Diagnosis of Sparse Adoption Data Using an Expert System-Guided Innovation Diffusion Simulation Model

Alan J. Thomson

Pacific Forestry Centre
506 West Burnside Road
Victoria, BC, Canada V8Z 1M5
email: athomson@pfc.cfs.nrcan.gc.ca
Phone: (250) 363-0632
Fax: (250) 363-0775

Abstract

Forecasting of the likely rate of uptake of an innovation, idea or product is at the heart of business case development, but adoption of simulation models to produce such forecasts is often low due to complexity faced by users in selecting appropriate sets of parameters that provide a best fit to available observations. In addition, when observations are sparse, correlation between observed and predicted values is not an appropriate guide to model appropriateness. The present study describes (a) a computer simulation of the innovation diffusion process, and (b) an analysis and diagnosis package that can be added to simulation models to interpret output and advise on parameter changes to bring model scenarios in line with observations. The approach is two-part. First, envelopes are created around different scenario outputs and the relationship of the observation set to the envelopes determined. Subsequently, a rule mapping table edited by an expert user, links parameter modification options to specific aspects of the observations-envelope relationship. The expert system thus acts as a change agent to reduce the complexity of simulation model use.

Keywords: change agent; computer simulation; diffusion of innovation; forecasting; user acceptance;

Diagnosis of Sparse Adoption Data Using an Expert System-Guided Innovation Diffusion Simulation Model

Introduction

Forecasting of the likely rate of uptake of an innovation, idea or product is at the heart of business case development.   Simulation models are important members of the repertoire of computer-based forecasting tools in a wide range of fields but often fail to meet their objectives in terms of user acceptance (Lynch and Gregor 2003). Failures are often attributed to system complexity (Jenkins and Chapman 1998, Newman *et al*. 2000).   In particular, complexity is very dependent on the number of features available: initial ready acceptance due to a simple-appearing user interface "soon fragments into complexity as the novice modeler starts to confront the tacit detail of the real world" (Jenkins and Chapman 1998: 1553). However, simulation models are often used as there is no other choice if one requires estimation of future states for decision making purposes (Thomson *et al*. 2007).

Smith and Guerrero (1994) indicate that a decent fit between field data and model predictions is almost the only test of model validity that wins general acceptance. However, correlations between observed and predicted values may be poor, especially where data are sparse, and other characteristics of the model and data may be more appropriate guides to a good fit, such as whether they are monotonic, asymptotic or non-monotonic. Users who are unfamiliar with the workings of a simulation may be at a loss regarding how to proceed when different combinations of parameter settings produce similar-appearing results, and this situation is exacerbated when the model user has only sparse data to guide their choices.

Linking Artificial Intelligence, particularly expert systems, to simulation models has been used in attempts to increase user acceptance. Greer *et al*. (1994) built a hybrid decision support system which first interpreted the output of a quantitative simulation model and then constructed explanations, which in turn were adapted to the particular user. Thomson and Ross (1996) developed an irrigation scheduling tool in which parameters were adjusted as the system was used, based on soil water sensor responses to drying. An expert system determined which sensor readings were valid before they were used to adjust parameters.

Doyle *et al.* (2003) recognized that many simulations are treated as "black boxes", with outputs being produced without any explanation of how the output was derived, and that a system must give plausible explanations for any decisions it makes. Greer *et al*. (1994) addressed simulation "black boxes" by developing a system that created explanations suitable for the agricultural community. They described a system consisting of a User Interface, a Simulation Model, an Explanation Model and a User Model. The user interface queries the user for suitable inputs and outputs results along with an explanation of the simulation to the user. The simulation model carries out the simulation on the users input, and passes the results to the explanation module. Inputs from the simulation model and user model are combined by the explanation module to create an individualised explanation of the simulation which is then passed back to the user via the interface.

As described above, demonstration of fit of existing observations, even if sparse or of uncertain quality, to model output, is often a prerequisite to model adoption. The present study

therefore takes a different perspective on providing explanations. Here, an expert system evaluates the relationship of sparse data to model output in different model scenarios and explains changes in model parameterization that would be required to bring model performance in line with the observations. As desire for increasing user acceptance and system adoption underlies research in this area, the simulation model used as the case study for development of the prototype expert system is a simulation model of Innovation Diffusion, specifically developed to explore success of computer system development (Thomson *et al.* 2004) and subsequently adapted to run over the World Wide Web.

System concepts

Situation analysis
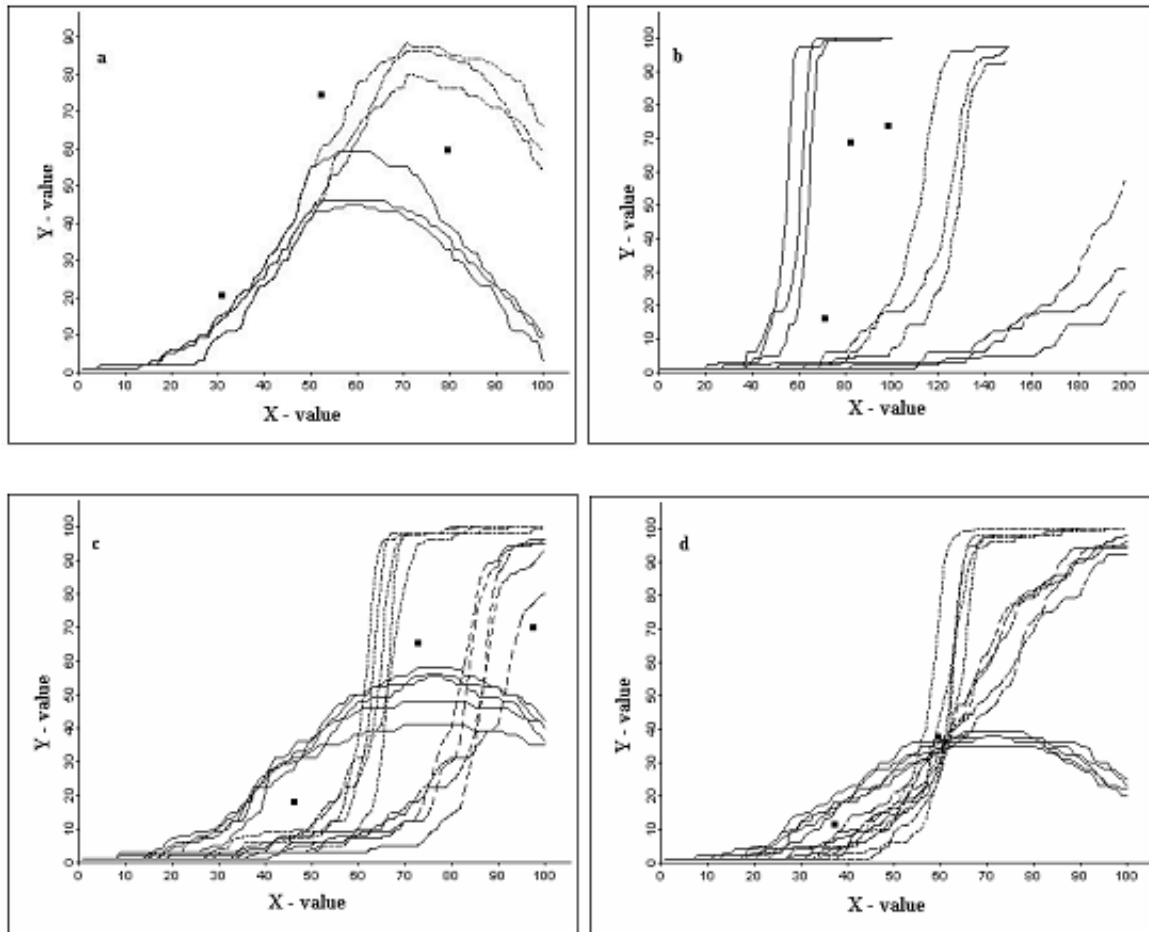
Figure 1. Simulation modelling with sparse data points plotted in relation to multiple runs of a stochastic simulation using different parameter values.

The situation often facing a simulation model user is illustrated in Figure 1, where multiple runs of a stochastic simulation are plotted on the same graph using parameter values for different scenarios. Observations, which may be sparse, may be relatively close to the estimated values (Fig. 1a), or at least to the form of the relationship (Fig. 1b). However, it may be hard to select among sets of observed and expected values (Fig. 1c), or there may be good fit to one or more

scenarios up to a point (Fig. 1d). Figure 1d also indicates a situation that might arise post-diagnosis, where model parameters have been changed based on the advice provided to bring the models in line with the observations, with the simulation results giving an idea of possible trajectories.

In the examples (Fig. 1), "sparse" data refers to only two or three observations, but "sparse" in relation to the multivariate "curse of dimensionality" can actually refer to a much larger number of points, with related problems of determining outliers (Rocke and Woodruff 1996). Outlier determination is not explored in the present study.

System design

Figure 2 illustrates the system design. An analysis and diagnosis system with associated database, developed in Java, can accept the output of a simulation model, compare the output with observations, and provide advice on adjusting parameters to bring model behavior in line with the observations.
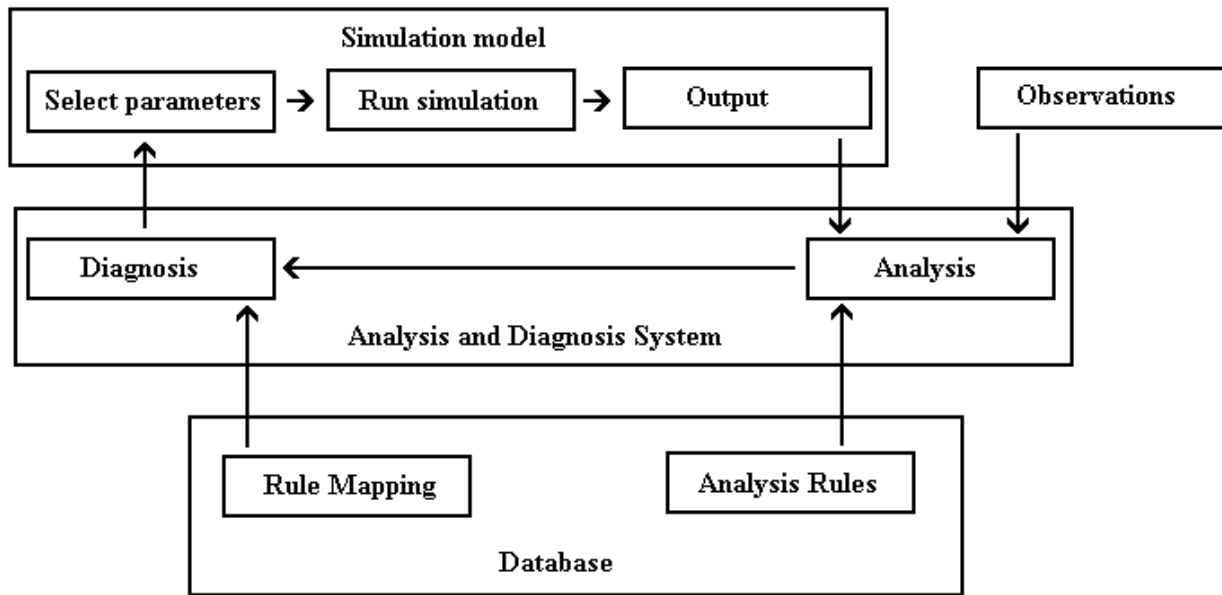

Figure 2. System components. The analysis and diagnosis system with associated database compares the observations with model output and advises on parameter adjustment.

Simulation model

The analysis and diagnosis functionality can be added to any simulation model or be used as a stand-alone system to explore libraries of previously saved system output files. The present study illustrates a stochastic simulation specifically designed to compare multiple sets of parameters (up to three sets) with up to five replicates (different random sequences), accept entry of observations, and apply the analysis and diagnosis functions.

The model used is a simulation of the Innovation Diffusion process (Thomson *et al*. 2004), based

on Rogers (1995). The original model was a system dynamics model developed in an early version of Simile[1], a visual modelling environment for ecological, biological and environmental research (Fig. 3). The present version was redesigned in Java in the form of a simulation model to link with the analysis and diagnosis package and provided with a web-based interface that included an integrated graphics package.
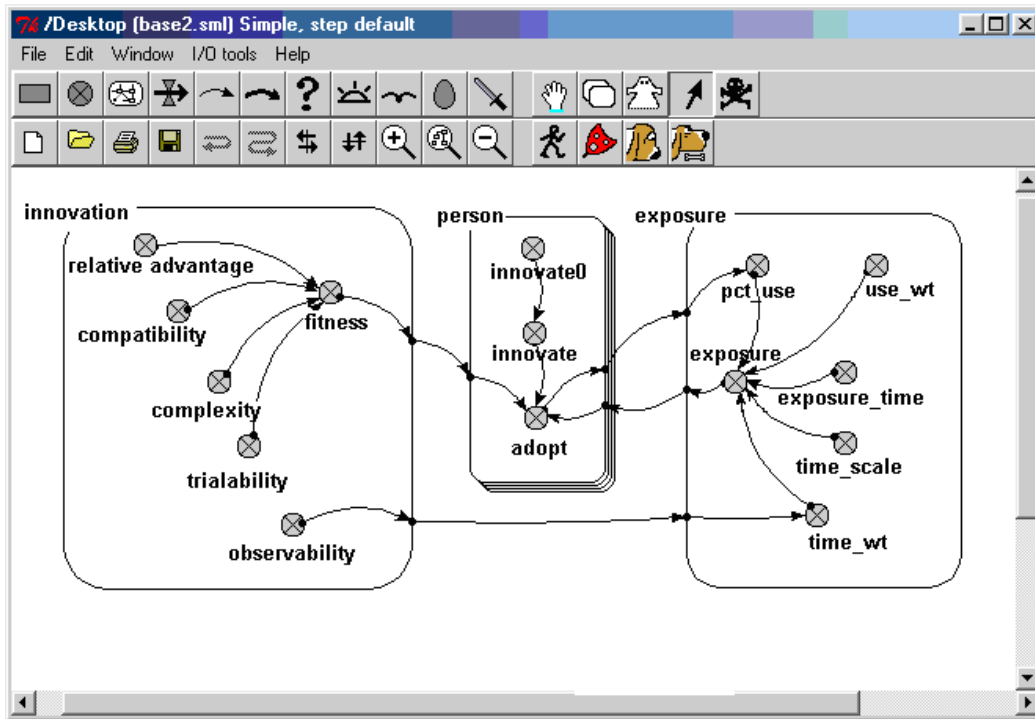

Figure 3. System diagram in the Simile modelling environment. Influence arcs connect the system variables, which are grouped in submodels (from Thomson *et al.* 2004).

Rogers (1995:5) defined diffusion as "the process by which an **innovation** is **communicated** through certain **channels** over **time** among **members of a social system**". He describes innovations in terms of five attributes: relative advantage, compatibility, complexity, trialability and observability. Relative advantage is the degree to which an innovation surpasses the idea it supersedes; compatibility is the degree to which the innovation is compatible with the values, needs and norms of the population; complexity is the extent to which the innovation is perceived as difficult to understand or use; trialability measures the ease with which people can try out the innovation, while observability is the degree to which others can observe the innovation.

The model of Thomson *et al.* (2004) therefore contained three principal submodels (Fig. 3): (a) the innovation; (b) person, in which the stack structure (Fig. 3) indicates an array of individuals simulated; i.e. the social system, and (c) exposure, which combines communications, channels and time. Individuals within a population who are exposed to innovations are characterized as falling into five main categories in a continuous distribution (Rogers 1995). Innovators, who are the first to take up a new idea, comprise less than 3% of the population. The

---

[1] http://www.simulistics.com/

rest of the population is made up of 13% early adopters, 34% early majority, 34% late majority and the remaining 16% being laggards.   Individuals are assigned a random level of innovativeness based on this distribution.

Figure 3 illustrates two unique features of the model. Firstly, a measure of innovation fitness was developed in order to be able to compare innovations with different characteristics. The fitness measure combined four of the five attributes of innovations - relative advantage, compatibility, complexity and trialability - in a single measure (see Appendix). The fifth attribute, observability, had a separate effect, acting through the second feature of the model, "exposure".

Rogers' (1999: 5) definition of diffusion was based on communication through different channels. Communication channels were not explicit in the model. Instead it was assumed that the communication pattern within the social system would be determined by (a) the number of previous adopters, expressed as percent use (*pct_use* in Fig. 3) and (b) the length of time since introduction of the innovation (*exposure_time* in Fig. 3), with the relative importance of (a) and (b) being set by weighting factors (*use_wt* and *time_wt* respectively in Fig. 3). It was also assumed that, other things being equal, more observable innovations would have higher exposure, thus Observability modifies the weighting factor for time exposed. An individual adopts once the exposure level multiplied by the fitness exceeds their individual innovativeness level.

In the web-based simulation (Fig. 4), up to three scenarios can be explored simultaneously, with up to five iterations (different random sequences) per scenario. Innovation and exposure parameters are as described above. A time scale is set in terms of days, weeks or years: days are used in the examples of the Figures. The population parameters set the number of individuals in the population and establish the distribution of innovativeness values, with "3" setting the distribution to give the 3-13-34-34-16 percentages above; i.e. a normal distribution with mean: 0.5, standard deviation: 0.1667. The other selections provide a more innovative or less innovative population. The advanced features allow for a specified rate of population turnover, abandonment of the innovation, or the introduction of a change agent who can influence the complexity, trialability and/or observability for a specified period (normally early in the process). Model results are displayed on a graph as in Figure 1: observations can then be entered by clicking on the graph at appropriate locations on the graph and selecting the "Analyze" button once all observations have been entered. This initiates the analysis and diagnosis processes.
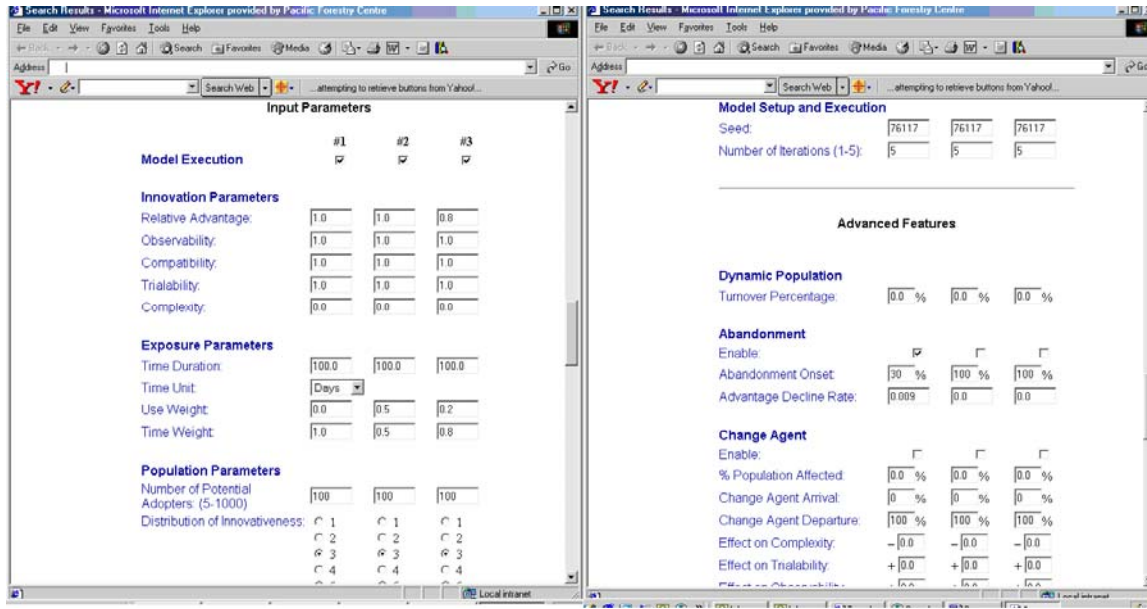
Figure 4. The input screen of the web-based implementation of the simulation model: the view on the right is obtained by scrolling down from the view on the left. Parameter settings are those used to generate Figure 1d.
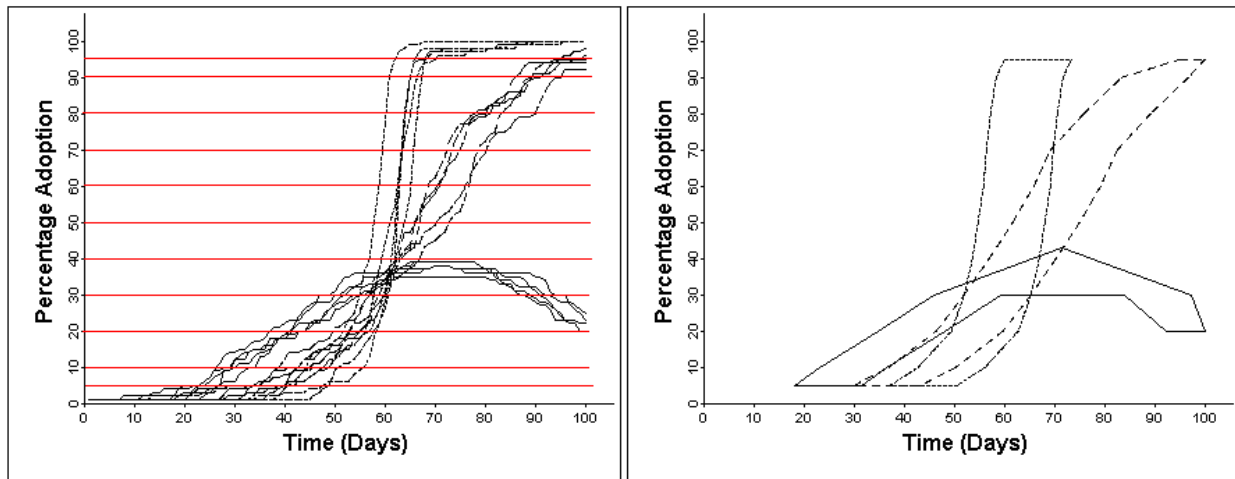
Analysis

Figure 5. Envelopes created around the average line for each scenario of Figure 1d. Left: the Y-values at which the envelope is determined. Right: the envelopes.

Analysis begins by creation of an envelope around the simulation output values for each scenario (set of parameter values). Where there are multiple replications of a scenario output, such as with a stochastic model, an average curve is first created. For the Innovation Diffusion model, analysis centers on the change in percentage adoption with time (Fig. 5). The envelope is created by calculating the X-value of the average curve at specified Y-values (5, 10, 20, 30, 40, 50, 60, 70, 80, 90, and 95 percent adoption in this case) (Fig. 5), and a pre-set value is added to

and subtracted from that X-value. A wider envelope reflects willingness to accept a model formulation for which the simulated values do not exactly correspond with the data.

In non-monotonic cases where there is a peak in the data, this is performed for both the ascending and descending parts of the curve. In addition to the maximum Y-value and the X-value at which it occurs, the maximum positive slope of each envelope is determined as well as a boolean value to record if a downturn occurs. If a downturn occurs, the maximum negative slope is also determined.

The properties of the observation set are then determined and evaluated before proceeding: points must be within the 5-95% range covered by the envelopes. Existence of any peak or asymptote in the data is determined and the observations are then compared to the envelope properties. For each model, the number of observations lying within the ascending and descending arms of each model envelope is determined. For those observations lying outside the envelope, the cumulative X-values to left or right of the envelopes are determined in order to guide advice-giving. Where the envelope has a peak or an asymptote at less than 95% adoption, the number of observations lying above the envelope is determined. Finally, the ratio of the maximum slope of the observation point set to the maximum slope of the ascending portion of each envelope is determined. A similar ratio of minimum slopes is established for the descending portions, where these occur.

Diagnosis

Diagnosis is conducted by a Java program that processes rules in a Microsoft Access database. Each rule has a header record and a variable number of clause records. Both header and clause records use the same table, but not all fields are used with each type. Fields include a Predicate name (header), Rule number, Clause number, "or" usage, Modifier/Variable/Condition/Value (not used in header), and four message fields. The four fields Modifier/Variable/Condition/Value are used together in a clause   The Modifier "not" allows negation of the fact that the Variable has a logical Condition related to a particular Value. The message fields include a Success Message (Table 1) and a Failure Message that are embedded in the processing history as appropriate. As systems produced by the Canadian government must be bilingual, a French version of each message is provided.

To this point, observations have been evaluated in relation to properties of a generalized envelope. To provide advice, however, these generalized properties must be linked to system-specific attributes. This is accomplished by use of the Rule Mapping Table (Fig. 2; Table 2). Actions in the table all relate to changes in the parameters in Figure 4.

Table 1. Predicate name of rules and associated Success Messages. The "R_" convention distinguishes rules from various forms of variable names, as a rule can be used as a clause within another rule. Note that the Success Messages end in colons, preparing for addition of text from clause Success Messages.

Predicate	Success Message
R_DecreaseAsymptote	The asymptote must be reduced using one of the following methods:
R_DecreasePeak	The peak must be reduced using one of the following methods:
R_DropPeak	The peak must be removed by one of the following methods:
R_FasterDecline	A faster decline must be produced using one of the following methods:
R_FasterIncrease	A faster rate of increase in the model should be produced using one of the following methods:
R_IncludePeak	A peak must be introduced by one of the following methods:
R_IncreaseAsymptote	The asymptote must be increased using one of the following methods:
R_IncreasePeak	The peak must be increased using one of the following methods:
R_MoveLeft	The model must be moved to the left using one of the following methods:
R_MoveRight	The model must be moved to the right using one of the following methods:
R_SlowerDecline	A slower decline must be produced using one of the following methods:
R_SlowerIncrease	A slower rate of increase in the model should be produced using one of the following methods:

Table 2. The Rule Mapping Table (RP is Rule Precedence, OP is Option Precedence). This table is edited by an expert user of the specific simulation model. Values shown are for the Innovation Diffusion simulation model.

RuleMapping			
Predicate	RP	OP	OptionMessage
R_IncludePeak	1	1	Enable the Abandonment feature to create a peak in the model to match the data.
R_DropPeak	2	1	Disable the abandonment feature from the model as the data do not yet exhibit a peak followed by decline.
R_MoveLeft	3	1	Move the model towards the left by selecting a shorter time duration.
R_MoveLeft	3	2	Selecting an index of Distribution of Innovativeness values closer to 1 will move the model further towards the left.
R_MoveLeft	3	3	Increasing the Time weight and decreasing the use weight will result in an earlier start.
R_MoveLeft	3	4	Increasing the Relative Advantage, Compatibility, Observability or Trialability towards 1.0 or decreasing the Complexity towards 0.0 will move the model to the left.
R_MoveLeft	3	5	A change agent early in the adoption process can move the model to the left.
R_MoveLeft	3	6	A novelty effect (not yet included in the system) may move the model to the left.
R_MoveRight	4	1	Move the model towards the right by selecting a longer time duration.
R_MoveRight	4	2	Selecting an index of Distribution of Innovativeness values closer to 5 will move the model further towards the right.

RuleMapping			
Predicate	**RP**	**OP**	**OptionMessage**
R_MoveRight	4	3	Decreasing the Time weight and increasing the use weight will result in a later start.
R_MoveRight	4	4	Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will move the model to the right.
R_IncreaseAsymptote	5	1	Increasing the Relative Advantage, Compatibility, Observability or Trialability towards 1.0 or decreasing the Complexity towards 0.0 will result in more adoptions in a given time.
R_DecreaseAsymptote	6	1	Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will result in fewer adoptions in a given time.
R_DecreaseAsymptote	6	2	Enabling the dynamic population option with higher turnover percentages will decrease the asymptote
R_IncreasePeak	7	1	Increase the peak by moving the Abandonment onset closer to 100%.
R_DecreasePeak	8	1	Decrease the peak by moving the Abandonment onset further from 100%.
R_FasterIncrease	9	1	Decreasing the Time weight and increasing the use weight will result in a faster increase.
R_FasterIncrease	9	2	Increasing the Relative Advantage, Compatibility, Observability or Trialability towards 1.0 or decreasing the Complexity towards 0.0 will result in a faster increase in the model results.
R_SlowerIncrease	10	1	Increasing the Time weight and increasing the use weight will result in a slower increase.
R_SlowerIncrease	10	2	Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will result in a slower increase in the model results.
R_FasterDecline	11	2	Increasing the advantage decline rate will produce a faster decline in the model results.
R_SlowerDecline	12	2	Decreasing the advantage decline rate will produce a slower decline in the model results.

For each rule of Table 1, the Rule Mapping Table (Table 2) provides one or more possible approaches to obtaining the desired result. An expert user defines the order in which rules are most appropriately applied (Rule Precedence). For that rule, options are then ranked (Option Precedence). For each option, the message is expressed in terms of adjusting specific parameters available through the user interface (Fig. 4). Note that it is also possible to provide guidance to possibilities not included in the simulation: R_MoveLeft (RP 3) includes (OP 6) "A novelty effect (not yet included in the system) may move the model to the left." In practice, the system user will have knowledge of the extent to which particular options may be appropriate for their specific situation. Table 3 shows the text displayed in the web page for the case illustrated in Fig. 1c.

Table 3. Diagnosis of the case illustrated in Figure 1 c. The text is displayed in the web-based interface when the "Analyze" button is selected.

Model: 1

The peak must be removed by one of the following methods:

1. Disable the abandonment feature from the model as the data do not yet exhibit a peak followed by decline.

The model must be moved to the left using one of the following methods:

1. Move the model towards the left by selecting a shorter time duration.
2. Selecting an index of Distribution of Innovativeness values closer to 1 will move the model further towards the left.
3. Increasing the Time weight and decreasing the use weight will result in an earlier start.
4. Increasing the Relative Advantage, Compatibility, Observability or Trialability towards 1.0 or decreasing the Complexity towards 0.0 will move the model to the left.
5. A change agent early in the adoption process can move the model to the left.
6. A novelty effect (not yet included in the system) may move the model to the left.

A slower rate of increase in the model should be produced using one of the following methods:

1. Increasing the Time weight and increasing the use weight will result in a slower increase.
2. Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will result in a slower increase in the model results.

Model: 2

A slower rate of increase in the model should be produced using one of the following methods:

1. Increasing the Time weight and decreasing the use weight will result in a slower increase.
2. Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will result in a slower increase in the model results.

Model: 3

A slower rate of increase in the model should be produced using one of the following methods:

1. Increasing the Time weight and decreasing the use weight will result in a slower increase.
2. Decreasing the Relative Advantage, Compatibility, Observability or Trialability towards 0.0 or increasing the Complexity towards 1.0 will result in a slower increase in the model results.

Discussion

An expert system has been developed to guide the users of simulation models through the selection of parameter combinations when observations available to guide the process are sparse. The expert system thus acts as a change agent reducing system complexity. The process first creates envelopes around the outputs of different simulation scenarios, and then evaluates the observation set in relation to generalized envelope properties. Advice on parameter adjustment is then provided based on expert user-defined Rule Mapping table which prioritizes rules and options appropriate for the specific simulation model.

As mentioned in the Introduction, Smith and Guerrero (1994) indicate that a decent fit between field data and model predictions is almost the only test of model validity that wins general acceptance. With sparse data, correlations between observed and predicted values are not appropriate guides and visual comparison of observed and predicted values is the approach often used. The present approach uses an objective comparison of observed and predicted values through the use of slope comparisons, and shape comparisons such as peaks or asymptotes.

However, although objective, some issues must still be resolved. Figure 1c and the diagnosis in Table 3 indicate that the objective comparison determined that a) the observations had not reached an asymptote, and b) all three models had faster rates of increase at some point than the maximum of the observations' slopes (between the first two points in the example). The objective comparison is based on arbitrarily-set threshold values, but a 1:20 and a 1:40 slope are visually similar with regard to determining an asymptote, while a 50:1 slope and a 30:1 slope appear similar with regard to maximum slopes. The key point is that setting the thresholds may have system-specific features. For the Innovation Diffusion system, thresholds that allow for considerably more variability than used in the case study (Fig. 1c, Table 3) may be appropriate.

In Table 3, Model 1 (Fig 1c, the model with the peak) appears to be the poorest fit to the observations, requiring three different forms of adjustment to fit the observations, whereas there is no difference between the appropriateness of Models 2 and 3, both with the same single adjustment required. If a rule were developed that advised on the possibility of a peak between observations that flatten (the two rightmost points of Fig. 1c), then Model 1c might actually be the one most appropriate to pursue.

Similarly, Figure 1c shows the spread of observations crossing the S-shaped curves for models 2 and 3. The diagnostic system therefore did not advise moving either left or right. However, with a different threshold, an asymptote to the observations would have been found, leading to differentiation of Models 2 and 3. An alternate approach to right/left advice might be based only on points in the lower half of the Y-scale, or by comparing X-values at the positions of maximum slope.

As with the thresholds, width of the envelope around the line representing the model scenario is arbitrary, and this will influence the diagnosis; decision on width may have system-specific features. The line representing the model scenario was also arbitrarily determined; i.e. by the average of the multiple stochastic runs. With stochastic simulations and threshold effects, this may not be the best approach (Fig. 6). An extension of this concept is that no simulation model is actually required to generate advice: the X/Y coordinates of the lines could come from a library of hand-drawn curves, with the Rule Map being purely based on an experts' understanding of the dynamics of the system.
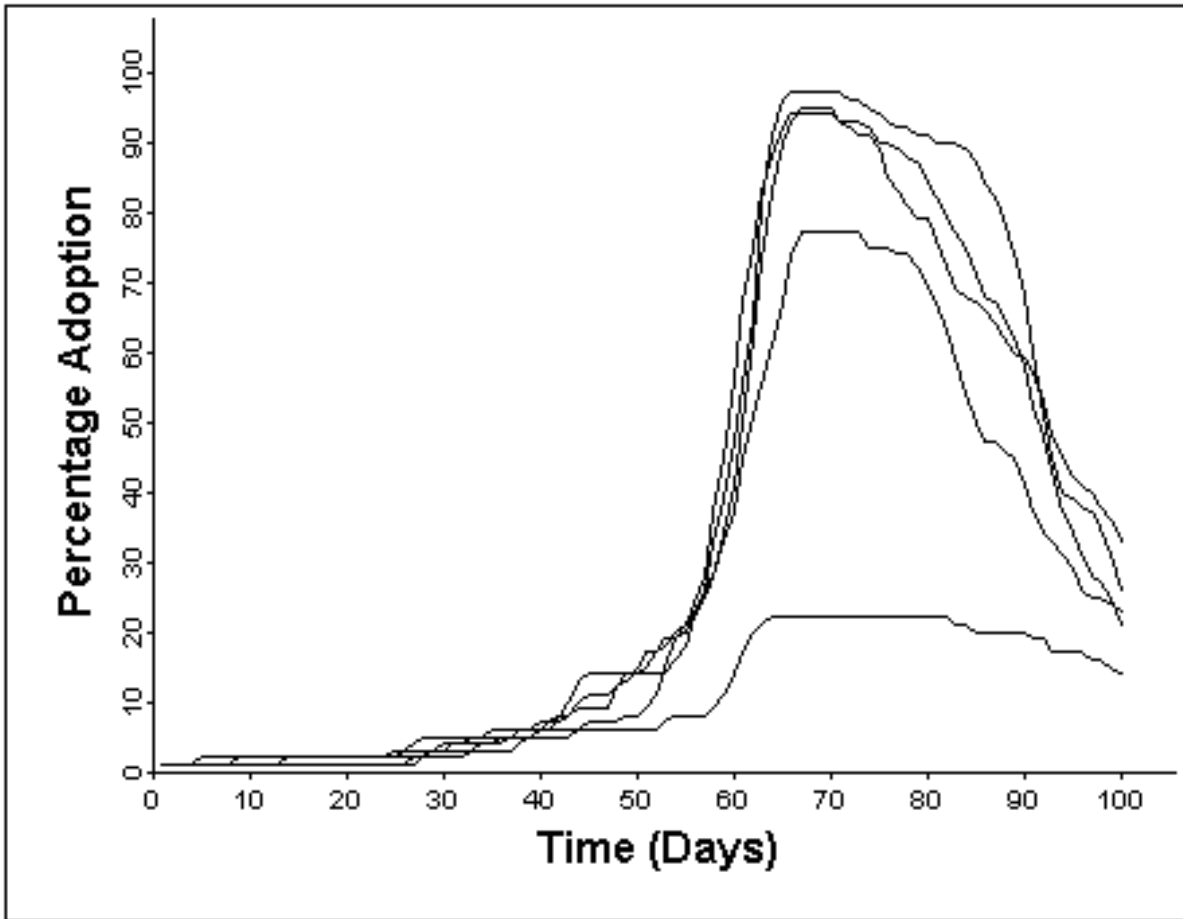
Figure 6. A scenario with a high degree of variability due to stochastic processes in relation to thresholds. Use of averages at a fixed set of Y-values (Fig. 5) to create the envelope may not be the best approach for all stochastic simulations.

If different combinations of parameters give similar envelopes and equivalent fit to observations, the approach provides some guidance as to the value of obtaining further information to refine the fit. However, some combinations of parameters may have multiple effects. For example, changing time weight and use weight in the present study not only moves the graph it also affects the steepness, so guidance should be provided so the user does not get caught in an endless cycle of changes.

Table 2 illustrates the combination of multiple possible approaches in one rule; e.g. "Increasing the Relative Advantage, Compatibility, Observability or Trialability towards 1.0 or decreasing the Complexity towards 0.0 will move the model to the left." The appropriateness of combining approaches within one rule will be system specific. Table 2 also illustrates use of possibilities not in the model (e.g. novelty effect). It is impossible to include all aspects of the real world in a model, but this approach at least makes it feasible to provide advice on the possibility to the model user.

It is not only possibilities not in the model that can be addressed, but also possibilities in the observations. The present system allows only a single peak, with no inflections following it, otherwise no diagnosis is provided. However, such a case could arise in the Innovation Diffusion situation as a result of changes in the innovation properties causing a renewed cycle of adoption, or else the population has changed in a manner beyond that simulated. In addition, the observations could be based on samples, not absolute values, requiring further enhancement of the advice provided.

Finally, adding expert system guidance to parameter selection in simulation modelling appears to have potential for increasing system adoption. The analysis and diagnosis code developed in this study can be used with any simulation model, although, as discussed above, a number of issues remain to be resolved. Even if the diagnostic code is not actually added to the model, consideration of the issues related to graph interpretation raised in this study forces system developers to address the problems faced by users in interpreting output and deciding how to proceed.

About the Author

Alan Thomson is a Senior Research Scientist at Pacific Forestry Centre, Canadian Forest Service, Victoria, British Columbia. He holds a B.Sc. in Zoology from Glasgow University and a Ph.D. in Ecology from McMaster University. His principal area of research is the development of multi-disciplinary, multi-agency knowledge-based systems and decision support systems for natural resource management. In recent years, this has included the use of Artificial Intelligence approaches, customized document generation, and web-based system deployment. His interest in innovation diffusion and adoption stems from his experiences with acceptance and use of these systems. He serves on the editorial board of the journal Computers and Electronics in Agriculture.

References

Doyle, D., A. Tsymbal and P. Cunningham. 2003. *A Review of Explanation and Explanation in Case-Based Reasoning.* Technical Report TCD-CS-2003-41, Department of Computer Science, Trinity College, Dublin. Available at https://www.cs.tcd.ie/publications/tech-reports/reports.03/TCD-CS-2003-41.pdf. Accessed on February 28, 2007.

Greer, J.E., S. Falk, K.J. Greer and M.J. Bentman. 1994. "Explaining and justifying recommendations in an agriculture decision support system". *Computers and Electronics in Agriculture* 11:195–214.

Jenkins, R. and Chapman, R. 1998. "The process of adoption". In: *Proceedings of 1998 Winter Simulation Conference,* December 13-16, 1998, Washington DC, USA. ACM, 1998: 1547-1554.

Lynch, T. and S. Gregor. 2003. "Technology-push or user-pull? The slow death of the transfer-of-technology approach to intelligent support systems development". In: S. Clarke, E. Coakes, M. G. Hunter, and A. Wenn (eds.), *Socio-Technical and Human Cognition Elements of Information Systems.* Information Science Publishing, Hershey, 158-180.

Newman, S., T. Lynch and A.A. Plummer. 2000. "Success and Failure of Decision Support Systems: Learning as We Go". *Proceedings of the American Society of Animal Science*, 12 pp. Available at http://www.asas.org/jas/symposia/proceedings/0936.pdf. Accessed on February 28, 2007.

Rocke, D.M. and D.L. Woodruff. 1997. "Robust estimation of multivariate location and shape". *Journal of Statistical Planning and Inference*, 91:245-255.

Rogers, E,M. 1995. *Diffusion of Innovations*. Fourth Edition. The Free Press, New York.

Smith, G. and J. Guerrero. 1994. "Potential for application of current models for the improvement of helminth control: advantages, limitations, shortcomings". In: B.D. Perry and J.W. Hansen (eds), *Modelling Vector-Borne and other Parasitic Diseases*. Irland, Nairobi, Kenya, 1994, 309-323. Available at http://www.fao.org/wairdocs/ilri/x5441e/x5441e0b.htm. Accessed on February 28, 2007.

Thomson, A., M. Haggith and R. Prabhu. 2004. "Innovation diffusion: predicting success of system development". *Proc. 15th International Workshop on Database and Expert Systems*, Zaragoza, Spain, 30 Aug-3 Sept 2004: 627-631. IEEE Computer Society.

Thomson, A.J., H.M. Rauscher, D.L. Schmoldt, and H. Vacik. 2007. "Information and Knowledge Management for Sustainable Forestry". In: Reynolds, K., Thomson, A., Köhl, M., Shannon, S., Ray, D., and Rennolls, K. (eds.) *Sustainable Forestry: From Monitoring and Modelling to Knowledge Management and Policy Science*. CABI Publishing, Wallingford, UK.

Thomson, S.J., and B.B. Ross, 1996. "Model-based irrigation management using a dynamic parameter adjustment method". *Computers and Electronics in Agriculture*, 14: 269-290.

Appendix: Equations used in the innovation diffusion model.

1) Fitness: "fitness" was defined as "an assumed property of a system that determines the probability that that system will be selected" (the purpose of the model was to explore system adoption), and combines four of the five attributes of an innovation into a single measure. Weightings could be applied to the different components.

fitness = (advantage x compatibility x (1.0-complexity) x trialability)

2) Exposure: the "exposure" variable integrates time since introduction of the innovation and numbers of previous adopters to modify the adoption rate. Exposure was determined as:

exposure =
$$\frac{[(time\_wt * exposure\_time / time\_scale) + (use\_wt * last(pct\_use)/100.0)]}{(time\_w\ t + use\_wt)}$$

where *time_wt* and *use_wt* define the relative importance of exposure duration (*exposure_time*), and exposure to prior adopters, expressed as percent use (*pct_use*).