MIXED STRUCTURAL AND BEHAVIORAL MODELS FOR PREDICTING THE FUTURE BEHAVIOR OF SOME ASPECTS OF THE MACROECONOMY

by

Mukund Moorthy

A Thesis Submitted to the Faculty of the

DEPARTMENT OF ELECTRICAL AND COMPUTER ENGINEERING

In Partial Fulfillment of the Requirements
For the Degree of

MASTER OF SCIENCE

In the Graduate College

THE UNIVERSITY OF ARIZONA

1 9 9 9

STATEMENT BY AUTHOR

This thesis has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the head of the major department or the Dean of the Graduate College when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

SIGNED:______________________________

APPROVAL BY THESIS DIRECTOR

This thesis has been approved on the date shown below:

Dr. François E. Cellier
Professor of
Electrical and Computer Engineering

Date

ACKNOWLEDGMENTS

I would like to express my gratitude to Prof. François E. Cellier for his invaluable guidance, assistance, and suggestions throughout this research effort. His encouragements and wise counseling helped me solve many problems during the course of this thesis effort. I would also like to thank Prof. Jeffrey T. LaFrance of University of Berkeley for his continued guidance.

I would like to express my sincere thanks to Prof. Larry C. Schooley and Prof. Lester Taylor (Economy) for reviewing the thesis and serving on my Committee.

I would also like to acknowledge the contributions made by Prof. B. Wilburn (Optical Sciences), and Prof. E. Zajac (Political Economy) towards the thesis.

I also wish to gratefully acknowledge the financial support obtained for this research from the U.S. Department of Defense under grant number MDA 904–97–C–4054.

I am also thankful to my parents, my brother and my sister without whose support, I would never have accomplished so much.

Dedicated to Bhimanagari Amman, my Grandparents and my Parents.

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

A new methodology for predicting the behavior of macro-economic variables is described. The approach is based on System Dynamics and Fuzzy Inductive Reasoning. A four-layer pseudo-hierarchical model is proposed. The bottom layer makes predictions about population dynamics, age distributions among the populace, as well as demographics. The second layer makes predictions about the general state of the economy, including such variables as inflation and unemployment. The third layer makes predictions about the demand for certain goods or services, such as milk products, used cars, mobile telephones, or internet services. The fourth and top layer makes predictions about the supply of such goods and services, both in terms of their volume and their prices. Each layer can be influenced by control variables, the values of which are only determined at higher levels. In this sense, the model is not strictly hierarchical. For example, the demand for goods at level three depends on the prices of these goods, which are only determined at level four. Yet, the prices are themselves influenced by the expected demand. The methodology is exemplified by means of a macroeconomic model that makes predictions about U.S. food demand during the 20th century.

CHAPTER 1

Introduction

During the late 1950's, some very unusual, miniature size economy cars were imported to the United States. One of these was the German Zündapp, a 9.5 foot, 16 horsepower four-seater with two doors : one at the front of the car and the other at the rear. The model designation was Janus, for the two-faced Roman God who guarded the gates with the help of peripheral vision. The Zündapp's driver and front passenger faced forward, while the two rear passengers faced the back.

Being a rear-seat passenger in the Zündapp was a lot like engaging in the time series based forecasting approaches. The only thing you would see was where you had been, and you often had little idea where you were going.

A navigator sitting in the rear of the Zündapp and having to give instructions to a blindfolded driver in the front would have an easy task, if the future were not much different from the past. Economic Forecasting is a similar task. Unless hampered by unpredictable events such as technological advances and changes in either the economy or consumer demands, accurate forecasts can be made.

Economics is a decision science with an objective of trying to explain the decisions of the agents that make up the economy. The U.S. economy consists of over 120 million families, each of which is making individual decisions designed to optimize their own well-being or utility, each having different family characteristics and histories and with decisions based largely on different information. The macroeconomy is then the aggregate of the series produced by the individual families, an example being consumption. A subset of this example, namely food consumption is the subject of this study. Whether or not a non-linear relationship at the micro–level produces non-linearity at the macro–level depends on various conditions, on the type of non-linearity, and the common features of the various micro information sets.

As an example of a family decision making which is not a rule-driven autonomous, or deterministic mechanism, consider the consumption of poultry products by a family over some period, such as a month. The family starts a month with a stock of food products and periodically buys some fresh food products over the course of the month. In most cases, the total consumption of food products changes little over a month. For a longer period, the total consumption will change and probably become more efficient. The consumption of poultry products can be considered to be a constant proportion of the consumption of total food products. Some forecasting models of consumption quantity of poultry products are based on total food consumption and a forecast of this quantity. However, throughout the month, the family makes a sequence of decisions about whether or not to consume a particular variety of food, in this case poultry food products on any occasion. These decisions will depend on family characteristics, such as family size and age composition, and its economic variables such as the family income, the price of poultry food products and other potential purchases. The economists will emphasize income and prices, even though many families do not actually know these variables with any precision. The decisions are made frequently and not independently and differ across families. One aspect of the economy is that any of its variables dislikes being thought of as being forecastable. At the next possible instant, the economy reacts to this action and probably ensures that the forecast is incorrect.

Making informed predictions about macroeconomic quantities is a very difficult task. The same facts are interpreted in a number of different ways by various economists leading to diametrically opposite opinions about the state of the economy. The economy actively defies any attempt at being understood. This is called the "*efficiency*" of the economy. A totally *efficient* economy is one that it is totally unpredictable. The reasoning behind this is also convincing. As long as the economy is predictable, someone will take advantage of it and make a lot of money. But the economy reacts to this action also. In short, it reacts vehemently to any small change. This aspect of its behavior is what makes its study interesting. Under this situation, the best that we can hope for, would be to get as close as possible to predicting the behavior of the economy and at the same time allow the economy to be as elusive as it wants to be about its behavior.

Making predictions about the state of the economy is not a difficult task [1]. What is more significant is the reliability of these predictions. Predictions about the state of the economy are not valuable unless the quality of these predictions can be assessed. The heart of the modeling/simulation environment is the part that estimates the error of the prediction. An economic modeling tool that is not self-critical, that does not check the validity of its own predictions is essentially worthless.

1.1 Economic Modeling

A variety of different techniques have been proposed and are in use for economic modeling [2]. Commonly employed methods use statistical techniques. Most of the time-series based forecasting techniques are statistical techniques. The majority of statistical procedures for economic modeling are designed to be used with data originating from a series of independent experiments or survey interviews. The resulting data or sample is taken as representative of some population. The statistical analysis that follows is largely concerned with making inferences about the properties of the population from the sample. With a time series data set, which is usually the case for economic models, the sequence of occurrence of the data plays considerable importance as it represents the time–history of the data.

Statistical procedures usually involve using past values of the data series to fit a straight line or an exponential curve to build a model. Extrapolation is then adopted to make a forecast. These forecasts have no insight into the economic forces that drive the model. It is also a rather naïve model since, sometimes readily available information is ignored. Statistical models may suffice if the model is less complex and fairly straight-forward.

The other types of models that are available, then fall into two categories. They are knowledge-based deductive models and pattern-based inductive models.

1.2 Knowledge-based Deductive Models

The most widely used knowledge-based deductive economic models are System Dynamics models [3]. *System Dynamics* starts out with the selection of a number of so-called level variables (state variables). *Levels* are variables that accumulate over time. For each level, a number of rate variables are defined. *Rates* are divided into inflows and outflows. Inflows contribute to the growth of the associated level, whereas outflows contribute to their decline. For each rate variable, a set of the most important factors is written down that influence the value of the rate. This is consistent with other modeling approaches. According to George Klir, the selection of variables must be the first step in any modeling effort. It constitutes the *level 0* of the epistemology of levels in his General System Problem Solving framework [4].

The factor variables of the System Dynamics methodology are then grouped into four classes : levels, rates, external inputs, and auxiliary variables. The factors may themselves be levels or rates, i.e., variables whose dynamics are already covered by the description provided earlier, or they may be external driving functions, whose values

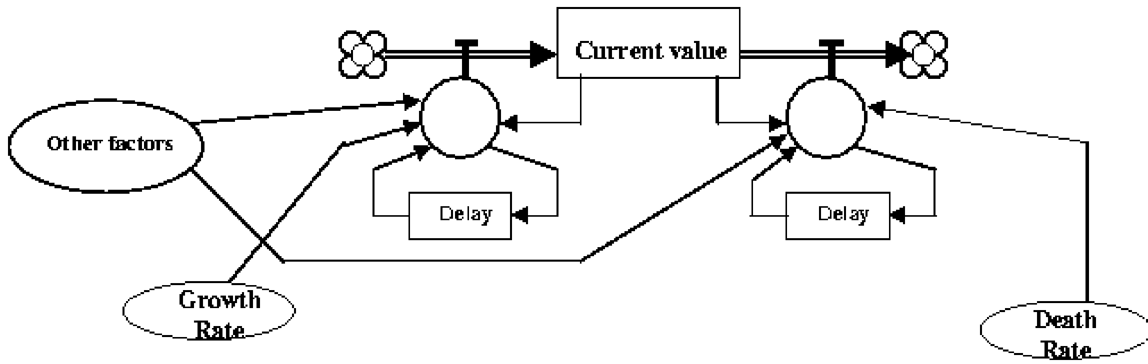

Figure 1.1: A Typical System Dynamics Model

as functions of time must be known, or finally they can be auxiliary variables. Any variable that does not fall into one of the first three categories is called an auxiliary variable. It is simply added to the set of variables, for which influencing factors must be determined.

A typical System Dynamics model in Fig. 1.1 shows how a growth variable can be represented using System Dynamics. The value of the variable at any time instant t, is the sum of the value of the variable at any time instant, t-k and the accumulation during the time period k. The net addition during the time k, is the sum of the amount added due to growth and the amount added as the input to the system. Therefore, in this example, the current value is the "*level*" variables, the additions are the "*rate*" variables and the growth rate is an "*external function*".

Classical System Dynamics then goes about proposing equations that link the factors together, and this is where it becomes unreasonable. System Dynamics assumes independence of different influencing factors. This assumption is not justifiable. The disadvantage with this model is that it makes structural assumptions about the system to be modeled, assumptions that are difficult to verify or refute, yet its predictions depend heavily on these silent assumptions, and the method is totally blind to its own gullibility.

A deductive model, derived from physical principles, has a high degree of validity in the sense that somewhat valid results can be obtained for a large range of values of its parameters. However, the results may not be very accurate due to the uncertainty of the actual parameter values and due to unmodeled dynamics.

1.3 Pattern based Inductive Models

The most commonly used pattern-based inductive economic models are Neural Network models [5]. A Neural Network is a highly non–linear set of (either static or dynamic) equations of arbitrary complexity. Using enough neurons distributed over a sufficient number of layers, basically *any* functional relationship can be approximated with arbitrary accuracy. Since the precise structure of the Neural Network doesn't matter, the structural assumptions are essentially harmless. The Neural Network assumes a very rich structure and can basically represent any behavioral pattern.

Parametric models suffer from the following drawbacks :

- Since the Neural Network assumes a structure in advance, it becomes very difficult at a later stage, to assess the error of this assumption. But there is also an advantage due to this assumption - since the assumed model structure is very rich, structural error is usually negligible.
- Since most parametric models are deterministic in nature, it becomes difficult to assess the error of the predictions made, because this error cannot be estimated in a deterministic sense. Once a Neural Network has been trained, it will make a prediction for any input pattern it is presented with, irrespective of how unlikely the correctness of the prediction may be. This is a serious drawback of all parametric models as we can never evaluate the genuineness of our prediction.
- Optimization is a part of any inductive model. The larger the number of parameters to be optimized, the slower the optimization will be. Training will thus take a very long time before we can start making a prediction.

Parametric models make an assumption about the structure of the relationship and then optimize parameter values to obtain an optimal curve fit between the observed and predicted output trajectories. Non-parametric models do not make any assumptions about the underlying model structure, and thus restrict themselves to characterize and record in the most efficient manner previously observed input/output patterns for use in the future and for interpolation.

Non-parametric models offer solutions to the aforementioned problems. They also have their drawbacks. Extrapolation in a non-parametric inductive model is totally impossible. A non-parametric model can only reproduce what it has been shown earlier in terms of input/output patterns. It cannot extrapolate beyond the range of input values that it has previously encountered.

In the case of parametric models, the extrapolation power lies precisely in the structural assumption made i.e., the more incorrect that structural assumption is, the more likely it will be that the extrapolated predictions are wrong. The enhanced validity of a parametric model hinges upon a correct guess of the underlying model structure.

Hence the real issue is not whether to use a structural or a behavioral model. The problem of gullibility can only be overcome by a *non–parametric model* that preserves the training data during the simulation (prediction) phase, comparing the current input patterns with those that had been used in the modeling (training) phase. Structural information should be added where available in order to reduce the need for training data, but should be limited to assumptions that can be verified.

Fuzzy Inductive Reasoning is a non-parametric modeling technique. Fuzzy Inductive Reasoning is a modeling methodology that preserves the best of both worlds, by enabling the user to mix deductive and inductive models in a single modeling and simulation environment.

1.4 Fuzzy Inductive Reasoning

Inductive Modeling tries to deduce a relationship between observational patterns, from a set of observations of input-output behavior of the system. The exact relationship between the variables is never calculated. The relationship should be capable of reproducing the observed output patterns when the model is presented with the observed input patterns. It should also be capable of producing sensible predictions of the output that are not totally incorrect when the model is presented with previously unobserved and different input patterns. Fuzzy Inductive Reasoning (FIR) is one such modeling approach. FIR has a history of success in identifying dynamic models of complex systems [6, 7, 8].

Fuzzy Inductive Reasoning not only allows modeling the system, but also allows modeling the error of the simulation. However, the same methodology used to model the output cannot be used to model its error. This is because of the fact that, if indeed it were possible to calculate a deterministic value for the inaccuracy, then the error could be subtracted from the prediction and a precise value of the output could be obtained. Hence, the error can only be modeled in a statistical sense.

Fuzzy Inductive Reasoning has an inherent self-validation capability. It rejects making predictions that are not justifiable on the basis of the available facts [9], a feature that can also be used to estimate the horizon of predictability [10].

Fuzzy Inductive Reasoning is a modeling and simulation methodology that generates a qualitative input/output model of a system by finding the best possible fuzzy infinite state machine between discretized (fuzzified) input and output states of the system. The methodology is composed of the following main engines :

- Fuzzification (Recoding)
- Qualitative Modeling
- Qualitative Simulation
- Defuzzification (Regeneration)

Inductive reasoning models the behavior of time-dependent phenomena by a pure pattern-matching approach. All inductive modeling methodologies involve studying input/output patterns followed by an optimization phase. To minimize the time taken for optimization, the real–valued signals are recoded into a smaller set of discrete classes. For example, instead of calculating the price of a food item in dollars, we classify the price as being either 'very cheap', 'cheap', 'moderate', 'costly', or 'very costly'. What recoding achieves is that it converts the quantitative trajectory behavior into a qualitative episodical behavior. Evidently some amount of knowledge is lost in the process of recoding. This is minimized by the use of fuzzy rules. Hence the process of recoding is also called fuzzification.

Fuzzification of real-valued variables is done for the following reasons :

- It speeds up the optimization dramatically. Suppose a relationship between n inputs and 1 output is given. Rather than searching through a n dimensional continuous search space to find the optimal input-output patterns, the search is limited to the very coarse n dimensional continuous discrete search space of class values. In this way, class values are used for determining the neighborhood of the optimal solution, whereas the fuzzy membership information is then used for interpolation in the vicinity of the optimal solution.

- The Optimization in the discrete space of the class values is deterministic. However, the subsequent interpolation in the continuous space of fuzzy membership values is stochastic. This approach is better capable of coping with model uncertainty than a purely deterministic approach.

- An almost identical technique to the one that is used to predict, in a statistical sense, the fuzzy membership value of the output, can also be used to assess, again in a statistical sense, the accuracy of the prediction made.

The second module is the *qualitative modeling* engine. Once the quantitative trajectory behavior has been recoded into a qualitative episodical behavior, the process of modeling consists of finding finite automata relations between the recoded variables that make the resulting state transition matrices as deterministic as possible. Such a relation is called a mask.

The mask consists of positive, negative and zero elements. The negative elements in the matrix denote inputs of the qualitative functional relationship, so-called *m-inputs* [11]. The positive value represents the *m-output*. A mask has the same number of columns as the episodical behavior to which it is applied, and it has a certain number of rows. The number of rows of the mask matrix is called the depth of the mask. The mask can be used to flatten a dynamic relationship out into a static relationship. A mask candidate matrix is an ensemble of all possible masks, from which the best one is chosen by a mechanism of exhaustive search. The mask candidate matrix contains -1 elements where the mask has a potential *m-input*, it contains a +1 element where the mask has its *m-output*, and it contains 0 elements to denote forbidden connections.

Each of the possible masks is compared to the others with respect to its potential merit. The optimality of the mask is evaluated with respect to the maximization of its forecasting power. The Shannon entropy measure is used to determine the uncertainty associated with the forecasting of the desired output state, for given feasible input states.

The third module is the *qualitative simulation* engine. FIR makes predictions by comparing the newly observed input pattern with all the input patterns in the experience data base (the training data), and finds the *five nearest neighbors*. It then predicts the most likely class and side values, and calculates the membership value as a weighted average of the membership values of the five nearest neighbors. In this way, reasoning is done using the (discrete) class and side values only, whereas the concrete quantitative information is preserved by interpolating among the (real–valued) membership functions of the five nearest neighbors.

The fourth and final module is the *defuzzification* module. Here, the predicted class, side, and membership values are converted back to real–valued quantitative predictions using the inverse operation to the fuzzification.

A schematic representation of the two primary engines of the FIR methodology, namely the qualitative modeling and qualitative simulation, as well as the two interface engines, fuzzification and defuzzification is shown in Fig. 1.2

1.5 Hybrid Model

Since the non-parametric method seems like a viable alternative for forecasting, it could then be employed for economic forecasting. However as in the case of the statistical techniques and other methods, the methodology in its pure state does not allow for incorporation of economic policies. Therefore, if the economic policies can somehow be forced into the methodology, a smart forecasting technique may result.

Economists insistently talk about the equilibrium of demand and supply, yet the prices tend to fluctuate and never really settle down. This phenomenon has a number of features.

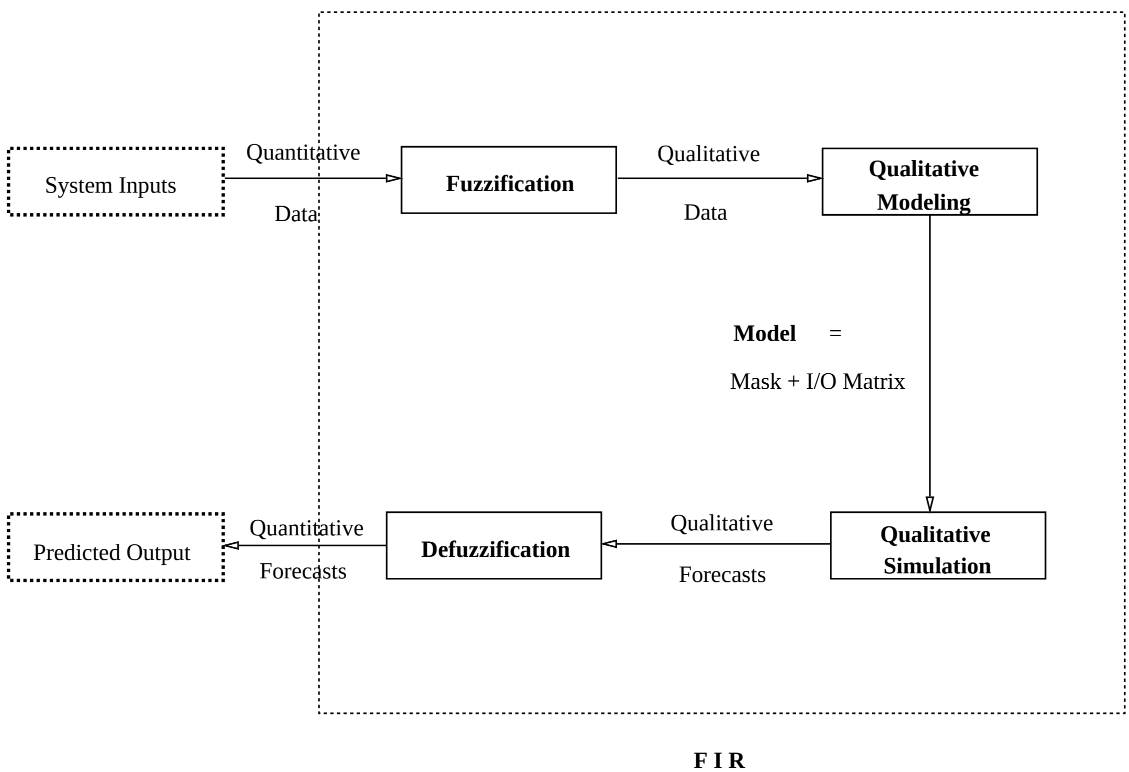

Figure 1.2: Schematic of the Fuzzy Inductive Reasoning Methodology

The systems that generate such erratic behavior consist of individual parts that interact with each other. In a market place, demand and supply interactions involve exchange of goods and services between producers and consumers. The interactions among the individual components do not occur instantaneously, but in a time delayed manner. Producers who offer their goods and services on the market may generate excess supply that leads to a drop in price. As a result, they may restrict production in the next period, leading to shortages and subsequent price increases. In many real-world dynamic processes, the response of one system component may not occur in direct proportionality to a stimulus that it receives. Rather, the responses may be related to the square of the initial stimulus or take place in some other non–linear relationship.

Understanding the economy requires knowledge about the role of complex feedback processes and the way in which their strengths change over time. A typical modeling methodology in such cases is to compartmentalize systems into subsystems for which it is possible to specify cause–effect relationships that have closed–form solutions.

Unfortunately, the methods that achieve such solutions may limit the extent to which one is able to accommodate timelags and non–linear relationships. By placing emphasis on finding closed–form solutions, we run the risk of eliminating from our models the very features that make them interesting.

In our day to day lives, we develop several models of dynamic processes. We notice and collect a number of details. In our mind, we abstract many details that we consider inconsequential. Then, we relate the remaining pieces of information to each other and make a projection of the possible outcome of the dynamic process. If we have drawn the right conclusion, we will use our model again in similar situations. If we are wrong and can afford to be wrong, we will revise our model the next time.

While for some decisions mental models are sufficiently simple and accurate to provide a basis for action, this is usually not the case. The larger the number of system components and the more time lags and non–linearities there are in the system, the more difficult it is for us to develop mental models for decision making as it becomes a tedious task to identify and throw away the inconsequential details.

To model and better understand non–linear dynamic systems requires identification of the main system components and their interactions. Such a model should then optimize the available facts by mixing knowledge-driven structural information with data-driven behavioral information. This suggests that we would need a System Dynamics model to identify our economic forces and then use Fuzzy Inductive Reasoning to make a forecast. This research work proposes such a model. The main system components are identified using System Dynamics and the consequence and inconsequence of variables and their interactions with one another is determined using Fuzzy Inductive Reasoning.

This research effort departs from the classical System Dynamics approach at the point where the System Dynamics model assumes one-to-one relationship between its variables. Instead, each rate equation is identified from observational data as an Inductive Reasoning model.

This research is the first attempt ever to combine System Dynamics with Fuzzy Inductive Reasoning in a mixed quantitative and qualitative modeling effort.

Data deficiency is a major problem in all economic predictions. The data rate is dictated by the natural time constants of the processes to be predicted. For example, population dynamics change over years, not over days. Hence, providing data more frequently than about once a year doesn't help. However, how relevant are data collected before the invention of reliable contraceptives for making predictions about the number of children per woman of childbearing age today? How many data points can be used for predicting the number of mobile telephones? Mobile telephones simply haven't been around very long. The data deprivation problem is the major stumbling block of *any* economic model. The proposed model shall address this issue, and put the problem into proper perspective.

Finally, any prediction is an act of *extrapolation.* Clearly, an external event, such as a war or a new invention, that is not foreseeable, cannot be predicted, and when it occurs, it immediately invalidates any prediction made prior to the event. A useful side product of being able to produce an estimate of the prediction error, is that this same estimate can be used to estimate the *horizon of predictability*, i.e., the time window into the future, for which meaningful predictions can be made, assuming that no unpredictable external events interfere.

Some of the conventionally used econometric modeling techniques are discussed in Chapter 2. Chapter 3 describes in detail the System Dynamics methodology. The Fuzzy Inductive Reasoning methodology is described in Chapter 4. Chapter 5 tries to elaborate on the structure for the proposed model and looks into two such models - a simple model that is specific to the U.S. Food and Agriculture industry, and a complete econometric model that can be applied to any U. S. industry. The complex model was run with the same food data and the results are explained in Chapter 6.

CHAPTER 2

Review of Economic Forecasting Techniques

For over fifty years, experts and novices, learned scholars and naïve upstarts have searched for statistics that signal future changes in economic activity. It is a very difficult task to identify the perfect indicator in forecasting future economic activity. Statistical techniques exist that attempt to determine this relationship to the best possible extent, but none have been outstandingly successful. Most time–series forecasts are based on statistical theory, not economic theory. Therefore, the time–series forecaster has no insight into the economic forces that drive the numbers. The time–series forecaster can adjust the predictions in the light of current economic events and may use economic theory to do so. Still, economic theory is not formally incorporated into the procedure.

In recent years, it has become fashionable to employ *Neural Network* models for predicting economic quantities. In some situations, Neural Networks are better suited than econometric models or NARMAX models [12], because they make less stringent structural assumptions about the relationships of the variables among each other.

A basic review of different techniques available for economic forecasting follows.

2.1 Time Series Data

Time Series Data are a sequence of observations over regular intervals of time. The time–series may consist of weekly production output, monthly sales, annual growth, or any other variable, the value of which is observed or reported at regular intervals. Analyzing a time–series helps identify patterns and tendencies that explain variation in past sales, growth rate or any variable of interest. This understanding contributes to our ability to forecast future values of the variable.

2.1.1 Time Series Components

Analysis of a time series involves identifying the components that have led to the fluctuations in the data. This approach, known as the classical approach, identifies the following components :

- **Trend(T)** – This is an overall downward or upward tendency. To the extent that the trend component is present, a regression line fitted to the points on the time–series will have either a positive or a negative slope.

- **Cyclical(C)** – These are the fluctuations that repeat over time, with the periodicity usually greater than one year. A business cycle is an example of this type of fluctuation.

- **Seasonal(S)** – These are also periodic fluctuations, but their periodicity is at the maximum, one year.

- **Irregular(I)** – This component represents random, or noise fluctuations that are the result of chance events, such as work stoppages, oil embargoes, equipment malfunction, or other happenings that either favorably or unfavorably influence the value of the variable of interest. Random variation can make it difficult to identify the effect of the other components.

In the classical time–series model, these components may be combined and represented in various ways, i.e., either as a product or as a sum. The classical time–series model can be represented by :

$$y = T \cdot C \cdot S \cdot I \tag{2.1}$$

$$\begin{aligned} \textit{where} \quad y &= \textit{observed value of the time series variable} \\ T &= \textit{trend component} \\ C &= \textit{cyclical component} \\ S &= \textit{seasonal component} \\ I &= \textit{irregular component} \end{aligned}$$

The model assumes that any observed value for y is the result of influences exerted by the four components, and here it has been assumed that the effect can be represented by a product of all the four components. The observed value of y is assumed to be the trend value, T adjusted upward or downward by the combined influences of the cyclical, seasonal and irregular components.

Several alternative models exist, including an additive model, with

$$y = T + C + S + I \tag{2.2}$$

It can also be assumed that the value of y depends on the value for y for the preceding period. In other words, y_t (observed value for time period t) may be expressed as a function of T, C, S, I and y_{t-1} (observed value for the preceding time period). Also the multiplicative and additive models may be combined as :

$$y = (T + C) \cdot \; S \; \cdot \; I \tag{2.3}$$

In general, the most important component of most time series is the trend, which may be examined by,

- using regression techniques to fit a trend to the data or
- using smoothing techniques to moderate the peaks and valleys within the series.

2.1.2 Curve Fitting

Regression techniques are used to fit the available data to an equation. The equation may be linear or non–linear. Non–linear equations may be exponential or polynomial equations.

While, the linear equation fitting technique is an extension of the Simple Linear Regression technique, the polynomial equation fitting is an extension of the Multiple Regression technique.

2.1.2.1 Linear Trend Equation

The linear trend equation fitting technique is a simple linear regression model, having a y intercept and a slope, with estimates of these population parameters based on sample data and determined by standard formulas. The model is described in terms of the population parameters as :

$$y_i = \quad b_0 + b_1 \cdot x_i + \epsilon_i \tag{2.4}$$

$$\begin{aligned} where \quad y_i &= a\ value\ of\ the\ dependent\ variable, y \\ x_i &= a\ value\ of\ the\ independent\ variable,\ x \\ b_0 &= the\ y\ intercept\ of\ the\ regression\ line \\ b_1 &= slope\ of\ the\ regression\ line \\ \epsilon_i &= random\ error,\ or\ residual \end{aligned}$$

The equation can then be written as :

$$\widehat{y_i} = \quad b_0 + b_1 \cdot x_i \tag{2.5}$$

$$\begin{aligned} where \quad \widehat{y_i} &= estimated\ value\ of\ the\ dependent\ variable, y \\ x_i &= a\ value\ of\ the\ independent\ variable,\ x \\ b_0 &= the\ y\ intercept\ of\ the\ regression\ line \\ b_1 &= slope\ of\ the\ regression\ line \end{aligned}$$

The *best fit* equation is then determined using the Least Squares criterion. The equation is then moved one step forward in time to make a forecast.

2.1.2.2 Exponential Trend Equation

A non–linear equation may prove to be more appropriate than a straight line for some time–series. One of the possibilities is to fit an exponential equation, a method that is advantageous whenever the time–series tends to increase at an increasing rate.

The exponential trend equation is expressed as

$$\widehat{y}_i = b_0 \cdot (b_1)^{x_i} \tag{2.6}$$

The equation can also be expressed using logarithms as :

$$log\ \widehat{y}_i = log\ b_0 + x_i \cdot log\ b_1 \tag{2.7}$$

In this case, a simple linear regression model is obtained similar to the previous case. A forecast can then made by moving the model ahead one time instant into the future.

2.1.2.3 Polynomial Trend Equation

A non–linear time series can also be fitted by a polynomial equation. This involves an equation estimating y as a function of x, but the method treats x, x^2, . . . , x^n as independent variables instead of just one.

The polynomial equation can be expressed as

$$\widehat{y}_i = b_0 + b_1 \cdot x_i + b_2 \cdot x_i^2 + \ldots + b_n \cdot x_i^n \tag{2.8}$$

The polynomial equation includes a constant component(b_0), a linear component($b_1 x$) and non–linear components($b_2 x^2$, …,$b_n x^n$). As such, it is especially appropriate when an upward trend in the time–series is followed by a downward trend or vice–versa. This type of equation can also be used whenever the time series either grows or decays at an increasing rate. This capability rivals that of the exponential equation, but whether the polynomial equation fits the time series better than the exponential equation will depend on the data to which they are being applied.

2.1.3 Smoothing Techniques

The techniques described earlier fit an actual equation to the time series. However most time series have short term fluctuations. There are several techniques available that smoothen out these short-term fluctuations. They dampen the sudden upward and downward jolts that occur over the series. There exist several forecasting procedures that can be classified as using *smoothing techniques*. These procedures all conceive of a time series possessing locally elements of level, trend, and possibly seasonality. Some of these techniques are described below

2.1.3.1 Moving Average

The moving average replaces the original time series with another series, each point of which is the center and the average of N points from the original series. For this reason, this technique is also known as the centered moving average. The purpose of the moving average is to take away the short-term seasonal and irregular variation, leaving behind a combined trend and cyclical movement. Including more periods in the moving average dampens the original fluctuations to a greater extent. This is because a larger base reduces the impact of any single data point.

2.1.3.2 Exponential Smoothing

The basic exponential smoothing equation replaces an observed series X by a smoothed series $\bar{X}$ - an exponentially weighted moving average of current and past values of X, e.g.

$$\bar{X}_t = aX_t + (1-a)X_{t-1}, \qquad 0 \leq a \leq 1 \tag{2.9}$$

In the simplest version of exponential smoothing, the latest available smoothed value is used to forecast all future observations.

In practice, the simplest approach is rarely employed, and several modifications designed to take into account local trend and seasonality have been introduced. The following is due to Holt [13] and Winters [14]. Writing the trend factor at time t as T_t, the seasonal factor as S_t, and again denoting the smoothed series as X_t, the local trend is estimated as

$$T_t = C(\bar{X}_t - \bar{X}_{t-1}) + (1-C)T_{t-1}, \qquad 0 \leq C \leq 1 \tag{2.10}$$

If the seasonal cycle has period L, then the seasonal factor at time t is

$$S_t = B\,\frac{X_t}{\bar{X}_t} + (1-B)S_{t-L}, \qquad 0 \leq B \leq 1 \tag{2.11}$$

and the smoothed series is given by

$$\bar{X}_t = A\,\frac{X_t}{S_{t-L}} + (1-A)(\bar{X}_{t-1} + T_{t-1}), \qquad 0 \le A \le 1 \tag{2.12}$$

The forecast of X_{n+h} made at time n is then

$$\hat{X}_n(h) = (\bar{X}_n + hT_n)\;S_{n-L+h}, \qquad h = 1, 2, \ldots, L \tag{2.13}$$

$$\hat{X}_n(h) = (\bar{X}_n + hT_n)\;S_{n-2L+h}, \qquad h = L+1, L+2, \ldots, 2L \tag{2.14}$$

The Holt-Winters method assumes additive trend and multiplicative factors, but can be modified to deal with a multiplicative trend or an additive seasonal component.

It remains to determine suitable values for the smoothing constants A, B, C. This can be achieved by calculating *forecasts* of the known observations X_n, X_{n-1}, X_{n-2},... over a grid of possible values of the smoothing constants, and selecting the set which performs best in terms of average squared forecast error.

Two alternative exponential forecasting procedures in common use are due to Brown [15] and Harrison [16].

2.1.4 Box-Jenkins Method

Box and Jenkins [17] described a forecasting procedure based on fitting a stochastic model to an observed time series. The class of models considered by these authors for non-seasonal time series can be written as

$$(1-\phi_1\ B-\phi_2\ B^2-\ldots-\phi_p\ B^p)\ (1-B)^d\ X_t=\theta_0+(1-\theta_1\ B-\theta_2\ B^2-\ldots-\theta_q\ B^q)a_t \tag{2.15}$$

where the a_t's are uncorrelated deviates, each with the same variance, and B is a back-shift operator on the index of the time series, so that $BX_t = X_{t-1}$, $B^2X_t = X_{t-2}$, and so on. It is assumed that the roots of the polynomial equations

$$(1-\phi_1\ B-\phi_2\ B^2-\ldots-\phi_p\ B^p)=0 \tag{2.16}$$

and

$$(1-\theta_1\ B-\theta_2\ B^2-\ldots-\theta_q\ B^q)=0 \tag{2.17}$$

all lie outside the unit circle.

Model-building is viewed by Box and Jenkins as an iterative process of identification, estimation and diagnostic checking. At the identification stage, tentative values for p,d and q are chosen. The coefficients of the selected model are then estimated using statistical techniques. Finally, diagnostic checks on the adequacy of representation of the model are employed. These may suggest modifications to the model, in which case the whole cycle is repeated until a satisfactory model is found.

For seasonal time series of period L, Box and Jenkins proposed the model

$$(1-\Phi_1B^L-\Phi_2B^{2L}-\ldots\Phi_pB^{PL})(1-\phi_1B-\phi_2B^2-\ldots-\phi_pB^p)(1-B)^d(1-B^L)^DX_t$$

$$=\theta_0+(1-\Theta_1B^L-\Theta_2B^{2L}-\ldots-\Theta_qB^{QL})\ (1-\theta_1B-\theta_2B^2-\ldots-\theta_qB^q)a_t \tag{2.18}$$

The assumptions of multiplicativity in this equation can be relaxed if such a course is suggested at the identification stage of the model-building cycle.

Forecasts of future values of a time series can be obtained by projecting forward the Box-Jenkins models. Unlike the exponential smoothing methods, the Box-Jenkins forecasting procedure is not fully automatic, as a good deal of judgment is required in constructing the particular forecasting model to be employed. This fact illustrates both the principal strength and a slight drawback of the procedure. The rigidity inherent in the exponential smoothing methods is considerably slackened, but forecasts take rather longer to generate when the Box-Jenkins approach is employed.

2.1.5 Step-wise Autoregressive Procedure

A vast majority of the economic time series are non-stationary i.e., they do not have a fixed mean over time. These series can generally be reduced to stationarity by first differencing. Accordingly, to minimize the danger of *discovering* spurious relationships it is advisable to analyze changes rather than levels of such series.

Consider a model of the form

$$x_t = \alpha + \sum_{j}^{K} \beta_j x_{t-j} + a_t \tag{2.19}$$

where

$$x_t = X_t - X_{t-1} \tag{2.20}$$

and K is some, possibly quite large, positive integer. One possibility, of course, is simply to fit the above equation by ordinary least squares. However, unless K is very small, such an approach would yield estimates of the coefficients that are no where close to the actual original values. An alternative is to employ a step-wise regression routine. In the first step, the lagged value x_{t-j} which contributes most towards explaining variance in x_t is introduced. In the second step, the lagged value which best improves the fit is added, and so on until introduction of further lagged values fails to produce a significant improvement in fit. Regressors introduced at an earlier stage, whose later contribution to the fit turn out to be insignificant can be dropped from the regression. Forecasts of future observations are then obtained by projecting forward the final fitted model.

Like exponential smoothing, this approach is fully automatic. Unlike exponential smoothing, however, the choice of forecast function is quite wide i.e., as in the Box and Jenkins method, an identification process is built into the forecasting mechanism.

2.2 Neural Networks

A Neural Network can be used to make forecasts just like any of the other statistical techniques. Although the simple exponential smoothing is an efficient statistical technique for forecasting, more accurate forecasts can be obtained by the Box and Jenkins technique. It has been reported that simple Neural Networks can outperform conventional methods, sometimes by several orders of magnitude [18]. It was also shown that the Neural Network model could forecast just as well as the Box Jenkins forecasting system [19].

Most of the stock market forecasting systems are based on some kind of statistical models such as Box-Jenkins methodologies. In recent years, artificial neural networks have also been applied to stock market forecasting and have achieved a certain degree of success [20], [21]. Wu *et al* developed a neural network model that makes forecasts of the stock market [22]. It was shown by Wu *et al* that the forecasts made by the Neural Networks were much better than the Box-Jenkins forecasts for the same system under consideration. For short memory systems, it was shown by Tang *et al* [23] that Neural Network models are superior to Box Jenkins models.

To monitor the forecasting system and determine whether one has to adjust parameters to reduce forecasting errors, a recurrent Neural Network can be used. The weight adjusting strategies of the recurrent Neural Network can be used to reduce the forecasting errors. Therefore, we can obtain forecasts efficiently based on simple exponential smoothing without having to monitor the forecasting system constantly and adjust the parameters manually. This proves to be a very effective tool in forecasting.

When the general Box-Jenkins methodologies of time series are applied, the tasks of tentative model identification, parameter estimation and diagnostic will have to be performed before the model can be used to forecast. They usually suffer from the huge computational cost of handling large volume data. However, in order to have a more accurate prediction, these systems usually are required to forecast based on the data of the last ten or twenty years. Therefore, most of them can only forecast based on monthly data of past years to be computationally feasible. Moreover, every time we have a new time series of historical data, the Box Jenkins forecasting methodology has to be reapplied to identify the model and estimate the parameters to produce future forecasts. Therefore, this modeling is a labor-intensive task because the model identification can only be judged by the researchers from the available candidate models.

A brief detail over how a Recurrent Neural Network can be used for forecasting using the same technique as exponential smoothing, but yielding better results follows. The equation used for simple exponential smoothing technique can be written as:

$$\hat{Y}_t = \alpha \cdot Y_t + (1 - \alpha) \cdot Y_{t-1} \tag{2.21}$$

$$where \quad \hat{Y}_t = estimated\ value\ of\ the\ variable, Y\ at\ time\ t$$

$$Y_t = value\ of\ the\ variable,\ Y\ at\ time\ t$$

$$Y_{t-1} = value\ of\ the\ variable,\ Y\ at\ time\ t-1$$

$$\alpha = weight\ of\ the\ neural\ network$$

It is obvious that the recurrent neural network will produce exactly the same results as the simple exponential technique as long as the value of α is known.

In simple exponential smoothing, the value of α is chosen in such a way as to minimize sums of squares of forecast errors. A usual method of finding an appropriate value of α is through a search on the interval (0,1) using the historical data. Once a value for α is determined, it is incorporated into the smoothing equation to generate future forecasts. Since no forecasting system will provide perfect forecasts, it is essential to monitor the system in some fashion to determine whether or not an adjustment to α is necessary. Hence, the method of simple exponential smoothing still runs into the problem of manually monitoring and adjusting the system.

In the artificial neural network as shown in eq. 2.21, the smoothing parameter now becomes the weight of the connection between the input node Y_{t-1} and the output node $\hat{Y}_t$ in the recurrent neural network. The weight adjustment formula of the artificial neural network can be used to monitor the smoothing parameter of simple exponential smoothing. Since the recurrent neural network self-adjusts the weight of its connections as described by Wasserman [24], it does not need any manual monitoring and adjusting. Although this method is similar to the simple exponential smoothing, the recurrent neural network has the advantage of automatically adjusting itself to the new data and simultaneously producing forecasts.

Neural Network models are better than Structural Models because they make less stringent structural assumptions about the relationships of the variables among each other. While this is good, this turns out to be its own undoing. The reason is the following. A Neural Network is a *parametric* model. The Neural Network being a highly non-linear set of equations can be used to represent any functional relationship with arbitrary accuracy. The structural assumptions are essentially harmless because during the training period of the Neural Network, the structural assumptions about the system hardly matter. But once the model has been trained, the system knowledge is totally contained in the network parameter values. The training data is no longer preserved. So now, if the input variable driving the system and the model suddenly leaves the range of input values used in training, the Neural Network never gets to know this. The Neural Network will still continue to make predictions on the output although the output has no significance. This is its undoing. The reason here being that the structural assumptions did not matter before training, but they do, once the network parameters have been fixed.

Finally, Neural Networks like other statistical techniques never incorporate the economic policies within the model. Therefore, a model that does incorporate such features is always a better alternative.

CHAPTER 3

System Dynamics

In a primitive society, all systems were those arising in nature and man adapted to these systems without ever feeling compelled to understand them. With the onset of the Industrial Revolution, these systems developed complexity and began to affect and eventually dominate life with their impact on all facets of life - economic cycles, political turmoil, unstable prices and fluctuating unemployment. Their behavior was so confusing that a general theory to explain their behavior seemed impossible. System Dynamics arose out of the need to find an orderly structure for explaining the cause and effect relationships. System Dynamics tries to establish the basic principles behind system behavior.

Most social systems are information feedback systems. The system behavior is affected by information flow forward and backward into the system. Feedback systems deal with the manner in which information is used for the purpose of control. It helps us understand how the amount of corrective action and the time delays in interconnected components can lead to unstable fluctuations. The study of feedback systems occupies a special place in economic systems because in economic systems, interactions between system components strongly affect the system behavior. As was seen in the previous chapter, Mathematics has been used to structure knowledge in science, but has not been adequate for handling the essential realities of important social systems.

In an information feedback system, it is the presently available information about the past that is used as a basis for deciding future action. The behavior of all feedback systems can be explained by *Structure*, *Delays* and *Amplification*. The structure of a system tells how the parts are related to one another. Delays exist in the availability of information and taking decisions and actions based on this information. Amplification exists all throughout the system and is manifested when an action is more forceful than might at first seem to be implied by the information given as input to the governing decisions. The behavior of an economic system arises from the different ways in which structure, timelags and amplification interact in the system.

Economic systems are abundant with fragments of knowledge. We have such fragments of knowledge, but have no way to structure this knowledge. To effectively interrelate and interpret observations in any field of knowledge, we require a theory. This theory could help put these fragments of knowledge together, to help learn the system. System Dynamics is a methodology or theory that helps assemble these fragments into a unified structure. With the help of this structure, it now becomes possible to interrelate facts and observations and learn from experience. It is this structure that helps to interpret past information, to understand it and to predict the future.

System Dynamics is based on the fact that all processes in nature are feedback systems, having many components, each of which can itself be a feedback system.

The System Dynamics (SD) methodology can be considered to be superior over conventional mathematical models because it looks at the general concepts that govern system behavior rather than an exclusive examination of any single part, which may not yield vital information. The SD methodology tries to fit a structure reflecting the information and material flow paths as they naturally occur in the system, rather than force a structure upon the model that satisfies an arbitrary set of mathematical equations.

3.1 Modeling Dynamic Systems

Most mathematical models of economic systems are steady-state, stable, and linear. In reality, most economic systems are unstable and restrained only by their non–linearities. Time and time rate changes form the essence of these models. Therefore, a successful model would be one that is dynamic and capable of adequately generating its own evolution over time.

System Dynamics models start with a structure, meaning the general nature of the interrelationships within it. Assumptions about the structure are not related to the data that drive the model. Plausible numerical values are assigned to the coefficients after a reasonable structure matching the descriptive knowledge of the system is obtained. The model is then accordingly altered such that its behavior resembles that of the real world system. The model tries to relate individual rules and characteristics of the system to the consequences that they imply. System Dynamics models are chiefly represented using information feedback loops, levels, rates and decision functions[25] [26].

3.1.1 Information Feedback Loops

Since economic systems are closed loop, information feedback systems, models of such systems, must preserve the closed loop structure that gives rise to so much of the interesting behavior. Information is continually fed back into the system as decisions are made. These decisions are also fed back into the system as information. A distinction between information and decision is very difficult as there is no specific difference between the start or the end of the loop. A representation for a feedback loop is shown in Fig. 3.1.

Economic fluctuations are one manifestation of the time-varying interactions that occur in the information feedback loops. When rising sales exceed plant capacity, this leads to expansion plans, which in turn restores a balance of supply and demand. Therefore the decision ultimately affects the environment that causes the decision.

The SD methodology might seem to suggest that the model should lack interest in the microscopic separate events that eventually build up the model. The study of

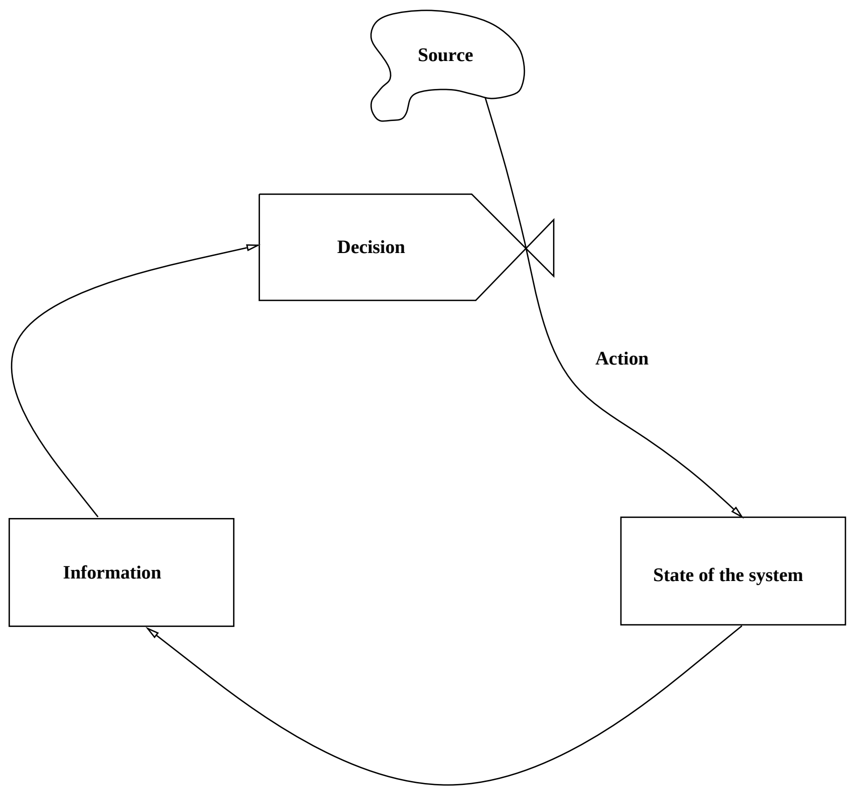

Figure 3.1: A Feedback Loop

individual events gives an idea of the decision structure and delays in the system, and is one of the richest sources of information about the way the model should be built. The methodology appreciates this fact, and simply suggests that more attention be paid to the system as a whole rather than concentrating at one single specific aspect of the system.

The fundamental nature of any SD model is then one that alternates between the following two entities namely, *Information Reservoir* and *Decisions*. The information reservoirs are called "*levels*" in the System Dynamics terminology, and the basic structure of the model can then be an alternating structure of reservoirs or levels interconnected by controlled flows. One such simple model is shown in Fig. 3.2.

The essential features of the model are

- Several levels, that each accumulate over time
- Flows that transport the contents of one level to another
- Decision functions that control the rates of flow between levels
- Information channels that connect the decision functions to the levels.

3.1.2 Levels

The levels are accumulations within the system. Levels are the present values of those variables that have resulted from the accumulated difference between inflows

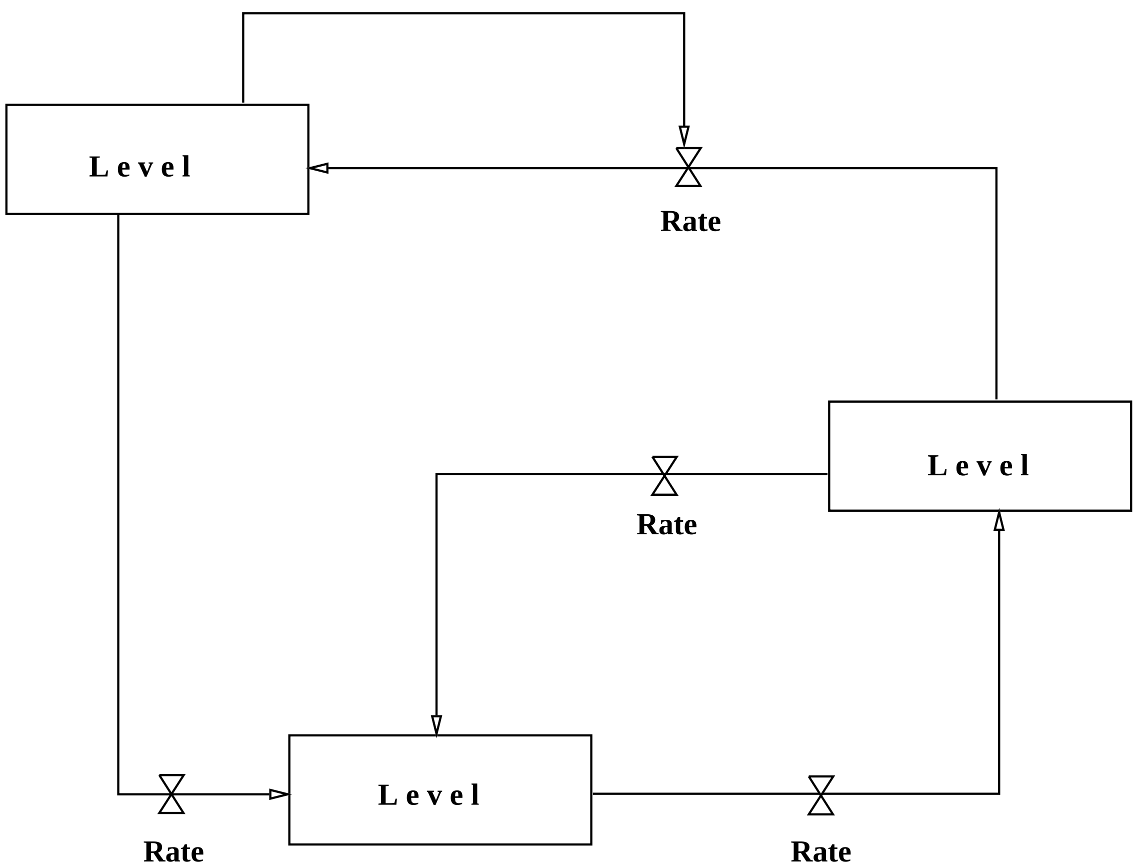

Figure 3.2: An Embryonic System Dynamics Model

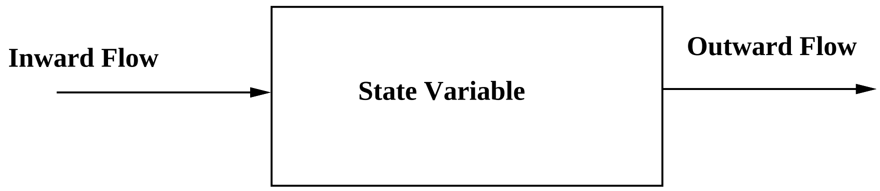

Figure 3.3: Representation for a Level Variable

and outflows. A level may have any number of inflow or outflow channels. In short, levels are the time integrals of the net flow rates. These variables are analogous to state variables in physical systems. A graphical representation for a level variable is shown in Fig. 3.3.

3.1.3 Flow Rates

Rates define the present, instantaneous flows between the levels in the system. The rates correspond to activity, while the levels measure the resulting state to which the system has been brought by the activity. The rates of flow are determined by the levels of the system according to the rules defined by the decision functions. The rates in turn determine the levels. The levels determining a particular flow rate will usually include the level from which the flow itself comes.

Rate variables and Level variables, although very similar in appearance to one another are in reality, quite different. If a system has been brought to rest, i.e., all activity in the form of flows were to cease, the levels would still exist but not the rates. The inputs to the rate variable could be a decision, an auxiliary variable or

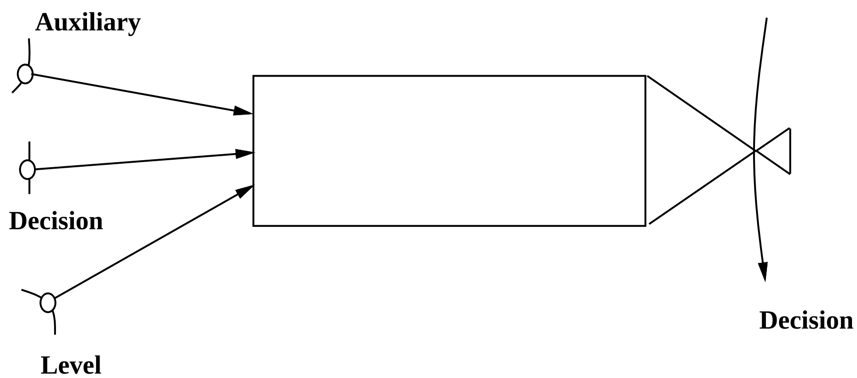

Figure 3.4: Representation for a Rate Variable

a level variable itself. These input variables help the rate variables control the flow from the source to the system. This graphical representation for a rate variable is shown in Fig. 3.4.

3.1.4 Decision Functions

A decision function is a simple equation that determines, in some elementary way, a flow in response to the condition of one or two levels. The decision functions determine how the available information about levels leads to the current rates or decisions. All decisions pertain to impending action and are expressed as flow rates.

There also exist decision functions that are long and elaborate. They might involve evaluation of a number of subdecisions. They can then be broken down into a number of smaller stages. Each of the smaller stages is evaluated by a separate decision function.

Decision functions determining the rates are dependent only on information about the levels. Rates are not determined or influenced by other rate variables. The present instantaneous rates are not available as inputs to the making of other decisions. Two variables can be connected by a level variable. System Dynamics then goes about determining a set of equations to relate different components in the model.

(i)**Level Equations** – New values of levels are calculated at each of the closely spaced solution intervals. Levels are assumed to change at a constant rate between solution times, but no values are calculated between such times. Level equations are independent of one another. A level at time t depends on its previous value at time t-k and on rate of flow during the interval k.

(ii)**Rate Equations** – The rate equations define the rates of flow between the levels of the system. The rate equations are the decision functions. A rate equation is evaluated from the current value of levels in the system, very often including the level from which the rate arises and the one into which the rate enters. The rates in turn cause the changes in levels. The rate equations in a broad sense, decide what happens next in the system. A rate equation evaluated at time t, determines the decision governing the rate of flow over time k, until the next time instant t+k.

Rate equations are evaluated independently of one another within any particular time step, just as level equations are. Interactions occur by their ensuing effect on levels that then influence other rates at later times.

(iii)**Auxiliary Equations** – When a rate equation is rather complicated, it can be broken down into a number of component equations called auxiliary equations. The auxiliary equation is of great help in keeping the model formulation in close correspondence with the actual system, it can be used to define separately many factors that make up the decision function. The auxiliary equations can be substituted forward into one another and then finally into the rate equation. They disappear with the appearance of complexity into the rate equations. The auxiliary equations are evaluated at time $t+k$, but after evaluation of the level equations and before the evaluation of the rate equations for the same time instant, $t+k$.

(iv)**Supplementary Equations** – Supplementary equations are used to define variables which are not part of the model structure, but help understand the system behavior and eventually aid a better and efficient representation of the system behavior.

(v)**Initial Value Equations** – Initial Value equations are used to define initial values of all levels that are necessary before starting to run the model. They are also occasionally used to compute some constants from other constants. The initial value equations are evaluated before the start of each model run.

The model is then run assuming a closed boundary system. Any action of the system is explained in terms of its components and their interactions. The System Dynamics methodology overcomes the disadvantages of other economic modeling techniques as it tries to make a more sensible explanation about the behavior of the system. It places more emphasis on the components of the system, and recognizes the overall behavior to be a by-product of the interactions among its components.

3.2 Laundry List

As a first step towards deriving a set of state equations, we try to enumerate all the factors that influence the rate variables. Such an enumeration is called a laundry list. The influencing factors may be levels, rates or converters. Laundry lists are the first step towards deriving state equations. Care must be taken to avoid algebraic loops in the equations.

For example, it might be claimed that the birth rate depends on: the population, the material standard of living, the available food (both quantity and quality), education, the availability of contraceptives, as well as religious beliefs, to just mention the more important factors. This can be written as :

$$birth\ \ rate \longleftarrow \begin{cases} \text{population} \\ \text{material standard of living} \\ \text{available food quantity} \\ \text{available food quality} \\ \text{education} \\ \text{availability of contraceptives} \\ \text{religious beliefs} \end{cases}$$

Such an enumeration is called a laundry list[27].

Dubious relations such as :

$$death\ \ rate \longrightarrow birth\ \ rate$$
$$birth\ \ rate \longrightarrow death\ \ rate$$

should be avoided to eliminate algebraic loops among the rate variables within the structure.

3.3 Structure Diagram

Forrester suggested a method for the representation of levels, rates and other variables occurring in a System Dynamics model[25]. The Structure Diagram clearly distinguishes between levels, rates and converters or auxiliary variables. Level variables are represented by square boxes. Most levels are bracketed by two little clouds that represent the sources and sinks of the level. Sources provide an infinite supply of the material which is stocked up or accumulated in the level variable. Sinks provide an inexhaustible dumping place for the same material. The double lines from the source cloud to the level and further to the sink cloud symbolize the flow of the material. Single lines symbolize the flow of information or the decisions.

Rate variables are denoted by circles with an attached valve. The rate variables control the flow into and out of the storages, i.e., the levels and stocks, symbolically by opening/closing the valve that they are responsible for. The rate variables are governed by the decision functions to decide how much of material should flow. Each rate variable is influenced by several other decision functions. This is symbolized by several decisions entering a rate variable. Great care must be taken to avoid algebraic loops if the rate variables influence themselves.

Converter variables are denoted by circles without an attached valve. They collect information from sources and deliver it to the rate variables after processing it. They can be used to explain additional dynamics. A structure diagram that represents the total population is shown in Fig. 3.5. The total number of people at any instant is dependent on the birth rate and the death rate. The birth rate and death rate are influenced by factors such as war. During a war, the death rate increases while the birth rate decreases. However, the rate variables themselves usually change only

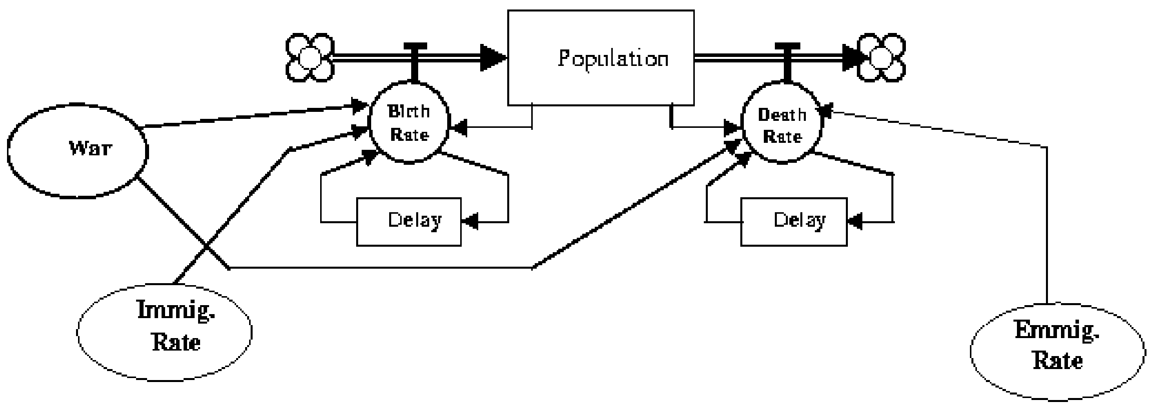

Figure 3.5: A typical Structure Diagram

slowly; e.g., this year's birth rate is positively correlated with last year's birth rate. This inertance of the rate variables is modeled using the delay boxes placed around the valves. The net population is also dependent on immigration and emmigration. These two factors were, in the model of Fig. 3.5 lumped together with births and deaths.

3.4 Forrester's World Model

The highly successful nature of the System Dynamics approach prompted Jay Forrester to propose a model of the world based on the System Dynamics theory[28].

Jay Forrester proposed the World Model, hoping to answer the questions posed by the "Club of Rome", a group that tried to decide whether there was a way to determine the destiny of the human race. The world model was one of the most famous SD models ever developed and published. Most of the concepts in the world model reflect the attitudes and motivations of the recent past and present. In these aspects, the model is a very primitive one.

The world model proposed by Forrester addressed only the broad aspects, and assumed that the present course of human events is not altered. This though, is a valid assumption and most models make a similar assumption, that the working is conditional to the persistence of a certain phenomenon or feature. The aggregration of the world model is at such a high level that the distinctions between developed and underdeveloped economies do not appear explicitly. The structure diagram for the world model is shown in Fig. 3.6.

The world model was built with five levels chosen as cornerstones for building the structure, namely

- Population
- Capital Investment
- Unrecoverable Natural Resources
- Fraction of Capital invested in the Agricultural sector
- Pollution

Each level was assumed to represent a principal variable in a major subsystem of the world structure. Forrester had decided that the world could be captured by these

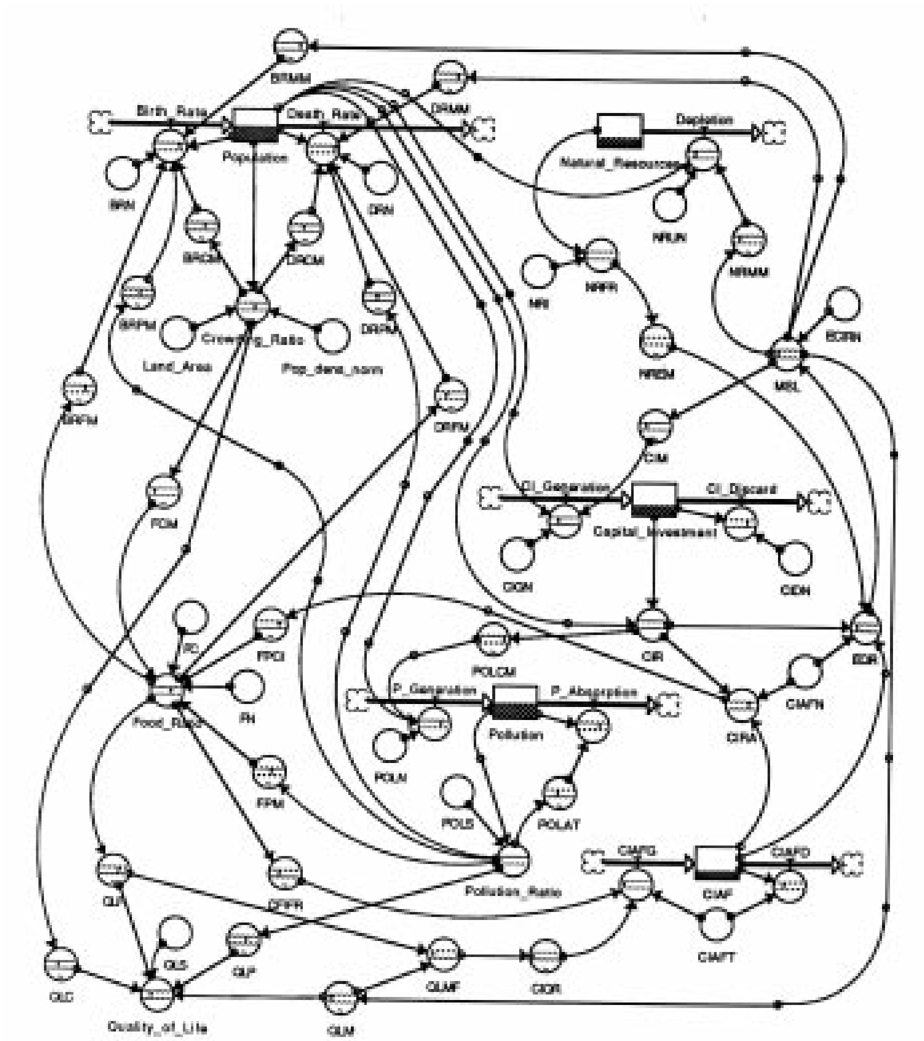

Figure 3.6: Forrester's World Model

five levels. The interactions between these five levels and the levels themselves were expected to answer most of the phenomena arising in the world.

The world model used the year 1970 as reference for defining constants and variables i.e., the world conditions are described relative to the conditions in 1970. Initial values starting from the year 1900 were determined and the model was run. The model was run beyond the year 1970 to make predictions on the values of the different levels and the rates.

The world model proposed by Forrester brought about a revolution in the field of modeling, with a sudden spurt in the use of System Dynamics for modeling. What followed his work, was a number of related works that used the methodology of System Dynamics to model different social systems whose modeling was earlier thought impossible.

A bibliography of System Dynamics written in the early eighties [29] is more than 30 pages long and contains already more than 600 entries. System Dynamics seemed unstoppable. Unfortunately, for reasons explained later, the world model just made explicit and very obvious some of the ridiculous assumptions that the SD methodology used as a rule–base for its model building.

3.5 Shortcomings of the World Model

The World Model eventually brought about the downfall of the System Dynamics models because it made explicit and very clear the drawbacks of using a System Dynamics model. The validity of a SD model is questionable.

In a physical system, like an electrical system or a mechanical system, we are aware of the variables that contribute to the power of the system. But in a SD model, we are never sure of this. The given model may be able to reflect the already observed behavior of the real system, but we may not be able to justify the model predictions of future behavior to reflect the true future of the real system under study.

There always exists the possibility of unmodeled dynamics. When the world model is run backwards in time, which can be easily achieved by reversing the signs of all rate equations governing the model, the results are quite amazing. The population decreases for a short period, before rising to an unacceptable value and eventually reaching infinity, when in reality it should near zero. This is a very good illustration of the failure of the world model and System Dynamics in general. The model may be able to capture most of the activities in terms of the interactions and attributes that contribute to the power of the system, but the model is never complete unless the unmodeled dynamics are negligible.

For example, one can write that:

$$BR = f(P, MSL, FQn, FQl, Ed, Co, RB) \tag{3.1}$$

to encode the aforementioned laundry list for the birth rate, and although this equation is an approximation of reality, it is an acceptable approximation, because the implicit interactions among the influencing factors sufficiently account for the dynamics of the system. Any modeling effort is inherently reductionistic in nature, and more influencing factors can always be added to the laundry list, if it should turn out that important dynamics have been overlooked.

Unfortunately, it is at this point where the SD methodology, as proposed by Forrester, becomes questionable. Because Forrester didn't know how to handle an equation as complex as the one presented above, he proposed to "rewrite" the above equation as follows:

$$BR = BRN * P * f_1(MSL) * f_2(FQn) * f_3(FQl) * f_4(Ed) * f_5(Co) * f_6(RB) \quad (3.2)$$

i.e., he pulled out the *normal birth rate constant (BRN)* as well as the population, and then described the "small signal behavior" of the variations imposed by the remaining factors. His assumptions are that each rate equation is purely static in nature, and that the influencing factors are independent of each other. These are a lot of assumptions that can hardly ever be justified, and that almost invariably lead to behavioral patterns in simulations that have little in common with reality.

First, it should not be assumed that the rate equations are static in nature. There exist natural state variables in systems that don't fall into the category of things that accumulate. For example, in mechanics:

$$\frac{dx}{dt} = v \tag{3.3}$$

$$\frac{dv}{dt} = a \tag{3.4}$$

The position, x, and the velocity, v, are natural state variables, whereas the acceleration, a, is not. Yet, it doesn't make sense to proclaim that positions and velocities accumulate, or that velocity is an inflow or outflow of position. These secondary dynamics can be captured by allowing the rate equations to be themselves dynamic.

Second, the assumption of independence of different influencing factors is ludicrous. This assumption cannot be justified on any grounds. However, it was precisely this assumption that made many economists embrace this technology at first, because it makes the modeling effort tractable. It eliminates the need for large quantities of measurement data.

The world model was very explicit in bringing out the drawbacks of SD because SD was trying to represent a very large system, where the possibility of unmodeled dynamics was very high. System Dynamics works wonderfully well when the system is small and of limited complexity, as there exist sufficiently few actions and interactions, and modeling this is not a difficult task. But when a system is large, complex and comprised of a number of actions and counteractions, the SD methodology fails as there still exist important unmodeled dynamics even after a high-level SD model is proposed for the said system.

CHAPTER 4

Fuzzy Inductive Reasoning

System Identification consists of two chief steps – in the first step, the structure that best characterizes the observed input-output behavior patterns is identified and in the next step, the parameters associated with the selected structure are identified, such that the observed input-output patterns are optimally reproduced by the synthesised system.

Fuzzy Inductive Reasoning (FIR) can be used as a tool for structure identification. An abstraction concept called *optimal mask* is introduced by FIR to help describe the optimal structure of the system to be identified. Since the system is qualitative in nature, the *optimal mask technique* used by FIR does not describe the complete structure. It selects a subset of variables from all observed potential input variables that are best suited to be used in reproducing the observed input-output patterns in the best possible manner.

Inductive Modeling was invented by George Klir [4] in the 1970's. It was used in his General System Problem Solving (GSPS) methodology as a tool to describe conceptual modes of system behavior. A first implementation of the Klir methodology was the *System Approach Problem Solver (SAPS)* [30]. A more practical application is SAPS-II, an implementation of a large subset of the GSPS methodology developed at the University of Arizona[31]. Fuzzy measures were introduced into the GSPS methodology in 1989 [32], and the kernel functions of fuzzy inductive reasoning were incorporated in SAPS-II, a year later [33]. SAPS-II is now available as a Matlab toolbox. The syntax of the SAPS-II function calls and the proper use of these functions are presented in [34].

4.1 Fuzzification

Since a qualitative model is required by FIR, this is achieved by fuzzy recoding. The process of converting quantitative variables into qualitative equivalents is called recoding. Recoding usually means that some amount of knowledge is lost. This is minimized in the *triplet* methodology. All real–valued variables are recoded into qualitative triplets. The first component of the triple is the class value, the second is the fuzzy membership function value, and the third is the side value [34, 35, 36]. The real–valued variable is recoded into a class value that is a coarse discretization of the original real–valued variable. The fuzzy membership value denotes the level of confidence expressed in the class value chosen to represent a particular quantitative value. The side value, tells us whether the quantitative value is to the left or to the right of the peak value of the associated membership function. The side value, which is a speciality of FIR, is responsible for preserving the complete knowledge in the qualitative triplet that had been contained in the original quantitative value.

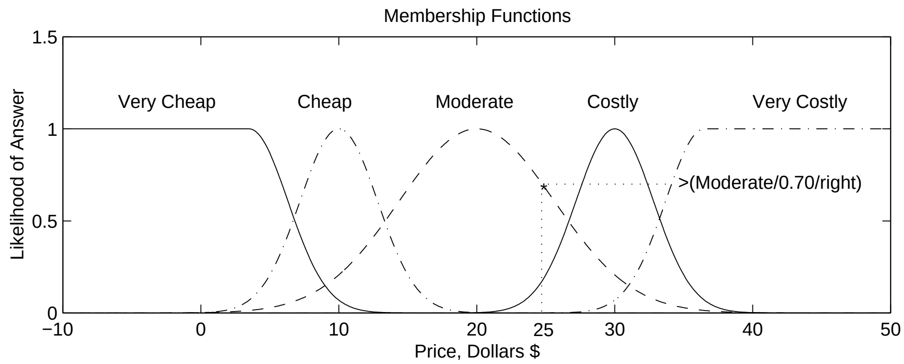

Figure 4.1: Membership functions of the price of food variables

Fig. 4.1 shows the fuzzy recoding of a quantitative variable (price of a commodity) into the five classes 'very cheap', 'cheap', 'moderate', 'costly' and 'very costly' using, in the shown example, popular knowledge to determine the so-called *landmarks*, i.e., the borders between neighboring classes. A quantitative value of price = 25, would in this case be recoded into a class value of 'moderate', a membership value of 0.70 and a side value of 'right'. Evidently, the qualitative triple contains exactly the same information as the original quantitative value.

The process of recoding is applied to each observed variable (trajectory) separately. The recoded qualitative episodical behavior is stored in three matrices, one containing the class values, the second storing the membership function values, and the third keeping the side values. Each column of these matrices represents one of the observed variables, and each row represents recorded states. The quantitative trajectory behavior can thus be mapped into qualitative episodical behavior.

4.1.1 Number of levels

One consideration during the process of fuzzification is the choice of number of *levels* into which the episodical behavior is to be captured. Any qualitative prediction made on the basis of a qualitative model will be limited to the same set of classes. One cannot hope to make predictions that are more precise than the qualitative model that one is working with. Therefore selecting the number of levels is an important decision. The selection of the number of discrete classes for representing each of the variables in the system relates to the struggle between generality and specificity. The more levels are chosen, the larger will be the expressiveness or specificity of the qualitative model. However, then making predictions becomes a difficult task as there is a need for more data. The smaller the number of levels chosen, the predictiveness(generality) of the model becomes better, but the less useful the predictions will be. If every variable is recoded into exactly one level, then the model will be highly predictive, but also at the same time completely useless.

From statistical considerations, in any class analysis, ideally each possible discrete state should be recorded at least five times[37]. The number of discrete states depends on the number of data points and is given by:

$$n_{rec} \geq \; 5 \;.\; n_{leg} = 5 \;.\; \prod_{\forall i} k_i \tag{4.1}$$

where n_{rec} denotes the total number of recordings, i.e., the total number of observed states, n_{leg} denotes the total number of different legal states, i is an index that loops over all variables, and k is an index that loops over all levels. If each variable assumes the same number of levels, this reduces to:

$$n_{rec} \geq \quad 5 \ . \ (n_{lev})^{n_{var}} \tag{4.2}$$

where n_{var} denotes the number of variables and n_{lev} denotes the chosen number of levels for each variable. The number of variables is usually known and the number of recordings is frequently predetermined. The optimum number of levels is then:

$$n_{lev} = round \ \ (\sqrt[n_{var}]{\frac{n_{rec}}{5}}) \tag{4.3}$$

For reasons of symmetry, usually an odd number of levels is preferred over an even number of levels. Choosing an odd number of levels will allow grouping anomalous levels symmetrically around the normal state.

Once the trajectory behavior has been recoded into episodical behavior, the inductive reasoning methodology can be applied. Inductive reasoning consists of a step of inductive modeling followed by a step of deductive simulation. In the inductive modeling step, a qualitative model is induced in the form of a finite state machine relating qualitative inputs to qualitative outputs. The selection of variables is the qualitative model. An abstraction mechanism is employed that determines which input variables to look at when one wishes to conclude something about a particular output variable.

4.2 Inductive Modeling

Inductive modeling consists of studying the input–output behavior to obtain a good representation of the system behavior in terms of the most significant inputs for any output. The episodical behavior from the recoding phase is stored in a raw data matrix. The task at hand is to identify a model that best describes the relationship between the different rows of the raw data matrix at each time instant. Each row of the raw data matrix represents one variable and each column represents one time instant. The values of the raw data matrix are in the set of legal levels that the variable can assume. For a five variable system, the episodical system might look like:

$$\begin{array}{ll} \textit{time} & \quad\; u_1 \quad\; u_2 \quad\; y_1 \quad\; y_2 \quad\; y_3 \\ \begin{array}{l} 0.0 \\ \delta t \\ 2.\delta t \\ 3.\delta t \\ \vdots \\ (n_{rec}-1).\delta t \end{array} & \left(\begin{array}{ccccc} \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \ldots & \ldots & \ldots & \ldots & \ldots \end{array} \right) \end{array} \tag{4.4}$$

4.2.1 Input-Output Behavior

The modeling process involves determination of finite automata between the re-coded variables that are as deterministic as possible. A state transition matrix that relates the input and output variables along with their probabilities of occurrence is built. This state transition matrix is a possible relationship between the qualitative variables. A possible relation among the qualitative variables of a five-variable system could be of the form:

$$y_1(t) = \tilde{f}(y_3(t-2\delta t), u_2(t-\delta t), y_1(t-\delta t), u_1(t)) \tag{4.5}$$

where $\tilde{f}$ denotes a qualitative relationship. The $\tilde{f}$ does not stand for any explicit formula relating the input arguments to the output arguments, but only represents a generic causality relationship that, in the case of the inductive reasoning methodology, will be encoded in the form of a tabulation of likely input/output patters, i.e., a state transition table. In SAPS-II, the equation (4.5), is represented by the following matrix :

$$\begin{array}{c} {}_t\backslash^x \\ t-2\delta t \\ t-\delta t \\ t \end{array} \begin{array}{c} \begin{array}{ccccc} u_1 & u_2 & y_1 & y_2 & y_3 \end{array} \\ \begin{pmatrix} 0 & 0 & 0 & 0 & -1 \\ 0 & -2 & -3 & 0 & 0 \\ -4 & 0 & +1 & 0 & 0 \end{pmatrix} \end{array} \tag{4.6}$$

The negative elements in this matrix are referred to as *m-inputs*. *m-inputs* denote input arguments of the qualitative functional relationship. They can be either inputs or outputs of the subsystem to be modeled, and they can have different time stamps. The above example contains 4 *m-inputs*. The sequence in which they are enumerated is immaterial. They are usually enumerated from left to right and top to bottom. The single positive value denotes the *m-output*. The terms *m-input* and *m-output* are used in order to avoid a potential confusion with the inputs and outputs of the system. In the above example, the first *m-input* corresponds to the output variable y_3, two sampling intervals back, $y_3(t - 2\delta t)$, whereas the second *m-input* refers to the input variable u_2 one sampling interval into the past, $u_2(t - \delta t)$, etc.

In inductive reasoning, such a representation is called a mask. A mask denotes a dynamic relationship among qualitative variables. A mask has the same number of columns as the episodical or recoded behavior to which it should be applied, and it has a certain number of rows, the depth of the mask. The mask helps to flatten the dynamic relationships between variables as shown in Fig. 4.2.

The optimal mask is the one abstraction that leads to the most deterministic input/output behavior. The problem of finding the optimal mask once again relates to the struggle between the generality and specificity. If more *m-inputs* are added to the mask, the observed patterns become more and more specific. Yet, chances are that a newly observed input pattern has never been seen before, making a prediction

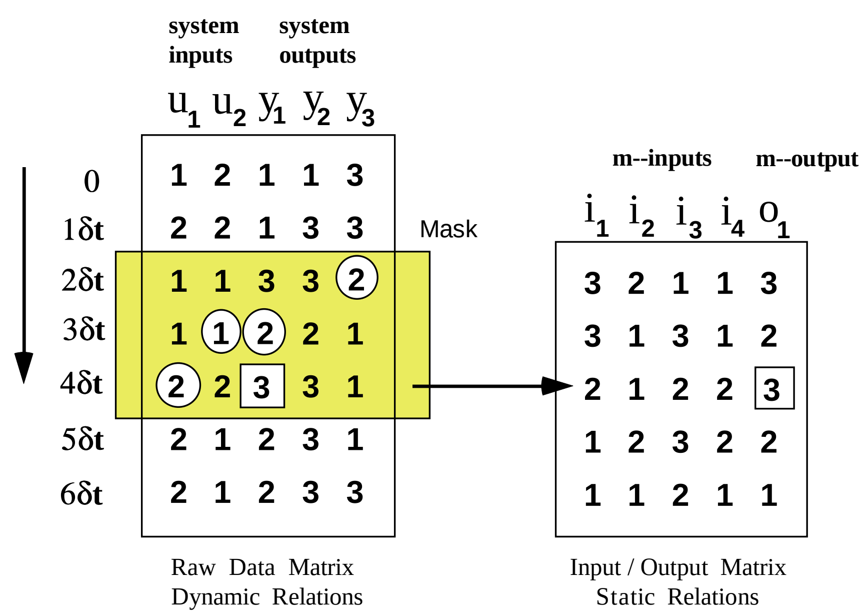

Figure 4.2: Flattening dynamic relationships through masking

impossible. Removing *m-inputs* from the mask leads to bolder, less specific, patterns that are likely to be ambiguous. The model no longer represents the true dynamics of the system, leading to non-deterministic input/output behavior, i.e., to ambiguities in the predictions made.

4.2.2 Optimal Masks

Determination of the optimal masks involves a definition of an ensemble of all feasible masks from which the optimal mask will be selected. This is accomplished by means of defining a *mask candidate matrix.* A mask candidate matrix for the previous example might be defined as :

$$\begin{array}{c} {}_t\backslash^x \\ t-2\delta t \\ t-\delta t \\ t \end{array} \begin{array}{c} \begin{array}{ccccc} u_1 & u_2 & y_1 & y_2 & y_3 \end{array} \\ \begin{pmatrix} -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & -1 & -1 \\ -1 & -1 & +1 & 0 & 0 \end{pmatrix} \end{array} \tag{4.7}$$

In the mask candidate matrix, -1 elements indicate potential inputs, whereas the +1 element still indicates the true mask output. 0 elements denote forbidden connections. From the mask candidate matrix, the optimal mask is often determined in a process of exhaustive search, but more effective suboptimal search strategies have also been devised[38].

A general rule for constructing the mask candidate matrix is to cover with the mask the largest time constant of the system that we wish to capture in the model. The mask depth should be chosen as twice the ratio between the largest and the smallest time constants to be captured.

In order to determine the optimal mask from the episodical behavior, constraints limiting the search space can be specified through the mask candidate matrix by blocking out forbidden connections by means of 0 elements, and by specifying the maximum tolerated mask complexity, i.e., the largest number of non-zero elements that the optimal masks may contain.

The exhaustive search process starts by evaluating first all legal masks of complexity two, i.e., all masks containing a single input and then proceeds by evaluating all legal masks of complexity three, i.e., all masks with two inputs and finds the best of those and this is continued until the maximum complexity has been reached. In all practical examples, the quality of the masks will first grow with increasing complexity, eventually reach a maximum, and then decay rapidly. A reasonable value for the maximum complexity is usually five or six.

Each of the possible masks is compared to all others with respect to its potential merit. The optimality of the masks is evaluated with respect to the maximization of their forecasting power. The Shannon entropy measure is used to determine the uncertainty associated with forecasting a particular output state given any legal input state. The Shannon entropy relative to one input is calculated from the equation :

$$H_i = \sum_{\forall i} p(o/i) \; . \; log_2 \; p(o/i) \tag{4.8}$$

where $p(o/i)$ is the conditional probability of a certain output state o to occur, given that the input state i has already occurred. The term probability is meant in a statistical rather than in a pure probabilistic sense. It denotes the quotient of the observed frequency of a particular state divided by the highest possible frequency of that state.

The overall entropy of the mask is then calculated as the sum:

$$H_m = -\sum_{\forall i} p(i) \; . \; H_i \tag{4.9}$$

where $p(i)$ is the probability of that input to occur. The highest possible entropy H_{max} is obtained when all probabilities are equal, and a zero entropy is encountered for relationships that are totally deterministic. A normalized overall entropy reduction H_r is defined as :

$$H_r = 1.0 - \frac{H_m}{H_{max}} \tag{4.10}$$

H_r is obviously a real-valued number in the range of 0.0 to 1.0, where higher values indicate an improved forecasting power. The optimal mask among a set of mask candidates is defined as the one with the highest entropy reduction. In the computation of the input/output matrix, a confidence value can be assigned to each row, and this value indicates how much confidence can be expressed in the individual rows of the input/output matrix.

The *basic behavior* of the input/output model can now be computed. It is defined as an ordered set of all observed distinct states together with a measure of confidence of each state. Rather than counting the observation frequencies, which is what is done in the case of a probabilistic measure, the individual confidences of each observed state are accumulated. If a state has been observed more than once, more and more confidence can be expressed in it. Thus, the individual confidences of each observation of a given state are simply accumulated. A normalized confidence of each input-output state can then be calculated by dividing the accumulated confidence in that input-output state by the sum of confidences for all input-output states sharing the same input state.

Application of the Shannon entropy to a confidence measure is a somewhat questionable undertaking on theoretical grounds since the Shannon entropy was derived in the context of probabilistic measures only. For this reason, some scientists prefer to replace the Shannon entropy by other types of performance indices that were derived in the context of the particular measure chosen [32], [39].

The size of the input/output matrix grows with the increasing complexity of the mask, and consequently, the number of legal states of the model grows quickly. Since the total number of observed states remains constant, the frequency of observation of each state shrinks rapidly, and so does the predictiveness of the model. The entropy reduction measure does not account for this problem. With increasing complexity, H_r simply keeps growing. Soon, a situation is encountered where every state that has ever been observed has been observed precisely once. This obviously leads to a totally deterministic state transition matrix, and H_r assumes a value 1.0. Yet, the predictiveness of the model will be dismal, since in all likelihood already the next predicted state has never been observed, and that means the end of forecasting. Therefore, this consideration must be included in the overall quality measure.

From statistical considerations, in any class analysis, ideally each possible discrete state is recorded at least five times[37]. Therefore, an observation ratio, O_r, is introduced as an additional contributor to the overall quality measure[33] :

$$O_r = \frac{5 \cdot n_{5x} + 4 \cdot n_{4x} + 3 \cdot n_{3x} + 2 \cdot n_{2x} + n_{1x}}{5 \cdot n_{leg}} \tag{4.11}$$

where :

n_{leg} = number of legal input states

n_{1x} = number of input states observed only once

n_{2x} = number of input states observed twice

n_{3x} = number of input states observed thrice

n_{4x} = number of input states observed four times

n_{5x} = number of input states observed at least five times

If every legal input state has been observed at least five times, O_r is equal to 1.0. If no input state has been observed at all, i.e., no data is available, O_r is equal to 0.0. Thus, O_r can be used as a quality measure. With the observation ratio, O_r, the overall quality of a mask, Q_m, is then defined as the product of its uncertainty reduction measure, H_r, and its observation ratio, O_r as:

$$Q_m = H_r \cdot O_r \tag{4.12}$$

The optimal mask is the mask with the largest Q_m value.

Once the optimal mask has been found and the corresponding Finite State Machine (FSM) generated, forecasting future system behavior is almost trivial. All that needs to be done is compare newly observed input patters with those stored in the FSM, and read out the corresponding output patterns. If the FSM is not totally deterministic, i.e., if for the same input pattern several different output patterns have been observed in the past, then there are several choices :

- All possible outcomes along with their previous relative observation frequencies can be reported.
- Reporting is limited to the most probable outcome.
- A random number can be generated and any one of the previous observations can be predicted with the correct statistical probability of occurrence.

The relative frequency of occurrence of an output pattern for any given input pattern can be used as a measure of correctness of the prediction made. The prediction contains an error and also, the quantity that measures the correctness of the prediction contains an error itself. The FIR methodology has been successfully used in a number of different applications [40, 41].

The optimal mask is applied to the given raw data matrix resulting in an input/output matrix. Since the input/output matrix contains functional relationships within single rows, the rows of the input/output matrix can be sorted in alphanumerical order. The result of this operation is called the input/output behavior matrix of the system. The input/output behavior matrix is a finite state machine. For each combination of input values, it shows which output is most likely to be observed.

4.3 Inductive Simulation

The next module is the *qualitative simulation* engine. FIR makes predictions by comparing the newly observed input pattern with all the input patterns in the experience data base (the training data), and finds the *five nearest neighbors*. It then predicts the most likely class and side values, and calculates the membership value as a weighted average of the membership values of the five nearest neighbors. In this way, reasoning is done using the (discrete) class and side values only, whereas the concrete quantitative information is preserved by interpolating among the (real–valued) membership functions of the five nearest neighbors.

4.3.1 Five Nearest Neighbors

The second stage of Qualitative Simulation is a statistical stage. In this stage, a prediction is made of a most likely fuzzy membership value for the output. The fuzzy membership value is a distance weighted average of the five nearest neighbors in the experience database, whereby the distance function is computed in the input space, and the interpolation is done in the output space. The mathematical details of what really happens are explained below.

Since FIR deals with multi-input/single-output (MISO) systems exclusively, each state consists of *m-input* variables and a single *m-output* variable. The first problem to be considered is one of normalization. Since the different *m-inputs* can represent arbitrary physical or other quantities, their absolute values can be vastly different from one another. In order to create a meaningful metric of proximity in the input space, it is necessary to normalize the input variables. This is accomplished using a pseudo-regeneration of the fuzzified input variables:

$$pos_i = class_i + side_i \; . \; (1.0 - Memb_i) \tag{4.13}$$

where the class values are assumed to be integers starting from '1' representing the lowest class, and the side values are also integers assuming the values '-1' representing the logical value 'left', 0 representing the value 'centre', and '+1' representing the value 'right'. The index i represents the i^{th} input variable in the input state of the current observation. The position value, pos_i can be viewed as a normalized pseudoregeneration of the i^{th} input variable. Irrespective of the original values of the input variable, pos_i assumes values in the range $[1.0, 1.5]$ for the lowest class, $[1.5, 2.5]$ for the next higher class, etc.

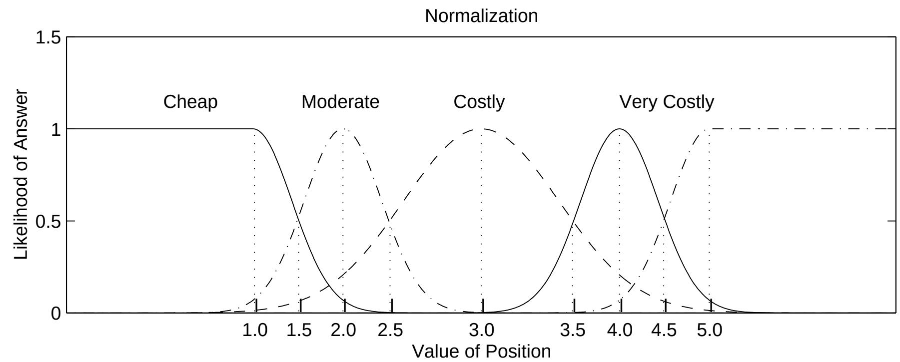

Figure 4.3: Normalization of the input variables

Fig. 4.3 shows the normalization of an input variable. Here the price of a food variable is normalized.

Similarly,

$$pos_{ij} = class_{ij} + side_{ij}\ .\ (1.0 - Memb_{ij}) \tag{4.14}$$

represents the normalized pseudo-regeneration of the i^{th} input variable of the j^{th} nearest neighbor in the experience database.

$$\mathbf{pos} = [pos_1, pos_2, \ldots, pos_n] \tag{4.15}$$

is the position vector representing the current input state, assuming that the system to be modeled contains n m-inputs, and

$$\mathbf{pos_j} = [pos_{1j}, pos_{2j}, \ldots, pos_{nj}] \tag{4.16}$$

represents the corresponding position vector of the j^{th} nearest neighbor.

The distance between the current input state and its j^{th} nearest neighbor is computed as:

$$dis_j = \| \ \mathbf{pos} - \mathbf{pos_j} \ \| \tag{4.17}$$

It is necessary to avoid distance values of 0.0:

$$d_j = max(dis_j, \epsilon) \tag{4.18}$$

where ϵ is the smallest number that can be distinguished from 1.0 in addition.

$$s_d = \sum_{j=1}^{5} d_j \tag{4.19}$$

is the sum of the distances of the five nearest neighbors, and:

$$d_{rel_j} = d_j / s_d \tag{4.20}$$

are the relative distances. By applying this algorithm either to the entire experience data base or a suitable subset thereof, the five nearest neighbors can be determined while simultaneously computing their distance function.

The interpolation is done in the output space. Absolute weights are computed as

$$w_{abs_j} = 1.0 / d_{rel_j} \tag{4.21}$$

and

$$s_w = \sum_{j=1}^{5} w_{abs_j} \tag{4.22}$$

is the sum of the absolute weights. Hence the relative weights can be computed as :

$$w_{rel_j} = w_{abs_j} / s_w \tag{4.23}$$

Using this information, the membership value of the predicted output is determined as:

$$Memb_{out} = \sum_{j=1}^{5} w_{rel_j} \ . \ Memb_{out_j} \tag{4.24}$$

If one of the observations in the experience database coincides, by chance, with the new observation, its relative distance value will be very close to 0.0, whereas that of the other four neighbors will be considered larger. Consequently, this data record will have a predominant influence on determining the membership value of the output. On the other hand, if the five nearest neighbors are all approximately equally far away from the new observation, the relative distance values will all be approximately 0.2, and each of the corresponding records in the experience database will have equal weight in determining the membership value of the new output.

4.4 Defuzzification

In the *defuzzification* module, the predicted class, side, and membership values are converted back to real–valued quantitative predictions using the inverse operation to the fuzzification.

FIR's *confidence* measure has two components. FIR measures the distance between the new data point to be predicted from its five nearest neighbors in the input space. If the distance is small, only little interpolation needs to be done, and FIR is more confident that the proposed prediction is accurate. Secondly, it looks at the dispersion among the outputs of the five nearest neighbors. If the dispersion is small, FIR is confident that it can predict accurately the new output value. If it is large, it cannot know which of the neighbors is right, and therefore, assigns a lower value to the confidence measure [9].

FIR, although internally operating as a *qualitative* technique (reasoning about discrete class values), looks from the outside like a *quantitative* technique thanks to its fuzzification and defuzzification engines. FIR is therefore compatible with quantitative approaches, such as System Dynamics. FIR submodels can be easily embedded in SD models, as proposed in the advocated methodology.

CHAPTER 5

Food Demand Model

The proposed methodology is a blend of the knowledge based models and the pattern recognition models. The methodology is based on System Dynamics. The problem concerning the determination of relationships between different variables which is not very easy to determine using conventional System Dynamics is overcome by the use of Fuzzy Inductive Reasoning.

5.1 The Naïve Model

The methodology was put to test with a simple SD model [42]. Fig. 5.1 shows the highly simplified *System Dynamics* macroeconomic model that was used to predict aspects of U.S. food demand in the 20^{th} century. This is a rather naïve model, but it was found to serve for illustrative purposes.

The amount of food available on the market, a *level* variable, depends on food production and consumption, two *rate* variables. Both food production and consumption are heavily influenced by the food prices, an *auxiliary* variable. The food prices depend primarily on the amount of food currently on the market, but also on the state of the economy, here reflected by another auxiliary variable, the inflation. Both the economy (here represented by the unemployment rate) and the population

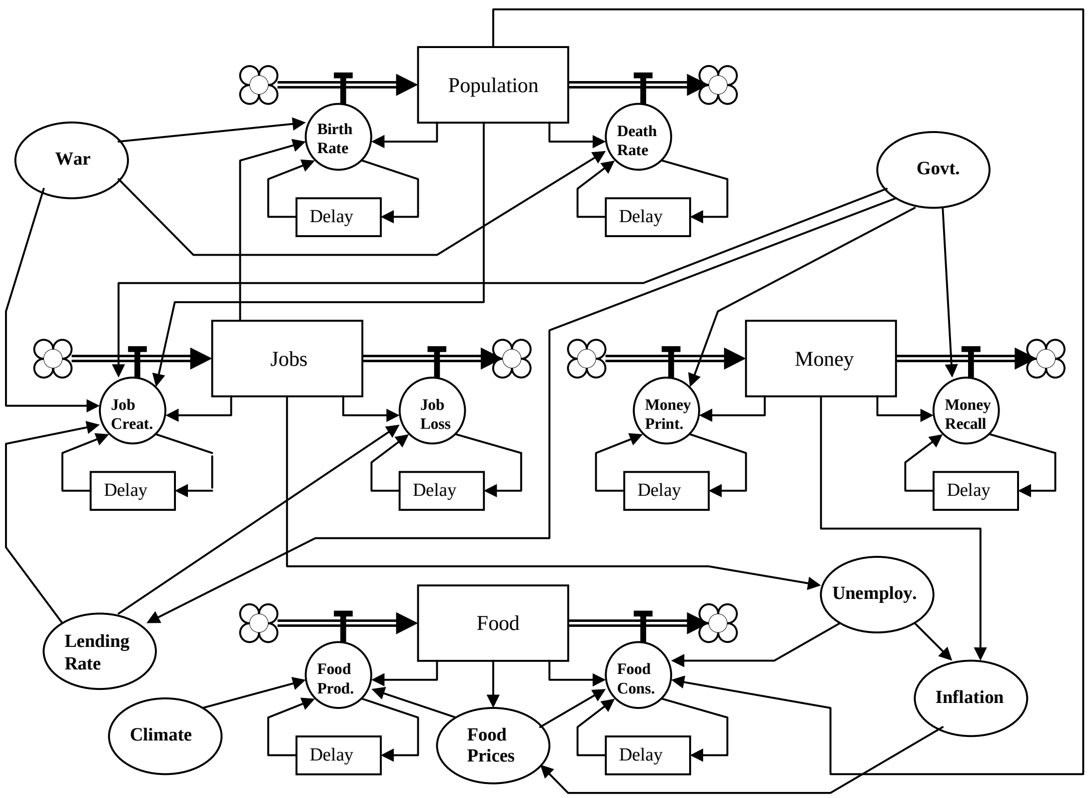

Figure 5.1: System Dynamics Model of U.S. Food Demand

(the prospective buyers) influence food consumption. The two delay boxes represent secondary dynamics. They reflect the inertia of both food production and consumption, as neither of them will change abruptly over night.

In order to be able to predict food demand and supply, it is necessary to know something about the state of the economy and population dynamics [43]. The economy is represented by two level variables, the number of jobs, and the volume of money. Although there exist tight interactions between these variables, they can be modeled almost independently. The reason is that the number of jobs (or rather the unemployment rate) is a *controlled variable*. The government tries to keep it always around 5%. If the unemployment rate climbs, the lending rate is lowered; this provides incentive for the construction industry, which absorbs the surplus unemployed. If the unemployment rate decreases much below the 5% level, the lending rate is increased, which dampens the construction industry, which makes more workers unemployed. The reason is simple: if there are too few unemployed workers, the employers have to raise the salaries in order to attract employees, which leads to an increased inflation rate.

Knowledge of the population dynamics, and in particular the number of young adults, helps with predicting unemployment. Once the unemployment is predicted, it can in turn be used as a driver for predicting the inflation rate.

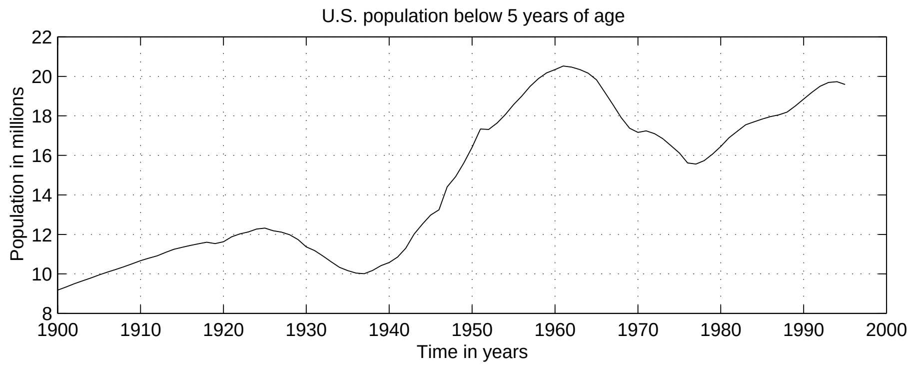

Figure 5.2: U.S. Toddler Population Dynamics in the 20th century

The population dynamics are influenced by the state of the economy, but only to a minor extent. Only during the great recession of the 30s, could a notable change in birth rate patterns be observed as a consequence of a poor economy. The other anomaly are the few years starting close to 1960, which was around the time when the birth control pill was introduced. It also coincides with the years of the Vietnam war, when many potential fathers were abroad for years in a row, and therefore, could not sire children in the U.S. The U.S. toddler population dynamics are shown in Fig. 5.2.

To summarize, it makes sense to postulate a hierarchical model, whereby the population dynamics (level 1) are explained only from their own past, whereas the economy (level 2) is explained from its own past *and* the already predicted population dynamics. Finally, the food model (levels 3/4) is predicted by its own past *as well as* all the previously predicted variables of the population dynamics and economy layers. The three–layer hierarchy is shown in Fig. 5.3. For reasons to be explained

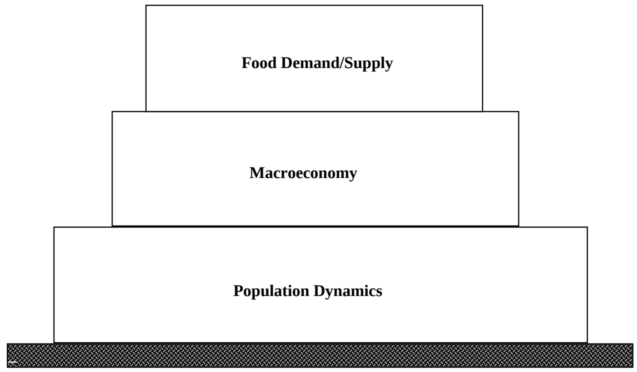

Figure 5.3: The Layered Model Architecture

later, the top two layers (food demand and supply) were lumped into a single layer for the purpose of this research effort.

An additional advantage of the layered architecture is that the models at levels 1 and 2 are *generic models* that can be created once and for all. They do not depend on the application to be predicted at level 3, i.e., if there should suddenly be a need to predict the demand for used cars rather than powdered milk, the bottom two layers of the architecture will remain the same. Only the top layer changes.

5.1.1 The Population Dynamics Layer

The *System Dynamics* methodology stipulates that equations need to be found that predict the rate variables. The level variables then follow by integration. Yet in practice, it is much easier to gain access to good measurement data for level variables than for rate variables. Therefore, it may be easier to predict the level variables directly.

The data available to us for this study include the total U.S. population recorded (estimated) annually since 1910. They also include the percentages of the population in different age brackets, as well as the demographic distribution.

The idea was to create a *Fuzzy Inductive Reasoning* model that predicts each of these variables from its own past and from past values of the other population dynamics variables. For example the number of newly borns depends on the population of childbearing age, whereas the population of teenagers depends on previous values of the population of toddlers, etc. The data from 1910 until 1970 were to be used as training data, whereas the remaining 25 years should be used to validate the model.

One problem was that FIR is totally pattern–based, i.e., FIR can only predict what it has been shown before. Yet, the population is non–stationary. It grows almost exponentially. FIR certainly is not capable of predicting such a growth variable. This problem was solved with a simple trick. Since the population grows almost exponentially, it makes sense to postulate the model:

$$\frac{dP}{dt} = k(t) \cdot P \tag{5.1}$$

If k were a positive constant, the population, P, would grow exponentially. By allowing $k(t)$ to be time–dependent, the actually observed population dynamics can be modeled. However, whereas P is a growth variable, k is essentially stationary.

The population derivative can be approximated as:

$$\frac{dP}{dt} \approx \frac{P(n) - P(n-1)}{\Delta t} \tag{5.2}$$

and since $\Delta t = 1$:

$$k(n) \approx \frac{P(n) - P(n-1)}{P(n)} \tag{5.3}$$

Since $k(t)$ is stationary, k can be predicted from its own past and from past values of other population dynamics variables. However, once $k(n+1)$ is predicted, Eq.(5.3) can be shifted by one year into the future and solved for $P(n+1)$:

$$P(n+1) \approx \frac{P(n)}{1.0 - k(n+1)} \tag{5.4}$$

Fig. 5.4a shows one, three, and five year predictions of the U.S. toddler population plotted together with the observed data for the years 1970 until 1995. Fig. 5.4b shows the average relative error as a function of the number of years to be predicted. Similarly good results have been obtained for the populations in age brackets and for the demographics.

The forecast is non–trivial, because the onset of the forecasting period coincides with the anomaly of the Vietnam war and the introduction of the birth control pill. Several trivial predictions were tried also. Their prediction errors were always larger by at least a factor of four in comparison with the FIR predictions. The errors that were calculated were not relative errors.

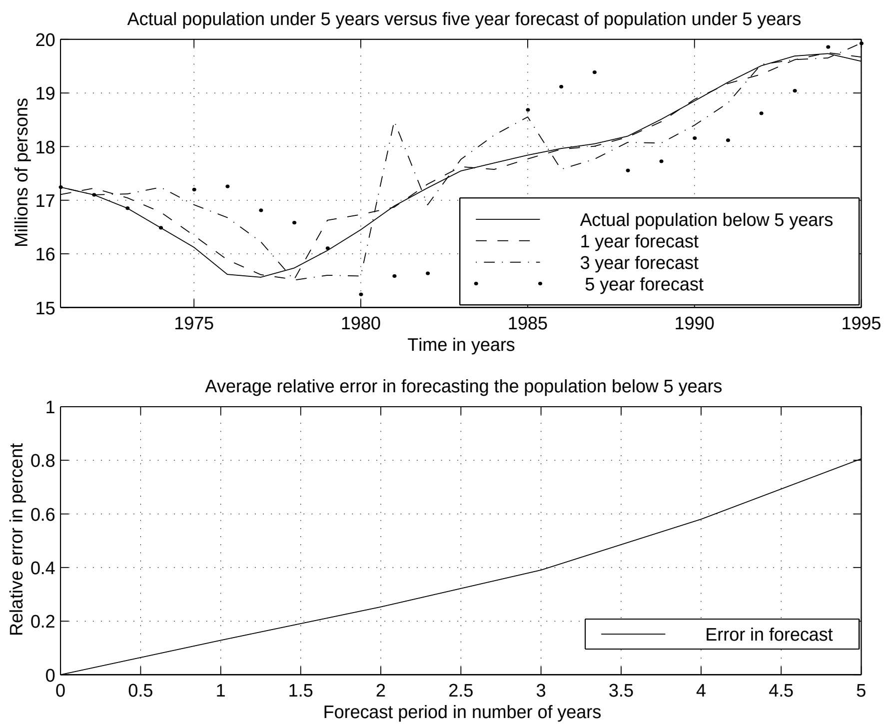

Figure 5.4: Toddler Population Forecast and Error Curves

5.1.1.1 Prediction Error

The errors that were calculated consist of four different components. The first component measures the accuracy with which the forecast predicts the *mean value* of the time series:

$$err_{mean_i} = \frac{abs(mean(y(t)) - mean(\hat{y}_i(t)))}{max(abs(mean(y(t))), abs(mean(y_i(t))), \epsilon)} \tag{5.5}$$

It is a relative error with a fudge factor, ϵ that is used to make sure that the formula never fails. But in reality, the fudge factor will never come into play, because it will only be applied when:

$$mean(y(t)) = mean(y_i(t)) = 0.0 \tag{5.6}$$

in which case the numerator is exactly equal to zero.

The second component measures the accuracy, with which the *standard deviation* of the time series is being predicted:

$$err_{std_i} = \frac{abs(std(y(t)) - std(\hat{y}_i(t)))}{max(abs(std(y(t))), abs(std(y_i(t))), \epsilon)} \tag{5.7}$$

The same applies as above w.r.t the fudge factor.

For the third and fourth component, the time series and its prediction are jointly normalized to the range [0.0,1.0]. Let:

$$y_{max} = max(y(t), y_i(t)) \tag{5.8}$$

where the *max*-operator is applied to the concatenated series consisting of $y(t)$ and $y_i(t)$, and similarly:

$$y_{min} = min(y(t), y_i(t)) \tag{5.9}$$

Normalized time series can be computed as:

$$y_{norm}(t) = \frac{y(t) - y_{min}}{max(y_{max} - y_{min}, \epsilon)} \tag{5.10}$$

and similarly:

$$y_{norm_i}(t) = \frac{y_i(t) - y_{min}}{max(y_{max} - y_{min}, \epsilon)} \tag{5.11}$$

Since the two curves have been normalized, it is now possible to use the absolute errors. The pointwise absolute error between the two curves $y(t)$ and $y_i(t)$ can be computed as:

$$err_{abs_i}(t) = abs(y_{norm}(t) - y_{norm_i}(t)) \tag{5.12}$$

It is also possible to define the pointwise *similarity* between the two curves as:

$$sim_i(t) = \frac{min(y_{norm}(t), y_{norm_i}(t))}{max(y_{norm}(t), y_{norm_i}(t), \epsilon)} \tag{5.13}$$

where, this time around, the *min*- and *max*-operators are being applied element-wise rather than to the concatenated time series.

Using pointwise similarity, a pointwise *similarity error* can be defined as:

$$err_{sim_i} = 1.0 - sim_i(t) \tag{5.14}$$

The averaged absolute and similarity error is then defined as:

$$err_{avg_i} = mean(err_{abs_i}(t)\ +\ err_{sim_i}(t)) \tag{5.15}$$

Finally, the total error, in percentage, is the sum of the four components multiplied by 25.0:

$$err_{tot_i} = 25.0\ \cdot\ (err_{mean_i}\ +\ err_{std_i}\ +\ err_{avg_i}) \tag{5.16}$$

The factor 25.0 is justified, because there are four separate components that all measure different aspects of one and the same thing. Each one of them is usually in the range of [0.0, 1.0] (although some errors could be larger than 1.0 at times), thus, the accumulated total error should be somewhere in the range [0%,100%] most of the time.

The error formula is a compromise that was developed over a long period of time by J. López [44], and that saw many revisions over the years. It results in a quantification of success or failure of predictions that is quite consistent with the intuitive understanding of success or failure that a human observer would have when comparing $y(t)$ and $y_i(t)$ by the naked eye.

5.1.2 The Economy Layer

Whereas the population dynamics layer is quite sophisticated, the economy layer is still rudimentary. Until now, *unemployment rate* and *inflation* were used to represent the state of the economy. Other important economic drivers, such as import/export statistics and national debts were ignored.

Fig. 5.5a shows the three year prediction of the inflation, represented by the total cost of food spent per person per year, predicted from its own past alone and predicted using the already predicted population dynamics as additional input. Fig. 5.5b shows the corresponding error curves as a function of the years of prediction. The lower level aided the prediction. When FIR made a forecast with variables from the population layer along with variables from layer two, the error was much lower than the forecast made with layer two variables alone. This justifies the layered approach. The same strategy was used for the inflation (another growth variable) that had been used in the case of the population dynamics.

Fig. 5.6a shows the three year prediction of the unemployment rate predicted only from its own past, and also predicted using the predicted population dynamics as well as the predicted inflation rate as additional inputs. Fig. 5.6b shows the corresponding relative error curves as a function of the years of prediction. As postulated, knowledge of the lower levels helped reduce the prediction error.

5.1.3 The Food Demand/Supply Layer

Food data were available in four different categories [45]:

1. dairy products,

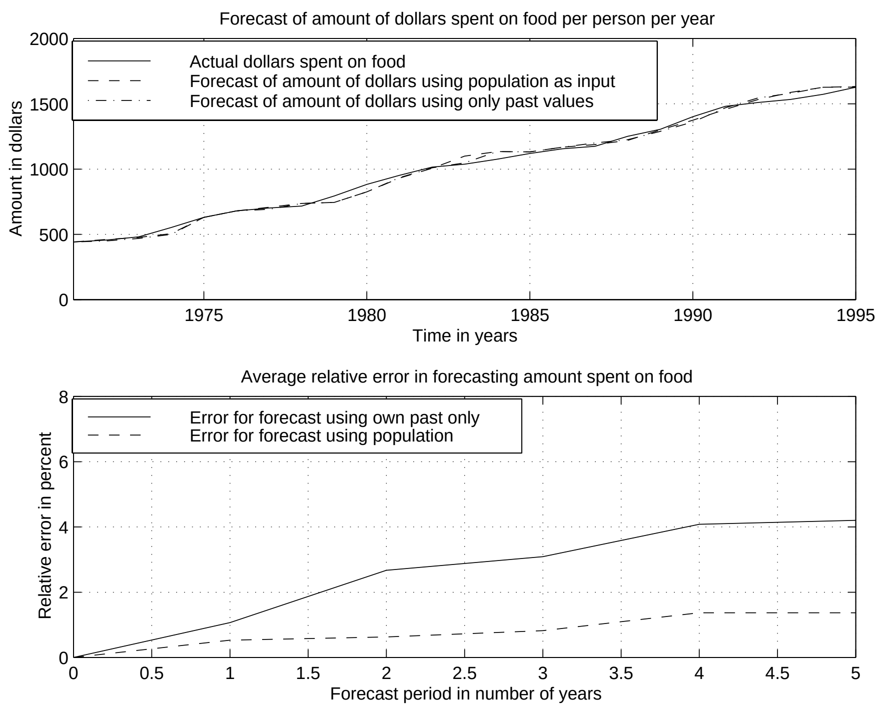

Figure 5.5: Inflation Forecast and Error Curves

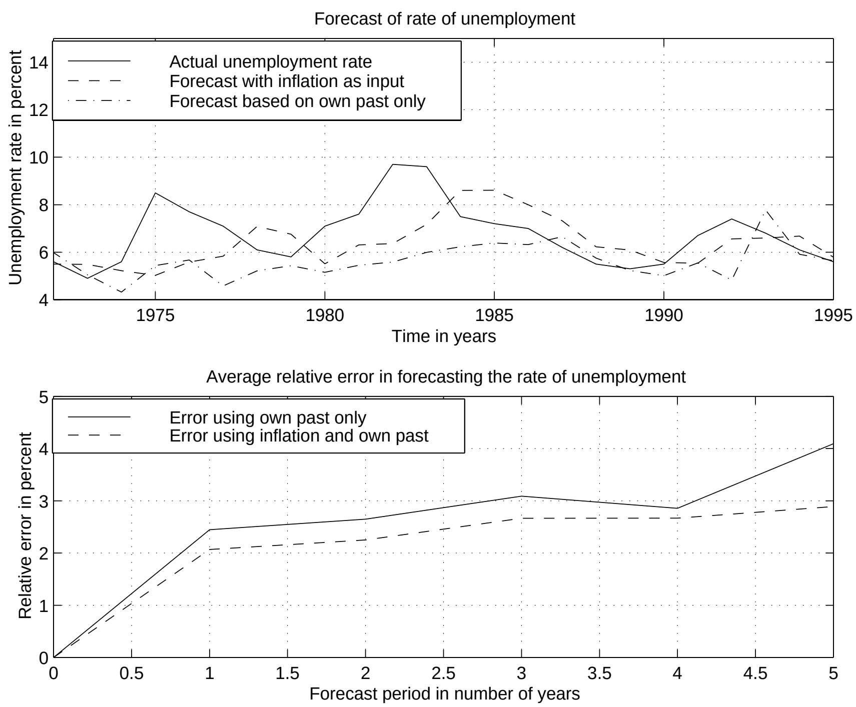

Figure 5.6: Unemployment Forecast and Error Curves

2. meats, fish, and poultry,
3. fruits and vegetables, and
4. miscellaneous foods.

Each category is subdivided further into individual products. For each product, there are available data about the amount of consumed goods as well as the prices that they were sold at. Hence there are lots of data available to base the food–layer model upon.

Since this was a first study to ascertain the working of the methodology, only a few individual products, one from each category, were more or less arbitrarily picked out. Models were obtained for:

1. fresh milk and cream,
2. fish,
3. fresh vegetables, and
4. cereal, grains and bakery products.

Food prices and volume cannot be decoupled from each other. The food prices depend on the amount of food available, yet the food consumption depends on the price. One problem, of course, is that the time constants for food prices and consumption are much shorter. The food prices change over the year, and so does consumption. With food consumption being lumped and food prices being averaged over an entire year, important dynamics are being lost.

In order to come up with a decent model at such a high aggregation level, the relationship between prices and volume must be considered to be immediate. The food prices depend on the *current* food consumption, and vice versa. In order to tackle this interdependence problem, the following approach could be used. The prices are first fixed at the previous year's level, and the amount of food sold at those prices, given the current population dynamics and economy data, can be calculated. From the estimated consumption, the profit of the producers can be obtained. This model is embedded in an optimization layer, in which the food prices are treated as parameters, and the profit is the performance index to be maximized.

Yet, in this first study, a much simpler approach was chosen. It was assumed that:

$$\begin{aligned} price(n) &= f_p(price(n-1), volume(n-1), inflation(n)) \\ volume(n) &= f_v(volume(n-1), price(n), inflation(n), population(n)) \end{aligned} \tag{5.17}$$

where n means "up to year n", and $n-1$ signifies "up to year $n-1$," i.e., dependences on earlier years are always allowed as well. In this way, the costly optimization was avoided. However, the assumption that this year's food prices depend on last year's volume is certainly not justifiable, and will probably lead to a significant reduction in overall prediction quality.

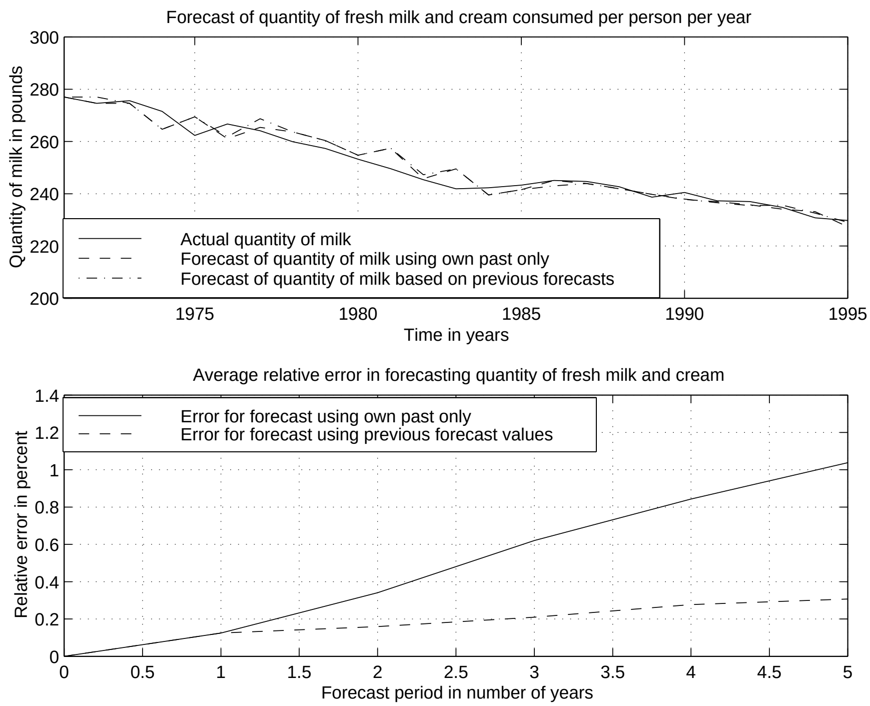

Figure 5.7: Fresh Milk and Cream Forecast and Error Curves

As in the case of the economy layer, predictions using only the past of the variable to be predicted, i.e.:

$$volume(n) = f(volume(n-1)) \tag{5.18}$$

were used as reference. Fig. 5.7a shows the three year prediction of the volume of fresh milk and cream predicted using Eqs.(5.17) on the one hand, and Eq.(5.18) on the other. Fig. 5.7b shows the prediction error as a function of the number of years of prediction for the two models.

Again, the more complex model performed better in the case of fresh milk and cream, but this wasn't the case with all variables. Also, the improvement is not very large.

In order to obtain yet better results, it will be necessary to implement the aforementioned optimization scheme. Also, it may make sense to allow past values of the overall amount of food in a given category to be used as additional input for predicting the consumption of a particular food item within the same category. This modification will allow to take into account replacement foods. Whereas the total intake of calories per person per year is constant, people may replace one food item by another within the same category if the prices of the individual items change.

Annual data from 1910 to 1970 were used as training data, i.e., only 61 data records were available for training the model. The model could be expected to work better if we had a bigger training set.

The data deprivation problem was overcome by interpolating between the measurement data points using e.g. spline interpolation (available in Matlab). Clearly, adding more data points by means of interpolation doesn't add any additional information to the data set. The new data are *derived data*, and one shouldn't expect that the model would improve as a consequence of such data. Yet, they may actually help for two reasons.

First, FIR uses the five nearest neighbors for predicting fuzzy membership values. Because there are only 61 data records in the training data base, the neighbors will be far away, and therefore, a lot of interpolation has to be done. By generating e.g. three new artificial data points for each measured one, the system would now have 241 data points to work with. Therefore, the five nearest neighbors will be much closer to the current data point, and consequently, much less interpolation will be needed.

Second, FIR tries to have at least five recordings for each discrete state. With 61 data records, only 12 different states can be recorded 5 times each. This means that FIR will always pick extremely simple models with one to three ternary input variables only. If FIR is offered four or five possible input variables, it will pick the most relevant ones, and discard the others, although they might carry useful information. If the data deprivation problem is reduced, FIR might pick a mask of higher complexity, and thereby also exploit the information contained in less important variables.

5.2 The Enhanced Macroeconomic Model

The flexibility offered by FIR might then be used by us to describe an enhanced macroeconomic model that can be used for any commodity in general. A new revised model structure that could be used for any industry is now proposed. Most of the drawbacks of the simple model were overcome here. The overall three-tier layer structure and the dependence of the different variables on each other can be described by Fig. 5.8.

The advantage of such a model structure is that the third layer can be replaced by any commodity or service without affecting the other two layers. The population layer in the new model is the same as the one that was used in the simple model. The enhanced model differs from the simple model in layers two and three. In layer two of the simple model, the total dollars spent on food products was taken as a measure of the inflation rate. While this may hold true for making forecasts in the food products industry, this might not necessarily be the case for other industries. The newly proposed model has a much more stable economy layer consisting itself in a pseudo two–layer structure. The per capita income, wage rate and unemployment rate are dependent on the population layer and the inflation rate is measured from the consumer price index and the producer price index, which are dependent on the population layer and the other economy layer variables also. The Consumer Price Index is a measure of the changes in price over time. The Consumer Price Index reflects the effects of inflation and government action to control it. The Producer Price Index is a measure of the prices of commodities in the market. The third layer makes prediction on the price of food products and the quantity of food consumed. External functions such as the total dollars spent on food and the quantity of food consumed in a particular group are input into the structure. The proposed model structure is not specific to any industry. The first two layers are the generic layers

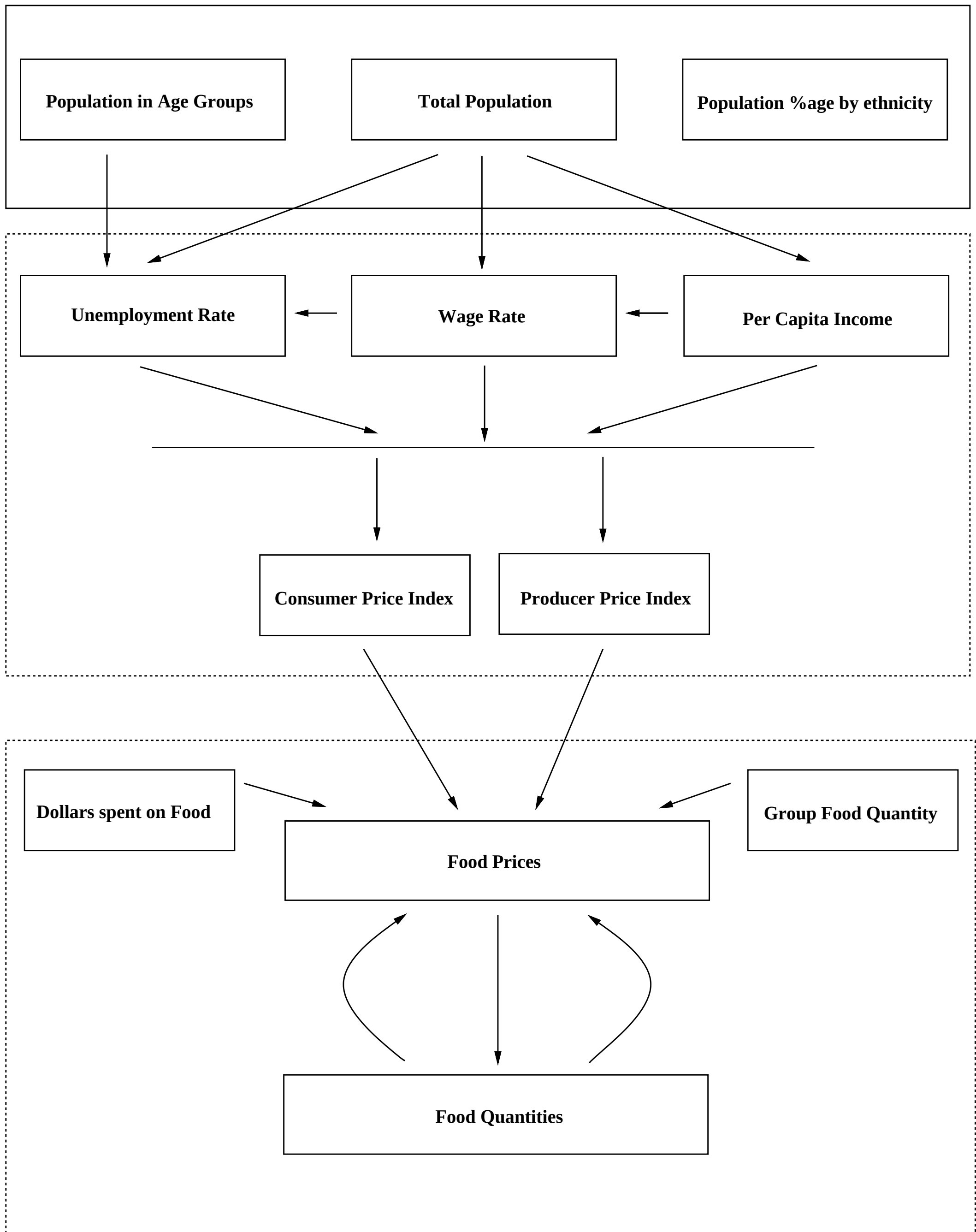

Figure 5.8: Hierarchical Structure of Food Demand Model

and the third layer can be changed to fit any industry of choice. If we needed to make a prediction on the number of cellular phones in use in the United States, the third layer needs to be replaced with variables that are specific to the cellular market in the U.S.

5.2.1 Population Layer

The first layer of the model is the Population layer. The population layer is broken up into 3 sections. The first section makes predictions on the population in different age groups. The second, on the total population and the third on the ethnicity in population. Fig. 5.9 shows the SD model for the distribution of population in different age groups.

The toddler population is dependent on the young adult population. The young adult population has been assumed to be the population in the age group of 15-24 years. As the average marriage age grew, the population between the ages of 25-34 years started to also contribute to the toddler population in recent years. This has not been accounted for here for the sole reason that this trend is quite recent.

The population in each age group is dependent on the population in the earlier age group and also on its own value. The dependency on its own value is shown with the help of the delay box. Each level variable is driven by the previous level variable.

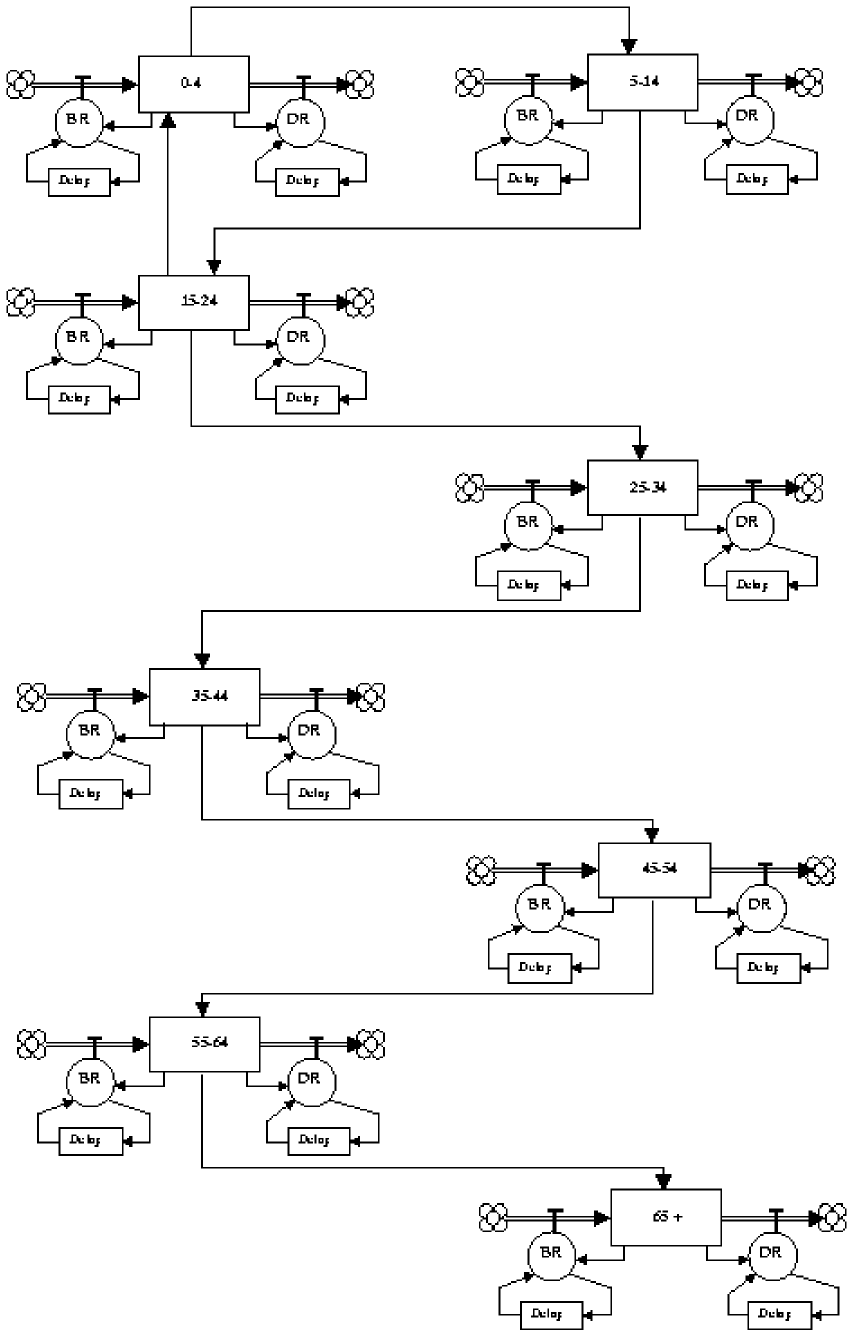

Figure 5.9: System Dynamics Model for Population in Different Age Groups

The total population can be represented as the net sum of the total input into the population and the outflow out of the population. The inputs are the immigration rate and birth rate and the outputs are the emmigration rate and the death rate. Obviously, the total population is the sum of the population in the different age groups. This is shown in Fig. 5.10.

A SD model was also developed for the distribution of population based on the demographics. This helped make forecasts based on percentages of population that were white, black and neither black nor white. These results were not used in the example described, as the consumption of food products is affected very little by the demographics. Since the average income levels are still different among different ethnic groups, one might think that demographics would affect food consumption, but the simulations performed showed this influence to be of second order small. However, demographics probably would affect other economic variables, such as the percentage of used cars bought relative to new cars.

5.2.2 Economy Layer

The proposed economy layer is driven by the population layer. The economy is represented by seven level variables. The variables are split into two pseudo–layers. The per capita income, wage rate and the rate of unemployment are directly dependent on the population layer; the consumer price index and the producer price index,

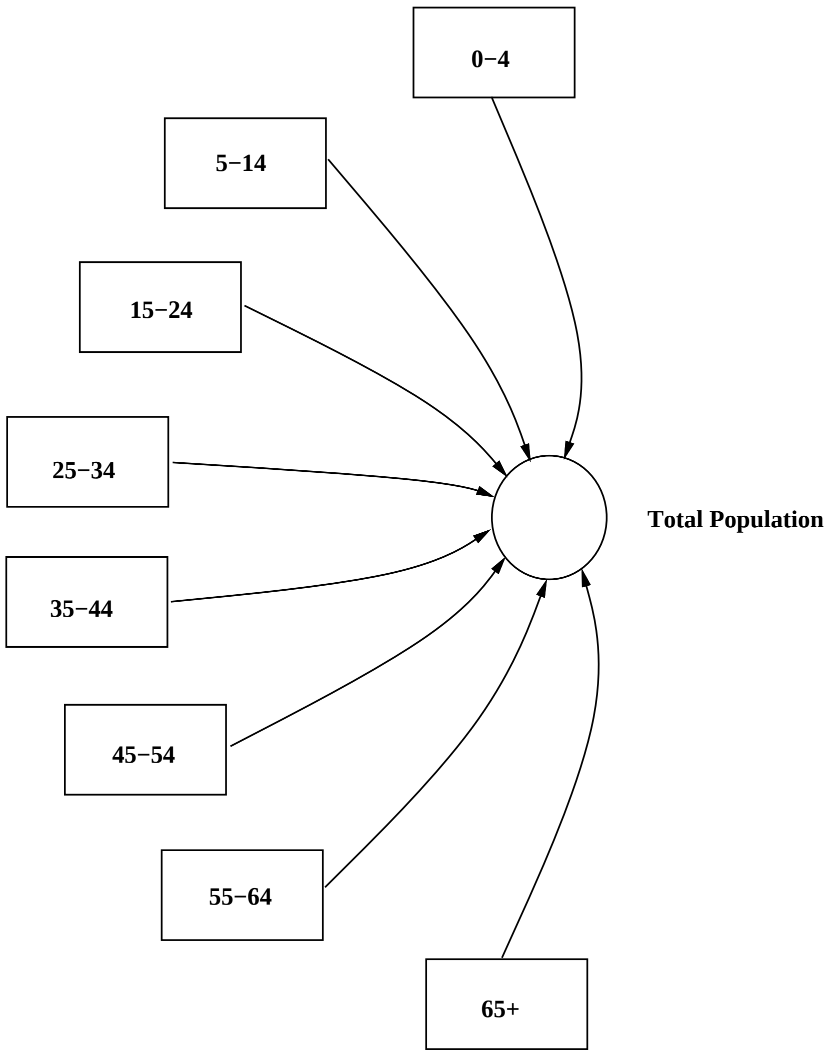

Figure 5.10: System Dynamics Model for Total Population

which are indicators of the state of the economy, are dependent on the population layer as well as the earlier forecast economy variables.

The per capita income is a good measure of the buying capacity of the consumer, and the wage rate and the rate of unemployment represent the growth of the economy and are very closely related to each other. Together, they represent the state of the economy from a consumer's perspective.

The Consumer Price Index (CPI) describes the change in prices from one time period to another for a fixed 'market basket' of goods and services. It is usually generated monthly or annually. The CPI can be used as an economic indicator. By reflecting changes in prices, the Consumer Price Index is the primary measurement of inflation in the United States and of the perceived success or failure of government policies to control inflation.

The Producer Price Index (PPI) is another variable used by the federal government and other institutions as a special purpose index. The PPI is the oldest continuous statistical series published by the U.S. Bureau of Labor Statistics. It reflects the prices of approximately 3200 commodities. Price data are collected from sellers of these commodities and usually apply to the first large-volume transaction for each commodity. Like the CPI, it is based on shipments of commodities as measured in the industrial census and from other data.

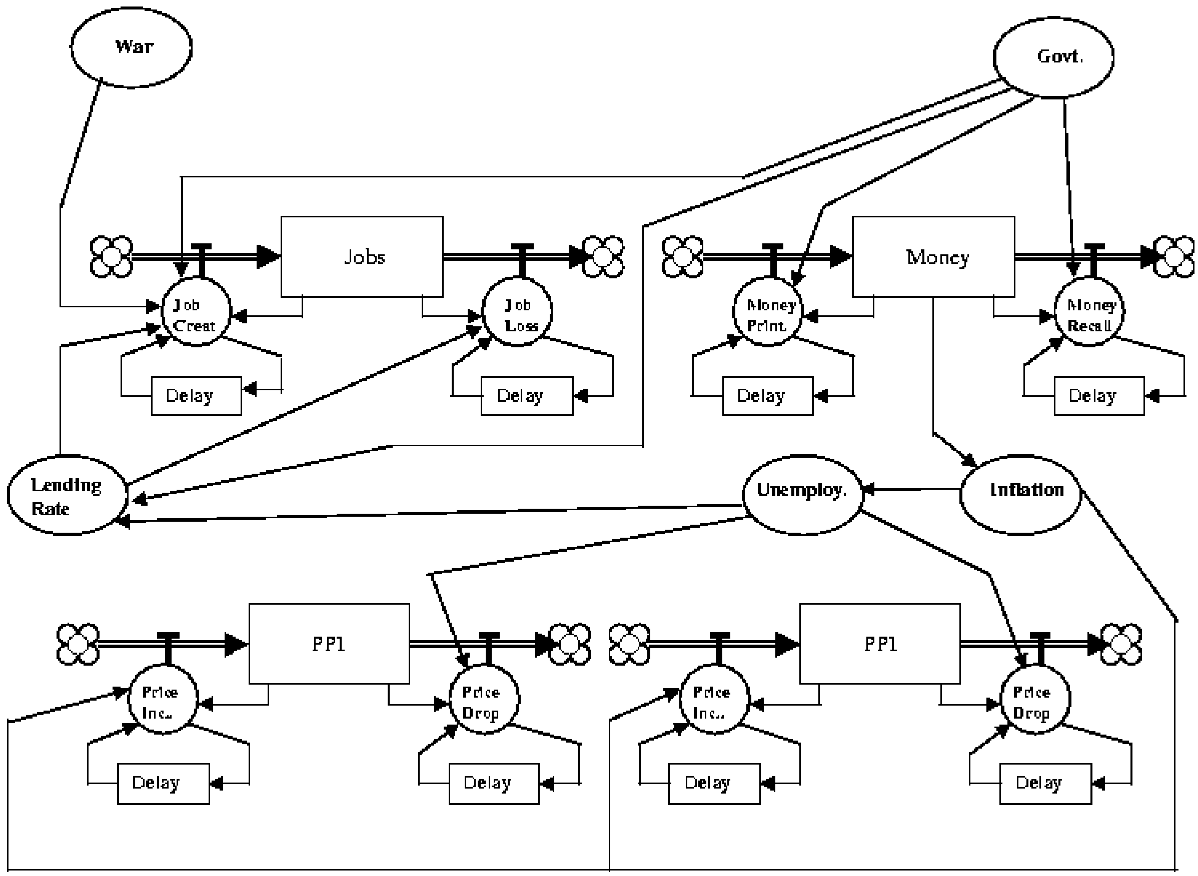

Figure 5.11: System Dynamics Model for the Economy Layer

The CPI and the PPI represent the state of the economy from a supplier's perspective. Fig. 5.11 shows the SD Model that represents the economy layer.

5.2.3 Food Demand/Supply Layer

The amount of food available in the market is a level variable that depends on another level variables, namely the price of the food items. The food items were classified into four categories and for each of these categories, the consumption quantity of any item would be proportional to the consumption quantity of the other items. This is because food items with similar calorie content and within the same price range would be consumed in almost similar quantities. Therefore, the consumption amount of the total group quantity is an external function to the level variables. These variables are also influential on the food prices. The SD model for the topmost layer is shown in Fig. 5.12.

All food products cannot be lumped together as each individual product has its own dynamics. In some cases, the price of food products is not very influential on the consumption quantity. For example, the demand for fresh milk is not very responsive to price changes, due to the perishability and high transport costs of raw whole milk. However, the demand for manufactured milk products is more responsive to price changes because these products are less bulky and perishable, and hence more easily stored and transported. This difference in the responsiveness of demand to price is the basis for the classified pricing [46].

Definitely, the prices of food components that are forecast earlier depend on the forecast of the food quantity. This loopback is achieved by an Optimization layer. In the Optimization layer, it is assumed that the food suppliers act rationally and that they have the same information available i.e., that they are able to make predictions on the demand on the basis of the current prices. It can be assumed that they all try to maximize their profits in terms of dollars earned.

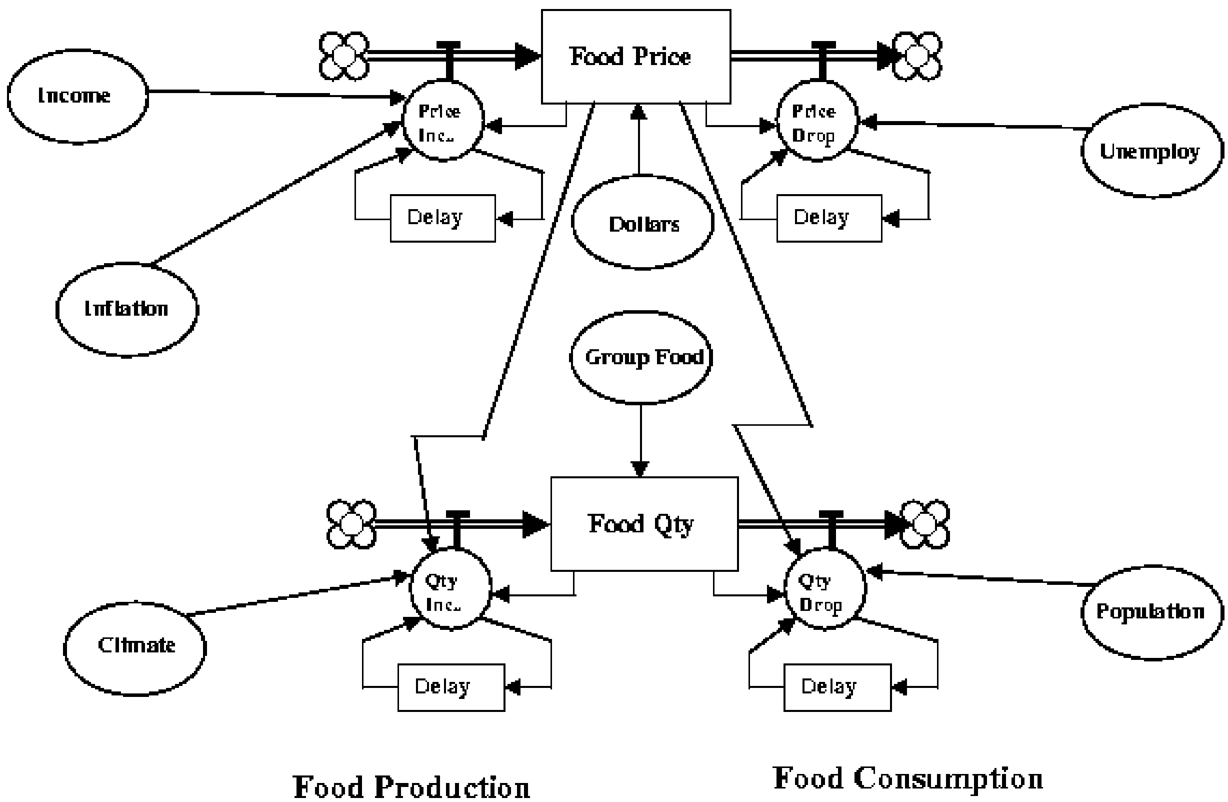

Figure 5.12: System Dynamics Model for Food Demand/Supply Layer

The food suppliers are introduced into the model. The model would predict the prices of food items given the expected demand and the pricing structure of other competitive food items. The pricing structure of other competitive food items must carry a small weight, because otherwise, the overall model may become unstable, yet it is an important incentive function that cannot be totally ignored.

This model will then be combined with the demand model. In the combined model, the prices needed by the demand model are computed using the supply model and the demand predictions needed by the supply model are computed by the demand model.

CHAPTER 6

Experimental Simulation and Results

6.1 Annual Data

The proposed model was run with annual data for each year from 1910 to 1994. The data set consists of annual time series observations over the period 1910-1994. Per-capita consumption of twenty-one food items and corresponding average retail prices for those items were constructed from several USDA and Bureau of Labor Statistics (BLS) sources. The quantity data are aggregates taken from the USDA series Food Consumption, Prices and Expenditures. Estimated retail prices corresponding to the quantity data were constructed as follows. Detailed retail price estimates that are available for 1967 were used along with the respective quantity observations to construct an average retail price per pound in 1967 for each food category (e.g., milk). For all other years, the fixed 1967 quantity weights, together with either consumer price indices for food items or average retail food prices were combined to construct a consistent retail price series for each commodity. The consumer price index (CPI) for all nonfood items is used for the price of nonfood expenditures. The income variable is per capita disposable income.

The demographic factors included in the data are the empirical age distribution for the U.S. population and proportions of the U.S. population that are white, black and neither white nor black. The estimated age distribution is based on ten-year age intervals, plus categories for children less than five-years of age and adults who are sixty-five years and older. The ethnic variables are linearly interpolated estimates of Bureau of Census figures reported on 10-year intervals.

6.2 Quarterly Data

As suggested in [42], the annual data was interpolated using the built-in *splines* interpolator in Matlab, and data was generated every three months. This data is in reality "*dummy*" data, because we have not added any knowledge to the model by increasing the number of data points. but this will help improve the forecasts as has been explained earlier.

6.3 Layer One

Fig. 6.1 shows the forecast results for the toddler population. The toddler population depends on its own value in the previous year as well as the population capable of bearing children, namely the young adult population.

Correction of Forecast Values

Since the sum of population in individual age groups must be equal to the total population, there is a possibility to correct the forecasts made in the bottommost layer. The total population can be written as sum of all the population in the

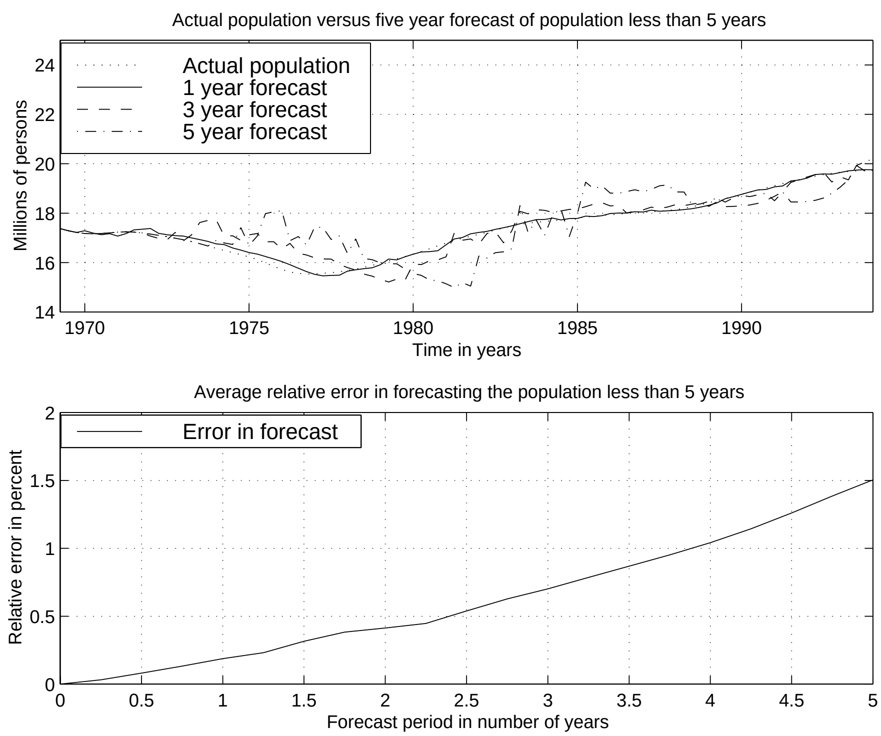

Figure 6.1: Forecast Results of Population below 5 years Using Quarterly Data

different age groups.

$$P_{sum} = P_{0-4} + P_{5-14} + P_{15-24} + P_{25-34} + P_{35-44} + P_{45-54} + P_{55-64} + P_{65+} \qquad (6.1)$$

A correction factor 'k' can then be introduced as :

$$k = \frac{P_{total}}{P_{sum}} \qquad (6.2)$$

Then the corrected forecast in each age group is

$$P_{corr} = P_{agegroup} \cdot k \qquad (6.3)$$

Fig. 6.2 shows the forecast results for the toddler population after making the necessary correction. The error in forecast dropped by a small amount after correction was incorporated. Forecasts were also made in the other age groups and all error rates were well under ten percent. As in the previous cases, the error is not a relative error measure.

Fig. 6.3 shows the forecast results for the total population.

Forecasts were also made on the percentage distribution of population based on the ethnicity. The error in the forecast is shown in Fig. 6.4. It is interesting to note that the error remains at a very low value for the first few forecast periods and then increases very steeply.

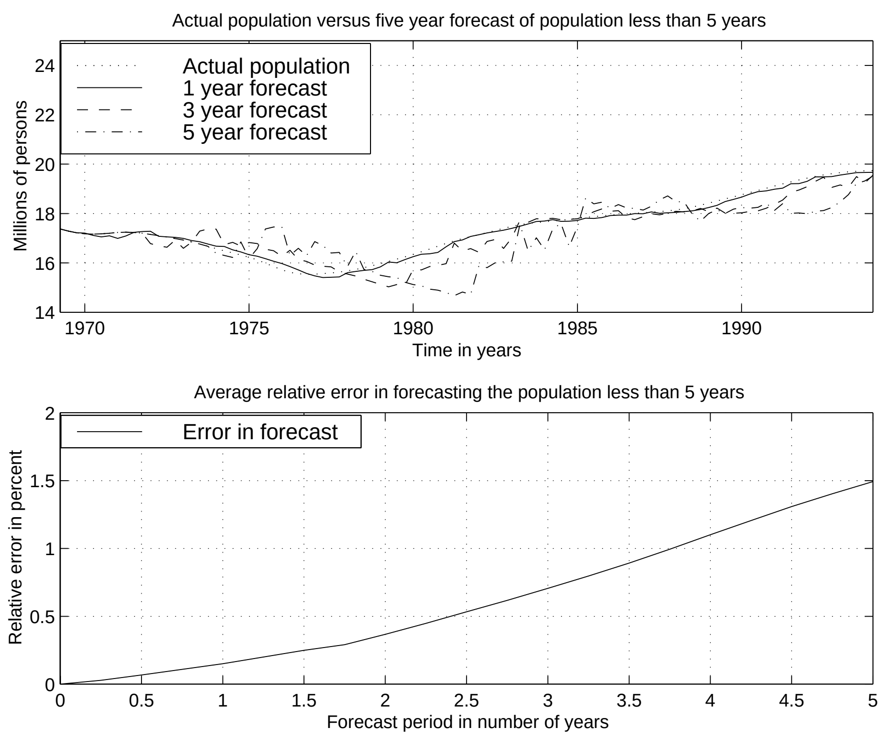

Figure 6.2: Forecast Results of Population below 5 years Using Quarterly Data and Correction

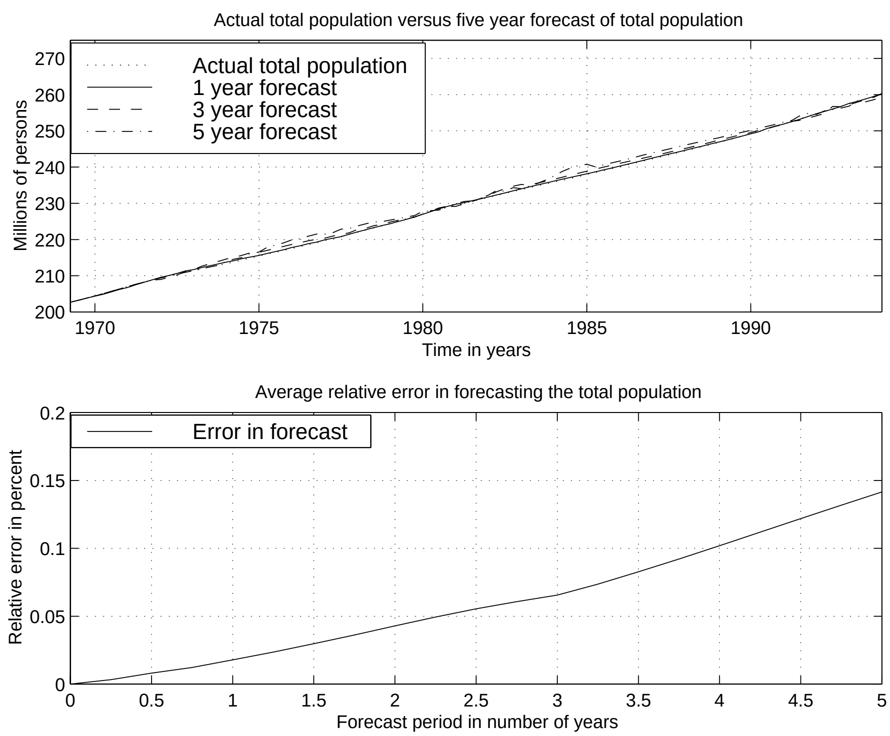

Figure 6.3: Forecast Results of Total Population Using Quarterly Data

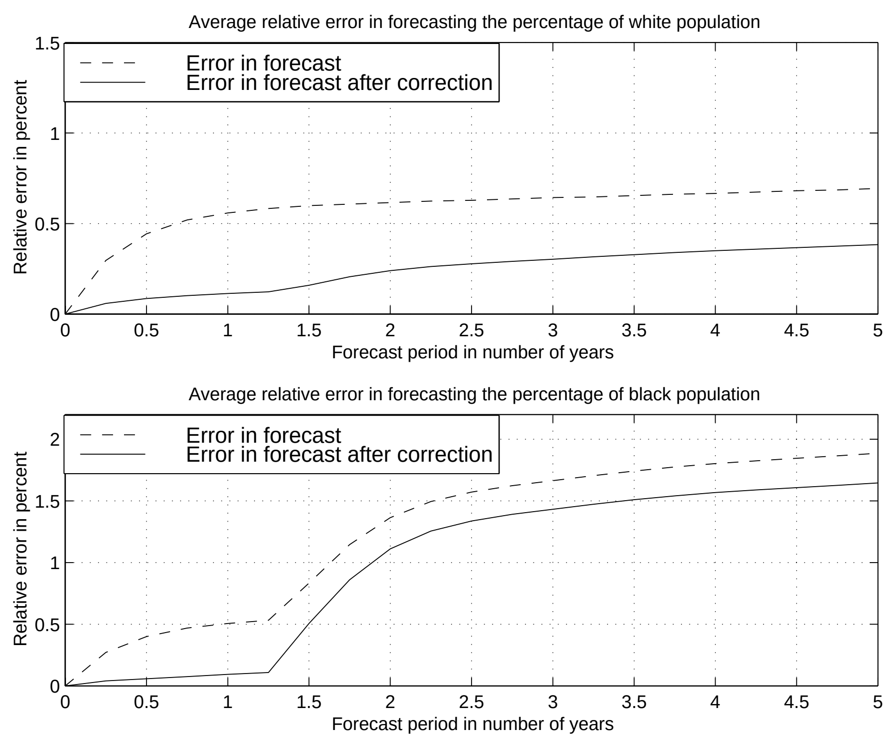

Figure 6.4: Error in Forecast of Percentage of White and Black Population Using Quarterly Data

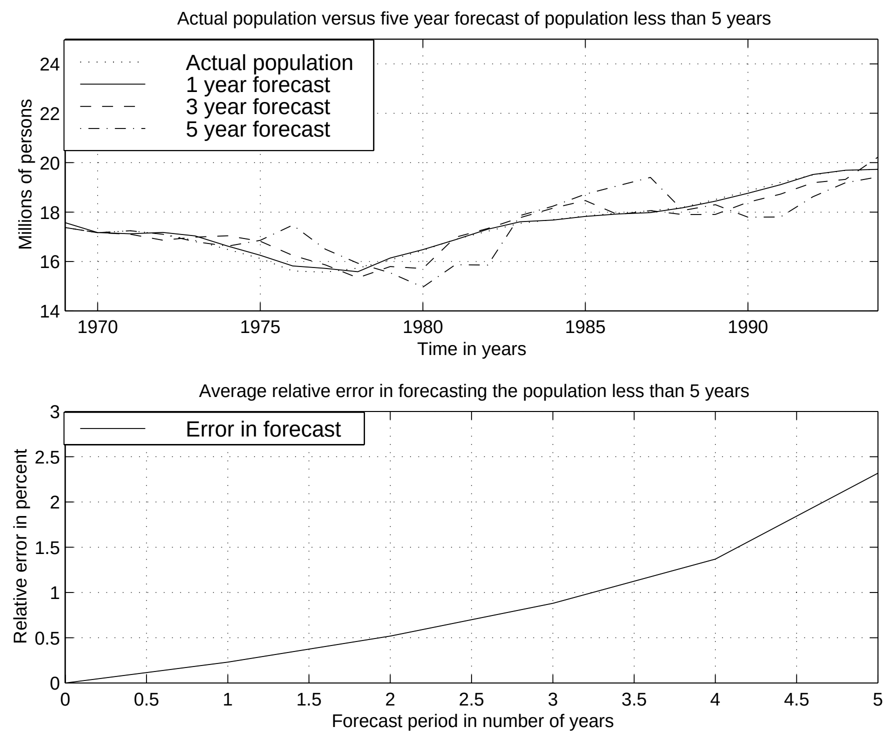

Figure 6.5: Toddler Population Forecast with Annual Data

To illustrate the better performance of the model using quarterly data over the model using annual data, the results of similar forecasts using annual data is shown in Fig. 6.5, Fig. 6.6, Fig. 6.7 and Fig. 6.8.

Since the total population was an almost linear curve, forecasts were also made on the total population without incorporating the growth variable strategy discussed in the earlier section. This was possible because forecasts were only made for a short period into the future, and FIR sometimes tolerates a (very small) amount of extrapolation. The forecast results are shown in Fig. 6.9. There is improvement in the forecasts when a growth variable is used instead of the variable itself. The

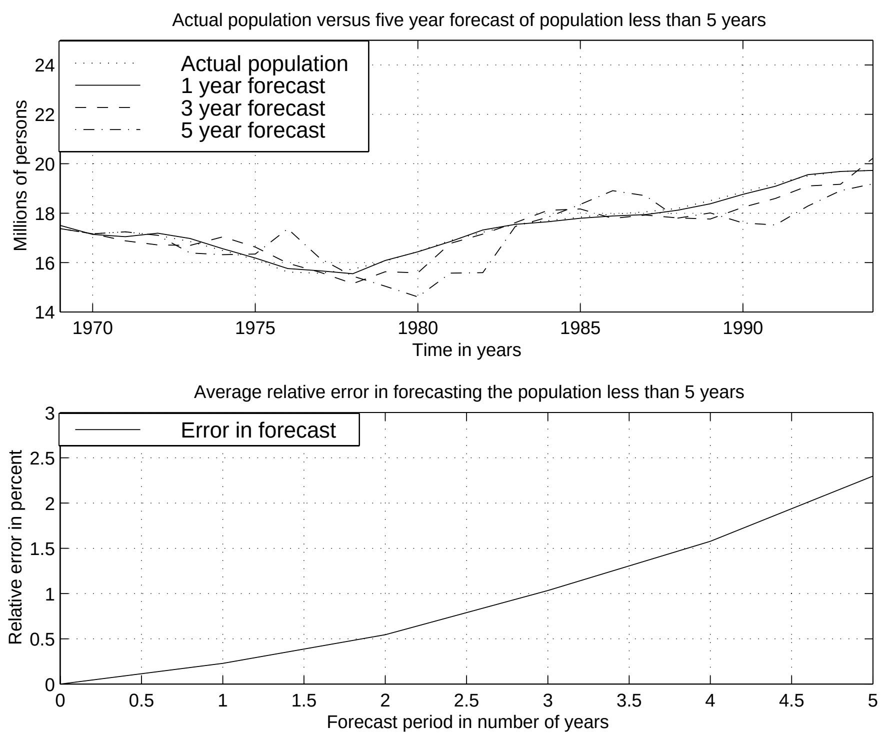

Figure 6.6: Toddler Population Forecast after correction using Annual Data

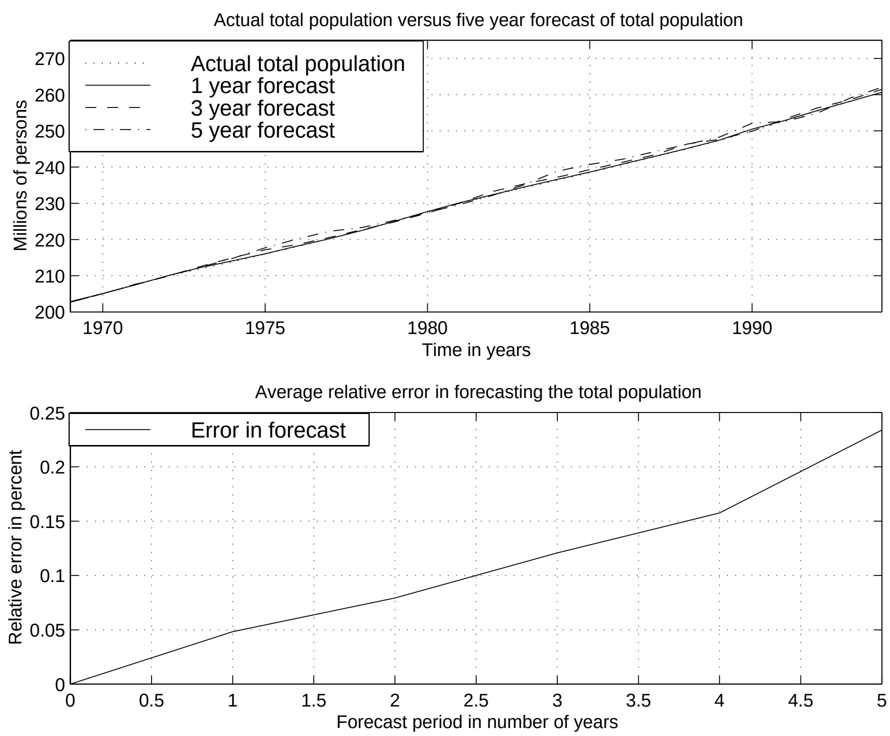

Figure 6.7: Total Population Forecast and Error Curves using Annual Data

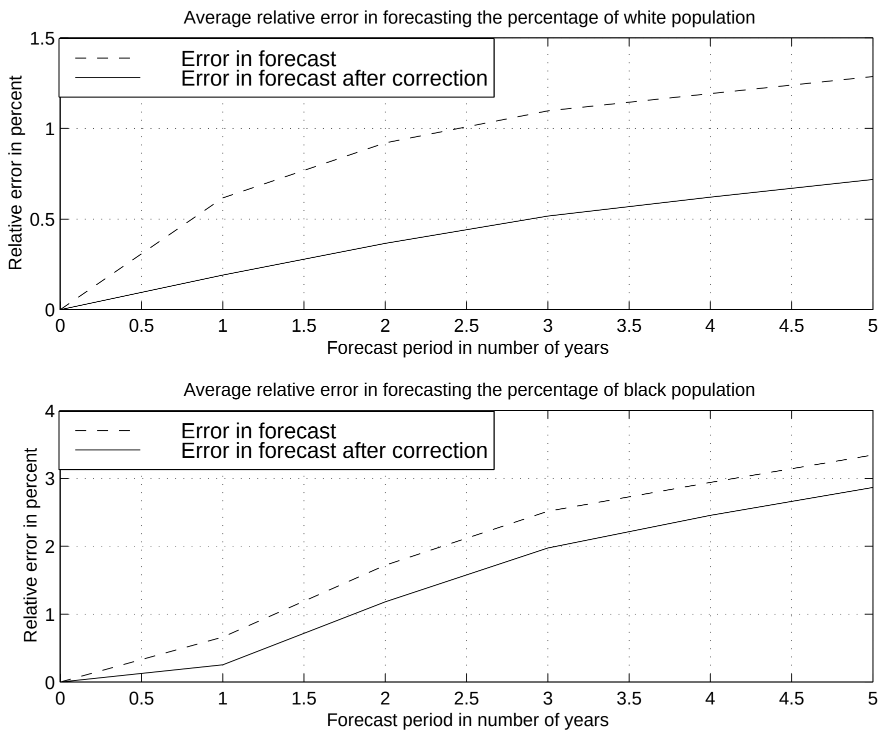

Figure 6.8: Error in Forecast of White and Black Population using Annual Data

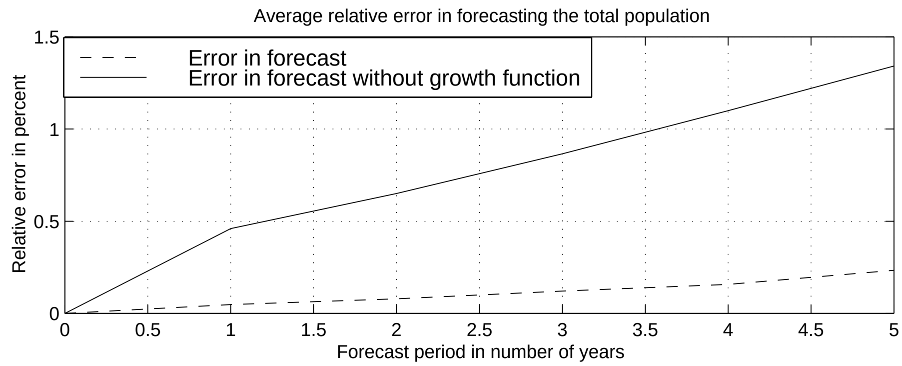

Figure 6.9: Error in Forecast without the Growth Function

total population curve is almost linear. When direct forecasts were attempted on variables that capture more non–linearity, the variables in layers two and three for instance, FIR simply refused to make any prediction without proper use of the growth functions.

The approximation of the growth variable as a first order equation helps reduce the error values. This might suggest that if the growth variable is approximated as a still–higher order equation, yet better forecasts can be obtained. This turned out not to be true. Fig. 6.10 shows that, while for a smaller horizon, the error values using first and second order approximations are comparable, the first order approximation performs better for longer horizons.

With quarterly data, two different mask candidate matrices were used to check for any improvement in the forecast, namely a simple and a complex mask candidate matrix.

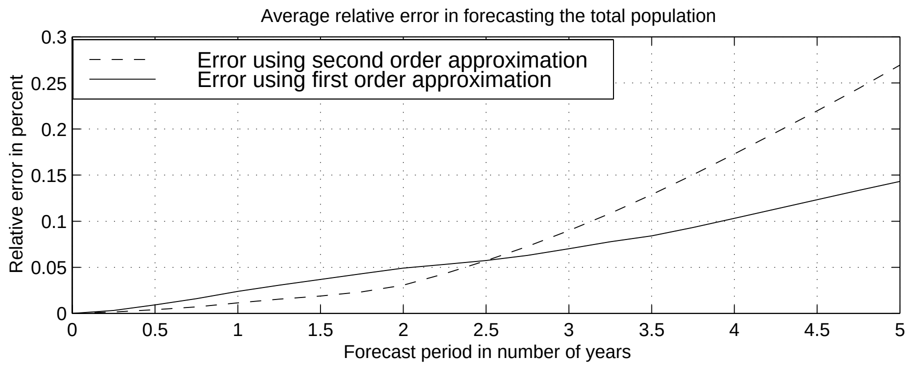

Figure 6.10: Error in Forecast by Approximating the Growth Function as Second Order Function

The simple mask candidate matrix that was used for the forecast of a three variable system using quarterly data looks like :

$$\mathbf{mcan} = \begin{pmatrix} -1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \\ +1 & 0 & 0 \end{pmatrix}$$

The rationale for the simple mask candidate matrix is that since the added data is only *dummy*, those entries can be masked out in the search for an optimal mask. This helps speed up the modeling process.

The complex mask candidate matrix that was used for the forecast of a three variable system using quarterly data looks like :

$$\mathbf{mcan} = \begin{pmatrix} -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ -1 & -1 & -1 \\ +1 & 0 & 0 \end{pmatrix} \tag{6.4}$$

The rationale for the complex mask candidate matrix is that it provides FIR with optimal flexibility in choosing the best possible *m-inputs*. However, the forecasts were similar in both cases, as this was expected since the extra data added had no knowledge in them. If the added data were original data, i.e., if we had raw quarterly data to feed into the model, then better forecast results could be expected.

6.4 Economy Layer

Fig. 6.11 and Fig. 6.12 show the forecast results for the disposable per capita income and the wage rate. These variables depend on the total population, a variable that was forecast in layer one. Other forecast models that take as input only their own past values were also tried. The model that used data from layer one outperformed the simple model.

Fig. 6.13 shows the forecast results for the rate of unemployment. The forecast results of the first layer were included to make a forecast of the rate of unemployment.

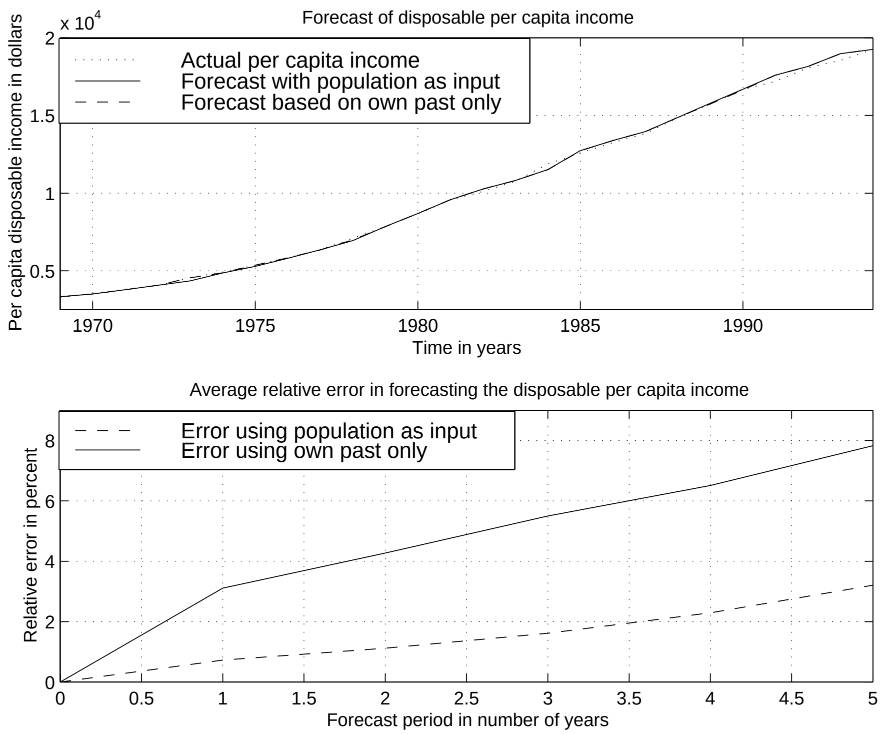

Figure 6.11: Disposable Per Capita Income Forecast and Error Curves

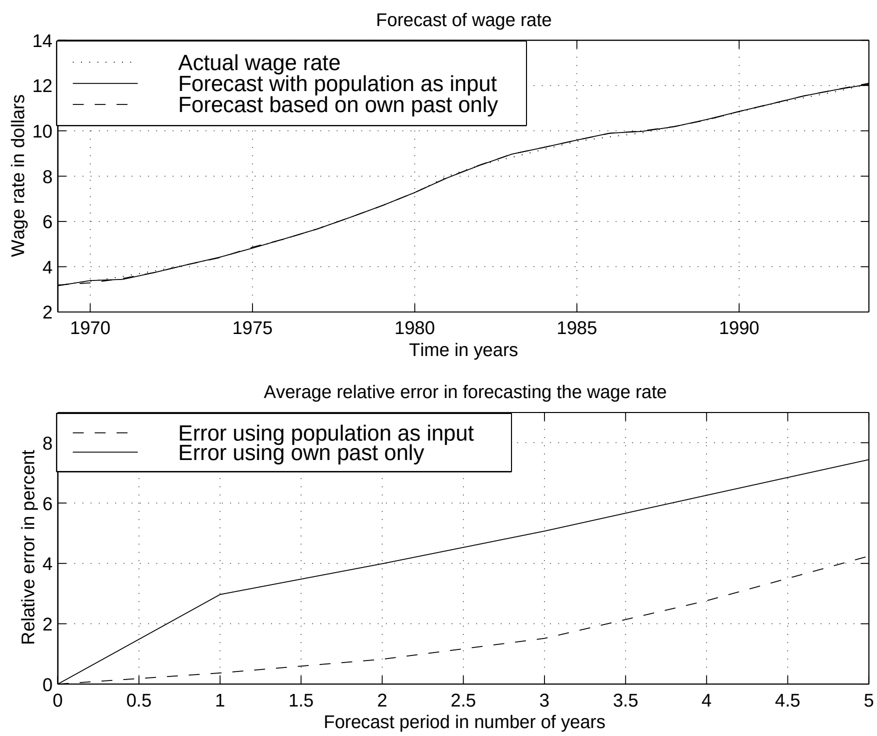

Figure 6.12: Wage Rate Forecast and Error Curves

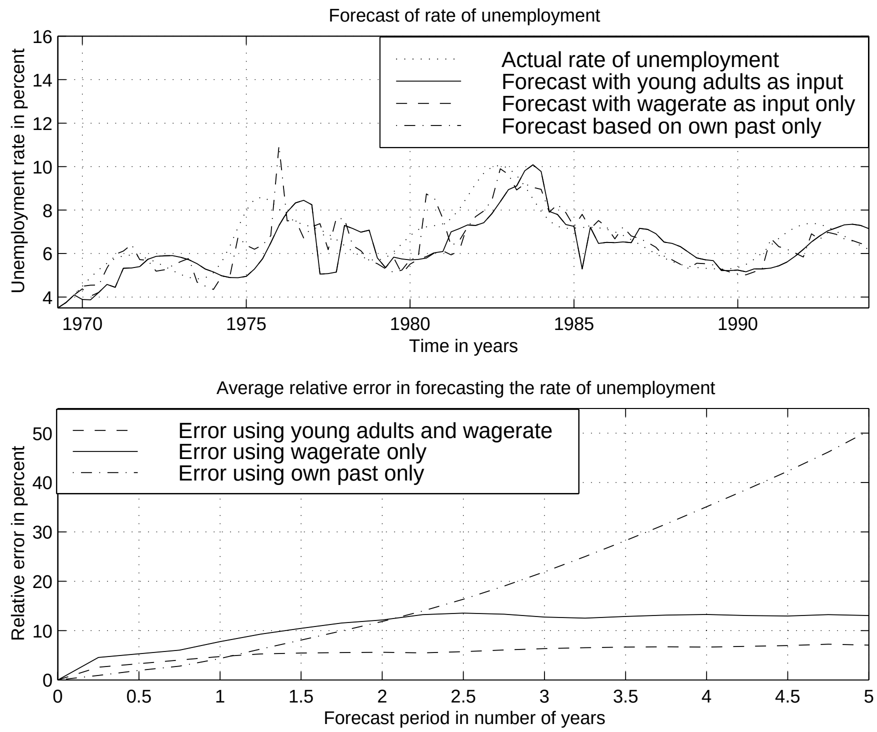

Figure 6.13: Unemployment Forecast and Error Curves

The forecast also included the young-adults who are the young working population, the total population and the wage rate. Similar forecast were also performed using only the rate of unemployment and using the rate of unemployment and the wage rate only. The forecast error in the simpler models were much higher than that of the composite model.

Fig. 6.14 shows the forecast and error curves for the Consumer Price Index for all items, namely food and non–food. Various other forecast models were also used that used partial inputs from the first and second layers. The errors in the partial models were much higher than those of the proposed model.

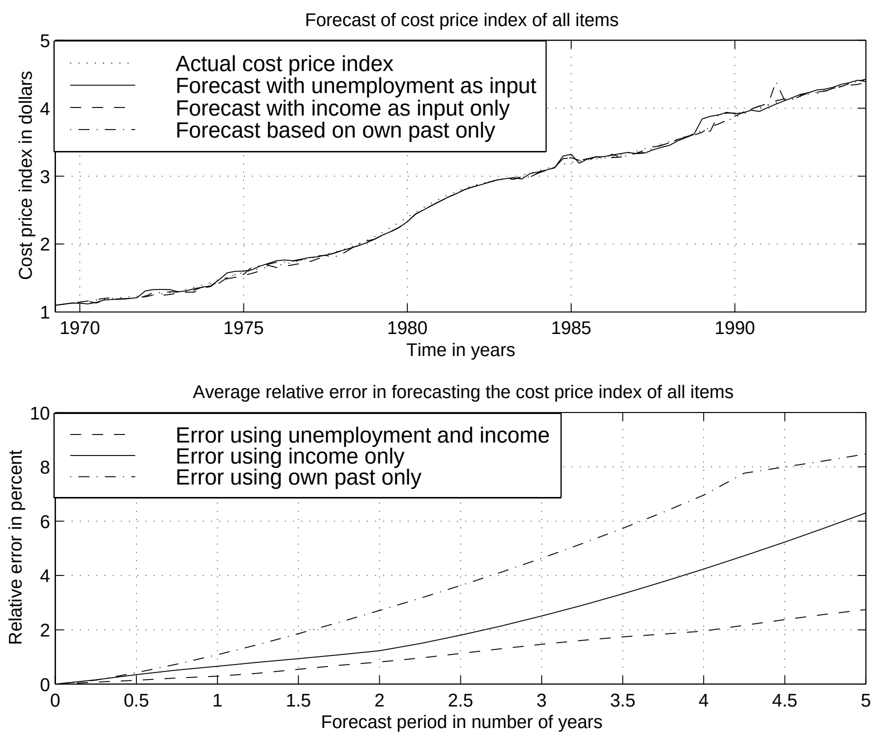

Figure 6.14: Consumer Price Index Forecast and Error Curves

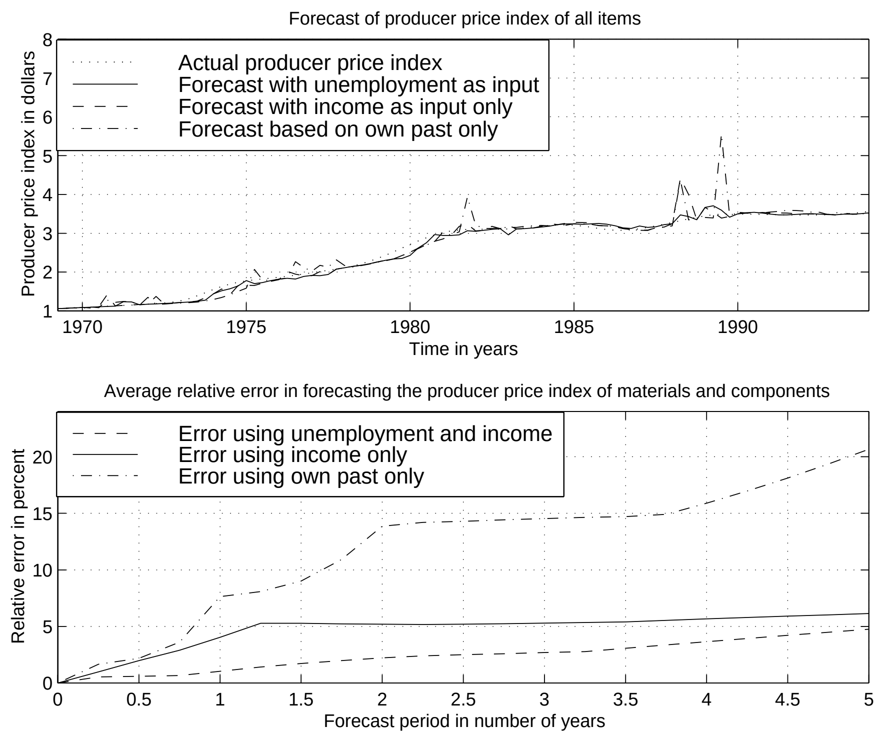

Figure 6.15: Producer Price Index Forecast and Error Curves

Fig. 6.15 shows the forecast and error curves for the Producer Price Index for materials and components. Various other forecast models were also used that used partial inputs from the first and second layers. Once again, the more complex model performed better than the simpler models.

6.5 Food Supply Layer

Fig. 6.16, Fig. 6.17, Fig. 6.18 and Fig. 6.19 show the forecast and error values for the price of poultry products, fruit produce, cheese products and fat and oil products, respectively. As in earlier cases, various other models that depend only on layer three

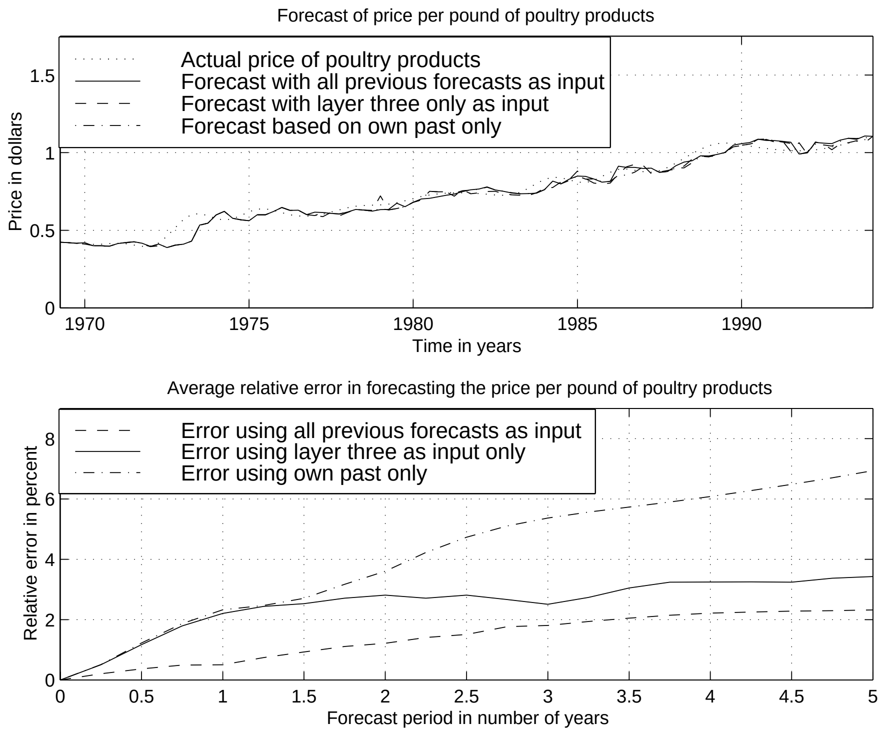

Figure 6.16: Poultry Products Price Forecast and Error Curves

alone, or layer two and three alone or their own past alone were tried. The complex model that depended on all three layers produced the best results.

6.6 Food Demand Layer

Fig. 6.20, Fig. 6.21, Fig. 6.22 and Fig. 6.23 show the forecast and error values for the consumption quantity of poultry products, fruit produce, cheese products and fat and oil products, respectively. As in earlier cases, various other models that depended only on layer three alone, or layer two and three alone or their own past

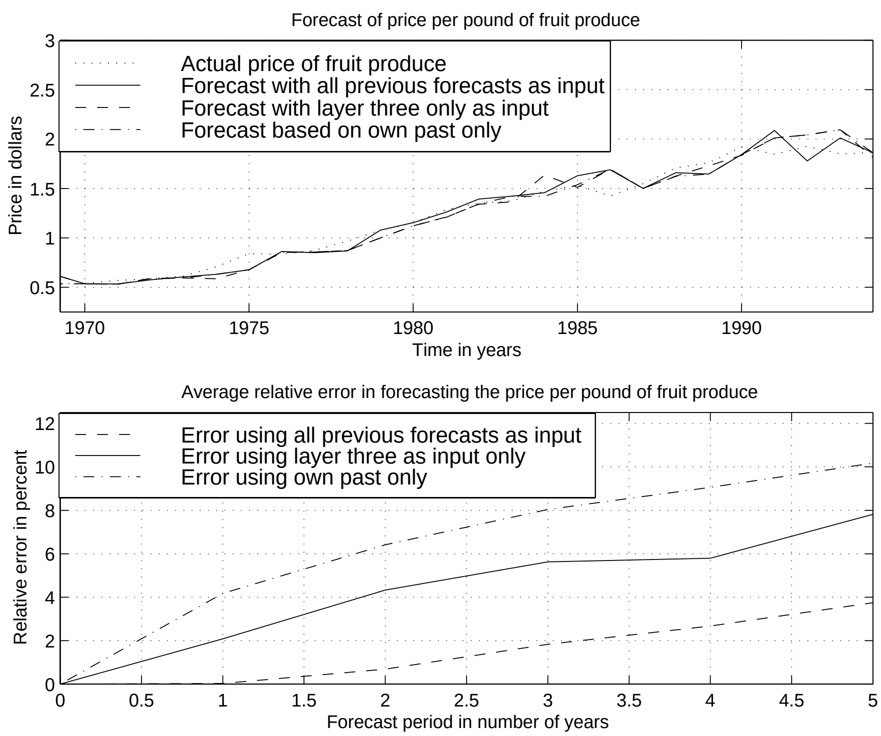

Figure 6.17: Fruit Produce Price Forecast and Error Curves

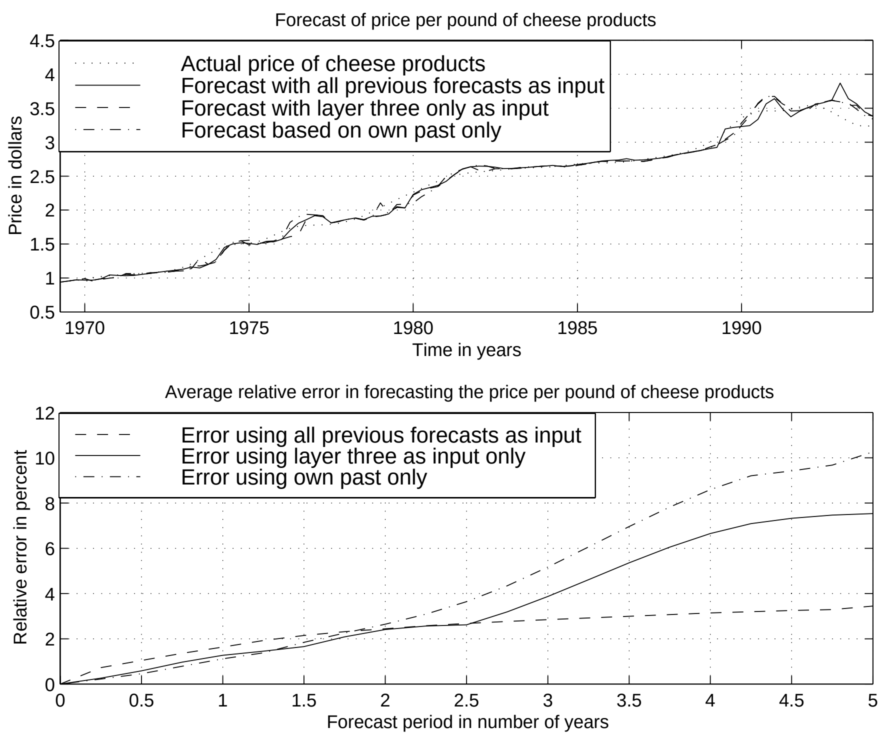

Figure 6.18: Cheese Price Forecast and Error Curves

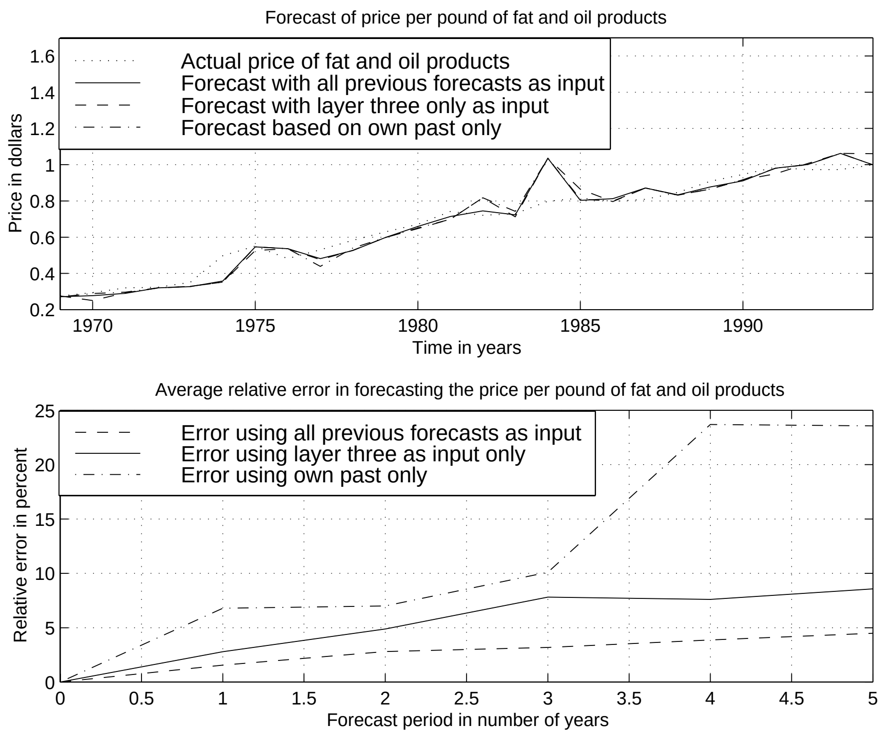

Figure 6.19: Fat and Oil Products Price Forecast and Error Curves

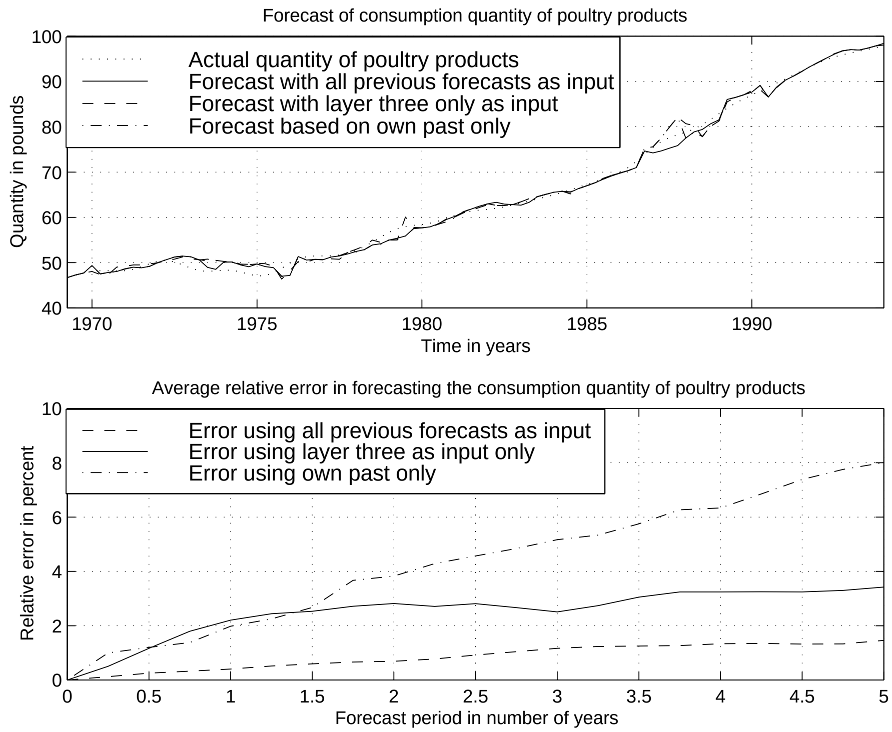

Figure 6.20: Poultry Products Quantity Forecast and Error Curves

alone were tried. The complex model that depended on all three layers produced the best results.

The use of quarterly data helped make better forecasts. Only in some cases were the forecast results similar. As explained earlier, the consumption quantity of some food products, like fresh milk and cream, are not affected very much by changes in their price. In these cases, the forecasts using annual data and quarterly data are almost the same. The results are shown in Fig. 6.24 and Fig. 6.25 respectively.

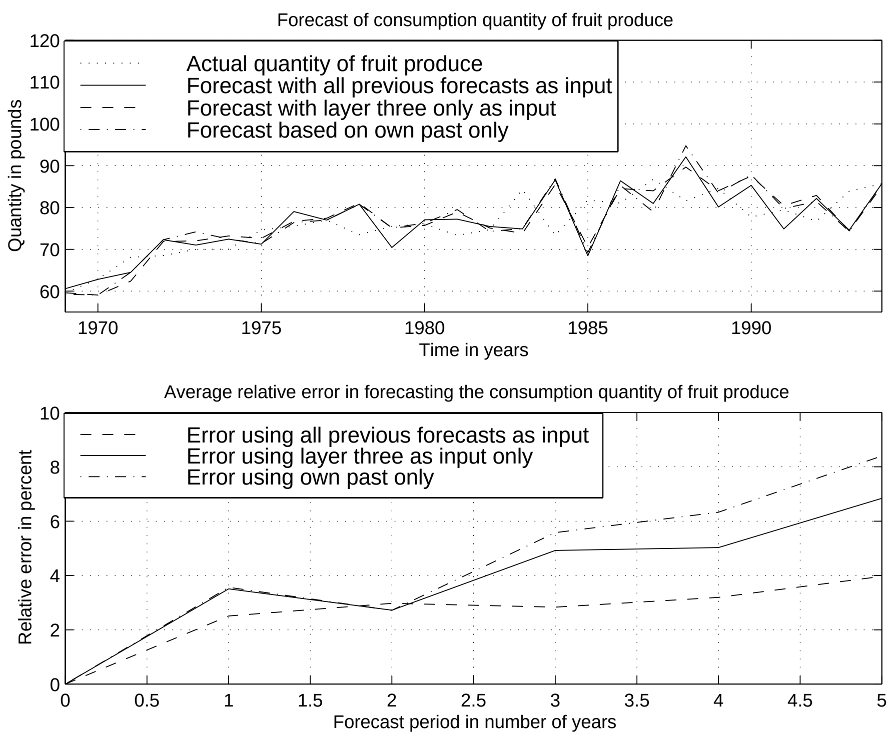

Figure 6.21: Fruit Produce Quantity Forecast and Error Curves

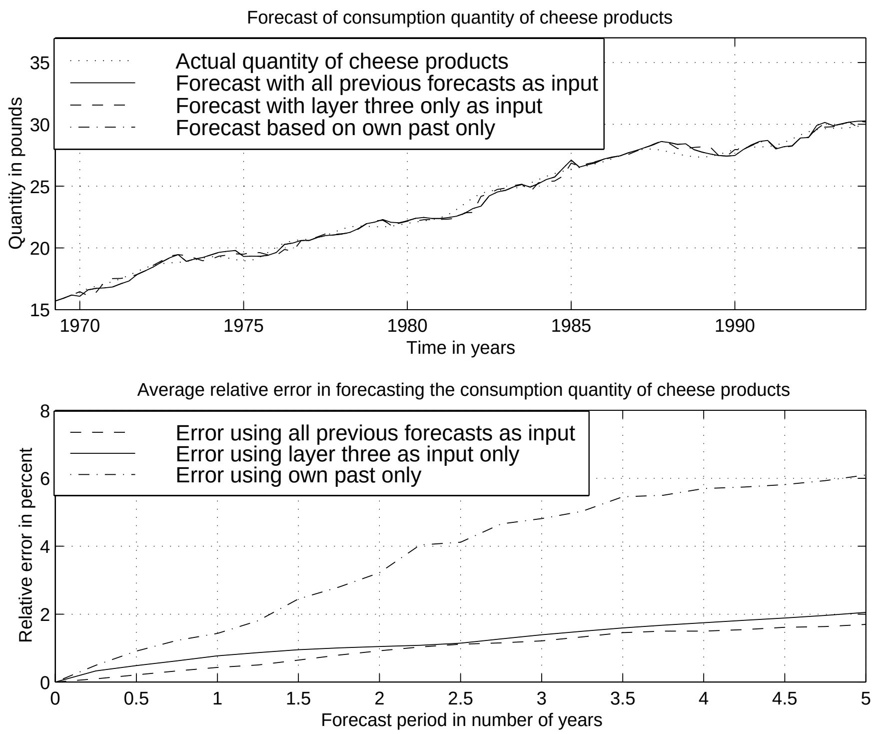

Figure 6.22: Cheese Quantity Forecast and Error Curves

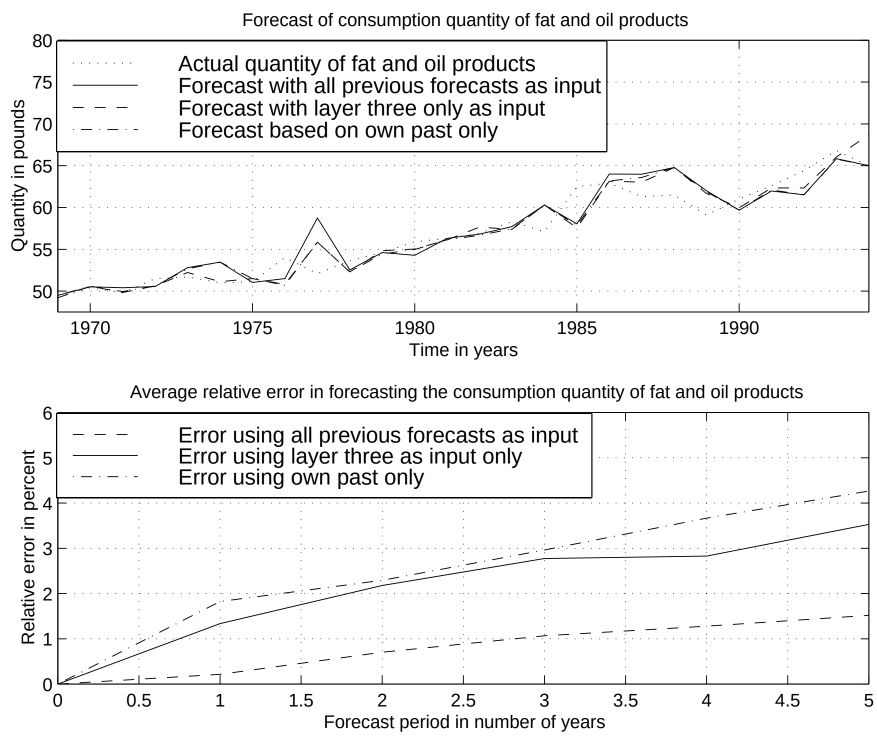

Figure 6.23: Fat and Oil Products Quantity Forecast and Error Curves

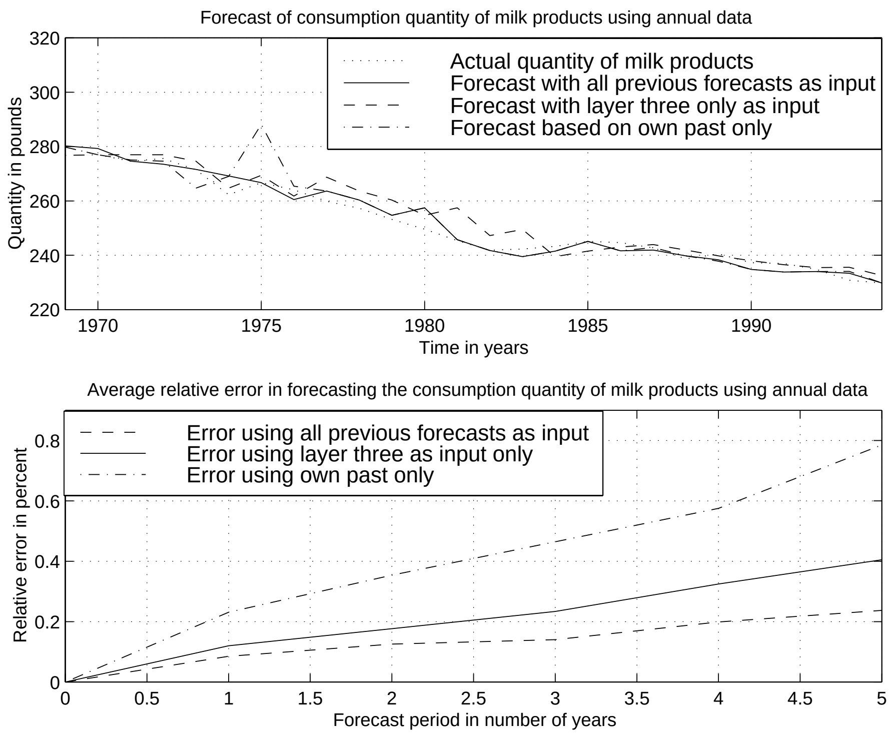

Figure 6.24: Milk Quantity Forecast and Error Curves using Annual Data

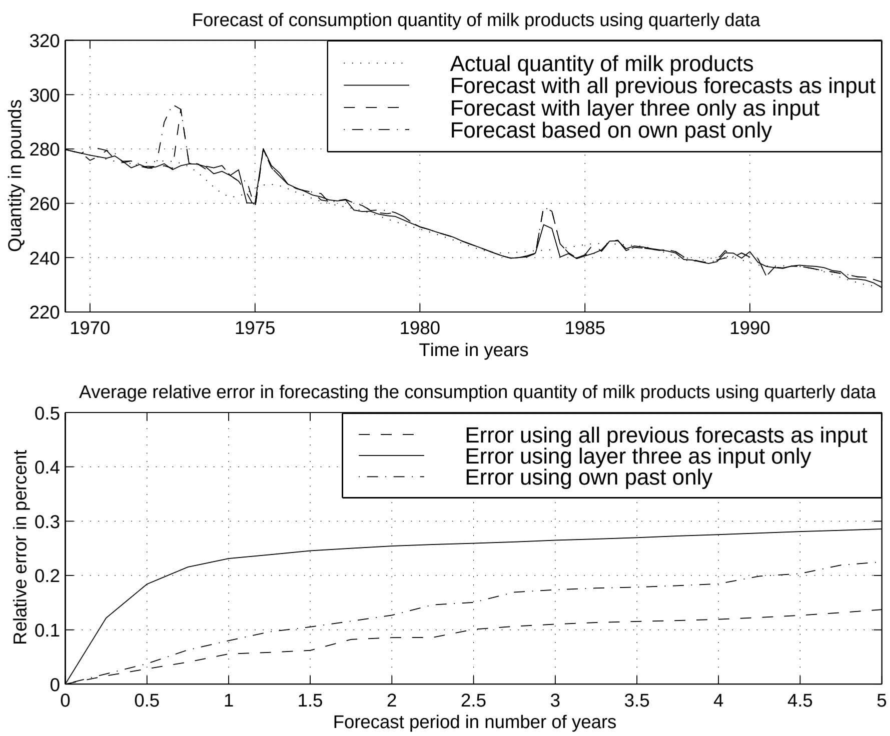

Figure 6.25: Milk Quantity Forecast and Error Curves using Quarterly Data

6.7 Optimization Scheme

In this scheme, a distinction is made between the demand and supply layers which were earlier combined together as a single layer. As explained earlier, the relationship between the prices and consumption volume are considered to be immediate. The food prices depend on the current food consumption volume and vice–versa. The optimization scheme tries to model this interaction.

The prices are first fixed at the previous year's level. The volume of food sold at those prices, given the current population dynamics and economy data, can be calculated. From the estimated consumption, the profit of the producers can be obtained. This model can then be embedded in an optimization layer, in which the food prices are treated as parameters, and the profit is the performance index to be maximized.

The algorithm to be followed is described below for making forecasts on a few diary products :

Using the food supply model, a forecast of the price of milk and fresh cream, butter products and cheese products is made. Let this be represented as P_1^0, P_2^0 and P_3^0 and let

$$\mathbf{P^0} = \begin{pmatrix} P_1^0 \\ P_2^0 \\ P_3^0 \end{pmatrix}$$

Using the food demand model, a forecast of the consumption quantity of milk and fresh cream, butter products and cheese products is made. Let this be represented as A_1^0, A_2^0 and A_3^0 and let

$$\mathbf{A^0} = \begin{pmatrix} A_1^0 \\ A_2^0 \\ A_3^0 \end{pmatrix}$$

The Profit Index can then be calculated as

$$PI^0 = \quad P_1^0 \cdot A_1^0 \quad + \quad P_2^0 \cdot A_2^0 \quad + \quad P_3^0 \cdot A_3^0 \tag{6.5}$$

Adjusting the prices from the food supply layer in the following manner:

$$\mathbf{P^1} = \begin{pmatrix} P_1^0 \ * \ 1.05 \\ P_2^0 \\ P_3^0 \end{pmatrix}$$

$$\mathbf{P^2} = \begin{pmatrix} P_1^0 \\ P_2^0 \ * \ 1.05 \\ P_3^0 \end{pmatrix}$$

$$\mathbf{P^3} = \begin{pmatrix} P_1^0 \\ P_2^0 \\ P_3^0 \quad * \quad 1.05 \end{pmatrix}$$

we can obtain from the food supply model, the corresponding

$$\mathbf{A^1} = \begin{pmatrix} A_1^1 \\ A_2^1 \\ A_3^1 \end{pmatrix}$$

$$\mathbf{A^2} = \begin{pmatrix} A_1^2 \\ A_2^2 \\ A_3^2 \end{pmatrix}$$

$$\mathbf{A^3} = \begin{pmatrix} A_1^3 \\ A_2^3 \\ A_3^3 \end{pmatrix}$$

The corresponding Price Indices are then

$$PI^1 = \quad P_1^1 \cdot A_1^1 \quad + \quad P_2^0 \cdot A_2^1 \quad + \quad P_3^0 \cdot A_3^1 \tag{6.6}$$

$$PI^2 = \quad P_1^0 \cdot A_1^2 \quad + \quad P_2^2 \cdot A_2^2 \quad + \quad P_3^0 \cdot A_3^2 \tag{6.7}$$

$$PI^3 = \quad P_1^0 \cdot A_1^3 \quad + \quad P_2^0 \cdot A_2^3 \quad + \quad P_3^3 \cdot A_3^3 \tag{6.8}$$

The sensitivity of the consumption quantity to the price is calculated as

$$\frac{\delta PI}{\delta P_1} = \frac{PI^1 - PI^0}{P_1^1 - P_1^0} \tag{6.9}$$

$$\frac{\delta PI}{\delta P_2} = \frac{PI^2 - PI^0}{P_2^2 - P_2^0} \tag{6.10}$$

$$\frac{\delta PI}{\delta P_3} = \frac{PI^3 - PI^0}{P_3^3 - P_3^0} \tag{6.11}$$

and is written as

$$\delta \mathbf{PI} = \begin{pmatrix} \frac{\delta PI}{\delta P_1} \\ \\ \frac{\delta PI}{\delta P_2} \\ \\ \frac{\delta PI}{\delta P_3} \end{pmatrix}$$

A fresh value for the price is then calculated as

$$\mathbf{P^4} = \begin{pmatrix} P_1^4 \\ P_2^4 \\ P_3^4 \end{pmatrix}$$

where

$$\mathbf{P^4} = \mathbf{P^0} + \gamma \cdot \delta\mathbf{PI} \tag{6.12}$$

The value of γ is set at a small value, say 0.01. This value of $\mathbf{P^4}$ can be used as the price from the food supply model and all the above steps can be repeated again, until the norm of $\delta\mathbf{PI}$ drops to a value almost close to zero. The value of $\mathbf{P^4}$ is then the best price that will maximize the profits for the suppliers.

It was not possible to perform the above mentioned optimization scheme using the data that was available because of the small size of the training data. When the price was moved beyond a certain margin, FIR started to complain that it had not seen the input pattern earlier and so refused to make a forecast. This can be overcome, if we had true data that were available in more frequent intervals. For instance, the price of tomatoes changes on a daily basis. The dynamics in such a model could be sufficient to run the optimization scheme.

CHAPTER 7

Conclusion and Future Research

In this thesis, a new mixed quantitative and qualitative approach to modeling macroeconomic systems for the purpose of short–term prediction (one to five years) was presented. A three–layer architecture was proposed. It was shown that fairly accurate predictions of macroeconomic variables can be obtained using this layered architecture.

What are the main advantages of a mixed SD/FIR approach? Pure SD is attractive because it requires very few training data, but the methodology is treacherous, because it doesn't offer any self–assessment capabilities. Pure FIR is attractive because of its sheer generality and ease of use, but it has the problem of requiring lots of data, and without any underlying structure, the modeling effort needs to be started from scratch for each new application. The mixed SD/FIR approach combines the best of both worlds. It saves of SD what is worth saving, but adds FIRs self–assessment capabilities and reduces drastically the assumptions made on the model.

It is amazing how well the methodology worked, given the fact that the model had to operate in a mode of *severe data deprivation.* Annual data from 1910 to 1970 were used as training data, i.e., only 61 data records were available for training the model.

The data deprivation problem was overcome by interpolating between the measurement data points using spline interpolation (available in Matlab). Clearly, adding more data points by means of interpolation did not add any additional information to the data set. The new data are *derived data*, and one shouldn't expect that the model would improve as a consequence of such data. Yet, they actually helped for two reasons.

First, FIR uses the five nearest neighbors for predicting fuzzy membership values. Because there are only 61 data records in the training data base, the neighbors will be far away, and therefore, a lot of interpolation has to be done. By generating e.g. three new artificial data points for each measured one, the system would now have 241 data points to work with. Therefore, the five nearest neighbors will be much closer to the current data point, and consequently, much less interpolation will be needed.

Second, FIR tries to have at least five recordings for each discrete state. With 61 data records, only 12 different states can be recorded 5 times each. This means that FIR will always pick extremely simple models with one to three ternary input variables only. If FIR is offered four or five possible input variables, it will pick the most relevant ones, and discard the others, although they might carry useful information. If the data deprivation problem is reduced, FIR might pick a mask of higher complexity, and thereby also exploit the information contained in less important variables.

The improvement in forecasts by use of interpolated data also suggests that if it were actually possible to obtain true data collected at more frequent intervals, then better forecasts can result.

How does this study help with predicting other variables, such as the demand for used cars, or the prices of telephone calls? The two bottom layers of the architecture are independent of the application at hand. Only the top layer (demand and supply) needs to be reidentified for each new application. This certainly helps.

Many technological variables have a much shorter history. For example, cellular phones or the world wide web simply haven't been around very long. Yet, their time constants are considerably shorter also, and therefore, data can be recorded more frequently. If monthly data are meaningful (because the time constants are months rather than years), only about six years worth of data would be needed to get 60 data points. In this case, a spline interpolation on the lower levels serves an additional purpose. It can be used to provide the intermediate data points needed to feed the faster changing technological variables of the top layer.

Finally, although the bottom layers are certainly specific to the U.S., the three–layer architecture itself is not. For each new country or region, the two bottom layers will have to be reidentified, yet, the structure of the architecture will remain the same.

Future work should include obtaining data at more frequent intervals of time such that the Optimization between the Food Demand and Food Supply layers can be performed. Data sets pertaining to a different U.S. industry could be used to make forecasts in that industry. If it were possible, it would also be helpful to obtain a break-up of the population based on income levels. This would help bind the population and macroeconomy layers much closer together.

The modeling approach chosen in this research represents a drastic deviation from the approaches pursued by mainstream economic research. In order to better assess the usefulness of the advocated methodology, a thorough and fair comparison between the two approaches should be made.

This thesis work is a first step towards a mixed deductive and behavior modeling approach and hopes to pave the way for newer and bolder economic modeling techniques.

REFERENCES

[1] P.W. Anderson, K.J. Arrow and D. Pines, *The Economy as an Evolving Complex System*, Eds. Redwood City, CA, Addision-Wesley, 1989.

[2] R. A. Pollak, and T. J. Wales, *Demand system Specification and Estimation*, Oxford University Press, 1992.

[3] B. Hannon, M. Ruth, *Dynamic Modelling*, Springer–Verlag, New York, 1994.

[4] G. J. Klir, *Architecture of Systems Problem Solving*, Plenum Press, New York, 1985.

[5] K. N. Pantazopoulos, L. H. Tsoukalas *et al.*, "Financial Prediction and Trading Strategies Using Neurofuzzy Approaches" *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 28, No. 4, pp.520-531., August 1998.

[6] A. Nebot, *Qualitative Modeling and Simulation of Biomedical Systems Using Fuzzy Inductive Reasoning*, Ph.D. Dissertation, Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 1994.

[7] F. Mugica, *Diseño Sistemático de Controladores Difusos Usando Razonamiento Inductivo*, Ph.D. Dissertation, Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 1995.

[8] A. de Albornoz, *Inductive Reasoning and Reconstruction Analysis: Two Complementary Tools for Qualitative Fault Monitoring of Large–Scale Systems*, Ph.D. Dissertation, Llenguatges i Sistemes Informàtics, Universitat Politècnica de Catalunya, Barcelona, Spain, 1996.

[9] F. E. Cellier, J. López, A. Nebot, and G. Cembrano, "Confidence Measures for Predictions in Fuzzy Inductive Reasoning," *Intl. J. General Systems*, manuscript in review for a special issue on Fuzzy Inductive Reasoning.

[10] J. López, F. Tjäarnström, F. E. Cellier, G. Cembrano and L. Ljung, "Estimating the Horizon of Predictibility in Time Series Predictions Using Inductive Modeling Tools," *Intl. J. General Systems*, manuscript in review for a special issue on Fuzzy Inductive Reasoning.

[11] F. E. Cellier and J. López, "Causal Inductive Reasoning: A New Paradigm for Data-Driven Qualitative Simulation of Continuous-Time Dynamical Systems," *Systems Analysis Modelling Simulation,* 18(1), pp.27-43, 1995.

[12] T. Johansen and B. Foss. "Constructing NARMAX models using ARMAX models," *International Journal of Control*, vol. 58, pp.1125-1153, 1993.

[13] C. C. Holt, "Forecasting Sales and Trends by Exponentially Weighted Moving Averages," *US Office of Naval Research Memorandum*, no.57, 1957.

[14] P. R. Winters, *Forecasting Sales by Exponentially Weighted Moving Averages*, Management Science, vol. 6, no. 3, pp.324-342, 1960.

[15] R. G. Brown, *Smoothing, Forecasting and Prediction of Discrete Time Series*, Prentice-Hall, 1963.

[16] P. J. Harrison, "Short Term Sales Forecasting", *Journal of the Royal Statistical Society* series C, vol. 14, part 2, pp.102-139, 1960.

[17] G. E. Box and G. M. Jenkins, *Time-Series Analysis, Forecasting and Control*, Holden-Day, 1970.

[18] A. Lapedes and R. Farber, "Nonlinear Signal Processing Using Neural Networks:Prediction and System Modelling", *Los Alamos National Lab Technical Report* LA-UR-87 - 2662, 1987.

[19] R. Sharda and R. B. Patil, "Combining Artificial Neural Networks and Statistics for Stock Market Forecasting" *Proceedings of the IJCNN Meeting*, Washington, pp.491-494., 1990.

[20] K. N. Pantazopoulos, L. H. Tsoukalas, and E. N. Houstis, "Neurofuzzy characterisation of financial time series in an anticipatory framework," *Proc. of IEEE/IAFE 1997 Conf. Computational Intelligence Financial Engineering*, pp.50-56., New York, March 1997.

[21] J. Ahmad and H. Fatmi, "A quadratic neural network system for prediction of time-series data," *Proc. of IEEE World Congr. on Neural Networks*, Orlando, FL, pp.3667-3670, 1994.

[22] S. I. Wu and R. P. Lu, "Combining Artificial Neural Networks and Statistics for Stock Market Forecasting" *Proceedings of the 21st Computer Science Conference*,

Indianapolis, Indiana, Feb 16-18, Association for Computer Machinery, New York, pp.257-264., 1993.

[23] Z. Tang, C. de Almeida, and P.A. Fishwick, "Time Series Forecasting Using Neural Networks vs. Box-Jenkins Methodology," *Simulation*, vol.57, no. 5, Nov. 1991, pp. 303-310.

[24] P. D. Wasserman, *Neural Computing : Theory and Practice*, Van Nostrand Reinhold, NewYork, 1989.

[25] J. W. Forrester, *Industrial Dynamics*, MIT Press, Cambridge, Mass, 1964.

[26] J. W. Forrester, *Principles of Systems*, Wright-Allen Press, Cambridge, Mass, 1968.

[27] B. Richmond, S. Peterson, and P. Vescuso, *An Academic User's Guide to STELLA* High Performance Systems, Inc., Lyme, N.H., 1987

[28] J.W. Forrester, *World Dynamics*, Wright-Allen Press, Cambridge, Mass, 1971.

[29] J. D. Lebel, "System Dynamics," *Progress in Modelling and Simulation* (F.E. Cellier, Ed.), Academic Press, London, U.K., pp.119-158, 1982.

[30] H. J. Uyttenhove, *SAPS - System Approach Problem Solver,* Ph.D dissertation, SUNY Binghampton, New York, 1979.

[31] F. E. Cellier and D. W. Yandell, "SAPS-II: A New Implementation of the Systems Approach Problem Solver," *Intl. J. General Systems*, **13**(4), pp.307-322, 1987.

[32] G. J. Klir, "Inductive Systems Modelling : An Overview." In M. S. Elzas, T. I. Oren, and B. P. Zeigler (Eds.), *Modelling and Simulation Methodology: Knowledge Systems' Paradigms* Elsevier Science Publishers B.V. (North-Holland), Amsterdam, The Netherlands, 1989.

[33] D. Li and F. E. Cellier, "Fuzzy Measures in Inductive Reasoning," *Proc. 1990 Winter Simulation Conference*, New Orleans, La., pp.527-538, 1990.

[34] F. E. Cellier, *Continuous System Modeling,* Springer–Verlag, New York, 1991.

[35] F. E. Cellier, A. Nebot, F. Mugica, and A. de Albornoz, "Combined Qualitative/Quantitative Simulation Models of Continuous-Time Processes Using Fuzzy Inductive Reasoning Techniques," *Intl. J. General Systems*, **24**(1-2), pp.95-116, 1996.

[36] A. Nebot, F. E. Cellier, and M. Vallverdú, "Mixed Quantitative/Qualitative Modeling and Simulation of the Cardiovascular System," *Computer Methods and Programs in Biomedicine*, 55(2), pp. 127-155.

[37] A M. Law and W. D. Kelton, *Simulation Modeling and Analysis,* second edition, McGraw-Hill, New York, 1990.

[38] R. C. Conant, "Extended Dependency Analysis of Large Systems," *Intl. J. General Systems*, vol.14, pp.97-141, 1988.

[39] M. Sugeno and G. T. Kang, "Structure Identification of Fuzzy Model," *Fuzzy Sets and Systems*, vol.28, no. 1, pp.15-33, 1988.

[40] F. E. Cellier, "Qualitative Simulation of Technical Systems Using the General System Problem Solving Framework," *Intl. J. General Systems*, **13**(4), pp.333-344, 1987.

[41] P. J. Vesanterä and F.E. Cellier, "Building Intelligence into an Autopilot Using Qualitative Simulation to Support Global Decision Making," *Simulation*, **52**(3), pp.111-121, 1989

[42] M. Moorthy, F. E. Cellier and J. T. LaFrance, "Prediction of U.S. Food Demand in the 21st Century - A New look at System Dynamics," in *Proceedings of the SPIE Conference*, Orlando, FL, 1998, vol. 3369, pp 343-354.

[43] S. R. Johnson, Z. A. Hassan, and R. D. Green, *Demand Systems Estimation*, Iowa State University Press, 1984.

[44] J. López, *A Contribution to the Prediction of Time Series Using Fuzzy Inductive Reasoning*, Ph.D. Dissertation, Institut de Cibernètica, Universitat Politècnica de Catalunya, Barcelona, Spain, 1999.

[45] J. T. LaFrance, "U.S. Food and Nutrient Demand and the Effects of Agricultural Policies," *Econometrica*, manuscript in review.

[46] J.D. Farber and J.J. Sidorowich, “Can new approaches to nonlinear modeling improve economic forecasts ?,” in *The Economy As An Evolving Complex System*, P. W. Anderson, K. J. Arrow, and D. Pines, Eds. Redwood City, CA: Addison-Wesley. 1988, pp. 99-115