Classification of Tree Species by **Federated Learning**

Michal Staňo, Ladislav Hluchý, Peter Krammer, Štefan Dlugolinský,

Viet Tran and Marcel Kvassay

michal.stano@savba.sk, ladislav.hluchy@savba.sk, peter.krammer@savba.sk,
stefan.dlugolinsky@savba.sk, viet.tran@savba.sk, marcel.kvassay@savba.sk

Federated Learning

Peter Richtárik

Jakub Konečný

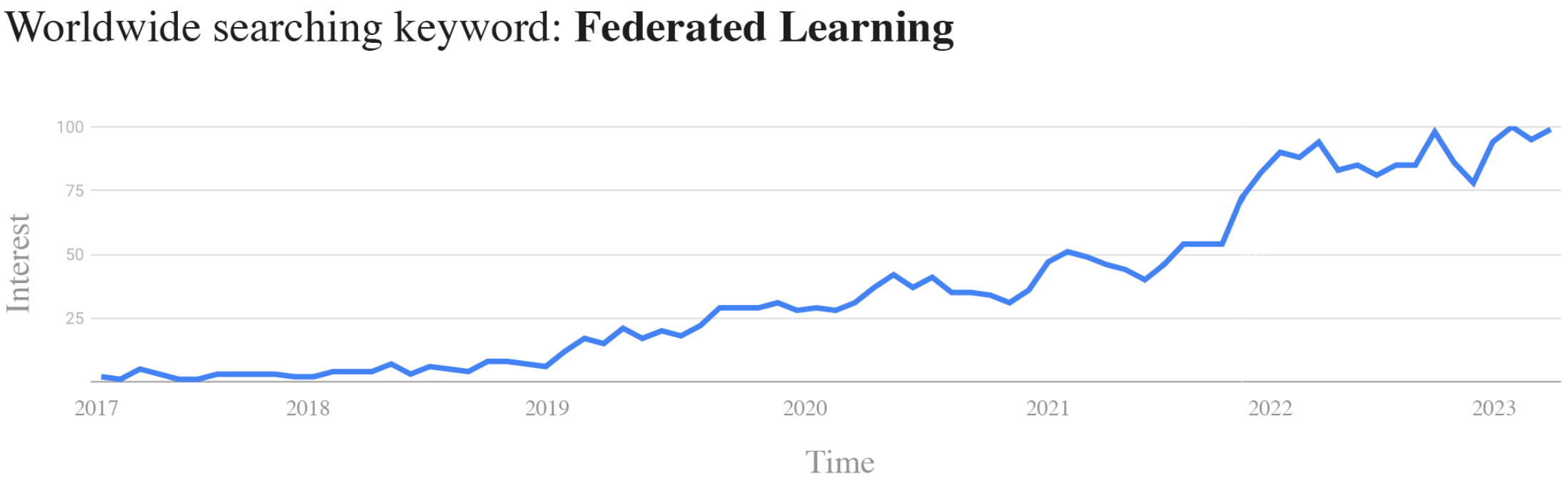
Google

H Brendan McMahan

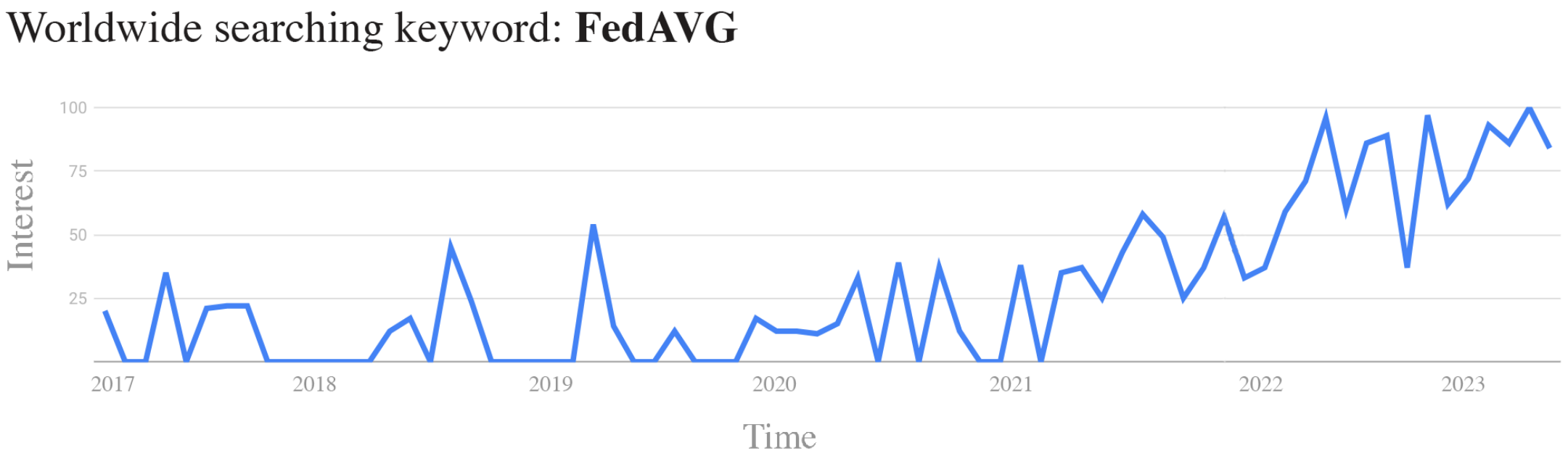
THE UNIVERSITY of EDINBURGH

Federated Learning was developed in 2016 in a collaboration between University of Edinburgh and Google

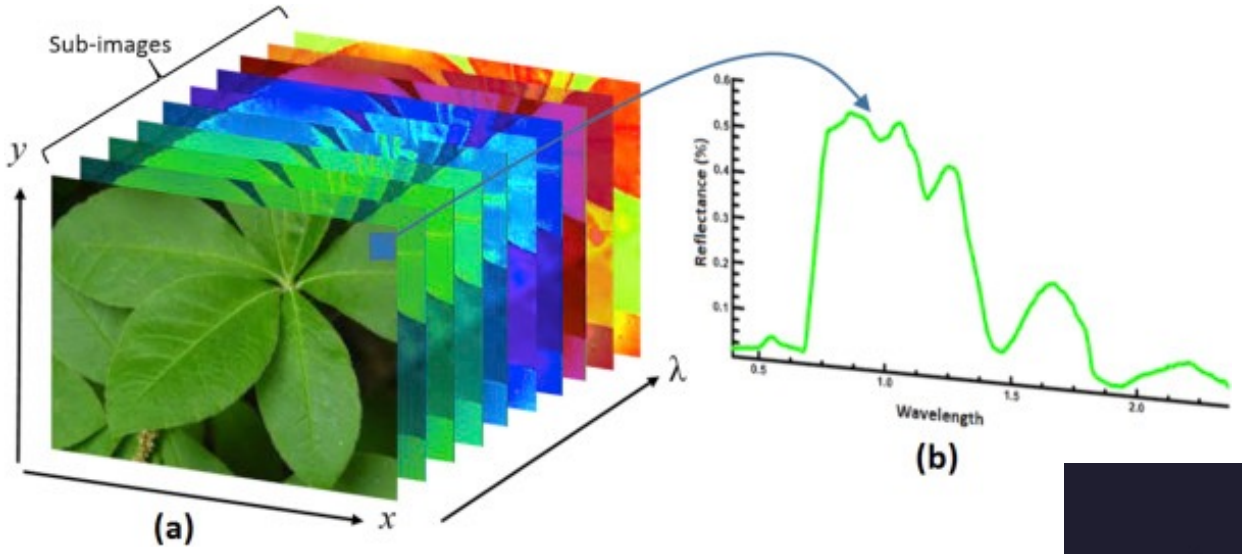
Federated Learning



Federated Learning Techniques

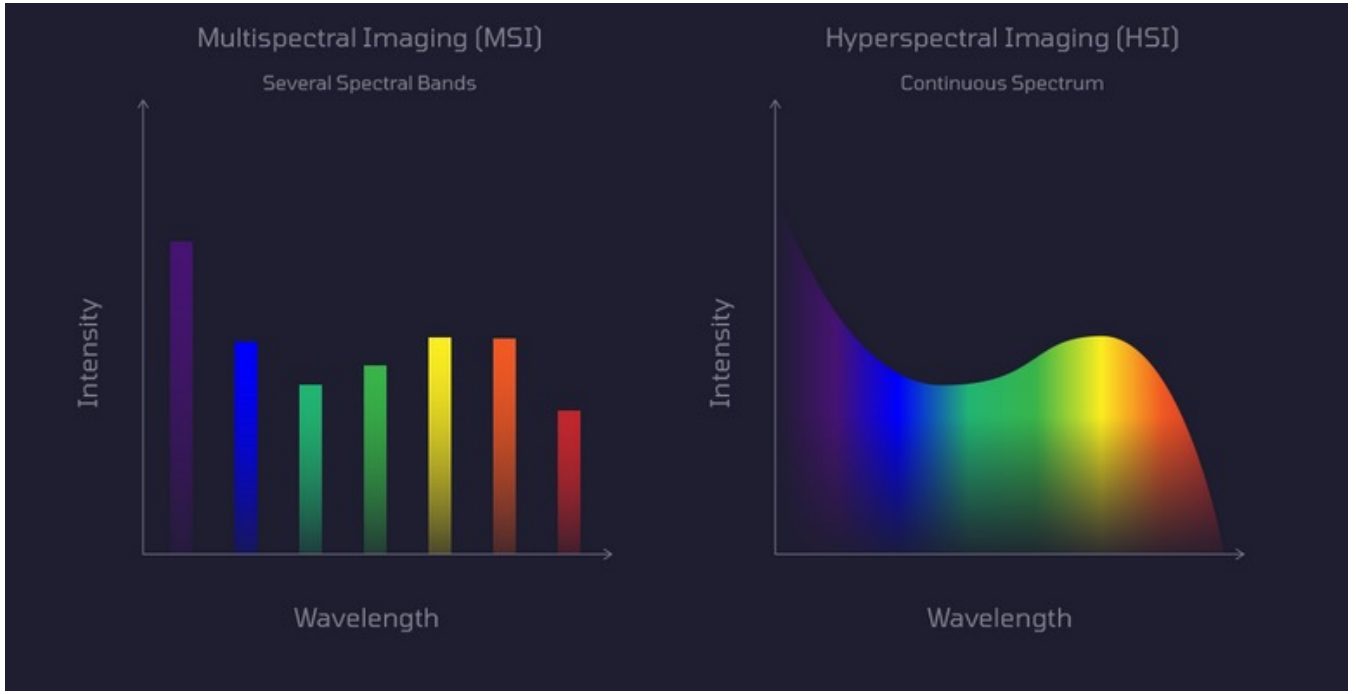


Hyperspectral Data & Analysis



- Multiple input frequencies (wavelengths) - 474
- For each available image which represents specific characteristic

- Lot of options for preprocessing and scaling / normalization
- Removing irrelevant frequencies which contain always same values
- Task : classification tree types



Dataset Properties

The currently available dataset consists of 33720 records (each record represents one image pixel) obtained from 3 independent helicopter flights and scans of a given site. The dataset contains 14 different types of trees, representing individual classification classes.

- 474 input attributes (frequency bands)
- 186 relevant attributes carrying information

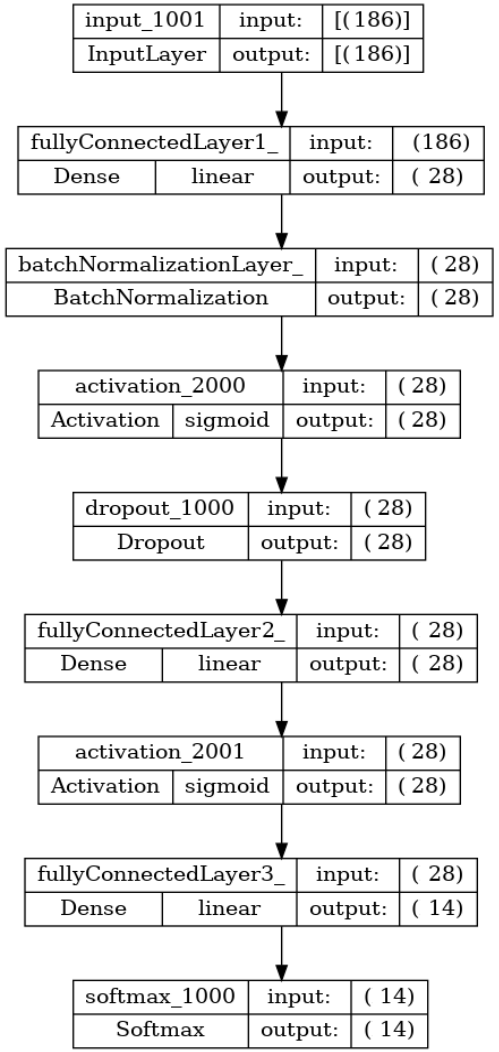
TABLE I. DISTRIBUTION OF TREE TYPES IN THE DATASET.

Type of tree	Count	Type of tree	Count
Pine Tree	24	Rowan	577
Birch	655	Limba	491
Cedar	2982	Aspen Tree	5608
Quince	2624	Spruce	16531
White Oak	112	Willow	322
Hornbeam	35	Hazel	200
Sycamore Maple	49	Chokecherry	138

Structure of Neural Network

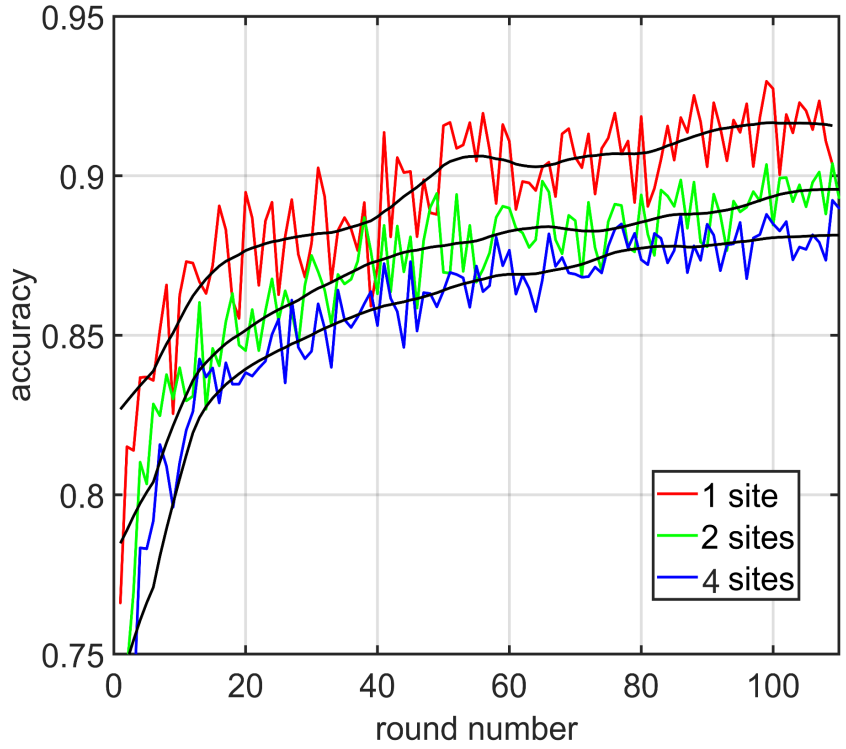
A deep neural network with a very simple and universal structure was used to classify the trees species. It consisted of the following layers: FeatureInputLayer, BatchNormalizationLayer, fullyConnectedLayer, sigmoidActivationLayer, DropOutLayer, SoftMaxLayer and ClassificationLayer or their Python alternatives.

The number of neurons in most layers - 28 was determined as twice the number of classes for classification. The probability of signal disabling in the DropOut layer was set to 0.15. The InitialLearnRate parameter was set to 5.2e-5 and ADAM was used as the optimization algorithm. The BatchSize parameter was set to 32.

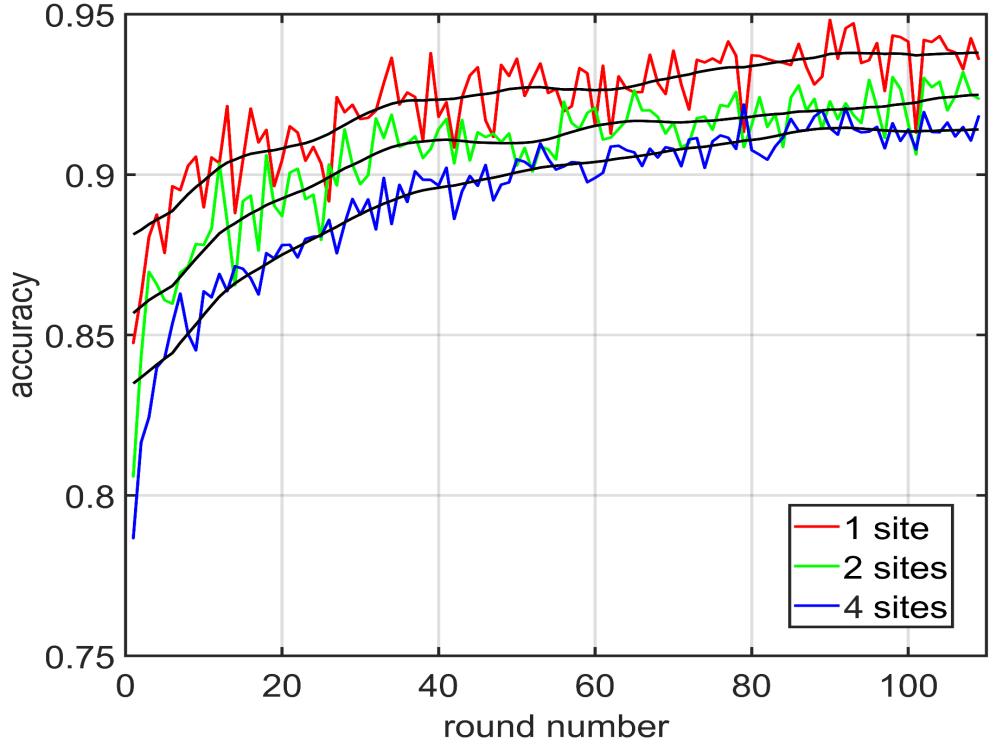


Graphs of training proces

Model accuracy in training process evaluated on independent testing data for 1 site (red), 2 sites (green), 4 sites (blue).



Synchronization with the server is realized once per 2 epochs



Synchronization with the server is realized once per 10 epochs

With a limited - predetermined total volume of available data, the accuracy of the model decreases as the number of sites increases.

Standard Deviation comparison

TABLE II. STANDARD DEVIATION VALUES OF DIFFERENCE BETWEEN THE ACTUAL ACCURACY AND THE DOUBLY FILTERED ACCURACY FOR FEDERATED LEARNING WITH SERVER SYNCHRONISATION ONCE PER 2 EPOCHS.

sites num.	**Rounds**			
	1 - 27	***28 - 54***	***55 - 81***	***82 - 108***
1 site	0.017409	0.012499	0.009163	0.008030
2 sites	0.029608	0.009858	0.008094	0.006968
4 sites	0.043599	0.006955	0.005824	0.005455

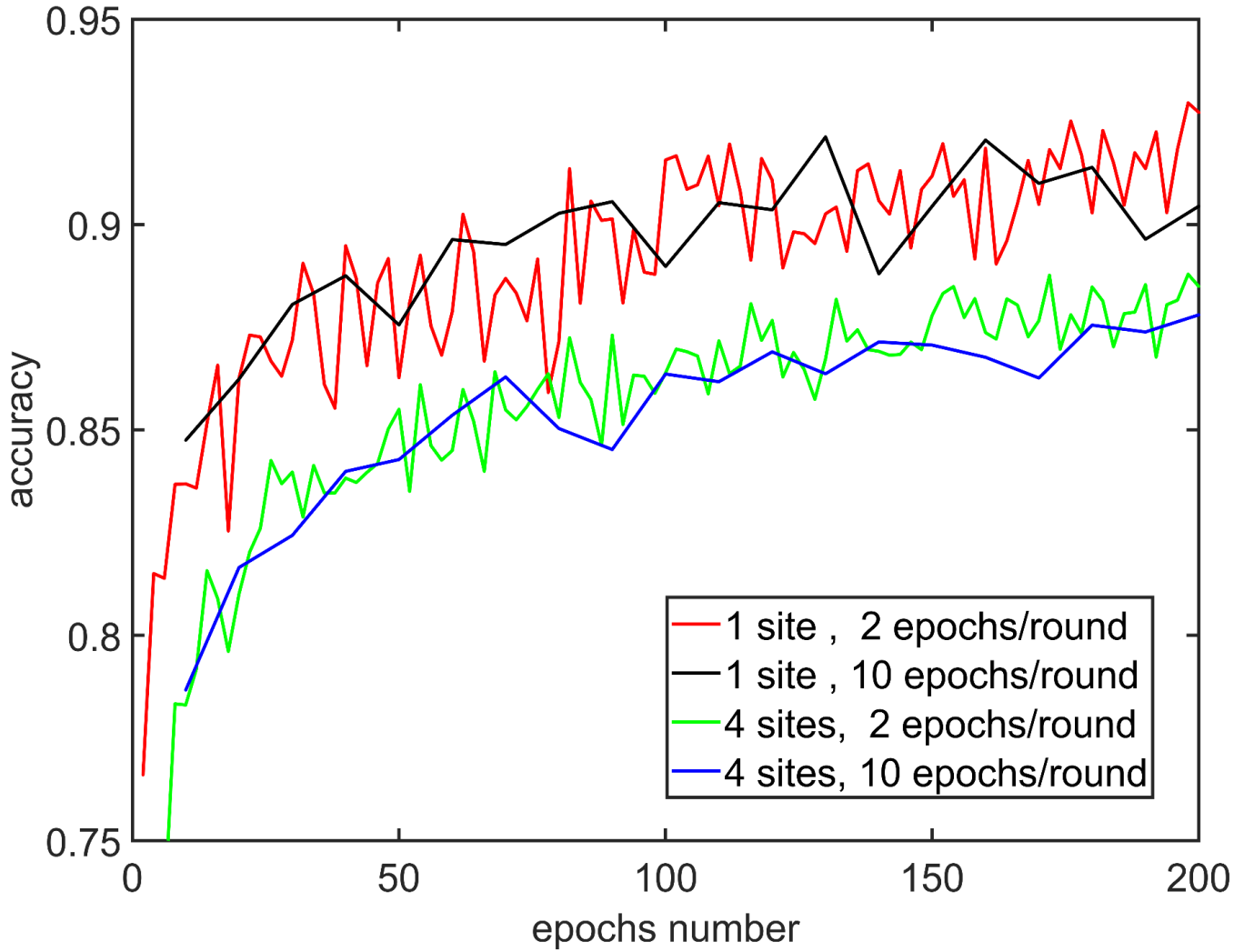
TABLE III. STANDARD DEVIATION VALUES OF DIFFERENCE BETWEEN THE ACTUAL ACCURACY AND THE DOUBLY FILTERED ACCURACY FOR FEDERATED LEARNING WITH SERVER SYNCHRONISATION ONCE PER 10 EPOCHS.

sites num.	**Rounds**			
	1 - 27	***28 - 54***	***55 - 81***	***82 - 108***
1 site	0.011845	0.007610	0.007227	0.007144
2 sites	0.013522	0.005742	0.005263	0.006283
4 sites	0.011713	0.004814	0.003910	0.003410

An interesting aspect is that as the number of sites increases, the accuracy of the model decreases, but the standard deviation also decreases.

Training graph with depends on epochs

- For limited volume of data, number of sites is very important parameter
- The number of epochs per round is not so key
- Solid matching for red vs black curve, and blue vs green curve
- Number of epochs is more representative than rounds number



Model accuracy plotted against the number of epochs for both server synchronization settings

Conclusions

- With a limited amount of data, the claim that increasing the number of workers used reduces the accuracy of the model was confirmed.
- With the volume of data, which increases along with the number of workers, the accuracy of the model should naturally increase.
- The decisive factor is the influence of the number of epochs during training (the number of rounds is not so decisive).
- We managed to verify the stability as well as the convergence of the learning process using federated learning.
- Creation of a prototype of a classification model, intended for tree species recognition, using hyperspectral data.
- Achieving a relatively high accuracy of classification, given the relatively high number of recognized tree species.

Thank you for your attention