INTERNATIONAL JOURNAL OF CLIMATOLOGY
Int. J. Climatol. **29**: 1872–1880 (2009)
Published online 23 December 2008 in Wiley InterScience
(www.interscience.wiley.com) DOI: 10.1002/joc.1819

Short Communication

Spatial variability in correlation decay distance and influence on angular-distance weighting interpolation of daily precipitation over Europe

Nynke Hofstra[a,b]* and Mark New[a]
[a] *Oxford University Centre for the Environment, South Parks Road, Oxford, OX1 3QY, England, UK*
[b] *Environmental Systems Analysis Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands*

ABSTRACT: Angular-distance weighting (ADW) is a common approach for interpolation of an irregular network of meteorological observations to a regular grid. A widely used version of ADW employs the correlation decay distance (CDD) to (1) select stations that should contribute to each grid-point estimate and (2) define the distance component of the station weights. We show, for Europe, that the CDD of daily precipitation varies spatially, as well as by season and synoptic state, and is also anisotropic. However, ADW interpolation using CDDs that varies spatially by season or synoptic state yield only small improvements in interpolation skill, relative to the use of a fixed CDD across the entire domain. If CDDs are optimized through cross validation, a larger improvement in interpolation skill is achieved. Improvements are larger for the determination of the state of precipitation (wet/dry) than for the magnitude. These or other attempts to improve interpolation skill appear to be fundamentally limited by the available station network. Copyright © 2008 Royal Meteorological Society

KEY WORDS interpolation; precipitation; correlation; Europe

Received 27 November 2007; Revised 12 May 2008; Accepted 24 October 2008

1. Introduction

Gridded climate data derived from meteorological measurements are important for climate change detection (e.g. Karoly *et al.*, 2003), validation of regional and global climate models (Caesar *et al.*, 2006), and to drive many models used in global and regional change studies (e.g. hydrology, biodiversity, biogeochemical cycling). One of the most common methods used to interpolate from an irregular network of observations to a regular grid is angular-distance weighting (ADW). ADW has been used to interpolate climatic means (Legates and Willmott, 1990a,b) or monthly anomalies (New *et al.*, 2000), daily data (Caesar *et al.*, 2006) and extreme climate indices (Kiktev *et al.*, 2003; Alexander *et al.*, 2006). Various forms of ADW have been used; all have a common approach that an estimate at a particular point is a weighted average of nearby station data, where individual station weights are a function of inverse distance from the point to be estimated and angular isolation from other data points (Shepard, 1968; Willmott *et al.*, 1985; New *et al.*, 2000).

A common implementation of ADW (New *et al.*, 2000, henceforth NEW2000) uses the concept of correlation decay distance (CDD), also called correlation length scale or decorrelation length, for (1) selection of stations to average when estimating a grid value and (2) formulation of the inverse-distance component of the station weight. The CDD is defined as the distance where the correlation between one station and all other stations decays below $1/e$ (Briffa and Jones, 1993; Jones *et al.*, 1997, see Section 3). The search radius for the selection of stations used for the interpolation is set equal to the CDD, because the best result is expected when we use only stations that are correlated with the target grid point. The use of the CDD in the distance weighting is explained in Section 4.

The use of CDD in the NEW2000 formulation of ADW assumes that CDD is isotropic. Previous work has shown that for temperature and rainfall CDD can be anisotropic. Briffa and Jones (1993) and Jones *et al.* (1997) studied annual temperature and conclude that correlations decay much more rapidly in the meridional than zonal direction; they also found that in the Northern Hemisphere CDDs are lower at higher latitudes. On the other hand, Caesar *et al.* (2006) and Alexander *et al.* (2006), who examined daily temperature and daily temperature extremes respectively, conclude that CDDs are higher at higher latitudes, and that CDDs in summer

* Correspondence to: Nynke Hofstra, Environmental Systems Analysis Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands. E-mail: nynke.hofstra@wur.nl.

are lower than in winter. Osborn and Hulme (1997) found, in a study of daily rainfall over Europe, that CDDs are larger in the north and the west of Europe, and lowest in Northern Italy. Alexander *et al*. (2006) found for indices of daily precipitation extremes that CDDs are generally greater at lower latitude and, similarly to Osborn and Hulme (1997), that the CDD is lower in summer than in winter due to the larger percentage of frontal rain in winter.

Gridded datasets of daily climate by Alexander *et al*. (2006) and Caesar *et al*. (2006) were created using ADW interpolation with CDDs varied by latitude band and season. These studies however, did not fully explore the spatial variability in CDD within latitude bands or assess the relative improvement in interpolation skill arising from the use of a variable CDD. The purpose of this study is therefore twofold. We first make use of a new dataset of daily precipitation for meteorological stations over Europe to explore in detail the spatial pattern of CDD of daily rainfall as a function of seasonality, anisotropy and synoptic state. Having identified how CDDs vary as a function of these factors, we then explore the extent to which the use of varying CDDs influences the accuracy of ADW interpolation.

2. Data

We use a new dataset of meteorological station daily precipitation measurement, collated in collaboration with over 50 partners from European countries (Klok and Klein Tank, 2008) as a part of the EU FP6 ENSEMBLES project. The version of the dataset used for this study is a preliminary version containing 1768 stations (Figure 1); the station distribution is best over the Netherlands, and good over Switzerland, Ukraine, Belarus, the Baltic States and Portugal. The distribution is poor over Poland, the Balkans, Scandinavia and Northern Africa. The dataset covers the period 1950 or earlier to present, but for this study we use the period 1961–1990, because the data availability is best in this period. The data have been quality controlled, so potentially erroneous outliers have been removed.

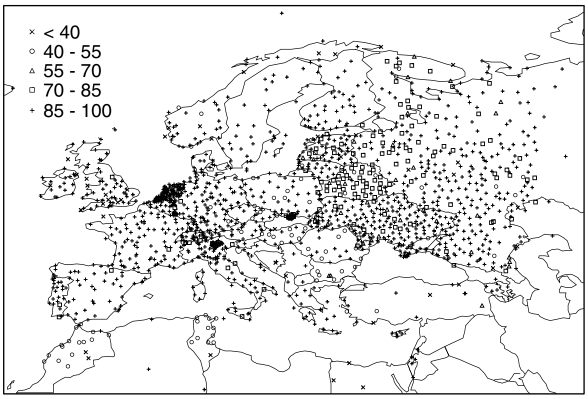

Figure 1. Station availability for the period 1961–1990. The symbols represent the percentage of available data for each station.

3. Variability in CDD

3.1. Calculation and subsetting of CDDs

CDDs can be calculated from a correlation matrix for all the stations in the dataset. For each station, correlations (r) with all other stations are extracted, and plotted against distance from the target station (x). An exponential decay function of the form is then fitted by least squares through the points.

$$r = e^{-x/CDD} \tag{1}$$

The CDD therefore corresponds to the distance where r equals 1/e, approximately the 0.05 significance level for r with large samples (Figure 2). A large CDD indicates that more distant stations retain a significant correlation (Briffa and Jones, 1993; Osborn and Hulme, 1997).

As a base case we calculate CDDs using all available data. We then calculate CDDs using subsets of the data as follows:

– *Seasonality*. Data are divided into four seasons, DJF, MAM, JJA, SON and separate CDDs are calculated at each station for each season.
– *Anisotropy*. Stations are subdivided into quadrants originating at the target station, and extending north, south, east and west (for example the northern quadrant extends from NW to NE). CDDs are then calculated separately for stations in the meridional and zonal pairs of these quadrants.
– *Synoptic state*. We use self organizing maps (SOMs) to identify 12 'archetypal' synoptic states from daily mean sea level pressure from the ERA40 reanalysis data (see Hewitson and Crane, 2002 for a description of the use of SOMs for synoptic classification). Each day is assigned to a SOM node (synoptic state number) and the CDDs are then calculated using station data only from days associated with each synoptic state. In

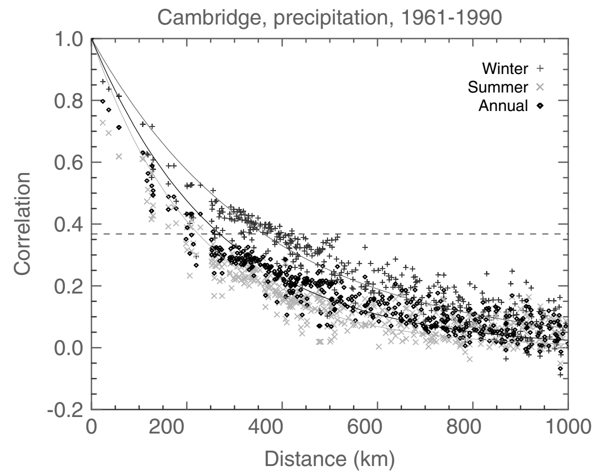

Figure 2. Example of computation of CDD, for Cambridge, UK using all data, and also winter and summer. Exponential curves are fitted to the data by least squares, and the CDD is defined as the distance where the fitted curve has a correlation of $1/e$ (dashed line): 269 km for all data, 361 km for winter and 244 km for summer.

DOI: 10.1002/joc

this study we use 12 nodes, which is a user defined number, dependent on data availability and the degree on differentiation required.

3.2. Variability in CDD: results

The spatial pattern of CDDs calculated using all the data are shown in Figure 3; they are similar to those reported by Osborn and Hulme (1997), with longer CDDs in NW and W Europe (up to 400 km) and shorter CDDs in Mediterranean Europe, especially in Italy, (as low as 150 km). These patterns are consistent with greater frontal influence on rainfall in the NW and W, and more convective and local scale influence in the E and Mediterranean. CDDs are also lower over the Alps, especially the Southern Alps, where the Alps act as a barrier to frontal influence. CDDs tend to be lower at the edge of the analysis domain, at least partly because these stations can only be correlated with stations within the domain.

When subsetting the data by season (Figure 3(b)–(e)), quite large contrasts in CDD between winter and summer can be seen, with intermediate differences in spring and autumn (not shown). In winter, CDDs at any station are nearly always larger than those using annual data, and significantly larger than in summer. This effect is greatest in NW and W Europe, where frontal rainfall is more important in winter compared to summer; differences here between summer and winter CDDs can be as much as 150 km.

If stations are divided according to azimuth, some differences in the patterns of CDDs in an EW and NS direction emerge (Figure 3(f) and (i)), but they are not particularly distinct. In Western Europe, EW CDDs tend to be larger than NS, and also CDDs derived using all station data. In central Europe the opposite tends to occur, and in Eastern Europe there is little difference. Discrepancies are largest over the northern part over the Alps, where EW CDDs are higher than NS CDDs. We also evaluated whether a lagged correlation in an EW direction would increase CDDs, based on the idea of EW movement of frontal systems, but find no significant increase in CDDs when this is taken into account.

Figure 4 shows the patterns of CDDs under different synoptic states. Node 1 corresponds to a situation with low pressure over Scandinavia; in the nodes below that (nodes 5 and 9), the low pressure moves towards Greenland and becomes stronger. These patterns occur mostly in winter and also occasionally in spring and autumn. From node 9 to the right the low-pressure system becomes less strong and is located progressively closer to the United Kingdom; these patterns are also mostly associated with winter, spring and autumn weather. Nodes 2, 3, 4 and 7 correspond to high pressure close to the Spanish coast and a low-pressure system over the Middle East and are more common in spring, summer and autumn.

The CDDs associated with each node show quite marked variation with pressure pattern. Patterns with deep low pressure to the west and NW tend to have the largest CDDs, while CDDs associated with high-pressure patterns are lower. These results are similar to those achieved using seasonal subsetting, because different nodes tend to be preferentially associated with a particular season or seasons (e.g. node 1 in winter). However, the use of synoptic patterns has the advantage that it can account for occurrence of particular pressure patterns at any time of the year, whereas different patterns are lumped together if seasonal subsetting is used.

4. Influence of the CDD on ADW

4.1. Angular-distance interpolation

The ADW interpolation scheme used in this study is a modified version of Shepard's algorithm (Shepard, 1968). All stations within the search radius (which equals the CDD) from the point at which an interpolated estimate is needed (henceforth L) are selected. Then, weights are assigned to each station, which are function of the distance between L and the stations and the angular separation between the stations. The distance weight is equal to:

$$w_i = (e^{-x/CDD})^m \tag{2}$$

where w_i is the weight for station i and x the distance between station i and point L. m is a constant, set to 4, through cross validation. This value has also been used by New *et al.* (2000) and Caesar *et al.* (2006). The angular-distance weight for each station i out of k contributing stations is:

$$W_i = w_i \left\{ 1 + \frac{\sum_k w_k[1 - \cos(\theta_k - \theta_i)]}{\sum_k w_k} \right\}, i \neq k \tag{3}$$

where the position of the ith station is defined in terms of its distance, x_i and its angle to north, θ_i, relative to the specified point L (Caesar *et al.*, 2006). These weights ensure that stations closest to L and/or more isolated stations have a greater weight. Note that weights are standardized to sum to 1.0, regardless of the number of contributing stations.

Anomalies, in the case of rainfall the percentage of the monthly total, are used to reduce the influence of factors such as elevation (Widmann and Bretherton, 2000). Initially, all stations within the search radius are selected, but if there are more than ten stations, only the ten stations with the highest weights are used for the actual interpolation. If less than three stations are present within the search radius, interpolation for this grid point is not deemed possible, and the point is assigned a missing value.

4.2. Comparisons

In Section 3, we have shown that CDD exhibits considerable spatial variability, and that the degree of spatial

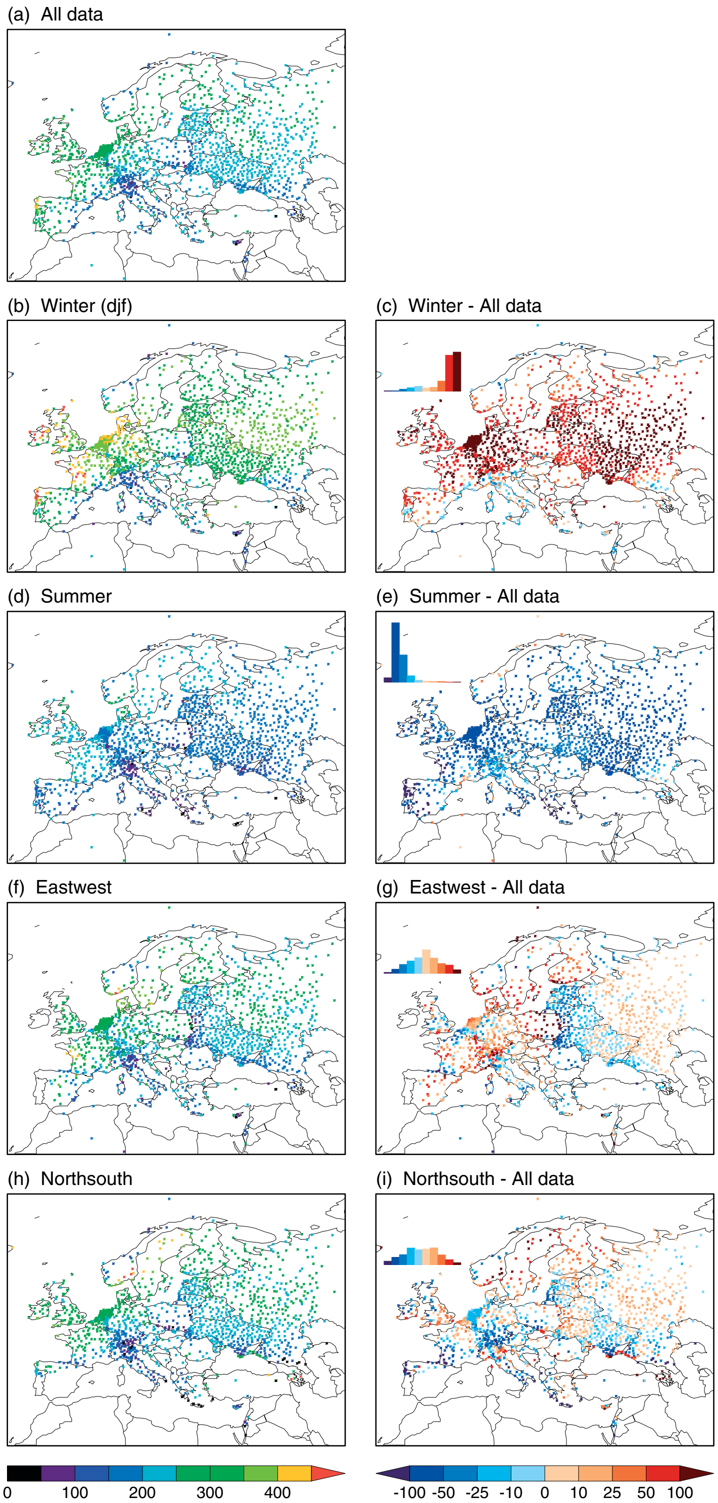

Figure 3. Spatial pattern of CDDs over Europe, for the period 1961–1990. (a) Calculated using all data; (b) and (d) for winter and summer; (f) and (h): split into east–west and north–south quadrants; (c), (e), (g), and (i): difference from a, with the frequency distribution of differences as inset. Figures (f)–(i) loose stations, e.g. Portugal and Ireland in (f) and (g) and Spain and Africa in (h) and (i), due to a lack of station data in each quadrant available for the calculation of the CDD. This figure is available in colour online at www.interscience.wiley.com/ijoc

variability is also dependent on season, synoptic state, and to a smaller extent, subsetting by azimuth (EW and NS quadrant pairs). Therefore, using a constant CDD within ADW might in particular locations be either too generous, or too conservative. For example, CDDs are typically larger in NW Europe and using an average (smaller) CDD for the whole of Europe would exclude stations that have useful information. Similarly, for Mediterranean Europe, CDDs are generally smaller, so using an average CDD would potentially include stations with no useful information, contaminating the interpolated value. These effects can be even larger if season and/or synoptic state

DOI: 10.1002/joc

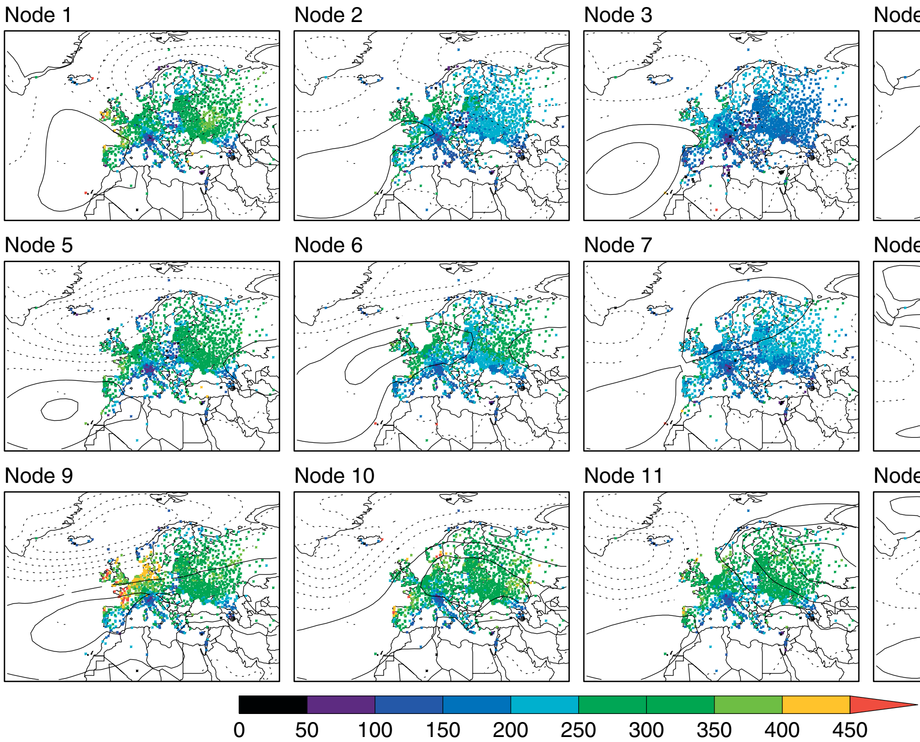

Figure 4. Pattern of CDDs for different synoptic states over 1961–1990, defined using self organizing maps. Contours show the 'archetypal' pressure patterns for each SOM node, from ERA40 mean sea level pressure data (solid lines show high-pressure, dotted lines low pressure; contour interval is 50 hPa). This figure is available in colour online at www.interscience.wiley.com/ijoc

are considered, for example, in winter over NW Europe, where CDDs are up to 200 km larger than the average for Europe.

We therefore test the sensitivity of ADW interpolation skill using a variable CDD. Our base case is interpolation using a constant CDD of 250 km, approximately the average for all stations over Europe (henceforth ADW0). We then allow CDD to vary spatially on a station-by-station basis within the ADW as follows:

– CDD at each station calculated using all data (ADW1);
– CDD varies as function of season (ADW2);
– CDD varies as function of synoptic state (ADW3);
– CDD optimized through cross validation at each station (ADW4);
– CDD optimized through station cross validation and varying as a function of season (ADW5);
– CDD optimized through station cross validation and varying as a function of synoptic state (ADW6).

ADW4–ADW6 represent the best possible interpolation that can be achieved with the existing method. For ADW4, at each target station we iterate through a range of possible CDDs (50–500 km) and choose the CDD that produces the best skill score (R, Compound Relative Error (CRE), Critical Success Index (CSI) and proportion correct (PC), see below). Two different optimal CDDs are obtained for each station: (1) state (wet/dry) and (2) magnitude. ADW is then run first to estimate state; if the value is $\geq$0.5 mm (wet) then the magnitude is estimated using the second CDD. The same procedure is adopted for ADW5 and ADW6, but there the iteration is undertaken separately for each season or each synoptic state.

4.3. Skill evaluation

We evaluate interpolation skill through station cross validation. Each station is excluded from the dataset in turn, and its daily values estimated using the remaining stations. Interpolation skill is then calculated for the excluded station by comparing estimated and observed values. We note that the interpolation scheme we are evaluating is not an exact interpolator, so cross validation against station data is not strictly appropriate; however, we expect the relative scores for cross validation to correspond to the relative skill of the interpolation method in estimating area average values at high spatial resolution.

Five skill scores are used: the Pearson correlation (R), the mean absolute error (MAE) and the CRE for the magnitude of rainfall, and the CSI and the PC for the state (wet or dry, where a wet day is quite simply defined as having a value $\geq$0.5 mm). These skill scores evaluate different characteristics of the interpolated data; we use a range of scores so that we can explore the extent to which improvement in skill due to interpolation method is consistent across different aspects of the observations.

– R is a statistic that removes the effect of any bias in the interpolated data and highlights just problems

DOI: 10.1002/joc

with modelling the daily variability. Problems with correctly capturing the variance will not be emphasized, because the measure normalizes the observed and modelled values by their standard deviations. The correlation coefficient should be considered as a measure of *potential* skill, because of its insensitivity to biases and errors in variance (Murphy and Epstein, 1989; Wilks, 2006). It is bounded below by −1 and above by 1 (best case).

– MAE is, according to Willmott and Matsuura (2006), a natural, unambiguous, measure of average error. MAE shows the errors in the same unit as the climate variable itself and is bounded below by 0 (best case) and unbounded above.

– CRE is a measure of similarity between the interpolated and observed values. The correspondence of two variables is measured in terms of relative departures from the means, and in terms of the means and absolute variances of the two series. It is bounded below by 0 (best case) and unbounded above and can be calculated as follows (Murphy and Epstein, 1989; Schmidli *et al*., 2001):

$$CRE = \frac{\sum_{k=1}^{n}(y_k - o_k)^2}{\sum_{k=1}^{n}(o_k - \overline{o})^2} \quad (4)$$

where y is the series to evaluate, o the observed series. A disadvantage of the CRE is that this skill score tends to favour interpolations that are too smooth. The method is sensitive to outliers (Murphy and Epstein, 1989; Schmidli *et al*., 2001).

– PC is the proportion correctly predicted, and is quite simple defined as the number of correct state predictions divided by the total number of predictions; it is the most direct and intuitive measure of accuracy:

$$PC = \frac{w_c + d_c}{n} \quad (5)$$

where w_c and d_c are the number of correct predictions of wet and dry days, respectively. The PC may not be the best measure to use when one of the correct events is less common than the other, such as for rainfall in very wet or dry regimes.

– In this case, CSI is a more appropriate method (Wilks, 2006):

$$CSI = \frac{w_c}{w_c + w_i + d_i} \quad (6)$$

where w_i and d_i are the number of days incorrectly predicted as wet and dry respectively.

4.4. ADW: results

As shown in Table I, the average skill across the European domain shows only small differences between the various approaches to using CDD for ADW interpolation. The skill of each method tends to be similar across the scores, but for some versions, the rank differs for magnitude (R, MAE and CRE) and state (CSI and PC). All but ADW1 performs better than our base case, ADW0. Thus using a CDD that varies as a function of season or synoptic state, and at the same time varies spatially produces better results than using a constant CDD or CDD that varies spatially, but not as a function of season of synoptic state. Although these improvements in average performance are larger for state than magnitude, in both cases they are not particularly notable. Using more or fewer SOM nodes does not produce the change in skill markedly (not shown).

Figures 5 and 6 show the differences in skill between the different ADW methods, for state (CSI) and magnitude (CRE) respectively. To reduce the number of maps, we do not include ADW5 and ADW6, as the patterns of these 'optimized' methods are similar to ADW4. The figures show that ADW1–ADW3 have a higher CRE skill than ADW0, mostly for areas in Eastern Europe. This area benefits from a lower CDD in general for ADW1 and a lower CDD in summer for ADW2 and ADW3. For CSI improvements in skill when using ADW1–ADW3 are mainly seen in the SE, whereas the skill in the NW deteriorates compared to ADW0. The interpolation of the state of rainfall generally benefits from lower CDDs and for ADW1–ADW3 the CDDs are higher in the NW than the SE. ADW4 nearly always has a better skill than ADW0. In areas with a sparser

Table I. Skill scores and their ranks for seven different versions of ADW (Section 4). Results of 1471 stations have been used for the comparison. The 220 (1) stations are lost because the CDD of ADW1 (ADW0) is too low to do the interpolation, the other 76 stations do not have enough data available.

ADW version	Average rank	Rank	R	Rank	MAE	Rank	CRE	Rank	CSI	Rank	PC
0	5.6	5	0.761	5	1.237	4	0.417	7	0.660	7	0.838
1	6.6	7	0.760	7	1.231	7	0.421	6	0.662	6	0.840
2	4.8	5	0.761	5	1.237	5	0.419	4	0.664	5	0.841
3	4.2	4	0.762	4	1.225	5	0.419	4	0.664	4	0.842
4	2.6	3	0.772	3	1.152	3	0.398	2	0.691	2	0.862
5	2	2	0.773	2	1.147	2	0.396	2	0.691	2	0.862
6	1	1	0.775	1	1.146	1	0.394	1	0.693	1	0.863

DOI: 10.1002/joc

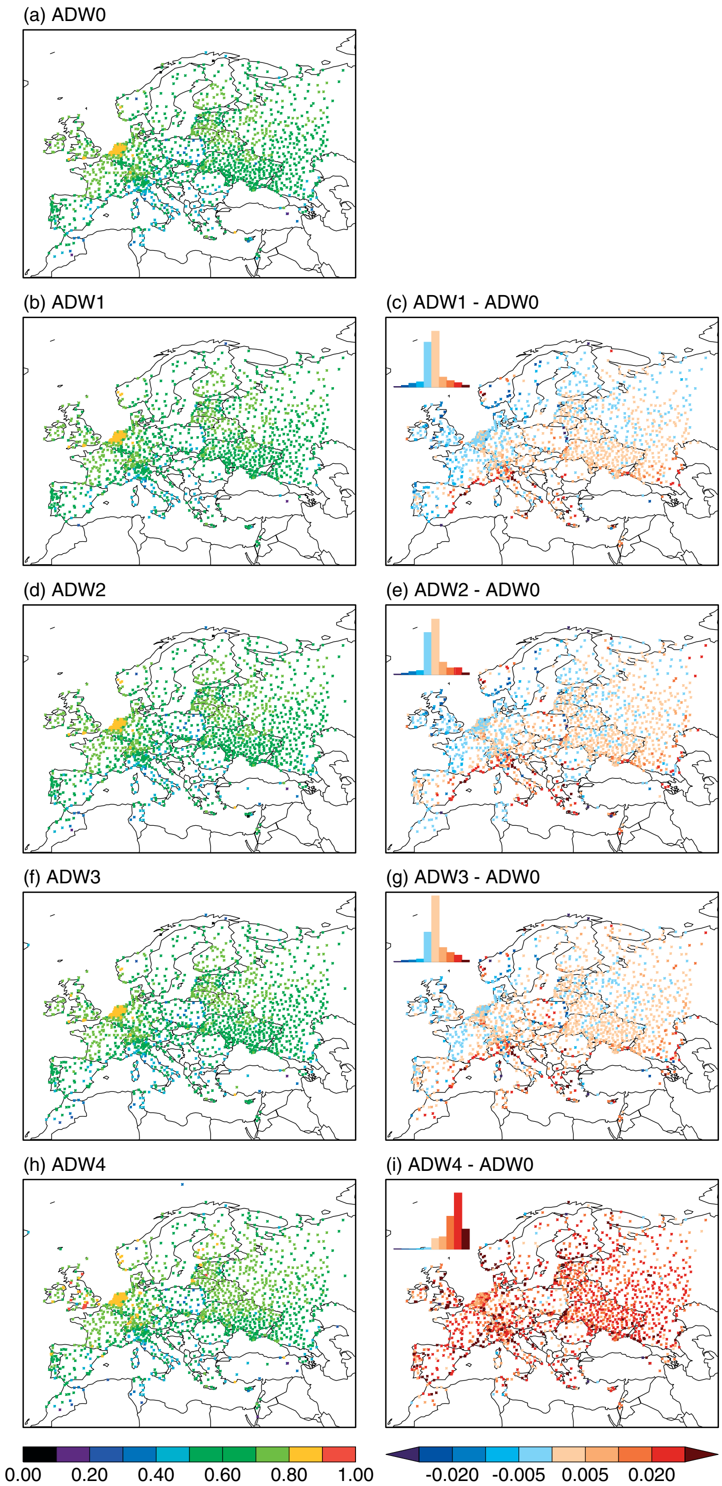

Figure 5. Spatial pattern of CSI for each ADW method (left) and difference from the base case, ADW0 (right); also shown is the frequency distribution of differences in CSI across the dataset. This figure is available in colour online at www.interscience.wiley.com/ijoc

station network, differences are less consistent; here ADW0 shows better skill at some stations. As with the spatially averaged statistics, improvements in skill are quite small, typically less than 1% for CRE and 2–5% for CSI.

One consequence of using variable CDDs is that for smaller CDDs the number of stations (or with gridding the number of grid points) where an estimate can be made decreases. This is because our implementation of ADW requires at least three stations within the CDD before an estimate can be made. Thus ADW1 is less effective from this point of view, as it has 197 and 220 fewer stations that have three or more neighbouring stations within the CDD than ADW0 and ADW4 respectively. ADW4 is penalized least in this way, because the optimization process chooses the best

DOI: 10.1002/joc

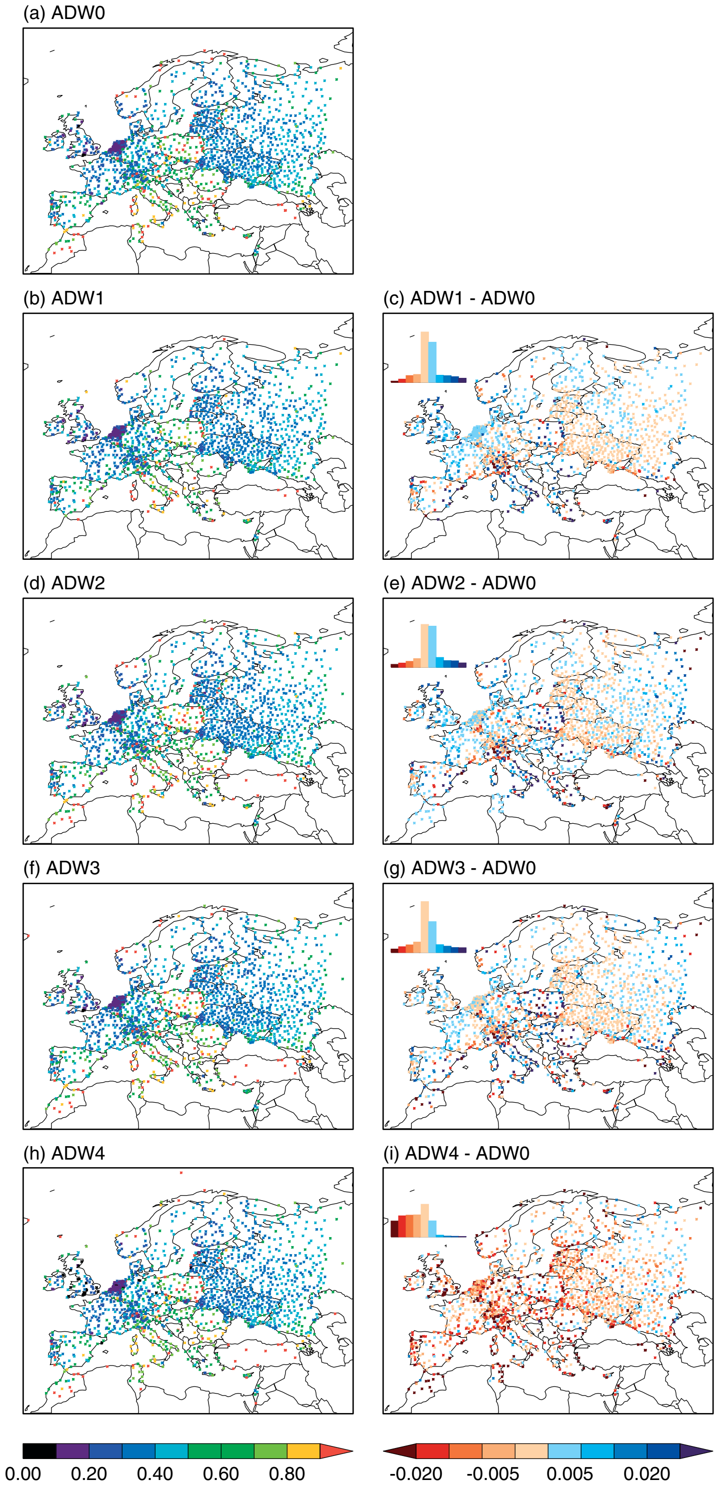

Figure 6. Same as in Figure 5, but for the CRE skill score. This figure is available in colour online at www.interscience.wiley.com/ijoc

CDD for which there are at least three stations within the CDD radius. This has the advantage that estimates can be made for nearly all station locations, but for locations with very distant neighbours, the interpolation will have lower skill. An analysis of the skill at these stations shows that the interpolation skill is reduced between about 5 and 30% for each of our skill scores. However, compared to other interpolation methods (Hofstra *et al*., 2008), ADW4 performs better for these specific stations and in a gridded dataset based on this station network (Haylock *et al*., 2008) uncertainty of grid-point estimates for these stations will be correspondingly higher.

5. Discussion and conclusions

The underlying rationale for the use of CDD in ADW is that only stations that are correlated with the target location should be used in the interpolation, and that using uncorrelated stations simply interpolates noise or error to the target location. Thus the use of CDDs is intuitively appealing. If they are used it would appear logical to use a spatially variable set of CDDs as this maximizes their benefit.

Our analysis has shown that the CDD varies with season, synoptic state and, to a lesser degree, are anisotropic, varying with azimuth. Similarly to Osborn and Hulme (1997) we find lower CDDs in the Mediterranean and much higher correlation distances in the United Kingdom (and also NW Europe generally), indicating that the CDD increases from SE to NW. Osborn and Hulme (1997) find CDDs that are generally slightly lower than the CDDs we find, which may be in part due to their use of a less dense station network.

Our SOM analysis has shown that CDD varies markedly with synoptic state. Highest CDDs are found when there is a high-pressure area just west of Spain and a low-pressure area over Iceland, corresponding to frontal conditions over NW Europe. The lowest CDDs are found when both pressure systems are much less strong and there is a low-pressure system over the Middle East.

Latitudinal subsetting has been used in the production, by ADW, of grids of daily temperature (Caesar *et al*., 2006) and daily temperature and rainfall extremes (Alexander *et al*., 2006). For rainfall in Europe, this subsetting would not be optimal because CDD generally changes in a NW–SE direction. Seasonally different CDDs have been used by Alexander *et al*. (2006), but Caesar *et al*. (2006) concluded that the annual mean interpolation error was slightly lower using an annually constant CDD than using different CDDs for each month. This study shows that the introduction of seasonal or synoptic state dependent CDDs does improve the skill of ADW, but only very slightly.

Despite the intuitive appeal of using a variable CDD in ADW interpolation, only small improvement gained from implementing this approach over Europe for daily precipitation. The primary limitation on interpolation accuracy appears to be station network density, as shown by Hofstra *et al*. (2008) in a comparison of several interpolation methods. In that comparison, however, use of ADW4 produces interpolation skill extremely close to the best performing interpolation method (global kriging) while use of a constant CDD reduced skill by between 1 and 5%. Improvement of ADW with CDDs optimized through cross validation should be considered in further studies on the interpolation of daily rainfall data, especially in areas with relatively dense networks.

Acknowledgements

We would like to thank all institutes that made data available for the study (see Appendix 1 of Klok and Klein Tank, 2008). Financial support for this research comes from the EU project ENSEMBLES (WP 5.1, contract number: GOCE-CT-2004-50539), the Dutch Talentenbeurs and the Dutch Prins Bernhard Cultuurfonds beurs.

References

Alexander LV, Zhang X, Peterson TC, Caesar J, Gleason B, Klein Tank AMG, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Rupa Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson DB, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre JL. 2006. Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* **111**: D05109, DOI: 10.1029/2005JD006290.

Briffa KR, Jones PD. 1993. Global surface air temperature variations during the twentieth century: Part 2, implications for large-scale high-frequency paleoclimatic studies. *Holocene* **3**: 77–88.

Caesar J, Alexander L, Vose R. 2006. Large-scale changes in observed daily maximum and minimum temperatures: Creation and analysis of a new gridded dataset. *Journal of Geophysical Research* **111**: D05101, DOI: 10.1029/2005JD006280.

Haylock MR, Hofstra N, Klein Tank AMG, Klok EJ, Jones PD, New M. 2008. A European daily high- resolution gridded data set of surface temperature and precipitation for 1950-2006. *Journal of Geophysical Research* **113**: D20119, DOI:10.1029/2008JD010201.

Hewitson BC, Crane RG. 2002. Self-organizing maps: applications to synoptic climatology. *Climate Research* **22**: 13–26.

Hofstra N, Haylock M, New M, Jones PD, Frei C. 2008. Comparison of six methods for the interpolation of daily, European climate data. *Journal of Geophysical Research* **113**: D21110, DOI:10.1029/2008JD010100.

Jones PD, Osborn TJ, Briffa KR. 1997. Estimating sampling errors in large-scale temperature averages. *Journal of Climate* **10**: 2548–2568.

Karoly DJ, Braganza K, Stott PA, Arblaster JM, Meehl GA, Broccoli AJ, Dixon KW. 2003. Detection of a human influence on North American climate. *Science* **302**: 1200–1203, DOI: 10.1126/science.1089159.

Kiktev D, Sexton DMH, Alexander L, Folland CK. 2003. Comparison of modeled and observed trends in indices of daily climate extremes. *Journal of Climate* **16**: 3560–3571.

Klok L, Klein Tank AMG. 2008. Updated and extended European dataset of daily observations. *International Journal of Climatology* (in press).

Legates DR, Willmott CJ. 1990a. Mean seasonal and spatial variability in gauge-corrected, global precipitation. *International Journal of Climatology* **10**: 111–127.

Legates DR, Willmott CJ. 1990b. Mean seasonal and spatial variability in global surface air temperature. *Theoretical and Applied Climatology* **41**: 11–21.

Murphy AH, Epstein ES. 1989. Skill scores and correlation coefficients in model verification. *Monthly Weather Review* **117**: 572–581.

New M, Hulme M, Jones PD. 2000. Representing twentieth-century space–time climate variability. Part II: development of 1901–1996 monthly grids of terrestrial surface climate. *Journal of Climate* **13**: 2217–2238.

Osborn TJ, Hulme M. 1997. Development of a relationship between station and grid-box rainday frequencies for climate model evaluation. *Journal of Climate* **10**: 1885–1908.

Schmidli J, Frei C, Schär C. 2001. Reconstruction of mesoscale precipitation fields from sparse observations in complex terrain. *Journal of Climate* **14**: 3289–3306.

Shepard D. 1968. A two-dimensional interpolation function for irregularly-spaced data. *ACM National Conference*, New York, Association for Computing Machinery, 517–523.

Widmann M, Bretherton CS. 2000. Validation of mesoscale precipitation in the NCEP reanalysis using a new gridcell dataset for the northwestern United States. *Journal of Climate* **13**: 1936–1950.

Wilks DS. 2006. *Statistical Methods in the Atmospheric Sciences*, *International Geophysics Series*, Vol. 91, 2nd edn. Academic Press: USA; 627.

Willmott CJ, Clinton MR, Philpot WD. 1985. Small-scale climate maps: A sensitivity analysis of some common assumptions associated with grid-point interpolation and contouring. *American Cartographer* **12**: 5–16.

Willmott CJ, Matsuura K. 2006. On the use of dimensioned measures of error to evaluate the performance of spatial interpolators. *International Journal of Geographical Information Science* **20**: 89–102.