Simulation of precipitation and temperature for generating long synthetic discharge series for use in spring flood scenarios.

Norges vassdrags- og energidirektorat
2004

Report no 1, 2004

Simulation of precipitation and temperature for generating long synthetic discharge series for use in spring flood scenarios.

Utgitt av: Norges vassdrags- og energidirektorat

Redaktør:

Forfatter: Jens Erik Rindal og Thomas Skaugen

Trykk: NVEs hustrykkeri
Opplag: 30 +16
Forsidefoto:
ISSN 1502-3540
ISBN 82-410-0501-6

Emneord: precipitation, temperature, simulation, scenario

Norges vassdrags- og energidirektorat
Middelthuns gate 29
Postboks 5091 Majorstua
0301 OSLO

Telefon: 22 95 95 95
Telefaks: 22 95 90 00
Internett: www.nve.no

Mars 2004

Table of Contents

Preface

This report describes some of the work and results of the project Flomrelatert FoU, financed by the Norwegian Water Resources and Energy Directorate. This particular study describes the methodology put forward in an effort to generate long time series of precipitation and temperature, which can be used in order to simulate long time series of discharge. Besides being a, hopefully, helpful tool in design work, the project is a contribution of the NVE and Hydrology department to the effort to provide Predictions for Ungauged Catchments (PUB) which is an international initiative put forward by the International Association of Hydrological Sciences (IAHS).

Oslo, March, 2004

Kjell Repp

head of Department

Sverre Husebye

head of section

Summary

Three individual models and their calibration are presented in this study. The Bartlett-Lewis precipitation generator, The Onof-Rindal temperature generator and the Swedish rainfall-runoff model HBV. These three models form the system for generating discharge scenarios and synthetic time series of precipitation, temperature and discharge. The main issue is to generate long time series in order to make reliable inferences on the probabilities of extreme events. The synthetic time series of the respective hydrometeorological variables are promising in respect to reproducing the extreme values from observed series and operational use of the system for spring flood scenarios will be initiated in the spring of 2004 by the flood forecasting centre at the NVE.

1 Introduction

For the great majority of the Norwegian water basins, the annual spring flood is the major hydrological event. The Norwegian winter is historically cold and snowy, and snowmelt will produce significant discharge in spring, when the temperature rises. Naturally it is of great importance to model and forecast these events as accurate as possible in order to mitigate the social and economic effects of floods.

The flood forecasting service at the Norwegian national flood forecasting service (NVE) provides each spring flood-scenarios for the melting season. The customary way to compute these scenarios is to use the most updated state of the Nordic rainfall-runoff model, HBV, for a certain catchment, and then run the model forward as many times, N, as you have years of historical data of precipitation and temperature. This provides us with N outcomes, scenarios, of the future melting season. From the N outcomes, we can assess the probability of exceedance of certain critical discharge values like the mean annual- or ten year flood. But, in general, we do not have sufficient data to make reliable inferences on the probabilities of exceedance, which would ideally require several hundreds of years of data, whereas we usually only have from one to a few decades of observations.

To remedy this problem it was decided to find and use models to simulate precipitation and temperature, and from these generate synthetic discharge data of typical 1000 years length. This would then provide the flood-scenarios with 1000 outcomes, and enough data to make more reliable estimates of probabilities of exceedance. Stochastic models are used to generate synthetic data of precipitation and temperature. These data act as input for the HBV-model, calibrated for our catchment of interest, to generate daily discharge.

The 3 models for generating precipitation, temperature and discharge will be presented and validated, and the results for 8 catchments in Norway will be presented and discussed in this report.

2 Model descriptions

2.1 Precipitation model

The model chosen to simulate precipitation is the Randomized Bartlett-Lewis Rectangular Pulse Model (RBLRPM). RBLRPM models rainfall as a clustering of rainfall cells within larger structures or storms. These arrive as an independent point process, which leads to the Poisson process for which the inter-arrival times are exponential distributed.

Storms arrive according to a Poisson process with parameter λ. Each storm is followed by a Poisson process of cell arrivals with randomised rate β $(\beta = \kappa \cdot \eta)$ which has a finite duration *V*. *V* is chosen as an exponentially distributed random variable with the randomised parameter γ $(\gamma = \phi \cdot \eta)$. The precipitation is then added to this wet/dry picture in the form of precipitation pulses *P* of exponentially distributed intensity *x* (mean μ_x) and independently exponentially distributed duration *W* with the randomised parameter η, which is Gamma distributed with shape parameter α and scale parameter $1/\nu$, $\eta \sim (\Gamma(\alpha, 1/\nu))$. This model thus has 6 parameters to estimate: $\lambda, \kappa, \phi, \nu, \alpha, \mu_x$. Details of the RBLRPM can be found in Onof (2000) and references therein.

2.2 Temperature model

Temperature is usually an auto correlated, non-stationary process. A common and simple model for temporary data is the autoregressive model of order p.

$$x_t - m(X) = \sum_{i=1}^{p} w_i \left(x_{t-i} - m(X)\right) + \varphi \tag{1}$$

where φ is an independent and identical distributed (iid) random shock with mean 0 and variance σ^2, w_i is the i'th autoregressive (AR) parameter , x_t is X at time t and $m(X)$ is the mean level of the process

This model assumes a second order stationary process as well as independent shock-terms. It is also common to assume normally distributed shocks with mean 0 and variance σ^2, in which case they are referred to as white noise.

The chosen model assumes white noise and is limited to a maximum order of 4, which results in the following:

$$x_t - m(X_\delta) = \sum_{i=1}^{4} w_i \left[x_{t-i} - m(X_\delta)\right] + \varepsilon \tag{2}$$

where ε is the residual (deviation between modelled and observed data), and $\varepsilon \sim N(0, \sigma^2)$
$m(X_\delta)$ is the mean level for the process X over a window size δ

Since the spring flood is generated by the melting of snow and ice, it is critical to be able to simulate temperatures correctly in the melting season. The model above produced a too early start of the spring flood. An example of this is seen in figure 2.1. As can be seen, the discharge is too high during winter, which also removes volume from the spring flood. Investigations indicated that this is a result of skewness in the data (Skaugen et. al.,2002; Onof 2003), and that this skewness is non-stationary. Since an AR model cannot model skewness, too many warm days in late winter were generated resulting in the early spring flood. To ensure stationarity of the third order, it proved necessary to remove any skewness in the data with a Box-Cox transformation, which is defined as:

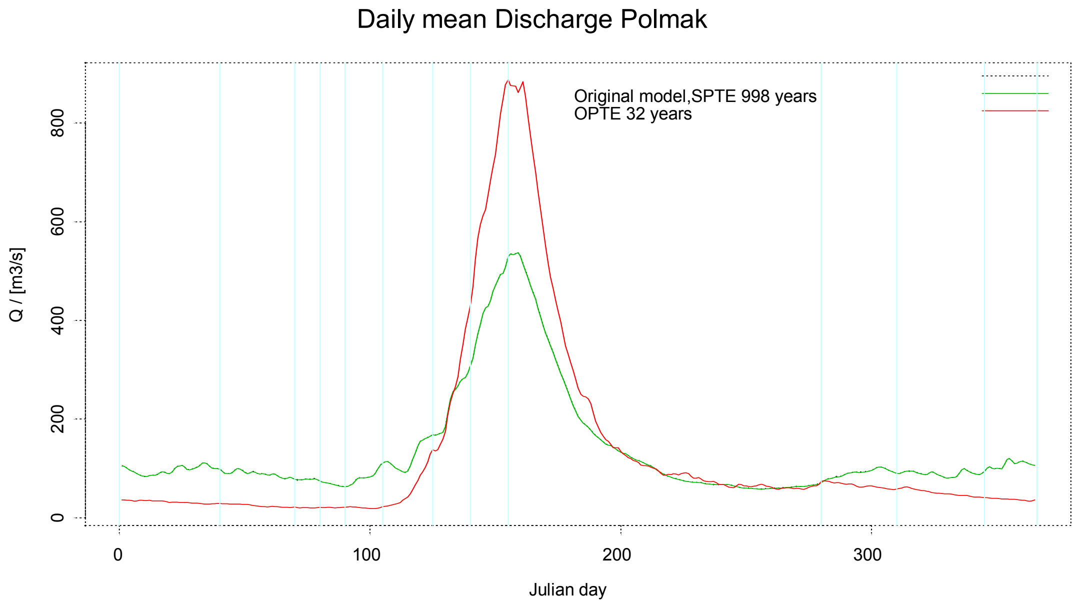
$$x' = \frac{(x+b)^{\alpha_S} - 1}{\alpha_S} \quad , \alpha_S \neq 0$$

$$x' = \ln(x+b) \quad , \alpha_S = 0$$

where x is observed data, x^t is Box-Cox transformed data, b is a constant chosen to ensure $x + b > 0$ and α_S is shape parameter to be estimated within each season.

This model thus has 7 parameters to be estimated.

To meet the demand of third order stationarity, the year is divided into 12 seasons of choice, where we within each season can assume near-stationarity in variance and skewness. The Box-Cox transformation is performed within each season. The seasons are chosen by trial-and-error.



F*igure 2.1 Example of a too early melting season*

2.3 The HBV model

The HBV model was developed at the Swedish Meteorological and Hydrological Institute, and has been the dominant operative rainfall-runoff model in Scandinavia for over two decades. It can be classified as a semi-distributed conceptual model, using sub basins as primary hydrological units, area-elevation distribution and a simple classification of land use (forest, open and lakes). The model can be described consisting of three main components: subroutines for snow accumulation and melt, subroutines for soil moisture accounting and subroutines for response- and river routing. NVE has calibrated this model for 60 drainage basins, which forms the basis for the forecasting of runoff and flood warnings in Norway. The model accepts precipitation and temperature data as input, and estimates daily runoff.

3 Overview of the simulation procedure

Simulation is done using FORTRAN programs in UNIX, while most of the parameter estimation is done in Splus and Excel. A detailed description of these steps is presented in the manual Rindal and Skaugen (2004).

3.1 Precipitation simulation

A FORTRAN program named **prepPrecip** is used to obtain statistical information regarding the precipitation and initial values for an optimalization of parameters of the precipitation model. The output file provides input values to the optimalization program written in the Visual Basic programming language, and contained in an Excel worksheet. The estimated parameters are, along with some other info, written to a file called *input_bartlewtempsim*, which will act as input file for the simulation program **PTsimulation**. The main output of this program is a file containing the simulated precipitation.

For catchments with more than one meteorological station, the observed data will be a weighted sum from these stations. We thus simulate only one precipitation series for each catchment.

3.2 Temperature simulation

The FORTRAN program **PTsimulation** is used to calculate statistics of observed data, and makes an input file for Splus. The script **yearanalysis** plot these statistics. The year is then divided into 12 seasons, in which we can assume stationarity in variance and skew. Using **PTsimulation** again, the Box-Cox parameter is estimated for each season, and another input file for Splus is created. Running the script **ar4v22** on this will estimate optimal AR model for each season. Finally, **PTsimulation** will simulate temperature data based on these estimated parameters.

For catchments with more than one meteorological station, the observed data will be a weighted sum from these stations. We thus simulate only one temperature series for each catchment.

3.3 Discharge simulation

Simulated temperature and precipitation is combined into an input file for the HBV-model. Along with a catchment-specific parameter file, this model will generate daily estimates of discharge.

4 Model validation and case study

Because our goal is to estimate discharge based on Simulated Precipitation and Temperature (SPTE-Simulated Precipitation and Temperature Estimation), we must validate not only simulated temperature and simulated precipitation against its observed data, but also validate SPTE against generated discharge based on Observed Precipitation and Temperature (OPTE - Observed Precipitation and Temperature Estimation). This is done by comparing the annual hydrograms based on daily average values. An especially important feature of the hydrogram in this context is the spring flood. Simulating this feature correctly has been our main concern.

1000 years of precipitation and temperature data were simulated for 8 catchments in Norway, and both these and the simulated discharge were compared against observed meteorological data and discharge simulated from OPTE respectively. These catchments are presented in table 1 along with their meteorological stations and catchment area. Figures for comparison of precipitation, temperature and discharge are presented below.

Table 1. The catchments used in model validation.

Discharge station	Meteorological station(s)	Area / km^2
Atnasjø	Sørnesset	465
Aulestad	Venabu	866
	Vest-Torpa	
Austenå	Byglandsfjord	274
	Tveitsund	
Eggedal	Nesbyen	304
	Lyngdal	
Gaulfoss	Berkåk	3 085
	Selbu	
Gjerstad	Nelaug	235
Knappom	Flisa	1 625
Polmak	Karasjok	14 169

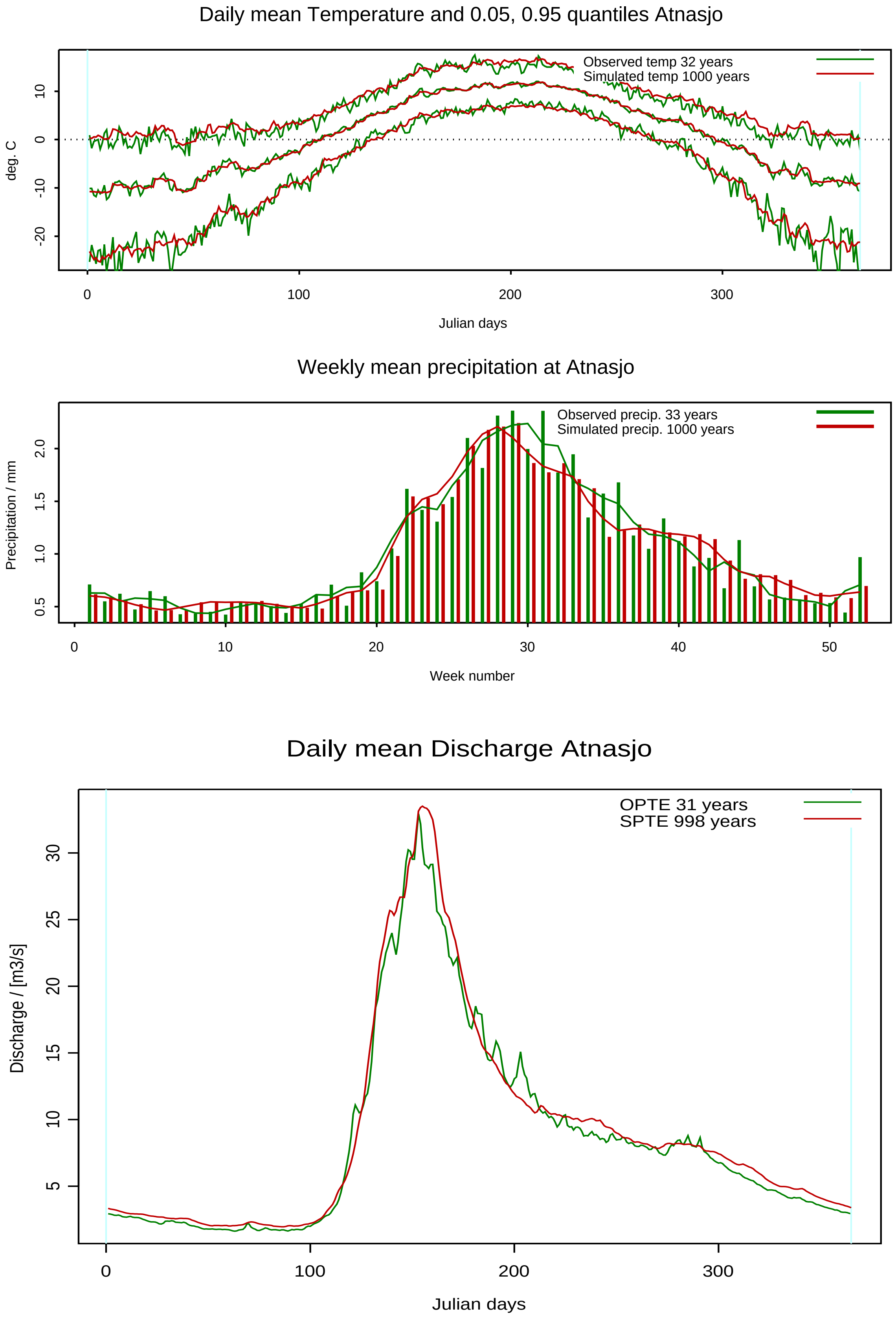
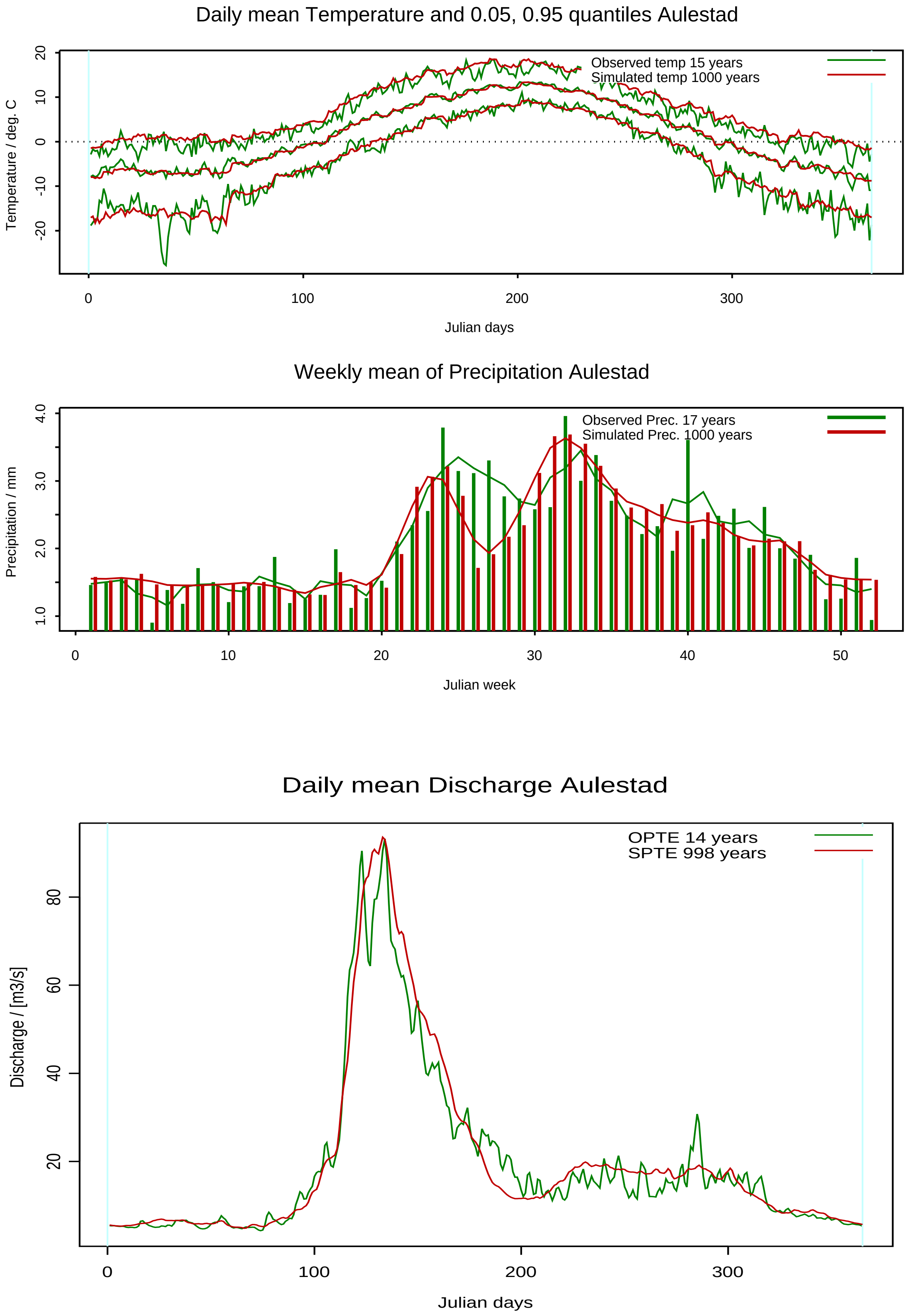


Figure 1. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Atnasjø. "Observed" disharge is simulated with observed precipitation and temperature.


Daily mean Temperature and 0.05, 0.95 quantiles Aulestad

Weekly mean of Precipitation Aulestad

Daily mean Discharge Aulestad

Figure 2. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Aulestad. "Observed" disharge is simulated with observed precipitation and temperature.

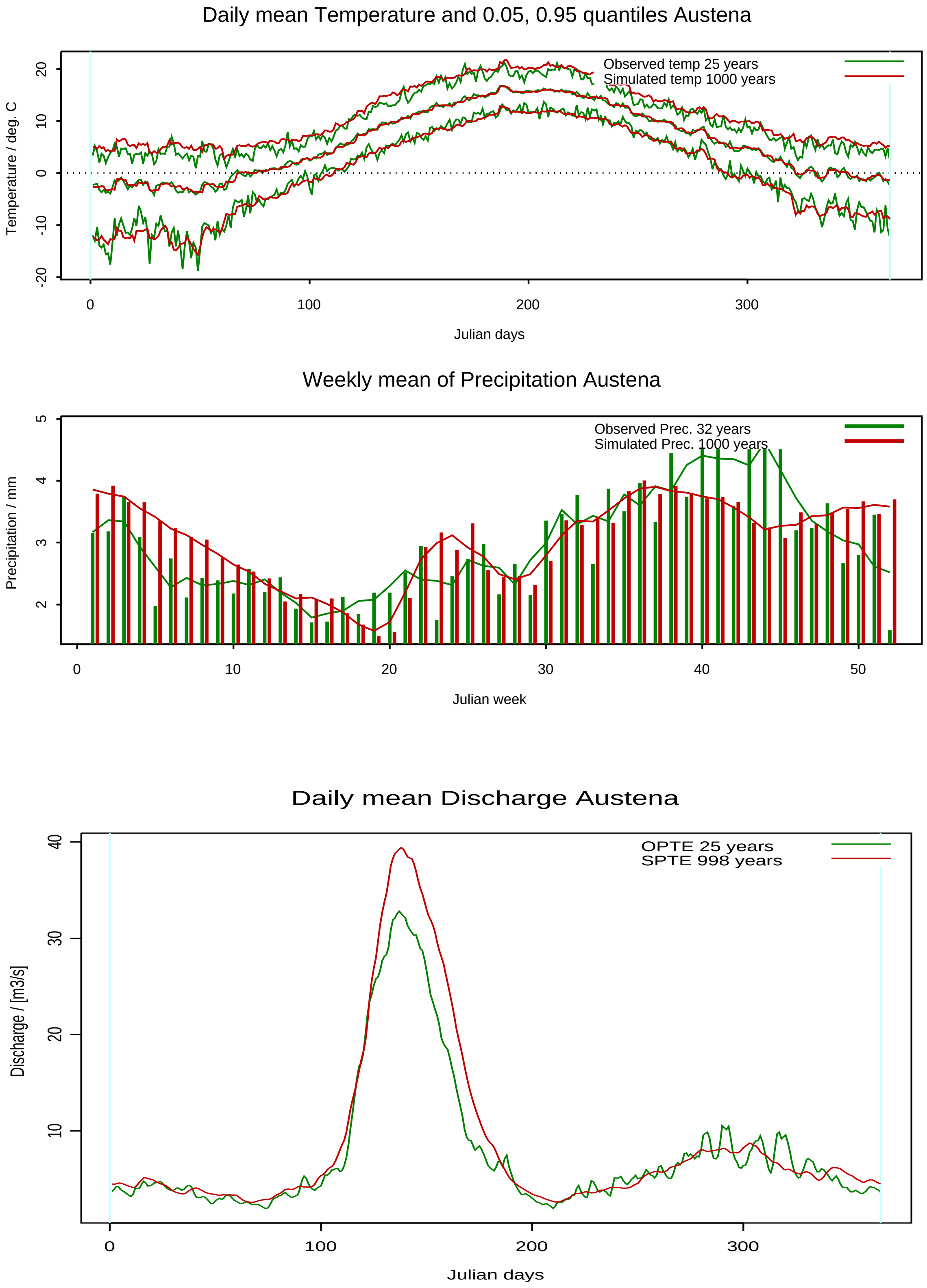


Figure 3. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Austenå. "Observed" disharge is simulated with observed precipitation and temperature.

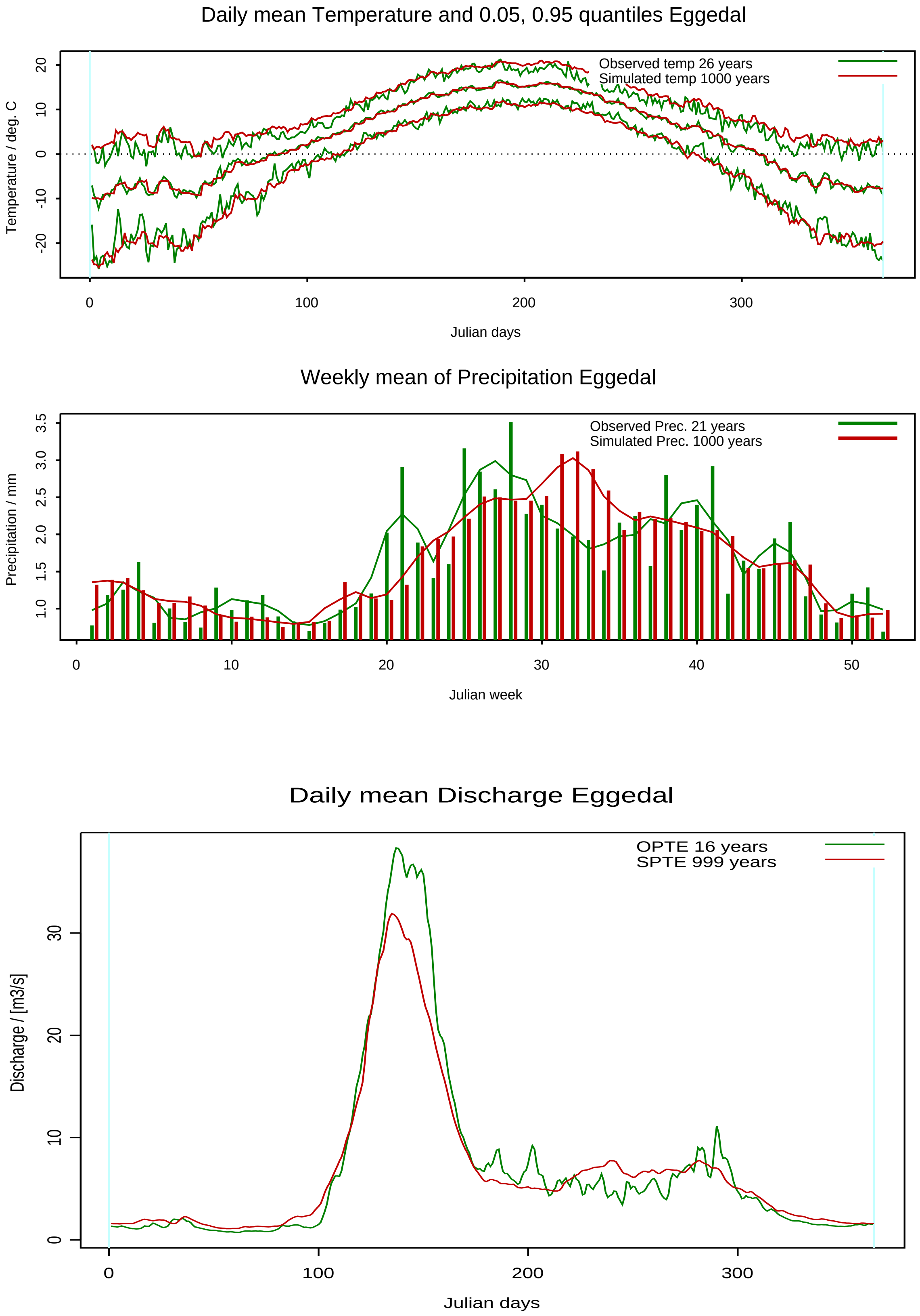


Figure 4. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Eggedal. "Observed" disharge is simulated with observed precipitation and temperature.

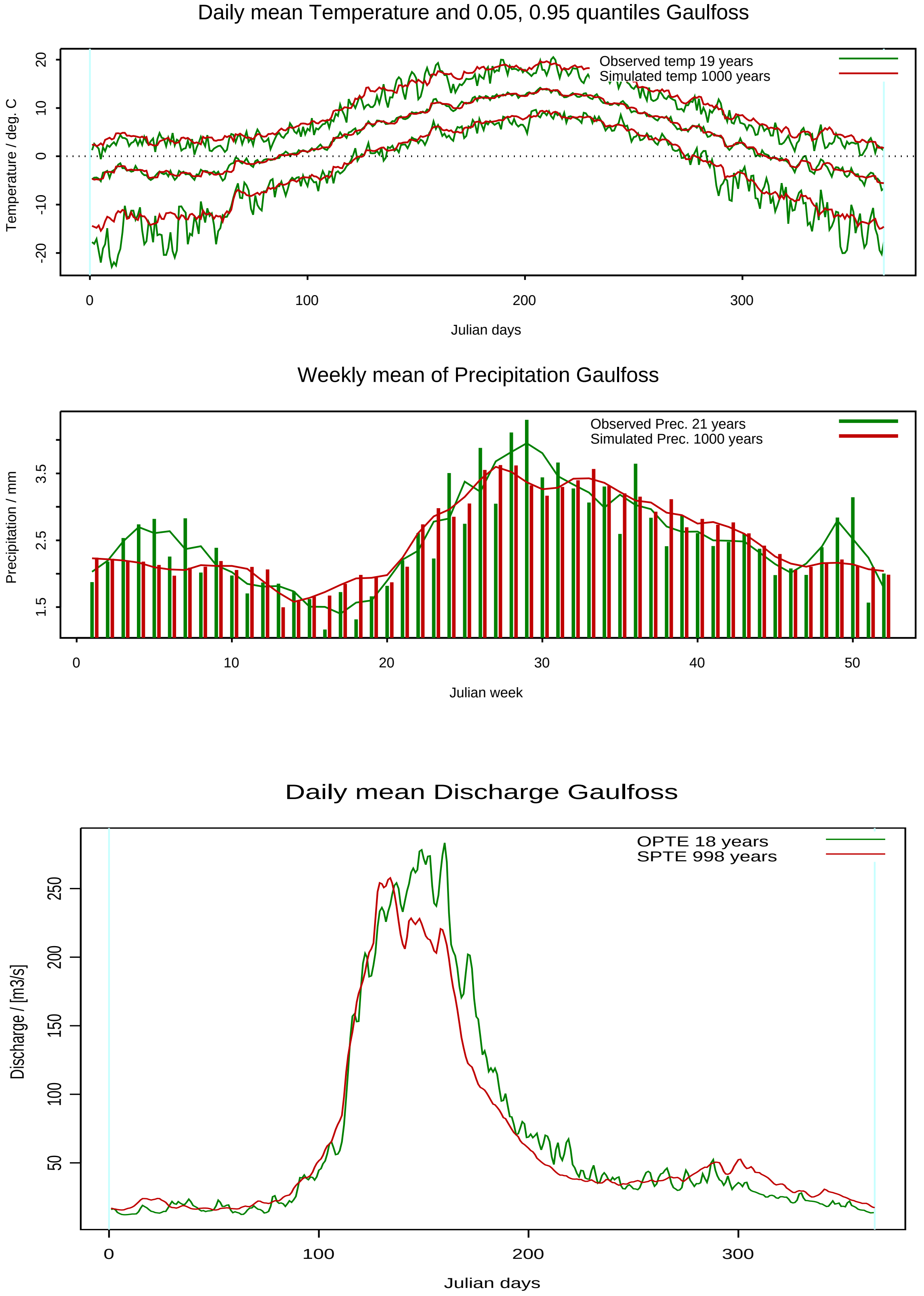


Figure 5. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Gaulfoss. "Observed" disharge is simulated with observed precipitation and temperature.

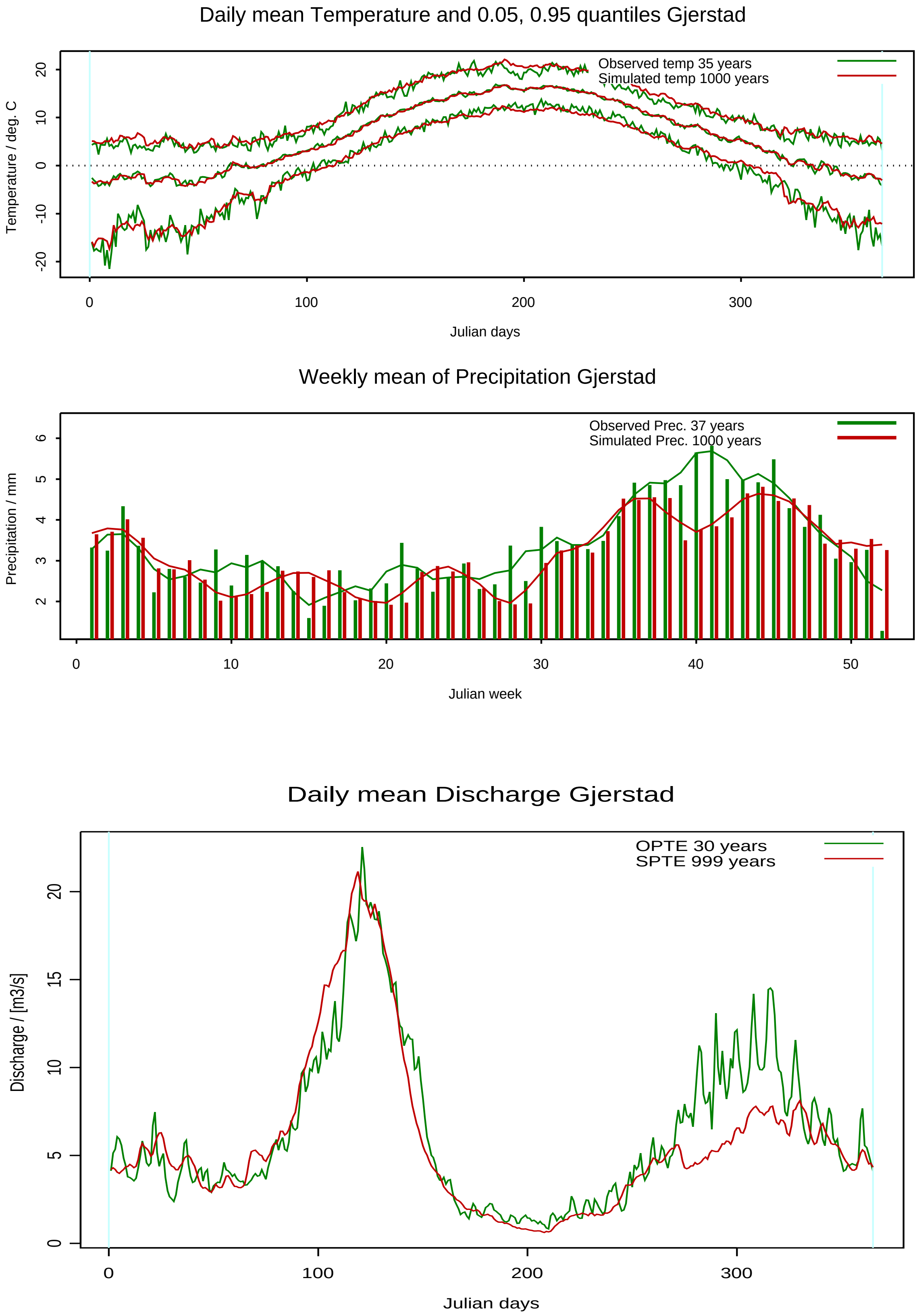


Figure 6. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Gjerstad. "Observed" disharge is simulated with observed precipitation and temperature.

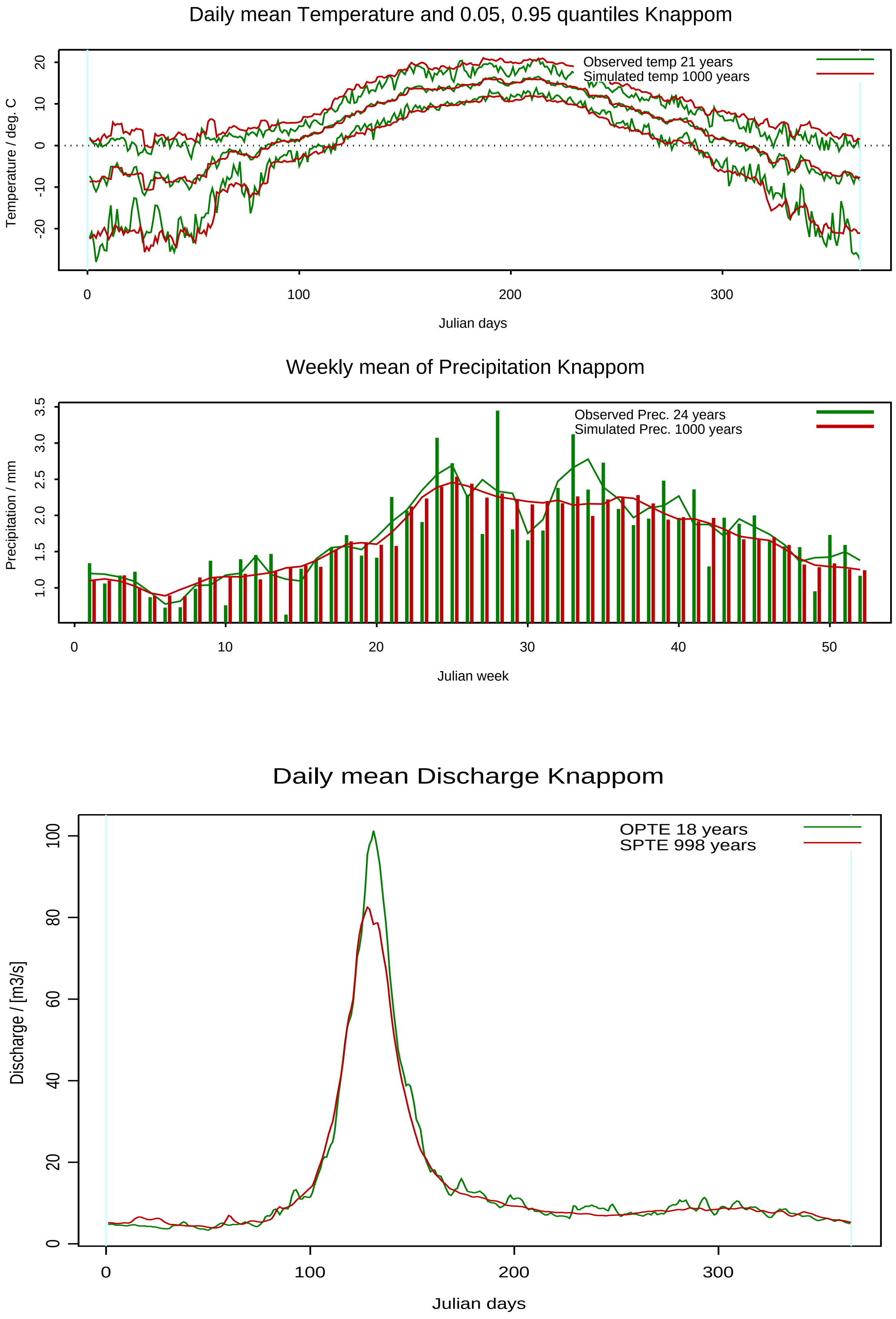
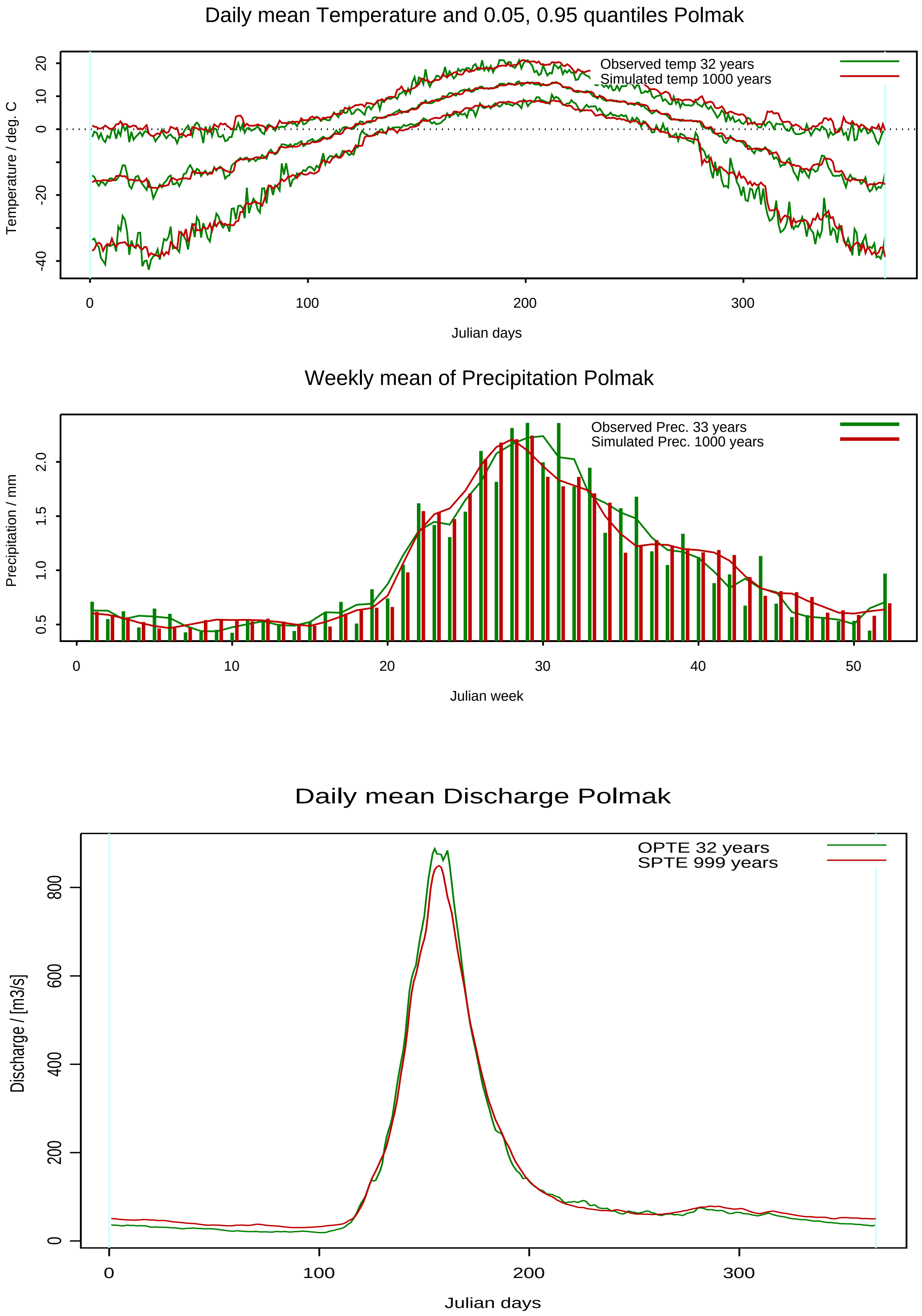


Figure 7. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Knappom. "Observed" disharge is simulated with observed precipitation and temperature.



Daily mean Temperature and 0.05, 0.95 quantiles Polmak

Weekly mean of Precipitation Polmak

Daily mean Discharge Polmak

Figure 8. Observed and simulated temperature (top), precipitation (middle) and discharge (bottom) for Polmak. “Observed” disharge is simulated with observed precipitation and temperature.

5 Discussion

5.1 Temperature model

As can be seen from the above figures, the daily mean temperature during the year is very well simulated as well as the 0.05 and 0.95 quantiles. However, discrepancies are shown when studying daily max/min values. Figure 9 is an example of this taken from the Polmak catchment. Both daily maximum and minimum clearly deviates from their observed counterpart, especially during the cold seasons. This effect is visible for all catchments, and may help explain the marginal higher runoff during winter for some of the above catchments. However, since we are simulating 1000 years of data, it is an expected effect, but one can question the probability of observing +19 degree Celsius in January for the Polmak catchment. If one is uncomfortable with this, another distribution with smaller tails than the normal could be used in our AR model. Also, introducing some sort of threshold may be a possibility, but it may not be clear how to set this threshold. For our purpose of running flood-scenarios for the melting season, and calculate probability of exceedance, the effect can probably be ignored since we are interested in quantiles less than and included 0.95 which is shown to be rather good.

Another visible effect in Figure 9 is the abrupt change in minimum temperature values around days 130 and 270. This is caused by a change of seasons, and represents our inability to continuously transform the data into strictly second order stationarity. By choosing optimal seasons, this effect is reduced.

The last effect to be noticed is the staircase pattern of the daily mean temperature. This is a result of using a window mean; in this case the window size is 5 days. The rationale behind using window means instead of daily means is to reduce the number of parameters. The effect can be ignored if not too large window size is used.

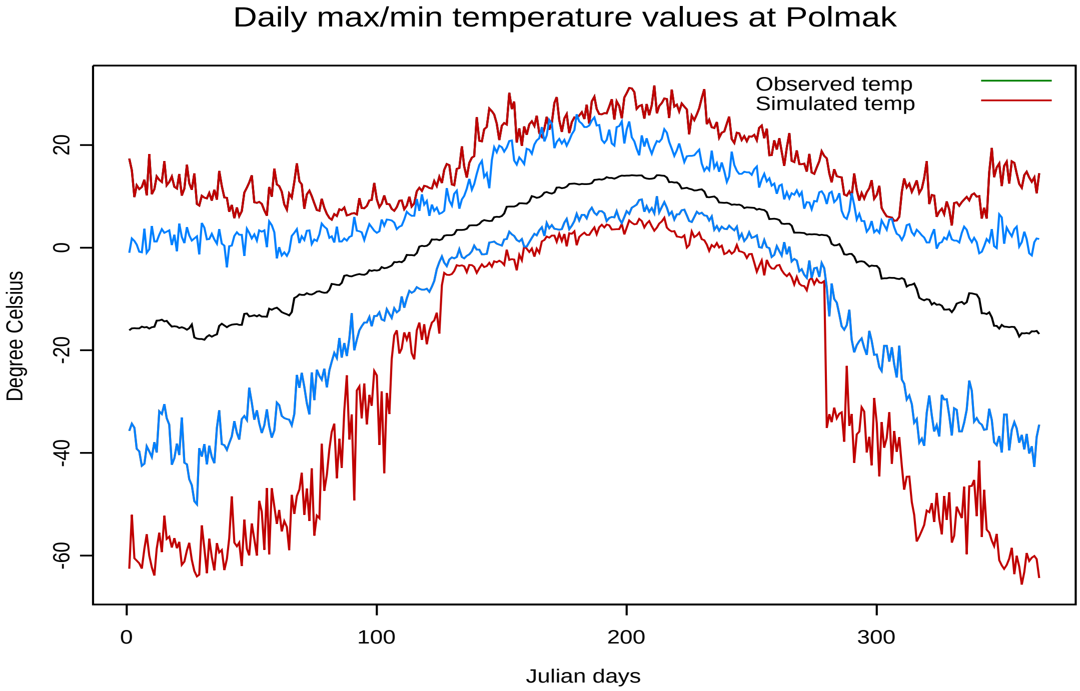


Figure 9 Daily maximum and minimum temperatures for the Polmak catchment.

5.2 Precipitation model

Most of the significant deviations between observed and simulated discharges observed in the figures in chapter 4 can be directly linked to discrepancies between observed and simulated precipitation. As an example take the Austena catchment (Fig. 3). The simulated discharge has too much volume as can be seen from the higher flood-peak. This is caused by to much snow accumulation during the winter. The Gjerstad catchment (Fig. 6) has an opposite effect in the autumn. Here, the simulated discharges are too low as we are unable to simulate enough rain.

6 Flood-scenarios

The simulated precipitation and temperature data is first used to calculate return levels for each catchment. This is done under the assumption that the annual maximum discharge follows the Gumble distribution. Then a spring flood scenario is performed for each catchment producing N outcomes, where N is the number of simulated years of precipitation and temperature data. Based on these outcomes, we get N forecasted annual maximums that can be used to calculate probabilities of exceedance for a specified return level. The return levels used are for the mean, 5 years, 10 years, 20 years, 50 years and 100 years return period. Details of how to run the scenario can be found in Rindal and Skaugen (2004). Table 2 below shows the estimated probabilities of exceedance for both simulated data and observed data from the period 1961-2003. For some catchments, however, missing data are found, and these are interpolated from neighbouring stations. Since the simulation procedure interpolates if there are more than 10 consecutive missing data, the underlying data used in simulated scenario runs may be rather different from those used in the standard scenario runs with observed data.

Table 2. Probabilities of exceedance as estimated from scenario run.

	Probability of exceedance (state March 23 2004)						
Catchment	Data source	mean	5 years	10 years	20 years	50 years	100 years
Atnasjo	Simulated	0,286	0,133	0,069	0,034	0,017	0,012
	Observed	0,167	0,071	0,071	0,024	<0,005	<0,005
Aulestad	Simulated	0,061	0,018	0,006	<0,005	<0,005	<0,005
	Observed	0,048	0,024	<0,005	<0,005	<0,005	<0,005
Austena	Simulated	0,329	0,104	0,036	0,022	0,007	<0,005
	Observed	0,429	0,119	0,048	0,024	<0,005	<0,005
Eggedal	Simulated	0,629	0,277	0,116	0,044	0,019	0,012
	Observed	0,690	0,286	0,071	0,024	0,024	<0,005
Gaulfoss	Simulated	0,707	0,372	0,189	0,083	0,039	0,027
	Observed	0,476	0,095	0,048	<0,005	<0,005	<0,005
Gjerstad	Simulated	0,253	0,109	0,055	0,018	0,005	<0,005
	Observed	0,310	0,095	0,024	0,024	<0,005	<0,005
Knappom	Simulated	0,760	0,371	0,163	0,073	0,021	0,007
	Observed	0,833	0,357	0,119	0,048	<0,005	<0,005
Polmak	Simulated	0,783	0,439	0,227	0,089	0,023	0,010
	Observed	0,762	0,476	0,190	0,095	0,048	<0,005

For some of the catchments the agreement between simulated and observed probabilities is very good; especially for the Polmak catchment. For the Gaulfoss catchment in particular, we observe a rather large discrepancy. It is not apparent why, as the above analysis show good agreement between simulated and observed discharge data. One should bear in mind, however, that since the underlying data is not identical, and observed probabilities are based on 42 years of data compared to 1000 simulated years, we expect discrepancy between the simulated and observed probabilities.

7 Conclusion

For the purpose of generating flood-scenarios, the simulations are adequate as long as the precipitation simulation is improved for some catchments. The increased number of data is likely to increase the precision of our estimates of exceedance compared to the classical method of using observed time series.

As a method of simulating temperatures specifically, the chosen model is not entirely satisfactory. It is not possible to ensure third order stationarity throughout the year. A possible solution would be to introduce more seasons, so that approximate stationarity could be better achieved within the seasons. The trade-off with this is increase uncertainty in the parameter estimation. Also, it is an ad hoc solution as it does not address the real issue; our inability to model higher order statistics like skewness. If temperature simulation was the main issue, a more complex model could be out forward.

References

Onof, C., 2000 A precipitation and temperature generator for the nordic climate.Consultancy report for NVE, May 22, 2000

Onof, C., 2003 Time-series modelling of temperatures. Consultancy report for NVE, March 12 2003.

Rindal, J. E. and T. Skaugen, 2004. Manual for generating synthetic series of precipitation and temperature for use in spring flood scenarios. NVE Report March 2004.

Skaugen, T., M. Astrup, E. G. Langsholt , C. Onof and H-C. Udnæs, 2002. Using simulated time series of rainfall and temperature for design and long term flood warning. In: M. Spreafico and R. Weingartner (eds.) *International Commission for the Hydrology of the Rhine basin*, Proceedings for the International Conference on Flood Estimation, March 6-8, 2002, Berne, Switzerland, CHR Report II-17, 247- 253.

This series is published by Norwegian Water Resources and Energy Directorate (NVE)

Published in the Report series 2004

No. 1 Jens Erik Rindal, Thomas Skaugen: Simulation of precipitation and temperature for generating long synthetic discharge series for use in spring flood scenarios. (22 pp.)