Inductive Revision of Quantitative Process Models

Keywords: Scientific discovery

Abstract

Most research on computational scientific discovery has focused on developing an initial model, but an equally important task involves revising a model in response to new data. In this paper, we present an approach that represents candidate models as sets of quantitative processes and that treats revision as search through a model space which is guided by time-series observations and constrained by background knowledge cast as generic processes. We demonstrate our system's ability on two different scientific domains and associated data sets with encouraging results. We also discuss its relation to other work on model revision and consider directions for additional research.

1. Introduction and Motivation

Most research on computational scientific discovery (e.g., Langley, 2000) has emphasized the generation of entirely new models to describe or explain data. However, human scientists spend much of their time not developing new accounts of phenomena but rather revising and improving an existing model that already has credibility. If we desire computational tools that help scientists in practice, we should examine seriously the issues that arise in the revision of scientific models.

There are other excellent reasons for focusing on model revision rather than generation. Most approaches to scientific discovery carry out search through the space of hypotheses or models, which can be quite large. By starting from an existing candidate, we can reduce the effective size of this space, making some tasks tractable that would otherwise be too difficult. This approach also places constraints on the search mechanism that can reduce variance and thus decrease chances of overfitting the data. Finally, model revision gives the domain user more control over the region of the model space explored, and increases the chances that he will find the new model comprehensible.

In this paper, we report one approach to the problem of scientific model revision. Unlike earlier work on this topic, we focus on the modification of quantitative process models, a representation of knowledge that we believe offers additional advantages related to search and interpretability. We review this formalism briefly in the section that follows, along with IPM/R, an algorithm that revises an initial process in response to time-series data. After this, we report experimental results on two environmental domains, one involving water behavior in a Danish fjord and the other concerning phytoplankton growth in an Antarctic sea. In closing, we discuss related work on model revision and outline directions for future research.

2. Process Models and their Revision

In this section, we review our previous work on process models and their induction, after which we present our approach to revising models within this framework.

2.1. Quantitative Process Models

Scientific models are often stated formally as sets of equations, but they are often described informally in terms of the processes that determine those equations. We have developed the formalism of *quantitative process models* to encode both aspects of scientific knowledge. In this framework, a model consists of a set of processes, each of which specifies one or more equations that represent causal relations among variables. These can take the form of algebraic equations for instantaneous effects or differential equations for changes over time. Processes can include threshold conditions on variables that characterize when they are active.

A process model specifies not only a set of processes, but also the variables which they connect. A given variable may be labeled as observable, meaning it is present in the data, or it may play the role of a theoretical term that serves mainly to link processes. Each variable may also be labeled as exogenous, in that it

Preliminary work. Under review by the International Conference on Machine Learning (ICML). Do not distribute.

Table 1. A process model for the Ross Sea ecosystem.

```
model Ross_Sea_Ecosystem;
variables phyto, residue, nitrate, light, growth_rate,
          nitrate_rate, light_rate, n_to_c_ratio, r_max;
observable phyto, nitrate, light;
exogenous light;
process phyto_loss;
  equations d[phyto,t,1] = -0.1 * phyto;
            d[residue,t,1] = 0.1 * phyto;
process phyto_growth;
  equations d[phyto,t,1] = growth_rate * phyto;
process phyto_absorbtion_nitrate;
  equations
   d[nitrate,t,1] = −1 * n_to_c_ratio * growth_rate * phyto;
process growth_limitation;
  equations
     growth_rate = r_max * min(nitrate_rate, light_rate);
process nitrate_availability;
  equations nitrate_rate = nitrate/(nitrate + 5);
process light_availability;
  equations light_rate = light/(light + 50);
process global_parameters;
  equations n_to_c_ratio = 0.204;
            r_max = 0.23;
```

influences other variables but is not influenced in return, or as endogenous, which means it appears in the left side of one or more equations. Our framework is a quantitative variant of Forbus' (1984) qualitative process theory, from which we have borrowed many ideas.

Table 1 shows a process model for the acquatic ecosystem of the Ross Sea in Antartica, which is based on an earlier differential equation model developed by Earth scientists (Arrigo et al., 2003). The model includes three observables – the exogenous variable light and the endogenous concentrations of phytoplankton (phyto) and nitrate. The first process characterizes loss of phyto due to unknown sources, along with an increase in residue, whereas the second specifies growth of phyto as a function of its current concentration and its growth_rate. The third process concerns the decrease in nitrate associated with its absorption by phytoplankton. Next, growth_limitation indicates that growth_rate is the product of the maximum rate r_max and the minimum of two theoretical terms, nitrate_rate and light_rate, which are functions of nitrate and light. A final process specifies two parameters that occur across processes.

We can utilize a process model of this sort, together with initial values, to simulate the model's behavior over time and thus predict values for each endogenous variable. We have implemented a simulator that determines which processes are active on each time step

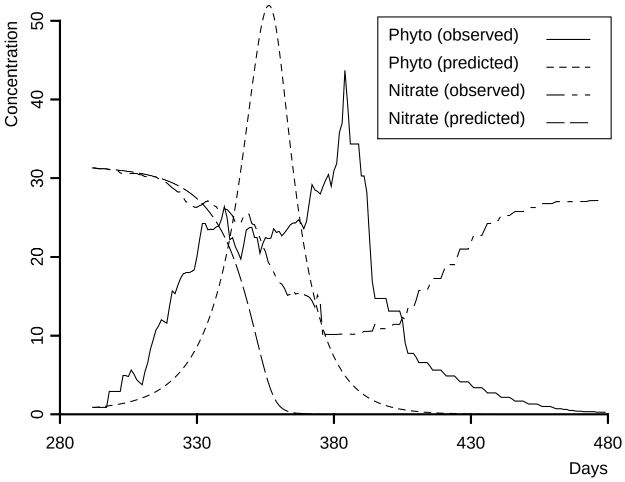

Figure 1. Observed phytoplankton and nitrate concentrations in the Ross Sea, along with concentrations predicted by the process model in Table 1.

and then invokes standard methods for solving ordinary differential equations to generate predicted trajectories. Figure 1 plots observations for phyto and nitrate, collected from ocean samples, over a period of six months, along with the trajectories the model predicts during the same period from the initial values.

Process models of this sort provide an explanation of observations in that they offer a causal account in terms of processes and equations that are familiar to domain specialists. For example, Table 2 presents some generic processes relevant to acquatic ecosystems that serve as background knowledge. These differ from *specific* processes in that they do not commit to particular variables or parameter values. However, they can indicate constraints, such as stating that the variable P in phyto_loss must have type p_species and that its coefficient alpha must fall between 0 and 1. The table also specifies that the same parameter must appear in multiple equations within some processes.

In a previous paper, we proposed the task of inducing process models like the one in Table 1 from time-series data and from background knowledge like the generic processes in Table 2. We noted that this task differs from those typically studied in machine learning, in that process models characterize the behavior of dynamical systems with continuous variables that change over time, and thus are not independently and identically distributed. Moreover, such models are explanatory in nature, in that processes themselves are not observable, processes can interact to produce complex behavior, and process models can include theoretical variables that are also unobservable. However, these complicating factors are offset by the assumption that the dynamical systems are deterministic (although the

Table 2. Some generic processes for aquatic ecosystems with constraints on their variables and parameters.

```
generic process zoo_grazes_phyto;
  variables P{p_species}, Z{z_species},
            R{residue}, G{gz_rate};
  parameters gamma[0, 1];
  equations d[P,t,1] = −1.0 * G * Z;
            d[R,t,1] = gamma * G * Z;
            d[Z,t,1] = (1 − gamma) * G * Z;
generic process Ivlev_rate;
  variables G{gz_rate}, P{p_species};
  parameters delta[0, 10], rho[0, 10];
  equations G = rho * (1 − exp(−1 * delta * P));
generic process residue_loss_to_remineralization;
  variables RES{residue}, REM{r_rate};
  equations d[RES,t,1] = −1 * REM * RES;
generic process nitrate_remineralization;
  variables N{n_nutrient}, REM{r_rate},
            RES{residue}, NtoC{n_const};
  equations d[N,t,1] = REM * NtoC * RES;
generic process phyto_growth;
  variables P{p_species}, GR{gw_rate};
  equations d[P,t,1] = GR * P;
generic process phyto_absorbtion_nitrate;
  variables N{n_nutrient}, GR{gw_rate},
            P{p_species}, NtoC{n_const};
  equations d[N,t,1] = −1 * NtoC * GR * P;
generic process growth_limitation;
  variables GR{gw_rate}, RMAX{r_const},
            NR{n_rate}, LR{l_rate};
  equations GR = RMAX * min(NR, LR);
```

observations may contain noise), since scientists often make this assumption. We have also made arguments, but will not repeat here, that existing methods for machine learning and knowledge discovery do not solve the task of inductive process modeling.

In response, we developed an initial algorithm, called IPM, that addresses this task. The system carries out constrained search through the space of process models, first instantiating the known generic processes with specific variables, then combining subsets of these instantiated processes into generic models that specify candidate explanatory structures. For each such structure, IPM utilizes a gradient descent search through the space of parameter values that attempts to minimize the mean squared error over observed variables. The system then selects the parametrized model that produces the best score on an evaluation criterion that incorporates both error and model complexity.

Experiments with IPM demonstrated its ability to induce process models that involved both conditions and theoretical terms, and also that it had some robustness in the presence of noise. We obtained encouraging results on real-world data collected from batteries on the Space Station, but our studies with environmental models, which had motivated our research on process model induction, dealt only with synthetic data and involved a target model with only five processes. The availability of both new environmental data sets and codification of additional ecosystem processes has encouraged us to extend the IPM framework and evaluate its behavior in this new context.

2.2. The IPM/R Revision Algorithm

Although IPM produced encouraging results, it had drawbacks that limited its applicability. The system constrained its search space by utilizing background knowledge about generic processes, but the space of models could still be large. Also, IPM provided no way to guide the search toward models a scientist might find more plausible. As argued earlier, model revision seems an appropriate response to both issues, so we developed an extended system, IPM/R, that adopts this approach to process model induction.

The IPM/R algorithm requires the user to specify four inputs. These include: an initial model that encodes the user's beliefs about the processes that are most likely involved; a set of allowed changes that specify which of the initial processes can be removed or have their parameters changed; a set of generic processes that may be added to the initial model; and observations to which the revised model model should be fit. The initial model constitutes the user's best guess about the processes that are present in the system, whereas the allowed changes indicate his areas of greatest uncertainty. These criteria, together with candidate processes for addition, provide IPM/R with a heuristic that guides search toward parts of model space that are consistent with domain knowledge.

As output, the algorithm generates a set of revised models that are sorted by their distance from the initial model and presented with their mean squared error on the training data. The distance between a revised model and the initial model is defined as the number of processes that are present in one but not in the other. This output format lets one observe the tradeoff between performance of revised models and their similarity to the initial model. This leaves the user to determine the best compromise between the factors and to select a model from the suggested set.

IPM/R's search mechanism operates in two main stages. The first involves finding all model structures that are consistent with the specified constraints. Each candidate model structure includes a subset of processes from the initial model and instantiated pro-

Table 3. Initial process model for Ringkobing Fjord.

```
model Fjord_Height_Dynamics;
variables A, H, hsea, n, Q, Wdir, Wvel, WaterFlow;
observable A, H, hsea, n,Q, Wdir, Wvel;
exogenous A, hsea, n, Q, Wdir, Wvel;
process waterFlowThroughGates;
 equations WaterFlow = −1000000 * 0.5 * n * (H − hsea);
process freshWaterInput;
  equations WaterFlow = Q;
process flowHeightRelation;
  equations d[H,t,1] = WaterFlow / A;
process windForcingSine;
  equations d[H,t,1] = 0.05 * Wvel * sin(Wdir);
```

Table 4. Generic processes for Ringkobing Fjord.

```
generic process gate_water_flow;
  variables F{flow}, H{level}, HSEA{level}, N{integer};
  parameters k [0,1];
  equations F = −1000000 * k * N * (H − HSEA);
generic process fresh_water_flow;
  variables F{flow}, Q{flow};
  equations F = Q;
generic process flow_height_relation;
  variables F{flow}, H{level}, A{surface};
  equations d[H,t,1] = F / A;
generic process wind_forcing_sine;
  variables WVEL{speed}, WDIR{direction}, H{level};
  parameters c [-0.1,0.1];
  equations d[H,t,1] = c * WVEL * sin(WDIR);
generic process wind_forcing_cosine;
  variables WVEL{speed}, WDIR{direction}, H{level};
  parameters b [-0.01,0.01];
  equations d[H,t,1] = b * WVEL * cos(WDIR);
generic process wind_forcing_cosine_squared;
  variables WVEL{speed}, WDIR{direction}, H{level};
  parameters b [-0.01,0.01];
  equations d[H,t,1] = b * (WVEL * cos(WDIR)) ∧ 2;
generic process wind_forcing_simple_damping;
  variables H{level};
  parameters hbar [0,1], k [0,1];
  equations d[H,t,1] = -k * (H − hbar);
generic process wind_forcing_proportional_damping;
  variables H{level}, WVEL{speed}, WDIR{direction};
  parameters hbar [0,1], k [−1,1];
  equations
     d[H,t,1] = k * (H − hbar) * WVEL * cos(WDIR);
```

cesses from the addition set. The constraints consist of the changes allowed by the user and conventions of the process modeling framework, such as requirements that models form connected graphs and include all observable variables. The second stage focuses on each model structure to determine parameter values for new processes and to revise values for processes allowed to change. To this end, IPM/R utilizes gradient descent, combined with a randomized hill-climbing scheme, to attempt to find parameter values that minimize the mean squared error.

3. Experimental Evaluation

We decided to evaluate IPM/R by applying it to two actual modeling problems: predicting the height of water within a Danish fjord and modeling phytoplankton dynamics in the Antarctic Ocean. Because we were interested in the advantages of revising models over generating them, we compared the revisions suggested by IPM/R to the best results returned by the original IPM, which induces a process model from scratch without an initial model as its starting point. Naturally, we were also interested in whether the revised models constituted improvements over the initial ones.

3.1. Water Dynamics in Ringkobing Fjord

The Ringkobing Fjord, on the Danish west coast, is a shallow body of water that is fed by tributaries from the land and that exchanges water with the North Sea through a narrow channel on its western edge. A barrier with 14 gates has been constructed across this channel in order to regulate water flow between the fjord and the sea. Officials would like to predict in advance the water level at the gates, so they can be ready to open or close them as needed.

When the gates are closed, the dominant influences on the water level H inside the gates to the fjord are the wind direction Wdir and wind speed Wvel, as well as the inflow Q of fresh water per unit time. When the gates between the fjord and the sea are open, the difference between H and sea level (hsea) is also an important factor. All of these influences are incorporated in the initial model presented in Table 3.

However, domain experts are uncertain of some model components, such as the wind forcing function, so we provided our system with four other generic processes shown in Table 4. One incorporates the cosine of wind direction in an effort to capture influences of the wind's other components, whereas another (rather arbitrarily) takes the cosine squared. Two other processes reflect the idea that, as wind pushes water to one side of the fjord, gravitational potential energy builds up in the water and causes it to resist.

We have access to almost one year's observations of the fjord, sampled every five hours, for all of the veriables in the initial process model. To evaluate IPM/R in this domain, we treated the first 1322 observations as a training set and used the remaining 330 data points

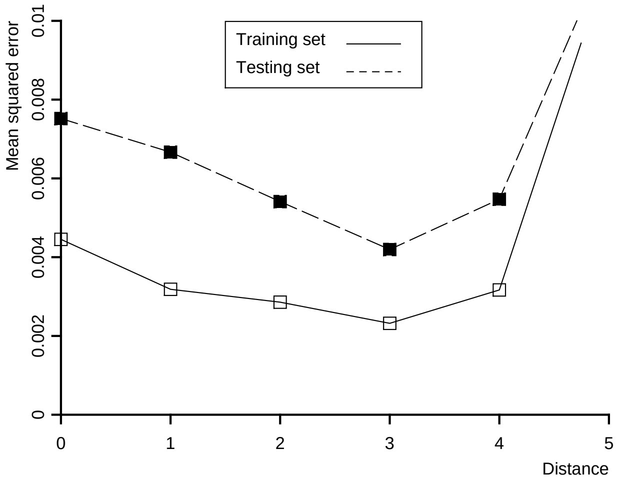

Figure 2. Errors for best IPM/R models for Ringkobing Fjord as a function of distance from initial model.

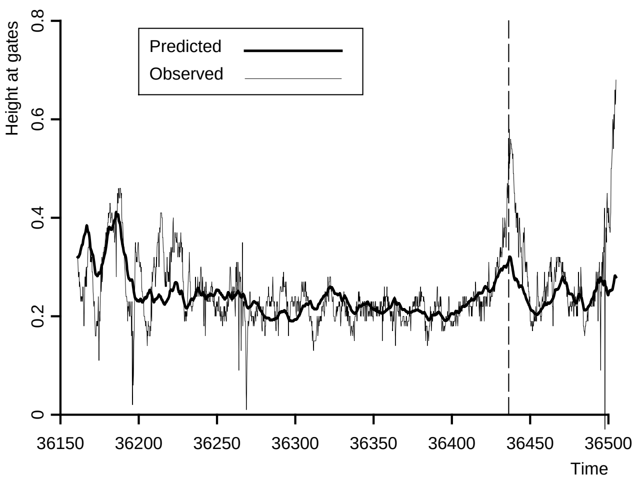

Figure 3. Observed water height in Ringkobing Fjord over a year, along with heights predicted by the revised model.

as a test set to measure generalization error. We provided the initial model, the generic processes, and the training data to IPM/R, which searched the revision space and returned the best model at each distance from the initial one. The system considered 32 distinct model structures, which took eight hours on a Linux PC with a 1.5 Ghz Pentium 4 processor. We used 200 random restarts in the parameter-fitting stage.

Figure 2 summarizes the results by plotting the training and test error against the distance from the initial model for the revisions suggested by IPM/R. We measured this distance as the symmetric difference between the processes in the initial and revised models. The U shapes reveal that error decreases initially as one moves away from the initial model, but after a distance of three the errors increase again. This is not surprising if the 'true' model falls at an intermediate distance from the initial one. More interesting is that the curves for training and test set error are similar, which suggests that our model-revision method is not overfitting the training data, at least in this domain.

Inspection suggests that the revisions IPM/R proposes are physically reasonable. The best model at distance one simply has the process wind_forcing_sine removed, which means that north-south component has no effect on height. At distance two, the best model adds the process wind_forcing_simple_damping, which incorporates resistance to the wind force. The next revision adds wind_forcing_proportional_damping, which introduces an east-west component into the model.

The best revised model's predictions for water height are shown in Figure 3, with points to the left of the vertical line denoting the training set and those to its right the test set. As the figure shows, IPM/R tracks the qualitative behavior of the fjord dynamics in a reasonable way. These results indicate that the system can revise an initial model in ways that extrapolate effectively to new observations.

3.2. Population Dynamics in the Ross Sea

The Ross Sea in Antarctica has been the focus of many studies (Arrigo et al., 2003) because it has a relatively simple food web in comparison to open ocean systems, and thus provides a tractable starting point for modeling efforts. The most important organism in this ecosystem is phytoplankton, which undergoes repeated cycles of population increase and decrease.

Together the Earth scientist on our team, we developed the initial process model shown in Table 1, which relates resources such as light and nutrients to phytoplankton growth. Although the model encodes much existing knowledge about how the variables interact, uncertainty about several components led us to consider alternative generic processes like those in Table 2. These include mechanisms for zooplankton grazing on phytoplankton, nitrate remineralization, and residue loss. Because zooplankton was not measured, IPM/R treats it as an unobservable variable.

We have daily measurements of phytoplankton and nitrate concentrations in the Ross Sea, as well as light levels, for 188 days. However, in this study, we did not divide the observations into training and test sets because we currently have data from only one yearly cycle. Earth scientists believe that different regimes are active during the year, making it necessary to have data from all parts to infer an accurate model. For ex-

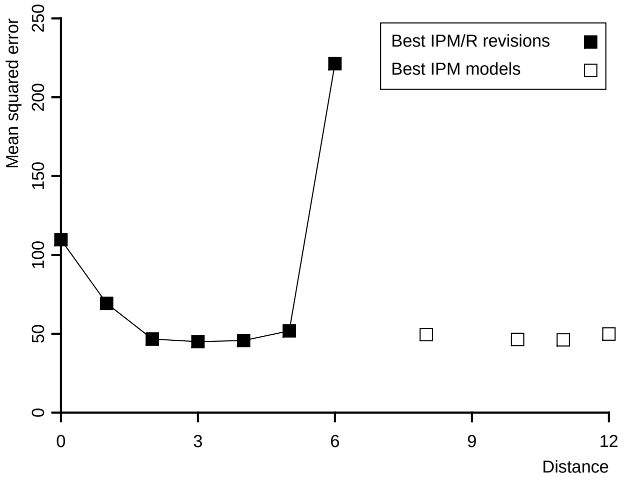

Figure 4. Errors for best IPM/R and IPM induced models for Ross Sea as a function of distance from initial model.

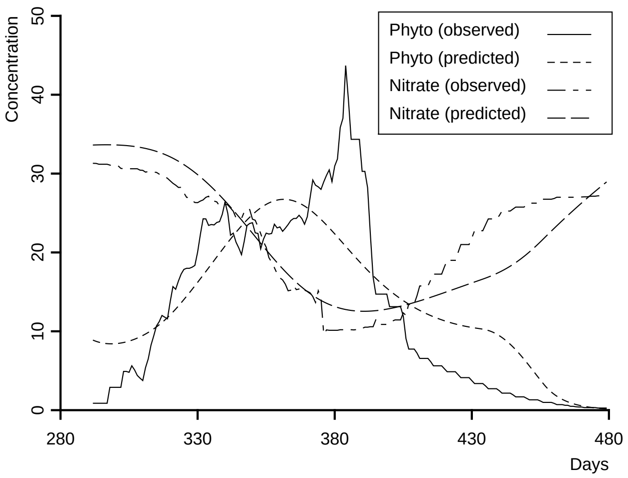

Figure 5. Observed phytoplankton and nitrate concentrations, along with trajectories predicted by revised model.

ample, phytoplankton would initially experience unlimited growth, followed by nutrient limitation and then light limitation. Splitting the data would risk testing on different regimes from the training set, so we report instead resubstitution error. Such training set error can be biased downward, but this is mitigated by the fact that the models are highly constrained.

We believe obtaining error estimates from a dynamical system that experiences multiple regimes is an open problem. Block cross validation with many folds is not an adequate solution. Many folds implies only a few items in each fold, which makes it likely that every regime will be present in both the training and testing sets, but it fails to measure two important properties. First, small block sizes means that predictive error can only be measured for short time periods, whereas long-term predictions hold much greater interest. Second, small blocks make it unlikely that the model will be tested as it moves through different regimes. In the future, we plan to investigate the bias of resubstitution error in the process modeling framework by obtaining independent cycles of data from other years.

As before, we ran both IPM/R and IPM on the Ross Sea data. For this task, IPM/R took 15 hours to explore 84 model revisions, whereas the IPM algorithm took over 25 hours to examine 516 different models. In this case, we used 300 random restarts during the parameter-fitting stage. Figure 3 presents the main results by plotting the training error against distance from the initial model for the revisions suggested by IPM/R and the four best models returned by IPM. The line's shape indicates that IPM/R finds models that reduce error substantially with only a few revisions, with error eventually rising with greater changes. In contrast, IPM produces models with comparable error, but these are very different from the initial model.

The training set error for IPM/R has a U shape, as the error is minimized by revised models that are neither very close nor too far from the initial model. This is a somewhat surprising behaviour, as normally one would expect that more complicated models should have lower training set error. We believe IPM/R exhibits this behavior because process models are not arbitrary function approximators. Rather, they are highly constrained by the allowed generic processes and the restrictions on parameter values. In IPM/R, adding processes does not guarantee training error will decrease, as it does with most modeling formalisms.

Detailed inspection revealed that IPM/R's revisions were ecologically plausible. The best model at distance one added the process nitrate_remineralization, which involves restoring nitrate ions into the water from the residue of dead phytoplankton. The best model at distance two added the process residue_loss_to_remineralization, which is a conservation term to balance mass transfer from residue to nitrate ions. In contrast, the best model found by IPM dramatically changed the structure of the model. It posited the presence of zooplankton grazing on the phytoplankton and removed all nutrient consumption processes. The levels of zooplankton are known to be low in the Ross Sea, but a grazing effect is possible. However, the removal of the nutrient consumption by phytoplankton is not ecologically plausible.

Figure 5 shows the actual predictions of the best revised model, along with the original observations.

Compared with the predictions made by the initial model in Figure 1, it is clear that the revised model does a much better job of tracking the trajectories of both variables. The results from our experiment show that, by using IPM/R to revise an initial model instead of learning from scratch, one can generate a small set of revisions that improve the model greatly and yet remain consistent with domain knowledge.

4. Related and Future Work

Computational methods for model revision are certainly not a new idea. Early research (e.g., Towell, 1990) focused on supervised learning for classification tasks, but it supported modification of models with theoretical terms and offered a general framework from which we have borrowed. For instance, Ourston and Mooney's (1990) Either utilized search operators for adding and deleting rules, which correspond to our operators for adding and removing processes. However, they also included ones for adding and removing conditions on rules, which are quite different from our scheme for parameter revision. Still, their view of model revision as searching from an initial model, guided by fit to training data, is a general one that is relevant to revision of quantitative process models.

A different tradition has explored methods for revising qualitative causal models of scientific phenomena. Early systems altered thir models incrementally, but later work has utilized nonincremental methods with simpler representations of causal connections. Recent examples include Bay et al.'s (2002) method for revising qualitative models of gene regulation, which carries out greedy search guided by candidate models' fits to the data. Bryant et al. (2001) report a different approach that uses abductive logic programming to extend a qualitative model of metabolic control.

A parallel line of research has addressed the revision of quantitative models. Chown and Dietterich (2000) dealt with improvement of the parameters in a complex ecosystem model, whereas Glymour et al.'s (1987) Tetrad system revised the structure of models cast as sets of linear equations. Other work has focused on revising the structure and parameters of nonlinear models, often drawing on knowledge about the domain to constrain search (e.g., Bay et al., 2002; Todorovski & Džeroski, 2001). These two efforts focused on differential equation models for time-series data, and thus come closest to the work reported here. The main difference is that our approach introduces the notion of processes, which provide a useful framework for encoding domain knowledge that constrains search and produces more interpretable results.

Although most research has emphasized automated revision methods, a few groups have developed interactive systems. For instance, Mitchell et al.'s (1997) program encourages metallurgists to actively direct its search for quantitative relations, whereas Mahidadia and Compton (2001) report an environment for the interactive revision of qualitative causal models in neuroendocrinology. We believe that most scientists will prefer computational discovery tools that keep them involved in the revision effort, and we intend to embed future versions of IPM/R in such an interactive environment for scientific model development.

Naturally, we also intend to extend our revision algorithm itself on various fronts. One drawback of the current system is that the parameter estimation takes 99.9 percent of the computation time, which limits the number of models with different structure it can consider effectively. Preliminary studies with a hierarchical method for multiple shooting (Horbelt et al., 2001) have shown promise for speeding this component. More rapid techniques for parameter fitting should also let us augment IPM/R to consider new conditions on processes during the revision effort, thus increasing its flexibility.

We should also improve the system's ability to evaluate alternative model structures. The current implementation attempts only to minimize error on the training data; this can produce overfitting, although we believe this effect is mitigated by the process modeling formalism and our focus on model revision. However, some method for variance reduction would seem appropriate. Our earlier results on model generation incorporated a variant of minimum description length that preferred models with simpler structures. We intend to modify this for model revision by incorporating a bias toward candidates with structures similar to that of the initial model, much as in Saito et al.'s (2002) work on revision of qualitative models.

Finally, IPM/R's current reliance on exhaustive search through the space of model structures lets it handle only relatively small models. Future versions should replace this approach with a heuristic method that scales reasonably to more complex models and to more extensive changes. Ideally, the generation of model structures should be more closely linked to parameter estimation, thus making both more efficient. We should also extend the framework to support revision of models with subsystems, which ecosystem modelers often utilize when dealing with complex domains. This would reduce effective search by drawing on knowledge about likely subsystems rather than isolated processes, and thus further improve scaling ability.

5. Concluding Remarks

In this paper, we reported an approach to computational scientific discovery that focuses on the revision of existing models rather than on their construction. Unlike earlier work in this area, we utilized the formalism of quantitative process models, which support explanatory accounts of continuous time series in terms of unobservable variables and processes. Our specific algorithm, IPM/R, lets users specify an initial model, a data set, and a set of allowed revisions. The latter can include specific processes that may be deleted, specific processes for which the parameters may be altered, and generic processes that may be added. We demonstrated this system's abilities in two environmental domains, one involving changes in water height in a Danish fjord and another concerning population dynamics in the Ross Sea.

Our experimental results were generally encouraging. In both domains, we showed that IPM/R found meaningful revisions that had lower error than the original model. Moreover, we found that some of these models were just as accurate as those produced by IPM, which composes models entirely from generic processes, but included additional information about distance from the initial model, which is important to domain scientists. At the same time, we identified some limitations of the current system that we should address in future work. In summary, our approach builds on previous research in model revision but extends it in ways that make it useful for domains like Earth science, which focus on quantitative models of dynamical systems.

References

Arrigo, K. R., Worthen, D. L. & Robinson, D. H. (2003). A coupled ocean-ecosystem model of the Ross Sea: 2. Iron regulation of phytoplankton taxonomic variability and primary production. *Journal of Geophysical Research*, *108*, C7, 3231, 10.1029/2001JC000856.

Bay, S. D., Shapiro, D. G., & Langley, P. (2002). Revising engineering models: Combining computational discovery with knowledge. *Proceedings of the Thirteenth European Conference on Machine Learning* (pp. 10–22). Helsinki, Finland.

Bay, S. D., Shrager, J., Pohorille, A., & Langley, P. (2003). Revising regulatory networks: From expression data to linear causal models. *Journal of Biomedical Informatics*, *35*, 289–297.

Bryant, C. H., Muggleton, S. H., Oliver, S. G., Kell, D. B., Reiser, P., & King, R. D. (2001). Combining inductive logic programming, active learning and robotics to discover the function of genes. *Electronic Transactions in Artificial Intelligence*, *5-B1(012)*, 1–36,

Chown, E., & Dietterich, T. G. (2000). A divide and conquer approach to learning from prior knowledge. *Proceedings of the Seventeenth International Conference on Machine Learning* (pp. 143–150). San Francisco: Morgan Kaufmann.

Forbus, K. D. (1984). Qualitative process theory. *Artificial Intelligence*, *24*, 85–168.

Glymour, C., Scheines, R., Spirtes, P., & Kelly, K. (1987). *Discovering causal structure: Artificial intelligence, philosophy of science, and statistical modeling*. San Diego: Academic Press.

Horbelt, W., Voss H. U., & Timmer, J. (2001). Parameter estimation in nonlinear delayed feedback systems from noisy data. *Physics Letters A*, *299*, 513–521.

Langley, P. (2000). The computational support of scientific discovery. *International Journal of Human-Computer Studies*, *53*, 393–410.

Mahidadia, A., & Compton, P. (2001). Assisting model discovery in neuroendocrinology. *Proceedings of the Fourth International Conference on Discovery Science* (pp. 214–227). Washington, D.C.: Springer.

Mitchell, F., Sleeman, D., Duffy, J. A., Ingram, M. D., & Young, R. W. (1997). Optical basicity of metallurgical slags: A new computer-based system for data visualisation and analysis. *Ironmaking and Steelmaking*, *24*, 306–320.

Ourston, D., & Mooney, R. (1990). Changing the rules: A comprehensive approach to theory refinement. *Proceedings of the Eighth National Conference on Artificial Intelligence* (pp. 815–820). Boston: AAAI Press.

Saito, K., Bay, S., & Langley, P. (2002). Revising qualitative models of gene regulation. *Proceedings of the Fifth International Conference on Discovery Science* (pp. 59–70). Lubeck, Germany: Springer.

Todorovski, L., & D̃zeroski, S. (2001). Theory revision in equation discovery. *Proceedings of the Fourth International Conference on Discovery Science* (pp. 389–400). Washington, D.C.: Springer.

Towell, G. (1991). *Symbolic knowledge and neural networks: Insertion, refinement, and extraction.* Doctoral dissertation, Computer Sciences Department, University of Wisconsin, Madison.