1 Summary

I appreciate the opportunity to review Shadrick and others (submitted to Earth Surface Dynamics, 2021). In this study the authors estimate millennial scale cliff erosion rates using a coupled model of rocky coast evolution and cosmogenic radionuclide production. The authors implement a multi-objective optimization approach to understand the relative roles of ^{10}Be concentrations and topographic profiles on constraining model parameters. Because the authors consider two sites, they are able to provide insight into the extent to which results are general or site specific. The study is well designed and well described. The work reflects a novel application of optimization to interpreting geologic data and it represents a big advance in using this type of methodology to interpret geologic data and geomorphic models. The authors nicely explore a variety of important topics including the relative contribution of different types of data to parameterize a coupled model, and the role of parameter covariance and equifinality.

I recommend it for publication after minor revision.

2 Narrative comments

1. Provide more explanation and/or justification for the 8 kyr model duration. Was this chosen because it is the extent of GIA modeling, needed for RSL as a boundary condition, or because of some other reason? On this same point, I'd recommend providing a bit of context to the reader regarding typical timescales to reach a steady state topographic profile under steady forcing. It is clear that this is not your aim, since you use nonsteady RSL forcing, and will of course depend on parameter values... however, it will be helpful context for how far away from equilibrium the coastal profile would be over a 8 kyr duration.

2. Similarly, provide additional information regarding what initial conditions were used in the model, and whether the model shows sensitivity to the initial conditions over the timescale simulated.

3. Recommend that most of section 3.4 (MCMC inputs) be moved into section 3.2 (The coastal evolution model). Specifically, I would recommend adding to section 3.2 an introduction to the parameters y and K (as F_R is already introduced) as well as a brief description of other parameters in the model presented by Matsumoto, Dickson, and Kench (2016) which are not considered here. It is certainly reasonable to not consider these parameters, but let the reader know a bit more about what they are.

   I would then recommend putting most of section 3.4 at the end of a revised section 3.2. This will help the reader understand what the model and parameters are before you begin section 3.3 and discussion of Dakota.

4. I was surprised not to see a plot of the pareto front itself (that is, a plot with the scaled and weighted topographic RMSE on one axis and the ^{10}Be RMSE on the other axis, with one line for each of the two sites). The simulated topography and ^{10}Be concentrations are provided in Figure 4, but the pareto trade-off itself is important to visualize in a multi-objective function study. As the multi-objective nature of this study is so novel, I'd strongly recommend adding such a subplot to Figure 4. It would support the text from lines 435–448, namely that the hump in the Bideford ^{10}Be data is best fit only when the topographic fit is mostly ignored.

5. None of the simulations seem to capture well the increase in slope in the most shore-ward 50 m of the the Scalby site. Could you discuss this more? Are there field observations from Scalby that might provide more context to this topographic feature?

6. Your results show that cliff retreat rates match RSL rise closely. But they also show (Figure 6) that scaled RSL is on the lower end of the Bideford retreat rates, while it is on the upper end of the Scalby rates (dashed line is below(above) the mean line for Bideford (Scalby)). Is this an interpretable result?

Are there differences in the wave climate, geomorphology, or geology of the sites that could explain this?

7. You nicely discuss parameter covariance near the end of the manuscript. However, there were two points related to parameter covariance that I think are important to discuss. First is related to your specific implementation. You set $K = 5^c F_R = 5^c 10^b$, removing some parameter covariance that would have existed otherwise. Why not treat them as fully independent and document the nature of the covariance as you do for a and b in section 5.2.

   A second, bigger picture comment that you are well poised to make is about the “meaning” of these model parameters. Namely, when model authors write models, we often think of the parameters as independent and meaningful—and sometimes linked to field or laboratory measurements (Dietrich et al. 2003). But it is not uncommon to find these sort of covariance issues, most often because of how these parameters are used. In the case of the model presented by Matsumoto, Dickson, and Kench (2016) is this something that could be anticipated because of the mathematical form of the model? Or is it fully emergent.

8. One theme that emerges from both sites is the overwhelming influence of the RSL boundary condition. Looking at the different cliff retreat rates in Figure 4 it seems reasonable to conclude that no matter the topography vs ^{10}Be weighting the estimated cliff retreat rates match the RSL forcing. It might be worth commenting in the discussion on the relative role of the boundary condition vs parameter values for this type of analysis.

9. Overall all figures are well designed and clear. However, my black and white printer led to this comment: consider a grayscale safe color scales (e.g., viridis used in Fig 9) and/or different symbols for Bideford and Scalby sites. Colorbrewer is a good resource for this.

3 Line level comments

Bullet points in this Section indicate “<LineNumber>”, “T<Table Number”, or “F<Figure Number>”.

15 Perhaps add a statement about the sort of timeframes that would be possible without this method. Such a statement would make clear by contrast the benefit of this approach.

185 The model represents a cross-section, yes? Recommend using the term “cross section” if this is the case.

220 At the end of this paragraph I wanted a sentence or two introducing the concept of the pareto front.

231 I think you need to add something like the phrase “with different weights” to the end of this sentence, because it is specifically the nature of the different weights that allows you to explore the pareto front.

235 I find it helpful to include a section or subsection “Model Implementation” that puts together information like the cell size, as well as a few items not stated. For example, what timestep was used? This would also be where you might state the simulation duration and why it was chosen.

242 You mention measurement error here, but it is not clear whether measurement error was incorporated into the RMSE. In your case, I’d expect you could have a different measurement error for each ^{10}Be observation, and that including/excluding this might have an impact on the Bideford results because the highest ^{10}Be measurements also have the highest error.

246 Why is w_i in a square root? I would expect just w_i here such that $\sum_{i=1}^{N_i} w_i = 100$.

255 Your discussion and approach to scaling makes sense when I read it here. However, after reading this portion of the methods I was confused when I got to Figures 7, 8, and 9 because I was expecting the topographic and ^{10}Be values to be of similar magnitude (which they are not). I think my confusion could be addressed with minimal revision to sections 3.3.2 and 3.3.3. Specifically, I would recommend talking only about constructing the topo and ^{10}Be RMSE values in the first of these sections and introduce an equation that looks something like

$$RMSE_i = \sqrt{\frac{1}{N_j} \sum_{j=1}^{N_j} \left(\frac{Mod_{i,j} - Meas_{i,j}}{\sigma_{i,j}} \right)^2} \tag{1}$$

Here I've also added a term for measurement error $\sigma_{i,j}$. Then, after discussing how the two RMSEs were each constructed, in the second section introduce scaling/weighting and constructing the pareto front with an equation like

$$TotalRMSE_p = \sum_{i=0}^{N_i} \frac{w_{i,p}}{s_i} RMSE_i \tag{2}$$

This would more clearly separate the construction of your two objective functions from the scaling and weighting of them for the pareto analysis since you don't need to discuss scaling/weighting until you get to the discussion of multiobjective and the pareto front. Similarly, I would recommend adding a subscript of some sort (I've used $_p$ above) to denote that the weights change depending on which pareto set is being used.

317 Recommend introducing y and K earlier, as discussed in the narrative comments.

392 This is the first point in the text where the simulation duration of 8 kyr is mentioned. I would recommend mentioning it earlier as well as providing context regarding why that duration was used at that point in the text.

458 Since you only consider one point on the pareto front, I think this section title is misleading. Consider revising.

519 The last sentence of the figure caption is confusing to me. I think you mean to say that the gray shaded region corresponds to cliff retreat rates which occurred when the cliff was further offshore than your topographic profile measurements. Revise for clarity.

683–685 Recommend framing this differently. Rather than making a statement about confidence, I'd recommend pointing out that it is rare to formally evaluate how different data sources (^{10}Be, topo) constrain a model differently. Each of these are valid data, and if you had only one (most commonly, no ^{10}Be) it would be totally fine to parameterize the model with only the available data (see, for example, any paper fitting river long profiles to some sort of fault history and/or value of fluvial erosion coefficient). But because you have both types of data you have the opportunity to evaluate the relative *information* provided by each data source. If, as is not the case here, both data sources yielded the same parameter estimates, we would learn we don't gain anything from the second dataset (and it is thus not necessary to make those observations). In contrast, as you find, if the two data sources contain different information, these datasets pull the coupled model in different directions. An potentially relevant reference here is Furbish (2003).

T1 Are the weights listed w_i or $\sqrt{w_i}$. Please clarify.

T2 Why is the base for K 5 rather than 10?

F3 I'd recommend adding one more loop here to denote that this workflow is done for each location on the pareto front. You nicely emphasize how your approach could be generalized to include additional objective functions, and this could help emphasize that it can be generalized to additional points on the pareto front.

F5 Because the right column represents a zoom into a portion of the left column, a gray rectangle or similar that indicates this region in the left column would guide the reader.

F6 Similarly a box in the upper right of each panel showing the extent of the inset would be helpful.

F6 Why only show 7 kyr when the simulation duration is 8 kyr. If the first 1 kyr is a "burn in period" to forget the initial conditions, that is reasonable but should be stated.

F7,8,9 Rather than plotting the objective function itself in Figures 8 and 9 I would recommend plotting the posterior. It is often easier to interpret because it does not have the same issues with overplotting that the objective function provides. One nice tool for this is corner.py (Foreman-Mackey 2016). It nicely also shows the marginal distributions.

F10 It is not clear if these three different sets (orange, blue, pink) come from different pareto sets, or different samples from the 50-50 evaluation. Clarify.

756 This is a very nice subsection and analysis.

References

Dietrich, W. E., D. G. Bellugi, L. S. Sklar, J. D. Stock, A. M. Heimsath, and J. J. Roering (2003). "Geomorphic Transport Laws for Predicting Landscape form and Dynamics". In: *Prediction in Geomorphology*. American Geophysical Union (AGU), pp. 103–132. DOI: `https://doi.org/10.1029/135GM09`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/135GM09`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/135GM09`.

Foreman-Mackey, D. (2016). "corner.py: Scatterplot matrices in Python". In: *Journal of Open Source Software* 1.2, p. 24. DOI: `10.21105/joss.00024`. URL: `https://doi.org/10.21105/joss.00024`.

Furbish, D. J. (2003). "Using the Dynamically Coupled Behavior of Land-Surface Geometry and Soil Thickness in Developing and Testing Hillslope Evolution Models". In: *Prediction in Geomorphology*. American Geophysical Union (AGU), pp. 169–181. DOI: `https://doi.org/10.1029/135GM12`. eprint: `https://agupubs.onlinelibrary.wiley.com/doi/pdf/10.1029/135GM12`. URL: `https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1029/135GM12`.

Matsumoto, H., M. E. Dickson, and P. S. Kench (2016). "An exploratory numerical model of rocky shore profile evolution". In: *Geomorphology* 268, pp. 98–109. DOI: `https://doi.org/10.1016/j.geomorph.2016.05.017`. URL: `https://www.sciencedirect.com/science/article/pii/S0169555X16303385`.