

DATA-DRIVEN MODELLING IN RADAR HYDROLOGY

by
Reinhard Teschl

A dissertation submitted to the
Faculty of Electrical and Information Engineering
of
Graz University of Technology

in partial fulfilment of the
requirements for the degree of
Doctor of Technical Sciences

approved by

Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Walter Randeu
Ao. Univ.-Prof. Dipl.-Ing. Dr. techn. Martin Friedrich

Graz, Austria
2010

STATUTORY DECLARATION

I declare that I have authored this thesis independently, that I have not used other than the declared sources / resources and that I have explicitly marked all material which has been quoted either literally or by content from the used sources.

.....
Date

.....
Signature

ABSTRACT

Although data-driven models and particularly Artificial Neural Networks (ANNs) are established in the field of hydrology, they are rarely trained with weather radar data. But especially in radar hydrology data-driven models are promising. The data volume produced by weather radar networks is considerably large. Reams of gigabytes of data are stored in the archives. Trends, patterns, and regularities are hidden in the data and data-driven approaches aim to extract and model them.

This thesis presents data-driven models for the two principal purposes of radar hydrology: runoff prediction and radar rainfall estimation. The first approach aims to predict the runoff of a small Alpine catchment. ANNs and model trees are the data-driven models used for this purpose. Several input configurations were investigated and it became apparent that they influence the performance and the time lag of the predictions. The models were trained on various lead times and the ANNs consistently perform better than the model trees. When forecasting 45 minutes ahead, the ANN model reaches an efficiency coefficient of 97.4 % compared to 90.9 % of the model tree. Data-driven approaches were also used to improve weather radar estimates of rainfall. The modelled relationship between the measurements of a rain gauge and weather radar data above was tested at a different location. The deviations could be decreased and the correlation coefficient increased compared to applying the standard Z-R relationship. The relative improvements range from 7 to 34 % depending on model and performance measure.

Keywords: Artificial Neural Network (ANN); Model tree; Weather radar; Rainfall-runoff prediction; Radar rainfall estimation

ZUSAMMENFASSUNG

Obwohl datenbasierte Modelle und unter ihnen vor allem Künstliche Neuronale Netze (KNN) mittlerweile vermehrt in der Hydrologie Anwendung finden, beziehen nur die wenigsten Modelle Wetterradardaten mit ein. Doch gerade auf dem Gebiet der Radar-Hydrologie ergeben sich vielfältige Einsatzmöglichkeiten. Wetterradaranlagen produziert große Mengen an Daten, und dementsprechend viele Gigabytes sind in den Archiven gespeichert. Trends, Muster und Regelmäßigkeiten sind in den Daten verborgen, und datenbasierte Methoden versuchen diese zu finden und abzubilden.

Diese Arbeit stellt datenbasierte Modelle für die zwei Hauptaufgaben der Radar-Hydrologie vor: Abflussvorhersage und Niederschlagsermittlung. Mit Hilfe von KNN und stückweise linearen Approximationen, sogenannten *Model-Trees*, wurde der Abfluss eines kleinen alpinen Einzugsgebiets vorhergesagt. Verschiedene Input-Konfigurationen wurden untersucht und dabei zeigte sich, dass diese einen nicht unwesentlichen Einfluss auf Effizienz und Zeitversatz der Prognose haben. Die datenbasierten Modelle wurden für verschiedene Prognosezeiträume trainiert, dabei arbeiten KNN durchwegs besser als Model-Trees. Bei einem Prognosezeitraum von 45 Minuten erreichen KNN einen Effizienz-Koeffizienten von 97.4 % verglichen mit 90.9 % beim Model-Tree. Datenbasierte Modelle wurden auch eingesetzt um die Wetterradardaten selbst zu verbessern. Der an einer Stelle ermittelte Zusammenhang zwischen Regenmesser- und den darüberliegenden Wetterradardaten erwies sich auch an einem anderen Standort als aussagekräftig. Im Vergleich zur Standard Z-R Beziehung konnten die Abweichungen verringert und der Korrelationskoeffizient der Messreihen erhöht werden. Je nach Modell und Effizienz-Parameter konnten relative Verbesserungen im Bereich von 7 bis 34 % erzielt werden.

Schlüsselwörter: Künstliche Neuronale Netze (KNN); Model-Tree; Wetterradar; Niederschlags-Abfluss Vorhersage; Niederschlagsmessung mittels Radar

CONTENTS

1	Introduction	1
1.1	Motivation	1
1.2	Explanation of terms	1
1.3	Current research and contributions of this study.....	3
1.4	Structure of the work.....	5
2	Precipitation and its measurement by rain gauges.....	7
2.1	Forms of precipitation	7
2.2	Precipitation generation types	7
2.3	Measurement of precipitation.....	9
2.3.1	Types of rain gauges	9
2.3.2	Measuring errors of rain gauges.....	10
2.4	Variability of precipitation	12
2.4.1	Spatial and temporal variability of precipitation in the study area	13
2.5	Point measurements and areal estimation	18
3	The weather radar	21
3.1	Principle	21
3.2	Drop-size distribution.....	23
3.3	Radar reflectivity factor	25
3.4	Measuring rain rate with weather radar.....	26
3.5	Weather radar frequencies.....	26
3.6	Scanning modes and radar displays	27
3.6.1	Range Height Indicator (RHI).....	28
3.6.2	Plan Position Indicator (PPI).....	29
3.6.3	Constant Altitude Plan Position Indicator (CAPPI).....	29
3.7	Sources of possible errors	32
3.7.1	Beam blockage	32
3.7.2	Beam attenuation.....	34
3.7.3	Overshooting beam	36
3.7.4	Evaporation	36
3.7.5	Ground clutter	37
3.7.6	Anomalous propagation	37
3.7.7	Non weather targets.....	38
3.7.8	Bright band.....	38
3.8	The Austrian weather radar network.....	38
4	Rain gauge and weather radar comparison.....	39
4.1	Sampling characteristics.....	39
4.2	Quantisation of the weather radar data.....	40
4.3	Case study	42
5	Data-driven modelling.....	47
5.1	Artificial neural networks.....	47
5.1.1	Neuron model.....	48
5.1.2	Network architectures	49
5.1.3	Learning rules.....	52

5.2	Model trees	54
5.2.1	Building the tree	55
5.2.2	Pruning the tree	56
5.2.3	Smoothing the tree	56
6	Modelling the rainfall-runoff relationship	57
6.1	Study area	57
6.2	Available data	58
6.2.1	Rain gauge and flow meter data	58
6.2.2	Radar data	58
6.3	Performance measures	59
6.4	Preprocessing	61
6.4.1	Plausibility check	61
6.4.2	Data analysis	61
6.4.3	Test-, training-, and validation-data	69
6.5	Feed-forward neural network approach	71
6.6	Model tree approach	73
6.7	Results	74
6.7.1	Results of standard configuration	74
6.7.2	Results of the cumulative rainfall configuration (CR)	81
6.7.3	Results of the previous runoff configuration (PR)	82
6.8	Summary and discussion	84
7	Improving weather radar estimates of rainfall with data-driven models	89
7.1	Introduction	89
7.2	Data available	91
7.3	Preprocessing	91
7.3.1	Data analysis	91
7.3.2	Input/output parameters	92
7.3.3	Data separation	93
7.4	Data-driven approaches	94
7.4.1	Regression analysis	94
7.4.2	Neural network approaches	97
7.4.3	M5 model tree approach	99
7.4.4	Instance-based learner	99
7.5	Performance evaluation	100
7.6	Results	101
7.7	Summary and discussion	104
8	Conclusion and outlook	107
	References	109

FIGURES

Figure 2.1.	Convective precipitation	8
Figure 2.2.	Orographic precipitation	8
Figure 2.3.	Two forms of cyclonic frontal precipitation	9
Figure 2.4.	Positions of 11 available rain gauges in the geographic coordinate system	14
Figure 2.5.	Mean annual precipitation versus elevation of the rain gauges	14
Figure 2.6.	Gauge-to-gauge correlogram for 2002-2003	15
Figure 2.7.	Quarterly gauge-to-gauge correlograms	16
Figure 2.8.	Auto-correlation of the rain gauge stations with respect to the 15-minute time lag.....	17
Figure 2.9.	Quarterly auto-correlation plots of the rain gauge stations with respect to the 15-minute time lag	17
Figure 2.10.	Arithmetic mean of measured rainfall depths	18
Figure 2.11.	Thiessen polygon method	19
Figure 2.12.	Isohyetal method	19
Figure 3.1.	Azimuth, elevation and range of the weather radar echoes	22
Figure 3.2.	Marshall-Palmer drop-size distribution for various rain rates	23
Figure 3.3.	Gamma drop-size distribution.....	25
Figure 3.4.	Schematic RHI plot.....	28
Figure 3.5.	Conversion of raw data into a cubic grid	28
Figure 3.6.	RHI plot of a convective precipitation event	29
Figure 3.7.	Schematic PPI scan	30
Figure 3.8.	Procession of a CAPPI display from PPI scans at successive elevation angles ..	30
Figure 3.9.	CAPPI-plots of one precipitation event at different altitudes	31
Figure 3.10.	Total beam blockage	32
Figure 3.11.	Identification of beam blockage in radar images.....	33
Figure 3.12.	Beam attenuation by heavy rain close to the radar site.....	34
Figure 3.13.	Identification of beam attenuation due to an intense shower cell in radar images	35
Figure 3.14.	Overshooting beam	36
Figure 3.15.	Virga - evaporation of precipitation above the ground.....	36
Figure 3.16.	Ground clutter	37
Figure 3.17.	Anomalous propagation of the radar beam.....	37
Figure 4.1.	Idealized sampling volume of a tipping rain gauge	39
Figure 4.2.	Idealized sampling volume of the weather radar	40
Figure 4.3.	Quantisation of rain rates	41
Figure 4.4.	Consideration of surrounding radar volume elements	44
Figure 5.1.	Illustration of an ANN with nodes and connections between them.....	48
Figure 5.2.	Mode of operation of a single neuron	49
Figure 5.3.	Schematic curve of a sigmoid transfer function.....	49
Figure 5.4.	One layer of an ANN. The layer consists of S neurons and an R -dimensional vector forms the input	50
Figure 5.5.	Fully connected feed-forward neural network with two hidden layers	50
Figure 5.6.	Schematic curve of the radial basis transfer function	51
Figure 5.7.	Mode of operation of a radial basis neuron.....	51
Figure 5.8.	Architecture of an RBF neural network.....	52
Figure 5.9.	Schematic illustration of supervised learning.....	53
Figure 5.10.	Example of a classification tree. The tree classifies forms of precipitation	54

Figure 6.1.	Map of the Wernersdorf catchment	58
Figure 6.2.	Exemplary rainfall and runoff curves of one rainfall event	65
Figure 6.3.	Runoff curve and time shifted precipitation measurements.....	66
Figure 6.4.	Runoff curve and averaged precipitation measurements	67
Figure 6.5.	Runoff curve and precipitation measurements with flexible integration time....	68
Figure 6.6.	Subsets used for testing (numbered 1 to 5).....	70
Figure 6.7.	Network architecture used for the rainfall-runoff prediction.....	72
Figure 6.8.	Examples of good and poor runoff predictions.....	76
Figure 6.9.	Two examples of time lags between predicted and observed runoff curves	78
Figure 6.10.	Method to visualise the timing lags of (a) ANN and (b) MT predictions.....	80
Figure 6.11.	Forecasts 3 time steps ahead of ANN model and MT	85
Figure 6.12.	Visualisation of the mapping of the parameter <i>actual runoff</i> in ANN ensemble models for various configurations	86
Figure 7.1.	Three-dimensional structure of precipitation given by the weather radar	90
Figure 7.2.	Study area in the province of Styria, Austria, showing radar and rain gauge locations	91
Figure 7.3.	Visualisation of the inputs of the models.....	93
Figure 7.4.	Fitting the Z - R data with non-linear regression	95
Figure 7.5.	Fitting the $\log(Z)$ - $\log(R)$ data with linear regression	96
Figure 7.6.	Relative improvement of the performance of the FNN models.....	103

TABLES

Table 3.1.	Frequency bands of weather radars.....	27
Table 3.2.	Maximum particle diameter of Rayleigh targets for weather radar frequencies.....	27
Table 4.1.	Quantisation intervals of the Austrian weather radar system	41
Table 4.2.	Deviations due to quantisation.....	42
Table 4.3.	Comparison of rain gauge and shifted weather radar time series CAPPI 3 km..	43
Table 4.4.	Comparison of rain gauge and shifted weather radar time series CAPPI 4 km..	43
Table 4.5.	Comparison of rain gauge and modified weather radar time series CAPPI 3 km	44
Table 4.6.	Comparison of rain gauge and time shifted weather radar time series CAPPI 4 km	45
Table 6.1.	Time lag with maximum cross correlation coefficient between rainfall and runoff series.....	62
Table 6.2.	Time lags and cross correlation coefficients between runoff and precipitation measurements.....	63
Table 6.3.	Time lags and cross correlation coefficients between runoff and integrated precipitation	67
Table 6.4.	Cross correlation coefficients between runoff and precipitation measurements with flexible integration time.....	68
Table 6.5.	MSE on the validation dataset as model selection criterion	72
Table 6.6.	Model selection for the cumulative rainfall configuration.....	73
Table 6.7.	Model selection for the previous runoff configuration	73
Table 6.8.	Performance of ANNs and MTs in standard configuration with different configurations of input data	75
Table 6.9.	Weighted average performance measurements.....	77
Table 6.10.	Performance of ANNs and MTs for the effective forecast.....	81
Table 6.11.	Performance of ANN and MT in cumulative rainfall configuration for the effective forecast.....	82
Table 6.12.	Performance of ANN and MT in previous runoff configuration for the effective forecast.....	82
Table 6.13.	Performance measures of ANNs when trained on different lead times.....	83
Table 6.14.	Performance measures of MTs when trained on different lead times.....	83
Table 7.1.	Overview of the applied DDMs and their input parameters	93
Table 7.2.	Determined coefficients of the Z-R relationship.....	96
Table 7.3.	Applied FNN architecture and input/output configurations	97
Table 7.4.	Applied RBF architecture	98
Table 7.5.	Applied MTs and their input/output configurations	99
Table 7.6.	Applied IBk learner configuration	99
Table 7.7.	Performance measures of the various data-driven approaches compared to standard Z-R outcome and location optimized Z-R relationship.....	102

1 INTRODUCTION

1.1 MOTIVATION

Today's weather radar systems provide data with high temporal and spatial resolution. Every five minutes the Austrian weather radar system generates a new radar image, covering not only the area of Austria but also transborder regions. Reflectivity measurements are obtained with a resolution of $1 \text{ km} \times 1 \text{ km} \times 1 \text{ km}$ for altitudes up to 16 km. Keeping in mind that Austria covers an area of more than 83,000 km², and that more than one decade of data are stored, one can get an idea of the large amount of archived data.

What happens with all these data – not only weather radar data, but generally all data that are stored for operational purposes? Often the data are just stored. To that aspect Witten and Frank (2005, p. 4) highlight the role of the computer that makes it “too easy to save things that previously we would have trashed” and comment with respect to data: “Inexpensive multigigabyte disks make it too easy to postpone decisions about what to do with all this stuff – we simply buy another disk and keep it all“.

Archives do not only contain data, but also hidden information e.g. patterns, trends, regularities. Usually the longer the available time series, the more can be learned about the system that the dataset describes. Large amounts of data, however, can not be investigated manually. Established methods to extract the information are data-driven models.

1.2 EXPLANATION OF TERMS

This work introduces data-driven models for applications in the field of radar hydrology. The thesis is aimed at readers in the fields of machine learning and modelling, as well as weather radar technology and hydrology. As readers from different fields of science are expected, their diverse previous knowledge is taken into consideration.

This section will give an explanation of the most important technical terms used in the course of this work. This will make it easier for the reader to find access to hitherto possibly unfamiliar techniques utilised in the present work.

The science of extracting potentially useful information from data sets is called data-mining. Data-mining is often associated with commercial applications and intelligence services and in this regard it has also raised privacy concerns, but it is also increasingly used in the scientific world to extract information from the huge data sets produced by today's automated measuring and detection systems.

A field closely related to data mining is machine learning. Machine learning is a branch of artificial intelligence and a collective term for techniques which allow computers to “learn” from experience, more precisely to modify their transaction as a result of new information.

According to Solomatine (2002, p. 758) most of the machine learning problems “can be formulated as problems of function approximation”, and the author sees “machine learning as the main source of methods for data driven modelling”. Indeed, data-driven modelling applies techniques and approaches used in data mining and machine learning. One of the most popular examples is the Artificial Neural Network (ANN). ANNs are computational models based on neural networks in neuroscience. As the biological example, an ANN consists of simple processing elements – the neurons – operating in parallel. The global behaviour of an ANN is determined largely by the connections between the neurons. ANNs automatically adjust the values of the weighted connections to perform a particular function. Artificial neural networks have been considered as the mainstream technology for data-driven modelling and several types of them exist. In addition to ANNs there are a number of other data-driven modelling approaches. Solomatine (2002) comprises amongst others *fuzzy logic* methods, *decision trees*, *Bayesian classification*. Janes and Yaffe (2006) rank *clustering*, *principal components analysis*, and *partial least squares* in this category.

In the previous paragraph the term data-driven modelling was introduced. In the following the term *radar hydrology*, the other expression in the title of this thesis, is derived from more general terms.

Hydrology is defined as “the branch of geology that studies water on the earth and in the atmosphere: its distribution and uses and conservation” (Miller, 2009). Due to solar radiation all water continuously circulates within the hydrosphere. Water evaporates from land and the sea, condensation forms clouds, precipitation brings the water back to the earth, it accumulates in the soil or in reservoirs and runoff brings it back to the sea, and with the re-evaporation the whole process starts again.

This hydrologic cycle is vital for our life. Its exploration is one principal task of hydrology. Dyke and Peschke (1995) name in addition the calculation of the water balance for catchments, regions, continents, and up to the global scale as another major task. Furthermore, the evaluation of the water resources dealing with the accessibility and quality of the water is a basic challenge for drinking water and food production. These examples demonstrate that many fields of sciences are related to hydrology: the broad field of civil engineering for example, especially the part that deals with hydroelectric power plants and flood control as well as water supply and sewage; or the field of climatology, the quantitative description of the climate of a region. Another principal task of hydrology is the hydrologic prediction where observations of hydrologic processes are used to make forecasts.

It becomes clear that hydrology is a term of wide comprehension. Precipitation is an important if not the most important parameter. For a long time rain gauges were the only means to record precipitation, and they are still among the most exact instruments for this purpose. During the late 1930s another technology emerged. The first military radar applications mainly for aircraft detection were used. One decade later scientists mainly from the United States and Canada began to use the radar for meteorological purposes. Later, all over the world researchers studied the atmosphere using the radar as a means of observation and measurement. Thus the technology that was used for military purposes emerged to be valuable for precipitation measurement. Henceforth the term radar meteorology refers to the branch of meteorology that uses radars for weather observations and forecasts.

Since the beginnings of radars in meteorology the detection and measurement of precipitation has been a field of intense research. The estimation of the rain rate from radar reflectivity measurements still is of particular importance. Over the years rainfall estimation techniques improved and yielded more reliable data. The more accurate the data became, the more rewarding they were for hydrological applications.

This area of research is known as radar hydrology. Radar hydrology can be seen as a branch of hydrology. The *Glossary of Meteorology* of the American Meteorological Society (AMS, 2000) defines the term radar hydrology as

“The use of meteorological radar measurements for hydrological purposes, particularly for estimating the current precipitation intensity as a function of location over a region and the total precipitation during a prescribed time interval, and for deriving estimates of runoff and streamflow.”

According to Uijlenhoet (2001, p. 616) “At the heart of the problem of radar hydrology lies the conversion of the radar reflectivity factor Z [...] to rain rate R ”. The rain rate is indeed essential for virtually all hydrological applications. Therefore, radar rainfall estimation is of high significance in the radar community.

1.3 CURRENT RESEARCH AND CONTRIBUTIONS OF THIS STUDY

For the past 15 years, data-driven models have been used in many areas of sciences (cf. Cherkassky *et al.*, 2007). The number of applications has been growing steadily. This includes approaches in climate change studies, satellite meteorology and oceanography, weather forecasting, and applications in hydrology.

It is noteworthy that data-driven approaches and especially artificial neural networks undergo a high period now, especially since researchers developed neural networks already in the 1950s and the models experienced their first heyday in the 1960s. In the 1970s the confidence of many scientists in neural networks was on the decline. Often the book “Perceptrons; an introduction to computational geometry” by Minsky and Papert (1969) that highlighted also drawbacks of the perceptron (an important kind of artificial neural networks, introduced by Rosenblatt, 1958) is seen at least jointly responsible for the decline. In German-language literature, Meyer (2004) describes the controversy in the study of neural networks, and mentions that after 1970 the study of neural networks was no longer accepted as mainstream-science.

Till the mid 1980s, neural networks did not attract much attention, but practical applications in pattern recognition and signal processing put them back on the scene.

Focusing on hydrological applications, especially the rainfall-runoff relationship has been modelled with artificial neural networks since the mid 1990s. Hsu *et al.* (1995) showed the potential of ANNs for modelling the nonlinear behaviour of watersheds. The authors showed that the ANN approach provides “a better representation of the rainfall-runoff relationship” (p. 2517) of a medium-size basin than a linear time series approach or a conceptual model. Hsu *et al.* (1995) admitted though that the ANN approach “is by no means a substitute for conceptual watershed modelling”, because such models “do not have physically realistic

components and parameters” (p. 2517). Shamseldin (1997) tested neural networks with different external input scenarios and states that “neural networks can provide more accurate discharge forecasts than some of the traditional models” (p. 292). But he was also aware of the fact that his neural network yields variable results on different catchments.

Some scientists report that their ANNs underestimate flood peaks or perform poorly at high levels (See *et al.*, 1997; Dawson and Wilby, 1998; Campolo *et al.*, 1999). ANNs do not extrapolate well. Imrie *et al.* (2000) described the problem of underestimating peaks when ANNs encounter events containing hitherto unobserved values and presented a method for improved generalisation. Sudheer *et al.* (2002) utilised with cross-, auto-, and partial-auto-correlation statistical properties of data series for identifying an input vector that best represents the basin processes.

Many papers cover feed-forward neural networks, but also other architectures have been investigated. Mason *et al.* (1996) used a Radial Basis Function Neural Network (RBFNN) in rainfall-runoff modelling. Senthil Kumar *et al.* (2005) compared the radial basis function network with the multi-layer architecture with sigmoid transfer functions, a judgement, however, on which type is superior was not possible. Pan and Wang (2004) used a dynamic recurrent neural network (State Space Neural Network, SSNN; introduced by Zamarreño and Vega, 1998) to perform short term runoff forecasting and report a satisfactory performance.

Also alternatives to artificial neural networks for rainfall-runoff modelling have been investigated. Solomatine and Dulal (2003) compared Model Trees (MTs) and neural networks and found that both performed very well with short lead times, while they fail to produce good results with longer lead times.

Notably, most of the hydrological applications using data-driven models rely on rain gauge data only. Often so called lumped-models are built where the rainfall is assumed to be uniform over the whole catchment. For a long time, precipitation measurements from other sources were neglected in neural network models, but recent papers also used other precipitation data. Pereira Filho and Dos Santos (2006) and Teschl and Randeu (2006) utilised weather radar data in neural network models. The difficulty both approaches faced was the increase of input parameters of the ANN by using numerous gridded radar data instead of one or a few rain gauge measurements. Too many input parameters complicate the training process and therefore a compromise to limit their number had to be found. Pereira Filho and Dos Santos (2006) divided the watershed area into eight isochrones of 30 minutes each and obtained the areal precipitation by averaging over all corresponding 2 x 2 km² rainfall accumulations. Teschl and Randeu (2006) calculated the time lag between precipitation and runoff measurements by cross correlation analysis, and combined these 1 x 1 km² radar measurements that showed the same time lag to clusters. Chiang *et al.* (2007) used merged satellite-derived precipitation and rain gauge measurements in flood forecasting. Here the number of measurements over the 204 km² drainage area is manageable because of a coarser grid resolution of the satellite images.

One of the main challenges in data-driven runoff modelling at present is the question of lagged runoff predictions. Often previous runoff measurements are used in runoff models to indirectly represent the hydrological state of a catchment. These data are seen responsible for the timing error. Radar data in data driven runoff models might be helpful to minimise timing errors. Abrahart *et al.* (2007) conjectured that additional inputs such as radar data “might

provide the necessary drivers to help correct the timing problem” (p. 429). However, the issue of timing errors in conjunction with radar data has not been explored. This study applies radar data in neural network models and model trees and particularly investigates timing errors in these models. The main research questions are: Do different input configurations have an influence on the timing error of the model? Are model trees similarly affected by timing errors? How do ANNs and MTs perform with the same input configuration?

Furthermore, another important question of radar hydrology is addressed and this is the accuracy of weather radar measurements. The problem of estimating ground rainfall using radar measurements aloft has already been investigated with data driven approaches. So far neural networks have been used. Xiao and Chandrasekar (1997) applied a feed-forward backpropagation neural network for rainfall estimation from weather radar data. Liu *et al.* (2002) developed a RBFNN to estimate ground rainfall using the vertical profile of reflectivity as input vector. Li *et al.* (2003) compared several input vector configurations and showed that the radar reflectivity from 1 to 4 km height above the rain gauge is the best input vector to a RBFNN for estimating the ground rainfall. Xu and Chandrasekar (2005) also used this vertical profile as input vector for their RBFNN.

The studies above utilised weather radar data near ground level. Because of the orography in Austria such radar data are rarely available. The key question is: Does it nevertheless make sense to apply such models? Besides neural networks also model trees and instance-based learners which have not been used for radar rainfall improvement are investigated. It is examined what input configuration the models work best with.

1.4 STRUCTURE OF THE WORK

The more general chapters at the beginning are aimed at making the access to the work more easily. Readers with a background in general data-driven modelling may need to know the basics of precipitation measurement by rain gauge and weather radar. On the other hand, radar scientists and hydrologists may not be familiar with data-driven techniques like neural networks and model trees. The first chapters are aimed at describing the fundamentals of the involved disciplines.

Chapter 2 deals with precipitation and its measurement with rain gauges and addresses precipitation generation types as well as sources of possible errors of rain gauges. Subsequently, the spatial and temporal variability of precipitation in the area of interest is investigated. Chapter 3 describes the principle of the weather radar and gives details of the weather radar station utilised. Furthermore, advantages and drawbacks of different radar scanning modes and possible sources of errors are given. Chapter 4 discusses the different sampling characteristics of rain gauge and weather radar that has to be taken into consideration when comparing both measurements. Before describing the models, a short overview of the used data driven approaches is given in Chapter 5 to introduce terms and principles to newcomers in this domain.

Chapters 6 and 7 form the main part of the work. In Chapter 6 the rainfall-runoff relationship of a small catchment is modelled. Weather radar data, together with data of one rain gauge and the actual runoff, form the input of the model that is trained to predict the future runoff. In

detail the preprocessing steps to format the data are described. Later, modifications are presented that led to a better performance.

Chapter 7 addresses the improvement of weather radar measurements. Rain gauges measure the rain rate quite accurately but the measured rate is only relevant for a certain location. An approach is described that accepts radar data as input and predicts the rain rate as measured by the rain gauge. This relationship is also applied at another site and it is investigated if the model yields better estimates on ground level. In the following several different data-driven approaches are compared.

Chapter 8 discusses what configurations work satisfactory and gives an outlook in what fields of radar meteorology data-driven models could be applied in the future.

2 PRECIPITATION AND ITS MEASUREMENT BY RAIN GAUGES

Precipitation are water particles in solid or liquid form “that originate in the atmosphere and fall to the earth's surface” (AMS, 2000). This definition makes it clear that precipitation is one of the main components of the hydrologic cycle and the key element in hydrology. Precipitation can be divided into falling precipitation e.g. rain or snow, deposited precipitation e.g. dew or fog, and accumulated precipitation e.g. accumulated snow or hail. In this work the focus lies on falling precipitation. This chapter describes forms of precipitation and weather situations that lead to certain types of precipitation. Later, rain gauges – the typical instruments to measure falling precipitation – are described. Finally methods to estimate the areal rainfall based on several point measurements are introduced.

2.1 FORMS OF PRECIPITATION

Many forms of precipitation exist. Rain and drizzle are forms of liquid precipitation; snow, hail and graupel are forms of frozen precipitation. But also mixed forms occur, for example freezing rain. The freezing level is the altitude in the atmosphere at which the air temperature drops below 0° C. Precipitation that consists of solid and liquid forms of water is especially challenging for weather radar systems. Bright band is a catchword in this context and it will be explained in the next chapter. A closer disquisition on how precipitation forms and on processes like saturation, condensation and coalescence goes beyond the scope of this work. Suggested readings on these processes are Baumgartner and Liebscher (1996) (in German) and Davie (2003).

2.2 PRECIPITATION GENERATION TYPES

In the previous paragraph the forms of precipitation have been addressed. Forms of precipitation should not be mixed up with precipitation types. The latter term usually refers to the weather situations or mechanisms that lead to precipitation. In the literature this term is not unambiguously defined, sometimes it is also used in the sense of forms of precipitation. To make it clear and to avoid misunderstanding here the term “precipitation generation type” is applied.

Precipitation generation types are commonly categorised as convective, orographic or cyclonic. In Europe convective precipitation events occur mainly in the summer months caused by increased solar heating of the surface during the day. The columns of air above the warm surface become instable and in an updraft the air masses ascend. Convective precipitation tends to be associated with lightning and heavy rain, also hail is likely to occur. Convective clouds often are relatively small, and therefore the duration of convective precipitation is rather short. Figure 2.1 illustrates the development of convective precipitation.

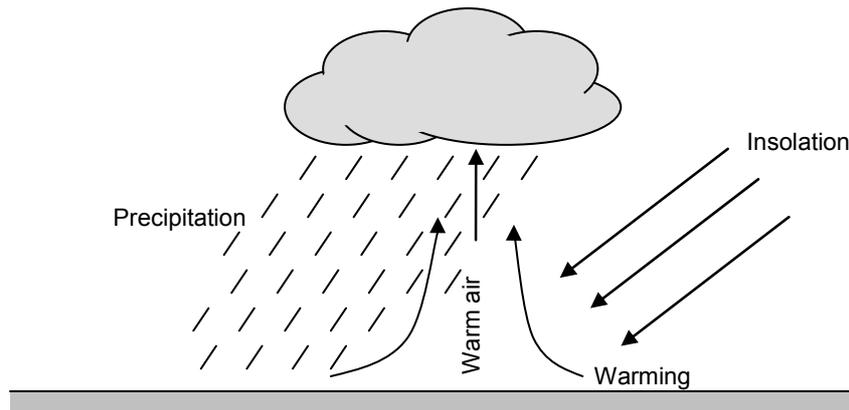


Figure 2.1. Convective precipitation.

Orographic precipitation occurs on the weather or windward side of a mountain range when air masses with high moisture content are forced to a higher elevation. The now cooler air cannot hold the humidity. As the air passes over the mountains precipitation occurs mainly on the windward slopes. Less or even no precipitation develops on the lee or leeward slope because the moisture is released as rain before. Moreover, the descending air increases its temperature and therefore more water can be held (Figure 2.2).

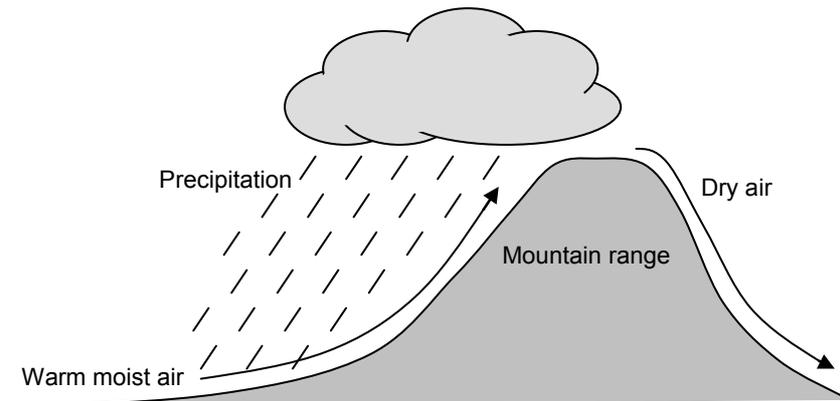


Figure 2.2. Orographic precipitation.

Cyclonic precipitation is generated by the general weather situation. A cyclone is a low-pressure area in contrast to an anticyclone which is an area with relatively high pressure. Sometimes the terms cyclonic and frontal precipitation are used synonymously (Deka, 2006), because cyclonic precipitation in many cases is bound to frontal events (Baumgartner and Liebscher, 1996), but in the strict sense cyclonic precipitation can either be frontal or nonfrontal. Here frontal cyclonic precipitation is described. A front is a boundary between air masses of different properties e.g. different temperature or different water vapour concentration. Cyclonic frontal precipitation occurs when either warm air masses are moving into a colder zone and are pushed upward by the cold air, or when cold air masses are moving into warm air masses and force them to raise (Figure 2.3). If the warm air moves towards the cold air it is called a “warm front”. In contrast if cold air moves towards the warm air masses it is called a “cold front”. According to Chang (2006, p. 141) cold fronts “usually

move faster, the frontal surfaces are steeper, their upward movements are more rapid, and precipitation rates are much greater than those of warm fronts”.

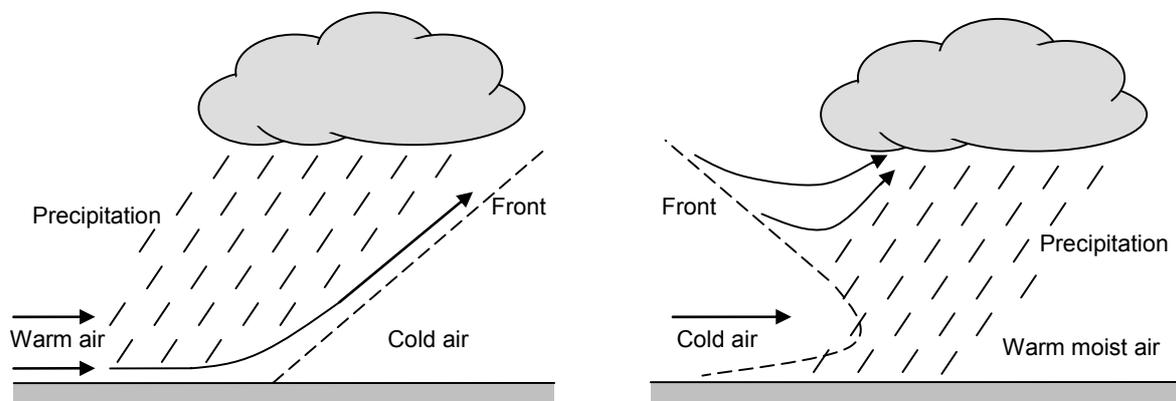


Figure 2.3. Two forms of cyclonic frontal precipitation.

2.3 MEASUREMENT OF PRECIPITATION

In hydrology it is indispensable to quantitatively record precipitation. The typical instrument used for this purpose is the rain gauge. Strictly speaking a rain gauge is an instrument to measure liquid precipitation as opposed to a snow gauge to measure frozen precipitation. But today's rain gauges are equipped with a de-icing system, effectively a heater, and so they also detect the liquid equivalent of frozen precipitation. Thus a typical rain gauge measures the height of water in millimetres (or litres per square metre).

2.3.1 TYPES OF RAIN GAUGES

One of the simplest types of rain gauges is the so-called ombrometer. It consists of a funnel as collector with a certain orifice area (typically 200 cm²) and a measuring cylinder. An observer determines the amount of precipitation at particular points of time, as a general rule daily, at 7 o'clock. Thus the observer gets a rain rate (height of precipitation per 24 hours). The temporal resolution is therefore rather poor and provides not much insight in the event that caused the precipitation. A rain rate of 30 mm per 24 hours for instance can be caused by a short but intense convective precipitation event as well as long-lasting orographic rainfall. With this type of rain gauge it is not possible to determine parameters like beginning and end of precipitation, duration of the rainfall event or progression of the rain rate, furthermore, real-time access to the data is not possible. A technology to overcome this drawback is the tipping-bucket rain gauge which is described below.

From the outside a tipping-bucket rain gauge is not much different from the simple ombrometer. It too consists of a funnel, but unlike the ombrometer the precipitation falls onto one of two buckets. The two buckets are arranged much like a pair of balances. If one bucket is filled with a predefined amount of water, the device tips and discharges the water. Because of the construction, now the other bucket is filled up until again the predefined amount of water is reached and the device tips again. It is the tips that contain the information. One tip is

equivalent to a certain height of precipitation, typically 0.1 mm. When the bucket tips an electric pulse is sent. This pulse can be recorded and transmitted to an external evaluation site. With this technology it is possible to monitor a network of telemetric rain gauges. Unlike the ombrometer, a tipping-bucket rain gauge allows determining beginning and end time of a precipitation event quite accurately and the course of the precipitation intensities can easily be monitored. But the technology of tipping-bucket rain gauges also has its disadvantages. The accuracy is an important point and therefore the next paragraph is dedicated to this issue.

2.3.2 MEASURING ERRORS OF RAIN GAUGES

First of all it has to be stated that rain gauges like the ones described in the previous paragraph are among the most exact instruments to measure precipitation. Rain gauges are a direct method to determine rainfall and they are used to validate and calibrate other methods such as weather radars. But even though rain gauges are used as calibration standard it is obvious that they are not free from errors. Typically it is distinguished between random and systematic errors. Random errors are due to fluctuations and due to changes in the vicinity of the rain gauge. Rainfall is a discrete process and therefore even two rain gauges side by side may produce different results. The most important source of errors according to Sevruck (1986 cited in ASCE, 1996, p. 27) are systematic errors. An installed rain gauge deforms the wind field and this is the main systematic error. Subsequently the systematic errors of rain gauges are described. Thereafter we take a closer look at the specific errors of the tipping-bucket rain gauge, since rain gauges of this type provided the rainfall data for this study.

Impact of the wind

Wind has an effect on the determination of the rainfall. The main source of error is that the wind field is deformed by the presence of an obstacle. The rain gauge is such an obstacle that the wind has to give way. This leads to higher wind speeds around the instrument. At the funnel also small turbulences can be observed. This brings about that small precipitation particles may not fall into the collector and therefore the actual precipitation is underestimated. The underestimation depends on wind speed, particle size distribution and design and placement of the rain gauge. The placement is important inasmuch as the wind speeds change significantly with height. It has often been observed that rain gauges near the ground level report more precipitation than neighbouring gauges that are elevated. This is also the reason why all rain gauges of a network are usually placed on the same height with respect to the ground. According to Dyck and Peschke (1995, p. 149) the error due to the deformation of the wind field can be 2 to 5 % for rain. But for snow the underestimation is especially pronounced and can lead to errors of 15 to 35 %.

To mitigate the error due to wind, gauges in windy areas are sometimes equipped with a wind deflector. Burton and Pitt (2001) report that these so-called shielded rain gauges perform slightly better when measuring rainfall but that they have “about half the magnitude of errors as unshielded gauges when monitoring snowfalls” (p. 386). For this study no specially equipped shielded gauges are used.

Wetting losses

Not every raindrop or snowflake which falls into the funnel is detected immediately. Precipitation that adheres to the funnel may never reach the measuring cylinder or the bucket or may reach it with a delay. This is known as wetting loss. The wetting loss depends on the material and design of the funnel. It is evident that the surface area of the collector is an important factor. Rain gauges with a larger collector may have a larger wetting loss than small ones. But as the orifice area is standardised, there is not much room for improvement. A more important factor is the material that is used for the funnel. It should be noted that by nature the funnel is exposed to the elements and therefore deterioration is an issue. Normally the wetting losses are rather small but they can approach 10 % depending on the form and frequency of precipitation. Wetting losses are most significant when it starts to drizzle and the surface of the funnel is dry.

Evaporation losses

Evaporation loss refers to precipitation that falls into the collector but fully evaporates before it is detected. This term is also used for water that evaporates from the measuring cylinder (ombrometer). But the last case is negligible because of good design of the instrument itself that minimises ventilation. Most significant are evaporation losses when light rain falls into a warm collector. In the winter season heated gauges can also pose a problem. Hanson *et al.* (1983 cited in ASCE, 1996, p. 31) investigated tipping-bucket gauges and recommends to maintain the collector temperature as low as possible to minimise the undercatch by heated tipping-bucket gauges.

Rain splash

The phenomena described so far lead to an underestimation or undercatch of the actual rainfall amount by the rain gauge. Rain splash can lead to under- as well as overestimation. Therefore, two forms of rain splash are distinguished according to ASCE (1996): splash-in and splash-out. Splash-in may occur if the gauge orifice is placed close to a surface that produces splashing. Splash-out is likely when the gauge orifice is large in relation to the depth of the gauge. Thus rain splash can be minimised by a proper design and placement of the rain gauge. An elevated rain gauge will not suffer from splash-in but on the other hand wind effects will become more significant. So often a compromise must be found. One possibility is a low elevation above the bare ground together with the use of special carpets that minimise splashing around the rain gauge.

Specific errors of the tipping-bucket rain gauge

The previously described errors are related to all kinds of rain gauges, because rain splash, evaporation, wetting and wind effects concern the collector of the gauge – the funnel. But some sources of errors are due to the measuring mechanism of the tipping bucket rain gauge, and they are characterised in the next paragraphs.

This device tips and discharges once the bucket is filled with a defined amount of water. If the rainfall stops, there is normally some water in one bucket but too little to cause the mechanism to tip. This amount of water may evaporate if there is a longer dry period, and the water is not recorded. But because of the design that minimises ventilation and thus evaporation, there is very likely still water in one of the buckets and if the next precipitation event starts a smaller amount of water is sufficient to cause the mechanism to tip. The consequence is that certain amounts of water are assigned to the wrong precipitation events.

Another systematic error of tipping bucket rain gauges can occur during heavy rainfall events. Davie (2003) stresses the importance of the correct size of the tipping-buckets for the prevailing conditions. “If the buckets are too small then a very heavy rainfall event will cause them to fill too quickly and water be lost through overspill while the mechanism tips” (p. 22).

Sevruk (1996) who has investigated the differences in precipitation values between the tipping bucket system and a standard ombrometer reports of higher wind speeds above the centre of the tipping bucket gauge because of its bigger body.

In addition to the above mentioned systematic errors also some random errors are more likely to occur because of the more complex mechanism of the tipping-bucket rain gauge. Sevruk (1996) mentions “clogging of the tipping-bucket gauge outflow” and “mechanical and electronical disturbances” (p. 243).

Besides systematic errors which can not be eliminated but minimised by the design of the rain gauge, especially the random errors listed in the previous paragraph have to be taken seriously. Thus careful maintenance and servicing is indispensable to obtain reliable data. The rain gauge data used in the course of this work originate from the Office of Styrian Regional Government, Hydrographic Department 19A in Graz, Austria. This institution also monitors and verifies the data.

2.4 VARIABILITY OF PRECIPITATION

In general precipitation is characterised by a high spatial and temporal variability. The variability fundamentally depends on the precipitation generation type. A convective precipitation event with a small shower cell has the highest variability. Within minutes the situation can change. High rainfall intensities are measured while the centre of the shower cell is located aloft a certain location, and the minute the cell moves away the rainfall is practically over. A site close to this location may not be affected by rainfall at all. Cyclonic precipitation events do not show such a high variability. The precipitation distribution in this case is more uniform. Naturally the orography also has an influence. This becomes apparent during orographic precipitation events with much rainfall on the windward side of the mountain and less or even no precipitation on the leeward slope.

What is necessary to record the variability of a precipitation event? To determine the temporal variability obviously the temporal resolution of the measuring system is crucial. An ombrometer that is only emptied once per day is certainly inadequate to study the variations during a single rainfall event but may be used to study variations during the year (between winter and summer months for instance). The tipping-bucket rain gauge tips after a defined amount of precipitation has fallen (0.1 mm for the gauges used here), thus the temporal

resolution depends on the rainfall intensity. During heavy rainfall the gauge may tip several times per minute which is certainly enough to capture the temporal variability of single precipitation events.

To record the spatial variability one single rain gauge, regardless of technology, falls far short. A network of rain gauges is necessary to capture the spatial variability of a certain region. The number of gauges depends on the prevailing meteorological conditions and the orography. As a matter of course the intended meteorological or hydrological purpose is decisive. For the measurement of rainfall amounts for agricultural purposes a smaller number of gauges will be satisfactory than for a hydrological investigation area. The closer the rain gauges are to each other the more information about the fine variations of precipitation is captured.

2.4.1 SPATIAL AND TEMPORAL VARIABILITY OF PRECIPITATION IN THE STUDY AREA

In this paragraph the variability of precipitation in the study area is investigated. This area lies in the south eastern part of Austria. Merz *et al.* (2001) report that short convective storms are dominant flood producing processes in this part of Austria. Wölfelmaier and Zwatz-Meise (2005) who analysed the life cycle of convective cells with satellite and radar data state that the southeastern Austrian province of Styria has a high frequency of thunderstorms and that “Heavy thunderstorms are often organised in multi-cell storms” (sec. 2). These references suggest that spatial and temporal variability of precipitation are quite high in the study area. Here this presumption is investigated with data of a rain gauge network.

The density of the rain gauge network in the study area is approximately of the order of one rain gauge station per 100 km². Data from eleven rain gauges are available.

Figure 2.4 shows their location. The rain gauges are situated at altitudes between 320 and 1245 m above the mean Adriatic Sea level. The mean annual precipitation obtained from these gauges in the years 2002 and 2003 varies from about 740 to 1130 mm. The relationship between annual precipitation and altitude of the rain gauge is shown in Figure 2.5. There is considerable scatter also because of the limited observation period, but the trend that the precipitation amount increases with the elevation can be seen.

Next the correlation of the time series of the different rain gauges is discussed. Correlograms of gauge-to-gauge rainfall can provide an informative basis for the variability of precipitation in an area. Figure 2.6 shows a gauge-to-gauge correlogram over a two-year period. A least squares, two-coefficient exponential function of the form

$$f(x) = a \cdot \exp(-x/b) \quad (2.1)$$

was fit to the data according to Young *et al.* (2000). It can be seen that the agreement between the rain gauges decreases by trend with increasing inter-gauge distances. Even between neighbouring stations which are less than 10 km apart, the correlation coefficient rarely lies above 0.5.

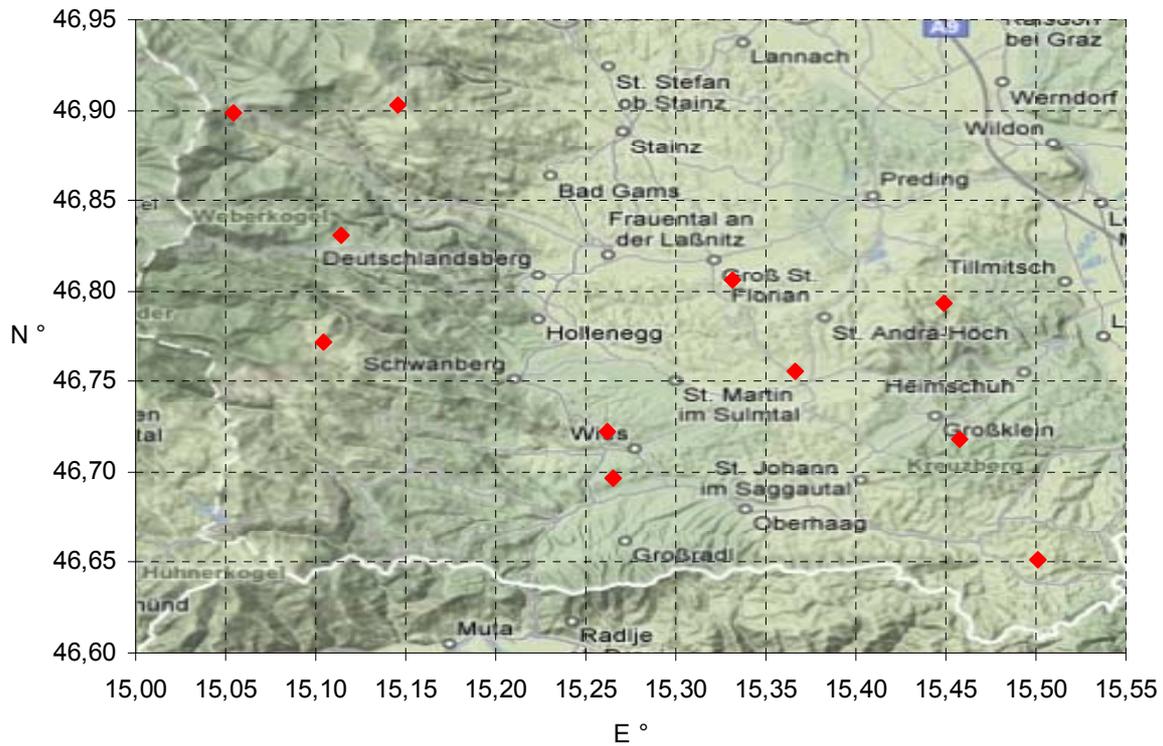


Figure 2.4. Positions of 11 available rain gauges in the geographic coordinate system (Underlay: Google © 2008).

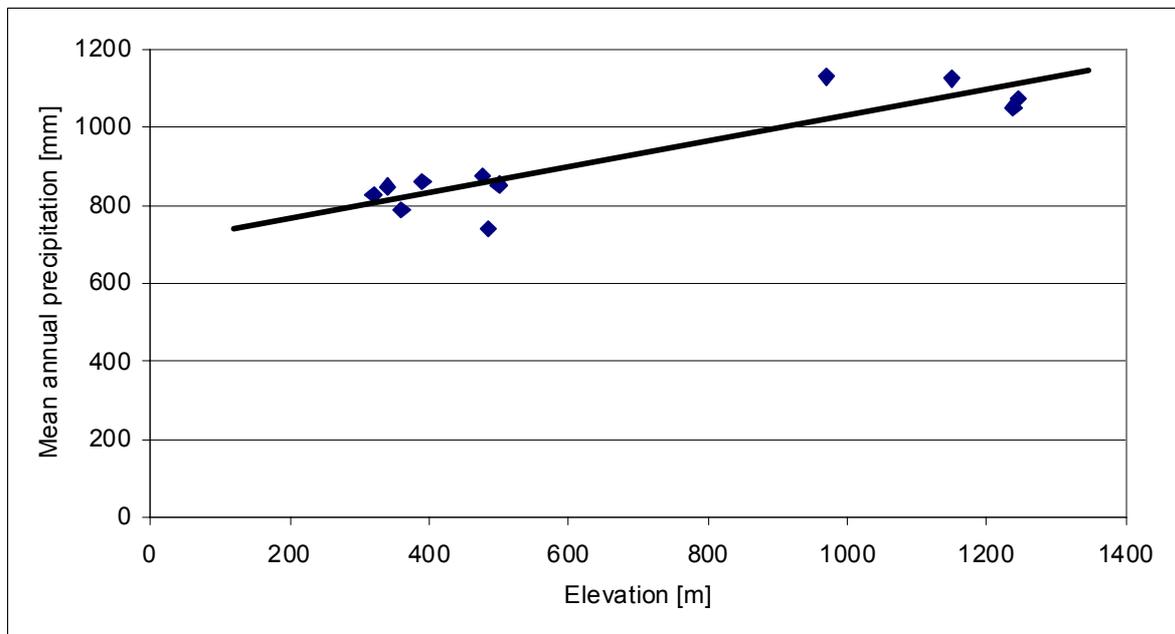


Figure 2.5. Mean annual precipitation versus elevation of the rain gauges. The data cover the years 2002 and 2003.

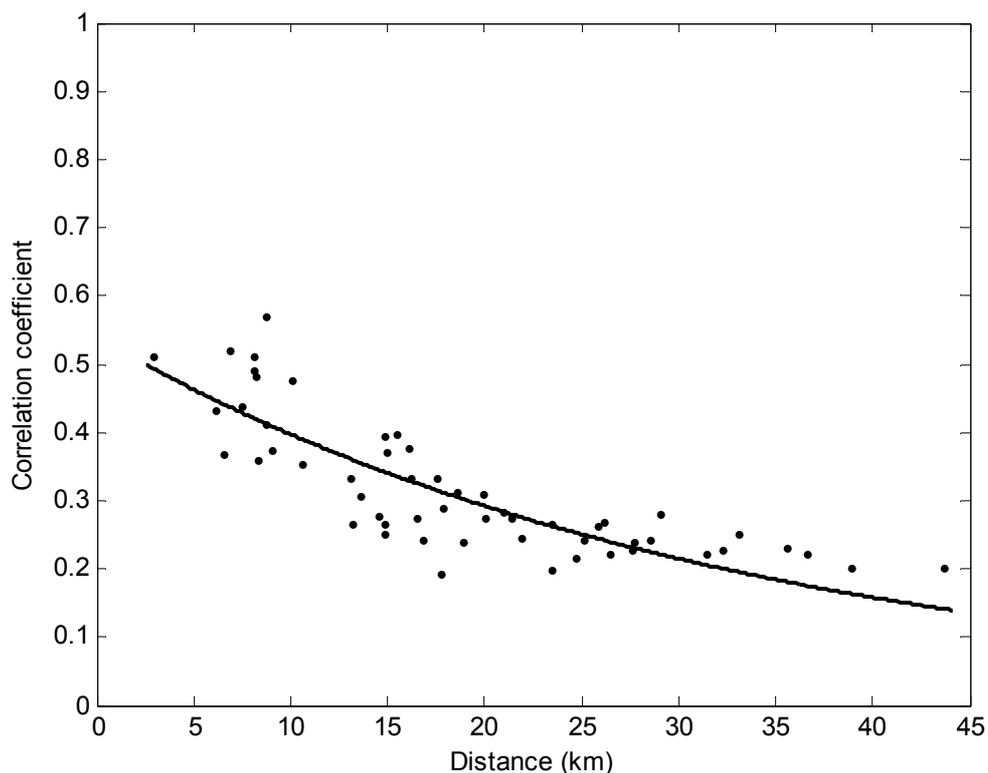


Figure 2.6. Gauge-to-gauge correlogram for 2002-2003. The integration interval of the rain gauge data is 15 minutes. Each point represents the correlation coefficient between the time series of two rain gauges separated by the inter-gauge distance. The solid line represents a least squares fit.

The existence of seasonal variations is investigated next. As stated above the variability also depends on the precipitation generation type. Thus one can expect seasonal changes in the correlogram. The variability of convective precipitation that often occurs in summer will be higher, thus the correlation between the rain gauges is supposed to be lower. Figure 2.7 shows the quarterly correlograms. It can be seen that the seasonal variations are quite pronounced. On average the lowest correlation between the gauges was measured in the second and third quarter. This is ascribed to convective precipitation events that often occur in this time of the year.

The highest correlation coefficients were measured in the last quarter. Here even at distances over 30 km the correlation coefficient frequently exceeds 0.5. This is all the more worth mentioning as the integration time is 15 minutes. It is known that “temporal differences tend to average out over longer time periods” (Quina, 2003, p. 31) and therefore one would expect such high values in hourly or daily data. The values point out a rather uniformly distributed rainfall over the whole study area.

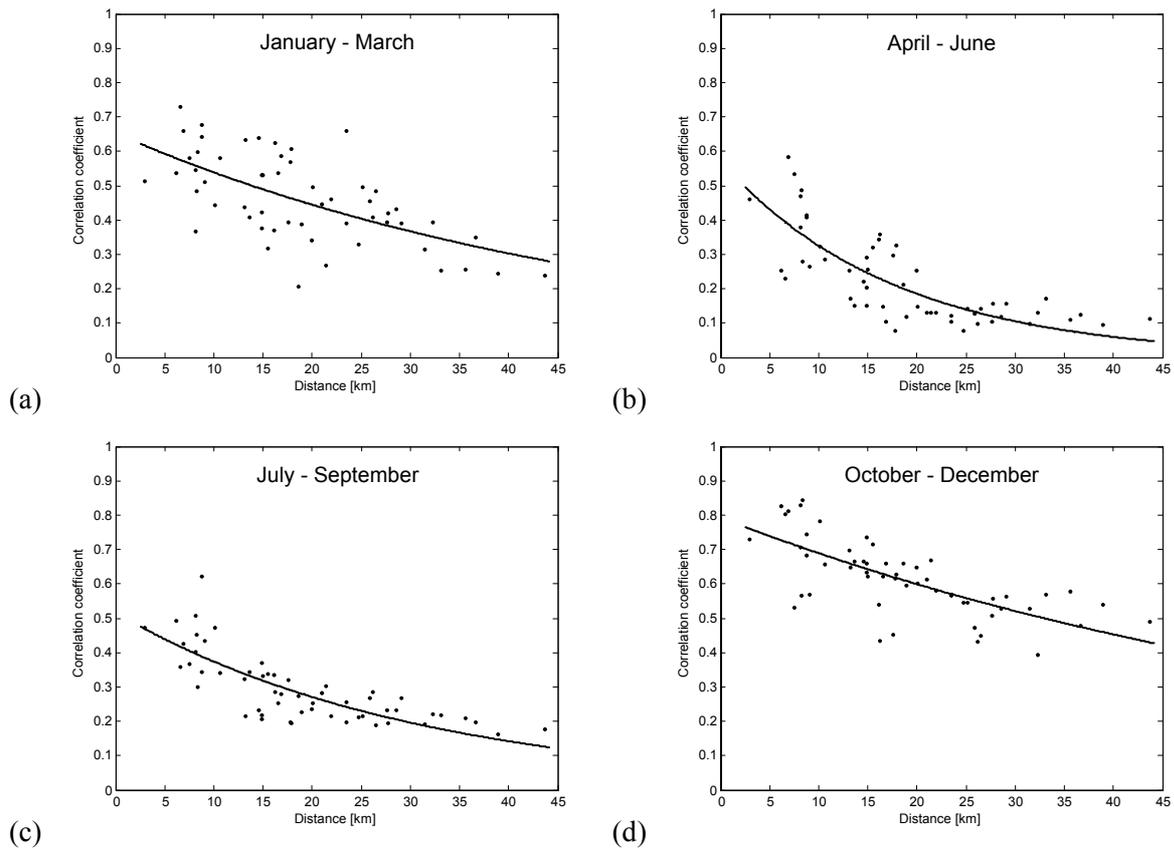


Figure 2.7. Quarterly gauge-to-gauge correlograms. The data represent (a) the first, (b) the second, (c) the third, and (d) the fourth quarters of 2002 and 2003.

The temporal variation of precipitation is investigated next. This investigation is based on auto-correlation plots of rain gauge data series. Figure 2.8 shows the auto-correlation of the eleven rain gauge stations with respect to the time lag. It can be seen that the auto-correlation coefficient decreases quickly with time. At the 15 minute time lag the mean correlation coefficient lies at about 0.55. This indicates a high temporal variation. As it was again expected that there are significant seasonal changes, the auto-correlation functions were plotted separately for the four quarters. In

Figure 2.9 it becomes obvious that the auto-correlation coefficient declines most in the second and third quarter. This means that the precipitation amount at the same station changes quite rapidly. Even after only 15 minutes the mean correlation lies at about 0.5. In the first and the last quarter, the decrease is not so dramatic. The auto-correlation is much higher. It lies around 0.8 after 15 minutes and in the mean over 0.5 after one hour.

Summing up it can be said that the spatial and temporal variability of precipitation in the study area is high and shows significant seasonal changes. Between April and September the variability is highest. This is mainly ascribed to convective precipitation with rather small shower cells or multi-cell storms.

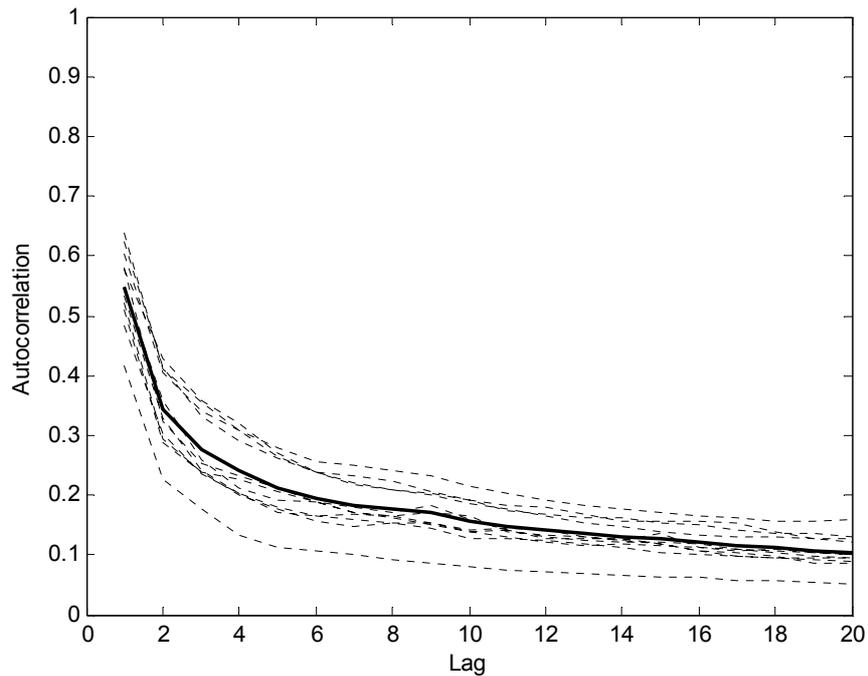


Figure 2.8. Auto-correlation of the rain gauge stations with respect to the 15-minute time lag. The data represent the years 2002 and 2003. The solid line illustrates the mean auto-correlation of all rain gauges time series.

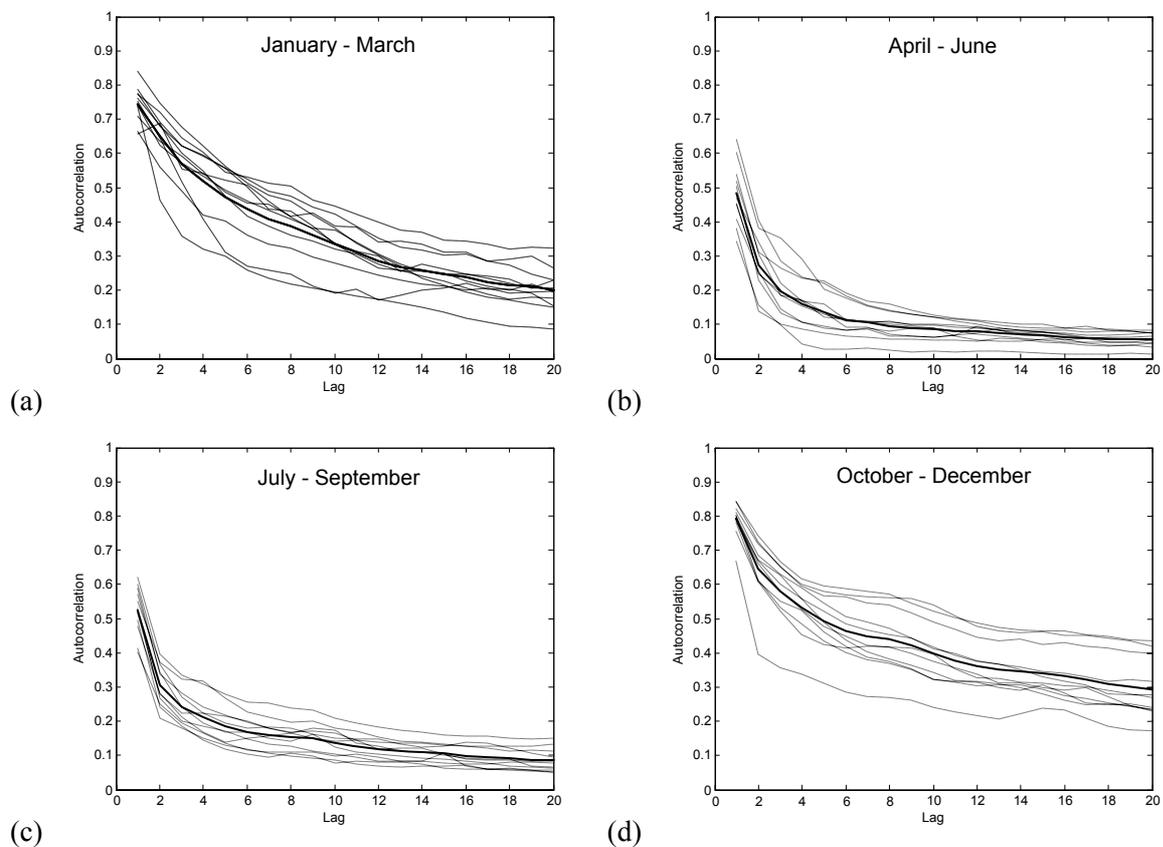


Figure 2.9. Quarterly auto-correlation plots of the rain gauge stations with respect to the 15-minute time lag. The data represent (a) the first, (b) the second, (c) the third, and (d) the fourth quarters of 2002 and 2003.

2.5 POINT MEASUREMENTS AND AREAL ESTIMATION

The precipitation measurement technique with rain gauges described so far is called point measurement, because the precipitation amount is determined at a selected location. Hydrologists are often interested in the amount of precipitation over a larger region. Therefore, the information that is available from the rain gauges has to be used and combined in a proper way to estimate the areal precipitation. This section gives an overview of standard methods for this purpose.

The most basic approach is the simple arithmetic mean. The rainfall depth of all available rain gauge stations is added and divided by the number of gauges thus one uniform value is assigned to the whole area (Figure 2.10). This approach is only advisable if the rain gauges are evenly spread over the area of interest and do not show considerable fluctuations.

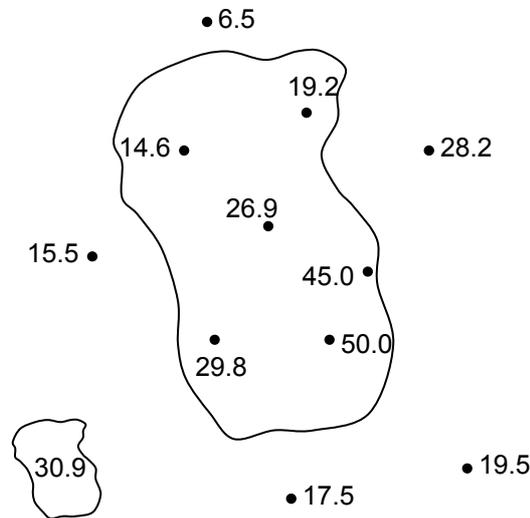


Figure 2.10. Arithmetic mean of measured rainfall depths. Adapted from Kainz (2002).

Generally rain gauges are not evenly distributed over the region. A method to meet this concern is the Thiessen polygon method named after the American climatologist Alfred H. Thiessen. This method is characterised by representative areas which are assigned to the rain gauge stations. This method can be easily shown on a map (Figure 2.11). All rain gauge stations are connected by straight lines. Now the perpendicular bisectors of these lines between gauges are drawn. The polygons are finally formed by the boundaries that are equidistant from two rain gauges. Thus representative areas emerge. All points of this area are closest to the assigned rain gauge relative to all others. The rainfall depth measured at one rain gauge is assigned to the surrounding area formed by the polygon. According to Davie (2003, pp. 23-24) “This technique is only truly valid where the topography is uniform within each polygon”.

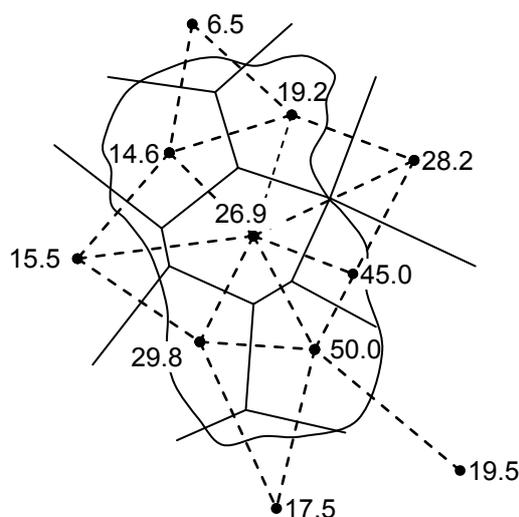


Figure 2.11. Thiessen polygon method. Adapted from Kainz (2002).

If the rain gauge network is dense enough, lines of equal precipitation (isohyets) can be interpolated from the point measurements. Hence, this technique is known as the isohyetal method. It involves determining the area between successive isohyets for which a uniform rainfall value is assigned (Figure 2.12). The advantage of this method is that known geologic realities (e.g. orographic effects) can be accounted for and therefore the isohyetal method is sometimes labelled the most accurate of the methods presented so far. But the result largely depends on the skills of the hydrologist who draws the isohyets.

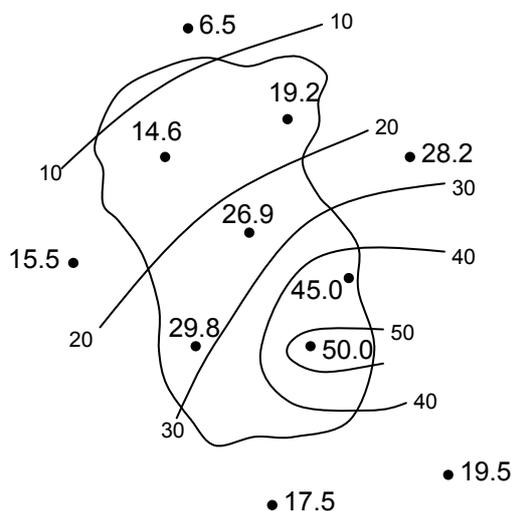


Figure 2.12. Isohyetal method. Adapted from Kainz (2002).

Another method to determine the rainfall amount at unobserved locations and thus to calculate the areal rainfall from point measurements by rain gauges is Kriging. It is a geostatistical method for spatial interpolation named after the South African mining engineer D. G. Krige. Initially conceived to evaluate mineral resources, Kriging has been applied to

various scientific disciplines including hydrology and environmental sciences. Spatial interpolation is important in these fields, since it is impossible to get data at every desired point because of practical reasons. Compared to the methods described so far, Kriging is a computationally demanding approach. The key element is to calculate the semivariogram to model the spatial dependency. Kriging can also assess the quality of prediction with estimated prediction errors. There are different variations of Kriging. Cokriging for example is Kriging with more than one variable, it combines spatial data on several variables to make the interpolation. A detailed description of Kriging and its variations can be found in Cressie (1993).

3 THE WEATHER RADAR

During World War II military microwave radars detected – besides echoes from aircraft – also echoes from rainstorms. As a consequence of this finding, the first radars specifically designed for meteorological purposes were developed after the war. Since that time major milestones on the way to today's weather radars were the development of Doppler and dual-polarization radars. Nowadays the weather radar is an indispensable tool in meteorology and hydrology.

This chapter describes the principles of the weather radar, gives an introduction to drop-size distribution models which are essential to determine the relationship between radar echo and rain rate, and lists common errors and limitations of weather radar measurements. Finally the specifications of the Austrian weather radar network that provided the data for this study are listed. The description is by no means comprehensive. However, the presentation shall give, together with Chapter 4, an indication for the complexity of the extraction of rain rates from radar data.

The reader is referred to Atlas (1990) for further reading about the history of weather radars, Doviak and Zrnic (1984) for Doppler radars, and Bringi and Chandrasekar (2001) for dual-polarization radars.

3.1 PRINCIPLE

Radar stands for radio detection and ranging. A weather radar is a special type of radar that is used to detect and quantify tropospheric precipitation. Weather radars send out pulsed directional microwaves. When the microwave encounters precipitation particles one part of the wave's energy is absorbed and the other part is scattered – typically in all directions. In general only a small fraction of the wave's energy is reflected back to the radar receiver. The larger the particles and the higher their concentration, the stronger is the reflected signal.

The duration τ of the transmitted pulse is on the order of microseconds. The pulse length h can be calculated by

$$h = c \cdot \tau \tag{3.1}$$

where

c speed of light [m s^{-1}], speed of light in vacuum: $c_0 = 2.99792458 \cdot 10^8 \text{ m s}^{-1}$
 τ pulse duration [s]

After a pulse is transmitted, the radar acts as a receiver and detects signals reflected back to the radar. The duration of this cycle is on the order of one millisecond. Given an operational range of the radar of 300 km the cycle must be long enough for the wave to propagate from the radar to a possible target 300 km away and again back to the radar, hence 2 ms.

At a certain time after transmitting a microwave pulse, the weather radar receives backscattering contributions from scatterers at certain range limits.

Assuming the radar starts to send out a microwave pulse at the time $t = 0$ with a duration τ . At the time $t = \tau$, the radar receives contributions from scatterers in a range

$$\left\{ 0 \dots \frac{c \cdot \tau}{2} \right\} = \left\{ 0 \dots \frac{h}{2} \right\}$$

Generally at a time $t = t_e$ the range limits are

$$\left\{ \left(\frac{c \cdot t_e}{2} - \frac{h}{2} \right) \dots \frac{c \cdot t_e}{2} \right\}$$

The distance between upper and the lower range limit is always the half pulse length ($h/2$) which is defined as the length of the range gates.

The weather radar scans a large volume of the atmosphere by pointing the antenna in all directions of interest. In the radar raw data the position of targets is stored in polar coordinates (azimuth and elevation angle, and range) as illustrated in Figure 3.1.

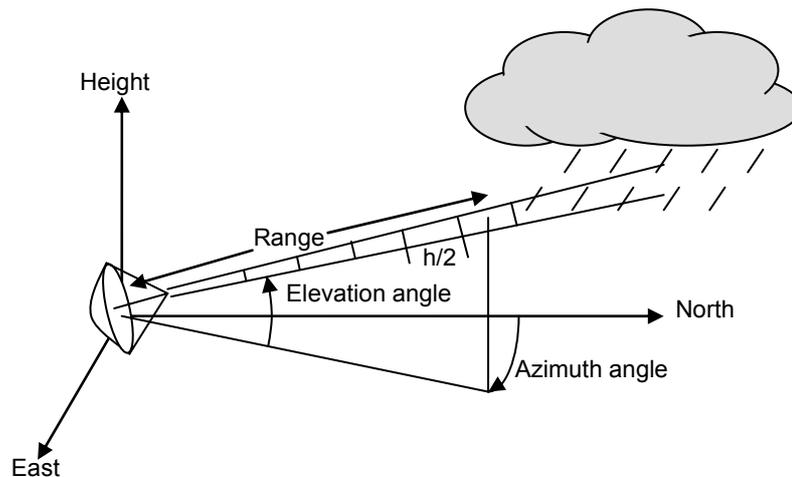


Figure 3.1. Azimuth, elevation and range of the weather radar echoes.

For each range gate the radar measures the radar reflectivity which is proportional to the ratio of the reflected to incident energy. When focusing only on precipitation particles as targets, the radar reflectivity depends primarily on the material (water or ice), on the size, shape and orientation of the particle and on the total number of particles per unit volume. Thus, relating the measured reflectivity to an amount of precipitation hinges on several assumptions. One of these assumptions is a simplified drop-size distribution. Below it is explained how 2-parametric and 3-parametric drop-size distributions are defined.

3.2 DROP-SIZE DISTRIBUTION

Drop-size distributions give the frequency of occurrence of drops of a certain size per unit volume of precipitation. The drop-size distribution is different for different precipitation generation types (e.g. convective and orographic precipitation). Marshall and Palmer (1948) experimentally found following exponential form of the drop-size distribution:

$$N(D) = N_0 \exp(-A D) \quad (3.2)$$

where

D drop diameter [mm], $0 \leq D \leq D_{\max}$

$N(D)$ number density of drops per unit volume [$\text{mm}^{-1} \text{m}^{-3}$]

N_0 zero intercept [$\text{mm}^{-1} \text{m}^{-3}$]

A slope of the distribution [mm^{-1}]

This distribution is called exponential or 2-parametric. Marshall and Palmer (1948) suggested that zero intercept N_0 was constant (3.3) and they related the parameter A to the rain rate R (3.4).

$$N_0 = 8000 \text{ mm}^{-1} \text{ m}^{-3} \quad (3.3)$$

$$A = 4.1 R^{0.21} \text{ mm}^{-1} \quad (3.4)$$

This Marshall-Palmer drop-size distribution for various rain rates is given in Figure 3.2.

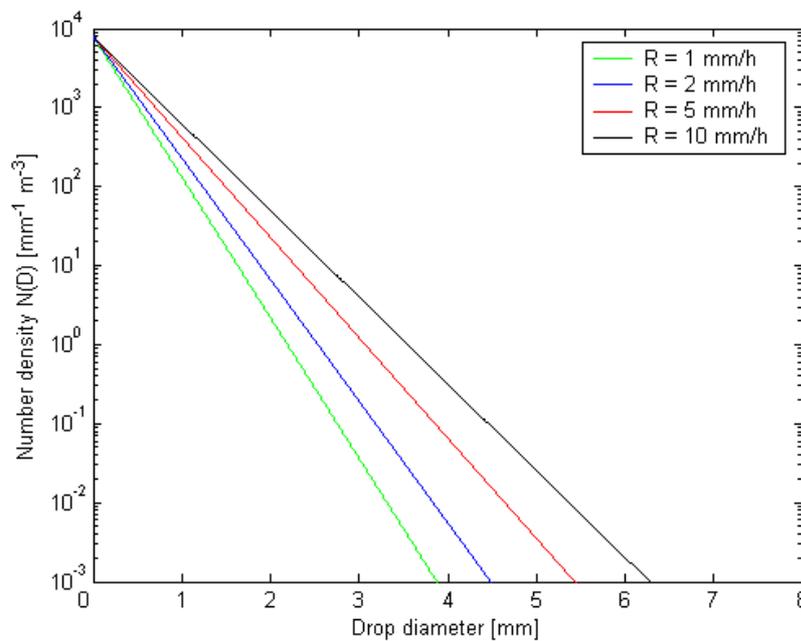


Figure 3.2. Marshall-Palmer drop-size distribution for various rain rates ($R = 1, 2, 5$ and 10 mm/h , $N_0 = 8000 \text{ mm}^{-1} \text{ m}^{-3}$, $A = 4.1 R^{0.21} \text{ mm}^{-1}$). Adapted from Xie (1988).

Later it has been found that sudden changes in N_0 can occur within a given type of rainfall (cf. Ulbrich, 1983). To account for these changes a more complex 3-parametric drop-size distribution was proposed in the form:

$$N(D) = N_0 D^\mu \exp(-A D) \quad (3.5)$$

where

D drop diameter [mm], $0 \leq D \leq D_{\max}$

$N(D)$ number density of drops per unit volume [$\text{mm}^{-1-\mu} \text{m}^{-3}$]

N_0 [$\text{mm}^{-1-\mu} \text{m}^{-3}$]

A slope of the distribution [mm^{-1}]

μ parameter that can take positive and negative values

This gamma function with its three parameters (N_0 , A and μ) can describe variations in the distribution more precisely than the 2-parametric Marshall-Palmer drop-size distribution. The parameter μ ranges approximately from -3 to 5 and N_0 can reach values up to the order 10^{10} (cf. Ulbrich, 1983)

A meaningful parameter for a given drop-size distribution is the median volume diameter D_0 . It is “that diameter for which the total volume of all drops having greater diameters is just equal to the total volume of all drops having smaller diameters” (AMS, 2000). The median volume diameter can be related to parameters of the drop-size distribution. Ulbrich (1983) found the following approximate expression

$$A D_0 = 3.67 + \mu \quad (3.6)$$

As a consequence the gamma distribution can be written as

$$N(D) = N_0 D^\mu \exp(-(3.67 + \mu) D/D_0) \quad (3.7)$$

Figure 3.3 shows gamma distributions with a varying parameter μ . A zero value of μ effectively reduces the gamma distribution (Eq. 3.5) to a Marshall-Palmer distribution (Eq. 2.3). A value $\mu > 0$ results in a distribution with relatively fewer small and large drops than in Eq. 3.2 (cf. Brandes, 2000). On the other hand a negative value of μ leads to a distribution with relatively more small and large drops.

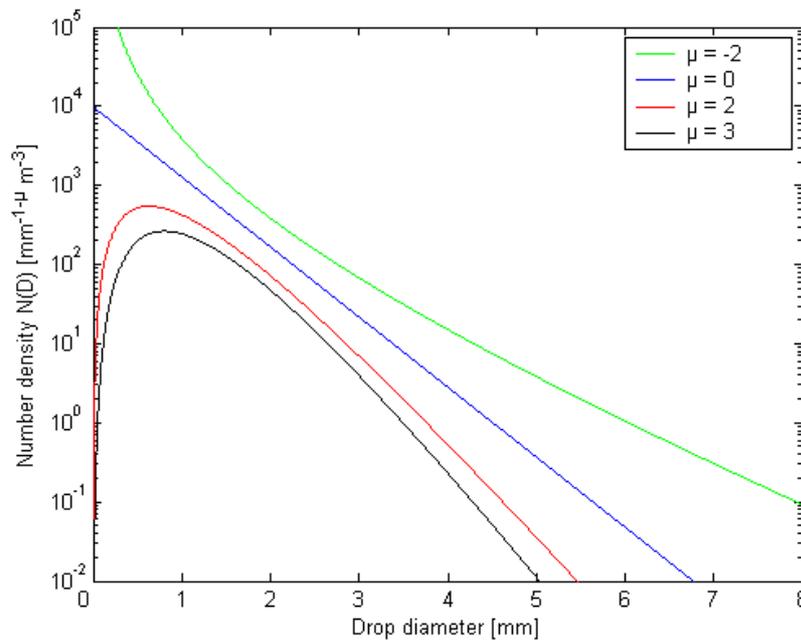


Figure 3.3. Gamma drop-size distribution ($N_0 = 10000 \text{ mm}^{-1-\mu} \text{ m}^{-3}$, $D_0 = 1.8 \text{ mm}$, and varying μ). Adapted from Xie (1988).

3.3 RADAR REFLECTIVITY FACTOR

The weather radar measures the radar reflectivity factor Z . This factor equals the sum of the sixth-powers of the diameters of all rain drops contained in a unit volume.

$$Z = \sum_{i=1}^N D_i^6 \quad [\text{mm}^6 \text{ m}^{-3}] \quad (3.8)$$

where

N number of raindrops per unit volume

Given the drop-size distribution as a continuous function $N(D)$, e.g. an exponential or gamma distribution, the reflectivity factor Z can be written as the integral

$$Z = \int_0^{D_{\max}} N(D) D^6 dD \quad [\text{mm}^6 \text{ m}^{-3}] \quad (3.9)$$

Thus, Z is proportional to the number of rain drops and proportional to the sixth power of the drop diameters. Therefore, Z is very sensitive to the size of the precipitation particles.

3.4 MEASURING RAIN RATE WITH WEATHER RADAR

A weather radar determines the rain rate R from the measured reflectivity factor Z . Also the rain rate R can be defined based on a drop-size distribution. According to Brandes (2000) R can be computed from

$$R = 6\pi \times 10^{-4} \int_0^{D_{\max}} D^3 v_t(D) N(D) dD \quad [\text{mm h}^{-1}] \quad (3.10)$$

where

D drop diameter [mm]
 $N(D)$ drop spectrum [$\text{mm}^{-1-\mu} \text{m}^{-3}$]
 $v_t(D)$ drop terminal velocity [m s^{-1}]

In radar meteorology empirical relationships between Z and R of the following form were established:

$$Z = a \cdot R^b \quad (3.11)$$

where

Z reflectivity factor [$\text{mm}^6 \text{m}^{-3}$]
 R rainfall rate [mm h^{-1}]
 a and b empirical parameters

A variety of Z-R relationships have been derived for various locations and weather conditions. Several dozens of them are listed in Battan (1973). The parameter a varies from 16.6 to 730 and b from 1.16 to 2.87.

For the different types of rainfall Battan (1973, p.89) considers following Z-R relationships to be “fairly typical”:

$$\text{stratiform rain} \quad Z = 200 R^{1.6} \quad (3.12)$$

$$\text{orographic rain} \quad Z = 31 R^{1.71} \quad (3.13)$$

$$\text{convective rain} \quad Z = 486 R^{1.37} \quad (3.14)$$

In case the radar measures the reflectivity of falling snow above the melting layer the relationship

$$Z = 2000 R^2 \quad (3.15)$$

was found by Gunn and Marshall (1958). In this relationship R is the rain rate of the melted snow.

3.5 WEATHER RADAR FREQUENCIES

The frequencies of the radio waves used for weather radars range from 1.5 GHz to above 30 GHz (cf. Battan, 1973). Frequencies around 22.2 GHz are not used since this is the resonance wavelength of water vapour where absorption can significantly reduce the range (cf. Skolnik,

1990). Frequencies higher than 30 GHz are mainly used to detect cloud particles. In Table 3.1 the frequency bands of weather radars are listed.

Table 3.1. Frequency bands of weather radars.

Frequency range [GHz]	Wavelength range [cm]	Band designation*
1 – 2	15 – 30	L
2 – 4	7.5 – 15	S
4 – 8	3.75 – 7.5	C
8 – 12	2.5 – 3.75	X
12 – 18	1.67 – 2.5	K _u
18 – 27	1.11 – 1.67	K
27 – 40	0.75 – 1.11	K _a

*according to IEEE Standard 512-2002

The proper choice of the frequency or wavelength respectively for a certain radar depends on the size of the precipitation particles that should be detected. For a reliable estimation of the rain rate R the wavelength of the weather radar has to be significantly longer than the size of the precipitation particles. It was stated above that the radar reflectivity factor Z is proportional to the sixth-power of the diameters of the particles. Strictly speaking this is only true if the particles are at least 10 times smaller than the wavelength λ (Rayleigh approximation). Table 3.2 lists the maximum diameters of Rayleigh targets for common radar frequencies.

Table 3.2. Maximum particle diameter of Rayleigh targets for weather radar frequencies.

Frequency [GHz]	Wavelength λ [cm]	$D_{\max} = \lambda / 10$ [cm]
2.8	10.7	1.07
5.625	5.3	0.53
9.6	3.1	0.31
14	2.1	0.21
35	0.9	0.09

3.6 SCANNING MODES AND RADAR DISPLAYS

From above we know that weather radars measure the reflectivity with a directive antenna. By changing elevation and azimuth angle the weather situation up to several hundred kilometres around the radar site and up to a height of approximately 16 km is detected.

In radar meteorology there are two main types of scanning modes. These are the Range Height Indicator (RHI) scan and the Plan Position Indicator (PPI) scan. Since each of these scans produces a certain radar display also the corresponding displays are named RHI and PPI respectively.

3.6.1 RANGE HEIGHT INDICATOR (RHI)

When scanning in RHI mode, the radar only varies the elevation angle of the antenna with the azimuth angle fixed. The antenna normally scans from angles near the horizon to angles near the zenith. The RHI display shows the return echoes on a vertical plane. RHI scans provide a useful insight in the vertical profile of clouds and precipitation. On an RHI display e.g. the height of the cloud tops and the height of the melting layer can be seen. Figure 3.4 shows a schematic RHI plot. The raw radar data yield the reflectivity for every range gate. In today's weather radar systems the raw data are converted into a cubic grid as demonstrated in Figure 3.5. An RHI plot of an observed convective precipitation event is shown in Figure 3.6.

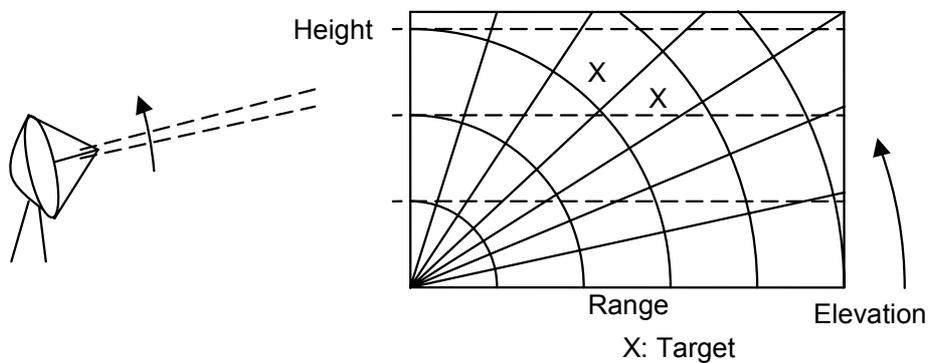


Figure 3.4. Schematic RHI plot.

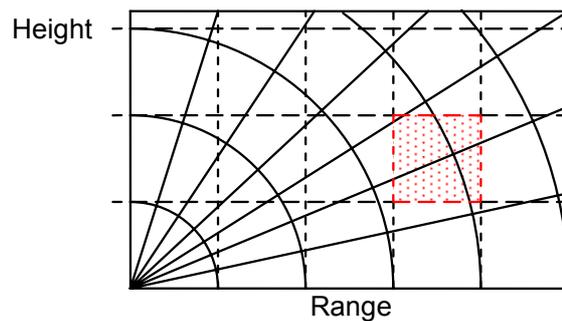


Figure 3.5. Conversion of raw data into a cubic grid.

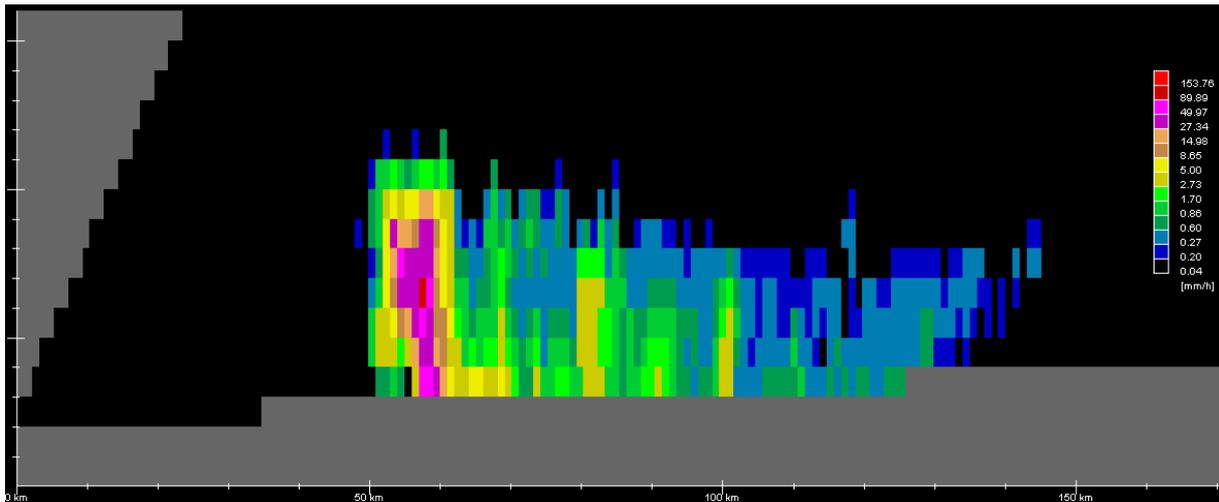


Figure 3.6. RHI plot of a convective precipitation event. The weather situation occurred on June 25, 2001, and was recorded by the weather radar station on Mt. Zirbitzkogel (Source: Austro Control GmbH).

3.6.2 PLAN POSITION INDICATOR (PPI)

When scanning in PPI mode, the radar only varies the azimuth of the antenna with the elevation angle fixed. The return echoes are mapped onto a horizontal plane. In the middle of the plane usually the position of the radar is indicated. Concentric circles indicate the range from the radar. Usually, north is at the top of the image. Figure 3.7 shows a schematic PPI scan.

3.6.3 CONSTANT ALTITUDE PLAN POSITION INDICATOR (CAPPI)

A further development of the PPI is the CAPPI. CAPPI is a composite radar display that shows precipitation on a constant altitude above ground. A CAPPI display is calculated from PPI scans at successive elevation angles. In regions near the radar site where there are typically gaps between the radar beams (see Figure 3.8) interpolation algorithms are used to process a continuous CAPPI display. PPI and CAPPI displays give a good overview of the general weather situation. Figure 3.9 presents CAPPI plots of two different altitudes. Because the lowest radar beam has an elevation angle of about one degree, the maximum range increases with the altitude of the CAPPI levels.

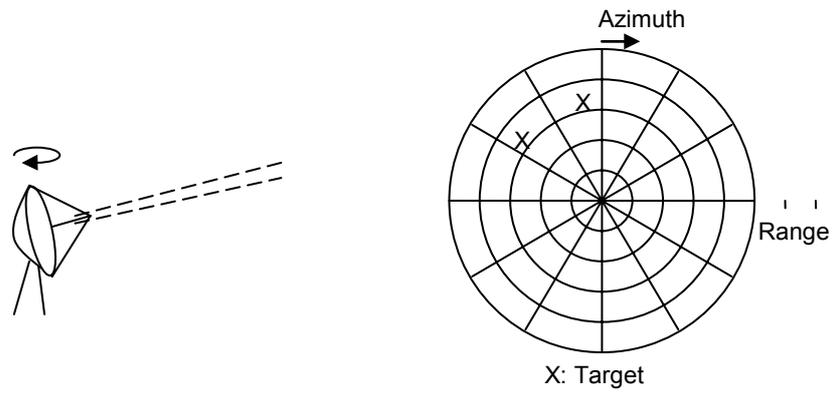


Figure 3.7. Schematic PPI scan.

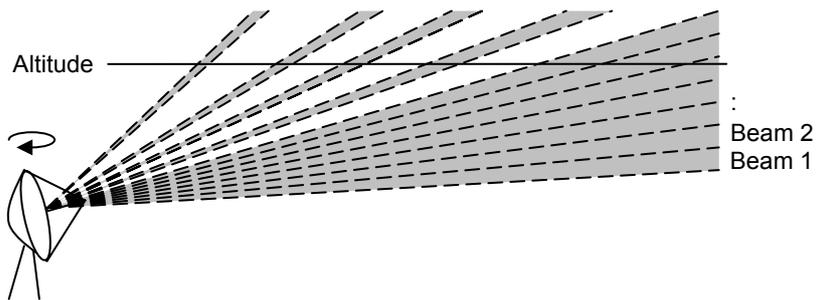


Figure 3.8. Procession of a CAPPI display from PPI scans at successive elevation angles.

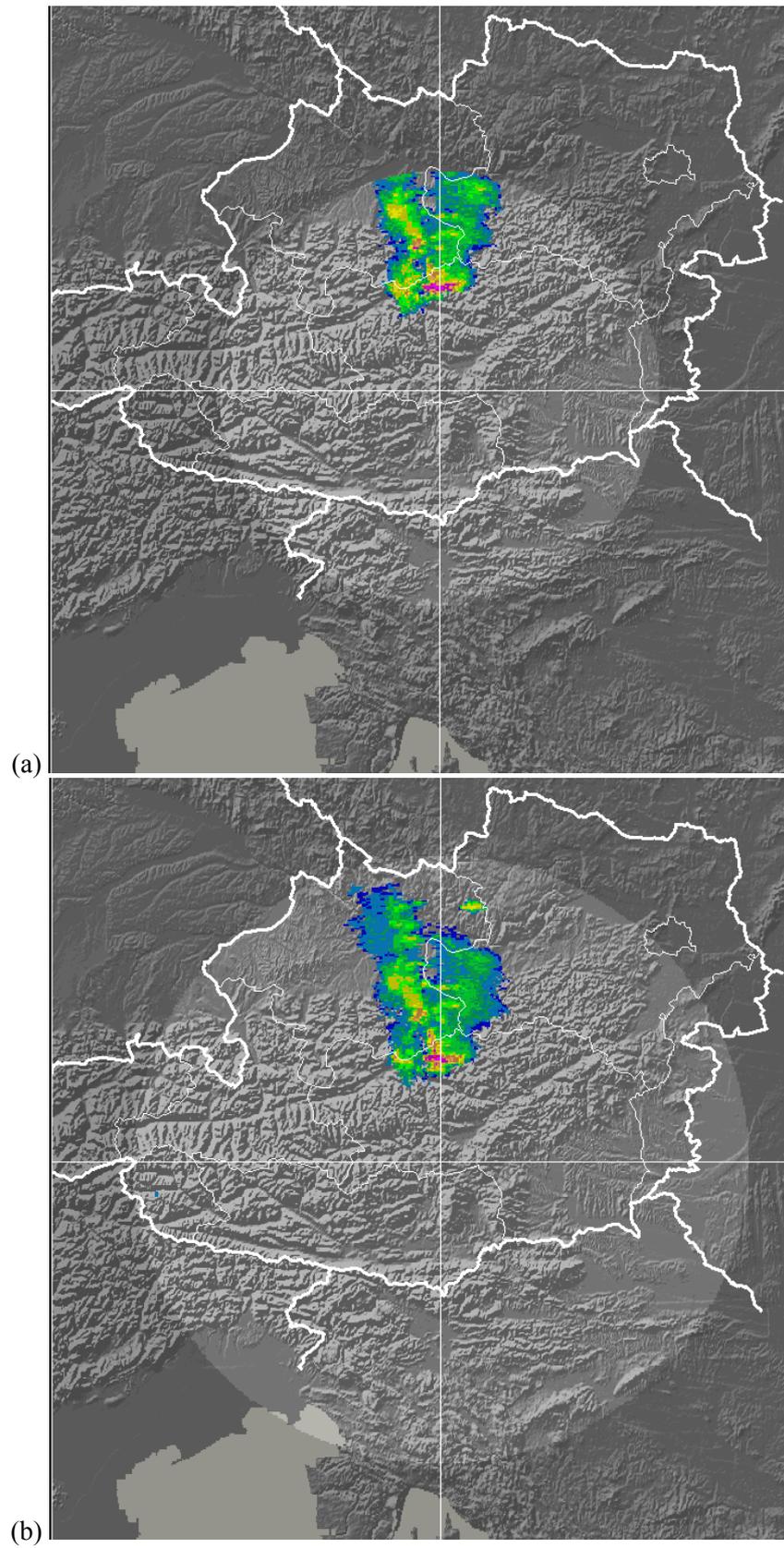


Figure 3.9. CAPPI-plots of one precipitation event at different altitudes, detected by the weather radar station on Mt. Zirbitzkogel on June 25, 2007, at 16:30. The constant altitude is (a) 3.5 km and (b) 4.5 km respectively (Source: Austro Control GmbH).

3.7 SOURCES OF POSSIBLE ERRORS

It has been mentioned that relating the measured reflectivity to an amount of precipitation hinges on several simplifying assumptions. This section emphasises on severe sources of errors weather radar data are likely to contain. Most of them are not correctable.

3.7.1 BEAM BLOCKAGE

At lower elevations the beam of weather radars operating in mountainous terrain is often blocked. The blockage can be total or partial. As a result, precipitation behind the obstacle cannot be detected or is underestimated (Figure 3.10).

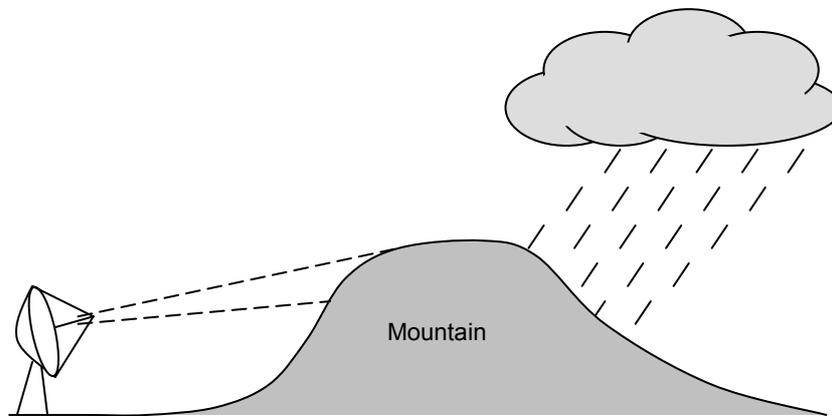


Figure 3.10. Total beam blockage. No information about the precipitation is gained.

Beam blockage is very common in Austria due to the Alpine mountain range. Figure 3.11 shows two radar images where this phenomenon is visible. It can be seen that the partial blockage leads to a small sector with a decreased predicted rain rate. The location of the sector does not change with time.

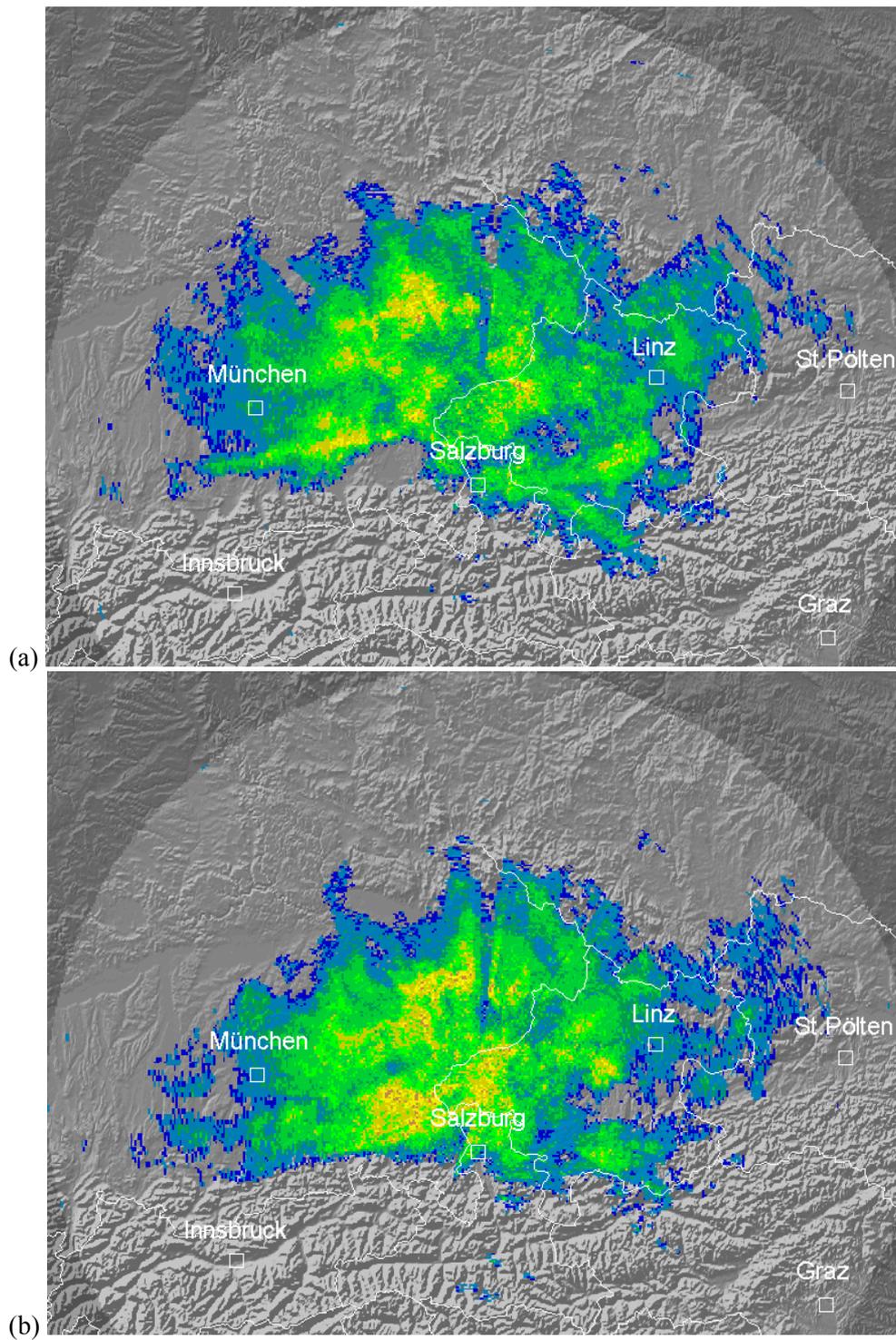


Figure 3.11. Identification of beam blockage in radar images. Two images originating from different recording times show the same sector of decreased rain rate. The recording time is (a) 09:25 and (b) 09:45 h on March 12, 2008 (Source: Austro Control GmbH).

3.7.2 BEAM ATTENUATION

The radar beam is attenuated by all obstacles and therefore also by precipitation particles. The attenuation occurs twice. It affects the transmitted wave as well as the reflected wave. The attenuation is especially significant when an intense convective cell is located close to the radar site as illustrated in Figure 3.12. In general attenuation leads to an underestimation of precipitation since it weakens the signal that the radar receives.

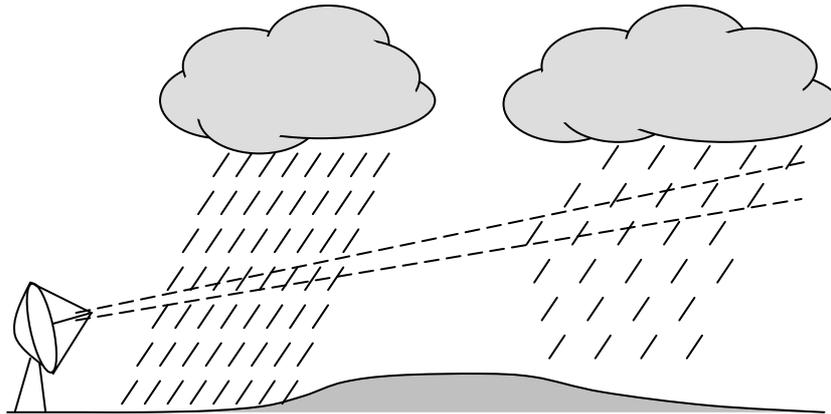


Figure 3.12. Beam attenuation by heavy rain close to the radar site.

Figure 3.13 shows three radar images where the phenomenon of beam attenuation due to an intense convective cell can be seen. Based on one single radar image, beam attenuation can hardly be distinguished from beam blockage. If the sector of decreased rain rate moves with time, it is a strong indication for beam attenuation.

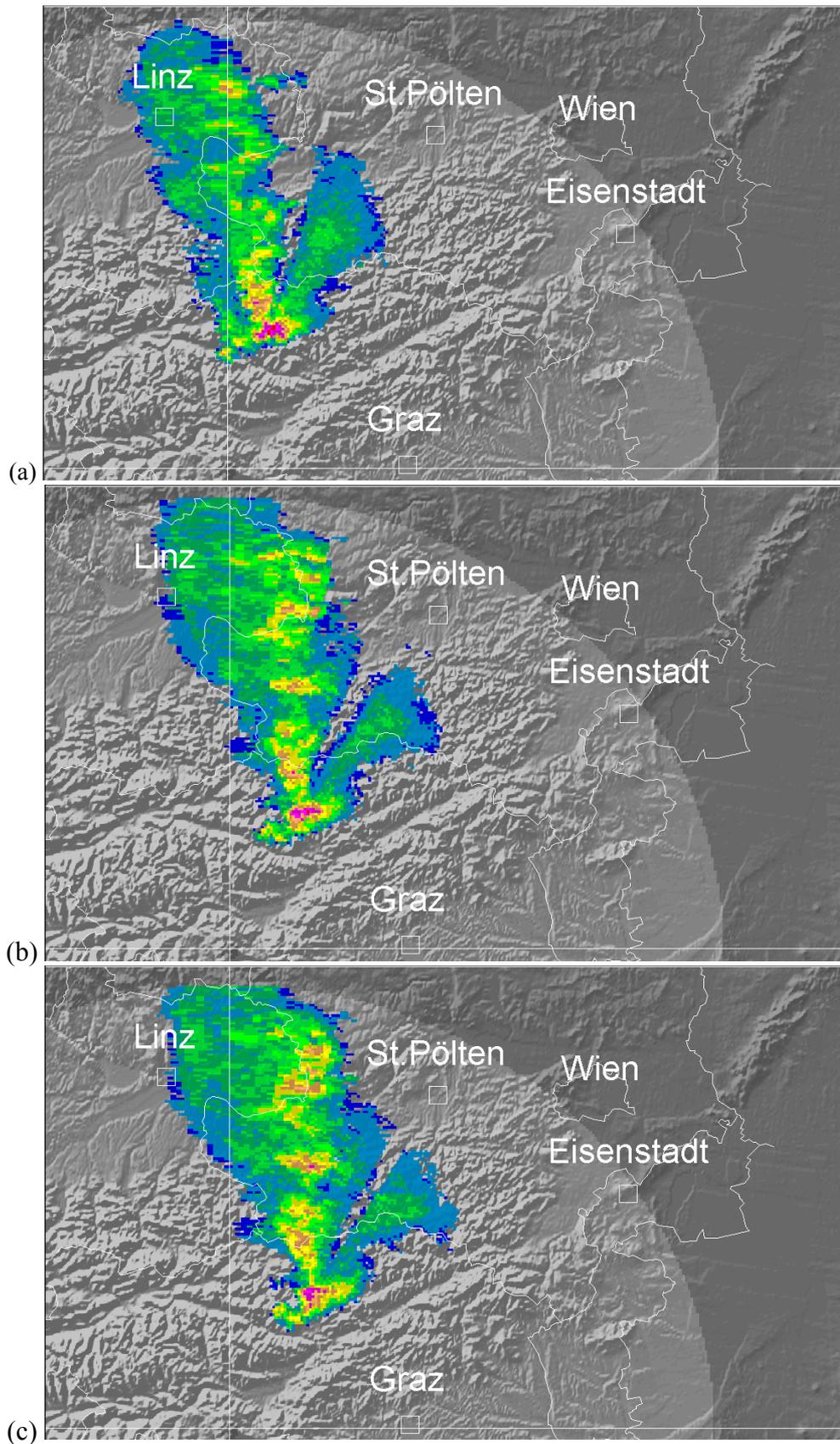


Figure 3.13. Identification of beam attenuation due to an intense shower cell in radar images. The sector without rainfall is not constant. The images originate from (a) 16:50, (b) 17:05, and (c) 17:10 h on June 25, 2007 (Source: Austro Control GmbH).

3.7.3 OVERSHOOTING BEAM

Overshooting beam likely occurs when rain clouds are close to the ground and when the radar is positioned on high altitude as illustrated in Figure 3.14. Weather radars have a minimum elevation angle of around 1° . Lowering the elevation angle below 1° would lead to capturing too much echoes from the ground (ground clutter, see below). Therefore, and due to the curvature of the earth, the beam likely overshoots clouds at long ranges. Overshooting beam leads to an underestimation of precipitation.

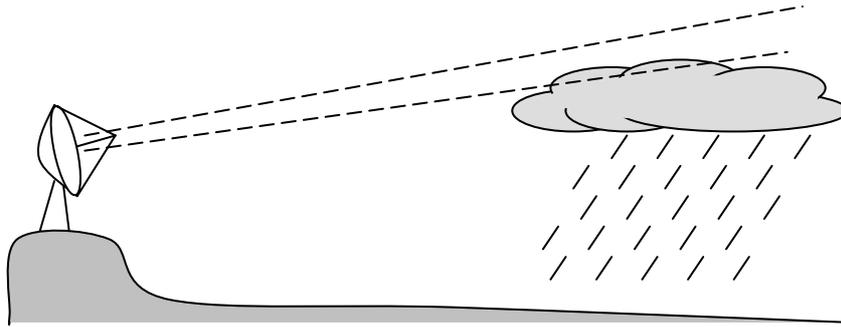


Figure 3.14. Overshooting beam. The precipitation below the radar beam is not detected.

3.7.4 EVAPORATION

Evaporation of precipitation can lead to significant overestimation of precipitation by weather radars. This is because the radar measures the precipitation aloft and due to evaporation this can be significantly more than actually reaches the ground. An extreme scenario is virga (see Figure 3.15). Virga is rain or ice “falling out of a cloud but evaporating before reaching the earth's surface as precipitation” (AMS, 2000).

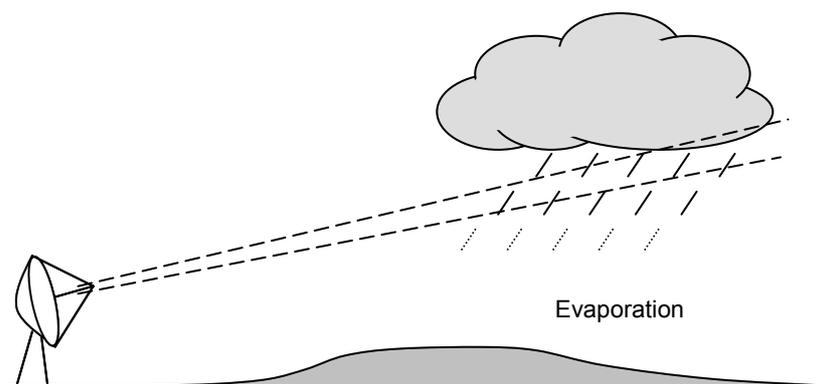


Figure 3.15. Virga - evaporation of precipitation above the ground.

3.7.5 GROUND CLUTTER

Ground clutter are radar echoes from objects on the ground like buildings or trees and from the ground itself. “Such echoes may be caused by the reflection of energy back to the radar in the main lobe or sidelobes of the antenna pattern and, in weather radar applications, interfere with the meteorological echoes at the same range” (AMS, 2000). In weather radar systems usually clutter filters are used to eliminate these unwanted echoes but often there is some residual clutter.

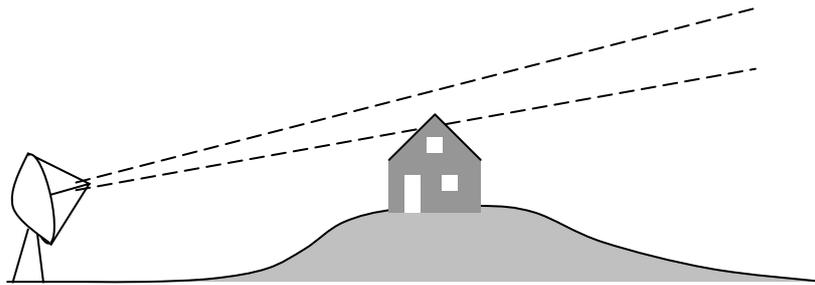


Figure 3.16. Ground clutter.

3.7.6 ANOMALOUS PROPAGATION

A reason for anomalous propagation of the radar beam is typically a strong temperature inversion near the ground. When there is a layer of cold air near the ground covered by a layer of warm air the radar beam can be bent to the ground as illustrated in Figure 3.17. As a consequence a strong signal is reflected to the radar. These echoes typically occur at large distances from the radar and they can move around due to changes of temperature and pressure in the atmosphere.

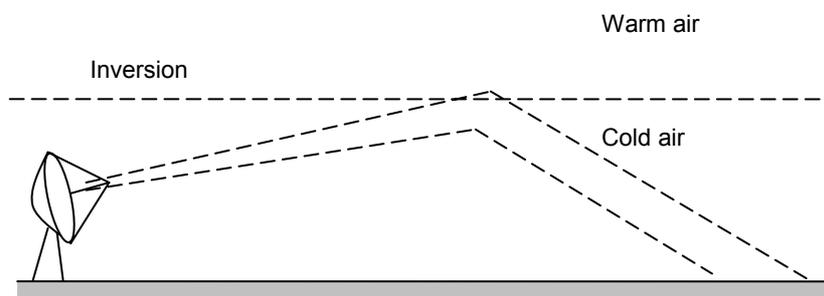


Figure 3.17. Anomalous propagation of the radar beam.

3.7.7 NON WEATHER TARGETS

Not all targets detected by weather radars are precipitation particles. For example birds and insects are often misinterpreted as rain. Especially flocks of birds can lead to strong false echoes.

False echoes can also occur when the weather radar antenna points directly at the sun. This occurs mainly at sunrise and sunset. Then the radar receives electromagnetic radiation, which can be seen in the radar image as a thin echo line from the radar site in the direction of the sun.

3.7.8 BRIGHT BAND

On an RHI-display anomalously high reflectivity is shown in the height where frozen precipitation particles melt. This layer is called bright band. Melting first affects the surface of the frozen precipitation particles and results in a water coating. Because of the 10 times higher index of refraction of water than of ice, the water coated particles appear to the radar as big raindrops. This effect can lead to an overestimation of precipitation.

3.8 THE AUSTRIAN WEATHER RADAR NETWORK

The Austrian weather radar network operated by Austro Control GmbH (the Austrian air navigation service provider) today consists of five C-band radars working at a frequency of 5.6 GHz. Four single polarization and the most recent dual-polarization radar. The radar data utilised in this work originate from the weather radar station on Mt. Zirbitzkogel which operates with single polarization. The radar has the following specifications:

- Altitude of the radar-station above Mean Sea Level (MSL): 2372 m
- Time interval between measurements: 5 minutes
- 3-dB-Beamwidth: 1°
- Minimum elevation angle: 0.8°
- Spatial resolution of the volume element: $1 \text{ km} \times 1 \text{ km} \times 1 \text{ km}$
- Resolution in measured reflectivity: 14 levels of rain-rate, converted from reflectivity Z by using the fixed relationship $Z = 200 R^{1.6}$
- Instrumented range: 220 km

Additional radar display

A particularity of the Austrian weather radar network is the *maximum value projection* which can be considered another radar display (see Section 3.6). Based on limitations of the data transmission link, initially the radar data were not transmitted in three dimensions. Instead, only the highest echo above the ground was transmitted. The information on what altitude the echo was detected is lost. Visually this radar display is similar to a CAPPI plot with measurements from different altitudes though. In Chapter 6 such data are utilised.

4 RAIN GAUGE AND WEATHER RADAR COMPARISON

When comparing rain gauge and weather radar measurements, their working principles and sources of possible errors should be kept in mind. Rain gauges are a direct method to determine rainfall. They measure the rainfall directly on the ground. Thus they are normally used to validate and calibrate indirect methods like weather radar measurements. When doing this, besides the sources of possible errors discussed before, the different sampling characteristics of rain gauge and weather radar must be taken into consideration. Another aspect is the data quantisation of operational weather radars. Quantisation is the process of approximating a continuous range of values by a defined number of levels. This approximation always introduces an error.

In this chapter the different sampling characteristics of rain gauge and weather radar are addressed as well as the effect of quantisation of the weather radar. In a case study at the end of the chapter it is described how these issues are dealt with in the course of this work.

4.1 SAMPLING CHARACTERISTICS

The rain gauge provides a point measurement. The “point” is the orifice area of the funnel, and the sampling interval of the tipping bucket rain gauges used here depends on the rain rate R . Although it is not common to refer to a rain gauge as a volume measurement, in the strict sense, the rain gauge scans the volume of a very thin cylinder. If we assume a fall velocity of the raindrops of 6 m/s and a tipping rate of 1 min⁻¹ then the cylinder the rain gauge scans is 360 m high. With crosswind the rain gauge will not capture the raindrops directly above the instrument but rain drops laterally displaced as indicated in Figure 4.1.

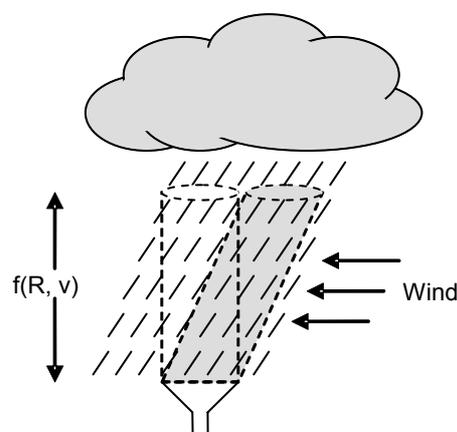


Figure 4.1. Idealized sampling volume of a tipping rain gauge. The height of the cylinder depends on the rain rate R and the falling velocity v of the precipitation particles.

The radar on the other hand measures the reflectivity of a volume determined by the main lobe of the antenna and the range gate length and from it estimates the rain rate (Figure 4.2). The precipitation particles that are detected at a considerable height are supposed to arrive (unchanged) at the ground later, depending on the fall speed of the particles.

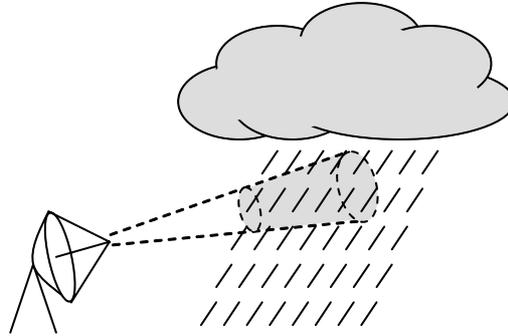


Figure 4.2. Idealized sampling volume of the weather radar. The sampling volume increases with the distance from the radar.

There is a large discrepancy in sampling volumes as Austin and Seed (2005) call it. Villarini and Krajewsky (2008) point out that the sampling discrepancy between rain gauge and weather radar can be nine orders of magnitude. As rain gauge and weather radar do not observe the same precipitation particles, there will be differences between rain gauge and radar measurements even if certain sources of possible errors can be ruled out.

4.2 QUANTISATION OF THE WEATHER RADAR DATA

An effect that is often not considered when comparing rain gauge and radar data is the influence that the quantisation intervals of the operational weather radars have. Quantisation is the process of approximating a continuous range of values by finite steps or levels. This approximation always introduces an error that is not correctable. Comparing quantised weather radar data with quasi continuous rain gauge data leads to deviations. This phenomenon is discussed below.

The Austrian weather radar network archives the data with 14 levels of inferred rain rate. Each of these quantisation intervals represents a range of rainfall depths reaching from the lower to the upper level border. These quantisation intervals increase with the rainfall intensity. For data processing all values lying anywhere within the quantisation interval are assigned the same magnitude. As a general rule the mean value is taken because the quantisation error can thus be minimised to the half width of the quantisation interval. For lower rain rates the quantisation error can be neglected. However, for higher rain rates the error can be considerable. Level 13 of the Austrian weather radar network for example represents rain rates from 89.9 mm/h to 153.8 mm/h (see Table 4.1). By assigning the mean value (121.85 mm/h), the maximum error due to quantisation is almost 32 mm/h.

Table 4.1. Quantisation intervals of the Austrian weather radar system.

Interval no.	From Rain rate [mm/h]	To Rain rate [mm/h]	Assigned Rain rate [mm/h]
1	0.0	0.2	0.00
2	0.2	0.3	0.25
3	0.3	0.6	0.45
4	0.6	0.9	0.75
5	0.9	1.7	1.30
6	1.7	2.7	2.20
7	2.7	5.0	3.85
8	5.0	8.6	6.80
9	8.6	15.0	11.80
10	15.0	27.3	21.15
11	27.3	50.0	38.65
12	50.0	89.9	69.95
13	89.9	153.8	121.85
14	153.8	∞	153.80

Generally the mean values of the intervals are used except for the intervals 1 and 14. For interval 1, zero is used for data processing. For interval 14, which represents rain rates greater than 153.8 mm/h, no mean or maximum value can be given, thus the minimum value is taken.

The deviations that can occur because of the quantisation intervals can be considerable. Figure 4.3 shows ten consecutive rain rates measured by a rain gauge and the respective quantisation intervals an ideal and error-free radar would yield. Table 4.2 gives the Root Mean Squared Error (RMSE) and bias that can occur when different values within the radar quantisation interval are taken for comparison.

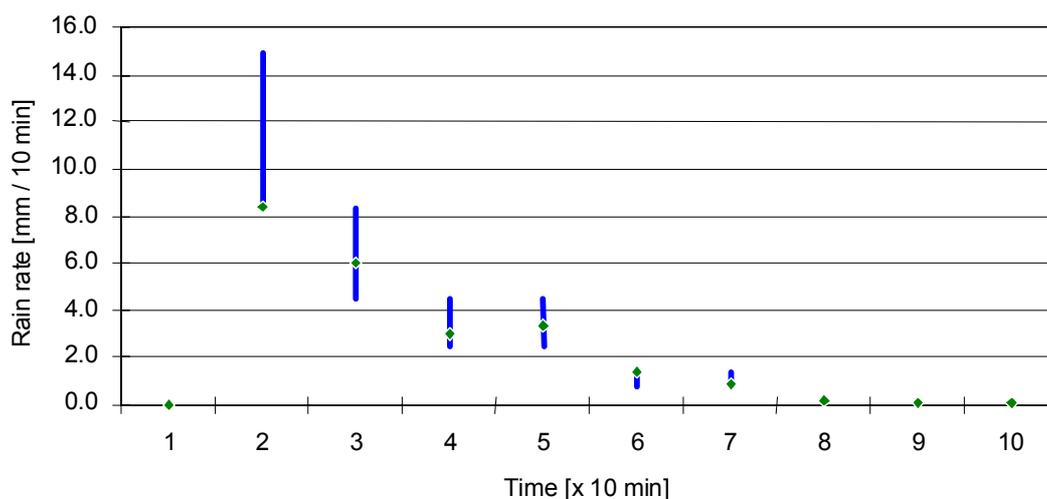


Figure 4.3. Quantisation of rain rates. Rainfall event as recorded by a rain gauge in Graz, Austria on July 24, 2004 (green diamonds) and the respective ideal and error-free quantisation intervals of the Austrian weather radar system.

Table 4.2. Deviations due to quantisation. Rain gauge values compared with minimum, mean and maximum values of the radar intervals of Figure 4.3.

	RMSE [mm/10 min]	BIAS [mm/10 min]
Minimum values	0.59	-0.36
Mean values	1.06	0.44
Maximum values	2.30	1.24

It becomes clear that by comparing rain gauge measurements with minimum, mean, and maximum values considerable deviations can occur. In Figure 4.3 it is noticeable that especially the bigger intervals associated with higher rain rates contribute to these deviations. The highest precipitation level that occurred in the example was Level 12 representing rain rates from 50.0 to 89.9 mm/h (8.3 to 15.0 mm/10 min). Here the error due to quantisation are significant.

4.3 CASE STUDY

In order to compare rain gauge and radar measurements the first idea was to estimate the areal precipitation by interpolating between the rain gauge measurements (as described in Chapter 2.5). Thus, not the point measurements of single rain gauges should be compared with the weather radar measurements, but the areal rainfall determined from the measurements of several gauges.

The most striking fact is that the density of the rain gauge network is low compared to the areal resolution of the weather radar (see Chapter 2.4.) The rain gauge density is approximately on the order of one station per 100 km², whereas the weather radar provides one measurement per 1 km². Thus, the weather radar captures much more spatial information than the rain gauge network. The problem with calculating the areal precipitation based on rain gauge measurements is that the density of the rain gauge network is not high enough to make substantial contributions. Calculating the areal precipitation definitely makes sense if there are several rain gauges per radar pixel as in Villarini and Krajewsky (2008) with up to eight gauges per radar pixel. If the next rain gauge is several kilometres away and is used together with the rain gauge within the radar pixel to estimate the rainfall on a basis of the pixel size of 1 km², the danger is that the precipitation situation at the other rain gauge is completely different (no precipitation or an other shower cell) and the areal estimate is even worse. Thus, in this work only the rain gauge station within the radar pixel is used as a reference for the radar.

It was analysed if the different sampling characteristics of rain gauge and weather radar result in a time lag of the two time series. Rain gauge and weather radar data over a two-year period from 2002 to the end of 2003 were investigated. Precipitation data were available from the rain gauge station Laßnitzhöhe close to the city of Graz. The radar data used originate from the weather radar station on Mt. Zirbitzkogel. The distance between the radar and the rain gauge station is about 75 km. The lowest elevation level above the rain gauge site visible to the radar is 3 km (MSL).

First the data were transferred into the same temporal resolution. Originally the integration time was 10 minutes for the rain gauge, and 5 minutes for the weather radar. Thus, in order to compare the two time series, also the weather radar data were integrated to the same 10 minute intervals. The radar data which originate from the lowest elevation level above the rain gauge (CAPPI 3 km) were shifted with respect to the rain gauge data due to the delayed detection of precipitation by the rain gauge. The best agreement with the rain gauge time series showed the 10 minutes shifted radar data. This time series shows the best figures in terms of correlation coefficient and RMSE, see Table 4.3. The same comparison was done with weather radar data from the layer above CAPPI 4 (Table 4.4). Again the 10 minutes shifted weather radar data showed the best agreement. It can be assumed that there is a time shift between the radar data of the different CAPPI levels, but the temporal resolution of the weather radar does not allow a more precise temporal analysis. Noticeable is the lower correlation of the weather radar data of CAPPI level 4 with the rain gauge data than those of level 3. Generally speaking – and not surprisingly – the lower the elevation of radar data the better is the agreement with rain gauge data.

Table 4.3. Comparison of rain gauge and shifted weather radar time series CAPPI 3 km.

Time Shift [min]	Correlation coeff.	RMSE [mm/10 min]
0	0.4451	0.1591
5	0.4753	0.1574
10	0.5310	0.1539
15	0.4313	0.1598

Table 4.4. Comparison of rain gauge and shifted weather radar time series CAPPI 4 km.

Time Shift [min]	Correlation coeff.	RMSE [mm/10 min]
0	0.3631	0.1669
5	0.3972	0.1661
10	0.4516	0.1640
15	0.3817	0.1660

The different place of observation of rain gauge and weather radar results in a time lag of 10 minutes between both time series. In this examination usually the mean values of the radar intervals were used for the comparison as shown in Table 4.1. Below the influence of the quantisation interval is investigated.

The approach is that not automatically the mean value of the quantisation interval is taken, but that the measurements of surrounding radar measurements of the same elevation (CAPPI level) in a $5 \text{ km} \times 5 \text{ km}$ radar grid influence the choice (Figure 4.4).

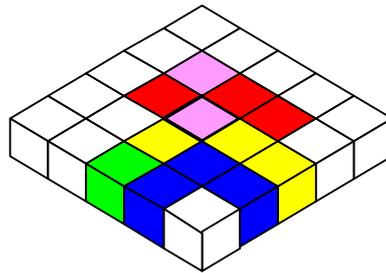


Figure 4.4. Consideration of surrounding radar volume elements. These measurements co-determine the value of the quantisation interval of the cell in the middle (cell above the rain gauge).

The assumption behind this idea is that it is more probable that the reflectivity value of one volume element lies near the lower interval border if the adjacent elements show a lower reflectivity. On the other hand if the radar measures higher values for the surrounding volume elements than for the volume element in the centre, the actual reflectivity is supposed to lie near the upper level border.

The chosen method is to interpolate between the mean values of the $5 \text{ km} \times 5 \text{ km}$ radar grid with cubic interpolation and integrate over the pixel in the middle. The adapted value thus determined is assigned to the volume element. With this technique not always the mean value of the quantisation interval is taken but the value will vary. It will be biased towards the upper interval limit if the adjacent values are higher, and it will be biased towards the lower interval limit in the other case. It is assured that the values in every case lie within the quantisation interval. With this technique modified weather radar data are obtained that do not stick to the fixed number of quantisation values.

The values in Table 4.5 show that a choice other than the mean value of the quantisation interval yields slightly better results concerning correlation coefficient and RMSE. The results for the higher elevation level are similar (Table 4.6).

Table 4.5. Comparison of rain gauge and modified weather radar time series CAPPI 3 km (The values in brackets show the pristine data of Table 4.3)

Time Shift [min]	Correlation coeff.	RMSE [mm/10 min]
0	0.4641 (0.4451)	0.1583 (0.1591)
5	0.4978 (0.4753)	0.1564 (0.1574)
10	0.5506 (0.5310)	0.1533 (0.1539)
15	0.4610 (0.4313)	0.1584 (0.1598)

Table 4.6. Comparison of rain gauge and time shifted weather radar time series CAPPI 4 km (The values in brackets show the pristine data of Table 4.4)

Time Shift [min]	Correlation coeff.	RMSE [mm/10 min]
0	0.3745 (0.3631)	0.1666 (0.1669)
5	0.4036 (0.3972)	0.1660 (0.1661)
10	0.4566 (0.4516)	0.1641 (0.1640)
15	0.3892 (0.3817)	0.1658 (0.1660)

In summary it can be said that the different sampling volumes of rain gauge and weather radar result in a significant time shift. Also the quantisation interval and the choice of the value for processing have an influence. For single measurements, this influence can be high, as one example showed, but it averages out in longer term comparisons. Given that the interpolation procedure is very time consuming and does not in every case bring an improvement concerning RMSE, it is not applied in the further course of this work. The consideration of the time shift between radar and rain gauge data, however, is practicable and therefore considered in further analyses and comparisons.

5 DATA-DRIVEN MODELLING

When we look at the data weather radars and rain gauge networks produce we will notice that the volume is particularly large. In the archives reams of gigabytes of environmental data are stored. Today's technology allows us to save all these data. But it is not just data that is stored in the archives – information is hidden. Data-driven approaches aim to extract this information and find interrelationships.

Artificial neural networks and decision trees are data-driven approaches applied in this work. A basic knowledge of these techniques is important for the understanding of development, application, and outcome of these paradigms. For this reason here a general survey is given. Please note that the description below is only as detailed as necessary to understand the subsequent chapters. It is not a comprehensive disquisition on data-driven approaches. For a detailed explanation of artificial neural networks the reader is referred to Peretto (1992) and to Rojas (1993) for a textbook in German. Witten and Frank (2005) comprehensively discuss decision trees and various other models.

5.1 ARTIFICIAL NEURAL NETWORKS

Artificial neural networks got their name from their analogy with biological neural networks. Biological neurons are cells that receive, process, and transmit information using biochemical reactions. The human brain consists of about 10 billion neurons which are highly interconnected. One neuron is connected via dendrites to a thousand neighbouring neurons.

Artificial neural networks (ANNs) are far from reaching such dimensions. Moreover, artificial neurons simplify the processes that are going on in biological ones. In machine learning ANNs try to simulate some properties of biological neural networks. The neurons in ANNs are simple processing elements operating in parallel. Because of the numerous connections between these elements, ANNs can exhibit complex global behaviour.

ANNs can perform pattern recognition, identification and classification tasks, system control, and other complex functions. Often they are trained in function approximation, so that a particular input leads to a specific output. This form of training also allows making predictions. The relationship found in the training process between input and output variables can be used to predict the output of unseen examples by knowing only the input data. Here such a configuration is used with weather radar data. Below the neuron model, network architectures, and learning rules are described. The software tool that is used in the work is the MATLAB *Neural Network Toolbox*. The description below and the terminology are based on Demuth and Beale (1998).

5.1.1 NEURON MODEL

An artificial neural network is usually presented as an interconnected group of nodes as shown in Figure 5.1. The first step to understand the behaviour of an ANN is to understand how the nodes are defined.

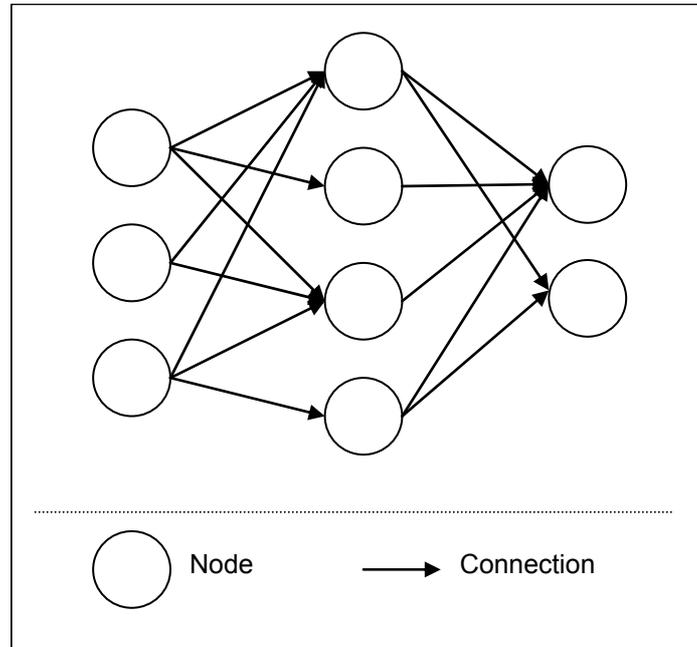


Figure 5.1. Illustration of an ANN with nodes and connections between them.

Nodes in an ANN are simple processing elements that are connected to others. Some of the nodes in Figure 5.1 are target of only one arrow. The neuron model is illustrated on the basis of this special case.

Figure 5.2 shows what is inside the circle that is frequently used to visualise a neuron. First, the input p is multiplied by the scalar w to form the product wp . Then another scalar, the bias b is added. The transfer function f takes the argument n which is the sum of the weighted input and the bias. The transfer function which often is a nonlinear function maps n to the output a .

In general, applications using ANNs require nonlinear mappings. That's why nonlinear transfer functions are needed. In this work linear and nonlinear transfer functions are used. The applied nonlinear function is the sigmoid transfer function shown in Figure 5.3. This transfer function is commonly used in ANNs. It takes the input, which can have any value between plus and minus infinity, and maps the output into the range 0 to 1. Often the concept of a sigmoid functions in the hidden layer (see next section) and linear transfer functions in the output layer yields the best results (e.g. Loukas, 2000). This combination is also applied here.

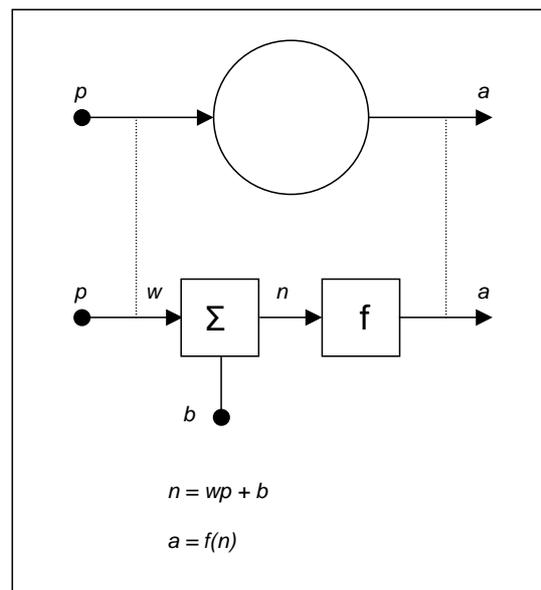


Figure 5.2. Mode of operation of a single neuron.

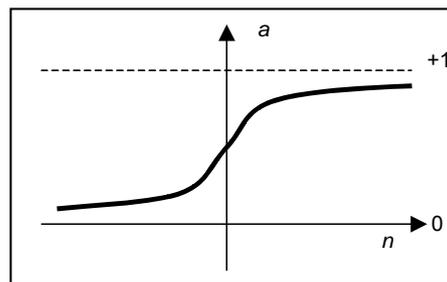


Figure 5.3. Schematic curve of a sigmoid transfer function.

Above the mode of operation of a single neuron with a scalar as input was shown. Generally, the input is a vector rather than a scalar. In this case the operation is not much different, except that the transfer function input n changes to:

$$n = w_1 p_1 + w_2 p_2 + \dots + w_R p_R + b \quad (5.1)$$

where R is the number of elements in the input vector.

5.1.2 NETWORK ARCHITECTURES

In ANNs the neurons usually are arranged in layers. An ANN can contain several of these layers. In Figure 5.4 each element in the input vector p is connected to each neuron. The weights of this layer with S neurons and R inputs can be described with a weight matrix \mathbf{W} with S rows and R columns.

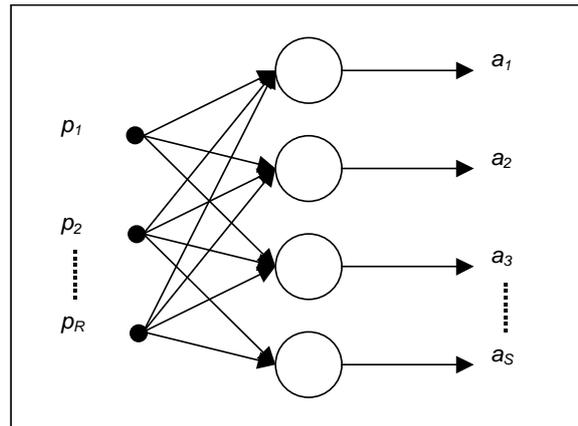


Figure 5.4. One layer of an ANN. The layer consists of S neurons and an R -dimensional vector forms the input.

A network with several layers has a weight matrix and a bias vector for each layer. The ANN of Figure 5.1 is a multilayer network with three layers. They are referred to as input, hidden, and output layer. The input layer accepts the inputs, the output layer produces the outputs of the network, and any layers in between are called hidden layers.

In networks as described above, the information always goes in one direction from the inputs to the outputs. There are no feedback cycles. Networks of this type are called feed-forward neural networks. Feed-forward neural networks are the standard type of ANNs. To better illustrate the information flow, they are often drawn with one layer upon the other as in Figure 5.5.

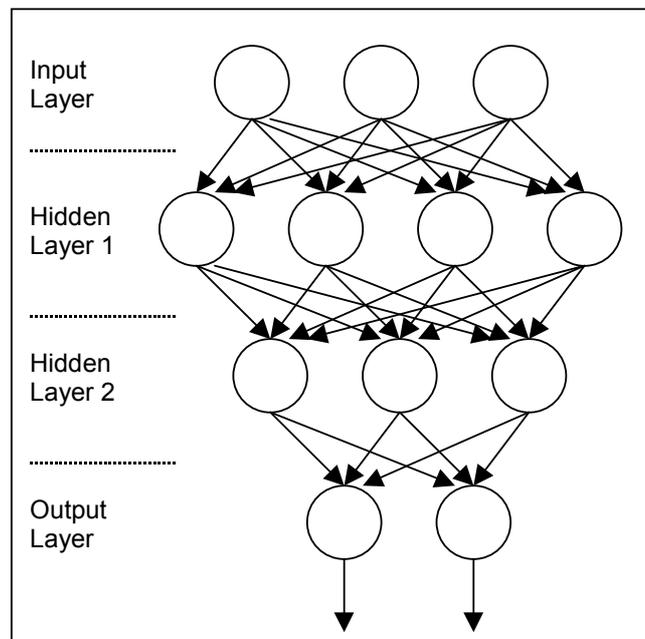


Figure 5.5. Fully connected feed-forward neural network with two hidden layers.

Feed-forward neural networks are presumably the most widely used type of ANNs. Also many practical neural network applications in the field of hydrology use this type of network.

But setting up such a network can be very time consuming. Another type of ANNs, the Radial Basis Function (RBF) neural network is an alternative architecture that has advantages concerning setup time. However, it involves a different architecture. For the task of rainfall estimation (Chapter 7) RBF networks are alternatively applied. Below an overview is given.

Radial basis function neural networks apply a different transfer function. As the name implies it is the radial basis function. It is defined as:

$$RBF(n) = e^{-n^2} \quad (5.2)$$

The *RBF* has its maximum of 1 when the input is 0 (see Figure 5.6).

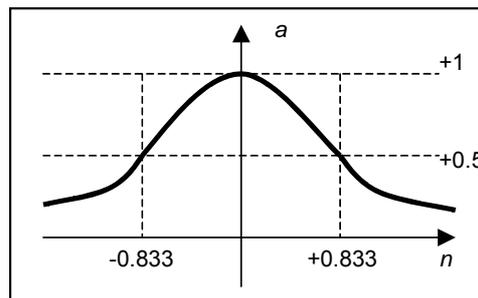


Figure 5.6. Schematic curve of the radial basis transfer function.

But not only the transfer function is different compared to the functions normally used in feed-forward neural networks. The neuron also has another entry stage. In the former type the input p is multiplied by the scalar w and a bias b is added. In radial basis neurons, the *Euclidean distance* between the input vector p and the weight vector w is calculated and multiplied by the bias b . The outcome n is the argument of the *RBF* (see Figure 5.7).

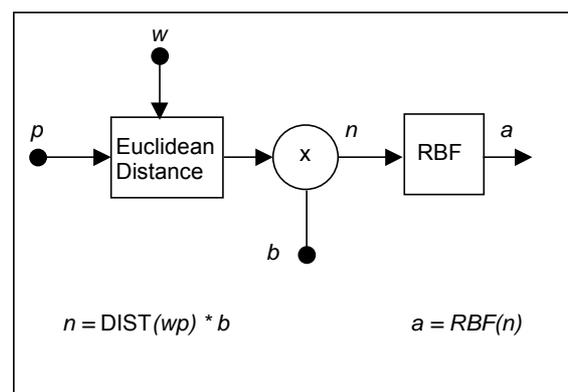


Figure 5.7. Mode of operation of a radial basis neuron.

The output of the neuron increases with decreasing distance between the input vector p and the weight vector w . The output is 1 when p and w are identical. Thus, Demuth and Beale (1998) equate the radial basis neuron with a detector. The sensitivity of the neuron can be

adjusted by the bias. A bias $b < 1$ virtually broadens the function, a bias $b > 1$ on the other hand narrows it (makes it react with a smaller output to the same input).

Now the mode of operation of a single radial basis function is known. The architecture of a radial basis network consists of two layers. The output layer is a simple linear layer. It is the same type as introduced above for feed-forward networks, but with a linear in place of a sigmoid transfer function (see Figure 5.8).

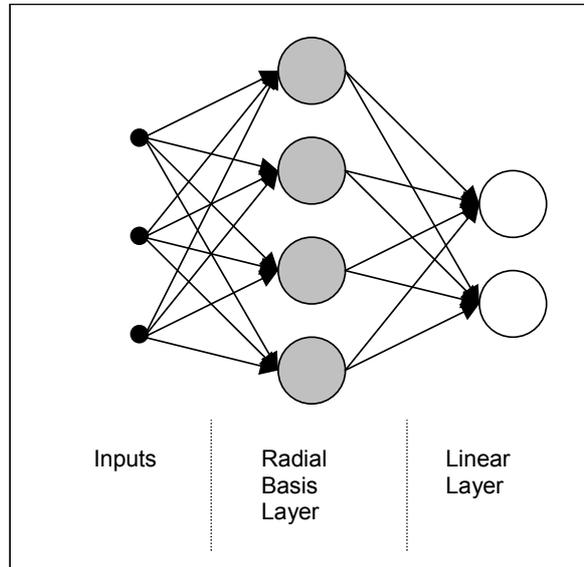


Figure 5.8. Architecture of an RBF neural network.

As the architecture of feed-forward and radial basis function neural networks are established, the question is: How the weights and biases are determined? This is the task of the training process which is described in the next section.

5.1.3 LEARNING RULES

The *learning rule* as referred to in Demuth and Beale (1998) is a procedure for modifying the weights and biases of a network. This procedure also is known as *learning paradigm* and under the more general term *training algorithm*. Often it is distinguished between two categories of learning rules: supervised and unsupervised learning. Supervised learning is applied in the applications of Chapter 6 and Chapter 7 and thus described below.

In supervised learning, together with the input vector p another vector is given: the target vector t . The target represents the desired output of the neural network when supplied with the input p . Note that the dimension of the vectors p and t is not necessarily the same. In hydrological applications one may have precipitation data from several points within a catchment as input vector and the target is the runoff at the watershed outlet (thus a scalar). Initially the weights and biases are chosen randomly. The network supplied with the input data will thus produce a random output a . Then the learning rule is applied to adjust weights and biases in a way to move the output a closer to the target t . Figure 5.9 demonstrates the supervised learning procedure.

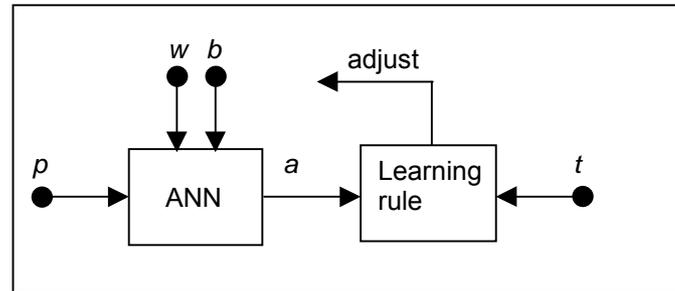


Figure 5.9. Schematic illustration of supervised learning.

The network is provided with Q examples of input (p) and target (t) vectors. Let us assume, the input p_k leads to the output a_k while the desired target is t_k . The error is usually calculated as the mean squared error and should be minimised in the learning process.

$$\text{MSE} = \frac{1}{Q} \sum_{k=1}^Q (a_k - t_k)^2 \quad (5.3)$$

The MSE will be different for different weights w and biases b , and it is up to the learning rule to determine w and b in a way to minimise the error. Apparently it is not a trivial task to optimise w and b , especially in a multilayer network with dozens of weights and biases. Many learning rules or training algorithms exist for this purpose. The standard algorithm used in the applications of this work is the Levenberg-Marquardt algorithm. It is a fast training algorithm. Levenberg-Marquardt for neural network training is described in Hagan and Menhaj (1994).

Determination of the proper weights and biases for the training dataset is only one part of the whole training process. In order to make predictions, it is essential to test the network with unseen data (data different to the data the network was trained on). It is very likely that the performance deteriorates. This is understandable since the network parameters w and b are optimised on the basis of the training dataset. The goal is a network that is not overtrained (which is also known as overfitting), but able to generalise to new data. That's why an extensive test process is necessary.

5.2 MODEL TREES

Tree-like structures are frequently used in machine learning and data mining. The simple tree diagram with nodes, branches, and leaves forms the basis of a decision tree as a predictive model. Two main types of decision trees exist: classification and regression trees. Trees that predict a symbolic or categorical attribute are called classification trees whereas regression trees predict a numeric value. An example of a classification tree for forms of precipitation can be seen in Figure 5.10.

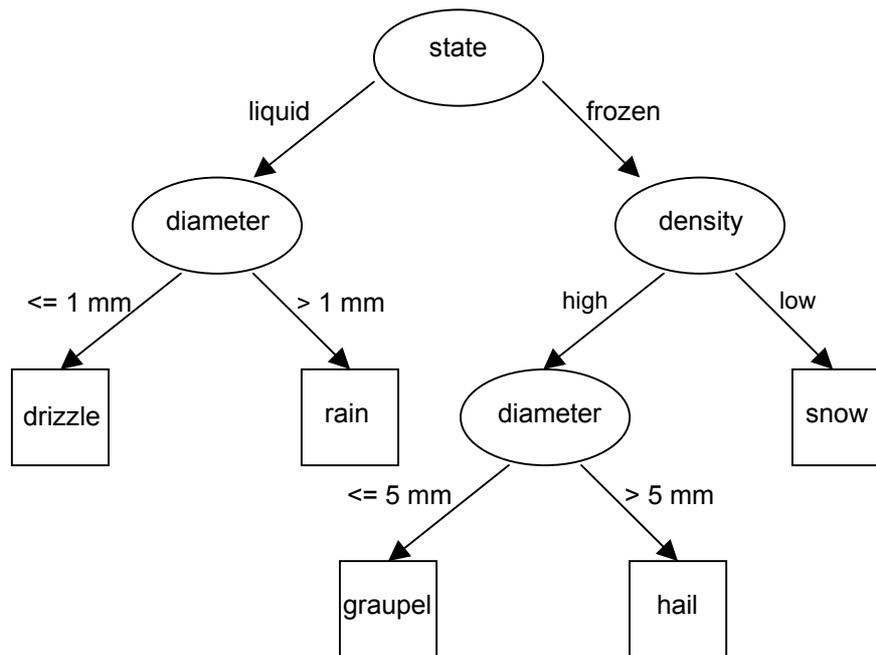


Figure 5.10. Example of a classification tree. The tree classifies forms of precipitation.

The regression tree on the other hand gives an average numeric prediction for each leaf of the tree. There is a special type of regression trees: If all leaves contain linear regression models, rather than average values, then the tree is called model tree. Model trees can be seen as piece-wise linear approximations. According to Bhattacharya and Solomatine (2005a) the major advantages of model trees are that they are “much smaller than regression trees, the decision strength is clear, and the regression functions do not normally involve many variables” (p. 386). Furthermore, it should be mentioned that all kinds of decision trees are simple to understand and interpret, even by someone not familiar with data-driven approaches. Stravs and Brilly (2007) mention that regression and model trees “can give a structural insight into the hydrological process that is being modelled” (p. 467) and accentuate this as a difference to neural networks.

With Stravs and Brilly (2007) a recent paper uses model trees in the field of hydrology, but this data-driven approach is not very common in hydrology. Compared to the numerous neural network approaches in this field of science, model trees live a shadowy existence. No more than a few years ago, model trees were presented as alternatives to neural network

approaches in rainfall-runoff modelling (Solomatine and Dulal, 2003) and flood forecasting (Solomatine and Xue, 2004).

The tree model that is used in this work is the M5 model tree. The algorithm was invented by Quinlan (1992). Witten and Frank (2000) implemented it in their WEKA software package (Waikato Environment for Knowledge Analysis).

The idea behind building a model tree is straightforward. Wang and Witten (1997) name three stages of developing the tree. In the first stage an induction algorithm is used to build the tree. Here a splitting criterion is used that minimises the intra-subset variation. The second stage is pruning the tree back from each leaf to help to generalise the tree. And the final stage is a smoothing process.

5.2.1 BUILDING THE TREE

The following consideration is the beginning of building the tree. Split the data and build a linear regression model for each of the subspaces. Suppose that T is a portion (or subset) of training data and consequently the question is: Should T be associated with a leaf, or should T be split again? The splitting criterion in the *Weka* software package is the Standard Deviation Reduction (*SDR*; Witten and Frank, 2000). *SDR* is used to determine which attribute value of T is the best to split the portion T . It is defined as:

$$\text{SDR} = sd(T) - \sum_i \frac{N_i}{N} \times sd(T_i) \quad (5.4)$$

where $sd(T)$ is the standard deviation of T . T_i are the sets that result from splitting the node according to the chosen attribute. N and N_i are the numbers of instances in the trees T and T_i respectively.

For splitting, the attribute that maximises the expected reduction is used. According to Witten and Frank (2005) the splitting process terminates when the class values of the instances that reach a node, vary only slightly or when only a few instances remain at one node. Instances are the examples to be classified.

Now the attribute values have been chosen to make the routing decisions at each note. The instances are routed following the tree down to a leaf. Each leaf contains a linear regression model based on attribute values. The regression model has the form:

$$w_0 + w_1 a_1 + w_2 a_2 + \dots + w_k a_k \quad (5.5)$$

where a_1, a_2, \dots, a_k are attribute values and w_1, w_2, \dots, w_k are the weights which are calculated using standard regression (cf. Witten and Frank, 2005).

Finally the tree is built. The nodes for the routing have been chosen and at each leaf a linear regression model predicts a value for the instances that reach the leaf. However, the tree is built using training data only. The pruning procedure introduced in the next chapter accounts for this disadvantage.

5.2.2 PRUNING THE TREE

The unpruned tree has been expressly built for the training dataset. It can be said that the accuracy of the tree concerning this dataset increases with each leaf as the tree grows. This is not necessarily true for the unseen cases of the test dataset. It is again the overfitting problem known from neural networks. In the *Weka* software the overfitting problem for model trees is addressed as follows. According to Wang and Witten (1997, sec. 2.2) “the absolute difference between the predicted value and the actual class value is averaged for each of the training examples that reach that node”. For the unseen test data the error will be higher. That’s why it is multiplied by a factor for compensation. Because of the compensation factor, the linear model can often be further simplified. Details to the pruning process are given in Witten and Frank (2005). The result is that the tree gets smaller because some subtrees may be replaced by single leaves.

5.2.3 SMOOTHING THE TREE

The pruned tree consists of linear models at each leaf. There will be unavoidable discontinuities between adjacent linear equations of the tree. Smoothing is the process updates the adjacent linear equations. The update takes place in a way that the predicted outputs for neighbouring input vectors corresponding to different equations are aligned (cf. Bhattacharya and Solomatine, 2005b). For more information see Witten and Frank (2005).

6 MODELLING THE RAINFALL-RUNOFF RELATIONSHIP

Rainfall-runoff relationships play an important role in hydrology. Their purpose is the conversion of rainfall on the catchment area into runoff at the watershed outlet. The objective is to predict the runoff and – in consequence – flood waters.

Many processes are involved in the conversion of rainfall into runoff, such as interception, infiltration, and surface retention. In principle these processes are computable, but often sufficient information about the catchment is not at hand. The more data are available, the more detailed a rainfall-runoff model can be. Only in hydrological investigation areas with very dense measuring networks and spatial information about the orography and geology, the above listed processes can be accounted for, but even here considerable simplifications have to be made.

In this chapter data-driven approaches such as neural networks and model trees are used to model the rainfall-runoff relationship of a small catchment in Austria. Below it is described how weather radar and rain gauge data are utilised in these models. Furthermore, performance measures to evaluate the predictions are given. Finally the rainfall-runoff models are specified and the results are presented and discussed.

6.1 STUDY AREA

The study area is the Sulm catchment in the south-west of Styria, Austria. The whole basin includes an area of 1105.7 square kilometres. Elevations range from 263 m (MSL) at the watershed outlet (Leibnitz) to 2125 m (MSL) on Koralpe mountain. The average watershed slope is 11.9 %.

The scope of this analysis is the sub-catchment Wernersdorf. This small catchment (less than 35 km² in area) is of particular interest. Since there are no more flow meters upstream, the possibilities for flood warnings for this area are limited. The flow-meter data show that high peaks at the site Wernersdorf are usually followed by high peaks downstream. Therefore the discharge measurements at Wernersdorf can be a helpful indicator for severe situations that subsequently may lead to hazards downstream. Figure 6.1 presents a map of the Wernersdorf catchment, showing the radar grid and the location of rain gauge and flow meter.

In summer this region is often affected by rain showers. The catchment response of this part of Austria can be considered as flashy. The floods and the annual maximum daily precipitation traditionally occur in late summer. Short convective storms are the dominant events producing floods (cf. Merz *et al.*, 2001). Sometimes the spatial extension of these showers is so small that their detection is only possible by weather radar, while none of the rain gauges in the area reports any precipitation.

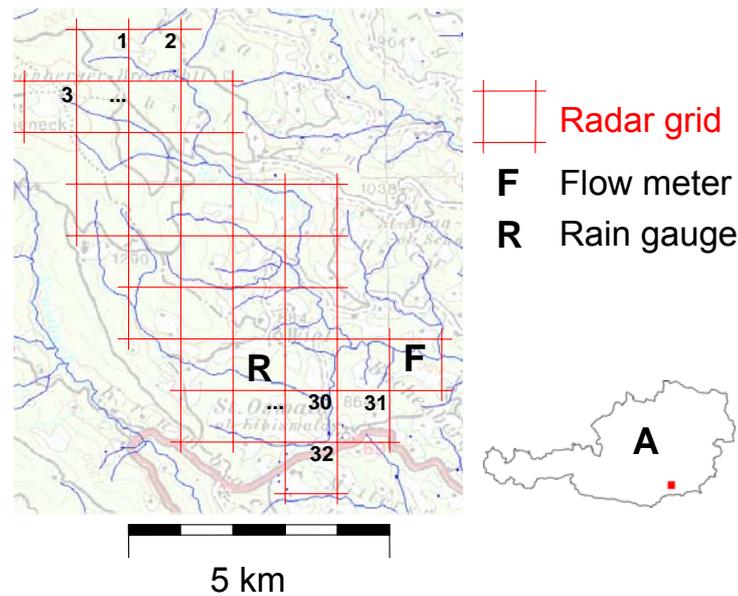


Figure 6.1. Map of the Wernersdorf catchment. Numbering of the radar pixels is from 1 to 32 top down, line by line.

6.2 AVAILABLE DATA

6.2.1 RAIN GAUGE AND FLOW METER DATA

Because of the specific geographic and climatic situation, the rain gauge and flow meter density in the Sulm catchment is quite high compared to other parts of Austria. The runoff [m^3/s] is known for all tributaries in the Sulm basin at 13 different sites. The time interval between the outflow-measurements is 15 minutes. Precipitation data are available from a network of rain gauges. The rain gauges are working on the tipping bucket principle with a resolution of 0.1 mm. The temporal resolution is 15 minutes. Data from 10 rain gauges are available.

One rain gauge station is located in the Wernersdorf sub-catchment. Thus for the development of the rainfall-runoff models, measurements from one rain gauge and a flow meter as well as radar data are available. The datasets extend over a 1-year period from January to December 2000.

6.2.2 RADAR DATA

Radar data from the Doppler weather radar station on Mt. Zirbitzkogel are used to improve the spatial coverage. The used radar is a high-resolution C-band weather radar. Radar measurements in *maximum value projection* are utilised (see Section 3.8). The distance between radar station and catchment is between 42 km (Koralpe mountain) and 80 km (Leibnitz, watershed outlet). The temporal resolution of all datasets was assimilated to the temporal resolution of the runoff data which is 15 minutes.

6.3 PERFORMANCE MEASURES

Visual comparisons of predicted and observed runoff curves can give a first impression of the performance of a rainfall-runoff model. But in order to quantitatively compare several models, performance measures are indispensable.

The correlation coefficient can be a first evidence for the quality of a prediction. It measures the statistical correlation between two time series. In our case the time series are the predicted Q_p and the observed runoff Q_o . The correlation coefficient is defined as:

$$\text{CORR} = \frac{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)(Q_{p,i} - \bar{Q}_p)}{\sqrt{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2} \cdot \sqrt{\sum_{i=1}^n (Q_{p,i} - \bar{Q}_p)^2}} \quad (6.1)$$

where \bar{Q}_o is the mean of the observed runoff measurements and \bar{Q}_p is the mean of the predicted runoff values.

The correlation coefficient, often denoted by r , ranges from -1 to $+1$. The value $+1$ stands for perfect correlation between predicted and observed runoff time series. The value -1 marks perfectly negative correlation (inverse correlation). The closer r is either to $+1$ or -1 respectively, the more closely the time series are related. A value of 0 means no correlation at all. The correlation coefficient is a scale-independent measure in that the error does not change if one time series is multiplied by a constant factor or added to a bias. In other words if the predicted runoff is throughout twice as high as the observed, r will still be 1 .

In data mining the correlation coefficient is an important measure, in hydrological applications, however, a scale-independent measure as the correlation coefficient has its disadvantages as the example above shows. Therefore, in hydrological modelling another coefficient – namely the efficiency coefficient – is used. The efficiency coefficient (cf. Nash and Sutcliffe, 1970) is defined as:

$$E = 1 - \frac{\sum_{i=1}^n (Q_{o,i} - Q_{p,i})^2}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2} \quad (6.2)$$

This coefficient can range from $-\infty$ to $+1$. An efficiency of 1 stands for a perfect match between predicted and observed data. An efficiency of 0 indicates a prediction as accurate as the mean of the observed runoff. An efficiency <0 is not desirable as this suggests a worse prediction than the mean value.

Fundamentally, the efficiency should lie close to 1 . An efficiency value “greater than 0.9 indicates a very satisfactory model performance”, while a value “in the range $0.8 - 0.9$ indicates an acceptable (or good) model, and values less than 0.8 indicate an unsatisfactory model” (Coulibaly and Baldwin, 2005, p. 170).

The Mean Squared Error (MSE) is a widely used measure to quantify the amount the predicted value differs from the observed value on average. Here also the Root Mean Squared Error (RMSE) is used to give the measure the same dimensions as the predicted values. An RMSE of 0 means that predicted and observed time series are identical. The RMSE is defined as:

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (Q_{o,i} - Q_{p,i})^2} \quad (6.3)$$

Because of taking the squared errors, the RMSE uprates outliers. Large errors are weighted much more heavily than small ones. The Mean Absolute Error (MAE) on the other hand treats all sizes of errors evenly (cf. Witten and Frank, 2005). The MAE is defined as:

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |Q_{o,i} - Q_{p,i}| \quad (6.4)$$

The Relative Absolute Error (RAE) is expressed as a percentage. The lower the percentage value the better the performance of the model. An RAE of 100 % indicates the same performance as predicting the mean value.

The RAE is defined as:

$$\text{RAE} = \frac{\sum_{i=1}^n |Q_{o,i} - Q_{p,i}|}{\sum_{i=1}^n |Q_{o,i} - \bar{Q}_o|} \quad (6.5)$$

The Root Relative Squared Error (RRSE) is defined similar to the efficiency coefficient. The principal modification is that the root is applied to the relative squared error. In machine learning this measure is more common than the efficiency coefficient E .

RRSE is defined as:

$$\text{RRSE} = \sqrt{\frac{\sum_{i=1}^n (Q_{o,i} - Q_{p,i})^2}{\sum_{i=1}^n (Q_{o,i} - \bar{Q}_o)^2}} \quad (6.6)$$

Many performance measures have been introduced so far. In machine learning often several indicators like correlation coefficient, RMSE, MAE, RAE, and RRSE are taken as performance measures but “there is no single performance measure that has become standard” (Trigg, 1998, p. 1). Witten and Frank (2005) comment on this question that “in most practical situations the best numeric prediction method is still the best no matter which error measure is used” (p. 179). Here the efficiency coefficient is taken as the reference in case of doubt.

6.4 PREPROCESSING

Before applying the available dataset to data-driven models, it is important to examine it. This process is usually referred to as preprocessing. It includes checking for data samples that deviate conspicuously from neighbouring values (e.g. plausibility check) as well as defining an appropriate input vector. It is also advisable to divide the data into sets for training, validating, and testing the data-driven models. Below the applied preprocessing steps are described.

6.4.1 PLAUSIBILITY CHECK

The examination of the rain gauge data was done by an official institution as described in Section 2.3. The same applies for the runoff data. Furthermore, it was checked that the data were transformed correctly into the format necessary for the software tools applied. The weather radar data were automatically read out from the archive with manual inspections.

6.4.2 DATA ANALYSIS

The radar and rain gauge rainfall datasets as well as the runoff datasets have been investigated with statistical methods and the correlation between rainfall and runoff data was determined. This analysis has two main reasons:

- to define for what time span the runoff can be predicted in advance for a certain catchment i.e. the forecast time and
- to find an appropriate input vector

6.4.2.1 Defining the forecast time

A significant measure for this analysis is the cross correlation between rainfall and runoff. The cross correlation is a measure of similarity of two different signals. It is a function of the relative time between the signals. The cross correlation coefficient between rainfall and runoff has been used to identify the time lag (offset) where the similarity is highest.

Rain gauge as well as weather radar series were investigated and the analysis showed that the time lags with the highest cross correlation coefficients between rainfall and runoff series lie between 10 to 16 time steps (150 and 240 minutes), depending on the position within the catchment where the rainfall was measured. Table 6.1 shows the time lags in detail.

The analysis revealed that the correlation coefficients of rain gauge and radar measurements vary significantly. None of the radar pixels achieved the maximum correlation of the rain gauge (30.08 %). The poorer correlation coefficients of the radar measurements are believed to originate mainly from two causes. Firstly, the radar data are archived with 14 quantisation intervals. Because of these fixed levels of rain rate the changes do not appear as directly as in quasi continuous rain gauge data. This effect, however, is relativised by the 15-minute integration time of the radar data. Secondly, the indirect measurement and the consequential sources of possible errors are believed to have an impact. The effect of overshooting beam

(see Section 3.7.3) was detected on several occasions where the radar did not report much precipitation because of low altitude cloud tops. However, high reaching convective rain cells, the dominant source of high water and floods in this area, can be detected with good visibility by the weather radar station on Mt. Zirbitzkogel.

Table 6.1. Time lag with maximum cross correlation coefficient between rainfall and runoff series.

Measuring site	Cross correlation coeff. [%]	Time lags [15 min]
Rain gauge site	30.08	11
Weather radar		
Pixel* 1	13.19	13
2	13.14	13
3	13.93	13
4	13.22	13
5	13.52	13
6	14.18	13
7	13.73	13
8	15.20	15
9	13.65	13
10	15.38	13
11	14.25	13
12	13.72	13
13	16.50	15
14	9.82	16
15	14.91	13
16	16.76	14
17	14.43	13
18	10.55	14
19	15.46	12
20	17.72	13
21	17.85	12
22	14.69	11
23	15.78	12
24	13.80	11
25	13.54	11
26	14.41	11
27	14.92	11
28	13.74	11
29	14.57	11
30	14.54	11
31	16.14	11
32	16.16	10

*The pixel number refers to the position in Figure 6.1.

The cross correlation analysis also yields important information concerning the output. The question is: For what time span a forecast can be given? Naturally one is interested in a long forecast period, but a small catchment often does not allow this. The shortest time lag between precipitation and runoff series that could be found in the cross correlation analysis is 10 time steps (150 minutes). Thus the models were trained on the runoff 150 minutes ahead of time ($Q_{t+150 \text{ min}}$) which is henceforth named the target runoff.

6.4.2.2 Defining the input vector

This section covers the issue how to integrate the (distributed) radar measurements into the data driven models. It demonstrates why it is not feasible to use all single radar pixels as coordinates of the input vector and introduces an approach to reduce the dimension of the input vector without losing too much information.

One may intuitively include all available precipitation data in the input vector. But taking the rain gauge data and all of the 32 weather radar measurements would lead to a huge number of input parameters even for such a small catchment. An effective training would not be possible because the more inputs the more network parameters (weights) are to be determined. A fully connected feed-forward network with six hidden neurons, one neuron in the output layer, and six inputs has 42 weights to be determined. With 36 inputs the number of weights increases to 222. The more weights the more training data are necessary. Furthermore, it is a far more complex task for the training algorithm to update the weights. Moreover, the risk of getting stuck in a local minimum and not finding the global minimum is much higher. That's why the dimension of the input vector has to be reduced.

An established method to eliminate redundancy in the input vector is the principal component analysis (e.g. Demuth and Beale, 1998) which removes those coordinates of the input vector contributing the least to the variation in the data set. But this means that not all radar measurements would be part of the input vector and the main advantage of the radar – the gapless spatial coverage – would be lost.

The method used to reduce the dimension of the input vector without losing too much information was to group several radar pixels. All radar pixels showing correlation maxima with the runoff at the same time lag were grouped to one cluster. (The pixels forming one cluster, however, not necessarily adjoin.) As input the average rain rate of each cluster is taken. This leads to a smaller input vector. The initial 32 radar pixels were reduced to seven clusters. Table 6.2 shows correlation coefficient and time lag of the rain gauge time series and the clusters which form the input vector. The rain gauge time series is left unmodified. The advantage of this technique is that information of each pixel above the catchment is still represented in the dataset. Because of the bigger clusters, the information where exactly a small convective shower cell occurred is lost but the rainfall amount within the area represented by the cluster is available and the time lag when the rainfall shows the highest correlation with the runoff series is known.

Table 6.2. Time lags and cross correlation coefficients between runoff and precipitation measurements.

Measuring site	Cross correlation coeff. [%]	Time lag [15 min]
Rain gauge site	30.08	11
Cluster 1 (1 pixel)	16.16	10
Cluster 2 (9 pixels)	16.12	11
Cluster 3 (3 pixels)	16.90	12
Cluster 4 (14 pixels)	17.41	13
Cluster 5 (2 pixels)	14.32	14
Cluster 6 (2 pixels)	17.67	15
Cluster 7 (1 pixel)	9.82	16

Thus, in this study, the radar data are represented in clustered form in the input vector. For a fine-tuning of the input radar data the input-output behaviour is studied in detail.

Input-output behaviour

The proposed model aims to predict the future runoff with rain gauge, radar, and actual runoff as inputs. Thus, the model is not a pure rainfall-runoff model because it uses the actual runoff (Q_t) as additional input coordinate. This is important since rainfall data typically contain numerous zero values, but “the condition of rainfall and no-rainfall is difficult to identify by DDM [annotation: data-driven models] with only rainfall time series as inputs” (Solomatine and Dulal, 2003, p. 404).

So far in the literature mainly rain gauge data have been used in ANNs to predict the future runoff, often on a daily or hourly basis. However, for the considered catchment a higher temporal and spatial resolution is needed, since it typically lasts only a few hours between incipient rain and peak runoff. This high temporal and spatial resolution of the precipitation data make the input variables jagged and result in a complex input-output behaviour.

The rainfall and runoff curves of one rainfall event are exemplified in Figure 6.2. The rain rate from the rain gauge and from Cluster 7 are contrasted with actual and target runoff. Typically the weather radar reports precipitation before it can be detected by the rain gauge. At the peak of the radar and rain gauge measurements the runoff level may not even have increased as shown in this example. Here the runoff peak occurs when the precipitation is essentially over. The target runoff is the actual runoff 10 time steps (150 minutes) shifted against the time axis. This is the curve to be predicted. In the present case it can be seen that the target runoff starts to increase even before any precipitation has been detected which suggests how difficult a precise forecast is for any hydrological model.

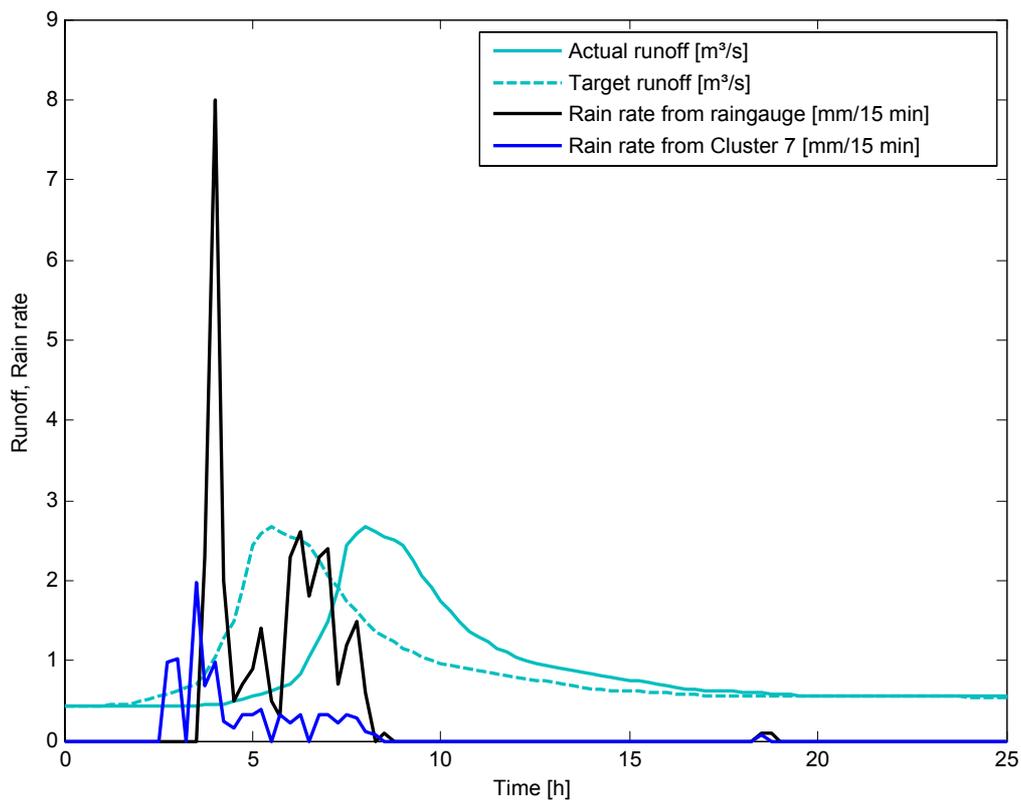


Figure 6.2. Exemplary rainfall and runoff curves of one rainfall event.

The example above points out the highly nonlinear relationship between precipitation measurements from radar and rain gauge and the target runoff. Below simple data manipulation methods are introduced that aim to increase the correlation between target runoff and precipitation measurements and thus presumably improve the quality of the forecast. These methods are time shifting of the precipitation curves and applying a moving average on them.

Time shifting

The seven clusters differ in the time lag with respect to the runoff. The time lags range from 10 to 16 time steps. Cluster 1 exhibits its maximum correlation with the runoff when shifted 10 time lags to the right. The other clusters and the rain gauge measurement, however, do show their maxima at higher lead times. Thus, the idea is to use forward-shifted clusters and rain gauge measurements as input vector coordinates. These clusters are shifted forward just as far as to show a maximum correlation with the target runoff. The implementation of a time shift can also be argued from a physical point of view. In fact it takes a certain time for the precipitation fallen at a remote part of the catchment to contribute to the runoff. Actually it seems reasonable to shift precipitation inputs from remote locations in time since they can by no means immediately contribute to the runoff. Figure 6.3 shows the same rainfall event as above with time shifted precipitation measurements to coincide on average with the desired target runoff.

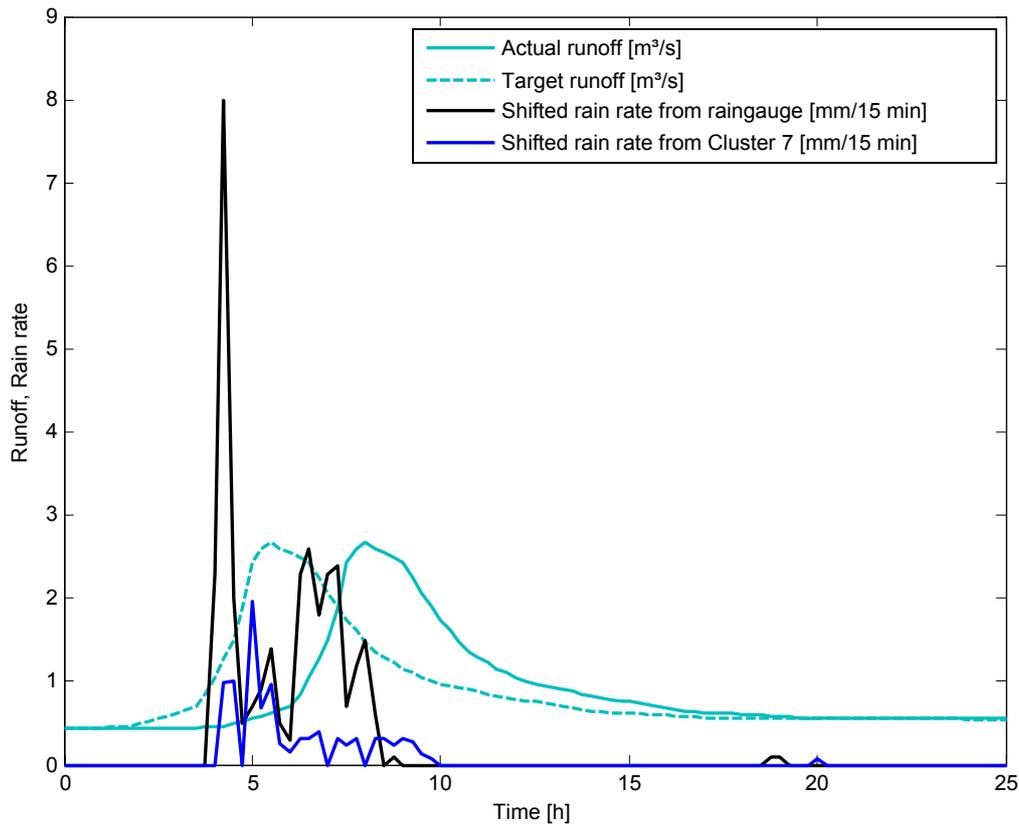


Figure 6.3. Runoff curve and time shifted precipitation measurements. The single precipitation measurements are shifted forward just as far as to show a maximum correlation with the target runoff.

Moving rainfall average

The precipitation curves in Figures 6.2 and 6.3 are rather jagged, whereas the runoff time series is quite smooth at the given temporal resolution of 15 minutes. This is also reflected in the poor correlation coefficients between these signals. The correlation becomes better when averaging the precipitation measurements over a certain time span. Averaging the rain rate seems reasonable since short term variations of the rainfall are not relevant for the runoff. An example with a moving rainfall average over a preceding period of 10 time steps can be seen in Figure 6.4. The precipitation curves thus averaged much better match the target runoff. Table 6.3 shows that the correlation coefficients increase by about 10 percentage points. It has to be considered though that applying a moving rainfall average on the precipitation curve reduces its time lag with the runoff curve. This may not be desirable because in the figures the averaged precipitation curves move to the right. Most averaged precipitation curves now have their correlation maximum less than 10 time steps ahead which is the time lag aimed to forecast.

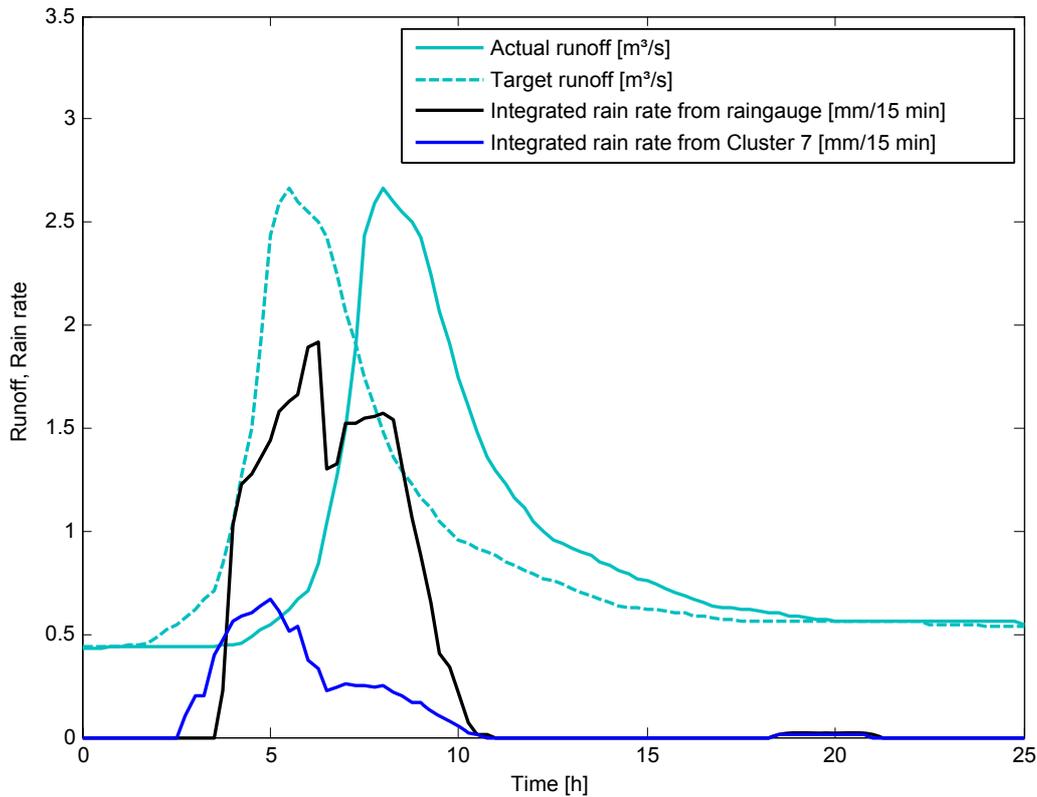


Figure 6.4. Runoff curve and averaged precipitation measurements. Moving rainfall average over a preceding period of 10 time steps.

Table 6.3. Time lags and cross correlation coefficients between runoff and integrated precipitation measurements (moving rainfall average over 10 preceding time steps).

Measuring site	Cross correlation coeff. [%]	Time lag [15 min]
Rain gauge site	43.29	7
Cluster 1 (1 pixel)	26.29	8
Cluster 2 (9 pixels)	25.58	8
Cluster 3 (3 pixels)	26.69	8
Cluster 4 (14 pixels)	25.40	10
Cluster 5 (2 pixels)	25.65	9
Cluster 6 (2 pixels)	27.44	10
Cluster 7 (1 pixel)	19.30	10

The relocation of the averaged curve is a cause for concern. Theoretically the relocation to the right can be avoided by averaging over an equal number of preceding and following rainfall measurements. But from the forecast perspective the following measurements lie ahead and thus are not available at this particular time. To meet these timing concerns, the modification proposed here uses the preceding moving rainfall average with such an integration time that the maximum correlation occurs at a time lag of 10. Thus, the integration time varies for the different precipitation curves (Table 6.4). Because the integration time is often shorter than

10, as in the example before, the cross correlation coefficients often deteriorate. However, on average the timing of every single of the thus averaged precipitation curves complies with target runoff. Figure 6.5 shows the rainfall and runoff curves with this technique applied. Admittedly, the correct timing holds for the data on average and does not fit for every single runoff peak.

Table 6.4. Cross correlation coefficients between runoff and precipitation measurements with flexible integration time.

Measuring site	Cross correlation coeff. [%]	Integration time
Rain gauge site	35.61	3
Cluster 1 (1 pixel)	21.57	4
Cluster 2 (9 pixels)	22.20	5
Cluster 3 (3 pixels)	24.20	6
Cluster 4 (14 pixels)	25.40	10
Cluster 5 (2 pixels)	24.99	9
Cluster 6 (2 pixels)	27.44	10
Cluster 7 (1 pixel)	19.75	11

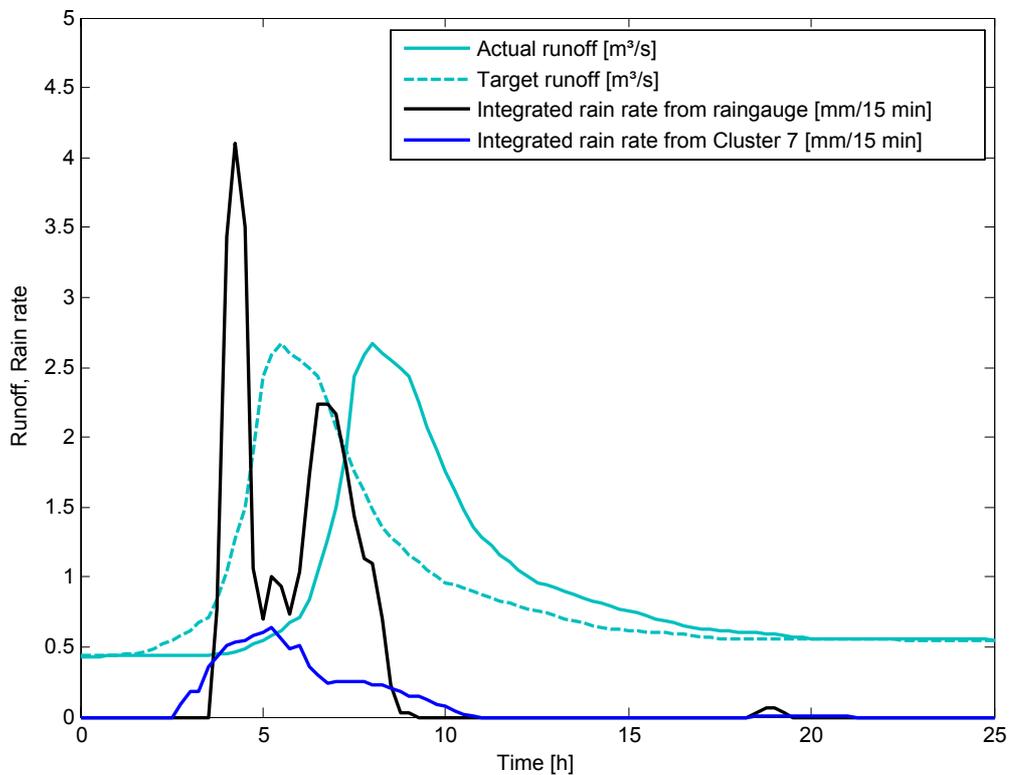


Figure 6.5. Runoff curve and precipitation measurements with flexible integration time.

So far different kinds of simple precipitation data manipulation methods have been proposed. Below they are listed with the abbreviations used in the results section.

- Original chronological sequence (O)
- Time shifted (TS)
- Averaged with fixed integration time of 10 time steps (I10)
- Averaged with flexible integration time (IFlex)

Representing the Hydrological state

So far the actual hydrological state of the catchment has not been considered apart from the actual runoff (Q_t) that is taken as additional input coordinate. The actual hydrological state is important to estimate the effect of the rainfall on the runoff. Because of the lack of more direct indicators of the hydrological state such as soil moisture, many studies use indirect indicators like previous runoff values or cumulative rainfall over a longer period before the rainfall event. Here these two methods of hydrological state representation are applied. Below the different hydrological state representations are described.

Cumulative rainfall configuration

In addition to the standard configuration, the cumulative rainfall of all precipitation measurements over a longer period is taken as additional input parameter. Three additional input coordinates are added which differ in the length of the integration time: one day, one week and two weeks. The integration time here is much longer than the 10 time steps used above because the purpose is different. As a data manipulation method, the input was averaged to better fit the target. Here, the longer integration period is aimed to account for the rainfall history and thus the hydrological state of the catchment.

Previous runoff configuration

This configuration differs from the standard configuration given above in the number of runoff measurements presented to the model. In addition to the actual runoff (Q_t) also three preceding runoff measurements are part of the input vector ($Q_{t-15 \text{ min}}$, $Q_{t-30 \text{ min}}$, and $Q_{t-45 \text{ min}}$). The choice of the number of preceding runoff measurements was a trade-off between additional information and increasing complexity with each added input coordinate.

6.4.3 TEST-, TRAINING-, AND VALIDATION-DATA

The next step of the preprocessing was to divide the available dataset into sets for training, validating, and testing the data-driven model. These three subsets have different functions. The training dataset is necessary for determining the model's parameters (weights and biases of the neural network; and leaves with their linear regression functions and splitting criteria of model trees). When building a data-driven model, the engineer can also influence the outcome (e.g. by simulating different numbers of nodes in the hidden layer of ANNs, or by deciding to

split an existing leaf of a decision tree again). Thus several slightly different models result. The function of the validation dataset is to simulate these models and to evaluate them according to significant performance measures. Often this is seen as the test run. But in the strict sense the validation data can not be considered as independent, because they are involved in the model selection process. To avoid any form of influence, here three datasets are used. The third set is the test dataset. It is not involved in any form in the training or model selection process and is applied strictly after the whole model has been built. The test data are the unseen examples that are used to rate the models in the end.

Essential for a good performance of the model is careful selection of these datasets. Hydrological parameters show considerable seasonal variations (see Section 2.4.1). That's why it is important that the different datasets do not originate from one season only. A data-driven model that was trained to predict the runoff with data originating from the winter season only will not perform well in predicting the reaction of the same catchment to a convective summer storm. In the present case where data of a period of only one year are available, the selection of the subsets is even more critical.

Here a method was used that ensures that each of the three subsets contains random data from the different seasons. Therefore, the whole data set was divided into rainfall events and their corresponding runoff hydrographs. These events were classified into the seasons they belong. Thereafter, training, validation, and test subsets were formed by randomly assigning events from different times of the year to the subsets.

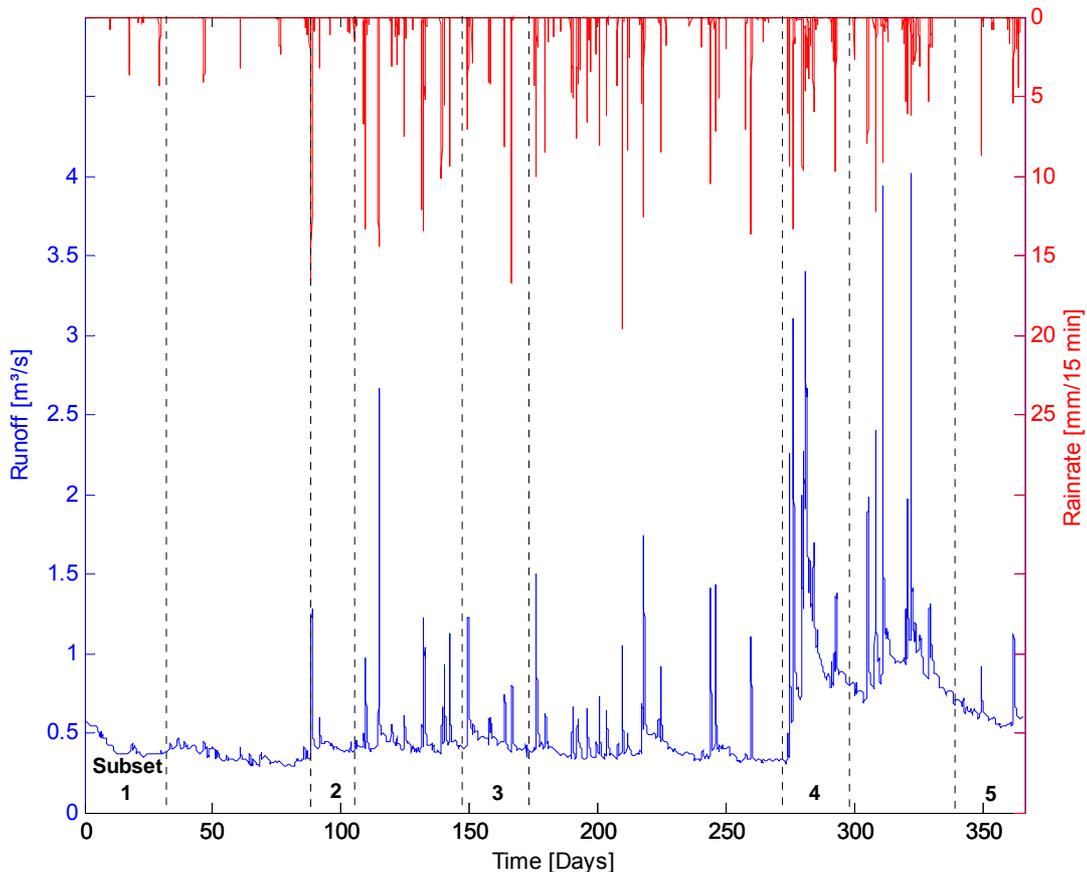


Figure 6.6. Subsets used for testing (numbered 1 to 5).

6.5 FEED-FORWARD NEURAL NETWORK APPROACH

In this section the architecture of the ANN for rainfall-runoff prediction is derived. For identifying the architecture of a multilayer neural network associated with determining the number of neurons in each layer, the trial-and-error approach is still the most common (e.g. Imrie *et al.*, 2000; Pan and Wang, 2004; Toth *et al.*, 2000). The architecture is dependent on the problem to be solved and that's why no general solution can be given. Here the setup of the ANN for the different configurations is shown.

An area of conflict is that a small network may have insufficient degrees of freedom (weights and biases) to realistically represent the relationship between rainfall and runoff. A large network on the other hand with many weights to be adapted may memorise particularities in the training data and is therefore not able to generalise. Thus, the method used to determine the architecture of the ANN was to start with a small network with one node in the hidden layer. The number of nodes in the input and output layer is given by the number of input and output coordinates. Because of nine nodes in the input layer, one node in the hidden layer, and one node in the output layer, the network architecture is called 9-1-1. During the training process the error on the validation set was monitored. Here the Mean Squared Error (MSE) was taken. When the validation error increased the training was stopped and the minimum of the validation error was taken as indicator for best performance. This approach is sometimes called *early stopping*. It was applied to prevent the network from overfitting.

The networks were trained several times with randomly chosen network weights and biases. The performance of these networks showed variations because they did not lead to exactly the same parameters. This phenomenon is well known and related to local and global minima in the error surface. Sometimes the training algorithm may get stuck in a local, instead finding a global minimum. Moreover, because of the early stopping approach the training may come to a halt due to an increasing validation error before the global minimum is found. This is why some authors do not present one single training trial but an ensemble over several training trials (Gaume and Gosset, 2003; de Vos and Rientjes, 2005). An ensemble is believed to be a better indicator for the average performance than a single training trial. This is why here also an ensemble of network trials is used. The ensemble size is 20 and the median of the ensemble is given in the tables and time series plots below.

Step by step the number of nodes in the hidden layer was increased and the ensemble of network trials was again trained with randomly chosen parameters. Table 6.5. gives the MSE concerning the validation dataset. The precipitation data in original chronological sequence (O) were used. As Table 6.5 shows, the MSE of the smallest network with one hidden node is highest. With increasing number of hidden neurons the MSE decreases. The best configuration concerning the MSE exhibits the 9-4-1 architecture. A further increase of the number of neurons does not necessarily lead to a decrease of the MSE. Networks with more hidden layers can theoretically also perform the same function, but it turned out that practically the training function did not always find such good solutions in the higher dimensional space. Because the number of input coordinates is constant 9 and this model selection procedure is very time consuming it was not repeated with the other techniques of precipitation data manipulation (TS, I10, and IFlex). For all these configurations the 9-4-1 architecture (see Figure 6.7) was used.

Table 6.5. MSE on the validation dataset as model selection criterion.

Architecture	MSE [(m ³ /s) ²]
9-1-1	0.0109
9-2-1	0.0099
9-3-1	0.0094
9-4-1	0.0090
9-5-1	0.0091
9-6-1	0.0091
9-7-1	0.0090
9-8-1	0.0091
9-9-1	0.0090

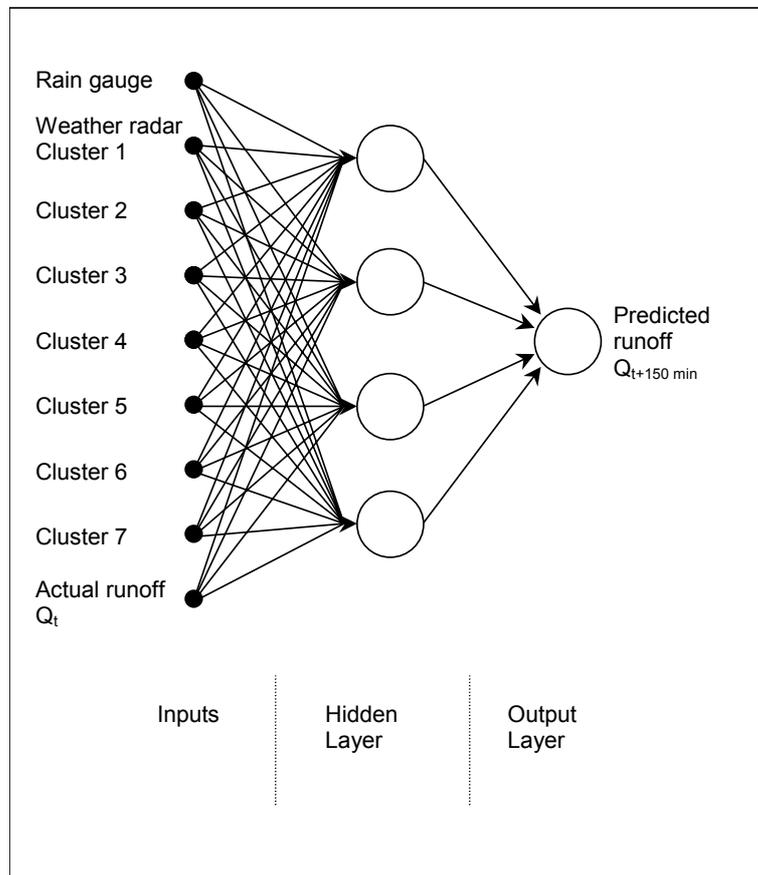


Figure 6.7. Network architecture used for the rainfall-runoff prediction.

The two other configurations which account for the hydrological state, *Previous runoff configuration* and *Cumulative rainfall configuration* comprise more input coordinates (a total of 12 in each case). The input parameters are different though. Therefore, the model selection procedure was carried out separately for these configurations.

Table 6.6 gives the figures of the model selection procedure for the ANN in cumulative rainfall configuration. The model with 5 hidden neurons works best on the validation dataset. Hence, this architecture is used for testing.

Table 6.6. Model selection for the cumulative rainfall configuration.

Architecture	MSE [(m ³ /s) ²]
12-1-1	0.0105
12-2-1	0.0103
12-3-1	0.0099
12-4-1	0.0098
12-5-1	0.0096
12-6-1	0.0099
12-7-1	0.0096
12-8-1	0.0100
12-9-1	0.0098

In previous runoff configuration the model with 3 hidden neurons works best on the validation dataset as shown in Table 6.7.

Table 6.7. Model selection for the previous runoff configuration.

Architecture	MSE [(m ³ /s) ²]
12-1-1	0.0089
12-2-1	0.0087
12-3-1	0.0085
12-4-1	0.0087
12-5-1	0.0086
12-6-1	0.0086
12-7-1	0.0088
12-8-1	0.0087
12-9-1	0.0086

6.6 MODEL TREE APPROACH

In neural network training, identifying the proper architecture means determining the number of neurons in the hidden layer. When setting up model trees, which are piece-wise linear models, the question is: How many linear equations should be included into the tree? In the used *Weka* software, the number of linear equations can be indirectly controlled with the parameter *minNumInstances* which determines the minimum number of instances at which one node is considered for splitting into two.

The method to set up the model tree was to vary the parameter *minNumInstances* and monitor the performance on the validation dataset. The configuration which exhibited the best MSE concerning the validation dataset was taken and simulated with the hitherto uninvolved test dataset.

When the precipitation data in original chronological sequence (O) were used as input configuration a *minNumInstances* of 60 leads to the best performance on the validation dataset. The resulting model tree consists of 12 linear equations. This model selection procedure was executed for every input configuration anew. Henceforth the number of linear equations is given in brackets as in “MT (12)” to indicate the complexity of the model tree.

6.7 RESULTS

This section gives the results of the ANNs and MTs for rainfall-runoff modelling. First, the performances of the machine learning approaches in standard configuration with the different manipulations of input data are presented. Subsequently the benefits of the two other input vector configurations for hydrological state representation are shown.

6.7.1 RESULTS OF STANDARD CONFIGURATION

The data-driven models that have been optimised in the training process are now tested with the hitherto unseen test dataset. Their capacity in terms of the performance figures can be seen in Table 6.8. The ANN and the MT approaches were simulated with four variations of input data respectively (O, TS, I10, and IFlex; see Section 6.4.2.2). The figures of the MT represent the performance of one single model, whereas the figures of the ANN show the median of an ensemble of 20 test trials.

The ANN with time shifted input data (TS) performs best according to all performance measures. The MTs do not show such a uniform pattern with respect to the different input variations. The MT with the input data in original chronological sequence (O) performs best according to most measures including efficiency coefficient, correlation, and MSE. Solely concerning the absolute error measures (MAE and RAE) other configurations perform better.

Table 6.8. Performance of ANNs and MTs in standard configuration with different configurations of input data (The best performance figures of the ANN and MT approaches are underlined).

Architecture		RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 9-4-1 / O	Total	0.1146	0.0131	0.0307	95.2	15.0	34.8	87.9
	Subset 1	0.0100	0.0001	0.0049	99.1	8.5	14.4	97.9
	2	0.0444	0.0020	0.0142	91.1	31.5	43.0	81.6
	3	0.0546	0.0030	0.0200	90.0	32.3	51.2	73.8
	4	0.2424	0.0587	0.0982	92.8	28.4	47.3	77.6
	5	0.0319	0.0010	0.0160	91.6	33.5	45.5	79.3
ANN 9-4-1 / TS		<u>0.1113</u>	<u>0.0124</u>	<u>0.0288</u>	<u>95.3</u>	<u>14.1</u>	<u>33.8</u>	<u>88.6</u>
	Subset 1	0.0093	0.0001	0.0045	99.1	7.8	13.3	98.2
	2	0.0473	0.0022	0.0137	89.9	30.3	45.7	79.1
	3	0.0541	0.0029	0.0188	90.2	30.3	50.7	74.3
	4	0.2344	0.0549	0.0932	92.7	26.9	45.7	79.1
	5	0.0324	0.0011	0.0141	90.8	29.4	46.2	78.6
ANN 9-4-1 / I10		0.1231	0.0152	0.0307	93.9	15.0	37.4	86.0
	Subset 1	0.0077	0.0001	0.0044	99.5	7.6	11.0	98.8
	2	0.0499	0.0025	0.0137	88.1	30.4	48.2	76.8
	3	0.0708	0.0050	0.0214	86.1	34.6	66.4	55.9
	4	0.2573	0.0662	0.0986	90.4	28.5	50.2	74.8
	5	0.0329	0.0011	0.0154	90.6	32.2	46.9	78.0
ANN 9-4-1 / IFlex		0.1159	0.0134	0.0302	94.8	14.8	35.2	87.6
	Subset 1	0.0086	0.0001	0.0046	99.4	8.0	12.3	98.5
	2	0.0461	0.0021	0.0144	90.7	32.0	44.6	80.1
	3	0.0606	0.0037	0.0208	88.3	33.6	56.8	67.8
	4	0.2440	0.0595	0.0952	91.7	27.5	47.6	77.3
	5	0.0303	0.0009	0.0161	92.3	33.7	43.3	81.3
MT (12) / O		<u>0.1202</u>	<u>0.0145</u>	<u>0.0418</u>	<u>94.8</u>	<u>20.4</u>	<u>36.5</u>	<u>86.7</u>
	Subset 1	0.0125	0.0002	0.0084	99.0	14.6	18.0	96.8
	2	0.0503	0.0025	0.0157	87.5	34.8	48.7	76.3
	3	0.0550	0.0030	0.0192	88.5	31.0	51.6	73.4
	4	0.2498	0.0624	0.1150	91.5	33.2	48.7	76.3
	5	0.0569	0.0032	0.0477	85.3	99.7	81.1	34.2
MT (14) / TS		0.1281	0.0164	0.0413	93.8	20.1	38.9	84.9
	Subset 1	0.0121	0.0001	0.0085	99.1	14.7	17.4	97.0
	2	0.0519	0.0027	0.0163	88.2	36.0	50.2	74.9
	3	0.0521	0.0027	0.0192	89.8	31.0	48.8	76.2
	4	0.2702	0.0730	0.1194	89.4	34.5	52.7	72.2
	5	0.0495	0.0024	0.0406	87.3	84.8	70.5	50.3
MT (10) / I10		0.1250	0.0156	<u>0.0388</u>	94.0	<u>18.9</u>	38.0	85.6
	Subset 1	0.0139	0.0002	0.0075	98.6	12.9	20.1	96.0
	2	0.0637	0.0041	0.0159	78.8	35.1	61.6	62.1
	3	0.0549	0.0030	0.0175	90.6	28.3	51.5	73.5
	4	0.2603	0.0678	0.1076	90.0	31.1	50.8	74.2
	5	0.0516	0.0027	0.0434	86.0	90.6	73.6	45.9
MT (9) / IFlex		0.1240	0.0154	0.0406	94.1	19.8	37.7	85.8
	Subset 1	0.0121	0.0001	0.0074	99.5	12.8	17.4	97.0
	2	0.0478	0.0023	0.0130	89.8	28.9	46.2	78.6
	3	0.0760	0.0058	0.0224	81.3	36.2	71.2	49.3
	4	0.2535	0.0643	0.1080	91.0	31.2	49.5	75.5
	5	0.0575	0.0033	0.0484	86.5	101.2	82.0	32.8

On the total test dataset, the variations between the different approaches are not very pronounced. Especially CORR and E vary only about a few percentage points. To show the differences more distinctly, the performance measures are listed for the five subsets separately. These figures reveal a more differentiated picture. Throughout, Subset 1 is predicted best by all models. On this subset the ANN models exhibit an efficiency coefficient E of almost 98 to 99 %. The model trees reach 96 to 97 %. The prediction of the model that performs best on this subset (ANN/I10) is shown in Figure 6.8(a). The runoff of this dataset does not show much variation. The ANN prediction features an efficiency coefficient E of

98.8 %. The worst forecast in terms of E exhibits the MT in IFlex configuration on Subset 5 (Figure 6.8b). The poor 32.8 % result from a continuing underestimation of the runoff curve.

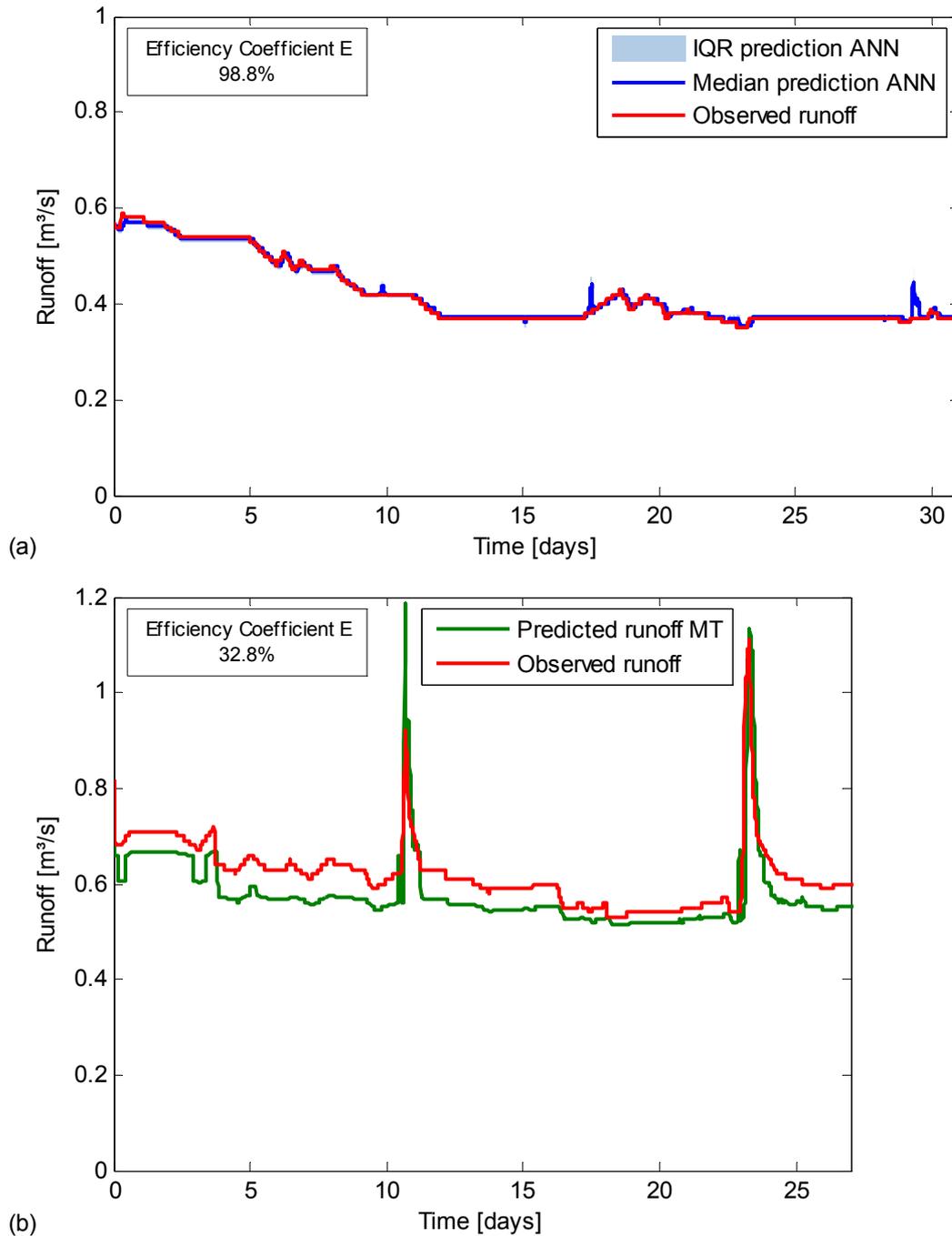


Figure 6.8. Examples of good and poor runoff predictions. The best efficiency coefficient on a subset exhibits an ANN model (a) and the poorest a model tree (b). IQR stands for interquartile range and gives the difference between the third and first quartile of the ANN ensemble.

A comparison between the total figures and the figures for the associated subsets in Table 6.8 often reveals a high discrepancy. The efficiency coefficient E of the MT in IFlex configuration for instance of nearly 86 % seems out of scale considering the measures of the subsets with an E of only 33 % in Subset 5 and 49 % in Subset 3. On the other hand, the total figure of the ANN (IFlex) is not much higher than that of the MT although the performance on all subsets is far better. The reason for this behaviour is that the performance measures E , RAE, and RRSE are made relative to the mean observed value. The five subsets, originating from different seasons, exhibit quite different mean values. When calculating the performance of the whole composite dataset, the mean value over all subsets is taken. Although mathematically correct, the information value of the measures applied on composite datasets is limited.

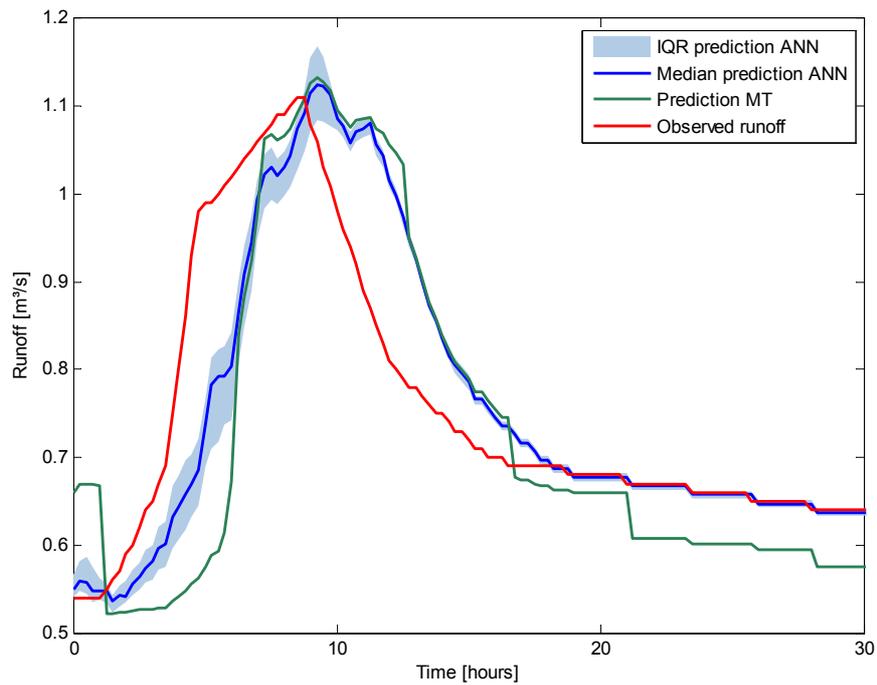
Thus, in the following the weighted average figures of the test dataset's performance measures are given (Table 6.9). These values are weighted with the length of the subset. The weighted averages represent a more conservative measure and better reflect the overall performance than the measures for the total dataset. Differences between the performance of the different approaches emerge more explicitly. Obviously MSE and MAE do not show any difference between total and weighted average calculation. This is the reason why the MSE was chosen as stopping criterion in the training process. Its value is significant for composite datasets as well. The MAE was not used because large discrepancies should weigh much more than small ones.

Table 6.9. Weighted average (w. av.) performance measurements.

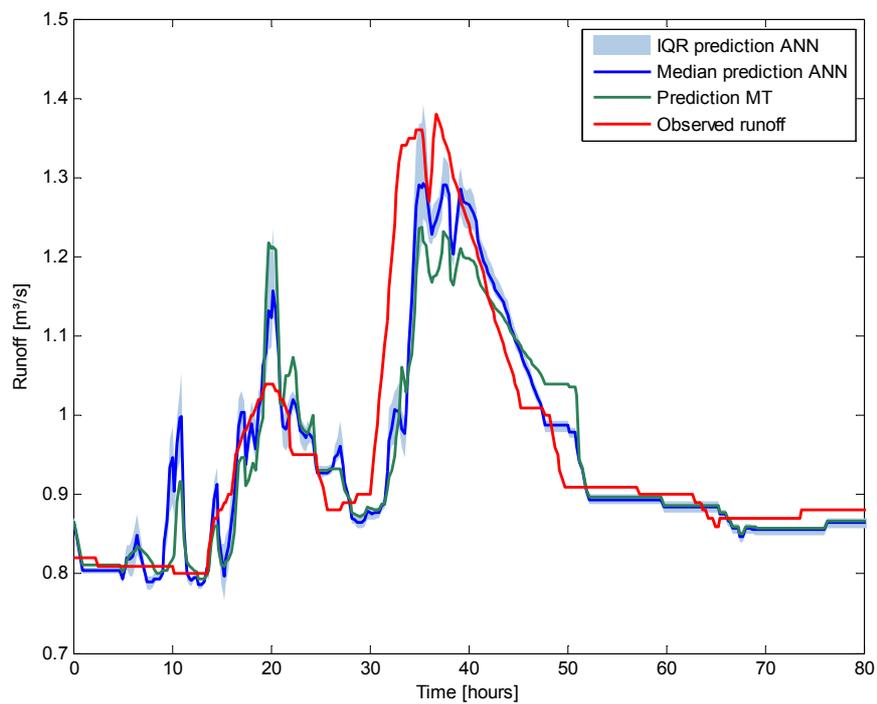
Architecture		RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 9-4-1/ O	w. av.	0.0759	0.0131	0.0307	<u>93.3</u>	25.8	39.1	<u>82.7</u>
ANN 9-4-1/ TS	-	<u>0.0745</u>	<u>0.0124</u>	<u>0.0288</u>	93.0	<u>23.9</u>	<u>38.9</u>	<u>82.7</u>
ANN 9-4-1/ I10	-	0.0826	0.0152	0.0307	91.5	25.7	42.9	77.8
ANN 9-4-1/ IFlex	-	0.0770	0.0134	0.0302	92.9	25.9	39.5	81.8
MT (12) / O	w. av.	<u>0.0842</u>	<u>0.0145</u>	0.0418	90.9	42.6	48.7	71.7
MT (14) / TS	-	0.0863	0.0164	0.0413	<u>91.3</u>	39.9	<u>46.7</u>	<u>74.8</u>
MT (10) / I10	-	0.0873	0.0156	<u>0.0388</u>	89.9	<u>39.3</u>	49.7	71.8
MT (9) / IFlex	-	0.0890	0.0154	0.0406	90.0	42.4	52.6	66.7

As in Table 6.8, here again the best performance figures of the ANN and MT approaches are underlined. It becomes clear that the calculation of the weighted average performance measurements has an effect on the ranking of the input variations of the data-driven models. The ANN with time shifted input data (TS) no longer performs best according to all performance measures. The MT (TS) now performs best according to CORR, RRSE, and E .

Possible time offsets between predicted and observed runoff curves are neither directly readable in the tables, nor can they be seen in Figure 6.8 because of the poor temporal resolution. Below this problem is addressed.



(a)



(b)

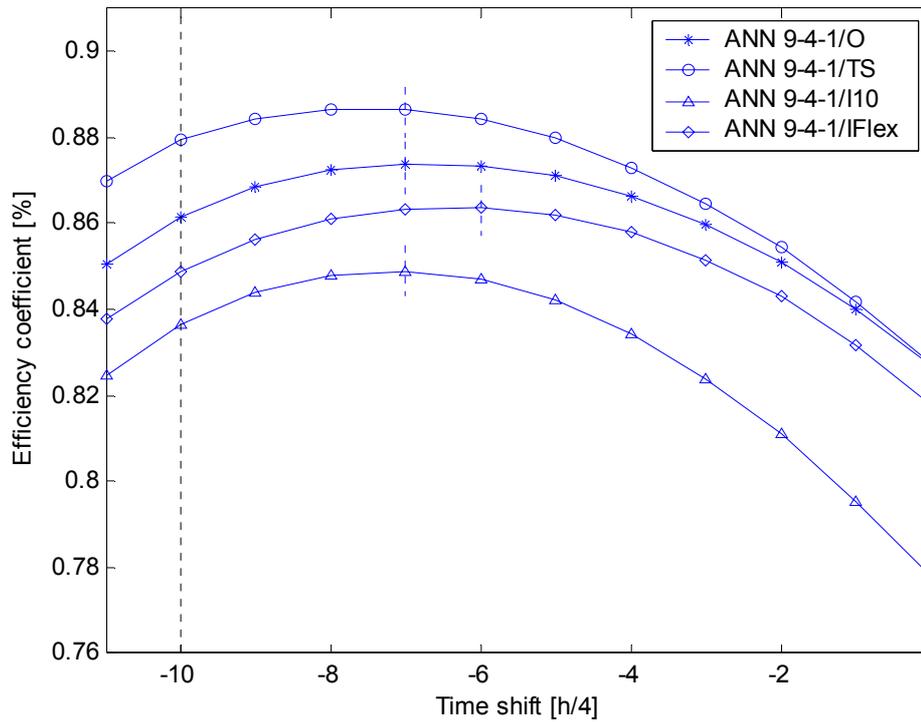
Figure 6.9. Two examples of time lags between predicted and observed runoff curves. The predictions originate in both cases from ANNs and MTs with averaged precipitation input data (IFlex configuration).

Figure 6.9 shows peaks of the test dataset in fine temporal resolution. Here time lags between predicted and observed runoff are identifiable. Figure 6.9(a) shows a lagged prediction of both data-driven approaches (ANN and MT). Although both models reproduce the peak quite well (quantitatively as well as temporally), to a large extent the predictions follow the actual runoff of the input data. The second higher peak of Figure 6.9(b) again shows a pronounced time lag in the prediction. The prediction of the first lower peak shows considerable fluctuations and thus a constant time lag is not detectable.

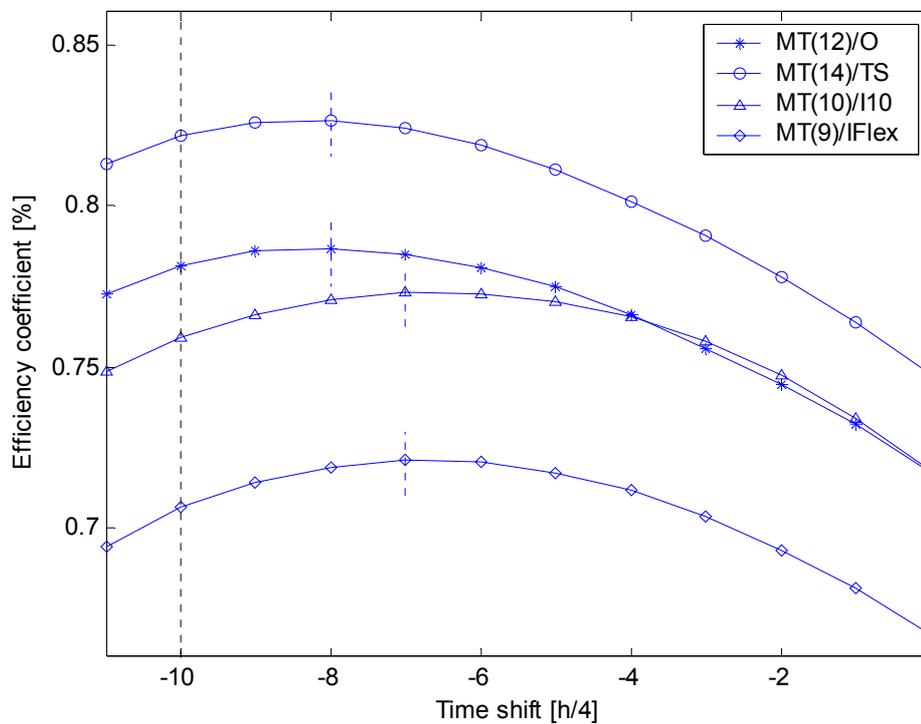
The examples make clear that the predictions are often behind the observed runoff and they also reveal that it is not trivial to decide which of the methods performs better in terms of timing lags. Should the agreement at the peak be the determining factor or the prediction of the rising slope? The method used here to evaluate the time lag was described by de Vos and Rientjes (2005) after an idea by Conway *et al.* (1998). It is based on repeated calculations of the efficiency coefficient E . So far E was calculated between predicted and observed runoff. Henceforth the predicted runoff is repeatedly shifted against the time axis to the observed time series, and E is recalculated for each step. The time shift at which E exhibits a maximum is a measure of the mean time lag of the model. Figure 6.10 shows the time shifts thus calculated of ANN and MT for the various input vector configurations.

The values in Figure 6.10 at zero time shift correspond to the values in Table 6.9. Ideally, the curves should have their maxima at zero time shift, saying the predictions needed not to be shifted against the time axis to exhibit the maximum efficiency coefficient with the target runoff. Effectively, the predicted curves have to be shifted in order to exhibit their maxima. The dashed vertical line at a time shift of -10 represents the limit for a practical prediction. The maximum should lie to the right because otherwise there is no advantage over the actual runoff which is 10 time steps lagged. The offset between -10 and the time shift where a model exhibits its maximum can be seen as the time span the model effectively predicts ahead. It is obvious that ANNs and MTs exhibit a significant time lag. The ANN models effectively predict 3 to 4 time steps ahead, instead of 10, on which they were trained on. The MT perform worse and predict only 2-3 time steps ahead.

Although timing errors “appear to be a common problem in most NN rainfall-runoff forecasting models” (Abrahart *et al.*, 2007, p. 415), the issue of lagged predictions in data-driven models has rarely been addressed in the literature. Often the performance measures are given for the time step the network was trained on, and it is not investigated if the prediction is better for an other time shift. But it seems more straightforward, sincere, and practical to communicate the effective lead time and the corresponding performance measures. In Table 6.10 the performance figures are given for the effective lead time where the efficiency coefficients in Figure 6.10 exhibit their maxima.



(a)



(b)

Figure 6.10. Method to visualise the timing lags of (a) ANN and (b) MT predictions. The graphs show how many time steps the predictions need to be shifted against the time axis to exhibit the maximum efficiency coefficient with the observed runoff.

Table 6.10. Performance of ANNs and MTs for the effective forecast.

Architecture	Trained on	Time shift	Eff.	RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 9-4-1/ O	10	-7	3	0.0669	0.0112	0.0264	95.7	20.7	32.3	87.4
ANN 9-4-1/ TS	10	-7	3	<u>0.0628</u>	<u>0.0098</u>	<u>0.0234</u>	<u>96.0</u>	<u>17.5</u>	<u>30.4</u>	<u>88.7</u>
ANN 9-4-1/ I10	10	-7	3	0.0692	0.0117	0.0244	95.2	18.9	33.2	84.9
ANN 9-4-1/ IFlex	10	-6	4	0.0664	0.0106	0.0258	95.3	21.0	32.6	86.4
MT (12) / O	10	-8	2	0.0731	<u>0.0117</u>	0.0368	94.6	36.5	41.0	78.8
MT (14) / TS	10	-8	2	<u>0.0727</u>	0.0125	0.0353	<u>95.4</u>	<u>33.3</u>	<u>37.9</u>	<u>82.6</u>
MT (10) / I10	10	-7	3	0.0772	0.0119	<u>0.0335</u>	93.2	34.6	44.7	77.3
MT (9) / IFlex	10	-7	3	0.0788	0.0123	0.0361	92.8	37.3	46.5	72.2

The figures in Table 6.10 show that there is a discrepancy between the lead time the models were trained on and the steps ahead they effectively forecast. This discrepancy appears in the time shift. Generally the ANNs perform better than the MTs. There are also differences concerning the input vector configuration. The model trees with averaged precipitation input configurations (I10 and IFlex) exhibit a better time shift than the ones with non integrated precipitation measurements. However, the performance in terms of E and most other measures is worse. The ANN models show a similar behaviour. The ANN/IFlex effectively predicts farthest ahead, but the performance measures are only average. In Table 6.10 the best performance values of the ANN and MT approaches are underlined. Among the ANNs, the configuration with time shifted input data (TS) performs best according to all performance measures. Speaking about MTs, the calculation of the weighted averages leads to the strange situation that the MSE of the MT in TS configuration is worst, whereas the RMSE is best. Nevertheless, concerning most performance measures including E again the TS configuration is superior.

The calculation of the performance figures for the effective lead time has shown that the input configuration both data-driven models work best with, is the one with time shifted input data (TS). The former comparisons based on the lead time the models were trained on were not so clear. Below this TS configuration is taken and it is investigated what improvements the consideration of the hydrological state has.

6.7.2 RESULTS OF THE CUMULATIVE RAINFALL CONFIGURATION (CR)

In this configuration 3 additional input coordinates are added to the TS configuration to account for the hydrological state of the system. They are the cumulative rainfall of all precipitation measurements over one day, one week, and two weeks. Thus the models now have 12 input coordinates instead of 9. In Table 6.11 the performance of data-driven models in this configuration can be seen. As in Table 6.10 the performance measures are given for the effective lead time.

Table 6.11. Performance of ANN and MT in cumulative rainfall configuration for the effective forecast.

Architecture	Trained on	Time shift	Eff.	RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 12-5-1/CR	10	-7	3	0.0686	0.0122	0.0266	95.2	21.4	31.8	87.6
MT (23)/CR	10	-9	1	0.0896	0.0235	0.0433	93.1	31.9	37.6	83.0

The ANN model exhibits a time shift of -7 which is equivalent to the ANN in TS configuration. However, the performance measures, which are again evaluated based on an ensemble of 20 test trials, decreased slightly. Noticeable the time lag of the MT has increased. Now the MT effectively predicts merely one time step ahead. The performance measures of the MT are given for the effective lead time of 1 time step. Some measures (RAE, RRSE, and *E*) seem to have changed for the better compared with the MT in TS configuration in Table 6.10, but it has to be kept in mind that the former performance figures referred to an effective lead time of 2 time steps. Hence, the former MT exhibited an *E* of 82.6 % when forecasting 2 time steps ahead, while this MT reaches an *E* of 83 % when forecasting only 1 step ahead. Altogether the cumulative rainfall configuration does not lead to a real improvement.

6.7.3 RESULTS OF THE PREVIOUS RUNOFF CONFIGURATION (PR)

In the previous runoff configuration 3 additional former runoff measurements are incorporated. Thus, the models have 12 input coordinates as in the cumulative rainfall configuration above. Table 6.12 gives the results. It can be seen that the time lag of the prediction has decreased for the ANN and the MT, moreover, the performance measures increased. The ANN exhibits a satisfactory *E* of more than 91 %, the model tree 85 %. Worth mentioning is that the models now feature better *E* values even when effectively forecasting longer ahead than before.

Table 6.12. Performance of ANN and MT in previous runoff configuration for the effective forecast.

Architecture	Trained on	Time shift	Eff.	RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 12-3-1/PR	10	-6	4	0.0564	0.0079	0.0204	97.0	14.9	27.0	91.5
MT (18)/PR	10	-7	3	0.0718	0.0127	0.0343	95.4	31.1	35.8	85.0

So far all models were trained on the runoff 10 time steps ahead and all showed significant time lags, some configurations more than others. Below it is investigated how the time lag changes when the models are trained on different lead times. ANNs and MTs were trained on lead times between 10 and 1 time steps, again in previous runoff configuration. The results of the ANN models are given in Table 6.13, the figures of the MTs in Table 6.14. It can be seen that the ANN trained on the runoff 9 time steps ahead ends up with the same effective forecast of 4 time steps than the ANN showed above trained on 10 time steps lead time, however, all performance measures are better. The same is true for the MTs. The MT trained on 10 time steps lead time which was presented above effectively predicts 3 time steps ahead. When training this model on 7 to 9 time step lead time, the network also effectively predicts 3

time steps ahead, although with better performance figures. In the tables below the best effective forecasts in each case appear in bold print.

Table 6.13. Performance measures of ANNs when trained on different lead times.

Architecture	Trained on	Time shift	Eff.	RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
ANN 12-3-1/PR	1	0	1	0.0080	0.0001	0.0025	99.9	2.4	4.2	99.8
-	2	-1	1	0.0115	0.0003	0.0040	99.8	3.8	6.0	99.6
-	3	-1	2	0.0171	0.0007	0.0052	99.6	4.8	8.8	99.2
-	4	-2	2	0.0215	0.0010	0.0072	99.4	6.5	11.5	98.6
-	5	-3	2	0.0258	0.0015	0.0090	99.2	7.6	13.4	98.1
-	6	-3	3	0.0316	0.0024	0.0111	98.9	9.2	15.6	97.4
-	7	-4	3	0.0370	0.0034	0.0129	98.7	9.9	17.8	96.5
-	8	-5	3	0.0426	0.0045	0.0157	98.2	12.3	20.6	95.2
-	9	-5	4	0.0491	0.0061	0.0178	97.8	13.1	23.3	93.8
-	10	-6	4	0.0564	0.0079	0.0204	97.0	14.9	27.0	91.5

Table 6.14. Performance measures of MTs when trained on different lead times.

Architecture	Trained on	Time shift	Eff.	RMSE [m ³ /s]	MSE [(m ³ /s) ²]	MAE [m ³ /s]	CORR [%]	RAE [%]	RRSE [%]	E [%]
MT (11)/PR	2	-1	1	0.0136	0.0004	0.0067	99.8	7.4	8.0	99.3
MT (17)/PR	3	-2	1	0.0223	0.0011	0.0098	99.5	10.0	12.3	98.4
MT (32)/PR	4	-3	1	0.0358	0.0028	0.0136	98.5	12.2	19.0	95.7
MT (21)/PR	5	-3	2	0.0389	0.0030	0.0146	97.7	13.5	22.2	93.9
MT (17)/PR	6	-4	2	0.0421	0.0040	0.0198	97.9	21.5	23.5	92.8
MT (16)/PR	7	-4	3	0.0501	0.0060	0.0258	97.7	25.3	26.7	90.9
MT (31)/PR	8	-5	3	0.0569	0.0078	0.0283	97.3	26.7	29.0	90.1
MT (16)/PR	9	-6	3	0.0649	0.0104	0.0320	96.1	27.8	32.1	87.9
MT (18)/PR	10	-7	3	0.0718	0.0127	0.0343	95.4	31.1	35.8	85.0

The effective lead time of the forecast increases with the lead time the models were trained on. This phenomenon can be observed for data-driven models trained on the runoff 1 to 10 times steps ahead and can also be expected for higher lead times. ANNs generally exhibit a higher ratio between effective and trained-on time lags than MTs. The ANN trained on 1 time step lead time is the only model where overall no time shift between predicted and observed runoff could be observed. The MT trained on 1 time step lead time does not appear in Table 6.14 because the effective lead time of the prediction is zero and hence it features no outlook.

Below in Figure 6.11 the forecasts of the ANN and the MT for effectively predicting 3 time steps ahead are shown. The comparison between predicted and observed runoff is given for each subset separately. To the right of each subset its highest observed runoff peak is shown in higher temporal resolution (except for Subset 1 where no pronounced runoff peaks can be observed). On the whole dataset both approaches reach an efficiency coefficient E of more than 90 %. The ANN exhibits an E of >90 % in all five subsets. The MT exhibits E values <90 % for the Subsets 4 and 5. Generally the ANN performs better than the MT, except in Subset 2 where the model tree exhibits the better efficiency coefficient. A visual comparison

of the highest peak in this subset shows that the ANN prediction has more variations and thus substantiates the performance measures. Subset 4 exhibits the highest runoff peak of the entire test dataset. This peak is underestimated by all data-driven models. It can also be seen that the interquartile range (IQR) of the ANN ensemble is very small, indicating that the variations of the single ANNs trained with randomly chosen initial parameters are marginal. The poorest performance exhibits the MT in Subset 5 with an E just under 75 %. The poor efficiency originates from a continuing underestimation of the base runoff rather than a divergence in predicting the runoff peaks.

6.8 SUMMARY AND DISCUSSION

Previous runoff measurements are frequently used in runoff models to indirectly represent the hydrological state of a catchment but normally at the cost of a timing error. However, the problem of lagged predictions in ANN models is often not investigated. So far only a few papers report approaches to eliminate or reduce timing errors (e.g. De Vos and Rientjes, 2005; Abrahart *et al.*, 2007). Unfortunately the elimination of the timing error was not possible without an increase of the RMSE. It seems to be a trade-off between timing errors and statistical measures. The above mentioned papers used rain gauge data only but there are indications that radar data might be helpful to correct the timing problem (cf. Abrahart *et al.*, 2007). The results of the present work that utilised radar data show that the choice of the input parameters has indeed an influence on the timing error, but radar data alone are not the ultimate solution to this problem.

The various input configurations of the models lead to different timing errors and statistical measures. When calculating the performance measures of composite datasets, weighted averages have a higher information value. Concerning the efficiency coefficient E , ANNs and MTs performed best with time shifted input data. This configuration was expanded with 3 additional input coordinates leading to the cumulative rainfall and previous runoff configuration, respectively. The cumulative rainfall configuration does not lead to a real improvement. The previous runoff configuration on the other hand features better statistical measures and a reduced timing error. The reason of this instance is believed to lie in the way the runoff values are mapped by the model. De Vos and Rientjes (2005) described the problem that ANNs give most of the weight to the runoff input and thus undervalue the precipitation inputs. Therefore, it is discussed below how the input parameter *actual runoff* is treated in the various configurations. When simulating the models with constant values except for the *actual runoff* the mapping of this parameter can be visualised. Figure 6.12 shows the output of the ANN ensemble-models when all inputs except *actual runoff* are kept constant at their median occurring value. It can be seen that the *actual runoff* is mapped similarly in standard and cumulative rainfall configuration. The mapping is in both cases quite linear. The additional cumulative rainfall inputs have not much influence on the mapping of the *actual runoff*. When three preceding runoff measurements are included, however, the response curve changes considerably. The additional preceding runoff parameters change the weight given to the *actual runoff* significantly. This is also the configuration that exhibits the minimum time lag.

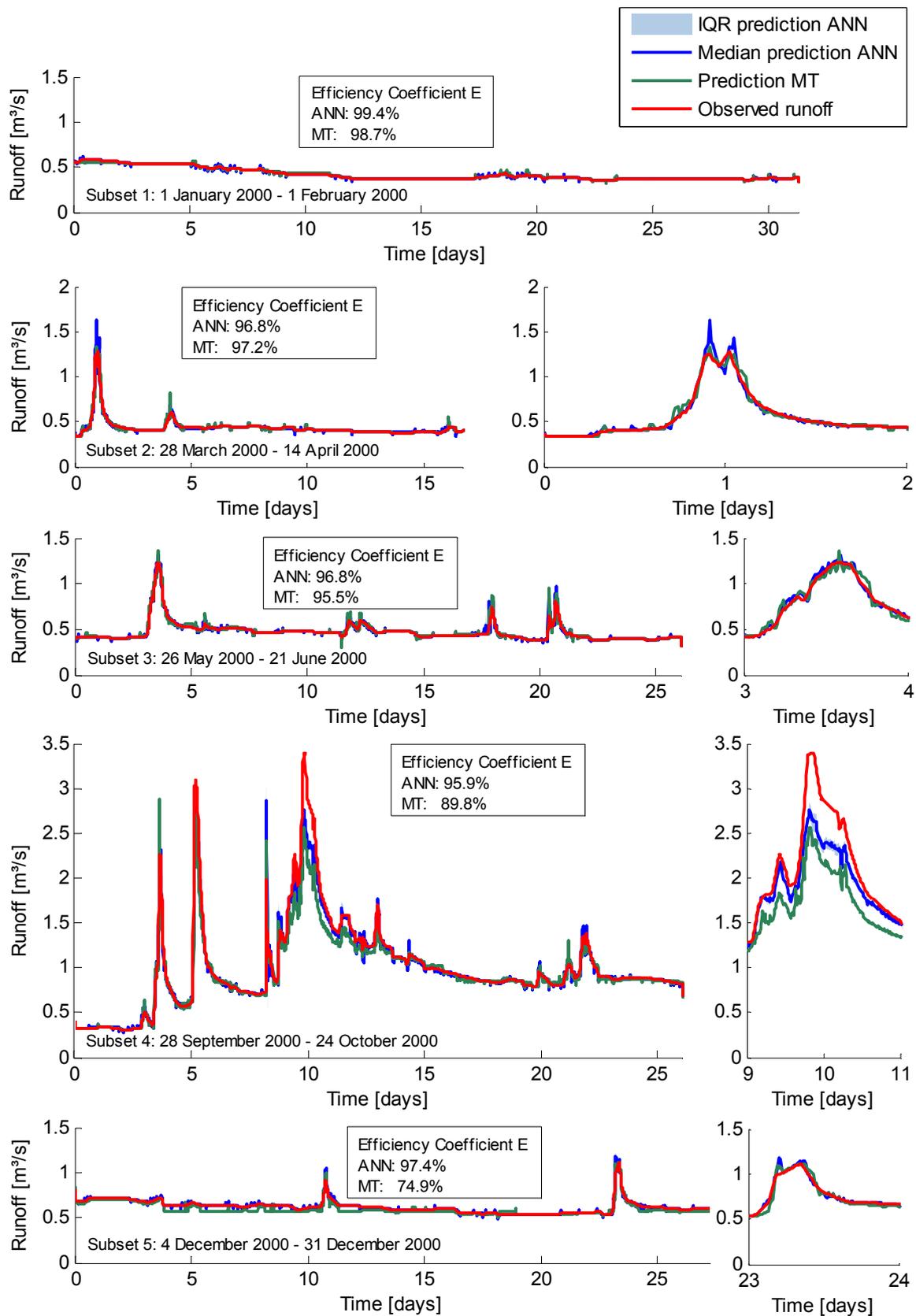


Figure 6.11. Forecasts 3 time steps ahead of ANN model and MT. The comparison between predicted and observed runoff is given for each subset separately. On the right hand side the highest observed runoff peak of each subset is given in higher temporal resolution.

The investigation of different lead times in model training revealed that the lead times the models are trained on are rarely the lead times the models effectively forecast. Improvements in the efficiency coefficient of up to 6 percentage points were observed when training the models on shorter lead times but with constant effective lead time. This phenomenon was observed on ANN models as well as MTs and it can help to get qualitatively better forecasts. The observation that the “prediction lag effect is especially significant in forecasts with small lead times“ in De Vos and Rientjes (2005, p.119) can not be supported by the present data. This can be due to the fact that here a 15-minute integration time was applied and not hourly data were used. Here the ANN model trained to predict the runoff 15 minutes ahead did not show a time lag. The results indicate that it is worthwhile to train data-driven models on different lead times and examine the effective lead time because the performance may increase. It also became apparent that the input configuration has a considerable influence on the outcome of the model. A certain way how the data are presented to the model can make it easier or contrariwise also more difficult for the data-driven approaches to extract the relevant information.

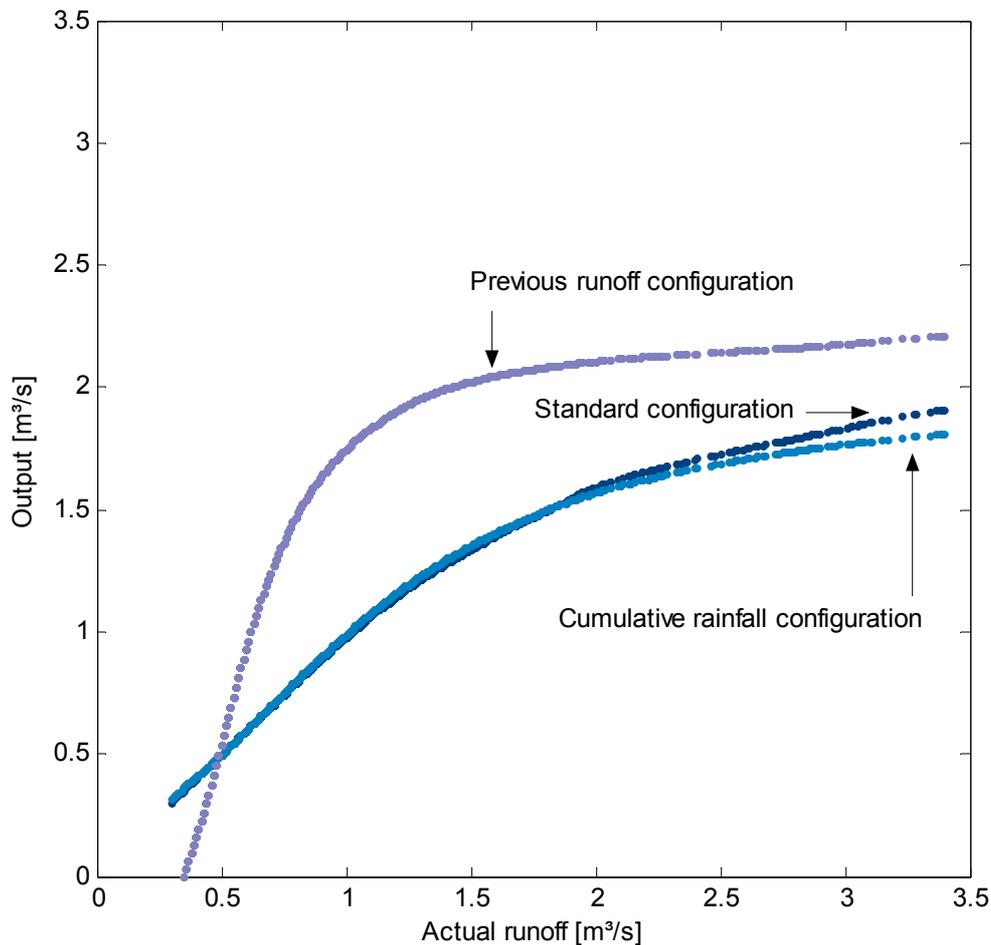


Figure 6.12. Visualisation of the mapping of the parameter *actual runoff* in ANN ensemble models for various configurations. The curves show the output of the ANN ensemble models when all inputs except *actual runoff* are kept constant at their median occurring value.

One objection to ANN models that is occasionally raised is that these models are almost intransparent with their many weights and complex nonlinear functions. The figure above has shown that there are methods to make the network behaviour more transparent but for many users of neural network tools the ANN is still a black box where inputs lead to outputs but it is not traceable how the relationship materialises. Model trees allow a little more insight how they treat the inputs but they also become unclear the more linear models are included. The presented MTs exhibit 9 to 32 linear models. The parameters do not allow an interpretation of the hydrological processes involved. Moreover, only a few clusters are included in the linear models and as a consequence the desired gapless spatial coverage is lost.

Concerning the statistical measures ANNs typically perform better than MTs with the same configuration. Worth mentioning is that in each case the median prediction of an ensemble of 20 ANNs is taken for the purpose of comparison. The variations among the ensemble are small though.

The hydrographs indicate that predicting high peaks with data-driven models is demanding. One reason is that high peaks occur rarely and that's why the models do not gain so much information about them in the training process. DDMs are known to perform poorly on data beyond the range they were trained on. Moreover, "during extreme events data usually contain a significantly higher level of observation noise" (Cherkassky *et al.*, 2006, p. 116). Some approaches, therefore, propose several local models instead of one global one (e.g. Solomatine and Siek, 2006). This can also be a future optimisation of the present ANN approach. When more data of high peaks are available, separate models for high and low flows can be advantageous.

The issue of poor performance beyond the range ANNs were trained on is not unique to hydrological applications. Also authors in other fields of sciences report about inadequate extrapolation capabilities of their ANNs and sometimes propose ways to overcome this problem. Landauer (2010) for instance defines a *strange range* besides a range amply covered by data. For the *strange range* another extrapolation mechanism is used for the ionospheric model. The investigation of such approaches from other fields of sciences with regard to technical feasibility might lead to unconventional new approaches in hydrology.

7 IMPROVING WEATHER RADAR ESTIMATES OF RAINFALL WITH DATA-DRIVEN MODELS

7.1 INTRODUCTION

Although the weather radar can more likely be considered a qualitative than a quantitative instrument, it has always been used to estimate rainfall amounts. The main advantage of weather radars when it comes to estimating rainfall is their high spatial and temporal resolution and the gapless spatial coverage. Today's weather radars make a full volume scan within about 5 minutes and thereby cover an area of tens of thousands of square kilometres. Perhaps the biggest advantage of weather radars is that they reveal a three-dimensional structure of precipitation (RHI-plot, see Figure 7.1). By displaying in what height what intensity of precipitation can be assumed, weather radars give an impression of the precipitation generation type: whether it is a cyclonic or a convective event.

Because of their advantages "radar measurements of precipitation have enjoyed wide-spread operational usage and will remain so in the future" (Bringi and Chandrasekar, 2001, p. 534). But other than rain gauges which measure precipitation directly on the ground, the radar determines the reflectivity aloft. Due to this principle, several sources of errors occur (see Section 3.7). Therefore, estimating the rain rate on the ground from radar reflectivity measurements is still a hot topic in radar meteorology and hydrology.

Bringi and Chandrasekar (2001) divide rainfall estimation techniques by radar into physically and statistical/engineering based approaches. Physically based approaches attempt to estimate rain rates by radar measurements together with an underlying rain model. The rain model generally describes the shape and material parameter of the rain drops, their size distribution and fall speed. Such an approach is used in nearly all operational weather radars. The rain rate is estimated from the measured reflectivity factor by non-linear relationships. Physically based approaches do not include any feedback e.g. from rain gauges as opposed to statistical/engineering based approaches where a feedback is used (cf. Bringi and Chandrasekar, 2001).

Rain gauges and weather radars have different advantages. Rain gauges measure the rain rate quite accurately but the measured rate is only relevant for a limited area. A few kilometres away the situation can be quite different, especially during convective rain events and in Alpine terrain. On the other hand, weather radars provide data with good spatial coverage and temporal resolution. They typically determine the rain rate for every square kilometre, one weakness, however, is their often poor metering precision limiting the applicability of the radar for quantitative purposes. An obvious goal is to adjust radar measurements to the measurements of a rain gauge, combining the main advantage of radar namely gapless spatial coverage with that of the rain gauge – more accurate measurements on ground level.

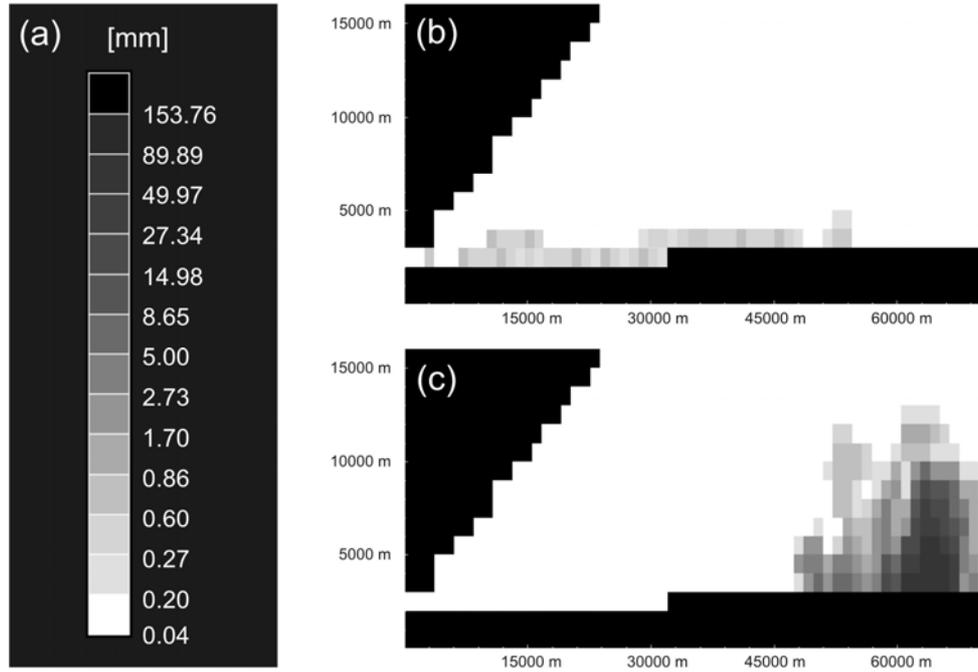


Figure 7.1. Three-dimensional structure of precipitation given by the weather radar. Levels of rain rate as distinguished by the Austrian weather radar network operated by Austro Control GmbH (a). Range-height indication for an advective event (b). Range-height indication for a convective event (c).

Here it is investigated if the vertical reflectivity profile as input of data-driven models can help to achieve better rainfall estimates for the ground level. For this purpose weather radar data were adjusted to rain gauge measurements using several data-driven approaches.

Operational weather radars generally apply a Z-R relationship as in Eq. (3.11) to relate the measured variable – the reflectivity Z – to the desired rain rate R . In the Austrian weather radar system the “standard Z-R relationship”

$$Z = 200 R^{1.6} \quad [\text{mm}^6/\text{m}^3], R \text{ in } [\text{mm}/\text{h}] \quad (7.1)$$

is implemented. Typically the measured reflectivity at the lowermost elevation is related to the rain rate. In an Alpine environment like in Austria, low level radar measurements are rarely available. In the study area the lowest elevation where radar data aloft the rain gauge are available is 3 km (MSL).

The objective of the present study is to find a relationship that is not only valid for the particular site on which the model was trained, but can also be applied to other comparable terrain where no rain gauge data are available. Henceforth only radar data should be necessary for the model to give a better estimate for the rain rate on the ground.

Data driven models are efficient methods to model complex input-output relationships. In this chapter several approaches are described to improve weather radar estimates of rainfall by using data-driven models. The models accept radar data as inputs and are trained to predict the rain rate as measured by the rain gauge.

7.2 DATA AVAILABLE

For this study, rain gauge and radar data from the province of Styria, Austria, were available. The data sets extend over a two year period (2001 and 2002). The two available rain gauges use the tipping bucket principle with a resolution of 0.1 mm. Their temporal resolution is 15 minutes. Reflectivity measurements are obtained from the Doppler weather radar station on Mt. Zirbitzkogel (see Section 3.8). The distance between the rain gauges and the weather radar is about 70 km, see Figure 7.2.

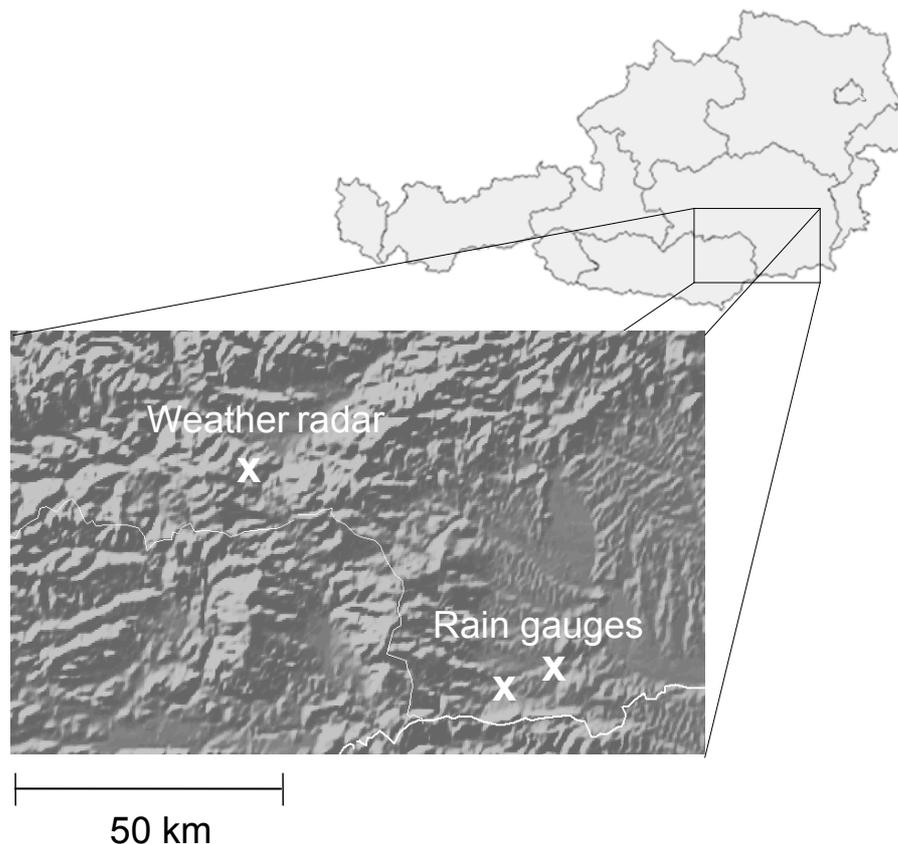


Figure 7.2. Study area in the province of Styria, Austria, showing radar and rain gauge locations.

7.3 PREPROCESSING

As described in Section 6.4, in data-driven modelling the examination of the data is important. Below the applied preprocessing steps are described.

7.3.1 DATA ANALYSIS

Precipitation is detected by the radar earlier than by the rain gauge. The weather radar on Mt. Zirbitzkogel can not measure the reflectivity below 3 km (MSL) over the study area. However, the two rain gauges there detect the precipitation later when the raindrops reach the ground. The radar and rain gauge time series were compared based on 5-minute time steps.

The highest correlation between these two time series was observed when the rain gauge series was shifted one time step against the time axis. This can be explained with the mean falling velocity of raindrops. Therefore, this time lag was considered in the models.

7.3.2 INPUT/OUTPUT PARAMETERS

As the model should predict the rain rate on the ground, the output vector consists of one variable namely the rain rate measured by a rain gauge on ground level. For the input vector the question is: Which radar measurements of the three-dimensional space aloft should be chosen as input vector? Liu *et al.* (2001) used the reflectivity at 1 km height (nine adjacent measurements with the rain gauge at the centre of the quadratic grid) as input vector for their neural network scheme for radar rainfall estimation. For the present study this approach is not possible because of the Alpine terrain and the location of the radar at more than 2300 m (MSL), low-level radar data are not available. Other authors (Li *et al.*, 2003; Xu and Chandrasekar, 2005; Teschl *et al.*, 2007) used superposed reflectivity measurements, the so-called vertical profile of reflectivity as input vector. The vertical profile of reflectivity also yields information about the nature of the precipitation event. Convective precipitation is characterised by high reaching clouds up to 10 km and more, whereas cyclonic rainfall events generally are characterised by lower level radar echoes. This information is considered important as the rain rate is also related to the type of rainfall.

In this study, first only the lowest measured radar reflectivity Z_3 (above 3 kilometres) was taken as input parameter and the Z-R relationship was recalibrated using regression analysis techniques. Feed-forward neural networks and model trees were also trained with this setting. Moreover, these DDMs were also trained with several superposed reflectivity measurements as input vector to examine what improvement the knowledge of the vertical structure of precipitation can bring when estimating ground rainfall. The applied weather radar measures the vertical profile in 1 km steps up to a height of 16 km (MSL). The lowermost visible measurement lies at 3 km (MSL; see Figure 7.3). To represent the vertical profile the reflectivities Z_3 , Z_4 , Z_5 , and Z_6 (from 3 to 7 km) have been included in the input vector. The highest level where precipitation was detected h [km] is an additional parameter, because there is a significant correlation between measured rainfall on the ground and the highest level with precipitation as measured by the radar. Table 7.1 lists the input and output vectors for all applied DDMs. The log transformation is used to linearise the pristine non-linear Z-R relationship.

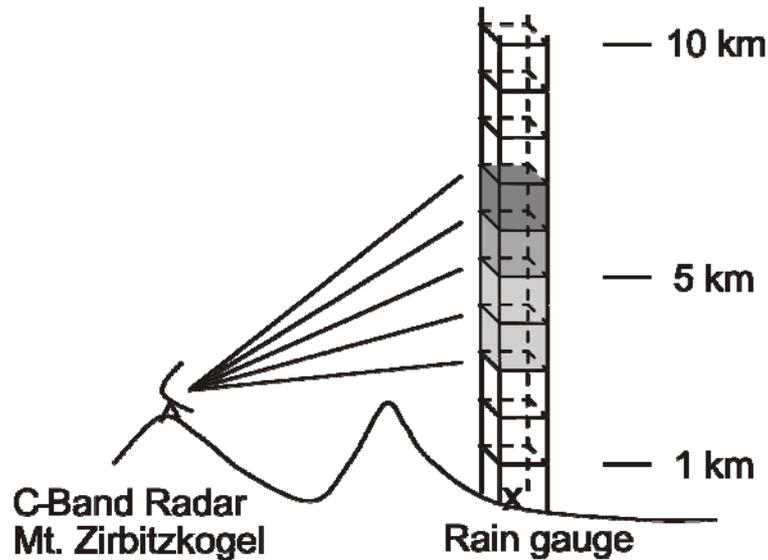


Figure 7.3. Visualisation of the inputs of the models. Radar measurements above the rain gauge form the input parameters of the data-driven models.

Table 7.1. Overview of the applied DDMs and their input parameters.

DDM	Inputs	Output
Non-linear regression	Z_3	R
Linear regression	$\log(Z_3)$	$\log(R)$
Feed-forward NN	Z_3	R
.	$\log(Z_3)$	$\log(R)$
.	Z_3, Z_4, Z_5, Z_6, h	R
.	$\log(Z_3), \log(Z_4), \log(Z_5), \log(Z_6), h$	$\log(R)$
MT	Z_3	R
.	$\log(Z_3)$	$\log(R)$
.	Z_3, Z_4, Z_5, Z_6, h	R
.	$\log(Z_3), \log(Z_4), \log(Z_5), \log(Z_6), h$	$\log(R)$
Radial basis NN	a_3, a_4, a_5, a_6, h	R
IBk	a_3, a_4, a_5, a_6, h	R

Z_x : Reflectivity ($\text{mm}^6 \text{m}^{-3}$) measured from $x \cdot 1000 \text{ m}$ to $(x+1) \cdot 1000 \text{ m}$ altitude

a_x : Radar rain rate ($\text{mm}/15 \text{ min}$) measured from $x \cdot 1000 \text{ m}$ to $(x+1) \cdot 1000 \text{ m}$ altitude

h : altitude of highest radar echo

7.3.3 DATA SEPARATION

For the development process of the data driven models, several sources of data are necessary: data for training and validating the model and data for testing the completed model. The dataset was divided into training, validation, and test datasets. However, the test datasets originate from another location 10 km away. The training dataset was used to determine the parameters of the model. The validation dataset is used to determine the best model configuration based on performance measures. Note that for regression models this validation step is not required thus in this case training and validation data were used to determine the model parameters. In order to examine if the relationship found between radar reflectivity and

rain gauge is representative for other sites, data from the second rain gauge and the associated radar data were used as test dataset. This dataset was used to determine the performance of the model. The partitioning between training and validation dataset from the first rain gauge site was done randomly. Each of these datasets comprises 449 pairs of radar and corresponding rain gauge measurements. The test dataset taken from the other rain gauge site contains 1310 pairs of measurements. The test data was formatted in the same way as the training and validation data. Again the rain gauge time series was shifted 5 minutes with respect to the radar measurements. The elevation of this rain gauge station is approximately the same as the first one.

7.4 DATA-DRIVEN APPROACHES

In this section the configuration of the various data-driven approaches for improving weather radar estimates of rainfall is described and the details for the setup of the models are given.

7.4.1 REGRESSION ANALYSIS

Regression analysis is an example of data-driven modelling and frequently used for data analysis and prediction. Here regression analysis is used to readjust the parameters of the Z-R relationship. For this purpose the weather radar and rain gauge data were investigated. Below the recalibration is described.

In general the standard Z-R relationship is used to transform the radar reflectivity measurements into rain rate. As stated in Chapter 3.3, the Z-R relationship is linked to the drop size distribution of the precipitation event. As the drop size distribution is far from being uniform, not during one single precipitation event and even less over various seasons and locations, the assumption of a fixed Z-R relationship is always only an approximation.

The regression analysis is applied to readjust the coefficients a and b of the power-law function in Eq. (3.11). For that purpose the data are plotted in two dimensions. The Z-data from the weather radar on the abscissa and the R-data originating from the rain gauge on the ordinate. The weather radar measurements are all but the predictor data and the rain gauge data the desired response. Therefore, the Z-R relationship is written in the form below:

$$R = \left(\frac{Z}{a} \right)^{\frac{1}{b}} \quad (7.2)$$

The scatter plot was approximated by this power-function. The method of least squares was used to fit Z and R values, i.e. to determine the coefficients a and b . For this purpose non-linear regression techniques are necessary. The calculation is much more complex than in linear regression because an iterative approach is required to determine the model parameters. First an adequate start value for each of the parameters is necessary. Subsequently the parameters are adjusted and it is determined whether or not the fit improves. In this process the fitting algorithm determines direction and magnitude of the adjustment. The Levenberg-Marquardt method which is used as training algorithm in ANN training can be used as fitting algorithm in non-linear regression. The fitting curve in Figure 7.4 was obtained using the non-linear least-squares fitting method. The applied software tool was the MATLAB *Curve Fitting Toolbox*. The residuals are

also presented. The residual of a specific data point is defined as the difference between the response value and the predicted response value by the fitted model.

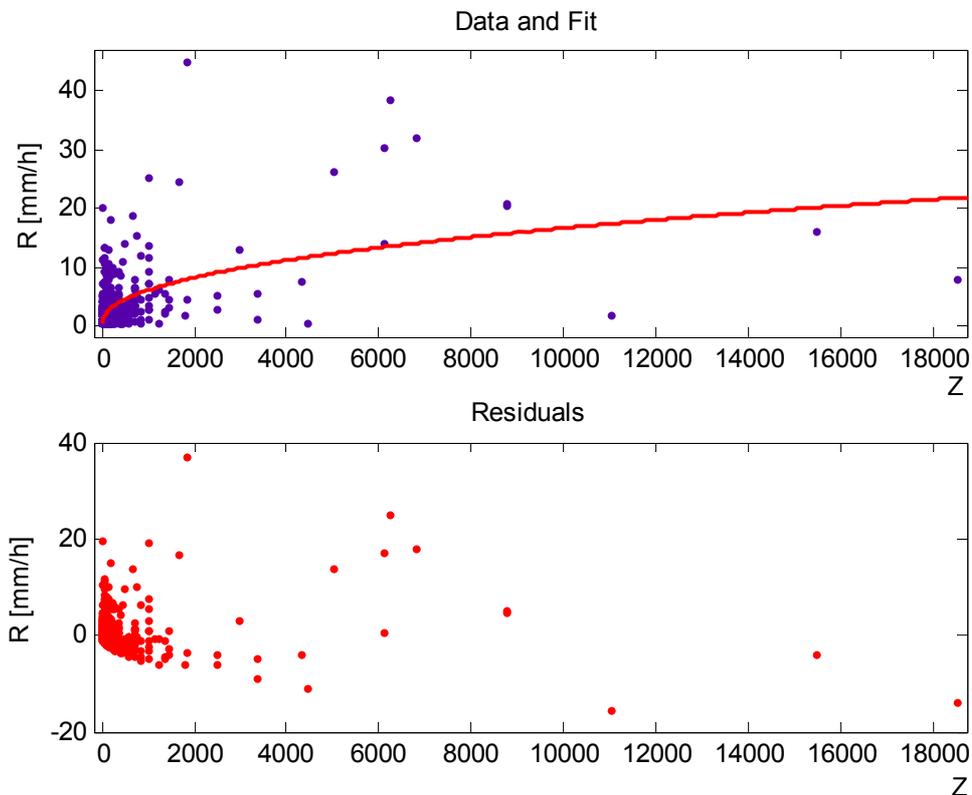


Figure 7.4. Fitting the Z-R data with non-linear regression. Below the fitting curve the residuals of the data points are given.

In the field of weather radar research the complex non-linear regression techniques are sometimes avoided by transforming the data prior to the regression. The logarithm lends itself for this transformation because the original Z-R relationship is a power function. When applying the log transformation on both sites, the Z-R relationship reduces to a linear equation in the bi-logarithmic plot.

$$\log Z = \log a + b \log R \quad (7.3)$$

In other words, the log transformation linearises the original non-linear model. This also simplifies the regression analysis. Now a linear regression can easily be carried out. Simple matrix techniques can be used to determine the parameters a and b . The fitting was carried out applying the linear least-squares fitting process. This process minimises the summed square of the residuals. Figure 7.5 shows the linear regression on the logarithmised data.

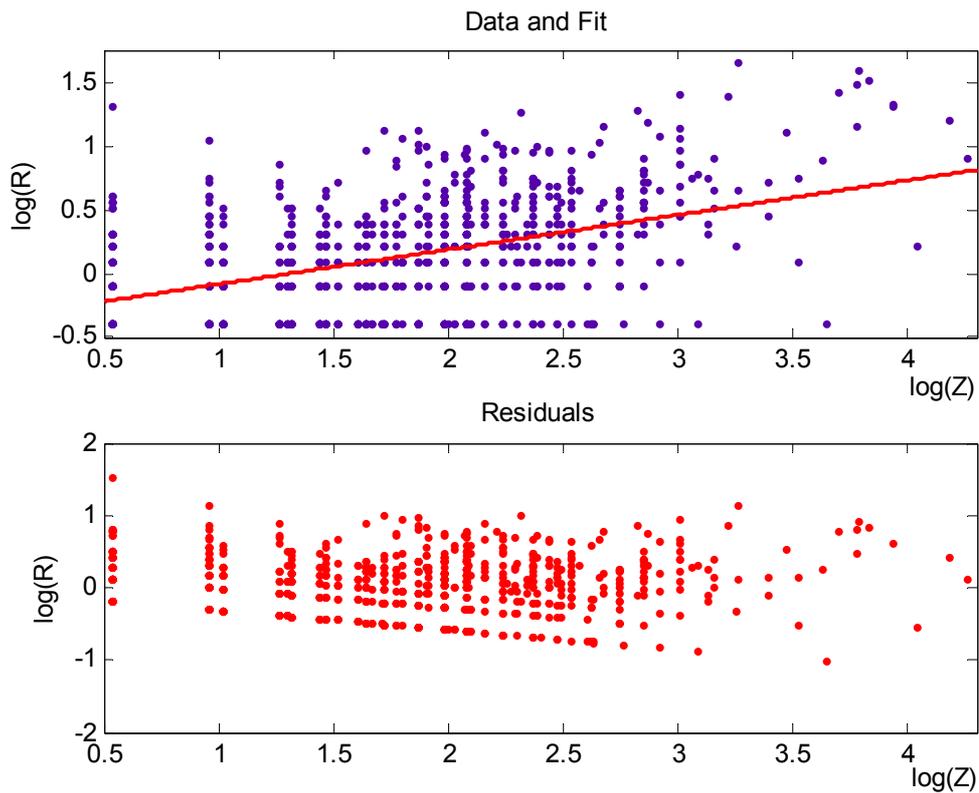


Figure 7.5. Fitting the $\log(Z)$ - $\log(R)$ data with linear regression. Below the fitting curve the residuals of the data points are given.

Besides the linearisation another effect of the log transformation implicates. The Z data that originally range over more than four orders of magnitude are thus evenly spread out.

The coefficients a and b , the two regression methods ended up with, are different. The reason is that different data are optimised. In case of the linear regression in the bi-logarithmic plot, the minimised sum of squares of the residuals (SSR) refers to logarithmic data. (Table 7.2).

Table 7.2. Determined coefficients of the Z-R relationship.

Method	Data	Coefficients		SSR
		a	b	
Nonlinear Regression	Z, R	15.80	2.29	8869.1
Linear Regression	$\log(Z), \log(R)$	20.09	3.69	132.2

7.4.2 NEURAL NETWORK APPROACHES

Feed-forward neural networks and radial basis function neural networks have been used for weather radar rainfall estimation. Both network types can solve the desired complex approximation task. The feed-forward neural network is one of the most common types of neural networks. Xiao and Chandrasekar (1997) developed a neural network of this type for radar rainfall estimation. According to Liu *et al.* (2001) a disadvantage of feed-forward neural networks is the computationally demanding training process. They have chosen a radial basis function neural network for their weather radar rainfall estimation task.

Here both types of neural networks – feed-forward and radial basis function – are applied to the available data (originally and logarithmised).

7.4.2.1 Feed-forward neural network approach

The network function of a feed-forward neural network (FNN) is largely determined by the number of neurons in the different layers and the weighted connections between them. Here a fully connected network is used where all neurons of two successive layers are connected to each other (see Chapter 5). An area of conflict is that a small network may have insufficient degrees of freedom (weights and biases) to realistically represent the relationship between radar reflectivity and ground measured rainfall, and a large network with many weights to be adapted may memorise fluctuations in the training data and is consequently not able to generalise. The generalisation capability of the network is very important for this particular task, because it is assumed that weather radar measurements contain noise that should not be memorised by the network.

Thus the method used to determine the architecture of the FNN was to start with a small network with one neuron in the hidden layer and to train it 20 times with randomly chosen initial parameters. Thus here an ensemble of 20 network training trials is used. As in Chapter 6 this is to indicate the average performance of the network. When the mean squared error on the validation dataset increased the training was stopped and the minimum of the validation error was taken as indicator for the performance. Thereafter the number of hidden nodes is increased and the network architecture yielding the best performance is chosen. The architectures thus determined for the various input/output data are given in Table 7.3. Besides the original Z - R data the network is also trained with logarithmised data and with additional input parameters. Sigmoid transfer functions were applied for all layers except the output layer where an unbounded linear function was used.

Table 7.3. Applied FNN architecture and input/output configurations

Input	Output	Architecture
Z_3	R	1-6-1
$\log(Z_3)$	$\log(R)$	1-2-1
Z_3, Z_4, Z_5, Z_6, h	R	5-6-1
$\log(Z_3), \log(Z_4), \log(Z_5), \log(Z_6), h$	$\log(R)$	5-8-1

7.4.2.2 Radial basis neural network approach

Radial basis function (RBF) neural networks have also been utilised for weather radar rainfall estimation. This neural network type has advantages concerning setup time. In the standard form the RBF network has as many radial basis neurons as training vectors (here 449). But not only the number of vectors is identical, also the weights of the RBF neurons are identical to the training vectors. The RBF which calculates the Euclidean distance between input vector and weight vector (see Section 5.1.2) will thus produce 1 whenever a new input vector is identical to a training vector. The advantage of this architecture is that it is not necessary to find the proper number of neurons by trial and error.

The parameter that has to be adapted is the *spread* of the radial basis function. This parameter determines “the width of an area in the input space to which each neuron responds” (Demuth and Beale, 1998, p. 6-6). If the *spread* constant is too small the network will only react to the test vectors close to the training vectors. Test vectors which exhibit too large an Euclidean distance from the training vectors are not detected. On the other hand if the *spread* is too large all neurons respond to the whole input space.

It is advisable to scale the input parameters such that they have similar ranges since the Euclidean distance is crucial. The applied data transformation method normalises the mean and standard deviation of the training dataset. It is a scaling of the network inputs. And this scaling has also to be executed for new data since the network was trained to perform on transformed data.

The proper *spread* of the radial basis function was determined based on the validation data. The MSE on the validation dataset was calculated for different spread values. It appeared that the RBF network with input data used so far did not yield satisfactory results. Even the applied data transformation was of no avail. Instead, the *Z*-values were transformed to *R*-values prior to the training of the network. For that purpose the standard *Z*-*R* relationship was applied.

The proper spread value of the network with scaled inputs was determined based on the validation data. The MSE on the validation dataset was calculated for different spread values. A minimum was determined at a spread value of 3.1. Table 7.4 shows the configuration of the RBF network.

Table 7.4. Applied RBF architecture. The variable a_x refers to the radar measurement Z_x transformed with the standard *Z*-*R* relationship into rain rate, h is the highest level where precipitation was detected.

Input	Output	RB neurons	<i>spread</i>
a_3, a_4, a_5, a_6, h	R	449	3.1

7.4.3 M5 MODEL TREE APPROACH

The adaptation of model trees to certain data is equivalent to the optimisation of the linear equations. The number of the linear equations can be controlled in the applied *Weka* software indirectly by the *minNumInstances* parameter. This parameter determines the minimum instances one node is considered for splitting into two. The MSE on the validation dataset is the criterion to decide which of the model trees is chosen for testing with the independent test dataset. Besides the original *Z-R* data the model trees were also trained with logarithmised data and with additional input parameters (Table 7.5).

Table 7.5. Applied MTs and their input/output configurations.

Input	Output	# equations
Z_3	R	11
$\log(Z_3)$	$\log(R)$	2
Z_3, Z_4, Z_5, Z_6, h	R	27
$\log(Z_3), \log(Z_4), \log(Z_5), \log(Z_6), h$	$\log(R)$	2

7.4.4 INSTANCE-BASED LEARNER

The instance-based learners have a similarity to the radial basis network. In instance-based learners the training data are stored and again the Euclidean distance is used to determine the training vectors that are closest to a test vector. These kinds of data-driven models are also known as memory-based or “lazy learners” because they basically store training examples to make predictions.

The IBk , which is used here, is a k -nearest-neighbour classifier (cf. Witten and Frank, 2005). The number of nearest neighbours k can be specified. This is the equivalent to the *spread* constant in RBF neural networks. The difference though is that the number of neighbours contributing to the result is fixed. The tuning of the IBk learner is done via the k value. The k parameter was determined based on the validation dataset.

Another parameter of tuning is to adapt the influence of the neighbours according to their distance to the test vector. When no weighting is applied, all k neighbours contribute equally to the result. Options are to obtain the weight from the inverse distance ($1/\text{dist}$), or from the complement of the distance ($1-\text{dist}$) (cf. Gómez *et al.*, 2007). All of these alternatives have been applied. The best results obtained the configuration $1/\text{dist}$ (Table 7.6). Again the Z -values were transformed to R -values prior to the training.

Table 7.6. Applied IBk learner configuration. The IBk uses 15 inverse-distance-weighted nearest neighbours for classification.

Input	Output	Weighting	k
a_3, a_4, a_5, a_6, h	R	$1/\text{dist}$	15

7.5 PERFORMANCE EVALUATION

The performance of the models was evaluated by comparing their outputs with the rain rates measured by the rain gauge. In case the model was trained to predict the logarithmised value $\log(R)$ the inverse transformation was applied prior to the evaluation. The statistical measures used are defined below.

$$\text{CORR} = \frac{\sum_{i=1}^n (r_i - \bar{r}) \cdot (a_i - \bar{a})}{\sqrt{\sum_{i=1}^n (r_i - \bar{r})^2} \cdot \sqrt{\sum_{i=1}^n (a_i - \bar{a})^2}} \quad (7.4)$$

$$\text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_i - a_i)^2} \quad (7.5)$$

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (r_i - a_i)^2 \quad (7.6)$$

Note: The MSE is used in many data-driven approaches as performance measurement on the validation data. The RMSE is used to give the measure the same dimensions as the predicted values.

$$\text{MAE} = \frac{1}{n} \sum_{i=1}^n |r_i - a_i| \quad (7.7)$$

$$\text{RAE} = \frac{\sum_{i=1}^n |r_i - a_i|}{\sum_{i=1}^n |r_i - \bar{r}_i|} \quad (7.8)$$

In Equations (7.4) to (7.8) n represents the number of pairs of variates, r the rain rate from the rain gauge, and a the estimated rainfall from the radar measurements. The values r and a are given in mm per 15 minutes. These measures were calculated for the rain gauge site that was not involved in the training and validation process.

7.6 RESULTS

The data-driven models as defined in Section 7.4 and trained on the first rain gauge site are now simulated with the new input data from the second site 10 km away from the location they were trained on. Their performance is given in Table 7.7.

The weather radar data at this location show a poor performance with respect to the reference rain gauge data. The study reveals what improvements are possible with the various DDMs and input configurations.

When applying the original Z -data of the lowest elevation as input, the FNN leads to the best performance measures apart from the correlation coefficient which is better in the non-linear regression model. The model tree, however, did not yield a better performance than the standard relationship except for MAE and RAE.

With logarithmised inputs the situation is reverse. Here the model tree is superior to the FNN in all aspects. The regression model also performs better than before. All approaches exhibit lower absolute errors than with original data. The RAE is up to 11 percentage points lower.

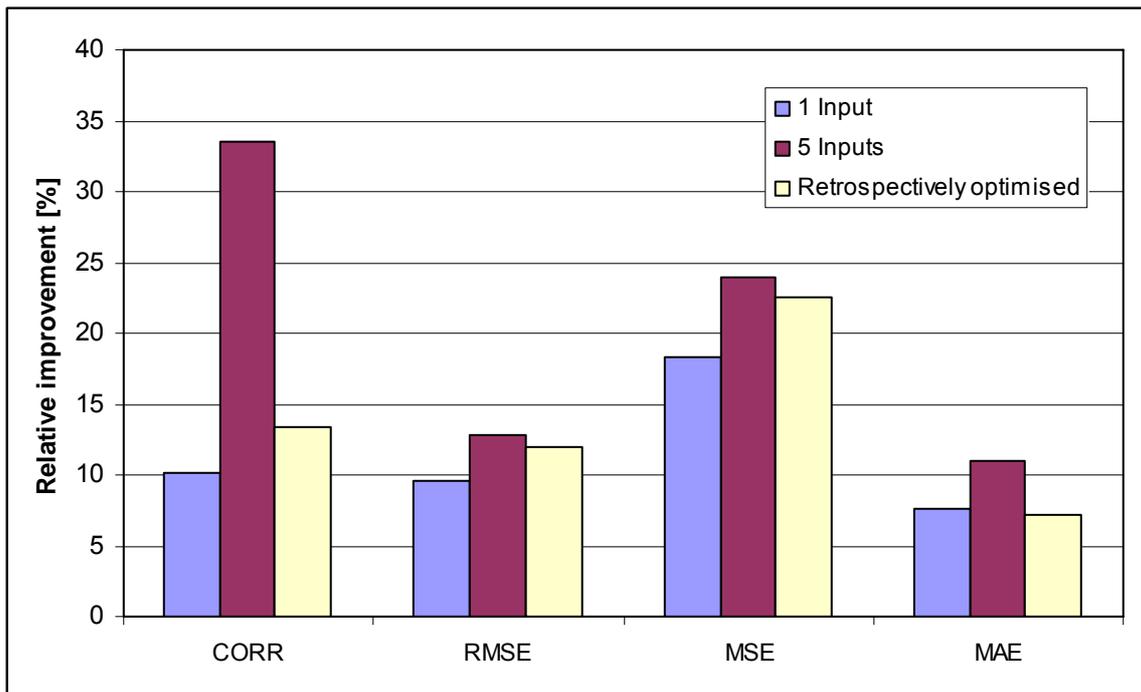
Both feed-forward neural network and model tree were also trained with additional inputs to account for the vertical profile of precipitation. It appears that the performance of the FNN is typically better than with one input parameter only. Figure 7.6 shows the relative improvement of the performance of the models with 1 and 5 inputs based on the standard Z - R relationship. The measures are often better than the retrospectively optimised Z - R relationship for this very site.

However, the MT still performs poorer than the standard relationship with original input data. With logarithmised data the additional inputs are not reflected in the model equations. The model tree builds the same model as before with one input only and that's why the performance measures remain the same.

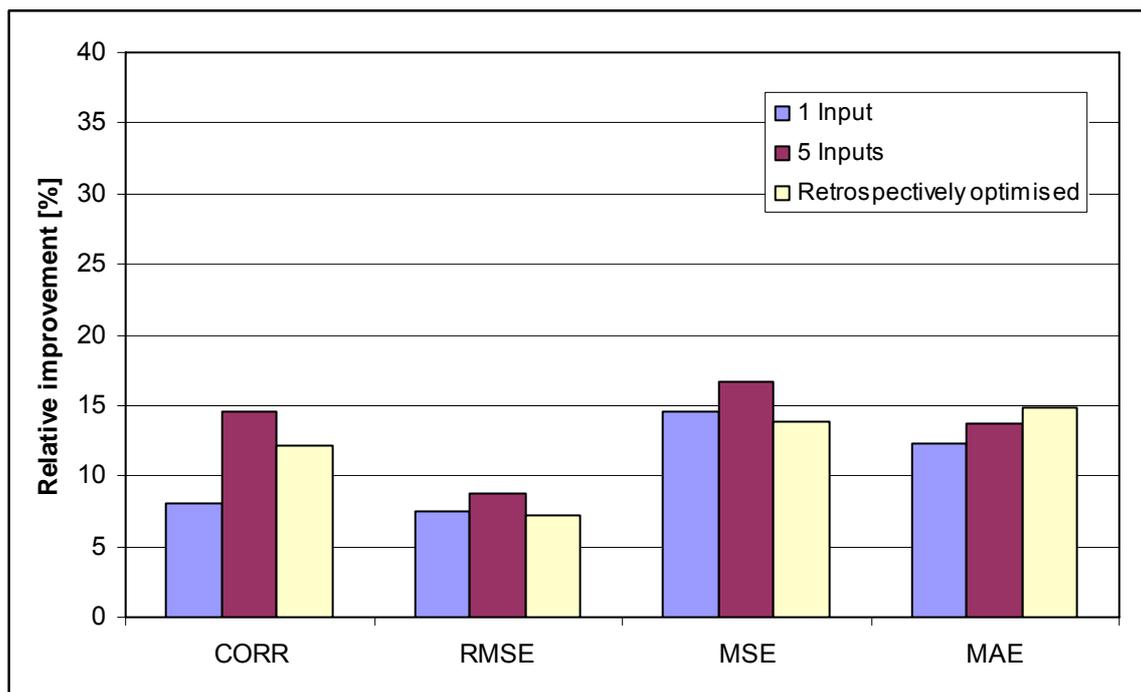
The memory-based approaches RBF and IBk were also trained with a five-dimensional input vector, the original Z input data were converted into rain rate though. When applied to the new data both models exhibit a correlation $>34\%$. The RBN shows lower absolute errors, the IBk on the other hand, lower squared errors.

Table 7.7. Performance measures of the various data-driven approaches compared to standard Z-R outcome and location optimized Z-R relationship. The best performance measures of each block appear bold.

DD Model	# inputs	CORR [%]	RMSE [mm/15 min]	MSE [(mm/15 min) ²]	MAE [mm/15 min]	RAE [%]
Weather radar data Standard Z-R relationship ($a = 200, b = 1.6$)	-	28.76	1.0909	1.1901	0.5391	100.30
Optimised retrospectively based on						
Original data ($a = 2.43, b = 3.78$)		32.62	0.9604	0.9223	0.5004	93.11
Logarithmised data ($a = 7.91, b = 4.93$)		32.26	1.0127	1.0256	0.4588	85.37
Original data						
Regression	1	31.93	1.0249	1.0505	0.5121	95.29
FNN	1	31.66	0.9858	0.9719	0.4984	92.73
MT	1	26.88	1.2117	1.4681	0.5260	97.87
Logarithmised data						
Regression	1	32.65	1.0103	1.0208	0.4611	85.78
FNN	1	31.09	1.0085	1.0172	0.4729	87.98
MT	1	31.89	1.0004	1.0009	0.4656	86.65
Original data						
FNN	5	38.41	0.9516	0.9055	0.4799	89.28
MT	5	28.66	1.1124	1.2373	0.5175	96.29
Logarithmised data						
FNN	5	32.93	0.9960	0.9920	0.4655	86.62
MT	5	31.89	1.0004	1.0009	0.4656	86.65
Original radar rainrate data						
RBN	5	34.22	0.9762	0.9529	0.4788	89.08
IBk	5	35.12	0.9658	0.9328	0.4918	91.51



(a)



(b)

Figure 7.6. Relative improvement of the performance of the FNN models. Models with 1 and 5 inputs are compared to the standard Z-R relationship for original data (a) and logarithmised data (b). For the purpose of comparison the measures of the retrospectively optimised Z-R relationship are given.

7.7 SUMMARY AND DISCUSSION

By means of data-driven models the radar reflectivity above a rain gauge was mapped to the measurements on the ground. This is challenging because of errors and limitations of the radar data. The errors of the training data set can theoretically be driven to a very low value with sufficiently capacious models. What is important, however, is the performance on new data. Here the test dataset comes from another rain gauge site. This was thought to be important in order to test the ability of the models to generalise.

First it was investigated what improvements are possible when readjusting the parameters a and b of the Z-R relationship. The determined parameters often do not fall into the ranges found in the literature (Battan, 1973). The sized a value is often less than the 16.6 while b frequently exceeds 2.87. This is ascribed to the special situation of the study area: the lack of low-level data and the related sources of radar errors. Moreover, the differences between the a values determined at the two rain gauge sites are noteworthy. This shows that the transfer of the determined relationship to the other site is challenging.

The results show that the style the data are presented to the models (in original or logarithmised form) influences the outcome as well as the choice of the model. Striking is the fact that the model tree shows acceptable performance on the logarithmised data but yields often poorer performance values than the standard Z-R relationship when trained with original data. The MT which produces piecewise linear models has problems to yield stable results based on the non-linear Z-R relationship. That does not mean though that the MT is not able to fit the original data appropriately. With 11 and 27 linear equations (for the 1 and 5 input version, respectively) the MT is able to approximate the non-linear relationship of the training data properly, but the model is not able to generalise to new data. The linear relationship of the logarithmised data, however, is fitted with 2 equations and produces acceptable results. Worth mentioning is the fact that the MT weights the lowest elevation in particular. Even when another three logarithmised Z inputs and the height of the highest radar echo are presented, the model tree only applies the lowest reflectivity. Thus here more inputs do not lead to a better performance.

Contrariwise, FNNs do recognise the additional input parameters. They exhibit better performance measures with the vertical reflectivity profile as input vector. This shows that additional measurements are helpful even if their distance to the ground is several kilometres. Nevertheless it can not be recommended to include all available measurements. Li *et al.* (2003) use equispaced reflectivity inputs from 1 to 4 km above the gauge and report that a resolution better than 1 km “did not significantly improve the estimation accuracy while it added to the computational load” (p. 2349).

In contrast to the MTs the data are based upon an ensemble of models. Twenty FNNs are applied for each configuration. With an ensemble the proper size of the network could be optimised in the validation process on a well-founded basis. The findings are more reliable than using one single training trial only. The networks thus determined exhibit one hidden layer only. The performance of bigger networks decreased on average. In earlier publications on neural network based rainfall improvement two hidden layers were used (Xiao and Chandrasekar, 1997; Teschl *et al.*, 2007). In these papers the network size was determined based on several trial and error experiments.

The RBF neural network turned out to be very sensitive concerning the input data. Scaling the data is advisable even though it was sometimes not enough. Finally the transformation of the Z -values into rain rate brought performance values better than the standard relationship. Here a network with as many neurons as training vectors was used and the model features low absolute errors and a high correlation compared to the other approaches. To ensure the comparability the RBF neural network was trained with the same datasets as the other models but this neural network approach is also suitable for adaptive training. Liu *et al* (2001) and Xu and Chandrasekar (2005) presented RBF networks that can be updated when new data become available. In this process it is not necessary to retrain the whole neural network merely some RFB neurons are added. It is not required that there are as much neurons as training samples, neurons can also be removed. Therefore, RBF networks proved to be a strong alternative to FNNs. Moreover, the training process of a RBF neural network is often shorter compared to a FNN.

The instance-based learner surprised with a good performance. This type of memory-based model is also not very time-consuming to set up. An input scaling is also recommended when using this type of models.

The validation process proved to be essential for the ability of the model to generalise. In this phase the best configuration of the data-driven model is chosen. It also prevents FNNs from overfitting. The results have shown that the so validated models succeeded in predicting new data. Thus they generalised from the training data. The consideration of the vertical profile of precipitation in FNN models leads to better results even in complex terrain and thus can be recommended.

8 CONCLUSION AND OUTLOOK

Meanwhile data-driven models are quite common in hydrology but the applications in radar-hydrology are still rare. But particularly in this field of science data-driven approaches like neural networks, model trees, and instance-based learners are promising. On the one hand lots of data are stored in the archives, which is essential to train these approaches, and on the other hand physically based models are often not applicable or do not produce satisfactory results because of a limited knowledge of the processes and parameters involved.

The use of meteorological radar measurements for estimating the current precipitation and for deriving estimates of runoff is the principal duty of radar hydrology. Both tasks are addressed in this thesis. When using radar data in rainfall-runoff models in general much more data are involved than using rain gauges only. This is challenging in the training process because it increases the parameters in the input model. Too many input parameters complicate the training process. Here measurements that showed the same time lag were combined to clusters and this reduced the number of input parameters.

Neural networks and model trees were used to model the relationship between rainfall (measured by rain gauge and weather radar) and runoff at the watershed outlet. Datasets from different seasons were used for training, validating, and testing. It turned out that calculating several common performance measures for independent subsets en bloc can lead to anomalous high values which do not reflect the performance on the subsets. The average values, weighted in respect of the size of the subset, turned out to be a more meaningful measure.

Simple data manipulation methods like time shifting and calculating moving rainfall averages were applied to increase the correlation between target runoff and precipitation measurements. A higher correlation between input and target data does not assure better model forecasts though. The consideration of the actual hydrological state is important for estimating the rainfall response. Several previous runoff measurements appeared to work best both with ANNs and MTs. It was shown that additional cumulative rainfall measurements in the ANN models did not significantly change the performance and the weight given to the *actual runoff*.

The training of the models on different lead times revealed an approach that improved the performance measures of both ANN and MT models. Training on shorter lead times did often not change the effective lead time but improved the statistical measures and thus the quality of the forecast.

The other task of radar hydrology, the estimating of the current precipitation was also investigated with data-driven models. First the different sampling characteristics were analysed. The different place of observation of rain gauge and weather radar measurements results in a time shift that was considered in the data-driven approaches. Feed-forward neural networks, radial basis networks, model trees, and instance-based learners were used to

estimate the rainfall and the results were compared with the outcome of both linear and non-linear regression analysis and the standard Z-R relationship. The model tree which approximates the data with linear relationships worked best on the by log-transformation linearised Z-R relationship. Feed-forward networks performed better on the original non-transformed data. Worth mentioning is the relative improvement of the results when training the FNN with additional inputs reflecting the vertical profile of reflectivity.

Feed-forward neural networks – the mainstream technology for data-driven modelling – performed best in both applied tasks of radar hydrology.

For the future promising improvements can be expected by the use of polarimetric radar parameters. The radar data in this study originate from a conventional weather radar that transmits and receives electromagnetic waves in horizontal polarization. Dual-polarization is becoming the standard for new weather radar systems. In contrast to conventional weather radars, dual-polarization radars transmit and receive in horizontal and vertical polarization. Thus, they determine polarimetric parameters like *differential reflectivity*, *linear depolarization ratio*, and *differential propagation phase*. By knowing these parameters additional information on size and shape of precipitation can be obtained. Also the accuracy of the rain rate estimation should be better compared to standard Z-R relationships of conventional radars.

Data-driven approaches can use the additional parameters to distinguish between forms of precipitation. Probabilistic neural networks seem suitable for this task. If polarimetric parameters and target data are available the network can determine which precipitation type is most likely. And this in turn shall lead to better rainfall predictions and better rainfall-runoff models.

REFERENCES

Abrahart, R. J., Heppenstall, A. J., and See, L. M. (2007): Timing error correction procedure applied to neural network rainfall-runoff modelling, *Hydrolog. Sci. J.*, 52(3), 414-431, doi:10.1623/hysj.52.3.414.

AMS (2000): *Glossary of Meteorology, Second Edition* [Online], American Meteorological Society. Available at: <http://amsglossary.allenpress.com/glossary>, [Accessed 25 February 2010].

ASCE (1996): *Hydrology Handbook (ASCE Manuals and Reports on Engineering Practice No. 28), Second edition*, American Society of Civil Engineers, ISBN: 0-7844-0138-1.

Atlas, D., editor (1990): *Radar in Meteorology*. American Meteorological Society, Boston, MA, ISBN: 0-933876-86-6.

Austin, G. and Seed, A. (2005): Special issue on the hydrological applications of weather radar - guest editors' preface, *Atmos. Sci. Lett.*, 6(1), 1, doi:10.1002/asl.103.

Battan, L. J. (1973): *Radar Observation of the Atmosphere*, Chicago University Press, Chicago, IL.

Baumgartner, A. and Liebscher, H-J. (1996): *Allgemeine Hydrologie - Quantitative Hydrologie, Lehrbuch der Hydrologie, Band 1*, Gebrüder Borntraeger, Berlin Stuttgart, ISBN: 3-443-30002-2.

Bhattacharya, B. and Solomatine, D. P. (2005a): Neural networks and M5 model trees in modelling water level–discharge relationship, *Neurocomputing*, 63, 381-396, doi:10.1016/j.neucom.2004.04.016.

Bhattacharya, B. and Solomatine, D. P. (2005b): Modelling harbour sedimentation using ANN and M5 model trees, *Proceedings of International Joint Conference on Neural Networks*, Montreal, Canada, July 31 –August 4, 2005, pp. 2643-2648, vol.4, doi:10.1109/IJCNN.2005.1556320.

Brandes, E. A. (2000): Dual-Polarization Radar Fundamentals and Algorithm Prospects, NEXRAD program report—operational support facility WSR-88D commerce-defence-transportation, [Online]. Available at: <http://www.roc.noaa.gov/app/sta/algorithm00.pdf> [Accessed 25 February 2010].

Bringi, V. N. and Chandrasekar, V. (2001): *Polarimetric Doppler weather radar – Principles and applications*, Cambridge, United Kingdom, Cambridge University Press.

- Burton, G. A. and Pitt, R. (2001): *Stormwater Effects Handbook: A Toolbox for Watershed Managers, Scientists, and Engineers*, CRC Press, ISBN: 0-87371-924-7.
- Campolo, M., Andreussi, P., and Soldati, A. (1999): River flood forecasting with a neural network model. *Water Resour. Res.*, 35(4), 1191-1197.
- Chang, M. (2006): *Forest Hydrology: An Introduction to Water and Forests*, Second Edition, CRC Press, ISBN: 0-8493-5332-7.
- Cherkassky, V., Hsieh, W., Krasnopolsky, V., Solomatine, D. P., and Valdes, J. (2007): Computational intelligence in earth and environmental sciences, *Neural Networks*, 20(4), 433, doi:10.1016/j.neunet.2007.05.001.
- Cherkassky, V., Krasnopolsky, V., Solomatine, D. P., and Valdes, J. (2006): Computational intelligence in earth sciences and environmental applications: Issues and challenges, *Neural Networks*, 19(2), 113-121, doi:10.1016/j.neunet.2006.01.001.
- Chiang, Y.-M., Hsu, K.-L., Chang, F.-J., Hong Y., and Sorooshian, S. (2007): Merging multiple precipitation sources for flash flood forecasting, *J. Hydrol.*, 340, 183-196, doi:10.1016/j.jhydrol.2007.04.007.
- Conway, A. J., Macpherson, K. P., and Brown, J. C. (1998): Delayed time series predictions with neural networks, *Neurocomputing* 18, 81-89.
- Coulibaly, P. and Baldwin, C. K. (2005): Nonstationary hydrological time series forecasting using nonlinear dynamic methods, *J. Hydrol.*, 307, 164-174, doi:10.1016/j.jhydrol.2004.10.008.
- Cressie, N. A. C. (1993): *Statistics for spatial data - revised edition*. Wiley series in probability and mathematical statistics: Applied probability and statistics section, New York, ISBN: 0-471-00255-0.
- Davie, T. (2003): *Fundamentals of Hydrology*, Routledge Fundamentals of Physical Geography, ISBN: 0-4152-2028-9.
- Dawson, C. W. and Wilby, R. (1998): An artificial neural network approach to rainfall-runoff modelling, *Hydrolog. Sci. J.*, 43(1), 47-66.
- De Vos, N. J. and Rientjes, T. H. M. (2005): Constraints of artificial neural networks for rainfall-runoff modelling: trade-offs in hydrological state representation and model evaluation, *Hydrol. Earth Syst. Sc.*, 9, 111-126.
- Deka, P. (2006): *Geography: Physical and Human. 2nd edition*, New Age Publishers, New Delhi, ISBN: 81-224-1912-7.
- Demuth, H. and Beale, M. (1998): *Neural network toolbox user's guide — For use with MATLAB, Version 3*, The MathWorks, Inc, Natick, MA.

-
- Doviak, R.S. and Zrníc, D. S. (1984): *Doppler radar and weather observations*, Academic Press, San Diego, CA. and London, ISBN: 0-12-221420-X.
- Dyck, S. and Peschke, G. (1995): *Grundlagen der Hydrologie. 3., stark bearb. Auflage*, Verlag für Bauwesen, Berlin, ISBN: 3-345-00586-7.
- Foresee, F. D. and Hagan, M. T. (1997): Gauss-Newton approximation to Bayesian learning, in *Proc. of the 1997 International Conference on Neural Networks*, 3, 1930-1935.
- Gaume, E. and Gosset, R. (2003): Over-parameterisation, a major obstacle to the use of artificial neural networks in hydrology ?, *Hydrol. Earth Syst. Sc.*, 7(5), 693-706.
- Gómez, O., Morales, E. F., and González, J. A. (2007): Weighted Instance-Based Learning Using Representative Intervals, *Lecture Notes in Computer Science. MICAI 2007: Advances in Artificial Intelligence*, pp. 420-430, doi:10.1007/978-3-540-76631-5.
- Gunn, K. L. S. and Marshall, J. S. (1958): The Distribution with Size of Aggregate Snowflakes. *J. Atmos. Sci.*, 15(5), 452-461.
- Hagan, M. T. and Menhaj, M. (1994): Training feed-forward networks with the Marquardt algorithm, *IEEE T. Neural Networ.*, 5(6), 989-993.
- Hanson, C. L., Zuzel, J. F., and Morris, R. P. (1983): Winter precipitation catch by heated tipping-bucket gages. *Transactions - American Society of Agricultural Engineers*, 26(5), 1479-1480.
- Hsu, K., Gupta, H. V., and Sorooshian, S. (1995): Artificial neural network modeling of the rainfall-runoff process, *Water Resour. Res.*, 31(10), 2517-2530.
- Imrie, C. E., Durucan, S., and Korre, A. (2000): River flow prediction using artificial neural networks: generalisation beyond the calibration range, *J. Hydrol.*, 233, 138-153.
- Janes, K. A. and Yaffe, M. B. (2006): Data-driven modelling of signal-transduction networks, *Nat. Rev. Mol. Cell Bio.*, 7, 820-828, doi:10.1038/nrm2041.
- Kainz, H. (2002): *Hydrologie und Wasserwirtschaft, Vorlesung Teil 1 Hydrologie*. Lecture notes, Institute of Urban Water Management and Landscape Water Engineering, Graz University of Technology.
- Landauer, G. (2010): *An extension of the ionospheric model IMAZ for large absorptions*, Project thesis, Institute of Communications Networks and Satellite Communications, Graz University of Technology.
- Li, W., Chandrasekar, V., and Xu, G. (2003): Investigations in radar rainfall estimation using neural networks, in *Proc. IGARSS '03*, pp. 2347-2349, doi:10.1109/IGARSS.2003.1294437.
- Liu, H., Chandrasekar, V., and Xu, G. (2001): An adaptive neural network scheme for rainfall estimation from WSR-88D observations, *J. Appl. Meteorol.*, 40(11), 2038-2050.

- Loukas, Y. L. (2000): Artificial neural networks in liquid chromatography: Efficient and improved quantitative structure-retention relationship models. *J. Chromatogr. A*, 904(2), 119-129, doi:10.1016/S0021-9673(00)00923-7.
- Marshall, J. S. and Palmer, W. McK. (1948): The distribution of raindrops with size, *J. Meteorol.*, 5, 165-166.
- Mason, J. C., Price, R. K., and Temme, A. (1996): A neural network model of rainfall-runoff using radial basis functions. *J. Hydraul. Res.*, 34(4), 537-548.
- Merz, R., Blöschl, G., and Piock – Ellena, U. (2001): *Flash-flood Risk Assessment under the impacts of land use changes and river Engineering*, WORKs EC contract ENV4-CT97-0529 Programme Environment and Climate 1994-1998 Topic 2.3.1 (Hydrological and Hydrogeological Risks), Final partner report 1998 - 2000, March 2001, Institut für Hydraulik, TU Wien.
- Meyer, U. (2004): *Die Kontroverse um Neuronale Netze – Zur sozialen Aushandlung der wissenschaftlichen Relevanz eines Forschungsansatzes*, Deutscher Universitäts-Verlag/GWV Fachverlage GmbH, Wiesbaden, ISBN: 3-8244-4555-7.
- Miller, G. A. (2009): *WordNet - About Us*, WordNet. Princeton University, [Online]. Available at: <http://wordnet.princeton.edu>, [Accessed 13 April 2010].
- Minsky, M. and Papert, S. (1969): *Perceptrons; an introduction to computational geometry*. MIT Press, Cambridge, Mass.
- Nash, J. E. and Sutcliffe, J. V. (1970): River flow forecasting through conceptual models I: A discussion of principles. *J. Hydrol.*, 10, 282-290.
- Pan, T.-Y. and Wang, R.-Y. (2004): State space neural networks for short term rainfall-runoff forecasting, *J. Hydrol.*, 297(1-4), 34-50.
- Pereira Filho, A. J. and Dos Santos, C. C. (2006): Modeling a densely urbanized watershed with an artificial neural network, weather radar and telemetric data, *J. Hydrol.*, 317, 31-48, doi:10.1016/j.jhydrol.2005.05.007.
- Quina, G. S. III (2003): *Statistical and hydrological evaluations of rain gauge and radar-derived precipitation for the Florida peninsula*. M.Sc. Thesis, Department of Meteorology, The Florida State University.
- Quinlan J. R. (1992): Learning with continuous classes, *Proceedings of the Fifth Australian Joint Conference on Artificial Intelligence*, Hobart, Tasmania, 16-18 November 1992, pp. 343-348.
- Rosenblatt, F. (1958): The perceptron: a probabilistic model for information storage and organization in the brain, *Psychol. Rev.*, 65, 386-408.

- See, L., Corne, S., Dougherty, M., and Openshaw, S. (1997): Some initial experiments with neural network models of flood forecasting on the River Ouse, *Proceedings of the 2nd International Conference on GeoComputation*, University of Otago, New Zealand, 26-29 August 1997, pp. 15-22.
- Senthil Kumar, A. R., Sudheer, K. P., Jain, S. K., and Agarwal P. K. (2005): Rainfall-runoff modelling using artificial neural networks: comparison of network types, *Hydrol. Process.*, *19*(6), 1277–1291.
- Sevruk, B. (1986): Correction of precipitation measurements, summary report. In: Sevruk, B. (Ed.), *Correction of Precipitation Measurements*. ETH/IASH/WMO Workshop on the Correction of Precipitation Measurements, Zürich, April 1–3, 1985. *Zürcher Geographische Schriften* 23, ETH, Geographisches Institut, Zürich, pp. 13–23.
- Sevruk, B. (1996): Adjustment of tipping-bucket precipitation gauge measurements. *Atmos. Res.*, *42*(1-4), 237-246.
- Shamseldin, A. Y. (1997): Application of a neural network technique to rainfall-runoff modeling. *J. Hydrol.*, *199*, 272-294, doi:10.1016/S0022-1694(96)03330-6.
- Skolnik, M. I. (1990): *Radar Handbook - 2nd edition*, McGraw-Hill, New York.
- Solomatine, D. P. (2002): Data-driven modelling: paradigm, methods, experiences. *Proc. 5th Int. Conference on Hydroinformatics*. Cardiff, UK, pp. 757-763.
- Solomatine, D. P. and Dulal, K. N. (2003): Model trees as an alternative to neural networks in rainfall–runoff modelling, *Hydrolog. Sci. J.*, *48*(3), 399-411.
- Solomatine, D. P. and Siek, M. B. (2006): Modular learning models in forecasting natural phenomena. *Neural Networks*, *19*(2), 215-224, doi:10.1016/j.neunet.2006.01.008.
- Solomatine, D. P. and Xue, Y. (2004): M5 Model Trees and Neural Networks: Application to Flood Forecasting in the Upper Reach of the Huai River in China, *J. Hydrol. Eng.*, *9*(6), 491-501, doi:10.1061/(ASCE)1084-0699(2004)9:6(491).
- Stravs, L. and Brilly, M. (2007): Development of a low-flow forecasting model using the M5 machine learning method. *Hydrolog. Sci. J.*, *52*(3), 466-477, ISSN: 0262-6667.
- Sudheer, K. P., Gosain, A. K., and Ramasastri, K. S. (2002): A data-driven algorithm for constructing artificial neural network rainfall-runoff models, *Hydrol. Process.*, *16*(6), 1325-1330.
- Teschl, R. and Randeu, W. L. (2006): A neural network model for short term river flow prediction, *Nat. Hazard. Earth Sys.*, *6*, 629-635.
- Teschl, R., Randeu, W. L., and Teschl, F. (2007): Improving weather radar estimates of rainfall using feed-forward neural networks. *Neural Networks*, *20*(4), 519-527, doi:10.1016/j.neunet.2007.05.005.

- Toth, E., Brath, A., and Montanari, A. (2000): Comparison of short-term rainfall prediction models for real-time flood forecasting, *J. Hydrol.*, 239, 132-147.
- Trigg, L. (1998): *An entropy gain measure of numeric prediction performance*, Working Paper 98/11, University of Waikato, Department of Computer Science, Hamilton, New Zealand.
- Uijlenhoet, R. (2001): Raindrop size distributions and radar reflectivity–rain rate relationships for radar hydrology. *Hydrol. Earth Syst. Sc.*, 5(4), 615-627.
- Ulbrich, C. W. (1983): Natural variations in the analytical form of the raindrop size distribution. *J. Clim. Appl. Meteorol.*, 22, 1764-1775.
- Villarini, G. and Krajewski, W. F. (2008): Empirically-based modeling of spatial sampling uncertainties associated with rainfall measurements by rain gauges. *Adv. Water Resour.*, 31(7), 1015-1023.
- Wang, Y. and Witten, I. H. (1997): Induction of model trees for predicting continuous classes. *Proc European Conference on Machine Learning Poster Papers*, pp. 128-137, Prague, Czech Republic.
- Witten, I. H. and Frank, E. (2000): *Data mining: practical machine learning tools and techniques with Java implementations*, Morgan Kaufmann, San Francisco.
- Witten, I. H. and Frank, E. (2005): *Data mining: practical machine learning tools and techniques*, 2nd edition, Morgan Kaufmann, San Francisco.
- Wölfelmaier, F. and Zwatz-Meise, V. (2005): Life cycle of convective cells with rapid scan satellite and radar data in the eastern Alpine region. *Proc. of World Weather Research Programme's symposium on nowcasting and very short range forecasting (WSN05)*, Toulouse, France, 5-9 September 2005.
- Xiao, R. and Chandrasekar, V. (1997): Development of a neural network based algorithm for rainfall estimation from radar observations, *IEEE T. Geosci. Remote*, 35, 160-171.
- Xie, L. (1988): *Untersuchungen der atmosphärischen Dämpfung und Depolarisation mittels eines doppelt linear polarisierten Wetter-Radars*, Doctoral Thesis, Graz University of Technology.
- Xu, G. and Chandrasekar, V. (2005): Operational feasibility of neural-network-based radar rainfall estimation, *IEEE Geosci. Remote S.*, 2(1), 13-17.
- Young, C. B., Bradley, A. A., Krajewski, W. F, Kruger, A., and Morrissey, M. L. (2000): Evaluating NEXRAD Multisensor Precipitation Estimates for operational Hydrologic Forecasting, *J. Hydrometeorol.*, 1, 241-254.
- Zamarreño, J. M. and Vega, P. (1998): State space neural network. Properties and application *Neural Networks*, 11(6), 1099-1112.