Review of Freudenberg, Schnell and Magdon "A Sentinel-2 Machine Learning Dataset for Tree Species Classification in Germany"

Freudenberg, Schnell and Magdon have carefully revised their manuscript, taking into account my comments and suggestions and those of the other reviewer. I still believe this dataset is unique in its size and spatial extent and therefore potentially very useful for scientists, forest managers and policy makers alike. I appreciate the addition of the separability analysis in the results section. Besides some minor changes in the wording of some sentences (see minor comments) I have no further objections to the publication of this data paper. The paper would benefit from a thorough copy editing and typesetting revision by the journal to improve clarity and correct language use.

Minor comments

L26-L29 "Machine learning, particularly deep learning […]" As I understand it, deep learning is a form of machine learning that uses many data layers and artificial neural networks in classification tasks, often applied in image recognition. It is also a buzzword. Please add references to this sentence to studies in which deep learning was used for tree species classification or similar. Good to shortly explain the difference between deep learning and machine learning here.

L77 "but due" indicates a decline in forest area from the 32% in 2012, but a decline in growing stock does not necessarily mean a decline in forest area, for example thinning. Please rephrase.

L109 "The growing space "approximately corresponds to the crown projection area"(Riedel et al., 2017, pp. 39, author's translation), so we use these terms interchangeably in the following" No this should not be used interchangeably because it is very confusing to the reader. "approximately corresponds" is not sufficient to use the terms interchangeably. The "growing space" here is not defined, neither are its units. Is it cubic meter? This should really be changed, be careful with the use of terms and their units.

L113-L114 still very unclearly written, not sure how trees were selected as visible. Please rephrase.

L196 "Obviously, broadleaf trees exhibit a much stronger seasonal pattern", not if they are evergreen broadleaf trees… please separate leaf shape (broadleaf or needleleaf) and phenology (evergreen and deciduous) more clearly. For example, holly (*Ilex*) is an evergreen broadleaf tree/shrub in Europe that does not likely show an obvious seasonal pattern in reflectance. On the other hand, Larch (*Larix*) is a common deciduous coniferous needleleaf tree. Suggested edit: "Obviously, deciduous broadleaf trees exhibit a much stronger seasonal pattern than the evergreen coniferous trees in our dataset."