Skoltech
Skolkovo Institute of Science and Technology

Jury Member Report – Doctor of Philosophy thesis.

Name of Candidate: Dmitry Shadrin

PhD Program: Computational and Data Science and Engineering

Title of Thesis: Data-driven modeling of plant growth dynamics in controlled environments

Supervisor: Professor Maxim Fedorov

Name of the Reviewer:

I confirm the absence of any conflict of interest	Signature:
	David Reis
	Date: 23-09-2020

Reviewer's Report

The Ph.D. Thesis by Dmitry Shadrin proposes a variety of complementary ICT techniques and solutions to support the increase of productivity and sustainability of precision agriculture (e.g. based on hydroponic techniques). Thus, the potential social and economic impact of the research is very high. Special attention is devoted to plant growth monitoring and prediction in controlled environments and to the use of machine learning techniques applied to different datasets and in multiple contexts.

The technical content of the Thesis is broad, as it ranges from agronomy and environmental science issues to information technology subjects. Both theoretical and implementation aspects are considered. The amount of experimental work is impressive, which confirms the scientific maturity of the research.

The Thesis structure is well balanced and consists of six Chapters, covering complementary aspects of the research. The writing style is quite good, but not always clear. The whole thesis should be carefully checked and revised to fix a large number of minor typos, English mistakes or improper sentences. Some paragraphs and explanations are verbose and entangled. As a result, the Thesis is not always easy to follow. Also, the style chosen for citations is very weird and it should be changed. The Author refers indeed to other papers

using a format like "Name Surname [Year]" in the middle of sentences. However, in this way the Author continuously interrupts his writing, which is quite annoying for the reader. I recommend that the Author changes the style of all references, including Authors' names between brackets (e.g. [John Smith 2001] or simply [Smith 2001]) and moving them to the end of the paragraph they refer to. This would avoid breaking sentences, making reading smoother.

The Introduction is well focused on research objectives and on the underlying motivations. The second Chapter about the Background provides an exhaustive overview of the state of the art, as confirmed by the large number of references, which is definitely adequate for a Ph. D. dissertation. Maybe some Sections (e.g. Sections 2.2) should be broken into smaller, more focused subsections as the current ones are very long and it is not so easy for the reader to extract the novelty of the proposed work out of such a large amount of information.

Chapters 3 and 4 present probably the main technical and scientific contributions done by the Ph.D. candidate. The study about plant growth monitoring and prediction is presented thoroughly with plenty of details about the adopted techniques, the experimental setups and the most relevant implementation issues. In all cases, the research methodology looks sound. The experimental results confirm the research goals and are presented rigorously. As a possible improvement, I think that the Author should highlight more clearly why different kinds of models are used to assess plant growth dynamics as well as the advantages and disadvantages of the investigated techniques so as to guide other researchers towards the best option for a given scenario. For instance, it is not so clear why or under what conditions the Kalman filter-based models are preferable to data-driven approaches. Maybe a final table summarizing pros and cons of different techniques (to be added in the Conclusions) could help. In general, it is not straightforward for the reader to understand quickly which techniques are preferable in different contexts and what their requirements and limitations are. I also found a bit strange that a Kalman filter is used not really to estimate the state of growth of the plant (which is indeed dynamic), but the parameters of the model (i.e., S_{max} and μ that look static or quasi static). Indeed, not surprisingly the system matrix is the identity matrix.

Chapter 5 presents additional research achievements in precision agriculture. In essence, the Author applies various machine learning techniques to two crucial problems, i.e. water quality assessment and specific pollutant detection. The study is rigorous from the statistical point of view. However, this part of the Thesis lacks generality, as it seems to be driven by too specific case studies. Thus, the generality of results and conclusions is questionable. More specifically, I appreciate the great work done by the Author to investigate such complex environmental issues as those mentioned above, by using machine learning techniques. However, the replicability/applicability of such techniques on a large scale in conditions and contexts different from those of the case studies presented can be hardly assessed due to the large variety and variability of parameters and factors that may affect this kind of analysis. Maybe the Author could try to add one or more Sections where the aforementioned problems are formalized in more abstract and general terms, e.g. providing a list of the quantities that in general must be considered for monitoring, prior to focusing on the reported case studies.

Ultimately, I think that the Ph.D. Thesis by Mr. Dmitry Shadrin is an extensive research work in a challenging and interdisciplinary research area. The overall quality is very good. The issues highlighted in the report are mainly suggestions and they can be addressed within a short time. Globally, the Thesis is broad and deep enough to admit the Ph.D. candidate to a formal thesis defense. The high quality of the research is further confirmed by the Ph.D. Candidate's publication records, which includes 3 well-known international peer-reviewed journal papers and 5 international conference papers as a first-author, plus 5 further journal and conference publications as a co-Author.

☒ *I recommend that the candidate should defend the thesis by means of a formal thesis defense*

☐ *I recommend that the candidate should defend the thesis by means of a formal thesis defense only after appropriate changes would be introduced in candidate's thesis according to the recommendations of the present report*

☐ *The thesis is not acceptable and I recommend that the candidate be exempt from the formal thesis defense*