**Comments from the Editor (EC) for Moderate Revisions, prior to sending out to Reviewers 1 (R1) and 2 (R2):**

EC1) Please don't use abbreviations (e.g., SSY) in the highlights, please expand in full.  As you have 2 highlights spare this should be OK.

Done.

Common abbreviations were used to keep under the limit of 85 characters, including spaces. They have been rewritten.  
  
EC2) L38 (and elsewhere), can you please number all heading, sub-headings etc.?  This will enable reviewers to more easily navigate your manuscript when they are reviewing it.

Done.

My mistake leaving these out.  
  
EC3) L128, rather than have 'questions addressed' it would be preferable to have specific objectives or aims for your study. In the last paragraph of your Introduction can you please explicitly state what your 'aim(s)' or 'objective(s)' or 'hypothesis (hypotheses)' is (are)?  That is, specifically use one of these words.  While you have a purpose, this is a little broader than having specific aims or objectives.  Consider using a bulleted sentence structure to list these.  Note the word 'question' is used in the following to generically mean aim / objective / hypothesis.  It is common practice to list the aims/objectives at the very end of the last paragraph of the Introduction section.  
  
Note the grammar of such a sentence follows (please pay careful attention to the use of colons and semi-colons):  
(i)     question 1 is interesting;  
(ii)    question 2 is really interesting; and  
(iii)   my Mum thought I should write something about question 3.  
  
Implementing this point makes it much easier for scientists from all language backgrounds to easily understand what you aim to do.  This dove-tails into the comment directly below.

Done.

The word “Objectives” has been inserted to describe the numbered items in the last paragraph of the Introduction that summarize the two main objectives of the study.

We feel it is critical to explicitly state the research questions that the paper seeks to answer, as questions ending with a question mark. These questions are explicitly answered later in Results and Discussion, demonstrating that we have learned something about this hydrological system and our hypotheses were confirmed or not. We state the objectives of the data analysis (to quantify and model sediment yield) but more importantly we pose the questions that motivate this analysis, and structure key take-away messages of the paper.

EC4) Improved structure: once you've explicitly used one the following words to state what your 'aim(s)' or 'objective(s)' or 'hypothesis (hypotheses)' is (are), then, assuming you have objectives, use these objectives to provide structure to your revised MS. For example, let's assume you have three objectives, then use them to structure your Methods section, Results section and Discussion sections, as follows.  
1 Introduction  
2 Study Site and Materials (have as many sub-headings as needed to introduce all the datasets used, their pre-processing - or maybe this needs to be 2 main headings, noting you might also need a "2 Theoretical Background" section too, in which case this would heading #3, and all others would increment by 1)  
3 Methods  
            3.1 Objective 1 (4-8 words to summarise objective 1)  
            3.2 Objective 2 (4-8 words to summarise objective 2)  
            3.3 Objective 3 (and so on)  
4 Results  
            4.1 Objective 1 (same words as 3.1)  
            4.2 Objective 2 (same words as 3.2 and so on)  
            4.3 Objective 3  
5 Discussion  
            5.1 Objective 1  
            5.2 Objective 2  
            5.3 Objective 3  
6 Conclusion  
  
Currently I'm up to page 16 and given your structure and possible lack of numbered heading, I'm finding it very challenging to know what section of the (rather long) manuscript I'm reading.  
  
Using the aims/objectives at end of your Introduction section to structure the rest of the paper makes it easy to read (and review).

Done.

The manuscript has been significantly shortened, and reorganized following the above template

EC5) L178, units of annual potential evapotranspiration need to be mm/yr, as you have correctly provided for the mean annual precipitation a few lines earlier.

Done.  
  
EC6) As the JoH Guide for Authors is currently being updated and does not state the following can you please implement? Can you please provide all Figures in a WORD document with the figure captions directly following each figure? This means a reviewer only has to flick back-n-forth between 2 pages (text and figure).

Done.

The JoH and Elsevier Guide for Authors specifically calls for Figures and Tables to be submitted as separate files, with Figure Captions in separate document. This procedure required a fair bit of time to accomplish using the online submission system.

However, in the interest of making the review process as easy as possible, we have also compiled a new document including all figures, followed by the caption.  
  
EC7) Plus you may wish to embed your figures and tables (and their captions) into the text directly following the paragraph where they are first mentioned. This will make it even easier for a reviewer.  If you don't do this please put  
< Figure 1 here please >  
< Table 1 here please >  
on new lines to highlight to reviewers (and the layout people) where the non-text elements should be located.

Done.

We added these tags in the appropriate places in the manuscript text, in addition to compiling a document with figures and captions, and uploading separately. We did not choose to embed figures in the text.  
  
EC8) L749, what discharge metrics?  It seems there is at least 1 word missing from this sentence.

Revised text:

Similar to other studies the highest correlations with SSYEV at Faga'alu were observed for discharge metrics Qsum and Qmax (Basher et al., 2011; Duvert et al., 2012; Fahey et al., 2003; Hicks, 1990; Rankl, 2004; Rodrigues et al., 2013).  
  
EC9) The submitted manuscript is long, very long.  The PDF I see has 87 pages - that is big bordering on huge. As I've noticed that long manuscripts usually suffer harsh reviews I strongly urge you to seek to reduce the length of the submitted manuscript.

Can you move the Appendices into the Supplementary Material? If you can reduce the PDF that reviewers download to be 45 to 55 pages your manuscript will be much less likely to put reviewers offside from the start.

Appendices have been moved to Supplementary Material. We wanted to include these additional materials for an interested reviewer but they can be moved wherever you think is best. The manuscript text has also been shortened significantly.

This, and EC4, are the reasons I'm requesting moderate revision, as opposed to minor revision, as positively implementing these will take some time, however, it will be worth it as your manuscript will be clearer and much more likely to be viewed favourably by JoH reviewers.

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
COMMENTS FROM EDITORS AND REVIEWERS  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
Editor Comments:  
  
1) Thanks for making improvements as requested by my pre-screening of your original submission.

No problem, thank you for taking the time to provide constructive reviews for improving the paper  
  
2) Both reviewers in their confidential comments to me stated they thought the MS was too long. Assuming this MS is ultimately published in JoH, if a paper is too long it will not receive the attention it deserves, both you and the JoH do not want that.  Once you've addressed all reviewer comments go through the revised MS with a fine tooth comb and critically ask yourself is this sentence needed and/or can it be tightened to reduce the overall word count.  Please let me know how many words were contained the version the assessed by the reviewers, and how many words the next iteration is.  Please also calculate a percentage reduction (if possible, if you can reduce it by 15% or thereabouts that would be ideal - and I know Editors/Reviewers are seeking more detail with less words, its tough, yet I sense you're up for the challenge - good luck).

The MS Addressed by R1 and R2 was 9,425 words (not counting Acknowledgments, References, Figure Captions, or Supplementary Material)

Every effort was made to reduce the length of the MS. The revised MS is now 9,280 words (not counting Acknowledgments, References, Figure Captions, or Supplementary Material); we only cut by 3% but added several new components suggested by R1, including more information on T-SSC relationships, a sensitivity analysis of T-SSC relationships, and Smearing estimate corrections for log-bias transformation in the storm metric-SSYEV relationships.

Cut by 15% would be 8,177 (not counting Acknowledgments, References, Figure Captions, or Supplementary Material)

3) Please respond to all comments using codes for each point. That is RXCY for reviewer comments where X is the reviewer number and Y is the reviewer comment number, for example R2C3 means Reviewer 2 Comment 3, AE1 is Associate Editor Comment 1, EC1 means Editor Comment 1. Using such codes will allow you to easily perform inter- and intra-reviewer crossing referencing which may allow your response letter to be more integrated, and also means that navigating through your response letter is easier. Provide your response directly after each comment and do not edit or shorten the comment. That is, even if the comment contains spelling errors do not correct them.

Do not change the order of the comments, the reviewers / editors have provided you with comments in a specific order, and expect your responses to be in the same order. If multiple comments are similar, so your response is similar, then please refer back to the detailed response, for example, "Please see our response to R1C3 for full details" as a response works.  
  
In your response letter please make all of the following comments blue and your interleaved responses black. Where you wish to highlight line numbers in the revised MS please highlight with the yellow highlight option.   
  
If the comment is editorial in nature (e.g., identifying a typographical error) then your response is a simple 'Done', obviously for comments of a scientific nature a more considered response will be required.   
  
If you do not implement a reviewer's comment, then I expect that a well-reasoned scientific rebuttal is provided in your response letter. This is critically important.   
  
Good luck making your revisions and substantially improving your manuscript; I look forward to seeing your revised manuscript and responses to the following comments when you're ready.  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
Dear Dr Messina  
as you can see below, I have received two rather contrasting reviews! After reading the reviews, and re-reading your MS, I would suggest somewhere in the middle of the two reviews is about right. I do agree with R1 that there are some missing information on sampling protocols, and that it can be a little hard to follow the logic of the analyses at times. R1 has done a very thorough job, and I would urge you too consider the comments around these issues carefully. There are a number of specific comments that you will need to consider also.  
  
There are some questions around citations that you should think about. I might add that there has been some work done on how different road surfaces affect erosion rates (e.g., Sheridan and Noske, 2007, Hydrol. Procs, and others). I look forward to seeing your revision.  
  
Patrick Lane  
Associate Editor  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
Reviewer #1:   
  
R1C0:

Review Summary  
I recommend this study for publication after major revision.  
  
The manuscript (MS) presents a thorough case study of the effects of human disturbance on suspended sediment dynamics in a small, steep tropical watershed over event to interannual time scales.  This study is of high technical merit.  The work is justified by a lack of such characterizations of watersheds in this geophysical setting, as well as its applicability to sediment management in these regions, which is critical to coral reef management.  Reporting effectively situated this work within the context of both the technical and applied canon in most cases.  
  
However, there are many aspects of the reporting, from trivial to substantial, as well as several minor to moderate technical issues that will require a substantial amount of effort before this MS is ready for publication.  These issues necessitate a recommendation for major revision.  Major points of issue are detailed below, followed by line-specific comments.  
  
It is my hope that the authors will take the thoroughness of this review as an encouragement, for only interesting pieces are worth such effort.  I look forward to seeing a revised MS for review, and eventually in print.

R1C0 Response:

Thanks to the reviewer for a thorough and thoughtful review.

R1C1: Seven Substantial problems were identified with the reporting of methods and results.  It is suggested that you:  
  
1. Present effective overviews at the beginning of each section/subsection, including the development of a table or flowchart presenting your experimental overview.  
Please try to remember that each section and sub-section are a unit of reporting.  Each unit should begin with definitive statements providing an overview of the contents of the section/sub-section.  At the manuscript scale I would suggest pulling together a table or flowchart presenting your experimental overview.  This was briefly and generally presented at the end of the introduction, but the beginning of the methods section really demands a more technical break down of the experimental overview.  At the section and sub-section scale I found that initial overviews were often incomplete, and in some cases followed by motivations, and then detailed accounts of each component, without ever introducing each component in the first place.  It is important to give the reader the skeleton of what you are reporting at each turn, particularly when multi-step procedures, with multiple approaches as each step, have been employed.

R1C1 response:

Agreed--method overviews were added to the first paragraph of the Methods section, and all subsections, with substantial rewriting of the MS. A table or flowchart of methods would add to the length of the manuscript; we hope that the additional paragraphs are sufficient for clarity.  
  
R1C2. Report on data at the beginning of the Methods section rather than the end.  
On to sequencing: the reader needs to be informed about certain components of your work before others.

In almost all work, one should report on data first or very early in the methods section. All analyses depend on the data, so the details of their collection and processing must come first to forestall unnecessary uncertainty and questioning on the part of the reader.  
[Editor: I agree here, this is what the proposed "Section 2 Study Site and Materials" refers to in my previous review.]

R1C2 response:

Agreed- the field data collection section has been moved to the beginning of the Methods section.

(we concluded that including all method information in the Study Site section would make the section too long).

Our internal reviewers previously commented that the data collection section belonged at the end of the Methods so that the development of the stage-Q and T-SSC relationships would not distract from the overall SSY modeling that is the main objective of the paper. They argued that the overall analytical approach should be described first so that it makes sense why each dataset was collected and how it is used to answer the research questions or accomplish the research objectives. I think this problem was mainly due to the uncertainty of how to treat the stage-Q and T-SSC components since each one has its own Methods, Results, and Discussion. Are they Methods with Results and Discussion that should be organized as such? Or were they just a component of Data Collection? We have now treated the stage-Q and T-SSC components as Data Collection (in Methods) since these are standard methods that have been used for many years and moved the technical details to the Appendices. Many papers simply don’t include any of the technical details since they download the data from a website or the methods and results are detailed in a previous paper.

R1C2a. Furthermore, performing a flux based approach to investigating human impacts on natural processes, such as sediment production dynamics, is by necessity a series of stepwise procedures, which provides a basis for reporting structure.  Of course the linearity of these operations (i.e. Step 2 dependent on the results of Step 1, etc.) can be complicated by recursive operations, whereby information gleaned from later steps can inform the reoperation of earlier steps.  This phenomenon in concert with choices, such as the employment and comparison of multiple approaches to obtain the estimate at any given step, can complicate the process of crafting an effective methodological report.  However, one must try to navigate an effective path through these complications to effectively communicate your program of research.

R1C2a response:

See R1C2 response. Yes, the question of treating the T-SSC relationship as simply a method to sample and calculate SSY or is it a preliminary result in and of its own. Since it is a widely applied method we have chosen to treat it as a sampling method to calculate SSY and treated it solely as Data Collection.  
  
R1C3. Discussion elements incorporated in the Methods and Results sections.  
You have elected to apply a Methods/Results/Discussion for the overall structure of the MS.  However, Discussion material has been interdigitated with Methods and Results material.  Reorganize as needed.

R1C3 response:

Done. Significantly reorganized the MS.

R1C4. Clear presentation of the relationship between sampling protocol and estimation of sub-watershed signals.  
Sub-watersheds along a mainstem channel are nested, rather than discreet, i.e. the watershed of FG2 includes the watershed of FG1 rather than only the additional drainage area between FG1 and FG2 as implied in Fig.1 and in the text (p. 206-210).  Make sure that you nail this point when introducing your study design/study region.

R1C4 response:

Done. The nested watersheds/subwatersheds are described in detail in 2nd paragraph of Study Area and the calculation of SSY from each subwatershed is described in the Methods overview and Methods section 3.2.1

R1C5. Discuss all findings in the context of relevant literature.  
The findings were very well discussed in the context of the canon of relevant literature, with the exception of section 5.1.

R1C5 response:

We added citations to relevant literature for section 5.1 where possible. Some Discussion findings are simply Discussion of observations ie interesting patterns in Q-SSC plots as a result of quarry operations (Section 5.1.1).

R1C6. Include the Appendix in your revision.  
The Appendix was missing from the MS.R1.  These materials contain a lot of information, some of which should be moved to the main body of the MS, and should certainly be included in the revision package. (I was able to access them from your first submission).

R1C6 response:

In EC9, the Editor requested the Appendices be submitted separately so the page count of the manuscript would be lower. In the interest of shortening the MS, as suggested by the Editor, the information in Appendices has been left there.

R1C7. Streamline sub-section headings.  
On a minor note, I would suggest rewriting the sub-section heading by removing the 'Objective #' component - the continuity in the heading phrases is sufficient.  
[Editor: good suggestion as it will provide a structure that is 100% clear.]

R1C7 response:

The subheadings have been revised and “Objective #” have been removed following R1C7 and the Editor comment agreeing with this change.

Technical   
  
R1C8. Minor to moderate issues were found in some technical aspects of the work related to data/sample collection, accounting for log-bias, and error estimates.  More information on the collection of some data/samples are needed, and a figure displaying the temporal distribution of monitoring/sampling efforts at each station would be a welcome addition.

R1C8 response:

We have clarified the sample collection method in the Methods overview. Please let us know if additional detail is necessary.

As for the temporal distribution of monitoring: Figure 3 shows when discharge was monitored at all stations and Tables 2 and 4 show when SSC was monitored and by which method. L216-218 also describe when monitoring was conducted. We opted against a separate figure showing sampling periods due to space constraints.

R1C8 cont. Log-linear (i.e. power law) relationships were used for some estimates of SSY and SSC, but the issue of log-bias was not investigated/corrected.  See Duan (1983), Ferguson (1986), Gray et al. (2015), and others for guidance on this issue.  I applaud the use of error estimates and the relatively transparent reporting of their computation.  However, some important details are missing (see specific comments.)

R1C8 cont. response:

- SSC was predicted by a linear regression of T vs. SSC, so no log-bias correction was needed.

-Annual sediment yield was estimated by calculating the SSYEV for storms without measured Q and SSC using the SSYEV-Qmax model. Model performance metrics (r2 and RMSE) are in Table 6. We thought it would be too confusing or cluttered to add error estimates to the Annual SSY estimates since there are several different estimation methods, and their respective error estimation methods would be different.   
  
Specific Comments (below these are mainly identified by line number)  
R1C9. L8-11 Over what time period were the observations collected?

R1C9 response:

Done. Added “(2012-2014)” to text L10  
  
R1C10. L16 Recommend stating that the human-disturbed watersheds 'have been estimated in increase loads'.  For presumably that is what you did, rather than monitor the situation over the course of the initiation of human disturbance.  If so, that should have been stated earlier.

R1C10 response:

Revised text: The human-disturbed subwatershed (5.2% disturbed) accounted for an average of 71-87% of SSYEV from the watershed. Observed sediment load to the coast, including human disturbed subwatersheds, was 3.9x the natural background.

R1C11. L22 A yield is area normalized by standard definition.  Here you are reporting loads (simply mass flux) and yields.  Such detail is not needed in an abstract - just report the yields.

R1C11 response:

Done. Only reporting yields L22-23.  
  
R1C12. L22 You defined SSYEV as event specific suspended sediment yield.  How does one have an 'annual SSYEV'?  Perhaps just 'annual SSY'?

R1C12 response:

Done. Changed to “SSY” L22-23.  
  
R1C13. L26 Replace 'sediment yield increased significantly' with 'sediment yield from this area was 3.9X higher than undisturbed areas' or the like.  Using terms like 'increased' in the abstract implies to many that you monitored the region over the period of change, which you did not.

R1C13 response:

This sentence was removed in rewriting the MS.  
  
R1C14. L36 Recommend changing 'Sediment yield' to Suspended sediment yield' as this study did not address bedload.

R1C14 response:

Done.  
  
R1C15. L45 Citation?

R1C15 response:

Revised MS text: “Anthropogenic sediment disturbance can be particularly high on volcanic islands in the humid tropics, where erosion potential is high due to high rainfall, extreme weather events, steep slopes, and erodible soils”

Cited Milliman, J.D., Syvitski, J.P.M., 1992. Geomorphic/tectonic control of sediment discharge to the ocean: the importance of small mountainous rivers. J. Geol. 100, 525–544.   
  
R1C16. L47 Citation?

R1C16 response:

Revised MS text: “Sediment yield in densely-vegetated watersheds can be particularly sensitive to land clearing, which alters the fraction of exposed soil more than in sparsely-vegetated regions.”

Sentence removed.   
  
R1C17. L48 Citation? i.e. Milliman & Syvitski 1992 or a more tropical specific pub?

R1C17 response:

Revised MS text:

“The steep topography and small floodplains on small volcanic islands further limits sediment storage and the buffering capacity of the watershed against increased hillslope sediment supply (Walling, 1999)”

Cited Walling, D.E., 1999. Linking land use, erosion and sediment yields in river basins. Hydrobiologia 410, 223–240  
  
R1C18. L79 Word choice: stating that an environment is 'challenging' does not inform one as to the 'challenges of…monitoring.'

R1C18 response:

Done. Removed “challenging”

Revised MS text:

“Knowledge of suspended sediment yield (SSY) under both natural and disturbed conditions on most tropical, volcanic islands remains limited, due to the challenges of in situ monitoring in these remote environments.”

R1C19. L85 Define SSYEV here, as it is used in the body of the MS for the first time.

R1C19 response:

Removed “EV” since SSY is sufficient  
  
R1C20. L85-88 Awkwardly worded sentence, revise.

R1C20 response:

Done.

Revised MS text:

“Models that predict SSY from small, mountainous catchments would establish baselines for change-detection, and improve regional-scale sediment yield models (Duvert et al., 2012).”

R1C21. L92 Citation for end of the sentence to illustrate 'interannual…relationship'.  Recommend Gray et al. 2014 or the like.

R1C21 response:

Cited Gray et al. 2014 and added citation to Stock and Tribble 2010 for hysteresis in Hawaiian watersheds and Kostachuk 2002 for Fijian watershed.

R1C22. L104 Now only use the SSYEV acronym.

R1C22 response:

Done, except when presenting annual SSY.  
  
R1C23. L133 Erosivity Index should be introduced before mention here.

R1C23 response:

The Erosivity Index was previously introduced with a reference in L105 (Kinnell 2013), and in the Abstract L22.   
  
R1C24. L139 Section 2. Identify your study sub-watersheds in this section.  I would suggest turning away from your FILE\_BASED naming convention to something that isn't so jarring on the page.

R1C24 response:

Names were changed to lower case: Lower\_Quarry instead of LOWER\_QUARRY, and dentified sub-watersheds in the Study Area section, L139-L148, in the reference to Figure 1.

Now there are descriptions of the subwatersheds in Figure 1, in the second paragraph of Study Site, second paragraph of Methods section 3.2 L311-319, and in Table 1.

R1C25. L142 '(~3 km)' presumably stream length: label as such.

R1C25 response:

Done.  
  
R1C26. L201 The opening paragraph of the Methods section should provide a more comprehensive overview.  I realize that the Introduction ended with an overview of objectives and goals, but here the technical skeleton can be drawn.

R1C26 response:

Added MS text:

An additional paragraph was added to the Method section that introduce each of the subsequent sections:

The field methods used to calculate event-wise suspended sediment yield (SSYEV) are described in section 3.1. The equations and analytical methods used to accomplish Objectives 1-3 are described in sections 3.12-3.3, and 4. Briefly, the in-stream suspended sediment load (tons) and yield (SSY, tons/km2) were calculated for individual storm events (SSYEV) at three locations in Faga’alu watershed using calculated discharge (Q) and suspended sediment concentration (SSC)(Figure 1) during four field campaigns (Section 3.1). Each subwatershed had distinct land cover (forest at FG1, quarry and forest at FG2, and village and forest at FG3). Precipitation was recorded with a tipping bucket raingage (Section 3.1.1). Q was calculated from continuously recorded stage and a stage-discharge relationship calibrated with field measurements (Section 3.1.2). SSC was measured directly from grab samples or modeled from continuously monitored turbidity (T) and T-SSC relationships calibrated to in-stream SSC (Section 3.1.3). Storm events were identified using automated hydrograph separation, and SSYEV calculated for each monitored location with Q and SSC data (Section 3.2.1). The subwatersheds were nested, so SSYEV contributions from subwatersheds were calculated by subtracting SSYEV at the upstream subwatershed from SSYEV at the given downstream subwatershed. The sediment yield from disturbed surfaces was calculated assuming a uniform yield from forested parts of disturbed subwatersheds (Section 3.2.2). The cumulative probable error of SSYEV was calculated for each storm to incorporate errors in Q and SSC (Section 3.2.3). Log-linear regression models were developed to predict SSYEV from storm metrics for the undisturbed and disturbed subwatersheds (Section 3.3). Annual SSY was estimated from the regression models and the ratio of annual storm precipitation to the precipitation during storms where SSYEV was measured (Section 3.4).

R1C27. L204-206 Start by informing us about your monitoring program.  The three metrics listed here: Q, SSC, and SSY were not measured during the storms - you collected measurements and samples which were used to compute estimates later.  What did you actually measure?  This should be the leading section here, rather than cropping up after sections detailing the calculations of loads, yield and indices that rely on your monitoring data.

R1C27 response:

Descriptions of the data collection have been moved to the beginning of the Methods section.

See also responses to R1C2 and R1C2a

R1C28. L206-210 Here you introduce your watershed labels for the first time.  This belongs in the previous section.  More importantly, the wording of this section implies that you are sampling only the additional sub-watersheds at FG2 and FG3 rather than the integrated expression of all sub-watersheds upstream as all of your gage sites are situated along a mainstem drainage.  This is an important distinction, as your analysis then relies on subtracting the loads calculated for upstream gage stations in order to arrive at estimates of sediment loads from given sub-watersheds.  The entire study is designed around this concept, so it must be clearly communicated.

R1C28 response:

Done. Subwatersheds are defined in Study Area L139-147. SSY calculations are described in Section 3.2 L311-319. Here, in the overall description of Methods, we indicate L210-211 “SSYEV contributions from subwatersheds were calculated by subtracting the contribution of the upstream subwatershed.”  
  
R1C29. L221 What is 'quickflow?'

R1C29 response:

Hewlett and Hibbert (1967) proposed the term “quick flow” (also appears as “quickflow” in Dunne and Leopold (1978)) to describe the portion of the hydrograph caused by direct surface runoff during storms, to differentiate it from what they termed “delayed flow” or “baseflow”, or the portion of the hydrograph that remains more or less constant during non-storm periods, fed by subsurface water. The EcoHydRology package in R separates the observed hydrograph into “quickflow” and “baseflow”, we use the quickflow to define when storms occurred.

Revised text: Due to the large number of storm events and the prevalence of complex storm events observed at the study site, we used a digital filter signal processing technique (Nathan and McMahon, 1990) in the R-statistical package EcoHydRology (Fuka et al., 2014), which separates the hydrograph into quickflow, or direct surface or subsurface runoff that occurs during storms, and baseflow or delayed flow. Quickflow and baseflow components are not well defined in terms of hydrologic flowpath; here we use the separation operationally to define storm events.  
  
R1C30. L221 You must mean 110% rather than 10%.

R1C30 response:

Done. Revised text:

“Spurious events were sometimes identified due to instrument noise, so only events with quickflow lasting at least one hour and peak quickflow greater than 10% of baseflow were included (See Appendix C for example).”

R1C31. L225 I am confused by your notation with the introduction of a new term 'specific sediment yield' (sSSY), which you identify as having units of mass/volume, while again identifying SSY as having units of volume. I suppose this is fine, as long as you are consistent (proof the MS diligently for this). But keep in mind that most literature identifies the mass flux of sediment as 'sediment load' or 'sediment discharge', while sediment yield usually denotes mass/area. I would suggest sticking with the more widely used convention.

R1C31 response:

The literature has varied definitions of yield and load. Our use of the terms SSY and sSSY is in line with Walling and Fang (2003) and Walling in other related publications who refer to “specific sediment yield” with units of tons/km2, and several others who also refer to “specific sediment yield” (e.g. Lenzi et al, 2003) as tons/km2. Milliman and Meade (1983) and Syvitski et al (2005) use yield as tons/km2 as suggested by the reviewer.

In L225 we define specific suspended sediment yield as “sSSY” with units of mass per area (tons/km2). We tried to keep from using “load” and “yield” so as not to confuse the reader, and use “specific yield” and define the units

R1C32. L230 Here you use the sediment load.  Again, I would suggest sticking with convention (i.e. sediment load (mass), sediment yield (mass/area)).

R1C32 response:

See above response to R1C31.  
  
R1C33. L230-244 Section 3.1.3:  This section belongs in the discussion.

R1C33 response:

I agree that this section may be out of place here but I disagree that it belongs in Discussion. This section describes critical Methodological assumptions about our calculation of sediment yield and justifies our rationale for ignoring some components of the complete sediment budget like floodplain and in-channel deposition. If this was moved to Discussion, the reader may think our Methods are inappropriate or inadequate and be skeptical of our Methods and Results until much later in the Discussion section.

We moved this to the beginning of Methods section L197-207 so that we say up front that we’re only calculating suspended yield and why. Hopefully the reader will then not be looking for how we calculated bedload. The MS then outlines the overall sampling and analytic methods before moving on to describing Field Data Collection.  
  
R1C34. L231 Yes, an at-a-station fluvial suspended sediment budget, the counterpart to which is the fluvial bedload budget, which you should mention here as the component that was not measured.

R1C34 response:

Done. See response to R1C33.  
  
R1C35. L247 (second to last sentence) Define 'Erosivity Index'.

R1C35 response:

Added MS text: “The Erosivity Index describes the erosive power of rainfall (Kinnell, 2013).”  
  
R1C36. L247 (last sentence) Isn't SSYEV 'normalized by watershed area' sSSYEV by your notation system?

R1C36 response: Yes, that’s correct. We kept the notation SSY to match with other literature, especially Duvert et al (2012), but wanted to explicitly point out to the reader that both SSY *and* the Q metrics are normalized by watershed area for the purpose of comparing the different sized watersheds.  
  
R1C37. L254-273  Section 3.3: Please begin such sections with an overview of what was done in the first paragraph.  In the first paragraph of this section you write broadly what was done, followed by motivation, followed by two of the approaches (Psum-SSYEV; Qmax-SSYEV), without ever introducing the third approach (SSYEV x PEVann/PEVmeas).  Tell us upfront what your approaches were, and then report them in detail.

R1C37 response:

Done.

Added MS text:

“Annual SSY and sSSY were estimated using 1) the developed storm metric-SSY models, and 2) the ratio of annual storm precipitation to precipitation measured during storms with SSYEV data.”  
  
R1C38. L255 Style: That 'estimates' were 'estimated' is clear, without the clunky construction.

R1C38 response:

Done.  
  
R1C39. L272-273 What does 'low' mean?  Back up qualitative statements with quantitative examples.

R1C39 response:

Revised MS text:

“Equation 6 also ignores sediment yield during non-storm periods, which is justified by the low SSC (typically under 20 mg/L) and Q (baseflow) observed between storms.”  
  
R1C40. L274-287 Section 3.4:  This should be the first paragraph of your methods section.

R1C40 response:

Done. The Methods section has been significantly reorganized. See our responses to R1C2 and R1C27  
  
R1C41. L275-278 A figure illustrating the temporal distribution of sample/monitoring data collection by station would be appreciated.

R1C41 response:

It is difficult to visualize data from over 100 storms at three stations, or over the monitoring period given the various instrument errors at various locations. Instead we elected to show an example of one storm (Figure 4) to convey the general pattern of data collection and sediment dynamics during storms, and provide record of the temporal distribution of monitoring data in Figure 3, Tables 2 and 4. See our response to R1C8.

R1C42. L282-283 Report rain gage specifications (size, resolution).

R1C42 response:

Done, though the size certainly seems like a technical detail that could be looked up since we provided the make/model.

Added MS text:

P was measured in Faga'alu watershed from January, 2012, to December, 2014, using two tipping-bucket rain gages (RG1 and RG2; 20cm dia., 1 minute resolution) and a Vantage Pro Weather Station (Wx; 20cm dia. 15 min resolution) (Figure 1).   
  
R1C43. L305-306 Assumed to be the same rather than just 'similar'. Is this a valid assumption? How do the major factors thought to contribute to specific water discharge differ between these sub-watersheds (average slope, vegetation)? A brief statement supporting this assumption would suffice.

R143 response:

Added MS text:

“The specific water discharge at FG2 is assumed to be the same as above FG1 since average slopes, vegetation, and soils of the watersheds are extremely similar.”

We also describe the caveat that the quarry surface is continually disturbed and these are conservative estimates of SSY as a result.  
  
R1C44. L320-322 Where was the ISCO's inlet positioned? Fixed or varied btw storms?

R1C44 response:

Inlet was fixed throughout the study.

Added MS text:

“The Autosampler inlet tubing was oriented down-stream, just below the water level sensor, approximately 30 cm above the stream bed, on rebar positioned midstream.”

R1C45. L327-328 Why would you assume that SSC would ever be zero - particularly if you have enough baseflow samples over the course of your study to characterize this condition?

R1C45 response:

A similar approach was used in the literature (Lewis et al, 2001), and from field observations at the beginning of many storms it seemed like this was a good enough assumption based on the median SSC values during Non-Storm periods (varying from 5-60 mg/L at FG1-3). Even if SSC at the storm start was assumed to be higher it wouldn’t make much of a difference in the total SSY calculation since it only spanned one or two intervals until a sample was collected, usually with far higher SSC in excess of several hundred mg/L.  
  
R1C46. L341-343 The T-SSC relationship is too important to completely burry in your appendix.  The basic details should be summarized here: model (i.e. linear, log-linear, etc.), were model assumptions met, etc.

R1C46 response:

Indeed, whole papers detailing the calibration of a T-SSC relationship for a single instrument and site are published (Minella et al., 2008). We felt that these methods are well-established at this point and we didn’t have any novel contributions.

The T-SSC models were linear, added MS text: “A unique*, linear* T-SSC relationship…”. We also added key assumptions:

The critical assumption in our application is that the parameters of the T-SSC relationship are stable over time and among storm events. The T-SSC relationship can be unique to each region, stream, instrument or even each storm event (Lewis et al., 2001), and can be influenced by water color, dissolved solids, organic matter, temperature, and particle shape, size, and composition.

Here the Methods section summarizes the r2 values of the T-SSC relationships (which were high: 0.79-0.99), and RMSE % errors were added to Methods section. RMSE (mg/L) and % errors are also in Results section 4.2. Cumulative Probable Error. We thought this was a better place to describe the T-SSC relationship errors, showing the reader not only the error in the T-SSC relationships but how those errors propagate to the final SSY estimates to increase uncertainty. The reader is likely only interested in the details and errors of the T-SSC relationships for their impact on the final SSY calculations and the conclusions drawn from those calculations. Rather than simply provide the raw estimates of model error and leave the reader to infer how those errors impact the final calculations, we provide the error estimates in the context of the final SSY estimates.

A possible criticism could be that maybe SSY was higher at FG3 because of the steeper T-SSC relationship for the YSI at that location, compared to the YSI at FG1. We conducted a sensitivity analysis where we substituted the T-SSC relationship developed at FG1 for the steeper T-SSC relationship developed at FG3, to see how much the DR and % contributions would change. Using the same T-SSC relationship for both locations we found the DR decreased from 3.9 to 2.5, and the percent contributions changed from 13% Upper, 87% Total, to 20% Upper, 80% Total. We concluded these were relatively small changes to the overall story and maintained our overall conclusion that human disturbance has significantly increased sediment yields to coast.

If the reader requires more information on the T-SSC relationships, they are provided in the Appendix and we have added subplots zoomed in on the lower T-SSC values to increase confidence in the line-fitting. The intention was to not distract the reader from the main objectives or to lengthen the text with too much detail about the intermediary steps of the T-SSC relationship. The details of the T-SSC were provided in the Appendix because they are so critical, as R1 points out. We tried to strike a balance of placing the full technical detail in an Appendix while providing relevant information for addressing the research questions in the MS.

R1C47. L350 Equation 2: I applaud the transparent reporting of your method for sediment load uncertainty estimations.  However, this equation appears to only relate to one approach to obtaining SSC (T-SSC rating curve), whereas your study employed a number of methods (linear interpolation of between SSC values, T-SSC rating, and assumed values for base flow).  Also, if I understand your equation correctly, you don't need the summation symbol, as you indicate this operation with '+' signs.  
Taking a step back, I would also suggest reconsidering your approach to error estimation.

R1C47 response:

We removed the summation symbol, and added error for interpolated grab samples to SSC measurement error.

Yes, the T-SSC model errors are included explicitly, the interpolated grab sample error was taken from the DUET-H/WQ look up table (5%) and included as measurement error for those SSYEV estimates. Only 3/42 storms used interpolated grab samples to calculate SSYEV.

As far as reconsidering the approach, there are definitely many different ways to estimate uncertainty so any suggestion of which method would be more appropriate are welcome. We chose this method since it was straightforward to calculate and well-documented in previous literature and a software tool (Harmel, Topping etc.).   
  
R1C48. L356-359 Why not calculate the uncertainty associated with this estimate?  It is after all just another estimate, merely arrived at by arithmetic processing of two other estimates…

R1C48 response:

Yes it could be simple arithmetic to take the mean of the UPPER and TOTAL error estimates, or following the cumulative probable error method it would be necessary to calculate PE based on all measurement and model errors at FG1 and FG3, thereby increasing the uncertainty in FG2 even more. There was no clear evidence that either was the appropriate method and it would significantly complicate the error estimation to include another section in the MS on comparing error estimates from different methods.

R1C49. L365 No difference at all?  Would be nice to see those data, at least in an appendix.

R1C49 response:

Revised MS text:

“Linear relationships between daily P at RG1 and Wx (slope=0.95, r2=0.87) and RG1 and RG2 (slope=0.75, r2=0.85) were observed. Higher P was expected at higher elevation at RG2 so lower P at RG2 was assumed to be caused by measurement error, as the only available sampling location was a clearing with high surrounding canopy.”  
  
R1C50. L419-427 Discussion material. Some interpretation of results certainly belongs in the Results section, but there is a line, and this appears to be over it.

R1C50 response:

Done.

Moved to Discussion section 5.1.1 Compare SSC for disturbed and undisturbed watersheds in Faga’alu  
  
R1C51. L476 It appears that you have chosen your SSC grab sample error from a table.  Do you think this is a realistic error in terms of the depth-concentration profile expected for the flow fields and particle size distribution characteristics that you encountered?

R1C51 response:

Without any data to estimate this, I can only assume this is a realistic error estimate. The study site stream is extremely well mixed given the high slope and energetic flow. Our motivation was mainly to quantify sediment yields to coral reef areas, and so our focus was particles that would stay in suspension in fairly calm water, long enough to be transported from the stream mouth to corals. Hence, we did not do any depth-integrated sampling.  
  
R1C52. L485-486 How did you arrive at 5% error associated with interpolation between SSC samples?  Also, you have estimated your grab sample error as 3% - do you think this is a realistic error in terms of representation of the average SSC in consideration of the depth-concentration profile expected for the flow fields and particle size distribution characteristics that you encountered?

R1C52 response:

The 5% error was taken from the DUET-H/WQ lookup table, default value (values can range from 0-21%) (Harmel et al., 2009). Considering SSC values of grab samples regularly exceeded 1,000 mg/L we’re considering 3% errors of 30 mg/L or more. See our response to R1C52.   
  
R1C53. L486-487 Was sample analysis error estimated from replicates, on the basis of known compounded measurement precision, or simply assumed?

R1C53 response:

Errors were taken from DUET-H/WQ software Lookup Tables, using default values, unless they were from the models provided in the Appendices. We did not do replicate sample analysis.  
  
R1C54. L487-489 Did your SSC-turbidity models meet the basic assumptions for attributing a single error estimate to dependent variable estimates produced over the entire independent variable domain (i.e homoscedastic and normally distributed residuals)?  Also, as these models are buried in an appendix, the reader has no knowledge at this point as to their basic formulation (log-linear? Basic assumptions for regression met?).  See Helsel and Hirsch (2002) for guidance on regression assumptions and error estimates from regression.  As the MS is long, it makes sense to shunt methodological details to an appendix, but such fundamental details should still be reported in the main text.

R1C54 response:

See our response to R1C46.  
  
R1C55. Table 2. Ditch the 'Upper' and 'Lower' and just report the gage codes, as that is what one can reference off of the study location map.

R1C55 response:

We also struggled with how to refer to the subwatershed areas versus the sampling locations. As pointed out in R1C28 all gages on the mainstem channel receive flow from upstream subwatersheds. If referred to as gage codes FG1, FG2, FG3 these would imply the SSY measured at these gages. The terms LOWER, LOWER\_QUARRY, etc. mean these SSY values in Table 2-5 are calculated as the difference of the measurements at the gages to isolate the runoff from particular subwatersheds by subtracting runoff from upstream subwatersheds. These watershed labels are referenced on the study location map just as the gages are (Fig. 1). These labels are also intended to quickly convey which sediment sources are being referenced, ie Upper means natural forest, Lower\_Quarry means from the quarry in the Lower, human-disturbed watershed etc.  
  
R1C56. L492-493 Reference the relevant figure/table.

R1C56 response:

Done.  
  
R1C57. L519-520 Reiterate what you are comparing here (between stations) for the sake of clarity - this is a long MS with multiple components and many readers would have to flip back to the methods here to refresh.  Also, reference the relevant table/figure.

R1C57 response:

Done.  
  
R1C58. L565-580 Incorporate previous studies into your discussion (as you have done in 5.2).

R1C58 response:

Done. See our response to R1C5.  
  
R1C59. L597-599 Note that with Asselman (2000) you are referring to Q-SSC relationships rather than Qmax-SSY.

R159 response:

Yes, Asselman (2000) was just a good reference for interpreting slope values in the Q-SSC relationship, but we feel it would extend to Qmax-SSY as well.  
  
R1C 60. L628-639 Nice follow-up with Milliman and Syvitski (1992) relationship.  Also note the size of your headwater drainage area relative to their study areas, and how sediment yield generally scales with drainage area, which would further support your findings.  Also note the short duration of your study.

R160:

Done.

Added MS text:

“Faga’alu is also a much smaller watershed and the study period was relatively short (3 years) compared to others included in their models.”  
  
R1C61. L640-665 Note the monitoring base period for these studies in comparison to each other and yours.

R1C61 response:

Done. To shorten the MS in response to Editor and R1 request, this section was significantly cut down.

Revised MS text:

“Sediment yield was measured from two Hawaiian watersheds which are physiographically similar though much larger than Faga’alu: Hanalei watershed on Kauai (“Hanalei”, 54 km²), and Kawela watershed on Molokai (“Kawela”, 14 km²) (Table 8) (Ferrier et al., 2013; Stock and Tribble, 2010). Hanalei had slightly higher rainfall (3,866 mm/yr) than Faga’alu (3,247 mm/yr) but slightly lower SSC (mean 63 mg/L, maximum of 2,750 mg/L) than the Total Faga’alu watershed (mean 148 mg/L, maximum 3,500 mg/L) (Ferrier et al., 2013; Stock and Tribble, 2010). Kawela is drier than Faga’alu (P varies with elevation from 500-3,000 mm) and had much higher SSC (mean 3,490 mg/L, maximum 54,000 mg/L) than the Total Faga’alu watershed. SSY from Hanalei was 369 ± 114 tons/km2/yr (Ferrier et al., 2013), which is higher than the undisturbed subwatershed in Faga’alu (45-68 tons/km2/yr) but similar to the disturbed (430-441 tons/km2/yr) subwatersheds. Stock and Tribble (2010) estimated SSY from Kawela was 459 tons/km²/yr, similar to the disturbed Lower Faga’alu watershed, but nearly twice as high as the Total Faga’alu watershed. Overall, both Hawaiian watersheds have higher SSY than Faga’alu, which is consistent with the low Qmax-SSYEV intercepts and suggests Faga’alu has relatively low erosion rates for a steep, volcanic watershed. Precipitation variability may contribute to the difference in SSY, so a more thorough comparison between Hanalei and Faga’alu would require a storm-wise analysis of the type performed here.”

R1C62. Appendix 4 As the T-SSC relationships are a fundamental pillar of your SSY estimations, they need to be detailed here in terms of model formulations, model assumptions, and statistical descriptors of fit.  More realistic assessments of total model (SSY) error is contingent upon the error associated with these estimates.

R1C62 response:

Statistical descriptors of fit: we provide r2, RMSE and % error, similar to Minella et al. (2008) who wrote a whole paper on the calibration of a single T-SSC relationship.

References  
  
Duan, N., 1983. Smearing estimate - a nonparametric retransformation method. J. Am. Stat. Assoc. 78 (383), 605-610.  
  
Ferguson, R.I., 1986. River loads underestimated by rating curves. Water Resour. Res. 22 (1), 74-76.  
  
Gray, A.B., Warrick, J.A., Pasternack, G.B., Watson, E.B., Goñi, M.A., 2014. Suspended sediment behavior in a coastal dry-summer subtropical catchment: Effects of hydrologic preconditions. Geomorphology, 214(0): 485-501. DOI:http://dx.doi.org/10.1016/j.geomorph.2014.03.009  
  
Gray, A.B., Pasternack, G.B., Watson, E.B., Warrick, J.A., Goni, M.A., 2015. The effect of El Nino Southern Oscillation cycles on the decadal scale suspended sediment behavior of a coastal dry-summer subtropical catchment. Earth Surface Processes and Landforms, 40(2): 272-284. DOI:10.1002/esp.3627  
  
Helsel, D.R., Hirsch, R.M., 2002. Statistical Methods in water resources - hydrologic analysis and interpretation.  
  
Milliman, J.D., Syvitski, J.P.M., 1992. Geomorphic/Tectonic Control of Sediment Discharge to the Ocean: The Importance of Small Mountainous Rivers. The Journal of Geology, 100(5): 525-544.  
  
Rouse, H., 1937. Modern conceptions of the mechanics of fluid turbulence. Trans. Am. Soc. Civ. Eng. 102, 463-541.  
\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*  
  
  
Reviewer #2:

R2C0. Generally this paper is now in very good shape and I recommend it for publication more or less as it stands. I note that it has already gone through the review process so it seems well polished. I am notoriously bad at picking up typos and only picked up a couple of problems.

R2C0 response: Thank you, we appreciate the positive review.

R2C1. The conversion factor in equation 1 from mg to tons which should be 10^-9 .  
R2C1 response:

Done. Good catch.

R2C2. The units drawn on figure A2.2 need to use superscripts for square (2) and cubic (3) meters

R2C2 response:

Done.  
  
R2C3. In line with modern trends I think the Introduction section and Discussion section are both too long. In my view the results stand by themselves and need little elucidation. But this is probably more a reflection of my age and mathematical bent where equations are my friend and words are not, than a problem with the paper. I'll leave that one to the editor.

R2C3 response:

We cut the MS down by 3%.  
  
R2C4. I was particularly happy that there was lots of data used in this paper and there was a decent error analysis. To show just how pedantic a reviewer can be, I object to the three significant figure accuracy claimed for the RMSE used in line 484 (16.3%). Perhaps leave this at 16% and similarly in line 479 (8.5  to 8%).

R2C4 response:

Done.

References

Lenzi, M. A., Mao, L., & Comiti, F. (2003). Interannual variation of suspended sediment load and sediment yield in an alpine catchment. *Hydrological Sciences Journal*, *48*(6), 899–915.

Milliman, J. D., & Meade, R. H. (1983). World-wide delivery of river sediment to the oceans. *The Journal of Geology*, 1–21.

Minella, J.P.G., Merten, G.H., Reichert, J.M., Clarke, R.T., 2008. Estimating suspended sediment concentrations from turbidity measurements and the calibration problem. Hydrol. Process. 22, 1819–1830. doi:10.1002/hyp.6763

Syvitski, J. P. M., Vörösmarty, C. J., Kettner, A. J., & Green, P. (2005). Impact of Humans on the Flux of Terrestrial Sediment to the Global Coastal Ocean. *Science*, *308*, 376 – 380, 10.1126/science.1109454.

Walling, D. E., & Fang, D. (2003). Recent trends in the suspended sediment loads of the world’s rivers. *Global and Planetary Change*, *39*(1-2), 111–126.