

# PREDICTING PEAK ABOVE-GROUND BIOMASS IN (SEMI-)NATURAL GRASSLANDS WITH SENTINEL 2A

Stien Heremans<sup>1,\*</sup>, Sam Medart<sup>2</sup>, Laura Vanierschot<sup>2</sup>, Ben Somers<sup>2</sup>

<sup>1</sup>: Research Institute Nature and Forest, Kliniekstraat 25, Brussels, Belgium

<sup>2</sup>: Dept. Earth and Environmental Sciences, University of Leuven, Belgium

\*: corresponding author: stien.heremans@inbo.be



## INTRODUCTION

**Grasslands** are key ecosystems in Europe because of their:

- High biodiversity
- Large number of ecosystem services:
  - ✓ Carbon storage
  - ✓ Flood mitigation
  - ✓ Water purification
  - ...

**Sentinel 2** is a recent constellation of (currently) two satellites with good temporal and radiometric resolutions for vegetation studies.

## RESEARCH QUESTION

What is the potential of Sentinel 2 images for predicting **peak grassland biomass**?

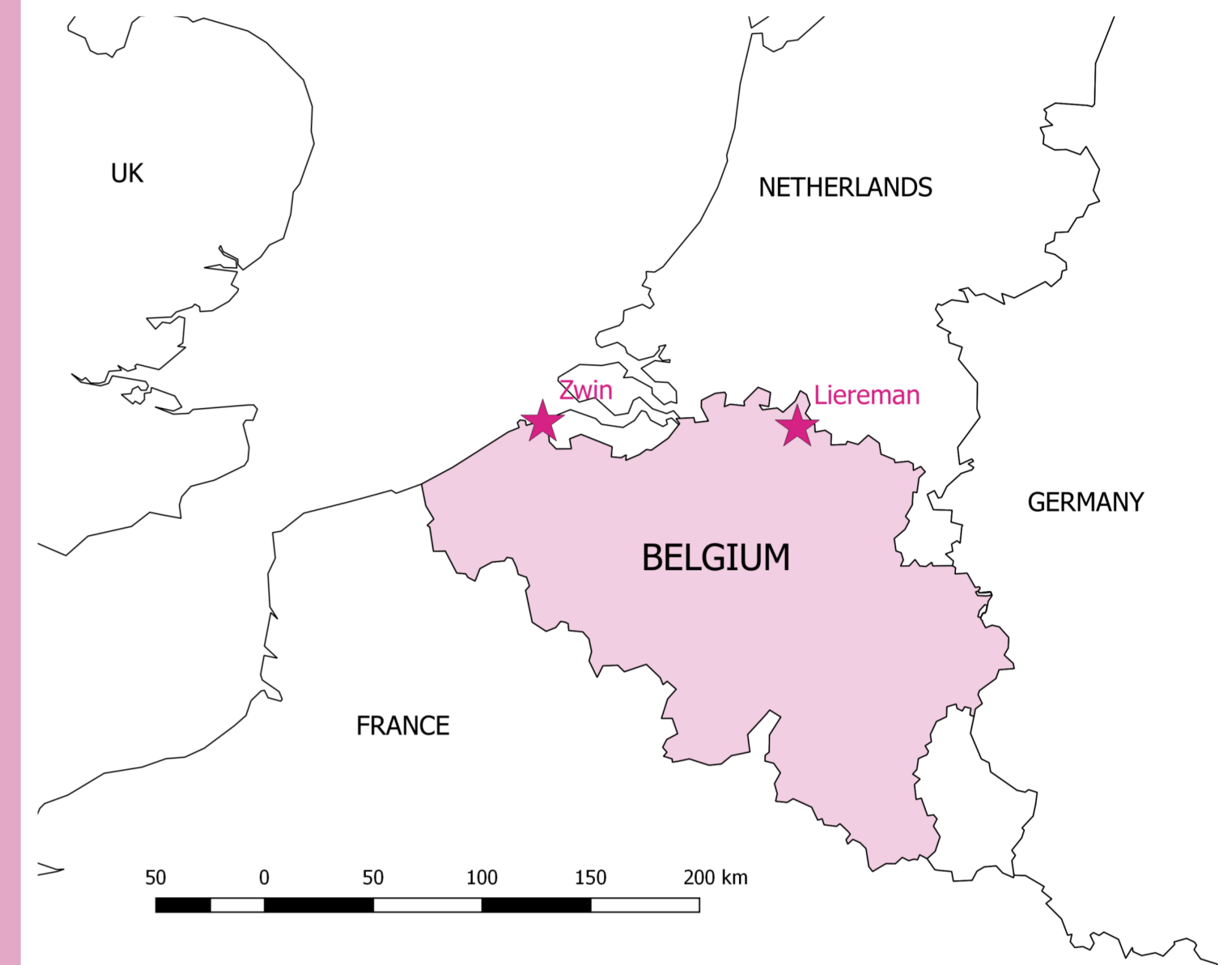
Peak biomass = maximum in the growing season, timing depends on the study area

Biomass is important as an indicator of biodiversity

## STUDY SITES

Two Natura 2000 nature reserves in Belgium (Liereman and Zwin, see map) with:

- ✓ a large number of (semi-)natural grasslands
- ✓ gradients in soil fertility and moisture,
- ✓ thus also gradients in peak biomass.



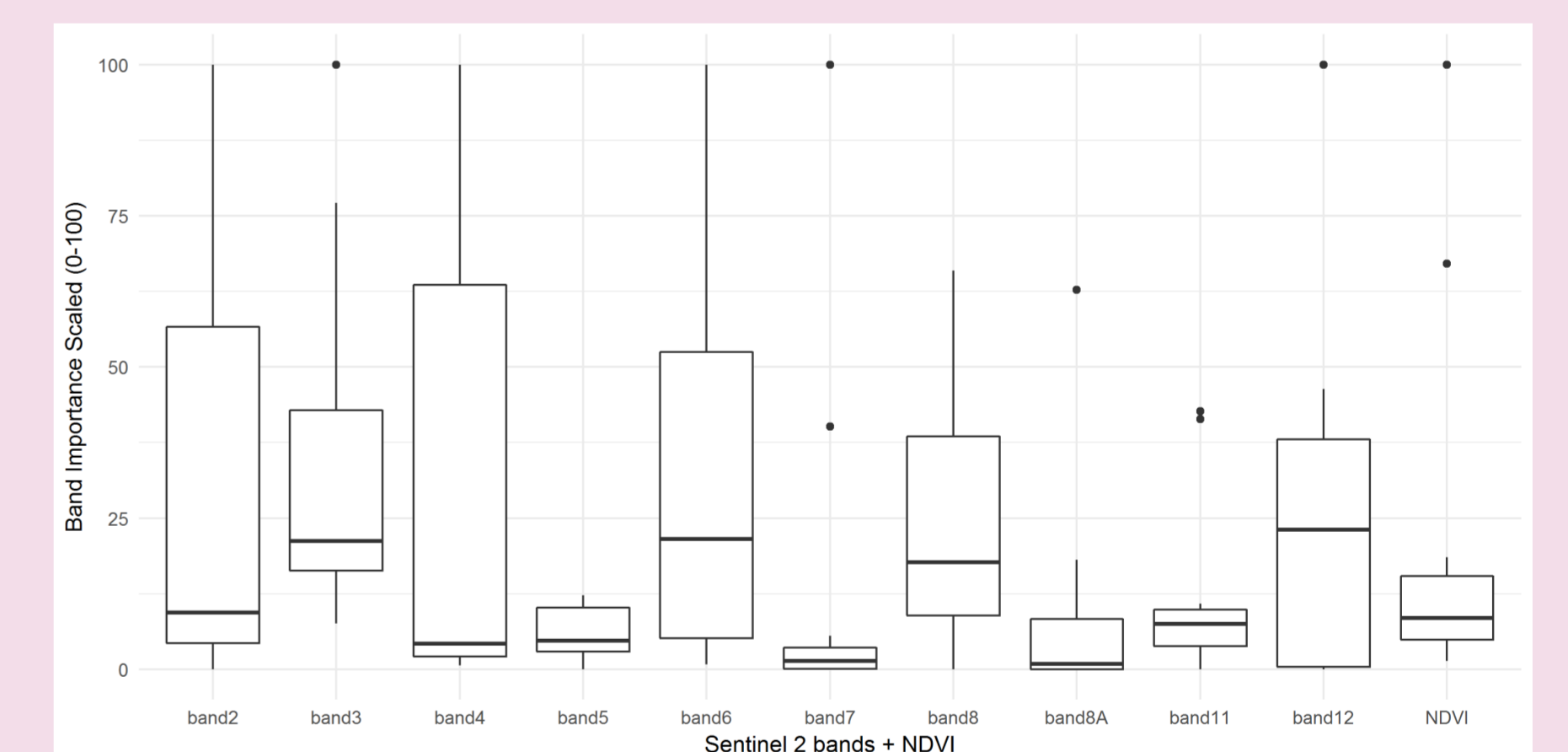
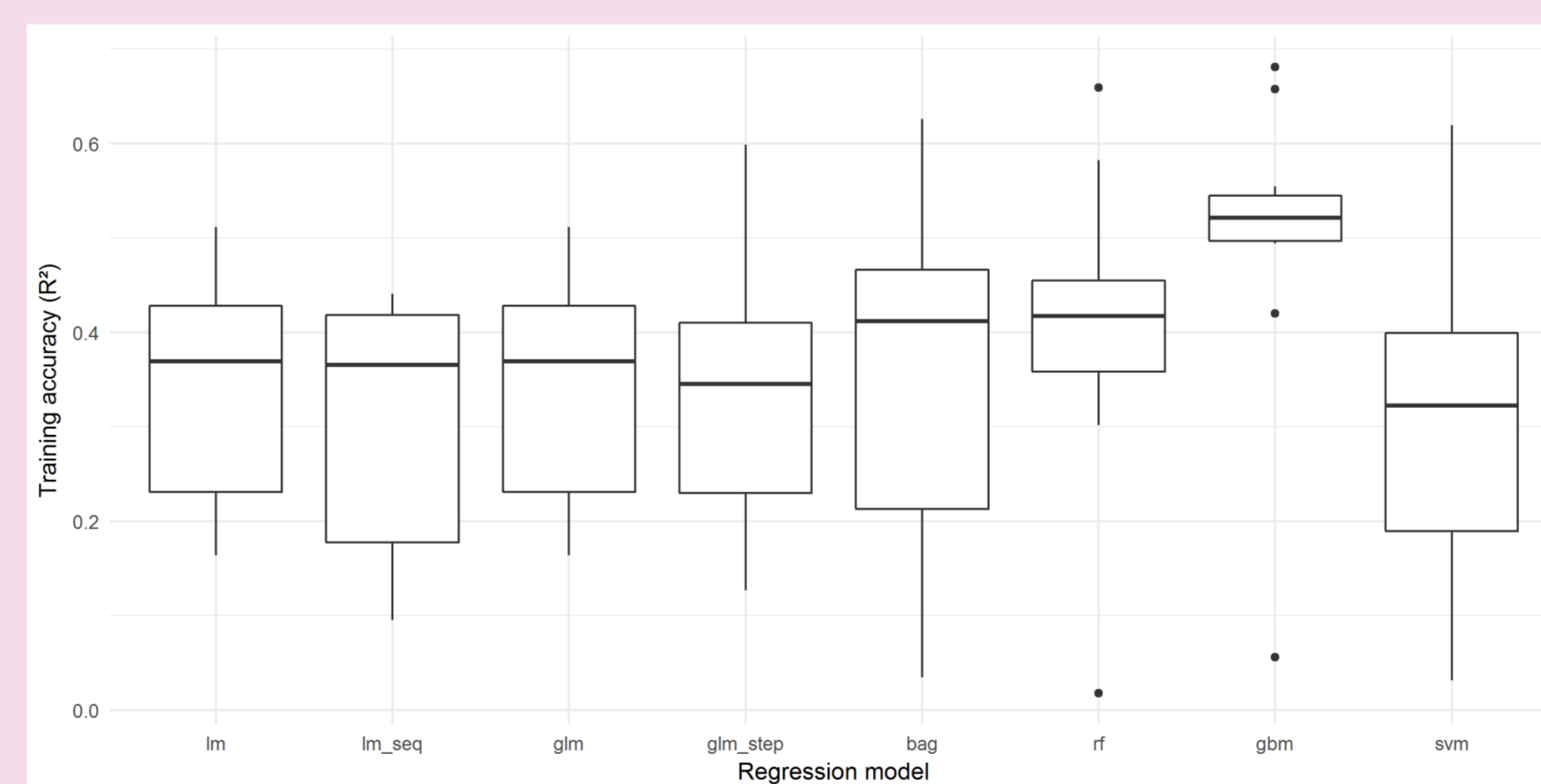
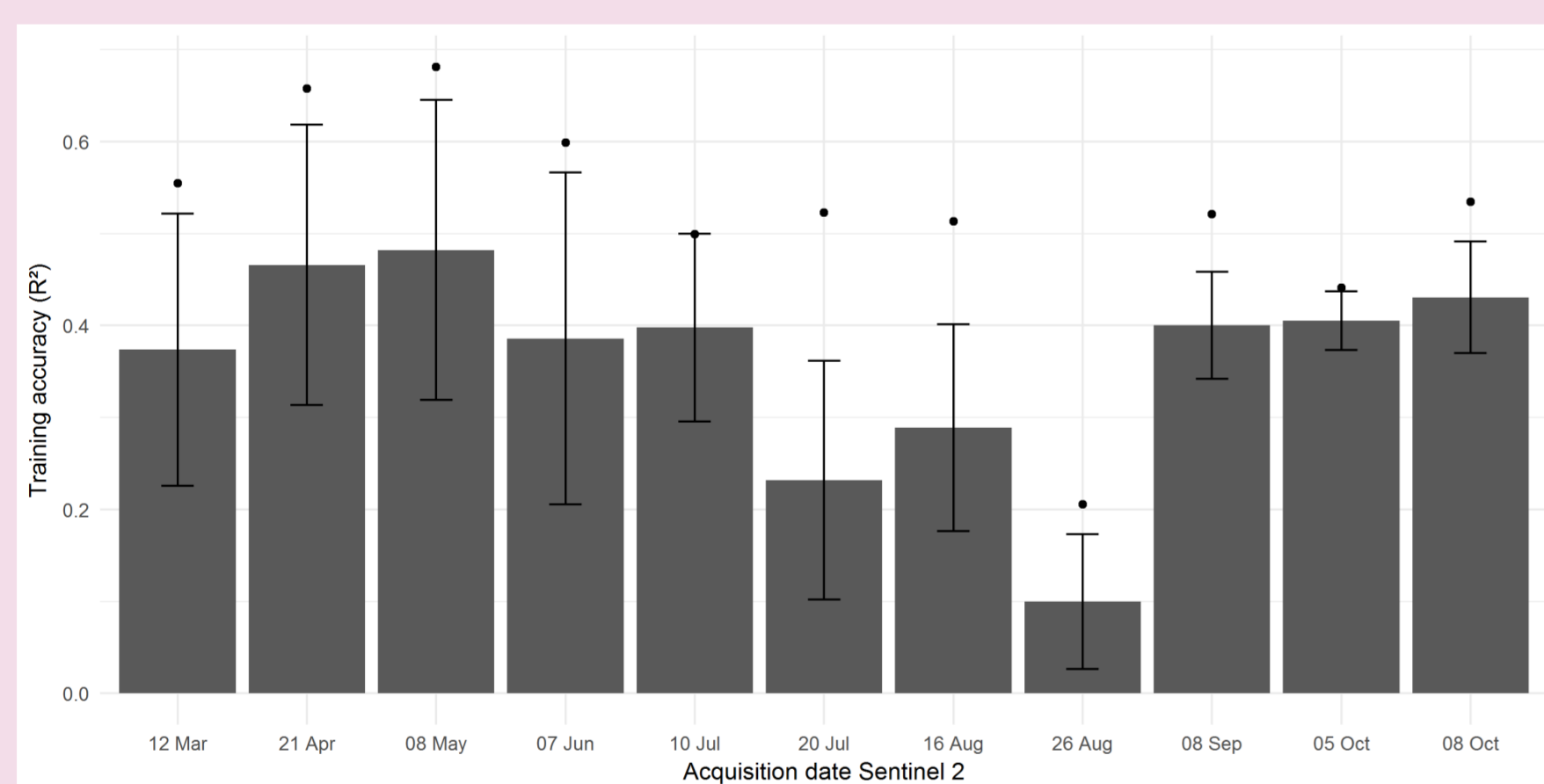
## METHODS

- **Peak biomass** (ground truth) was measured in the field in May-July for 30 plots per study site by cutting the biomass down to mowing level (above-ground only).
- All cloud-free **Sentinel 2A** images from March-Oct 2016 were downloaded and preprocessed (atmospheric correction using Sen2Cor and coregistration in ENVI)
- **8 types of regression models (see table)** were compared for predicting the peak biomass, per study site, ranging from simple linear regression to advanced machine learning methods (in Caret package in R)
- Models were trained using **leave-one-out cross-validation** (only 30 samples), the resulting  $R^2$  was used to compare Sentinel 2 acquisitions and model types.
- **Variable (band) importance** was assessed for the best model (gbm)

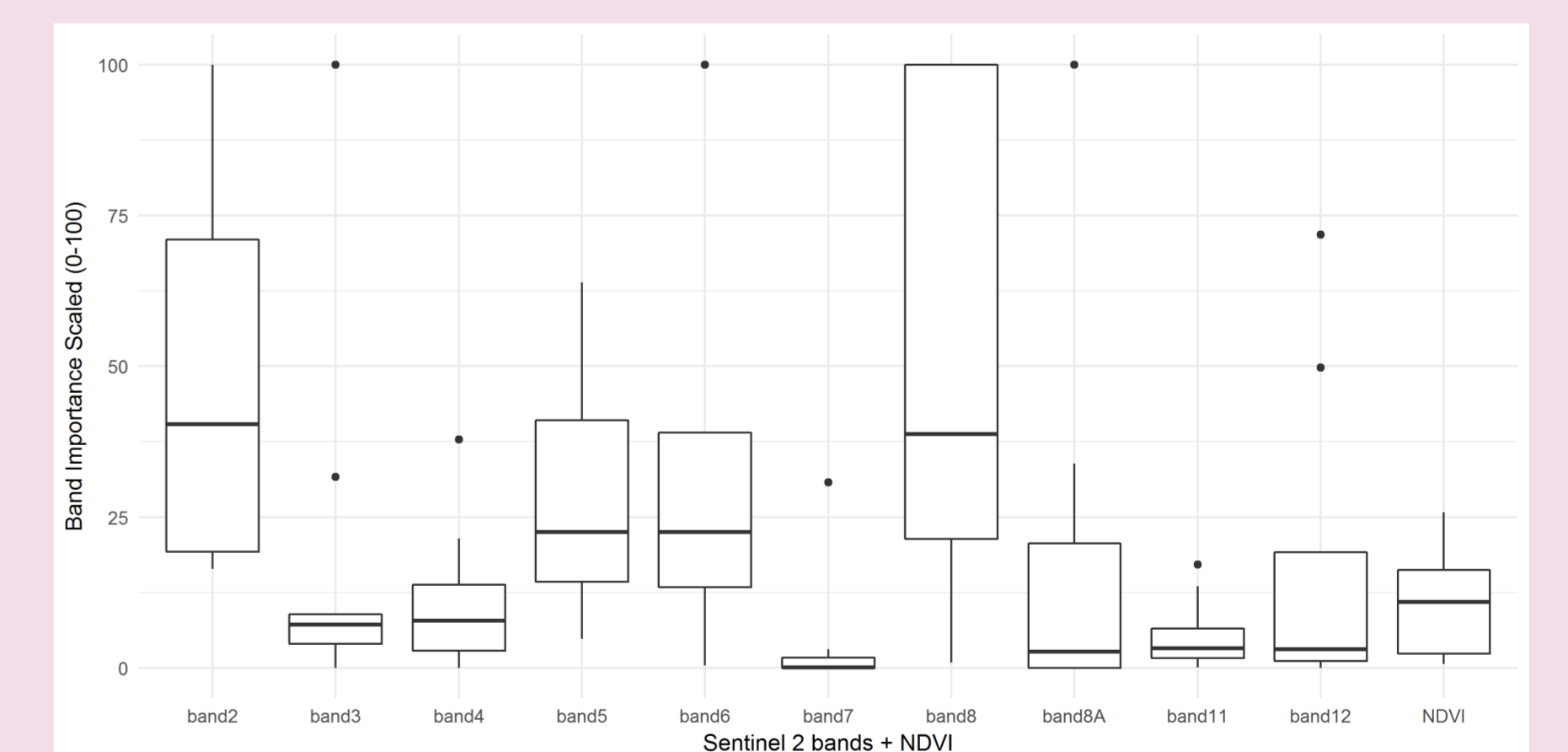
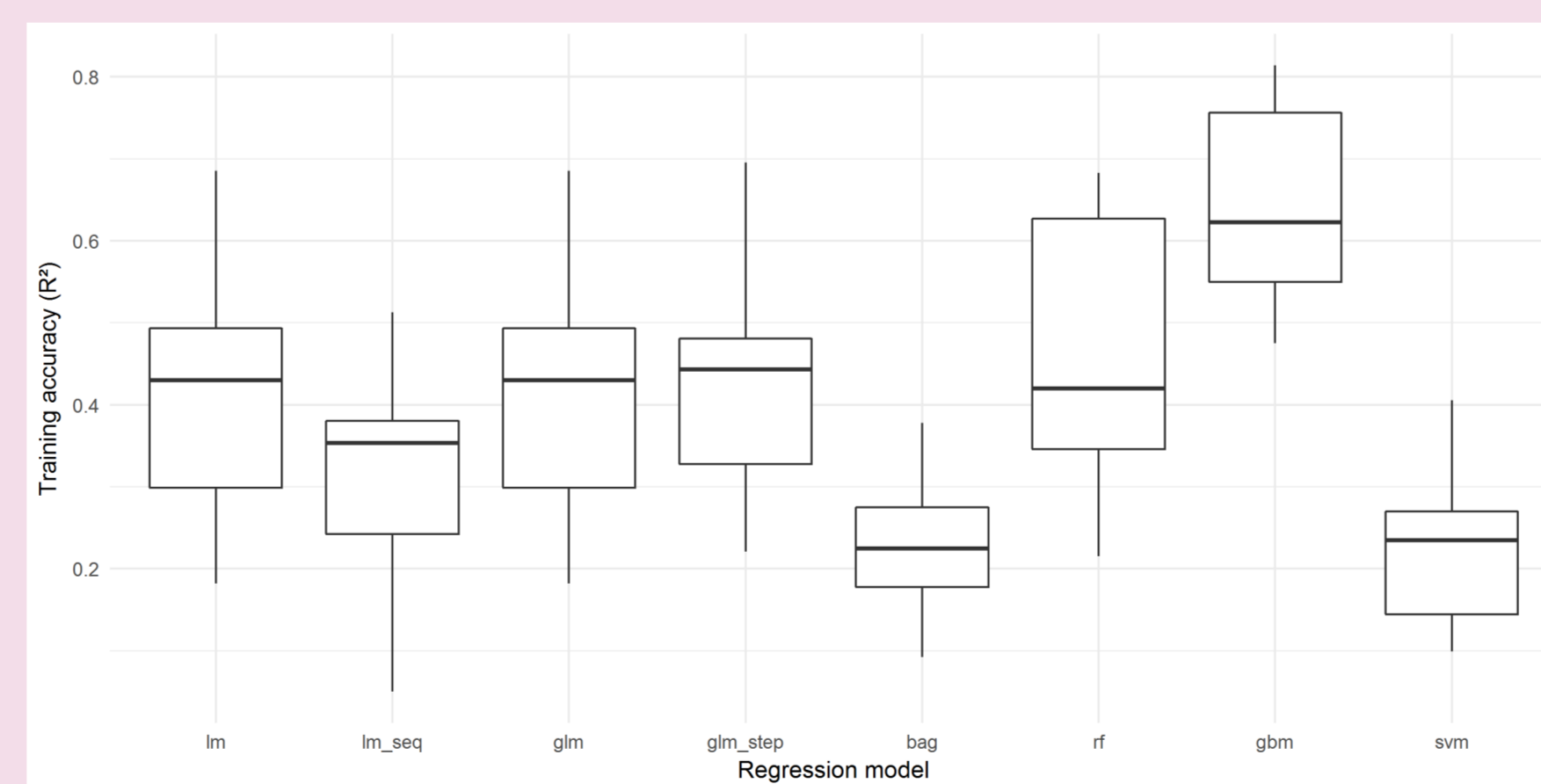
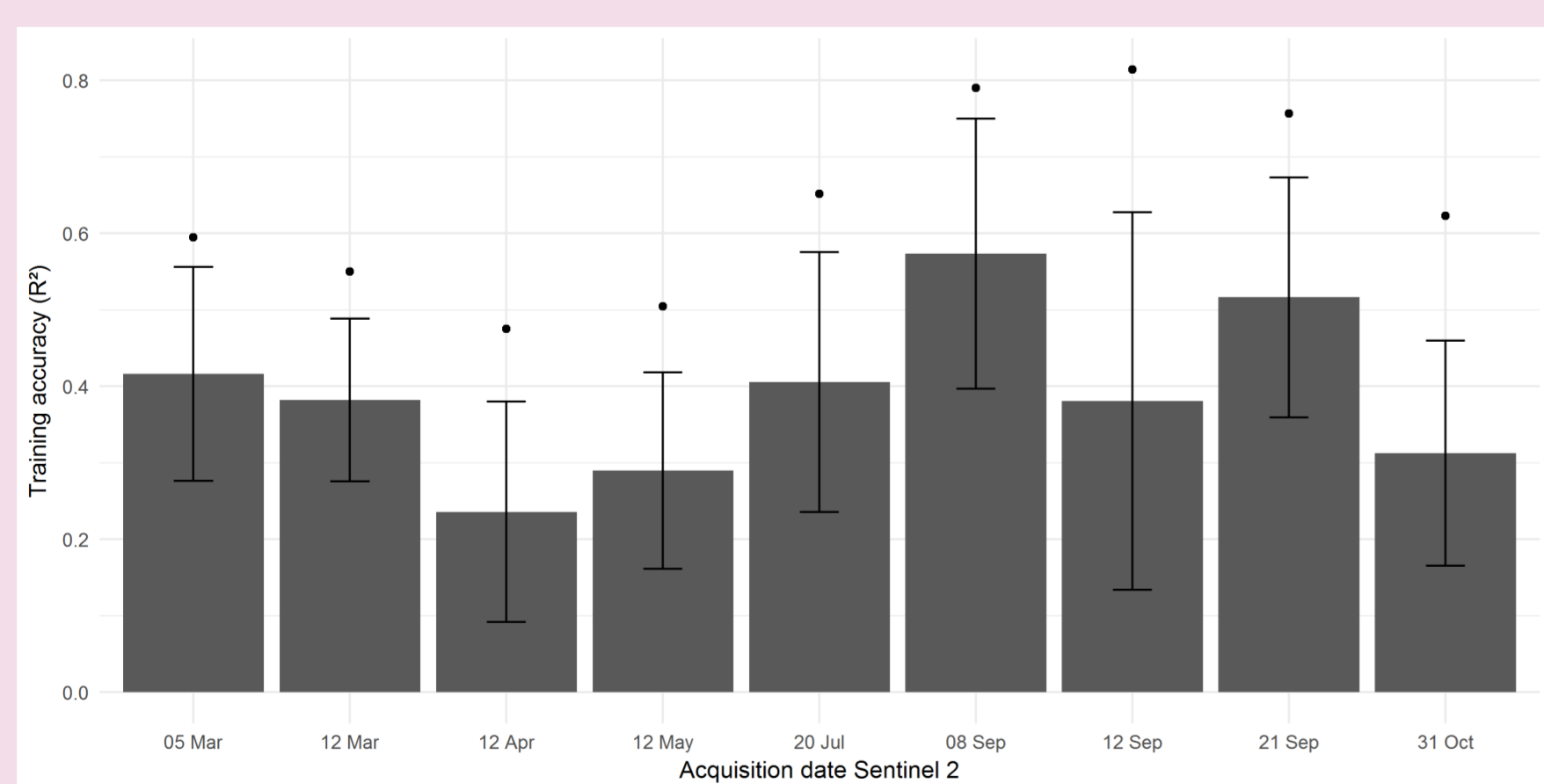
Regression models	abbreviation
Linear Regression	lm
Linear Regression With Stepwise Selection	lm_step
Generalized Linear Models	glm
Stepwise Generalized Linear Models	glm_step
Bagged Regression Trees	bag
Random Forest	rf
Boosted Regression Trees	gbm
Support Vector Machines	svm

## RESULTS

Liereman



Zwin



Training accuracy throughout growing season

Boxplots of training accuracy per model type

Importance of the Sentinel 2 bands in gbm (best model)

## CONCLUSION & FOLLOW-UP

- The results show that **accurate peak biomass prediction of grasslands with Sentinel 2 is possible within a single nature reserve and boosted regression trees are the best model type**, but the accuracy fluctuates throughout the season
- As a follow-up, we will try to **unravel the physical processes** that drive the relationship between biomass and reflectance
- For 2017, we will also test the potential of Sentinel 2 to predict the currently standing biomass (not peak only) and the **transferability of the prediction models in space and time**.