Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Federal Department of Economic Affairs,
Education and Research EAER

Agroscope

Module 5- Concept for a final CAS project

Nicolas Vuille-dit-Bille

Agroscope

SMART objectives

- Predict winter wheat nitrogen (N) status in different environments and at different crop stages using hyperspectral indicators (spectral vegetation indices) in order to optimize N fertilization strategy (economic and environmental cost), crop yield and grain quality (recommendation to farmers).

- The prediction should perform accurately at 3 hierarchical levels:

1) Site (soil N status and climatic condition) → e.g. Site 1 needs more N than Site 2

2) Field (spatial heterogeneity) → e.g. One part of the field needs more N than the other part

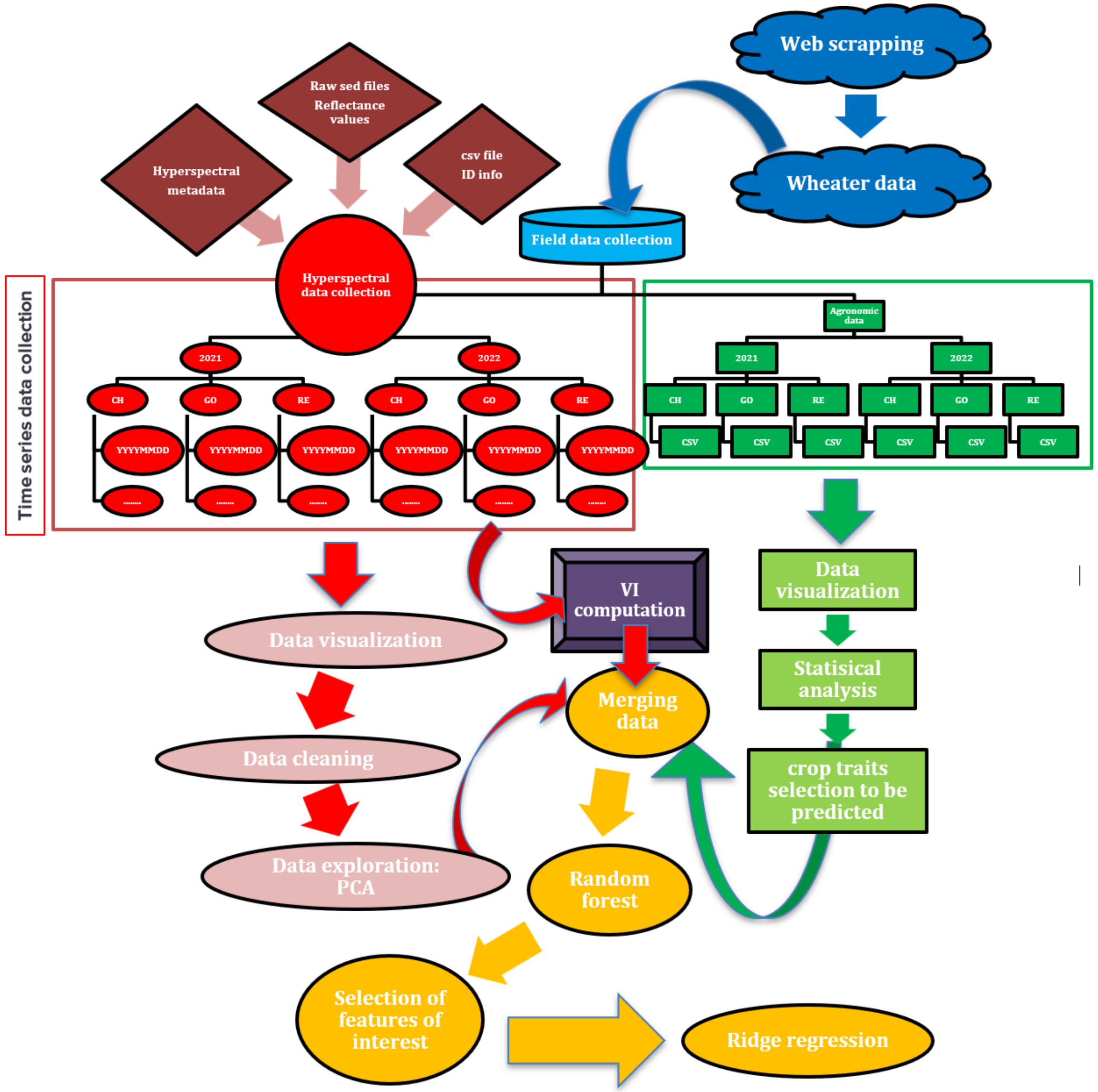
3) Wheat variety (N use efficiency (NUE)) → e.g. This specific variety needs more N because it as lower NUE

- At the end, a selection of specific spectral band should allow the model to predict accurate amount of N to apply at specific environment, field and variety level in order to optimize fertilization application (maximum grain yield for minimum N loss)

Planned data

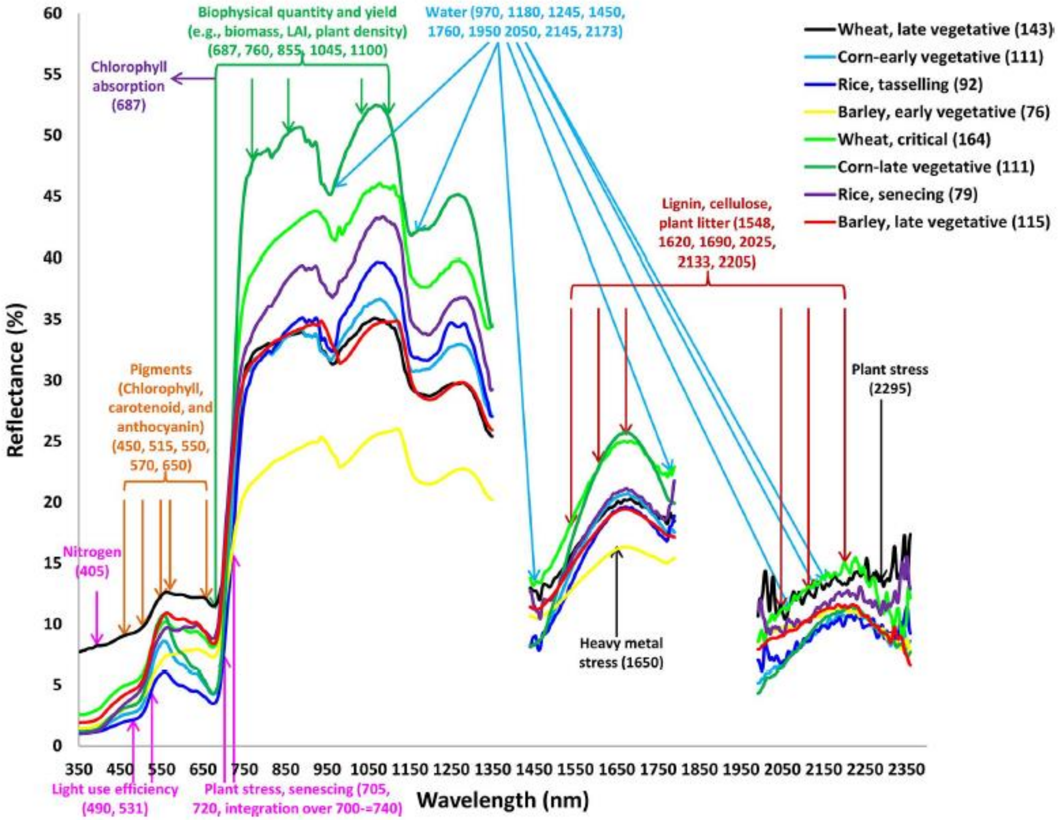
- 5 winter wheat varieties
- 2 years: 2021 to 2022
- 3 sites: Changins, Goumoens and Reckenholz
- 3 main treatments: none, reduced and conventional
- Measurement to estimate varieties performance:

➢ Grain yield
➢ Straw yield
➢ Grain protein
➢ Other physiological parameters (harvest index, leaf area index, chlorophyll content, canopy cover…)
➢ **Hyperspectral data**



Data quality

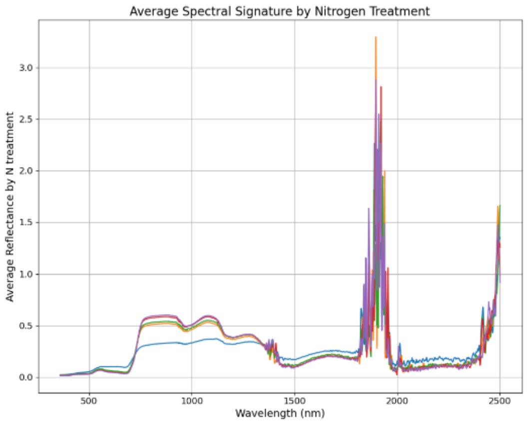
THEORY



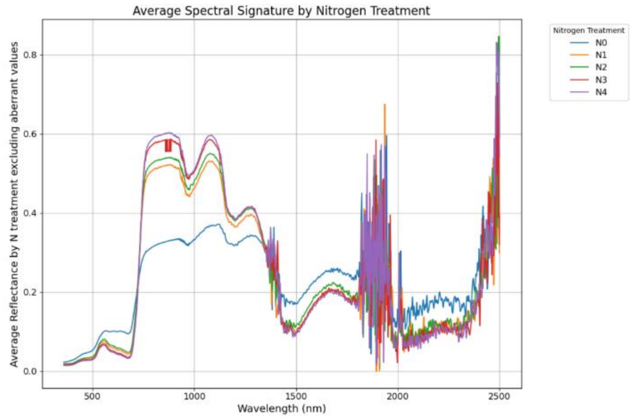
Thenkabail, P. S., Mariotto, I., Gumma, M. K., Middleton, E. M., Landis, D. R., & Huemmrich, K. F. (2013). Selection of hyperspectral narrowbands (HNBs) and composition of hyperspectral twoband vegetation indices (HVIs) for biophysical characterization and discrimination of crop types using field reflectance and Hyperion/EO-1 data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 6(2), 427-439. https://doi.org/10.1109/JSTARS.2013.2252601

FIELD DATA

A



B



C

rflt_1897	rflt_1898	rflt_1899	rflt_1900	rflt_1901	rflt_1902	rflt_1903	rflt_1904	rflt_1905	rflt_1906	rflt_1907	rflt_1908	rflt_1909	rflt_1910	rflt_1911	rflt_1912	rflt_1913	rflt_1914	rflt_1915
3.65109	1.57072	0.11207	0	0	0.49875	3.40531	6.31836	6.93534	5.42574	3.44232	2.08599	0.8421	0.0401	0.09552	0.68635	0.99688	0.70513	0.22659
0	0	0	0	0	0.20351	1.30842	2.54547	3.29907	3.65119	3.90309	4.32029	4.82436	5.0141	4.6142	3.82203	2.90764	1.75334	0.56256
5.4275	2.55384	0.48193	0	0	0	0	0	0.32878	1.30063	2.57771	4.41397	6.88327	8.03605	6.11016	2.43515	0	0	0
0	0	0	0	0	0.04905	0.22445	0.6276	1.59903	3.30271	4.76955	5.17186	4.84432	4.08858	2.79168	1.08563	0	0	0.01463
5.37194	2.27108	0.0801	0	0	0.3443	2.32134	4.40049	5.14362	4.61088	4.00978	4.28905	5.07373	5.92674	6.98027	8.06941	7.88266	5.33062	1.95038
0.81859	2.04315	2.59596	1.91616	0.76257	0.07337	0	0	0.27699	1.21087	1.94526	1.68752	0.6835	0.00747	0.48706	1.51293	1.99315	1.3634	0.38286
1.93819	0.74591	0	0	0	0.12372	0.73757	1.55642	2.52785	3.81587	4.84425	5.10155	4.79747	4.08797	2.79606	1.09488	0	0	0
3.27589	1.56713	0.33642	0.07524	0.17031	0.17755	0.0956	0.02367	0	0	0	0	0	0	0	0	0.0064	0.0616	0.18455
1.14922	0.42958	0	0	0	0.20863	1.36616	2.60572	3.19384	3.23986	3.05809	2.77345	2.2949	2.00312	2.3005	2.87672	2.96343	2.00092	0.77056
6.31177	2.53065	0	0	0	0.3682	2.50787	4.58384	4.73626	2.9105	1.17578	1.2578	2.49347	3.88858	5.28418	6.71098	6.92894	4.96177	2.18082
0	0	0.29913	1.14134	2.11909	2.53421	2.28112	1.75126	1.24262	0.62538	0.09923	0	0	0	0	0	0.02468	0.46145	0.99958
0.78897	0.31633	0	0	0	0.18593	1.27455	2.31343	2.34271	1.37021	0.32577	0	0	0	0.8807	2.24004	3.02623	2.84011	2.19731
0.54719	0.00091	0.05426	1.35364	2.97649	3.44338	2.39589	0.97024	0.25355	0.05448	0.01941	0	0	0	0.89127	2.24505	3.02574	2.80263	2.14106
1.8161	1.2263	0.71446	0.38403	0.12306	0	0	0	0.006	0	0.4497	2.78224	6.7009	8.98935	7.06027	2.82781	0	0	0
5.44003	3.69178	2.44344	2.3637	2.59375	2.32541	1.35329	0.31996	0	0	0	1.03008	2.87957	3.99043	3.14675	1.24665	0	0	0.10897
3.01297	1.19763	0	0	0	0.2584	1.76293	3.21643	3.25489	1.74557	0.56125	1.42127	3.72405	5.85795	7.24715	8.21736	7.8752	5.18782	1.64394
5.28126	3.29602	1.94438	1.96428	2.41014	2.42069	1.89147	1.24672	0.75402	0.21031	0.12636	1.3102	3.59699	4.9896	3.93647	1.57768	0	0	0
0	0	0.02033	0.41413	0.86717	0.97242	0.60299	0.13592	0	0.02957	0.13389	0.12389	0.05917	0.00291	0	0	0.03497	0.67788	1.44723
3.82576	1.52524	0	0	0	0	0	0	0	0	0	0.83589	2.60319	3.93772	4.10994	3.62621	2.92847	1.86369	0.70435
1.51907	0.54738	0	0.12876	0.60934	1.19642	2.14235	2.92237	2.73378	1.50242	0.31528	0	0.11688	0.88174	2.99495	5.83466	6.96641	4.94827	1.75135
9.26216	8.25811	6.86398	4.98083	2.85051	1.44894	0.91119	0.71853	0.45186	0.20684	0.03585	0	0	0	0	0	0.03759	0.59401	1.38275
0.78021	0.29783	0	0.01707	0.17226	0.26313	0.27788	0.28465	0.29655	0.24524	0.37956	1.02185	2.09111	2.95082	3.18501	3.08764	3	3	3
4.01407	10.0492	13.0255	10.5126	5.76396	2.60768	1.12044	0.29945	0	0	0	1.4921	4.64609	6.91176	6.81081	5.38885	3.86886	2.12624	0.461
0.52924	1.41811	2.03425	2.22447	2.16642	1.88139	1.22152	0.51891	0.21235	0.26766	0.37709	0.30411	0.14125	0.00744	0	0	0.09983	1.93012	4.12489
0.95261	2.37456	2.99046	2.10132	0.65867	0	0	0.23182	0.24433	0	0	0.89976	2.5685	3.92348	4.39194	4.35785	3.93777	2.66389	1.25786
1.74158	2.69089	2.97404	2.04297	0.65633	0	0	0	0	0	0	0	0	0	3.49292	8.70671	10.9645	7.66269	2.35185
2.36691	0.94899	0	0	0	0.04842	0.31588	0.60499	0.75857	0.81762	0.78215	0.60118	0.27782	0.01482	0	0	0.04005	0.78142	1.66402
0	0	0.00998	0.19301	0.41249	0.50578	0.42347	0.31326	0.33885	0.50708	0.62806	0.5209	0.24479	0.01253	0	0	0.04067	0.82827	1.73434
0.49318	1.42973	2.02426	2.03145	1.75393	1.37434	0.78361	0.19059	0	0	0	0.65621	2.0429	2.97404	2.6921	1.74414	0.94708	0.4879	0.15627
4.94584	5.88809	5.93987	4.42599	2.30051	1.0106	0.70185	0.72965	0.62906	0.51624	0.41081	0.22911	0	0	0.46711	1.27388	2.07387	2.97309	3.78122

- ➢ Remove bands with aberrant values : <0 or >1
- ➢ Remove bands with 0 value

Agroscope

Methods

Unsupervised learning

- Data exploration and features selection with PCA

Supervised learning

- Random forest
- Partial least square (PLS)
  - Not covered in the course
  - Best approach to deal with hyperspectral data
  - Performs dimensionality reduction and handles multicollinearity

Performance metrics (PLS)

- **Root Mean Squared Error (RMSE)**

➢ Model error with same unit as the response variable

- **Compare Mean Absolute Error (MAE) and Mean Squared Error (MSE)**

➢ MAE: less sensitive to ouliers

- **R^2 (Coefficient of Determination) or adjusted R^2**

➢ Proportion of variance in the response variable predictable from the predictors

➢ Adjusted R^2 penalizes the inclusion of irrelevant predictors (more realistic in this context

- **Cross-Validation Metrics:**

➢ Q^2: similar to R^2 but on unseen data

- **Predicted Residual Sum of Squares (PRESS):**

➢ Sum of squared prediction errors from cross-validation

➢ Lower PRESS indicates better predictive performance

Other performance metrics (random forest)

- **Mean Absolute Percentage Error (MAPE):**

➢ Expresses prediction errors as a percentage of actual values

- **Prediction Error Variance (Uncertainty):**

➢ Identify regions where predictions are less certain by comparing variance in predictions across trees in the forest

- **Out-of-Bag (OOB) Error:**

➢ Cross-validation-like estimate of prediction error computed during model training (out of bootstrap sample)

- **Cross-Validation Metrics:**

➢ k-fold cross-validation to estimate performance metrics like RMSE, MSE, or R^2 on unseen data (model generalization estimate)

Thank you for your attention

Nicolas Vuille-dit-Bille

nicolas.vuille-dit-bille@agroscope.admin.ch

Agroscope good food, healthy environment

www.agroscope.admin.ch