CHAPTER 2

LITERATURE REVIEW

2.1 System, model and simulation in agriculture

2.1.1 System define

Many definitions related to "system" term, which were argued by former authors (Conway, 1986; Dillon, 1990; Dent and Blackie, 1979). System was a complex set of related components within autonomous framework (Dent and Blackie, 1979), or collection interacting elements that formation together for some purposes or was a limited part of reality that contains interrelated elements such as agricultural systems consists of crop, animal, and human (Jintrawet, 1990). Likewise, simple system can be considered to compose small components and relationships among them and these relationships may be modeled by mathematic formulas/functions. This was concerned and studied by researchers and pioneer modelers.

2.1.2 Simulation

In the system analysis, at first, a set of logical statements and mathematic formulas on the real observed system is set up. Secondly, the stage to mimic the behavior of the complex system is developed. Finally, an experiment can be performed using the model (Nix, 1986). Using the model is so called as simulation. System approach and simulation tools have been used by engineers for over 30 years (Jintrawet, 1990), and presently are being applied in agricultural system research (Lemon, 1986). The approach is characterized by three categories: system, model and simulation. Crop or plant growth model belong sub symbolic model, mechanistic model, can be divided into classes: preliminary, comprehensive models and summary models (Penning de Vires, 1982). Modeling, in practice, represents a real object, in which, crop growth is also such object. Modeling crop growth interested by researchers and agronomists, dealt with prediction of growth in association with environment factors such as climate and soil etc. The practice of agriculture is based on knowledge, tradition and conjecture, and agricultural research improves the knowledge that provides the basis for decision-making. Traditional disciplinary research methods have been used to deal with biological and economic problems but have not been entirely successful in handling the inherent complexities of agricultural activities. However, as knowledge is accumulated, results obtained from observation change from being qualitative to being quantitative and mathematics can be adopted as the tool to express biological hypotheses (France and Thornley, 1984). Furthermore, Penning de Vries *et al.* (1991) indicated that the advantages of simulation models are directly related to the mechanistic approach in the sense: (i) they can help researchers to gain better understand the systems. This leads to either finding gap in knowledge and data, or to determine opportunities for improving management of the real system. In both cases, simulation models help focus research and experimentation; (ii) they help to improve extrapolation of research findings to new environments whether existing or not e.g. global climatic change. Greater extrapolation allows for more extensive use of experimental data and reduces the need for additional experiments. This increases the efficiency of adaptive research in similar extrapolation domains; (iii) they also provide means for communicating within and among organization for accelerated knowledge transfer and application. Before Simulation process done it must be identify model for it. Some crop growth models and their feature are described in next part.

2.1.3 Crop growth models in agriculture

At present, crop model has been developed under the different boundaries of production levels. Such model has been developed through the processes, formulation, validation and sensitive analysis (Dent and Blackie, 1979). Moreover, they have being widely used in agriculture research such as CERES model (Richie *et al.*, 1998), CROPGRO model (Boote *et al.*, 1998), APSIM model. Over the last 20 years, scientists have made considerable progress in development of computer models that simulate the interactive effects of weather, soil and management factors on the growth and yield of crop. Two major goals of most of these modeling efforts were (i) to better understand the processes that contribute to the growth and yield of crop, (ii) to apply the models to improve crop management (Jones and Ritchie, 1991)

Agricultural models are mathematical equations that represent the reactions that occur within the plant and the interactions between the plant and its environment. Owing to the complexity of the system and the incomplete status of present knowledge, it becomes impossible to completely represent the system in mathematical terms and hence, agricultural models are but crude images of the reality (Passioura, 1973, 1996).

Features of crop models

The main aim of constructing crop models is to obtain an estimate of the harvestable (economic) yield. According to the amount of data and knowledge, that is available within a particular field, models with different levels of complexity are developed. Grouping of models have been attempted by various authors (Brockington, 1979; France and Thornley, 1984; Brown and Rothery, 1994), but strong demarcations cannot be made since a model generally possesses the characteristics of more than one group. The most pertinent aspects of crop models are described below.

Empirical model

Empirical models are direct descriptions of observed data and are generally expressed as regression equations (with one or a few factors) and are used to estimate the final yield. Examples of such models include the response of crop yield to fertilizer application, the relationship between leaf area and leaf size in a given plant species and the relationship between canopy alone or coupled with stem number and / final yield in the soybean (Analla, 1998). These models are crude and are good means for interpolation at the location and the range over which they have been derived (Sinclair and Seligman, 1996) but it is advisable to avoid extrapolation.

Mechanistic model

A mechanistic model is one that describes the behavior of the system in terms of lower-level attributes. Hence, there is some mechanism, understanding or explanation at the lower levels. These models have the ability to mimic relevant physical, chemical or biological processes and to describe how and why a particular response results. The modeler usually starts with some emprise and as knowledge is gained additional parameters and variables are introduced to explain crop yield. Thus, the modeler adopts an approach. Most crop growth models, namely those mentioned in sections.

Static and dynamic models

A static model is one that does not contain time as a variable even if the end products of cropping systems are accumulated over time, e.g., the empirical models. In contrast dynamic models explicitly incorporate time as a variable and most dynamic models are first expressed as differential equations:

$$dy/dt = f(X)$$

Where y = an attribute of the system (animal live weight)

t = time variable

f = some function, possibly of y, t and other parameters.

The integration of the above equation will give the actual behavior of the system over time. It may be possible that at some stage, the rate of change of the system becomes zero such that $dy/dt = 0$ and therefore $f(X) = 0$, and the model is then static (France and Thornley, 1984). This continuum from dynamic to static state of dynamic models was also reported by Brown and Rothery (1994).

Deterministic and stochastic models

A deterministic model is one that makes definite predictions for quantities (e.g., animal live weight, crop yield or rainfall) without any associated probability distribution, variance, or random element. However, variations due to inaccuracies in recorded data and to heterogeneity in the material being dealt with are inherent to biological and agricultural systems (Brockington, 1979). In certain cases, deterministic models may be adequate despite these inherent variations but in others they might prove to be unsatisfactory e.g. in rainfall prediction. The greater the uncertainty in the system, the more inadequate deterministic models become.

When variation and uncertainty reaches a high level, it becomes advisable to develop a stochastic model that gives an expected mean value as well as the associated variance. However, stochastic models tend to be technically difficult to handle and can quickly become complex. Hence, it is advisable to attempt to solve the problem with a deterministic approach initially and to attempt the stochastic approach only if the results are not adequate and satisfactory (Thornley and Johnson, 1990).

Simulation and optimizing models

Simulation models form a group of models that is designed for the purpose of imitating the behavior of a system. They are mechanistic and in the majority of cases they are deterministic. Since they are designed to mimic the system at short time intervals (daily time-step), the aspect of variability related to daily change in weather and soil conditions is integrated. The short simulation time-step demands that a large amount of input data (climate parameters, soil characteristics and crop parameters) be available for the model to run. These models usually offer the possibility of specifying management options and they can be used to investigate a wide range of management strategies at low costs. Most crop models that are used to estimate crop yield fall within this category. Optimizing models have the specific objective of devising the best option in terms of management inputs for practical operation of the system. For deriving solutions, they use decision rules that are consistent with some optimizing algorithm. This forces some rigidity into their structure resulting in restrictions in representing stochastic and dynamic aspects of agricultural systems. Linear and non-linear programming was used initially at farm level for enterprise selection and resource allocation.

Later, applications to assess long-term adjustments in agriculture, regional competition, transportation studies, integrated production and distribution systems as well as policy issues in the adoption of technology, industry re-structuring and natural resources have been developed (Wegener, 1994).

Optimizing models do not allow the incorporation of much biological detail and may be poor representations of reality. Using the simulation approach to identify a restricted set of management options that are then evaluated with the optimizing models has been reported as a useful option (Swartzman and Van Dyne, 1972; Crabtree, 1972; Trebeck and Hardaker, 1972).

Model-building capabilities are developed and it becomes possible to adopt a holistic and quantitative approach to problem solving within the agricultural field. Owing to the inherent complexity of agriculture, modeling studies started only in the 1970s. Rapid accumulation of knowledge in the agricultural field and the increased accessibility to information technology has contributed to the development of a wide number of agricultural models over the last three decades.

2.2 Testing model

The model testing stage involves the confirmation that the calibrated model closely represents the real situation. The procedure consists of a comparison of simulated output and observed data that have not been previously used in the calibration stage. Ideally, all mechanistic models should be tested both at the level of overall system output and at the level of internal components and processes. The latter is an important aspect because due to the occurrence of feedback loops in biological systems, good prediction of system's overall output could be attributed to compensating internal errors (van Keulen, 1976). However, testing of all the components is not possible due to lack of detailed datasets and the option of testing only the determinant ones is adopted. For example, in a soil-water-crop model, it is important to test the extractable water and leaf area components since biomass accumulated is heavily dependent on these.

Testing procedures involve both qualitative and quantitative comparisons. Before starting the quantitative tests, it is advisable to qualitatively assess time-trends of simulated and observed data for both internal variables and systems outputs. Major discrepancies can be detected visually and these can be corrected before any quantitative tests are attempted. Quantitative comparison is generally restricted to a linear regression of the observed on simulated data (or vice versa), the expectation being a regression line with slope = 1 and intercept = 0 in the ideal case (Jones and Kiniry, 1986; Jones *et al*., 1989; Hammer and Muchow, 1991; Carberry and Muchow, 1992; Keating *et al*., 1999; Cheeroo-Nayamuth *et al*., 1999). Adjusted R^2 and root mean square deviation are usually adopted to assess the goodness-of-fit despite objections raised by Thornton and Hansen (1996), Mitchell (1997) and Analla (1998). Inadequate predictions of model outputs may require "re-fitting" of the regression curves or fine-tuning of one or more internal variables. This exercise should be undertaken with care because arbitrary changes may lead to changes in model structure that may limit the use of the model as a predictive tool.

In some cases, it is best to seek more reliable data through further experimentation than embarking on extensive modification of model parameters to achieve an acceptable fit to doubtful data. This decision relies on the modeler's expertise and rigor as well as on human resources and time available to invest in fine-tuning model predictions.

Evaluation is involved comparison of the outputs of fully calibrated model to real data and a determination of suitability for an intended purpose (Lemon, 1977), if it is desired to predict grain yield, the evaluation end point should encompass information on relationship between predicted and actual grain yields, on the environments involved and on specific aspects that could affect interpretation. Environment could be specified in a general way by using environment index such as grain yield as used in some plant breeding analysis, in term of some agronomic factors such as planting date, plant population, or major physical aspects of the environment e.g. soil texture, soil depths, mean temperature, day length. In principle, in model evaluation, providing of each specific reasons, there are three levels of model testing and evaluation: informational model testing (presumes model has been developed), (ii) minimal model testing (collecting enough season data to check model performance for a new region or cultivars), (iii) maximum-model calibration or detail testing (Tsuji, 1998). Also, for evaluation of model, based on fundamental principle, the determination of suitability of model for intended used, they common compare simulated data from the model and measured data from real experiments, popular method is to use formula error sum of square between simulation and observed data (Hunt, 1988).

There were so many authors who interested in model evaluation in diversified and various aspects by using previous models and developed, modified, testing and evaluation, applying the model to new and potential areas and new cultivars. Alagarswamy *et al.* (2000) studied the evaluation and application of the CROPGRO soybean simulation model in specific soil condition namely a Vertic Inceptisol. The model predicted reasonably the temporal changes in leaf are index, biomass and grain yield, was used to develop yield - ET relationship, and to assess the influence of soil water storage capacity on yield. Panya (1993) with his experiment of some rice varieties and based on five planting dates, using CERES-rice, validated the model. It was shown that simulated results and experimental data had close relationship, concluded the successful modified coefficient. The CERES-wheat was tested for the phasic development of using 113 independent data sets from a diversity of location throughout the world (Otter-Nacke *et al.*, 1987) and for accuracy of planting dates of various phonological events for winter wheat in Kansas (French and Hodges, 1985). Ruiz-Nogueira *et al.*, (2001) calibrated the CROPGRO-soybean model for growth and yield under rain fed conditions in Galicia, Spain, and then use the calibrated model to establish the best sowing dates for three cultivars at three locations in this region.

2.3 Previous studies related to determine genetic coefficients

There were some previous studies studying to determine or estimate the genetic coefficient i.e. for rice, and in general they had to be based on fundamental principles were defined that characteristics for one cultivars have been termed the "Genetic coefficients" for that cultivars. They can be defined that coefficients expressed summary way in which a specific crop cultivars divided up its life cycle, responds to different aspects of its environment, or appear/changes morphologically (IBSNAT, 1990). The crop genetic coefficients with each crop, consists of various terms and divided up into different phonological stages, i.e. soybean and maize, genetics coefficients of soybean are more so many than maize (IBSNAT, 1990).

There are previous studies on additional reproductive stages development one of the problems in obtaining robust and accurate parameters estimates for predicting soybean phenological stages is the lack of adequate data sets that include a wide range of night length and temperature. Development stages of soybean was divided into vegetative and reproductive, through such description soybean research in aspect of plant development is based on the standard descriptions of soybean development stages. In the past there were some systems for soybean development stages; also thanks separating into development stages, Jone *et al.*, (2001) in their study they applied this for phenology module that was run in DSSAT software.

The genetic coefficients need to be estimated for new cultivars. Many authors by various methods and approaches calculate the genetics coefficients for new varieties or genotypes. Jintrawet (1991), Panya (1993) estimated the genetic coefficients for rice, by using CRES-rice and indicated that crop model can determine the genetic coefficients successfully although it need time and precise of recorded data from the experiments or related trials. Singh and Virmani (1996) used experiment data of the 1984 and 1986 seasons, the model calibrated for cultivars-specific parameters of Annigeri and JG74 chickpea varieties. The model validated that result show CHICKPEA can be used to predict potential and water of chickpea in the Indian plateau. Wilson *et al.* (1995) used modified coefficient for maize to compare effects of temperature and solar radiation on growth and yield to be simulated in both warm and cool climate by using modified coefficients in terms of phenology, growth and final harvest. Alagarswamy *et al.* (2000) in their study, has determined the genetic coefficients for soybean cultivars PK470 that continued to be used for simulating in next steps, model testing of CROPGRO-soybean, and commonly the genetics coefficients were determined by GENCAL (Hunt *et al.*, 1993) in the DSSAT v3.5.

2.4 Methods of model evaluation

Evaluation of a crop model simulation model involves establishing confidence in its capability to predict outcomes experiences in the real world. A frequently used method for evaluation of models involves comparing observed values with simulated results in a scatter diagram. Normally a linear regression is used to fit a straight line between observed and simulated values (Ohnishi *et al.*, 1997). Then either parametric (Hammer and Muchow, 1991) used to determine whether the intercept of linear regression is equal to zero and the slope is equal to or not significantly different from unity. Mitchell (1997) argued against using linear regression as a testing tool because of its inherent inappropriateness and violation of assumptions associated using regression as a tool, and difficulties experienced in accrediting a true null hypothesis. Mitchell and Sheehy (1997) provided an alternative objective and simple method, free of a priory assumption. This method uses the deviations (prediction minus observation) plotted against the observed values and specific two criteria for adequacy of the model. They are the envelope of acceptable precision and proportion of points that must lie within the envelope. In this method, no statistical tests are involved and hence the problem of satisfying assumptions is avoided. Testing is one of some components of model evaluation, despite extensive literature dealing with testing procedures, validating simulation models remains a difficult and elusive task (Shannon, 1975). Different testing methods have been applied ranging from simple visual comparison of model predictions with field observations to highly sophisticated statistical tests. Some of the testing procedure, however, violate the basic assumption of statistical independence and cannot be legitimately used (Curry and Feldman, 1987). A distribution- free, non parametric test for the regression slope described by Hollander and Wolfe (1973) has been suggested by Welch *et al*. (1981) for model testing in pest management. The method consists of plotting observed versus predicted values, and testing whether the points deviate significantly from a line of unit slope. Commonly, goodness of fit was evaluated visually and by computing a standardized bias (R) and a standardized mean square error (V).

2.5 Environmental factors effect on soybean development and yield

2.5.1 Environment factors and crop yield

Some environment factors effect to soybean yield such as elevated CO_2, increased temperature, and altered rainfall patterns. Past studies to address effects of climatic changes have been conducted with crop models of varying abilities, some with somewhat empirical adjustments for the CO_2 fertilization effect. Continued evaluation is needed, particularly using models with mechanistic processes and sensitivity to CO_2 and temperature. With climatic change, the weather is proposed to be more variable, but present in weather simulators poorly reproduce rainfall patterns in same areas of the world. It may be better to use long term sequences of historical weather and to modify the temperatures and rainfalls proportionately to correspond to the monthly temperature and rainfall offsets predicted by the GCMs, as was done by Curry and Feldman (1987).

Based on factors affecting to soybean growth i.e. management conditions such as planting date, row spacing, plant population, irrigation and cultivars choice, (Boot *et al*., 1998) model was formed to simulate. In practice, yield response to long-term historical weather records for a region, and to optimize planting date, planting density, row spacing, choice cultivars, and fertilization application for different soil types. Egli and Bruening (1992) used the SOYGRO model to predict soybean response to sowing date in Kentucky. Based on model evaluations, they concluded that lower yield with later sowing date could be attributed to lower solar irradiance during late plantings and, in some cases, to lower temperature during grain filling for later cultivars.

2.5.2 Sowing date and crop yield

Main characteristics to consider in selecting a soybean varieties include maturity, lodging resistance, and resistance to disease and insects. Planting date is one of factors that effect on soybean yield due to soybean response to environment and so it declined an average of 3.6% as planting date was delayed from early May to the last of May or the first of June (Walter, 1983). Seed quality can influence yield in two ways indirectly by influencing emergence and final stand or directly through its influence on plant vigor, if inadequate plant population are obtained as a result of the use of low quality planting seed; yield will be reduced (Wilson *et al.*, 1995).

Aggarwal and Kalra (1994) used a wheat simulation model to show that delay in sowing date decreased wheat yield in India, in part by subjecting the crop to warmer temperatures during the grain filling. Crop models were used by Muchow *et al.* (1994) to access climatic risks relative to planting date decisions for sorghum in a subtropical rain fed region. In many tropical and subtropical regions, planting decisions await the onset of a short summer rainy season, and the available soil water reservoir is often only partially recharged in winter season. Singh *et al.* (1994) evaluated peanut growth model, PNUTGRO, and found good predictions of soil water dynamics and pod yield in response to seasonal variation of rainfall.

Sowing date closed soybean crop development process due to soybean is sensitive with temperature, photoperiod. Vegetative growth response to temperature, vegetative processes that are sensitive to temperature include rate of germination and emergence, rate of vegetative node formation, duration of vegetative growth, specific leaf area, photosynthesis, and maintenance respiration. Reproductive growth responses to temperature, it is important to describe appropriately temperature effects on the duration of seed growth phase, seed growth rate, pod addition, and portioning/pod abortion.

Two environment variables, photoperiod and temperature, strongly affect on soybean development. Soybean is a quantitative, short day plant. Most cultivars flower sooner under long night than under short night (Borthwick and Parker, 1938).

Gallegos *et al*. (1996) found significance effect on the durations of phases from flowering, pod set, and end of flowering to maturity when photoperiod was increased at the beginning of each of those phases.

2.5.3 Cultivars and crop yields

Cultivars proved that with various cultivars crop yields also ranged and changed though under the same climatic and soil conditions. Cultivars also related to number of grain seed in harvesting as well as seed quality and nutrition. Each cultivars is characterized by its genetics coefficients (IBSNAT, 1990), and brings various grain yields, responds with specific climate and soil situations. Due to relationship between cultivars and crop grain yield, researchers and breeders always search new crop lines and varieties selection, and expand new cultivars suitably to specific agro ecological zones to obtain the potential crop yield. Providing each cultivars in different condition, grain yields can be appeared differently from among others, and also extend more differences from various crop cultivars of several thousand breeds. Crop model simulation can be employed to select new improved cultivars based on modeling and simulation for crop responds with climatic and soil conditions, farm practices, and genetics coefficients.