Intelligent Clustering Techniques for Prediction of Sugar Production

V.G. Kaburlasos(1), V. Spais(1), V. Petridis(1), L. Petrou(1), S. Kazarlis(1),
N. Maslaris(2), and A. Kallinakis(2)

(1) Aristotle University of Thessaloniki, GR-540 06 Greece
Department of Electrical and Computer Engineering
Automation and Robotics Laboratory

(2) Hellenic Sugar Industry S.A.
34 Mitropoleos Street
GR-541 10 Thessaloniki, Greece

Abstract

The accurate-, and timely prediction of the annual sugar-beet crop yield is important to Sugar Industry because, based on it, the "harvest campaign" can be scheduled efficiently. This work presents intelligent clustering techniques for effecting efficient-, small error prediction of the annual sugar-beet crop yield for the Hellenic Sugar Industry based on production- and meteorological- data acquired during a period of eleven years. The experiments here demonstrate that intelligent clustering techniques can provide with better estimates of sugar production than alternative prediction methods including an "energy conservation" system model.

1 Introduction

The sugar production in Europe stems from an annual (in farm practicing) plant, namely *Beta Vulgaris L* or simply *sugar-beet*. The sugar-beet is planted in early spring, and sugar is produced from processing the roots of the plant which (roots) are harvested, in Greece, during late summer /early fall time period.

The capacity to predict both accurately and timely the annual sugar-beet crop yield is of vital importance to Sugar Industry because the deployment of the "harvest campaign" can be scheduled efficiently resulting in an increased margin of profit.

Root and sugar yield forecasts are based on pre-harvest sugar-beet and other data samples taken from randomly selected fields. Forecasting is achieved using a prediction model on the basis of the trend in the current year compared with the corresponding trend in previous years [16]. For instance, linear and multiple regression models have been employed [1, 2, 16] based on historical data. Attempts have been made to incorporate in the prediction models various aspects of climate and environmental parameters [10, 12, 15].

At an individual- "farm level", decision support systems have been developed for improving sugar-beet crop yield based on expert knowledge and symbolic data manipulation rather than on numeric data processing [18]. A "systems approach" to the problem of sugar production and prediction has been adopted by several researchers aiming at an understanding, as well, of the interaction among various underlying mechanisms and factors [9, 16]. There exists an extended bibliography presenting the results of various studies for modeling and forecasting of both the sugar-beet crop yield and the sugar production [4, 17, 19].

The algorithmic prediction of sugar production in Greece is not trivial for two reasons. First, prediction models which have been applied successfully elsewhere in Europe are not valid in Greece due to differences in both climate and soil fertility [8]. Second, the number of data available for developing as well as for validating various prediction models barely spans the last decade. Therefore previous attempts to develop prediction models such as interpolation-, polynomial-, linear autoregression-, and neural- predictors for Hellenic Sugar Industry (HSI) have met with only limited success [13]. More specifically, in [13] a prediction error around 6% is reported using data of six years. This work builds on previous experience to improve prediction accuracy.

Section 2 presents the sugar prediction problem in terms of the data which are available. Section 3 presents two forecasting models. Section 4 details the experiments and the corresponding results. Finally, section 5 concludes by discussing comparatively the merits of the proposed techniques and by delineating the potential for further improvement.

2 Data Acquisition

The sugar-beet growing area in Greece, stretches from Florina to Orestiada along central and northern Greece. The production is organized around five principal factories located at Larisa, Platy, Serres, Xanthi, and Orestiada. In the beginning of the year in every factory a set of representative fields, namely *pilot fields*, is defined by agriculturalists of Hellenic Sugar Industry (HSI) to obtaining sample measurements from. Different factories might define different numbers of pilot fields. Moreover, even for the same factory, the number of pilot fields might differ from year to year. Note that it is not necessary, in a factory, one pilot field to be "pilot field" again for any future year. Nevertheless, as long as a set of pilot fields is defined in the beginning of a year for a factory, then all samples during the year in question are obtained from the defined pilot fields.

Sample measurements for several variables of agricultural interest were available in this study where each variable was being sampled for a number of consecutive years; the last year any variable sampled was 1999. Several variables have been dropped for various reasons. In particular, for some variables there were samples available for a period of only 2 years, therefore those variables were ignored from the outset. Other variables, pertaining to "soil", were ignored because i) those variables were being sampled for only 6-7 years, and ii) a data preprocessing procedure showed that the values of those variables were fairly stable from year to year. Note that a variable pertaining to soil is sampled only once a year.

Samples of *production variables* as well as of *meteorological* (*climate*) *variables* have been considered in this work. The *production variables* were recorded from year 1987 to 1999 including, and they are summarized in Table 1.

Table 1 Production variables

	Variable name	Description (where necessary)	Unit
1	average root weight		g
2	POL (sugar content)	sugar in fresh root weight	%
3	α-amino-Nitrogen (α-N)		meq*/100g root
4	potassium (K)		meq*/100g root
5	sodium (Na)		meq*/100g root
6	Leaf Area Index (LAI)	leaf area per field area	dimensionless
7	TOP	plant top weight	kg/stremma**
8	Roots Yield (RY)		kg/stremma**
9	Nitrogen-test (N-test)	NO_3-N content in pedioles	$mg.kg^{-1}$
10	planting date		

* meq: milli-equivalent
** 1 stremma = 1000 m^2

The *sugar content* or, equivalently, *sugar production* is calculated by the product POL×RY. The *meteorological* (*climate*) *variables* were recorded from year 1989 to 1999 including, and they are summarized in Table 2.

Table 2 Meteorological (climate) variables

	Variable name	Unit
1	average daily temperature	^oC
2	maximum daily temperature	^oC
3	minimum daily temperature	^oC
4	relative humidity	%
5	wind speed	miles /hour
6	daily precipitation	mm
7	daily evaporation	mm
8	sunlight	hours /day

Sample measurements of both *production variables* and *meteorological* (*climate*) *variables* have been used in this work for eleven consecutive years from 1989 to 1999 including. An additional number of samples has been dropped as explained in the following.

The meteorological data here were recorded by the National Meteorological Service of Greece (EMY) in local stations nearby the factory areas. Data for the Orestiada factory did not exist, whereas the data for the Xanthi factory were sparse and practically unusable. Note that there were available complete meteorological data sets from EMY stations at Xrisoupolis and at Alexandroupolis. Nevertheless due to the considerable distances of Xrisoupolis and Alexandroupolis from, respectively, Xanthi and Orestiada as well as due to the considerable geographic /climate differences between the aforementioned districts it was decided to use jointly *production* and *meteorological* data from only three factories, these are Larisa, Platy, and Serres. The last difficulty regarding the data arose from the fact that the sampling of the production- and the meteorological- variables was not uniform as explained below.

On the one hand, the *meteorological variables* were sampled every day during the whole year. On the other hand, the *production variables* were sampled at different rates in different years and, in addition, the sampling was lasting for different lengths of time. In particular until year 1997, starting on June 1, or June 11, or June 21, samples of the production variables were taken in all factories every 10 days until either September or early October. For the last two years, starting on June 20, the production variables were sampled in all factories every 20 days until either September or early October. Another problem stemmed from the fact that the "pilot fields" were actually "commercial fields" cultivated for profit, therefore when harvest was beginning in late August the pilot fields were bound to be harvested as well, hence the measurements were distorted. We decided to consider measurements up until September 10 including. In conclusion, regarding the *production variables*, samples have been considered from June 20 through September 10 every 20 days, whereas, regarding the *meteorological variables*, daily samples have been considered from June 1 through September 10. Note that for the aforementioned data the "missing values" represented around 1% of the total amount of data.

In conclusion, the "raw" data considered in this work have been samples of 10 *production*- and 8 *meteorological- variables* for the three factories of Larisa, Platy, and Serres for eleven consecutive years from 1989 to 1999 including. The Larisa factory included always 50 pilot fields, the Platy factory 54, 58, or 60 pilot fields, and the Serres factory included 40 or 45 pilot fields.

3 Forecasting Models

The goal has been to predict accurately at the end of July the *sugar content* to be on September 10. Recall that at the end of July a set of measurements (samples) of the *production variables* is available immediately before the beginning of the harvest campaign. Note that in this work the prediction of sugar content is aimed at the factory level, that is there is no interest here for prediction either at the individual field- or at the plant-level.

3.1 "First Principles" Model

A "system modeling" approach has been elaborated for the prediction of sugar-beet crop growth. The model is based on "first principles" where the sugar-beet plant growth process is modeled as an energy- and matter-conversion process at a macroscopic scale. The *process inputs* (exogenous variables) are mainly weather related. In particular these variables include the daily average temperature, the daily sunshine (in hours), and the date of planting which also depends on both the weather and plant variety. The system *state variables* include the leaf weight, total root weight, and weight of sugar in the plant. The most significant internal derived variable is the leaf area index (LAI) which determines the plant energy intake and subsequent growth.

The model is described by a set of non linear differential equations that are numerically integrated using the Euler method at one day sampling intervals. The energy input from sunshine, modulated by the daily average temperature and current LAI, is converted into solid matter for the leaf and root (non-sugar) parts of the plant and into sugar. The daily average temperature determines leaf loss and metabolic loss of solid matter.

3.2 "Computational Intelligence" based Prediction

By "Computational Intelligence" is meant an array of tools such as neural networks, fuzzy systems, and evolutionary computation techniques, which attempt to simulate computationally various aspects of human intelligence [14]. In this work computational intelligence techniques have been used for building a model which best fits the data. The *sugar production* has been the *output* of the predictor, whereas the corresponding *inputs* are selected among the *production*- and the *meteorological*- variables as explained below.

In the first place, a "clustering" procedure was carried out in order to group together similar arithmetic values of the output (sugar production) in a factory. Note that in a previous work on sugar prediction [13], "clustering" has also been employed for grouping together various inputs to a predictor model in order to train one model with "similar", in a "cluster belonging" sense, inputs. Moreover note that the prediction of sugar production is treated as a time-series classification problem in [13] using data of six years from 1989 to 1994 without including any meteorological data. In this work clustering was carried out on the output (sugar production) values for every factory. The decision for this sort of clustering was taken in line with the common practice by agriculturalists of HSI that the correct classification of a year as "good", "medium", or "poor" (sugar production year) is sufficient in practice. After having clustered the sugar production (of September 10) in a factory in three clusters, a training /testing computation was carried out as explained in the following.

The leave-1-out paradigm has been used for training /testing, that is one year was left out for testing and the remaining years were used for training. Both "training" and "testing" used the *L1-distance* between two sets of samples which (L1-distance) is computed as follows. Let S_1 and S_2 be two sets of numbers, then the L1-distance between S_1 and S_2 is defined to be the absolute value of the difference of the average values of the two sets S_1 and S_2. Note that a set of numbers emerges in this work from either the *production*- or the *meteorological- variables*. In particular for production variables, a set of numbers emerges from the set of pilot fields, whereas for meteorological variables a set of numbers emerges from the set of the daily samples taken during the previous 20 days.

During the "training phase" one square symmetric matrix with the L1-distances between different "training" years was calculated for each input variable. Let $X=\{x_1,\ldots,x_N\}$ be the set of input variables and let

$I=\{1,\ldots,N\}$ be a set of indices. Furthermore, let D_i denote the square symmetric matrix with the L1-distances for input variable x_i. A weighted matrix D was defined by

$$D= w_i D_i+\ldots+w_k D_k, \text{ where } i, \ldots,k\in I, \text{ and } w_i, \ldots w_k \text{ are real number weights.}$$

A 'training year" was associated with another one which corresponded to the shortest distance in D. A *contradiction* occurs if the two years (associated with the shortest distance) are in different categories. A *cost function* C(D) was defined as the sum of all contradictions. Hence, the following optimization problem had to be solved:

Find 1) indices $i, \ldots,k\in I$, and
2) weights $w_i, \ldots w_k$
such that C(D) is minimized.

In case more than one optimal solutions are computed, then one solution is selected randomly.

Having computed indices $i,\ldots,k\in I$ as well as weights $w_i, \ldots w_k$, the "testing phase" is carried out by computing the enhanced matrices $D'_i,\ldots,D'_k$ and D' such that the "testing year" is also accounted in the computation of matrices $D'_i,\ldots,D'_k$ and D'. More specifically, each dimension of square matrices $D'_i,\ldots,D'_k$, D' is "by one" larger than the dimension of square matrices $D_i,\ldots, D_k$, D. Finally, the "testing year" is assigned to the category of its shortest-distance-year as computed from the enhanced weighted matrix D'.

The optimization of the aforementioned cost function C(D) was carried out using a Genetic Algorithm (GA) [5, 6]. The encoded *solution vector* (genotype) of the GA was a binary string containing representations of the weights $w_i,\ldots w_k$ of variables $x_i,\ldots x_k$, respectively. Note that the weights $w_i,\ldots w_k$ were being varied in the closed interval [0,2], where weight value '0' implies that the corresponding input variable is "missing" from the solution.

4 Experiments and Results

4.1 "First principles" Model

A number of modeling relationships and parameters are defined based on Greek Sugar Industry experience as well as on previous experimentation. Four to twenty parameters (depending on the simulation experiment) are left free for estimation by algorithmic techniques based on process data. Gradient free search techniques are used for this purpose and in particular the Simplex method of Nelder-Mead as implemented in Matlab [3] and a variation of a genetic algorithm [7].

Five previous years of plant growth data from the Larisa factory are employed. There are fifty different measurement fields per production region. Each year of data is considered *in isolation*, i.e. data from different years are not used to try to estimate a given year's model parameters. The first three measurements (at 170, 190 and 210 days from the beginning of the year respectively) are used in a system identification manner while the measurement for day 230 is estimated. The resulting relative identification and prediction errors are shown in Fig.1. The average prediction error in five years was 15.63%. These errors, while within reasonable bounds given the variability of the underlying process, are too large for the intended purposes of prediction.

The large discrepancies between predicted and measured values for the sugar content are attributed to the small number of data available for model identification. Therefore, no further predictions have been attempted using this type of predictor model for the factories of Platy and Serres.

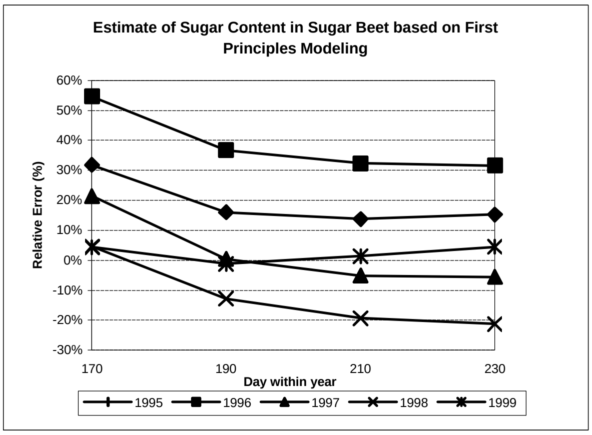

Fig.1 Prediction errors during five consecutive years of sugar production for the Larisa factory based on "first principles" system modeling techniques

4.2 "Computational Intelligence" based Prediction

Section 3.2 presented the arguments for employing three cluster levels, namely "good", "medium", and "poor" levels, for the annual sugar production in a factory. Due to the different capacities of the sugar growing districts of Larisa, Platy, and Serres, the levels for "good", "medium", and "poor" have been assigned by an expert agriculturalist at different levels for each factory (district) as shown in Table 3. Note that, during the procedure of defining the levels of three clusters per factory among the sugar production data of eleven years, one year was left out as an "outlier" for each factory because its corresponding sugar production was far beyond, either above or below, the sugar production rate of other years. In conclusion, data for ten years have been used for each factory.

For different reasons three production variables as well as two meteorological variables were ignored in the experiments. More specifically, 1) production variable *LAI* was ignored because there were not enough data; 2) meteorological variable *average daily temperature* was ignored as "redundant" since both the *maximum*- and the *minimum- daily temperatures* had been included in the experiments; 3) meteorological variable *wind strength* was ignored because a data preprocessing procedure showed an insignificant correlation with the annual sugar production; 4) production variable *planting date* was ignored because a "Markovian" assumption was made such that the planting date had no relevance, as soon as the first measurement (sample) was available; 5) production variable *potassium* (*K*) was ignored due to its high correlation with production variable *sodium* (*Na*), the latter variable has been included in the experiments. Hence, in conclusion, 7 production variables and 6 meteorological variables have been considered in the experiments. Two additional variables have been used namely PTOP and QR. Variable PTOP is defined as "the percentage TOP of the total plant's weight", and variable QR is defined by equation QR= POL/Na, where 'TOP', 'POL' and 'Na' are known production variables. Note that both variables PTOP and QR are known *indices* for sugar quality [11].

Table 3 Sugar production levels in Kg/stremma for "good", "medium", and "poor" years, in three agricultural districts.

Sugar Production level	Agricultural District		
	Larisa	Platy	Serres
"good"	1040	1045	1164
"medium"	970	961	1064
"poor"	890	925	982

Recall that the problem is, using the data available on July 30, to assign a year the correct label from the set of labels {"good", "medium", "poor"}. In other words, the problem of "prediction" has been treated here as a problem of "classification". It might be useful to point out that the minimum prediction error, ever expected, by the proposed prediction-via-classification method is 1.08%, 1.44%, and 1.46%, respectively, for the Larisa, Platy, and Serres factories as shown in the first row of Table 4. Note that the "minimum prediction error" is attained when each year is classified to its correct cluster "good", "medium", or "poor".

Table 4 also shows the average errors for the three factories of HSI when each year is assigned randomly a value in the set {"good", "medium", "poor"}. The error rates 8.45%, 5.87%, and 7.47% in Table 4 for the Larisa, Platy, and Serres factories, respectively, have been calculated by carrying out a number of random experiments in the computer. It should be pointed out that, even with a random prediction /classification, the prediction error is significantly smaller than the error of the previous "first principles" model. Hence, prediction-via-classification, as presented here, is a "well posed" technique in the sense that a small prediction error is expected from the outset.

Various techniques have been employed for reducing further the classification error. For instance a "medium" (sugar production) has been assumed for every year in a factory. In such a case the prediction error is further reduced as shown in Table 4. In particular prediction errors of 6.13%, 3.44%, and 5.54% have been calculated for Larisa, Platy, and Serres, respectively. This is yet another indication that the technique of prediction-via-classification is "well posed".

A further reduction of the prediction error was achieved using a genetic algorithm in order to identify the input variables which can better classify years as "good", "medium", and "poor". The GA was applied as described in section 3.2. The GA implementation here was a *simple GA*, that is no problem-specific-operators or other techniques were employed. Ten experiments were carried out for each factory area (Larisa, Platy and Serres), that is one experiment for each year of data. The GA encoded the 15 input variables' weight parameters using 3 bits per parameter and thus a total genotype length of 45 bits. A population of 200 genotypes (solutions) was employed and it was left to evolve for 500 generations. For each experiment (year) 10 independent GA runs were carried out and the best task was taken as the final result for the specific year. The search space contains $2^{45} \approx 3.5 \times 10^{13}$ solutions. The average number of solutions considered and evaluated by the GA in each experiment was only 10^5. The average time for the GA runs was about 10 seconds (for each run) on a Pentium III 667 Mhz platform. For the Larisa pilot fields the training error was in the range 0.91% to 2.05% and the testing error in the range 0.22% to 16.98%. For the Platy pilot fields the training error was in the range 1.4% to 2.4% and the testing error was in the range 1.05% to 13.3%. Finally for the Serres pilot fields the training error was in the range 1.38% to 2.34% and the testing error was in the range 0.82% to 10.43%. The average testing (prediction) errors for the three factories are summarized in Table 4.

Table 4 Average % prediction error rates using various methods for three factories of HSI.

Prediction Method	Larisa	Platy	Seres
minimum prediction error	1.08	1.44	1.46
GA with L1-distances	5.83	4.42	4.58

"medium" selection	6.13	3.44	5.54
random prediction	8.45	5.87	7.47

5 Discussion and Conclusion

There is a need for Hellenic Sugar Industry (HSI) to develop algorithms for predicting accurately the annual sugar production. Prediction models established in other countries can not be used due to differences in both environmental conditions and soil fertility.

A system modeling approach was used for predicting sugar production based on "first principles", in particular the sugar-beet plant growth was modeled as an energy conversion process. The large size of the prediction error (it was larger than 15% for the Larisa factory) was attributed to the lack of both detailed knowledge of the underlying and sufficient measured data. We intend to improve the algorithm by integrating into the model plant leaf aging which was found to be a significant in the overall beet growth process and is believed to be insufficiently approximated in the current model. Furthermore, as more data will become available in the future, better estimates of the model parameters could be estimated.

It should be pointed out that a system model for sugar production, despite its poor performance as a predictor in the context of this work, has been hailed firmly by expert agriculturalists of HSI as an effective tool for the education of young-, practicing- agriculturalists.

"Intelligent clustering" techniques for prediction of sugar production have demonstrated quite promising results here. In particular there have been reported good results for all three factories for which there was a sufficient number of data. An advantage of the intelligent clustering techniques presented here is that they are "model free". That is they do not require the expensive ("expensive", at least, in terms of time) procedure of developing a prediction model. The experiments in this work have shown that a "small" number of ten years of dependable measurements is enough for reducing the prediction error to around 5%.

A promising research direction, currently pursued, is to use a distance between the distributions of samples (measurements) of the pilot fields in a factory, in order to calculate the proximity of a year to previous years, instead of using the L1-distance. A further reduction of the sugar production prediction error could be achieved by taking into account the effects of potential diseases including cercospora, rhizomania disease, etc. as soon as such data become available.

References

[1] D. Bruandet, "Les prévisions de récolte à l' IRIS: forme actuelle et perspectives d' amélioration", *Sucrerie française*, pp. 297-302, 1989.

[2] B.M. Church, and A. Gnanasakthy, "Estimating sugar production from preharvest samples", *British Sugar Beet Review*, vol. 53, no. 3 pp. 9-11, 1983.

[3] T. Coleman, M.A. Branch, A. Grace, "Optimization Toolbox For Use with Matlab", The Mathworks Inc., 1990-1999.

[4] R.P. Freckleton, A.R. Watkinson, D.J. Webb, and T.H. Thomas, "Yield of Sugar-Beet in Relation to Weather and Nutrients", *Agricultural and Forest Meteorology*, vol. 93, no. 1, pp. 39-51, 1999.

[5] D.E. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*, Reading, MA: Addison Wesley, 1989.

[6] J.H. Holland, *Adaptation in Natural and Artificial Systems*, Ann Arbor, Michigan: The University of Michigan Press, 1975.

[7] S. Kazarlis, *Application of Genetic Algorithms on Real-World Constrained Optimization Problems and the Varying Fitness Function Technique* (in Greek), Ph.D. Dissertation, Dept. of Electrical and Computer Engineering, Aristotle University of Thessaloniki, Greece, 1998.

[8] A. Kallinakis, "Criteria and factors of sugar beet quality" (in Greek), Technical Report, Hellenic Sugar Industry S.A., Direction of Agricultural Services, Serres 1991.

[9] S. Krupa, M. Nosal, and A. Legge, "Modeling plant response to tropospheric ozone: concepts and strategies", in H.J. Jäger, M. Unsworth, L. De Temmerman and P. Mathy (Eds.), *Proceedings of the final Symposium of the European Open-Top Chambers project*, pp. 131-150, 1992.

[10] S. Mambelli, G. Vitali, G. Venturi, and M.T. Amaducci, "Simulation model for predicting sugar beet growth and sugar yield under Po-valley conditions", *Proceedings of the HRB 55th Winter Congress*, pp. 305-314, 1992.

[11] N. Maslaris, P. Christodoulou, and G. Zountsas, "Quantitative effect of root/leaf growth rate on root yield and quality parameters in sugar beet", *Zuckerindustrie*, vol. 122, no. 3, pp. 208-211, 1997.

[12] S.A. Modig, "Swedish Forecasts of sugar beet yields – Some Regression Models", *Proceedings of the IIRB 55th Winter Congress*, pp. 189-210, 1992.

[13] V. Petridis, A. Kehagias, L. Petrou, A. Bakirtzis, N. Maslaris, S. Kiartzis, H. Panagiotou, "A Bayesian Multiple Models Combination Method for Time Series Prediction", *Intl. Journal of Intelligent and Robotic Systems*, 2000.

[14] X. Yao, "Evolving Artificial Neural Networks", *Proceedings of the IEEE*, special issue on Computational Intelligence, vol. 87, no. 9, pp. 1423-1447, 1999.

[15] C.J.T. Spitters, H. van Keulen, and D.W.G. van Kraailingen, "A simple and universal crop growth simulator: SUCROS87", in R. Rabbinge, S.A. Ward, and H.H. van Laar (Eds.), *Simulation and systems management in crop protection*, Simulation Monographs 32, Pudoc, Wageningen, pp. 147-181, 1989.

[16] H. Vandendriessche, “Modeling Growth and Sugar Accumulation of Sugar Beet (*Beta Vulgaris L.*) for Potential Production Conditions”, Biologische Wetenschappen, Katholieke Universiteit Leuven, Belgium, Ph.D. Dissertation, September 1995.

[17] H. Vandendriessche, M.K. van Ittersum, “Crop models and decision support systems for yield forecasting and management of the sugar beet crop”, *Agricola*, vol. 4, no. 3, pp. 269-279, 1995.

[18] J.M. Vion, “SAKARA: Système expert pour l’ amelioration de la productivité des betteraves sucrières”, *Proceedings of the IIRB 55th Winter Congress*, pp. 17-31, 1992.

[19] C.R. Webb, A.R. Werker, and C.A. Gilligan, “Modeling the Dynamical Components of the Sugar-Beet Crop”, *Annals of Botany*, vol. 80, no. 4, pp. 427-436, 1997.