Algorithms for Predictive Plant Monitoring

STUDENTS / UNIVERSITIES

Umut Çakır / Boğaziçi University

M.Orhun Gale / Sabancı University

SUPERVISOR(S)

A. Ozan Biçen

Nihal Öztolan Erol

Stuart J. Lucas

ABSTRACT

Pressure on food production increases day by day because of population growth and climate change. To overcome this pressure, increasing the yield in every arable land is essential. Since osmotic stress limits crop yield, in this project, selecting *T. Durum* wheats which have osmotic stress resistance by various statistical methods and Machine Learning algorithms is aimed. A data set which includes some morphological data of T. Durum wheats such as root length, surface area, root volume, average root diameter and number of tips are examined. After these processes, we selected AS37, AS74 and AS131 as osmotic stress resistant samples.

Figure 1: T. durum and its relatives [1]

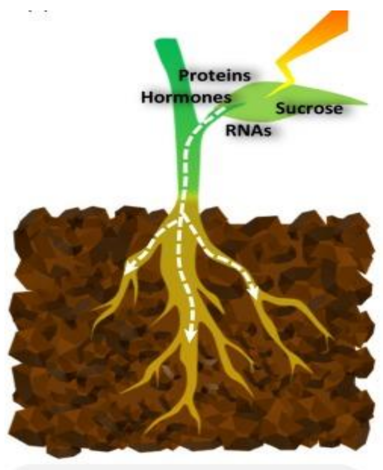
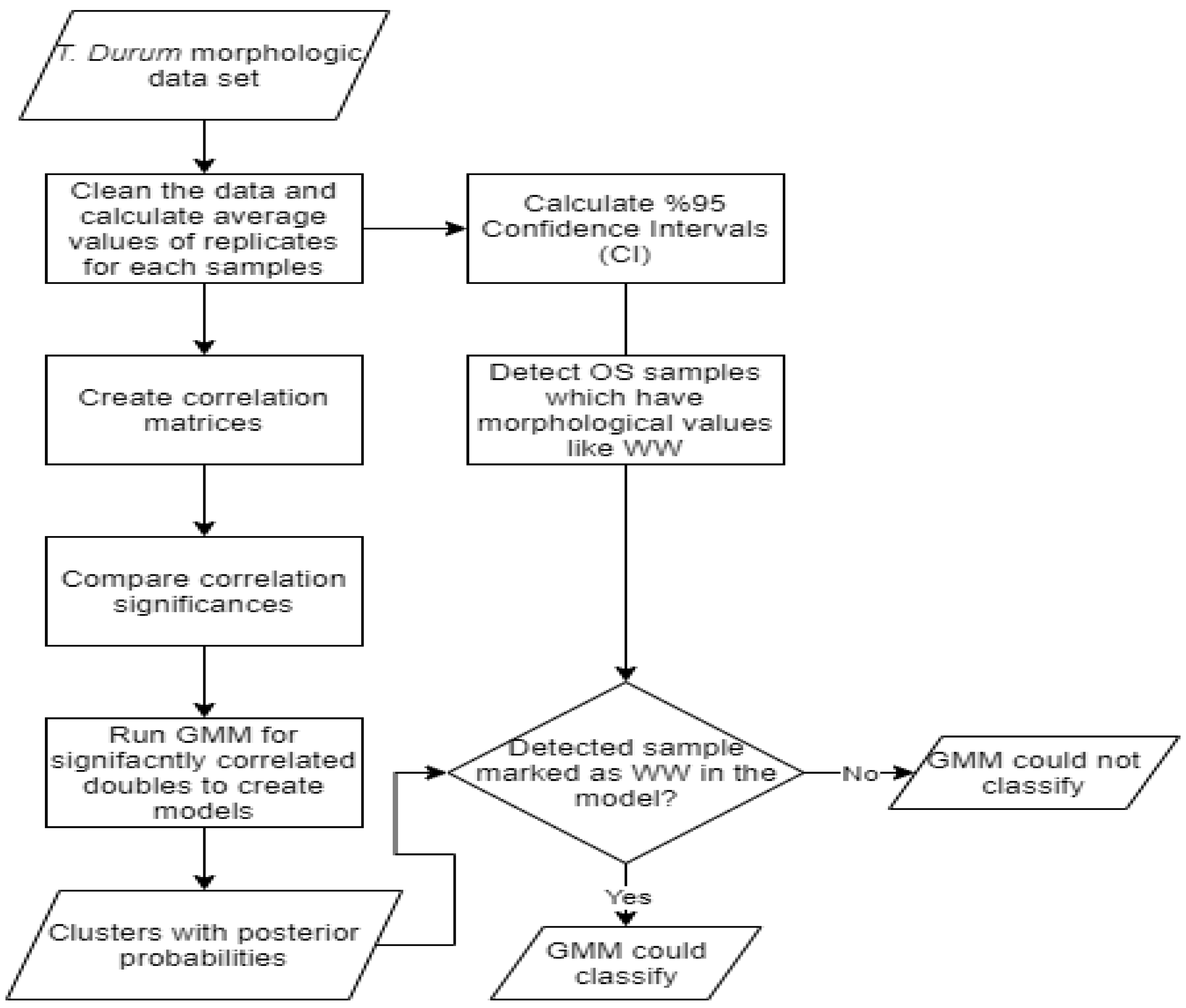


Figure 2: Illustration of wheat root [2]

OBJECTIVES

Selection of *T. Durum* wheats which have osmotic stress resistance among various samples by analyzing phenotypical data of the samples by using statistical methods and Gaussian Mixture Model (GMM), which is a machine learning algorithm.

METHODOLOGY



DATA ANALYSIS

Comparing Days in the Same Treatment:

Table 1 shows median value of root parameters in OS3 and OS5. According to Table 1, length, surface area, and root volume, and tips significantly increase in time, but average dimeter does not change significantly.

Likewise, Table 2 shows median value of root parameters in WW3 and WW5. According to Table 2, length, surface area, and root volume also significantly increase in time and average diameter does not change significantly. Unlike OS3/OS5, tips does not change significantly in time.

Table 1: Comparison of OS3 and OS5; *** : $p < 0.001$

Median	Lcm	SAcm2	AvDmm	RVcm3	Tips
OS3	10.34 ***	2.02 ***	0.65	0.03 ***	12.20 ***
OS5	16.82 ***	3.57 ***	0.66	0.06 ***	21.25 ***

Table 2: Comparison of WW3 and WW5; *** : $p < 0.001$

Median	Lcm	SAcm2	AvDmm	RVcm3	Tips
WW3	15.67 ***	3.53 ***	0.68	0.06 ***	15.00
WW5	21.84 ***	4.95 ***	0.71	0.09 ***	13.67

Comparing Treatments in the Same Day:

We can also compare treatments in the same day. OS3/WW3 and OS5/WW5 comparisons are shown in Table 3 and Table 4, respectively. In both table, in WW condition length, surface area, and root volume are significantly higher than OS condition and average diameter is not significant. On the other hand, while at 3rd day tips is not significant, at 5th day it is significant, that is, tips is significantly higher in OS5 than WW5. Then, average diameter seems to non-informative for both days or treatment and tips is also non-informative only for OS3/WW3.

Table 3: Comparison of OS3 and OS5; *** : $p < 0.001$

Median	Lcm	SAcm2	AvDmm	RVcm3	Tips
OS3	10.34 ***	2.02 ***	0.65	0.03 ***	12.20
WW3	15.67 ***	3.53 ***	0.68	0.06 ***	15.00

Table 4: Comparison of OS3 and OS5; *** : $p < 0.001$

Median	Lcm	SAcm2	AvDmm	RVcm3	Tips
OS5	16.82 ***	3.57 ***	0.66	0.06 ***	21.25 ***
WW5	21.84 ***	4.95 ***	0.71	0.09 ***	13.67 ***

Tips effect on root development:

If a plant has small number of roots (tips), it does not have chance to develop "good" root like WW condition. Small tips always mean small length, small surface area, or small root volume. To select osmotic stress resistance samples first we look high number in tips.

Selecting Sample which have resistance against osmotic stress:

We determine which OS samples develop their root like WW samples. In other words, osmatic stress resistance samples develop their roots like WW condition even though water is limited. Confidence intervals (CI) for each parameter were calculated. OS samples whose parameters fall into respective WW CI were selected. Still most sample could not be candidate according to all parameters, so we gave one nonmatch privilege. In general, OS samples should have high number in tips, length, surface area, and root volume. Following list was obtained (Table 5).

Table 5: Samples whose tips are higher

Sample ID	L (cm)	SA (cm2)	AvD (mm)	RV (cm3)	Tips
AS37	22.53612	5.1372	0.71842	0.0948	35.6
AS74	20.96398	4.92894	0.72938	0.1002	25.4
AS131	25.59553	4.907775	0.61125	0.07675	28.25

Cluster Analysis:

GMM is used to create covariance structure options and posterior probability models in order to create cluster models. Posterior probability models are created by using significantly correlated doubles and detected samples are found in these models to find out whether these models are marked as WW or not. In half of the models, selected samples are marked as WW. In the rest of the models, after data normalize to a better form, samples may be marked as WW as well.

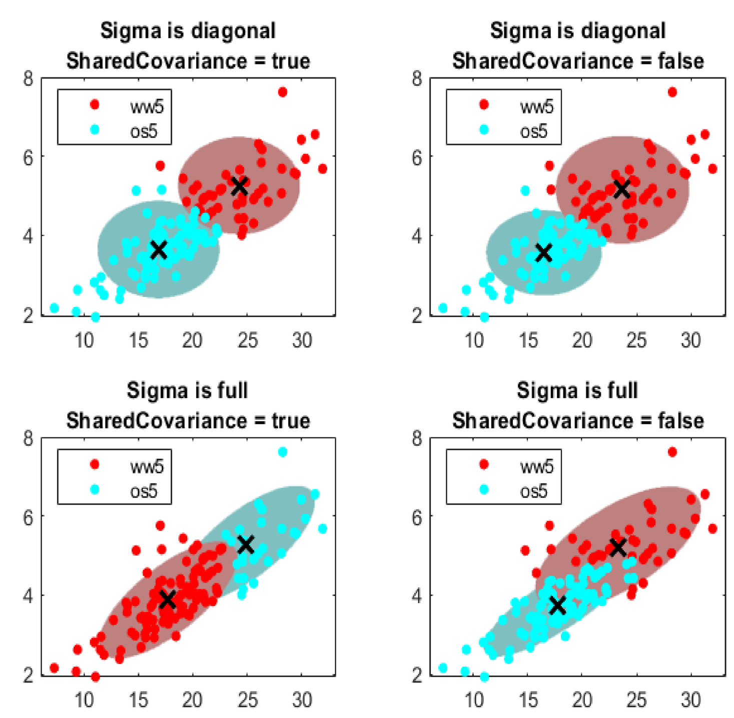


Figure 3: Possible covariance options for Length vs. Surface Area clusters

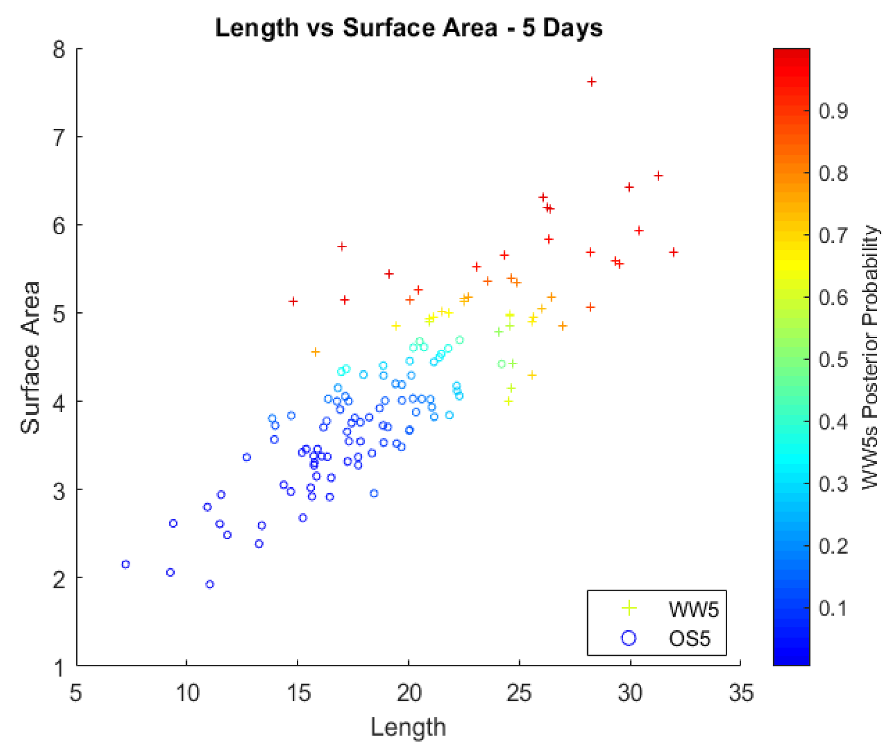


Figure 4: Posterior probabilities for Length vs. Surface clusters

DISCUSSION AND CONCLUSION

- AS37, AS74, and AS131 samples were selected as osmotic stress resistance candidates by using confidence intervals.
- GMM algorithm mark the selected samples as statistical methods remark in the cases which are properly modelled.
- Adjustments in data set may fix problematic models too and algorithm may mark samples completely as expected. Therefore, it can be said that selecting plants which have osmotic stress resistance by ML algorithms is possible.
- For future works with this data set, biologically irrational samples must be removed in accordance with the biological information. It may make ML algorithms work more efficient.
- In order to find which genes/proteins have a role in osmotic stress resistance, transcriptome, proteome, single nucleotide polymorphism or mutation analysis can be conducted.

REFERENCES

- Budak, H. (2017). High-throughput SNP genotyping of modern and wild emmer wheat for yield and root morphology using a combined association and linkage analysis. *Functional & Integrative Genomics*, *17*. https://doi.org/10.1007/s10142-017-0563-y
- Chlingaryan, A., Sukkarieh, S., & Whelan, B. (2018, August 1). Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture*. Elsevier B.V. https://doi.org/10.1016/j.compag.2018.05.012
- Dhanda, S. S., Sethi, G. S., & Behl, R. K. (2004). Indices of Drought Tolerance in Wheat Genotypes at Early Stages of Plant Growth. *Journal of Agronomy and Crop Science*, *190*(1), 6–12. https://doi.org/10.1111/j.1439-037X.2004.00592.x
- Fao, & Dwfi. (n.d.). *Yield gap analysis of field crops: Methods and case studies*. Retrieved from www.fao.org/publications
- Genetic markers signal increased crop productivity potential | News | CORDIS | European Commission. (n.d.). Retrieved 25 August 2020, from https://cordis.europa.eu/article/id/118823-genetic-markers-signal-increased-crop-productivity-potential
- Lezoche, M., Panetto, H., Kacprzyk, J., Hernandez, J. E., & Alemany Díaz, M. M. E. (2020, May 1). Agri-food 4.0: A survey of the Supply Chains and Technologies for the Future Agriculture. *Computers in Industry*. Elsevier B.V. https://doi.org/10.1016/j.compind.2020.103187
- Miranda, J., Ponce, P., Molina, A., & Wright, P. (2019). Sensing, smart and sustainable technologies for Agri-Food 4.0. *Computers in Industry*, *108*, 21–36. https://doi.org/10.1016/j.compind.2019.02.002
- Oktem, H., Eyidoøan, F., Selçuk, F., da Silva, J., & Yücel, M. (2006). Osmotic Stress Tolerance in Plants: Transgenic Strategies (pp. 194–208).
- Singh, A., Ganapathysubramanian, B., Singh, A. K., & Sarkar, S. (2016, February 1). Machine Learning for High-Throughput Stress Phenotyping in Plants. *Trends in Plant Science*. Elsevier Ltd. https://doi.org/10.1016/j.tplants.2015.10.015
- [1] Image Credit: Mona Schreiber. [2] Image Credit: doi: 10.1016/j.tplants.2017.06.009