Safety and Reliability – Bedford & van Gelder (eds)
© 2003 Swets & Zeitlinger, Lisse, ISBN 90 5809 551 7

Graphical models for the evaluation of multisite temperature forecasts: comparison of vines and independence graphs

U. Callies
GKSS Research Centre, Geesthacht, Germany

D. Kurowicka & R.M. Cooke
Delft University of Technology, Delft, Netherlands

ABSTRACT: Vine and independence graphs are employed to extract conditional independence relations from multivariate meteorological data so as to construct a simple graphical model which adequately represents the interrelationships between observations and corresponding model results at different sites. The independence graph approach identifies partial correlations of maximal order. Statistically negligible partial correlations are set to zero. Iterative proportional fitting is used to find a maximum likelihood distribution satisfying the stipulated zero partial correlations. The deviance between the fitted distribution and the original distribution measures the goodness of fit. The vine approach constructs a regular vine in which negligible partial correlations are set to zero. No proportional fitting is required. Again, deviance is used to measure goodness of fit. The connection between vines and continuous belief nets, where an arc from node i to j is associated with a (conditional) rank correlation between i and j is presented.

1 INTRODUCTION

This paper investigates interaction structures between observed December mean temperatures at four European stations and corresponding 'forecasts' of a regression model based on sea-level pressure (SLP) fields as explanatory variables. With temperatures at different stations being correlated each local forecast can be expected to be informative to a certain extent about observations at other locations as well. The following three questions naturally arise: 1) Is the local forecast really more informative about local conditions than forecasts delivered for other sites are? 2) Knowing the local forecast, can additional information be obtained from consulting other forecasts? 3) If other forecasts are found to contain additional information, how much of this incremental information can be attributed to correlations between observations at different sites?

The statistical notion of conditional independence allows these questions to be formalized. Callies (2000) contrasted the concepts of conditional independence and sufficiency, which both refer to the complete statistical information embodied in the joint distribution of forecasts and subsequent observations. Brooks & Doswell (1996) contrasted the distribution-oriented approach to forecast verification with the more conventional approach of calculating global measures of correspondence between forecasts and observations. The purpose of this study is to explore the use of different techniques of graphical modelling for forecast evaluation. Throughout the paper a multinormal joint distribution of observations and corresponding forecasts is assumed.

2 DATA

Observations of local December mean temperatures at 14 European stations were taken from the World Monthly Station Climatology (WMSC) of the National Center for Atmospheric Research (NCAR). Corresponding diagnostic "forecasts" between 1900 and 1993 were produced by regressing local temperatures on monthly mean regional-scale atmospheric sea-level pressure distributions as represented by $5° \times 5°$ analyses (Trenberth & Paolino 1980) at 60 gridpoints covering the region 40°N to 64°N and 20°W to 25°E. Data are made available through NCAR. The regression scheme was calibrated for 1960–1980. This period was excluded from the analysis below.

Prior to regression, both regional and local data were filtered by standard principal component analysis

to reduce the number of degrees of freedom and to avoid overfitting. In both data sets only four degrees of freedom were retained. Accordingly the complete amount of information contained in the forecasts is available after selecting any four of the predictions. The following (8×8) sample correlation matrix, S, embodies all information about the interactions between observations, θ, and the corresponding forecasts, F, at the four stations Geneva (G), Innsbruck (I), Budapest (B) and Copenhagen (K):

$$S = \begin{pmatrix} 1 & .35 & .50 & .49 & .68 & .38 & .50 & .59 \\ & 1 & .79 & .69 & .12 & .64 & .62 & .49 \\ & & 1 & .72 & .18 & .61 & .58 & .43 \\ & & & 1 & .05 & .46 & .47 & .43 \\ & & & & 1 & .33 & .51 & .71 \\ & & & & & 1 & .97 & .77 \\ & & & & & & 1 & .90 \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} \theta^K \\ \theta^G \\ \theta^I \\ \theta^B \\ F^K \\ F^G \\ F^I \\ F^B \end{matrix} \quad (1)$$

The values for the observation–forecast correlations vary from 0.43 for Budapest to 0.68 for Copenhagen.

3 FITTING INDEPENDENCE GRAPHS

The graphical modelling approach elaborated by Whittaker (1990) identifies genuine variable interactions that are not mediated through any third variable in the data set. For this purpose maximum order pairwise partial correlations given all remaining variables in the data set are analysed. The two independence graphs in Figure 1, for instance, are made up by only 11 edges. In a process of recursive model simplification, which starts from the saturated graph, 17 edges have been discarded by setting corresponding partial correlations to zero. The two different graphs represent structural ambiguity caused by the very high correlation (0.97) of F^G and F^I. The maximum likelihood correlation matrix $V^{(a)}$ in Eq. (3) corresponds with graph (a) in Figure 1. It arises from a constrained fit for the non-zero partial correlations and minimises the following entropy measure (deviance) of the distance between S and V as function of V (N denotes sample size):

$$\text{dev}(S, V) = N\,[\text{tr}\,(SV^{-1} - I) - \log\det(SV^{-1})] \quad (2)$$

The graphical constraint manifests itself by the fact that all elements of $(V^{(a)})^{-1}$ corresponding with missing links in the graph must be zero. It turns out that only elements (given in bold type) of $V^{(a)}$ in the same position differ from the sample correlation matrix S (Whittaker 1990):

$$V^{(a)} = \begin{pmatrix} 1 & \mathbf{.38} & \mathbf{.41} & .49 & .68 & \mathbf{.38} & \mathbf{.48} & \mathbf{.58} \\ & 1 & .79 & \mathbf{.56} & \mathbf{.18} & .64 & \mathbf{.62} & \mathbf{.48} \\ & & 1 & .72 & \mathbf{.12} & \mathbf{.50} & \mathbf{.48} & \mathbf{.37} \\ & & & 1 & .05 & \mathbf{.35} & \mathbf{.32} & \mathbf{.23} \\ & & & & 1 & .33 & .51 & \mathbf{.72} \\ & & & & & 1 & .97 & .77 \\ & & & & & & 1 & .90 \\ & & & & & & & 1 \end{pmatrix} \begin{matrix} \theta^K \\ \theta^G \\ \theta^I \\ \theta^B \\ F^K \\ F^G \\ F^I \\ F^B \end{matrix} \quad (3)$$

The total deviance of graph (a) in Figure 1 is 41.9 for sample size 63 (years with incomplete data have been discarded). All but three of these incremental deviances are larger than the total deviance of 41.9 due to 17 missing links. Therefore, the truncated interaction

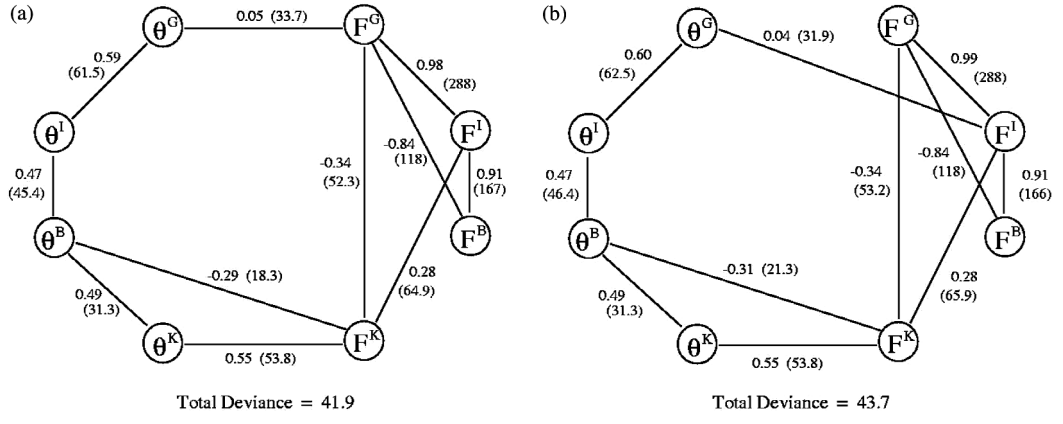


Figure 1. Two conditional independence graphs for observations, θ, and corresponding forecasts, F, at four locations. 17 edges have been discarded. Numbers indicate partial correlations and EEDs (in parentheses).

structure portrayed by the graph seems to provide a reasonable model for the data.

For each link the corresponding partial correlation and the increase of the graph's total deviance, dev (S, V) that would arise from the link's omission (Edge Exclusion Deviance, EED) are given. Note that the partial correlations θ^G–F^G in graph (a) and θ^G–F^I in graph (b), respectively, are very small due to the high correlation between F^G and F^I (cf. matrix S). Nevertheless large deviances (33.7 and 31.9, respectively) indicate that the corresponding links are relevant.

The matrices $V^{(a)}$ (cf. Eq. (3)) and $V^{(b)}$ have been specified numerically using iterative proportional fitting (Whittaker 1990). Generally, analytic solutions of the maximum likelihood problem are not available. This is related to the fact that partial correlations in the independence graph cannot be prescribed independently without violating the positive definiteness of the correlation matrix. Graphical vines are an alternative representation of correlation matrices, for which this is possible.

4 FITTING C- AND D-VINES

Vines have been introduced recently by Bedford & Cooke (2001, 2002). Basically, a *vine* on N variables is a nested set of trees, where the edges of tree j are the nodes of tree $j + 1$; $j = 1, \ldots, N - 2$, and each tree has the maximum number of edges (Cooke 1997). A *regular* vine on N variables is a vine in which two edges in tree j are joined by an edge in tree $j + 1$ only if these edges share a common node, $j = 1, \ldots, N - 2$. There are $(N - 1) + (N - 2) + \cdots + 1 = N(N - 1)/2$ edges in a regular vine on N variables. Figure 2 shows a regular vine on 5 variables. The four nested trees are distinguished by the line style of the edges; tree 1 has solid lines, tree 2 has dashed lines, etc. The conditioned (before |) and conditioning (after |) sets associated with each edge are determined as follows: the variables reachable from a given edge are called the constraint set of that edge. When two edges are joined by an edge of the next tree, the intersection of the respective constraint sets are the conditioning variables, and the symmetric difference of the constraint sets are the conditioned variables. The regularity condition ensures that the symmetric difference of the constraint sets always contains two variables. Note that each pair of variables occurs once as conditioned variables.

We recall two generic vines, the *D-vine* D(1, 2, …, n) and C-vine C(1, 2, …, n), shown on Figures 2 and 3. Both C and D-vine are determined by the choice of the first tree.

In contrast with independence graphs vines are made up by partial correlations of varying order (the order of a partial correlation is the number of conditioning variables). Generally, the lower the order of an independence relation (i.e. partial correlation equals zero), the stronger this relation is. Whereas in an independence graph all partial correlations are all order $N - 2$, in a regular vine there is only one correlation of order $N - 2$, and $N - 1$ correlations of order zero. Hence, setting K partial correlations in a regular vine equal to zero imposes stronger independence than setting K partial correlations of maximal order equal to zero, as with independence graphs. Vines are derived from tree structures and therefore (again in contrast with independence graphs) assume a particular ordering of variables.

There is a one to one relationship between correlation matrices and partial correlation–vine specification so that the deviance defined in Eq. (2) can be used for the assessment of vines as well. Since the deletion of any partial correlation from a vine does not change other partial correlations no proportional fitting algorithm is needed to guaranty consistent correlation

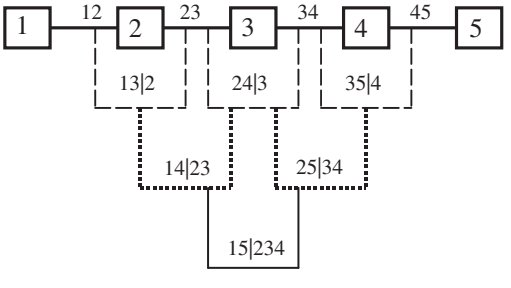


Figure 2. The D-Vine on 5 variables D(1, 2, 3, 4, 5) showing conditioned and conditioning sets.

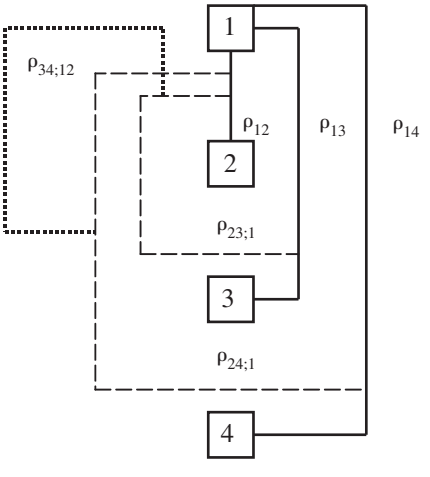


Figure 3. The C-vine on 4 variables C(1, 2, 3, 4) with associated to the edges partial correlations.

matrices. However, additional efforts are needed to search for that permutation (ordering) of the set of variables, which allows for the most effective model simplification.

Generally to try all models with different degrees of simplification based on different orders of the variables is a tremendous task. Here it was assumed *a priori* that the number of edges being discarded from the vine should be 17 as it is in Figure 1. Then, for each of 40320 (=8!) permutations of variables EEDs of all individual links were calculated and the edge with the minimum EED was excluded from the graph (for the symmetric D-vines it was sufficient to test 8!/2 orderings of variables). This procedure was iterated until 17 links had been discarded. Even though other partial correlations are not affected by the removal of individual links, this is not true for the EEDs of the remaining links. Thus a strict optimisation would have needed the testing of all possible combinations of edges rather than using the sequential edge exclusion scheme. This, however, would have been too computationally demanding.

Figure 4 depicts the total deviances, which emerge from the 60 most successful orders of the eight variables when using C-vines and D-vines, respectively. In our example D-vines allow for a better fit than C-vines (note that the two panels are scaled differently), the deviance of the best D-vine is smaller than the deviances of the independence models in Figure 1. In addition, the optimum choice of the ordering of variables is clearly better determined for D-vines. For C-vines the deviances tend to be more clustered and show a very flat minimum.

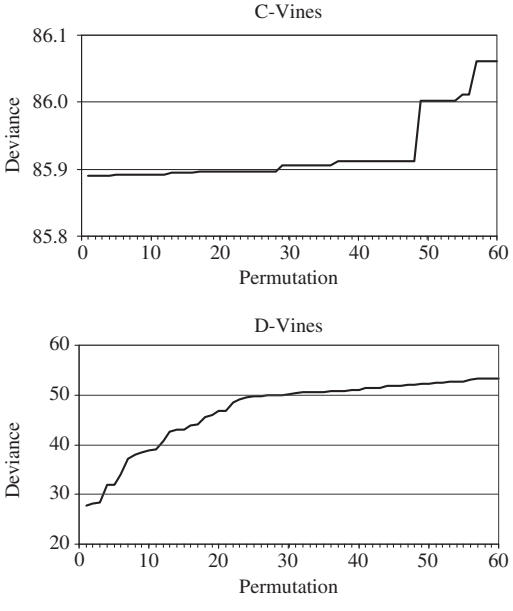


Figure 4. Deviances of graphical vine models with 17 missing edges as function of the 60 most successful permutations of variables.

Figure 5 depicts the best fitting C-vine (one among four options with the same deviance) and D-vine, respectively, both being made up by 11 edges. It should be noted that the notation used in the graphs differs from the notations that was used in Figures 2 and 3. Instead of emphasizing the nature of a vine as a nested set of trees, links in Figure 5 indicate pairs of correlated variables. Numbers between variables specify marginal correlations. The C-vine allows also for marginal correlations between non-neighboring variables. These are indicated by dashed lines above the variable names. Partial correlations indicated below the variable names are to be interpreted differently for the two vines. In the C-vine partial correlations are conditioned on all variables to the left of the first variable connected by the respective solid line. In the D-vine the set of conditioning variables comprises all variables in between the two correlated variables.

Similar to the independence graphs both the C-vine and the D-vine employ three edges for establishing the link between observations and forecasts. Like the independence graph (b) in Figure 1 both vines include edges F^K–θ^K and F^I–θ^G.

In particular the D-vine tends to establish links conditioned on no or a small number of variables. However, (in contrast with the C-vine) no marginal correlation between observations and forecasts is maintained. The alignment of observations is the same as in the independence graphs.

Generally edges corresponding with strong partial correlations have been retained. There are, however, exceptions. In the C-vine the marginal correlation

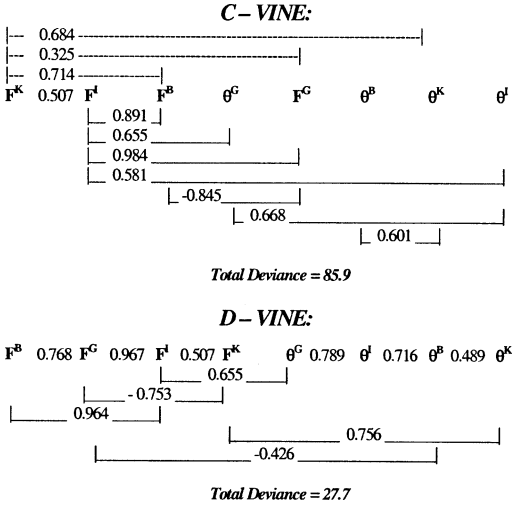


Figure 5. Optimized vines with 11 links (see text).

Table 1. Edge exclusion (inclusion) deviance for the C-vine in Figure 5. Existing links are written in bold. The first column ("Level") gives the number of variables that are held fixed for the respective partial correlation. The last column specifies the order in which the 17 links have been removed.

Level		EED	EID	Corr	
0	$\mathbf{F^K–F^I}$	**3254.0**		**0.507**	
0	$\mathbf{F^K–F^B}$	**651.3**		**0.714**	
0	$F^K–\theta^G$		−2.36		(7)
0	$\mathbf{F^K–F^G}$	**760.8**		**0.325**	
0	$F^K–\theta^B$		0.24		(6)
0	$\mathbf{F^K–\theta^K}$	**55.8**		**0.684**	
0	$F^K–\theta^I$		0.12		(8)
1	$\mathbf{F^I–F^B}$	**272.0**		**0.891**	
1	$\mathbf{F^I–\theta^G}$	**60.3**		**0.655**	
1	$\mathbf{F^I–F^G}$	**477.0**		**0.984**	
1	$F^I–\theta^B$		14.31		(17)
1	$F^I–\theta^K$		−9.29		(12)
1	$\mathbf{F^I–\theta^I}$	**43.8**		**0.581**	
2	$F^B–\theta^G$		−0.72		(1)
2	$\mathbf{F^B–F^G}$	**78.9**		**−0.845**	
2	$F^B–\theta^B$		6.14		(15)
2	$F^B–\theta^K$		1.59		(4)
2	$F^B–\theta^I$		3.77		(11)
3	$\theta^G–F^G$		0.30		(3)
3	$\theta^G–\theta^B$		0.91		(16)
3	$\theta^G–\theta^K$		−8.71		(13)
3	$\mathbf{\theta^G–\theta^I}$	**37.5**		**0.668**	
4	$F^G–\theta^B$		−0.76		(5)
4	$F^G–\theta^K$		1.97		(10)
4	$F^G–\theta^I$		0.09		(2)
5	$\mathbf{\theta^B–\theta^K}$	**31.0**		**0.601**	
5	$\theta^B–\theta^I$		−2.02		(14)
6	$\theta^K–\theta^I$		4.70		(9)

between forecasts F^K and F^G, 0.325, has been retained despite of its relatively small value. According to Table 1 the EED of this edge, 761, is much higher than, for instance, the EED of the link $F^I–\theta^G$, 60, although the partial correlation for the latter pair of variables is 0.655. We may conclude once more that partial correlations do not necessarily provide information about the statistical significance of links in graphical models. A surprising result is that in contrast with independence graphs for vines the decrease of total deviance resulting from a link's inclusion (Edge inclusion deviance, EID) can be negative telling that the removal of a link would result in an improvement (!) of the model.

The upper panel of Figure 6 depicts for the two vines in Figure 5 the total deviances as functions of the number of discarded links. Obviously the total fit of the C-vine could be much improved by re-establishing the link $F^I–\theta^B$, which has been removed last (according to Table 1 the total deviance would decrease by a value of 14.3). When doing so, however, the EID of the link $\theta^G–\theta^B$ turns out to increase from its former value of 0.91

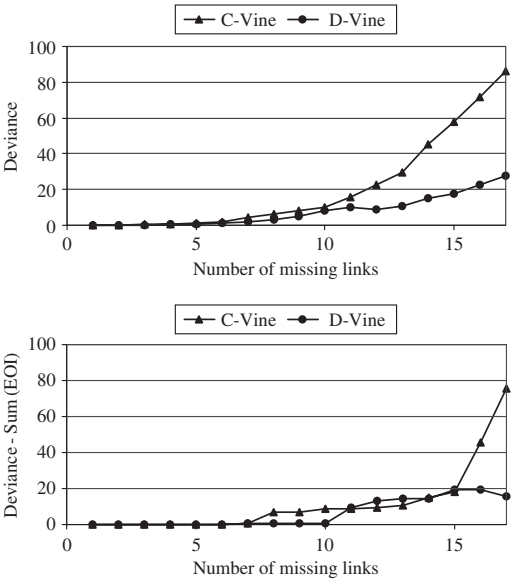


Figure 6. Upper panel: Total deviance as function of the number of edges discarded from the optimal C- and D-vine, respectively. Lower panel: Differences between total deviances and sums of individual edge inclusion deviances.

(cf. Table 1) to 13.4. The lower panel of Figure 6 reveals that for both kinds of vines the differences between the total deviances and the sums of individual EIDs increase with model truncation, i.e. the assessment of interaction structures becomes less local. Beyond 15 missing links, this non-locality increases dramatically for the C-vine while it still remains on the same level for the D-vine.

5 ASSOCIATING A D-VINEWITH A BELIEF NET

The idea of parameterizing the correlation structure with algebraically independent partial correlations, as in a regular vine, can also be extended to belief nets. In fact, the idea of capturing "influence" via conditional sampling leads to a natural homomorphism between regular vines and belief nets, which we now describe.

A belief net is a directed acyclic graph with nodes representing random variables and arcs representing "influence". In continuous belief nets we associate nodes in a belief net with continuous univariate random variables, and arcs with (conditional) rank correlations (Kurowicka & Cooke 2002). The following steps indicate how translate this into partial correlations on vines, and hence compute the entire correlation structure. We number the nodes in a belief net 1, …, n.

Step 1 *Sampling order*

We construct a *sampling order* for the nodes, that is, an ordering such that all ancestors of node i appear before i in the ordering. A sampling order begins with a source node and ends with a sink node. Of course the sampling order is not in general unique.

Step 2 *Factorize joint*

We first factorize the joint in the standard way following the sampling order. If the sampling order is 1, 2, …, n, write:

$$P(1, \ldots, n) = P(1)P(2|1)P(3|12)\ldots P(n|1, 2, \ldots, n-1).$$

Next, we underscore those nodes in each condition which are not necessary in sampling the conditioned variable. This uses (some of) the conditional independence relations in the belief net. For each term, we order the conditioning variables, i.e. the variables right of the "|", such that the underscored variables (if any) appear right-most and the non-underscored variables left-most.

Step 3 *Quantify D-vine for node n*

Suppose the last term looks like:

$$P(n \,|\, n-1, n-3, \ldots, \underline{n-2, 3, 2, 1}).$$

Construct the D-vine with the nodes in the order in which they appear, starting with n (left) and ending with the last underscored node (if any).

If the D-vine D(n − 1, n − 3, …, 1) is given, the D-vine D(n, n − 1, …, 1) can be obtained by adding the edges:

$$(n, n-1), (n, n-3 \,|\, n-1), \ldots, (n, \underline{1} \,|\, n-1, \ldots, \underline{2}).$$

For any underscored node $\underline{k}$, we have

$(n \perp \underline{k} \mid$ all non-underscored nodes ∪ any subset of underscored's not including k).

The conditional correlation between n and an underscored node will be zero.

For any non-underscored node j, the bivariate distribution

(n, j | non-underscored nodes before j)

will have to be assessed. The conditioned variables (n, j) correspond to an arc in the belief net.

Write these conditional bivariates next to the corresponding arcs in the belief net. Note that we can write the (conditional) correlations associated with the incoming arcs for node n without actually drawing the D-vines. If the last factor is $P(5|1, 2, 3, \underline{4})$, we have incoming arcs (5, 1), (5, 2) and (5, 3) which we associate with conditional correlations (5, 1), (5, 2|1) and (5, 3|12).

Step 4 *Quantify D-vine for node n − 1, for node n − 2 etc.*

Proceed as in step 3 for nodes 1, 2, …, n − 1. Notice that the order of these nodes need not be the same as in the previous step. Continue until we reach the D-vine D(1 2) or until the order doesn't change in smaller subvines, i.e., if for node 4 the D-vine is D(4 3 2 1) and for node 3 it is D(3 2 1) then we can stop with node 4; or better, we can quantify the vine D(3 2 1) as a subvine of D(4 3 2 1).

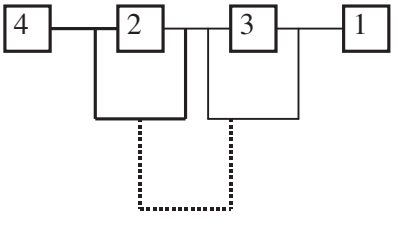
Step 5 *Construct partial correlation D-vine (1, …, n)*

As a result of steps 1–4 each arc in the belief net is associated with a (conditional) bivariate distribution. These conditional distributions do not necessarily agree with the edges in D(1, …, n) since the orders in the different steps may be different. However, if the conditional bivariate distributions are given by partial correlations, then given D(1, …, k) we can compute $D(\pi(1)\ldots\pi(k))$ where $\pi \in k!$ Is a permutation of 1, …, k. Repeatedly using this fact, we compute the partial correlations for D(1, …, n).

Since the values of conditional correlations on regular (sub)vines are algebraically independent and uniquely determine the correlation (sub)structure, the above algorithm leads to an assignment of correlations and conditional correlations to arcs in a belief net which are algebraically independent and which, together with the "zero edges" in the corresponding D-vines, uniquely determine the correlation matrix.

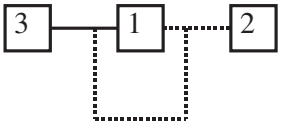
EXAMPLE 1

Sampling order: 1234
Factorization: $P(1)P(2|\underline{1})P(3|1\underline{2})P(4|23\underline{1})$
D-vine 4231:

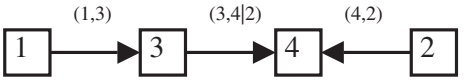


The dotted edge has partial correlation zero, the bold edges correspond to (4, 2) and (4, 3|2). These are written on the belief net and must be assessed.

We now consider the term P(3 | 1 2). The order is different than for the term P(4 | 2 3 1). We construct D(3 1 2):

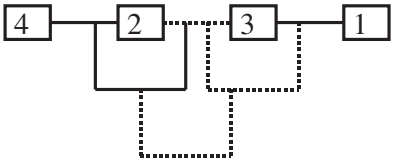


The dotted edges have partial correlation zero, the bold edge must be assessed, it is added to the belief net. The belief net is now quantified:



With the partial correlations in D(3 1 2) we compute using the recursive relations D(2 3 1). In fact, we find $\rho_{23} = \rho_{23,1} = 0$.

We now have D(4 2 3 1) which corresponds to the belief net:



The distribution having specified univariate margins and satisfying rank correlation structure of the regular vine above can be obtained and sampled (Kurowicka & Cooke 2001).

6 ASSOCIATING A BELIEF NET WITH A D-VINE

Starting with a D-vine, we associate a belief net by identifying the sampling order of the belief net with the ordering of the vine. Let us assume that D(n, n − 1, …, 1) is given, hence the sampling order is 1, 2, …, n. Start with the variable 1, draw arcs to all variables j such that partial correlation $\rho_{1j;2,\ldots,j-1}$ on the D-vine is not equal to zero. Proceed the same way with the next variable in the ordering, that is 2 etc.

We illustrate this procedure using the D-vine in Figure 5. The ordering of this vine is: θ^K, θ^B, θ^I, θ^G, F^K, F^I, F^G, F^B. We start with θ^K. Both the marginal correlation θ^K and θ^B and the partial correlation θ^K and F^K given θ^B, θ^I, θ^G are non-zero. Therefore we

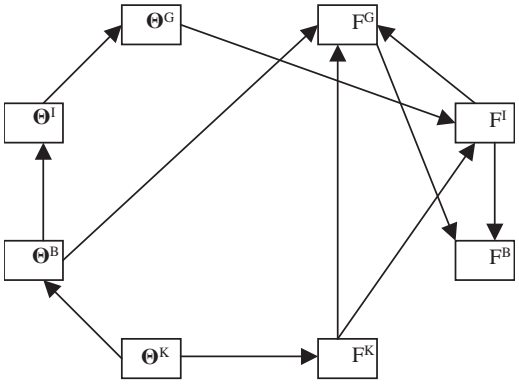


Figure 7. A belief net corresponding to the D-vine in Figure 5.

draw arcs from θ^K to θ^B and from θ^K to F^K. Similarly arcs must be drawn from θ^B to θ^I and from θ^B to F^G etc. Following this procedure the belief net in Figure 7 can be created that corresponds to the D-vine in Figure 5.

The graphs in Figure 1 and 7 are quite similar. Only the link θ^B–F^G is surprising since it was discarded in both cases in Figure 1. It must be emphasized, however, that in order to derive an undirected conditional independence graph from a directed belief net, generally new links must be added connecting all variables with a common child (moral graph; cf. Lauritzen & Spiegelhalter 1988, Whittaker 1990). This formal rule does not include a statement about whether or not the new links are statistically relevant, so that that the comparison of Figures 1 and 7 is not straightforward.

7 DISCUSSION

Numerical climate models are important tools for climate change studies. A general problem, however, is that global atmospheric models are not able to resolve the local scales at which most users require information so that statistical models must be employed to relate large-scale climate model output to regional scale observations (cf. Zorita & von Storch 1999). The temperature forecasts analyzed in this paper have been produced by such a “downscaling” method using observed large scale atmospheric pressure data as explanatory variables. A main objective of the study was to explore the application of two different techniques of graphical modeling for the assessment of the amount of genuine local information conveyed by the results of downscaling.

In the specific example it turned out that, for instance, the partial correlation between predictions for the stations Budapest, F^B, or Innsbruck, F^I, and

observations at any of the four locations of interest are quasi zero given the values of forecasts for the two stations Geneva, F^G, and Copenhagen, F^K. This means that regression on forecasts F^G and F^K could be used as a surrogate for the forecasts delivered for Innsbruck or Budapest. In this sense the latter two forecasts, F^I and F^B, could be skipped without significant loss of information.

Another conclusion from the graphical independence model was that none of the four forecasts contained a significant amount of incremental information about temperature in Innsbruck as soon as the temperatures at Geneva and Budapest are known. This exemplifies how the analysis of interaction patterns can be used to identify sets of local stations, for which the downscaling scheme provides relevant independent information. Given the forecasts for these key stations observations at other stations might be simulated by an interpolating regression scheme.

A by-product of the graphical analysis is an estimate of the dimensionality of the interaction between the set of observations and the set of forecasts. Such information could also be achieved by considering, for instance, the number of pairs of correlated patterns that turn out to be relevant in conventional canonical correlation analysis. The main advantage of the graphical methods studied here is that they refer to the original set of variables and do not introduce derived artificial variables. Generally the exclusion of irrelevant derived variables (low correlation and/or explained variance) does not result in an reduction of the number of original variables the analysis is based on.

Graphical modeling intends to eliminate unimportant relations that may be present in empirical data in order to define parsimonious models with maximal explanatory power. Whittaker's method of independence graphs is currently the state of the art. The results of this study indicate that vines *might* offer improvements by imposing stronger independence relations (i.e. zeroing partial correlations of lower order) while deviating less from the original distribution. However, the computational burden is formidable.

Independence graphs discussed by Whittaker (1990) are defined by sets of pairwise conditional independence restrictions of maximum order. Non-zero partial correlations cannot be prescribed but must be estimated from data in a consistent way. In contrast, partial correlations in a vine can be prescribed independently so that the maximum likelihood estimate of non-zero partial correlations is trivial (simply choose the observed value). However, the occurrence of negative EIDs only for vines seems to be related to this fact that edges of vines are defined without adjustment of other links.

The main difficulty when fitting vines lies in the particular tree structure that each vine imposes onto the set of variables. It is an open question whether methods can be designed that circumvent the need for testing all possible permutations of variables. Fitting independence graphs does not impose any ordering onto the set of variables, which makes fitting independence graphs computationally much less demanding.

The notation of D-vines is in closer correspondence with independence graphs than the notation of C-vines. Both use the very intuitive notation of context variables put in between those variables, the partial correlation of which is calculated. For instance, the orderings of observations in the graphical model and in the D-vine were found to be identical. In the present example D-vines performed clearly better than C-vines. This is an interesting result because due to its asymmetric topology for C-vines the number of different options for ordering variables is twice as large as for D-vines.

Fitting a C-vine, however, might be advantageous when a particular variable is known to be a key variable that governs variable interactions in the data set (maximum average correlation with other variables?). In such a situation one may decide to locate it at the root of the canonical vine.

All graphical models discussed in this study portrayed the mechanism of interactions between observations and forecasts in a consistent way. All methods used three edges for establishing interaction between observations and forecasts. Both the independence graph and the D-vine concluded that forecasts do not contain a significant amount of site specific information about θ^I, which cannot be replaced by using observations θ^B and θ^G as a surrogate. The C-vine with forecasts at its root has not the proper structure for portraying this fact.

ACKNOWLEDGEMENTS

For their open provision of data we are indebted to NCAR.

REFERENCES

Bedford, T. & Cooke, R. M. 2001. Probability density decomposition for conditionally dependent random variables modeled by vines. *Annals of Mathematics and Artificial Intelligence* 32: 245–268.

Bedford, T. & Cooke, R. M. 2002. Vines – a new graphical model for dependent random variables. *Annals of Statistics* 30(4): 1031–1068.

Brooks, H. E. & Doswell, C. A. 1996. A comparison of measure-oriented and distribution-oriented approaches to forecast verification. *Weather and Forecasting* 11: 288–303.

Cooke, R. M. 1997. Markov and entropy properties of tree and vines- dependent variables, Proceedings of the ASA Section of Bayesian Statistical Science.

Callies, U. 2000. Comparative forecast evaluation: Graphical Gaussian models and sufficiency relations. *Mon. Wea. Rev*. 128: 1912–1924.

Kurowicka, D. & Cooke, R. M. 2001. Conditional, Partial and Rank Correlation for Elliptical Copula, *Dependence Modeling in Uncertainty Analysis, Proc. of ESREL 2001*.

Kurowicka, D. & Cooke, R. M. 2002. The vine copula method for representing high dimensional dependent distributions: application to continuous belief nets. In E. Yücesan, C.-H. Chen, J.L. Snowdon & J.M. Charnes (eds), Proceedings of the 2002 Winter Simulation Conference, 263–269.

Lauritzen, S. L. & Spiegelhalter, D. J. 1988. Local computations with probabilities on graphical structures and their application to expert systems (with discussion). *J. Roy. Statist. Soc. B.* 50(2): 157–224.

Trenberth, K. E. & Paolino (jr.), D. A. 1980. The northern hemisphere SLP-dataset: trends, errors and discontinuities, *Mon. Wea. Rev.* 108: 855–872.

Whittaker, J. 1990. Graphical Models in applied multivariate statistics. John Wiley & Sons, Chichester.

Zorita, E. & von Storch, H. 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods, *J. Climate* 12: 2374–2489.