University of São Paulo
"Luiz de Queiroz" College of Agriculture

High-throughput phenotyping via UAS: the optimization within a breeding
program and a new validation method based on simulation

Giovanni Galli

Thesis presented to obtain the degree of Doctor in
Science. Area: Genetics and Plant Breeding

Piracicaba
2020

Giovanni Galli
Agronomist

High-throughput phenotyping via UAS: the optimization within a breeding program and a new validation method based on simulation

Advisor:
Prof. Dr. **ROBERTO FRITSCHE NETO**

Thesis presented to obtain the degree of Doctor in Science. Area: Genetics and Plant Breeding

Piracicaba
2020

RESUMO

Fenotipagem de alto rendimento via VANTs: a otimização em um programa de melhoramento e um novo método de validação baseado em simulação

A Fenotipagem de alto rendimento (HTP), ou simplesmente fenômica, tem chamado a atenção da comunidade científica como uma área com potencial de aumentar a custo-efetividade e acurácia de fenotipagem. Entretanto, a viabilidade deste conjunto de abordagens ainda precisa confirmação. Neste contexto, dois grandes desafios para a seu emprego são a otimização do uso de dados (*data-to-decision*) e a validação de procedimentos para cenários específicos de seleção. Nós acrescentamos a este tema reportando resultados de dois estudos que objetivaram a otimização e validação de HTP para experimentos a campo baseada em veículos aéreos não tripulados (VANTs). No primeiro, apresentamos uma prova de conceito usando dados de sorgo granífero com o objetivo de identificar quando os dados de HTP devem ser coletados e como devem ser processados para a otimização da predição de dois caracteres de importância agronômica, produtividade de grãos e sanidade de planta. Nossos resultados sugerem que não há incremento da capacidade preditiva quando múltiplos índices vegetativos e voos são combinados. Adicionalmente, um único índice e voo pode ser usado para predizer ambas características sem perda expressiva de acurácia. No segundo, apresentamos uma nova ferramenta para validação de abordagens de HTP baseadas em imagens aéreas com uso de simulações de computador. A ferramenta foi exemplificada com um estudo de caso de mensuração de altura de plantas em milho. Nossos resultados sugerem que os experimentos gerados *in silico* podem ser adequadamente reconstruídos com algoritmos de estrutura-por-movimento usando imagens renderizadas, permitindo a realização de inferências sobre os fatores testados. Este estudo também trouxe novos conhecimentos sobre o efeito de fatores experimentais sobre a acurácia da mensuração de altura de plantas usando HTP. Por fim, acreditamos que nossos resultados permitirão a compreensão mais profunda da prática da HTP, auxiliando os melhoristas na busca por uma seleção mais confiável e custo-efetiva.

Palavras-chave: Fenômica, In silico, Imageamento aéreo, Sorgo, Structure-from-motion

ABSTRACT

High-throughput phenotyping via UAS: the optimization within a breeding program and a new validation method based on simulation

High-throughput phenotyping (HTP), or simply phenomics, has drawn the attention of the scientific community as a field with the potential to increase phenotyping cost-effectiveness and accuracy. Nevertheless, the feasibility of this set of approaches is yet to be confirmed. In this sense, two major challenges to its application are optimizing the data-to-decision process and the validation of procedures and pipelines for specific selection scenarios. We add to this matter by reporting on two studies aimed at the optimization and validation of field HTP based on unmanned aerial systems (UAS). In the first, we presented a proof-of-concept investigation using a grain sorghum dataset with the intent of identifying when HTP data should be collected and how it should be processed for the optimization of prediction of two major traits, grain yield and plant health. Our findings suggest that there is no predictive ability increase when combining multiple vegetation indices and flight dates. Additionally, a single index and flight can be used for predicting both traits without expressive accuracy loss. In the second, we presented a new tool for validating aerial image-based HTP approaches with computer simulations. The approach was exemplified with a comprehensive study case of plant height estimation in maize. Our results show that the *in silico* experiments could be adequately reconstructed with structure-from-motion algorithms using UAS-like rendered images, enabling inference-making about tested factors. This study also brought new insights into the effect of experimental factors over the accuracy of plant height assessment using HTP. At last, we believe that our findings allowed the promotion of a deeper understanding of the HTP practice, enabling breeders to work towards a more reliable and cost-effective selection.

Keywords: Phenomics, In silico, Aerial imagery, Sorghum, Structure-from-motion

1. INTRODUCTION

Genetic gain is the rate of performance increase achieved by artificial selection per unit time. This key concept of plant breeding is defined as a function of selection intensity, selection accuracy, genetic variance, and time per selection cycle (Araus et al., 2018). High-throughput phenotyping (HTP), also known as phenomics, has taken the spot as a field with potential to act over these parameters, enabling breeding programs to increase the rates of gain (White et al., 2012; Araus and Cairns, 2014; Araus et al., 2018). Among the realm of available HTP devices, special attention has been paid to unmanned aerial systems (UAS) due to its relatively low cost and ability to carry out large-scale screenings at high temporal/spatial resolution. Nevertheless, its effectiveness still needs proof. Therefore, new approaches, pipeline optimizations, and proof-of-concept studies are highly valuable for the improvement and feasibility assessment of UAS-based HTP, producing results of immediate and direct benefit to the scientific community.

One of the major challenges for the effective application of HTP is overcoming the data-to-decision barrier. Recent technological advancements have made data collection no longer a bottleneck as the high temporal/spatial resolution of HTP is capable of generating vast amounts of data (Araus and Kefauver, 2018). Nevertheless, given that the set of phenotypes of an organism is highly conditional and dynamic across its life cycle (Cobb et al., 2013), identifying meaningful genetic signals from data and efficiently translating the information to target traits is key (Araus and Kefauver, 2018). One approach towards optimized phenotyping is by finding the most informative phenological stage(s) and high throughput traits for a given selection purpose. This could work towards overcoming the technical data-to-decision difficulty enabling the application of truly high-throughput phenotyping.

Another key challenge to the application of HTP in plant breeding is the validation of procedures and pipelines for specific selection scenarios. In this sense, controlled and representative experimental conditions are required for meaningful and generalist inferencing. However, these conditions are sometimes not easy to achieve. A usual way to bypass the limitations of real datasets is by implementing simulations (Barth et al., 2018; Müller et al., 2018). This could provide the process of validating HTP methodologies with adequate conditions, potentially improving the reliability of estimates. Additionally, by employing simulations into the HTP context, one would be able to test new methods, experimental setups, measure the potential of a given approach, and dimension equipment, all of this without performing any field experiments or acquiring phenotyping gear. However, to this date, no report on the use of simulations for HTP validation has been made.

In the following sections, we report on two studies aimed at validating and optimizing HTP of field experiments. In the first, we presented a research as to when UAS imagery data should be collected and how it should be processed for the optimization of the phenotyping for grain yield and overall plant health (based on anthracnose scores) in grain sorghum. In the second, we proposed a novel way of designing and validating HTP approaches based on aerial imagery with computer-generated tridimensional experiments. This new tool is exemplified with a comprehensive case study of plant height estimation in maize under contrasting experimental setups in the context of plant breeding.

2. GENERAL CONCLUSIONS

We reported on two studies regarding the optimization and validation of UAS-based HTP. Our findings indicated the possibility of optimizing HTP based on the identification of the most informative flight dates and secondary traits. The results suggest that using a single flight date (119 DAS) and secondary trait (NDVI or MSR) can maximize prediction accuracy for both grain yield and plant health (anthracnose resistance) in sorghum hybrids.

Furthermore, we proposed a novel way to validate and design UAS HTP pipelines based on computer-generated simulations. Our results showed that synthetic experiments can be adequately reconstructed from rendered images using structure-from-motion algorithms. Using this tool, we inferred that treatment variance, experimental repeatability and the chosen percentile of points from dense clouds affect the accuracy of HTP.

REFERENCES

Araus, J.L., and J.E. Cairns. 2014. Field high-throughput phenotyping: the new crop breeding frontier. Trends Plant Sci. 19(1): 52–61. doi: 10.1016/j.tplants.2013.09.008.

Araus, J.L., and S.C. Kefauver. 2018. Breeding to adapt agriculture to climate change: affordable phenotyping solutions. Curr. Opin. Plant Biol. doi: 10.1016/j.pbi.2018.05.003.

Araus, J.L., S.C. Kefauver, M. Zaman-Allah, M.S. Olsen, and J.E. Cairns. 2018. Translating High-Throughput Phenotyping into Genetic Gain. Trends Plant Sci. 23(5): 451–466. doi: 10.1016/j.tplants.2018.02.001.

Barth, R., J. IJsselmuiden, J. Hemming, and E.J. Van Henten. 2018. Data synthesis methods for semantic segmentation in agriculture: A Capsicum annuum dataset. Comput. Electron. Agric. 144(December 2017): 284–296. doi: 10.1016/j.compag.2017.12.001.

Cobb, J.N., G. DeClerck, A. Greenberg, R. Clark, and S. McCouch. 2013. Next-generation phenotyping: requirements and strategies for enhancing our understanding of genotype–phenotype relationships and its relevance to crop improvement. Theor. Appl. Genet. 126(4): 867–887. doi: 10.1007/s00122-013-2066-0.

Müller, M., V. Casser, J. Lahoud, N. Smith, and B. Ghanem. 2018. Sim4CV: A Photo-Realistic Simulator for Computer Vision Applications. Int. J. Comput. Vis. 126(9): 902–919. doi: 10.1007/s11263-018-1073-7.

White, J.W., P. Andrade-Sanchez, M.A. Gore, K.F. Bronson, T.A. Coffelt, et al. 2012. Field-based phenomics for plant genetics research. F. Crop. Res. 133: 101–112. doi: 10.1016/j.fcr.2012.04.003.