

Reviewing Graphical Modelling of Multivariate Temporal Processes

Matthias Eckardt

Abstract Graphical models provide a suitable approach of dealing with uncertainty and complexity by using conditional independence statements and factorizations of joint densities. Static undirected as well as directed graphical models have been applied frequently to pattern analysis, decision modelling, machine learning or image filtering. Several temporal extensions have been published including dynamic Bayesian networks or temporal Markov random fields. Although, graphical models are most commonly used within computer science there has been a growing interest in adjacent disciplines. Recently, a few temporal extensions have been applied to multivariate time series data and event histories.

1 Introduction

Statistical models for the analysis of highly complex data and processes have gained strong attraction within the last decade. Graphical models have been proven to be a sufficient approach of dealing with high dimensionality and uncertainty. A reduction of complexity is achieved by using factorizations of joint densities. Plotting the graph offers a simple and intuitive visualization of the underlying dependence structure. Besides, several statistical models can be seen as special cases of a general graphical model formalism (e.g. mixture models, factor analysis).

Most research on graphical models have focussed on cross-sectional data, henceforth referred to as *static* graphical models. As most prominent cases, this class includes *directed acyclic graphs* (also known as *Bayesian networks*) introduced by Pearl (1988) as well as *undirected graphical models*, which are also labelled as *conditional independence graphs* or *Markov random fields* (cf. Lauritzen 1996). Applications of static models can be found in different scientific fields including machine learning, decision modelling, artificial intelligence or image analysis. Markov random fields have played a prominent role in spatial econometrics and spatial statistics with regard to lattice data. Currently, a strong increase of interest

M. Eckardt (✉)

Institute of Computer Science, Humboldt-Universität zu Berlin, Berlin, Germany
e-mail: eckardtm@cms.hu-berlin.de

in resp. usage of (social) network analysis emerged in economics as well as risk assessment or natural sciences. A profound treatment of static graphical models is presented in Lauritzen (1996), Cowell et al. (1999), Edwards (2000), Cox and Wermuth (1996), Pearl (1988), Spirtes (2000), Whittaker (2008) and Koller and Friedman (2010).

Recently, several extensions have been introduced aiming to model temporal dynamics and changing dependence structures including time series data as well as event histories. The objective of this paper is to review these extensions.

2 Graph Theoretic Preliminaries

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ denote a graph with $\mathcal{V} = \{v_1, \dots, v_k\}$ as finite set of *vertices* and $\mathcal{E} \subseteq \mathcal{V} \times \mathcal{V}$ as set of *edges*—joining the vertices where $\mathcal{E}(\mathcal{G}) \cap \mathcal{V}(\mathcal{G}) = \emptyset$. Two vertices v_i, v_j are called *adjacent* if and only if $(v_i, v_j) \in \mathcal{E}(\mathcal{G})$. Otherwise, v_i, v_j are *non-adjacent*. If an edge e_i consists of a pair of identical nodes (v_i, v_i) we refer to e_i as a loop of \mathcal{G} . Furthermore, if a distinct pair (v_i, v_j) is joined by more than one edge, we label this *multiple edges* or *parallel edges*. In order to emphasize graphs containing multiple edges we term any such graph *multigraph*. Similarly, a graph is *simple* if it is not a multigraph. An *undirected* or *unoriented* edge exists if the pairs (v_i, v_j) and (v_j, v_i) are both in the edge set $\mathcal{E}(\mathcal{G})$ given $v_i \neq v_j$. We use $v_j \sim v_i$ to indicate undirected edges. To specify adjacency in case of undirected edges we define the *neighbourhood* as $\text{ne}(v_j) = \{v_i : v_j \sim v_i\}$. If all edges in $\mathcal{E}(\mathcal{G})$ are undirected \mathcal{G} is an *undirected graph*. In contrast, an edge is called *directed* or *oriented* if the ordered pair $(v_i, v_j) \in \mathcal{E}(\mathcal{G})$, $v_i \neq v_j$. Thus, only $(v_i, v_j) \in \mathcal{E}(\mathcal{G})$ while $(v_j, v_i) \notin \mathcal{E}(\mathcal{G})$. Directed edges are also referred to as *arcs* and we write $v_i \rightarrow v_j$ if there is a directed edge from v_i to v_j . Formally, for $v_i \neq v_j$ we define $\text{pa}(v_j) = \{v_i : v_i \rightarrow v_j\}$ as the *parents*. In addition, if $v_i \rightarrow v_j$ we call $\{v_j\}$ *children* of v_i and define $\text{ch}(v_i) = \{v_j : v_i \rightarrow v_j\}$ (see Kolaczyk 2009; Bondy and Murty 2008). Additionally, let $\text{co-pa}(v_i) = \{v_j : \text{ch}(v_j) \cap \text{ch}(v_i) \neq \emptyset\}$ be the *co-parents* of v_i . Hence, if $\text{co-pa}(v_i) = v_j$ it follows that v_i and v_j share a common child whereas v_i and v_j have not necessarily to be joined by an edge. Consequently, a *directed graph* or *digraph* is a graph exclusively build on directed edges. A detailed discussion on digraphs is given in Bang-Jensen (2001). In addition, a *multiple digraph* is a multiple graph exclusively build on directed edges. Following Kolaczyk (2009), *mutual edges* are directed parallel edges with diametrically direction (e. g. if $v_i \rightarrow v_j$ and $v_j \rightarrow v_i$ are in $\mathcal{E}(\mathcal{G})$) whereas *multi-arcs* are multiple edges with identical orientation.

Let $(v_0, e_1, v_1, e_2, \dots, v_{k-1}, e_k, v_k)$ be a sequence of vertices and edges of \mathcal{G} with endpoints v_0 and v_k such that $\forall e_i, 1 \leq i \leq k$ the pair v_{i-1} and v_i is joined by e_i . We call this sequence w_{ik} of potentially repeating pairs of vertices a *walk* of length k in \mathcal{G} . If a walk passes through every node of a sequence exactly once we label this as a *path*. A path with identical endpoints is a *cycle* and a cycle of length one is a loop. Thus, a *directed acyclic graph* is a digraph without any directed cycles which

is linked to undirected graphs by an operation termed *moralization*. In a first step an undirected edge is inserted joining every co-parents in \mathcal{G} . Hereafter, every directed edge in \mathcal{G} is substituted by an undirected edge.

A graph build on directed as well as undirected edges is called a *mixed graph*. As a special case hereof, a *chain graph* is a simple mixed graph without any partially directed cycle where $\mathcal{V}(\mathcal{G})$ is partitioned into k blocks \mathcal{B}_k such that $\mathcal{V}(\mathcal{G}) = \mathcal{B}_1 \cup \mathcal{B}_2 \cup \dots \cup \mathcal{B}_{k-1} \cup \mathcal{B}_k$ and

1. $v_i \longrightarrow v_j$ if and only if $v_i \in \mathcal{B}_i$ and $v_j \in \mathcal{B}_j, i < j$
2. $v_i \sim v_j$ if and only if $v_i \in \mathcal{B}_i$ and $v_j \in \mathcal{B}_i$.

Hence, undirected and directed acyclic graphs are special cases of chain graphs.

3 Causality in Graphical Models

Graphical models relate conditional independence statements among random variables of a multivariate probability distribution to graphs such that the nodes represent random variables and edges encode the dependence structure. Recently, these models have been extended to the time domain based on different definitions of causality. Eichler (2013), Didelez (2011) as well as Eichler and Didelez (2010) have discussed alternative formulations including *Granger causality*, *Sims causality*, *local dependence* and causality in terms of *interventions*. Hereof, Granger causality and local dependence have been most prominent with regard to graphical models. As shown in Florens and Fougere (1996) local dependence can be seen as a continuous time version of Granger causality.

Definition 1 (Granger Causality). Let $\{X(t)\}$ and $\{Y(t)\}$ be stochastic processes on $(\Omega, \mathcal{F}, \mathbb{P})$ where $t \in T \subset \mathbb{Z}$. Given $\{\Omega(t)\}$ as all information in the universe $\{X(t)\} \subset \{\Omega(t)\}$ is causal with respect to $\{Y(t)\}$ if the prediction is less precise based on $\{\Omega(t)\} \setminus \{X(t)\}$ (ct. Granger 1969; Lütkepohl 2005).

Different levels of Granger causal relations have been discussed in Florens and Fougere (1996). Obviously, conditional on all information in the universe seems impracticable and might be replaced by all available information with regard to a vector valued process.

Local dependence has been introduced by Schweder (1970) who focussed on transition intensities in discrete state-space Markov processes. As an extension hereof, Aalen (1987) considered non-causality in continuous time restricted to processes satisfying the Doob–Meyer decomposition.

Definition 2 (Local Independence). Let \mathcal{F}_t denote the information which is available at time $t \in T \subset \mathbb{R}_+ = [0, \infty)$. Then, $\{Y(t)\}$ is locally independent of $\{X(t)\}$ given $\{Z(t)\}$ if the compensator $\Lambda^Y(t)$ of $\{Y(t)\}$ remains unchanged whether conditional on \mathcal{F}_{t-}^{XYZ} or conditional on \mathcal{F}_{t-}^{YZ} (ct. Schweder 1970; Aalen 1987; Florens and Fougere 1996).

4 Graphical Modelling of Temporal Processes

Several graphical models have been developed to approach temporal stochastic processes either in the time as well as in the frequency domain. Misleadingly, different models have been named identical. Generally, two different classes can be differentiated with regard to the underlying definition of the nodes. Firstly, the vertex set can encode random variables at different times. Secondly, the components of a vector valued process can be represented by individual nodes which lead to a coarser modelling of the graphical structure.

4.1 Time Series Data

Most of the research regarding graphical modelling of time series data focussed on vector valued stationary processes in discrete time. Different approaches of graphical modelling towards autoregressive processes are summarized in Songsiri et al. (2010).

Static graphical models have been applied to time series data manifold. Chain graph models in which time slices are represented by blocks have been discussed in Lynggaard and Walther (1993) and Dahlhaus and Eichler (2003). Furthermore, Queen and Smith (1992) and Anacleto and Queen (2013) introduced a *dynamic chain graph* model based on multivariate Bayesian dynamic models.

Extensions of directed acyclic graphs (so-called *dynamic Bayesian networks*) have extensively been treated by Murphy (2002). Defined as a sequence of directed acyclic graphs such that time is displayed in form of stacked time slices these models include hidden Markov models and Kalman filters besides other latent state-space models as special cases. An overview of different models belonging to this class of graphs is given in Barber and Cengil (2010). Anacleto et al. (2013a,b) extended the multiregression dynamic model as introduced in Queen and Smith (1993) and used directed acyclic graphs to analyse multivariate time series from traffic flows. These graphs have been called *Bayesian dynamic graphical models*.

Several papers have focussed on structural vector autoregressive processes and discussed the sufficiency of learning structural constraints from static graphical models. Moneta (2008) proposed the usage of a structural learning algorithm with regard to directed acyclic graphs in order to obtain the contemporaneous dependence structure. Alternatively, Oxley et al. (2008), Meurk et al. (2007) as well as Penny and Reale (2004) presented a two-step estimation procedure based on a novel graph linkage called *demoralization*. Thereby, the optimal directed acyclic graph is chosen from a list of suitable graphs obtained from an undirected graphical model by the inverse moralization operation.

Additionally, sequences of undirected Gaussian graphical models arranged in independent and identically distributed blocks have been used by Talih and Hengartner (2005), Talih (2003) and Cai and Li (2012) to capture changing dependence

structures of multivariate time series. Thus, a new block only emerges if a new edge is included or excluded into $\mathcal{E}(\mathcal{G})$ such that consecutive blocks contain different graphs. Xuan and Murphy (2007) presented a similar approach. Gao and Tian (2010) proposed a mixed graphical model called *latent ancestral graph* to model latent variables in case of structural vector autoregressive processes.

A first approach focussing on graphical models related to components of vector valued stationary time series in discrete time has been presented by Brillinger (1996) in frequency domain. Dahlhaus (2000) introduced a refined version hereof called *partial correlation graph* based on partial spectral coherence between components of a multivariate time series. The resulting graph is a simple undirected graph in which the nodes correspond to the components of a vector valued process. Partial spectral coherence measures the dependence of two components after removing linear time invariant effect of the remaining series (cf. Brillinger 1981). Thus, two components are conditional orthogonal if the partial error processes are uncorrelated after filtration. This is equivalent to conditional independence only in cases of Gaussian time series. Efficiently, instead of computing the correlation between partial error processes conditional orthogonality can equivalently be achieved from the partial spectra coherence and similarly been read of zero entries of the inverse spectral matrix. Partial correlation graphs have been applied in various fields by Gather et al. (2002), Fried et al. (2004), Feiler et al. (2005) and Allali et al. (2008). Avventi et al. (2013) discussed the usage of these graphs in case of autoregressive moving average processes. Additionally, structural learning has been treated in Bach and Jordan (2004).

Eichler (1999, 2012) has introduced a mixed graph called *Granger causality graph* since it encodes Granger causal as well as contemporaneous relationships between time series components. This graph has also been called *dynamic chain graph* by Murphy (2002). Corander and Villani (2003, 2006) discussed Granger causal graphs from a Bayesian perspective. Additionally, Marttinen and Corander (2009) dealt with the task of Bayesian learning of such graphs. Application of Granger causal graphs are presented in Wild et al. (2010), Allali et al. (2011) and Arnold et al. (2007). Focussing on latent variables Eichler (2010) recently introduced extended Granger causality graphs related to autoregressive moving average processes which he called *dynamic maximal ancestral graphs*. In difference to Granger causality graphs these graphs consist of one additional edge type.

4.2 Event History Data

Based on the concept of local dependence as described in Sect. 3 Didelez (2000) introduced extensions of directed acyclic graphs with regard to counting processes which she termed *local dependence graphs*. These models are defined in case of marked point processes (Didelez 2008) as well as composable finite Markov processes (Didelez 2007). Similarly, Gottard (2007) presented so-called *graphical duration models* as extensions of chain graph models displaying marked point

processes. Dreassi and Gottard (2007) dealt with Bayesian estimation of this duration models. This graphs might also allow to model frailty terms besides the hierarchical structures discussed in Gottard and Rampichini (2007). Additionally, Fosen et al. (2006) derived *dynamic path analysis* models based upon Aalens additive hazard model. A further discussion of these models is given in Aalen et al. (2008), whereas the large sample properties are derived in Martinussen (2010). Alternatively, dynamic Bayesian networks build on two time slices have been used to model duration data. Confusingly, Donat et al. (2008, 2010) named these models also *graphical duration models*.

Graphical modelling of counting processes in the frequency domain has been described in Dahlhaus et al. (1997) and Eichler et al. (2003). In both papers partial correlation graphs are used to model intensity functions.

4.3 Beyond Discrete Time

Besides discrete time a limited number of papers have been published with regard to stochastic processes that evolve in continuous time. Shelton et al. (2010) and Nodelman et al. (2002, 2003) introduced continuous time Bayesian networks (CTBN). Alternatively, El-Hay et al. (2006) presented an extended version of undirected graphical models which they called continuous time Markov network (CTMN).

5 Discussion

This review has shown the state of the art and the great variety of graphical models in the temporal domain. Mostly, they are closely related to classical graphical models and satisfy the traditional Markov properties. The main difference between the approaches arises with regard to the inherent definition of the nodes. Relating the vertices to the components of a multivariate process leads to a coarser modelling of the dependence structure. Oppositely, the dimension of the graphical model strongly increases in case of temporally separated nodes as $\dim(\mathcal{G}) = N \times T$. This might negatively impact the costs of estimation and computational efficiency, especially in case of high dimensional data structures evolving in time. Contrary, parameter as well as structural estimation can be built on already existing algorithms when applying static models to time series data.

Additionally, Granger causality graphs are strongly effected by the choice of the correct time intervals taking into account. Hence, larger intervals correspond to marginalization over time and create additional correlation. Nonetheless, the graphs are suitable models for structural learning in high dimensional time series and might easily be extended to more complex dimensions. Furthermore, extensions to non-linear relationships can be achieved conditioning on σ -algebras.

References

- Aalen, O. O. (1987). Dynamic modelling and causality. *Scandinavian Actuarial Journal*, 1987, 177–190.
- Aalen, O. O., Borgan, Ø., & Gjessing, H. K. (2008). *Survival and event history analysis: A process point of view*. Berlin: Springer.
- Allali, A., Oueslati, A., & Trabelsi, A. (2008). The analysis of the interdependence structure in international financial markets by graphical models. *International Research Journal of Finance and Economics*, (15), 283–298.
- Allali, A., Oueslati, A., & Trabelsi, A. (2011). Detection of information flow in major international financial markets by interactivity network analysis. *Asia-Pacific Financial Markets*, 18, 319–344.
- Anacleto, O., & Queen, C. (2013). *Dynamic chain graph models for high-dimensional time series*. Technical Report, The Open University, Milton Keynes.
- Anacleto, O., Queen, C., & Albers, C. J. (2013a). Multivariate forecasting of road traffic flows in the presence of heteroscedasticity and measurement errors. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 62, 251–270.
- Anacleto, O., Queen, C., & Albers, C. J. (2013b). Forecasting multivariate road traffic flows using Bayesian dynamic graphical models, splines and other traffic variables. *Australian and New Zealand Journal of Statistics*, 55, 69–86.
- Arnold, A., Liu, Y., & Abe, N. (2007). Temporal causal modelling with graphical granger methods. In P. Berkhin, R. Caruana & X. Wu (Eds.), *Proceedings of the 13th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 66–75). New York: ACM.
- Avventi, E., Lindquist, A. G., & Wahlberg, B. (2013). ARMA identification of graphical models. *IEEE Transactions on Automatic Control*, 58, 1167–1178.
- Bach, F. R., & Jordan, M. I. (2004). Learning graphical models for stationary time series. *IEEE Transactions on Signal Processing*, 52, 2189–2199.
- Barber, D., & Cengil, A.T. (2010). Graphical models for time series. *IEEE Signal Processing Magazine, Special Issue on Graphical Models*, 27, 18–28.
- Bang-Jensen, J. (2001). *Digraphs: Theory, algorithms and applications*. Berlin: Springer.
- Bondy, J. A., & Murty, U. S. R. (2008). *Graph theory*. Berlin: Springer.
- Brillinger, D. R. (1981). *Time series: Data analysis and theory*. New York: Holt, Rinehart and Winston.
- Brillinger, D. R. (1996). Remarks concerning graphical models for time series and point processes. *Revista de Econometria*, 16, 1–23.
- Cai, F., & Li, Y. (2012). Modelling changing graphical structure. *International Journal of Engineering and Manufacturing*, 2, 50–57.
- Corander, J., & Villani, M. A. (2003). *Causality in vector autoregressions: A Bayesian graphical modelling approach*. Technical Report. Statistiska Institutionen, Stockholms Universitet.
- Corander, J., & Villani, M. A. (2006). Bayesian approach to modelling graphical vector autoregressions. *Journal of Time Series Analysis*, 27, 141–156.
- Cowell, R. G., Dawid, A. P., Lauritzen, S., & Spiegelhalter, D. J. (1999). *Probabilistic networks and expert systems*. Berlin: Springer.
- Cox, D. R., & Wermuth, N. (1996). *Multivariate dependencies: Models, analysis and interpretation*. Boca Raton: Chapman & Hall CRC.
- Dahlhaus, R. (2000). Graphical interaction models for multivariate time series. *Metrika*, 51, 157–172.
- Dahlhaus, R., & Eichler, M. (2003). Causality and graphical models in time series analysis. In P. J. Green, N. L. Hjort & S. Richardson (Eds.), *Highly structured stochastic systems* (pp. 115–137). Oxford: Oxford University Press.
- Dahlhaus, R., Eichler, M., & Sandkühler, J. (1997). *Identification of synaptic connections in neural ensembles by graphical models*. Maastricht: Maastricht University.

- Didelez, V. (2000). *Graphical models for event history analysis based on local independence*. Universität Dortmund: Dortmund.
- Didelez, V. (2007). Graphical models for composable finite Markov processes. *Scandinavian Journal of Statistics*, 34, 169–185.
- Didelez, V. (2008). Graphical models for marked point processes based on local independence. *Journal of the Royal Statistical Society, Series B*, 70, 245–264.
- Didelez, V. (2011). Graphical models for stochastic processes. In P. J. Green, N. L. Hjort & S. Richardson (Eds.), *Highly Structured Stochastic Systems* (pp. 138–141). Oxford: Oxford University Press.
- Donat, R., Bouillaut, L., Aknin, P., & Leray, P. (2008). A dynamic graphical model to represent complex survival distributions. In T. Bedford, J. Quigley, L. Walls, B. Alkali, A. Daneshkhah, & G. Hardman (Eds.), *Advances in mathematical modelling for reliability* (pp. 17–24). Amsterdam: IOS Press.
- Donat, R., Leray, P., Bouillaut, L., & Aknin, P. (2010). A dynamic Bayesian network to represent discrete duration models. *Neurocomputing*, 73, 570–577.
- Dreassi, E., & Gottard, A. (2007). A Bayesian approach to model interdependent event histories by graphical models. *Statistical Methods and Applications*, 16, 39–49.
- Edwards, D. I. (2000). *Introduction to graphical modelling*. Berlin: Springer.
- Eichler, M. (1999). *Graphical models in time series analysis*. Heidelberg: Universität Heidelberg.
- Eichler, M. (2010). Graphical Gaussian modelling of multivariate time series with latent variables. *Journal of Machine Learning Research - Proceedings Track*, 9, 193–200.
- Eichler, M. (2012). Graphical modelling of multivariate time series. *Probability Theory and Related Fields*, 153, 233–268.
- Eichler, M. (2013). Causal inference with multiple time series: Principles and problems. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 371(1997), 20110613.
- Eichler, M., & Didelez, V. (2010). On Granger-causality and the effect of interventions in time series. *Lifetime Data Analysis*, 16, 3–32.
- Eichler, M., Dahlhaus, R., & Sandkühler, J. (2003). Partial correlation analysis for the identification of synaptic connections. *Biological Cybernetics*, 89, 289–302.
- El-hay, T., Friedman, N., Koller, D., & Kupferman, R. (2006). Continuous time Markov networks. In R. Dechter & T. Richardson (Eds.), *Proceedings of the Twenty-Second Conference Annual Conference on Uncertainty in Artificial Intelligence (UAI-06)* (pp. 155–164). AUAI Press.
- Feiler, S., Müller, K., Müller, A., Dahlhaus, R., & Eich, W. (2005). Using interaction graphs for analysing the therapy process. *Psychotherapy and Psychosomatics*, 74, 93–99.
- Florens, J.-P., & Fougere, D. (1996). Noncausality in continuous time. *Econometrica*, 64, 1195–1212.
- Fosen, J., Borgan, Ø., Weedon-Fekjær, H., & Aalen, O. O. (2006). Dynamic analysis of recurrent event data using the additive hazard model. *Biometrical Journal*, 48, 381–398.
- Fried, R., Didelez, V., & Lanius, V. (2004). Partial correlation graphs and dynamic latent variables for physiological time series. In D. Baier & K.-D. Wernecke (Eds.), *Innovations in classification, data science, and information systems* (pp. 259–266). Berlin: Springer.
- Gao, W., & Tian, Z. (2010). Latent ancestral graph of structure vector autoregressive models. *Journal of Systems Engineering and Electronics*, 21, 233–238.
- Gather, U., Imhoff, M., & Fried, R. (2002). Graphical models for multivariate time series from intensive care monitoring. *Statistics in Medicine*, 21, 2685–2701.
- Gottard, A. (2007). On the inclusion of bivariate marked point processes in graphical models. *Metrika*, 66, 269–287.
- Gottard, A., & Rampichini, C. (2007). Chain graphs for multilevel models. *Statistics and Probability Letters*, 77, 312–318.
- Granger, C. (1969). Investigating causal relations by econometric models and cross-special methods. *Econometrica*, 37, 424–438.
- Kolaczyk, E. D. (2009). *Statistical analysis of network data: methods and models*. Berlin: Springer.
- Koller, D., & Friedman, N. (2010). *Probabilistic graphical models*. Cambridge, MA: MIT Press.

- Lauritzen, S. (1996). *Graphical models*. Oxford: Oxford University Press.
- Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin: Springer.
- Lynggaard, H., & Walther, K. H. (1993). *Dynamic modelling with mixed graphical association models*. Aalborg: Aalborg University.
- Martinussen, T. (2010). Dynamic path analysis for event time data: Large sample properties and inference. *Lifetime Data Analysis*, 16, 85–101.
- Marttinen, P., & Corander, J. (2009). Bayesian learning of graphical vector autoregressions with unequal lag-lengths. *Machine Learning*, 75, 217–243.
- Meurk, C. S., Brown, J. A., & Reale, M. (2007). Graphical modelling of ecological time series data. In L. Oxley & D. Kulasiri (Eds.), *MODSIM 2007 International Congress on Modelling and Simulation. Modelling and Simulation Society of Australia and New Zealand* (pp. 1393–1398).
- Moneta, A. (2008). Graphical causal models and VARs: An empirical assessment of the real business cycles hypothesis. *Empirical Economics*, 35, 275–300.
- Murphy, K. (2002). *Dynamic Bayesian networks: Representation, inference and learning*. Berkeley: University of California.
- Nodelman, U., Shelton, C. R., & Koller, D. (2002). Continuous time Bayesian networks. In A. Darwiche & N. Friedman (Eds.), *Proceedings of the 18th Conference on Uncertainty in Artificial Intelligence*. San Francisco: Morgan Kaufmann.
- Nodelman, U., Shelton, C. R., & Koller, D. (2003). Learning continuous time Bayesian networks. In C. Meek & U. Kjærulff (Eds.), *Proceedings of the 19th Conference on Uncertainty in Artificial Intelligence* (pp. 451–458). San Francisco: Morgan Kaufmann.
- Oxley, L., Reale, M., & Wilson, G. T. (2008). *Constructing structural VAR models with conditional independence graphs*. Technical Report, University of Canterbury, Canterbury.
- Pearl, J. (1988). *Probabilistic reasoning in intelligent systems: Networks of plausible inference*. San Francisco: Morgan Kaufmann Publishers Inc.
- Penny, R. N., & Reale, M. (2004). Using graphical modelling in official statistics. *Quaderni di Statistica*, 6, 31–48.
- Queen, C. M., & Smith, J. Q. (1992). Dynamic graphical models. In J. M. Bernardo, J. O. Berger, A. P. Dawid & A. F. M. Smith (Eds.), *Bayesian Statistics 4. Proceedings of the 4th Valencia International Meeting* (pp. 741–751). Oxford: Oxford University Press.
- Queen, C. M., & Smith, J. Q. (1993). Multiregression dynamic models. *Australian and New Zealand Journal of Statistics*, 49(3), 221–239.
- Shelton, C. R., Fan, Y., Lam, W., Lee, J., & Xu, J. (2010). Continuous time Bayesian network reasoning and learning engine. *Journal of Machine Learning Research*, 11, 1137–1140.
- Schweder, T. (1970). Composable Markov processes. *Journal of Applied Probability*, 7, 400–410.
- Songsiri, J., Dahl, J., & Vandenberghe, L. (2010). Graphical models of autoregressive processes. In D. P. Paloma & Y. C. Eldar (Eds.), *Convex optimization in signal processing and communications* (pp. 89–116). Cambridge: Cambridge University Press.
- Spirtes, P. (2000). *Causation, prediction, and search*. Cambridge, MA: MIT Press.
- Talih, M. (2003). *Markov random fields on time-varying graphs, with an application to portfolio selection*. Yale University: Yale.
- Talih, M., & Hengartner, N. (2005). Structural learning with time-varying components: Tracking the cross-section of financial time series. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 321–341.
- Whittacker, J. C. (2008). *Graphical models in applied multivariate statistics*. New York: Wiley.
- Wild, B., Eichler, M., Friederich, H. C., Hartmann, M., Zipfel, S., & Herzog, W. (2010). A graphical vector autoregressive modelling approach to the analysis of electronic diary data. *BMC Medical Research Methodology*, 10, 28.
- Xuan, X., & Murphy, K. (2007). Modeling changing dependency structure in multivariate time series. In Z. Ghahramani (Ed.), *Proceedings of the 24th International Conference on Machine Learning* (pp. 1055–1062). New York: ACM.