Addendum to: A dataset mapping the potential biophysical effects of vegetation cover change

G. Duveiller, J. Hooker & A. Cescatti

The original dataset (Duveiller, Hooker, and Cescatti 2018) provides spatialised estimates of how a series of biophysical environmental variables (e.g. daytime land surface temperature, latent heat, etc.) would locally change following potential changes in vegetation cover. The data-driven methodology imposed a limitation on the spatial extent over which these estimates could be made, as it required both source (what is transitioned from) and target (what is transitioned to) vegetation cover types to co-exist within a local moving window. To provide a more widespread spatial coverage as required by some users, this second version of the dataset has been expanded to places where only one of the two vegetation classes involved in a given transition occurs. This is achieved using a Random Forests (Breiman 2001), a non-linear regression tree approach, to predict the change in biophysical variables based on local climate. The method employed is described in (Duveiller et al. 2020), but is here applied using the updated version ts4.04 of the CRU climate dataset (Harris et al. 2020) and an extra climatic predictor describing snow presence (as the cumulated precipitation during days with negative temperatures in degree Celsius) The other predictors are mean temperature, temperature range, cumulated precipitation and an aridity index.

Both the original and updated datasets are now hosted in the JRC Data Catalogue: http://data.europa.eu/89h/f97e4216-e81c-4dfb-b42d-d0c19070d029

More specifically, the new version (v2.0) of the dataset is available at: https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/ECOCLIM/Biophysical-effects-vgt-change/v2.0/ This new version supercedes the previous version (v1.0), which can still be found either at either: https://jeodpp.jrc.ec.europa.eu/ftp/jrc-opendata/ECOCLIM/Biophysical-effects-vgt-change/v1.0/; or at: https://doi.org/10.6084/m9.figshare.c.3829333

Code to necessary to transform v1.0 to v2.0 is available at: [add URL of zenodo repo handle]

The updated data have the same format as before, but now contains an extra variable with an *_ext* suffix, which is the spatially extended version. The original data (without the *_ext* suffix) are also available in each file. Note that over the same location both will differ as the extended version is in effect performing a smoothing operation due to the climate-based interpolation. Finally, the file sizes have been substantially reduced by using the internal compression

within the netCDF files.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45 (1): 5–32.

Duveiller, Gregory, Luca Caporaso, Raul Abad-Viñas, Lucia Perugini, Giacomo Grassi, Almut Arneth, and Alessandro Cescatti. 2020. "Local Biophysical Effects of Land Use and Land Cover Change: Towards an Assessment Tool for Policy Makers." *Land Use Policy* 91: 104382.

Duveiller, Gregory, Josh Hooker, and Alessandro Cescatti. 2018. "A Dataset Mapping the Potential Biophysical Effects of Vegetation Cover Change." *Scientific Data* 5: 180014.

Harris, Ian, Timothy J Osborn, Phil Jones, and David Lister. 2020. "Version 4 of the Cru Ts Monthly High-Resolution Gridded Multivariate Climate Dataset." *Scientific Data* 7 (1): 1–18.