

Supplementary Information for the manuscript

“Understanding deep learning in land use classification based on Sentinel-2 time series”

Manuel Campos-Taberner, Francisco Javier García-Haro, Beatriz Martínez, Emma Izquierdo-Verdiguier, Clement Atzberger, Gustau Camps-Valls, María Amparo Gilabert

Table S1 shows the overall accuracy (OA) of the evaluated classifiers. Algorithms using different number of BiLSMT layers reported the best results except in the case of 4-BiLSTM, in which the accuracy was reduced to 94.5% being outpaced by the RF algorithm.

Table S1. Overall Accuracy (OA) obtained over the validation data set with different machine and deep learning algorithms.

Algorithm	Overall Accuracy (OA)
2-BiLSMT	98.7
1-BiLSMT	97.4
3-BiLSMT	95.6
RF	94.9
4-BiLSMT	94.5
SVM	92.7
NN	92.6
k-NN	88.6
DT	85.9

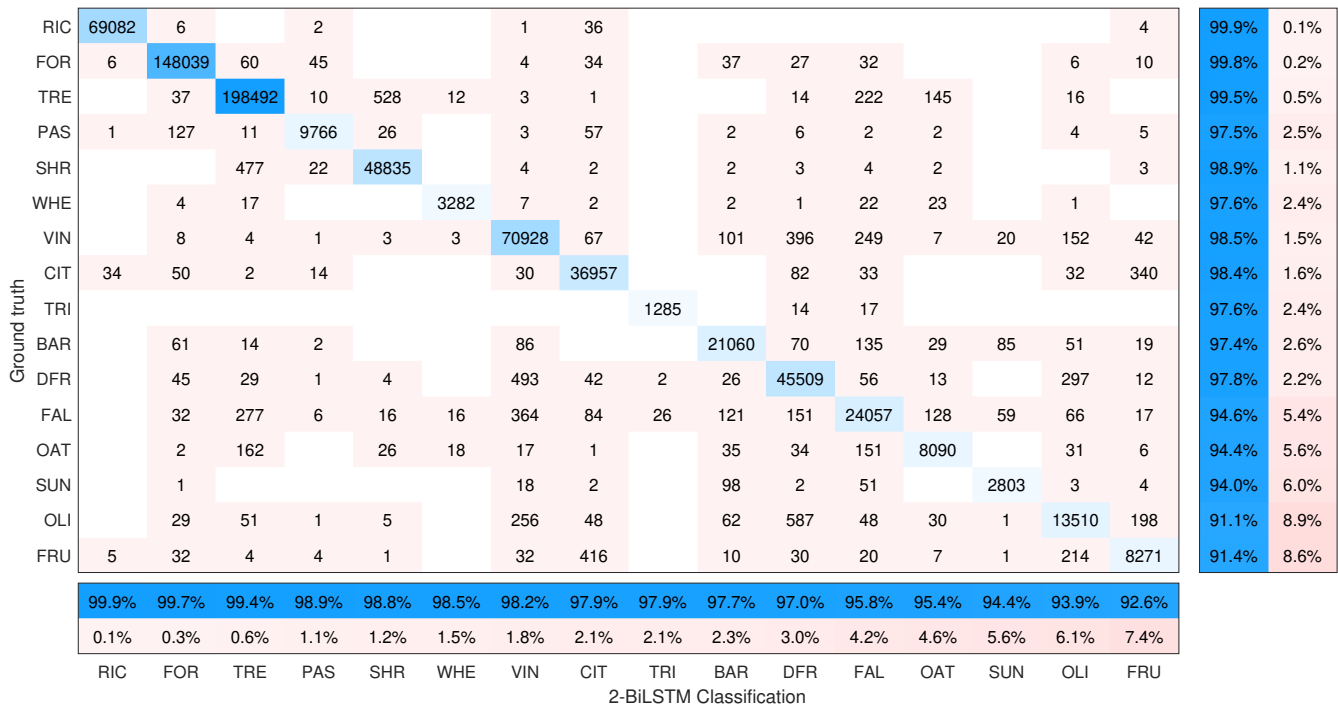


Figure S1. Confusion matrix obtained with the 2-BiLSTM. The matrix is sorted according to the true positive rate. The percentages of correctly classified observations (off matrix bottom rows) can be thought of as class-wise precisions, and the percentages of correctly classified observations (off matrix right columns) can be thought of as class-wise recalls.

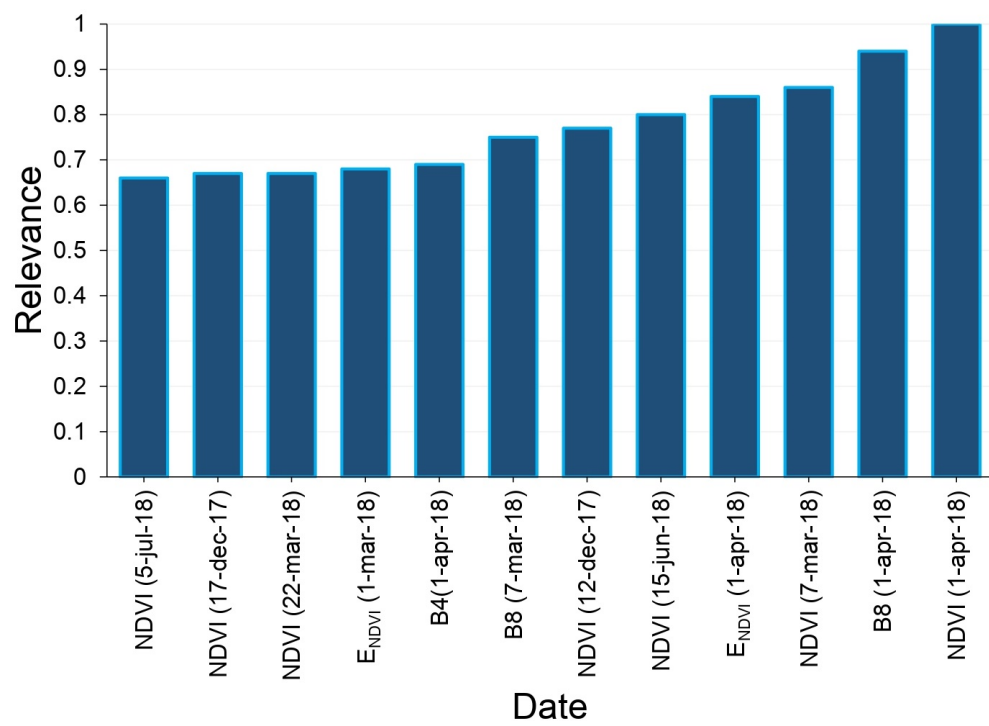


Figure S2. Relevance of the top ten predictors provided by the RF classification algorithm.