18th World IMACS / MODSIM Congress, Cairns, Australia 13-17 July 2009
http://mssanz.org.au/modsim09

A database application and maturity index prediction model to improve supply chain operations of pea production

Prestwidge, D.B.[1], A.J. Higgins[1], D. Stirling[2], J. Weeks[3] and J. Yost[3]

[1] *CSIRO Sustainable Ecosystems, 306 Carmody Road, St. Lucia, Queensland, 4067, Australia*
[2] *Simplot Australia Pty. Ltd., 85 Devonport Main Road, Quoiba, Tasmania, 7310*
[3] *Simplot Australia Pty Ltd., Airport Road, East Devonport, Tasmania, 7310*
Email: di.prestwidge@csiro.au

Abstract: The harvest timing of green peas is vital in determining "softness" and retail value. "Softness" is measured as Maturity Index (MI), and increases rapidly in the final 3 to 4 days of pea maturity. This affords a very narrow time window for arranging harvests amongst the many wide-spread paddocks accessed by commercial pea processors. Achieving the ideal processed MI requires efficient prediction and communication across the field, harvest, delivery and processing sectors within the supply chain between the field and factory. Accuracy in predicting paddocks that are approaching ideal MI requires integrated information systems that allow data to be efficiently made available for planning purposes. Currently a manual method using hard copy sampling cards and communicating results to regional field officers verbally by phone; and a manual historical prediction method for daily MI gain are used. In collaboration with a large food processing factory in Tasmania, we have developed and implemented a database application and model that enables:

- electronic entry of field sampling data and linkage to the factory production databases;
- real time transfer of paddock maturity rate to regional field officers via email-to-SMS technology;
- Integration with an MI prediction model based on an Artificial Neural Network (ANN) developed using historical production and weather information; and
- efficient planning for harvest and transport operations when there are short lead times.

Figure 1 shows the flow of MI sampling and prediction data using the database application across the pea supply chain. When comparing the manual method to the ANN model for 2 day lead time predictions, the average error for the ANN was 30.3 compared to 36.6 for the manual method. The use of the ANN prediction model enables MI predictions 4 to 7 days prior to optimal harvest rather than the manual method allowing only a 2 day lead time. We show how the application and the model can be used as planning tools for forecasting production and improve planning activities in a green pea supply chain. The sampling and harvesting scheduling database system can also be applied to other crops, such as beans and broccoli.

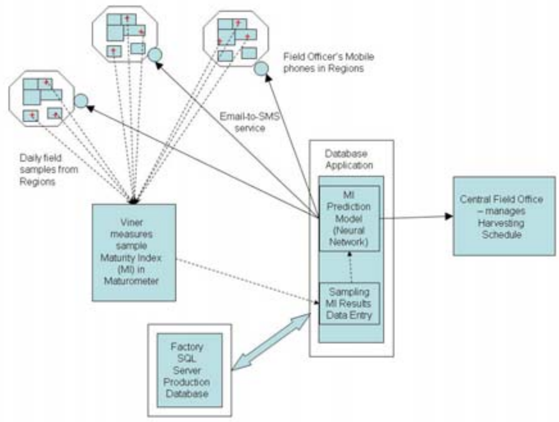

Figure 1. Flowchart of Maturity Index (MI) sampling and prediction data using the database application across the pea supply chain.

Keywords: *Database model, Field sampling, SMS, Harvest planning, Neural networks*

1. INTRODUCTION

Green peas are grown in Tasmania, Australia, because they are particularly suited to the cooler temperate regions with a high annual rainfall. There are a total 19 sub-districts across the five major growing regions, separated by differing climatic conditions and soil types. The major processing factory is located in Devonport in the North West region. The pattern of pea planting and harvesting is dictated by climate, primarily rainfall and temperature. Since the growing regions in the south are drier than the north, planting begins in the southern region around late July with a growing time of about 140 days. In the northern regions, planting begins in about October, with a growing time of about 80 days. The shorter growing time in the northern regions is also due to greater exposure to higher average monthly temperatures. Varieties used by farmers also differ between the northern and southern growing regions of Tasmania. Interestingly, while climatic conditions dictate pea growth and development, they are not currently used to predict or organise harvesting schedules.

Maturity Index (MI) is a key determinant of pea tenderness and retail value of pea products. MI is typically measured using a maturometer (described by Mitchell and Lynch, 1952) which provides a score of the tenderness of peas. The optimal MI range is between 230 and 250, and peas in this range are used in premium brand products. Peas supplied with high MIs (e.g. >300) are often used in soups, for example, which are lower value per unit mass of peas. MI is difficult to predict and rapidly increases in the 2 days prior to optimal maturity. Currently field sampling is extensively used to monitor the change in MI as the paddock approaches its anticipated harvest date. Field sampling is not always accurate enough as a prediction tool to ensure harvest occurs in the narrow optimum window. Whilst sampling itself does not provide a prediction of MI, field officers at the factory try to use the MI measures to estimate MI one or two days hence, usually by adding an historical daily MI gain factor. The accuracy of this method declines rapidly when trying to predict more than one day ahead, and the food processor would ideally like to predict MI several days ahead. This method is time-consuming; costly; has short lead times; and currently recorded via an inefficient paper-based system.

Difficulty in MI prediction is exacerbated by the farm to factory supply chain characteristics. The pea supply chains in Tasmania are comprised of several hundred farmers, geographically dispersed, contracted to supply peas to a food processor in any given harvest season. The food processor aims to harvest the pea paddocks when the MI values are close to their optimal time. The harvest season extends from November to February. Variability of MI values within a paddock is due to the time of harvest, given that MI can increase by as much as 30 points in a 24 hour period and a paddock can take more than a day to harvest. Secondly MI can vary significantly geographically within a single paddock, usually with higher MI readings from the outer rim.

With the current manual paper-based field sampling process, the field samples are measured for their MI value along with a count of flat pods and a weight of small peas. These measurements are currently recorded onto paper sampling cards that are delivered to the central field office then recorded onto a sampling sheet containing seasonal sampling results. The field officers in each of regions are informed of these results by phone. This information is manually correlated with the current and predicted climate data along with historical MI values to predict MI gain for the following 3 to 4 days. The disadvantage of this method is that it does not provide timely information back to the field officers and harvesting planners and it is only hard recorded via a paper card system. Recording information electronically will be useful in making the whole supply chain more efficient.

Consequently benefits could be gained from streamlining the MI prediction process by making it more accurate; reduce the lead times; and making it less labour intensive. These benefits would be improved infrastructure utilization; reduced sampling costs; and improved average product value. These improvements will enable a greater number of pea paddocks to be harvested at the optimal MI range. Extending the lead time provided by MI prediction enables harvesting and transport logistics to be planned with greater efficiency and lower cost. It also provides better planning in factory throughput and cold storage, particularly if the factory processes multiple types of crop.

In this paper, we describe a recently developed and implemented a database application and MI prediction model developed in collaboration with a large food processing factory in Tasmania. The application allows the electronic entry of field sampling data, which is instantly transferred from the field to the factory via email-to-SMS technology. This is essential for planning harvest & transport operations when there are short lead times. It also integrates with an MI prediction model based on an Artificial Neural Network (ANN) developed using historical production and weather information. In Section 2, we describe the development of the sampling database application. Section 3 describes the integration of harvest planning tools with the sampling database, including the construction of the ANN maturity prediction model.

2. DEVELOPMENT OF THE SAMPLING DATABASE APPLICATION

A need to improve the efficiency of the pea supply chain and enhance the maturity prediction methods has lead to the development of an application linking of efficient databases, models, and communications systems. The current methods of paper-based sample recording, communicating sampling results via phone, and manual prediction methods showed large inefficiencies. The more efficient integrated application would require the ability to link with current factory SQL Server databases; record and aggregate sub-sample data over time; and link to immediate communication systems. Field officers need to constantly monitor the maturity of paddocks from each of the growing regions. The regional field officers are responsible for making daily decisions on which paddocks to harvest and when. Email-to-SMS technology is a quick and simple method for transferring data to field officer's mobile phones whilst they are traveling around their regions.

A Microsoft ACCESS database application using Visual Basic code was developed to record the field sample measurements and electronically communicate the results to the appropriate field officers out in the harvesting regions. The database is linked to the factory SQL Server production databases. The Viner (the person measuring the samples) enters the results into a data entry form as shown in Figure 2. For each paddock sampled the Viner selects the appropriate paddock; updates the estimated tonnage; and records 4 individual sub-sample values for MI, total pods, and flat pods; as well as the weight of large and small peas. The application calculates the average MI, flat pod and small pea percentages. After all the daily sampling card results are recorded they can be immediately sent to the field officer's mobile phones via an Email-to-SMS service. Each SMS message can contain up to 165 characters which allows five individual paddock sampling results as shown in Figure 3. The SMS messages are summarized and coded to fit in enough useful information. For example:- "03/12 60452A MI202 FP52% SP33%" means that the sample taken on 03/12/2008 for paddock 60452A had an MI value of 202, a flat pod number as a percentage of the total pod number of 52% and a small pea weight as a percentage of the total pea weight of 33 %. Table 1 shows a report of all the sample results for different Field Officers for a specific day. This report is useful for the central field office to monitor maturity trends within and across the regions.

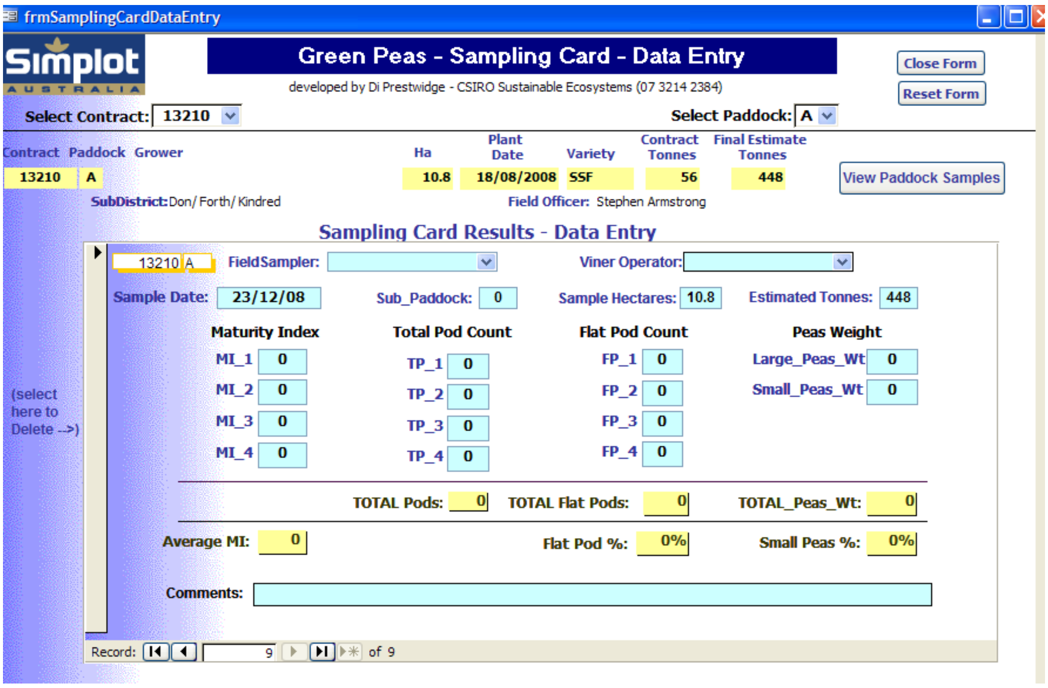

Figure 2. Sampling Card Data Entry form from the database application for entering sub-sample results.

The following data will be sent by SMS to: Jon Doust at: 0419834414 ---> 03/12 60452A MI202 FP52% SP33% _ 03/12 60202D MI184 FP35% SP40% _ 03/12 60452B MI180 FP69% SP45% _ 03/12 60452B MI180 FP69% SP41% _ 03/12 60202D MI151 FP46% SP52% _ SEND SMS? Yes/No

Figure 3. An SMS message containing 5 paddock sample results sent to a field officer's mobile phone.

Table 1. A daily sample results report for all paddocks sampled on 3/12/2008 with average maturity index (MI) value greater than 150.

	Contract_Pdk	Grower	Sample_Date	AvgMI	FlatPod	SmalPea%
Jon Doust	**Sub_District: Broadmarsh/ Plenty/ Derwent Valley**					
	60202 D		3/12/2008	151	46%	52%
	60202 D		3/12/2008	184	35%	40%
	60452 A		3/12/2008	202	52%	33%
	60452 B		3/12/2008	180	69%	45%
	60452 B		3/12/2008	180	69%	41%
Stephen Armstrong	**Sub_District: Don/ Forth/ Kindred**					
	10442 B		3/12/2008	192	12%	23%
	12608 A		3/12/2008	163	55%	28%
	13210 A		3/12/2008	186	15%	17%
	13210 A		3/12/2008	155	39%	32%
	21005 A		3/12/2008	187	29%	26%
	21005 B		3/12/2008		60%	61%

The central field office controls the order in which the paddocks are harvested depending of the availability of harvesting gangs, the maturity of the pea paddocks, and the factory tonnage requirement for processing different grades of peas. The database application provides an ability to view all paddocks sampling results for a specific period as shown in Figure 4.

									Sample_Date		
									3/12/2008	2/12/2008	1/12/2008
DistrictName	SubdistrictName	Contract	PlantDate	F	For	Variety	Ha	Tonnes	AvgMI	AvgMI	AvgMI
N-West	Don/ Forth/ Kindred	10442B	6/08/2008	G	Yes	SPANDINO	3	16	192	194	200
		12411A	6/08/2008	G	Yes	SSF	4.2	25		172	
		12608A	29/07/2008	A	Yes	SSF	5.2	29	162	152	
		13210A	18/08/2008	G	Yes	SSF	10.8	56	186 154	186	179
		Total									
Quoiba	Wesley Vale/ Sassafra	21005A	14/08/2008	A	Yes	SSF	6	27	186	165	178
		23201A	14/08/2008	G	Yes	SSF	13	78	159	151	153
		24208A	7/08/2008	G	Yes	SSF	8	43	165	162	160
		24209A	16/08/2008	A	Yes	SSF	4	18	172	161	153
		24801A	16/08/2008	G	Yes	SSF	12	70	210	213	207
		Total									

Figure 4. Sampling results from all paddock's with average maturity index (MI) values greater than 150 for 1/12/2008 to 3/12/2008.

3. INTEGRATION WITH HARVEST PLANNING TOOLS – PREDICTION OF MATURITY

With a database application in place to electronically record field information, there are now new opportunities to integrate decision support tools for planning. An obvious opportunity is to optimise harvesting and transport logistics, which allows increased lead time for allocation of harvest gangs and trucks to paddocks. These harvesting plans will ensure a more consistent flow of raw product to the processing factory along with the ability to plan for efficiencies within the cold storage facilities. In this section, we focus on a modeling approach to predicting MI, which was integrated into the database application. Currently predictions are generated at the central field office by taking an MI sample value of the paddock and adding expected MI gains per day, which are based on predicted pea heat units (described below). Once the field sample MI value reaches about 150, a prediction is generated for up to three days ahead by adding an estimated MI gain of 22 MI units per day. For example, if a field sample of a paddock had an MI of 160 units, the factory would predict MI to be 204 units two days later.

A maturity prediction model based on an artificial neural network (ANN) approach was developed using historical climate and production data. An ANN is an information processing model inspired by the way the interconnected structure of the brain processes information. ANNs are simplified mathematical models of biological neural networks. ANNs are non-linear statistical data modelling tools. They can be used to model complex relationships between inputs and outputs or to find patterns in data. A widely used ANN structure is the multi-layer perception, which was have employed in this model. It contains one input layer, two hidden layers and one output layer. Each layer employs several neurons and each neuron in the layer is connected to Neurons in the adjacent layer through various weights. The ANN was coded in NeuroSolns 5. The standard learning algorithm, back propagation, was used as the learning algorithm. The development of the ANN model approach used is described by Higgins (pers. comm.).

As part of building the ANN model, 4 years of historical MI data were used and collated into a database. This database contains information about each paddock that was processed across that time period. For each paddock processed, it contains the paddock details (farm ID, sub-region, variety, plant date) and MI details for each bin transported from the paddock to the factory during harvest (MI reading, time of harvest, time of processing, weight of peas in bin). A large amount of data pre-processing was required before it could be used to build the ANN. Firstly, extreme values of paddocks with MI of <150 or >400 at harvest were removed. The project team felt that the ANN didn't need to predict these extreme values, and there were external events affecting these MI values that could not be explained by an ANN. Also the purpose of the model is to predict the desirable MI of 200-250 rather than the very small number (<2%) of extreme values. Minor varieties were also removed (< 30 occurrences), since variety is categorical and there were not enough occurrences for the ANN training set. These filters reduced the data set by 7%.

The second part of the data pre-processing was the addition of the "cumulative pea heat units", "cumulative radiation", and "cumulative rainfall" variables. For each day of growth the pea plant is exposed to pea heat units which are calculated using the following formula: -

$$((\text{maximum temperature} + \text{minimum temperature})/2) - 4.4$$

These pea heat units were accumulated daily between the planting date and the date of harvest to create the variable "cumulative heat units". A base temperature of $4.4^{o}C$ was used to remain consistent with calculations by the central field office, though the accuracy of the ANN was independent of any value used. The same daily accumulation was done for daily radiation and rainfall to create the "cumulative radiation" and "cumulative rainfall" variables. Past data was obtained form the Bureau of Meteorology SILO website (www.bom.gov.au/silo) for four official weather stations located near the pea growing regions. The model used data from the nearest weather station in each growing region.

A comparison was made between the predictions by the central field office staff and those by the ANN model. Paddocks were selected that had a field estimate two days prior to harvest. For these paddocks, the field estimate of MI were used and added an expected MI gain for two days. This represented the manual method performed by the central field office staff. When comparing the manual method to the ANN for 2-day predictions, the average error for the ANN was 30.3 compared to 36.6 for the manual method.

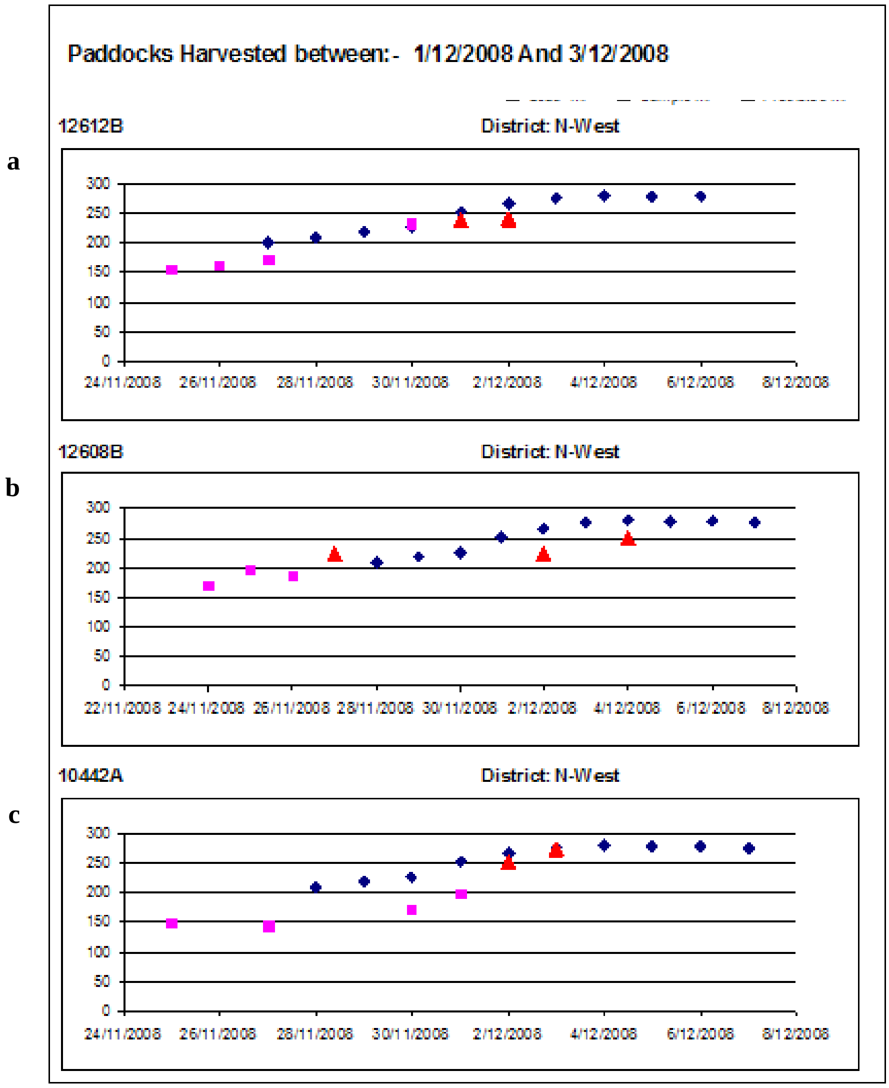

Figure 5. Graphs for pea paddocks comparing sample average maturity index (MI) (square), predicted average maturity index (MI) (diamond), and harvested (load) average maturity index (MI) (triangle).

The graphs in Figure 5 show the trend of Predicted MI (diamonds) over a ten day period for three paddocks harvested between 1st and 3rd December. For comparison purposes, the ANN was used to generate MI estimates five days before and after the first harvest date. Overlaid are field samples MI values (squares) with the sampling date, along with the actual MI of the paddock when harvested (Load MI - triangles). Each graph indicates a rapid gain in MI towards the harvest date, which was typical for the vast majority of paddocks. Whilst Figure 5 shows that the model does predict the rapid gain in MI, a significant weakness of the ANN model is also highlighted. The ANN will substantially overestimate lower MI's, which is typically the case when the ideal harvest time window is more than about 5 days away. This is indicated in Figure 5c where the ANN substantially overestimates the sample MIs. Such a problem is caused by the model being developed using historical data of harvested paddocks. In such data, there is little representation of paddocks being harvested several days prior to their optimal window. The problem is not a major concern for application of the model in practice, which is to predict the optimal time window of harvest.

For the 2008-2009 season, paddock field sampling data will be captured in the sampling database application along with daily observed and predicted climate data. The application links with the ANN MI Prediction model to predict daily MI values for paddocks that have been recording field sample MI values of higher than 150 MI for the previous week. This provides an opportunity to focus on paddocks that are approaching maturity and devise efficient harvesting schedules for harvesting gangs across the various districts. These harvesting plans will ensure a consistent flow of raw product to the processing factory along with ability to plan for efficiencies within the cold storage facilities.

The ANN MI Prediction model was integrated into the sampling database application so that is can be used as a harvest planning tool in day-to-day operations. Once all the appropriate decisions have been made for selecting paddocks that are ready to harvest, a daily harvest plan (or "bible") is created for the harvesting gangs to follow. The "bible" accumulates the tonnage of the selected paddocks to ensure manageable tonnages for the harvesting gangs and the factory.

4. CONCLUSIONS AND FURTHER DEVELOPMENTS

This study has shown ways to increase efficiencies in the pea supply chain from a harvesting perspective by improving the recording and communication of field sampling results to field officers and harvesting planners as well as improving the ability to predict MI of green peas using an ANN model. This includes the development of a database application to electronically record sampling results and communicating via Email-to-SMS technologies. Implementation of the sampling database and the ANN model has taken place during the 2008/2009 harvest season and a full evaluation of its usefulness is being conducted. Training of field officers at the processing factory has taken place to help ensure a smooth transition for model adoption.

The ANN model improves the capacity to predict MI of green peas, without being a data intensive burden. The model will be complementary to the field sampling, which is an accurate method for estimating MI at a given point in time. Whilst the ANN model requires more data than the manual method (historical data), it can easily be updated and linked to the database application. A single ANN model can be generated for an entire growing region in Tasmania, without any additional manual parameter settings by the user for specific districts or varieties. There are opportunities for improving the model. The four climate stations available do not accommodate micro-climates that exist within some regions. MI can also vary substantially within a single paddock, which means there is also a large amount of variability around the mean MI measured at the factory for the paddock. Currently, the ANN model estimates MI for a given day, which may be too coarse a time interval since the MI of a paddock can increase by 30 MI units in a single day. A future development of the model could be to predict MI at a date and time of day, rather than date only.

Other planning tools will be integrated into the database application to provide the ability to view a planting and harvest schedule for the season by districts, showing the tonnes harvested and due to be harvested for each week of the season. The sampling and harvesting scheduling database system can also be applied to other crops, such as beans and broccoli. By doing this, the processing factory can develop multi-crop intake plans, accommodating the processing capacity of the factory, that maximise profitability across its portfolio of food products. A next step in the research could be to link the ANN modelling system to an optimisation model for harvesting and transport logistics. This will give the processor a better capability to plan harvesting and transport crews to maximise efficiency.

ACKNOWLEDGMENTS

The authors thank Terry Dennis, Keith Jenkins, Petra Novak, Kim Arney and Josh Cables for their contributions to developing, installing and testing the sampling database application and maturity prediction model. Thanks also are extended to Dr Peter Stone for helpful comments in reviewing this paper and sharing his knowledge about pea growth modeling.

REFERENCES

Mitchell, R.S., and L.J. Lynch, (1952), The Maturometer, an instrument for the objective assessment of maturity in canning peas. *Food Technology in Australia*, 4, 195-196.