Dear Reviewer,

We provide hereunder an overview of how we address the comments, with explanations and references, on a point-by-point basis. Reviewers' comments are in black and our response in blue. In the manuscript, all the new/changed text is highlighted in yellow.

Reviewer 1 Comments to Author:

De Gelder et al., present a novel inversion approach to estimate marine terrace formation parameters and sea level histories from marine terrace topographic data. They use a set of test scenarios to demonstrate the feasibility of the approach and apply it to well-studied marine terrace sequences in Santa Cruz and the Gulf of Corinth. The topic and scope of the study are very promising. Using inverse schemes to study marine terrace sequences is a logical next step for the community and will be of interest to many people. This study therefore holds a lot of promise and comes with nice figures, however, the authors omitted a lot of critical detail that is necessary when describing a new method.

Thank you, we appreciate the comment. For more details on the method, see comments below.

Unfortunately, the authors uploaded a preprint without line numbers, therefore, I added some more descriptive text locations.

Our apologies, we forgot to add these.

The authors measure the horizontal misfit with topography and therefore focus on the width of terraces. This is interesting because terrace width is probably an underused terrace parameter that holds information. At the same time, terrace width (and topography I general) is often highly variable along-strike. It would be interesting to include a real case, where uplift rate is roughly constant, and several topographic profiles of the same sequence are jointly inverted. This would show whether this approach is robust enough to cope with along-strike topographic variability.

Actually the horizontal misfit is measured at every vertical step (istep), and so the misfit takes into account both the horizontal and vertical variability of the terrace sequence. We gave more details about the misfit calculation (Line 146-158):

"In this work, the data vector is defined as a set of points measured on the shoreline with a vertical step size (ipstep; Fig. 2a). The misfit between this observed topography and the modelled paleo-shoreline sequences is calculated on the horizontal axis. The elevation axis is therefore divided in a regular grid, and the level of data fit is measured by comparing the horizontal distance between observed and simulated horizontal position at each grid point. In this way, the terrace width is used as an observed parameter. Uncertainties about the observed shorelines account for the inability of our numerical model to explain observations, and can be due to both observational or modelling errors. These errors are supposed Gaussian and described by a standard deviation (σ; Fig. 2a) and the level of spatial correlation related to the spatial resolution of the data (corrl; Fig 2a). In preliminary tests (Fig. S2) we also tested inversions with the misfit calculated along vertical axes (i.e. comparing elevations between observed and estimated profiles), but found it harder to reproduce realistic marine terrace sequences with a few m of terrace height variability, and a few hundred meter of terrace width variability (e.g. Regard et al., 2017; De Gelder et al., 2020). We note that changing the misfit calculation from horizontal to vertical misfit does not change the paleo sea-level posterior distribution (Fig. S2)."

To go further into horizontal and vertical uncertainties, we added 1) the horizontal and vertical ranges for the 2.5-97.5 percentiles visually in Fig. S4, as well as the average horizontal and vertical ranges for the different parameter variations, 2) an extra figure about the Santa Cruz terraces, showing how the chosen ranges approximately match the horizontal and vertical uncertainties in a real case, and 3) additional discussion about horizontal and vertical uncertainties in the text, also referring to this new supplementary figure (lines 440-449):

"To apply this model procedure to other sites, we can make several recommendations based on our findings in this paper. Any coastal area with marine terraces will have lateral variability

in terms of terrace width and terrace elevation. Given that the posterior ranges for the model parameters (paleo sea-level, uplift rate, etc.) will directly depend on the chosen inversion parameters (ipstep, σ and corrl), it is important to adjust those inversion parameters to the naturally observed variability in the geometry of a marine terrace sequence. In the Santa Cruz example the similarity in terrace width/height variability between the modeled posterior ranges (Fig. S4), and the observed variability within a section of the marine terrace sequence (Fig. S5) justifies our choice of inversion parameters. But, if we wanted to characterize a larger area along the Santa Cruz coastline, we would either have to increase σ and/or corrl to account for the higher degree of variability, or jointly invert multiple cross-sections in which every cross-section has a similar variability in terrace height/width."

The authors should include the equation of the forward model to make the methods easier to follow.

We added some text and the main equation to make the details of the forward model more clear (Lines 109-116):

"The initial erosion rate ER is expressed as an effective eroded volume per unit of time and coastal length, in which a fraction of erosional residual power ERr erodes the foundation at each location along the profile. This fraction ERr depends on the water level, h, the water depth for wave base erosion WB, and a coefficient for sea bed erodibility K, so that:

Then a residual power (ER - dERr /dt) carves out a 1 m high notch and all its overhanging material to form a cliff. Following previous studies (Anderson et al., 1999; Pastier et al., 2019), for K we use 0.1 as bedrock erodibility and 1 for notch carving."

A supplementary table summarizing all the paleo-sea level ranges, including a justification, should be added to the supplement.

Good suggestion, we added such a supplementary table (Table S1), and in the supplementary text we added the sentences "For MIS 10b and 10c we used the minimum from Spratt & Lisiecki (2016) and the maximum from the MIS 10a box." (which we had previously forgotten to mention) and "We summarize the ranges and justifications in Table S1 on the following page."

The methods require significantly more detail. Currently, the authors do not provide an adequate description of the algorithm. For instance, there is no mentioning of the type of sampler used in the MCMC. How many individual random walks are performed per scenario? Currently, only the number of forward simulations is mentioned. What acceptance ratio was aimed for? How was burn-in defined? What about autocorrelation? Effective sample size? ETC. This paper describes an exciting new inverse algorithm for marine terraces, and as such is lacking a lot of critical information. Citing an inverse problem review paper is not sufficient. This lack of detail also makes it harder to assess the presented study. None of the standard diagnostic plots of an MCMC inversion are presented, and it is therefore, hard to assess the performance of the algorithm.

More details have been given to describe the Bayesian inversion procedure and the McMC algorithm (Lines 125-141):

"Reconstructing the sea-level history from present day observations of marine terrace sequences can be mathematically formulated as a highly non-linear inverse problem where the solution is non-unique. To embrace this non-uniqueness, the problem can be cast in a Bayesian (probabilistic) framework where the solution is a posterior probability distribution describing the probability of the model parameters (here the past sea-level variations), given the observed data (here the geometry of marine terraces). We use a Markov chain Monte Carlo algorithm to sample the posterior distribution and explore the range of models that can explain the observed topography within errors. In the Monte Carlo exploration of the model space, nodes can either be fixed at certain ages and elevations, or left free to move within a prescribed range (e.g. red boxes in Fig. 2b,d,e). The 4 main erosion model parameters (IS, ER, WB, UR) can also be fixed to chosen values, as done in the synthetic tests below, or left free within chosen ranges, as done for the Santa Cruz and Corinth examples below. The algorithm samples the parameter space as a random walk, where at each iteration a new sea level history model is proposed by perturbing the current one. The proposed model is then

either accepted or rejected following an acceptance rule based on the level of data fit of the current and proposed models. The final solution is a large ensemble of paleo sea-level models that approximates the probabilistic solution. That is, the distribution of models follows the posterior probability solution. Of the 1 million model runs per figure, the first 50 000 models were discarded as burn-in models. To verify that the random walk samples the target distribution, we show a number of convergence diagnostics in Fig. S1, which includes parameter acceptance ratios and likelihood evolution for all model runs in this article."

In order to show that the MCMC sampler converged towards the posterior distribution, we added in supplements some figures showing standard convergence diagnostics: Supplementary Figure (Fig. S1) was added, with plots of parameter acceptance ratios and log likelihood (misfit) values in relation to the model runs for all the main figures of the paper.

Priors: The authors describe the model parameters, but do not specify how these are treated in the inversion. The authors refer to a prescribed range, and therefore, I assume, use flat priors. But this should be specified. Also, it would help to stick to common terms in Bayesian frameworks, such as priors.

Yes, we treat them as flat priors and have specified as such now (Lines 158-160):

"The assigned ranges for input parameters are all treated as flat uniform priors, i.e. a prior distribution that assigns equal probability to all the values within the specified ranges."

The authors measure the difference between observed and modelled topography in the horizontal axis, putting much emphasis on the width of terraces. How do the presented results change for measuring the vertical misfit? Commonly, terrace studies focus on terrace elevations and not their width.

Although the misfit is calculated on the horizontal axis, this does not imply there is no importance attributed to terrace elevations. We added Figure S2, and also specified in the text (Lines 154-158):

"In preliminary tests (Fig. S2) we also tested inversions with the misfit calculated along vertical axes (i.e. comparing elevations between observed and estimated profiles), but found it harder to reproduce realistic marine terrace sequences with a few m of terrace height variability, and a few hundred meter of terrace width variability (e.g. Regard et al., 2017; De Gelder et al., 2020). We note that changing the misfit calculation from horizontal to vertical misfit does not change the paleo sea-level posterior distribution (Fig. S2)."

Why are so many model parameters fixed in the test scenarios? Would be interesting to explore additional scenarios, to see which information is required for a given terrace sequence and when the inversion is not able to recover the parameters. Another important question is, how the inversion performs when the problem gets under-constrained. Does the algorithm converge on a (fake-) solution in a local minimum or does the MCMC roam the parameter space as it should.

We fixed most of the model parameters in the test scenarios to keep things simple initially, and increase the complexity later in the paper. Nevertheless, it is an interesting suggestion, so we added a similar test to the one in Figure 2, but with larger ranges on the priors for initial slope, uplift rate, erosion rate and wave base height (see Figure S3). This allowed us to verify that the inversion performs well and is able to sample a wide posterior distribution without getting trapped in local minima. We added some lines on this figure in the main text as well (Lines 212-217):

"To understand what happens in scenarios in which more parameters are unknown, we repeated the same tests (from Figures 2b-d) with broad prior ranges for the uplift rate, initial slope, erosion rate and wave base height (Fig. S3). Also in this case the joint inversion provides much narrower posterior ranges for paleo sea-level compared to the two individual profile inversions, and not too different from the cases with fixed model parameters (Figs. 2b-d). The posterior ranges for all parameters are consistently smaller for the joint inversions, compared to the individual inversions (Fig. S3)."

The authors do not define how they report inversion results. E.g., for the Santa Cruz inversion, they state the posterior range of parameters but do not explain, whether these are confidence intervals, a standard deviation around the mean, or similar. Also, the authors state that the model limits the uplift rate to 1.35-1.65 mm/yr, but this is the prior range and was therefore set beforehand. The word “limits” could also imply that the authors are in fact referring to the prior range. However, posterior range results are following, such as the range for initial slope. I know I am being nitpicky with this sentence, but here and elsewhere, imprecise language concerning aspects of the inversion creates confusion.

We appreciate the comment, and adjusted the text at several points for clarification (Lines 164, 237, 257, 258, 261, 262, 270, 277-279, 306, 307, 318, 326, 341, 342). In addition, we clearly differentiate the prior and posterior ranges now in Figures 3 and 5.

There seems to be a degree of circularity in the approach. Narrow uplift rate prior ranges are defined for the Corinth and Santa Cruz models, based on the elevation of dated marine terraces. These prior uplift rates ranges are then used to reproduce the stair-case morphology and invert for uplift rate, which was already an implicit input. Is this OK because the focus of the study is on the width of terraces and resolving parameters other than uplift? The authors should address this.

We added a small paragraph on the uplift rate in the discussion now (Lines 457-464):

“In the cases of Corinth and Santa Cruz we assigned relatively narrow windows for the uplift rate, based on chronological information that was already available. As such we only obtained refinements, rather than ‘new’ information about the uplift rate. For the Santa Cruz case, we tested four uplift rate scenarios matching different possible chrono-stratigraphies (Fig. S6) that could all explain the morphology equally well, albeit with different ranges for the other parameters. We also inverted the Santa Cruz terrace sequence morphology with a broad range between 0.3 and 1.5 mm/yr, but the inversion would converge on only one of the four scenarios in Fig. S6. This suggests that at least an approximate prior idea on the uplift rate is a prerequisite for a reliable inversion of a marine terrace sequence morphology.”

We agree that concerning the uplift rates, the inversion mainly re-fines previous estimates, which we feel is appropriate for the current study, given that it focuses on sea level variations. Of course, alternative explorations on uplift rates -or even time varying uplift rates- could have been carried out, but that was beyond our scope here.

For the Corinth case, the posterior distributions of wave base depth in profile 2 & 3 have their maximum values at the boundary of the prior range, suggesting the algorithm would like to go to even deeper wave base depths. The authors, do not mention the results for the IS, ER, WB, UR parameters. These results should also be described and the implications of the posterior distribution ramping up against the prior boundary should be discussed, since this may be a problem.

We now expanded the description of IS, ER and WB parameters in Section 5 now (Lines 321-326):

“The three profiles show variations in initial slopes that are in line with the overall morphology, i.e. present-day profile 1 is steeper than 2, which is steeper than 3, which is also what we find for their initial slopes. The three profiles do have similar posterior distributions for wave base depths and erosion rates (Fig. 5e). Although we might expect lateral differences in these rates given variability in sediment types, catchment area and coastal orientation, the broad ranges for the posterior distributions indicate we cannot quantify these lateral differences from the profile morphology alone. “

… and also added more discussion about the posterior distributions ramping up against the limits we set (Lines 450-456):

“The prior information required to obtain reliable results, will also be dependent on the context of a specific marine terrace sequence, and on the parameter cut-off choices deemed realistic. For Santa Cruz and Corinth, the posterior distributions for wave base depth suggest that, purely based on marine terrace sequence morphology, this cut-off could have been deeper than 10-12 m (Figs. 3 and 5). However, based on observations and models of cliff erosion it

seems unlikely that the wave base for bedrock erosion is more than 10 m for Santa Cruz (Kline et al., 2014), and wave base depths/heights should be smaller for the calmer Gulf of Corinth. This choice in cut-off in turn affects the posterior distribution of other parameters, and should thus be chosen carefully."

Currently, the discussion ends with a lengthy paragraph of the Gulf of Corinth case. As a reader, this is a bit weird. Until here, the focus of this paper was the inversion method. However, the long Corinth section hangs at the end like an afternote. To improve flow and readability, I'd suggest to condense this section.

We feel like the implications of the Corinth inversion are very interesting, and merit a deeper discussion beyond the methodological advances that this paper presents. We think some readers will appreciate such a discussion, and have thus decided to keep this section as detailed as it is. We did change the order of sections 6.1 and 6.2, as to end the discussion in a broader context.

To emphasize more that we find the discussion of Corinth a key part of the paper, we modified the title, which is now;

"Bayesian reconstruction of sea-level and hydroclimates from coastal landform inversion: application to Santa Cruz (US) and Gulf of Corinth"

and added several sentences throughout the paper to help that emphasis (Lines 76-81, 225-229, 293-296, 346-349, 355-356, 503-506).

The authors present an exciting new tool and there are many things that could be discussed, but currently are not. What about typical lateral variability of coastline morphology and its influence on inversion results? What prior knowledge is typically needed to recover reliable results? What parameter trade-offs typically exist? Etc.

We've added some discussion about lateral variability (Lines 440-449):

"To apply this model procedure to other sites, we can make several recommendations based on our findings in this paper. Any coastal area with marine terraces will have lateral variability in terms of terrace width and terrace elevation. Given that the posterior ranges for the model parameters (paleo sea-level, uplift rate, etc.) will directly depend on the chosen inversion parameters (ipstep, σ and corrl), it is important to adjust those inversion parameters to the naturally observed variability in the geometry of a marine terrace sequence. In the Santa Cruz example the similarity in terrace width/height variability between the modeled posterior ranges (Fig. S4), and the observed variability within a section of the marine terrace sequence (Fig. S5) justifies our choice of inversion parameters. But, if we wanted to characterize a larger area along the Santa Cruz coastline, we would either have to increase σ and/or corrl to account for the higher degree of variability, or jointly invert multiple cross-sections in which every cross-section has a similar variability in terrace height/width."

… and prior knowledge (Lines 450-456):

"The prior information required to obtain reliable results, will also be dependent on the context of a specific marine terrace sequence, and on the parameter cut-off choices deemed realistic. For Santa Cruz and Corinth, the posterior distributions for wave base depth suggest that, purely based on marine terrace sequence morphology, this cut-off could have been deeper than 10-12 m (Figs. 3 and 5). However, based on observations and models of cliff erosion it seems unlikely that the wave base for bedrock erosion is more than 10 m for Santa Cruz (Kline et al., 2014), and wave base depths/heights should be smaller for the calmer Gulf of Corinth. This choice in cut-off in turn affects the posterior distribution of other parameters, and should thus be chosen carefully."

We already mentioned trade-offs between timing and elevation of sea-level peaks (Lines 197-200, 251-253), and correlation between erosion rate and wave base depth (Lines 254-256) and did not feel we needed to go further into parameter trade-offs; in a way, this study can be regarded as a proof of concept, also designed to release the code and develop technical aspects, before it can be used for broader parametric studies and regional applications, by anyone interested.

Section 2, second paragraph: There seems to be a typo in the reference (REEF).

It was not a typo, but the name of the code that we used. We clarified that now (Line 104):

"The landscape evolution model we use (the code 'REEF'; Husson et al., 2018; Pastier et al., 2019) "

We very much appreciate the time and effort by the reviewer to go through our manuscript, as well as the good suggestions made to improve the manuscript.

We hope the applied changes will be appreciated by the reviewer,

Kind regards,

Gino De Gelder, Navid Hedjazian, Laurent Husson, Thomas Bodin, Yannick Boucharat, Kevin Pedoja, Tubagus Solihuddin and Sri Yudawati Cahyarini