On the Choice of the Temporal Aggregation Level for Statistical Downscaling of Precipitation

T. A. BUISHAND, M. V. SHABALOVA, AND T. BRANDSMA

Royal Netherlands Meteorological Institute, De Bilt, Netherlands

(Manuscript received 18 September 2002, in final form 27 August 2003)

ABSTRACT

The merits of daily and monthly downscaling models for precipitation are compared using data from Bern, Switzerland; Deuselbach, Germany; and De Bilt, the Netherlands. For each station, generalized linear models are developed to describe rainfall occurrence, the wet-day precipitation amounts, and the monthly precipitation totals. The predictor dataset includes dynamical variables and atmospheric moisture (relative humidity for rainfall occurrence and specific humidity for rainfall amount).

Fitting a generalized linear model to daily rainfall data generally results in larger regression coefficients than fitting the same model to monthly data. For rainfall occurrence this can be attributed mostly to the nonlinearity of the function that links the wet-day probabilities to the predictor variables, whereas for rainfall amounts there is, apart from nonlinearity, also a bias in the estimated regression coefficients of the monthly models caused by averaging predictor variables over both wet and dry days. Because of this bias a monthly rainfall amount model is less sensitive to an increase in the specific humidity than a daily rainfall amount model.

Although the squared correlation coefficient r^2 between the observed and predicted values of the daily models is low ($\approx$0.40 for rainfall occurrence and $\approx$0.15 for wet-day rainfall), aggregating the results from these models to monthly values gives r^2 values comparable to those in the direct fit to the monthly data ($\approx$0.65 for the number of wet days and $\approx$0.50 for rainfall totals). The temporal variations in the predicted annual amounts using monthly relationships are similar to those obtained from daily relationships. Daily models are preferable, however, for the generation of climate change scenarios for impact studies, because the significance of the predictor variables is generally stronger in these models and because the effect of a change in specific humidity is underestimated by the monthly models.

1. Introduction

The term statistical downscaling refers to statistical techniques that are used to obtain climate variables at the required temporal and spatial resolution for climate change impact studies. In particular for precipitation, the direct output of climate change simulations from general circulation models (GCMs) is inadequate for most impact studies and needs to be enhanced.

The statistical downscaling approach makes use of relationships between the observed local precipitation and atmospheric predictor variables. The predictors should be realistically modeled by the GCM and should fully represent the climate change signal. Practically, this implies that both circulation-based and humidity predictors are included in the downscaling model (Giorgi et al. 2001). There is a risk that a vital predictor for climate change is discarded as statistically not significant. The statistical significance of a predictor in the present climate depends, among other factors, on the chosen temporal aggregation level of the data. Some authors have developed downscaling relationships for monthly data (e.g., Kilsby et al. 1998; Murphy 1999, 2000), whereas others have considered daily data (e.g., Wilby et al. 1998; Beckmann and Buishand 2002). Typically, the aggregation level is simply determined by the temporal resolution needed for the climate change application of interest. Seldom has it been realized that the aggregation level influences the choice of predictors and the estimated changes under future climate conditions.

In this paper the differences between statistical downscaling models for daily and monthly precipitation data are studied in detail. This is done with data from three stations in the River Rhine basin: Bern in Switzerland, Deuselbach in Germany, and De Bilt in the Netherlands (Table 1). For each of these stations, generalized linear models (GLMs) are used to describe precipitation occurrence, the wet-day precipitation, and the monthly precipitation totals. GLMs extend the classical linear regression model to cases where the data come from an exponential family other than the normal distribution, and allow for a nonlinear link between the expected response and the predictors. Such a link is needed to ensure that

Corresponding author address: Dr. T. A. Buishand, Royal Netherlands Meteorological Institute (KNMI), P.O. Box 201, 3730 AE De Bilt, Netherlands.
E-mail: Adri.Buishand@knmi.nl

TABLE 1. Geographical position, altitude, mean annual rainfall, and mean number of wet days per year for Bern, Deuselbach, and De Bilt. Annual means refer to the period 1968–95.

	Bern	Deuselbach	De Bilt
Lat (N)	46°56′	49°46′	52°06′
Lon (E)	7°25′	7°03′	5°11′
Alt (m)	565	480	2
Annual rainfall (mm)	1058	792	802
No. of wet days	170	196	194

the probabilities of precipitation lie in the interval (0, 1) and that the wet-day precipitation amounts are positive. The influence of the aggregation level is studied for the estimated regression coefficients, their standard errors, and the proportion of explained variance. The effects of systematic changes in a predictor variable on the number of wet days and the rainfall amounts are studied for the fitted downscaling models.

Models for daily precipitation occurrence and for the monthly number of wet days are compared in section 2. Section 3 presents a similar comparison for precipitation amount models. In section 4 the estimated monthly and annual precipitation totals from the statistical models for daily and monthly data are compared. Section 5 closes the paper with a discussion and conclusions.

2. Rainfall occurrence models

A wet day is defined here as a day with 0.1 mm of precipitation or more. For daily data, the predictand Y_t can take only two values: $Y_t = 1$ if day t is wet or $Y_t = 0$ if it is dry. The number M_W of wet days per month is the predictand in the model for monthly data. In both cases, logistic regression is used to link the predictand to atmospheric variables. The logistic model belongs to the class of GLMs. There are many applications of the logistic model in the climatological and meteorological literature such as time series modeling (e.g., Coe and Stern 1982), medium-range weather forecasting (e.g., Lemcke and Kruizinga 1988), identification of weather codes (Merenti-Välimäki and Laininen 2002), and detection of trends in rare events (Frei and Schär 2001). Kilsby et al. (1998) and Wilby (2001) applied an empirical logistic transformation to observed frequencies of wet and dry days.

In this section the logistic model is introduced first for daily rainfall occurrence and then for the monthly number of wet days. The fit to rainfall occurrence data from Bern, Deusselbach, and De Bilt (Table 1) is discussed thereafter.

a. Logistic model for daily rainfall occurrence

The key parameter in the logistic model is the probability π of a day being wet. The model assumes that the logistic transformation of π is a linear function of the predictors $x_1, \ldots, x_p$:

$$g(\pi) = \log\left[\frac{\pi}{1 - \pi}\right] = a_0 + a_1 x_1 + \cdots + a_p x_p. \quad (1)$$

The function $g(\cdot)$ is known as the link function in the literature on GLMs. The regression coefficients a_0, a_1, . . . , a_p are usually estimated by the method of maximum likelihood (McCullagh and Nelder 1989). In GLMs, the maximum likelihood estimates can be obtained via an iteratively reweighted least squares (IRLS) procedure.

The mean and variance of the rainfall occurrence indicator Y are completely determined by the wet-day probability π:

$$\mathrm{E}(Y) = \pi \quad \text{and} \quad (2a)$$

$$\mathrm{var}(Y) = \pi(1 - \pi). \quad (2b)$$

As distinct from the classical linear regression model, the variance depends on the expected response and thus on unknown regression coefficients.

b. Logistic model for monthly rainfall occurrence

The logistic model for monthly rainfall occurrence assumes that the mean and variance of the number M_W of wet days can be represented by expressions similar to Eq. (2):

$$\mathrm{E}(M_W) = M\Pi \quad \text{and} \quad (3a)$$

$$\mathrm{var}(M_W) = \phi M\Pi(1 - \Pi). \quad (3b)$$

Here M stands for the total number of days in the month of interest, ϕ is a dispersion parameter, and the wet-day probability Π is given by the same logistic equation as Eq. (1), but now with monthly mean values $\overline{X}_1$, . . . , $\overline{X}_p$ of the predictors.

In contrast to the daily rainfall occurrence model, the distribution of the predictand is not entirely specified. For $\phi = 1$, Eq. (3b) gives the variance for the binomial distribution, which would arise if the Y_t values were independent with the same wet-day probability π. Because π varies with $x_1, \ldots, x_p$ and because of persistence in the daily rainfall occurrence, the variance of M_W differs from the binomial variance. Positive autocorrelation is a well-known cause of overdispersion; that is, $\phi > 1$ (McCullagh and Nelder 1989).

Even though the distribution of M_W is not completely defined, the regression coefficients a_0, a_1, . . . , a_p can still be estimated using the IRLS procedure for maximum likelihood estimation for data from a binomial distribution. The resulting estimates are known as quasi-likelihood estimates. The theory of quasi likelihood only requires that the relationship between the variance and the mean is specified (McCullagh and Nelder 1989).

In this study, the estimate of ϕ was based on Pearson's χ^2 statistic:

$$\hat{\phi} = \chi^2/(n - p - 1), \quad (4)$$

where

$$\chi^2 = \sum_{\tau=1}^{n} \frac{(M_{W,\tau} - M_\tau \hat{\Pi}_\tau)^2}{M_\tau \hat{\Pi}_\tau (1 - \hat{\Pi}_\tau)}, \quad (5)$$

with $M_{W,\tau}$ the number of wet days in month τ, M_τ the total number of days in month τ, $\hat{\Pi}_\tau$ the estimated wet-day probability for month τ, and n the total number of months. Note that the χ^2 statistic weights the squared residuals inversely proportional to the estimated variances.

c. *Results*

Daily data for the 28-yr period 1968–95 were considered. The same predictor variables were used for each rainfall station: the relative humidity rh at 700 hPa, the sea level pressure slp, and the west component u as well as the south component v of the geostrophic flow, both derived from the sea level pressure. The choice of these predictor variables was based on earlier downscaling studies for Bern (Buishand and Beckmann 2000), and the Netherlands and north Germany (Beckmann and Buishand 2002). Relative humidity and sea level pressure were obtained from the National Centers for Environmental Prediction–National Center for Atmospheric Research (NCEP–NCAR) reanalysis dataset (Kalnay et al. 1996; Kistler et al. 2001). Six-hourly values of numerous weather variables were available on a 2.5° × 2.5° grid. The grid point nearest to the rainfall station was considered. Daily averages of the predictor variables were calculated from the four 6- hourly values in the reanalysis data that were within the sampling interval of the daily rainfall measurements. For the monthly rainfall occurrence model these daily averages were converted to monthly averages.

The full 28-yr records, without seasonal stratification, were used to explore the effects of the temporal aggregation level at the three stations. Because the assumption of a constant statistical relationship over the year could be questioned, the data for De Bilt were also analyzed for a 5-month extended winter season (November–March) and a 5-month extended summer season (May–September). Extended seasons rather than standard 3-month seasons were chosen here to help ensure that a reasonable number of the predictors are statistically significant.

Table 2 summarizes for the three sites the results for the logistic models fitted to the rainfall occurrence data from the entire year. The correlation coefficients between the observed and predicted daily (r_D) and monthly (r_M) values are measures of model performance. The square of these coefficients is almost equal to the proportion of the explained variance. In the daily model all predictors are significant at the 1% level. The predictors in the monthly model are significant at the 5% level, except sea level pressure for Deuselbach and the v velocity for De Bilt.

Table 2 shows a clear dependence of the estimated regression coefficients and their standard errors on aggregation level. The absolute values of the estimated regression coefficients in the daily model are nearly always larger than those in the monthly model. The standard errors of the estimated regression coefficients of rh, slp, and u and v approximately double with the change from the daily to the monthly aggregation level. Despite the marked persistence in daily rainfall occurrence, the lag 1 autocorrelation coefficient r_{res} of the residuals ($Y_t - \hat{\pi}_t$) does not exceed 0.18. Somewhat surprisingly, the values of r_{res} for the monthly rainfall occurrence model are not smaller than those for the daily rainfall occurrence model. This may partly be due to nonhomogeneities in the reanalysis data (see section 4). The values of r_M^2 in Table 2 appear to be much larger than the corresponding values of r_D^2.

TABLE 2. Estimated regression coefficients and correlation coefficients in the logistic models for daily and monthly rainfall occurrence (r_{res} is the lag 1 autocorrelation coefficient of the residuals, and r_D and r_M are the correlation coefficients between the observed and predicted values of daily and monthly rainfall occurrence, respectively). For the model for monthly rainfall occurrence, $\hat{\phi}$ is the estimate of the dispersion parameter ϕ in Eq. (3b). The standard error se of an estimated regression coefficient is the average standard error for the three stations as obtained from Eq. (4.21) in McCullagh and Nelder (1989).

Variable	Bern	Deuselbach	De Bilt	se
		Daily rainfall occurrence		
rh (%)	0.056	0.053	0.046	0.0016
slp (hPa)	−0.055	−0.023	−0.081	0.0040
u (m s^{-1})	0.086	0.156	0.121	0.0044
v (m s^{-1})	−0.197	−0.041	−0.081	0.0058
r_{res}	0.17	0.18	0.15	
r_D^2	0.43	0.36	0.40	
r_M^2	0.63	0.63	0.64	
		Monthly rainfall occurrence		
rh (%)	0.037	0.040	0.029	0.0046
slp (hPa)	−0.020	−0.013	−0.054	0.0093
u (m s^{-1})	0.097	0.150	0.101	0.0077
v (m s^{-1})	−0.073	−0.028	−0.017	0.0122
r_{res}	0.23	0.15	0.18	
$\hat{\phi}$	1.04	1.40	1.44	
r_M^2	0.66	0.65	0.65	

The observed differences between the results for daily and monthly rainfall occurrence are due partly to the nonlinearity of the logistic model and partly to the autocorrelation of the daily variables (appendixes A and B). In the classical linear regression model with independent errors, there is no dependence of the estimated regression coefficients on temporal aggregation. To demonstrate the autocorrelation effect, the days in the 1968–95 period were randomly permutated and the logistic model was then fitted to the monthly number of wet days again. This was done 20 times. The results, averaged over all 20 permutations, are presented in Table 3. The regression coefficients in Table 3 are closer to those for the original monthly model than to those for the daily model, thus indicating that the differences between the regression coefficients in Table 2 are due mainly to nonlinearity. The standard errors of the es-

TABLE 3. Estimated regression coefficients and other parameters in the logistic model for monthly rainfall occurrence after random permutation of the days in the 1968–95 record (averaged over 20 different permutations; the standard errors se are also averaged over the three stations).

Variable	Bern	Deuselbach	De Bilt	se
		Monthly rainfall occurrence		
rh (%)	0.036	0.036	0.033	0.0050
slp (hPa)	−0.028	−0.018	−0.047	0.0130
u (m s^{-1})	0.051	0.095	0.070	0.0134
v (m s^{-1})	−0.117	−0.020	−0.043	0.0182
r_{res}	−0.01	0.01	−0.01	
$\hat{\phi}$	0.61	0.65	0.63	
r_M^2	0.40	0.35	0.39	

timated regression coefficients for the monthly model with permutated data are about three times as large as those for the daily model. The change in standard error with aggregation level is smaller if the predictor variable exhibits positive autocorrelation (appendix A) as observed in Table 2. The values of r_M^2 for the monthly model with permutated data are close to those of r_D^2 for the daily model. This indicates that the seemingly better performance of the monthly model in Table 2 is a result of persistence in the predictors. For the linear regression model with independent errors, it can easily be shown that aggregation leads to an increase in the proportion of explained variance if there is positive autocorrelation in the predictors (appendix B).

The estimated dispersion parameter $\hat{\phi}$ for monthly rainfall occurrence in Table 3 is smaller than 1 because there is no persistence, whereas the wet-day probabilities vary within the month. The values of $\hat{\phi}$ in Table 2 for the real number of wet days are larger than those in Table 3 because of persistence. Still $\hat{\phi}$ does not deviate much from 1 in Table 2 because a large part of the persistence in daily rainfall occurrence is explained by the persistence in the predictor variables.

Table 4 shows for De Bilt the results for the winter and summer rainfall occurrence data. There is seasonal variation in the values of various regression coefficients. The seasonal differences between the regression coefficients of slp and u in the daily rainfall occurrence models are even statistically significant at the 1% level. Despite the seasonal variation of the regression coefficients, the differences between the monthly and daily rainfall occurrence models are similar to those found for the models without seasonal stratification in Table 2: the estimated regression coefficients are smaller in the monthly models fitted to the seasonally stratified data, their standard errors are about twice as large as those in the daily models, and the values of r_M^2 are much larger than the corresponding values of r_D^2.

d. The effect of a systematic change in a predictor variable

Table 5 illustrates the sensitivity of the downscaling models to changes in the predictors. Shown are the

TABLE 4. Same as in Table 2 but for rainfall occurrence at De Bilt for the extended winter (Nov–Mar) and extended summer (May–Sep) seasons.

Variable	Winter	se	Summer	se
		Daily rainfall occurrence		
rh (%)	0.047	0.0025	0.053	0.0026
slp (hPa)	−0.071	0.0046	−0.115	0.0078
u (m s^{-1})	0.096	0.0054	0.139	0.0078
v (m s^{-1})	−0.099	0.0068	−0.088	0.0093
r_{res}	0.12		0.14	
r_D^2	0.42		0.40	
r_M^2	0.72		0.68	
		Monthly rainfall occurrence		
rh (%)	0.030	0.0079	0.043	0.0073
slp (hPa)	−0.051	0.0103	−0.072	0.0180
u (m s^{-1})	0.075	0.0077	0.088	0.0139
v (m s^{-1})	−0.064	0.0146	−0.007	0.0210
r_{res}	0.20		0.04	
$\hat{\phi}$	1.09		1.16	
r_M^2	0.72		0.68	

changes in the average number of wet days per year for De Bilt, resulting from a constant change in the sea level pressure, the u velocity or the v velocity, and assuming that the rainfall occurrence models are constant over the year. The monthly rainfall occurrence model is much less sensitive to a systematic change in the v velocity than is the daily rainfall occurrence model. For the perturbations of slp and the u velocity, the use of the monthly rainfall occurrence model does not lead to a smaller change in the average number of wet days, even though the regression coefficients are larger in the daily model. This can be understood as follows.

The change ΔN_W in the total number of wet days resulting from a perturbation Δx_k of the kth predictor variable can be approximated as (appendix C)

$$\Delta N_W \approx a_{k,D} C_D \Delta x_k \quad \text{(daily model)} \quad \text{and} \qquad (6a)$$

$$\Delta N_W \approx a_{k,M} C_M \Delta x_k \quad \text{(monthly model)}, \qquad (6b)$$

where $a_{k,D}$ and $a_{k,M}$ are the regression coefficients of the kth predictor. The quantities C_D and C_M depend on the wet-day probabilities in the present climate, but not on the perturbation Δx_k. The quantity C_M is larger than C_D. For the same perturbation Δx_k, the daily rainfall occurrence model provides a larger value of $|\Delta N_W|$ than the monthly model if $|a_{k,D}| > |a_{k,M}| \times C_M/C_D$ (=1.53 $\times$ $|a_{k,M}|$ for De Bilt). This condition is satisfied for the v velocity, but not for the u velocity. For slp the ratio $a_{k,D}/a_{k,M}$ is very close to 1.53 resulting in the same sensitivity of the daily and monthly rainfall occurrence models to a constant perturbation of this variable. For the other two stations similar values are found for the critical ratio between the regression coefficients, C_M/C_D = 1.64 for Bern and C_M/C_D = 1.45 for Deuselbach.

The monthly rainfall occurrence model for the summer data at De Bilt also shows a low sensitivity to a constant change of the v velocity. This is not the case

TABLE 5. Change in the average number of wet days per year at De Bilt for a constant increase in the sea level pressure of 1, 2, and 4 hPa, respectively, and for a constant increase in the u velocity or v velocity of 1, 2, and 4 m s^{-1}, respectively.

Variable	Daily model	Monthly model	Variable	Daily model	Monthly model	Variable	Daily model	Monthly model
slp + 1	−4	−4	$u + 1$	7	8	$v + 1$	−4	−1
slp + 2	−9	−9	$u + 2$	13	17	$v + 2$	−9	−3
slp + 4	−18	−18	$u + 4$	26	33	$v + 4$	−18	−6

for the model fitted to the monthly rainfall occurrence data from the winter season.

3. Rainfall amount models

Wet-day rainfall amounts have a highly skewed distribution, which has often been described by the gamma distribution (e.g., Katz 1977; Buishand 1978). The gamma distribution belongs to the exponential family and is the standard distribution for GLMs with constant coefficient of variation CV. Coe and Stern were the first to fit such a model to wet-day precipitation amounts (Coe and Stern 1982; Stern and Coe 1984). They demonstrated the use of GLMs to describe the seasonal variation of the mean wet-day rainfall amount and its dependence on rainfall occurrence. Buishand and Klein Tank (1996) applied a GLM with constant CV to link wet-day precipitation at De Bilt to temperature and sea level pressure. In the present paper this model is fitted not only to the individual wet-day precipitation amounts but also to the monthly average wet-day precipitation and to the monthly precipitation totals.

The GLM is defined first for daily data. Then its application to monthly data is discussed. The results for Bern, Deuselbach, and De Bilt are presented at the end of this section.

a. Generalized linear model for individual wet-day rainfall amounts

As in previous applications of GLMs to wet-day precipitation R, a log-link function is used here to avoid negative values for the expected amounts E(R). This implies that

$$E(R) = \exp(a_0 + a_1 x_1 + \cdots + a_p x_p), \quad (7)$$

where $x_1, \ldots, x_p$ are the predictor variables. The variance of R can be represented as

$$\mathrm{var}(R) = \phi[E(R)]^2, \quad (8)$$

where $\phi = \mathrm{CV}^2$ is a dispersion parameter. The specification of a constant CV is sufficient to obtain quasi-likelihood estimates of the coefficients $a_0, a_1, \ldots, a_p$ in Eq. (7) by the IRLS procedure. These estimates are the maximum likelihood estimates if a gamma distribution for R is assumed. The estimate of the dispersion parameter ϕ is based on the generalized χ^2 statistic:

$$\hat{\phi} = \frac{\chi^2}{N_W - p - 1} = \frac{1}{N_W - p - 1} \sum_{t=1}^{N_W} \frac{(R_t - \hat{\mu}_t)^2}{\hat{\mu}_t^2}, \quad (9)$$

with R_t the observed precipitation amount on the tth wet day, $\hat{\mu}_t$ the expected precipitation amount from the fitted model for the tth wet day, and N_W the total number of wet days.

b. Generalized linear models for monthly rainfall amounts

Two methods are considered to describe monthly rainfall. In the first method, the estimate of the number of wet days from Eq. (3a) is supplemented with an estimate of the monthly average wet-day precipitation amount $\overline{R} = R_M/E(M_W)$, where R_M is the monthly total precipitation amount. It is assumed that R_M is the sum of E(M_W) independent wet-day precipitation amounts. The mean of $\overline{R}$ can then be described by Eq. (7) and its variance is given by

$$\mathrm{var}(\overline{R}) = \phi[E(\overline{R})]^2/E(M_W), \quad (10)$$

where the dispersion parameter ϕ is again equal to the squared CV of the individual wet-day precipitation amounts.

In the second method, the GLM is fitted directly to R_M itself. Then the mean and variance are given by Eqs. (7) and (8), but with ϕ equal to the squared CV of the monthly totals. This model is virtually identical to that used for monthly precipitation totals in Kilsby et al. (1998), except that in our study the logarithm of E(R_M) is linked to the predictor variables rather than to the logarithm of R_M itself.

c. Results

The same 28-yr period as for the rainfall occurrence data was considered. Four predictor variables were chosen: the logarithm of the specific humidity q_s near the surface, slp, and u and v. In contrast to rainfall occurrence, the amount of precipitation is thus related to a measure of the absolute humidity of the atmosphere rather than relative humidity. An attempt to use the 700-hPa specific humidity q_{700} failed because its effect on the wet-day precipitation amounts at Deuselbach was quite different from that at Bern and De Bilt. Moreover, several remarkable outliers, all of them equal to 32.67 g kg^{-1}, were found in the q_{700} NCEP–NCAR reanalysis data. Therefore the specific humidity q_s near the surface was considered instead. The use of $\log q_s$ rather than q_s in the rainfall amount models facilitates the determination of the effect of a relative change in q_s (see section 3d).

TABLE 6. Estimated regression coefficients and other parameters in the generalized linear models for daily and monthly wet-day rainfall. Here, $\hat{\phi}$ is the estimate of the dispersion parameter ϕ in Eqs. (8) and (10), and r_D and r_M are the correlation coefficients between the observed and predicted values of the daily and monthly average wet-day rainfall amounts, respectively. The standard error se of an estimated regression coefficient is the average standard error for the three stations, based on the approximate covariance matrix in section 8.3.6 of McCullagh and Nelder (1989).

Variable	Bern	Deuselbach	De Bilt	se
Individual wet-day rainfall amounts				
log q_s (g kg^{-1})	0.811	0.879	0.887	0.0423
slp (hPa)	−0.045	−0.050	−0.050	0.0022
u (m s^{-1})	0.022	0.031	0.019	0.0025
v (m s^{-1})	−0.052	0.004	−0.016	0.0033
$\hat{\phi}$	1.46	1.50	1.43	
r_D^2	0.17	0.13	0.16	
Monthly average wet-day rainfall amounts				
log q_s (g kg^{-1})	0.424	0.283	0.270	0.0676
slp (hPa)	−0.015	−0.031	−0.031	0.0059
u (m s^{-1})	0.019	0.029	0.012	0.0071
v (m s^{-1})	0.002	0.001	−0.015	0.0113
$\hat{\phi}$	2.69	2.42	2.52	
r_M^2	0.15	0.14	0.15	

TABLE 7. Same as in Table 6 but for wet-day rainfall at De Bilt for the extended winter (Nov–Mar) and extended summer (May–Sep) seasons.

Variable	Winter	se	Summer	se
Individual wet-day rainfall amounts				
log q_s (g kg^{-1})	1.039	0.1127	0.806	0.1359
slp (hPa)	−0.044	0.0021	−0.061	0.0046
u (m s^{-1})	0.018	0.0031	0.003	0.0049
v (m s^{-1})	−0.009	0.0032	−0.027	0.0055
$\hat{\phi}$	1.37		1.50	
r_D^2	0.19		0.13	
Monthly average wet-day rainfall amounts				
log q_s (g kg^{-1})	0.222	0.2331	0.191	0.2900
slp (hPa)	−0.024	0.0059	−0.048	0.0139
u (m s^{-1})	0.021	0.0080	0.003	0.0150
v (m s^{-1})	−0.010	0.0118	−0.017	0.0216
$\hat{\phi}$	2.06		2.60	
r_M^2	0.26		0.10	

Table 6 shows the results for daily and monthly wet-day rainfall amount models with constant regression coefficients over the year. Most predictors in the daily model are significant at the 1% level. The only exception is the v velocity in the model for Deuselbach. The statistical significance of the regression coefficients is weaker for the monthly wet-day averages. Nevertheless, logq_s, slp, and u are significant at the 5% level for each station. As for the rainfall occurrence data, the estimated regression coefficients and their standard errors depend on the aggregation level. The absolute values of the estimated regression coefficients of logq_s, slp, and u in the model for the individual wet days are always larger than those in the model for the monthly wet-day averages. The standard errors of the estimated regression coefficients of slp, u, and v increase by about a factor of 3 if the model is applied to the monthly wet-day averages instead of the individual wet-day rainfall amounts. For logq_s, the increase in standard error is smaller, which can be understood from its stronger seasonal cycle (appendix A). There is also a marked difference between the values of the dispersion parameter ϕ in the daily and monthly models. In contrast with Table 2 for the rainfall occurrence data, there is little difference between the values of r_D^2 and r_M^2. This is partly because of the weaker temporal correlation in the wet-day rainfall data.

Apart from the nonlinearity of the link function, the averaging of the predictor variables over all days in the month of concern also causes differences between the estimated regression coefficients in the daily and monthly models. Furthermore, for the regression coefficient of logq_s it matters whether the monthly wet-day averages are based on the observed number of wet days or the expected number of wet days from the logistic model. The systematic differences in the estimated regression coefficients for logq_s and slp are much smaller if the monthly totals R_M are divided by the observed number of wet days and the predictor values are averaged over these wet days only. This also leads to a large reduction of the dispersion parameter and an increase of the squared correlation coefficient r_M^2. However, the application of such a downscaling model requires that a GCM simulates the occurrence of wet and dry days adequately, which is as yet not the case.

The systematic differences between the estimated regression coefficients for logq_s and slp in the GLMs for daily and monthly wet-day rainfall are also found for the seasonally stratified data from De Bilt (Table 7). The regression coefficient of logq_s in the model for the monthly wet-day averages is no longer statistically significant at the 5% level, due to a disproportional increase in its standard error. Table 7 further shows that the u velocity has no effect on wet-day rainfall in the summer season. As is the case for rainfall occurrence models (section 2c), evaluation of seasonally stratified data does not lead to other conclusions regarding the differences between daily and monthly GLMs.

Table 8 presents the results for the fit to the monthly precipitation totals R_M, assuming constant statistical relationships over the year. With the exception of the v velocity in the model for Deuselbach, the predictors are

TABLE 8. Same as in Table 6 but for monthly precipitation totals R_M.

Variable	Bern	Deuselbach	De Bilt	se
Monthly precipitation totals				
log q_s (g kg^{-1})	0.430	0.365	0.403	0.0710
slp (hPa)	−0.063	−0.069	−0.070	0.0062
u (m s^{-1})	0.082	0.099	0.059	0.0076
v (m s^{-1})	−0.057	−0.017	−0.024	0.0124
$\hat{\phi}$	0.22	0.17	0.19	
r_M^2	0.40	0.46	0.48	

significant at the 5% level. As in the fit to the monthly wet-day averages, the estimated regression coefficients of the moisture predictor in Table 8 are much smaller than those in the fit to the individual wet-day precipitation amounts as a result of averaging over both wet and dry days. The regression coefficients of the dynamical predictors slp, u, and v are, however, relatively large for the monthly precipitation totals. These variables are powerful predictors of the number M_W of wet days, which is an important component of the monthly precipitation totals. A relatively large influence of the dynamical predictors in the model for the monthly precipitation totals is also found for the seasonally stratified data from De Bilt.

d. The effect of a systematic increase in atmospheric humidity

In a warmer climate, the atmosphere can contain more moisture. The saturated vapor pressure increases by about 7% °C^{-1} according to the Clausius–Clapeyron relation, and similar changes in the specific humidity have been found in GCM simulations with increased greenhouse gas concentrations (e.g., Mitchell and Ingram 1992; Semenov and Bengtsson 2002). From Eq. (7) it follows that a relative increase of q_s of 10% results in a relative change in E(R) of $f = 1.1^{a_1}$, where a_1 is the regression coefficient of log q_s. This factor applies both to the daily and monthly models. For De Bilt, $f = 1.088$ in the fit to the individual wet-day rainfall amounts from the entire year ($a_1 = 0.887$), $f = 1.026$ in the fit to the monthly average wet-day rainfall amounts from the entire year ($a_1 = 0.270$), and $f = 1.039$ in the fit to the monthly precipitation totals from the entire year ($a_1 = 0.403$). A similar dependence of f on aggregation level is found for Bern, Deuselbach, and the seasonally stratified data from De Bilt. Thus a daily model turns out to be much more sensitive to changes in specific humidity than any of the monthly models. The reason for this behavior is that the values for the monthly models are biased because of the impact of dry days on the estimate of the regression coefficient a_1. This implies that the generation of precipitation scenarios for a future climate should preferably be based on daily rather than monthly models.

4. Aggregated rainfall amounts from daily and monthly downscaling models

Table 8 shows that the model for the monthly precipitation totals explains nearly half of the variance of R_M ($r_M^2 \approx 0.48$). Apart from the direct fit, R_M can also be estimated from the daily models for rainfall occurrence and wet-day rainfall as

$$R_M = \sum_t \pi_t \mathrm{E}(R_t), \qquad (11)$$

where π_t is the probability that day t is wet as determined from Eq. (1), E(R_t) is the expected rainfall amount for day t from Eq. (7), and t runs over all days in the month concerned. Another estimate of R_M can be obtained by multiplying the expected number of wet days E(M_W) from Eq. (3a) with the predicted average monthly wet-day rainfall amount $\overline{R}$. Table 9 presents the squares of the correlation coefficient between the observed and predicted monthly precipitation totals. These correlation coefficients are somewhat larger for the predictions from the models fitted to rainfall occurrence and wet-day rainfall amount than are those for the direct fit to the monthly precipitation totals. The values of r_M^2 for the monthly precipitation totals in Table 9 are much larger than the corresponding values of r_D^2 for the wet-day precipitation amounts in Table 6. As for rainfall occurrence, the aggregation of the predicted daily precipitation amounts to monthly precipitation amounts leads to a considerable increase in the proportion of explained variance.

TABLE 9. Squared correlation coefficients, r_M^2, between the observed and predicted monthly precipitation totals from different downscaling models.

Downscaling model	Bern	Deuselbach	De Bilt
Daily occurrence and amount models	0.44	0.46	0.57
Monthly occurrence and wet-day amount models	0.44	0.49	0.55
Monthly totals model	0.40	0.46	0.48

Figure 1 presents the observed annual precipitation totals at the three stations, together with those obtained by aggregating the expected precipitation amounts from the downscaling relationships. There is little difference between the values from the various downscaling models. All of them show for Bern an overestimation of precipitation till the end of the 1970s and an underestimation in the 1990s. For De Bilt there is an overestimation in the 1970s and an underestimation in the 1990s. These discrepancies can partly be attributed to nonhomogeneities in the NCEP–NCAR reanalysis data. While the model process in the NCEP–NCAR reanalysis did not change over time, trends may still be found in the model output, mainly as a result of variations in the amount or quality of the input data (Reid et al. 2001). After 1985, a systematic decrease in the annual mean 700-hPa relative humidity of about 5% is observed at the grid point near De Bilt. This decrease corresponds with a similar decrease in the observed 700-hPa relative humidity at De Bilt due to the use of another humidity sensor in the radiosondes from February 1985. The NCEP–NCAR reanalysis data near Bern show a systematic decrease in rh from 65% in 1968 to 55% in 1995, and positive jumps in the mean sea level pressure (≈3 hPa) around 1981 and the v velocity (≈1 m s^{-1}) around 1977.

Although the various downscaling models show only slight differences with respect to the explained variance of the monthly totals, their response to a systematic

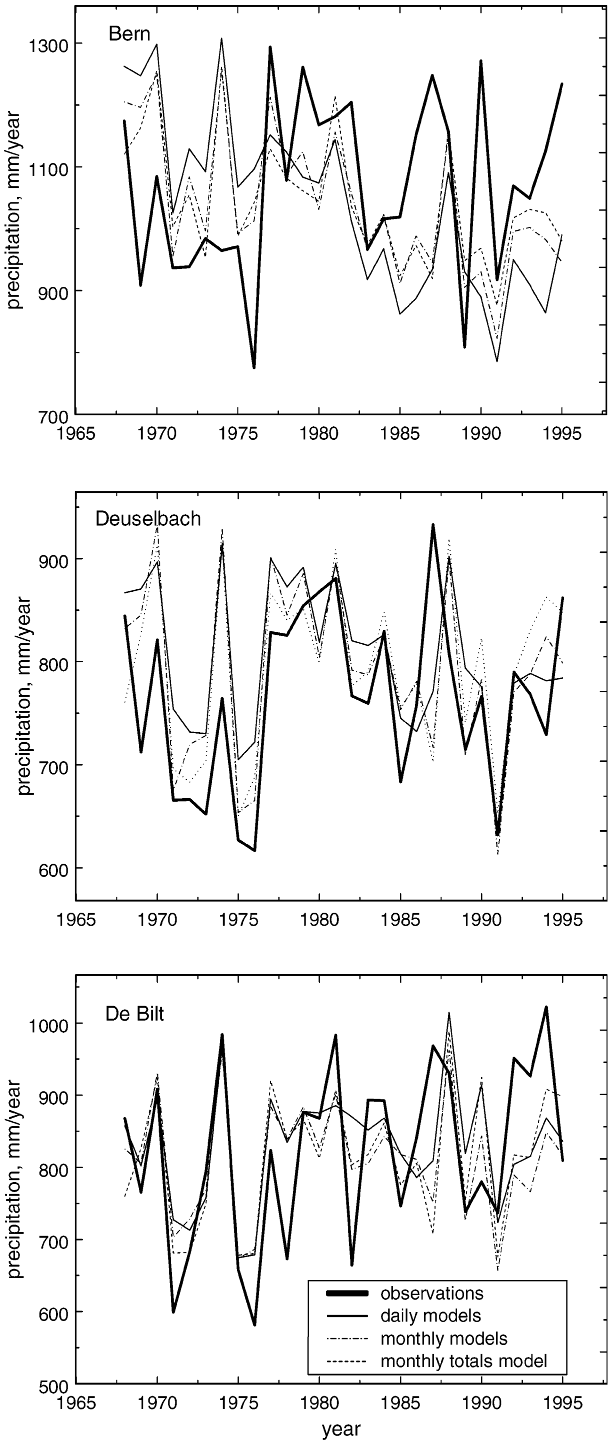

FIG. 1. Observed annual precipitation amounts at (top) Bern, (middle) Deuselbach, and (bottom) De Bilt, and predicted annual amounts from three downscaling models for these stations. Predictions represented by the thin and dashed–dotted lines are based on independent modeling of rainfall occurrence and wet-day rainfall amounts. The dotted line is based on the predictions from a model fitted to the monthly totals.

TABLE 10. Change in the average annual precipitation amount (mm) at De Bilt for a constant increase in the sea level pressure of 1, 2, and 4 hPa, respectively, and for a constant increase in the *u* velocity or *υ* velocity of 1, 2, and 4 m s^{-1}, respectively.

Variable	Daily models*	Monthly models*	Monthly totals model
slp + 1	−53	−42	−55
slp + 2	−103	−83	−106
slp + 4	−194	−159	−198
u + 1	36	44	49
u + 2	73	89	102
u + 4	148	179	216
υ + 1	−26	−18	−20
υ + 2	−52	−35	−39
υ + 4	−103	−69	−76

* Separate models for rainfall occurrence and wet-day rainfall amounts.

change of the predictor variables is not necessarily the same. In section 3d it was already shown that the daily model is much more sensitive to an increase in specific humidity than the monthly models. Table 10 illustrates the sensitivity of the modeled average annual precipitation amount at De Bilt to a constant change in the sea level pressure, the *u* velocity, or the *υ* velocity. The estimated changes in average annual precipitation refer to the models with constant regression coefficients over the year. For a given perturbation of slp, the estimated changes from the various downscaling models differ by not more than 25%. The effect of a perturbation of the *u* velocity on average annual precipitation is smallest for the daily downscaling models, partly because of the relatively low sensitivity of the model for daily rainfall occurrence to changes in *u* (section 2d). Both for the *u* velocity and the *υ* velocity the difference in the response can be as large as 50%.

5. Discussion and conclusions

GLMs were used to link the occurrence of wet and dry days, the wet-day precipitation amounts, and the monthly precipitation totals to dynamical variables and atmospheric moisture. It was shown that the standard errors of the regression coefficients are generally much smaller in the daily models than in the corresponding monthly models. As a consequence, the risk of discarding an important predictor as statistically not significant is relatively low in the daily models. Furthermore, the regression coefficients in the daily rainfall occurrence and amount models are larger than are those in the corresponding monthly models for the number of wet days and wet-day rainfall. This does not automatically result in a larger sensitivity of the daily models to a change in the predictor variables because of the nonlinearity of the link functions. The effect of a systematic increase in specific humidity on the amount of precipitation is, however, more than twice as large in the daily models as it is in the monthly models. The low sensitivity of the monthly precipitation amount models is mainly the result of bias due to the averaging of the predictors over both wet and dry days. Because the increase of absolute humidity is a major factor in greenhouse-gas-induced climate changes, downscaling models based on daily data are preferable for climate change impact studies.

Daily downscaling models explain only a small proportion of the variance in daily precipitation. This has been thought to be a serious drawback of regression-based downscaling (Wilby et al. 2002). However, the proportion of the explained variance is larger for the monthly aggregation level as a result of persistence in the predictor variables. Statistical models for daily and monthly values explain about the same proportion of the variance of the monthly number of wet days and the monthly precipitation totals. They also have similar skill in predicting annual precipitation amounts.

The regression models considered in this paper were linear on the logistic scale (rainfall occurrence) or the logarithmic scale (wet-day rainfall and monthly precipitation totals). Daily models can be improved by including nonlinearity in the predictors (Beckmann and Buishand 2002). Although this has little influence on the proportion of explained variance, it strongly influences the reproduction of extremes in the present climate and the predicted changes of extremes for future climates. Another improvement of daily models may come from the use of separate models for warm days (mainly convective precipitation) and cool days (mainly widespread frontal rainfall). For monthly totals, it is hardly feasible to apply this approach.

The predictor variables were generally significant at the 5% level, and in the daily models they were quite often also significant at the 1% level. This made it possible to demonstrate the effect of temporal aggregation clearly. Nevertheless, the choice of predictor variables could be questioned for real climate change applications at the three sites. Using precipitable water and relative humidity instead of specific humidity in the model for the wet-day rainfall amounts improved, for instance, the agreement between the changes in seasonal mean rainfall over the Netherlands and north Germany in a simulation experiment of the coupled ECHAM4/Ocean and Isopycnal Coordinates (OPYC3) atmosphere–ocean model and the estimated changes from the statistical relationships (Beckmann and Buishand 2002). The ability to reproduce the changes in precipitation in climate model simulations needs further research. It may also be advantageous to use different predictor variables for the three rainfall stations. A threshold of 0.1 mm was used to discriminate between wet and dry days. The sensitivity of the predicted changes in precipitation to the choice of this threshold needs some attention in climate change applications.

For two of the three stations the observed annual precipitation amounts were systematically underestimated or overestimated during some subperiods. These biases are partly caused by nonhomogeneities in the predictor variables. In fact, the downscaling relationships should be fitted to a homogeneous subset or the data should be corrected for nonhomogeneities. The first option is more easily achieved with statistical models for daily data because of the relatively low standard error of the estimated regression coefficients in these models. Corrections for nonhomogeneities are however, easier for monthly data.

Acknowledgments. The authors wish to thank G. P. Können for his help with the revision of an earlier version. They are also grateful to the Climatic Research Unit, University of East Anglia (Norwich), for extracting a subset of the NCEP–NCAR reanalysis data for a European window. The rainfall data for Bern were provided by the Swiss Meteorological Institute (Zürich). The rainfall data for Deuselbach were made available by the German Weather Service (Offenbach am Main) via the meteorological database of the International Commission for the Hydrology of the Rhine basin (CHR/KHR). This research was, in part, supported by the EU Environment and Climate Research programme (Contract EVK1-CT-2000-00075, SWURVE).

APPENDIX A

The Effect of Aggregation on Standard Errors

In this appendix, we explain how the standard error of an estimated regression coefficient may change with the temporal aggregation level. The change is explored in some detail for the classical linear regression model:

$$z_t = a_0 + a_1 x_t + e_t, \qquad t = 1, \ldots, N. \tag{A1}$$

It is assumed that the errors e_t are independent random variables with a mean of zero and standard deviation of σ_e.

The least squares estimates of the regression coefficients a_0 and a_1 are given by

$$\hat{a}_0 = \bar{z} - \hat{a}_1 \bar{x} \quad \text{and} \tag{A2a}$$

$$\hat{a}_1 = \frac{\sum_{t=1}^{N} (x_t - \bar{x})(z_t - \bar{z})}{\sum_{t=1}^{N} (x_t - \bar{x})^2}, \tag{A2b}$$

where $\bar{x}$ and $\bar{z}$ are the averages of the x_t and z_t values, respectively. For the variance of $\hat{a}_1$, it can be shown that (see, e.g., Weisberg 1985)

$$\text{var}(\hat{a}_1) = \frac{\sigma_e^2}{\sum_{t=1}^{N} (x_t - \bar{x})^2} = \frac{\sigma_e^2}{N s_x^2}, \tag{A3}$$

where s_x^2 is the sample variance of the x_t values.

We now aggregate the x_t and z_t values over blocks of length M. Let $n = N/M$ be the number of blocks. Then we have the following model for the aggregated values:

$$Z_\tau = a_0 + a_1 X_\tau + E_\tau, \qquad \tau = 1, \ldots, n, \quad \text{(A4)}$$

where

$$Z_\tau = \sum_{t=(\tau-1)M+1}^{\tau M} z_t, \quad \text{(A5a)}$$

$$X_\tau = \sum_{t=(\tau-1)M+1}^{\tau M} x_t, \quad \text{and} \quad \text{(A5b)}$$

$$E_\tau = \sum_{t=(\tau-1)M+1}^{\tau M} e_t. \quad \text{(A5c)}$$

The use of the aggregated values X_τ of the predictor variables instead of their averages $\overline{X}_\tau = X_\tau/M$ as in the main text does not affect the conclusions about the effect of aggregation on standard errors in this appendix and on the estimated regression coefficients in appendix B.

Because the errors e_t have been assumed to be independent with variance σ_e^2, we have

$$\text{var}(E_\tau) = \sigma_E^2 = M\sigma_e^2. \quad \text{(A6)}$$

For the least squares estimate $\hat{a}_{1,\text{aggr}}$ of the regression coefficient a_1 in Eq. (A4), an expression similar to Eq. (A2b) holds. The variance of $\hat{a}_{1,\text{aggr}}$ is given by

$$\text{var}(\hat{a}_{1,\text{aggr}}) = \frac{\sigma_E^2}{n s_X^2}, \quad \text{(A7)}$$

where s_X^2 is the sample variance of the X_τ values.

The change in standard error with the temporal aggregation level is determined by the relationship between the sample variances in the denominator of Eqs. (A3) and (A7). If the x_t values are independent realizations of a random variable (no temporal structure), then

$$s_X^2 \approx M s_x^2. \quad \text{(A8)}$$

Hence

$$\text{var}(\hat{a}_{1,\text{aggr}}) \approx \frac{\sigma_e^2}{n s_x^2} = M\,\text{var}(\hat{a}_1). \quad \text{(A9)}$$

So in this case the standard error of the estimated regression coefficient increases by a factor of $\sqrt{M}$ after aggregation over blocks of length M. When there is positive autocorrelation in x_t, then

$$s_X^2 > M s_x^2, \quad \text{(A10)}$$

and aggregation results in a smaller than $\sqrt{M}$ increase in the standard error. The stronger the autocorrelation, or the greater the temporal structure in the x_t values, the smaller will be the effect of temporal aggregation on the standard error of the estimated regression coefficient. An interesting case is that of a linear trend, $x_t = t$. Then $X_\tau = M^2\tau - M(M-1)/2$, and the sample variances of the x_t and X_τ values can be approximated as (Kendall et al. 1983, section 45.23)

$$s_x^2 = \frac{1}{12}(N^2 - 1) \approx \frac{1}{12}N^2 \quad \text{and} \quad \text{(A11a)}$$

$$s_X^2 = \frac{1}{12}(n^2 - 1)M^4 \approx \frac{1}{12}n^2 M^4 = M^2 s_x^2. \quad \text{(A11b)}$$

From Eqs. (A6), (A7), and (A11b), it follows that

$$\text{var}(\hat{a}_{1,\text{aggr}}) \approx \frac{\sigma_e^2}{nMs_x^2} = \text{var}(\hat{a}_1). \quad \text{(A12)}$$

Thus temporal aggregation has little effect on the standard error of the least squares estimate of a linear trend. Following the same arguments, it can be shown that the presence of a seasonal cycle in x_t also reduces the effect of temporal aggregation on the standard error of the estimated regression coefficient.

Similar effects of temporal aggregation may be expected in the case of multiple predictor variables. When these predictor variables are orthogonal, Eqs. (A3) and (A7) apply to each individual regression coefficient. For nonlinear models the magnitude of the change in the standard error due to aggregation may differ from that expected for the linear model. For instance, the observed increase of a factor 3 for the fit of the logistic model to the monthly number of wet days in the permutated daily record in section 2c is smaller than that expected for the linear model ($\sqrt{30} \approx 5.5$).

APPENDIX B

The Effect of Aggregation on Estimated Regression Coefficients

Under very general conditions, $\sum_{t=1}^{N} (x_t - \overline{x})(z_t - \overline{z})/N$ and $\sum_{t=1}^{N} (x_t - \overline{x})^2/N$ converge to the corresponding population moments. Thus, for sufficiently large N, the least squares estimate $\hat{a}_1$ in Eq. (A2b) is approximately

$$\hat{a}_1 \approx \frac{\text{cov}(z_t, x_t)}{\text{var}(x_t)} = \rho_{z,x}(0)\frac{\sigma_z}{\sigma_x}, \quad \text{(B1)}$$

where $\rho_{z,x}(0)$ is the correlation coefficient between z_t and x_t, and σ_z and σ_x are the standard deviations of z_t and x_t, respectively.

Similarly, we have for the aggregated values

$$\hat{a}_{1,\text{aggr}} \approx \frac{\text{cov}(Z_\tau, X_\tau)}{\text{var}(X_\tau)}. \quad \text{(B2)}$$

For the second-order moments on the right-hand side of Eq. (B2), we can write

$$\text{cov}(Z_\tau, X_\tau) \approx \sigma_x\sigma_z \sum_{k=-M}^{M} (M - |k|)\rho_{z,x}(k) \quad \text{and} \quad \text{(B3)}$$

$$\text{var}(X_\tau) \approx \sigma_x^2 \sum_{k=-M}^{M} (M - |k|)\rho_{x,x}(k), \quad \text{(B4)}$$

where $\rho_{x,x}(k)$ is the correlation coefficient between x_t and x_{t+k} (lag k autocorrelation) and $\rho_{z,x}(k)$ is the correlation coefficient between z_t and x_{t+k} (lag k cross-correlation). If the errors e_t in Eq. (A1) are mutually independent, and also independent of x_t, then it is easily verified that

$$\rho_{z,x}(k) = \rho_{x,z}(0)\rho_{x,x}(k). \quad \text{(B5)}$$

Substituting (B5) into (B3) results in

$$\frac{\text{cov}(Z_\tau, X_\tau)}{\text{var}(X_\tau)} = \rho_{z,x}(0)\frac{\sigma_z}{\sigma_x} \quad \text{(B6)}$$

and there is no systematic difference between $\hat{a}_1$ and $\hat{a}_{1,\text{aggr}}$. The estimate $\hat{a}_{1,\text{aggr}}$ tends to be smaller than $\hat{a}_1$ if the lag k cross-correlation coefficients decrease faster with increasing k than the lag k autocorrelation coefficients, and tends to be larger than $\hat{a}_1$ if the opposite occurs. The use of the classical linear regression model (A1) is, however, questionable in such situations. It may be necessary to enter lagged predictor variables in the regression. This had little effect on the daily rainfall models considered in this paper. The nonlinearity of the link function in these models is another cause of systematic differences between the estimated regression coefficients for distinct aggregation levels.

If the errors e_t are independent and if there is positive autocorrelation in the predictor x_t, then there is also positive autocorrelation in the predictand z_t. From Eq. (B4) it follows that in that situation $\text{var}(Z_\tau) > M\,\text{var}(z_t)$. Hence,

$$\frac{\text{var}(E_\tau)}{\text{var}(Z_\tau)} < \frac{\text{var}(e_t)}{\text{var}(z_t)}, \quad \text{(B7)}$$

which implies that the proportion of explained variance is larger for the aggregated data.

APPENDIX C

Change in the Number of Wet Days Due to a Constant Change in a Predictor Variable

For the probability π_t that day t is wet, it follows from Eq. (1) that

$$\pi_t = \frac{1}{1 + \exp(-a_{0,D} - a_{1,D}x_1 - \cdots - a_{p.D}x_p)}. \quad \text{(C1)}$$

The change in π_t resulting from a perturbation Δx_k of the kth predictor variable can be approximated as

$$\Delta\pi_t \approx \frac{\partial \pi_t}{\partial x_k}\Delta x_k = a_{k,D}\pi_t(1 - \pi_t)\Delta x_k. \quad \text{(C2)}$$

Summation of $\Delta\pi_t$ over all N days gives the expected change in the number of wet days:

$$\Delta N_W = \sum_{t=1}^{N} \Delta\pi_t \approx a_{k,D}C_D\Delta x_k, \quad \text{(C3)}$$

where

$$C_D = \sum_{t=1}^{N} \pi_t(1 - \pi_t). \quad \text{(C4)}$$

The product $\pi_t(1 - \pi_t)$ is relatively large if π_t is in the interval (0.3, 0.7) and small if π_t is close to 0 or close to 1. Days with a low or high probability of rain thus give a small contribution to ΔN_W.

In the same way, we have for the monthly rainfall occurrence model

$$\Delta N_W \approx a_{k,M}C_M\Delta x_k, \quad \text{(C5)}$$

where

$$C_M = \sum_{t=1}^{n} M_\tau\Pi_\tau(1 - \Pi_\tau), \quad \text{(C6)}$$

with M_τ the total number of days in month τ and n the total number of months. Because the most extreme values of Π_τ are farther away from 0 and 1 than those of π_t, the quantity C_M is larger than C_D.

REFERENCES

Beckmann, B.-R., and T. A. Buishand, 2002: Statistical downscaling relationships for precipitation in the Netherlands and north Germany. *Int. J. Climatol.,* **22,** 15–32.

Buishand, T. A., 1978: Some remarks on the use of daily rainfall models. *J. Hydrol.,* **36,** 295–308.

——, and B.-R. Beckmann, 2000: Development of daily precipitation scenarios at KNMI. *Climate Scenarios for Water-Related and Coastal Impacts,* J. Beersma et al., Eds., ECLAT-2 Workshop Rep. 3, Climatic Research Unit, University of East Anglia, 79–91. [Available from Climatic Research Unit, University of East Anglia, Norwich NR4 7TJ, United Kingdom.]

——, and A. M. G. Klein Tank, 1996: Regression model for generating time series of daily precipitation amounts for climate change impact studies. *Stochastic Hydrol. Hydraul.,* **10,** 87–106.

Coe, R., and R. D. Stern, 1982: Fitting models to daily rainfall data. *J. Appl. Meteor.,* **21,** 1024–1031.

Frei, C., and C. Schär, 2001: Detection probability of trends in rare events: Theory and application to heavy precipitation in the Alpine region. *J. Climate,* **14,** 1568–1584.

Giorgi, F., and Coauthors, 2001: Regional climate information—Evaluation and projections. *Climate Change 2001: The Scientific Basis,* J. T. Houghton et al., Eds., Cambridge University Press, 583–638.

Kalnay, E., and Coauthors, 1996: The NCEP/NCAR 40-Year Reanalysis Project. *Bull. Amer. Meteor. Soc.,* **77,** 437–471.

Katz, R. W., 1977: Precipitation as a chain-dependent process. *J. Appl. Meteor.,* **16,** 671–676.

Kendall, M., A. Stuart, and J. K. Ord, 1983: *The Advanced Theory of Statistics.* Vol. 3. 4th ed. Charles Griffin, 780 pp.

Kilsby, C. G., P. S. P. Cowpertwait, P. E. O'Connell, and P. D. Jones, 1998: Predicting rainfall statistics in England and Wales using atmospheric circulation variables. *Int. J. Climatol.,* **18,** 523–539.

Kistler, R., and Coauthors, 2001: The NCEP–NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.,* **82,** 247–267.

Lemcke, C., and S. Kruizinga, 1988: Model output statistics forecasts: Three years of operational experience in the Netherlands. *Mon. Wea. Rev.,* **116,** 1077–1090.

McCullagh, P., and J. A. Nelder, 1989: *Generalized Linear Models.* 2d ed. Chapman and Hall, 511 pp.

Merenti-Välimäki, H.-L., and P. Laininen, 2002: Analysing effects of meteorological variables on weather codes by logistic regression. *Meteor. Appl.,* **9,** 191–197.

Mitchell, J. F. B., and W. J. Ingram, 1992: Carbon dioxide and climate: Mechanisms of changes in cloud. *J. Climate,* **5,** 5–21.
Murphy, J., 1999: An evaluation of statistical and dynamical techniques for downscaling local climate. *J. Climate,* **12,** 2256–2284.
——, 2000: Predictions of climate change over Europe using statistical and dynamical downscaling techniques. *Int. J. Climatol.,* **20,** 489–501.
Reid, P. A., P. D. Jones, O. Brown, C. M. Goodess, and T. D. Davies, 2001: Assessments of the reliability of NCEP circulation data and relationships with surface climate by direct comparisons with station based data. *Climate Res.,* **17,** 247–261.
Semenov, V. A., and L. Bengtsson, 2002: Secular trends in daily precipitation characteristics: Greenhouse gas simulation with a coupled AOGCM. *Climate Dyn.,* **19,** 123–140.
Stern, R. D., and R. Coe, 1984: A model fitting analysis of daily rainfall data (with discussion). *J. Roy. Stat. Soc.,* **147A** (1), 1–34.
Weisberg, S., 1985: *Applied Linear Regression.* 2d ed. John Wiley and Sons, 324 pp.
Wilby, R. L., 2001: Downscaling summer rainfall in the UK from North Atlantic Ocean temperatures. *Hydrol. Earth Syst. Sci.,* **5,** 245–257.
——, H. Hassan, and K. Hanaki, 1998: Statistical downscaling of hydrometeorological variables using general circulation model output. *J. Hydrol.,* **205,** 1–19.
——, C. W. Dawson, and E. M. Barrow, 2002: SDSM—A decision support tool for the assessment of regional climate change impacts. *Environ. Modell. Software,* **17,** 147–159.