Time Series Adjustment Enhancement of Hierarchical Modeling of *Arabidopsis Thaliana* Gene Interactions

Edward E. Allen[1], John Farrell[2], Alexandria F. Harkey[3], David J. John[2(✉)], Gloria Muday[3], James L. Norris[1], and Bo Wu[1]

[1] Mathematics and Statistics, Wake Forest University, Winston-Salem, NC 27109, USA
{allene,norris,wub18}@wfu.edu
[2] Computer Science, Wake Forest University, Winston-Salem, NC, USA
{farrjt17,djj}@wfu.edu
[3] Biology, Wake Forest University, Winston-Salem, NC, USA
{harka14,muday}@wfu.edu

Abstract. Network models of gene interactions, using time course gene transcript abundance data, are computationally created using a genetic algorithm designed to incorporate hierarchical Bayesian methods with time series adjustments. The posterior probabilities of interaction between pairs of genes are based on likelihoods of directed acyclic graphs. This algorithm is applied to transcript abundance data collected from *Arabidopsis thaliana* genes. This study extends the underlying statistical and mathematical theory of the Norris-Patton likelihood by including time series adjustments.

Keywords: Gene interaction network modeling · Bioinformatics · Genetic algorithms · Bayesian methods · Time series adjustment

1 Introduction

Cell signaling is accomplished via networks of transcriptional changes that lead to synthesis of distinct sets of proteins, which cause changes in growth, development, or metabolism. Treatments that elevate levels of hormones result in cascades of changes in gene expression driven by activation and synthesis of transcription factors which are required to turn on downstream genes. One approach to model these gene regulatory networks is to collect measurements of changes in abundance of gene transcripts across a time course. The expression of a gene encoding a transcriptional activator or repressor protein may signal to the next gene to either turn on or turn off downstream genes and their encoded proteins. Thus, time course transcriptomic data sets contain important information about how genes drive these changes in biological networks. Yet genome-wide transcript abundance assays examine tens of thousands of genes so identification of patterns or networks within these large data sets is difficult. It is also

C. Martín-Vide et al. (Eds.): AlCoB 2020, LNBI 12099, pp. 143–154, 2020.
https://doi.org/10.1007/978-3-030-42266-0_11

critical to filter the meaningful transcript changes in these data sets to remove genes whose responses are not above background or that are dissimilar due to biological or technical variation. Yet even though the bioinformatics community has developed statistical methods to filter the data [9], additional approaches are needed to identify the networks and patterns in these large data sets.

An important modern approach to statistical modeling includes Bayesian techniques involving likelihoods and posterior probabilities. Here, we extend our previous work on this problem by incorporating time series adjustments in the computation of Bayesian likelihoods. We apply this method to time course data generated in response to treatments that elevate the levels of the hormone ethylene in *Arabidopsis thaliana*. We take advantage of a previously published genome-wide transcriptional data set [9], subjected to rigorous filtering and from which all the genes predicted to encode transcription factors have been identified. The goal is to predict gene regulatory networks that control time-matched developmental changes.

The results in this paper are novel for several reasons. First, the methods use the hierarchical nature of the data sets. For example, replicate data are not averaged. Rather, the method constructs a model over all of the data that uses each replicate as a source of information. The assumption is that at each level of the hierarchy there are commonalities in the data and parameters. Thus, the replicate data is not independent. Second, the addition of time series adjustment to improve the independence of the model's residuals gives these techniques stronger statistical foundations. Third, the combination of Bayesian model averaging with a cutting edge genetic algorithm provides rigorous estimates of posterior probabilities for edges. These computational modeling algorithms are derived using rigorous mathematical and statistical techniques and are computationally efficient. The models produced are easily understandable.

Many different techniques for modeling non-hierarchical data using gene expression data have been proposed. An excellent recent survey on this subject was given by Emily [4]. There are many techniques for modeling gene and protein networks–with various different properties–available in the literature. Our technique in this paper is a Bayesian regression type method. Variations of Bayesian modeling can be found in [7,11,19]. Other methods that use types of regression include [2,21] which focus on logistic regression techniques, and [22,23] which use Poisson regression. Other approaches to modeling these types of problems include differential equations [1] and Boolean modeling [14].

This Bayesian likelihood computational algorithm incorporates additional important features from earlier versions. Earlier variations included computing posterior probabilities for a single replicate [11] and for multiple replicates with both hierarchical [18] and independent [17] structures. Over the course of this research, the search procedure has changed from Metropolis Hastings to genetic algorithms. Genetic algorithms' execution times are typically polynomial rather than the doubly exponential execution time, in terms of the numbers of time points and genes, of Metropolis Hastings.

This variation also uses a Bayesian version of the *Cross generational elitist selection, Heterogeneous recombination, Cataclysmic mutation* algorithm (CHC) [6]. Genetic algorithms are motivated by the operators of selection, crossover, and mutation. The CHC variation does not allow the crossover of similar parents. Once the population becomes too homogeneous, then a cataclysmic mutation event regenerates the population from the current *most fit* parents. The Bayesian CHC (BCHC) implemented in this paper uses a hierarchical statistical construct (the Norris-Patton Likelihood) as the fitness function.

The hormone ethylene (ACC) is known to activate root growth in *Arabidopsis thaliana* [9]. Transcription factors (TFs) are cellular proteins that bind to DNA to turn genes either *on* (activation) or off (repression). Developmental changes are controlled by these genes. The data set used in this modeling process was the complete set of abundance levels of the twenty-six TFs believed/known to be involved in the activation of the growth of roots at eight time points after treatment with the ethylene precursor ACC [9]. Here, constructing an appropriate network model has potential agricultural applications in that it should lead to more complex understandings of root development.

2 Mathematical and Statistical Preliminaries

Three network modeling paradigms are generally considered in the literature: *cotemporal*, *next state one step* and *next state one and two steps*. A next state one step model predicts the transcript abundance relationships between genes at time j based on the transcript abundance at time $j-1$. In this paper, we will only consider next state one step models; for simplicity, we will refer to next state one step as next state. The time series adjusted (tsa) next state models are an amalgamation of next state modeling with standard time series adjustments [12]. The time series adjustment methodology makes the residuals (i.e., the estimated error terms) more independent.

A *directed graph* $G = (V, E)$ consists of a pair of collections: V a set of vertices (or nodes); and, E a collection of directed edges between pairs of vertices. A cycle is a sequence $v_1, e_1, v_2, e_2, \ldots, v_{n-1}, e_{n-1}, v_n = v_1$ where $v_i \in V$ and $e_j \in E$ is a directed edge from vertex v_{j-1} to vertex v_j. Directed acyclic graphs (DAGs) do not contain cycles. An example of a DAG is given in Fig. 1. In this modeling algorithm, DAGs form the mathematical foundation of our computational approach. The vertices of a DAG represent genes and the directed edges are one-way relationships between pairs of vertices. When there is a directed edge from v_i to v_j, then v_i is a parent of v_j and v_j is a child of v_i.

For any DAG D with vertex set $V = \{v_1, v_2, \cdots, v_n\}$, the vertices can be topologically sorted. This gives a total order $>$ on V such that if v_i is an ancestor of v_j, meaning that there is a directed path from v_i to v_j, then $v_i < v_j$. Without loss of generality, lets assume that $v_i < v_{i+1}$ for $1 \leq i \leq n-1$.

Conditional probability gives that for any two events A and B, the probability

$$P(A \text{ and } B) = P(A)\ P(B|A) = P(B)\ P(A|B).$$

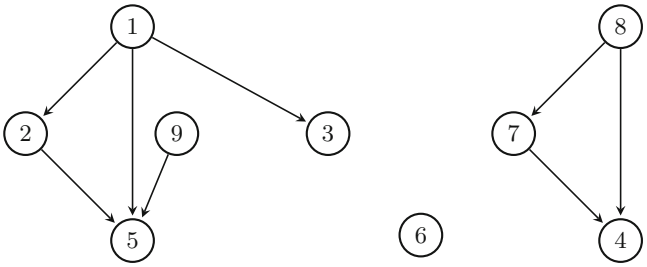


Fig. 1. A directed acyclic graph. Gene 1 affects genes 2, 3 and 5 but not genes 4, 6, 7, 8 and 9. Genes 1, 6 and 8 are not affected by any other gene.

Similarly, the density function f for two continuous variables y_1 and y_2 is

$$f(y_1 \text{ and } y_2) = f(y_1, y_2) = f(y_1)\ f(y_2|y_1) = f(y_2)\ f(y_1|y_2)$$

Recursively, using the order $<$ implied by topologically sorting the DAGs on the set of continuous variable $Y = \{y_1, y_2, \ldots, y_n\}$ (i.e., $y_i < y_j$ if and only if $v_i < v_j$), gives

$$f(y_1, y_2, y_3, \ldots, y_n) = f(y_1)\ f(y_2|y_1)\ f(y_3|y_1, y_2) \ldots f(y_n|y_1, y_2, \ldots, y_{n-1}).$$

Specific for a particular DAG D, let y_1 be the gene that cannot have any parents. Let y_2 be the gene that can have at most parent y_1. Similarly, let y_h be the gene that can have parents from the collection $\{y_1, \cdots, y_{h-1}\}$. Therefore, if we let y_i represent the data of child i for all of the r replicates, we have for D

$$f(y_1, y_2, \ldots, y_n|D) = f(y_1|D) f(y_2|y_1, D) f(y_3|y_1, y_2, D) \cdots f(y_k|y_1, \cdots y_{n-1}, D)$$

3 Time Series Adjustment

Statistical regression models of response (child) data from predictors (parents) data over time nearly always have correlated residuals over time. This is usually due to the remaining influence of the previous time's response data. In complicated modeling situations (e.g., like ours where we need to obtain closed form likelihoods of DAGs within a hierarchical structure in order to produce posterior probabilities of edges), it is common to derive results as if there were non-correlated residuals, as we have done in previous work. Our previous work has shown utility both for simulated and biological data, but we now rigorously incorporate a time series adjustment into our model. This should result in substantially less correlated residuals and thus more accurate likelihoods for the DAGs. Since these likelihoods are the foundations for the edges' estimated posterior probabilities, these estimates should also be improved.

Our time series adjustment is an integer autoregressive adjustment of order 1 in the commonly used family of Markov conditioning. It is a version of Kedem's

and Fokianos' autoregressive model [12, page 184]. In our setting, this simply adds the child's data at the previous time as an additional regressor for the child's data at the current time. Thus, much of the child's data at the previous time's influence would be *regressed out* leaving less correlated, closer to independent, residuals from one time to the next.

4 Next State Time Series Adjustment Computation

For each h, with $1 \leq h \leq n$, $f(y_h|y_1, y_2, \ldots, y_{h-1}, D)$ gives the density of y_h given y_h's parent's data for DAG D. Now, let ${}_iy_c$ be the data vector of any given child c from the i^{th} replicate. The vector ${}_iy_c$ has dimension t, the number utilized time points in the child c data set for a given replicate i. The symbol ${}_ix_c$ is the $t \times k_c$ *regressor* matrix for ${}_iy_c$. For next state with time series adjustment, t is the number of time points per replicate minus one since at time 1, the child data has no last previous parent data nor last previous child (tsa) data–so, the utilized child data starts at time 2. The value of k_c is the number of parents of c plus two since ${}_ix_c$ has a separate column for each of its parent's data at the previous time, a column of 1's for the intercept, and a column of the child's data at the previous time (the time series adjustment). A k_c dimensional slope vector for child c's regressors is ${}_i\beta_c$. The common within replicate residual variance of child c is σ_c^2.

Assumptions which detail the hierarchical structure include that for a given ${}_i\beta_c$ and σ_c^2, each ${}_iy_c$ is independent and normally distributed, and therefore ${}_iy_c|\,{}_i\beta_c\sigma_c^2 \sim N_t({}_ix_c\ \ {}_i\beta_c, \sigma_c^2 I)$. Note the ${}_iy_c$ (child) response and the underlying regression structure of the product of the ${}_ix_c$ matrix and the ${}_i\beta_c$ vector. We have ${}_i\beta_c|\sigma_c^2 \sim N_{k_c}(0, g\sigma_c^2(\bar{x}_c^T\bar{x}_c)^{-1})$ and $\sigma_c^2 \sim$ Inverse-gamma$(v_o/2, v_o\sigma_o^2/2)$. With these assumptions, we have the following result that gives the mathematical and statistical computation of the Norris-Patton likelihood (NPL), $L = f(y_1, \ldots, y_n|D)$, as the value of the density/likelihood function for D.

Theorem 1. *The closed form solution of the likelihood assuming a hierarchical structure among replicates is* $\prod_{c=1}^{n} f(y_c|$*parents of* $y_c, D)$ *where* n *is the number of genes and* D *is a DAG. Also,*

$$f(y_c|\textit{parents of } y_c, D) = (2\pi)^{-\frac{rt}{2}} \left(\frac{1}{2}\right)^{-\frac{(rt+v_0)}{2}} g^{-\frac{rk_c}{2}} \Gamma(v_0\sigma_0^2/2) \frac{\Gamma[(rt+v_0)/2]}{\Gamma(v_0/2)}$$

$$\times \frac{|\ \bar{x}_c^T\bar{x}_c\ |^{\frac{r}{2}}}{\prod_{i=1}^{r} |\ {}_ix_c^T\,{}_ix_c + {}_i\bar{x}_c^T\,{}_i\bar{x}_c\left(\frac{1}{g}\right)\ |^{\frac{1}{2}}}$$

$$\times \left[v_o\sigma_0^2 + \sum_{i=1}^{r} {}_iy_c^T\,{}_iy_c - ({}_ix_c^T\,{}_iy_c)^T[({}_ix_c^T\,{}_ix_c + {}_i\bar{x}^T\,{}_i\bar{x}\frac{1}{g})^{-1}]^T\,{}_ix_c^T\,{}_iy_c\right]^{-\frac{rt+v_0}{2}}$$

The proof of Theorem 1 uses the following lemmas whose computation can be found in [16] (a thesis from our research group). We include the proof of Lemma 2

to show how the computation of the likelihood includes the slope parameters ${}_i\beta_c$ of each of the replicates separately.

Lemma 2. *The contribution of child c to the likelihood to DAG D is*

$$\begin{aligned} f(y_c \mid D) &= f({}_1y_c, \ldots, {}_ry_c \mid D) \\ &= \int_{\sigma_c^2} \left[f({}_1y_c \mid \sigma_c^2) \cdots f({}_ry_c \mid \sigma_c^2) \right] f(\sigma_c^2)\ d\sigma_c^2 \end{aligned}$$

Proof. Using integration, we have

$$\begin{aligned} & f({}_1y_c|D) = f({}_1y_c, \ldots, {}_ry_c \mid D) \\ &= \int_{\sigma_c^2} \int_{r\beta_c} \cdots \int_{1\beta_c} f({}_1y_c, \ldots, {}_ry_c, {}_1\beta_c, \ldots, {}_r\beta_c, \sigma_c^2|D) d\,{}_1\beta_c\ \cdots d\,{}_r\beta_c\ d\sigma_c^2 \\ &= \int_{\sigma_c^2} \int_{r\beta_c} \cdots \int_{1\beta_c} f({}_1y_c, \ldots, {}_ry_c, |\,{}_1\beta_c, \ldots, {}_r\beta_c \sigma_c^2, D) \\ &\times f({}_1\beta_c, \ldots, {}_r\beta_c|\sigma_c^2)\ f(\sigma_c^2)\ d\,{}_1\beta_c\ \cdots d\,{}_r\beta_c\ d\sigma_c^2 \\ &= \int_{\sigma_c^2} \left[f({}_1y_c \mid \sigma_c^2) \cdots f({}_ry_c \mid \sigma_c^2) \right] f(\sigma_c^2)\ d\sigma_c^2 \end{aligned}$$

□

Letting $|M|$ denote the determinant of the matrix M, we have the following:

Lemma 3. *For a given replicate i and letting exp(x) represent the exponential function e^x, we have*

$$\begin{aligned} & f({}_iy_c \mid \sigma_c^2) \\ &= (2\pi\sigma_c^2)^{-\frac{t}{2}} \mid g\sigma_c^2 (\bar{x}_c^T \bar{x}_c)^{-1} \mid^{-\frac{1}{2}} \mid {}_iA_c \mid^{\frac{1}{2}}\ exp\left(-\frac{1}{2} \left[\frac{1}{\sigma_c^2}\, {}_iy_c - {}_im_c^T\, {}_iA_c^{-1}\, {}_im_c \right] \right) \end{aligned}$$

where

$${}_iA_c^{-1} = \frac{1}{\sigma_c^2} \left({}_ix_c^T\, {}_ix_c\ + \bar{x}_c^T \bar{x}_c \left(\frac{1}{g} \right) \right)$$

and

$${}_im_c = \left({}_ix_c^T\, {}_ix_c + \bar{x}_c^T \bar{x}_c \left(\frac{1}{g} \right) \right) {}_ix_c^T\, {}_iy_c.$$

Extending Lemma 2 to the product of density functions used in Lemma 1, we have:

Lemma 4

$$\begin{aligned} & f({}_1y_c \mid \sigma_c^2) f({}_2y_c \mid \sigma_c^2) \cdots f({}_ry_c \mid \sigma_c^2) \\ &= (2\pi)^{-\frac{rt}{2}} (\tau_c)^{\frac{rt}{2}} (g)^{-\frac{rk_c}{2}} \frac{\mid \bar{x}_c^T \bar{x}_c \mid^{\frac{r}{2}}}{\prod_{i=1}^{r} \mid {}_ix_c^T\, {}_ix_c + \bar{x}_c^T \bar{x}_c \frac{1}{g} \mid^{\frac{1}{2}}} \\ &\times\ exp\left(-\frac{1}{2} \tau_c \sum_{i=1}^{r} \left[{}_iy_c^T\, {}_iy_c - ({}_ix_c^T\, {}_iy_c)^T [({}_ix_c^T\, {}_ix_c + \bar{x}_c^T \bar{x}_c \left(\frac{1}{g} \right))^{-1}]^T\, {}_ix_c^T\, {}_iy_c \right] \right) \end{aligned}$$

Note that g, v_0 and σ_c^2 are positive free parameters. In our modeling algorithm, we set $g = v_0 = \sigma_c^2 = 1$. The use of the time series adjusted next state Norris-Patton likelihood, along with a tailor-made genetic algorithm and Bayesian model averaging, allows for the rigorous estimation of posterior probabilities for all gene pair interactions.

```
 1: procedure TBCHC
 2:     t ← 0
 3:     Archive ← {}
 4:     multi-step initialization of 400 DAG(s) for P(0)
 5:     indicator ← 50
 6:     while t < 600 do
 7:         t ← t + 1
 8:         X ← randomly reorderP(t−1)
 9:         Y ← {}
10:         for all parent pairs (X[2i], X[2i+1]) do
11:             if parent pair (X[2i], X[2i+1]) are dissimilar then
12:                 Y ← Y∪ {crossover-repair (X[2i], X[2i+1])}
13:             end if
14:         end for
15:         indicator ← indicator − (|P(t−1)| − |Y|)
16:         P(t) ← NPL fittest |P(t−1)| of P(t−1) ∪ Y
17:         if indicator < 0 then
18:             P(t) ← cataclysm(P(t))
19:             indicator ← 50
20:         end if
21:         Archive ← Archive ∪ P(t)
22:     end while
23:     return Archive
24: end procedure
```

Fig. 2. The TBCHC genetic algorithm searches for and returns an archive of unique DAGs (lines 3, 21 and 23). After applying Bayesian model averaging to the archive, the gene interaction model is formed. The initial population consists of 400 DAGs. The time series adjustment is applied in finding the fittest DAGs (line 16). The variable *indicator* triggers cataclysmic mutation (lines 17–19).

5 Genetic Algorithms

Simply put, a genetic algorithm (GA) takes the current population and produces the next generation using the operations of *selection*, *crossover*, and *mutation* [15]. Individuals (i.e., DAGS) are automatically moved to the next generation with preference given to those with the higher likelihoods (the *elitist* strategy). The first population must be initialized. The genetic algorithm terminates after a specified number of iterations.

The TBCHC genetic algorithm is an extension of BCH [13] which was heavily influenced by the CHC [5]. The TBCHC fitness function includes the next state

time series adjustment. The TBCHC operators of *selection*, *crossover*, *mutation*, and *repair* will be discussed in the following paragraphs.

The population of each generation consists of a fixed number of DAGs. Each DAG represents gene relationships. The genetic algorithm's aim is to move from the current population of DAGs to a new generation where the overall quality improves (as measured by the Norris-Patton likelihood). The elitist strategy only moves the top 10% of DAGs from the current generation to the next and the balance is filled by crossover. As TBCHC iterates, all distinct DAGs are archived. The final gene interaction model is produced from this archived collection.

Generally, the selection operator chooses which members of the current population can potentially contribute children to the next generation. In Fig. 2 selection is accomplished through a random pairing of all parents in the current population (lines 8–10). By assuming prior probabilities for the DAG, the likelihood of a given DAG D is proportional to the D's NPL [3]. Thus, the fitness of a candidate D can be computed using the NPL.

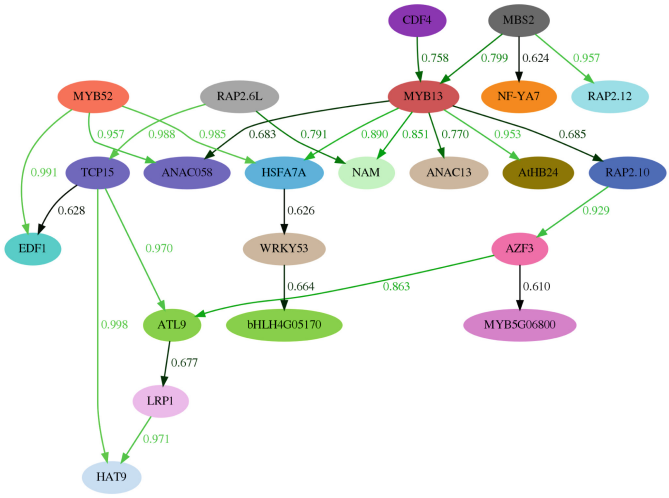
The crossover operator (line 12) exchanges genetic information (i.e., directed edges) between two parents producing two new offspring. The edges chosen to be exchanged are chosen randomly. There is one caveat: if the two parents are too similar–determined by the Hamming distance between them then the two selected parent DAGs are not allowed to produce offspring (line 11). In a simple genetic algorithm, all selected parents are allowed to produce offspring. This TBCHC prohibition of mating by similar parents may result in fewer DAGS in the next population than in the current population. Since the modeling process is based on DAGs, if the crossover operator introduces a cycle in the offspring, a repair operator is applied. Selection and crossover are used exclusively in TBCHC until the population becomes too similar. At that point, cataclysmic mutation (line 17) is applied to reset the population by creating a new population of DAGs from the top 10% NPL DAGs.

There are no known techniques for assigning the optimum values to the genetic algorithm parameters. However, experience and the literature give general criterion for appropriate values. Still, values are often determined on a case by case basis. The TBCHC algorithm parameters include the following: 20 parallel executions each with 600 generations; the number of initial DAGS is 400; the crossover probability is 0.30; and, the number of parents of any given node is limited to 3. Cataclysmic mutation causes the population of DAGs to be replaced by DAGs generated by crossover and mutation on the top 10% of the population to restore the candidate class to 400.

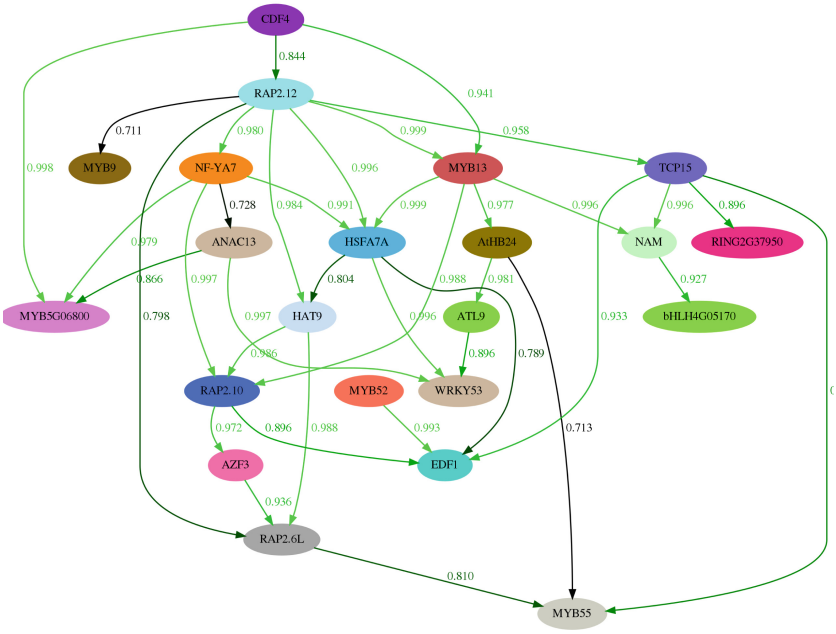
This TBCHC algorithm is implemented in *python 3.0* using the *NetworkX* [8] and *dispy* packages [20].

6 Gene Interaction Model and Bayesian Model Averaging

It is important to realize that each directed edge in the model is labeled by a number in the interval $[0, 1]$ indicating the posterior Bayesian probability that the associated relationship exists in the biological network. Using Bayesian statistics,

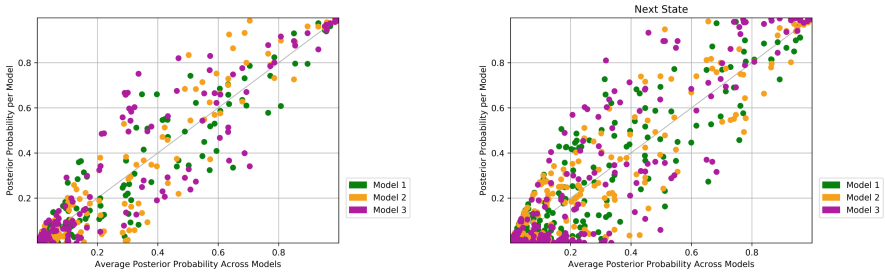


(a) NS TSA Gene Interaction Model.



(b) NS (without TSA) Gene Interaction Model.

Fig. 3. (a) is the next state time series adjusted model for ACC26 and its analysis. The numerical label on the directed edges is the posterior probability. For clarity, only edges with posterior probability greater than 0.3 are indicated. As a consequence, four genes (MYB55, MYB9, BYB93, and RING2G37950) are not shown in (a) and four genes (ANAC058, LRP1, MBS2, and MYB93) are not shown in (b).



(a) NS TSA Analysis. (b) NS without TSA Analysis.

Fig. 4. Across three independently generated similar gene interaction models, plotting the edges' average posterior probabilities versus the individual edge posterior probabilities provides a consistency analysis of the models given in Fig. 3.

these probabilities are estimated by a weighted sum over all of the models found in the archive AR. With $L = f(y_1, \ldots, y_n|D)$ the NPL of data $y_1, \ldots, y_n$, and $\chi_D(e) = 1$ if e is a directed edge in DAG D and $\chi_D(e) = 0$ otherwise, the posterior probability of an edge e is computed by the Bayesian model averaging formula [10]

$$\frac{\sum_{D \in AR} \chi_D(e) f(y_1, y_2, \ldots, y_n|D)}{\sum_{D \in AR} f(y_1, y_2, \ldots, y_n|D)},$$

which simply and appropriately weights each visited DAG D according to its likelihood. This methodology requires equally likely priors since in such a situation the posterior for D is proportional its likelihood [3]. In order for this estimate to reflect its true value, it is necessary that AR contain a large and varied collection of DAGs of high likelihood.

7 Next State Gene Interaction Models

Using the transcript abundance data for 26 *Arabidopsis thaliana* genes stimulated by ACC, gene interaction models for a next state with and without time series adjustment were computationally created, shown in Fig. 3. Each edge is labeled by its posterior probability. Figure 4 provides comparisons of three similar models to those given in Fig. 3. Figure 4(a) shows a stronger and tighter distribution of posterior probabilities than Fig. 4(b). There is significant agreement across the models for average posterior probabilities exceeding 0.8 and less than 0.2. However, for average posterior probabilities with values greater than 0.2 and less than 0.8 there is a great deal of variance, which reflects the lack of a strong posterior probability over this range.

8 Conclusion and Further Considerations

A typical underlying assumption of statistical analysis is that the residuals are independent [3, page 737]. It is well understood, however, that the residuals

associated with time course data are not usually independent. By incorporating time series adjustments into the modeling process, the residuals' independence is much improved; thus, yielding a less approximated, more accurate likelihood function.

The continuation of this research includes four tasks. First, the computational networks have been sent to the Muday lab for biological investigation, confirmation and interpretation. Second, in this paper, we investigated the enhancement of times series adjustment on a next state one step model. There are two other *time paradigms*, next state one and two steps and cotemporal, each of which has a time series adjustment analogue and a corresponding Norris-Patton likelihood. Comparing and contrasting the computational results of these three distinct modeling methods–as well as their biological interpretations–are important in understanding the gene interaction models developed using this methodology. Third, we will further consider higher order autoregressive adjustment to continue improving the independence of the residuals. Fourth, effort is underway to implement nonuniform priors in the modeling techniques. This would permit construction of gene interaction models that reflect relationships found in the literature.

Acknowledgments. The authors thank the National Science Foundation for their support with a grant, NSF#1716279. John Farrell thanks Wake Forest University for support as a Wake Forest Fellow for Summer 2019.

References

1. Cao, J., Qi, X., Zhao, H.: Modeling gene regulation networks using ordinary differential equations. In: Next Generation Microarray Bioinformatics, Methods in Molecular Biology, vol. 802, pp. 185–197. Springer (2012). https://doi.org/10.1007/978-1-61779-400-1_12
2. Cordell, H.: Detecting gene-gene interactions that underlie human diseases. Nat. Rev. Genet. **10**(2), 392–404 (2002)
3. DeGroot, M.H., Schervish, M.J.: Probability and Statistics, 4th edn. Addison-Wesley, Boston (2012)
4. Emily, M.: A survey of statistical methods for gene-gene interaction in case-control genome-wide association studies. J. Soc. Fr. Stat. **159**(1), 27–67 (2018)
5. Eschelman, L.J.: The CHC adaptive search algorithm: how to have safe search when engaging in nontraditional genetic recombination. In: Rawlins, G.J.E. (ed.) Foundations of Genetic Algorithms, pp. 265–283. Morghan Kaufmann, Burlington (1991)
6. Eshelman, L.J.: Genetic algorithms. In: Bäck, T., Fogel, D.B., Michalewicz, T. (eds.) Evolutionary Computation 1 - Basic Algorithms and Operators, Chapter 8, vol. 1, pp. 64–80. Institute of Physics Publishing, Bristol (2000)
7. Friedman, N., Linial, M., Nachman, I., Pe'er, D.: Using Bayesian networks to analyze expression data. J. Comput. Biol. **7**(3), 601–620 (2000). https://doi.org/10.1186/gb-2004-5-12-r100
8. Hagberg, A.A., Schult, D.A., Swart, P.J.: Exploring network structure, dynamics, and function using NetworkX. In: Varoquaux, G., Vaught, T., Millman, J. (eds.) Proceedings of the 7th Python in Science Conference, pp. 11–15, Pasadena (2008)

9. Harkey, A.F., et al.: Identification of transcriptional and receptor networks that control root responses to ethylene. Plant Physiol. **176**(3), 2095–2118 (2018). https://doi.org/10.1104/pp.17.00907. http://www.plantphysiol.org/content/176/3/2095
10. Hoeting, J.A., Madigan, D., Raftery, A.E., Volinsky, C.T.: Bayesian model averaging: a tutorial (with comments by M. Clyde, David Draper and E.I. George, and a rejoinder by the authors). Stat. Sci. **14**(4), 382–417 (1999)
11. John, D.J., Fetrow, J.S., Norris, J.L.: Continuous cotemporal probabilistic modeling of systems biology networks from sparse data. IEEE/ACM Trans. Comput. Biol. Bioinform. **8**(5), 1208–1222 (2011). https://doi.org/10.1109/TCBB.2010.95
12. Kedem, B., Fokianos, K.: Regression Models for Time Series Analysis. Wiley, Hoboken (2002)
13. LaPointe, B.A., et al.: A BCHC genetic algorithm model of cotemporal hierarchical Arabidopsis thaliana gene interactions. In: Proceedings of 2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 2701–2708 (2018)
14. Liang, J., Han, J.: Stochastic Boolean networks: an efficient approach to modeling gene regulatory networks. BMC Syst. Biol. **6**(113), 1–20 (2012). http://www.biomedcentral.com/1752-0509/6/113
15. Mitchell, M.: An Introduction to Genetic Algorithms. MIT Press, Cambridge (1998)
16. Patton, K.L.: Bayesian interaction and associated networks from multiple replicates of sparse time-course data. Master's thesis, Wake Forest University, Department of Mathematics (May 2012)
17. Patton, K.L., John, D.J., Norris, J.L.: Bayesian probabilistic network modeling from multiple independent replicates. BMC Bioinform. **13**(Supplement 9), 1–13 (2012)
18. Patton, K.L., John, D.J., Norris, J.L., Lewis, D., Muday, G.: Hierarchical Bayesian system network modeling of multiple related replicates. BMC Bioinform. **7**, 803–812 (2013)
19. Pe'er, D.: Bayesian network analysis of signaling networks: a primer. Sci. STKE **2005**, 1–12 (2005)
20. Pemmasani, G.: dispy: distributed and parallel computing with/for python (2016). http://dispy.sourceforge.net
21. Purcell, S., et al.: PLINK: a toolset for whole-genome association and population-based linkage analysis. Am. J. Hum. Genet. **81**, 559–575 (2007)
22. Wan, X., et al.: BOOST: a fast approach to detecting gene-gene interactions in disease data. Am. J. Hum. Genet. **87**, 325–340 (2010)
23. Yung, L.S., Yang, C., Wan, X., Yu, W.: GBOOST: a GPU-based tool for detecting gene-gene interactions in genome-wide case control studies. Bioinformatics **27**(9), 1309–1310 (2011)