Learning Continuous Time Bayesian Networks in Non-stationary Domains (Extended Abstract)[*]

Simone Villa[1] and **Fabio Stella**[2]

[1] LGT Investment Partners, Pfäffikon SZ, Switzerland

[2] Department of Informatics, Systems and Communication, University of Milan-Bicocca, Italy

simone.villa84@gmail.com, stella@disco.unimib.it

Abstract

Non-stationary continuous time Bayesian networks are introduced. They allow the parents set of each node in a continuous time Bayesian network to change over time. Structural learning of non-stationary continuous time Bayesian networks is developed under different knowledge settings. A macroeconomic dataset is used to assess the effectiveness of learning non-stationary continuous time Bayesian networks from real-world data.

1 Introduction

The identification of relationships and statistical dependencies between components in multivariate time-series, and the ability of reasoning about whether and how these dependencies change over time is crucial in many research domains such as biology, economics, finance, traffic engineering and neurology, to mention just a few. In biology, for example, knowing the gene regulatory network allows to understand complex biological mechanisms ruling the cell. In such a context, Bayesian networks (BNs) [Segal *et al.*, 2005; Nagarajan and Scutari, 2013], dynamic Bayesian networks (DBNs) [Vinh *et al.*, 2012] and continuous time Bayesian networks (CTBNs) [Acerbi *et al.*, 2014; 2016] have been used to reconstruct transcriptional regulatory networks from gene expression data. The effectiveness of discrete DBNs has been investigated to identify functional correlations among neuroanatomical regions of interest [Burge *et al.*, 2009], while a useful primer on BNs for functional magnetic resonance imaging data analysis has been made available [Mumford and Ramsey, 2014]. However, the mentioned applications require the time-series to be generated from a stationary distribution, i.e. one which does not change over time. While stationarity is a reasonable assumption in many situations, there are cases where the data generating process is clearly non-stationary. Indeed, in recent years, researchers from different disciplines, ranging from economics to computational biology, to sociology and to medicine have become interested in representing relationships and dependencies which change over time. More precisely, researchers have been interested

in analyzing the temporal evolution of genetic networks [Lebre *et al.*, 2010], the flow over neural information networks [Smith *et al.*, 2006], heart failure [Liu *et al.*, 2016], complications in type 1 diabetes [Marini *et al.*, 2015] and the dependence structure among financial markets during crisis [Durante and Dunson, 2014]. According to the specialized literature [Robinson and Hartemink, 2010], evolution models can be divided into two main categories: *structurally non-stationary*, i.e. those models which are allowed to change their structure over time, and *parametrically non-stationary*, i.e. those models which only allow the parameters' values to change over time.

In this paper, the *structurally non-stationary continuous time Bayesian network* model (nsCTBN) is introduced. A nsCTBN consists of a sequence of CTBNs which improves expressiveness over a single CTBN. Indeed, a nsCTBN allows the parents set of each node to change over time at specific *transition times* and thus it allows to model non-stationary systems. To learn a nsCTBN, the Bayesian score for learning CTBNs [Nodelman *et al.*, 2003] is extended.

The contribution of this paper is twofold. Firstly, it defines the model of *structurally non-stationary continuous time Bayesian networks*. Secondly, the effectiveness of learning non-stationary continuous time Bayesian networks from real-world data is assessed by analyzing macroeconomic data.

2 Model

In this section, continuous time Bayesian networks and non-stationary continuous time Bayesian networks are introduced together with three knowledge learning settings for structural learning from complete data.

2.1 Continuous Time Bayesian Networks

Continuous time Bayesian networks combine Bayesian networks and homogeneous Markov processes together to efficiently model discrete state continuous time dynamical systems [Nodelman *et al.*, 2002]. They are particularly useful for modeling domains in which variables evolve at different time granularities, such as to model the presence of people at their computers [Nodelman and Horvitz, 2003], to study reliability of dynamical systems [Boudali and Dugan, 2006], to model failures in server farms [Herbrich *et al.*, 2007], and to detect network intrusion [Xu and Shelton, 2008].

---

[*]This paper is an extended abstract of an article in the Journal of Artificial Intelligence Research [Villa and Stella, 2016].

The continuous time Bayesian network model is defined as follows.

Definition 1. *Continuous time Bayesian network [Nodelman et al., 2002]. Let $\boldsymbol{X}$ be a set of random variables $\boldsymbol{X} = \{X_1, X_2, \ldots, X_N\}$. Each X has a finite domain of values $Val(X) = \{x_1, x_2, \ldots, x_I\}$. A continuous time Bayesian network over $\boldsymbol{X}$ consists of two components: the first is an initial distribution $P_{\boldsymbol{X}}^0$, specified as a Bayesian network over $\boldsymbol{X}$, the second is a continuous time transition model specified as: a directed (possibly cyclic) graph $\mathcal{G}$ whose nodes are $X_1, X_2, \ldots, X_N$; a conditional intensity matrix, $\boldsymbol{Q}_X^{Pa(X)}$, for each variable $X \in \boldsymbol{X}$, where $Pa(X)$ denotes the set of parents of X in the graph $\mathcal{G}$.*

The conditional intensity matrix $\boldsymbol{Q}_X^{Pa(X)}$ consists of the set of intensity matrices

$$\boldsymbol{Q}_X^{pa_u} = \begin{bmatrix} -q_{x_1}^{pa_u} & \cdot & q_{x_1 x_I}^{pa_u} \\ q_{x_2 x_1}^{pa_u} & \cdot & q_{x_2 x_I}^{pa_u} \\ \cdot & \cdot & \cdot \\ q_{x_I x_1}^{pa_u} & \cdot & -q_{x_I}^{pa_u} \end{bmatrix},$$

where pa_u ranges over all possible configurations of the parents set $Pa(X)$, while $q_{x_i}^{pa_u} = \sum_{x_j \neq x_i} q_{x_i x_j}^{pa_u}$. Off-diagonal elements of $\boldsymbol{Q}_X^{pa_u}$, i.e. $q_{x_i x_j}^{pa_u}$, are proportional to the probability that the variable X transitions from state x_i to state x_j given the parents' state pa_u. The intensity matrix $\boldsymbol{Q}_X^{pa_u}$ can be equivalently summarized with two independent sets: $\boldsymbol{q}_X^{pa_u} = \{q_{x_i}^{pa_u} : 1 \leq i \leq I\}$, i.e. the set of intensities parameterizing the exponential distributions over *when* the next transition occurs, and $\boldsymbol{\theta}_X^{pa_u} = \{\theta_{x_i x_j}^{pa_u} = q_{x_i x_j}^{pa_u}/q_{x_i}^{pa_u} : 1 \leq i, j \leq I, j \neq i\}$, i.e. the set of probabilities parameterizing the multinomial distributions over *where* the state transitions. Note that the CTBN model assumes that only one single variable can change state at any specific instant, while its transition dynamics are specified by its parents via the conditional intensity matrix and they are independent of all other variables given its *Markov Blanket*[1].

Continuous time Bayesian networks are both *structurally stationary*, as the graph does not change over time, and *parametrically stationary*, as the conditional intensity matrices do not change over time. These stationarity assumptions are reasonable in many situations, but there are cases where the data generating process is intrinsically non-stationary and thus CTBNs can no longer be used. Therefore, the CTBNs have been extended to become structurally non-stationary, i.e. the CTBN's structure is allowed to change over continuous time.

2.2 Non-stationary CTBNs

In the non-stationary continuous time Bayesian network model, the graph of the CTBN is replaced by a *graphs sequence* $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \ldots, \mathcal{G}_E)$, where a graph $\mathcal{G}_e$ represents the causal dependency structure of the model for the epoch $e \in \{1, 2, \ldots, E\}$[2]. This model is structurally non-stationary

because of the introduction of the graphs sequence and it can handle transition times that are common to the whole network and/or node-specific. Following the notations and definitions used for non-stationary dynamic Bayesian networks (nsDBNs), we let $\mathcal{T} = (t_1, \ldots, t_{E-1})$ be the *transition times* sequence, i.e. the times at which the causal dependency structure $\mathcal{G}_e$, active at epoch e, is replaced by the causal dependency structure $\mathcal{G}_{e+1}$, which becomes active at epoch $e + 1$. An *epoch* is defined to be the period of time between two consecutive transitions, i.e. the epoch e is active during the period of time starting at t_{e-1} and ending at t_e. The graph $\mathcal{G}_{e+1}$, which is active during the epoch $e + 1$, differs from the graph $\mathcal{G}_e$, which is active during the epoch e, in a set of edges that we call the *set of edge changes* $\Delta\mathcal{G}_e$. A non-stationary continuous time Bayesian network is defined as follows.

Definition 2. *(Structurally) non-stationary continuous time Bayesian network. Let $\boldsymbol{X}$ be a set of random variables $X_1, \ldots, X_N$. Each X has a finite domain of values $Val(X) = \{x_1, \ldots, x_I\}$. A (structurally) non-stationary continuous time Bayesian network $\mathcal{N}_{ns} = (\mathcal{B}, \mathcal{M}_{ns})$ over $\boldsymbol{X}$ consists of two components:*

- *an initial distribution $P_{\boldsymbol{X}}^0$, specified as a Bayesian network $\mathcal{B}$ over $\boldsymbol{X}$,*

- *a non-stationary continuous time transition model $\mathcal{M}_{ns}$ specified as: a sequence of directed (possibly cyclic) graphs $\mathcal{G} = (\mathcal{G}_e)_{e=1}^E$ whose nodes are $X_1, \ldots, X_N$, where E represents the number of epochs; a conditional intensity matrix, $\boldsymbol{Q}_{X,H(X)}^{Pa_{\mathcal{G}}(X)}$, $\forall X \in \boldsymbol{X}$, where $Pa_{\mathcal{G}}(X)$ denotes the parents sets of X in $\mathcal{G}$, and $H(X)$ denotes the intervals associated with X.*

The conditional intensity matrix $\boldsymbol{Q}_{X,H(X)}^{Pa_{\mathcal{G}}(X)}$ consists of a set of intensity matrices

$$\boldsymbol{Q}_{X,h_m}^{pa_u} = \begin{bmatrix} -q_{x_1,h_m}^{pa_u} & \cdot & q_{x_1 x_I,h_m}^{pa_u} \\ q_{x_2 x_1,h_m}^{pa_u} & \cdot & q_{x_2 x_I,h_m}^{pa_u} \\ \cdot & \cdot & \cdot \\ q_{x_I x_1,h_m}^{pa_u} & \cdot & -q_{x_I,h_m}^{pa_u} \end{bmatrix},$$

one for each configuration pa_u of each parents set $Pa(X) \in Pa_{\mathcal{G}}(X)$ which is active during the interval $h_m \in H(X)$.[3]

Non-stationary continuous time Bayesian networks allow each node to have its own sequence of parents sets, each parents set being active at a given epoch. Therefore, we introduce the concept of *homogeneous interval* $H(X) = (h_1, \ldots, h_M)$ associated with node X, which is defined as the union of consecutive epochs during which the same parents set $Pa(X)$ is active for the node X. Note that if each epoch is associated with a different parents set, then M is equal to E. Figure 1 shows the graphs sequence $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4)$ of a nsCTBN consisting of four epochs ($E = 4$) with transition times $\mathcal{T} = (t_1, t_2, t_3)$.

Learning a non-stationary continuous time Bayesian network from a fully observed dataset can be done using the Bayesian learning framework. The following learning settings are considered; known transition times (KTT), where

---

[1]The set of its parents, children, and its children's other parents.

[2]It is worthwhile to mention that the first epoch, i.e. the epoch starting at time 0 and ending at time t_1 is associated with the graph $\mathcal{G}_1$, while the last epoch, i.e. the epoch starting at time t_{E-1} and ending at time T (the supremum of the considered time interval, i.e. [0,T]) is associated with the graph $\mathcal{G}_E$.

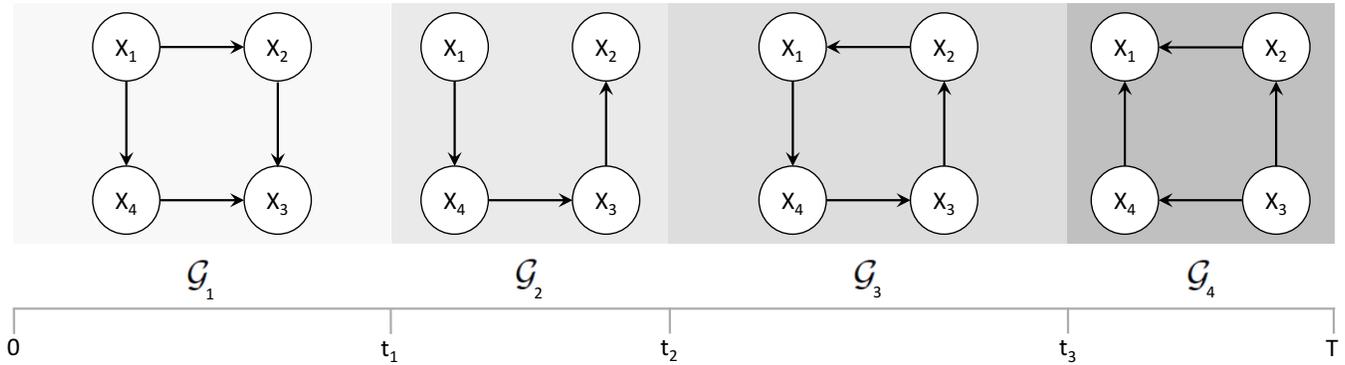[3]Note that the equation $q_{x_i,h_m}^{pa_u} = \sum_{x_j \neq x_i} q_{x_i x_j,h_m}^{pa_u}$ still holds.

Figure 1: Graphs sequence $\mathcal{G} = (\mathcal{G}_1, \mathcal{G}_2, \mathcal{G}_3, \mathcal{G}_4)$ of a nsCTBN with four epochs, $E = 4$, and three transition times, $\mathcal{T} = (t_1, t_2, t_3)$, where the edges are gained and lost over time. Each epoch is associated with a set of edge changes. Specifically, the graph $\mathcal{G}_2$ differs from the graph $\mathcal{G}_1$ by the following set of edge changes $\Delta\mathcal{G}_1 = \{X_3 \rightarrow X_2, X_2 \nrightarrow X_3, X_1 \nrightarrow X_2\}$, the graph $\mathcal{G}_3$ differs from the graph $\mathcal{G}_2$ by the following set of edge changes $\Delta\mathcal{G}_2 = \{X_2 \rightarrow X_1\}$ and the graph $\mathcal{G}_4$ differs from the graph $\mathcal{G}_3$ by the following set of edge changes $\Delta\mathcal{G}_3 = \{X_3 \rightarrow X_4, X_4 \rightarrow X_1, X_1 \nrightarrow X_4, X_4 \nrightarrow X_3\}$.

transition times between epochs are given, known number of epochs (KNE), where the number of epochs is given, and unknown number of epochs (UNE) where nothing is known about epochs. These learning settings have increasing difficulty, for each of them the Bayesian score decomposition has been formally derived and the corresponding learning algorithm has been designed [Villa and Stella, 2016].

3 Numerical Experiments

Numerical experiments have been performed on both synthetic and real-world datasets. Synthetic datasets were used to compare nsCTBNs to nsDBNs under three knowledge settings in terms of accuracy, precision, recall and F_1 measure. The following real-world datasets: *drosophila*, *saccharomyces cerevisiae* and *songbird*, were used to compare nsCTBNs to state-of-the-art algorithms, i.e. TSNI (a method based on ordinary differential equations), nsDBN [Robinson and Hartemink, 2010] and non-homogeneous dynamic Bayesian networks with Bayesian regularization (TVDBN) [Dondelinger *et al.*, 2013], under the UNE knowledge setting. Results show that nsCTBNs are highly competitive when compared to state-of-the-art algorithms for both synthetic and real-world datasets [Villa and Stella, 2016].

In this paper, for the sake of brevity, we illustrate results of nsCTBNs structural learning on the *macroeconomics dataset* which consists of 17 financial/economic time-series pertaining to the economy of the United States. Time-series have different time granularity and span from 1st January 1986 to 31st March 2015. More specifically, five time-series have daily granularity, namely Crude oil (OIL), USD to EUR spot exchange rate (USDEUR), Gold (GOLD), S&P500 equity index (S&P500) and the 10-years treasury bond yield rate (US10yrsNote). Eleven time-series have monthly granularity, namely production of total industry (PTI), real manufacturing and trade industries sales (RMTIS), personal income (PI), unemployment (UN), consumer price index (CPI), federal funds rate (RATE), producer price index (PPI), non-farm payrolls (NFP), new one-family houses sold (NHSold), new houses

for sale (NHSale) and new private house permits (NHPermit). Finally, the gross domestic product (GDP) time-series has quarterly granularity.

The goal of this study is to discover how the financial and economic environment evolve over time. In particular, we focused the attention to detect *business cycles*[4] and the associated change of relationships among financial and economic variables. Given that the duration of a business cycle is highly variable, the ability to identify the turning point of a cycle (i.e. when a recession starts) is of considerable importance to policymakers, financial companies as well as to individuals. A substantial literature is available about the business cycle turning points detection relying on Markov-switching models [Hamilton and Raj, 2005]. However, these models are not able to represent important features such as the dependence structure among variables in each business cycle. In order to use the nsCTBN model in such a context, we applied a binary discretization to the variable associated with each time-series. Discretization was performed using a look-back period of 1 year for daily data (i.e. if the current value is greater than the past one, then the binary variable is set to 1 otherwise it is set to 0), while a look-back of 1 month for monthly data and 3 months for quarterly data. The approach of looking back into the past is widely used in finance [Moskowitz *et al.*, 2012]. nsCTBNs learning was performed under the UNE setting using the following parameter values: $\lambda_c = \{0.5, 1, 2\}$, $\lambda_e = \{0.1, 1, 10\}$, 2 maximum parents per node, 300 iterations and 10 runs.

Figure 2 shows the probability of transition (left side, left axis) versus the S&P500 equity index used as a reference (left side, right axis) and the posterior probability over the number of epochs (right side). The nsCTBN consists of three epochs with transition times close to the end of July 2000 and the end of November 2007.

---

[4]Business cycles are fluctuations in aggregate economic activity, they are recurrent (i.e. it is possible to identify expansion-recession cycles), persistent and not periodic (i.e. they differ in length and severity).
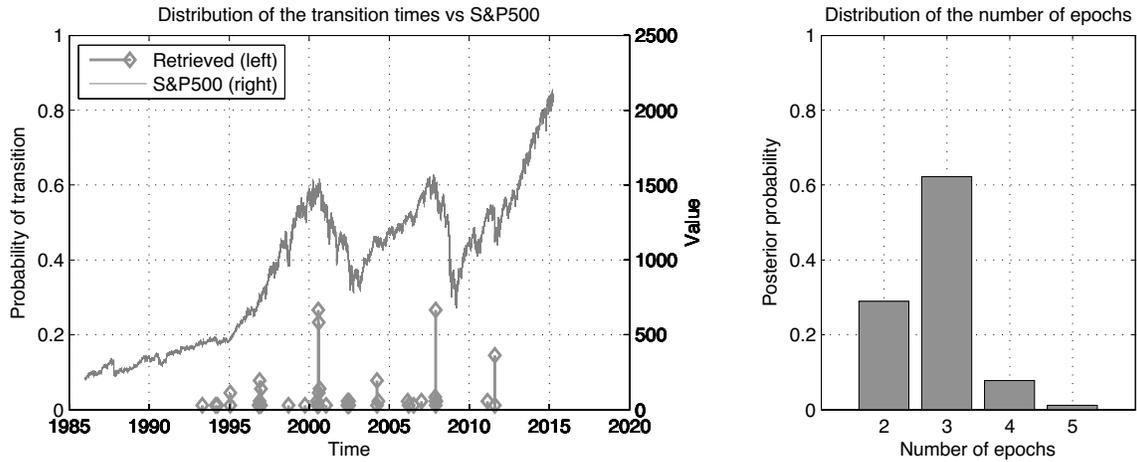
Figure 2: Distribution of the transition times versus the S&P500 equity index over time (left picture) and the posterior probability distribution of the number of epochs (right picture) for the learned nsCTBN under the unknown number of epochs setting.

(a) Epoch 1 (Jan 1986 - Jul 2000).   (b) Epoch 2 (Aug 2000 - Nov 2007).   (c) Epoch 3 (Dec 2007 - Mar 2015).
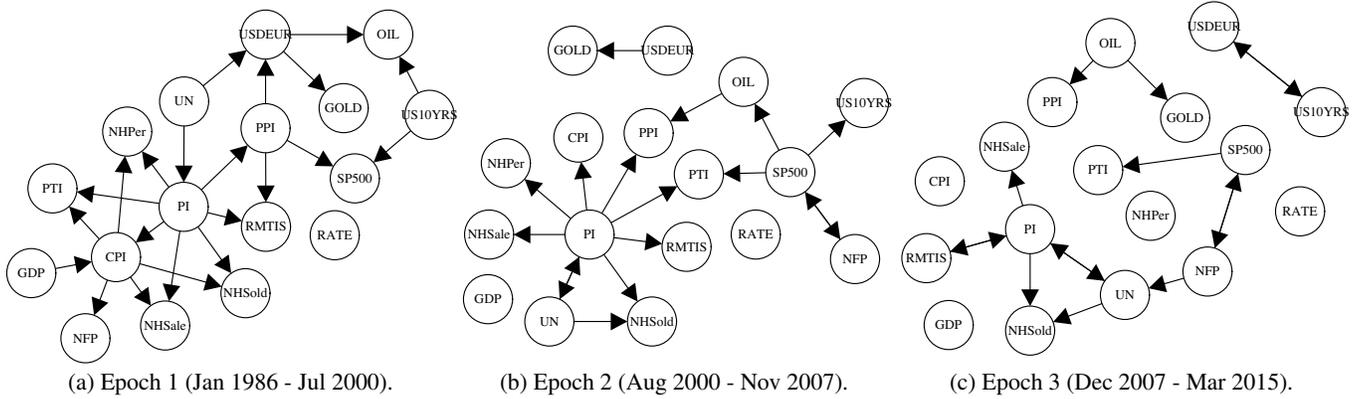
Figure 3: nsCTBN learned on the macroeconomics dataset under the unknown number of epochs setting. nsCTBN corresponds to the most probable number of epochs ($E = 3$). An arc is included in the nsCTBN model when it occurs in more than 75% of the runs in each epoch.

If we compare these dates to the turning points of the US business cycle reported by the National Bureau of Economic Research[5], then we see that we are not far from the turning point of March 2001 and very close to the one of December 2007, while the turning point occurred in July 1990 is missed, probably because of the limited length of the dataset.

Figure 3 shows the structure of the nsCTBN model corresponding to the most probable number of epochs, i.e. $E = 3$. An arc is included in the nsCTBN model when it occurs in more than 75% of the performed runs in each epoch. The retrieved networks correspond to the following time periods: from January 1986 to July 2000, from August 2000 to November 2007 and from December 2007 to March 2015.

The novelty of such approach opens the door to many considerations about the economic variables during business cycles, e.g. the learned models show the well-known central role of the personal income and its relation to the unemployment [Mankiw, 2014] and the less known relation of the nonfarm payrolls to the S&P500 equity index [Miao et al., 2014].

---

[5]See the official data at http://www.nber.org/cycles.html

4 Conclusions

We introduced non-stationary continuous time Bayesian networks (nsCTBNs), a probabilistic graphical model generalizing continuous time Bayesian networks to the case where the structure of the probabilistic graphical model is allowed to change over time. nsCTBNs have been shown to be competitive with the state-of-the-art algorithms when both synthetic and real-world datasets are considered. The results of applying the nsCTBN's structural learning algorithm when the number of epochs is unknown show the usefulness of the proposed approach but at the same time warn about the following issues; the variables are assumed to be discrete; specifically each variable of the dataset must take value over a countable number of states and finding the optimal parameter settings of structural learning is extremely difficult as it happens for non-stationary dynamic Bayesian networks. Possible directions for further research include the development of a modeling approach that goes towards the direction of allowing each node to change its parents set asynchronously as well as the objects classification task in a streaming domain.

References

[Acerbi *et al.*, 2014] Enzo Acerbi, Teresa Zelante, Vipin Narang, and Fabio Stella. Gene network inference using continuous time bayesian networks: a comparative study and application to th17 cell differentiation. *BMC Bioinformatics*, 15(1), 2014.

[Acerbi *et al.*, 2016] Enzo Acerbi, Elena Viganò, Michael Poidinger, Alessandra Mortellaro, Teresa Zelante, and Fabio Stella. Continuous time bayesian networks identify prdm1 as a negative regulator of th17 cell differentiation in humans. *Scientific Reports*, 6:23128, 2016.

[Boudali and Dugan, 2006] Hichem Boudali and Joanne Bechta Dugan. A continuous-time bayesian network reliability modeling, and analysis framework. *IEEE Transactions on Reliability*, 55(1):86–97, 2006.

[Burge *et al.*, 2009] John Burge, Terran Lane, Hamilton Link, Shibin Qiu, and Vincent P. Clark. Discrete dynamic bayesian network analysis of fmri data. *Human brain mapping*, 30(1):122–137, 2009.

[Dondelinger *et al.*, 2013] Frank Dondelinger, Sophie Lebre, and Dirk Husmeier. Non-homogeneous dynamic bayesian networks with bayesian regularization for inferring gene regulatory networks with gradually time-varying structure. *Machine Learning*, 90(2):191–230, 2013.

[Durante and Dunson, 2014] Daniele Durante and David B. Dunson. Bayesian dynamic financial networks with time-varying predictors. *Statistics & Probability Letters*, 93:19–26, 2014.

[Hamilton and Raj, 2005] James D. Hamilton and Baldev Raj, editors. *Advances in Markov-Switching Models: Applications in Business Cycle Research and Finance*. Studies in Empirical Economics. Springer-Verlag, 2005.

[Herbrich *et al.*, 2007] Ralf Herbrich, Thore Graepel, and Brendan Murphy. Structure from failure. In *The 2nd USENIX workshop on Tackling computer systems problems with machine learning techniques (SYSML 07), Cambridge, USA*, pages 1–6, 2007.

[Lebre *et al.*, 2010] Sophie Lebre, Jennifer Becq, Frederic Devaux, Michael Stumpf, and Gaelle Lelandais. Statistical inference of the time-varying structure of gene regulation networks. *BMC Systems Biology*, 4(1):130+, 2010.

[Liu *et al.*, 2016] Manxia Liu, Arjen Hommersom, Maarten van der Heijden, and Peter J.F. Lucas. Hybrid time bayesian networks. *International Journal of Approximate Reasoning*, pages 460 – 474, 2016.

[Mankiw, 2014] N. Gregory Mankiw. *Principles of Macroeconomics*. South-Western College Pub, 7th edition, 2014.

[Marini *et al.*, 2015] Simone Marini, Emanuele Trifoglio, Nicola Barbarini, Francesco Sambo, Barbara Di Camillo, Alberto Malovini, Marco Manfrini, Claudio Cobelli, and Riccardo Bellazzi. A dynamic bayesian network model for long-term simulation of clinical complications in type 1 diabetes. *J. of Biomed.l Inf.*, 57:369 – 376, 2015.

[Miao *et al.*, 2014] Hong Miao, Sanjay Ramchander, and J. Kenton Zumwalt. S&p 500 index-futures price jumps and macroeconomic news. *Journal of Futures Markets*, 34(10):980–1001, 2014.

[Moskowitz *et al.*, 2012] Tobias J. Moskowitz, Yao Hua Ooi, and Lasse Heje Pedersen. Time series momentum. *Journal of Financial Economics*, 104(2):228–250, 2012.

[Mumford and Ramsey, 2014] Jeanette A. Mumford and Joseph D. Ramsey. Bayesian networks for fmri: A primer. *Neuroimage*, 86:573–582, 2014.

[Nagarajan and Scutari, 2013] Radhakrishnan Nagarajan and Marco Scutari. *Bayesian Networks in R with Applications in Systems Biology*. Springer, New York, 2013. ISBN 978-1-4614-6445-7, 978-1-4614-6446-4.

[Nodelman and Horvitz, 2003] Uri Nodelman and Eric Horvitz. Continuous time bayesian networks for inferring users' presence and activities with extensions for modeling and evaluation. Technical Report MSR-TR-2003-97, Microsoft Research, 2003.

[Nodelman *et al.*, 2002] Uri Nodelman, Christian R. Shelton, and Daphne Koller. Continuous time bayesian networks. In *The 18th Conference on Uncertainty in Artificial Intelligence (UAI 2002), Edmonton, Canada*, pages 378–387, 2002.

[Nodelman *et al.*, 2003] Uri Nodelman, Christian R. Shelton, and Daphne Koller. Learning continuous time bayesian networks. In *The 19th Conference on Uncertainty in Artificial Intelligence (UAI 2003), Acapulco, Mexico*, pages 451–458, 2003.

[Robinson and Hartemink, 2010] Joshua W. Robinson and Alexander J. Hartemink. Learning non-stationary dynamic bayesian networks. *J. of Mach. Learn. Res.*, 11:3647–3680, 2010.

[Segal *et al.*, 2005] Eran Segal, Dana Pe'er, Aviv Regev, Daphne Koller, and Nir Friedman. Learning module networks. *J. of Mach. Learn. Res.*, 6:557–588, 2005.

[Smith *et al.*, 2006] V. Anne Smith, Jing Yu, Tom V. Smulders, Alexander J. Hartemink, and Erich D. Jarvis. Computational Inference of Neural Information Flow Networks. *PLoS Computational Biology*, 2(11):e161+, November 2006.

[Villa and Stella, 2016] Simone Villa and Fabio Stella. Learning continuous time bayesian networks in non-stationary domains. *Journal of Artificial Intelligence Research*, 57:1–37, 2016.

[Vinh *et al.*, 2012] Nguyen Xuan Vinh, Madhu Chetty, Ross Coppel, and Pramod P Wangikar. Gene regulatory network modeling via global optimization of high-order dynamic bayesian network. *BMC Bioinformatics*, 13:131, 2012.

[Xu and Shelton, 2008] Jing Xu and Christian R. Shelton. Continuous time bayesian networks for host level network intrusion detection. In *The European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECML PKDD 2008), Antwerp, Belgium*, pages 613–627, 2008.