Technical report; TR-230

KNMI

KNMI contribution to the European project WRINCLE: downscaling relationships for precipitation for several European sites

B.R. Beckmann and T.A. Buishand

De Bilt, 2001

Koninklijk Nederlands Meteorologisch Instituut

Technical Report = Technisch Rapport; TR-230

De Bilt, 2001

PO Box 201
3730 AE De Bilt
Wilhelminalaan 10
http://www.knmi.nl
Telephone +31 30 220 69 11
Telefax +31 30 221 04 07

Author: B.R. Beckmann and T.A. Buishand

UDC: 551.577.35
551.513.1
551.583.1
(4)

ISSN: 0169-1708

ISBN: 90-369-2190-2

© KNMI, De Bilt. All rights reserved. No part of this publication may be reproduced, stored in retrival systems, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without prior permission in writing from the publisher.

KNMI Contribution to the European Project WRINCLE:

Downscaling Relationships for Precipitation for Several European Sites

B.-R. Beckmann

T. A. Buishand

Technical Report TR-230

Summary

Statistical downscaling of precipitation refers to statistical techniques that have been used to obtain precipitation data with the required spatial resolution for climate-change impact studies. It is widely acknowledged that the direct precipitation output of climate change simulations from General Circulation Models (GCMs) is inadequate for such studies. Statistical downscaling techniques make use of fitted relationships between observed precipitation and other meteorological variables that can be extracted from GCM simulations. The recent availability of reanalysis data from numerical weather prediction models offers new oppurtunities to improve the meteorological basis of statistical downscaling models. In this report the statistical linkage of daily precipitation to NCEP reanalysis data is described for eleven stations across Europe: De Bilt and Maastricht (the Netherlands), Hamburg, Hanover and Berlin (Germany), Vienna (Austria), Berne, Neuchâtel and Payerne (Switzerland), and Salto de Bolarque and Munera (Spain). Daily data for the period 1968-1997 were considered. The work forms the KNMI contribution to the European project WRINCLE (Water Resources: the INfluence of CLimate change in Europe).

Two separate statistical models were used to describe the daily precipitation at a particular site: an additive logistic model for rainfall occurrence (1 for wet days and 0 for dry days) and a generalised additive model for wet-day rainfall. Both models are extensions of the standard linear regression model for data from a normal distribution. Typical predictor variables were the u-velocity (westerly flow), the v-velocity (southerly flow), the relative vorticity, the 1000-500 hPa thickness, the baroclinicity and atmospheric moisture. With respect to the latter two options for rainfall amount modelling were compared: (i) the use of the specific humidity at 700 hPa, and (ii) the use of both the relative humidity at 700 hPa and precipitable water. For rainfall occurrence modelling the relative humidity at 700 hPa was considered as moisture variable. The 1000-500 hPa thickness was only included in the model for rainfall occurrence.

For all stations the moisture variable appears to be one of the most significant predictors both for rainfall occurrence and rainfall amount. The u-velocity is an important predictor for the northern stations De Bilt, Maastricht, Hamburg and Hanover. Although precipitation usually has a strong link to relative vorticity, this is not the case for Vienna and the Swiss stations. The baroclinicity is an important predictor for these middle European stations. For the Spanish sites it was beneficial to derive the velocity components and the relative vorticity from the 850 hPa heights instead of the 1000 hPa heights or sea level pressure as was done for the other sites. Although the downscaling relationships for rainfall occurrence and wet-day rainfall are assumed constant over the year, they can reasonably reproduce the seasonal cycles in the probability of rain and mean wet-day rainfall.

An application is given with data from a time-dependent greenhouse gas forcing experiment using the coupled ECHAM4/OPYC3 atmosphere-ocean GCM for the periods 1968 - 1997 and 2070 - 2099. The fitted statistical relationships were used to estimate the changes in the mean number of wet days and mean rainfall amounts for the winter and summer halves of the year at De Bilt, Hanover, Berlin, Berne and Salto de Bolarque.

For all these stations a decrease in the mean number of wet days was found. This decrease is mainly due to the larger 1000-500 hPa thickness in the future climate. At Berlin this thickness effect is somewhat reduced by the changes in the velocity components. A marked influence of the velocity components was also observed at Salto de Bolarque, but here the effect of the change in the u-velocity is counterbalanced by that in the v-velocity. The largest decrease in the number of wet days occurs in the summer period. The mean wet-day precipitation amounts are sensitive to the larger atmospheric water content in the future climate (increase in specific humidity and precipitable water). This leads to an increase in the mean winter rainfall amounts at De Bilt, Berlin and Berne despite the decrease in the mean number of wet days. The estimated increase in mean winter rainfall for these stations is comparable with that in the simulated rainfall of the ECHAM4/OPYC3 model. The use of the rainfall amount model with specific humidity results in a larger increase than the model with precipitable water and relative humidity. The estimated change in mean winter rainfall at Hanover differs from that at De Bilt and Berlin, and from the direct climate model output. This can be attributed to an anomalous decrease in the relative vorticity near Hanover in the ECHAM4/OPYC3 experiment and a rather strong impact of vorticity on daily rainfall at this site. Consistent with the simulated rainfall of ECHAM4/OPYC3 a decrease in the mean winter rainfall amouts at Salto de Bolarque was found. The use of a rainfall amount model with specific humidity results in a slight increase in the mean summer rainfall amounts at De Bilt, Hanover and Berlin. In the other cases there is a decrease in mean summer rainfall, quite often comparable with that from the direct climate model output. Except for the changes in winter rainfall at Hanover and the number of wet days in Berlin and Salto de Bolarque, the changes in the circulation variables have little effect on the mean number of wet days and the mean rainfall amounts. Moreover, the changes in circulation variables are generally small compared with their bias. The anomalous change in vorticity during winter at Hanover is in fact mainly due to biases in the simulated sea level pressure.

The use of the downscaling models for scenario production was examined for Berne. Scenarios of daily precipitation for the 2070 - 2099 period were obtained by perturbation of the observed rainfall record and by stochastic time series simulation. In the case of time series perturbation the standard method of scaling the observed rainfall amounts was extended to allow for a decrease in the number of wet days. The 90^{th} percentile of the distributions of the N-day annual maximum precipitation amounts for $N = 1$, 3, 10 and 30 was considered to compare the resulting scenarios. The representation of the coefficient of variation of the wet-day precipitation amounts in the stochastic model for time series simulation strongly influences the reproduction of this extreme-value characteristic and its changes. Besides difficulties with the description of the coefficient of variation, the model for wet-day rainfall often overpredicts the mean rainfall amounts in situations where extreme rainfall could be expected. Interaction between predictor variables has to be incorporated to reduce this bias.

Contents

1 Introduction

A major problem associated with future global warming is the potential change in precipitation because of its impacts on hydrology and water resources. Simulations with General Circulation Models (GCMs) form the primary source of information about future climate change. These models provide time series of many climate variables on a rather coarse grid (current grid size about 300 km), which is not suitable for direct use in climate-change impact studies. A so-called downscaling technique is needed to obtain these variables at the desired local scale. The term statistical downscaling refers to the use of a statistical model for this purpose.

Statistical downscaling of precipitation has often been focussed on the link between local precipitation and atmospheric flow characteristics. Several techniques have been used ranging from multiple regression (KILSBY *et al.*, 1998; WILBY *et al.*, 1998) and canonical correlation analysis (VON STORCH *et al.*, 1993) to multivariate adaptive regression splines (CORTE-REAL *et al.*, 1995), kriging (BIAU *et al.*, 1999) and artificial neural networks (HEWITSON and CRANE, 1996), and from parametric time series modelling (Bárdossy and PLATE, 1992; CORTE-REAL *et al.*, 1999) to nonparametric resampling (ZORITA *et al.*, 1995; CONWAY and JONES, 1998). Although long-term variations in precipitation during the past may be explained by fluctuations in the atmospheric circulation, it is questionable whether potential future systematic changes in precipitation resulting from the global warming can be derived from the changes in the atmospheric circulation alone. Various GCM experiments suggest that this is not the case (MATYASOVSZKY *et al.*, 1993; JONES *et al.*, 1997; WILBY and WIGLEY, 1997).

Besides the atmospheric circulation, there is also an effect of temperature on precipitation, because warm air can contain more moisture than cold air. Shower activity also depends on temperature. For De Bilt (the Netherlands) these phenomena can easily be identified by plotting the mean wet-day precipitation amounts versus the daily mean or daily maximum 2-metre temperature (KÖNNEN, 1983; BUISHAND and KLEIN TANK, 1996). A similar analysis for a number of sites in England, Switzerland, Portugal and Italy showed, however, that the temperature effect on precipitation is often obscured by other meteorological factors (BRANDSMA and BUISHAND, 1996). Because the correlation between daily precipitation and temperature is generally weak in observed data, it is quite difficult to use temperature as a covariate in stochastic models for daily precipitation.

With the exception of a paper by KARL *et al.* (1990), a measure of the upper-air humidity has only recently been included in statistical downscaling models for daily precipitation. The correlation between precipitation and atmospheric moisture is generally stronger than that between precipitation and temperature. CRANE and HEWITSON (1998) and CAVAZOS (1999) considered the specific humidity at different heights as a measure of the atmospheric water content. On the other hand, CHARLES *et al.* (1999) used the 850 hPa dew point temperature depression in their model for daily rainfall occurrence, because the probability of rain is better related to the degree of saturation of the atmosphere than to the total amount of water.

In this report new downscaling relationships are developed for a number of sites in the Netherlands, Germany, Austria, Switzerland and Spain. The work was done under the framework of the European project WRINCLE (Water Resources: the INfluence of CLimate change in Europe). Separate models are used to describe the probability of rain and the wet-day rainfall amounts. Several circulation variables are included in these models. The probability of rain is further linked to the relative humidity at 700 hPa. Two measures of absolute humidity are compared for rainfall amount modelling. All predictor variables were derived from the NCEP reanalysis data (KALNAY *et al.*, 1996). Results from a time-dependent greenhouse gas forcing experiment with the ECHAM4/OPYC3 climate model are used to estimate the changes in precipitation for the period 2070-2099 for a selection of stations. These estimates are compared with the changes in precipitation in the climate model output. In addition, daily precipitation scenarios are constructed for one of the Swiss sites. Extreme N-day rainfall is considered to compare the scenarios with the record of observed daily rainfall.

This report is structured as follows. Section 2 introduces the statistical models for the probability of rain and the wet-day rainfall amounts. Section 3 presents a survey of the precipitation records and reanalysis data. Some properties of the climate model output are also described in that section. The results of the statistical analysis are discussed in Section 4. Section 5 deals with the estimation of the changes in the seasonal mean precipitation amounts for the period 2070 - 2099. The construction of daily precipitation scenarios is presented in Section 6. The statistical model for the wet-day rainfall amounts is re-examined in Section 7. A concluding discussion is given in Section 8.

2 Statistical models

Generalised additive models have been used to explore the relationship between daily precipitation and the predictor variables. These models are an extension of the standard linear model for data from a normal distribution, covering both non-linearity and a variety of distributions (HASTIE and TIBSHIRANI, 1990). They were successfully applied to describe the dependence of the wet-day precipitation amounts on 2-metre temperature and circulation variables in earlier work (BRANDSMA and BUISHAND, 1997). Here generalised additive models are also used to describe rainfall occurrence.

2.1 Rainfall occurrence modelling

The occurrence of wet and dry days can be represented as a binary sequence $\{y_t\}$, where $y_t = 1$ if day t is wet and $y_t = 0$ if day t is dry. A wet day is defined here as a day with a precipitation amount of 0.1 mm or more. The key parameter in rainfall occurrence modelling is the probability P of a day being wet. In this study P is described by the additive logistic model:

$$\text{logit}(P) = \ln\left(\frac{P}{1-P}\right) = a_0 + \sum_{i=1}^{p} f_i(x_i) \tag{1}$$

where the $f_i(x_i)$ are arbitrary smooth functions of the predictor variables x_i. The logit transformation ensures that P lies in the interval between 0 and 1. The functions $f_i(x_i)$ have been iteratively estimated by a locally weighted running-line (loess) smoother with a span of 0.5 (proportion of the data entered in the local fit), using the S-PLUS statistical software package (CHAMBERS and HASTIE, 1993). To exclude free constants the fitted values for each function $f_i(x_i)$ are adjusted to average zero. In the case of a linear function, this implies that $f_i(x_i) = a_i(x_i - \bar{x}_i)$, where a_i is a regression coefficient and $\bar{x}_i$ is the average of the variable x_i.

As an illustration Figure 1 shows the shapes of the two most significant predictor functions $f_i(x_i)$ for rainfall occurrence at Berne (Switzerland). In this example the function $f_i(x_i)$ is linear for the relative humidity at 700 hPa (rh) and non-linear for the southerly flow component at 1000 hPa (v_{1000}). The rainfall probability on the right-hand side of each panel is given by:

$$P = \frac{1}{1 + \exp\left[-a_0 - f_i(x_i)\right]} \tag{2}$$

As expected P increases with increasing rh. Furthermore, P tends to be high if v_{1000} is negative (northerly flow) and low if v_{1000} is positive (southerly flow).

Model selection involves the choice of atmospheric variables to be included and the choice between linear and non-linear predictor functions. Atmospheric variables having a relatively strong correlation with rainfall occurrence were considered. Inclusion of strongly correlated predictor variables was avoided. Different sets of predictor variables were compared using Akaike's information criterion (see the Appendix). This criterion puts a

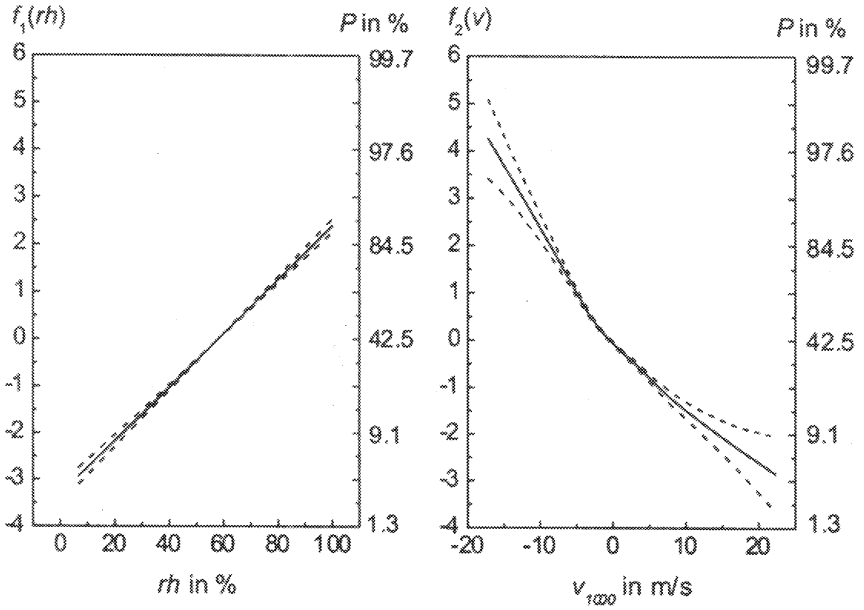

Figure 1: Estimates of the functions $f_i(x_i)$ for relative humidity at 700 hPa (rh) and southerly flow at 1000 hPa (v_{1000}) in the additive logistic model for rainfall occurrence at Berne (Switzerland). The dashed lines mark pointwise 2 x standard-error bands.

penalty on the number of predictor variables and the use of non-linear functions. The significance of every predictor variable and the non-linearity of predictor functions in the selected set were checked by an approximate χ^2-test (see the Appendix). In a final stage the physical realism of the shape of the predictor function was examined. A predictor variable was excluded if its effect on rainfall occurrence was at variance with that expected on physical grounds.

2.2 Rainfall amount modelling

The generalised additive model with constant coefficient of variation and log link function has been adopted to analyse the wet-day precipitation amounts. The log link function specifies the expected wet-day amount as:

$$R = \exp\left[a_0 + \sum_{i=1}^{p} f_i\left(x_i\right)\right] \tag{3}$$

Besides that R cannot become negative with this representation, the log link is also convenient for scenario construction (see Section 6.1). A constant coefficient of variation implies that the variance of the wet-day precipitation amounts increases with the expected value, which is for wet-day precipitation much more realistic than the more familiar assumption of a constant variance. As in Equation (1) the $f_i(x_i)$ are arbitrary smooth functions of the predictor variable x_i.

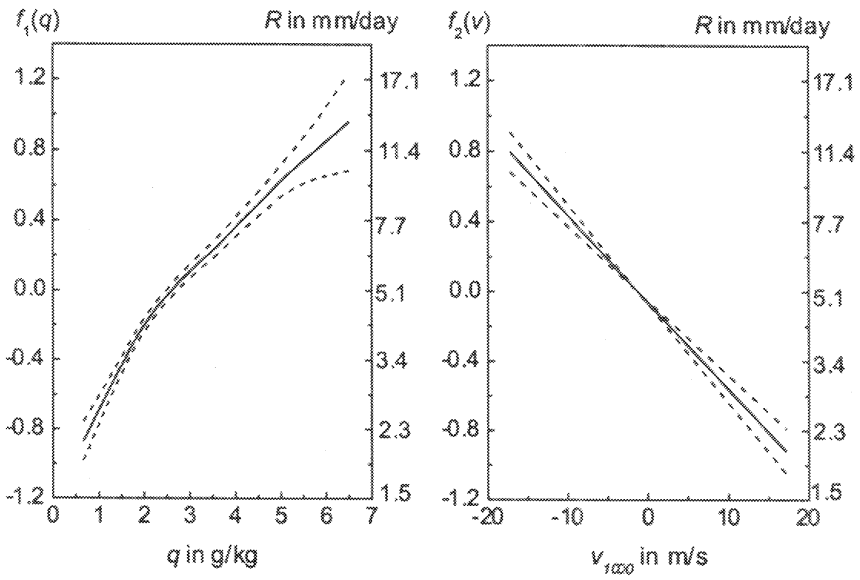

Figure 2: Estimates of the functions $f_i(x_i)$ for relative humidity at 700 hPa (rh) and southerly flow at 1000 hPa (v_{1000}) in a generalised additive model for wet-day rainfall at Berne (Switzerland). The dashed lines mark pointwise 2 x standard-error bands.

The shapes of the two leading predictor functions $f_i(x_i)$ for Berne (Switzerland) are presented in Figure 2. In this example the function $f_i(x_i)$ is non-linear for the specific humidity at 700 hPa (q) and linear for the southerly flow component at 1000 hPa (v_{1000}). The rainfall amount on the right-hand side of each panel is given by:

$$R = \exp\left[a_0 + f_i(x_i)\right] \tag{4}$$

Like the probability of rain, the mean wet-day rainfall amount at Berne tends to be high on days with northerly flow and low on days with southerly flow. Further, the expected wet-day rainfall amount increases with increasing specific humidity at 700 hPa (q).

Model selection was done in the same way as for rainfall occurrence. However, an approximate F-test (see the Appendix) was used instead of a χ^2-test to assess the significance of predictor variables and non-linearity of predictor functions.

3 Data

The statistical downscaling models were fitted to observed local precipitation data using predictor variables from the NCEP reanalysis. For the climate change application, a transient simulation with the coupled global atmosphere-ocean ECHAM4/OPYC3 model was considered. Details of the precipitation data are given first. Then the predictor variables are discussed. Finally, some properties of the ECHAM4/OPYC3 simulation are presented.

3.1 Observed precipitation data

Statistical models were developed for eleven European stations. These stations are listed in Table 1. The data were provided by the national weather services (either directly or via a WRINCLE partner). All time series data were available for the period 1968 - 1997 except the data for Vienna (1968 - 1994) and Munera (November 1969 - 1997). Precipitation data prior to 1968 were not examined because of a problem with the NCEP reanalysis data (see Section 3.2).

Table 1: Positions and characteristics of selected precipitation stations. MAR stands for mean annual rainfall and NWET for the mean annual number of wet days. Mean values refer to the period 1968 - 1997 except for Vienna (1968 - 1993) and Munera (1970 - 1997).

Station	Latitude	Longitude	Elevation asl in m	MAR in mm	NWET
The Netherlands					
De Bilt	51.10° N	5.18° E	2	802	193
Maastricht[a]	50.92° N	5.78° E	114	733	188
Germany					
Hamburg-Fuhlsbüttel	53.63° N	9.98° E	11	774	192
Hanover-Langenhagen	52.47° N	9.70° E	55	648	189
Berlin-Dahlem	52.47° N	13.30° E	51	581	176
Austria					
Vienna-Hohe Warte	48.25° N	16.37° E	203	601	150
Switzerland					
Berne-Liebefeld	46.93° N	7.42° E	565	1040	169
Payerne-Ville	47.27° N	6.94° E	450	878	145
Neuchâtel	47.00° N	6.95° E	485	960	169
Spain					
Salto de Bolarque	40.37° N	2.83° W	620	454	82
Munera	39.05° N	2.48° E	930	422	78

[a]Measurements at the airport in Beek

3.2 NCEP reanalysis data

An 'analysis' is an estimate of the state of the atmosphere based on observations and background information like a numerical weather forecast. Analyses form the basis for weather forecasts. In particular for climate research, reanalyses have been performed to get rid of nonhomogeneities due to developments in numerical weather modelling. In the 1990s a fifty-year reanalysis from 1948 was performed by the National Center of Environmental Prediction (NCEP), USA. Numerous weather variables from this reanalysis are available on a 2.5°x2.5° grid (KALNAY *et al.*, 1996). For the work described in this report the 6-hourly data in the period 1968 - 1997 were used. Earlier reanalysis data were not considered, because an error was identified in the conversion of sea level pressure data over Europe from 1948 till 1968. Because of this error most of the observed pressure values lower than 1000 hPa were discarded in the reanalysis (http://lnx21.wwb.noaa.gov/-images/psfc/psfc.html). For the calculation of the daily average of the reanalysis data, the four values were chosen, which were within the sampling period of a daily precipitation measurement. The predictor values of the grid point nearest to the rainfall measurement site were used (Table 2) for model development.

For rainfall occurrence modelling in the Netherlands and North Germany it emerged that the relative humidity at 700 hPa (rh)[1], the baroclinicity (b), the thickness of the layer between 1000 hPa and 500 hPa ($thick$), the sea level pressure (slp), the west component of the geostrophic wind (u_p), the south component of the geostrophic wind (v_p) and the geostrophic relative vorticity (ζ_p) are suitable predictor variables. The baroclinicity was taken as the absolute value of the temperature gradient on the 700 hPa level. Both the velocity components and the vorticity were derived from the sea level pressure at the four grid points nearest to the target grid point of the station of interest. For the more southerly sites it turned out to be better to use the geopotential heights at 1000 hPa (h_{1000}) or 850 hPa (h_{850}) instead of sea level pressure and the velocity components and vorticity derived from these heights. These flow variables are denoted as u_{1000}, v_{1000}, ζ_{1000} and u_{850}, v_{850}, ζ_{850}, respectively.

[1]The daily relative humidity was derived from the daily specific humidity and the daily temperature at 700 hPa.

Table 2: Selected NCEP reanalysis grid points.

Location	Latitude	Longitude
De Bilt	52.5° N	5.0° E
Maastricht	50.0° N	5.0° E
Hamburg, Hanover	52.5° N	10.0° E
Berlin	52.5° N	12.5° E
Vienna	47.5° N	15.0° E
Switzerland	47.5° N	7.5° E
Spain	40.0° N	2.5° W

For wet-day rainfall amount modelling the specific humidity at 700 hPa (q) or the precipitable water (pw) are more suitable predictors than the relative humidity at 700 hPa (rh). The correlation coefficient between q and *thick* ranges from 0.75 for wet days at Salto de Bolarque to 0.88 for wet days at the Swiss sites and that between pw and *thick* from 0.85 for wet days at Salto de Bolarque to 0.93 for wet days at Berlin and the Swiss sites. Hence, it is not possible to include both thickness and specific humidity or precipitable water in the statistical models. Inclusion of a moisture variable is preferable because of its stronger statistical and physical significance.

3.3 Climate model data

The estimated changes in the local precipitation in Section 5 and the daily precipitation scenarios in Section 6 are based on the results from the time-dependent greenhouse gas (GHG) experiment with the coupled global atmosphere-ocean model ECHAM4/OPYC3 of the Max-Planck-Institute of Meteorology and the Deutsche Klimarechenzentrum (both in Hamburg). A steadily growing concentration of greenhouse gases, as observed between 1860 and 1990 and according to IPCC scenario IS92a ('business as usual') from 1990 onward, was prescribed in that experiment (ARPE and ROECKNER, 1999; ROECKNER *et al.*,1999). Monthly values from the GHG experiment were made available. For each climate variable a climate change signal was obtained as the difference between the 30-year mean values for the periods 1968 - 1997 and 2070 - 2099. Note that the first period corresponds to that used for calibrating statistical relationships. Separate climate change signals were determined for the winter (October - March) and summer (April - September) halves of the year.

The climate change studies were done for De Bilt, Hanover, Berlin, Berne and Salto de Bolarque. The coordinates of the nearest ECHAM4/OPYC3 grid points to these sites are given in Table 3. The wind velocity components, the relative vorticity and the baroclinicity were obtained in the same way as those for the reanalysis data. The change in relative humidity in the GCM data was not used in this work (see Section 5.1). The precipitable water pw was not available in the climate model output. The winter and summer means of precipitable water were therefore derived from those of specific humidity at 700 hPa using linear regression relations fitted to the reanalysis data. Separate relations were used for winter and summer.

Table 3: Selected ECHAM4/OPYC3 grid points for climate change studies.

Location	Latitude	Longitude
De Bilt	51.63° N	5.63° E
Hanover	51.63° N	8.44° E
Berlin	51.63° N	14.06° E
Berne	46.04° N	8.44° E
Salto de Bolarque	40.46° N	2.81° W

Table 4: Winter means (1968 - 1997) of NCEP predictor variables, bias in the ECHAM4/OPYC3 winter mean relative to the NCEP mean, and changes for the period 2070 - 2099 for Hanover. Biases and changes printed in *italics* are significant at the 5% level (t-test).

Variable	NCEP mean 1968 - 1997	GHG bias 1968 - 1997	GHG change 2070 - 2099
q	1.59 g/kg	+ *0.18* g/kg	+ *0.57* g/kg
slp	1015.5 hPa	+ *1.7* hPa	+ 0.3 hPa
u_p	4.59 m/s	+ *4.18* m/s	+ 0.42 m/s
v_p	0.83 m/s	+ *0.43* m/s	- 0.01 m/s
ζ_p	-0.0083 kHz	+ *0.0138* kHz	- *0.0046* kHz
$thick$	5369.9 m	+ *37.9* m	+ *89.7* m
b	0.0076 K/km	- *0.0030* K/km	- 0.0003 K/km

Table 5: Summer means (1968 - 1997) of NCEP predictor variables, bias in the ECHAM4/OPYC3 summer mean relative to the NCEP mean, and changes for the period 2070 - 2099 for Hanover. Biases or changes printed in *italics* are significant at the 5% level (t-test).

Variable	NCEP mean 1968 - 1997	GHG bias 1968 - 1997	GHG change 2070 - 2099
q	2.62 g/kg	+ *0.40* g/kg	+ *0.87* g/kg
slp	1015.3 hPa	- *1.0* hPa	+ 0.2 hPa
u_p	1.60 m/s	+ *1.97* m/s	+ 0.03 m/s
v_p	- 0.77 m/s	+ *2.04* m/s	- 0.01 m/s
ζ_p	0.0021 kHz	+ 0.0006 kHz	- *0.0014* kHz
$thick$	5505.9 m	+ *54.6* m	+ *95.6* m
b	0.0066 K/km	- *0.0029* K/km	+ *0.0003* K/km

For the grid point close to Hanover Tables 4 and 5 compare the seasonal means of the predictor variables in the climate model output with those in the NCEP reanalysis for the winter and summer, respectively. The changes for the 2070-2099 period are also presented. There is almost always a significant difference between the ECHAM4/OPYC3 and NCEP mean values. The most notable biases are found for u_p and ζ_p in the winter season caused by a strong north-south gradient in sea level pressure over North Germany in the climate model output (see Figure 3). A discussion of the differences between the patterns of sea level pressure in the NCEP reanalysis and the ECHAM4/OPYC3 data can be found in Arpe and Roeckner (1999) and Knippertz *et al.* (2000). For both summer and winter the changes in specific humidity and thickness are statistically significant and exceed the bias. The increase in thickness of about 90 m is in agreement with the temperature increase of about 4.5°C in the climate model. The relative increase in specific humidity amounts to 7.5%/°C in winter and 6.4%/°C in summer, which is close

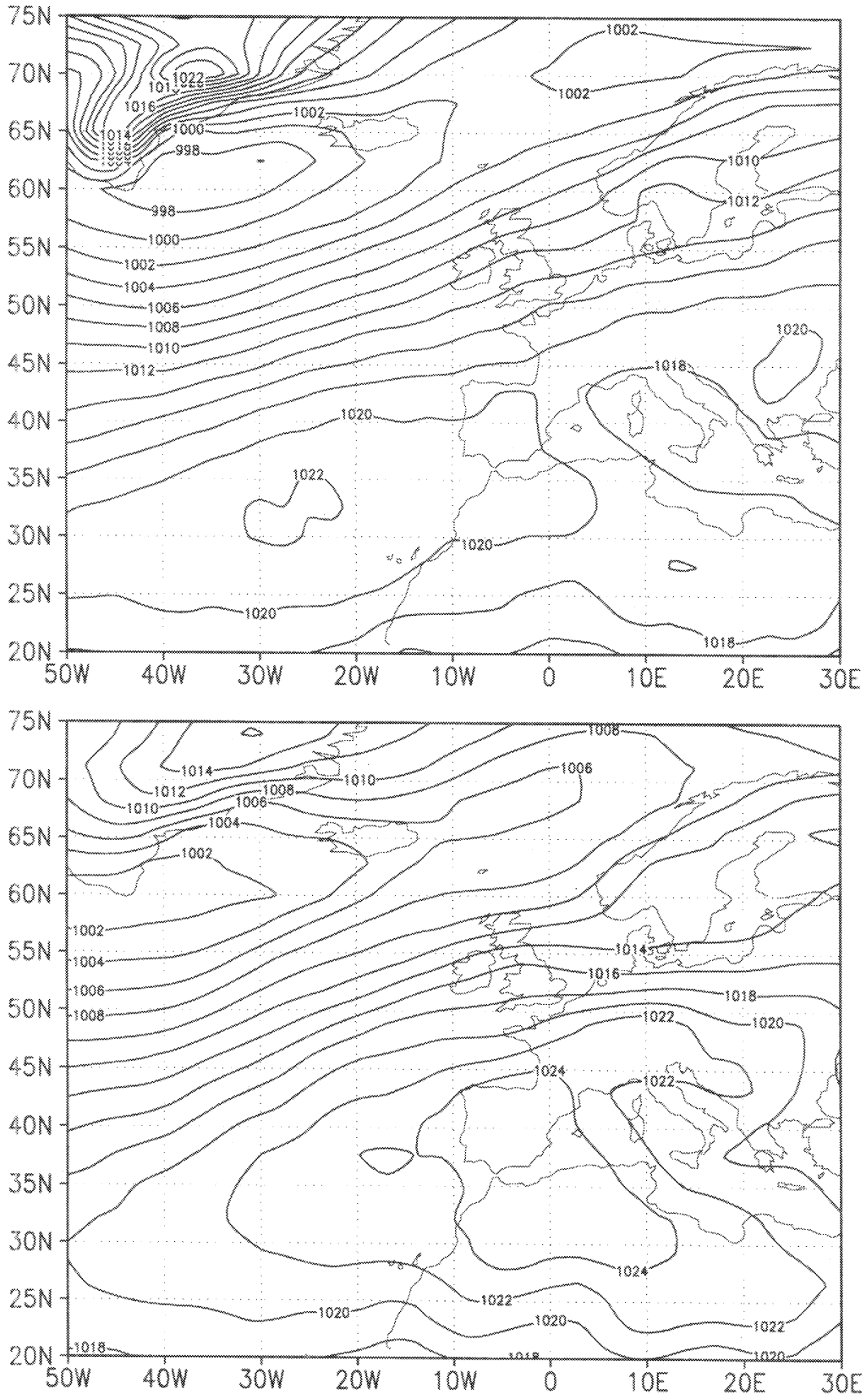

Figure 3: Mean sea level pressure in hPa for the winter months December to February (DJF) for the period 1968 - 1997 in the NCEP reanalysis (top) and the ECHAM4/OPYC3 GHG experiment (bottom).

to that expected from the Clausius-Clapeyron relation. Significant changes are also found for the vorticity but for the winter half of the year the change is smaller than the bias. The changes in the other variables are small compared with their bias.

The other four grid points show similar changes and biases in q and *thick*. Biases in dynamical variables are also observed at these grid points. Their sign and magnitude differ from grid point to grid point. Sometimes these biases are quite large because velocity components and vorticity are very sensitive to small changes in the seal level pressure and geopotential height patterns. Changes and biases in baroclinicity are more or less similar to those at Hanover.

4 Results for rainfall occurrence and rainfall amount

In this section the fitted relationships are discussed. Results for rainfall occurrence are presented in Section 4.1 and for wet-day rainfall amount in Section 4.2. Section 4.3 deals with some issues concerning seasonal variation.

4.1 Rainfall occurrence

For all precipitation stations in Table 1 additive logistic regression models were fitted to the rainfall occurrence data. Table 6 presents a survey of the chosen predictor variables. The proportion of explained variance in the third column of the table is given by

$$EV = 1 - \frac{\sum_{t=1}^{N} (y_t - P_t)^2}{\sum_{t=1}^{N} (y_t - \bar{y})^2} \tag{5}$$

where y_t is the observed rainfall occurrence at day t, $\bar{y}$ is the average of the y_t values (fraction of wet days), P_t is the estimated rainfall probability for day t and N is the number of days in the record. The residual autocorrelation in the last column refers to the lag 1 autocorrelation coefficient of the residuals $\{y_t - P_t\}$.

For all stations six or seven predictors were selected. About 50% of the predictor functions $f_i(x_i)$ are non-linear. The u-velocity is the most significant predictor for De Bilt, Maastricht, Hamburg and Hanover. For the more continental station Berlin relative humidity becomes the most significant predictor. This is also the case for the southern stations. A remarkable point is that vorticity has only a minor effect on rainfall occurrence at Vienna and the Swiss sites. The baroclinicity is an important predictor for these sites. For the Spanish sites the baroclinicity is not significant. Two different models are presented for Vienna. For the first one the dynamical predictors are based on the 1000 hPa heights and for the second one on the 850 hPa heights. Both models explain nearly 37% of the variance of daily rainfall occurrence. Something similar was found for the three Swiss stations. For the climate change applications for Berne in Sections 5.4 and 6 the model with the 1000 hPa data in Table 6 is used. For the Spanish sites the use of 850 hPa heights proved to be superior.

Relatively strong correlations were found between *slp* and vorticity ζ_p (-0.35 to -0.49), the 1000 hPa height h_{1000} and vorticity ζ_{1000} (-0.37 to -0.58), and the 850 hPa height h_{850} and 1000-500 hPa thickness *thick* (0.52 to 0.54). In most cases two separate predictor effects were clearly identified. Only was the thickness not included in the rainfall occurrence models with 850 hPa data for Vienna and the Swiss sites. The consequences of the use of both the thickness and the 850 hPa heights for the Spanish grid point are discussed in Section 5.5.

The sequence of wet and dry days exhibits moderate persistence. The lag 1 autocorrelation coefficient for Hanover is for instance 0.37. The residual autocorrelation coefficients are much smaller (see Table 6) but still statistically significant. The persistence of daily rainfall occurrence is thus not fully captured by the autocorrelation in the predictor variables. Inclusion of the wet-dry status of the previous day may improve the reproduction

Table 6: Predictors for rainfall occurrence for different European sites. The first variable in the row of predictors is the most significant predictor variable and the last variable the least significant one. For the predictors in bold face $f_i(x_i)$ is not linear.

Station	Predictors	Expl. Variance	Resid. Autocor.
De Bilt	$\boldsymbol{u_p}$, $\boldsymbol{rh}$, ζ_p, $\boldsymbol{v_p}$, $\boldsymbol{thick}$, b, slp	47.8%	0.111
Maastricht	$\boldsymbol{u_p}$, $\boldsymbol{rh}$, $\boldsymbol{v_p}$, ζ_p, $\boldsymbol{thick}$, b, slp	47.5%	0.117
Hamburg	$\boldsymbol{u_p}$, $\boldsymbol{rh}$, ζ_p, $\boldsymbol{thick}$, $\boldsymbol{v_p}$, b, slp	49.1%	0.138
Hanover	$\boldsymbol{u_p}$, $\boldsymbol{rh}$, ζ_p, v_p, $\boldsymbol{thick}$, b, slp	46.2%	0.138
Berlin	$\boldsymbol{rh}$, $\boldsymbol{u_p}$, v_p, ζ_p, $\boldsymbol{thick}$, b, slp	43.1%	0.125
Vienna	$\boldsymbol{rh}$, $\boldsymbol{v_{1000}}$, b, $thick$, u_{1000}, ζ_{1000}, h_{1000}	36.7%	0.139
	$\boldsymbol{rh}$, $\boldsymbol{v_{850}}$, b, $\boldsymbol{u_{850}}$, ζ_{850}, $\boldsymbol{h_{850}}$	36.9%	0.131
Berne	rh, $\boldsymbol{v_{1000}}$, $\boldsymbol{u_{1000}}$, b, $\boldsymbol{thick}$, $\boldsymbol{h_{1000}}$, ζ_{1000}	47.4%	0.146
Neuchâtel	$\boldsymbol{rh}$, $\boldsymbol{u_{1000}}$, $\boldsymbol{v_{1000}}$, b, $\boldsymbol{thick}$, h_{1000}, ζ_{1000}	45.6%	0.163
Payerne	rh, $\boldsymbol{u_{1000}}$, $\boldsymbol{v_{1000}}$, b, h_{1000}, ζ_{1000}, $thick$	44.2%	0.144
Salto de Bolarque	rh, $\boldsymbol{u_{850}}$, ζ_{850}, $\boldsymbol{thick}$, v_{850}, h_{850}	39.2%	0.129
Munera	rh, $\boldsymbol{u_{850}}$, ζ_{850}, $\boldsymbol{v_{850}}$, $\boldsymbol{thick}$, h_{850}	41.1%	0.123

of the persistence of daily rainfall occurrence. An other possibility is to develop separate models for the probabilities of a wet day following a wet day and that of a wet day following a dry day (CONWAY *et al.*, 1996; WILBY *et al.*, 1998).

Figure 4 shows the estimates of the functions $f_i(x_i)$ in Equation 1 for all seven predictor variables for Hanover. The rainfall probability increases with increasing wind speed from westerly and northerly directions. Weather conditions with strong westerly flow over Europe are favourable for rainfall. The occurrence of showers at the rear of cold fronts contributes to the relatively high rainfall probability in situations with strong northerly flow. A relatively high rainfall probability on days with northerly flow was also observed in Figure 1 for Berne. The opposite occurs at the two Spanish stations where rainfall probability increases with increasing southerly flow. Figure 4 further delineates the well-known facts from meteorology that the probability of rain increases with increasing vorticity (or cyclonality), relative humidity and baroclinicity, and that it decreases with increasing sea level pressure. The decrease of P with increasing 1000-500 hPa thickness is similar for all stations considered in this report.

4.2 Rainfall amount

The generalised additive model introduced in Section 2.2 was used to analyse the wet-day precipitation amounts at all sites. Two different models were created for every site: one model with specific humidity at 700 hPa (q) as atmospheric moisture variable and another one with precipitable water (pw) and relative humidity at 700 hPa (rh). It was not possible to include both specific humidity and relative humidity into one model because of the correlation between these two variables (0.39 for wet days at Payerne up to 0.59 for

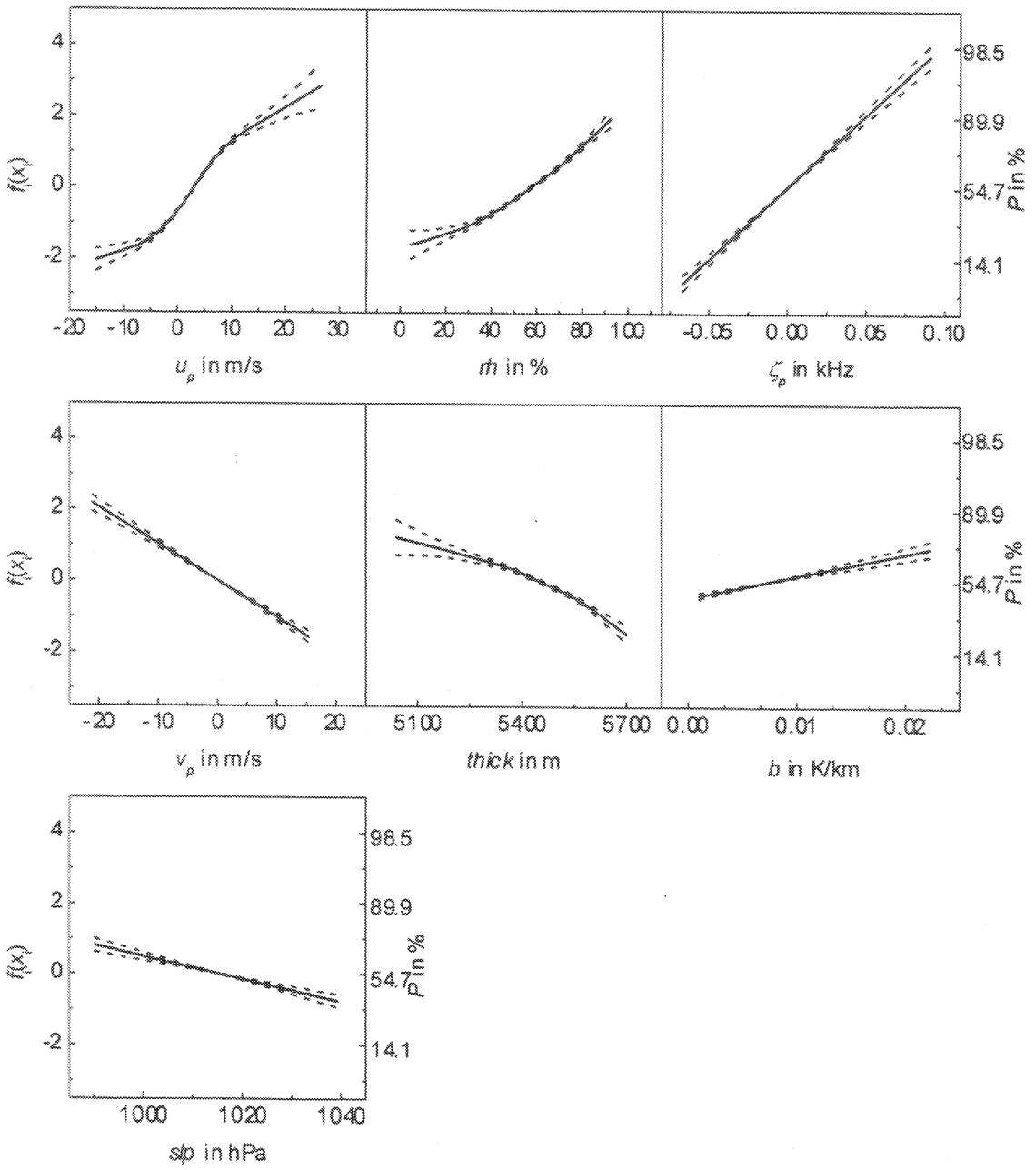

Figure 4: Estimates of the functions $f_i(x_i)$ for the seven predictor variables in the additive logistic model for rainfall occurrence at Hanover. The dashed lines mark pointwise 2 x standard-error bands.

Table 7: Predictors for rainfall amount for different European sites. The first variable in the row of the predictors is the most significant predictor variable and the last variable the least significant one. For the predictors in bold face $f_i(x_i)$ is not linear.

Station	Predictors	Expl. Variance
De Bilt	$\boldsymbol{q}$, $\boldsymbol{\zeta_p}$, u_p, $\boldsymbol{slp}$, b, v_p	29.7%
	$\boldsymbol{\zeta_p}$, u_p, $\boldsymbol{pw}$, $\boldsymbol{slp}$, rh, b, v_p	32.0%
Maastricht	$\boldsymbol{q}$, u_p, $\boldsymbol{v_p}$, $\boldsymbol{\zeta_p}$, $\boldsymbol{slp}$, b	30.2%
	u_p, $\boldsymbol{pw}$, v_p, rh, $\boldsymbol{\zeta_p}$, b, slp	33.3%
Hamburg	$\boldsymbol{u_p}$, $\boldsymbol{q}$, $\boldsymbol{\zeta_p}$, $\boldsymbol{slp}$, $\boldsymbol{v_p}$, b	28.3%
	u_p, $\boldsymbol{\zeta_p}$, $\boldsymbol{pw}$, rh, $\boldsymbol{v_p}$, $\boldsymbol{slp}$, b	29.6%
Hanover	$\boldsymbol{\zeta_p}$, $\boldsymbol{q}$, u_p, $\boldsymbol{slp}$, $\boldsymbol{v_p}$, b	29.4%
	$\boldsymbol{\zeta_p}$, $\boldsymbol{rh}$, $\boldsymbol{u_p}$, $\boldsymbol{pw}$, b, $\boldsymbol{v_p}$, $\boldsymbol{slp}$	31.3%
Berlin	$\boldsymbol{q}$, $\boldsymbol{\zeta_p}$, $\boldsymbol{v_p}$, $\boldsymbol{slp}$, b, $\boldsymbol{u_p}$	25.1%
	$\boldsymbol{pw}$, rh, $\boldsymbol{v_p}$, $\boldsymbol{\zeta_p}$, slp, b, $\boldsymbol{u_p}$	28.3%
Vienna	$\boldsymbol{v_{1000}}$, $\boldsymbol{q}$, ζ_{1000}, b, h_{1000}	22.1%
	$\boldsymbol{v_{1000}}$, $\boldsymbol{rh}$, $\boldsymbol{pw}$, ζ_{1000}, b, h_{1000}	25.4%
	q, $\boldsymbol{v_{850}}$, ζ_{850}, b, u_{850}, h_{850}	23.6%
	$\boldsymbol{v_{850}}$, $\boldsymbol{pw}$, rh, ξ_{850}, b, h_{850}	25.0%
Berne	$\boldsymbol{q}$, v_{1000}, b, $\boldsymbol{h_{1000}}$, $\boldsymbol{u_{1000}}$, ζ_{1000}	20.6%
	$\boldsymbol{pw}$, b, v_{1000}, $\boldsymbol{h_{1000}}$, $\boldsymbol{u_{1000}}$, rh, ζ_{1000}	22.7%
Neuchâtel	q, u_{1000}, b, h_{1000}, $\boldsymbol{v_{1000}}$, ζ_{1000}	22.3%
	pw, u_{1000}, b, $\boldsymbol{rh}$, h_{1000}, $\boldsymbol{v_{1000}}$, ζ_{1000}	25.6%
Payerne	$\boldsymbol{q}$, b, $\boldsymbol{h_{1000}}$, $\boldsymbol{u_{1000}}$, ζ_{1000}	20.1%
	pw, b, h_{1000}, $\boldsymbol{u_{1000}}$, $\boldsymbol{rh}$, ζ_{1000}	21.5%
Salto de Bolarque	q, ζ_{850}, h_{850}	11.8%
	pw, ζ_{850}, h_{850}	13.3%
Munera	ζ_{850}, q, h_{850}	8.1%
	pw, ζ_{850}, h_{850}	9.0%

wet days at De Bilt). The correlation between relative humidity and precipitable water (0.24 for wet days at Payerne up to 0.40 for wet days at Salto de Bolarque) is at every site weaker than that between relative humidity and specific humidity. The chosen predictor variables for both models and the explained variances are given in Table 7.

The models with precipitable water and relative humidity generally explain a somewhat larger percentage of the variance of wet-day rainfall than those with specific humidity. For the Netherlands, Hamburg and Hanover about 30% of the daily variance can be explained. These are the highest values in Table 7. The percentage of explained variance is slightly less for the more continental station Berlin in Germany. For Austria and Switzerland only 20 - 25% of the daily variance in rainfall amount can be explained by the statistical models. In these regions local effects like forced convection in the mountains become more important. The rainfall models for the two Spanish sites Salto de Bolarque and

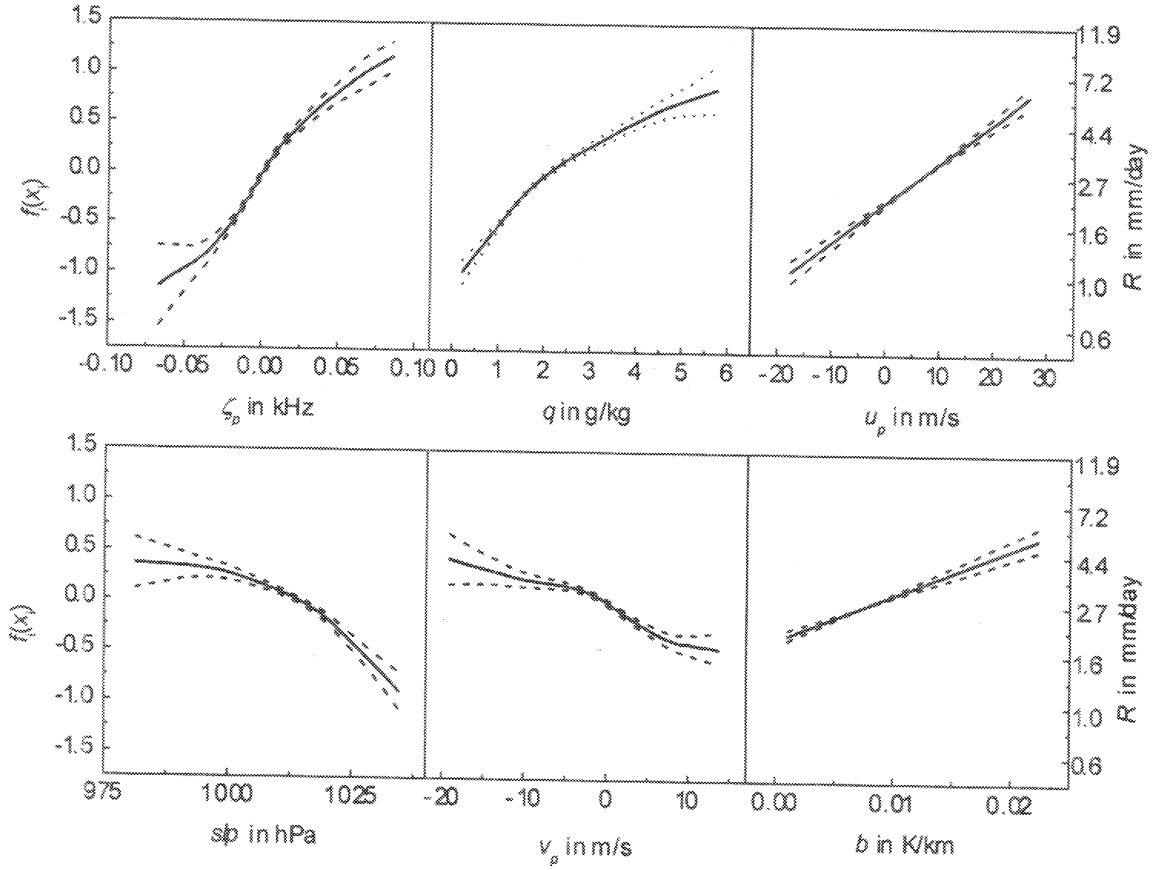

Figure 5: Estimates of the functions $f_i(x_i)$ for the six predictor variables in a generalised additive model for wet-day rainfall at Hanover. The dashed lines mark pointwise 2 x standard-error bands.

Munera explain substantially less of the daily variance (8.1 - 13.3%) than those for the other investigated sites.

It is not possible to include all predictors for rainfall occurrence modelling in the statistical model for wet-day rainfall. Because of its strong correlation with specific humidity and precipitable water, the 1000-500 hPa thickness cannot be entered into the statistical models for the wet-day rainfall amounts (Section 3.2). In only one of the four presented models for Vienna the u-velocity is significant. The v-velocity is, however, an important predictor variable for that station. For Payerne the v-velocity is not significant. This is a marked contrast with the wet-day rainfall amount models for Berne. For the statistical description of the rainfall amounts at the two Spanish sites Salto de Bolarque and Munera it turned out that beside an absolute moisture variable, the 850 hPa height and the relative vorticity at 850 hPa are the only appropriate predictors.

Figure 5 shows the estimates of the functions $f_i(x_i)$ in Equation 3 for all six predictor variables in the model with specific humidity for Hanover. For specific humidity the shape of $f_i(x_i)$ resembles that for Berne in Figure 2. In the alternative rainfall amount model for Hanover the expected rainfall amount increases with increasing precipitable

water and relative humidity. The remaining predictor functions for rainfall amount show qualitatively the same behaviour as those in rainfall occurrence modelling (compare Figure 4 with Figure 5).

4.3 Seasonal variation

4.3.1 Reproduction of the seasonal cycle

The relationships presented in Sections 4.1 and 4.2 were assumed constant over the year. A quick test of this assumption is possible by comparing the observed rainfall probabilities and mean wet-day rainfall amounts for each calendar month with those expected from the models.

Figure 6 shows the reproduction of the seasonal cycles for rainfall probability at Hanover and Berne. In both cases the models describe the observed seasonal cycle well. The errors are generally not more than twice the standard error of the observed monthly means. For Hanover there is a slight overestimation of the precipitation probability in the first half of the year and an underestimation in the second half of the year. This discrepancy was also found in the fits for the other German and the Dutch stations. For Berne there is some overestimation of the precipitation probability in the spring months March and April, which also occurs in the Neuchâtel and Payerne data. A standard one-sample t-test on

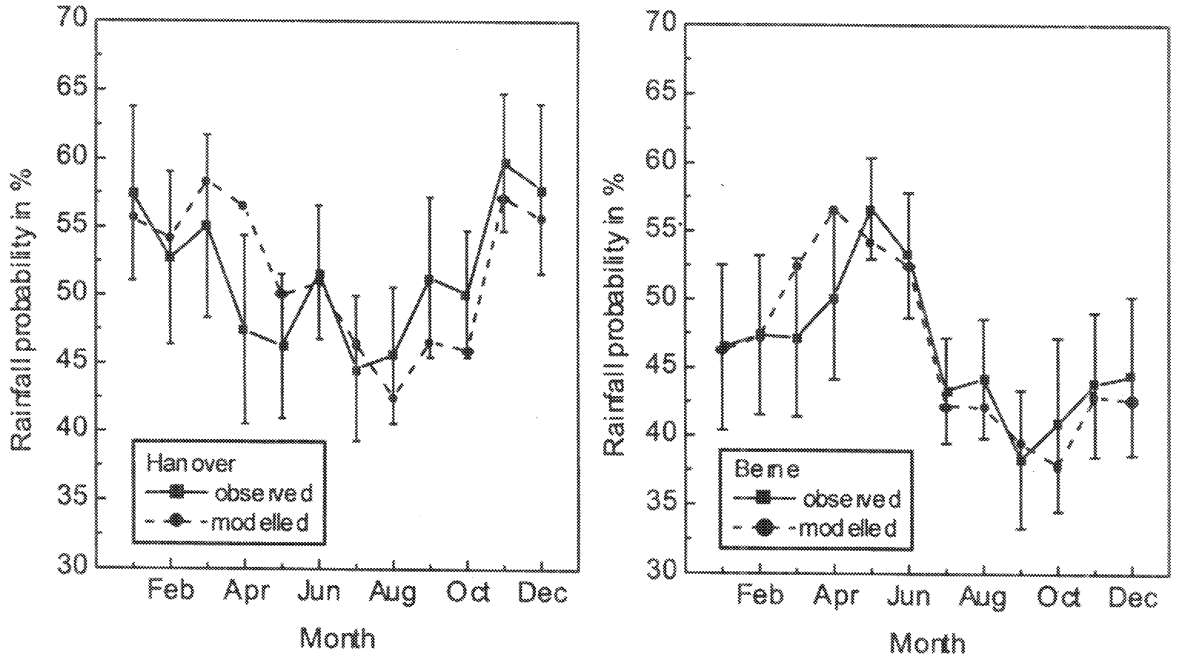

Figure 6: Observed and predicted seasonal cycles of the daily rainfall probability at Hanover and Berne for the period 1968 - 1997. The error bars indicate twice the standard error of the observed monthly means.

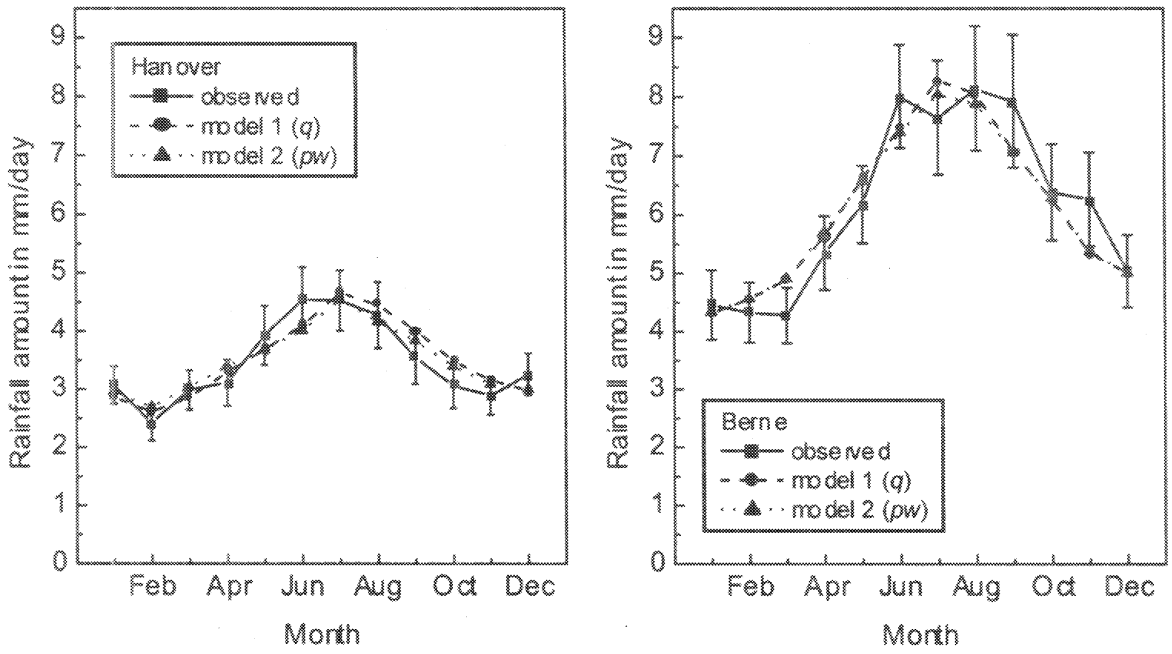

Figure 7: Observed and predicted seasonal cycles of the mean wet-day rainfall amount at Hanover and Berne for the period 1968 - 1997. In model 1 atmospheric moisture is represented as specific humidity, and in model 2 as relative humidity and precipitable water. The error bars indicate twice the standard error of the observed mean wet-day rainfall amounts.

the differences between the observed and modelled monthly values shows that several of these discrepancies are statistically significant at the 5% level. For Vienna the predicted monthly means correspond reasonably with the observed values, except for the month of April. The modelled values for the Spanish sites are all within the error bands of the observed values. The rainfall probability in the spring months March, April and May is overestimated at these sites.

Figure 7 compares the observed seasonal cycles of the mean wet-day rainfall amounts for Hanover and Berne with that predicted by the fitted models. The two wet-day amount models describe the seasonal cycles at these stations equally well. This holds also for most other stations. For De Bilt, Maastricht and Hamburg, however, the model with precipitable water and relative humidity performs somewhat better than the model with specific humidity. The differences between the observed and predicted mean wet-day precipitation amounts are sometimes larger than twice the standard error of the observed means[2]. For the three German and two Dutch stations the mean wet-day rainfall amount

[2]The standard error of the observed mean wet-day precipitation amount was derived here from the sample standard deviation of the wet-day amounts, assuming independence between the daily values. The use of the mean wet-day rainfall amounts for each month meets difficulties because of the random number of wet days in a month.

in May and June is underestimated. The largest underestimations (up to 0.9 mm per wet day) were found for De Bilt and Hamburg. For Berne there is a rather strong overestimation ($\approx$ 0.6 mm for both models) of the mean wet-day rainfall amount in March and an underestimation during the second half of the year. The other two Swiss stations show similar discrepancies. A considerable overestimation ($\approx$ 0.7 mm for the model with specific humidity and $\approx$ 0.9 mm for the alternative model) emerged for Vienna in April. The models further overestimate the mean precipitation amounts for the Spanish sites in July and August. For Munera the modelled values for these months are outside the error bars of the observed means.

4.3.2 Separate seasonal models

The discrepancies between the observed and modelled seasonal cycles suggest that the relationships are not constant over the year. For Hanover, the seasonal variation of the effects of the predictor variables was investigated by modelling the data for the winter and summer halves of the year separately. The results for the three most significant predictor variables in the rainfall occurrence model are presented in Figure 8. In particular, the effect of relative humidity is stronger in summer than in winter, whereas the effect of vorticity is somewhat stronger in winter. The approximate χ^2-test in the Appendix shows that the differences between two seasons are significant at the 1% level. For the wet-

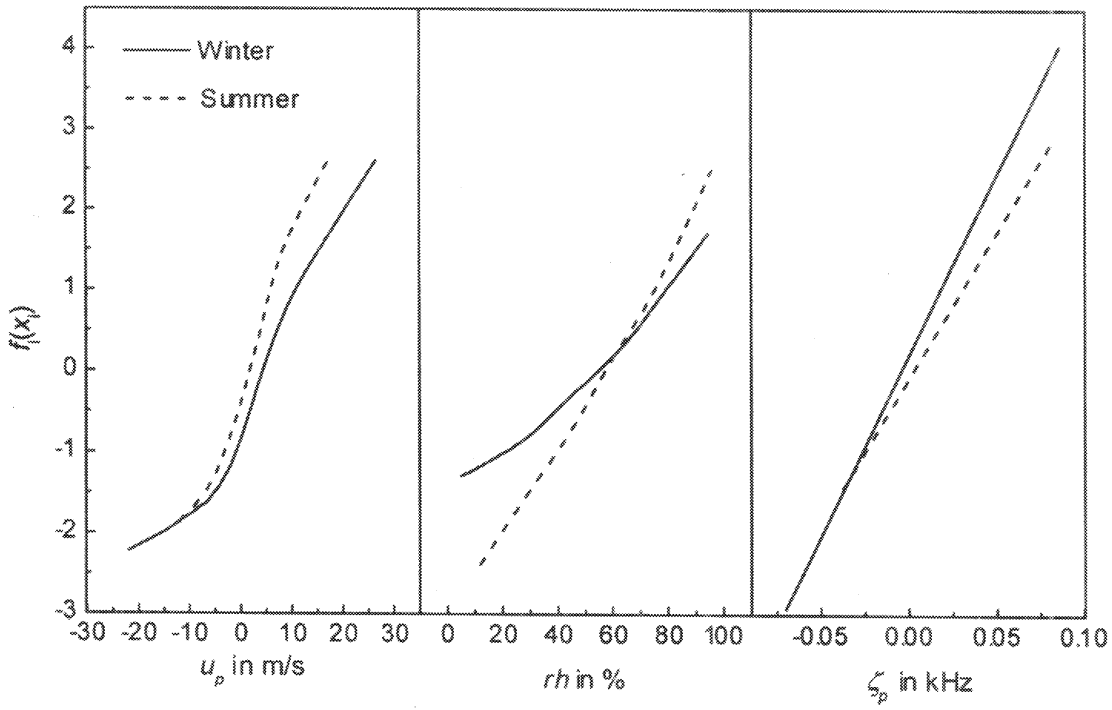

Figure 8: Estimates of the functions $f_i(x_i)$ for westerly flow (u_p), relative humidity at 700 hPa (rh) and relative vorticity (ζ_p) of the additive logistic model for rainfall occurrence at Hanover. Separate fits for the winter and summer halves of the year.

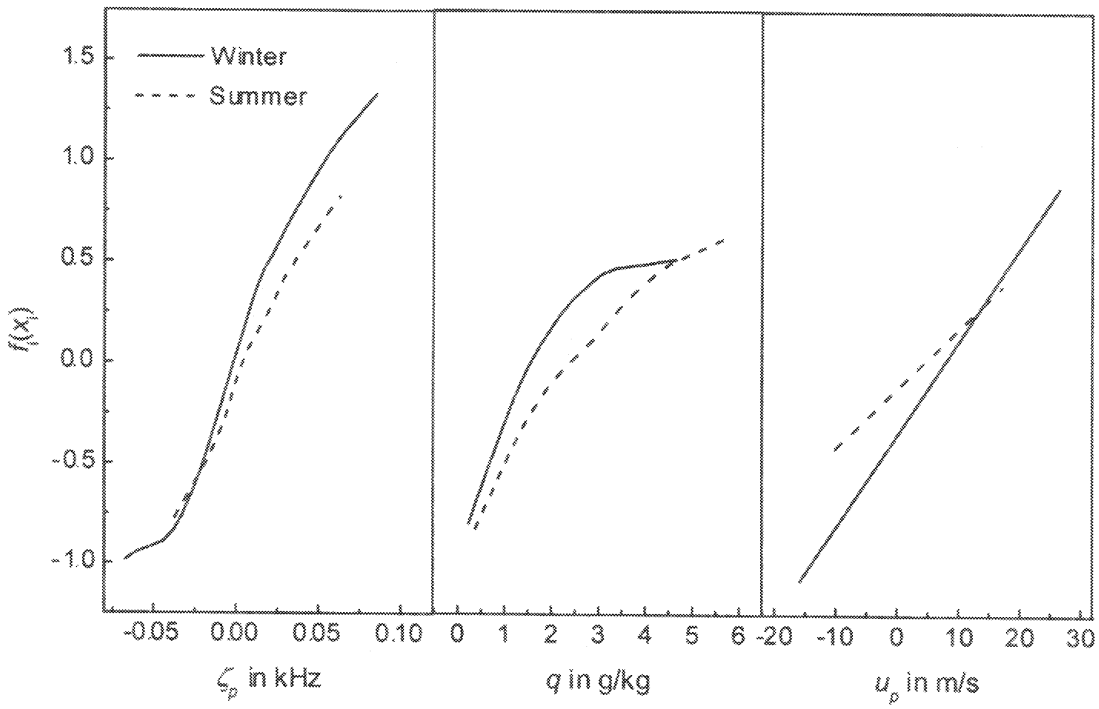

Figure 9: Estimates of the functions $f_i(x_i)$ for relative vorticity (ζ_p), specific humidity at 700 hPa (q) and westerly flow (u_p) of a generalised additive model for wet-day rainfall at Hanover. Separate fits for the winter and summer halves of the year.

day rainfall model with specific humidity, Figure 9 compares the seasonal estimates of the $f_i(x_i)$ for the relative vorticity, specific humidity and u-velocity. The effect of the u-velocity is much stronger in winter than in summer season. Like for rainfall occurrence, the effect of relative vorticity is also somewhat stronger in winter. The improvement due to using separate relationships for summer and winter halves of the year is significant at the 1% level (approximate F-test).

5 Changes in seasonal means

In this section the changes in the seasonal mean number of wet days and the seasonal mean precipitation amounts are estimated for the end of the 21^{st} century using the results of the ECHAM4/OPYC3 climate model described in Section 3.3. The seasonal mean values for the present climate refer to the period 1968 - 1997 and those for the future climate to the period 2070 - 2099.

5.1 Method of calculation

The mean number of wet days N_w and the mean precipitation amount R_{tot} for a particular season (winter and summer) were calculated as:

$$N_w = \frac{1}{J} \sum_t P_t \tag{6}$$

$$R_{tot} = \frac{1}{J} \sum_t P_t R_t \tag{7}$$

where P_t is the modelled probability of precipitation for day t, R_t is the expected precipitation amount for day t given that it is wet, and J is the number of years. The summation is over all days in the period concerned. The wet-day rainfall amount model is thus applied for every day t and not only for the wet days.

To estimate the seasonal mean number of wet days N_w^* and the seasonal mean precipitation amount R_{tot}^* for the end of the 21^{st} century, Equations 6 and 7 were applied to the perturbed predictor variables:

$$x_{ti}^* = x_{ti} + \Delta x_i \tag{8}$$

where x_{ti} is the value of the ith predictor for day t in the period 1968 - 1997 (from the NCEP reanalysis data) and Δx_i is the change in the mean of x_i between the periods 1968 - 1997 and 2070 - 2099 in the climate model experiment. However, the perturbed daily values of the relative humidity were not determined as $rh + \Delta rh$. They were derived from the perturbed daily values of the specific humidity and temperature at 700 hPa.

5.2 Model modifications for use in climate change studies

In the case of climate change it is obvious that the range of the predictor variables is somewhat different from that used in model construction. Unfortunately, the S-PLUS software only provides the smoothed functions within the range of the observed data. In order to make extrapolations, the smooth estimates of non-linear predictor functions were replaced by piecewise linear functions. As an example the approximations for sea level pressure (slp), specific humidity (q) and v-velocity (v_p) in a wet-day rainfall amount model for Hanover are presented here. The predictor function for slp in that model is splitted into a constant and a linear function:

$$f(slp) = \begin{cases} a_{slp0} & \text{if} \quad slp < c_{slp} \\ a_{slp0} + a_{slp1}\,(slp - c_{slp}) & \text{if} \quad slp \geq c_{slp} \end{cases} \tag{9}$$

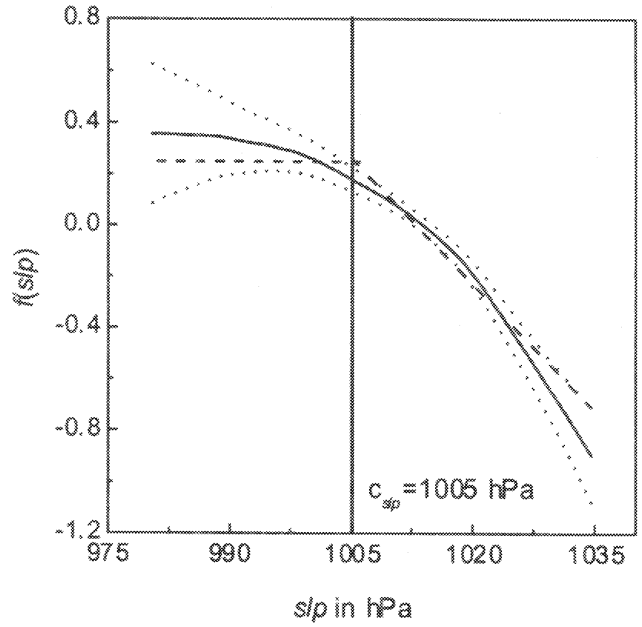

Figure 10: Smooth estimate of the function $f_i(x_i)$ for the sea level pressure *slp* in a generalised additive model for wet-day rainfall at Hanover (solid line). The dotted lines mark the accompanied 2 x standard error bands. The dashed line marks the piecewise linear approximation.

where the knot c_{slp} is located at 1005 hPa (see Figure 10). The predictor function for q is also splitted into two functions, both linearly increasing with increasing q (see Figure 11):

$$f(q) = \begin{cases} a_{q0} + a_{q1} q & \text{if} \quad q < c_q \\ a_{q0} + a_{q1} c_q + a_{q2} (q - c_q) & \text{if} \quad q \geq c_q \end{cases} \tag{10}$$

The piecewise linear splitting of the predictor function for v_p is shown in Figure 12. In this case $f(v_p)$ is approximated by a constant and two linear functions with two knots c_{v1} and c_{v2}:

$$f(v_p) = \begin{cases} a_{v0} & \text{if} \quad v_p > c_{v2} \\ a_{v0} + a_{v1} (c_{v2} - v_p) & \text{if} \quad c_{v1} < v_p \leq c_{v2} \\ a_{v0} + a_{v1} (c_{v2} - c_{v1}) + a_{v2} (c_{v1} - v_p) & \text{if} \quad v_p \leq c_{v1} \end{cases} \tag{11}$$

The use of these piecewise linear approximations worsens the fit to the observed wet-day rainfall amounts in terms of the deviance. The change in the deviance is, however, not significant according to the approximate F-test in the Appendix. The piecewise linear approximation for the relative vorticity predictor function is similar to that for the specific humidity in Figure 11 with a knot at 0.02 kHz.

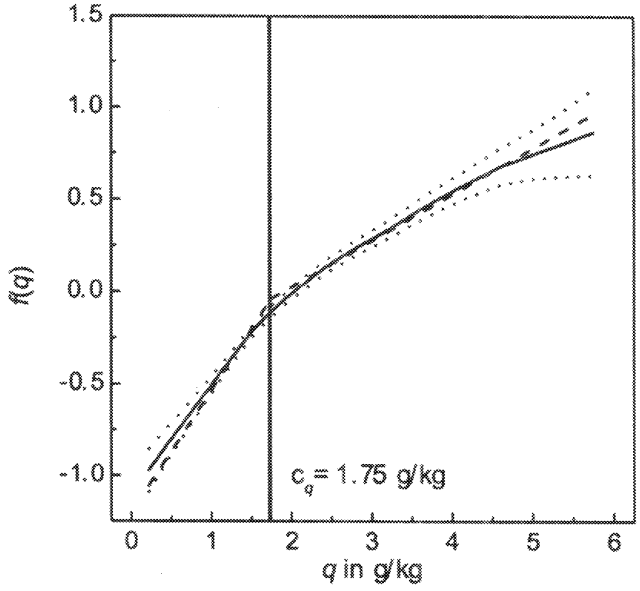

Figure 11: Smooth estimate of the function $f_i(x_i)$ for the specific humidity q in a generalised additive model for wet-day rainfall at Hanover (solid line). The dotted lines mark the accompanied 2 x standard error bands. The dashed line marks the piecewise linear approximation.

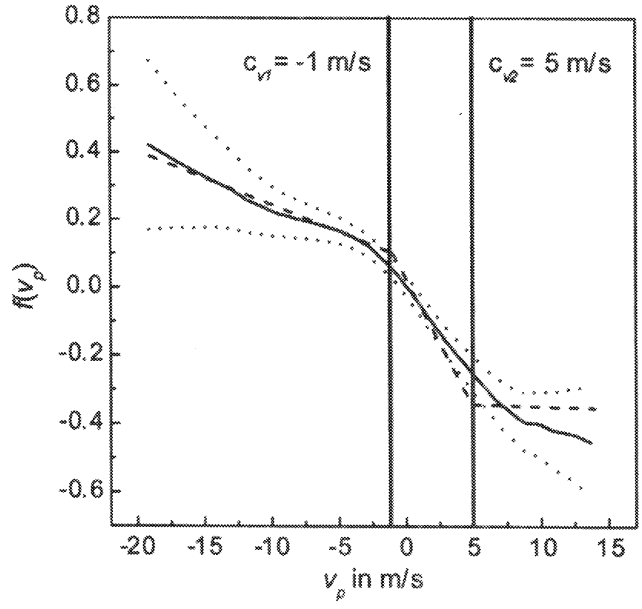

Figure 12: Smooth estimate of the function $f_i(x_i)$ for the southerly flow v_p in a generalised additive model for wet-day rainfall at Hanover (solid line). The dotted lines mark the accompanied 2 x standard error bands. The dashed line marks the piecewise linear approximation.

Table 8: Estimated coefficients of the linear and piecewise linear functions in the wet-day rainfall models with specific humidity (left) and in the rainfall occurrence models (right) for De Bilt, Hanover and Berlin.

	De Bilt	Hanover	Berlin
a_{slp1}	- 0.041	- 0.033	- 0.034
a_u	0.041	0.040	0.018
a_{v1}	- 0.048	- 0.075	- 0.076
a_{v2}	- 0.012	- 0.016	- 0.040
$a_{\zeta 1}$	19.92	23.38	18.78
$a_{\zeta 2}$	7.56	10.67	10.07
a_{q1}	0.63	0.67	0.83
a_{q2}	0.25	0.25	0.27
a_b	36.09	46.02	34.96

	De Bilt	Hanover	Berlin
a_{slp}	- 0.038	- 0.032	- 0.026
a_{thick1}	- 0.0063	- 0.026	- 0.0025
a_{thick2}	- 0.052	- 0.052	- 0.047
a_{u1}	0.20	0.24	0.18
a_{u2}	0.062	0.088	0.101
a_{v1}	- 0.15	- 0.11	- 0.13
a_{v2}	- 0.031	-	-
a_ζ	37.81	40.82	30.31
a_{rh1}	0.03	0.02	0.009
a_{rh2}	0.065	0.051	0.059
a_b	72.41	61.29	54.17

Such piecewise linear approximations to the non-linear predictor functions were developed for the rainfall occurrence and amount models of De Bilt, Hanover, Berlin, Berne and Salto de Bolarque. Approximations with more than two knots were not considered.

Although the resulting approximations are continuous, this does not hold for their derivatives. In BRANDSMA and BUISHAND (1997) natural cubic splines were used. This leads to smoother approximations but it makes the comparison between stations more difficult.

There are strong similarities between the piecewise linear approximations for the three northern sites De Bilt, Hanover and Berlin. For all three sites the knots in the non-linear functions of the rainfall amount model (i.e. $f(slp)$, $f(q)$, $f(v_p)$, $f(\zeta_p)$ in the model presented above) were chosen at the same location. The estimated regression coefficients do not vary much between the stations except for the westerly flow (see Table 8). For the more continental station Berlin the effect of the westerly flow on wet-day rainfall is much weaker than that for De Bilt and Hanover. Another point is that the regression coefficients for vorticity are relatively large for Hanover. In contrast with the other stations, vorticity is the most significant predictor variable for wet-day rainfall at Hanover (see Table 7). In the rainfall occurrence model two linear functions with a common knot at 5400 m were used for the thickness. There is, however, a significant increase in the deviance of the models if the knots for the velocity components and relative vorticity are taken to be the same for the three stations. The function $f(u_p)$ was approximated by two linear functions with a knot at 8.5 m/s (De Bilt) or by a constant and two linear functions with knots at -3.5 m/s and 11 m/s (Hanover and Berlin). For $f(rh)$ two linear functions with a knot at 60% (De Bilt) or 40% (Hanover and Berlin) were chosen. For De Bilt $f(v_p)$ was splitted into two linear functions with a knot at -1 m/s, whereas for Hanover and Berlin no splitting was necessary because $f(v_p)$ was already linear in the original models (see Table 6). The estimated regression coefficients in the linear and piecewise linear predictor functions for the three stations are presented in Table 8. The differences between the stations are

small. Again the regression coefficient for vorticity is relatively large for Hanover.

5.3 Changes for De Bilt, Hanover and Berlin

In this section the changes in the seasonal means (number of wet days, rainfall amounts) for the northern sites De Bilt, Hanover and Berlin are discussed. The seasonal means were derived from Equations 6 and 7 using the fitted models with linear and piecewise linear functions. The results are summarised in Table 9. The number of wet days decreases at all stations mainly due to the increased 1000-500 hPa thickness. This change in thickness has the strongest effect in summer because the slope of $f(thick)$ is steeper at large values of *thick* (Table 8). The decrease in vorticity during winter (Table 4) also contributes to the decrease in the number of wet days, in particular at Hanover which has the largest change in vorticity in the climate model output and the largest regression coefficient for vorticity in the rainfall occurrence model. The relatively small decrease in the number of wet days at Berlin is partly due to changes in the velocity components at the chosen grid point for that site. The effects of the changes in the other predictor variables are negligible.

Despite the decrease in the mean number of wet days, a larger seasonal mean rainfall amount is obtained if the wet-day rainfall model with specific humidity is applied. This is mainly caused by the increased specific humidity in the future, which leads to an increase of 20 - 25% in the mean wet-day precipitation amount. For De Bilt and Berlin the increase in mean summer rainfall is smaller than that in mean winter rainfall because of the stronger decrease in the number of wet days in summer. The changes in precipitation

Table 9: Observed seasonal mean number of wet days and seasonal mean rainfall amounts (1968 - 1997) at De Bilt, Hanover and Berlin, and their estimated changes for the end of the 21st century (2070 - 2099). In model 1 atmospheric moisture is represented as specific humidity in the relationship for the wet-day precipitation amounts, and in model 2 as relative humidity and precipitable water.

	De Bilt		Hanover		Berlin	
	Winter	Summer	Winter	Summer	Winter	Summer
Wet days						
Observed number	107	86	101	97	94	81
Change in %	- 9	- 17	- 14	- 18	- 6	- 13
Rainfall amount						
Observed in mm	416	385	299	349	258	324
Change in %						
• Model 1	+ 13	+ 6	0	+ 1	+ 19	+ 3
• Model 2	+ 3	- 9	- 9	- 10	+ 7	- 9
• ECHAM4/OPYC3	+ 10	-11	+ 12	- 7	+ 12	- 13

are quite different if the alternative wet-day rainfall model with precipitable water and relative humidity is applied. Then, there is no longer a general increase in the seasonal mean rainfall amounts. Precipitable water and relative humidity are positively correlated (Section 4.2). Each individual moisture variable then contributes less to the wet-day precipitation amounts than specific humidity in the other model. The smaller effect of the change in precipitable water than that in specific humidity and the fact that there is only little change in relative humidity leads to a smaller change in the mean wet-day precipitation amounts. The decrease in vorticity has a negative impact on the mean wet-day rainfall amounts but this impact is much smaller than that of the increase in specific humidity, in particular for De Bilt and Berlin. The perturbations of the other predictor variables were not large enough to cause a significant change in seasonal mean precipitation.

Especially the expected changes in the seasonal mean rainfall amounts from the statistical model with precipitable water and relative humidity are comparable with those from the ECHAM4/OPYC3 output. The only exception is the change in mean winter rainfall at Hanover. This can be attributed to the strong vorticity effect on the estimated change from the statistical model.

5.4 Changes for Berne

The results for the seasonal mean number of wet days and the seasonal mean rainfall amounts at Berne are summarised in Table 10. Like for the Dutch and German sites the mean number of wet days decreases and the largest decrease occurs in the summer half of the year. This can again be attributed to the increased 1000-500 hPa thickness. Vorticity is for Berne less important than for the three northern sites. However, the changes in the velocity components strengthen the decrease in the number of wet days for the summer half of the year. The other perturbed predictors have very little influence on the change in the number of wet days.

The mean winter rainfall amount increases because of the larger atmospheric water content in the future climate. For the summer half of the year the effect of the increased specific humidity is not strong enough to compensate the decrease in the number of wet days. Moreover, the changes in the velocity components have an adverse effect on the mean wet-day rainfall amounts in summer (Buishand and Beckmann, 2000). For the same reason as for the three northern sites, the results for the wet-day rainfall model with specific humidity differ from those for the model with precipitable water and relative humidity.

The changes in the simulated mean precipitation amounts of the ECHAM4/OPYC3 model are approximately in the same range as those derived from the statistical models. For the summer months June to August (JJA), Heimann and Sept (2000) found a considerable decrease in precipitation in the area around Berne using a statistical-dynamical downscaling method and the same GCM scenario.

Table 10: Observed seasonal mean number of wet days and seasonal mean rainfall amounts (1968 - 1997) at Berne, and their estimated changes for the end of the 21st century (2070 - 2099). In model 1 atmospheric moisture is represented as specific humidity in the relationship for the wet-day precipitation amounts, and in model 2 as relative humidity and precipitable water.

	Winter	Summer
Wet days		
Observed number	82	87
Change in %	- 9	- 21
Rainfall amount		
Observed in mm	417	623
Change in %		
• Model 1	+ 10	- 3
• Model 2	+ 3	- 7
• ECHAM4/OPYC3	+ 4	0

5.5 Changes for Salto de Bolarque

The last climate-change application in this section pertains to the Spanish station Salto de Bolarque. For this station only the wet-day rainfall model with specific humidity was considered to calculate the changes in the seasonal mean rainfall amounts. The results are presented in Table 11. The relative decrease in the number of wet days is larger than that at De Bilt, Hanover, Berlin and Berne. Due to the smaller rainfall frequency at Salto de Bolarque, the absolute decrease in the number of wet days is not much different from that at the other stations. Furthermore, the increased 1000-500 hPa thickness is not the only factor that causes the decrease in the number of wet days. There is also a substantial contribution from a weakening of the westerly flow in the future climate. These effects are somewhat counteracted by the change in the southerly flow. The stronger decrease of the summer rainfall probability is mainly due to the larger increase in thickness in summer and the steeper slope of the predictor function at large values of the thickness.

The rainfall occurrence model for Salto de Bolarque contains both the 1000-500 hPa thickness and the 850 hPa geopotential height. The correlation coefficient between these two variables is 0.52. The higher mean temperatures in the future climate do not only lead to an increase in *thick* but also in h_{850}. However, the increase in the latter has only a modest effect on the number of wet days. The impact of *thick* on the number of wet days becomes somewhat stronger if h_{850} is dropped from the model. For the winter half of the year this reduced model gives the same change in the number of wet days, whereas for the summer half of the year a somewhat larger decrease is found (42% instead of 39%).

Table 11: Observed seasonal mean number of wet days and seasonal mean rainfall amounts (1968 - 1997) at Salto de Bolarque, and their estimated changes for the end of the 21^{st} century (2070 - 2099). Atmospheric moisture is represented as specific humidity in the statistical model for the wet-day precipitation amounts.

	Winter	Summer
Wet days		
Observed number	49	33
Change in %	- 24	- 39
Rainfall amount		
Observed in mm	249	204
Change in %		
• Statistical model	- 15	- 21
• ECHAM4/OPYC3	- 15	- 13

Table 11 shows a decrease in the seasonal mean rainfall amounts at Salto de Bolarque. This decrease is the result of the strong decrease in the rainfall probability. The relative decrease in the seasonal mean rainfall amounts is smaller than that in the number of wet days because of the larger wet-day precipitation amounts in the future climate due to the increased specific humidity. There is further a negative effect of the increased geopotential height at 850 hPa on the wet-day precipitation amounts. This effect is questionable, however, because the change in h_{850} is more related to changes in the temperature than to changes in the atmospheric circulation, cf. ZORITA and VON STORCH (1999). Removing h_{850} from the model for wet-day rainfall does, however, not lead to larger wet-day precipitation amounts in the future climate, because the regression coefficient for specific humidity is smaller in the reduced model. The contribution of the vorticity to the change in seasonal mean rainfall amounts at Salto de Bolarque can be neglected.

The large-scale climate model also predicts a considerable decrease in the seasonal mean rainfall amounts at the grid point nearest to Salto de Bolarque. For the winter months this result corresponds exactly with the estimate from the downscaling technique.

6 Daily precipitation scenarios for Berne

Two different methods of scenario construction are presented for the Swiss station Berne:

- Perturbation of the observed daily precipitation record at Berne
- Conditional simulation of daily precipitation using a Monte Carlo method

The two methods are compared by computing the 90^{th} percentile of the distributions of the N-day annual maximum precipitation amounts ('10-year event'). Only one model for the wet-day rainfall amount is considered, namely that with specific humidity.

6.1 Time series perturbation

Daily precipitation scenarios have often been obtained by multiplying the observed precipitation amounts by a constant factor, for example, the ratio between the seasonal means of the future and present climate in a GCM experiment. The mean number of wet days then remains unaltered in the future climate. In order to produce a scenario with less wet days, KLEIN TANK and BUISHAND (1995) suggested randomly replacing wet days by dry days using the probabilities of precipitation from the logistic model. Let ΔM denote the expected decrease in the number of wet days from Equation 6. To achieve such a decrease under perturbed climate conditions, a wet day τ is replaced by a dry day with probability Π_τ, given by:

$$\Pi_\tau = \Delta M \cdot \frac{\left(1 - P_\tau^*\right)^\theta}{\sum_\tau \left(1 - P_\tau^*\right)^\theta} \tag{12}$$

where P_τ^* is the probability of rain in the future climate. Here, the summation is over all wet days in the observed record. The parameter θ was taken as 1 by KLEIN TANK and BUISHAND (1995). The probability Π_τ then increases linearly with the probability $1-P_\tau^*$ that day τ is dry. To ensure that $\Pi_\tau < 1$ for all τ, ΔM may not exceed $\sum_\tau \left(1 - P_\tau^*\right)^\theta$.

Future daily precipitation amounts R_{ft} on the remaining wet days are obtained by scaling the observed daily rainfall values R_{ot} by the ratio between the expected values in the future and present climate:

$$R_{ft} = R_{ot} \cdot \frac{R_t^*}{R_t} \qquad \text{with} \qquad \frac{R_t^*}{R_t} = \exp\left[\sum_{i=1}^{p} \left(f_i\left(x_i^*\right) - f_i\left(x_i\right)\right)\right] \tag{13}$$

Because of the log-link function in the model for the wet-day rainfall amounts the adjustment only depends on the changes in the values of the $f_i(x_i)$. The scaling factor reduces to a constant if these functions are linear.

Although the method is able to reproduce the changes in the mean number of wet days for Berne in Table 10, this is not necessarily true for the changes in mean seasonal or annual precipitation. The latter are influenced by the choice of θ in Equation 12. For Berne, a reasonable agreement with the estimated changes in the mean rainfall amounts could be achieved by setting θ equal to 1 for the winter half of the year and 0.5 for the summer half of the year.

6.2 Conditional simulation

Here the term conditional simulation implies that daily precipitation is generated conditional on the values of the predictor variables. This can be undertaken for both present and future climates.

For present climate conditions, a synthetic sequence can be generated as follows. For each day t a binary variable Y_t is drawn, taking the value 0 (dry day) with probability $1 - P_t$ and 1 (wet day) with probability P_t. In the latter case a random amount is generated as:

$$R_{sim,t} = \omega_t R_t \tag{14}$$

where R_t is the expected rainfall amount for day t from Equation 3 and ω_t is a positive random variable with mean 1, for example,

$$\omega_t = \frac{1}{\kappa}\text{Gam}(\kappa) \tag{15}$$

with $\text{Gam}(\kappa)$ a standard gamma variable with shape prameter κ. The gamma distribution has often been used in unconditional simulations of daily precipitation (WILKS and WILBY, 1999). It is also the standard distribution in generalised additive models with constant coefficient of variation CV. The following relationship exists between κ and CV:

$$\kappa = \frac{1}{CV} \tag{16}$$

It is therefore possible to preserve both the mean value and CV of the precipitation amounts.

The simulation of future climates is undertaken in the same way. The expected changes in the mean number of wet days and the seasonal mean rainfall amounts are the same as those derived from Equations 6 and 7. However, the seasonal cycles of the monthly means of the generated sequences show the same discrepancies as that observed in Figures 6 and 7 for Berne. In addition, extreme-value properties may not be preserved, as demonstrated in the following section.

6.3 Evaluation based on N-day annual maximum precipitation amounts

Percentiles of the distribution of annual maximum precipitation amounts over different time-intervals have often been used in hydrological design. From an engineering point of view, it is therefore more meaningful to study these percentiles than the frequency distribution or exceedance frequencies of the daily rainfall amounts. Here, the 90^{th} percentile of the distribution of the N-day annual maximum rainfall amounts is considered for $N =$ 1, 3, 10 and 30. In the hydrological literature, this percentile is usually denoted as the 10-year event because it is exceeded on average once in 10 years. The 10-year events were derived by fitting Gumbel distributions to the annual maxima using probability-weighted moments (LANDWEHR *et al.*, 1979).

Table 12: Estimates of the 90^{th} percentile (10-year event) of the N-day annual maximum precipitation amounts at Berne from the observed data and simulated sequences (average of 20 runs for each simulation method).

Method of generating ω_t	CV	10-year event in mm			
		$N = 1$	$N = 3$	$N = 10$	$N = 30$
HAY *et al.* (1991)	1.29	128	184	269	381
Gamma	1.16	96	133	218	328
Gamma	variable	70	106	188	299
Observed	1.16	65	99	149	258

Table 13: Percentage changes in the 10-year event of the N-day rainfall amounts at Berne for different scenarios for the end of the 21^{st} century (averages of 20 runs for each simulation method and ω_t generated from a gamma distribution).

Scenario	Change in 10-year event in %			
	$N = 1$	$N = 3$	$N = 10$	$N = 30$
Perturbed record	15	16	11	9
Simulated sequences with constant CV	15	14	12	11
Simulated sequences with variable CV	6	7	7	8

Table 14: Percentage changes in the 10-year event of the N-day rainfall amounts at Berne for different scenarios for the end of 21^{st} century (averages of 20 runs for each simulation method and ω_t generated from a gamma distribution) based on statistical models without baroclinicity as predictor variable.

Scenario	Change in 10-year event in %			
	$N = 1$	$N = 3$	$N = 10$	$N = 30$
Perturbed record	12	9	9	6
Simulated sequences with constant CV	13	12	8	5
Simulated sequences with variable CV	10	12	6	3

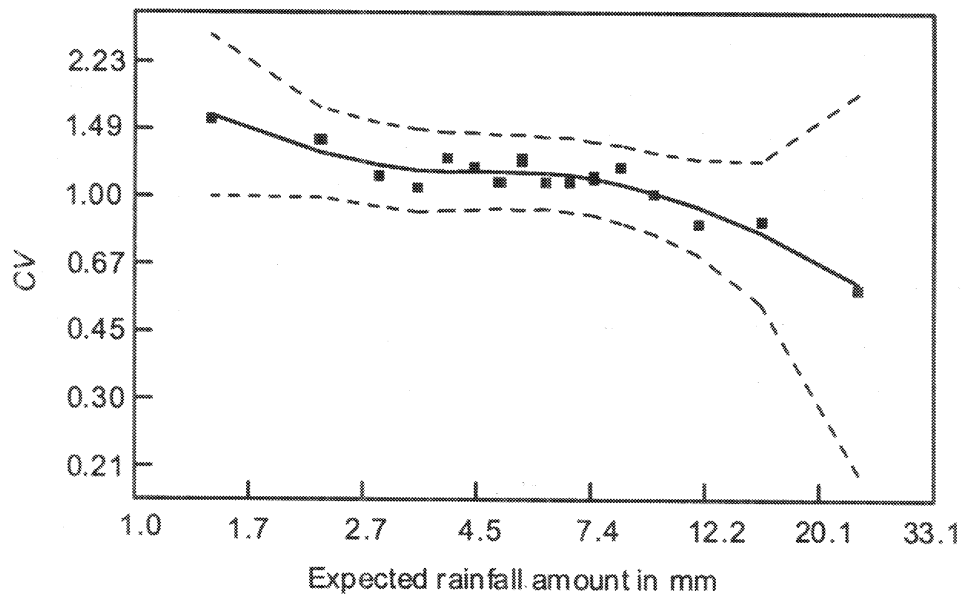

Figure 13: Coefficient of variation CV of the wet-day rainfall amounts at Berne as a function of the expected rainfall amount from Equation 3. Each dot represents an estimate of CV based on about 350 values, except for the last class where only 50 values were considered. The solid line is obtained using a locally weighted running line smoother with a span of 0.5; the dashed lines mark pointwise 2 x standard-error bands.

Table 12 compares the 10-year events of simulated sequences for present conditions with those of the observed data at Berne. In the method of HAY *et al.* (1991) ω_t is generated as the product of a uniformly distributed random number between 0 and 2 and a standard exponential variable. The extremes are too large in the resulting scenario, mainly because CV is not preserved (CV^2 is always 5/3 in that method). The 10-year events are still too large if ω_t is generated from a gamma distribution with the same CV as that of the observed data. The problem is that CV decreases with the expected wet-day rainfall amount R_τ (see Figure 13). A marked improvement in the reproduction of the 10-year events is achieved by incorporating this dependence in the simulation algorithm. For this application the smooth curve in Figure 13 was approximated by a piecewise linear function like the non-linear functions in the models for P and R.

Table 13 presents the estimated changes in the 10-year events between the periods 1968 - 1997 and 2070 - 2099. All scenarios show an increase in this extreme-value characteristic. This is because the effect of the higher wet-day rainfall amounts resulting from the increase in specific humidity dominates that of the decrease in the number of wet days. For time series perturbation, the changes are comparable with those for conditional simulation with a constant CV. Smaller changes are found if the decrease of CV with R_τ is taken into a account, in particular for $N = 1$ and $N = 3$.

Both the reproduction of the 10-year events and their changes under perturbed climate conditions are thus influenced by the representation of the coefficient of variation CV. The crucial aspect is the change in CV for large values of R_r and Figure 13 shows that this change is quite uncertain. The estimated CV for days with $R_r > 20\,\mathrm{mm}$ was 0.58, which is considerably lower than the values for other days with large expected rainfall amounts. The cause of this drop in CV is that the expected rainfall amount from the model is in fact too large (see Section 7). A crude way to get rid of this model deficiency is to remove the baroclinicity from the model for wet-day rainfall amounts. The CV then decreases less rapidly for large values of R_r than that observed in Figure 13. Table 14 presents the changes in the 10-year events for the case that there is no baroclinicity in the statistical models for rainfall occurrence and the wet-day rainfall amounts. The differences between the scenarios are smaller than those shown in Table 13.

7 Model deficiencies

In Section 6 it was noticed that the CV for Berne depends on the expected wet-day precipitation amount R_τ. It decreased rapidly for large values of R_τ. A similar drop was also found for Neuchâtel, De Bilt, Hanover and Berlin. For this reason the model for the wet-day rainfall amounts is explored further in this section.

In Figure 14 the observed wet-day rainfall amounts at De Bilt, Hanover and Berlin are plotted against their expected values from the fitted model with specific humidity as moisture variable and piecewise linear approximations to non-linear predictor functions. There is a good correspondence between the observed and predicted values except for the last class in which the models strongly overpredict the true wet-day rainfall amounts. The largest bias of 5.7 mm is found for Hanover. This bias also appears if the loess smoother is used to estimate the non-linear functions $f_i(x_i)$ instead of piecewise linear approximations. It further emerged that the bias could not be removed by fitting separate models to the data for the winter and summer halves of the year (Section 4.3.2).

A wrong choice of the link function in the model for the wet-day rainfall amounts could be another reason for the observed discrepancy in Figure 14. The use of a log link function implies that $\eta = \ln R$ is additive in the predictor variables. A more general link function

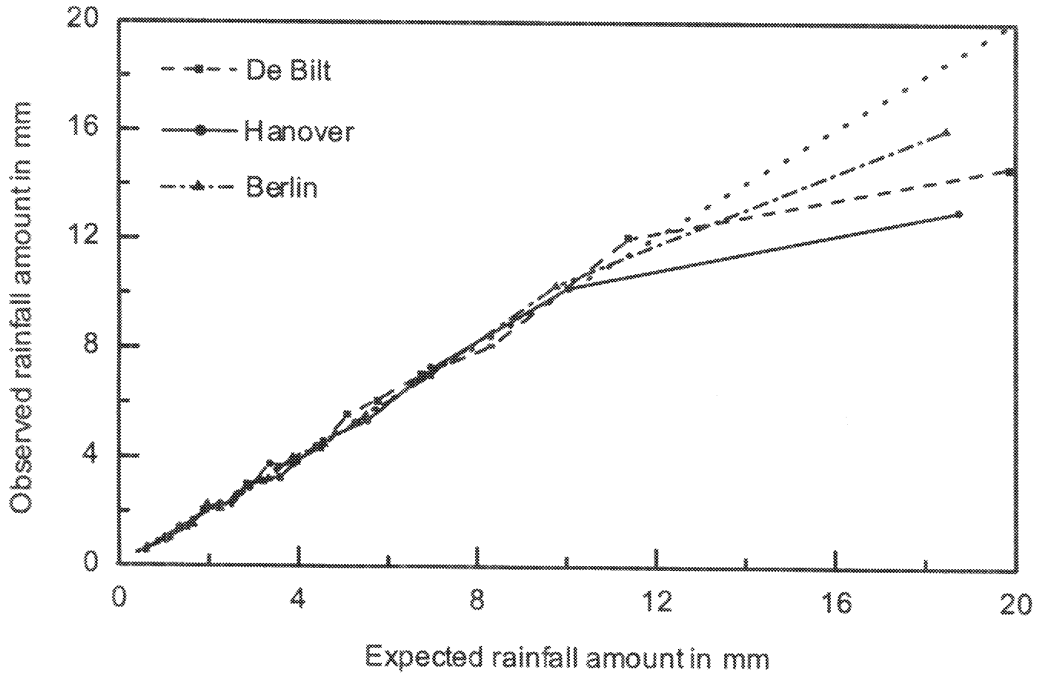

Figure 14: Observed mean wet-day rainfall amounts versus the expected wet-day rainfall amounts from fitted relationships for De Bilt, Hanover and Berlin. The fitted relationships contain specific humidity as moisture variable. Each mean value is based on about 350 values, except for the last class, where only about 50 values were considered.

is the power link function:

$$\eta = \frac{R^\lambda - 1}{\lambda} \tag{17}$$

The log link corresponds to the limit case $\lambda = 0$. A simple test of whether the log link is acceptable can be obtained from a Taylor series expansion of η about $\lambda = 0$ (PREGIBON, 1980; MCCULLAGH and NELDER, 1989):

$$\eta \approx \ln R + \frac{1}{2}\lambda (\ln R)^2 \tag{18}$$

Equation 18 can be rewritten as:

$$\begin{aligned} \ln R &\approx \eta - \frac{1}{2}\lambda (\ln R)^2 \\ &= \sum_{i=1}^{p} f_i(x_i) - \frac{1}{2}\lambda (\ln R)^2 \end{aligned} \tag{19}$$

Given the first estimate R_1 using the log link function, the analysis is repeated with the extra predictor variable $(\ln R_1)^2/2$. For Hanover it turned out that the contribution of this extra predictor is not significant, which supports the use of the log link function.

In the generalised additive model with log link function the contribution of the ith predictor to $\ln R$ is $f_i(x_i)$, independent of the values of other predictors. Interaction occurs if the value of a predictor determines the effect of other predictors. The fact that possible interactions are ignored seems to be the main reason for the overestimation of the mean wet-day precipitation amounts in extreme situations. For Hanover significant interactions were discovered by dividing the data into three vorticity categories:

1. $\zeta_p < -0.01$ kHz
2. $-0.01 \leq \zeta_p < +0.01$ kHz
3. $\zeta_p \geq +0.01$ kHz

Each class contains about the same amount of data. Table 15 shows that the regression coefficients for the u-velocity and baroclinicity in the third vorticity class strongly deviate from those in the other two classes. The use of separate models for the three vorticity

Table 15: Estimated coefficients a_u and a_b for the u-velocity and baroclinicity in wet-day rainfall amount models with specific humidity for Hanover. Separate fits for three different vorticity classes.

Vorticity class	a_u	a_b
1	0.038	36.13
2	0.041	41.16
3	0.023	52.81

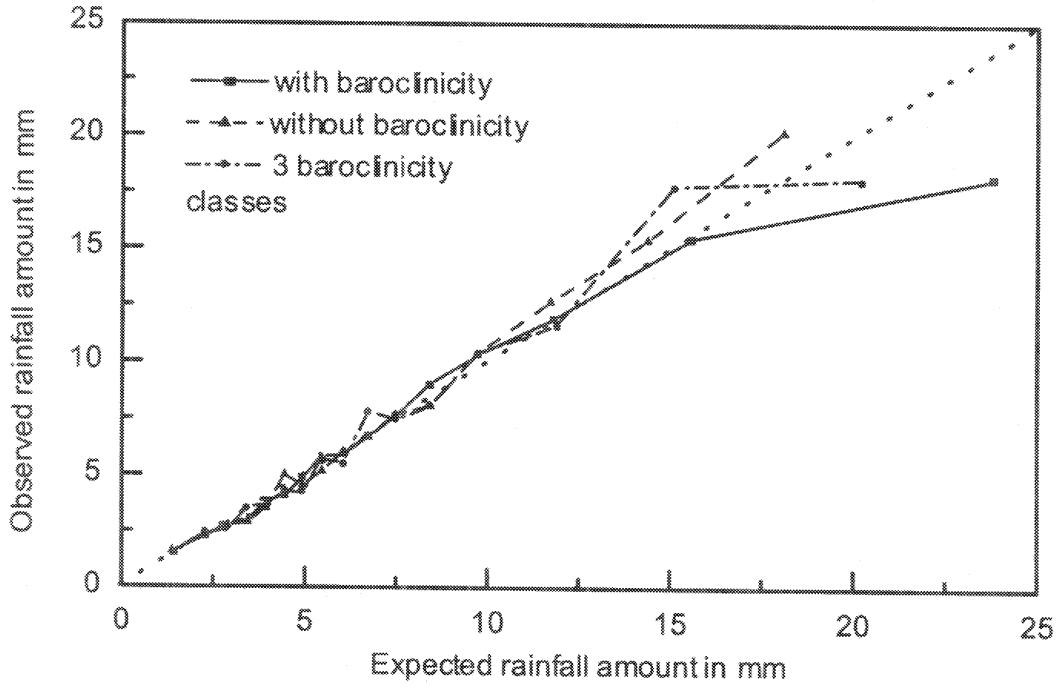

Figure 15: Observed mean wet-day rainfall amounts versus the expected wet-day rainfall amounts from three different fitted relationships for Berne. All fitted relationships contain specific humidity as moisture variable. Each mean value is based on about 350 values, except for the last class, where only 50 values were considered.

classes reduces the bias in extreme wet-day rainfall at Hanover from 5.7 to 2.4 mm.

Incorporation of interaction terms is a rather tedious job and makes the model quite complex. For Hanover, the interaction terms relate to vorticity that was poorly reproduced by the ECHAM4/OPYC3 model. Deletion of vorticity could therefore be an option. The bias in the upper wet-day rainfall amount class is then not more than 0.5 mm.

For Berne a large overprediction of the observed wet-day rainfall amount is found for days with $R_r > 20$ mm (Figure 15). Closer examination of these cases revealed that large residuals are quite often accompanied by large values of $f(b)$. The bias disappears if the baroclinicity is removed from the model.

In order to investigate possible interactions, three baroclinicity classes were distinguished:

1. $b < 0.006$ K/km
2. $0.006 \leq b < 0.010$ K/km
3. $b \geq 0.010$ K/km

with about the same number of days in each class. It emerged that the regression coefficients for the u- and v-velocities, the relative vorticity and the 1000 hPa height are

not constant over these three classes. Figure 15 shows that the bias in the upper rainfall amount class decreases if separate models for each baroclinicity class are used. It also provides a significantly better fit to the data than a single additive model (appoximate F-test).

8 Discussion and conclusion

Generalised additive models are flexible tools to explore the dependence of daily rainfall on other meteorological variables. Within this modelling framework it is convenient to analyse rainfall occurrence and the wet-day rainfall amounts separately. In this report daily rainfall at eleven sites in the Netherlands, North Germany, Austria, Switzerland and Spain was linked to atmospheric moisture and circulation variables from the NCEP reanalysis. For rainfall occurrence the relative humidity at 700 hPa is always one of the most significant predictors, whereas the specific humidity at 700 hPa and the precipitable water are powerful predictors for the wet-day rainfall amounts. This is in contrast with MURPHY (2000) who fitted linear regression equations to monthly rainfall amounts at 976 sites in Europe. Atmospheric moisture was rarely selected in that work.

Various circulation variables were considered. Their importance is region specific. Vorticity is one of the most significant predictors for stations in the Netherlands, North Germany and Spain but not for Vienna and the Swiss stations. A rather low influence of vorticity on daily rainfall has also been observed for a number of regions in the USA (WILBY and WIGLEY, 2000). Baroclinicity is a powerful predictor for Vienna and the Swiss sites but is not significant for the Spanish sites. The u- and v-velocity components are often important predictor variables. Wet-day rainfall at the Spanish sites could, however, not be linked to a velocity component. Circulation variables having a relatively strong influence on rainfall occurrence at a particular site often show a similar strong influence on the wet-day rainfall amounts.

Small discrepancies in the modelled seasonal cycles of rainfall occurrence and the mean wet-day rainfall amounts were observed. For five of the eleven stations estimates of the changes in the seasonal mean rainfall amounts for the end of the 21^{st} century were presented using the fitted statistical models with perturbed predictor variables. The perturbations were derived from the output of the time-dependent GHG experiment with the ECHAM4/OPYC3 model. In most cases the estimated changes agree well with those in the simulated rainfall of the climate model. The estimated change in seasonal mean rainfall from the statistical downscaling relationships is strongly determined by two factors: (i) the increase in the 1000-500 hPa thickness leading to a decrease in the number of wet days, and (ii) an increase in the specific humidity and precipitable water leading to larger wet-day rainfall amounts. The magnitudes of the changes in seasonal mean rainfall are sensitive to the moisture variables entered into the model for the wet-day precipitation amounts. The comparison with the ECHAM4/OPYC3 simulation does not provide sufficient evidence to make a choice between the inclusion of specific humidity or precipitable water and relative humidity. The fact that a model with precipitable water and relative humidity explains a somewhat larger proportion of the variance of the wet-day rainfall amounts is of little interest for daily precipitation scenario development. More comparisons with GCM simulations are needed to get a better insight into the moisture variables to use.

A decrease in the number of wet days at mid latitudes has also been observed in other GCM simulations with enhanced greenhouse gas concentrations (CUBASCH *et al.*, 1995;

HENNESSY, *et al.*, 1997). It has been attributed to a larger fractional contribution of convective precipitation to the total seasonal amounts. However, decreases in the number of wet days without an increase in the convective fraction have also been observed (JONES *et al.*, 1997). It is further not clear whether the change in the number of wet days is adequately described by the increased 1000-500 hPa thickness in the future climate.

A large decrease in the relative vorticity at the considered grid point for Hanover strongly influences the change in mean winter precipitation at this site. However, little confidence could be put on this decrease because of the large bias in the relative vorticity in the ECHAM4/OPYC3 simulation, resulting from a poor reproduction of the gradients in the mean sea level pressure at neighbouring grid points. Circulation variables should be calculated over a larger domain to cope with such biases. It is, however, not obvious how large this domain must be. A larger domain may reduce the performance of the regression models and the biases in the simulated circulation variables are model dependent. For Berlin and Salto de Bolarque the changes in the velocity components substantially contribute to the change in the mean number of wet days. The other changes in circulation variables have little influence on the mean number of wet days and the rainfall amounts. Moreover, these changes are often not significant and small compared with the bias in the climate model output.

A limitation of the presented downscaling relationships is that they are unable to take the effects of potential changes in the vertical atmospheric stability into account. A humid unstable vertical temperature gradient is a necessary condition for development of atmospheric convection and convective precipitation. For De Bilt the use of a number of stability indices was examined, among which HANSSEN's (1965) thickness index and the K-index (KARL *et al.*, 1990, MURPHY, 2000). The latter shows the strongest correlation with daily precipitation. However, the K-index is correlated with the moisture variables in this study, because it contains the temperature and the dew point temperature at 700 hPa. It was therefore not possible to enter both the K-index and relative humidity into the rainfall occurrence model. The relative humidity appears to be the better predictor. It was also not successful to use the K-index together with the specific humidity in the wet-day rainfall amount model. The Hanssen index is only marginally significant. One difficulty is that in very unstable situations there can be no or little precipitation at the target point and much precipitation somewhere else. The Hanssen index has a much better skill if one is interested in the largest rainfall amount in a region rather than local rainfall.

Another problem is the inclusion of the effect of forced convection in mountainous regions. For a single mountain range SINCLAIR (1994) estimated orographically forced vertical motion w as the scalar product of the low-level wind in coarse resolution analyses from the European Centre of Medium-Range Weather Forecasts (ECMWF) and the height gradient of the topography. This method is not applicable to areas with a complex structured orography like that studied in Switzerland. Wind fields of high spatial resolution may then be needed, which are neither in the NCEP reanalysis available nor in the GCM data.

Besides the choice of predictor variables the structure of the statistical models could be questioned, in particular that for the wet-day rainfall amounts. For Hanover and Berne it

was shown that the assumption of additivity of predictor effects leads to a poor reproduction of extreme daily rainfall. Moreover, stochastic simulation of daily precipitation for Berne showed that a good description of the coefficient of variation is necessary if extreme-value properties are of interest. In the case of a constant, or almost constant, coefficient of variation a simple time series perturbation technique can be a useful alternative to stochastic time series simulation.

Appendix

Deviance statistics

The deviance is a measure of the discrepancy of a fit. It is baed on the likelihood function. For the standard linear regression model the deviance reduces to the residual sum of squares. This appendix sketches the use of the deviance in model selection. A detailed discussion of the deviance is given in McCullagh and Nelder (1989). The approximate χ^2- and F-tests for generalised additive models are presented in Hastie and Tibshirani (1990).

Using the same notation as in Section 2.1, the deviance for the rainfall occurrence model reads:

$$D = 2\sum_{t=1}^{N}\left[y_t \ln\left(\frac{y_t}{P_t}\right) + (1-y_t)\ln\left(\frac{1-y_t}{1-P_t}\right)\right] \tag{A1}$$

For the wet-day rainfall amount model, the deviance is given by:

$$D = 2\sum_{\tau=1}^{N_{wo}}\left[\frac{R_{o\tau}-R_\tau}{R_\tau} - \ln\left(\frac{R_{o\tau}}{R_\tau}\right)\right] \tag{A2}$$

where $R_{o\tau}$ and R_τ are the observed and expected rainfall amount for the τth wet day, respectively, and N_{wo} is the observed number of wet days. Note that $D = 0$ for a perfect fit, i.e. $R_\tau = R_{o\tau}$ for all τ. The better the fit, the smaller the deviance will be.

Akaike's selection criterion is based on the statistic:

$$AIC = D + 2\nu\phi \tag{A3}$$

where ϕ is the dispersion parameter ($\phi = 1$ for the rainfall occurrence model and $\phi = CV^2$ for the rainfall amount model) and ν is the number of degrees of freedom in the fit. Inclusion of a linear predictor results in one extra degree of freedom and using a running-line smoother with a span of 0.5 corresponds to about three degrees of freedom for each predictor (Hastie and Tibshirani, 1990). The best model is the one with the lowest AIC statistic rather than that with the smallest deviance. The penalty term $2\nu\phi$ prevents from overfitting.

For the inclusion of an extra parameter or the use of a non-linear predictor, the approximate χ^2- and F-tests are more severe criteria than the AIC statistic (Hastie and Tibshirani, 1990). These tests are based on the decrease ΔD in the deviance:

$$\Delta D = D_1 - D_2 \tag{A4}$$

where D_1 is the deviance for the smaller model (with ν_1 degrees of freedom) and D_2 the deviance for the larger model (with ν_2 degrees of freedom). Under the null hypothesis that the smaller model is correct, the distribution of $\Delta D/\phi$ can be approximated by the χ^2-distribution with $\nu_2 - \nu_1$ degrees of freedom. This test assumes that the dispersion

parameter ϕ is known (rainfall occurrence model). The F-distribution has to be used if ϕ is replaced by an estimate (model for wet-day rainfall). Because of the large sample sizes in this study, the approximate F-test is almost identical to the χ^2-test.

Seasonal variations of relationships can also be tested with the approximate χ^2- and F-tests. For the example in Section 4.3.2 the test compares the deviance D_1 for the constant model with the deviance D_2 for the separate fits for the winter and summer halves of the year. The latter is obtained as the sum of the deviances for the two seasons.

Acknowledgements

The authors thank H. Thiemann and M. Lautenschlager of the Deutsche Klimarechenzentrum (Hamburg) for their efforts to make the ECHAM4/OPYC3 data available. They are also grateful to the Climatic Research Unit, University of East Anglia (Norwich) for extracting a subset of the NCEP reanalysis data for a European window. H. van de Brink prepared Figure 3. Daily rainfall data were kindly provided by the Deutsche Wetterdienst (Offenbach) for Hamburg, Hanover and Berlin, by the Swiss Meteorological Institute - MeteoSwiss (Zürich) for Berne, Neuchâtel and Payerne, by the Zentralanstalt für Meteorologie und Geodynamik (Vienna) for Vienna and by the Instituto Nacional de Meteorología (Madrid) for Salto de Bolarque and Munera.

The WRINCLE project was in part supported by the EU Environment and Climate Research Programme (contract: ENV4-CT97-0452, Climate and Natural Hazards).

References

Arpe, K. and E. Roeckner, 1999. Simulation of the hydrological cycle over Europe: model validation and impacts of increasing greenhouse gases, Advances in Water Resources, **23**, 105-119.

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

Biau, G., E. Zorita, H. von Storch and H. Wackernagel, 1999. Estimation of precipitation by kriging in the EOF space, Journal of Climate, **12**, 1070-1085.

Brandsma, T. and T.A. Buishand, 1996. KNMI contribution to the European project POPSICLE, Technical Report TR-194, Royal Netherlands Meteorological Institute, De Bilt, 44 pp.

Brandsma, T. and T.A. Buishand, 1997. Statistical linkage of daily precipitation in Switzerland to atmospheric circulation and temperature, Journal of Hydrology, **198**, 98-123.

Buishand, T.A. and B.-R. Beckmann, 2000. Development of daily precipitation scenarios at KNMI, In: (Beersma, J., Agnew, M., Viner, D. and Hulme, M., eds) Climate scenarios for water-related and coastal impacts, ECLAT-2 Workshop Report No.3 KNMI, the Netherlands, 10-12 May 2000, CRU, Norwich, UK, pp. 79-91.

Buishand, T.A. and A.M.G. Klein Tank, 1996. Regression model for generating time series of daily precipitation amounts for climate change impact studies, Stochastic Hydrology and Hydraulics, **10**, 87-106.

Cavazos, T., 1999. Large-scale circulation anomalies conductive to extreme precipitation events and derivation of daily rainfall in northeastern Mexico and southeastern Texas, Journal of Climate, **12**, 1506-1523.

Chambers, J.M. and T.J. Hastie, 1993. Statistical Models in S, Chapman and Hall, London, 608 pp.

Charles, S.P., B.C Bates, P.H. Whetton and J.P. Hughes, 1999. Validation of downscaling models for changed climate conditions: case study for southwestern Australia, Climate Research, **12**, 1-14.

Conway, D., and P.D. Jones, 1998. The use of weather types and air flow indicies for GCM downscaling, Journal of Hydrology, **212-213**, 348-361.

Conway, D. R.L. Wilby and P.D. Jones, 1996. Precipitation and air flow indices over the British Isles, Climate Research, **7**, 169-183.

Corte-Real, J., H. Xu and B. Qian, 1999. A weather generator for obtaining daily precipitation scenarios based on circulation patterns, Climate Research, **12**, 61-75.

Corte-Real, J., X. Zhang and X. Wang, 1995. Downscaling GCM information to regional scales: a non-parametric multivariate regression approach, Climate Dynamics, **11**, 413-424.

Crane, R.G. and B.C. Hewitson, 1998. Doubled CO_2 precipitation changes for the Susquehanna basin: downscaling from the Genesis general circulation model, International Journal of Climatology, **18**, 65-76.

Cubasch, U., J. Waszkewitz, G.C. Hegerl and J. Perlwitz, 1995. Regional climate changes as simulated in time-slice experiments, Climate Change, **31**, 273-304.

Hanssen, A.W., 1965. An objective method for forecasting thunderstorms in the Netherlands, Journal of Applied Meteorology, **4**, 172-177.

Hastie, T.J. and R.J. Tibshirani, 1990. Generalized Additive Models, Chapman and Hall, London, 335 pp.

Hay, L.E., G.J. McCabe, D.M. Wolock and M.A. Ayers, 1991. Simulation of precipitation by weather type analysis, Water Resources Research, **27**, 493-501.

Heimann, D. and V. Sept, 2000. Climate change estimates of summer temperature and precipitation in the Alpine region, Theoretical and Applied Climatology, **66**, 1-12.

Hennessy, K.J., J.M. Gregory and J.F.B. Mitchell, 1997. Changes in daily precipitation under enhanced greenhouse conditions, Climate Dynamics, **13**, 667-680.

Hewitson, B.C. and R.G. Crane, 1996. Climate downscaling: techniques and application, Climate Research, **7**, 85-95.

Jones, R.G., J.M. Murphy, M. Noguer and A.B. Keen, 1997. Simulation of climate change over Europe using a nested regional-climate model. II: comparison of driving and regional model responses to doubling of carbon dioxide, Quarterly Journal of the Royal Meteorological Society, **123**, 265-292.

Kalnay, E., M. Kanamitsu, R. Kistler, W. Collins, D. Deaven, L. Gandin, M. Iredell, S. Saha, G. White, J. Woollen, Y. Zhu, M. Chelliah, W. Ebisuzaki, W. Higgens, J. Janowiak, K.C. Mo, C. Ropelewski, J. Wang, A. Leetmaa, R. Reynolds, R. Jenne and D. Joseph, 1996. The NCEP/NCAR 40-year reanalysis project, Bulletin of the American Meteorological Society, **77**,437-471.

Karl, T.R., W.-C. Wang, M.E. Schlesinger, R.W. Knight and D. Portman, 1990. A method of relating general circulation model simulated climate to observed local climate, Part I: seasonal statistics, Journal of Climate, **3**, 1053-1079.

Kilsby, C.G., P.S.P. Cowpertwait, P.E. O'Connell and P.D. Jones, 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables, International Journal of Climatology, **18**, 523-539.

Klein Tank, A.M.G. and T.A. Buishand, 1995. Transformation of precipitation time series for climate change impact studies, Scientific Report WR 95-01, Royal Netherlands Meteorological Institute, De Bilt, 63 pp.

Knippertz, P., U. Ulbrich und P. Speth, 2000. Changing cyclones and surface wind speeds over the North Atlantic and Europe in a transient GHG experiment, Climate Research, **15**, 109-122.

Können, G.P. (ed.), 1983. Het weer in Nederland, Thieme, Zutphen, 143 pp.

Landwehr, J.M., N.C. Matalas and J.R. Wallis, 1979. Probability weighted moments compared with some traditional techniques in estimating Gumbel parameters and quantiles, Water Resources Research, **15**, 1055-1064.

Matyasovszky, I., I. Bogárdi, A. Bárdossy and L. Duckstein, 1993. Space-time precipitation reflecting climate change, Hydrological Sciences Journal, **38**, 539-558.

McCullagh, P. and J.A. Nelder, 1989. Generalized Linear Models, 2nd edition, Chapman and Hall, London, 511 pp.

Murphy, J., 2000. Predictions of climate change over Europe using statistical and dynamical downscaling techniques, International Journal of Climatology, **20**, 489-501.

Pregibon, D., 1980. Goodness of link tests for generalized linear models, Applied Statistics, **29**, 15-24.

Roeckner, E., L. Bengtsson, J. Feichter, J. Lelieveld and H. Rodhe, 1999. Transient climate change simulations with a coupled atmosphere-ocean GCM including the tropospheric sulfur cycle, Journal of Climate, **12**, 3004-3032.

Sinclair, M.R., 1994. A diagnostic model for estimating orographic precipitation. Journal of Applied Meteorology, **33**, 1163-1175.

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, Journal of Climate, **6**, 1161-1171.

Wilby, R.L. and T.M.L. Wigley, 1997. Downscaling general circulation model output: a review of methods and limitations, Progress in Physical Geography, **21**, 530-548.

Wilby, R.L. and T.M.L. Wigley, 2000. Precipitation predictors for downscaling: observed and general circulation model relationships, International Journal of Climatology, **20**, 641-661.

Wilby, R.L., H. Hassan and K. Hanaki, 1998. Statistical downscaling of hydrometeorological variables using general circulation model output, Journal of Hydrology, **205**, 1-19.

Wilks, D.S. and R.L. Wilby, 1999. The weather generation game: a review of stochastic weather models, Progress in Physical Geography, **23**, 329-357.

Zorita, E. and H. von Storch, 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods, Journal of Climate, **12**, 2474-2489.

Zorita, E., J.P. Hughes, D.P. Lettenmaier and H. von Storch, 1995. Stochastic characterization of regional circulation patterns for climate model diagnosis and estimation of local precipitation, Journal of Climate, 8, 1023-1042.

OVERZICHT VAN KNMI-PUBLICATIES, VERSCHENEN SEDERT 1999

KNMI-PUBLICATIE MET NUMMER

186-II Rainfall generator for the Rhine Basin: multi-site generation of weather variables by nearest-neighbour resampling / T. Brandsma a.o.
186-III Rainfall generator for the Rhine Basin: nearest-neighbour resampling of daily circulation indices and conditional generation of weather variables / Jules J. Beersma and T. Adri Buishand
186-IV Rainfall generator for the Rhine Basin: multi-site generation of weather variables for the entire drainage area / Rafal Wójcik, Jules J. Beersma and T. Adri Buishand
188 SODA workshop on chemical data assimilation: proceedings; 9-10 December 1998, KNMI, De Bilt, The Netherlands
189 Aardbevingen in Noord-Nederland in 1998: met overzichten over de periode 1986-1998 / [Afdeling SO]
190 Seismisch netwerk Noord-Nederland / [afdeling Seismologie]
191 Het KNMI-programma HISKLIM (HIStorisch KLIMaat) / Theo Brandsma, Frits Koek, Hendrik Wallbrink, Günther Können
192 Gang van zaken 1940-48 rond de 20.000 zoekgeraakte scheepsjournalen / Hendrik Wallbrink en Frits Koek

TECHNISCH RAPPORT = TECHNICAL REPORT (TR)

216 Evaluatierapport Automatisering Visuele Waarnemingen : Ontwikkeling Meestsystemen / Wiel Wauben en Hans de Jongh
217 Verificatie TAF en TREND / Hans van Bruggen
218 LEO - LSG and ECBILT coupled through OASIS: description and manual / A. Sterl
219 De invloed van de grondwaterstand, wind, temperatuur en dauwpunt op de vorming van stralingsmist: een kwantitatieve benadering / Jan Terpstra
220 Back-up modellering van windmeetmasten op luchthavens / Ilja Smits
221 PV-mixing around the tropopause in an extratropical cyclone / M. Sigmond
222 NPK-TIG oefendag 16 december 1998 / G.T. Geertsema, H. van Dorp e.a.
223 Golfhoogteverwachtingen voor de Zuidelijke Noordzee: een korte vergelijking van het ECMWF-golfmodel (EPS en operationeel), de nautische gidsverwachting, Nedwam en meteoroloog / D.H.P. Vogelezang, C.J. Kok
224 HDFg library and some HDF utilities: an extention to the NCSA HDF library user's manual & reference guide / Han The
225 The Deelen Infrasound Array: on the detection and identification of infrasound / L.G. Evers and H.W. Haak
226 2D Variational Ambiguity Removal / J.C.W. de Vries and A.C.M. Stoffelen
227 Seismo-akoestische analyse van de explosies bij S.E. Fireworks ; Enschede 13 mei 2000 / L.G. Evers en H.W. Haak
228 Evaluation of modified soil parameterization in the ECMWF landsurface scheme / R.J.M. IJpelaar
229 Evaluation of humidity and temperature measurements of Vaisala's HMP243 plus PT100 with two reference psychrometers / E.M.J. Meijer
230 KNMI contribution to the European project WRINCLE: downscaling relationships for precipitation for several European sites / B.-R. Beckmann and T.A. Buishand
231 The Conveyor Belt in the OCCAM model: tracing water masses by a Lagrangian methodology / Trémeur Balbous and Sybren Drijfhout
232 Analysis of the Rijkoort-Weibull model / Ilja Smits

WETENSCHAPPELIJK RAPPORT = SCIENTIFIC REPORT (WR)

99-01 Enhancement of solar and ultraviolet surface irradiance under partial cloudy conditions / Serdal Tunç
99-02 Turbulent air flow over sea waves: simplified model for applications / V.N. Kudryavtsev, V.K. Makin and J.F. Meirink
99-03 The KNMI Garderen experiment, micro-meteorological observations 1988-89: corrections / Fred C. Bosveld
99-04 ASGAMAGE: the ASGASEX MAGE experiment : final report / ed. W.A.Oost
00-01 A model of wind transformation over water-land surfaces / V.N. Kudryavtsev, V.K. Makin, A.M.G. Klein Tank and J.W. Verkaik
00-02 On the air-sea coupling in the WAM wave model / D.F.Doortmont and V.K. Makin.
00-03 Salmon's Hamiltonian approach to balanced flow applied to a one-layer isentropic model of the atmosphere / W.T.M. Verkley
00-04 On the behaviour of a few popular verification scores in yes-no forecasting / C.J. Kok
01-01 Hail detection using single-polarization radar / Iwan Holleman

KNMI-bibliotheek, postbus 201, 3730 AE De Bilt, tel. 030 - 2 206 855, fax 030 - 2 210 407; bibliotheek@knmi.nl