, or contacting the original data source (PI and student) directly. Under some circumstances (e.g. when you have manipulated data in a way that is beneficial to the third party) you may need to pass on someone else's data to a third party, you should obtain written consent from the original data source, for example through an email exchange that includes a discussion of terms of authorship and use.

Some data sharing will occur prior uploading the data to the CZO database. Data sharing at this early stage is encouraged and even necessary to enable students and PIs to conduct multidisciplinary research. The parties involved should establish authorship and use expectations at the time the data are shared. As discussed above, data should never be shared with a third party without first consulting and obtaining written consent from the original source of the data.

VII. Best practices for project management:

The Steering Committee shall be comprised of a subset of the PIs (some fixed one rotating), a subset of the staff, and one rotating student. The Steering Committee should send out updates after their meetings to keep co-PIs appraised of key decisions. The Steering Committee is an

appropriate outlet for all grievances related to the project. Discussions of sensitive issues (e.g. personnel) need not be shared, but decisions regarding allocation of resources and discussions about important changes affecting PIs should be shared.

As new PIs become involved in the CZO, the Steering Committee and all of the PIs will make every attempt to avoid the situation where more than one group is working on the same problem. However, some overlap will undoubtedly happen and some overlap is expected to be appropriate in some cases. The Steering committee will thus try to steer PIs toward collaborative approaches to overlap, or toward appropriate "competition". In this regard, "competition" means collegial testing of alternate hypotheses or alternate methodologies to understand functioning of the CZO. The CZO management ultimately has no authority to prohibit publication of ideas, data, or models for the CZO and in fact encourages competing ideas, data, and models.

In general, the CZO management will make every attempt to promote i) collegiality, ii) open communication, iii) excellence in research, iv) excellence in education, v) excellence in collaborative science, vi) excellence in outreach to the public.

A field crew comprised of a rotating group of students, postdocs, and staff supported by the project will assist with sample collection and general maintenance at the site and will help ensure that field sampling can always be conducted in pairs.

VIII. Best Practices for Advising Students:

In general, graduate and postdoctoral students who work at the CZO should be encouraged to appear as co-authors on joint publications as appropriate. Generally, a student will be first author on the project they spearhead (if they do most of the work), unless they do not move forward on publication in a timely manner. When students do not move forward on a project within one year of completion of their degree, the PI may write the paper and first-author the project.

It is the responsibility of PIs on the CZO to mentor their students regarding CZO best practices. Having your student sign this document is not enough; continuous mentoring regarding ethics and best practices is expected. PIs are expected to be aware of which data and models their students are using, which datasets originated from other CZO students or PIs, and to be engaged in all discussions regarding authorship and use of data, models, and infrastructure. Furthermore, PIs are expected to share relevant emails with their students including those related to infrastructure and site maintenance.

IX. Best Practices for Outreach:

The CZO has a commitment to complete outreach to nonscientists and the public in general. It is expected that everyone who works at the CZO will at some time (e.g. once per year) participate

in public outreach coordinated by the CZO. However, be aware that appropriate clearances are required before PSU faculty and staff can participate in outreach.

X. Best Practices for Reporting:

It is expected that everyone working at the CZO will provide reports of effort to the Program Coordinator in a timely manner. Lack of participation in reporting, if egregious, can be grounds for termination of collaboration at the CZO. Everyone working at the CZO will also be expected to cite the CZO appropriately (as indicated on the CZO website) and to provide copies of submitted, in press, and published papers to the Program Coordinator at the time of submission, acceptance for publication, or publication respectively.

Signature:

_____

Date: -

_____

*SSHCZO team members: David Eissenstat, Ken Davis, Li Li, Tess Russo, Roman DiBiase, Henry Lin, Margot Kaye, Yuning Shi, Li Guo, Elizabeth Hasenmueller, Tom Adams, Chen Bao, Weile Chen, Joanmarie Del Vecchio, Alison Denn, Xin Gu, Yuting He, Lillian Hill, Beth Hoagland, Warren Reed, Ismaiel Szink, Dacheng Xiao, Dan Arthur, and Brandon Forsythe