Referee #2

The impact of data noise on estimating two resilience metrics, variance and lag-1 autocorrelation, with satellite data was assessed in this study. The topic addressed is very importance, because satellite products are widely used to quantify the resilience of terrestrial ecosystems. My major concern is that it is within our expectation that data noise will affect the reliability of the metrics, what's the new finding of this study? I hope two aspects may be investigated in depth: 1. What's the uncertainty of the existing satellite products when used for quantifying resilience? For this purpose, the 'real noise' of the data needs to be quantified. 2. What's the uncertainty of using the products to depict the temporal changes in ecosystem resilience? For this purpose, the temporal changes in the noise are also need to be quantified.

Thank you for your comments and the time spent with the Manuscript. We will address your three main points individually.

1. What's the new finding of this study?

As you note, satellite products are widely used to quantify the stability and resilience of different natural systems. These estimates of stability changes are based on the assumption that changes to higher-order statistics (e.g. variance, autocorrelation) are due to changes in the system under observation, and not in the observation mechanism. However, this assumption is not always true for satellite data – many nominally continuous data sets are in fact made up of a constellation of sensors. You are clearly right that it is within our expectation that data noise will affect the reliability of resilience metrics. As satellite data is used more and more often in these contexts, however, we felt it was important to explore how exactly changes in measurement procedures could propagate into resilience estimates, and what strategies might be used to minimize this issue. This is why we set up a thorough investigation of the detailed effects of combining signals from different sensors, using synthetic time series constructed for a wide range of possible signal-to-noise ratios. We further compared our controlled synthetic experiment to three real-world data sets, showing that while individual time series might be reliable, averaging individual time series of resilience indicators over large regions will tend to enhance the effects of measurement changes and thus reduce the reliability of resilience estimates. We feel that this is an important contribution to best practices when analyzing resilience based on satellite data. We will clarify the above points in a revised manuscript.

2. What's the uncertainty of the existing satellite products when used for quantifying resilience? What's the uncertainty of using the products to depict the temporal changes in ecosystem resilience?

Answering this question is exactly the reason we conducted our investigation. We undertook our study using synthetic data because we cannot fully constrain real-world data – satellite data noise shifts drastically in space and time, and we have no 'perfect' measurement against which to compare as a ground truth. Satellite instruments have reported signal-to-noise ratios and are calibrated against known quantities (e.g., deep space, the Sahara); however, these do not account for all sources of noise in, for example, a vegetation measurement. Other factors – such as atmospheric water content, cloud cover, and satellite viewing angles – will also influence estimated surface parameters. Without a true reference, dis-entangling the time-variable (e.g., seasonal, annual) changes in noise from changes in system stability is not possible with real-world data. Hence, we focused here on a controlled synthetic system to explore the potential influences of changes in measurement procedures through time.

The lack of true reference also limits our ability to quantify the uncertainty in changes in resilience in real-world applications. We instead attempted to mimic a change in resilience proxies (autocorrelation and variance) using only changes in measurement noise; this could be thought of as a null model for whether or not there was a change in resilience. If changes in autocorrelation and variance are stronger than those implied by the changes in the underlying noise, than that could be interpreted as a more robust resilience change signal. Again, however, in a real-world system we would need a time-explicit estimate of how signal-to-noise ratios are changing to control for the influence of measurement procedure on our resilience estimates.

3. *the 'real noise' of the data needs to be quantified … the temporal changes in the noise also need to be quantified*

While we fully agree with your desire for time-explicit noise models for different satellite data, that is not the problem we set out to address with this publication. We instead approached the problem from a synthetic perspective, where we could control the amount and timing of noise, as well as how and when different sensors were mixed. Our work emphasizes the need for better constraints on the noise levels of satellite products and demonstrated the influence of varying signal-to-noise ratios on different ways to estimate resilience. It does not aim to provide a thorough accounting for different remote sensing products. We will revise our introduction to make this clearer.