Paper:

Jacucci, G., M. Foy, and C. Uhrik (1995). A SYBIL DSS: A Transportable Computerized Decision Support System for Plant Protection. In: *Proc. of the 1995 Annual Conference of the Associozione Italiana per L'Informatica ed il Calcolo Automatico (AICA'95)*, Cagliari, Italy, 27-29 September 1995. (Under Consideration for Publication)

Presented at:

The 1995 Annual Conference of the Associozione Italiana per L'Informatica ed il Calcolo Automatico (AICA'95), Cagliari, Italy, 27-29 September 1995.

A SYBIL DSS: A TRANSPORTABLE COMPUTERIZED DECISION SUPPORT SYSTEM FOR PLANT PROTECTION

Tipo: lavoro completo

Gianni Jacucci, Mark Foy and Carl Uhrik

Laboratorio di Ingegneria Informatica
Dipartimento di Informatica e Studi Aziendali
Università degli Studi di Trento
Via F. Zeni 8
I-38086 Rovereto - TN Italy

Tel: +39-464-443140
Fax: +39-464-443141

email: gianni@lii.unitn.it, mfoy@lii.unitn.it, carl@lii.unitn.it

Classificatione: Informatica in Agricoltura, DSS, Systemi Inteligenti, Informatica Grafica

ABSTRACT

EC Project SYBIL implements computerized decision support systems (DSSes) to help farmers intelligently manage crops (reducing environmental impact and increasing economic returns). Agro-meteorological computer models are used to assess the risk of the crop to pest and fungus damage. In particular, the DSS described here focuses on grape and apple models that help farmers predict when diseases and funguses will attack their plants, so they can make intelligent decisions on preventing these attacks.

To achieve computerized decision support in an intelligent manner, selected computer tools have been used and built. In particular, transportability has been one of the main goals, so this has greatly effected the types of tools which were included and developed. The components which have particular importance with respect to achieving the goal of transportability are: (1) the meteorological data management system which allows seamless access to many different types of data, and (2) the expert system component (which utilizes a genetic algorithm (GA), an *artificial intelligence* (AI) search technique) designed to adapt models to the region in which they are being used.

A SYBIL DSS: A TRANSPORTABLE COMPUTERIZED DECISION SUPPORT SYSTEM FOR PLANT PROTECTION

Gianni Jacucci, Mark Foy and Carl Uhrik

Laboratorio di Ingegneria Informatica
Dipartimento di Informatica e Studi Aziendali
Università degli Studi di Trento
Via F. Zeni 8
I-38086 Rovereto - TN Italy

email: gianni@lii.unitn.it, mfoy@lii.unitn.it, carl@lii.unitn.it

ABSTRACT

EC Project SYBIL implements computerized decision support systems (DSSes) to help farmers intelligently manage crops (reducing environmental impact and increasing economic returns). Agro-meteorological computer models are used to assess the risk of the crop to pest and fungus damage. In particular, the DSS described here focuses on grape and apple models that help farmers predict when diseases and funguses will attack their plants, so they can make intelligent decisions on preventing these attacks.

To achieve computerized decision support in an intelligent manner, selected computer tools have been used and built. In particular, transportability has been one of the main goals, so this has greatly effected the types of tools which were included and developed. The components which have particular importance with respect to achieving the goal of transportability are: (1) the meteorological data management system which allows seamless access to many different types of data, and (2) the expert system component (which utilizes a genetic algorithm (GA), an *artificial intelligence* (AI) search technique) designed to adapt models to the region in which they are being used.

1. INTRODUCTION

EC Project SYBIL (consisting of five partners from four countries) involves the implementation of computerized decision support systems (DSSes) to assist farmers in intelligently governing their crops such that environmental impact is reduced and economic returns are increased. Existing agro-meteorological computer models from multiple sources are integrated into the portable, user-friendly DSSes designed to assess the risk of a crop to pest and fungus damage. By evaluating this risk, the farmer has the option to apply pesticides and fungicides only when needed and avoid using these, often environmentally damaging, chemicals blindly on a regular basis or when the risk of pest and fungus damage is small. This evaluation has the potential to save the farmer both time and money because expensive chemicals will not be applied when they do not benefit the crop.

The project particularly emphasizes transportability of the tools/DSSes developed. That is, one of the main objectives is to to establish a framework/methodology for ensuring modularity, portability, and extensibility of DSSes between regions through the introduction of uniform software engineering, including uniform choices for data structures, computer platform, and user interface.

The discussion here will focus on one of the DSSes developed within the SYBIL project.

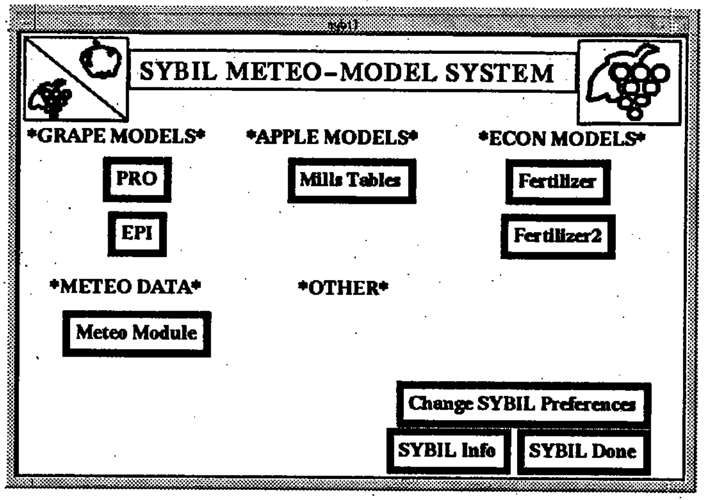

Figure 1. The System's Main Screen

This DSS is targeted mainly towards grape and apple agriculturalists. A brief description of this DSS will be given, which will be followed by a detailed discussion of two important DSS-computer aspects: the meteorological data management system and the expert system component.

2. DSS OVERVIEW

The SYBIL DSS we will describe here is targeted to grape and apple growers. Figure 1 displays the first screen of this DSS, and Figure 2 displays the structure of this DSS.

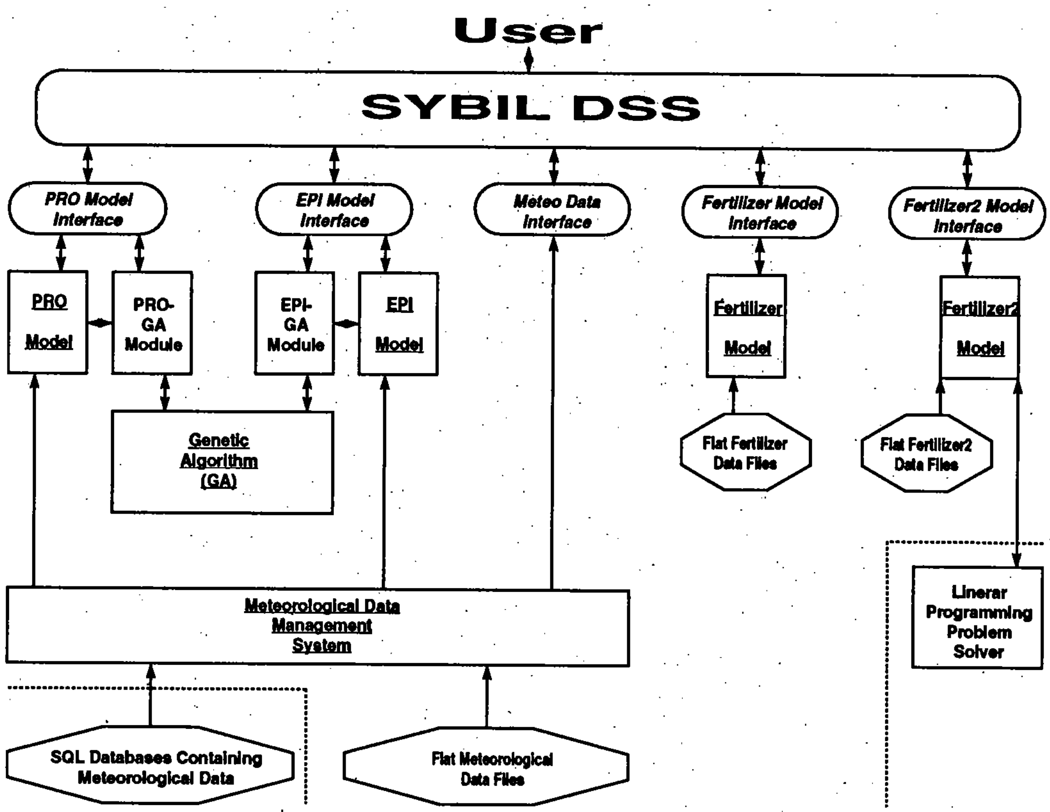

Figure 2. Structure of this DSS

This DSS is built on a common DSS platform so that the models and the DSS are modular, extendable, and can be run on 4 different computer platforms (IBM PCs under DOS, IBM PCs under MS-Windows, Macintosh machines, and many different UNIX machines with X-Windows) all with the same "look-and-feel" graphical user interface (GUI).

The current system includes six basic functionalities, which are listed below.

Meteo C Library
The Meteo C Library is a system which manages meteorological data. This library/system allows seamless access of meteorological data, whether stored in ASCII files, or in SQL databases. Overall, this component has the capability to handle many different types of meteorological data, but allows the models to access all types meteorological data in the same way so that models do not have to be re-written to handle different types of data (for example, hourly versus minutely or SQL database versus ASCII file). This library will be discussed in more detail in a following section.

P.R.O. Model
The P.R.O. (**P**lasmopara **R**isikoprognose **O**ppenheim or **P**lasmopara **R**isk **O**ppenheim) model for grapes is an analytical-biological life cycle model that simulates the infection and growth of downy mildew (Plasmopara viticola, also called peronospora) on grape vines based on meteorological conditions. The model was developed in Rheinhessen, Germany by Dr. Georg K. Hill (Hill 89) (Hill 90a) (Hill 90b) (Hill 90c) (Hill 93a) (Hill 93b). It was designed to help growers determine when it is necessary to spray grape vines against peronospora.

The model had been used by multiple Rheinhessen region farmers with good results; that is, the information provided to the farmer has assisted in making providing intelligent decisions about when to first spray against peronospora. The goal is to overcome the habit (which is not based on temporal information) of performing the first spray early in the season (possibly around May), which is often before it is necessary. This goal is approached by using the P.R.O. model to produce interpreted operational temporal information (i.e., useful up-to-date information about the status of the peronospora growth), then examining this information, and deciding if it necessary to spray at the current moment, or if spraying can be delayed (possibly many weeks) because the grape vines are not currently at risk to being damaged by peronospora. In the common cases where spraying can in fact be delayed beyond when an agriculturalist would normally spray, the overall number of interventions and amounts of chemicals sprayed on the crop are reduced. Agriculturalist using this model in the region around Rheinhessen have been able to save between one and four spraying applications per year, with an average saving of two (Hill 93b).

E.P.I. Model
Third, an implementation of the E.P.I. model from France is included. This E.P.I.-SYBIL model also addresses peronospora on grapes and assists a farmer in determining when infection of peronospora has occurred. The information given in this model is a bit more general, but with interpretation and/or localization (see below), it can assist the farmer in determining when it is necessary to spray against peronospora.

Mills Tables (Model)
Fourth, an implementation of the Mills Tables, including the extension proposed by

MacHardy, is included. This Mills Tables-SYBIL model addresses apple scab and predicts when infection of apple scab has occurred. This information allows the farmer to apply the first fungicide spraying only when needed.

Expert System Component
Fifth, an expert system component (utilizing an artificially intelligent search technique called a genetic algorithm) has been included in the SYBIL DSS. This component is specifically for helping users to localize/regionalize models to the region in which they want to use the model. This component will be discussed in more detail in a following section.

Economic Models
Last, there two different economic model for the application of fertilizer, which help the user/farmer decide how much fertilizer to apply to crops. The first model considers historical weather and crop performance data, costs, and profits, and uses a simple extrapolation technique to give advice. On the other hand, the second model, also considers all these issues, but additionally, this model can take into account environmental costs, and utilizes a linear programming technique to give advice.

In addition to these six major components, this SYBIL DSS has the ability to present text to the user in all four speaking languages of the participating EC countries (that is, Italian, German, French, and Danish), plus the English language. By allowing easy installation of multiple languages, multiple instantiations of the system will not be required because everyone can use the same system. All the user has to do is specifying which language they prefer, and that language will be used for the entire time the system is running.

The design of this SYBIL DSS application platform has focused on portable tools which can be utilized on many different standardized computer system platforms. User friendly graphical user interface tools and artificial intelligence tools have been utilized so the system would be easy to use. In the end, these design criteria have resulted in a SYBIL DSS that is user friendly and portable between many popular computers, and establishes a common platform which can be used in the future by other model programmers.

3. METEOROLOGICAL DATA MANAGEMENT SYSTEM

To satisfy the requirement to provide robust data management within the SYBIL DSSes, a flexible and modular data management component was implemented. Because the SYBIL DSSes deal almost exclusively with meteorological data, this data management component focuses on meteorological data and is called the **Meteo C Library**. This library allows seamless access to meteorological data in any type of ASCII file or from SQL databases, thereby providing "generic" access (i.e., multiple access methods) to meteorological data by all models within a DSS.

Figure 3 illustrates the structure of this Meteo C Library, and how models interact with the library.

Interaction between a model and the Meteo C Library begins when the user inserts a user preference file into a model (i.e., stipulates which user preference file the model should use). This file specifies the data source (an ASCII file or an SQL database), the year for model calculation, which meteorological station data should be used (stat#), and other user preferences regarding the model calculation. After this file is inserted, and loading of data or model calculation is initiated, the Meteo C Library is called to initialize the

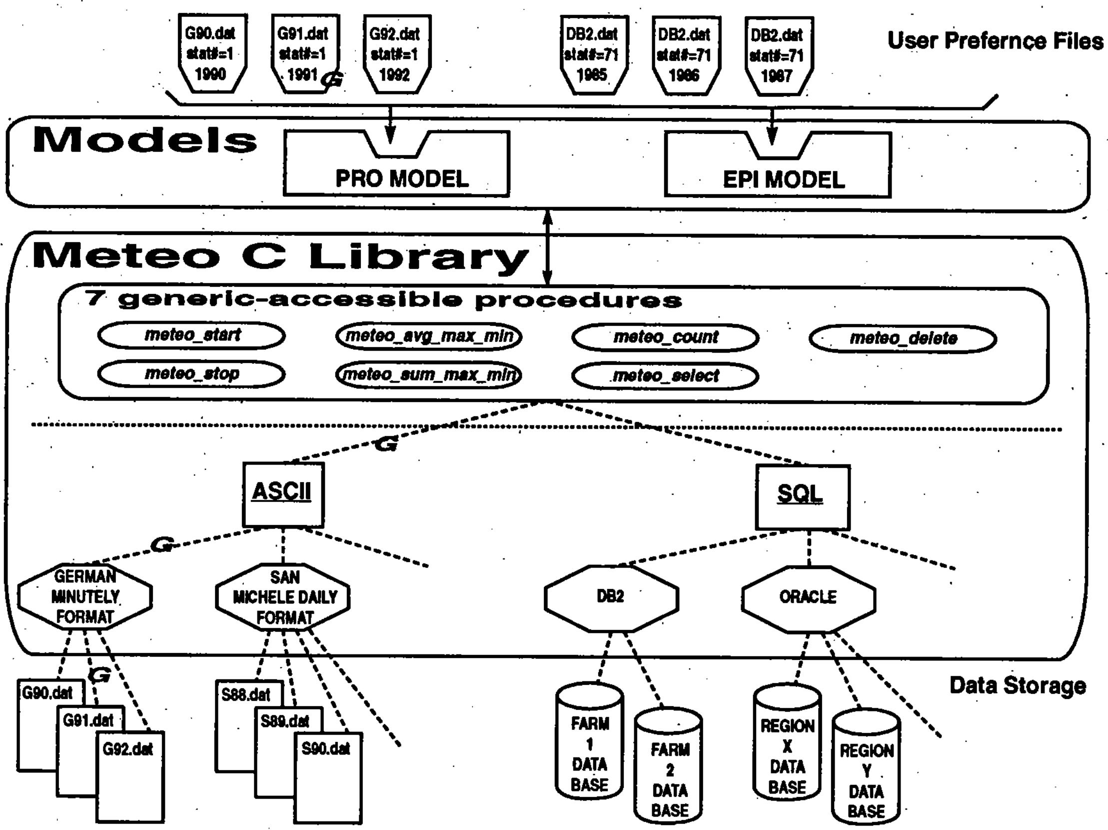

Figure 3. Structure of the Meteo C Library

meteorological data connection. Towards the goal of modular, extensible access to data, only seven "generic" library procedures have been made accessible to models. No matter which model or what type of data is being used, the model first calls the ***meteo_start*** procedure, passing to this procedure the data source file name, year, and station number. Upon the first call of this start procedure, no active data connections exist (i.e., in Figure 3, all dashed links below the seven accessible "generic" procedures are not formally connected). The start procedure then calls other procedures in the library that handle access to more specific types of data. These more deeply embedded procedures determine if the data source is a flat file (ASCII) or a database (SQL), and furthermore, the format type (e.g., German Minutely Format or San Michele Daily Format for ASCII sources, or Oracle or DB2 type databases for SQL). The library classifies the data source by looking directly into the top of the data source where keywords specifying type and format are placed. Each type of data has its own keyword which must be "registered" into the Meteo C Library. Using the keywords read from the data source, connections to the appropriate internal library procedures are made (and remembered) so that subsequent calls to the library (e.g., a call to ***meteo_avg_max_min***) will use the source of data specified in the last ***meteo_start*** call. When the model is finished with the data source, it calls ***meteo_stop***, and all connections are broken.

Figure 3 also gives an example of the data connection made when a user inserts a user preference file (shown in the figure as the file at the top, second from the left) which: (a) points to the *G90.dat* file, (b) specifies 1991 as the year for calculation, and (c) stipulates that data from station number 1 (stat#=1) be used. When model calculation is initiated with these parameters, links labeled **G** in the figure are "connected" and all subsequent requests for data use these links to supply data to the model.

In addition to "generic" data access to the library and dynamic adjustments to the type of data being used, a "standard" list of values/variables needs to be kept to achieve seamless data access. In the Meteo C Library, a list of all the possible values/variables (e.g., date, meteo station number, temperature, humidity, etc.) that can be handled by the library is maintained, and this list can be expanded as necessary if new data sources are "registered" into the library and they contain values which are not currently in the value/variable list. Furthermore, one agreement on the units of each value/variable is made (e.g., temperature in degrees Celsius), and then whenever data is read into the library, the data will be converted (if necessary) from the units stored in the data source, to the units used by the library. This ensures that data sent from the library to the models is of one particular type, and conversions do not have to be made in the model based on different data types. Overall, these features allow all data (no matter where they are from) to be stored in the same way inside the Meteo C Library, and only when reading the data from the data source into the library will it be necessary to perform conversions.

Lastly, because meteorological data is often found in many different frequencies (e.g., hourly, daily, etc.), the Meteo C Library handles frequency conversion from higher frequency to lower frequency (e.g., minutely to hourly) as needed, based on what data is requested from a model. This conversion is simply handled through averaging (as in the case of temperature) or summing (as in the case of rain) the more frequent values into the frequency or time period requested by the model.

In general, with these types of data management flexibility, when the DSS is to be transported to different regions where data is maintained in different ways, neither *models* within the DSS nor *data* will need to be changed so that the data can be used within the DSS (e.g., it will not be necessary to change model code nor to re-format data within an ASCII file nor to move data from an SQL database to an ASCII file or vice versa). From the user's point of view, data is automatically converted within this library between the units and granularity used to the units and granularity required. Therefore, computer models (agricultural or other) can access meteorological data in a standardized way (i.e., using the same calls and methods), no matter what type of data is being used (minutely, hourly, daily, weekly, monthly, or yearly - SQL database or ASCII file).

Under this scenario, the only action required when transporting the DSS is to *register* the meteorological data source within the Meteo C Library. Since the library is an independent service provider, it can be expanded to include additional data sources, and this will not have any effect on the model code. That is, new data sources (ASCII file formats or SQL database types) can be added without disrupting the model code which request meteorological data.

This independent characteristic is useful in the model development stage because it allows a model programmer to concentrate on developing their model without having to consider the origin of the meteorological data. When meteorological data is desired in the computer model, standardized calls to the Meteo C Library are made (asking for one of the standard values/variables which is handled by the library, and specifying the time period desired) and no matter what type of meteorological data is currently being used, the Meteo C Library will return the requested data after making any required frequency or unit conversions. This gives the programmer freedom from writing and re-writing code to access different types of meteorological data.

Overall, this library is a step towards establishing a common/standard meteorological data platform for the management of meteorological data within DSSes.

4. EXPERT SYSTEM COMPONENT BASED ON A HYBRID METHODOLOGY FOR LOCALIZING MODEL PARAMETER SETTINGS

4.1 Overview

One of the main problem we have encountered in transferring models is the problem of model accuracy in the new location. That is, often when a model developed in one region is used in a different region, the model outputs (such as, recommendations, results, and/or indicators) are inaccurate in the new region. We have categorized this situation into two cases:

- **(Case1)** the model is not transportable (i.e., it is too location-specific and not robust enough to allow different conditions to enter into the model) and
- **(Case2)** the model is transportable, but the model parameter settings need to be altered to allow the model to give accurate output values for the new region.

Case 1 is a simple situation (in that the model is just too specific to operate outside one area), but difficult to address and solve if it is strongly desired to transfer the model. This case will not be addressed within this discussion.

On the other hand, the dilemmas present in **Case2** (i.e., modifying/adapting/adjusting model parameter settings so that the model functions accurately in a new location) have lead us to develop an expert system component that determines location-specific model parameter settings. This component was developed based on a hybrid methodology specifically designed for models that can be placed into **Case2** (i.e., cases where it is beneficial to dynamically adapt the model parameter settings to the existing location). At the heart of this hybrid methodology is a genetic algorithm (GA) (an *artificial intelligence* search technique) linked with the agricultural risk assessment model engine. This is the combination that makes the methodology hybridized, and the general component created by this methodology is called an 'Agricultural Model-GA' or an AGMOD-GA.

4.2 Theory

The theory of this methodology is that by utilizing historical data from a particular region, a model's parameter settings can be *adapted* so that the new parameters allow the model to work well in the particular region. This adaptation is done by trying to match the model parameter settings (which should be easily changeable) to the particular region. That is, new model parameter settings are found which allows this model to give accurate output values (such as recommendations, results, and/or indicators) when run in the particular region. To find matching model parameter settings, intelligent search is performed which utilizes historical data as part of the objective function (i.e., the search attempts to *fit* the model parameter settings such that when used in the model, the model gives output near to the given historical data).

For example, suppose a model has been developed in **RegionA** that gives an assessment of the risk that a certain plant will be damaged by a certain fungus. This model uses an instantiation of model parameter setting values (such as values for thresholds and coefficients, etc.) called **ModelParmsA**. The model has proved quite useful in **RegionA**, and now agriculturalists in **RegionB** are interested to use this model. Upon first running this model in **RegionB** (with the original model parameter settings, i.e., **ModelParmsA**), the model gives output values (such as recommendations, results, and/or indicators) which are inaccurate. As an attempt to make this model work accurately in **RegionB**, we suggest that new model parameter settings (which are different from **ModelParmsA**) be derived. By applying the methodology discussed here (which requires historical data and an intelligent search method to be described below), it is possible to locate new model parameter settings, and once this is done, we can call this new instantiation of settings **ModelParmsB**. We propose that after applying this methodology, the model should produce accurate output values (about the same certain plant and the same certain fungus) in **RegionB** when using **ModelParmsB**.

Overall, by following this methodology, a component will be created which can search for good model parameter settings such that when the given model is applied and run at the location in question, the output values given will be consistent with the historical epidemiological data; moreover it is hoped that this will also the model to be generally used in this region, producing accurate output values on data which it has not seen. We can call this component that performs this search/adaptation an expert system component since it modifies and adjusts a model to work in a new location in the same way an expert would modify and adjust a model.

Additionally, it should be emphasized that this methodology is particularly appealing because it is not a strictly empirical or analytical, but both. That is, this methodology does not perform a search to *fit the historical data* from a particular location into an empirical algorithm; rather it performs the search in a larger context, *fitting the model parameter settings* to a particular location. Therefore, the resulting instantiation of the localized agricultural risk assessment model (with the new parameter settings inside) is as good (or as bad) as the original model; consequently, if the model is biologically significant (e.g., if it simulates biological events) then this is not lost by this localization methodology since the model is used in the same form (i.e., the structure of the model is left intact), only the model parameter settings are changed.

4.3 Elements

In general, this methodology prescribes the utilization of:
(i) historical situation data,
(ii) historical outcome data,
(iii) the agricultural risk assessment model, and
(iv) an intelligent search method (in this case, a genetic algorithm, also called a GA, which is an artificial intelligence search technique).

Historical Situation Data
Historical situation data is the basic data required by the model in question In the domain of agricultural risk assessment models, this often includes meteorological data since frequently this is an important input to the model. In most cases, the more historical model data that is available, the better.

Historical Outcome Data
The presence of historical outcome data plays a large part in how accurately a model will be adapted using this methodology. This is due to the basic fact that models accepts situations and computes outcomes. The historical outcome data will be used to fit the model parameter settings to the new region in question. Therefore, when constructing a component using the methodology described here, it must be possible to match model outputs to some combination of outcomes and/or events in the real-world (and there must be one-to-one correspondence). For risk assessment models, historical outcome data

regards the occurrence of fungus or pest problems in past years (epidemiological data); or for crop growth models, historical outcome data regards crop yield in past years.

Agricultural Risk Assessment Model

In this methodology, the agricultural risk assessment model (i.e., the engine or core of this model) is fundamental because it will be used to obtain evaluations of how well particular model parameter settings work in the given region (i.e., with the given data). In particular, the intelligent search method will repeatedly call upon this model engine as it constructs new model parameter settings that need to have their worth evaluated.

Intelligent search Method

The intelligent search method is an important part of this component because, in this particular domain of agricultural risk assessment models, knowledge of the domain is hard to codify (i.e., 'rules of thumb' are vague and difficult to construct), and the selection of an intelligent search method can help to alleviate this difficulty. This is due to the fact that intelligent search methods do not rely on 'rules of thumb', rather, rules are not required and these methods can actually facilitate the user in identifying 'rules of thumb'.

The selection of the actual intelligent search method to be employed was made among the following possible methods: hill-climbing, simulated annealing, and genetic algorithms (GAs). In the end, GAs were selected as the most desirable method because:

(a) they can perform unbiased search,
(b) they make no assumptions about the search space (i.e., the search space does not have to be smooth or regular),
(c) they carry out a more effective search of an irregular, multi-dimensional space because they search from a population of points rather than a single point,
(d) their search is not random, but intelligent (they utilize operators which are patterned after natural genetics), and
(e) they have been shown effective at finding optimal or near-optimal solutions to dynamic real-world problems (Goldberg, 1989) (Grefenstette, 1985) (Grefenstette, 1987) (Schaffer, 1989).

For a complete description of GAs, how they function, etc., refer to (Goldberg, 1989).

4.4 Agricultural Model-GA (AGMOD-GA)

Linking an Agricultural Risk Assessment Model to a Genetic Algorithm

To allow a GA to search the space of an agricultural risk assessment model's parameters, the agricultural model is linked to a GA, and the GA uses the model as the evaluation function. Furthermore, the model uses the historical situation data (as this is necessary to run the model in the given historical years), and the GA additionally uses the historical epidemiological data (discussed earlier) in combination with the output of the model. Whenever the GA wants to evaluate one instance of model parameter settings, the agricultural model is called, and the final outcome is returned to the GA so that a fitness can be computed. We call this resulting general component created by employing this hybridized methodology an 'Agricultural Model-GA' or an AGMOD-GA.

Figure 4 illustrates an agricultural model and a GA linked to form an AGMOD-GA.

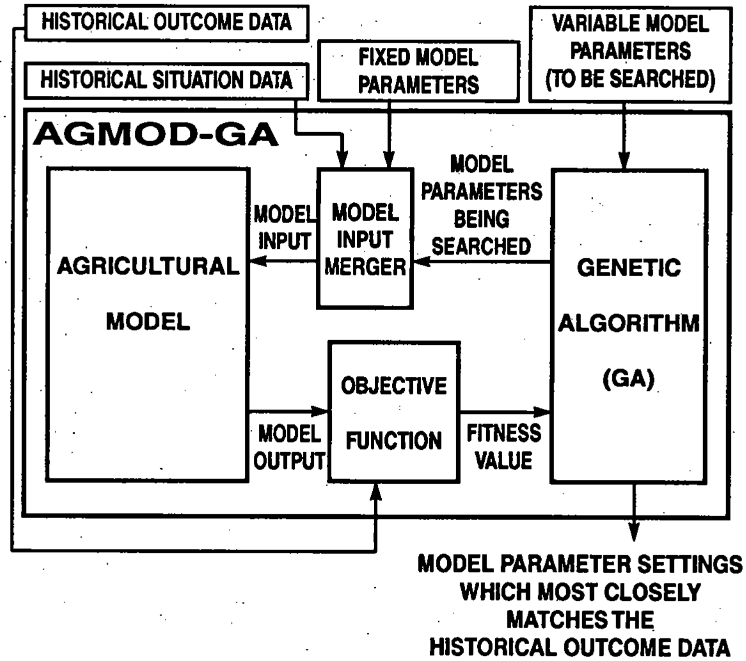

Figure 4. Structure of an AGMOD-GA

4.5 An Example - The P.R.O. Model for Grapes

First, an explanation of how the P.R.O. model was put into this SYBIL DSS is given. This is followed by the the problems with this inclusion, and how the AGMOD-GA was used to address these problems.

After deciding that the P.R.O. model (described earlier) was a good choice for inclusion into the DSS (and therefore a good choice for trying to transfer this model between countries), an instantiation of the model (with only small changes so that the model would accept other types of meteorological data) was programmed into the DSS, and test runs were made with various data from other regions (e.g., Würzburg, Germany and Trentino, Italy). Upon running these tests, it was found that the output values (which in the case of the P.R.O. model are: the primary infection date, the end of the incubation period, a list of special night occurrences, and a recommended spray date), were inaccurate in the new regions. That is, the P.R.O. model outputs were rejected by agricultural experts based on their historical epidemiological data (more generally, their historical outcome data) and general knowledge of when epidemiological events occur in their regions.

For example, Table 1 displays the results from running the P.R.O. model with data from an area inside the Trentino region of Italy. As this table shows, the dates produced by the original P.R.O. model using original model parameter settings (i.e., model parameter settings selected by Dr. Hill for the Rheinhessen area) (these dates shown in the column titled "Case 1") for data coming from Trentino, only approached the dates known to be correct from observations done by agricultural experts in Trentino (these dates shown in the column titled "Actual Dates") for the primary infection dates (rows titled "Prim Inf 19xx"). For the recommended spray dates (rows titled "Rec Spray 19xx"), the model could not even produce estimates of this date (with this data from Trentino) for two out of the three years in which actual dates were available for comparison. Therefore, the model in this state is of little use to agriculturalists in Trentino since it is generally not able to

Table 1. Results from running the P.R.O. model; Case 1 uses data from the Trentino region of Italy with original parameter settings used by Dr. Hill in Rheinhessen, Germany; actual dates in Trentino were only available for three years; date 12/31 is used to indicate the model never reached this date; difference is in absolute days

	Actual Dates	Case 1	Difference between Case 1 and Actual
Prim Inf 1988		5/3	
Rec Spray 1988		12/31	
Prim Inf 1989		5/13	
Rec Spray 1989		7/3	
Prim Inf 1990	5/24	6/5	12
Rec Spray 1990	6/7	12/31	207
Prim Inf 1991	5/11	5/11	0
Rec Spray 1991	6/7	12/31	207
Prim Inf 1992	5/5	5/23	18
Rec Spray 1992	6/1	7/10	39
Prim Inf 1993		4/30	
Rec Spray 1993		12/31	
Avg. Difference			80.50

produce an accurate estimate of the recommended spray date. Additionally, the difficulties observed in this case also held true for data taken from other regions, so overall the P.R.O. model was problematic because it did not give accurate output when run in regions external to where it was developed.

In addressing this problem, it is proposed that this model could be adapted to local conditions and that the problems stem from the fact that the model parameter settings were custom tailored to the region where it was developed (in this case, Rheinhessen, Germany). In addition, it is desirable to keep the same basic overall structure of the P.R.O. model because it had been proven valid and useful in the past. Therefore, the above described methodology was applied and an instantiation of an AGMOD-GA for the P.R.O. model was created, calling the new component the PRO-GA.

By running this PRO-GA component with data from a particular region, new model parameter setting values can be found which should allow the model to give accurate output values (i.e., recommendations, results, and/or indicators) for the region in question. Upon running the PRO-GA with three years (1990, 1991, and 1992) of historical situation data (in the case of the P.R.O. model, meteorological data) and historical outcome data (in this case, epidemiological data describing primary infection dates and recommended spray dates) from an area inside the Trentino region of Italy, new model parameter setting values were derived. Table 2 shows the results of using these new model parameter settings inside the P.R.O. model, again running with data from Trentino.

As expected, this table shows that the model parameter settings derived by the PRO-GA component allow the P.R.O. model to function with significantly increased accuracy on the Trentino data. The absolute average difference in days (i.e., the absolute average of how far off the model output is from actual dates) (in the row titled "Avg. Difference") between actual dates (column titled "Actual Dates") and dates produced by the running the P.R.O. model (column titled "Case 2") drops from 80.50 days to 0.83 days when new parameter settings derived by the PRO-GA are used instead of the original parameter settings. Generally, an average difference (i.e., an accuracy error) of 0.83 days is not significant or a critical inaccuracy, and this variation is well within the tolerable limits.

Table 2. Results from running the P.R.O. model; Case 1 is the same as shown in Table 1; Case 2 uses data from the Trentino region of Italy with parameter settings found by the PRO-GA using all available historical data; date 12/31 is used to indicate the model never reached this date; differences are in absolute days

	Actual Dates	Case 1	Difference between Case 1 and Actual	Case 2	Difference between Case 2 and Actual
Prim Inf 1988		5/3		5/3	
Rec Spray 1988		12/31		5/11	
Prim Inf 1989		5/13		4/19	
Rec Spray 1989		7/3		5/13	
Prim Inf 1990	5/24	6/5	12	5/24	0
Rec Spray 1990	6/7	12/31	207	6/6	1
Prim Inf 1991	5/11	5/11	0	5/12	1
Rec Spray 1991	6/7	12/31	207	6/7	0
Prim Inf 1992	5/5	5/23	18	5/4	1
Rec Spray 1992	6/1	7/10	39	5/30	2
Prim Inf 1993		4/30		4/30	
Rec Spray 1993		12/31		6/23	
Avg. Difference			80.50		0.83

Because this adaptation is performed with all available historical sets of data, this type of behaviour is generally expected since accuracy is tested on the same set of historical model data (in this case, meteorological data) that was used for adapting. This is still an important result because it shows that the model can be fit to the entire set of data, and that it is possible to find parameter settings that will give accurate results over many years. On the other hand, a better verification of this methodology is to adapt the model using one subset of historical model data, and then test the accuracy of the adapted model (i.e., the model with the new parameter settings) by running the model on a different subset of historical data.

Unfortunately, the currently available historical outcome (viz. epidemiological) data is very limited (three years of data from Trentino); therefore only a few small subsets of historical model data can be formed (in this case, only three interesting subsets: (1991, 1992), (1990, 1992) and (1990, 1991)). Therefore, it is not possible to test newly derived parameter settings as extensively as desired, but three tests have been performed, the results of which are shown in Table 3.

Even with the limited amount of historical data, these results are still interesting and significant. For example, when adapting the model using historical outcome data from 1990 and 1991 (shown in the last column of the table) the average difference between actual results and P.R.O. model produced results using the new parameter settings is only 2.00 (not shown in table, but computed separately). This is a lower accuracy than the average difference shown in Table 2 using an adaptation with all three historical data sets, but it is believed still to be quite adequate.

It is believed that if an increased amount of historical outcome data was available, and this was used in the adaptation process, that the adaptation should become even more robust, increasing the probability that accurate results are produced in the future when the model is running in real-time, giving output values to the agriculturalist for use in making intelligent crop management decision.

Table 3. Results from running the P.R.O. model; Case 1 is the same as shown in Table 1; Cases 3, 4, and 5 use data from the Trentino region of Italy with parameter settings found by the PRO-GA using three different subsets of the available historical data; Case 3 used only 1991 and 1992 data; Case 4 used only 1990 and 1992 data; Case 5 used only 1990 and 1991 data; bold results indicate results from data sets which were not used in adaptation; date 12/31 is used to indicate the model never reached this date

	Actual Dates	Case 1	Case 3	Case 4	Case 5
Prim Inf 1988		5/3	5/3	5/3	5/3
Rec Spray 1988		12/31	5/11	5/11	5/11
Prim Inf 1989		5/13	4/19	4/19	4/19
Rec Spray 1989		7/3	5/13	5/13	5/13
Prim Inf 1990	5/24	6/5	5/24	5/24	5/24
Rec Spray 1990	6/7	12/31	6/4	6/6	6/6
Prim Inf 1991	5/11	5/11	5/12	5/12	5/12
Rec Spray 1991	6/7	12/31	6/7	6/17	6/7
Prim Inf 1992	5/5	5/23	5/4	5/4	5/4
Rec Spray 1992	6/1	7/10	5/30	5/30	5/23
Prim Inf 1993		4/30	4/30	4/30	4/30
Rec Spray 1993		12/31	6/23	6/23	6/23

5. CONCLUSION

Agricultural computer models have been integrated into a decision support system (DSS) that was developed under an EC project called SYBIL. The DSS discussed in this paper addresses the plant protection needs of grape and apple growers by providing temporal information to assist these growers in the management of crops with respect to controlling fungus and pests.

Because the project is particularly interested transporting model technology between countries (i.e., the moving of functional and useful agricultural risk assessment models that are developed in one location to a new location so they can be used in this new location), and these transfers can be problematic, a transportable DSS platform has been created. The two main components of this transportable DSS platform are: (1) a meteorological data management system to handle data in varying formats (SQL databases and ASCII flat files) and (2) an expert system component which allows for the intelligent localization of models to the region in which they are being used.

ACKNOWLEDGEMENTS

Funding has been provided for the SYBIL project (project number PL 900615) by the EC CAMAR Program (Competitiveness in Agriculture). The authors wish to thank Dr. Val Reilly, SYBIL project supervisor from the European Commission, for his support and encouragement.

Furthermore, the authors would like to thank all EC Project SYBIL partners: Technische Universitaet Muenchen-Weihenstephan (Technical University of Munich) (TUM) (D), Bayerischen Staatsministerium für Ernährung, Landwirtschaft und Forsten (BALIS) (D), Association De Coordination Technique Agricole (ACTA) (F), and Danish Institute of Plant and Soil Science (DIPSS) (DK). In particular, thanks go Dr. Georg Wendl and Dr. Georg Froehlich at TUM; Institut und Bayerische Landesanstalt fuer Landtechnik (Institute of Agricultural Engineering) in Freising, Germany for their ideas, work, and collaboration on the meteorological data management component (Meteo C Library) and to EC Project SYBIL's coordinator Dr. Johann Bergermeier at BALIS in Munich, Germany and Dr. J.V. Herrmann at Bayerische Landesanstalt für Weinbau und Gartenbau

in Würzburg, Germany for their assistance with the P.R.O. model.

REFERENCES

Goldberg, D.E. 1989. Genetic Algorithms in Search, Optimization, and Machine Learning. Addison-Wesley, Reading, MA, USA.

Grefenstette, J.J. 1985, ed. Proceedings of an International Conference on Genetic Algorithms and Their Applications. Carnegie-Mellon University, Pittsburgh, PA, USA, July 24-26, 1985. U.S. Navy Center for Applied Research in Artificial Intelligence, USA.

Grefenstette, J.J. 1987, ed. Proceedings of the Second International Conference on Genetic Algorithms. Massachusetts Instituted of Technology, Cambridge, MA, USA, July 28-31, 1987. L. Erlbaum Associates, USA.

Hill, G.K. 1989. Mit P.R.O. System Spritzungen Sparen. Die Weinwirtschaft, 125, pp. 28-31.

Hill, G.K. 1990a. Das Peronospora-Riskoprognosemodell Oppenheim (P.R.O.) im Praxistest. Der Deutsche Weinbau, 44 (12/27 April 1990), pp. 514-517.

Hill, G.K. 1990b. Studies on Plasmopara Viticola Epidemics in the Vineyard. Proc. Atti del Convengo su Modelli Euristici ed Operativi in Agricoltora, Caserta, Italy, 27-29 Sept. 1990, Societa Italiana de Fitoiatria ed., pp. 266-273.

Hill, G.K. 1990c. Plasmopara Risk Oppenheim - A Deterministic Computer Model for the Viticultural Extension. Proc. Atti del Convengo su Modelli Euristici ed Operativi in Agricoltora, Caserta, Italy, 27-29 Sept. 1990, Societa Italiana de Fitoiatria ed., pp. 182-194.

Hill, G.K. 1993a. Severe Primary Infections of Plasmopara Viticola at bloom-Causal Climatic Conditions and Threshold of Damage on Bunches. Proc. OILB Meeting on Integrated Control in Viticulture, Bordeaux, France, 2-5 March, 1993.

Hill, G.K., Breth, K., Spies, S. 1993b. The Application of the P.R.O. - Simulator for Minimizing of Plasmopara Sprays in the Frame of an Integrated Control Project in Rheinhessen/Germany. Vitic. Enol. Sci., 48, pp. 176-183

Schaffer, J.D. 1989, ed. Proceedings of the Third International Conference on Genetic Algorithms. George Mason University, Fairfax, VA, USA, June 4-7, 1989. Morgan Kaufmann Publishers, Inc., USA.