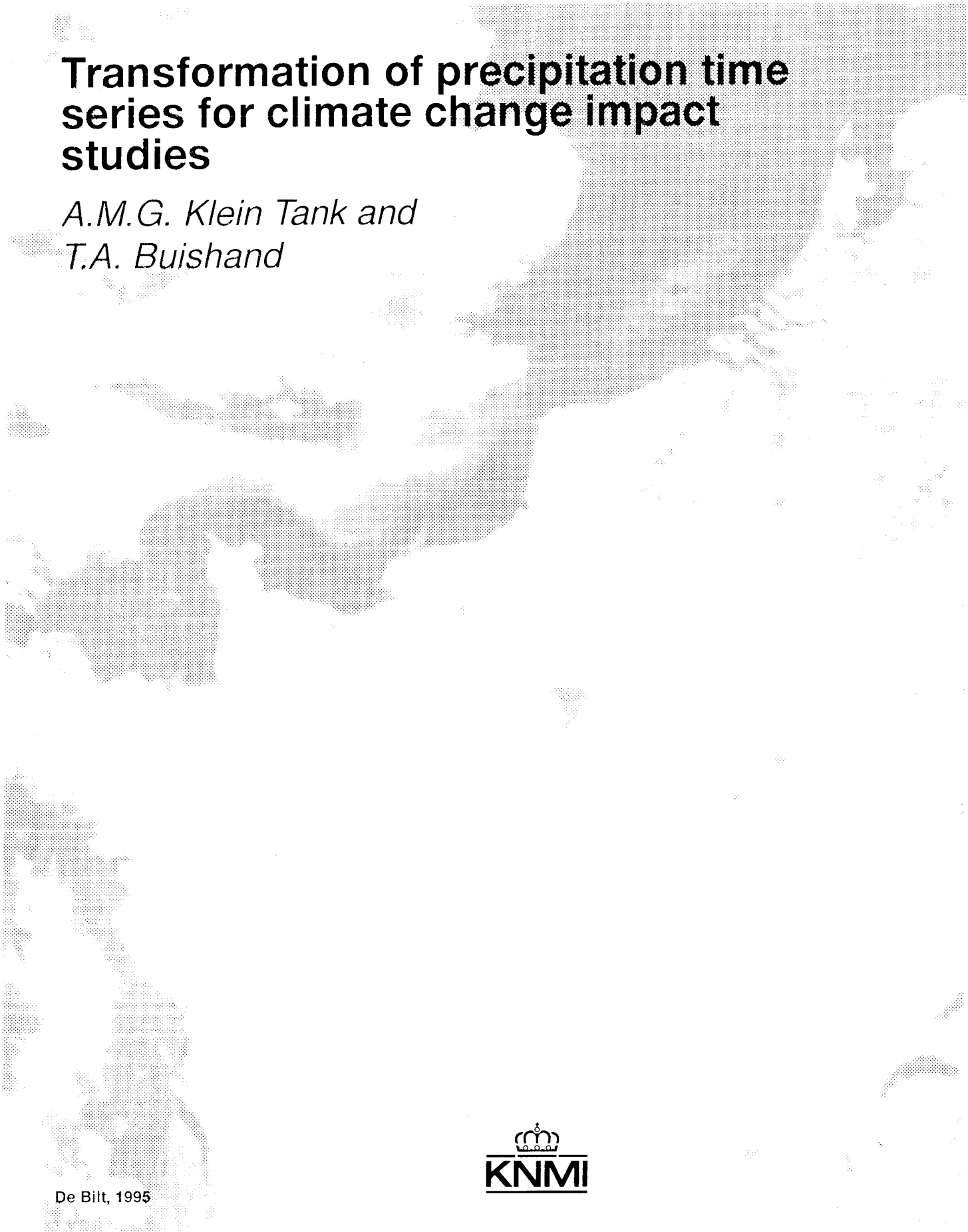
Scientific report
WR 95-01

*Ministerie van Verkeer en Waterstaat*
Koninklijk Nederlands
Meteorologisch Instituut

# Transformation of precipitation time series for climate change impact studies

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



De Bilt, 1995

Scientific report; WR 95-01

De Bilt 1995

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

UDC: 551.501.45
551.577.3
551.583.1

ISSN: 0169-1651

ISBN: 90-369-2070-1

# Transformation of precipitation time series 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*

# Preface

This report describes results of the Climate Scenarios Project at KNMI. The project is part of the Dutch National Research Programme on Global Air Pollution and Climate Change (NRP). The present report is a sequel to KNMI Scientific Report WR 93-02 presenting earlier work on the project. We thank the Canadian Climate Centre (CCC), Victoria, B.C., Canada for providing the output of their General Circulation Model. We also thank J.J. Beersma (KNMI), G.P. Können (KNMI), B.W.A.H. Parmet (RIZA, Arnhem) and F.W. Zwiers (CCC) for their comments on earlier versions.

February 1995

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

# Abstract

A new technique is developed to obtain suitable time series of daily precipitation for climate change impact studies by means of a transformation of an observed record. The transformation makes use of regression relations between precipitation and other meteorological elements. The method allows for a modification of the sequence of wet and dry days by assigning a probability of rain to each day.

Separate regression models for rainfall occurrence and the amount of precipitation on wet days are fitted to the historical data at De Bilt (the Netherlands). A logistic regression model is introduced to determine the probability of rain from the values of surface air pressure $P$, relative humidity $U$ and relative incoming solar radiation $Q$. Dry days with a high probability of rain are altered in wet days when the rainfall frequency increases in the future climate. A decrease in the rainfall frequency is achieved by changing wet days with a low probability of rain into dry days. The precipitation amounts on wet days are regressed on temperature $T$ and surface air pressure $P$. From this regression model a temperature and pressure dependent multiplying factor is derived. This factor gives the relative change in the precipitation amount resulting from a temperature and/or pressure perturbation and is used to transform the observed amounts. The regression model also provides the expected amounts for new wet days.

As an example, the 1961-1990 precipitation record at De Bilt is transformed using the seasonal mean changes in $T$ and $P$ over Western Europe for the $2\times CO_2$ climate in the Canadian Climate Centre General Circulation Model (CCC-GCM). $U$ and $Q$ were kept the same as in the present-day climate. The alternative to change these elements according to their statistical relations with $P$ is briefly discussed. The changes in the seasonal mean amounts in the resulting daily precipitation scenario are highest in winter (+44%) and lowest in summer (+8%). The annual mean amount increases by about 20%. These values differ considerably from the GCM predicted precipitation changes (+28% in winter, -26% in summer and no change on average over the year). Possible explanations for this disagreement are given. It is concluded that the simple transformation method is a good alternative to more advanced stochastic techniques of time series simulation.

# Contents

1 Introduction . . . . . 1

2 Statistical description of rainfall occurrence . . . . . 4

3 Statistical modelling of the relation between precipitation amount on rain days, daily mean temperature and surface air pressure . . . . . 13

4 GCM predicted climate changes . . . . . 25

5 Changes in the occurrence of rain based on statistical relations . . . . . 30

6 Changes in the amount of precipitation based on statistical relations . . . . . 35

7 Conclusions and discussion . . . . . 41

References . . . . . 45

Appendix A: Model validation . . . . . 50

Appendix B: Humidity and surface air pressure in the CCC-GCM . . . . . 59

Appendix C: Simulation of precipitation amounts . . . . . 61

Appendix D: Users guide . . . . . 62

# 1 Introduction

Studies of climate change have a long history in the scientific literature. Since the 1980's the main interest is focused on the potential changes of the Earth's climate caused by the increasing emissions of $CO_2$ and other trace gases. Advanced three-dimensional General Circulation Models (GCMs) have been used to simulate possible future climate conditions. These simulations indicate that the enhanced greenhouse effect resulting from the increased $CO_2$ concentrations can lead to substantial changes in temperature, precipitation and evapotranspiration. However, the magnitudes of the changes vary widely from model to model and for regional precipitation, for instance, there is even disagreement about the sign of the changes. Despite the large uncertainty about potential changes, it is important to assess their impacts on agriculture, water management and natural ecosystems. These studies need more detailed information about the future climate conditions for the site or region of interest than can be derived from the GCM simulations. Quite often climate change scenarios consisting of time series of daily precipitation, temperature and other variables (solar radiation, air humidity, wind) are required. Such scenarios should give a meteorologically consistent picture of a future climate with realistic statistical properties of the various elements (means, variances, occurrence of extreme events). They should not be considered as predictions, however.

The simplest and most widely used method of deriving time series suitable for impact studies is to transform an observed climate time series. For instance, a constant temperature increment of 1, 2 or 4°C may reflect the expected temperature rise from the enhanced greenhouse effect. An accompanying increase or decrease of 10% or 20% in the long term mean rainfall is usually achieved by multiplying the observed amounts by a constant factor. Both the monthly (Gleick, 1987; Mimikou et al., 1991; Kwadijk, 1993) and daily values (Bultot et al., 1988, 1992; Cooley, 1990; Panagoulia, 1991, 1992) of observed temperature and precipitation have been altered in this way. Other variables like air humidity and solar radiation have sometimes been considered by those authors to derive representative values of evapotranspiration. In Němec and Schaake (1982) and Arnell (1992) proportional adjustments were also applied to the estimates of potential evapotranspiration for the current climate.

Klein Tank and Buishand (1993) multiplied the daily precipitation amounts by a temperature dependent factor. This factor was derived from a mathematical description of the observed relation between the mean daily amounts on wet days and temperature. The preservation of this precipitation-temperature relation through the multiplying factor ensures that internally consistent sets of daily temperature and precipitation are obtained. The proportional adjustment of daily precipitation assumes that the number of rain days remains unchanged. In order to obtain a broader range of climate change scenarios it is desirable, however, to allow for changes in rainfall

occurrence. Some authors have modified the observed sequence of wet and dry days by a Monte Carlo technique (Bultot and Gellens, 1989; Robock et al., 1993). The dependence of rainfall occurrence on factors like the pressure distribution and cloudiness is, however, ignored in that work, which makes the internal consistency of the resulting scenarios questionable. For instance, there is a risk that a relatively large number of anticyclonic cloudless days will be considered to be wet in the perturbed climate.

Stochastic simulation of time series is another possibility to obtain the required input for impact studies. Especially in the hydrological literature several techniques can be found for generating synthetic rainfall sequences. Most of this work deals, however, with precipitation only. A notable exception is a paper by Richardson (1981) about the simultaneous generation of daily values of precipitation, maximum temperature, minimum temperature and solar radiation. Wilks (1992) discussed the use of his model for generating sequences for future climate conditions. Another application is given in Valdés et al. (1994). Semenov and Porter (1994) applied a slightly different stochastic model to assess the impacts of climate change on wheat crops. Simplifying assumptions about the distribution and correlation structure of the four variables may lead to physically unrealistic combinations of these variables or even impossible values. It is further unknown how far discrepancies between simulated time series for the current climate and the historical observations resulting from model deficiencies affect the final estimates of the climate change impacts.

Recently a lot of attention has been given to the generation of daily precipitation sequences conditional on the observed large-scale atmospheric circulation patterns or the simulated patterns in a GCM (Bárdossy and Plate,1992; Wilson et al., 1992; Hughes et al.,1993). This approach is known as statistical downscaling. The use of the method for generating realistic sequences of daily precipitation for impact studies of a $CO_2$ induced warming is hampered by the fact that the changes in the frequencies of circulation types in GCM runs with enhanced $CO_2$ concentrations are relatively small compared with the differences between the frequencies in the control run and the historical observations (Hughes et al., 1993; Matyasovsky et al., 1993). It is further not sufficient to relate precipitation to circulation patterns only. Matyasovsky et al. (1993) introduced the height of the 500 hPa level as an additional explanatory variable in the Bárdossy and Plate model. This height is strongly determined by the temperature in the lower atmosphere and thus the change in the amount of precipitation is coupled to temperature as in the approach of Klein Tank and Buishand (1993).

Deterministic simulation of regional climates using a high resolution Limited Area Model (LAM) nested in a GCM is popular among climate modellers to increase the resolution of GCMs over the area of interest. It assumes that the GCM can realistically simulate the large-scale atmospheric features and that the high resolution LAM is capable to take into account the effects of orography and regional-scale processes on

the various climate variables (Giorgi and Mearns, 1991). Due to computational expenses this deterministic downscaling has only been explored for relatively short periods (3.5 complete years) or selected seasons, which restricts its applicability in climate change impact studies (Giorgi et al., 1994).

From the above it is clear that the existing techniques need further development. In this report the method of Klein Tank and Buishand (1993) for the transformation of observed rainfall amounts on wet days is extended to allow for changes in the sequence of wet days. In Section 2 rainfall occurrence is related to surface air pressure, air humidity and solar radiation. As in Katz and Parlange (1993) daily mean surface air pressure at a point reflects the effect of the atmospheric circulation on precipitation. Humidity and solar radiation accommodate the effect of cloudiness on rainfall occurrence and are often required in climate change impact studies. In Section 3 surface air pressure is included in the regression model of Klein Tank and Buishand (1993). The multiplying factor for transforming the daily amounts then no longer depends solely on temperature, but also on surface air pressure. Section 4 deals with GCM predicted changes in the long term means of the various variables in the models for rainfall occurrence and amount of precipitation. The use of statistical relations between surface air pressure, relative humidity and solar radiation for supplementing the estimated changes from GCM simulations is discussed in Section 5. Two methods are presented to modify an observed sequence of wet and dry days given the changes in these three elements. Examples are given of the effects of the changes in surface air pressure, relative humidity and solar radiation on the mean number of rain days. In Section 6 the changes in the mean precipitation amounts are presented both for the situation of a temperature perturbation and that of an additional perturbation in surface air pressure. Section 7 gives a general discussion and the main conclusions.

# 2 Statistical description of rainfall occurrence

The number of rain days in a given month and daily probabilities of rain are related to surface air pressure $P$, relative humidity $U$ and relative incoming solar radiation $Q$. The latter is the quotient of incoming solar radiation at the surface and downward solar radiation at the top of the atmosphere (a function of date and geographical location) and is known as the clearness index or atmospheric transmittance in the literature on solar energy (Graham et al., 1988). By using $Q$ instead of absolute values of incoming solar radiation most of the annual cycle is eliminated. $Q$ is an adequate indicator of both cloud cover and thickness (Slob, 1985). High values of $Q$ correspond with little cloudiness. Two regression models are presented:

Model A: A multiple linear regression model that relates the percentage $N$ of rain days in each month to the monthly values of $P$, $U$ and $Q$. The percentage of rain days is chosen instead of the absolute number because of the differences in length between the various months.

Model B: A logistic regression model that relates the status of a day (wet or dry) to the daily values of $P$, $U$ and $Q$.

The effects of changes in the variables $P$, $U$ and $Q$, which may occur in a future climate, are most easily demonstrated with Model A (see section 5). The logistic model can also be used to alter wet days in an observed climate time series in dry days or to assign additional wet days to the record. An essential assumption in these applications is that the model remains unchanged in the perturbed climate.

Models A and B are described in detail in this section. Parameter estimates for the record at De Bilt (1961-1990) and their statistical significance are given. The percentage of days correctly predicted as wet or dry by Model B is also presented. The models' ability to reproduce the observed annual cycle of the monthly mean percentage of rain days is discussed. A rain day is defined as a day with a recorded precipitation amount $\geq 0.1$ mm. The effect of alternative definitions of a rain day is considered briefly.

Model A: Monthly values

In Model A the percentage $N_i$ of rain days in a month $i$ is linearly related to the monthly values $P_i$, $U_i$ and $Q_i$ of surface air pressure, relative humidity and relative incoming solar radiation:

$$N_i = a + bP_i + cU_i + dQ_i + \epsilon_i \qquad i = 1,\ldots,I \qquad (1)$$

Table 1 Estimated regression coefficients with standard error (se) and Student's $t$-statistic in the model given by equation (1) for rain days at De Bilt (1961-1990). The observed significance level $\alpha$ refers to a two-tailed test.

| Coefficent | Estimate | se | $t$ | $\alpha$ |
|---|---|---|---|---|
| $a$ | 1630.608 | 159.822 | 10.2 | < 0.01 |
| $b$ | -1.566 | 0.160 | -9.8 | < 0.01 |
| $c$ | 0.500 | 0.200 | 2.5 | 0.01 |
| $d$ | -0.708 | 0.153 | -4.6 | < 0.01 |

where $I$ is the total number of months of observation. It is assumed that the errors $\epsilon_i$ are independent and identically distributed with mean zero. $P_i$, $U_i$ and $Q_i$ are obtained by averaging the daily values of $P$, $U$ and $Q$. $P_i$ is expressed in hPa, while $U_i$ and $Q_i$ are expressed as percentages. At De Bilt the observed long term mean values of $P_i$, $U_i$, $Q_i$ and $N_i$ are 1015.2 hPa, 82%, 38% and 55% respectively. Table 1 presents the least squares estimates of the regression coefficients $a$, $b$, $c$ and $d$. These coefficients are statistically significant at the 5% level. The Student test in the table assumes that the errors come from a normal distribution. This assumption is reasonably satisfied for monthly data.

The three variables $P_i$, $U_i$ and $Q_i$ explain 50% of the variance of $N_i$. This value is only slightly lower when either $U_i$ or $Q_i$ is left out of the model, because of the rather strong correlation between these two variables (correlation coefficient -0.83). $P_i$ alone explains 40% of the variance of $N_i$. From Table 1 it follows that a change in $P_i$ of -3 hPa, while keeping other variables ($U_i$, $Q_i$) constant, would result in 4.7% extra rain days per month ($\approx$ 1.4 days). It should be noted, however, that it is not always allowed to change only one variable in the regression model independently of the others, since a decrease in $P_i$ would generally lead to systematic changes in $U_i$ and $Q_i$ as well. In Section 5 it is demonstrated from the relations between these variables in the observed climate of De Bilt that this effect is small, however.

Model B: Daily values

The occurrence of rain can be described by a binary variable which takes the value 1 when a day is wet and the value 0 when it is dry. A more advanced regression model than the standard linear model is needed to link this variable to other meteorological variables. The logistic model is popular for analysing binary data (Cox and Snell, 1989; McCullagh and Nelder, 1989). The model assigns to potential wet days a high probability of rain, whereas for potential dry days this probability is low.

The probability $p_j$ that day $j$ is wet is represented as:

$$p_j = \frac{1}{1+\exp\left[-h_j\right]} \qquad j=1,\ldots,J \tag{2}$$

$J$ is the total number of observation days and $h_j$ has the same form as the deterministic part in the standard linear regression model. In our application:

$$h_j = a + bP_j + cU_j + dQ_j \qquad j=1,\ldots,J \tag{3}$$

where $a$, $b$, $c$ and $d$ are the regression coefficients. Now $P_j$, $U_j$ and $Q_j$ are the daily values for surface air pressure (hPa), relative humidity (%) and relative incoming solar radiation (%), respectively. By definition the probability that day $j$ is dry equals $1-p_j$. In the logistic model $p_j$ is always between 0 and 1 regardless of the value of $h_j$. Figure 1 shows the relation between $p_j$ and $h_j$. This relation is non-linear and $p_j$ is therefore also non-linear in $P_j$, $U_j$ and $Q_j$.

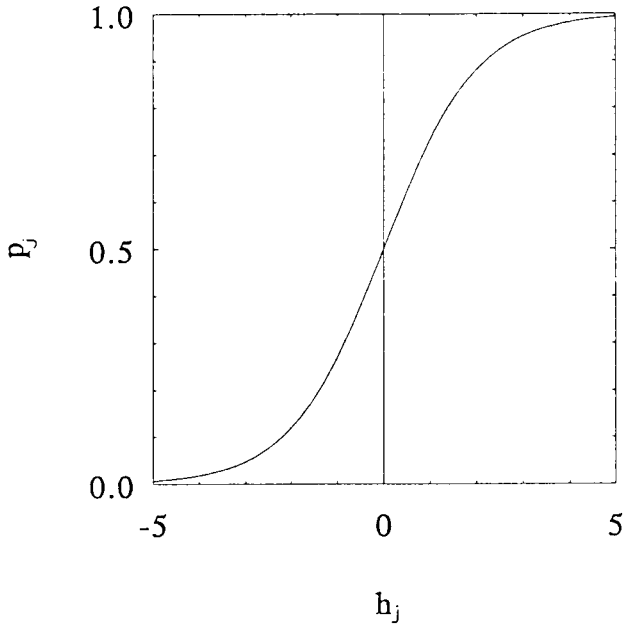



Figure 1 Relation between $p_j$ and $h_j$ in the logistic regression model.

Table 2 Maximum likelihood estimates of regression coefficients with standard error (se) and Wald-statistic in the logistic model for probability of rain using a subset that consists of every 10th day of the 1961-1990 record at De Bilt. The observed significance level $\alpha$ refers to a two-tailed test.

| Coefficient | Estimate | se | Wald | $\alpha$ |
|---|---|---|---|---|
| $a$ | 107.492 | 9.609 | 125.1 | < 0.01 |
| $b$ | -0.108 | 0.010 | 128.9 | < 0.01 |
| $c$ | 0.049 | 0.011 | 19.9 | < 0.01 |
| $d$ | -0.033 | 0.006 | 34.2 | < 0.01 |

The regression coefficients $a$, $b$, $c$ and $d$ were estimated by the maximum likelihood method. This is a standard method of parameter estimation in the logistic model. It is incorporated in a number of statistical packages (SPSS, SAS, BMDP). These packages assume, however, independence between successive cases. To exclude the effect of autocorrelation the model was fitted to subsets containing only every 10th day of the 1961-1990 record at De Bilt. Table 2 presents an example of estimated regression coefficients with their standard errors and the Wald statistic. This statistic can be used to test the statistical significance of a regression coefficient (Buse, 1982). It is comparable to the square of the $t$-value in Table 1 for the standard linear regression model. Table 2 shows that all regression coefficients in equation (3) significantly deviate from zero. Moreover, the values of the Wald statistic are all higher than the square of the $t$-values in Table 1 even though in this case 90% of the data is left out of the analysis. This clearly demonstrates that more accurate estimates of the regression coefficients are obtained with the use of daily values instead of monthly values.

Table 3 Means of estimated regression coefficients in the logistic model for 10 different subsets of the daily record at De Bilt (1961-1990).

| Coefficient | Estimate |
|---|---|
| $a$ | 101.412 |
| $b$ | -0.103 |
| $c$ | 0.060 |
| $d$ | -0.032 |

Because each subset contains 1 out of 10 days a total of 10 different logistic regression models was obtained from the complete historical time series. Table 3 shows the mean values of the parameters in these 10 models. Details of the fitted regression models are given in Appendix A.

The fit of the logistic model was tested with the Hosmer-Lemeshow statistic. The results are presented in Appendix A. From these results it can be concluded that systematic errors are small. The logistic model also rather successfully predicts the status of a day from the observed values $P_j$, $U_j$ and $Q_j$. Table 4 classifies the days in the record of De Bilt according to the observed rainfall occurrence and their predicted status from the fitted logistic model. The prediction for day $j$ is wet if the estimated probability of rain $\hat{p}_j \geq 0.5$ and dry otherwise, where $\hat{p}_j$ is given by:

$$\hat{p}_j = \frac{1}{1+\exp\left[-(\hat{a}+\hat{b}P_j+\hat{c}U_j+\hat{d}Q_j)\right]} \qquad j=1,\ldots,J \qquad (4)$$

with $\hat{a}$, $\hat{b}$, $\hat{c}$ and $\hat{d}$ the estimated regression coefficients in Table 3. From Table 4 it follows that 75% of the days (31.5 + 43.7) is correctly predicted, which is nearly as good as operational forecasts using advanced numerical weather prediction methods (Appendix A).

As a consequence of the non-linearity of the logistic model, the effect of a constant change in $P_j$, $U_j$ or $Q_j$ on $p_j$ depends on the values of $P_j$, $U_j$ and $Q_j$ and thus on the value of $p_j$ itself. This is illustrated in Figure 2, where the increase in $\hat{p}_j$ is presented when for each day $\Delta P_j$ = -3 hPa. The figure shows that a decrease in $P_j$ of 3 hPa has the highest impact on days with a probability of rain of about 0.5. The effect is only half as high on days with $\hat{p}_j = 0.1$ or $\hat{p}_j = 0.9$. When $\Delta P_j$ = -3 hPa, the observed long term mean probability of rain in De Bilt changes from 0.55 to 0.60, which gives 5% extra rain days per month (≈ 1.5 days). This is in good agreement with the 4.7% extra rain days per month found with Model A for the monthly values.

Table 4 Classification of days according to observed and predicted rainfall occurrence at De Bilt.

| | | Predicted | |
|---|---|---|---|
| | | Dry | Wet |
| Observed | Dry | 31.5% | 13.8% |
| | Wet | 11.1% | 43.7% |

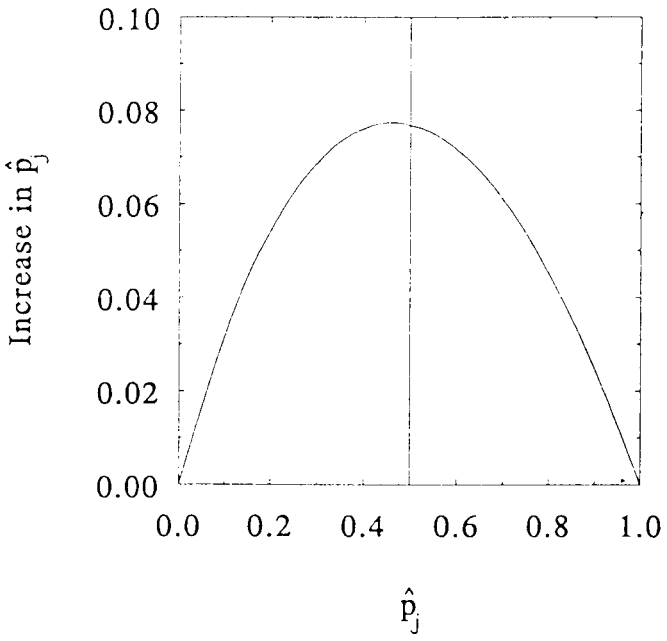



Figure 2 Increase in probability $\hat{p}_j$ of rain when $\Delta P_j$ = -3 hPa and the model parameters in Table 3 apply.

A slight improvement of model performance is possible by adding the status of the previous day (1 for a rain day and 0 for a dry day) to the right-hand side of equation (3). The values of the Hosmer-Lemeshow test statistic for this version of the model are presented in Appendix A. About 77% of the days is correctly predicted using again $\hat{p}_j = 0.5$ as a threshold for classifying a day as wet or dry.

Seasonal variation

The mean percentage of rain days at De Bilt is relatively large during winter and small during summer. Figure 3 shows that this annual cycle is reasonably reproduced when Model A is used to predict the percentage of rain days in each month from the observed values $P_i$, $U_i$ and $Q_i$. The largest discrepancy is found for February where the model overestimates the percentage of rain days on average by 6%. For the other winter months December and January the model also has a positive bias, which is, however, much smaller than that for February. Because the least squares method preserves the overall monthly mean percentage of rain days, a small negative bias is found for most of the rest of the year. A better descripition of the annual cycle requires a regression model with seasonally varying regression coefficients.

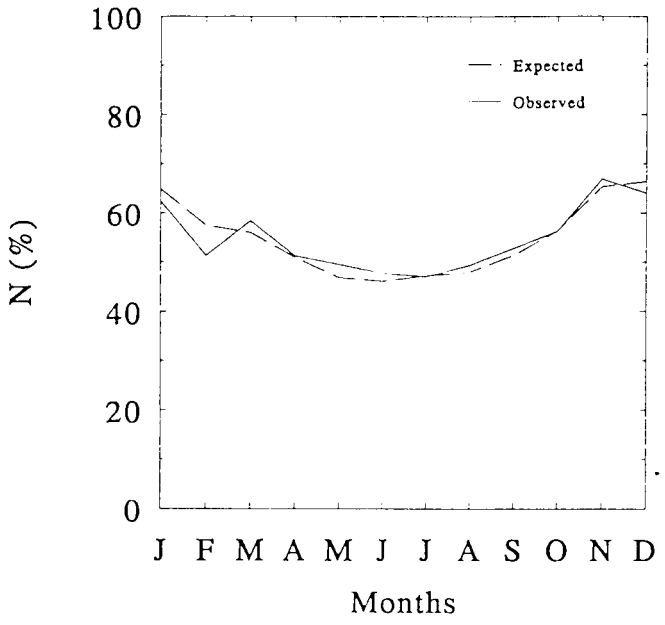



Figure 3 Annual cycle of expected and observed percentage of rain days at De Bilt (1961-1990). Expected percentages are based on the parameter values in Table 1.

For climate change scenario objectives the need for such a model is questionable, since there are large uncertainties in the perturbations $\Delta P_i$, $\Delta Q_i$ and $\Delta U_i$ (see Sections 4 and 5).

For Model B the total number $m$ of rain days for each calendar month can be estimated from the observed daily values $P_j$, $U_j$ and $Q_j$ as:

$$m = \sum_{j=1}^{K} \hat{p}_j \tag{5}$$

with $K$ the number of days in the month. The annual cycle in the percentage of rain days shows then similar discrepancies as in the case of Model A with a maximum difference of 8% between the predicted and observed percentage of rain days in February. Also for the daily values a seasonally varying model has to be introduced to obtain a better description of the annual cycle.

Sensitivity to the definition of a rain day

A rainfall threshold of 0.3 mm is sometimes preferred to the value of 0.1 mm. The reason for this is that small amounts (0.1 mm and 0.2 mm) in the observed record may well be caused by dew or fog. For the 0.3 mm threshold the estimates of $b$, $c$ and $d$ in Model A for the monthly values are almost the same as those in Table 1 for the 0.1 mm threshold, whereas the value of $a$ is smaller as a result of the lower number of rain days at the 0.3 mm threshold. For the daily model (Model B) the new parameter estimates are presented in Table 5. In this case the model fits the data worse than with the 0.1 mm threshold. For three out of the 10 subsets of the precipitation data at De Bilt the Hosmer-Lemeshow statistic is significant at the 5% level, indicating lack-of-fit (Appendix A). Nevertheless the status of 75% of the days is correctly predicted.

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 use of these precipitation amounts instead of the 0-0 UT amounts hardly effects the parameters in Model A for the monthly values. Table 6 presents the estimated coefficients for the logistic regression (Model B) of rainfall occurrence over the 24h interval ending at 8 UT on the values of $P_j$, $U_j$ and $Q_j$ for the 0-0 UT interval of the previous day. With the exception of $\hat{c}$ these estimates are only slightly different from those given in Table 3 for the 0-0 UT amounts. Again the status of about 75% of the days is correctly predicted, while the fit is not worse than for precipitation over the 0-0 UT interval (Appendix A).

Table 5 Means of estimated regression coefficients in the logistic model for 10 different subsets of daily rainfall data with a threshold of 0.3 mm defining a rain day instead of 0.1 mm as in Table 3.

| Coefficient | Estimate |
|---|---|
| $a$ | 114.480 |
| $b$ | -0.115 |
| $c$ | 0.046 |
| $d$ | -0.040 |

Table 6 Means of estimated regression coefficients in the logistic model for 10 different subsets of precipitation amounts over the interval from 8 to 8 UT instead of 0 to 0 UT as in Table 3.

| Coefficient | Estimate |
|---|---|
| $a$ | 100.867 |
| $b$ | -0.101 |
| $c$ | 0.045 |
| $d$ | -0.032 |

# 3 Statistical modelling of the relation between precipitation amount on rain days, daily mean temperature and surface air pressure

Precipitation amounts on rain days exhibit large variation and skewness. Days with small amounts predominate. A regression model to link such data to the daily temperatures is presented in Klein Tank and Buishand (1993). This section deals with the inclusion of surface air pressure in that model. As in the regression model for rainfall occurrence this variable partly accounts for the rather complicated influence of the atmospheric circulation.

First a short description of the regression model for the precipitation-temperature relation is given. A similar type of model is then used to relate the precipitation amounts on wet days to daily mean surface air pressure. From the two models for the daily precipitation amounts more comprehensive models are derived which contain both temperature and surface air pressure as explanatory variables. Parameter estimates and their statistical significance are presented. The value of a generalized chi-squared statistic is given to judge the adequacy of the fit. A non-parametric check is made on seasonal variation in the relation between precipitation, temperature and surface air pressure. The use of daily rainfall amounts over the 8-8 UT interval rather than the 0-0 UT interval is briefly considered at the end of this section.

Because of the large variation of the daily precipitation amounts long data records are needed to obtain accurate estimates of the mean amounts for given values of temperature and surface air pressure. As in Klein Tank and Buishand (1993) the 1906-1981 record at De Bilt is considered in this section.

Precipitation amount versus daily mean temperature

The model in Klein Tank and Buishand (1993) for the relation between precipitation amount $R$ on rain days and daily mean temperature $T$ is of the form:

$$R = \exp\,[g(T)] + \epsilon \qquad (6)$$

where $\epsilon$ is a random error term with mean zero. The function $g(T)$ consists of piecewise polynomials:

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

Table 7 Estimates of the coefficients in equation (8) with their standard error (se) for rain days at De Bilt (1906-1981). After Klein Tank and Buishand (1993).

| Coefficient | Estimate | se |
|---|---|---|
| $a$ | 0.7649 | 0.0233 |
| $b$ | 0.0829 | 0.0039 |
| $c$ | -0.0144 | 0.0015 |
| $d$ | 6.7E-4 | 9.9E-5 |

$$g(T) = a + bT + c(T-m_1)^2 + d(T-m_1)^3 \qquad T>m_1 \qquad (7)$$

or in the Heaviside function notation:

$$g(T) = a + bT + c(T-m_1)_+^2 + d(T-m_1)_+^3 \qquad (8)$$

where $(T-m_1)_+ = \max(0,T-m_1)$. The knot $m_1$ has been fixed a priori at $T = 7$°C. The model can be extended with a second knot at $T = 21$°C to preclude the possibility of a rapid change in the mean amounts at the end of the temperature range. Estimates of the coefficients $a$, $b$, $c$ and $d$ in equation (8) are presented in Table 7. Apart from the interval from 14°C to 18°C, the mean precipitation amount increases with increasing daily mean temperature. The coefficient of variation $v_R$ of the daily precipitation amounts (the standard deviation divided by the mean) is almost constant over the whole temperature range, $\hat{v}_R = 1.31$.

Precipitation amount versus daily mean surface air pressure

The mean precipitation amount generally decreases with increasing surface air pressure. Figure 4 presents the mean precipitation amount at De Bilt for daily mean surface air pressure class intervals of 6 hPa. Again only rain days are considered. A similar model as in the previous subsection can be used to describe the relation between $R$ and surface air pressure $P$:

$$R = \exp[h(P)] + \epsilon \qquad (9)$$

where $h(P)$ is a non-linear function of $P$ and $\epsilon$ has the same meaning as in equation (6). In this study the function $h(P)$ is taken as:

$$h(P) = a_0 + a_1(q_1-P) + a_2(q_1-P)_+^2 + a_3(q_1-P)_+^3 \qquad (10)$$

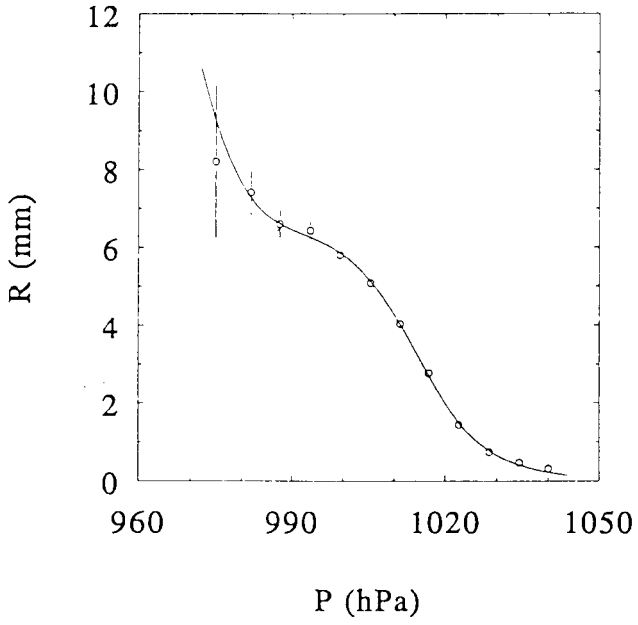



Figure 4 Mean precipitation amounts at surface air pressure class intervals of 6 hPa for rain days at De Bilt (1906-1981). The smooth curve is based on a fitted regression model. The error bars indicate the standard deviations of the means.

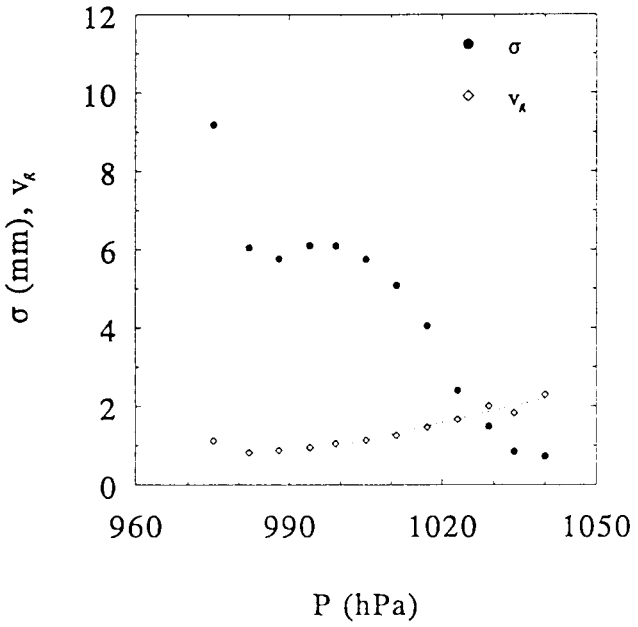



Figure 5 Standard deviation σ and coefficient of variation $v_R$ of daily precipitation $R$ at surface air pressure class intervals of 6 hPa at De Bilt (1906-1981). The dotted line is based on equation (11).

Table 8 Estimates of the coefficients in equation (10) with their standard error (se) for rain days at De Bilt (1906-1981). The statistic $Z$ is the ratio of these two quantities.

| Coefficient | Estimate | se | $Z$ |
|---|---|---|---|
| $a_0$ | 0.6862 | 0.0163 | 42.0 |
| $a_1$ | 0.1080 | 0.0032 | 33.3 |
| $a_2$ | -0.0036 | 2.7E-4 | -13.5 |
| $a_3$ | 4.3E-5 | 5.6E-6 | 7.7 |

where $(q_1\text{-}P)_+ = \max(0, q_1\text{-}P)$. The knot $q_1$ has been fixed a priori at $P$ = 1020 hPa. At this knot $h(P)$ changes from a cubic polynomial ($P \leq 1020$ hPa) to a linear function ($P > 1020$ hPa). The iteratively reweighted least squares technique described in Klein Tank and Buishand (1993) was used to estimate the parameters $a_0$, $a_1$, $a_2$ and $a_3$. The method was applied here to the mean precipitation amounts in the 12 air pressure class intervals in Figure 4. There were at least 8 rain days in each class interval. The final parameter estimates are presented in Table 8. These estimates have an asymptotic normal distribution. The statistic $Z$ is comparable to Student's $t$-statistic in Table 1 for the linear regression model. Since the number of rain days is large, the significance of the regression coefficients can be tested by comparing the values of $Z$ with the percentage points of the standard normal distribution. All parameters are significant at the 5% level.

The smooth curve in Figure 4 is based on the fitted model. The curve shows a monotone decrease in the mean precipitation amount with increasing surface air pressure. There is a point of inflexion near $P$ = 990 hPa. The occurrence of such a point is characteristic for a cubic polynomial. The coefficient $a_3$ is responsible for the rapid increase in the mean precipitation amount at very low values of $P$. This seems not very realistic since in general precipitation amounts are lower towards the centre of depressions. This feature of the model does, however, not lead to a significant value of the generalized Pearson $X^2$-statistic for lack-of-fit. The rapid rise of the mean precipitation amount at low values of $P$ can be avoided by supplying a second knot near the point of inflexion and imposing linearity of $h(P)$ beyond this knot. The coefficient of variation $v_R$ of the daily precipitation amounts is almost constant for $P \leq$ 1000 hPa, but increases with increasing $P$ for $P > 1000$ hPa (Figure 5). In the iteratively reweighted least squares procedure $v_R$ was described by:

$$\hat{v}_R = 0.959 + 0.031\,(P-q_2)_+ \qquad (11)$$

where $q_2$ = 1000 hPa. The value of $\hat{v}_R$ for $P \leq 1000$ hPa is considerably lower than the value 1.31 for the regression of precipitation on temperature in the previous

subsection. However, above 1000 hPa, $\hat{v}_R$ increases rapidly to values higher than 2 near the upper end of the pressure range.

Combined effects of temperature and surface air pressure on precipitation

Temperature and surface air pressure on rain days are only weakly correlated. The strongest correlation occurs in the winter season (correlation coefficient = -0.16). Equations (6)-(10) form therefore a useful starting point to formulate models for predicting the precipitation amount on rain days from daily mean temperature and daily mean surface air pressure. The models considered here are of the form:

$$R = \exp[k(T,P)] + \epsilon \qquad (12)$$

First the case is considered where $k(T,P)$ consists of the separate terms in equations (8) and (10):

$$k(T,P) = a_0 + a_1(q_1-P) + a_2(q_1-P)_+^2 + a_3(q_1-P)_+^3 + bT + c(T-m_1)_+^2 + d(T-m_1)_+^3 \qquad (13)$$

Thus in this model the parameter $a$ in equation (8) changes with surface air pressure according to equation (10). The parameters $b$, $c$ and $d$, that describe the change of the mean precipitation amount with temperature, do not depend on $P$. From another point of view, the parameter $a_0$ in equation (10) changes with temperature according to equation (8) and the parameters $a_1$, $a_2$, $a_3$, that describe the change of the mean amount with surface air pressure, do not depend on $T$. The effects of changes in $T$ and $P$ on the mean precipitation amount are multiplicative in this model (see Section 6). Table 9 gives the estimates of the coefficients $a_0$, $a_1$, $a_2$, $a_3$, $b$, $c$ and $d$. These were obtained by applying the iteratively reweighted least squares method to the mean amounts in a total of 119 temperature/pressure classes of at least 8 rain days each. Rain days with a temperature < -7°C were discarded for parameter estimation. The coefficient of variation $v_R$ was described by:

$$\hat{v}_R = 0.808 + 0.015\,T + 0.025\,(P-q_2)_+ \qquad (14)$$

where $q_2$ = 1000 hPa. As in the model for precipitation and pressure, $\hat{v}_R$ does not change with $P$ if $P \leq 1000$ hPa.

Figure 6 shows that this 7 parameter model reasonably reproduces the variation of the mean precipiation amount with temperature and surface air pressure. Nevertheless the value 184.9 of the $X^2$-statistic is significant at the 1% level. Figure 7 presents the standardized differences SDIF between the observed and predicted mean precipitation amounts in the various temperature/pressure classes.

Table 9 Estimates of the coefficients in equation (13) with their standard error (se) for rain days at De Bilt (1906-1981). The statistic $Z$ is the ratio of these two quantities.

| Coefficient | Estimate | se | $Z$ |
|---|---|---|---|
| $a_0$ | 0.2063 | 0.0230 | 9.0 |
| $a_1$ | 0.0882 | 0.0030 | 29.7 |
| $a_2$ | -0.0021 | 2.5E-4 | -8.4 |
| $a_3$ | 2.0E-5 | 0.5E-5 | 3.8 |
| $b$ | 0.0679 | 0.0034 | 20.2 |
| $c$ | -0.0088 | 0.0014 | -6.2 |
| $d$ | 4.2E-4 | 1.0E-4 | 4.1 |

Large symbols are used to mark classes with large residuals. Fourteen out of 119 values of SDIF in Figure 7 are either ≤ -2 or > 2. This is significantly larger than the expected value 6 assuming normally distributed residuals. There is no clear grouping of residuals with a particular symbol, indicating a rather complicated interaction between the effects of temperature and surface air pressure. For a better description of the mean precipitation, this interaction can be taken into account by adding several extra terms to the model. One possible form of $k(T,P)$ is:

$$\begin{aligned} k(T,P) = {} & a_0 + a_1(q_1-P) + a_2(q_1-P)_+^2 + a_3(q_1-P)_+^3 \\ & + \left[b_0 + b_1(q_1-P) + b_2(q_1-P)_+^2 + b_3(q_1-P)_+^3\right] T \\ & + \left[c_0 + c_1(q_1-P)\right](T-m_1)_+^2 + \left[d_0 + d_1(q_1-P)\right](T-m_1)_+^3 \end{aligned} \qquad (15)$$

In contrast to the previous 7 parameter model all parameters $a$, $b$, $c$ and $d$ in equation (8) vary now with surface air pressure.The isolines in Figure 8 represent the theoretical mean amounts according to the extended 12 parameter model. The 6 mm and 8 mm isolines are quite different from those for the 7 parameter model in Figure 6.

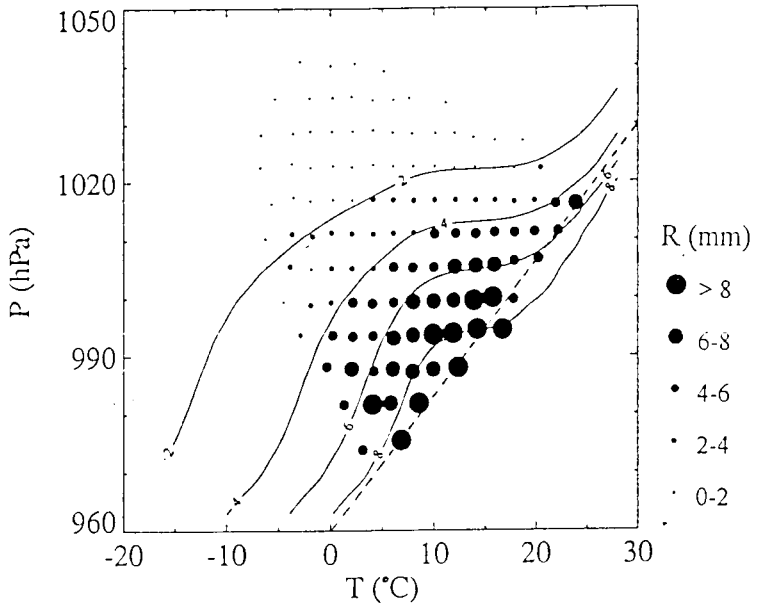



Figure 6 Mean precipitation amounts at temperature class intervals of 2 °C and air pressure class intervals of 6 hPa for rain days at De Bilt (1906-1981). Model values of the 7 parameter model (isolines) are compared to observed values (black circles). The dashed line $P = 960 + 2.3T$ is considered in Section 6.

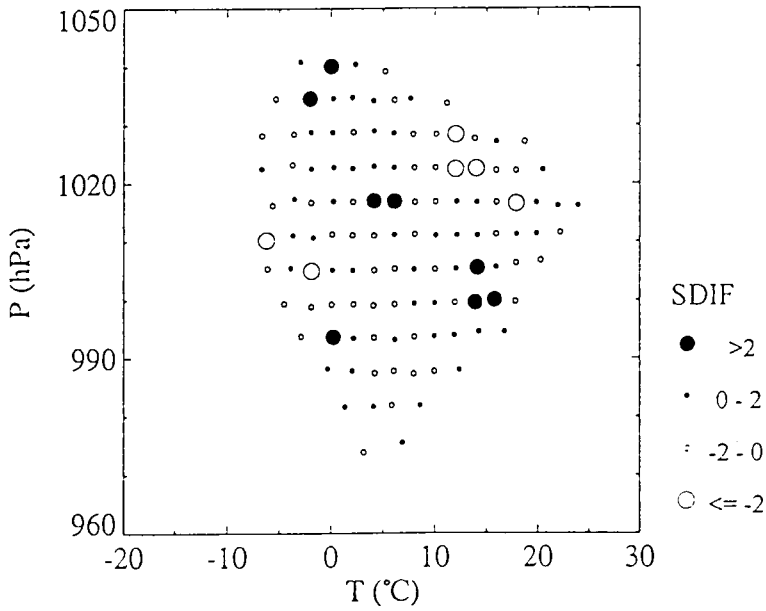



Figure 7 Standardized differences SDIF between the observed and predicted (7 parameter model) mean precipitation amounts in the various temperature/pressure class intervals in Figure 6. The difference in each class is standardized through division by the estimated standard deviation of the observed mean amount.

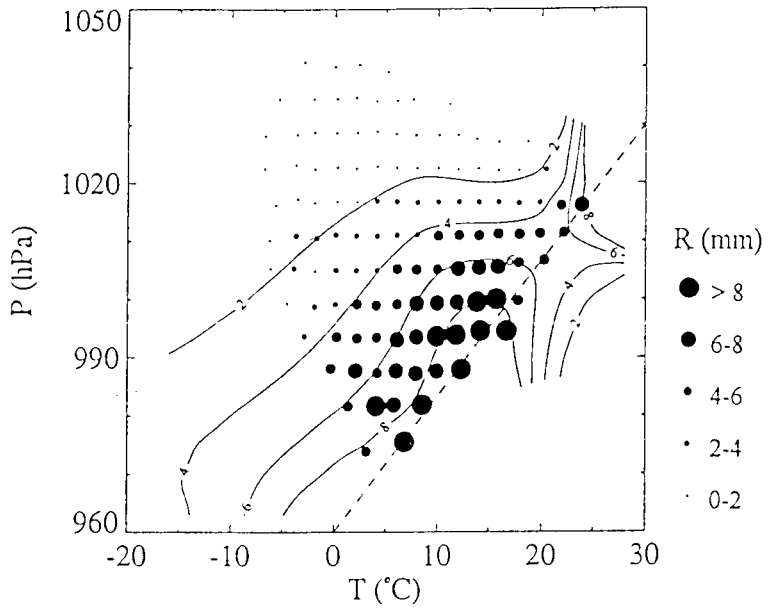



Figure 8 Mean precipitation amounts at temperature class intervals of 2 °C and air pressure class intervals of 6 hPa for rain days at De Bilt (1906-1981). Model values of the 12 parameter model (isolines) are compared to observed values (black circles). The dashed line $P = 960 + 2.3T$ is considered in Section 6.

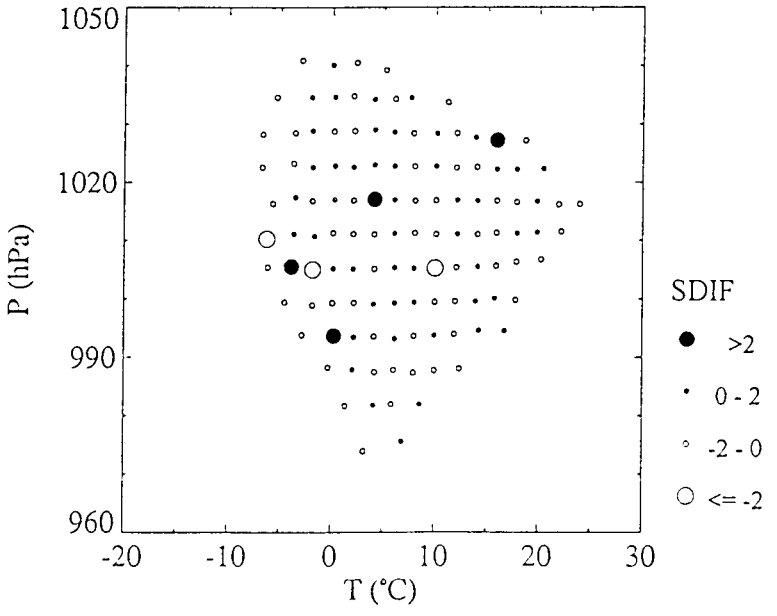



Figure 9 Standardized differences SDIF between the observed and predicted (12 parameter model) mean precipiation amounts in the various temperature/pressure class intervals in Figure 8.

Table 10 Estimates of the coefficients in equation (15) with their standard error (se) for rain days at De Bilt (1906-1981). The statistic $Z$ is the ratio of these two quantities.

| Coefficient | Estimate | se | $Z$ |
|---|---|---|---|
| $a_0$ | 0.2476 | 0.0302 | 8.2 |
| $a_1$ | 0.0739 | 0.0044 | 16.7 |
| $a_2$ | -0.0015 | 4.3E-4 | -3.4 |
| $a_3$ | 1.5E-5 | 1.0E-5 | 1.5 |
| $b_0$ | 0.0658 | 0.0053 | 12.3 |
| $b_1$ | 0.0025 | 6.4E-4 | 3.9 |
| $b_2$ | -1.6E-4 | 0.6E-4 | -2.6 |
| $b_3$ | 2.2E-6 | 1.6E-6 | 1.4 |
| $c_0$ | -0.0178 | 0.0024 | -7.5 |
| $c_1$ | 9.1E-4 | 2.2E-4 | 4.2 |
| $d_0$ | 0.0011 | 1.8E-4 | 6.2 |
| $d_1$ | -7.7E-5 | 1.9E-5 | -4.1 |

From Table 10 it is seen that, except for $a_3$ and $b_3$, the parameter estimates are significant at the 5% level. The coefficients $a_3$ and $b_3$ refer to the terms containing $(q_1 - P)^3_+$. This cubic term was strongly significant in the precipitation - pressure relation in the previous subsection. The value 127.6 of the $X^2$-statistic for the 12 parameter model is much lower than that for the 7 parameter model and no longer significant at the 5% level. The much better fit can also be seen from the standardized differences SDIF between the observed and predicted mean precipitation amounts in the various temperature/pressure class intervals as shown in Figure 9. The number of values in the extreme intervals is reduced to 7. Removing the two terms containing $(q_1 - P)^3_+$ from the regression model leads to a significant increase in the value of the $X^2$-statistic from 127.6 to 153.5. Further checks on the model are possible by estimating the parameters $a$, $b$, $c$ and $d$ in equation (8) for distinct intervals of surface air pressure or by estimating $a_0$, $a_1$, $a_2$ and $a_3$ in equation (10) for distinct intervals of temperature (Appendix A). The better description of the mean amounts by the 12 parameter model may be relevant for climate change scenario objectives, because the largest differences with the 7 parameter model are found for cases with relatively high mean amounts.

Seasonal variation

The 7 and 12 parameter regression models were kept constant over the year. As with the precipitation - temperature relation (Klein Tank and Buishand, 1993), a simple check on seasonal variation is possible without making assumptions about the form of the relation between precipitation, temperature and surface air pressure. Figure 10 compares the observed mean precipitation amounts with the expected means for a constant relation between those variables. The latter were obtained by replacing all observed daily amounts by the mean in their temperature/pressure class intervals. The constant relation seriously overestimates the observed monthly means during spring (with a maximum up to 23% in April) and underestimates the means for the summer months July and August. This pattern was also observed for the constant precipitation - temperature relation (Figure 5 in Klein Tank and Buishand, 1993). The main reason for the similarity between the two patterns is the small seasonal variation of the mean surface air pressure on rain days. The mean absolute difference between the observed and expected monthly means in Figure 10 is 8.4%. This is somewhat lower than the value 9.3% for the constant precipitation - temperature relation in Klein Tank and Buishand (1993). The slight improvement for the inclusion of surface air pressure is mainly caused by a better prediction of monthly mean rainfall during autumn.

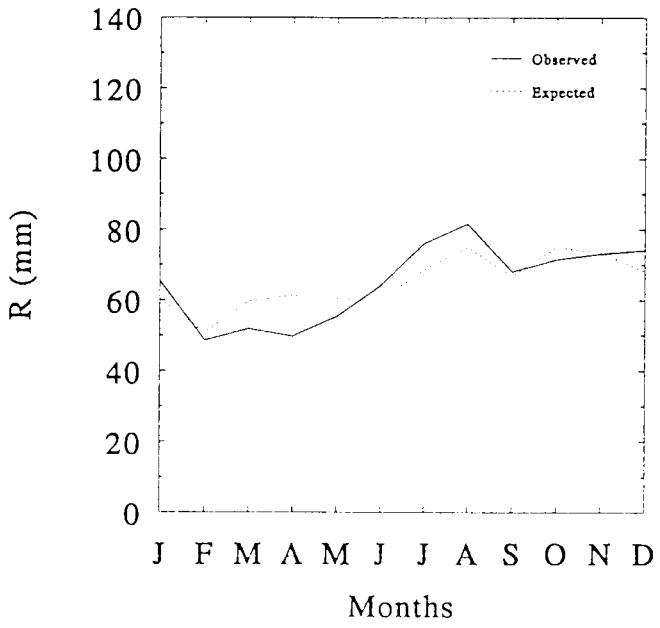



Figure 10 Expected monthly mean amounts for a constant precipitation - temperature/pressure relation (dotted) compared to the observed means at De Bilt (1906-1981).

Although mean surface air pressure shows little seasonal variation on rain days, surface air pressure still can contribute to certain aspects of the seasonal variation in the precipitation - temperature relation. In contrast to the other seasons there is a significant negative correlation between $T$ and $P$ for the winter season. A temperature rise in this season is therefore usually accompanied by a decrease in surface air pressure. The latter causes an additional increase in the expected mean precipitation amount. This partly explains the relatively high value of the coefficient $b$ in equation (8) for the winter season in Klein Tank and Buishand (1993). The above check on the monthly mean amounts is not suitable to detect this kind of seasonal variation in the precipitation-temperature relation.

Sensitivity to the definition of a rain day

The relation between the daily precipitation amount sampled at 8 UT and mean temperature/surface air pressure on the previous day (in the 0-0 UT interval) can also be described by the models in equations (12), (13) and (15). The choice of a different time interval for the precipitation amounts has very little effect on the coefficient of variation $v_R$, which was estimated as:

$$\hat{v}_R = 0.849 + 0.014\,T + 0.024\,(P-q_2)_+ \qquad (16)$$

Table 11 Estimates of the coefficients in equation (13) with their standard error (se) in case that precipitation is sampled at 8 UT over the past 24 hours instead of the 0-0 UT interval as in Table 9. The statistic $Z$ is the ratio of these two quantities.

| Coefficient | Estimate | se | $Z$ |
|---|---|---|---|
| $a_0$ | 0.4315 | 0.0234 | 18.4 |
| $a_1$ | 0.0759 | 0.0030 | 24.9 |
| $a_2$ | -0.0020 | 2.5E-4 | -7.8 |
| $a_4$ | 2.0E-5 | 0.5E-5 | 3.7 |
| $b$ | 0.0531 | 0.0034 | 15.7 |
| $c$ | -0.0057 | 0.0013 | -4.3 |
| $d$ | 2.7E-4 | 0.9E-4 | 2.9 |

where $q_2$ = 1000 hPa. Tables 11 and 12 give the estimates of the coefficients in equations (13) and (15) for the 7 and 12 parameter model, respectively. Only for the 12 parameter model is the value of the $X^2$-statistic not significant at the 5% level. The estimates of the regression coefficients $b$, $c$ and $d$ in the 7 parameter model and $b_0$, $b_1$, $b_2$, $b_3$, $c_0$, $c_1$, $d_0$ and $d_1$ in the 12 parameter model are almost all closer to zero for the 8-8 UT amounts, reflecting a less rapid increase of the mean with temperature. Such differences between 0-0 UT and 8-8 UT amounts do not exist for a change in surface air pressure. It can be concluded that somewhat different precipitation scenarios will be obtained for a warmer climate when only the 8-8 UT amounts are available instead of the 0-0 UT amounts.

Table 12 Estimates of the coefficients in equation (15) with their standard error (se) in case that precipitation is sampled at 8 UT over the past 24 hours instead of the 0-0 UT interval as in Table 10. The statistic $Z$ is the ratio of these two quantities.

| Coefficient | Estimate | se | $Z$ |
|---|---|---|---|
| $a_0$ | 0.5205 | 0.0310 | 16.8 |
| $a_1$ | 0.0604 | 0.0046 | 13.3 |
| $a_2$ | -0.0017 | 4.7E-4 | -3.6 |
| $a_3$ | 2.1E-5 | 1.1E-5 | 1.9 |
| $b_0$ | 0.0393 | 0.0054 | 7.2 |
| $b_1$ | 0.0033 | 6.7E-4 | 5.0 |
| $b_2$ | -1.4E-4 | 0.7E-4 | -2.0 |
| $b_3$ | 1.6E-6 | 1.7E-6 | 0.9 |
| $c_0$ | -0.0101 | 0.0022 | -4.5 |
| $c_1$ | 4.2E-4 | 2.1E-4 | 2.1 |
| $d_0$ | 0.0006 | 1.6E-4 | 3.9 |
| $d_1$ | -3.8E-5 | 1.7E-5 | -2.2 |

# 4 GCM predicted climate changes

General Circulation Models (GCMs) are the main tools available today for climate simulation. Most GCMs have a horizontal resolution of the order of 500 km and contain about 10 vertical levels. For some variables the GCM simulations provide a reasonable guess of possible large-scale changes. Other variables are, however, so poorly represented in the model simulations that they should not be considered at all. This will be illustrated here with output of the Canadian Climate Centre (CCC) second generation GCM (McFarlane et al., 1992; Boer et al., 1992).

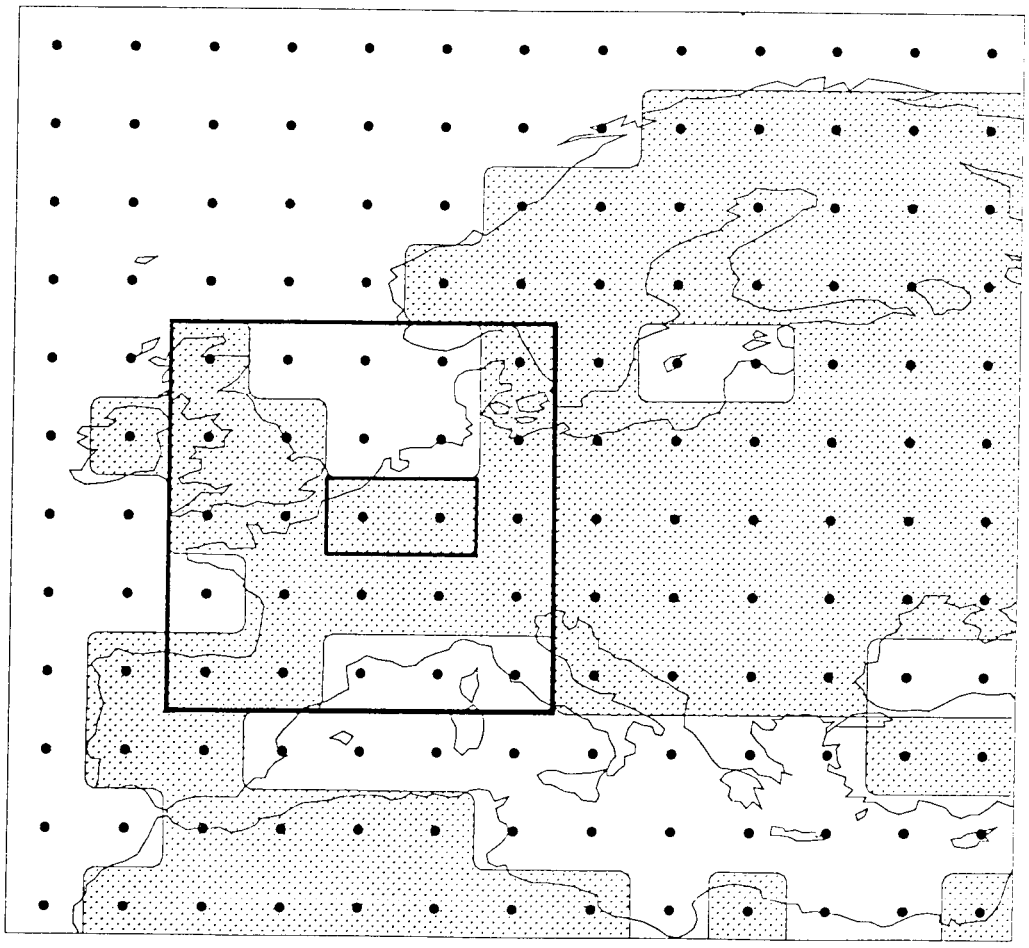


Figure 11 Geographic resolution and land-sea distribution of the Canadian Climate Centre (CCC) second generation GCM. The land gridpoints (shaded) in the small box (2 gridpoints) and the large box (16 gridpoints) are considered in this study.

Table 13 Mean percentage $\bar{N}$ of rain days, precipitation amount $\bar{R}$, temperature $\bar{T}$, surface air pressure $\bar{P}$ (reduced to sea level), relative humidity $\bar{U}$ and cloud cover $\bar{C}$ for the 10-year $1\times CO_2$ simulation of the CCC-GCM (small box in Figure 11) and for the 1961-1990 record at De Bilt†.

| Element | | Season | | | | Year |
|---|---|---|---|---|---|---|
| | | DJF | MAM | JJA | SON | |
| $\bar{N}$ (%) | | | | | | |
| | CCC-GCM | 91 * | 90 * | 92 * | 91 * | 91 * |
| | De Bilt | 60 | 53 | 48 | 59 | 55 |
| $\bar{R}$ (mm) | | | | | | |
| | CCC-GCM | 322 * | 263 * | 314 * | 331 * | 1243 * |
| | De Bilt | 194 | 176 | 214 | 220 | 803 |
| $\bar{T}$ (°C) | | | | | | |
| | CCC-GCM | 3.1 | 8.4 | 16.2 | 9.7* | 9.4 |
| | De Bilt | 2.6 | 8.4 | 16.2 | 10.2 | 9.4 |
| $\bar{P}$ (hPa) | | | | | | |
| | CCC-GCM | 1015.8 | 1014.8 | 1013.1* | 1015.8 | 1014.7 |
| | De Bilt | 1014.8 | 1014.4 | 1016.0 | 1015.3 | 1015.2 |
| $\bar{U}$ (%) | | | | | | |
| | CCC-GCM | 100 * | 95 * | 96 * | 98 * | 97 * |
| | De Bilt | 87 | 78 | 77 | 85 | 82 |
| $\bar{C}$ (%) | | | | | | |
| | CCC-GCM | 64 * | 63 | 66 | 66 | 65 |
| | De Bilt† | 74 | 62 | 66 | 67 | 67 |

† Cloud cover $\bar{C}$ only for the 5-year period 1986-1990.

* Significantly different from the values at De Bilt at the 5% level using the Welch-Aspin test (two tailed).

The output of a 10-year simulation with present-day atmospheric $CO_2$ concentrations and a 10-year simulation for an atmosphere with a doubled $CO_2$ concentration were made available on a 3.75° × 3.75° grid over Europe (Figure 11). For each grid box a number of important climate variables is archived with a time step of 12 hours or 1 day. These simulated climate data are generally not representative of the climate at a site within the gridbox. Quite often large differences are found between the long term local means and those for the $1 \times CO_2$ case. As an example Table 13 presents the seasonal and annual mean percentage $\bar{N}$ of rain days, the mean precipitation amount $\bar{R}$

and the means of the explanatory variables in the regression models for $N$ and $R$ for the 1961-1990 record at De Bilt and the two land gridpoints in the small box covering the Netherlands (Figure 11). Mean cloud cover $\bar{C}$ is considered instead of incoming solar radiation, because the latter was not available for the GCM. The relative humidity $U$ for the GCM was derived from the simulated values of temperature $T$ and specific humidity $q$ and the simulated values of surface air pressure $P$ were converted to mean sea level values (Appendix B). Because the assumption of equal variances in Student's $t$-test is not valid in a number of cases, the Welch-Aspin test (Pearson and Hartley, 1976) was used to determine the statistical significance of differences in observed and simulated means.

Most striking in Table 13 is the much larger simulated percentage of rain days than that observed. Large differences between the simulated and observed numbers of rain days have also been found for other GCMs (Reed, 1986; Beersma, 1992). It can partly be attributed to the fact that GCM output is considered to represent the mean precipitation over a grid box (Reed, 1986; Giorgi, 1990; Thomas and Henderson-Sellers, 1991). The number of rain days is not the same for each point within the grid box and even if this would be the case spatial averaging of precipitation amounts leads to an increase in the number of rain days. In the CCC-GCM the area of a grid box is about 100,000 km$^2$ at 50 degrees latitude. The annual number of days with mean rainfall $\geq$ 0.1 mm over a homogeneous region of that size is about 40% larger than that at a point (Buishand et al., 1993), giving an increase in $\bar{N}$ of about 20%. Thus about half of the differences between the simulated values of $\bar{N}$ and those observed at De Bilt should be attributed to spatial averaging rather than to the very crude description of the various physical processes in the GCM.

The poor representation of precipitation in GCMs often leads to substantial differences between the simulated and observed long term mean amounts. The simulated annual mean in Table 13 is about 50% larger than that at the De Bilt. The table further shows that the calculated mean values of the relative humidity are abnormally high for the GCM simulations. This anomaly is investigated further in Appendix B. Temperature and surface air pressure simulations are much more realistic than precipitation and humidity. The largest discrepancy for temperature in autumn (0.5°C) is just significant at the 5% level. Surface air pressure at De Bilt is only significantly underestimated during summer. The simulated mean cloud cover does also not differ much from the observed mean, except for the winter season for which $\bar{C}$ is strongly underestimated.

Because the simulated data at a single gridpoint are not comparable with an observed climate time series and the skill of GCMs to reproduce the climate at scales down to those of single grid boxes is questionable (von Storch et al., 1993) there is a large risk that differences between the means for the 2 × $CO_2$ and 1 × $CO_2$ cases at the nearest gridpoint are not representative of the true changes of the local or regional climate. Such differences not only exhibit systematic errors, but also random errors caused by

the limited lengths of the simulation runs. For precipitation, for instance, the relative standard deviation of the seasonal means for the GCM data in Table 13 is about 5%. Taking the average over a number of gridpoints lowers the random errors and it also reduces the risk of large systematic errors. Here the spatial averages of the 16 land gridpoints in the large box in Figure 11, covering Great Britain, Denmark, the Benelux, Germany, France, Switzerland and parts of Austria, Spain and Italy, are considered to obtain a better picture of the differences between the 2 × $CO_2$ and 1 × $CO_2$ climates. The choice of 16 gridpoints is rather arbitrary. It is impossible to derive an optimum number of gridpoints, because too little is known about the errors in the magnitude of the $CO_2$ effect over the region.

Table 14 Predicted mean changes in $\bar{N}$, $\bar{R}$, $\bar{T}$, $\bar{P}$, $\bar{U}$ and $\bar{C}$ for the CCC-GCM ($2\times CO_2$-$1\times CO_2$; 10-year runs) in the small box (2 gridpoints) and large box (16 gridpoints) in Figure 11.

| | Season | | | | Year |
|---|---|---|---|---|---|
| Element | DJF | MAM | JJA | SON | |
| $\Delta\bar{N}$ (%) | | | | | |
| 2 gridpoints | +5 * | -1 | -2 | -4 * | -1 |
| 16 gridpoints | +4 * | +1 | -7 * | -3 * | -2 * |
| $\Delta\bar{R}/\bar{R}$ (%) | | | | | |
| 2 gridpoints | +28 * | +21 | -26 * | -11 | +0 |
| 16 gridpoints | +25 * | +11 | -21 * | -1 | +4 |
| $\Delta\bar{T}$ (°C) | | | | | |
| 2 gridpoints | +3.2* | +2.3* | +3.5* | +3.2* | +3.1* |
| 16 gridpoints | +3.0* | +2.3* | +3.7* | +3.4* | +3.1* |
| $\Delta\bar{P}$ (hPa) | | | | | |
| 2 gridpoints | -3.9* | -0.9 | +0.5 | 0.0 | -0.7 |
| 16 gridpoints | -3.4* | -1.1 | +0.3 | -0.1 | -0.8 |
| $\Delta\bar{U}$ (%) | | | | | |
| 2 gridpoints | -2 * | -1 | 0 | 0 | -1 |
| 16 gridpoints | 0 | 0 | -2 * | -1 | -1 * |
| $\Delta\bar{C}$ (%) | | | | | |
| 2 gridpoints | +7 * | 0 | -7 * | -5 * | -1 * |
| 16 gridpoints | +5 * | 0 | -9 * | -3 * | -2 * |

* Changes significant at the 5% level using the Welch-Aspin test.

Table 14 presents the changes in the seasonal and annual means ($2 \times CO_2 - 1 \times CO_2$) of $N$, $R$, $T$, $P$, $U$ and $C$ in the CCC-GCM. The predicted mean changes for the small box (2 gridpoints) covering the Netherlands do generally not differ much from those for the large box (16 gridpoints). The temperature increase resulting from doubling the atmospheric $CO_2$ concentration in the GCM is about 3°C over Western Europe. The GCM also shows a sharp increase in mean precipitation over the region during winter and a decrease during summer (see also Houghton et al., 1990, Figure 5.6). The number of rain days changes in the same direction during these two seasons. The increase in the mean precipitation amount and the number of rain days during winter is consistent with the systematic changes in surface air pressure and cloud cover for that season. The lower rainfall frequency during summer is only accompanied by a significant decrease in cloud cover. The changes in the mean number of rain days and cloud cover are not related to the relative humidity in the GCM simulations. The mean relative humidity is, however, in both the $1 \times CO_2$ and $2 \times CO_2$ simulations close to 100% and should therefore not be considered further.

In the remainder of this report only the GCM predicted mean changes in $\bar{T}$ and $\bar{P}$ for the 16 land gridpoints over Western Europe are considered to obtain daily precipitation scenarios for a doubled $CO_2$ climate. The use of the changes in $\bar{T}$ and $\bar{P}$ is in line with the conclusions in Houghton et al. (1990, pp. 102, 108) that the recent higher resolution GCMs are capable of generally realistic simulations of surface air temperature and surface air pressure. This gives some confidence in the magnitude of changes predicted by these models.

# 5 Changes in the occurrence of rain based on statistical relations

In section 2 the dependence of rainfall occurrence at De Bilt on surface air pressure $P$, relative humidity $U$ and relative incoming solar radiation $Q$ was described. This dependence can be used to model the changes in rainfall occurrence given a systematic change in one or more of those variables. In the previous section it was suggested to base the change in the seasonal means of $P$ on GCM simulations. Regression analysis offers the possibility to estimate the accompanying systematic changes in $U$ and $Q$. This is demonstrated in this section with the monthly values in the 1961-1990 record at De Bilt. The linear Model A in Section 2 is used to estimate the changes in the percentage of rain days resulting from the systematic changes in $P$ over Western Europe in the CCC-GCM. Difficulties with the extension to daily values are identified. Methods to transform an observed sequence of wet and dry days with the logistic Model B in Section 2 are discussed and illustrated with the systematic changes in the percentage of rain days at De Bilt resulting from the earlier mentioned pressure perturbation.

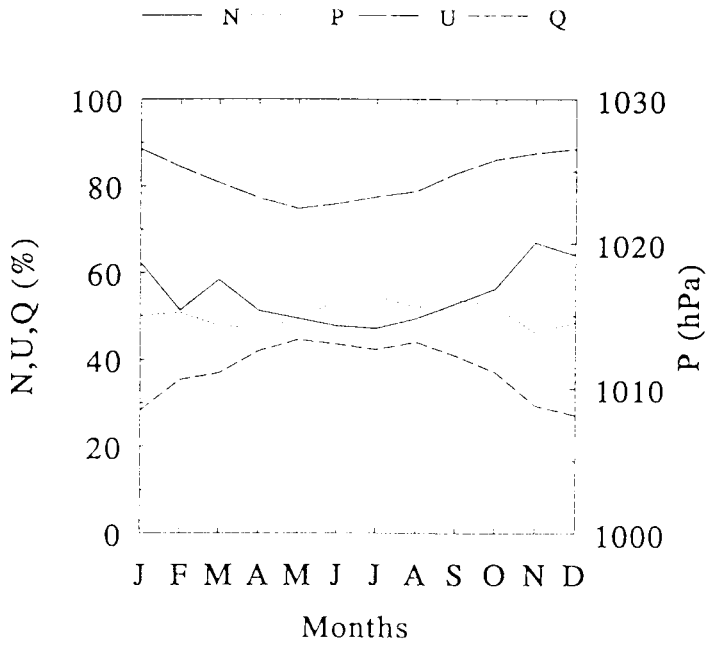



Figure 12 Annual cycle of $P$, $U$, $Q$ and $N$ at De Bilt (1961-1990).

Changes in monthly values

For the monthly values a linear regression model suffices to derive the changes in $U$ and $Q$ resulting from a pressure perturbation. The estimates of $U_i$ and $Q_i$ from the fitted models to the De Bilt data are given by:

$$\hat{U}_i = 260.94 + 6.67\cos(2\Pi i/12) - 0.18P_i \qquad i=1,...,I \qquad (17)$$

$$\hat{Q}_i = -457.39 - 7.83\cos(2\Pi i/12) + 0.49P_i \qquad i=1,...,I \qquad (18)$$

with $I$ the total number of months considered. A cosine term is included in these equations because there is a marked annual cycle in $U$ and $Q$ whereas $P$ shows little seasonal variation. From Figure 12 it is seen that $U$ is relatively high in winter. On the other hand the mean of $Q$ reaches its maximum in summer. When the relations (17) and (18) also hold in a perturbed climate, a constant decrease in $P_i$ of 3 hPa results in an increase in $\hat{U}_i$ of 0.54% and a decrease in $\hat{Q}_i$ of 1.47%.

The reported changes in $U$ and $Q$ are rather modest as a result of the weak correlation with $P$ and could perhaps be neglected for climate change scenario objectives. For the 1961-1990 record at De Bilt the estimated correlation coefficients between $P_i$ and $U_i$ and $P_i$ and $Q_i$ are -0.19 and 0.32, respectively. This does not improve when the annual cycles are removed. Nevertheless the values are statistically significant ($\alpha < 0.05$).

Although $P$ is the most important variable in the regression relation for rainfall occurrence, large changes in $U$ and $Q$ also lead to a considerable increase or decrease in the number of rain days. Because of the rather strong relation between the two variables, it is not justified to make a change in $U$ without changing $Q$. The estimated correlation coefficient between the monthly values $U_i$ and $Q_i$ in the record at De Bilt is -0.83 (-0.63 when the annual cycles in $U_i$ and $Q_i$ are removed). For the monthly values at De Bilt the change in $Q_i$ can be based on the following relation with $U_i$:

$$\hat{Q}_i = 128.85 - 1.11\,U_i \qquad i=1,...,I \qquad (19)$$

Provided that this regression relation holds for the perturbed climate, an increase in $U_i$ of for instance 5% results in a decrease in $\hat{Q}_i$ of 5.55%. According to the linear multiple regression model of Section 2 there will be an increase in the percentage of rain days of 6.43% ($\approx$ 2 days per month) with these changes in $U_i$ and $Q_i$, and no change in $P_i$. This is in sharp contrast with Bultot and Gellens (1989) who changed $U_i$ in their climate change scenario with 5% without changing $Q_i$ and the percentage of rain days.

Table 15 Changes in the mean percentage $\bar{N}$ of rain days in perturbed climates at De Bilt using Model A. The monthly values of $P$, $U$ and $Q$ in the 1961-1990 record were changed according to the GCM prediction in Table 14 ($2\times CO_2$-$1\times CO_2$, 16 gridpoints), equation (17) and equation (18), respectively.

| Perturbation | Season | | | | Year |
|---|---|---|---|---|---|
| | DJF | MAM | JJA | SON | |
| $\Delta\bar{N}$ (%) | | | | | |
| perturbation in $P$ only | +5 | +2 | 0 | 0 | +2 |
| perturbations in $P,U,Q$ | +7 | +2 | -1 | 0 | +2 |

Example: scenario in case of a perturbation in surface air pressure

As an example of the changes in the monthly mean percentage of rain days resulting from a pressure perturbation, the seasonal mean changes in surface air pressure $P$ due to doubling $CO_2$ in the CCC-GCM are considered for a large area surrounding the Netherlands (Table 14). $U$ and $Q$ were either kept the same as in the present-day climate or changed according to equations (17) and (18). Table 15 presents the changes in the mean percentage of rain days in the perturbed climates at De Bilt as obtained with Model A.

There is a large increase in the percentage of rain days during winter. A much smaller increase is found in spring and there is no or little change during the other seasons. Negating the changes in $U$ and $Q$ resulting from the pressure perturbation has only a slight effect on the predicted change in rainfall occurrence. Only for the winter and spring is there a reasonable agreement with the changes in the percentage of rain days predicted by the GCM (see Table 14). In contrast to the scenario in Table 15, the $2\times CO_2$ GCM run shows a strong decrease in the number of rain days during the summer and autumn. This decrease is accompanied by a systematic change in cloud cover $C$, whereas the changes in $P$ are small. The changes in $U$ and $Q$ from the climatological relation with $P$ in the second scenario in Table 15 are small and at variance with those expected from the change in $C$.

Changes in daily values and transformation of the sequence of wet and dry days

The linear models in equations (17), (18) and (19) always add a constant $\Delta U_i$ or $\Delta Q_i$ to the observed values $U_i$ and $Q_i$. For daily data this can be problematic because of the physical boundaries of these elements. Relative humidity $U$ is for instance quite often

close to its upper boundary 100%. Because in the Netherlands $U$ is always far away from its lower boundary, a change in the observed daily values $U_j$ can be achieved with the transformation $f(100\text{-}U_j)$. The factor $f$ can be estimated from the monthly mean values by fitting a linear regression model to $\ln(100\text{-}U_i)$. For the record at De Bilt the fitted relation is given by:

$$\ln(100 - \hat{U}_i) = -5.99 - 0.39\cos(2\Pi i/12) + 0.009 P_i \qquad i=1,...,I \tag{20}$$

The result of a decrease in $P$ of 3 hPa is now that the value of $(100\text{-}\hat{U}_i)$ changes with a factor $f = e^{-3\times0.009} = 0.97$. The same factor applies to the daily values. For De Bilt the long term annual mean humidity $\bar{U}$ is about 82%. Because $\Delta(100\text{-}\bar{U}) = (f\text{-}1)(100\text{-}\bar{U}) = -0.03\,(100\text{-}82) = -0.54$ the change in $\bar{U}$ is +0.54%. This corresponds well with the increase found earlier for the monthly values using equation (17).

Another transformation is necessary to change the daily values of $Q$. Especially during the winter season $Q$ is frequently close to its lower boundary zero in the Netherlands. One should further take into account that the temperature increase resulting from the higher concentration of greenhouse gases in the atmosphere has little effect on the values of $Q$ for days with clear skies. Rather advanced techniques are needed to transform the daily values of $Q$.

There are various methods to obtain a sequence of wet and dry days for a perturbed climate where the frequency of wet days differs from that in the present-day climate. If more rain days are expected in the perturbed climate, then dry days with high values of the probability $\hat{p}_j^*$ of rain in the perturbed climate can be transformed into rain days. It is possible to limit this transformation to days that are preceded or followed by a rain day. In case of an increasing frequency of dry days, rain days can be transformed in dry days in a similar way.

In Section 2 it was noted that for the logistic model the total number of wet days is estimated as:

$$m = \sum_{j=1}^{K} \hat{p}_j \tag{21}$$

with $K$ the number of days in the period considered. Now let us assume that, for instance in winter (DJF), the probability that day $j$ is wet changes from $\hat{p}_j$ to $\hat{p}_j^*$ due to a systematic change in surface air pressure in that period. The change in the number of wet days is then estimated as:

$$\Delta m = \sum_{j=1}^{K} \hat{p}_j^* - \sum_{j=1}^{K} \hat{p}_j \tag{22}$$

where $\Delta m$ is usually rounded to the nearest integer. For $\Delta m > 0$, the $\Delta m$ dry days

with the highest values of $\hat{p}_j^*$ can be transformed into wet days, whereas for $\Delta m < 0$, the $\Delta m$ wet days with the lowest values of $\hat{p}_j^*$ become dry days.

In general high values of $\hat{p}_j^*$ are found on days with low surface air pressure. The expected precipitation amount is relatively high for such days. Another selection procedure for the additional wet days will generally lead to a lower value of the expected precipitation amount for these days. Therefore, instead of choosing the dry days with the highest values of $\hat{p}_j^*$ a simple alternative based on simulation of binary variables is also considered. For each dry day a binary variable $Y_j$ is drawn giving the status of that day in the perturbed climate (1 = wet and 0 = dry). In order that the expected number of additional wet days equals $\Delta m$ the probability $\Pi_j$ that $Y_j = 1$ should be equal to:

$$\Pi_j = \Delta m \; \hat{p}_j^* \; / \; \sum_{j=1}^{L} \hat{p}_j^* \qquad j = 1,...,L \qquad (23)$$

with $L$ the total number of dry days in the period considered. The probability that day $j$ will be transformed into a wet day increases thus linearly with $\hat{p}_j^*$. Just as in the observed record there will be dry days with a high probability of rain in the transformed record. For the case $\Delta m < 0$ a similar procedure can be applied to the observed wet days.

Example: scenario of wet and dry days in case of a perturbation in surface air pressure

Equation (22) was used to obtain the change in the number of rain days at De Bilt from the seasonal mean changes in $P$ in the CCC-GCM. The daily values of $U$ and $Q$ were left unchanged. The probabilities of rain in the right-hand side of equation (22) were derived from the logistic model using the parameter estimates in Table 3. It was found that the value of $\Delta m$ varied between +4.7 days per year in the winter season ($\Delta \bar{N}$ = +5.2%) and -0.5 days per year in the summer season ($\Delta \bar{N}$ = -0.6%). As expected the change in the percentage of rain days is almost identical to that in Table 15 where the linear model was applied to monthly values.

# 6 Changes in the amount of precipitation based on statistical relations

The models in Section 3 for the dependence of precipitation amount on temperature and surface air pressure can be used to generate representative and internally consistent daily time series of these weather variables for a $CO_2$ induced warmer climate. Although a change in the $CO_2$ content of the atmosphere almost surely leads to a systematic change in $T$, it is much less certain whether this will also be the case for $P$. Therefore the situation that there is no systematic change in $P$ is discussed first. Then the changes in the seasonal means of $T$ and $P$ in the CCC-GCM are used to obtain a synthetic daily rainfall sequence for De Bilt in a doubled $CO_2$ climate from the 1961-1990 record at that location.

Changes in the amount of precipitation on rain days resulting from a temperature perturbation

When there is only a temperature perturbation it is reasonable to assume that the number of rain days remains the same as in the present-day climate. The precipitation amounts in the perturbed climate can then be obtained by multiplying all daily precipitation values in the observed record by a temperature and pressure dependent factor, representing the predicted change in the mean amount according to the regression model. The factor $F$ is given by:

$$F(T,T^*,P) = \exp\ [k(T^*,P) - k(T,P)] \tag{24}$$

with $T^*$ the temperature in the perturbed climate. For the 7 parameter model in equation (13) the factor no longer depends on $P$ and equation (24) reduces to:

$$F(T,T^*) = \exp\ [b\{T^*-T\} + c\{(T^*-m_1)_+^2-(T-m_1)_+^2\} + d\{(T^*-m_1)_+^3-(T-m_1)_+^3\}] \tag{25}$$

This expression is identical to that derived in Klein Tank and Buishand (1993) from the precipitation - temperature relation.

As an example, the situation of a constant temperature increase of 3 °C and no change in surface air pressure at De Bilt is considered. All daily amounts in the 1961-1990 record were multiplied by $F$ derived from the 4 parameter model in equations (6) and (8) with only temperature as explanatory variable, the 7 parameter model in equations (12) and (13) and the 12 parameter model in equations (12) and (15).

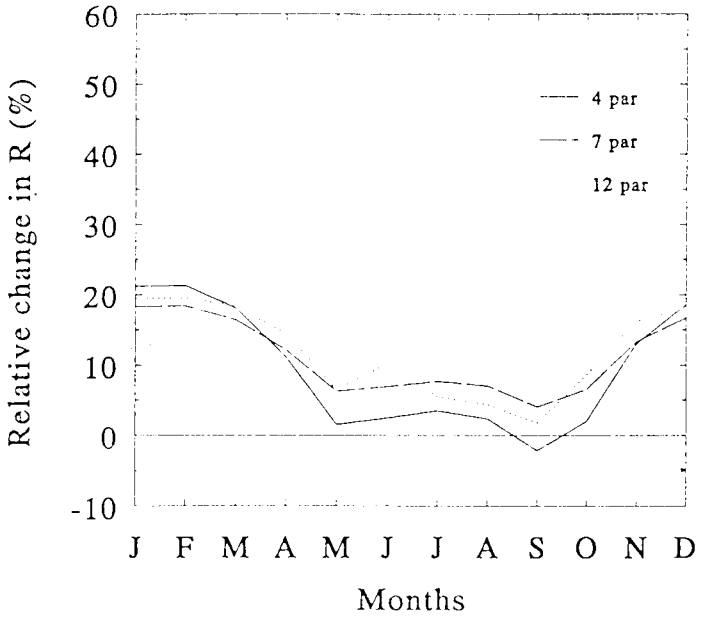



Figure 13 Changes in the mean precipitation amount $\bar{R}$ in a perturbed climate at De Bilt. All daily temperatures in the 1961-1990 record were increased by 3 °C. The multiplying factor $F$ was based on the 4 parameter model with only $T$ as explanatory variable (solid line), the 7 parameter model (dashed line) and the 12 parameter model (dotted line). In the last two cases surface air pressure was kept the same as in the present-day climate.

For all three cases Figure 13 shows the relative differences between the monthly means of the new record and the observed means. The multiplying factors were applied to all wet days in the record of De Bilt including wet days at very low temperatures ($T$<-7°C), which were discarded for parameter estimation.

The higher temperatures in the perturbed climate give rise to an increase in the mean precipitation amount compared to the historical observations at De Bilt. For all three models the relative increase in the monthly mean is about 20% for the winter season; the relative increase is nearly always less than 10% for the period May-October. The seasonal cycles for the 7 and 12 parameter models in $T$ and $P$ are almost identical. In comparison with the historical observations the annual mean increases with about 10%. The interannual standard deviation changes with the same factor as the annual mean.

The use of the 4 parameter precipitation - temperature relation does in fact assume that $P$ changes according to the present-day relation between $P$ and $T$ rather than that there is no change in $P$. Because there is a weak correlation between $P$ and $T$ the results

from this model slightly differ from the models containing both $P$ and $T$. The latter should be preferred.

Because in a warmer world temperatures are considered which have never occurred, the model for the relation between precipitation and temperature/surface air pressure is somewhat extrapolated beyond the range of observed data. Klein Tank and Buishand (1993) showed that for the 4-parameter precipitation-temperature relation the multiplying factor $F(T,T^*)$ in equation (25) rapidly grows at the upper end of the temperature range. For the case $T^* = T + 3$°C they found that $F$ varied between 0.95 at $T = 12$°C and 2.07 at $T = 24$°C. When the 7 parameter model in equations (12) and (13) for the precipitation-temperature/surface air pressure relation is used, the extreme values of $F$ are reached at the same temperatures (1.02 at $T = 12$°C and 1.82 at $T = 24$°C). The multiplying factor for the 12 parameter model in equations (12) and (15) shows a different behaviour.

Table 16 Multiplying factor $F$ (with its standard error se) for values of $T$ and $P$ along the line $P = 960 + 2.3T$ in Figures 6 and 8 when there is a constant increase in temperature of 3°C on every day and no change in surface air pressure.

| | | | 12 par model | | 7 par model | |
|---|---|---|---|---|---|---|
| $T$ | $T^*$ | $P$ | $F$ | se | $F$ | se |
| 3 | 6 | 967.0 | 1.20 | .33 | 1.23 | .01 |
| 4 | 7 | 969.3 | 1.17 | .27 | 1.23 | .01 |
| 5 | 8 | 971.7 | 1.18 | .22 | 1.22 | .01 |
| 6 | 9 | 974.0 | 1.23 | .19 | 1.19 | .01 |
| 7 | 10 | 976.3 | 1.30 | .17 | 1.14 | .01 |
| 8 | 11 | 978.7 | 1.34 | .16 | 1.10 | .01 |
| 9 | 12 | 981.0 | 1.32 | .14 | 1.07 | .01 |
| 10 | 13 | 983.3 | 1.26 | .11 | 1.04 | .01 |
| 11 | 14 | 985.7 | 1.17 | .09 | 1.03 | .01 |
| 12 | 15 | 988.0 | 1.07 | .07 | 1.02 | .01 |
| 13 | 16 | 990.3 | .97 | .07 | 1.02 | .01 |
| 14 | 17 | 992.7 | .88 | .07 | 1.03 | .02 |
| 15 | 18 | 995.0 | .81 | .08 | 1.04 | .02 |
| 16 | 19 | 997.3 | .77 | .09 | 1.07 | .03 |
| 17 | 20 | 999.7 | .74 | .10 | 1.10 | .05 |
| 18 | 21 | 1002.0 | .75 | .11 | 1.14 | .06 |
| 19 | 22 | 1004.3 | .79 | .12 | 1.19 | .08 |
| 20 | 23 | 1006.7 | .89 | .13 | 1.25 | .11 |
| 21 | 24 | 1009.0 | 1.06 | .15 | 1.33 | .14 |
| 22 | 25 | 1011.3 | 1.37 | .19 | 1.42 | .18 |
| 23 | 26 | 1013.7 | 1.93 | .30 | 1.53 | .23 |
| 24 | 27 | 1016.0 | 2.98 | .63 | 1.67 | .29 |

Consider for instance the values of $F$ in Table 16 along the dashed line $P = 960 + 2.3T$ in Figures 6 and 8. The lowest value of $F$ in the table is 0.74 ($T = 17°C$), the highest value 2.98 ($T = 24°C$). For the record of De Bilt very low values of $F$ can be reached on wet days with $P$ and $T$ values under the line $P = 960 + 2.3T$. For instance at $T \approx 18°C$ and $P = 995$ hPa the value of $F$ drops below 0.6. The cubic term in $T$ in equation (15) is then negative and dominates the factor. There is no clear physical explanation for this behaviour. To avoid such low values of the factor, a lower bound of $F$ can be introduced (for instance $F = 0.70$) or, for days with $T$ and $P$ values under the line $P = 960 + 2.3T$, the factor can be set equal to the value at the point on the line where the surface air pressure equals $P$. Since this is only necessary for a relatively small fraction of the wet days it hardly influences the changes in the long term monthly means given in Figure 13.

Apart from the very low values of $F$ for wet days with moderate temperatures under the line $P = 960 + 2.3T$, the multiplying factor for the 12 parameter model becomes extremely large at high temperatures if $P > 1006$ hPa. The cubic term in $T$ is then positive. In the record of De Bilt the multiplying factor reaches values as high as 6.3 at $T = 24°C$ and $P = 1022$ hPa. A less sharp rise of the factor at high temperatures can be achieved by introducing a second knot $m_2 > m_1$ after which $k(T,P)$ is taken again linear in $T$. When $m_2 = 21°C$ the highest value $F = 3.21$ in the record of De Bilt is reached at $T = 21°C$ and $P = 1030$ hPa. The large value of the factor is of little interest here because the precipitation amounts are small if $P = 1030$ hPa.

In the 7 and 12 parameter models the coefficient of variation varies with $T$ and $P$. The use of a factor does not preserve the relation between $v_R$ and these variables in equation (14). More advanced techniques are necessary to achieve that this relation also holds in the perturbed climate. The need for such techniques is questionable because the change in $v_R$ is rather modest (not more than 5% in case of a 3°C temperature increase and no change in $P$).

Precipitation scenario in case of perturbations in both temperature and surface air pressure

In case of a systematic change in both $T$ and $P$ the multiplying factor $F$ becomes:

$$F(T,T^*,P,P^*) = \exp\,[k(T^*,P^*) - k(T,P)] \qquad (26)$$

where $T^*$ and $P^*$ refer to the values of $T$ and $P$ in the perturbed climate. For the 7 parameter model $F$ can be written as the product of the temperature dependent factor in equation (25) and a pressure dependent factor of the form:

$$F(P,P^*) = \exp\,[a_1\{P-P^*\} + a_2\{(q_1-P^*)_+^2 - (q_1-P)_+^2\}$$

$$+ a_3\{(q_1-P^*)_+^3-(q_1-P)_+^3\}] \qquad (27)$$

The systematic change in $P$ is generally accompanied by a change in the number of rain days. The techniques in Section 5 should therefore be used to alter the observed sequence of wet and dry days. The precipitation amount on a new wet day can be based on the regression model for $R$. The simplest method is to use the predicted value $\exp[k(T^*,P^*)]$ from the model. A more realistic precipitation record that better reproduces the daily variability is obtained by generating these amounts from an exponential distribution (Appendix C). The various steps to obtain a scenario for the case of perturbations in both temperature and surface air pressure are summarized in Appendix D.

The scenario construction method was applied to the 1961-1990 record at De Bilt using the seasonal changes in temperature and surface air pressure over Western Europe in the CCC-GCM as presented in Table 14. Relative humidity and relative incoming solar radiation were left unchanged. Table 17 presents the relative differences between the seasonal and annual means of the transformed daily precipitation record and the original means. The multiplying factor $F$ was based on the 12 parameter model. For the determination of additional wet or dry days, both the method that changes the day status according to the highest or lowest values of $\hat{p}_j^*$ and the alternative Monte-Carlo method of generating binary variables were considered (see Section 5). On the new dry days (on average 0.5 day per year in the summer season) the precipitation amount in the scenario was set to zero. For the new wet days (on average 4.7, 1.7 and 0.2 days per year during winter spring and autumn, respectively) only the predicted mean amount $\exp[k(T^*,P^*)]$ was used for the computation of the seasonal and annual means.

Table 17 Changes in the mean precipitation amount $\bar{R}$ in perturbed climates at De Bilt. The daily values of $T$ and $P$ in the 1961-1990 record were changed according to the GCM predictions in Table 14 ($2\times CO_2$-$1\times CO_2$, 16 gridpoints). The multiplying factor $F$ was based on the 12 parameter model. A change in the number of rain days was achieved either by a change in the day status according to the highest or lowest values of $\hat{p}_j^*$ (I) or by generating binary variables (II).

| | Season | | | | Year |
|---|---|---|---|---|---|
| Method | DJF | MAM | JJA | SON | |
| $\Delta\bar{R}/\bar{R}$ (%) | | | | | |
| I | +47 | +22 | +9 | +11 | +22 |
| II | +44 | +20 | +8 | +11 | +20 |

The annual mean in this scenario increases with about 20%, whereas the interannual standard deviation increases with 14%. Due to the change in the number of rain days the mean and standard deviation no longer change with the same factor. The relative changes in the winter season are more than twice as high as in the other seasons. The method used to determine the additional wet or dry days has little effect on the changes in the seasonal and annual means. Almost identical results are obtained with the use of the 7 parameter model instead of the 12 parameter model.

As a result of the substantial systematic change in $T$ and $P$, this scenario is extremely wet. The percentage increase is much higher than that predicted by the CCC-GCM for the $2\times CO_2$ climate over Western Europe (Table 14). The predicted precipitation changes in the CCC-GCM are thus strongly at variance with those expected from the accompanying changes in $T$ and $P$. This discrepancy is partly caused by the fact that in the scenario $U$ and $Q$ were not altered, which is not consistent with the systematic changes in cloud cover in the GCM simulations. In contrast to the scenario there is a strong decrease in the number of rain days in the CCC $2\times CO_2$ summer and autumn climate, which can be attributed to the substantial decrease in cloud cover during these seasons. It is further possible that the relations between precipitation and other elements are not preserved in the GCM simulations as is assumed in the scenario. The predicted changes in mean precipitation from climatological relations are, however, a useful alternative to the GCM predictions because precipitation is very poorly reproduced in these models. Even if one wishes to base a scenario on the GCM predicted seasonal precipitation changes, one still has to find a method to obtain representative and internally consistent daily time series.

The unequal changes in $T$ and $P$ for each season result in different values of the temperature and pressure dependent factor $F$ in the scenario. It is possible to avoid abrupt changes in $F$ at the transition of seasons by describing the predicted changes by continuous periodic functions. Whether this is necessary is an open question because abrupt changes in $F$ are masked in the transformed series as a result of the day to day variation in $R$, $T$ and $P$.

# 7 Conclusions and discussion

Simple deterministic transformations of observed climate records have been popular to obtain time series for climate change impact studies. For daily rainfall data the most widely used method assumes that the number of rain days remains unchanged. Extensions that allow for changes in the sequence of rain days have recently been considered. A sensible approach is to base the inclusion or removal of rain days on an estimate of the probability of rain from the values of other meteorological variables. It is further important that the transformation of the precipitation amounts is consistent with the changes in other elements. This can be achieved by regressing the precipitation amounts on these elements.

For the record at De Bilt the percentage of rain days in a given month can be related to the monthly mean values of surface air pressure $P$, relative humidity $U$ and relative incoming solar radiation $Q$ by means of the standard linear regression model as presented in this report. The non-linear logistic model, however, is needed to obtain an estimate of the probability of rain for individual days. The four parameters in the model were taken constant over the year in this study. For a complete description of the annual cycle in the number of rain days, it is, however, necessary to incorporate seasonal variation in the parameters. The best fit to the data is obtained with a precipitation threshold of 0.1 mm.

The regression model for the relation between precipitation amount $R$ on rain days and daily mean temperature $T$ presented in Klein Tank and Buishand (1993) can be extended with surface air pressure to take partly into account the possible changes in atmospheric circulation that may accompany a temperature change in a future climate. There is a complicated interaction between the effects of temperature and surface air pressure on the amount of precipitation. Nevertheless a 7 parameter model which negates this interaction reasonably describes the variation of the mean precipitation amount with temperature and surface air pressure. A 12 parameter model, which includes several interaction terms, is needed to obtain a non-significant value of the $X^2$-statistic for lack-of-fit. Both models with constant regression coefficients over the year cannot fully explain the seasonal variation in the monthly means. Besides a temperature and pressure dependent multiplying factor to adjust the daily amounts on existing rain days, the models provide the expected rainfall amount on new wet days, i.e. observed dry days that were changed into wet days.

Sensible values for the changes in $T$, $P$, $U$ and $Q$ are required before the regression models for rainfall occurrence and precipitation amounts on rain days can be used to transform the observed time series of daily precipitation into a daily time series suitable for climate change impact studies. One possibility is to derive these changes from regional GCM output. This is only recommended, however, if the GCM can

reasonably reproduce the variable of interest. Statistical relations for the observed climate can also be used to obtain a consistent set of changes in $T$, $P$, $U$ and $Q$. Such a relation can be a linear regression model when the changes in monthly means are sufficient. More complex models will generally be needed to represent the relation between daily values, in particular for variables with clear physical boundaries ($U$ and $Q$).

A transformation of the observed daily rainfall sequences at De Bilt based on the seasonal mean changes in $T$ and $P$ as predicted by the CCC-GCM for the $2\times CO_2$ climate leads to an increase in the annual mean of about 20%. For the winter months (DJF) the increase is even twice as much. These results only slightly depend on how probabilities of rain are translated into wet and dry days and the choice between the 7 and 12 parameter model for the precipitation amounts. The GCM $2\times CO_2$ climate shows a less extreme increase in the mean amounts over Western Europe than the above transformed record. The GCM predicted changes in the mean amounts are not necessarily better because precipitation is very poorly represented in the model simulations. Different precipitation scenarios may be obtained from the regional predictions of $T$ and $P$ in other GCMs. In particular, recent transient experiments with coupled ocean-atmosphere models could be considered.

The derivation of a scenario from statistical relations for the present-day climate generally assumes that these relations remain unchanged. For this reason temperature was not considered in the logistic model for rainfall occurrence, since it is very unlikely that the relation between rainfall occurrence and temperature is preserved in the doubled $CO_2$ climate. For instance, as warm summer days are often dry, a temperature increase would lead to a decrease in the number of rain days during summer according to that relation. In general, it is, however, rather difficult to verify the constancy of statistical relations under climate change conditions. Regional GCM data are of limited use for this purpose. Verification of the statistical precipitation relations in this report is hampered by the high temporal variation of daily precipitation and the poor description of this element in the GCM simulations.

A marked distinction between our transformation method and the stochastic simulation procedure in Richardson (1981) is that the latter considers precipitation first. The other variables are conditionally generated on the sequence of wet and dry days. The relation between temperature and the amount of precipitation is ignored in that approach. Richardson's model assumes that the maximum temperature, the minimum temperature and the incoming solar radiation follow a three dimensional normal distribution. It is well-known that the normal distribution cannot adequately describe the distribution of the daily values of $Q$ (Bennett, 1975; Amato et al., 1985; Graham et al., 1988). In the Netherlands the distribution of the daily incoming solar radiation is negatively skewed for the dry days in the summer season and positively skewed with a relatively large coefficient of variation during the winter season. The normal distribution then overestimates $Q$ for clear summer days and occasionally leads to

negative values during winter (about 10% of the wet days and 5% of the dry days). Daily temperature distributions also may differ from the normal distribution (Toth and Szentimrey, 1990). The normal distribution is, however, the only suitable distribution to generate multivariate time series. Transformation of the original variables might be considered to achieve normality. But also then it is difficult to treat the variation of the incoming solar radiation within its physical boundaries well.

Richardson's model assumes first-order Markov dependence. The model reproduces the lag zero and lag one cross correlation coefficients between the maximium temperature, the minimum temperature and the incoming solar radiation. This does not guarantee, however, that all relations between these elements are preserved and thus there is a possibility that irrealistic combinations are generated. This risk grows if other variables (relative humidity, wind) are added to the model. It is further well-known that first-order Markov dependence is insufficient to reproduce the complex autocorrelation structure of climate time series (Gregory et al., 1993; Woolhiser et al.,1993). The variances of monthly and annual precipitation amounts are therefore often underestimated and prolonged droughts as well as periods of extreme wetness are not well preserved. This shortcoming cannot easily be remedied. The duration of heat waves and cold spells also strongly depends on the autocorrelation structure. The model's failure to reproduce long period variation may put limitations on its use for climate change impact studies.

Statistical downscaling of GCM simulations can lead to large errors in the long term seasonal and annual mean amounts, because present-day high resolution GCMs poorly reproduce the frequencies of the various circulation types, like the number of days with a cyclonic or anticyclonic circulation (Hulme et al., 1993; Beersma, 1994). It is also unlikely that the generated rainfall sequences will have the right autocorrelation structure due to simplifications in the stochastic precipitation model and because the GCMs have difficulties to describe the persistence of the large-scale circulation (Hughes et al., 1993). The classification of the circulation types might not be optimal for predicting rainfall occurrence, the amount of precipitation and their changes. Extension with other variables than precipitation seems desirable. This will, however, meet similar difficulties as in the Richardson approach.

Deterministic downscaling of GCM simulations has not yet developed far enough to be useful for a broad range of impact studies. Limited computer resources prevent simulations of sufficiently long climate time series. Moreover, this technique is unable to provide more detailed information than the resolution of the nested LAM, which is currently of the order of 50 km. Systematic errors in the large-scale GCM circulations are transmitted to this scale as the GCM provides the boundary conditions of the nested model.

It can be concluded that the grounds for using a relatively simple transformation of observed records instead of stochastic or deterministic simulation methods are rather

strong. Problems with dissimilarities between observed and simulated climates for the $1\times CO_2$ case do not exist in the transformation method. Nevertheless the regression of daily precipitation amounts on two or more weather variables causes already some difficulties. Further complications might be expected when seasonality is incorporated in the regression models. The necessity of this extension can, however, be questioned because of the large uncertainties in the changes of the covariates. Transformation of observed records becomes cumbersome when climate change is caused by a change in the mean direction of the air flow or by a change in the mean duration of blockings and depressions. Although there is some speculation that such changes may accompany the global greenhouse warming, their magnitude and direction are extremely uncertain. The transformation of the daily values of $U$ and $Q$ needs further attention. The fact that these two variables were kept the same precludes a scenario with a strong increase in extreme summer droughts, which may have important impacts.

# References


Amato, U., A. Andretta, B. Bartoli, B. Coluzzi, V. Cuomo and C. Serio, 1985. Stochastic modelling of solar-radiation data. Il Nuovo Cimento, **8 C**, 248-258.

Arnell, N.W., 1992. Factors controlling the effects of climate change on river flow regimes in a humid temperate environment. J. Hydrol., **132**, 321-342.

Azzalini, A., A.W. Bowman and W. Härdle, 1989. On the use of nonparametric regression for model checking. Biometrika, **76**, 1-11.

Bárdossy, A. and E.J. Plate, 1992. Space-time model for daily rainfall using atmospheric circulation patterns. Water Resour. Res., **28**, 1247-1259.

Beersma, J.J., 1992. GCM control run of UK Meteorological Office compared with the real climate in the NW European winter. KNMI Scientific Report WR 92-02, De Bilt.

Beersma, J.J., 1994. Storm activity over the North Sea and the Netherlands in two climate models compared with observations. KNMI Scientific Report WR 94-02, De Bilt.

Bennett, I., 1975. Variation of daily solar radiation in North America during the extreme months. Arch. Met. Geoph. Biokl., **23 B**, 31-57.

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

Buishand, T.A., B. van Mourik and A.M.G. Klein Tank, 1993. The effect of spatial averaging on threshold exceedances of daily precipitation amounts. KNMI Technical Report TR 154, De Bilt.

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. J. Hydrol. **99**, 319-347.

Bultot, F. and D. Gellens, 1989. Simulation of the impact of $CO_2$ atmospheric doubling on precipitation and evapotranspiration - Study of the sensitivity to various hypotheses. In: Proceedings Conference on Climate and Water, 11-15 September 1989, Valtion Painatuskeskus, Helsinki, Finland, 73-92.

Bultot, F., D. Gellens, M. Spreafico and B. Schädler, 1992. Repercussions of a $CO_2$ doubling on the water balance - a case study in Switzerland. J. Hydrol., **137**, 199-208.

Buse, A., 1982. The likelihood ratio, Wald, and Lagrange multiplier tests: An expository note. The American Statistician, **36**, 153-157.

Cooley, K.R., 1990. Effects of $CO_2$-induced climatic changes on snowpack and streamflow. Hydrol. Sci. J., **35**, 511-522.

Cox, D.R. and E.J. Snell, 1989. Analysis of Binary Data, Second Edition. Chapman and Hall, London.

Duvroye, L., 1986. Non-Uniform Random Variate Generation. Springer, Berlin.

Giorgi, F., 1990. Simulation of regional climate using a limited area model nested in a general circulation model. J. Climate, **3**, 941-963.

Giorgi, F. and L.O. Mearns, 1991. Approaches to the simulation of regional climate change: A review. Rev. Geophys., **29**, 191-216.

Giorgi, F., C.S. Brodeur and G.T. Bates, 1994. Regional climate change scenarios over the United States produced with a nested regional climate model. J. Climate, **7**, 375-399.

Gleick, P.H., 1987. Regional hydrologic consequences of increases in atmospheric $CO_2$ and other trace gases. Clim. Change, **10**, 137-161.

Graham, V.A., K.G.T. Hollands and T.E. Unny, 1988. A time series model for $K_t$ with application to global synthetic weather generation. Solar Energy, **40**, 83-92.

Gregory, J.M., T.M.L. Wigley and P.D. Jones, 1993. Application of Markov models to area-average daily precipitation series and interannual variability in seasonal totals. Clim. Dyn. **8**, 299-310.

Hanssen, A.W. and W.J.A. Kuipers, 1965. On the Relationship between the Frequency of Rain and Various Meteorological Parameters. KNMI Mededelingen en Verhandelingen No. 81, Staatsdrukkerij- en uitgeverijbedrijf, 's-Gravenhage.

Hosmer, D.W. and S. Lemeshow, 1980. Goodness of fit tests for the multiple logistic regression model. Communications in Statistics, **A9**, 1043-1069.

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

Hughes, J.P., D.P. Lettenmaier and P. Guttorp, 1993. A stochastic approach for assessing the effect of changes in synoptic circulation patterns on gauge precipitation. Water Resour. Res., **29**, 3303-3315.

Hulme, M., K.R. Briffa, P.D. Jones and C.A. Senior, 1993. Validation of GCM control simulations using indices of daily airflow types over the British Isles. Clim. Dyn., **9**, 95-105.

Katz, R.W. and M.B. Parlange, 1993. Effects of an index of atmospheric circulation on stochastic properties of precipitation. Water Resour. Res., **29**, 2335-2344.

Klein Tank, A.M.G. and T.A. Buishand, 1993. Modelling daily precipitation as a function of temperature for climate change impact studies. KNMI Scientific Report WR 93-02, De Bilt.

Kruizinga, S., 1983. Statistical medium range forecasting of the probability of precipitation in the Netherlands. In: Preprints Eighth Conference on Probability and Statistics in Atmospheric Sciences, Hot Springs, pp 181-185. American Meteorological Society, Boston, Massachusetts.

Kwadijk, J.C.J., 1993. The impact of climate change on the discharge of the River Rhine. PhD thesis, University of Utrecht.

le Cessie, S. and J.C. van Houwelingen, 1991. A goodness-of-fit test for binary regression models based on smoothing methods. Biometrics, **47**, 1267-1282.

Lemeshow, S. and D.W. Hosmer, 1982. A review of goodness of fit statistics for use in the development of logistic regression models. American Journal of Epidemiology, **115**, 92-106.

Matyasovsky, I., I. Bogardi, A. Bárdossy, L. Duckstein, 1993. Space-time precipitation reflecting climate change. Hydrol. Sci. J., **38**, 539-558.

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.

Mimikou, M., Y. Kouvopoulos, G. Cavadias and N. Vayianos, 1991. Regional hydrological effects of climate change. J. Hydrol., **123**, 119-146.

Němec, J. and J. Schaake, 1982. Sensitivity of water resource systems to climate variation. Hydrol. Sci. J., **27**, 327-343.

Panagoulia, D., 1991. Hydrological response of a medium-sized mountainous catchment to climate changes. Hydrol. Sci. J., **36**, 525-547.

Panagoulia, D., 1992. Impacts of GISS modelled climate changes on catchment hydrology. Hydrol. Sci. J., **37**, 141-163.

Pearson, E.S. and H.O. Hartley, 1976. Biometrika Tables for Statisticians, Volume I, Biometrika Trust, University College, London.

Reed, D.N., 1986. Simulation of time series of temperature and precipitation over eastern England by an atmospheric general circulation model. J. Climatol., **6**, 233-253.

Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour. Res., **17**, 182-190.

Robock, A., R.P. Turco, M.A. Harwell, T.P. Ackerman, R. Andressen, H.-S. Chang and M.V.K. Sivakumar, 1993. Use of general circulation model output in the creation of climate change scenarios for impact analysis. Clim. Change, **23**, 293-335.

Semenov, M.A. and J.R. Porter, 1994. The implications and importance of non-linear responses in modelling the growth and development of wheat. In: Grasman, J. and G. van Straten (eds.), Predictability and Nonlinear Modelling in Natural Sciences and Economics, pp. 157-171, Kluwer, Dordrecht.

Slob, W.H., 1985. Modelling daily global radiation in a number of spectral bands using measurements at Uccle (Belgium). KNMI Technical Report TR 65 (in Dutch), De Bilt.

Thomas, G. and A. Henderson-Sellers, 1991. An evaluation of proposed representations of subgrid hydrologic processes in climate models. J. Climate, **4**, 898-910.

Toth, Z. and T. Szentimrey, 1990. The binormal distribution: a distribution for representing asymmetrical but normal-like weather elements. J. Climate, **3**, 128-136.

Valdés, J.B., R.S. Seoane and G.R. North, 1994. A methodology for the evaluation of global warming impact on soil moisture and runoff. J. Hydrol., **161**, 389-413.

von Storch, H., E. Zorita and U. Cubasch, 1993. Downscaling of global climate change estimates to regional scales: an application to Iberian Rainfall in Wintertime. J. Climate, **6**, 1161-1171.

Wilks, D.S., 1992. Adapting stochastic weather generation algorithms for climate change studies. Clim. Change, **22**, 67-84.

Wilson, L.L., D.P. Lettenmaier and E.F. Wood, 1992. Simulation of daily precipitation in the Pacific Northwest using a weather classification scheme. Surv. Geophys., **12**, 127-142.

WMO, 1964. Note on the standardization of pressure reduction methods in the international network of synoptic stations. WMO publication 154, World Meteorological Organization, Geneva, Switzerland.

Woodcock, F., 1976. The evaluation of yes/no forecasts for scientific and administrative purposes. Mon. Wea. Rev., **104**, 1209-1214.

Woolhiser, D.A., T.O. Keefer and K.T. Redmond, 1993. Southern oscillation effects on daily precipitation in the southwestern United States. Water Resour. Res., **29**, 1287-1295.

# Appendix A: Model validation

The occurrence of rain: testing for lack-of-fit

Pearson's $X^2$-statistic and the likelihood ratio test are generally used to assess the adequacy of the fit of the logistic regression model (McCullagh and Nelder, 1989). The latter is connected with parameter estimation by the method of maximum likelihood. A problem with both test-statistics is that their asymptotic chi-squared distribution under the null hypothesis no longer applies when the independent variables are continuous. For this situation Lemeshow and Hosmer (1982) proposed a modified $X^2$-statistic based on a grouping of the data according to the value of the estimated probabilities of rain.

The sequence of wet and dry days can be represented as $y_1, y_2, ..., y_J$, where $y_j = 1$ if day $j$ is wet and $y_j = 0$ if it is dry. The estimated probability $\hat{p}_j$ of rain on day $j$ according to the logistic model is given by equation (4). First $g$ groups are formed with cutpoints $1/g, 2/g, ..., (g-1)/g$ and day $j$ is assigned to group $k$ if $(k-1)/g < \hat{p}_j \leq k/g$, $k = 1, ..., g$. In Table A1 the cutpoints are defined at 0.1, 0.2, ..., 0.9 ($g = 10$). For each group the numbers of dry and wet days are counted and these numbers are compared with the expected frequencies according to the fitted model. The Hosmer-Lemeshow goodness-of-fit statistic $\hat{H}_g^*$ is then obtained by calculating Pearson's $X^2$-statistic for this 2x$g$ table.

From asymptotic theory it follows that the null distribution of $\hat{H}_g^*$ can be approximated by the chi-square distribution with $g$-2 degrees of freedom. This has been confirmed by extensive simulation (Hosmer and Lemeshow, 1980). Table A2 presents $\hat{H}_g^*$ and parameter estimates in the logistic regression model for each of the 10 subsets of the observed daily rainfall data. The mean value of the test-statistic $\hat{H}_g^*$ is close to its theoretical value 8 under the null hypothesis. The largest value (13.56) of $\hat{H}_g^*$ in Table A2 is just significant at the 10% level. For this subset Figure A1 shows the relation between the predicted and observed number of rain days in the 10 groups (reliability diagram). Even for this worst case there is no serious bias.

In Table A3 $\hat{H}_g^*$ is given for the cases discussed in Section 2, where either the status $y_{j-1}$ of the previous day is added to the model, or the threshold defining a rain day is 0.3 mm instead of 0.1 mm, or precipitation is sampled over the interval from 8 to 8 UT instead of 0 to 0 UT. It is seen that for the 0.3 mm threshold three values of $\hat{H}_g^*$ are above the critical value 15.51 for a test at the 5% level. From this it can be concluded that there is a systematic departure from the logistic model in equations (2) and (3) with that threshold. For precipitation in the 8-8 UT interval there is no indication of lack-of-fit.

Table A1 Observed and expected frequencies of wet and dry days in intervals based on fixed cutpoints of the estimated probability of rain (with the parameter estimates in Table 2 for the 1961-1990 record at De Bilt).

| | Estimated probability of rain ($\hat{p}_i$) | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | [0-0.1] | (0.1-0.2] | (0.2-0.3] | (0.3-0.4] | (0.4-0.5] | (0.5-0.6] | (0.6-0.7] | (0.7-0.8] | (0.8-0.9] | (0.9-1.0] | Total |
| Observed | | | | | | | | | | | |
| Wet days | 5 | 14 | 26 | 33 | 51 | 61 | 68 | 95 | 125 | 119 | 597 |
| Dry days | 65 | 93 | 72 | 65 | 56 | 44 | 42 | 32 | 23 | 6 | 498 |
| Total | 70 | 107 | 98 | 98 | 107 | 105 | 110 | 127 | 148 | 125 | 1095 |
| Expected | | | | | | | | | | | |
| Wet days | 5 | 16 | 24 | 34 | 48 | 58 | 72 | 95 | 126 | 118 | 596 |
| Dry days | 65 | 91 | 74 | 64 | 59 | 47 | 38 | 32 | 22 | 7 | 499 |
| Total | 70 | 107 | 98 | 98 | 107 | 105 | 110 | 127 | 148 | 125 | 1095 |

Table A2 Values of the statistic $\hat{H}_g^*$ and estimated coefficients in the logistic regression model for each of the 10 subsets in the daily rainfall record at De Bilt (1961-1990).

| Subset | $\hat{H}_g^*$ | $a$ | $b$ | $c$ | $d$ |
|---|---|---|---|---|---|
| 1 | 2.09 | 107.492 | -0.108 | 0.049 | -0.033 |
| 2 | 8.55 | 102.893 | -0.104 | 0.049 | -0.043 |
| 3 | 8.07 | 114.537 | -0.118 | 0.076 | -0.024 |
| 4 | 5.04 | 96.549 | -0.097 | 0.047 | -0.040 |
| 5 | 7.86 | 91.196 | -0.094 | 0.065 | -0.026 |
| 6 | 8.25 | 100.421 | -0.104 | 0.082 | -0.025 |
| 7 | 13.13 | 101.571 | -0.104 | 0.066 | -0.033 |
| 8 | 13.56 | 93.507 | -0.095 | 0.050 | -0.031 |
| 9 | 4.19 | 100.558 | -0.101 | 0.051 | -0.037 |
| 10 | 4.11 | 105.392 | -0.108 | 0.062 | -0.027 |
| Average | 7.49 | 101.412 | -0.103 | 0.060 | -0.032 |

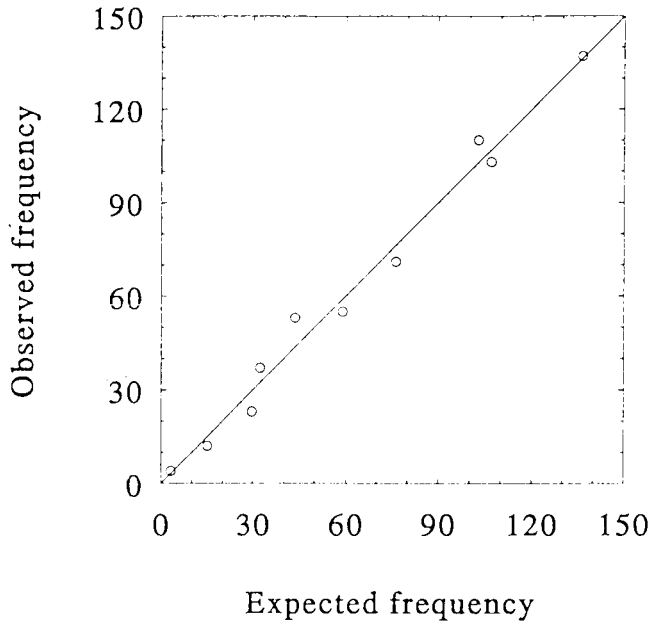



Figure A1 Reliability diagram for subset 8 in Table A2. The open dots refer to the numbers of wet days in the probability intervals used for the calculation of the $\hat{H}_g^*$-statistic.

Table A3 Values of the statistic $\hat{H}_g^*$ for each of the 10 subsets in the daily rainfall record at De Bilt (1961-1990) with either the status of the previous day added to the model ($y_{j-1}$), or with a precipitation threshold of 0.3 mm instead of 0.1 mm, or with precipitation amounts over the interval 8 to 8 UT instead of 0 to 0 UT.

| | $\hat{H}_g^*$ | | |
|---|---|---|---|
| Subset | $y_{j-1}$ | 0.3 mm | 8-8 UT |
| 1 | 16.04 | 8.12 | 16.96 |
| 2 | 5.50 | 9.09 | 9.34 |
| 3 | 5.67 | 18.09 | 10.07 |
| 4 | 4.35 | 7.97 | 7.19 |
| 5 | 6.22 | 15.12 | 1.39 |
| 6 | 8.47 | 10.82 | 5.45 |
| 7 | 5.07 | 17.98 | 3.77 |
| 8 | 9.84 | 16.13 | 10.51 |
| 9 | 6.84 | 5.74 | 11.65 |
| 10 | 4.46 | 4.36 | 2.85 |
| Mean | 7.25 | 11.34 | 7.92 |

More advanced goodness-of-fit tests using nonparametric kernel methods are given in Azzalini et al. (1989) and le Cessie and van Houwelingen (1991). These tests were not considered in the present study.

The occurrence of rain: performance indices

The logistic model provides for each day a probability $\hat{p}_j$ of rain. The model performs well if wet days ($y_j = 1$) are associated with high values of $\hat{p}_j$ and dry days ($y_j = 0$) with low values. Thus the score

$$B = \frac{1}{J} \sum_{j=1}^{J} [y_j - \hat{p}_j]^2 \qquad \text{(A1)}$$

should be small. In the meteorological literature $B$ is known as the Brier Score.

The observed long term mean probability or climatological probability of rain equals $\bar{y}$. This probability would have been obtained when the logistic model only contained the constant term $a$. For $\hat{p}_j = \bar{y}$, the score $B$ is equal to:

$$B_c = \frac{1}{J} \sum_{j=1}^{J} [y_j - \bar{y}]^2 = \bar{y}\,(1 - \bar{y}) \qquad \text{(A2)}$$

The Brier Relative Score (or Brier Skill Score) is defined as:

$$BRS = (1 - B/B_c) * 100 \qquad \text{(A3)}$$

For a perfect predictor ($B = 0$), $BRS = 100$, whereas $BRS = 0$ says that the prediction rule is not better than the use of $\bar{y}$.

For the 10 subsets in Table A2 the mean value of $BRS$ equals 32. This is slightly lower than the values reported in Kruizinga (1983; Figure 3) for an application of logistic regression to the output of the ECMWF numerical weather model (ECMWF = European Centre for Medium Range Weather Forecasting, Reading UK).

Performance indices can also be derived from the classification table in Section 2. A good one is the Hanssen and Kuipers' discriminant $V$ (Woodcock, 1976). $V = 0$ if days are randomly taken to be wet with probability $\bar{y}$, whereas for a perfect predictor $V = 1$. For Table 4 the value of $V$ is 0.49. In the original work of Hanssen and Kuipers (1965) about the same performance was achieved by using only the height of the 500 hPa level as a predictor ($V = 0.49$, p.33) and by a prediction rule based on surface air pressure, wind direction and the occurrence of precipitation on the previous day ($V = 0.50$, p. 45). For climate change scenarios the use of some of these predictors is questionable, however. It is for instance very likely that in a doubled $CO_2$ climate the

relation between precipitation occurrence and the height of the 500 hPa level will differ from that in the present-day climate. The higher temperatures in the $2\times CO_2$ climate lead to an increase in the height of this pressure level, but not necessarily to a change in the number of rain days (Section 7). When wind direction is incorporated in the model for precipitation occurrence, then the dependence of relative humidity and incoming solar radiation on wind direction must also be taken into account. Inclusion of predictors which are not required for the intended impact study should be avoided as much as possible.

Further checks on the model for daily precipitation amounts

For a further check on the 12 parameter model in Section 3 for the relation between the precipitation amount on rain days, daily mean temperature and daily mean surface air pressure, the parameters $a$, $b$, $c$ and $d$ in equation (8) were estimated for distinct pressure intervals. In Table A4 and Figure A2 six different intervals of surface air pressure are considered. Some notable features are the maximum of the parameter $b$ near $P = 1010$ hPa and the almost monotone change of the parameters $a$, $c$ and $d$ with $P$. The values of $a$, $b$, $c$ and $d$, derived from the 12 parameter model by substituting the mean of $P$ at each interval, are in good agreement with the direct estimates. Model values never differ more then 2.1 times the standard error se from the direct estimates.

In a similar way the parameters $a_0$, $a_1$, $a_2$ and $a_3$ in equation (10) were estimated for five distinct temperature intervals (Table A5 and Figure A3). Also for these parameters the values derived from the 12 parameter model are in good agreement with the direct estimates. Model values never differ more then 2.3 times the standard error se from the direct estimates. The statistical significance of a cubic term in $P$ is confirmed by the significant values of the estimate of $a_3$ in three out of the five intervals.

Table A4 Estimates of the coefficients $a$, $b$, $c$ and $d$ in the precipitation - temperature relation, equation (8), with their standard error (se) for various values of surface air pressure $P$ for rain days at De Bilt (1906-1981). Estimates are compared to the theoretical values according to the 12 parameter model in equation (15). SDIF gives the standardized difference between estimate and model value.

| | $a$ | | | | $b$ | | | |
|---|---|---|---|---|---|---|---|---|
| $P$ (hPa) | Estimate | se | Model | SDIF | Estimate | se | Model | SDIF |
| 990.6 | 1.572 | 0.056 | 1.538 | 0.6 | 0.049 | 0.010 | 0.051 | -0.2 |
| 1003.2 | 1.161 | 0.041 | 1.163 | 0.0 | 0.069 | 0.007 | 0.069 | 0.0 |
| 1011.1 | 0.768 | 0.050 | 0.791 | -0.5 | 0.076 | 0.008 | 0.077 | -0.1 |
| 1016.8 | 0.563 | 0.062 | 0.452 | 1.8 | 0.065 | 0.010 | 0.074 | -0.9 |
| 1022.6 | 0.095 | 0.074 | 0.048 | 0.6 | 0.054 | 0.013 | 0.061 | -0.5 |
| 1030.4 | -0.641 | 0.072 | -0.509 | -1.8 | 0.055 | 0.015 | 0.039 | 1.1 |

| | $c$ | | | | $d$ | | | |
|---|---|---|---|---|---|---|---|---|
| $P$ (hPa) | Estimate | se | Model | SDIF | Estimate | se | Model | SDIF |
| 990.6 | 0.001 | 0.009 | 0.010 | -1.0 | -0.8E-4 | 9.1E-4 | -1.2E-3 | 1.2 |
| 1003.2 | -0.003 | 0.003 | -0.002 | -0.3 | -3.0E-4 | 3.0E-4 | -2.0E-4 | -0.3 |
| 1011.1 | -0.009 | 0.003 | -0.010 | 1.7 | 4.0E-4 | 1.6E-4 | 4.0E-4 | 0.0 |
| 1016.8 | -0.013 | 0.003 | -0.015 | 0.7 | 7.0E-4 | 2.0E-4 | 9.0E-4 | -1.0 |
| 1022.6 | -0.023 | 0.006 | -0.021 | -0.3 | 17.0E-4 | 4.3E-4 | 13.0E-4 | 0.9 |
| 1030.4 | -0.006 | 0.012 | -0.028 | 1.8 | -4.0E-4 | 10.7E-4 | 19.0E-4 | -2.1 |

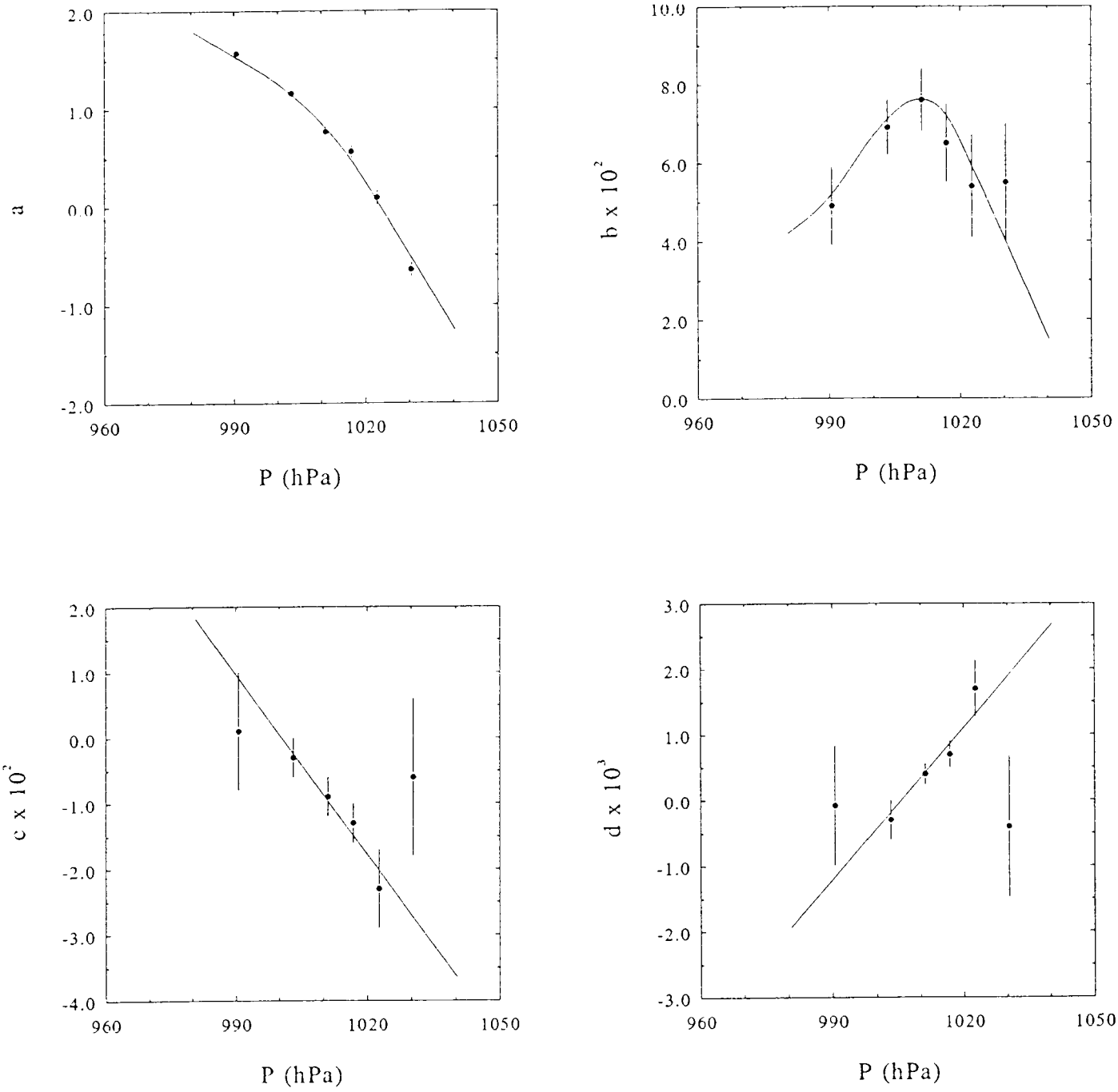



Figure A2 The functional dependence of the parameters *a*, *b*, *c* and *d* in the precipitation - temperature relation, equation (8), on surface air pressure for rain days at De Bilt (1906-1981). This relation was fitted to the precipitation amounts in six different intervals of surface air pressure. For each interval the estimated parameters are plotted at the mean value of *P*. The error bars indicate the standard errors of these estimates. The solid lines represent the theoretical values according to the 12 parameter model.

Table A5 Estimates of the coefficients $a_0$, $a_1$, $a_2$ and $a_3$ in the precipitation - pressure relation, equation (10), with their standard error (se) for various values of temperature $T$ for rain days at De Bilt (1906-1981). Estimates are compared to the theoretical values according to the 12 parameter model in equation (15). SDIF gives the standardized difference between estimate and model value.

| | $a_0$ | | | | $a_1$ | | | |
|---|---|---|---|---|---|---|---|---|
| $T$ (°C) | Estimate | se | Model | SDIF | Estimate | se | Model | SDIF |
| 0.4 | 0.279 | 0.036 | 0.259 | 0.6 | 0.078 | 0.005 | 0.073 | 1.0 |
| 5.2 | 0.639 | 0.024 | 0.584 | 2.3 | 0.087 | 0.004 | 0.086 | 0.3 |
| 8.9 | 0.718 | 0.029 | 0.779 | -2.1 | 0.101 | 0.006 | 0.099 | 0.3 |
| 13.1 | 0.703 | 0.032 | 0.695 | 0.3 | 0.144 | 0.009 | 0.125 | 2.1 |
| 17.1 | 0.765 | 0.059 | 0.680 | 1.4 | 0.117 | 0.021 | 0.135 | -0.9 |

| | $a_2$ | | | | $a_3$ | | | |
|---|---|---|---|---|---|---|---|---|
| $T$ (°C) | Estimate | se | Model | SDIF | Estimate | se | Model | SDIF |
| 0.4 | -1.6E-3 | 0.6E-3 | -1.1E-3 | -0.8 | 1.8E-5 | 1.5E-5 | 0.4E-5 | 0.9 |
| 5.2 | -2.4E-3 | 0.4E-3 | -2.2E-3 | -0.5 | 2.7E-5 | 0.8E-5 | 2.2E-5 | 0.6 |
| 8.9 | -3.1E-3 | 0.6E-3 | -3.0E-3 | -0.2 | 3.9E-5 | 1.4E-5 | 3.6E-5 | 0.2 |
| 13.1 | -5.9E-3 | 1.1E-3 | -3.9E-3 | -1.8 | 9.7E-5 | 3.2E-5 | 5.2E-5 | 1.4 |
| 17.1 | -4.2E-3 | 2.8E-3 | -4.8E-3 | 0.2 | 8.2E-5 | 10.0E-5 | 6.6E-5 | 0.2 |

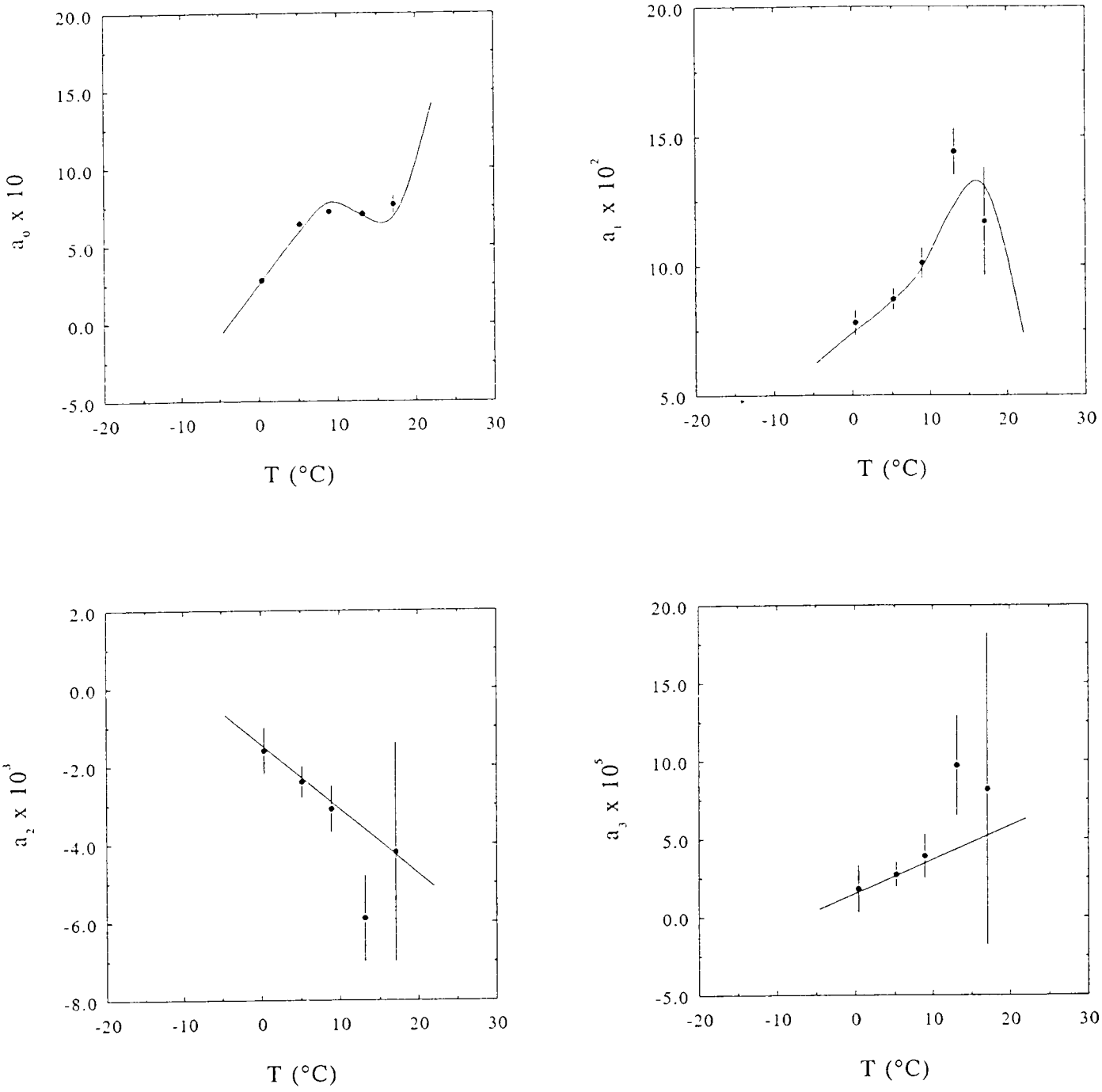



Figure A3 The functional dependence of the parameters $a_0$, $a_1$, $a_2$, $a_3$ in the precipitation - pressure relation, equation (10), on temperature for rain days at De Bilt (1906-1981). This relation was fitted to the precipitation amounts in five different temperature intervals. For each interval the estimated parameters are plotted at the mean value of $T$. The error bars indicate the standard errors of these estimates. The solid lines represent the theoretical values according to the 12 parameter model.

# Appendix B: Humidity and surface air pressure in the CCC-GCM

The CCC-GCM does not provide the relative humidity $U$ at screen level, but the specific humidity $q$. The latter was interpolated from the lowest prognostic level located approximately 200 meter above the surface. No checks were made on the internal consistency of the various sceen level quantities. In this study the daily values of $U$ were derived from the daily specific humidity $q$, surface air pressure $P$ and surface air temperature $T$, which were available as 12-hourly sampled data ($P$) or as 12-hourly means ($q$,$T$). From the mean of the two 12-hourly values the daily saturated vapour pressure $e_s$ was obtained using the Clausius Clapeyron relation (Magnus approximation):

$$e_s = 6.107 \exp[17.57\, T/(241.8+T)] \tag{B1}$$

and the vapour pressure $e$ was computed as:

$$e = qP/0.622 \tag{B2}$$

The daily relative humidity $U$ is given by:

$$U = 100\, e/e_s \tag{B3}$$

The daily relative humidity at De Bilt is, however, a mean value of the relative humidity at the 24 clock hours. The two daily values of $U$ differ because of the diurnal cycle of the temperature and the non-linearity of equations (B1) and (B3). To quantify the systematic difference the daily mean relative humidity at De Bilt was compared with that obtained from equations (B1) and (B3) using the daily means of $T$ and $e$. The latter is on average smaller than the values of $\bar{U}$ in Table 13 for De Bilt. The largest discrepancy 1% is found for summer. This discrepancy is much smaller and of opposite sign than those found between the model simulations and the values for De Bilt in Table 13. The abnormal mean values of the relative humidity for the model simulations should mainly be ascribed to deficiencies of the interpolation procedure to obtain the specific humidity at screen level.

Surface air pressure given by the CCC-GCM was reduced to mean sea level (msl) pressure according to WMO standards (WMO, 1964). A temperature correction was taken into account. The effect of air humidity could not be considered. The latter is, however, only important in (sub)tropical regions. The transformation formula reads:

$$P_{msl} = P_{surface} \times 10^{0.0148275H\ /\ (T+273.16+0.00325H)} \tag{B4}$$

with $H$ the surface height above mean sea level. The prescribed surface height of the two gridpoints in the small box of Figure 11 is 249 m and 351 m, respectively. The 16 land gridpoints in the large box vary in height between -9 m in the northwest and 593 m in the southeast.

# Appendix C: Simulation of precipitation amounts

The regression model provides a value for the mean precipitation amount $\mu(T^*,P^*) = \exp[k(T^*,P^*)]$ for each new wet day in the perturbed climate. The aim of simulation is to add random variation to this estimate in order to reproduce the variability of the daily precipitation amounts. Reasonable results are already obtained by generating from the one parameter exponential distribution. The density of the exponential variable $X$ is given by:

$$f(x) = \lambda e^{-\lambda x} \qquad x>0 \qquad \text{(C1)}$$

The parameter $\lambda$ determines the mean and the variance of $X$:

$$E(X) = 1/\lambda \quad ; \quad \text{var } X = 1/\lambda^2 \qquad \text{(C2)}$$

The variable $X$ takes values on the interval $(0,\infty)$. The precipitation amount $R^*$ has, however, a positive lower bound $\delta = 0.05$ mm because days with smaller amounts are recorded as dry days. Therefore $R^*$ should be generated as $R^* = X + \delta$. From $E(R^*) = E(X) + \delta$ it follows:

$$1/\lambda = \mu(T^*,P^*) - \delta \qquad \text{(C3)}$$

The variable $X$ can be written as $X = \tilde{X} / \lambda$ with $\tilde{X}$ the standard exponential variable (exponential with $\lambda = 1$). $\tilde{X}$ can be generated as $\tilde{X} = -\ln V$, where $V$ is a uniform (0,1)-variable. The precipitation amount $R^*$ is then:

$$R^* = (-\ln V)/\lambda + \delta \qquad \text{(C4)}$$

where $1/\lambda$ is given by equation (C3).

The coefficient of variation of $R^*$ equals $1/(1 + \delta\lambda) \approx 1$. This is in reasonable agreement with the values from equation (14). The value of $v_R$ can be exactly preserved by generating from the gamma distribution or another two parameter distribution on the interval $(0,\infty)$. The gamma distribution is a two parameter generalization of the exponential distribution and has the advantage that its parameters can simply be expressed in the mean and the variance of $X$. The generation of gamma variables requires a somewhat more advanced Monte Carlo procedure. A good survey of techniques can be found in Duvroye (1986).

# Appendix D: Users guide

Consider the situation that, in addition to a temperature effect on the precipitation amount, there is a change in the number of rain days as a result of a change in the pressure distribution (atmospheric circulation). The procedure outlined in this report to obtain a representative daily precipitation time series for that situation consists of the following 9 steps:

1. Prepare a file with time series of daily values of mean temperature $T$, precipitation amount $R$, mean surface air pressure $P$, relative humidity $U$ and relative incoming solar radiation $Q$. Records of several decades are necessary to assess the impacts of changes in the severity of droughts and other extreme events.

2. Determine the seasonal or monthly mean changes in $T$ and $P$, elements that are used as covariates in the models, and apply them to each individual day, giving $T'$ and $P'$. The changes may be derived from GCM experiments with enhanced $CO_2$ air concentration. In the present report $U$ and $Q$ were left unchanged, so that $U' = U$ and $Q' = Q$.

3. Define a wet day as a day with a recorded precipitation amount $\geq 0.1$ mm. The remaining days in the observed record are dry.

4. Compute for each day $j$ the probability of rain (equation (4)) as:

$$\hat{p}_j = \frac{1}{1+\exp\left[-(\hat{a}+\hat{b}P_j+\hat{c}U_j+\hat{d}Q_j)\right]} \qquad j=1,\ldots,K \qquad \text{(D1)}$$

and:

$$\hat{p}_j^* = \frac{1}{1+\exp\left[-(\hat{a}+\hat{b}P_j^*+\hat{c}U_j^*+\hat{d}Q_j^*)\right]} \qquad j=1,\ldots,K \qquad \text{(D2)}$$

for the actual climate and the perturbed climate, respectively, with $K$ the total number of days in the period considered. The estimated coefficients $\hat{a}$, $\hat{b}$, $\hat{c}$ and $\hat{d}$ are given in Table 3. For precipitation sampled at 8 UT over the past 24 hours instead of the 0-0 UT interval, the estimates in Table 6 should be used.

5. Compute for each of the four seasons or twelve months the change $\Delta m$ in the total number of wet days due to the perturbation in $P$ for that season or month. According to equation (22):

$$\Delta m = \sum_{j=1}^{K} \hat{p}_j^* - \sum_{j=1}^{K} \hat{p}_j \qquad \text{(D3)}$$

6. In the situation that the number of wet days increases ($\Delta m > 0$) transform the $\Delta m$ dry days with the largest values of $\hat{p}_j^*$ into wet days, where $\Delta m$ should be rounded to the nearest integer. As an alternative a binary variable $Y_j$ can be drawn for each dry day giving the status of that day in the perturbed climate (1 = wet and 0 = dry). In order that the expected number of additional wet days equals $\Delta m$ the probability $\Pi_j$ that $Y_j = 1$ should then be equal to:

$$\Pi_j = \Delta m \; \hat{p}_j^* \; / \sum_{j=1}^{L} \hat{p}_j^* \qquad j=1,\ldots,L \qquad \text{(D4)}$$

with $L$ the total number of dry days in the period considered. For the situation that the number of wet days decreases ($\Delta m < 0$) transform the $\Delta m$ wet days with the smallest values of $\hat{p}_j^*$ into dry days in a similar way.

7. Generate the amounts on the new wet days in the record from the exponential distribution (Appendix B). Set the amounts on the selected new dry days to zero.

8. Transform the amounts on days that remain wet in the scenario with the multiplying factor derived from the model for the relation between precipitation amount, temperature and surface air pressure:

$$F(T,T^*,P,P^*) = \exp\,[k(T^*,P^*) - k(T,P)] \qquad \text{(D5)}$$

The estimated coefficients in the model are given in Table 9 (7 parameter model) or Table 10 (12 parameter model). For precipitation sampled at 8 UT over the past 24 hours instead of the 0-0 UT interval, the estimates in Table 11 (7 parameter model) or Table 12 (12 parameter model) should be used.

9. Save the new daily time series of $R^*$, $T^*$, $P^*$, $U^*$ and $Q^*$ for use in impact studies.

Bijgaand treft U een overzicht aan van recentelijk in deze serie gepubliceerde titels. Een komplete lijst wordt U op verzoek toegezonden. U kunt Uw aanvraag richten aan de KNMI Bibliotheek, Postbus 201, 3730 AE De Bilt (030-206855). Hier kan men U tevens informeren over de verkrijgbaarheid en prijzen van deze publicaties.

89-05 Statistical forecasts of sunshine duration / Li Zhihong and S. Kruizinga.

90-01 The effect of a doubling atmospheric CO2 on the stormtracks in the climate of a general circulation model / P.C. Siegmund.

90-02 Analysis of regional differences of forecasts with the multilayer AMT-model in the Netherlands / E.I.F. de Bruin, Li Tao Guang and Gao Kang.

90-03 Description of the CRAU data-set : Meteosat data, radiosonde data, sea surface temperatures : comparison of Meteosat and Heimann data / S.H. Muller, H. The, W. Kohsiek and W.A.A. Monna.

90-04 A guide to the NEDWAM wave model / G. Burgers.

91-01 A parametrization of the convective atmospheric boundary layer and its application into a global climate model / A.A.M. Holtslag, B.A. Boville and C.-H. Moeng.

91-02 Turbulent exchange coefficients over a Douglas fir forest / F.C. Bosveld.

92-01 Experimental evaluation of an arrival time difference lightning positioning system / H.R.A.Wessels.

92-02 GCM control run of UK Meteorological Office compared with the real climate in the NW European winter / J.J. Beersma.

92-03 The parameterization of vertical turbulent mixing processes in a General Circulation Model of the Tropical Pacific / G. Janssen.

92-04 A scintillation experiment over a forest / W.Kohsiek.

92-05 Grondtemperaturen / P.C.T.van der Hoeven en W.N.Lablans

92-06 Automatic supression of anomalous propagation clutter for noncoherent weather radars / H.R.A. Wessels and J.H. Beekhuis.

93-01 Searching for stationary stable solutions of Euler's equation / R. Salden.

93-02 Modelling daily precipitation as a function of temperature for climatic change impact studies / A.M.G. Klein Tank and T.A. Buishand.

93-03 An analytic conceptual model of extratropical cyclones / L.C. Heijboer.

93-04 A synoptic climatology of convective weather in the Netherlands / Dong Hongnian.

93-05 Conceptual models of severe convective weather in the Netherlands / Dong Hongnian.

94-01 Seismische analyse van aardbevingen in Noord-Nederland : bijdrage aan het multidisciplinaire onderzoek naar de relatie tussen gaswinning en aardbevingen / H.W.Haak en T. de Crook.

94-02 Storm activity over the North Sea and the Netherlands in two climate models compared with observations / J.J. Beersema.

94-03 Atmospheric effects of high-flying subsonic aircraft / W. Fransen.

94-04 Cloud-radiation-hydrological interactions : measuring and modeling / A.J. Feijt, R. van Dorland, A.C.A.P. van Lammeren, E. van Meijgaard en P. Stammes.

94-05 Spectral ultraviolet radiation measurements and correlation with atmospheric parameters / F. Kuik and H. Kelder

95-01 Transformation of precipitation time series for climate change impact studies / A.M.G. Klein Tank and T.A. Buishand

95-02 Internal variability of the ocean generated by a stochastic forcing / M.B.H. van Noordenburg

95-03 Applicability of weakly nonlinear theory for the planetary-scale flow / E.A. Kartashova

95-04 Changes in tropospheric NOx and O3 due to subsonic aircraft emissions / W.M.F. Wauben, P.F.J. van Velthoven and H. Kelder