Project number: 6756
Funding source: Enterprise Ireland
Date: March 2017
Project dates: Sept 2014 – Feb 2017

Advanced predication for production and processing

Key external stakeholders:
Dairygold, ICBF and Glanbia

Practical implications for stakeholders:

- Long, medium and short term models were developed to help with milk prediction
- Prototype models developed when deployed will allow
  - Processing capacity planning to be more robust
  - Product Market planning to be more robust
  - Processing planning to be more robust
- The models will help farm financial planning through helping to plan milk deliveries

Main results:

- The model was between 95% and 96% accuracy at predicting peak milk production across two processors (Glanbia and Dairygold) when trained with historical data
- The model had between 94% with 98% peak supply identification accuracy and peak volume prediction with a 10 day forecast for one processor and for the other processor the corresponding numbers were between 98% and 99%
- The model had between 99% with 95% peak supply identification accuracy and peak volume prediction with a 20 day forecast for one processor and for the other processor the corresponding numbers were between 99% and 95%
- The model had between 98% with 89% peak supply identification accuracy and peak volume prediction with a 30 day forecast for one processor and for the other processor the corresponding numbers were between 93% and 90%

Opportunity / Benefit:

With the cessation of the milk quota system in 2015, producers and processors in the Irish dairy supply chain are challenged with a new business paradigm based on the global milk price and a free market approach to production and pricing. As the producers strive to maximise profitability followed by output, a key issue for processors is how to reliably plan transport, refrigeration, processing and finished product storage capacity. This is a particularly challenging (and costly) problem for the processors there is significant variance in milk production from one year to the next as well as the presence of significant short term (weekly) fluctuations in supply throughout the year.

A further challenge arises for the processors in seeking to increase profitability from the supply chain by directing milk with differing qualitative properties to the optimum processing facility. If production of milk with these qualitative characteristics could be accurately predicted in a reliable mathematical model, then milk with the relevant macro nutrients could be directed efficiently to the most suitable processing facility and thus optimise its profitability.

Contact Laurence Shalloo Email: Laurence.Shalloo@teagasc.ie

Collaborating Institutions:
Teagasc project team: Dr Laurence Shalloo (PL)
External collaborators: Eric Robson WIT,
Philip O Brien WIT,
Michael Murphy CIT,
John Fitzgerald Glanbia,
Liam O Flaherty Dairygold,
Sean Coughlin ICBF,
Mike Lynch ICBF

1. Project background:

The SMART APPI concept has evolved from discussions between Teagasc, TSSG, Glanbia and Dairygold Ltd. The publication of the food strategy document, "Food Harvest 2020" published by the Department of Agriculture, Food & the Marine, also gave the concept some background. Specifically, Food Harvest 2020 strives for a 50% increase in milk output. The dairy processors, Glanbia and Dairygold require to plan how they are going to deal with this expansion and were searching for predictive models for milk production that could accurately forecast and help manage fluctuations in the supply chain associated with the new production regime post quotas. This predictive problem exists both within and between years and creates challenges for both strategic and tactical planning at processor level. Teagasc had developed a model that was used by Dairygold and Glanbia for one to three year planning for milk. This model had also been implemented within the ICBF database and at individual processor level, while at the same time, Teagasc in conjunction with CIT had developed a dynamic model to predict short term fluctuation in milk supply. Through the development of these models, it was realized that there was the potential from a research program in this area. TSSG-WIT, with its expertise in data mining from multiple data sources, could provide the research and development capabilities required to build a software platform that would utilize predictive analytical engines, based on the Teagasc/CIT model backgrounds. As well as building new modelling frameworks for advanced data mining algorithms to deliver an accurate, reliable and easy to use digital modelling systems. This newly developed digital hub could then be used for future modeling research without the requirement to build the database architecture.

2. Questions addressed by the project:

- Developed a systems architecture for the deployment of milk prediction models for the Iirsh dairy industry.
- Support the development of robust algorithms with real time processing of multiple data input sources for site specific milk predictions on a short, long and medium term basis.
- Deploy a functional, robust and accurate milk supply chain variability prediction platform based on the robust algorithms developed that utilises advanced data management and predictive analytics to produce business optimisation information for both processors.
- Build into this prediction platform the ability to be expanded and enhanced through the use of new data sources in the future

3. The experimental studies:

The research team were tasked with scoping out the different model types to develop a plan for the potential models used in order to help with milk predictions. A software platform was then developed that was logically driven by these models to predict production levels based on the parameters taken from the data sources listed below. During the development phase the platform was hosted in TSSG's data centre with access available to all the project team. Long term, the hosting arrangements should be based in either the milk processors themselves or within ICBF database structure. In order for the systems to work, they will require a feed of live data on a continuous basis

The data connections required by the platform include;

- Glanbia and Dairygold – were both contributors to this project through the provision of data and through the contribution of financial support. The data included milk volume production from a random sample of producers which had firstly been anonymised, as well as including milk quality parameters such as fat and protein.
- ICBF – as the body charged with the maintenance and updating of the national herd registry, including the registry of ~80% of the animals in the state, ICBF provided data relating to the breed, output, progeny, parity, genetic makeup, EBI etc. make-up of the anonymised herds received from the individual processors.
- While meteorological data was tested in the models for its potential to increase model accuracy there

Contact Laurence Shalloo Email: Laurence.Shalloo@teagasc.ie

was no statistically significant relationship established which meant that in the final models developed climatic data was not included

However, as detailed above, the architecture for this platform will facilitate the rapid inclusion of future data sets not currently available. For instance reliable satellite imagery for more localised weather forecasting and data from sensor arrays deployed within the processors and on farms. The platform developed as part of this project will make it relatively easy to include new data at different time points in the future.

The key to success in a project like this centres on strong communication between the different parties at all times during the project. This ensures that there is no delays around data sharing and it also means that the outputs of the results are evaluated during and throughout the project.

4. Main results:

- The short term predictions require a much finer grained analytical approach and as such neural networks provide the capability to deliver high accuracy over volatile data. However before a neural network can be effectively used to predict values, a model must be fitted to the data by training the network with a significant portion of historical real data. For this scenario here a portion of the data was used to train the network and fit the model. The model accuracy was then verified by inputting the remaining data and comparing the resultant prediction with the real data for that time period. This mechanism proves (to the extent possible with this data set) that the results are not over fitted against the input data.
- The combination of the Autoregressive Moving Average (ARIMA) models and Long Short Term Memory Networks –Recurrent Neural Network (LSTM RNN) models allows us to accurately predict the peaks for the coming year, and then throughout the year, update the model with current data and accurately forecast in 10-30 day windows. Given the current state of the art of on farm data recording and management – the accuracies achieved in this research is substantial.

5. Opportunity/Benefit:

The problem for the processors and the opportunity for the proposed platform is the prediction of quantitative and qualitative milk production levels so as to enhance the dairy processors profitability in the face of global competition, while at the same time ensuring that there is adequate processing capabilities for the increased milk supply as well as ensuring that the investments made in additional processing capacity are appropriate and timely. Where this global competition is becoming most challenging is in the securing of large scale, long term contracts by Irish dairy processors in emerging markets, such as the Chinese market. Planning for the increased supply will ensure that the maximum returns can be achieved through the development of markets thus ensuring the maximum returns can be achieved for the dairy industry in Ireland.

6. Dissemination:

Open days:

Presented at Teagasc ICT Conference held at the Aviva Stadium in February 2016 and at the Moorepark Open Day as well a presentation at the National Ploughing Championships of the overall concept of the project.

Industry consultation

Dairygold and Glanbia are industry partners on this project and have had direct input to the overall project and have provided data for the model development process.

7. Compiled by: Dr Laurence Shalloo

Contact Laurence Shalloo Email: Laurence.Shalloo@teagasc.ie