Modelling daily precipitation as a function of temperature for climate change impact studies

A.M.G. Klein Tank and T.A. Buishand

Scientific reports; WR 93-02

Wetenschappelijke rapporten; WR 93-02

De Bilt, 1993.

De Bilt 1993
Scientific reports = wetenschappelijke rapporten; WR 93-02

Postbus 201
3730 AE De Bilt
Wilhelminalaan 10
Telefoon 030-206 911
Telefax 030-210 407

UDC: 551.501.777
551.524.31
551.577.31
551.583

ISBN: 90-369-2039-6

ISSN: 0169-1651

MODELLING DAILY PRECIPITATION AS A FUNCTION OF TEMPERATURE FOR CLIMATE CHANGE IMPACT STUDIES

A.M.G. Klein Tank and T.A. Buishand
Section of Climate Scenarios and Ozone
Royal Netherlands Meteorological Institute KNMI

Abstract

A method is presented to derive time series of daily precipitation and temperature in a possible future climate at a certain location from the observed time series of these variables. The method considers the relation between daily precipitation and temperature. From this relation a temperature dependent factor *F* is obtained which gives the relative change in daily precipitation due to a given temperature perturbation. Multiplication of the observed daily amounts with this factor yields a consistent scenario for daily precipitation and temperature. The number of rain days does not change in this method.

The observed precipitation-temperature relation for rain days at De Bilt (The Netherlands) is examined. This relation is described by a non-linear regression model using spline functions to reproduce the non-monotonous change of the mean amount with temperature. The standard deviation of the daily amounts increases with temperature, but the coefficient of variation is almost constant. This supports the use of a factor to transform the observed amounts toward a possible future climate. For a temperature increase the factor *F* derived from the model is generally greater than 1. *F* is, however, less than 1 when the daily mean temperature is around 12°C.

The effect of a constant temperature increase of 3°C on the mean precipitation amounts at De Bilt was studied with different versions of the model. During winter the monthly mean precipitation amounts increase by about 20% when seasonal variation in the precipitation-temperature relation is ignored. For a seasonally varying model the increase in mean winter precipitation amounts to 27%. The changes in the monthly means are much smaller for the other seasons. The increase in annual mean precipitation ranges between 7 and 9% for the various models fitted to the observed precipitation-temperature relation. This is consistent with the change in global mean precipitation predicted by present day general circulation models in case of a temperature increase of 3°C.

Contents

1 Introduction

There is a general concern about the global warming resulting from the increase of greenhouse gases in the atmosphere. General Circulation Models (GCMs) predict that besides this warming there are also considerable changes in precipitation. The increase of 3-15% in global mean precipitation predicted by these models for the $2\times CO_2$-case is usually attributed to the temperature effect on atmospheric moisture. In general, the greater the warming, the greater the enhancement of the hydrological cycle (Houghton et al., 1990).

Although the various GCMs give already quite different estimates for the increase in global mean precipitation, the range in their predictions is even much larger for the change in precipitation over a particular region. Furthermore, the actual regional precipitation climate is poorly reproduced in these models. The use of GCM predicted precipitation series for specific grid points is therefore not recommended. On the other hand local time series of daily precipitation are often required in climate change impact studies.

In this report the empirical relation between daily precipitation and air temperature is used for modelling the change in precipitation amount at the local scale. Given a certain temperature perturbation this relation provides a temperature dependent factor F by which observed daily precipitation amounts are multiplied to yield a time series for a possible future climate. The precipitation scenario derived this way assumes no change in the number and sequence of rain days. The scenario therefore precludes a systematic change in the atmospheric circulation. The time series analysed is the historical record of De Bilt in the Netherlands for the period 1906-1981. Unless stated otherwise a rain day is defined as a day receiving at least 0.1 mm of precipitation.

In Section 2 a statistical model for the relation between precipitation and daily mean air temperature is presented. The multiplication factor F for this model is derived in Section 3. As an example, a temperature perturbation of +3°C is considered. The monthly and annual mean amounts in the resulting precipitation scenario are compared with the observed means. In Section 4 the model is extended by incorporating the seasonal variation in the precipitation-temperature relation. The last two sections deal with the sensitivity of the results to the criterion for a rain day (Section 5) and the use of the daily maximum temperature as predictor instead of daily mean temperature (Section 6).

2 Statistical model for the relation between precipitation amount on rain days and daily mean temperature

The relation between daily precipitation and air temperature in the Netherlands has been discussed by Können (1983). Figure 1 presents mean precipitation at De Bilt for various values of the daily mean air temperature *T*. It is seen from the figure that the precipitation-temperature relation is rather complicated. At low temperatures, where precipitation is mainly caused by widespread frontal rain, the mean increases with temperature. The mechanism behind this is the earlier mentioned change in the maximum moisture content of the air. At very high temperatures there is also an increase only now convective precipitation (showers) prevails. In the interval from 14°C to 18°C the observed mean precipitation decreases with temperature because then fronts become less active with increasing temperature while the activity of showers is still low.

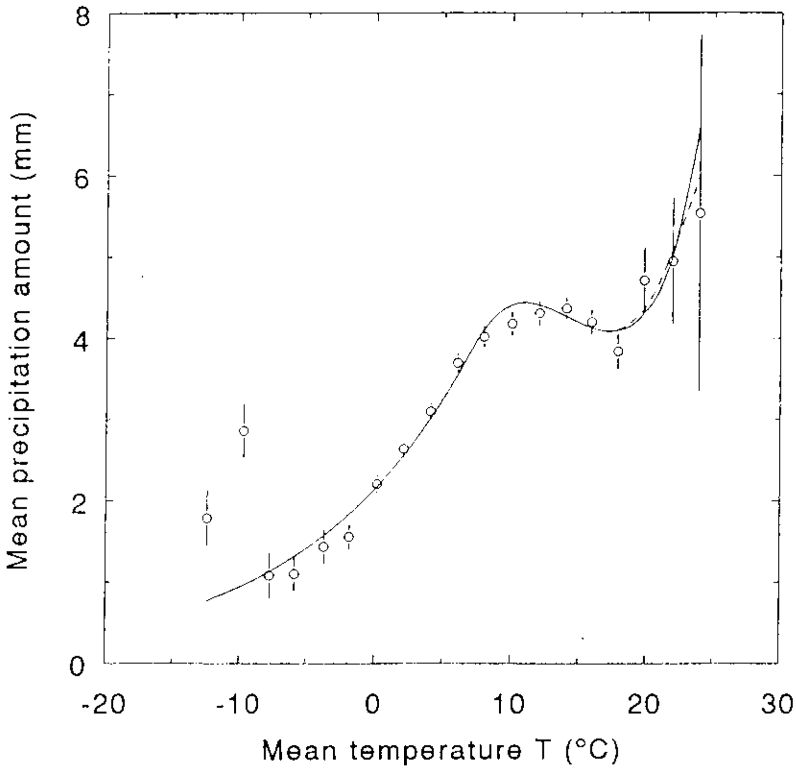

Figure 1 Mean precipitation amounts at temperature intervals of 2°C for rain days at De Bilt (1906-1981). The figure is based on 15,897 rain days (57% of the total number of days). The number of rain days in a temperature interval is as high as 2104 for T=6°C and decreases to about 10 at the extreme temperatures. The probability of rain is 70% for T between 4 and 8°C and only 30% or less for $T \leq$-4°C and $T \geq$20°C. The smooth curves are based on fitted regression models. The error bars indicate the standard deviations of the means (Appendix C).

Originally Können (1983) considered the daily maximum temperature T_{max} instead of the daily mean temperature. Mean temperature is, however, more frequently used in hydrological and crop growth models than maximum temperature. Publications on changes in temperature from GCM simulations also mainly deal with daily mean temperature. For these reasons in this report the emphasis is on daily mean temperature. The use of maximum temperature is discussed further in Section 6.

The observed relation between precipitation and temperature with local extremes around T=14°C and T=18°C can be described by a statistical model of the form:

$$R = \exp[g(T)] + \varepsilon \qquad (1)$$

where R is the precipitation amount on a rain day with mean temperature T, $g(T)$ is a non-linear function of T and ε is a random error term with mean zero. The first term in the right hand side (systematic part) represents the theoretical mean amount on a rain day with temperature T:

$$\mathrm{E}(R) = \exp[g(T)] \qquad (2)$$

The function $g(T)$ is necessary to describe the non-monotonous change of mean precipitation with temperature as discussed above. In the absence of an abrupt transition between widespread rainfall and convective precipitation $g(T)$ should also be a continuous function. The use of the exponential function in equation (1) avoids negative values for the mean amount.[1]

In this study the function $g(T)$ consists of piecewise polynomials (spline functions) that are flexible enough to obtain a satisfactory description of the relation between R and T. In the first model, Model 1, this function is taken as:

$$g(T) = a + bT \qquad T \le m_1$$

$$g(T) = a + bT + c(T-m_1)^2 + d(T-m_1)^3 \qquad T > m_1 \qquad (3)$$

The solid line in Figure 1 is based on this model. The function and its derivative are continuous over the full temperature range. The knot m_1 has been fixed a priori at T=7°C. The linear form of $g(T)$ for $T<m_1$ is approximately consistent with the relation between saturated vapour pressure e_s and temperature (Clausius Clapeyron relation). The cubic polynomial for

---

1 The true mean amount is in fact also larger than the threshold δ defining a rain day. Formally the mean precipitation amount should be treated as a conditional mean, here the mean of R given that $R \ge 0.1$ mm (or more precisely $R > 0.05$ mm, because daily rainfall has been recorded in units of 0.1 mm). In the literature on time series analysis of daily rainfall data probability distributions are usually fitted to the shifted amounts $\tilde{R} = R - \delta$ (Buishand, 1977; Stern and Coe, 1984). From a physical point of view it is, however, more elegant to model the mean of R by an exponential function rather than the mean of $\tilde{R}$. The use of the exponential function in the regression equation ensures that $\mathrm{E}(R)$ is well above the 0.1 mm threshold in the observed range of temperatures.

$T>m_1$ is necessary to reproduce the local maximum (near T=14°C) and the local minimum (near T=18°C). A consequence of this cubic polynomial is that $g(T)$ very rapidly increases at high temperatures. A less sharp rise can be achieved by introducing a second knot $m_2>m_1$ after which $g(T)$ is taken again linear in T (Model 2). Continuity of $g(T)$ and its derivative require that for $T>m_2$:

$$g(T) = a + bT + c(m_2-m_1)(2T-m_2-m_1) + d(m_2-m_1)^2(3T-2m_2-m_1) \qquad T>m_2 \tag{4}$$

Note that the additional knot does not lead to more unknown coefficients. The dashed line in Figure 1 is based on this second model; the knots m_1 and m_2 are fixed a priori at T=7°C and T=21°C.

In the statistical literature regression models with spline functions are known as regression splines (Smith, 1979; Wahba, 1989). These functions can be introduced in the regression without difficulties. It simply implies that the model contains a number of temperature dependent explanatory variables defined by the various terms in the right-hand sides of equations (3) and (4), whereas R is the dependent variable. There are, however, two factors that complicate parameter estimation. First, equation (1) is non-linear in the regression coefficients a,b,c and d because of the exponential function, and second, the standard deviation $\sigma(R)$ of R increases with temperature as is shown in Figure 2.

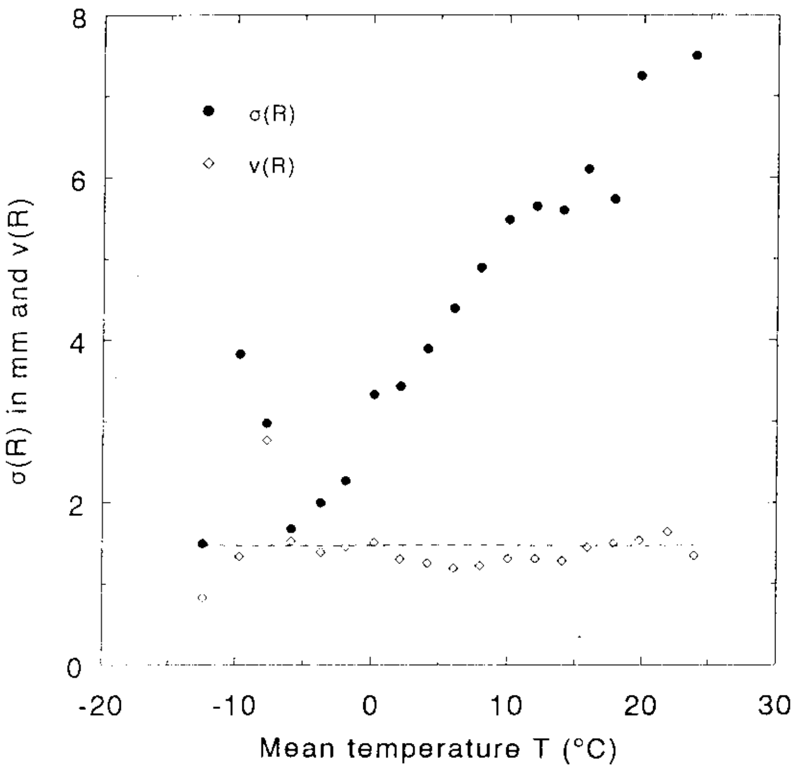

Figure 2 Standard deviation $\sigma(R)$ and coefficient of variation v(R) of daily precipitation R at temperature intervals of 2°C at De Bilt (1906-1981). The dashed line shows the mean of v(R).

From the figure it is also seen that the coefficient of variation $v(R)=\sigma(R)/E(R)$ remains constant. A popular method to get rid of the non-linearity in a situation like this is to fit the function $g(T)$ to the logarithms of the daily precipitation amounts. The method is, however, sensitive to roundings of small precipitation amounts and the mean of R then depends on the distribution of the error term in the regression model (Appendix D). Therefore, it was decided not to use this method but instead to fit the model directly to the mean precipitation amounts in Figure 1 by an iteratively reweighted least squares procedure (Appendix C). The use of temperature classes makes it possible to test for lack-of-fit. The results of the test in Appendix C reveal that the two lowest temperature classes should be discarded for parameter estimation. Long lasting snow events are responsible for the large mean precipitation amounts in these classes. The models also do not capture the relative extremes in Figure 1 very well, but this does not lead to a significant value of the test-statistic for lack-of-fit.

In Table 1 the estimated parameters for both Model 1 and Model 2 are presented. All parameter estimates are large compared to their standard errors. The largest relative standard errors (10-15%) are found for the estimates $\hat{c}$ and $\hat{d}$ of c and d, respectively. Figure 3 presents 95% and 99% confidence regions for these parameters, based on the quasi-likelihood function (Appendix C). The elliptic form of this region is in accordance with asymptotic theory. The oblique orientation of the axes is a result of the strong (negative) correlation between $\hat{c}$ and $\hat{d}$. Because of the elongated shape of the confidence region there is quite a large range of acceptable values of c and d. The consequences of the uncertainty in c and d are discussed in the next section.

Table 1 Estimated coefficients in equations (3) and (4) with their standard error (se).

	Model 1		Model 2	
Coefficient	Estimate	se	Estimate	se
a	0.7649	0.0260	0.7639	0.0261
b	0.0829	0.0043	0.0832	0.0044
c	-0.0144	0.0016	-0.0147	0.0017
d	6.7E-4	1.1E-4	7.0E-4	1.2E-4

It is interesting to compare the change in mean precipitation for $T<m$, with that in the saturated vapour pressure e_s. According to the Clausius Clapeyron relation, for this temperature range, the latter can be approximated by:

$$e_s \approx 6.107\exp\left[17.57\,T/(241.8+T)\right] \approx 6.107\exp\left[0.073\,T\right] \qquad (mbar) \tag{5}$$

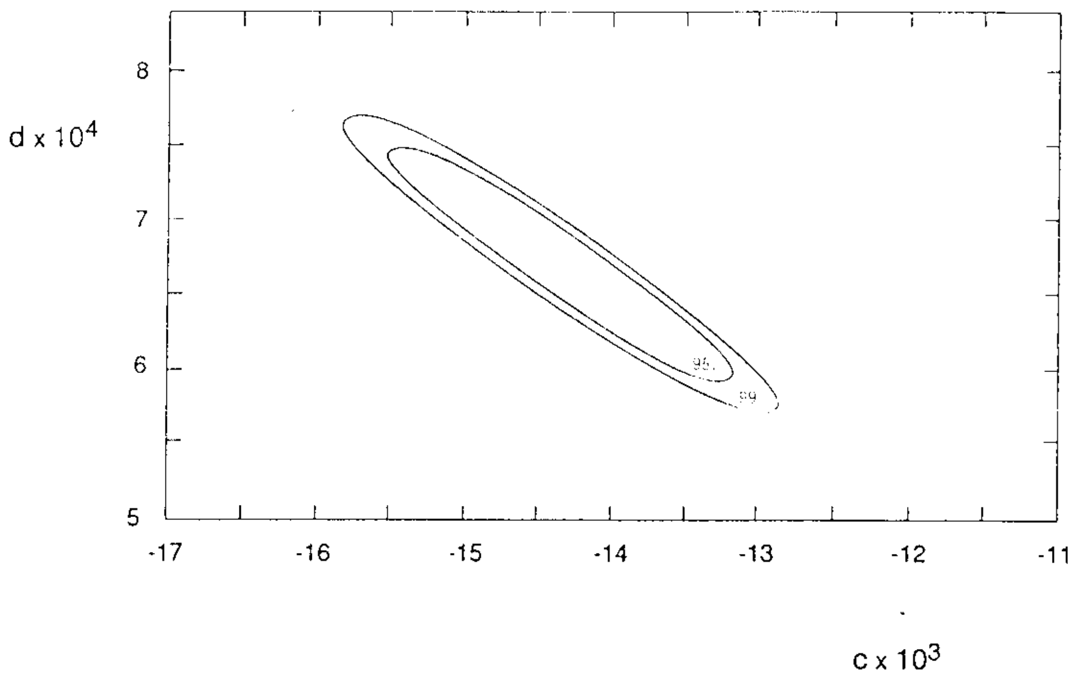

Figure 3 95% and 99% confidence regions for the parameters c and d in Model 1, based on the quasi-likelihood function (For details see Appendix C).

There is a reasonable correspondence between the estimate of *b* in Table 1 with the coefficient 0.073 in equation (5), which means that the relative changes in mean precipitation and e_s are of the same order of magnitude (about 8% for a temperature increase of 1°C). Equation (5) holds for the saturated vapour pressure over water; for the saturated vapour pressure over ice the coefficient in the exponent of equation (5) becomes 0.082, which is also consistent with the estimate of *b* in Table 1.

3 Precipitation scenario based on a temperature perturbation

In the previous section a model was fitted to the mean precipitation amounts at various temperatures. This model can be used to generate a time series of daily precipitation amounts for a possible future climate. In this study it is assumed that in the future climate $g(T)$ and the coefficient of variation remain the same. Because the latter is independent of temperature, the relative changes in the mean and standard deviation of the daily amounts are equal when daily temperature changes from T to T^*. A rainfall sequence for the perturbed climate can then be obtained by multiplying all daily values in an observed record by a temperature dependent factor F.

The factor is easily obtained by comparing the theoretical mean amount $\mathrm{E}(R)$ in equation (2) with the mean $\mathrm{E}(R^*)$ for the perturbed climate. The latter is given by:

$$\mathrm{E}(R^*) = \exp\left[g(T^*)\right] \tag{6}$$

and F becomes:

$$F = \exp\left[g(T^*) - g(T)\right] \tag{7}$$

The factor F is completely determined by the temperatures T and T^*. Substitution of the function $g(T)$ from equations (3) and (4) provides the value of F for any combination of T and T^*. For instance, in case of a warming climate and choosing Model 1 for $g(T)$ this factor becomes:

$$\begin{aligned}
F &= \exp\left[b(T^*-T)\right] && T,T^*\leq m_1 \\
F &= \exp\left[b(T^*-T) + c(T^*-m_1)^2 + d(T^*-m_1)^3\right] && T\leq m_1<T^* \\
F &= \exp\left[b(T^*-T) + c\{(T^*-m_1)^2 - (T-m_1)^2\} + d\{(T^*-m_1)^3 - (T-m_1)^3\}\right] && T,T^*>m_1
\end{aligned} \tag{8}$$

For Model 2 the factor is given in Appendix A.

As an example Table 2 presents the factor F for various values of T when there is a constant increase in mean temperature of 3°C. A value of 2.5 or 3°C is often chosen as the global average temperature increase after a doubling of the CO_2 concentration (Houghton et al., 1990, 1992). This value may be subject to change when results of more advanced GCMs become available. The current GCMs indicate that at high latitudes the temperature increase in winter is larger than in summer (Houghton et al., 1990). Such a seasonal difference on a regional scale or more local anomalies can also be taken into account. For the Netherlands one could for instance consider the seasonal changes in temperature over a part of Europe as given in Section 6.

Adding a constant to the daily temperatures as in Table 2, implies that the autocorrelation and the standard deviation in the temperature record remain unchanged. So far model experiments with doubled CO_2 give no clear indication of a systematic change in the variability of temperature on daily to interannual time-scales (Houghton et al., 1992). Recent analysis of daily temperatures in the Canadian Climate Centre (CCC) second-generation GCM (McFarlane et al., 1992; Boer et al., 1992) with newly developed test-statistics shows that there are no significant differences between the autocorrelation coefficients in the $1\times CO_2$ and $2\times CO_2$ cases for the grid points surrounding the Netherlands (Buishand and Beersma, 1993). This agrees with results in Rind et al. (1989) for a transient climate experiment.

Table 2 Multiplication factor *F* (with its standard error se) for various values of *T* when there is a constant increase in temperature of 3°C on every day ($T^* = T + 3°C$).

		Model 1		Model 2	
T	*T**	*F*	*se*	*F*	*se*
-12	-9	1.28	.02	1.28	.02
-10	-7	1.28	.02	1.28	.02
-8	-5	1.28	.02	1.28	.02
-6	-3	1.28	.02	1.28	.02
-4	-1	1.28	.02	1.28	.02
-2	1	1.28	.02	1.28	.02
0	3	1.28	.02	1.28	.02
2	5	1.28	.02	1.28	.02
4	7	1.28	.02	1.28	.02
6	9	1.22	.01	1.22	.01
8	11	1.08	.01	1.08	.01
10	13	.99	.02	.99	.02
12	15	.95	.02	.95	.02
14	17	.96	.02	.96	.02
16	19	1.02	.03	1.02	.03
18	21	1.13	.06	1.14	.07
20	23	1.32	.12	1.26	.10
22	25	1.61	.21	1.28	.11
24	27	2.07	.38	1.28	.11

Note that there is a temperature interval in which *F* is less than 1 and that *F* becomes very large (more than a doubling of precipitation) at extremely high temperatures in Model 1. The value of *F* does not change with *T* if $T<m_1$.

The standard error *se* in Table 2 quantifies the effect of sampling variability on the estimate of *F* as the result of the use of a finite number of observations (a 76 year time series). Because the correlation between daily precipitation amounts and temperature is very low (correlation coefficient in the order of 0.1) quite large samples are needed to keep this standard error within reasonable limits. Due to the great uncertainty in the parameters *c* and *d* large values of *se* are found at high temperatures. The rapid increase of the standard error in Model 1 at these temperatures is caused by the fact that the cubic term in equation (8) dominates the quadratic term. Details about the calculation of the standard error *se* are given in Appendix B. The standard error *se* refers to random errors. It is also likely that there are systematic errors because a regression model is no more than a tractable approximation to a complex real world relationship. The lack-of-fit test discussed in Appendix C shows no lack-of-fit for Models 1 and 2, but this test is unable to detect all systematic errors. Even quite large systematic errors near the ends of the observed temperature range can pass the test, because of the limited number of rain days at these temperatures. The differences in F between Models 1 and 2 give some indication about the magnitude of these errors. A further insight into the sensitivity of F to the form of the functional relation between R and T can be obtained by changing the values of the knots m_1 and m_2. Taking m_1=8°C instead of 7°C in Model 1 only results in a substantial increase in F at extremely high temperatures (e.g. at T=24°C F becomes 2.46 instead of 2.07). This is mainly due to the change in the coefficient d from 6.7E-4 to 8.6E-4. Since, for warm days, a scenario for a warmer world requires some extrapolation of the precipitation-temperature relation, the possibility of large systematic errors in the mean precipitation at high temperatures is of some concern. The consequences for monthly, seasonal or annual averages can, however, be neglected because there are only few rain days at high temperatures.

The daily amounts in the 1906-1981 record of De Bilt were multiplied by the temperature dependent factor F presented in Table 2. Figure 4 compares the monthly means of the new records with the observed means. Both Model 1 and Model 2 were considered to obtain a time series for the perturbed climate. The factor was also applied to the few rain days at very low temperatures, which were discarded in model calibration.

In September there is a slight decrease in the mean precipitation amount but in the other months the higher temperatures in the perturbed climate give rise to an increase in mean precipitation compared to the observations at De Bilt. In winter the increase amounts to 20 %, whereas in summer it is only 4%. These relative changes are in good agreement with those proposed by Bultot et al. (1988) for stations in Belgium and by Cole et al. (1991) for the UK. As expected Model 1 and Model 2 have almost the same annual cycle. In contrast to the minor differences in monthly means the differences in daily precipitation between the two models are quite substantial on warm days in summer.

The annual average increase in precipitation from 780 mm to 850 mm (9%) is considerably less than that expected from the Clausius Clapeyron relation (about 18% for a 3°C temperature increase which would give 920 mm). However, the smaller increase in the proposed scenario compares well with present day GCM simulations. The increase in global mean precipitation in these simulations is typically less than half that given by the Clausius Clapeyron relation (Mitchell, 1991).

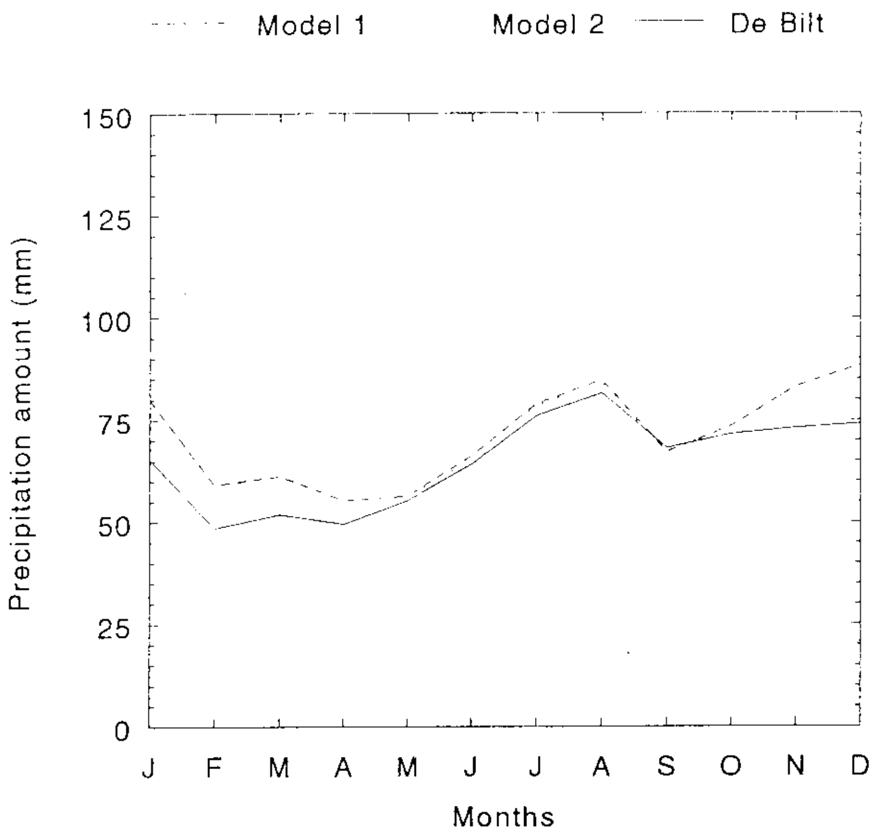

Figure 4 Monthly mean precipitation in a perturbed climate ($T^{*}=T+3°C$) from time series based on Model 1 and 2 compared to the climate of De Bilt (1906-1981).

Based on 17 GCMs, the Intergovernmental Panel on Climate Change indicates that precipitation increases with about 8% when there is a temperature increase of 3°C (Houghton et al., 1990). When mean temperature is increased by 2°C instead of 3°C, the use of the factor F in equation (8) yields a change in mean annual precipitation of 6%.

The major advantage of the method presented here to generate daily precipitation sequences for a possible future climate is that changes in daily precipitation are consistent with the temperature dependence of the two main precipitation mechanisms (frontal and convective precipitation) in the Netherlands. As a result it is possible that small changes in monthly mean precipitation during summer are accompanied by considerable increases in precipitation amounts on some specific days. The latter is necessary because a temperature increase leads to heavier summer showers when stability and relative humidity remain unchanged as is assumed in many scenarios for a double CO_2 climate.

4 Seasonal variation in the precipitation - temperature relation

Up till now it has been assumed that the relationship between precipitation and temperature is constant over the year. The absence of seasonal variation implies that the mean precipitation in the various temperature classes in Figure 1 remains the same throughout the year. This can be verified by comparing the expected mean precipitation amounts in each month to the observed means. The former are easily obtained by replacing all observed daily amounts by the average in their temperature class (circles in Figure 1). It is not necessary to select a function *g(T)* for this verification. From Figure 5 it is seen that the constant precipitation - temperature relation leads to an overestimate of the monthly mean precipitation amounts in spring (with a difference up to about 20% in April) and to relatively small monthly means in January, July, August and December.

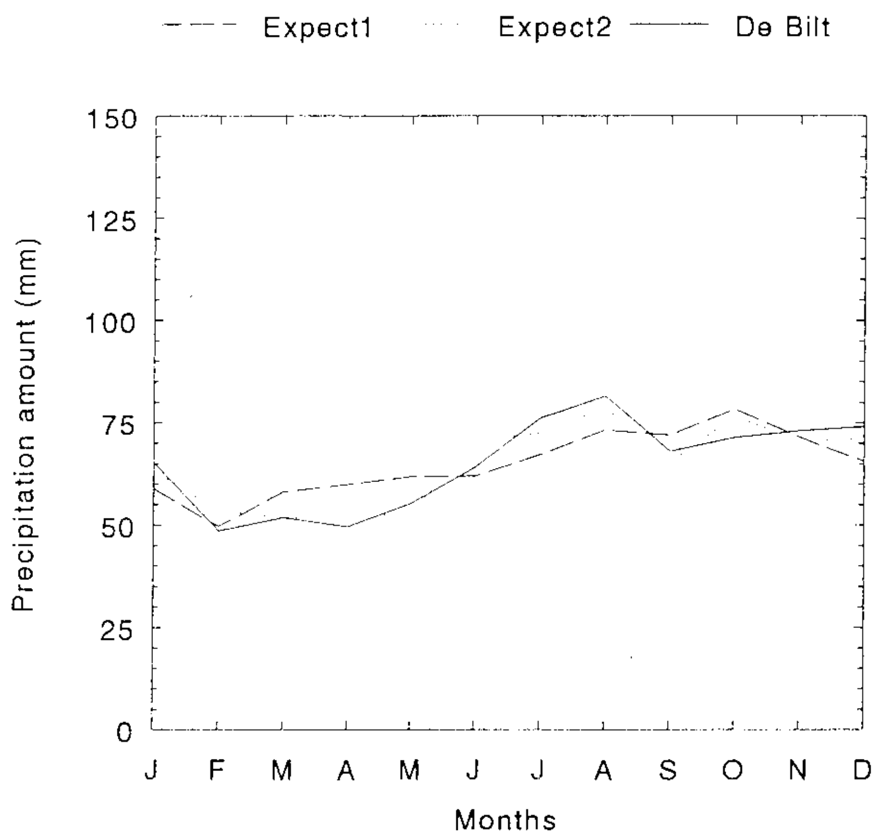

Figure 5 Expected monthly mean amounts for a constant precipitation-temperature relation (Expected 1) and for separate relations in three-month seasons (Expected 2) compared to the observed means at De Bilt (1906-1981).

Figure 5 also shows the expected monthly means for the case where all observed daily amounts are replaced by the seasonal average in their temperature class. These monthly means assume a separate relation between precipitation and temperature for each of the four seasons: winter (December, January, February), spring (March, April, May), summer (June, July, August), and autumn (September, October, November). The expected monthly mean precipitation is now much closer to the observations. The differences are always less than 10%. The seasonal dependence of the relation between precipitation and temperature is significant at the 5% level (Appendix C).

The seasonal dependence can be incorporated in the regression analysis by taking separate values a_{winter}, a_{spring}, a_{summer}, a_{autumn} and b_{winter}, b_{spring}, b_{summer}, b_{autumn} for the parameters a and b in equations (3) and (4). There is no indication of lack-of-fit for this 10 parameter model and the seasonal dependence in a and b is significant at the 5% level (Appendix C). Table 3 presents the estimated parameters with their standard errors. The estimate of a for the summer season is considerably larger than the estimates for winter, spring and autumn (about twice its standard error). The parameter b is relatively high in winter. The estimate of b_{winter} differs about five times its standard error from those for the three other seasons.

Table 3 Estimated regression coefficients with their standard error (se) in Model 1 with seasonal variation.

Coefficient	Estimate	se
a_{winter}	0.7663	0.0283
a_{spring}	0.7064	0.0550
a_{summer}	1.1657	0.2235
a_{autumn}	0.7733	0.0618
b_{winter}	0.1016	0.0056
b_{spring}	0.0752	0.0075
b_{summer}	0.0672	0.0175
b_{autumn}	0.0794	0.0080
c	-0.0167	0.0020
d	8.2E-4	1.3E-4

A consequence of a seasonally dependent b is that the factor F in equation (8) also varies over the year. In Table 4 this factor is considered when there is a constant increase in mean temperature of 3°C. F is only shown for the range of observed temperatures in each season.

The values of F from the seasonally varying model are rather different from those from the constant model in Table 2. Especially for the summer season the temperature interval where F is less than 1 is wider with lower values of F (down to 0.88). For $T<7$°C in winter F is higher than for the constant precipitation-temperature relation (1.36 compared to 1.28). Spring and autumn have quite similar values of F which are in between those for winter and summer.

Table 4 Multiplication factor *F* (with its standard error se) for various values of *T* when there is a constant increase in temperature of 3°C using Model 1 with seasonally varying parameters.

		Winter		Spring		Summer		Autumn	
T	*T**	*F*	*se*	*F*	*se*	*F*	*se*	*F*	*se*
-12	-9	1.36	.02						
-10	-7	1.36	.02						
-8	-5	1.36	.02						
-6	-3	1.36	.02						
-4	-1	1.36	.02	1.25	.03				
-2	1	1.36	.02	1.25	.03			1.27	.03
0	3	1.36	.02	1.25	.03			1.27	.03
2	5	1.36	.02	1.25	.03			1.27	.03
4	7	1.36	.02	1.25	.03			1.27	.03
6	9	1.28	.02	1.18	.02	1.15	.06	1.20	.02
8	11	1.11	.03	1.03	.02	1.00	.05	1.04	.02
10	13	1.01	.04	.93	.02	.91	.04	.94	.02
12	15	.97	.04	.90	.03	.88	.03	.91	.02
14	17			.92	.03	.89	.03	.93	.03
16	19			.99	.05	.97	.03	1.00	.05
18	21			1.14	.09	1.11	.06	1.15	.09
20	23			1.38	.16	1.35	.13	1.40	.16
22	25			1.79	.30	1.74	.25	1.81	.30
24	27					2.39	.49		

A consequence of the seasonally varying factors in Table 4 is that the summer months in the climate scenario become dryer and winter months wetter than in the scenario of Figure 4. Precipitation in winter months increases 27% (was 20%) whereas the period of decreased precipitation (up to 5%) lasts from May to October (was September only). For the summer season it is not possible to obtain very accurate estimates of *a* and *b* (Table 3) because of the limited number of rain days at low temperatures. The standard error of *F* is therefore relatively large at the lower end of the temperature range in summer.

The abrupt changes in the factor *F* at the transition of seasons would be avoided by fitting continuous periodic functions to the parameters *a* and *b*. Fourier series are popular in the literature (Buishand, 1977; Stern and Coe, 1984). An alternative is the sharply peaked oscillation in Batschelet (1981) which may describe the peaks in *a* and *b*. The approach leads, however, to a more complicated model and was therefore not considered in this preliminary study.

5 Sensitivity to the definition of a rain day

The results in the previous sections refer to days with a precipitation amount of at least 0.1 mm over the 0-0 UT interval. In this section two other definitions of a rain day are considered. The first refers to the height of the threshold, while the second deals with the choice of the starting point of the 24-hour interval.

A problem with the 0.1 mm threshold is that fog and dew are sometimes measured as small precipitation amounts (usually 0.1 or 0.2 mm). Therefore a rain day is often defined as a day with a precipitation amount of at least 0.3 mm. Table 5 gives the factor *F* based on Model 1 with this criterion for a rain day. Again the two lowest temperature classes were discarded for parameter estimation.

Table 5 Multiplication factor *F* (with its standard error se) for various values of *T* when there is a constant increase in temperature of 3°C, but the rain day threshold is chosen to be 0.3 mm instead of 0.1 mm.

		Model 1	
T	*T**	*F*	*se*
-12	-9	1.20	.01
-10	-7	1.20	.01
-8	-5	1.20	.01
-6	-3	1.20	.01
-4	-1	1.20	.01
-2	1	1.20	.01
0	3	1.20	.01
2	5	1.20	.01
4	7	1.20	.01
6	9	1.16	.01
8	11	1.07	.01
10	13	1.01	.01
12	15	.99	.01
14	17	1.00	.02
16	19	1.05	.03
18	21	1.14	.06
20	23	1.28	.11
22	25	1.50	.18
24	27	1.80	.30

The height of the threshold defining a rain day has a substantial effect on the parameter estimates (Appendix C) which explains the differences between the values of *F* in Table 2 and Table 5. There is now only one temperature class left where *F* drops below 1 (at *T*=12°C), because the relative extremes are less clear than in Figure 1. For the 0.3 mm threshold at low temperatures *F* is lower than for the 0.1 mm threshold because of a lower value of the coefficient *b* in equation (3). The decrease of the fraction of rain days with precipitation amounts of 0.1 and 0.2 mm with increasing temperature as shown in Table 6 is responsible for the higher value of *b* for the 0.1 mm threshold. From weather classification data it follows that only a small part (less than 1%) of this decrease can be attributed to days with fog that are misclassified as rain days.

Table 6 Fraction of rain days (threshold 0.1 mm) at various temperatures with precipitation amounts < 0.3 mm.

T	Fraction
-8	0.41
-6	0.42
-4	0.39
-2	0.39
0	0.30
2	0.21
4	0.19
6	0.16
8	0.14

A problem with the 0.3 mm threshold is that the coefficient of variation *v(R)* increases slowly with temperature (Appendix C and D). Strictly speaking, the factor *F* in Table 5 then only represents the relative change in the mean. The relative change in the standard deviation $\sigma(R)$ is, however, rather close to that in the mean. The factor can therefore still be used to generate a precipitation scenario. Figure 6 shows the monthly mean precipitation amounts for the perturbed climate obtained with the factor *F* in Table 5. *F* was also applied to the observed precipitation amounts of 0.1 and 0.2 mm which were excluded in the calibration of the regression model. The monthly means in the resulting precipitation scenario are almost identical to those obtained in Section 3 for the 0.1 mm threshold.

When the precipitation threshold of 0.3 mm is selected the annual mean increases from 780 mm to 850 mm (about 9%). This value is identical to that obtained for the 0.1 mm threshold.

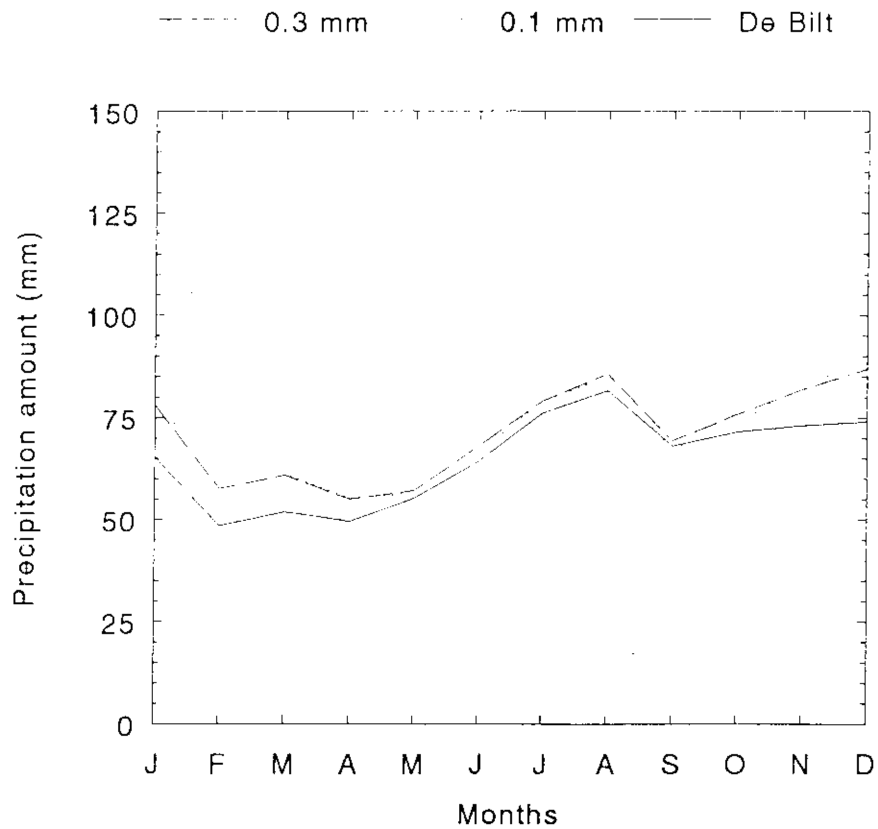

Figure 6 Monthly mean precipitation in a perturbed climate ($T^{*}=T+3°C$) from time series based on Model 1 compared to the actual climate of De Bilt (1906-1981). The precipitation thresholds for a rain day used in model calibration are 0.3 mm and 0.1 mm, respectively.

Long records of daily precipitation amounts over the 0-0 UT interval are generally derived from the registrations of self-recording raingauges (pluviographs). Such records are only available at a limited number of principal climatological stations. In the Netherlands most rainfall stations are only equipped with a standard raingauge and the precipitation amounts then refer to the totals over the 8-8 UT interval. The relation between precipitation sampled at 8 UT and mean temperature at the previous day (in the 0-0 UT interval) can also be described by the regression models in Section 2. From the results in Appendix C it is seen that the choice of a different time interval for the precipitation amounts does not affect the quality of the fit and the coefficient of variation $v(R)$. The estimates of the regression coefficients b, c and d are, however, closer to zero for the 8-8 UT amounts, reflecting the fact that the mean increases less rapidly with temperature. Table 7 gives the factor F derived from Model 1 for rain days with a precipitation threshold of 0.1 mm. Again the two lowest temperature classes were discarded in the regression analysis.

Table 7 Multiplication factor *F* (with its standard error se) for various values of *T* when there is a constant increase in temperature of 3°C, but precipitation is sampled at 8 UT over the past 24 hours instead of the 0-0 UT interval.

		Model 1	
T	*T**	*F*	*se*
-12	-9	1.23	.02
-10	-7	1.23	.02
-8	-5	1.23	.02
-6	-3	1.23	.02
-4	-1	1.23	.02
-2	1	1.23	.02
0	3	1.23	.02
2	5	1.23	.02
4	7	1.23	.02
6	9	1.18	.01
8	11	1.09	.01
10	13	1.02	.01
12	15	.99	.02
14	17	.99	.02
16	19	1.03	.03
18	21	1.10	.05
20	23	1.22	.09
22	25	1.39	.15
24	27	1.65	.25

The values of *F* in Table 7 for precipitation between 8-8 UT lead to somewhat higher monthly mean precipitation amounts in summer and somewhat lower values in winter compared to Figures 4 and 6. This is because the values of *F* at low and very high temperatures are smaller than those in Table 2, while *F* does not drop below 0.99 in the interval from 10°C to 14°C. Again the annual mean changes by 9% from 780 mm to 850 mm.

It can be concluded that the multiplication factor *F* is to some extent sensitive to the interval over which precipitation is sampled and to the threshold above which a day is classified as a rain day. This means that somewhat different precipitation scenarios are obtained when only the quantities over the interval 8-8 UT are available or the threshold of 0.3 mm is chosen for a rain day.

6 The use of maximum temperature as predictor

It was noted in Section 2 that originally Können (1983) considered the daily maximum temperature T_{max} instead of the daily mean temperature T. T_{max} is in fact a better indicator of convective precipitation than T. Because of this, and because of the larger diurnal variation in temperature on days with convective precipitation, the relative minimum in the relation between mean precipitation and T_{max} is more pronounced than in Figure 1. The results in Appendix C show that this is not accompanied by a smaller value of the constant coefficient of variation v(R). From the variation of the daily precipitation amounts within temperature classes it is not clear whether one should use T or T_{max}.

Model 1 of Section 2 was applied to the precipitation record of De Bilt, but now with T_{max} as predictor instead of T. Table 8 shows the resulting multiplication factor F for a constant increase in T_{max} of 3°C. Again the lowest temperature classes were discarded for parameter estimation. Further details about the regression analysis can be found in Appendix C.

Table 8 Multiplication factor F (with its standard error se), but now for various values of T_{max} instead of T. A constant increase in maximum temperature of 3°C is assumed (precipitation threshold 0.1 mm).

		Model 1	
T_{max}	T^*_{max}	F	se
-10	-7	1.25	.01
-8	-5	1.25	.01
-6	-3	1.25	.01
-4	-1	1.25	.01
-2	1	1.25	.01
0	3	1.25	.01
2	5	1.25	.01
4	7	1.25	.01
6	9	1.25	.01
8	11	1.24	.01
10	13	1.12	.01
12	15	1.00	.01
14	17	.93	.01
16	19	.90	.01
18	21	.90	.01
20	23	.94	.01
22	25	1.02	.02
24	27	1.15	.04
26	29	1.36	.07
28	31	1.66	.12
30	33	2.12	.21
32	35	2.81	.36

The interval with F-values < 1 is more pronounced than in Table 2. Transformation of the observed precipitation amounts with the factor in Table 8 results in a 7% increase in the annual mean precipitation for the perturbed climate. There is now not only a decrease in mean rainfall during September, but also for August and October.

The difference between the precipitation scenarios based on T_{max} and T will generally be larger when the predicted change in T_{max} is not the same as that in T. Karl et al. (1991) show that the increase in the observed mean temperature over the contiguous United States, the former Soviet Union and the Peoples Republic of China during the last century is mainly due to a trend in the minimum temperatures whereas there is no significant change in the maximum temperature. But for the 2×CO_2 experiment of a version of the UK Meteorological Office GCM (Cao et al., 1992) the increase in global annual mean surface temperature is accompanied by a comparable change in maximum temperature. Something similar holds for T and T_{max} in the CCC second-generation GCM over Northwestern Europe. Table 9 presents the changes in seasonal and annual averages of T and T_{max} for the two land grid points near De Bilt and for a larger region around the Netherlands (16 land grid points), covering Great Britain, Denmark, Germany, France, Switzerland and parts of Austria, Spain and Italy. The 1×CO_2 and 2×CO_2 cases in this table refer to a 10-year simulation. The difference in annual means between the 2×CO_2 and 1×CO_2 cases is about 3°C for mean temperature and 3.5°C for maximum temperature. The increase is relatively constant from June to February whereas the increase in spring is somewhat lower.

Table 9 Average daily mean temperature T and maximum temperature T_{max} observed at De Bilt (1906-1981) and 10 year simulations of the CCC-GCM (two nearest and 16 land grid points averaged).

		T (°C)				
		Winter	Spring	Summer	Autumn	Year
De Bilt		2.5	8.5	16.2	9.9	9.2
GCM 1×CO_2	2 grid points	3.1	8.4	16.2	9.7	9.3
	16 grid points	4.2	9.1	17.1	10.5	10.3
GCM 2×CO_2	2 grid points	+3.2	+2.3	+3.5	+3.2	+3.1
-1×CO_2	16 grid points	+3.0	+2.3	+3.7	+3.4	+3.0
		T_{max} (°C)				
		Winter	Spring	Summer	Autumn	Year
De Bilt		5.2	13.2	21.2	13.9	13.4
GCM 1×CO_2	2 grid points	4.6	11.1	19.2	12.2	11.6
	16 grid points	6.2	11.9	20.4	13.9	13.1
GCM 2×CO_2	2 grid points	+3.7	+2.4	+3.3	+3.8	+3.5
-1×CO_2	16 grid points	+3.2	+2.4	+4.1	+3.8	+3.4

The averages for the 76 year record of De Bilt are also shown in Table 9. The differences between the average temperatures at De Bilt and those for the GCM control run (1×CO_2-case) are small, in particular for the daily mean *T*.

Transforming time series of daily precipitation with a multiplication factor derived from T_{max} is a useful alternative to a factor based on *T*. The resulting precipitation scenario will not be very different when the changes in *T* and T_{max} are comparable as is the case in present day GCM simulations.

7 Conclusions and discussion

A time series of daily precipitation amounts in a future climate at a certain location can be generated using the precipitation-temperature relation for days with rainfall. A rather advanced regression model is needed to describe this relation. From the fitted model a temperature dependent factor F can be derived for the relative change in daily precipitation due to a given temperature perturbation. In the most extended version of the model this factor also depends on the time of the year. A series of daily precipitation amounts for the perturbed climate is subsequently obtained by multiplying the observed daily amounts by the factor F. In this method the number of rain days remains unchanged.

The factor F is to some extent sensitive to the height of the threshold defining a rain day and to the starting point of the 24-hour interval over which precipitation is sampled. A somewhat different factor is also found when the maximum temperature is used as predictor instead of the daily mean temperature.

The main advantage of the method is that on a daily basis the consistency between temperature and precipitation is not violated. For a 3 °C temperature increase at De Bilt the resulting precipitation scenario shows a marked increase in the monthly means during winter (20 to 27%, depending on the model used). The changes in monthly means are much smaller during late spring, summer and early fall. There is even a decrease in mean precipitation during part of that period, in particular when a seasonally dependent factor is used. For all models the increase in annual mean rainfall varies between 7 and 9%. This is consistent with the predicted change in global mean precipitation due to doubling CO_2 concentrations in present-day GCM simulations.

The use of the factor F in this study assumes that the precipitation-temperature relation is preserved in a changing climate. The verification of this assumption with GCM simulations needs further attention. The comparison with the increase in global mean precipitation may be criticized because the global mean is largely determined by the tropics and the oceans. The use of regional GCM data for this purpose may meet difficulties because daily precipitation is highly variable in time and poorly represented at the various grid points. There can also be quite substantial changes in the number of rain days in regional GCM data. Because precipitation occurrence and the amount of precipitation are partly influenced by the same factors, these changes might be accompanied by changes in the mean amount which are not accounted for by temperature. Because of these difficulties it is also questionable to study the precipitation-temperature relation directly in relatively short GCM runs.

The use of a factor implies that the constant coefficient of variation does not change. The verification of this assumption with GCM data may meet similar problems as with the relation between mean precipitation and temperature It is also possible to compare the values of $v(R)$

for different climatic regions in Europe. Little variation in these values supports the use of a factor. For the De Bilt record it is remarkable that $v(R)$ is not influenced by the increased contribution of convective precipitation at high temperatures.

The observed precipitation - temperature relation at De Bilt is strongly determined by the decreasing importance of frontal rain at daily mean temperatures above 10 °C and the occurrence of heavy showers at high temperatures. A separate study of the precipitation-temperature relation for frontal rain and convective precipitation may deepen the insight into the change of precipitation with temperature and may lead to a better choice of the systematic part of the regression model. Whenever hourly amounts or weather classification data are available there are many possibilities to make a subdivision of these two types of precipitation.

Although the precipitation-temperature relation was only examined for De Bilt, the resulting values of the multiplication factor are also applicable to other inland stations in the Netherlands. The use of the method for coastal regions and other countries requires further study of the precipitation-temperature relation.

Besides temperature and precipitation, other variables like solar radiation, humidity, wind speed and evapotranspiration are often required in climate change impact studies. To be consistent with a constant precipitation - temperature relation and no change in rainfall occurrence, it is advisable to maintain the relative humidity or the dew point depression in the observed record. The fact that the number of rain days remains the same in the perturbed climate also rules out large changes in the incoming solar radiation.

Other climate elements than temperature determine the precipitation amount on rain days as well. One of them is surface air pressure. The assumption that the sequence of rain days in the perturbed climate remains the same as in the present climate will generally no longer be valid when there is a systematic change in surface air pressure. The incorporation of surface air pressure in the regression model is the subject of further research.

Acknowledgements

The idea of using the empirical relation between daily precipitation amount and temperature arose from discussions with G.P. Können. It is due to him that we developed our temperature dependent factor. We thank J.J. Beersma, M.A.M. Mann and B.W.A.H. Parmet for their support during this study and A.J. Coops for his detailed comments on an earlier version. This work is part of the Dutch National Research Program on Global Air Pollution and Climate Change (NOP). The GCM data were kindly provided by the Canadian Climate Centre, Downsview, Ontario.

References

Batschelet, E., 1981. Circular Statistics in Biology. Academic Press, London.

Boer, G.J., N.A. McFarlane and M. Lazare, 1992. Greenhouse gas-induced climate change simulated with the CCC second-generation general circulation model. Journal of Climate, **5**, 1045-1077.

Buishand, T.A., 1977. Stochastic modelling of daily rainfall sequences. Communications Agricultural University Wageningen, Wageningen, 77-3.

Buishand, T.A. and J.J. Beersma, 1993. Jackknife tests for differences in autocorrelation between climate time series. Journal of Climate (in press).

Bultot, F., A. Coppens, G.L. Dupriez, D. Gellens and F. Meulenberghs, 1988. Repercussions of a CO_2 doubling on the water cycle and on the water balance - A case study for Belgium. Journal of Hydrology, **99**, 319-347.

Cao, H.X., J.F.B. Mitchell and J.R. Lavery, 1992. Simulated diurnal range and variability of surface temperature in a global climate model for present and doubled CO_2 climates. Journal of Climate, **5**, 920-943.

Cohn, T.A., L.L DeLong, E.J. Gilroy, R.M. Hirsch and D.K. Wells, 1989. Estimating constituent loads. Water Resources Research, **25**, 937-942.

Cole, J.A., S. Slade, P.D. Jones and J.M. Gregory, 1991. Reliable yield of reservoirs and possible effects of climatic change. Hydrological Sciences Journal, **36**, 579-598.

Davidian, M. and R.J. Carroll, 1987. Variance function estimation. Journal of the American Statistical Association, **82**, 1079-1091.

Duan, N., 1983. Smearing estimate: A nonparametric retransformation method. Journal of the American Statistical Association, **78**, 605-610.

Houghton, J.T., G.J. Jenkins and J.J. Ephraums (eds.), 1990. Climate Change, The IPCC Scientific Assessment. Cambridge University Press, Cambridge.

Houghton, J.T., B.A. Callander and S.K. Varney (eds.), 1992. Climate Change 1992, The Supplementary Report to the IPCC Scientific Assessment. Cambridge University Press, Cambridge.

Karl, T.R., G. Kukla, V.N. Razuvayev, M.J. Changery, R.G. Quayle, R.R. Heim, Jr., D.R. Easterling and C.B. Fu, 1991. Global warming: evidence for asymmetric diurnal temperature change. Geophysical Research Letters, **18**, 2253-2256.

Können, G.P., 1983. Het Weer in Nederland (The Weather in the Netherlands). Thieme, Zutphen.

McCullagh, P. and J.A. Nelder, 1989. Generalized Linear Models, Second edition. Chapman and Hall, London.

McFarlane, N.A., G.J. Boer, J.-P. Blanchet and M. Lazare, 1992. The Canadian Climate Centre second-generation general circulation model and its equilibrium climate. J. Climate, **5**, 1013-1044.

Mitchell, J.F.B., 1991. The equilibrium response to doubling atmospheric CO_2. In: Greenhouse-Gas-Induced Climatic Change: A Critical Appraisal of Simulations and Observations (M.E. Schlesinger, Ed.), Developments in Atmospheric Science, **19**, 49-61, Elsevier, Amsterdam.

Morrison, D.F., 1978. Multivariate Statistical Methods. McGraw-Hill, Auckland.

Rind, D., R. Goldberg and R. Ruedy, 1989. Change in climate variability in the 21st century. Climatic Change, **14**, 5-37.

Seber, G.A.F and C.J. Wild, 1989. Nonlinear Regression. John Wiley & Sons, New York.

Smith, P.L., 1979. Splines as a useful and convenient statistical tool. The American Statistician, **33**, 57-62.

Stern, R.D. and R. Coe, 1984. A model fitting analysis of daily rainfall data (with discussion). Journal of the Royal Statistical Society, Series A, **147**, 1-34.

Wahba, G., 1989. Spline functions. In: Encyclopedia of Statistical Sciences (S. Kotz, N.L. Johnson and C.B. Read, Eds.), Supplement Volume, 148-160, John Wiley & Sons, New York.

APPENDIX

A Factor *F* for Model 2

In case of a warming climate and choosing Model 2 (Section 2) for *g(T)* the multiplication factor *F* becomes:

$$F = \exp\left[b(T^*-T)\right] \qquad T,T^*\le m_1$$

$$F = \exp\left[b(T^*-T) + c(T^*-m_1)^2 + d(T^*-m_1)^3\right] \qquad T\le m_1<T^*<m_2$$

$$F = \exp\left[b(T^*-T) + c\{(T^*-m_1)^2 - (T-m_1)^2\} + d\{(T^*-m_1)^3 - (T-m_1)^3\}\right]$$
$$m_1<T,T^*\le m_2 \qquad \text{(A1)}$$

$$F = \exp\left[b(T^*-T) + c\{(m_2-m_1)(2T^*-m_2-m_1)-(T-m_1)^2\} + d\{(m_2-m_1)^2(3T^*-2m_2-m_1)-(T-m_1)^3\}\right] \qquad m_1<T\le m_2<T^*$$

$$F = \exp\left[b(T^*-T) + 2c(m_2-m_1)(T^*-T) + 3d(m_2-m_1)^2(T^*-T)\right] \qquad T,T^*>m_2$$

B Variance of the estimated factor

The factor *F* in equation (7) can be expressed as:

$$F = \exp\left[u^t\theta\right] \qquad \text{(B1)}$$

where *u* is a p′x1 vector depending only on *T* and *T**, θ is a p′x1 vector of unknown regression coefficients and *t* denotes the transpose of a vector (p′≤p). For instance for $T\le m_1<T^*$ in equation (8) the vectors *u* and θ are given by:

$$u^t = \left[T^*-T,(T^*-m_1)^2,(T^*-m_1)^3\right] \qquad \text{and} \qquad \theta^t = (b,c,d) \qquad \text{(B2)}$$

An estimate $\hat{F}$ of *F* is obtained by replacing the unknown elements of θ by the estimated regression coefficients:

$$\hat{F} = \exp\left[u^t\,\hat{\theta}\right] \qquad \text{(B3)}$$

Now $L=u^t\hat{\theta}$ is a linear combination of estimated regression coefficients. Its variance is given by (Morrison, 1978):

$$\mathrm{var}(L) = u^t\Sigma u \qquad \text{(B4)}$$

where $\sum$ is the covariance matrix of $\hat{\theta}$. An estimate of this matrix is easily obtained from the fitting procedure in Appendix C.

To obtain an approximation to the variance of $\hat{F}$ use can be made of a Taylor expansion of F about E(L)= $u^t\,\theta$ (delta method):

$$\hat{F} \approx \exp\left[u^t\,\theta\right] + \exp\left[u^t\,\theta\right] \cdot (L-u^t\,\theta) = F + F(L-u^t\,\theta) \qquad \text{(B5)}$$

giving:

$$\text{var}\,\hat{F} \approx F^2\,\text{var}(L) = F^2\,(u^t\,\Sigma u) \qquad \text{(B6)}$$

The square root of this approximation is given as the standard error se in Tables 2,4,5,7 and 8.

C Parameter estimation, testing for lack-of-fit and seasonal variation

<u>Non-seasonal models</u>

The unknown regression coefficients in the model were estimated by an iteratively reweighted least squares procedure. The procedure was applied to the mean precipitation amounts in the various temperature classes. This yields almost the same estimates as fitting the model to the individual daily precipitation amounts. The final weighted sum of squares in the procedure can be used to test for lack-of-fit.

Let n_k denote the number of rain days, $\overline{T}_k$ the mean (or maximum) temperature and $\overline{R}_k$ the mean precipitation amount for the k-th temperature class (k=1, ... ,K). Neglecting the slight variation of the temperature within the classes, the following regression equation is obtained for $\overline{R}_k$:

$$\overline{R}_k = \exp\left[g(\overline{T}_k)\right] + \varepsilon_k \qquad \text{(C1)}$$

The error term ε_k has zero mean and variance:

$$\text{var}(\varepsilon_k) = \sigma_k^{\,2}/n_k \qquad \text{(C2)}$$

where σ_k is the standard deviation of R for the k-th temperature class. It is the same standard deviation as that shown in Figure 2.

Equation (C2) assumes independence between the precipitation amounts within each temperature class. Buishand (1977, pp. 83,84) showed that in the Netherlands there is a significant correlation between the precipitation amounts on successive wet days during the winter period (correlation coefficient≈0.15) and that there is no correlation during the rest of the year. The small positive correlation for the winter period has little effect on the variance of ε_k and will therefore not impair the fitting procedure given below.

Fitting a model by least squares implies minimizing a sum of (weighted) squared residuals with respect to the unknown model parameters. In our application the sum

$$S(a,b,c,d) = \sum_{k=1}^{K} w_k \left[\overline{R}_k - \exp\left\{ g(\overline{T}_k) \right\} \right]^2 \tag{C3}$$

has to be minimized with respect to a,b,c and d. The weights w_k must be inversely proportional to var(ε_k). It is thus necessary to have an estimate of σ_k in equation (C2). A disadvantage of the sample standard deviations in the temperature classes is that they have a rather large standard error (in particular when n_k is small) which may invalidate the use of the chi-square distribution for testing lack-of-fit. It is therefore better to substitute a modelled value for σ_k, e.g. by making use of the fact that the coefficient of variation v_k hardly varies over the temperature classes. The latter can be approximated by a simple expression like:

$$v_k = h_1 + h_2 \overline{T}_k \tag{C4}$$

Table C1 shows estimates of the parameters h_1 and h_2. These estimates were obtained by a weighted least squares fit (with weights n_k) to the sample coefficients of variation in the various temperature classes. In most cases the coefficient h_2 is not significant (constant coefficient of variation). From equation (2) it follows that the standard deviation σ_k can be approximated by:

$$\sigma_k = v_k \exp\left[g(\overline{T}_k) \right] \tag{C5}$$

and thus the weights w_k become:

$$w_k = n_k \,/\, \left[v_k \exp\left\{ g(\overline{T}_k) \right\} \right]^2 \tag{C6}$$

Through its relation with $g(\overline{T}_k)$ the weights depend on the unknown regression coefficients. Iteratively reweighted least squares implies that these weights are adjusted to the current values of a,b,c and d in each iteration step. Initial estimates of these coefficients can be obtained by linear regression of ln(R_k) on $g(T_k)$ with weights n_k/v_k^2. The iteration procedure reduces to a sequence of weighted linear regressions of an adjusted dependent variable on the explanatory variables (McCullagh and Nelder, 1989). The final parameter estimates are denoted as $\hat{a},\hat{b},\hat{c}$ and $\hat{d}$. The asymptotic properties of these estimates are the same as in the case of known σ_k (Davidian and Carroll, 1987).

The fitting procedure also provides an approximate covariance matrix of the parameter estimates, cf. Seber and Wild (1989, p.89). The standard errors in Table 1 are based on this covariance matrix. The error bars in Figure 1 were obtained by substituting the final estimate of σ_k in equation (C2).

The generalized Pearson X^2-statistic $X^2=S(\hat{a},\hat{b},\hat{c},\hat{d})$ can be used to assess the goodness-of-fit. Under the null hypothesis X^2 has approximately a chi-square distribution with K-4 degrees of freedom. The quality of the approximation can be improved by grouping classes with few rain days. In this study there were at least n_k=8 rain days in each class. Table C1 presents the values of X^2 for the models considered in Sections 2, 5 and 6. From the table it is clear that an acceptable fit can only be achieved when the two lowest temperature classes are discarded.

Table C1 Estimated coefficients in the equations for $g(T)$ and v_k, number K of classes, and X^2-statistic for the non-seasonal models discussed in Sections 2, 5 and 6 (n.s. indicates that the parameter is not significant at the 5% level). The values of the test-statistic in brackets refer to the case where the two lowest temperature classes were included in the fitting procedure.

	Coefficients $g(T)$				Coefficients v_k		Goodness-of-fit	
	a	b	c	d	h_1	h_2	K	X^2
Model 1								
T;0-0 UT;>0.1 mm	0.7649	0.0829	-0.0144	6.7E-4	1.47	n.s.	17	14.8 (33.9)
T;0-0 UT;≥0.3 mm	1.0758	0.0619	-0.0098	4.9E-4	1.03	0.0097	17	6.9 (60.0)
T;8-8 UT;>0.1 mm	0.8473	0.0682	-0.0101	4.6E-4	1.42	n.s.	17	13.8 (37.9)
$Tmax$;0-0 UT;>0.1 mm	0.6406	0.0749	-0.0137	5.5E-4	1.42	n.s.	19	8.6 (30.5)
Model 2								
T;0-0 UT;>0.1 mm	0.7639	0.0832	-0.0147	7.0E-4	1.47	n.s.	17	14.4 (33.9)

Quasi-likelihood functions

The distribution of the parameter estimates in a non-linear regression model can strongly differ from the normal distribution. It is therefore not always safe to base statistical inference on the standard errors or the covariance matrix of the parameter estimates. A better alternative is to make use of likelihood ratio tests and likelihood confidence regions (Seber and Wild, 1989). The quasi-likelihood function is related to the log-likelihood. In contrast to the latter, it requires that only the first two moments are specified. The method is applicable to situations where the variance is the product of two terms, a constant dispersion parameter and a function of the mean. This is the case for the model with constant coefficient of variation.

Let $\mu_k=\exp[g(\bar{T}_k)]$ be the mean for the k-th temperature class (k=1, ... ,K) and v the constant coefficient of variation, then the variance of R_k is given by:

$$\mathrm{var}(\bar{R}_k) = v^2\mu_k^2/n_k \qquad (C7)$$

and the quasi-likelihood $Q\,(a,b,c,d)$ reads (McCullagh and Nelder, 1989):

$$Q\,(a,b,c,d) \;=\; -\sum_{k=1}^{K} n_k\left[\bar{R}_k/\mu_k + \ln\,\mu_k\right] \qquad (C8)$$

The quasi-likelihood estimates are obtained by maximizing (C8) with respect to a,b,c and d. These estimates are, however, equivalent to the iteratively reweighted least squares estimates. They also correspond to the maximum likelihood estimates for the gamma distribution. From the literature on stochastic rainfall modelling (Buishand, 1977; Stern and Coe, 1984) it is known that the distribution of R is close to the gamma distribution. A slight difference between these two distributions is, however, that the lower bound of R is not exactly zero.

The confidence regions in Figure 3 contain combinations of c and d for which the quasi-likelihood is close to its maximum $Q\,(\hat{a},\hat{b},\hat{c},\hat{d})$. To be more specific, let $\hat{a}(c,d)$ and $\hat{b}(c,d)$ be the quasi-likelihood estimates of a and b for fixed c and d. These are obtained by a similar iteratively reweighted least squares procedure as the full quasi-likelihood estimates.
The 100γ % confidence region consists of the values c and d for which

$$2Q\,(\hat{a},\hat{b},\hat{c},\hat{d}) \;-\; 2Q\left[\hat{a}(c,d),\hat{b}(c,d),c,d\right] \;\leq\; v^2 c_\gamma \qquad (C9)$$

with c_γ=-2ln(1-γ) the corresponding percentage point of the χ^2_2-variable. This region is easily obtained by calculating the differences between the quasi-likelihoods in (C9) for a grid of c,d values about $\hat{c},\hat{d}$.
A lack-of-fit test can also be derived from the quasi-likelihood function (McCullagh and Nelder, 1989). This test was not considered in the present study.

Seasonal models

The seasonal dependence of the precipitation temperature relation was investigated by classifying the daily precipitation amounts according to temperature and season. Let n_{ks} be the number of rain days, T_{ks} the mean temperature and $\bar{R}_{ks}$ the mean precipitation amount in the k-th temperature interval (k=1, ... ,K) for season s (s=1, ... ,S). Empty classes (e.g. at high temperatures in winter) were discarded.

The parameter estimates in Table 3 were obtained by applying the iteratively reweighted least squares procedure to the mean amounts $\bar{R}_{ks}$ in the various classes. To improve the chi-square approximation to the null distribution of the lack-of-fit statistic X^2, classes with few rain days were grouped within the season concerned until n_{ks} was at least 8. This resulted in X^2=42.4 with 35 degrees of freedom. The value of X^2 is not significant at the 5% level.

Besides that the coefficient of variation $v(R)$ hardly varies with temperature for the 0.1 mm threshold, it is also nearly constant over the various seasons. For the weighted mean $\bar{v}$ of $v(R)$ a value of 1.37 was found. The fact that this value is lower than that given in Table C1 for the non-seasonal classification is an indication of seasonal dependence of the precipiation temperature relation.

Seasonal dependence can be tested formally with a quasi-likelihood ratio statistic. Analoguous to (C8) the quasi likelihood function is defined as:

$$Q\ (\theta)\ =\ -\sum_{classes} n_{ks}\left[\overline{R_{ks}}/\mu_{ks}+\ln\ \mu_{ks}\right] \tag{C10}$$

where θ is the parameter vector and $\mu_{ks}=E(\overline{R_{ks}})$ is the theoretical mean amount in the k-th temperature interval for season s. The summation in equation (C10) is over all non-empty classes. Under the null hypothesis (H_0) θ equals θ_0, whereas under the alternative hypothesis (H_A) there are no restrictions on θ. The quasi-likelihood-ratio statistic is then given by:

$$D\ =\ 2\left[Q(\hat{\theta})\ -\ Q(\hat{\theta}_0)\right] \tag{C11}$$

where $\hat{\theta}_0$ and $\hat{\theta}$ are the iteratively reweighted least squares estimates of θ under H_0 and H_A respectively. Under H_0 the quantity $D/\bar{v}^2$ has an asymptotic chi-square distribution with r degrees of freedom, where r is the number of restrictions on θ to define θ_0.

The expected monthly means in Figure 5 were obtained without making use of a mathematical function for the precipitation-temperature relation. The theoretical means μ_{ks} in the non-empty classes form then the components of θ and $\overline{R_{ks}}$ is the estimate of μ_{ks} under H_A. Under H_0: $\mu_{k1}=\cdots=\mu_{kS}=\mu_k$ (k=1, ... ,K). The common mean μ_k is estimated as the average precipitation amount $\overline{R_{k.}}$ in the k-th temperature interval. The test-statistic in equation (C11) then reads:

$$D_1\ =\ 2\sum_{classes} n_{ks}\left[(\overline{R_{ks}}/\overline{R_{k.}}-1)\ +\ \ln(\overline{R_{k.}}/\overline{R_{ks}})\right] \tag{C12}$$

and $r=M-K$, with M the number of non-empty classes. The value of $D_1/\bar{v}^2$ in Table C2 is significant at the 5% level. Omitting classes with less than 8 rain days gives an identical result.

For the parametric model in Section 4 the means $\overline{R_{ks}}$ follow from $\overline{R_{ks}}=\exp[g(\overline{T_{ks}})]$ and the parameter vector is given by $\theta=(a_{winter}, \ldots ,a_{autumn}, b_{winter}, \ldots ,b_{autumn}, c, d)'$. The quasi-likelihood-ratio statistic for testing H_0: $a_{winter}=\cdots=a_{autumn}$ and $b_{winter}=\cdots=b_{autumn}$ is denoted as D_2. There are r=2(S-1)=6 degrees of freedom in this test. It is also possible to test for seasonal variation in b given that there is seasonal variation in a. The resulting test-statistic D_3 has r=S-1=3 degrees of freedom. The values of $D_2/\bar{v}^2$ and $D_3/\bar{v}^2$ in Table C2 are significant at the 5% level.

Table C2 Values of $D_1/\overline{v^2}$, $D_2/\overline{v^2}$ and $D_3/\overline{v^2}$ for testing seasonal dependence of the relation between precipitation and temperature (r=degrees of freedom).

	Model 1					
	$D_1/\overline{v^2}$	r	$D_2/\overline{v^2}$	r	$D_3/\overline{v^2}$	r
T;0-0 UT;>0.1 mm	137.3	34	102.5	6	8.8	3

D Fitting regression splines to the logarithms of daily precipitation amounts

The regression model given by equation (1) is almost equivalent to the following linear model for the logarithms of the precipitation amounts:

$$\ln R = g(T) + \varepsilon \qquad \text{(D1)}$$

The error term ε in this model has zero mean and its standard deviation $\sigma(\varepsilon)$ is constant or varies only slowly with temperature.

If the distribution of ε does not depend on temperature, then the mean of R is given by:

$$\mathrm{E}(R) = \exp\left[g(T)\right]\mathrm{E}\left[\exp(\varepsilon)\right] = C\exp\left[g(T)\right] \qquad \text{(D2)}$$

and R has constant coefficient of variation. In contrast with the non-linear model for R there appears now a constant factor C in the expression for the mean. This constant is determined by the distribution of ε. For the normal distribution $C=\exp[\sigma^2(\varepsilon)/2]$, cf. Cohn et al. (1989). Equation (7) for the factor F remains valid, since C appears both in $\mathrm{E}(R)$ and $\mathrm{E}(R^*)$. This is, however, no longer true when $\sigma(\varepsilon)$ changes with temperature as is the case for rain days with a precipitation amount $\geq$ 0.3 mm.

When $\sigma(\varepsilon)$ varies with temperature, the constant C in equation (D2) has to be replaced by a temperature dependent factor $C(T)=\mathrm{E}[\exp(\varepsilon)]$. Let η be the standardised error term, i.e. $\eta=\varepsilon/\sigma(\varepsilon)$, then $\exp(\varepsilon)$ can be approximated by:

$$\exp(\varepsilon) \approx \exp(\bar{\sigma}\eta) + \left[\sigma(\varepsilon) - \bar{\sigma}\right]\eta\exp(\bar{\sigma}\eta) \qquad \text{(D3)}$$

where $\bar{\sigma}$ is the average value of $\sigma(\varepsilon)$ for all rain days. Consequently:

$$C(T) \approx \bar{C} + \left[\sigma(\varepsilon) - \bar{\sigma}\right]\bar{D} \qquad \text{(D4)}$$

where

$$\bar{C} = \mathrm{E}\left[\exp(\bar{\sigma}\eta)\right] \quad and \quad \bar{D} = \mathrm{E}\left[\eta\exp(\bar{\sigma}\eta)\right] \qquad \text{(D5)}$$

The bootstrap procedure in Duan (1983) can be used to obtain distribution-free estimates of $\bar{C}$

and $\overline{D}$. To get an idea of the effect of a non-constant $\sigma(\varepsilon)$ on the factor F, it is, however, admissable to take the theoretical values $\overline{C}$=exp($\overline{\sigma}^2$/2) and $\overline{D}$=$\overline{\sigma}$exp($\overline{\sigma}^2$/2) for the normal distribution.

Assume now that $\sigma(\varepsilon)$ varies linearly with temperature:

$$\sigma(\varepsilon) = \overline{\sigma} + h(T-\overline{T}) \tag{D6}$$

with $\overline{T}$ the average temperature for all rain days. Substitution of (D6) in (D4) gives:

$$C(T) \approx \overline{C} + h(T-\overline{T})\overline{D} = \overline{C}\left[1+h(T-\overline{T})\overline{D}/\overline{C}\right]$$

$$\approx \overline{C}\exp\left[h(T-\overline{T})\overline{D}/\overline{C}\right] \tag{D7}$$

Consequently, the factor F in equation (7) should be multiplied by

$$f = C(T^*)/C(T) \approx \exp\left[h(T^*-T)\overline{D}/\overline{C}\right] \tag{D8}$$

For the rain days with a precipitation amount ≥ 0.3 mm $h\approx0.0087$ and $\overline{D}/\overline{C}=\overline{\sigma}\approx1.14$ giving $f\approx1.03$ for $T^*=T+3°C$. The correction is of the order of magnitude of the standard error of the estimated factor at low and moderately high temperatures (see Table 2).