Reply to the interactive comments of Anonymous Referee # 1 to our Discussion paper:

"Modelled potential forest area in the forest-steppe of central Mongolia is about three times of actual forest area"

Thank you very much for reviewing our work and your valuable comments that will help us to improve the manuscript. We will reply comment by comment.

Thanks for pointing out that the actual title of the manuscript might be interpreted in a misleading way. Instead of highlighting the modelling, it is more appropriate to point out to the parameters that were used as database for the delineating procedure. Thus, we propose to change the title to: "Geo-ecological parameters indicate explicit discrepancies between potential and actual forest area in the forest-steppe of central Mongolia"

We will increase the font size of all labels in the figures to a sufficient size.

Problems concerning the methods part:

1. We agree that an accuracy assessment for the forest delineation is required, when executing satellite-image classifications. Unfortunately, we did not describe the total forest mapping procedure detailed enough in the methods chapter. We will do this in a revised version of the manuscript (the procedure is mentioned in Figure 3, which is surely hard to read). A great portion of the definite forest area determination was done manually by visual interpretation of the satellite images when integrating the classification results of four different satellite scenes. In addition to the Sentinel 2 and Landsat 8 data, we used Spot 6 scenes with a spatial resolution of 1.6 x 1.6 m for the center part of the study area and high-resolution data from the base maps provided in ArcMap by ESRI.

Reference data for the production of a confusion matrix can be received by visual check of the satellite data to identify forest/no-forest areas for randomized points. However, we executed visual check during forest mapping. Thus, the mapping accuracy would reach up to 100%. Another way to check the mapping accuracy would be the comparison of the forest map with a forest classification of Spot 6 data. Intersection of both vector data provides information about area differences, mostly for mixed pixels caused by diverging spatial resolution of the datasets (30 x 30 m against 1.6 x 1.6 m), but not for the mapping accuracy within the satellite scene. Regarding the total mapping procedure, we propose to add more detailed information about forest mapping in the methods chapter without adding a confusion matrix.

2. The approach to generate potential forest biomass and forest area was based on the topographic and climatic parameters that are listed and explained in the methods chapter (lines 209-223). The way in which these parameters (which are ecological limitation factors, especially in the forest-steppe area) were used for delineation of potential forest area, is described in detail in the introduction chapter (lines 95-101). A main issue of our investigation was to check and proof the suitability of these parameters for the study area as presented in hypothesis III (lines 106-108).

3a. Spatial adjustments of field data to satellite data (e.g. NDVI) represent a general problem of remote sensing analysis. It would not be possible to match a measured area of 30 x 30 m exactly to one Landsat pixel in the field, because Mongolia (like most of the world) is a remote area with moderate GPS accuracy. Furthermore, the biomass plots (20 x 20 m) are bigger than the pixel size of Sentinel 2 (15 x 15 m), which also contributed to insufficient correlations between biomass and NDVI. The same is true for the Spot 6 data with its high spatial resolution where nearly every tree could be identified. In any case, the plot specific NDVI data needs to be interpolated or averaged from up to four pixels in the vicinity of the biomass plots. In the field, biomass plots were chosen that were most representative for a larger surrounding. If a slight spatial mismatch between field plots and satellite image pixel is a

relevant error source, then all research about NDVI and biomass done before would become truly obsolete. We can point out to this specific problem of potential error in a few sentences in the method chapter.

3b. It took 4 years (2014-2018) to collect the total biomass data. We presumed that the increase of tree biomass during this period was negligible and does not exceed the precision of the allometric method. The phenological status of the vegetation at the time of satellite passing / data collection for the NDVI seems more important than the temporal matching of field and satellite data. As an example for an extreme situation, biomass can be measured during winter, but NDVI is not representative at that time.

4. The forest distribution scenario was indeed insufficiently described in lines 256 to 258. We will change the sentence to:
"Thus, finally, we delineated the various forest types with their specific tree biomass (with respect to the AFA, 1986 (AFA plus BFA) and PFA) by multiplying the area of each forest type with the mean tree biomass of that forest type."

5. We will …
- ... increase the font size in the flowchart;
- … indicate the colors in Figure 2;
- … combine NDVI and LAI on the X axis;
- … omit some of the abbreviations like FDA, HMA, SDA. However, MAAT, MAP, MGST are common abbreviations for climatic parameters;
- … add information about the date of the Sentinel 2 images for NDVI average in Figure 3:
17th May 2018, 11th June 2018, 25th August 2018, 4th September 2018, 14th September 2018, 19th September 2017, 16th July 2016
- … change "site-representative" NDVI to "site-averaged" NDVI.