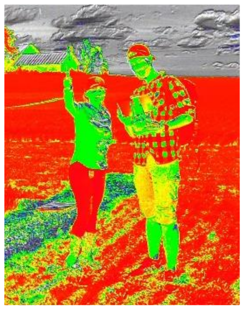
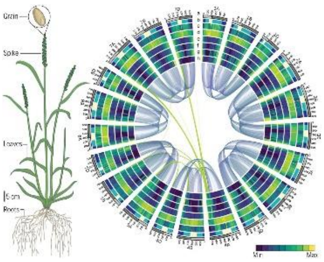
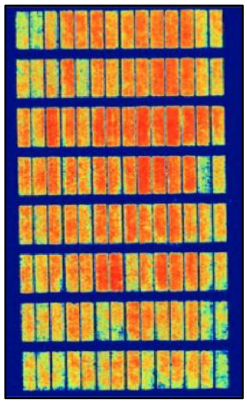
Master thesis topics in plant phenotyping and genetics – BIOVIT 2020



The following topics revolve around novel methodologies of high-throughput plant phenotyping (HTP), genomics and those two combined. HTP is a buzzing, recent topic addressing the need for time-and-workload-efficient screening of phenotypes of numerous plant genotypes in breeding programmes and basic research. It is a highly interdisciplinary domain – it involves robotics, drone technology, plant physiology and advanced data management schemes – data generated through HTP can be nearly considered "big data" thus needs dedicated tools for processing and analysis. Multispectral data ("normal" RGB channels + infrared reflectance) gathered with household-budget drones over the vegetation season in wheat has proven to carry enough information to allow us to predict the final yield of the crop through deep learning. Digital surface models built with RGB images can provide reliable plant height estimates, and close-up images of crop canopy show promise of extracting numerous variety features, such as spike density or architecture using image processing techniques and deep learning. By that, HTP not only allows us to do things faster, but also uncovers phenotypic traits unexplored before. However, as a young and promising discipline which found already its application niche, many aspects and especially the link between plant physiology and multispectral phenotype remains poorly understood. Doing your MSc project in HTP gives you an opportunity to change that!





We offer the following MSc thesis topics. Contact the supervisors if interested!

Best illumination geometry for wheat phenotyping using UAV Imagery

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Sahameh Shafiee, IPV: sahameh.shafiee@nmbu.no

Small unmanned aerial systems (UAS) have allowed the mapping of vegetation at very high spatial resolution for yield prediction, but a lack of standardization has led to uncertainties regarding data quality. For reflectance measurements and vegetation indices (VIs) to be comparable between sites and over time, careful flight planning and robust radiometric calibration procedures are required. One source of uncertainty that have received little attention is illumination geometry. We are aiming for a study to quantify and measure this effect on image indices and yield prediction model from Micasense RedEdge and Phantom4 multispectral camera. Generally, the following research questions need to be investigated:

- How sensitive are the individual bands measurements and indices to the light situation?
- Which camera is more stable? How consistent is reflectance measured by both sensors with ground measurements from a field spectrometer?
- What time during the day will be the best for flying regarding the Norwegian Weather condition?
- Is there any meaningful difference between different wheat varieties in response to the variations?

By doing this project, you will be involved in:

- Drone piloting
- Image processing
- Data analysis in QGIS, R or Python

Identification of genetic loci in wheat responsible for differences in multispectral reflectance captured by HTP drone imagery

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Sahameh Shafiee, IPV: sahameh.shafiee@nmbu.no

Plant canopy reflectance captured in the sub-visible region of the electromagnetic spectrum by satellites has been widely used since the 1980s to estimate green areas on Earth and physiological status of the plants. Recent development in affordable drone platforms has brought this technology to small-scale applications such as farms or orchards. Multispectral data captured over the season allows prediction of final crop output (such as yield) benefiting from advanced data management such as deep learning. However, the biological rationale behind the differences in canopy reflectance remains unclear. We propose to investigate the genomic regions consistently associated with the reflectance using multi-locus mixed linear modelling GWAS (genome wide association studies) on drone data captured over the vegetation season. This topic aims to answer the following questions:

- How does the wheat multispectral canopy reflectance behave throughout vegetation season for various genotypes?
- Are there any genomic regions consistently associated with multispectral canopy reflectance traits in wheat?
- If those regions exist, are they also associated with other plant traits like yield?
- Do those regions contain putative genes, allowing to form a rational link between plant physiology and multispectral phenotype?

By doing this project, you will be involved in:

- Drone piloting and imagery processing
- Data management and visualisation in R and Python
- Statistical analysis in quantitative genetics in R
- Bioinformatics – working with wheat reference genome

Integration of genotype, environmental and drone multispectral data to improve the accuracy of genomic prediction

Genomic prediction and selection have since the concept in 2001 been applied for both animal and plant breeding, yielding almost two-fold acceleration of genetic gains. However, cross-season and cross-environment prediction for plant traits does not yield satisfactory results due to intense genotype x environment (G × E) interactions present in plant populations. Drone multispectral images carry information on plant physiological performance, allowing prediction of final plant output using machine learning. Those models, however, works well only in the season they have been developed – predictions for other seasons are of low quality. We propose to improve the models developed on drone multispectral data by adding environmental and genomic covariates to enhance its across-season prediction ability. This topic is divided into two separate, standalone topics:

Sub-topic 1

Statistical methodology

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Tomasz Mroz, IPV: tomasz.mroz@nmbu.no

This topic will apply well-proven methods for genomic prediction such as Bayesian generalized linear regression and GBLUP, that will be extended to include G x E interactions and environmental covariables.

This topic aims to answer the following questions:

- Will adding environmental covariates improve the across-season prediction ability of the models?
- Will adding genotypic data (SNP markers) to the models improve the across-season prediction ability?

By doing this project, you will be involved in:

- Drone piloting and imagery processing
- Data management in R and Python
- Statistical analysis in R

Sub-topic 2

Deep learning methodology

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Sahameh Shafiee, IPV: sahameh.shafiee@nmbu.no

This topic will replace the classical statistical models for genomic prediction with deep learning methodology. It aims to answer the following questions:

- Will adding environmental covariates improve the across-season prediction ability of the model?
- Will adding genotypic data (SNP markers) to the model improve the across-season prediction ability?
- Does deep learning perform better in estimation of SNP effects than traditional statistical methods?

- Can deep learning models capture subtle genotype x environment interactions better than traditional statistical methods?

By doing this project, you will be involved in:

- Drone piloting and imagery processing
- Data management in R and Python
- Deep learning modelling in `keras/ tensorflow`
- Statistical analysis in quantitative genetics in R

Genetic basis for spike architecture in wheat mapping populations

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Tomasz Mroz, IPV: tomasz.mroz@nmbu.no

Yield is a complex trait controlled by numerous loci. Each genotype achieves its yield level as a combination of yield components (such as kernel weight, kernel size, number of kernels per spike or area, spike fertility, spike size to name but a few). Breeding progress in Nordic spring wheat has achieved yield improvements of around 20 kg per hectare per year over the past five decades. The main drive for this improvement was gradual increase of grains per spike, like other European breeding programmes. However, the basis of this grains per spike increase is unknown. We propose to investigate the contribution of spike architecture (spike length, number of spikelets and spike fertility) to historical yield increase in Norwegian wheat and to perform a genetic dissection of those traits using the MASBASIS association mapping population. The topic aims to answer the following questions:

- Is there a consistent trend over the past five decades of breeding in spike traits like spike length, number of spikelets per spike and spike fertility?
- What is the contribution of those spike traits to final grain yield?
- Are those traits affected by fertilization?
- Are there genomic regions consistently associated with those traits?
- Do those regions overlap with regions associated with grain yield?

By doing this project, you will be involved in:

- Field phenotyping for spike traits
- Data management in R/Python
- Statistical analysis in quantitative genetics in R
- Advanced data visualisation in R

Shadow removal from multispectral images using end-to-end Deep Learning model

Supervisors:
Morten Lillemo, IPV: morten.lillemo@nmbu.no
Sahameh Shafiee, IPV: sahameh.shafiee@nmbu.no

Clouds and their shadows are common features in UAV and satellite images, this causes serious problems for crop yield estimation in plant phenotyping. We propose to apply Deep Convolutional Neural Network to generate an automatic shadow removal algorithm for UAV multispectral images. The topic aims to answer to the following questions:

- Will the CNN model produce finer semantic information which leads to better process details in comparison to statistical analysis?
- Will an end-to-end CNN be able to remove the shadow automatically from multispectral images?

By doing this project, you will be involved in:

- UAV piloting
- Image processing in Python and QGIS
- Learning and applying Deep Convolutional Neural Network for image shadow removal

Do you want to know more?

https://doi.org/10.1016/j.ifacol.2017.08.1591