EIP

AI and Big Data in Agritech

First published by Process Engineering - February 2024

The world's population stands above 8 billion in 2024 and is still growing. That growth may be slowing down but nonetheless there may be another 2 billion mouths to feed in the next 50 years. Furthermore, it is predicted that the expected improvements in crop yields in that same time frame will be offset by increased disease pressure resulting from climate change.

The agricultural industry is increasingly turning to data-driven approaches to give them every advantage. What are researchers doing to take advantage of computers in the crop research and agritech space?

Bioinformatics

Computational genomics and other aspects of bioinformatics are incredibly important in understanding plant genomes. Typically plant research is performed on Arabidopsis thaliana, a model organism with only a small diploid genome of around 135 million base pairs which makes genetic analysis more straightforward. Unfortunately, crop plants tend to have much larger and more complex genomes. Bread wheat has a hexaploid genome of around 17 billion base pairs. Furthermore, the high percentage (close to 80%) of repetitive sequences makes gene sequence analysis more difficult. By comparison, about 50% of the human genome is repetitive sequences and that is on the higher side for mammals, let alone other eukaryotes. As such, developing new varieties and new transgenic crop plants can be very difficult.

This difficulty is enhanced in jurisdictions such as the EU where GMO crops are not authorised for human consumption; so bioinformatics techniques have to be used to enhance conventional crop breeding techniques rather than making direct modifications of the plant genome. Elsewhere, though there are sign of change: the UK passed the

Genetic Technology (Precision Breeding) Act for England last year – reducing the regulatory burden breeders face when making genetic modifications that could have occurred naturally through conventional breeding, but not introducing exogenous genes from other species. Similar legislation is being considered in the EU.

p2

Microbiome analysis is also a big part of the bioinformatics space. Rothamsted Research recently received a £4 million award for new computer infrastructure which will help analyse soil metagenomics, essential for understanding the soil in greater details as many of the microbes in soil cannot be readily cultivated in a lab.

Considering the patent side, we can see that many of the applications filed relate to the application of machine learning approaches to genomics or disease detection. We also see filings for prediction systems to optimise the chances of breeding success based on genomic data, whether via AI techniques or more classical statistical methods. Analysis of pathogens is common, with looking for pathogen DNA or identifying pesticide resistance genes.

~~Crop Modelling~~

Crop modelling continues to improve and develop, as it has for decades, but there are issues with lack of open data. Crop simulation has been developed since the 1960s but a renewed interest in the field has arisen from the improved sources of data (such as from remote sensing), the increasing demand from population growth, the impact of climate change, and the application of machine learning.

Looking at patent filings, development of these technologies seem to be in the application of machine learning but nonetheless, we still see filings for improvements to more classical state machines.

Interestingly, the majority of filings are in China and the US. The former is to be expected given the domestic patent policy in China but it is a little surprising to see hardly any PCT 1.applications being taken forward in Europe. This may be due to the understanding of the European Patent Office's requirements that computer software must have a technical effect in order to be patentable. As such, more abstract processes are less likely to be patentable in Europe. However, it should be understood that simulation can definitely be patented in Europe with the right claim wording, provided of course the invention has an inventive step and meets the other patentability criteria.

Remote Sensing

Remote Sensing generally refers to the collection of information about something from a distance but typically refers to the usage of Earth observation satellites and camera

drones. Some satellites are government-owned such as the European Space Agency's Sentinel satellites. Others are owned by private enterprise, such as Airbus's Pleiades Neo satellites.

Remote sensing from satellites has been used for decades to measure crop productivity. It has also been employed to detect and map plant diseases, which often alter the appearance and structure of the plants in such a way that is detectable through multispectral data; for example from yellowing leaves or from wilting crops altering the angular polarisation of radar scans.

Satellite technologies are also relevant to livestock monitoring. Albeit satellites are less effective at this given animals tendency to wander more than plants and because most satellites only pass over the same point at most a couple of times per day as the Earth rotates.

Great strides are also being made with aerial photography and the integration of that with satellite remote sensing. The massive proliferation of drone technology and improvement in digital camera technology allows farmers to survey huge areas with a high degree of accuracy. Drone footage not only provides a ground truth for human verification but can also be used more generally as a data source to integrate a big data approach.

Increasingly, we can expect a democratisation of this technology to smaller and smaller farmers, particularly as nearly all of the data captured by the Copernicus satellite constellation is freely available to download on the internet.

The number of patent applications filed for space-borne sensing and green applications has been rising rapidly in recent years, as shown in an EPO, ESA, ESPI joint study released in October 2022. The main developments that are being filed for are in signal processing on the software side, and AI/ML seems to be the core driver for this. Analysis and modelling of photosynthetic activity and crop productivity are also frequently mentioned in these applications, as well as weather and river/coastal observations.

In summary, agritech has always been an area of significant technological advancement but data driven technologies are becoming increasingly relevant in our interconnected world.

1. Patent Cooperation Treaty or 'international' patent applications offer simultaneous protection in a number of signatory states