



Plant engineering and systems biology

What we can probably learn from physics and engineering

An essay by Matthias Eberius, LemnaTec

LemnaTec GmbH
Matthias Eberius
Schumanstr. 18
52146 Würselen
02405 / 4126-0
matthias.eberius@lemnatec.de
www.lemnatec.de

For some of the older biologists, especially plant physiologists, it may just be a kind of déjà vu on a higher level, but for biological research in general a comprehensive assessment of plant growth and behaviour under defined conditions – using image processing complemented with genetic and biochemical measurements – opens up a whole new window to quantitative plant physiology, which may be one aspect or definition of systems biology.

The following essay shall highlight some aspects of the very successful reductionistic approach of science and how it can in our days – using state of the art science technology – be transferred to a new exciting level.

Plants and what distinguishes them from machines

Machines are designed by man. In many cases the design is based on empirical rules, but with increasing knowledge and computing capacities the options to model processes and constructions of machines based on physical models and physical laws are enlarged. In cases where structures are too complex to base the models directly on physical laws, numerical models are used, comprising up-to-date insights into materials and experiences. Without the superior knowledge of the humans that employ the machine, no machine would work for long.

Plants just exist as extremely complex functioning entities. They are not dependent on humans having insight into how they function. As a result from millions of years of evolution they have developed a high variability and complexity, depending on adaptations needed to survive in specific environments. While even complex machines show distinct and limited levels of complexity and interaction, in the post-genomic era man still has very restricted access to the construction plans of plants.

Machines have become more complex over time and with progress, but understanding plants is and has always been a complicated matter. From this point of view it is very significant and admirable what biological scientists have been able to elucidate about the function of plants during the last centuries.

What plants have in common with complex technical structures

When comparing plants with chemical production lines or physical machines, it becomes evident that plants are much more compact and complex and show an incredible flexibility and efficiency. Despite these differences they have to follow the same physical rules as all matter in the universe.

The explicit understanding and definition of these rules and laws are the result of the long-term efforts of physicists to understand what keeps atoms and the universe together. Incessantly making models and proving these models in bigger and bigger experiments leads to a quite sound but necessarily open set of theories and explanations for observed phenomena. Chemists increasingly learned how atoms can react and interact with each other, so that they are now able to use computational simulations based on physically based theories to model, calculate, predict and prove theories and observations.

Starting from very empirical first substance transformation experiments (fire, brewing alcohol), a chemical engineering has been developed that can now simulate and calculate the enormous mass of processes taking place in huge chemical plants. The same applies to construction engineers, who can now prove why (most of) their buildings last over long periods. This approach is far beyond the simple intuition and trial-and-error method of our ancestors.

In all cases scientists and engineers strive to reduce the amount of empirical settings in their models and base them as far as possible on a minimum number of physical laws and models.

This has always been a step-by-step process and it has made sense to start with the more general laws and models and then refine them over time wherever possible. As all these models have a mathematical basis, in any case a quantitative description makes sense to grant a better understanding.

Even though they are complex, it is obvious that plants have to follow the rules of conservation of mass and energy. As a result, any plant-changing process can be described based on sources, transformations and sinks of matter and energy. The inevitable charm of this kind of description is that it is mathematical on the one hand, but can on the other hand handle complex or poorly understood processes as mere black boxes, describing only input and output. While the burning process of a rocket engine is very complex if you take all spatial and chemical processes into account, it is quite easy to describe generally how fuel is transformed into CO₂ and water, setting free a great deal of energy. Even if we do not understand photosynthesis completely, we can develop a quite sufficient mathematical model on the transformation of light and CO₂ into sugar etc. The more we know about the single processes, the better we can model them.

What physicists do, engineers like and biologists may consider

Both physicists and engineers like to increase their knowledge, but with different consequences. While physicists have used proven theories and empirical observations to develop new models (and even bigger new experiments), engineers apply their knowledge to change things by providing better machines. Biologists did extremely well over the last centuries to describe additional details and general concepts of extremely complex biological systems. They needed to do this with basic empirical theories and models often not physically based. They never had a chance to start with simple physical models, reflecting the actual physical state of science, as their reality – life – had always been complex and needed complex models for any kind of understanding. But some things are changing now more than quantitatively.

Physical science and chemistry have reached a degree of development that should allow them to start describing plants with more physical models than ever before. All knowledge on genetics grants us now at least an insight into how mass and energy transfer is regulated. Computational and storage capacity make it possible these days to acquire data and calculate biological models and results complex enough to make sense in the biological way. It may sound strange, but building the first atomic bomb and flying to the moon was possible with a much lower computational capacity than just taking daily

images of 1,000 plants, putting them in a database and analysing their watering needs on account of hanging leaves. As these capacities are now available, biologists should develop concepts and apply them, learning as much as possible from the methods and experience of the physicists to get a better understanding themselves: Better understanding is the first step to better engineering.

As man started engineering plants long before thinking about cars or flying to the moon, it is rather obvious and urgent to continue this process with the tools of the 21st century. While different sciences developed separately from different origins for good reasons, we have now the chance and call to integrate them again. Instead of focussing on different sciences, we should take the chance to focus the best of all sciences on a specific aim: in this case plant sciences and plant engineering.

Maybe systems biology – having various definitions – may just mean this as well.

What can we probably learn from physics

When systems are complex to analyse, – like the universe or the elementary particles of atoms – huge collaborative projects like the Hubble telescope or the particle accelerator in Geneva (CERN) are initiated. In contrast to widely used biological approaches, having collaborations and a general concept can be very efficient, at least for a larger research unit with optimised work share, competence and data. The decoding of the human or Arabidopsis genome showed very impressively how big projects may produce vast amounts of data, providing a wealth of information to be analysed and understood even much later and without those researchers involved in the original experiment.

Such cooperative projects illustrate that a tendency actually started in physics has now expanded successfully to biological applications, and both have some important aspects in common:

1. Large number counts are needed. If you are looking for rare events and extraordinary occurrences or you need to get a survey on large, more or less heterogeneous populations: The number often makes the difference. Large numbers of plants have to be screened to establish correlations between genes and phenotypes or genes and reactions on external stressors. Furthermore, such information is needed to prove the reliability of theoretical models that have now reached a high complexity.
2. The infrastructure needs to be shared by many researchers. As tremendous efforts are necessary to establish the infrastructure around big screening experiments, it must become standard that specific institutions provide the infrastructure and scientists just bring their core issues there (e. g. their seeds). It makes much sense to establish a technical infrastructure and train people to maintain it, to fill thousands of pots, randomise them in greenhouses and water them according to the specified needs of the experiment. Having such highly skilled people at hand helps tremendously to perform research fast, efficiently and professionally. Being a good scientist does not necessarily imply being a good dealer and optimiser of logistics for large-scale experiments. (and this method of shared resources would surely enable researchers to concentrate on their intended work).

3. Detection systems should be handled professionally: The wealth and power of detection systems is the people who have the long-term experience to extract maximum information with “their” detection systems. This refers equally to all levels of research: from first screening detectors to highly specified detail analysts of single plants or tissues.
4. Reproducible quantitative data are needed: While biology was incredibly successful with qualitative assessment over the last hundreds of years, we are now at a stage where quantitative assessment of many kinds of data is possible and needed to feed models and statistics and to understand complex interactions. It is one of the big accomplishments of mathematics to allow handling of e. g. 7 parameters interacting while some of them include logarithmic or more complex interactions. There are few people who can work with such complex relations by just thinking about them, but when having the appropriate mathematical toolbox available, many more can handle and understand for example plant adaptation processes based on minimising (e. g. energy on a three-dimensional surface).
5. All data should be stored as comprehensively as possible. When starting an experiment, experience and theory may predict many parameters that should be monitored. As plants are complex, we should remain open for the unexpected. This means keeping raw data like images at hand for any later analysis or reanalysis. This approach is not very common in biology yet, but in the future there will be increasing numbers of biologists who do research by analysing existing data and not by doing their own experiments. Other faculties like physics have already done this successfully for quite a long time. Doing research on the hard drive instead of in the greenhouse can save costs and time and can be an ideal basis for the planning of new experiments without the need to repeat older ones, thus at least in part allowing researchers to handle raw data homogeneously.
6. Data should be stored in easily accessible databases. The more researchers have access, the more scientific results can be got out of experiments. This increases the efficiency of experiments dramatically and also reduces costs. In many cases data for proving new hypothesis could be available if test data – e. g. images of growth etc. – had been comprehensively stored and kept for reanalysis, and if future scientist would have access to such raw data.
7. Data mean nothing without the appropriate models. Producing huge amounts of data is one issue, being able to use them another one. To understand complex systems, quantitative data and models based largely on basic physical laws are needed. Where complexity is too high and too many matter properties need to be considered, biology can learn tremendously from other sciences such as chemical engineering, material science and geophysics. In the end everything is about transformation of mass, energy and information. Due to the historically based low affinity of biology to mathematics and physics, there may be many obstacles on the way, but on any field of plant sciences efforts should be taken to develop quantitative models and to interpret results on the background of quantitative systems analysis. Having data of large numbers of real plants available on your hard drive is the best basis to provide important facts about phenotypic parameters and variabilities, thus modelling plants completely in silico for a better

understanding and prediction of reaction patterns to genetic or environmental conditions.

What LemnaTec can do to help researchers to reach their challenging objectives

1. Both LemnaTec HTS and Scanalyzer 3-D systems with shorter or longer conveyor belts enable researchers to assess data for hundreds or thousands of plants every day, closely following their development.
2. Due to the supreme flexibility of the LemnaTec systems, they can be used parallel or in sequence by different research projects, as all parameters underlie complete control of the database-defined methods. Thus any imaging condition can be reproduced just by selecting a specific configuration from the database. An investment in LemnaTec systems is really an infrastructure feature (infrastructural advantage?) where different researchers can perform their specific experiments with only a minimal need to establish either imaging or data handling and analysis structures.
3. The LemnaTec systems can store the experience of the research projects in the shape of successfully used methods, allowing later projects to profit from their predecessors. Thus researchers can focus more on their specific interests instead of on the technology they need.
4. LemnaTec systems always perform exactly as defined by the stored methods, thus allowing the assessment of reproducible quantitative data. By later reanalysis it is possible to analyse all data homogeneously, even if new aspects appear during or after the primary data acquisition process. Extracting multiple parameters from individual plants allows for the development of a comprehensive quantitative fingerprint of each plant, which can then be related to all other conditions and data of the experiment.
5. As LemnaTec systems even allow storage of images taken at various wavelengths and under specific conditions, these images offer the most comprehensive databases for any further analysis of raw data. Keeping reanalysable images is the most basic approach to conserve raw data with a minimum reduction and loss of information, being open to many different aspects for later research. Having such data enhances the power and efficiency of the research to answer questions by first reanalysing existing experiments and then designing more specific new ones only if necessary.
6. The LemnaMiner concept allows the experimenting scientist to access data without specific database knowledge right after the analysis is completed, which is generally before all plants are brought back to the growth chamber or greenhouse. Even tedious calculations between measurement points (like with growth rates) can immediately be performed. Combined with a predefined classification or hit scheme setting specific threshold values for parameters in the database, this enables the researcher to react on measurement results almost at once, thus performing for example detailed biochemical or genetic analysis on plants showing specific phenotype parameters. Additionally, supervisors can easily get a survey on current

results without needing to wait for data analysis. And the LemnaMiner provides data mining tools for the creative playing around with data still directly linked to the images of the experiment.

7. The LemnaLauncher data analysis methods may produce a wealth of data without additional work, after being defined once in an overall configuration. More data may even be available by later reanalysis. Such wealth of quantitative data assessed homogeneously for all plants is the best basis for any high-level statistical analysis of phenotyping data, as all and not only the most outstanding plants are assessed comprehensively. Having such datasets is an excellent basis to feed in silico experiments with experimental data or – in later stages – prove the suitability of in silico models by using data from real experiments.

Some Conclusions

To proceed in plant research and engineering, changes of paradigms are necessary. Plant scientists need to learn more than ever how to adopt appropriate models from other fields of science and engineering. There is such a wealth of experience available that being curious about it will push plant science forward dramatically. This will soon result in more insights into which kind of data are needed most to develop better models, allowing to understand and thus set the basis for directed modification of plants. Initial projects on a large scale already suggest that quantitative high throughput screening combined with in-depth analysis of prominent representatives will increase our understanding dramatically. In all kinds of projects it should be taken into consideration that plants underlie mass, energy and information transfer as much as technical systems do. Including the physical perspective into plant science at all stages will open up new perspectives about what should be measured and analysed in more detail.

Only excellent understanding is the basis of directed development and acceptance and trust by the public. Very few people know how to calculate the statics of a bridge or can explain why a 300-ton airplane can fly, but there is huge trust in the engineers that their models are good enough for the real things to last and fly. Even while seeing that many more or less rationalised arguments against chemical engineered plants are really far away from any scientific basis, scientists need to understand the public's mistrust against the results of plant engineering. One point may be that plant science has not yet been able to communicate (or possess) a broad based theory and a set of widely accepted and proven models to describe what is actually going on. Engineers and physicists have showed massive proof of concept by bringing man to the moon, probes to the mars and building reliable airplanes and bridges.

Plant science still has to earn this trust of the public. This is an enormous challenge, as it looks like producing a genetically engineered plant resistant to diseases, using 50 % less water and giving 30 % higher yield is more difficult than flying to the moon.

Nevertheless, we should face this challenge now and ask any branch of plant research how it can contribute to produce, prove or falsify quantitative plant modelling.