The submitted manuscript by Peter van der Beek and Taylor Schildgen provides more detail on code that was published in a paper in 2018 (Schildgen et al., Nature, 2018). The method is already published and does not have any major flaws. Furthermore, it is based on earlier work by Mark Brandon (1997). Therefore, the paper could be published as is. However, I would like to highlight some issues with this approach in general and think that mentioning these ideas would strengthen the paper.

The authors use the unperturbed initial geothermal gradient as input for calculating geothermal gradients and the modern geothermal gradient is not really discussed. Furthermore, there is no way to link samples in space. One of the advantages of codes designed to exploit age-elevation relation relationships (Pecube, QTQt, GLIDE) is that the results should be less sensitive to geothermal gradient and there is a means to estimate the geothermal gradient. Basing the interpretation on an unperturbed geotherm does not provide that advantage and I feel that we are losing useful information.

It is also unclear how well known this unperturbed geotherm can ever be. Or even whether this makes sense. For example, if a sample has a known ZFT age, a geotherm that is consistent with an initial value and exhumation rate can be determined using the code. If this same sample also has a AHe age, why would it make sense to use the initial geothermal gradient before any exhumation as opposed to the geothermal gradient inferred the older ZFT age? This geothermal gradient is surely an improvement over the unperturbed gradient.

This concept raises the issue that a single sample may predict different exhumation rates for the modern period but also completely different geothermal gradients. This means that
if we are lucky enough to have an estimate of the modern geothermal gradient close by, we would have lots of internally inconsistent pieces of evidence on the exhumation rate. We would also have no way forward with this steady state approach.

The paper is very critical of GLIDE (Fox et al., Esurf, 2014) and does not really highlight our response to the earlier criticism of Schildgen et al., (Nature, 2018). For example, on line 125 please consider changing the language from “Moreover, it has been shown that the code translates abrupt spatial variations in thermochronological ages, such as across faults, into temporal increases in exhumation rates (Schildgen et al., 2018).” To “Moreover, it has been argued that the code translates abrupt spatial variations in thermochronological ages, such as across faults, into temporal increases in exhumation rates (Schildgen et al., 2018).” This is because, we have argued that this is not the case in Willett et al., (Esurf, 2020). Also, Fox and Carter (Geosciences, 2020) argued that Schildgen et al. (2018) did not show that most sampled regions on Earth may have insufficient data coverage for unbiased prediction of exhumation-rate histories using GLIDE. Instead, the synthetic data produced for analyses carried out by Schildgen et al., (2018) significantly changed the temporal and spatial resolution of the data with respect to the real data. In other words, the spatial and temporal resolution for the real ages from the Western Alps is sufficient because there are old and young ages either side of the fault. In contrast, the spatial and temporal resolution for synthetic ages from the Western Alps is insufficient because of the dramatically different age distributions. This is a complicated issue and it is not appropriate to simplify it to such a degree here. Finally, on line 129, the authors argue that GLIDE is as slow as Pecube. However this is almost certainly not the case if the areas are the same size and if Pecube is run in inverse mode. In fact, this is one of the reasons we developed GLIDE. GLIDE does not use the same half space solution as the Willett and Brandon code. It uses a numerical model with a flux boundary condition. Changing this to a fixed boundary condition requires uncommenting one or two lines.

It is unclear how \( \delta h \) is actually calculated. Where does \( \delta h \) appear in the flow chart in Figure 1B? This is crucial because the closure depth can change by over 100% during the iterative process. How is this accounted for during the whole process? The code requires that users extract \( \delta h \) from a different piece of software before the code can be used. It is unclear how long this takes so it is unfair to argue that the approach presented here is particularly fast when it is unclear how long it takes to run the other software.

I don’t understand why the upper boundary condition of the thermal model has a temperature evaluated at the sample elevation. Surely the thermal model should have an upper temperature boundary condition equated at \( h_{ave} \). Much more information is required to actually understand this part of the code. For example, if you have two samples from exactly the same geographic position but separated vertically by 2km, what is the thickness of the thermal model? If the surface temperature is used for the two different thermal models, the geothermal gradients will be different for the whole thickness of the crust. Why would the closure elevation (with respect to the centre of the earth) depend on the sample elevation? Given that most data are collected with the aim of producing age-elevation relationships, this needs to be discussed and considered. In addition, it is not clear why a constant thermal model thickness is used for the whole of the Himalayas. Surely if \( L \) represents a crustal thickness, the model should get much thicker as the topographic elevations increase to account for isostasy.
There is a lot of discussion on thermal boundary conditions. I don’t think detrital thermochronology tells us anything about how appropriate a fixed temperature at 30 km depth is over 10s of millions of years. Since the earliest studies in the Alps by Wagner, it has been shown that the focus of exhumation has been variable. Vernon et al., (EPSL, 2008) also showed that it was variable using iso-age surfaces. Even the results presented in Figure 4 show that exhumation rates have changed at specific locations. A constant lag time is more likely telling us that there is an area that is eroding at a common rate through time, but that that area might be variable. This seems reasonable given that the maximum erosion rates likely depend on long term tectonics and regional climate.

The thermal half space solution was also used by Moore and England 2001. They elegantly showed that by accounting for transient advection of heat, data that had been interpreted as recording accelerating erosion rates, were actually consistent with a constant exhumation rate. This study is not discussed at all but this result highlights the need to account for transient advection of heat. Without accounting for this, there will be the possibility of a systematic bias towards inferring accelerating exhumation rates as geotherms become higher during exhumation. This does not only happen when using a half space solution or a flux boundary condition, it is also predicted with a fixed boundary condition. We also discussed this in Fox and Carter (Geosciences, 2020).

Figure 4. I find the presentation of the results very hard to read. It is almost impossible to visualize whether rates are increasing or decreasing. It also requires people to be familiar with the closure temperatures of the different systems. The fact that a single sample can be associated with 5 different modern exhumation rates is clearly a problem and really defeats the purpose of collecting ages using different systems. Please make a map of the final geothermal gradients as well. How variable are they at a single location and is this a useful result?

Line 51: Please provide more details about what the original age2edot actually did. It is unclear whether it actually solved for ages given exhumation rates.

Figures 5 and 6. Please also show the actual values of the variables that are changing. It would be nice to know that a value changes from X to Y as opposed to just the percentage change.