Czech University of Life Sciences Prague

Faculty of Environmental Sciences

Diploma thesis

The analysis of drought indices forecast using neural network

Bc. Anna Morozova

Supervisor: doc. Ing. Petr Máca, Ph.D.

Prague 2017

CZECH UNIVERSITY OF LIFE SCIENCES PRAGUE

Faculty of Environmental Sciences

DIPLOMA THESIS ASSIGNMENT

Bc. Anna Morozova

Water in the Landscape

Thesis title

The analysis of drought indices forecast using neural networks

---

Objectives of thesis

The aim of the thesis is to describe the forecast of selected drought indices using the neural network models.

Methodology

The thesis deals with:

1. The selection of basins and the estimation of drought indices
2. Building drought forecasting models based on neural networks
3. Analysis of drought forecast using the neural network models with different input variables

The proposed extent of the thesis

standard

Keywords

drought, drought index, neural network, forecast

Recommended information sources

Bishop, C. M. Neural Networks for Pattern Recognition. Oxford University Press, Inc., 1995. ISBN 0198538642.

Máca, P. – Pech, P. – Pavlasek, J. Comparing the Selected Transfer Functions and Local Optimization Methods for Neural Network Flood Runoff Forecast. Mathematical Problems in Engineering. 2014.

Mishra, A. K. – Singh, V. P. Drought modeling – A review. Journal of Hydrology. JUN 6 2011, 403, 1-2, s. 157–175.

WoS a SCOPUS

Expected date of thesis defence

2016/17 SS – FES

The Diploma Thesis Supervisor

doc. Ing. Petr Máca, Ph.D.

Supervising department

Department of Water Resources and Environmental Modeling

Electronic approval: 4. 4. 2017

doc. Ing. Martin Hanel, Ph.D.

Head of department

Electronic approval: 5. 4. 2017

prof. RNDr. Vladimír Bejček, CSc.

Dean

Prague on 06. 04. 2017

Abstract

Drought forecasting is a critical component of drought risk management. The presented thesis analyses the forecasting drought indices with Artificial Neural Network (ANN). The indices used are the Standardized Precipitation Index (SPI), the Standardized Precipitation Evaporation Index (SPEI) and the Standardized Soil Moisture Index (SSI). Tested neural network was multilayer perceptron with two hidden layers and was trained using backpropagation algorithm. I used the data obtained from 13 meteorological stations, located in different parts of Czech Republic. The records were from the period 1.1.1982 – 1.12.2015. For all three indices, a number of different models with the lead time of 1 to 12 months have been tested out. The best models have the R^2 values of 0,83 – 0,98, it is also show that forecasts of SSI were superior. The result of drought indices forecast, explained by the values of four model performance indices, such as Mean Error (ME), Mean Square Error (MSE), Root Mean Square Error (RMSE) and coefficient of determination (R^2). Artificial Neural Network model show excellent results for forecasting drought.

Keywords: drought, drought index, neural network, forecast

Author's declaration

I declare that this submitted thesis "The analysis of drought indices forecast using neural networks,, is my own work, all co-authors of the manuscripts are properly named and only sources listed in the Bibliography were used.

Prague, 15th April 2017 ……

Anna Morozova

Dedication and acknowledgements

I would like to thank Petr Maca for his motivation, patience and toleration. Without his help, I would not make it that far. He was the best supervisor, he supported me and helped in many ways. I am very grateful to my family for their faith, help and the biggest support during my study. I want to thank to my friends for being there for me through the entire my study.

Table of contents

1 Introduction

Drought is a global and local problem that induce other problems in various ways. It causes huge losses in agriculture and other negative influences on Earth. Drought can last for many years, that's why calls degradation of soils and desertification and other social problems, for example famine, impoverishment (Nicholson et al., 1990; Pickup 1998).

Determination may be very difficult when a drought begins or ends. A drought can be short, lasting just a few months, or persist for years before climatic condition return to normal.

Drought has affected 50 % of the 2.8 billion between 1967 and 1992 people who suffered from all natural disasters (Obasi G., 1994). Many important applications have been made to develop methodologies to quantify different aspects related to drought. Most of it to develop drought indices, which allow an earlier identification of droughts, their intensity and surface extent.

During the 20th century, few of drought indices were developed, based on different parameters and variables. It is a challenge to predict the future drought periods and their intensity and extremity. Various tool, methods and statistical models for drought forecasting have been suggested in different countries. Application the last one has a long history in forecasting of drought. A wide usage of new statistical technique known as *Artificial Neural Network* (ANN) show the superior performance for any time period. Generally, forecasting of precipitation and drought with ANN is based on past observed data of these variables (Asrari E., et al., 2014; Morid S. et al., 2007).

1.1 Main goals

Forecasting drought indices using the artificial neural network was analysed within this thesis due to its advantages. The ANN was successfully used in many real-life case studies.

Main goals of this diploma thesis are:

- provide a literature review about impacts of drought in the world, drought indices, their limitation and comparison
- describe the forecast of selected drought indices using the neural network models.
- apply models of ANN on case study using MLP

2 Literary Research

2.1 Drought definition

Drought is an insidious natural hazard that results from a deficiency of precipitation from expected or "normal" that, when extended over a season or longer period of time, is insufficient to meet the demands of human activities and the environment. Drought must be considered a relative, rather than absolute, condition (Wilhite D. et al., 2009).

Not only on the atmosphere is the reason of occurrence of drought. Also on the hydrologic processes which feed moisture to the atmosphere.

There are several reasons why drought differ from other natural hazards:

- First, the start and the end of a drought is difficult to determine. That's why, a drought is often referred to as a creeping phenomenon.
- Second, the drought does not have the universal definition because it is difficult to define.
- Third, impacts of drought are not structural and spread over large geographic areas than damages that can result from other natural hazards. In compare with hurricanes, earthquakes, tornadoes a drought affects water bodies of water resources structure and it seldom results in structural damage.
- Fourth, human activities can cause a drought. Other natural hazards, which are like drought in terms of their rank, are tropical cyclones, regional floods, volcanoes and earthquakes. (Mishra A. K. et al., 2010).

2.2 Classifications of drought

Droughts are commonly classified by type as meteorological, agricultural, and hydrological. Droughts differ from one another in three essential characteristics: intensity, duration, and spatial coverage (Wilhite D., 2009).

The flow chart in Tab. 1 shows as the propagation of drought and how it is dependent on meteorological factors like precipitation and temperature from one region to another (Tallaksen et al., 2004).

There are four categories of the drought (Wilhite D., et al., 1985; American Meteorological Society, 2004):

- Meteorological drought
- Hydrological drought
- Agricultural drought
- Socio-economic drought

Tab. 1 Flow chart, based on Stahl K., et al.,(2001); Peters E., et al., (2003):

Meteorological situation	*"rain" season low Precipitation (and high ET)*	*"snow" season low T or high T + low P*
	↓	
Meteorological drought	Precipitation deficiency	↓
	↓	
Soil moisture drought		Low soil moisture
	↓	↓
Hydrological drought	Low discharge ↔ Low ground-water storage	

Meteorological drought

The first one, meteorological drought, is defined as a lack of precipitation over a region for some period of time. Some of the studies analysed drought using monthly precipitation data, other analyse drought duration and intensity in relation to cumulative precipitation shortages. It caused by persistent anomalies, for example, high pressure in large-scale atmospheric circulation patterns, which are often caused by anomalous tropical sea surface temperatures or it can be other conditions (Mishra A. K. et al., 2010; Dai A. et al., 2011).

Hydrological drought

The second one, hydrological drought, is related to a period with inadequate surface and subsurface water resources for established water uses of a given water resources management system, such as lakes, reservoirs. Hydrological drought is a global and

international phenomenon, with spatial and temporal characteristics that very significantly from one region to another. It develops more slowly because it involves stored water too. For hydrologic drought analysis, have been widely used a streamflow data (Tallaksen et al., 2004; Mishra A. K. et al., 2010; Dai A., 2011).

Agricultural drought

Soil moisture is often use as an indicator of agricultural drought monitoring, and sometimes in some studies used in different forms, for example, the soil moisture percentile or normalized soil moisture. On several factors depends the soil moisture, which usually affect meteorological, hydrological drought and the differences between Evapotranspiration (ET) and Potential Evapotranspiration (PET). Few of drought indices, based on a combination of precipitation, temperature and soil moisture, have been developed to study agricultural drought. SSI and SPI are very effective in showing seasonal trends in precipitations and that's why they are good indicators of agricultural drought (AghaKouchak A., 2014; Belayneh A. et al., 2012; Mishra A. K. et al., 2010).

Socio-economic drought

Socio-economic drought is driven by unbalances in supply and demand of economic goods due to the physical characteristics. Usually socioeconomic drought will not occur without other categories of drought. The socio-economic impacts of drought disasters can be a serious obstacle to the development of many less developed countries with losses equivalent to several years of national growth gains. (American Meteorological Society, 2004; UN. ESCWA, 2005).

Ground water drought

Usually these are just four categories of drought used to classify the drought. But exist one more type of drought, it is ground water drought. When groundwater systems are affects by droughts, first recharge is ground water, and later groundwater levels and groundwater droughts are generally occurred on time scale of years or months or even weeks. It is very difficult to define the total amount of water available. But even if it is defined, in most of them, negative impacts of storage depletion can be felt, long before the total storage is depleted (Mishra A. K. et al., 2010).

2.3 Impact of drought around the world

Before I began to summarize the impacts of drought, I want to mention that the drought is a serious socio-economic challenge to many countries in the world. Long-term droughts caused mass migration and humanitarian crises. Even thought, people are trying to minimalize the effects of drought to a greater extent by extending irrigation facilities and adopting crop rotation methods. Failure to develop strategies results in loss of production and livelihood to people. The assessment of future meteorological risks has become important to policymakers. Climate change may affect the energy, water and nutrient balance of forest ecosystems (Juana J.S. et al., 2014; Rebetez M. et al., 2006).

At least 330 million people are affected by drought in India. This number is likely to rise and temperature there crossing 40ºC for days. Every year hundreds of people, mainly the poor, die at the height of summer in India, but temperatures have risen earlier than normal, increasing concerns about this year's toll. Out of 795 million ha of geographical area in India about 260 million ha of land are subjected to different degrees of water stress and drought conditions (Mishra et al., 2006).

Another country that affected by drought is Botswana. Botswana is a semi-arid country in the centre of Southern Africa, with total area of 58 173 000 ha. This country is really water stressed with an average annual precipitation rate of 416 mm/year. The problem of drought in Botswana is explained by the persistent recurrence of drought. Drought is characterized by low precipitation, low humidity, high temperatures and high wind velocity. Drought causes low water supplies that are inadequate to support economic activities (Juana J.S. et al., 2014).

Summer 2003, was exceptionally hot over most of central and western Europe, from Spain to Hungary and from Iceland to Greece. This illation based on historical climatology data show that it was by far the hottest summer since at least AD 1500. Based on mean surface air temperatures, the hottest region was centred over Germany, France, northern Italy and western Switzerland. Besides July, two months June and August showed the largest deviation from mean air temperatures compared to the climate standard period (1961 - 1990). July was less extreme except Sweden, Finland and Norway. On the other part of the Europe both minimum and maximum air temperature were between 1 and 3 ºC above normal. May was hotter and maximum

temperatures were more than 4 ºC above normal from Italy to Bulgaria. Forest ecosystems were exposed to drought in many European countries (Rebetez M. et al., 2006).

For Czech Republic in the last 15 years, drought was not essential issue, besides periods 1993 – 1995 and 2003 – 2004. In 1997, 2002 and 2013 the large part of the country was significantly affected by extreme floods. A significant drought developed in the summer of 2015. August 16th could be identified as the peak of the summer 2015 drought. Rainfall total reached 353 mm from the beginning of January to the end of August. It makes it the second lowest rainfall total since 1961, except 2003, when the total precipitation was 335 mm (CHMI, 2015; UNISEF, 2008).

3 Material and Methods

3.1 Drought indices

In this chapter I want to describe drought indices, their usefulness, limitations and comparison between them. Before I began to describe indices, it is important to define what is the difference between *indicators* and *indices*. *Indicators* are variable of parameters to describe the drought condition, for example: precipitation, temperature, soil moisture, streamflow, groundwater and snowpack. *Indices* are computed numerical representations of drought severity, assessed using climatic or hydro meteorological inputs. They have been created for measurement the qualitative state of drought on the landscape for a given time period. In combination with additional information on exposed assets and their vulnerability characteristics are essential for tracking and anticipating drought-related impacts and outcomes. It plays critical role, especially depending on the index, that they can provide a historical reference for planners or decision-makers (Svoboda et at., 2016, Mishra A. K. et al., 2010).

In the past, scientists used just one indicator or index, because that was the only one measurement available to them and they had a limited time to acquire data and compute derivative indices or other deliverables (Svoboda et al., 2016).

Commonly, there are three main methods for monitoring drought and quidding early warning and assessment (Jacobi et al., 2013):

- Using a single indicator or index
- Using multiple indicators or indices
- Using composite or hybrid indicators

In recent years, a few different indices have been developed to identify a drought, but each of them has its own strengths and weaknesses. Usually they include a lot of them: The Palmer Drought Severity Index (PDSI; Palmer, 1965), Rainfall Anomaly Index (RAI; Van Rooy, 1965), Crop Moisture Index (CMI; Palmer, 1968), National Rainfall Index (NRI; Gommes And Petrassi, 1994), The Soil Moisture Index (SMDI; Hollinger Et Al., 1993), Standardized Precipitation Index (SPI; Mckee Et Al., 1993), Standardized Precipitation Evapotranspiration Index (Vicente-Serrano Et Al., 2010), Standardized Soil Moisture Index (Aghakouchak A., 2014). , but in my research I will describe just few of them: SPI, SPEI and SSI. Some of them are perceived to be more

useful and easy to implement and had a long history of usage within the drought community. In medium-range climate forecasting there are two predictands commonly used: El Nino Southern Oscillation (ENSO) and North Atlantic Oscillation indices (NAO). The ENSO is an anomalous large-scale ocean-atmosphere system associated with strong fluctuations in ocean currents and surface temperatures. NAO is a main mode of winter climate variability in the North Atlantic region ranging from central North America to Europe and much into Northern Asia. The ENSO is an excellent indicator to drought in Australia. (Mishra et al.,2010; Choi et al., 2012).

Almost all drought indices use precipitation either singly or in combination with other meteorological variables, such as temperature, soil moisture, but SPI use only precipitation (Mishra A.K., 2010).

3.1.1 The Standardized Precipitation Index (SPI)

The Standardized Precipitation Index (SPI), outlined by McKee et al. (1993) and Guttman (1999), measures normalized anomalies in precipitation and has been proposed as a key drought indicator by the World Meteorological organization (WMO) and universal meteorological drought index by the Lincoln Declaration on Drought (Hayes et al., 2011). SPI is recommended as a probabilistic drought index, which is simple and spatially consistent in its interpretation (Guttman et al., 1998; Hayes et al., 1999).

The standardized precipitation index based on the precipitation probabilistic approach and has found wide using for describing and comparing drought among defend time periods and regions with different climatic conditions. SPI has been used for studying in different aspects of drought, like forecasting, frequency analysis, climate impacts studies and temporal analysis, the reason for that, is comparable in time and space (Vicente-Serrano et al., 2009, 2010; Cancelliere et al., 2007; Mishra et al., 2010).

First step in calculating the SPI is to determine a probability density function that describes the long-term series of precipitation observations. The series can be for any time duration. Once it determined, the cumulative probability of an observed precipitation amount is computed. The inverse normal (Gaussian) function with mean zero and variance one, is then applied to the cumulative probability. The result is SPI (Guttman et al., 1998; Guttman et al., 1999).

Positive SPI values show wet conditions with greater median precipitation, and negative values of SPI shows dry conditions with lower than median precipitation. The SPI values split the range, as shown on Tab. 2 ((Belayneh et al., 2012; Maca P. et al., 2015).

Tab. 2 Drought severity classification, SPI

	Index value	*Class*
Non drought	SPI ≥ 2,00	Extremely wet
	1,50 ≤ SPI < 2,00	Very wet
	1,00 ≤ SPI < 1,50	Moderately wet
	-1,00 ≤ SPI < 1,00	Near normal
Drought	-1,50 ≤ SPI < - 1,00	Moderate drought
	-2,00 ≤ SPI < -1,50	Severe drought
	SPI < -2,00	Extreme drought

The calculating of the SPI is based on the long-term precipitation record (at least 30 years) for any region. Those long-term precipitation time series is then fitted to a gamma distribution, which is then transformed through an equal probability transformation into a normal distribution. There some of commonly used distributions, like: gamma distribution, Pearson Type III distribution, lognormal, extreme value and exponential distributions (Guttman et al., 1999; McKee et al., 1993).

In 2009, World Meteorological Organization (WMO) recommended this Index as the main meteorological drought index that countries should use to follow and monitor drought (Hayes et al., 2011).

3.1.1.1 Limitation of SPI

The main criticism of the SPI is that its calculations based only on precipitation data. The length of precipitation record and characteristic of probability distribution has a significant impact on the SPI values. The index does not include other variabilities like temperature, evapotranspiration (Mishra et al., 2010).

3.1.1.2 Comparison

In 1998 Gutman and Hayes et al. in 1999 compared the Standardized Precipitation Index with Palmer Drought Severity Index and they both concluded that SPI has advantages of statistical consistency and the ability to describe both, the short-term

and long-term impacts of drought through the different time scales of precipitation anomalies. Especially, for carrying out drought risk analysis, the SPI is considered as ideal candidate for that. At all six criteria of performance, like robustness, tractability, transparency, sophistication, extendibility and dimensionality, SPI shows just indicates strengths over PDSI. Eventually, often SPI is chosen due to its simplicity, ability to represent drought on multiple time scales and especially because it's probabilistic drought index (Guttmann, 1999; Cancelliere A. et al., 2007; Belayneh et al., 2012).

3.1.2 The Standardized Precipitation Evapotranspiration Index (SPEI)

The more recently recommended Standardized Precipitation Evapotranspiration Index (SPEI) was developed by Vicente-Serrano et al. in 2010 at the Instituto Pirenaico de Ecologia In Zaragoza, Spain. It is a relatively new index, that utilizes a similar concept as SPI, but instead normalizes accumulated climatic water balance anomalies, defined as the difference between precipitation and potential evapotranspiration (PET). SPEI can be used to identify and monitor conditions associated with a variety of drought impacts (Vicente-Serrano et al., 2010; Begueria et al., 2013).

The standardized precipitation evapotranspiration index very easy to calculate and is based on precipitation and PET. This new index is particularly suited to detecting, monitoring and exploring the consequences of global warming on drought condition (Vicente-Serrano et al., 2009, 2010).

In studying droughts using standardized drought indicators, most of them classified drought events into different categories. The SPEI values split the range, as shown on Tab. 3 (Maca P. et al., 2015; USDM, 2017).

Tab. 3 Drought severity classification, SPEI

	Index value	*Class*
Non drought	-1,00 ≤ SPEI < 1,00	Near normal
Drought	-1,50 ≤ SPEI < - 1,00	Moderate drought
	-2,00 ≤ SPEI < -1,50	Severe drought
	SPEI < -2,00	Extreme drought

Mathematically, the SPEI is similar to the SPI, but includes the role of temperature. The first step in calculation of the SPEI is evapotranspiration (ET), the most difficult thing, because of the involvement of numerous parameters, including surface temperature, ground-atmosphere latent, sensible heat fluxes etc. (Allen et al., 1998; Vicente-Serrano et al., 2010).

Actual term evapotranspiration is usually used to describe two processed of water loss from land surface to atmosphere, evaporation and transpiration. Evaporation is a simply process by which liquid turns into water vapor (vaporation) and removed from sources such as the soil surface, wet vegetation, pavements, water bodies and etc. Transportation is a discharge of water vapor from the leaves of plants and subsequent loss of water as vapor through leaf stomata. There is a difference between evapotranspiration and potential evapotranspiration. Potential evapotranspiration is value of maximum amount of water that would be evapotranspired, if enough amount of water is available. Actual evapotranspiration is how much water is actually evaporatranspired and it's limited by the amount of water that is available. Actual evapotranspiration is always less or it can be equal to potential evapotranspiration, so that's why PET is used for water demand component of the drought equation (NOAA, 2016; Zotarelli et al., 2010).

There are a lot of methods for calculating PET from meteorological parameters measured at weather stations. One of them for example is Penman-Monteith method (PM). This method requires a big amount of data because its calculation involves values for solar radiation, temperature, wind speed, evapotranspiration and soil water capacity. But in many regions these data are not available. The simplest method to calculate is Thornthwaite method (Vicente-Serrano et al., 2009).

3.1.2.1 The Thornthwaite Method

As was mentioned before the SPEI is an extension of the widely used SPI. The SPEI was create to count both precipitation and potential evapotranspiration in determining drought. The SPEI captures the main impact of increased temperatures on water demand and can be calculated on a big range of a timescales from 1-48 months. At longer timescales, especially more than 18 months, the SPEI index has been shown to correlate with the self-calibrating PDSI. But when only limited data are available, PET can be estimated with the simple Thornthwaite method. This method based on not

accounted variables, which can affect PET, for example wind speed, surface humidity and solar radiation (UCAR, 2017). Following this simple method, the monthly PET (mm) is obtained by (Vicente-Serrano et al., 2009):

$$PET = 16K\frac{10T}{I}^{m} \quad (1)$$

where T is the monthly-mean temperature (^{o}C),

I is a heat index, which is calculated as the sum of 12 monthly index values i,

K is a correction coefficient computed as a function of the latitude and month,

m is a coefficient depending on i.

The difference between the precipitation P and PET for the month i is calculated by:

$$D_i = P_i - PET_i \quad (2)$$

For calculation of the SPI at different time scales a probability distribution of the gamma family is used. Even though two-parameters distribution is using for calculation the SPI such as gamma distribution, a three-parameter distribution is needed to calculate the SPEI, such as Pearson III, Log-normal, General Extreme Value, Log-logictic. (Vicente-Serrano et al., 2009).

L-moments ratio diagrams are analogous to conventional central moments, but they are able to characterize a wider range of distribution function and are most robust in a relation to outliers in the data.

3.1.2.2 The Penman-Monteith method

Lately, there are many empirical methods have been developed to estimate evapotranspiration from different climatic variables. Some of them were derived from now well-known Penman equation to determine evapotranspiration from open water, bare soil and grass based on a combination of an energy balance and an aerodynamic

formula, given as (Penman, 1948; Zotarelli et al., 2010):

$$\lambda E = \frac{[\Delta(R_n - G) + (\gamma \lambda E_a)]}{(\Delta + \lambda)} \quad (3)$$

where λE – evaporative latent heat flux,

Δ – slope of the saturated vapor pressure curve,

R_n – net radiation flux,

G – sensible heat flux into the soil,

E – vapor transport of flux.

Various derivation of the Penman equation included a bulk surface resistance term and the resulting equation is now called the Penman-Monteith equation, which is expressed by following formula (Monteith, 1965; Zotarelli et al., 2010):

$$\lambda ET_0 = \frac{\Delta(R_n - G) + [86{,}400 \frac{\rho_a c_p (e_s^0 - e_a)}{r_{av}}]}{\Delta + \gamma (1 + \frac{r_s}{r_{av}})} \quad (4)$$

where ρ_a – air density,

Cp – specific heat of dry air,

e_s^0 – mean saturated vapor pressure,

r_{av} – bulk surface aerodynamic resistance for water vapor,

e_a – mean daily ambient vapor pressure,

r_s – the canopy surface resistance.

3.1.2.3 Limitations of SPEI

Main limitations of this index that is require more data than the precipitation SPI. It is also sensitive to the method to calculate potential evapotranspiration. And the last limitation, is that as with other drought indices should be used a long-base period (30-50 years) that samples the natural variability (NCAR, 2017).

3.1.2.4 Comparison

In calculation, the SPEI reminds of SPI, but important role play the temperature. As the SPEI is based on a water balance, it can be compared to the self-calibrated Palmer Drought Severity Index (sc-PDSI). Relative to the sc-PDSI, the SPEI has the advantage of being multi-scalar, which is crucial for drought analysis and monitoring. The SPEI combines sensitivity of PDSI to changes in evaporation demand with the simplicity of calculation and the multi-temporal nature of SPI (Vicente-Serrano et al., 2010).

3.1.3 The Standardized Soil Moisture Index (SSI)

As was describing before, the drought is classified as agricultural, meteorological and hydrological have been developed to describe different aspects of drought. Many of them are based on soil moisture, precipitation and runoff. Soil moisture is often used as an indicator of agricultural drought monitoring. The standardized soil moisture index can be defined in a same way to the popular standardized precipitation index, which is widely used in a variety of studies. The SSI is estimated using a nonparametric approach in which the empirical probability *(p)* of the historical soil moisture data is derived using the empirical Gringorten plotting position. Generally, instead of fitting a distribution function to soil moisture data, the probabilities are obtained empirically using the empirical Gringorten approach: $(i - 0.44)/(n + 0.12)$, where *n* denoted the sample size and *i* refers to the rank of soil moisture data from the smallest to the largest. (AghaKouchak, A., 2014; Gringorten, I., 1963).

Tab. 4 Drought severity classification, SSI (USDM, 2017).

	Index value	*Class*
Drought	-0,5 ≤ SSI< -0,7	Abnormally dry
	-0,8 ≤ SSI < -1,2	Moderate drought
	-1,3 ≤ SSI < -1,5	Severe drought
	-1,6 ≤ SSI < -1,9	Extreme drought
	SSI < -2,0	Exceptional drought

3.2 Artificial Neural Network (ANN)

The recent development of artificial neural network has a significant impact on the application of techniques for the forecasting of drought indices. An artificial neural network can be defined as a data processing system consisting of a large number of simple, highly interconnected processing elements, that calls artificial neurons, in an architecture inspired by the structure of the cerebral cortex of the brain (Tsoukalas et al., 1997; Maca et al., 2015).

3.2.1 Definition. What is a Neural Network?

Artificial Neural Network is an information processing approach that corresponds the structure and operation of the brain. In recent years, they have become a very popular thing for prediction in many of areas, for example forecasting in finance and medicine. First who introduce the concept of artificial networks were McCulloch and Pitts in 1943. ANNs are being used increasingly to predict and forecast water resources variables. One of the advantages of the ANN is that there is no need for the modeler to fully define the intermediate relationships, like physical processes, between inputs and outputs. This feature makes ANNs especially suitable for analysis of complex processes, like drought forecasting, where relationships of a large number of input variables with the output need to be explored. In my thesis, I will use Multilayer Perceptron, nowadays the most widespread topological tool (Maier et al., 1992, 2000; Morid S., et al., 2007).

3.2.2 The Neuron model

A neuron is an information-processing unit that is fundamental to the operation of a neural network. In a schematic drawing of a biological neuron (Fig. 1) we may see the tree-like communication network of nerve fibber called *dendrites*. They bring the inputs to *the cell body* (the soma) that contains the cell nucleus. Another received signal, from another neuron for example, passed through a *synapse* by way of a difficult complex chemical process, in which specific transmitter substances are released from the sending side of the junction. The aim is to lower or to raise the electrical protentional inside the cell body. If it reaches a threshold, an *action potential* of fixed strength and duration is transmitted down *the axon* of the neuron. Then we can say that neuron has “fired” (Haykin S., 2008; Prieto A., 1991).

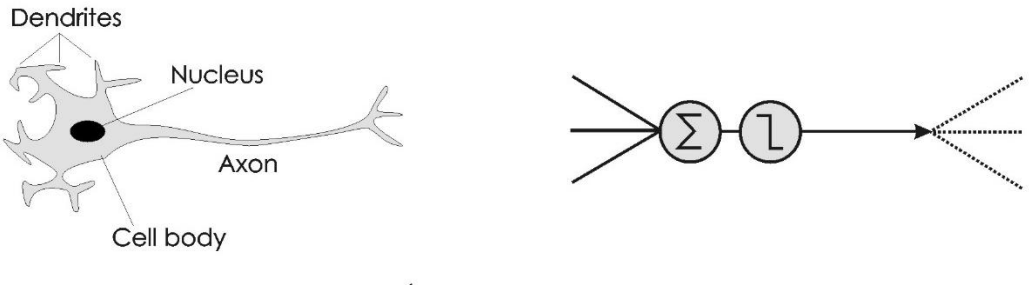

Fig. 1 Schematic comparison between a biological neuron and artificial (Leverington D., 2009).

A single artificial neuron (Fig. 1) can be described as a processing unit which gets a stimulation from its neighbours and respond to a given activation function. The operation of the artificial neuron is analogous to the operation of the biological neuron. Activations from other neurons are summed at the neuron and passed through an activation function, after which the value is sent to other neurons. The first static simple model of neuron is a binary threshold unit, which operates according to the following equation, was written by McCullogh-Pitts (Prieto A., 1991):

$$x_i(t+1) = stp(\sum_j w_{ij}x_j(t) - \theta_i) \tag{5}$$

where $x_i(t)$ is the output of neuron *i* at time *t*,

w_{ij} is the weight of the synapse between neurons *i* and *j*.

3.2.3 Network architectures

In general, there are three different classes of network architectures (Haykin S, 2008):

- Single-Layer Feedforward Networks
- Multilayer Feedforward Networks
- Recurrent Networks

3.2.3.1 Single-Layer Feedforward Networks

Network called a *single-layer network*, if it has the designation single layer referring to the output layer of computation nodes. The input layer of source nodes does not count because no computation is performed there (Haykin S., 2008). As illustrated in Fig. 2, the case of three nodes in both the input and output layers.

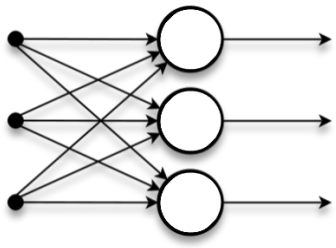

Fig. 2 Feedforward network with a single layer of neurons (Chrislb, 2005).

3.2.3.2 Multilayer Feedforward Networks

Usually processing elements are organized into a sequence of layers with full or random connections between them. A typical neural network is "*fully connected*". That means that there is a connection between each of the neurons in any given layer with each of the neurons in the next layer. Every neural network includes: input layers(s), hidden layer(s) and output layer(s). The input layer is a buffer which presents data to the network. This layer doesn't have weights and any activation function, and that's why it is not a neural computing layer. The top layer is the output layer which presents the output response to a given input. The rest of the layers, or it can be just one layer are called the intermediate or hidden layer because it usually has no connections to the outside world. The simple neural network is a fully connected, feedforward network with three neurons in the output layer, four in the middle or hidden layer, and two in the output layer. This architecture of the artificial neural networks called *Multilayer Perceptron* (MLP). Feedforward network means that there are no lateral connections between neurons in given layer and none back to previous layers. In all cases, these connections have weights and they had to be trained (Morid et al., 2007; Tsoukalas et al., 1997).

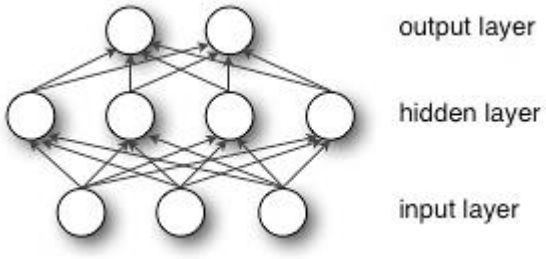

Fig. 3 Fully connected Multilayer Perceptron with one hidden layer (DeepLearning 0.1 Documentation, 2017)

The neural network in Fig. 3 is said to be fully connected, that means that every node in each layer of the network is connected to every other node in the adjacent forward layer. If some of the synaptic connections are missing, we called that network is "*partially connected"* (Haykin S., 2008).

3.2.3.3 Recurrent neural networks

A recurrent neural network, in a difference with a feedforward neural network is that it has at least one feedback loop. For example, a recurrent network may consist of a single layer of neurons with each neuron feeding its output signal back to the outputs of all other neurons (Haykin S., 2008).

3.2.4 Back-Propagation Algorithm

The most important tasks are to clearly define the choice of the numerical method used for determining optimal connection weights, or training algorithm. The Back-propagation (BP) algorithm has been used in many studies. Backpropagation is a systematic algorithm for training multiple layer artificial neural networks. The first who developed backpropagation training was Werbos (1974) as a part of his Ph.D. dissertation at Harvard University. The elucidation of this training algorithm in 1986 by Rumelhart, Hinton and Williams was the key step in making neural networks practical in many real-world situations. But, the BP algorithm can suffer from a few problems, the majority of which are the need for selection of several internal model parameters (momentum, learning rate and transfer function) and slow convergence speed. Despite this, 80% of all application utilize this backpropagation algorithm in one from another, because it has a strong mathematical foundation (Maier et al., 1998).

4 Results

4.1 Dataset description

Neural Network prediction for the drought indices, I used the data obtained from 13 meteorological stations (Tab. 5), located in different parts of Czech Republic. The records that I used from the period 1.1.1982 – 1.12.2015.

Tab. 5 Dataset used for prediction drought indices.

Identifier	*Name*	*Name of the station*	*Area [km^2]*	*Number of hydrological sequence*
1980	Berounka	Beroun	8 286,23	1-11-04-0560-0-00-30
2110	Teplá	Cihelny	262,58	1-13-02-0210-1-00-60
2400	Labe	Děčín	51 120,34	1-14-04-0010-0-00-70
2940	Odra	Bohumín	4 663,74	2-03-02-0220-0-00-30
3450	Morava	Raškov	349,79	4-10-01-0450-0-00-60
3511	Desna	Šumperk	240,63	4-10-01-0850-0-00-50
3540	Moravská Sázava	Lupěné	445,21	4-1-02-0420-0-00-70
3550	Morava	Moravičany	1 561,19	4-10-02-0650-0-00-70
4215	Morava	Strážnice	9 144,83	4-13-02-0340-0-00-30
4410	Svratka	Borovnice	127,97	4-15-01-0070-0-00-70
4530	Křetínka	Letovice	126,59	4-15-02-0340-2-00-30
4540	Svitava	Letovice	423,78	4-15-02-0350-0-00-30
4650	Jihlava	Dvorce	307,35	4-16-01-0270-0-00-50

4.1.1 Time series of drought indices

The indices of drought are essential tools for explaining the severity of drought events. Usually they are presented in a form of time series and used in drought modelling and forecasting. A time series also provides a framework for evaluating drought parameters of interest (Maca P. et al., 2015; Mishra A. et al., 2010). Time series of all three standardized indices for every meteorological station are shown in attachments, Fig. 4 – 16.

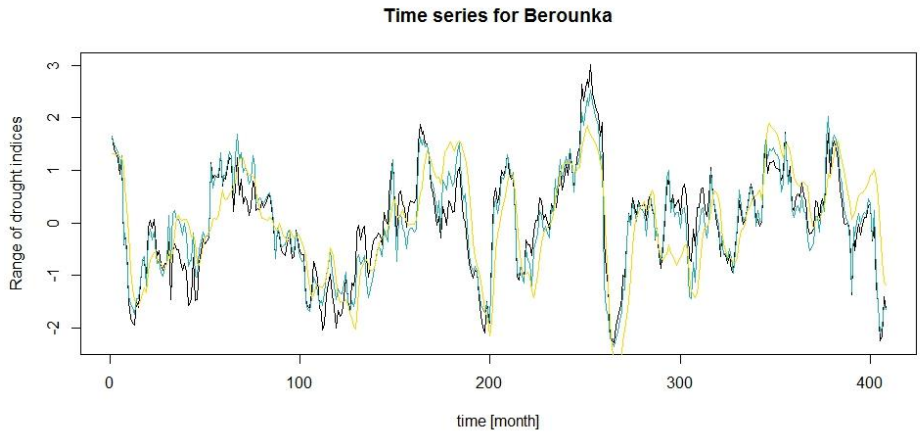

Fig. 4 Time series for Berounka, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

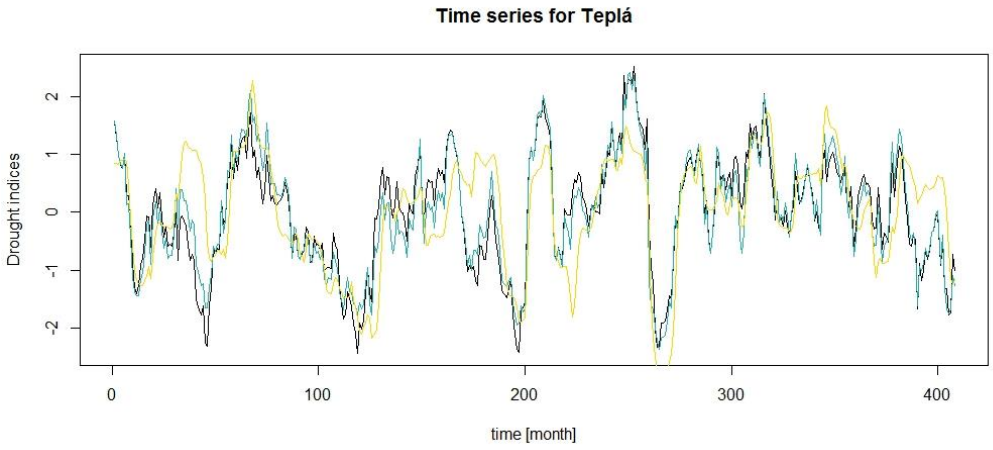

Fig. 5 Time series for Teplá, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

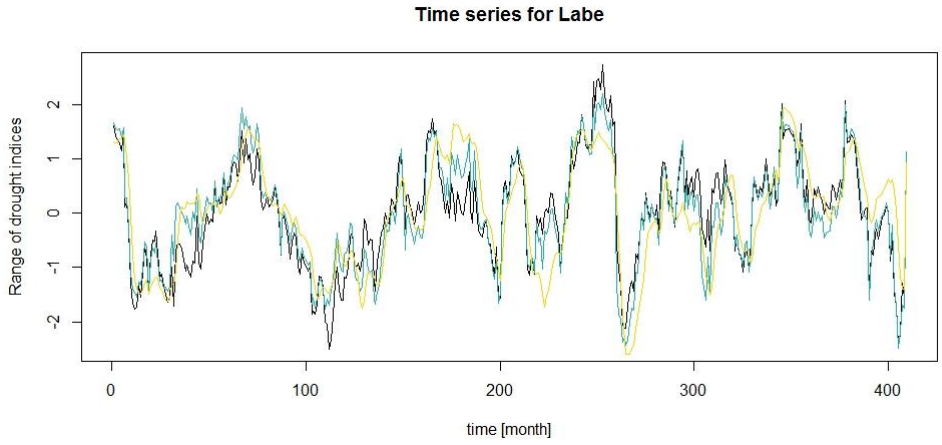

Fig. 6 Time series for Labe, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

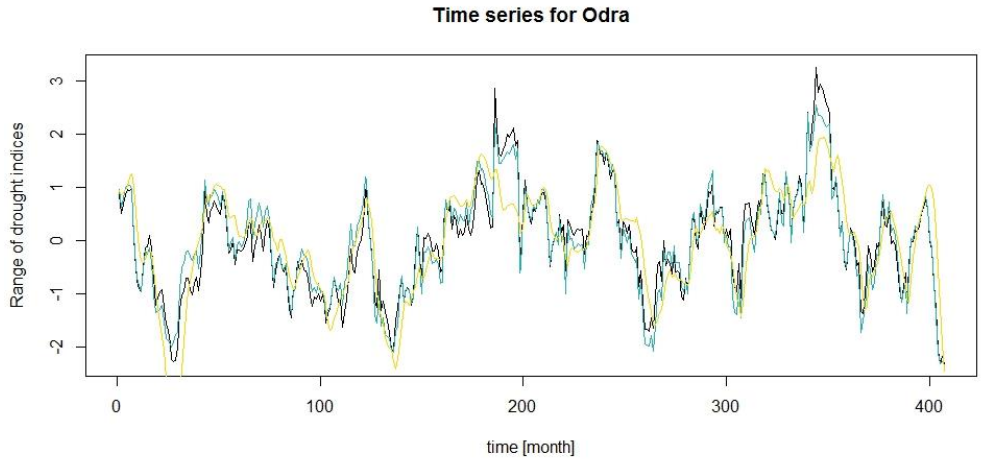

Fig. 7 Time series for Odra, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

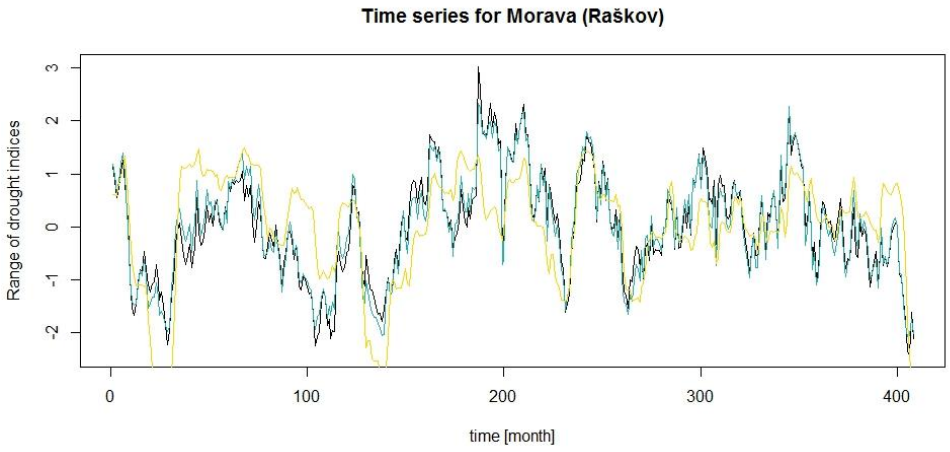

Fig. 8 Time series for Morava (Raškov), on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

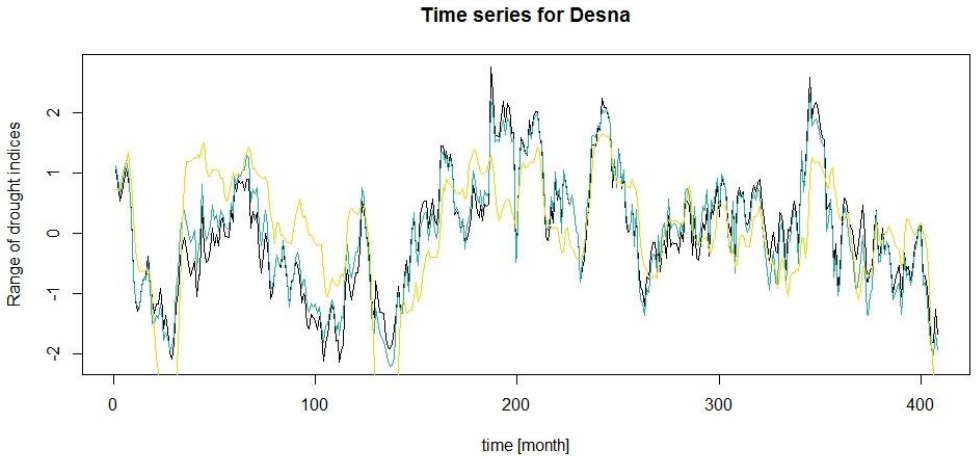

Fig. 9 Time series for Desna, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

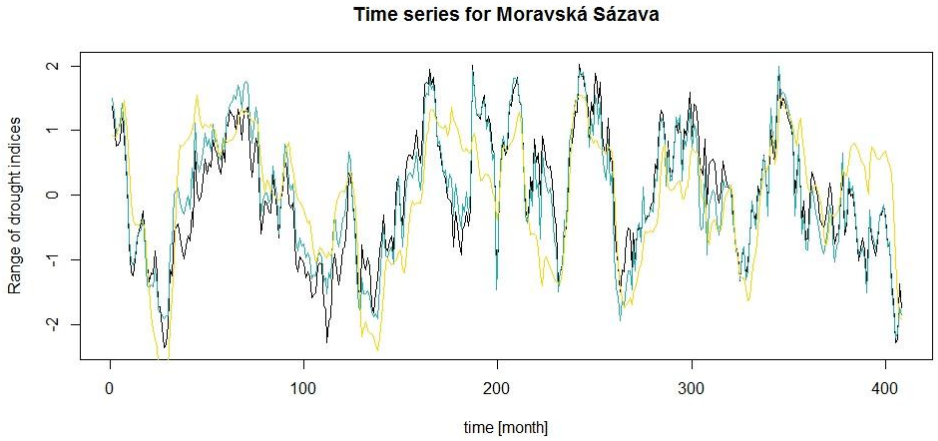

Fig. 10 Time series for Moravská Sázava, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

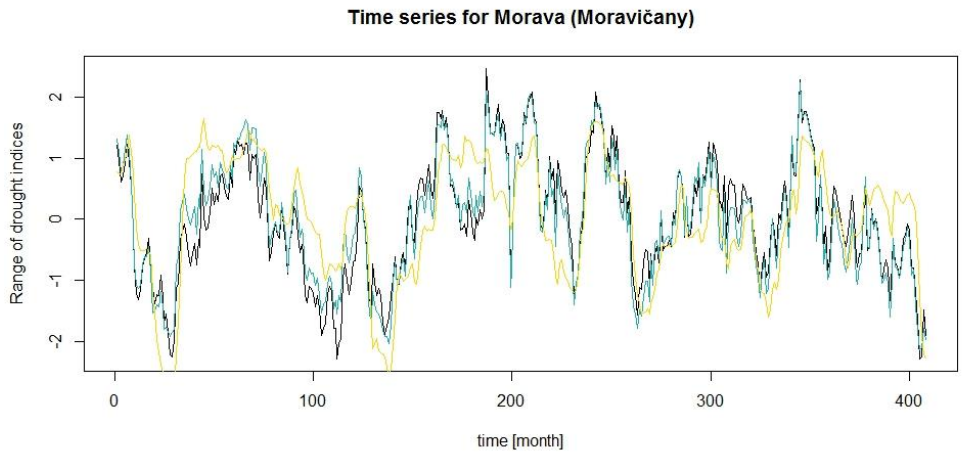

Fig. 11 Time series for Morava (Moravičany), on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

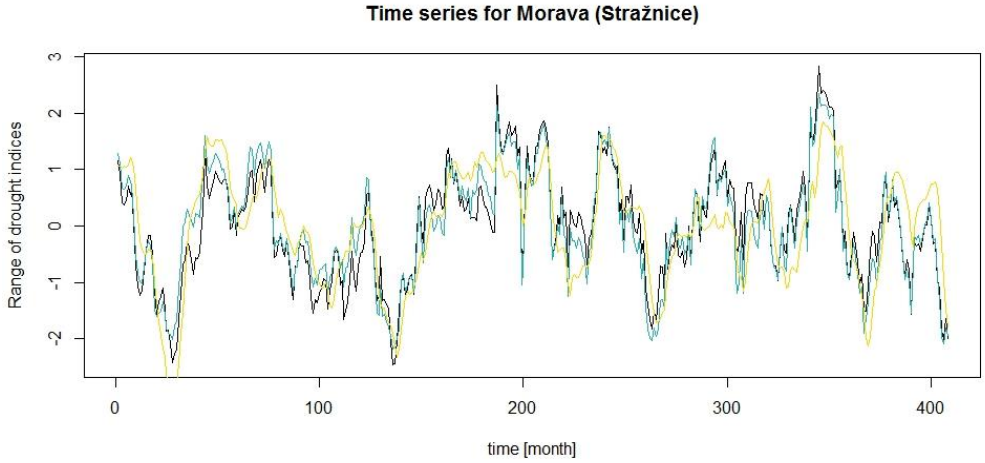

Fig. 12 Time series for Morava (Strážnice), on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

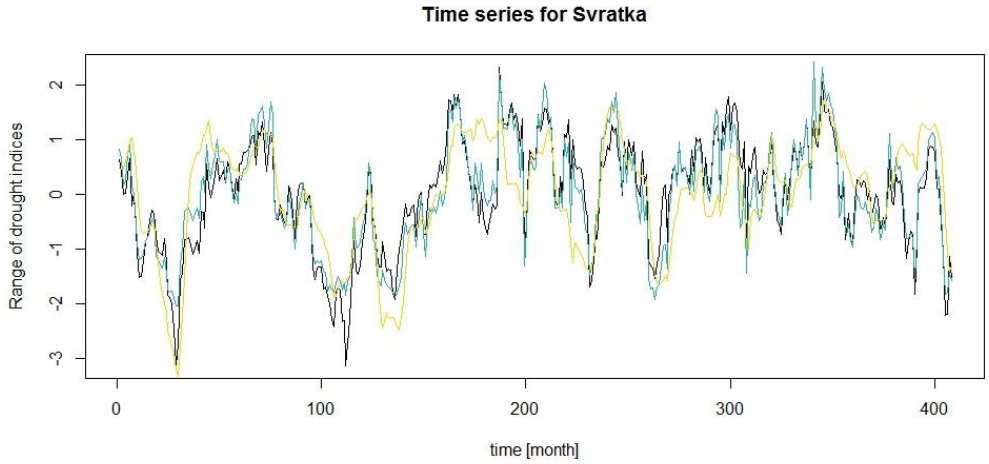

Fig. 13 Time series for Svratka, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

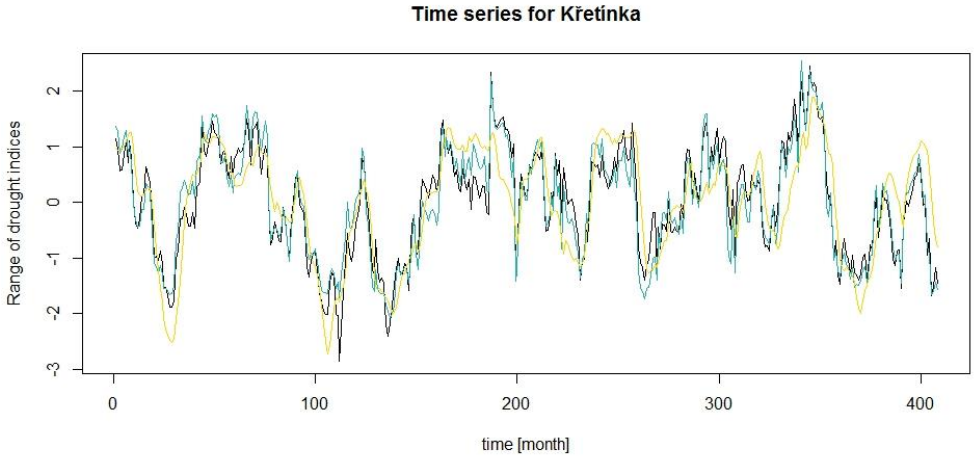

Fig. 14 Time series for Křetínka, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

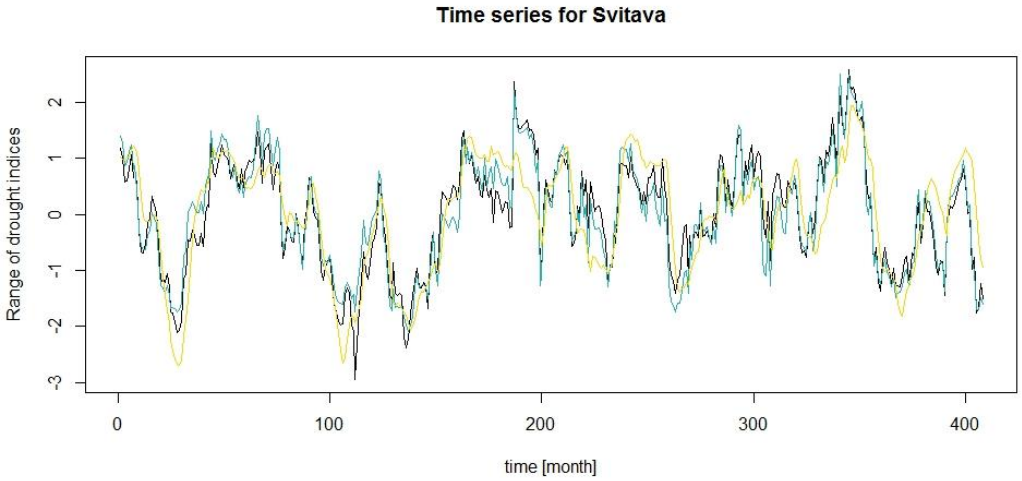

Fig. 15 Time series for Svitava, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

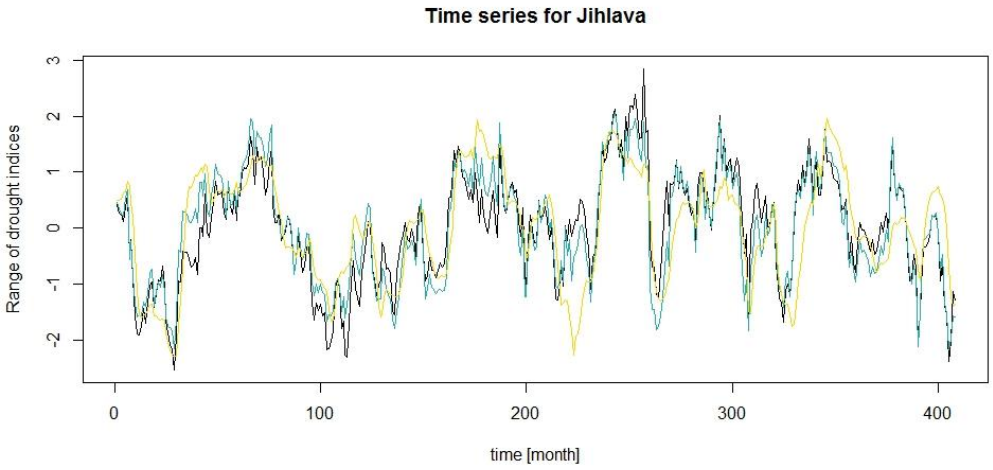

Fig. 16 Time series for Jihlava, on x axis is time in month, on y axis is a range of drought indices. The black line: SPI, the blue line: SPEI and the yellow line: SSI index.

4.1.2 Statistical evaluation of input data

Usually they are used to summarize a set of observations. The five-number summary is a useful measure of variation for observed data. This include the mean or median of numeric data, minimum and maximum, 25% of values fall below the 1st quartile and 25% of values fall above the 3rd quartile. I also included Standard Deviation, which shows a measure of dispersion of a set of data from its mean. This range called *the interquartile range* (Mangiafice S.S., 2016).

On Tab 6 represented basic statistics of precipitation for every meteorological station. Maximum and minimum value of precipitation was on 3450 Morava (Raškov).

On Tab. 7 to Tab. 9 represented basic statistics of every drought index and every meteorological station.

Tab. 6 Summary statistics of precipitation

DBC	*Rain*					
	Min	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*MAX*
1980	1,199	36,33	56,05	60,01	75,75	200,2
2110	1,434	3,08	47,24	52,17	67,6	205,6
2400	1,141	34,34	52,09	55,91	71,18	204,8
2940	0,550	37,11	55,34	65,52	85,84	393,5
3450	**0,433**	49,86	71,22	79,67	101,6	**455,8**
3511	0,508	49,9	72,16	78,6	100,1	431,5
3540	0,536	38,64	58,89	64,52	83,84	328,6
3550	0,536	42,99	62,61	68,2	87,33	352
4215	0,496	35,55	53,6	58,86	78,58	309,4
4410	0,649	39,28	61,56	67,11	86,9	361,2
4530	0,477	31,34	48,16	54,65	71,23	314,9
4540	0,469	31,35	46,55	54,11	69,84	324,3
4650	1,516	33,21	53,17	57,29	75,26	213,5

As seen on Tab. 7 the maximum value of Standardized Precipitation Index was on 2940 (Bohumin), the minimum value of SPI was on 4410 (Borovnice).

Tab. 7 Basic statistics for SPI

DBC	*SPI*						
	Min	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*MAX*	*sd*
1980	-2,2900	-0,6358	0,0528	-0,0027	0,5689	3,0240	0,9827
2110	-2,4440	-0,6412	0,0517	-0,0022	0,6882	2,5310	0,9798
2400	-2,5070	-0,6929	0,0526	0,0007	0,6054	2,7350	0,9746
2940	-2,3280	-0,6570	0,0097	-0,0041	0,5524	**3,2670**	0,9866
3450	-2,4330	-0,6256	-0,0061	-0,0009	0,6592	3,0350	0,9728
3511	-2,1460	-0,6579	-0,0194	0,0002	0,5397	2,7650	0,9728
3540	-2,3660	-0,7096	-0,0373	-0,0004	0,7092	2,0330	0,9633
3550	-2,2960	-0,7023	0,0111	-0,0006	0,6571	2,4800	0,9675
4215	-2,4600	-0,6429	0,0149	-0,0008	0,6521	2,8400	0,9736
4410	**-3,1460**	-0,6675	0,1113	-0,0014	0,7136	2,3400	0,9939
4530	-2,8630	-0,7693	0,0539	-0,0006	0,7550	2,4610	0,9724
4540	-2,9510	-0,7531	0,1207	-0,0009	0,7079	2,5920	0,9760
4650	-2,5390	-0,6948	0,0584	0,0008	0,6748	2,8550	0,9801

A histogram is a display of statistical information that uses rectangles to show the frequency of data items in successive numerical intervals of equal size. A histogram is the easiest method for determining the density of the data. In Fig. 17 there is a histogram of SPI.

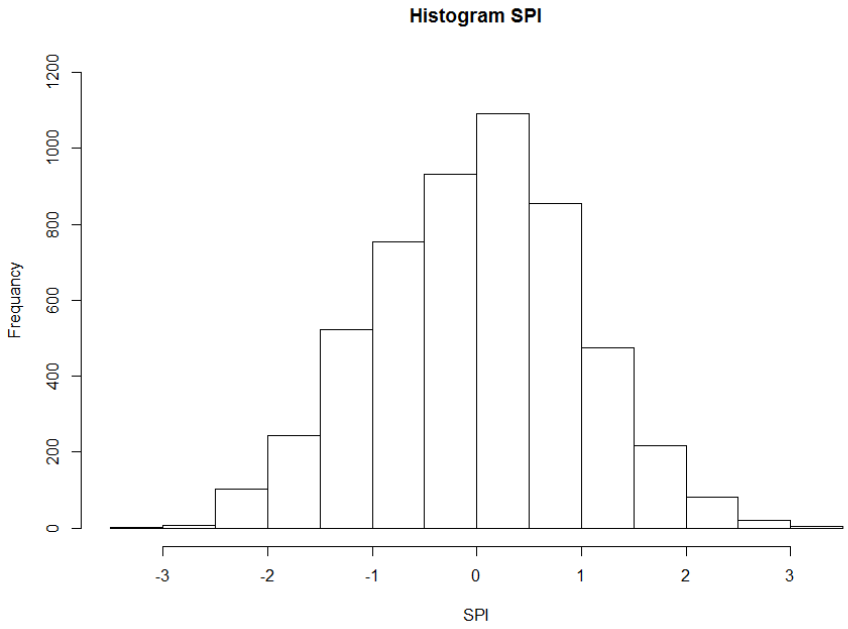

Fig. 17 Histogram of SPI, on x axis there is a range of SPI, on y axis is the frequency of SPI

Empirical distribution function is a formal direct estimate of the cumulative distribution function. In hydrology, empirical distribution functions are commonly labelled duration curves, in which case all the values in a time series are included (Tallaksen et al., 2004). For SPI, empirical distribution is shown in Fig. 18.

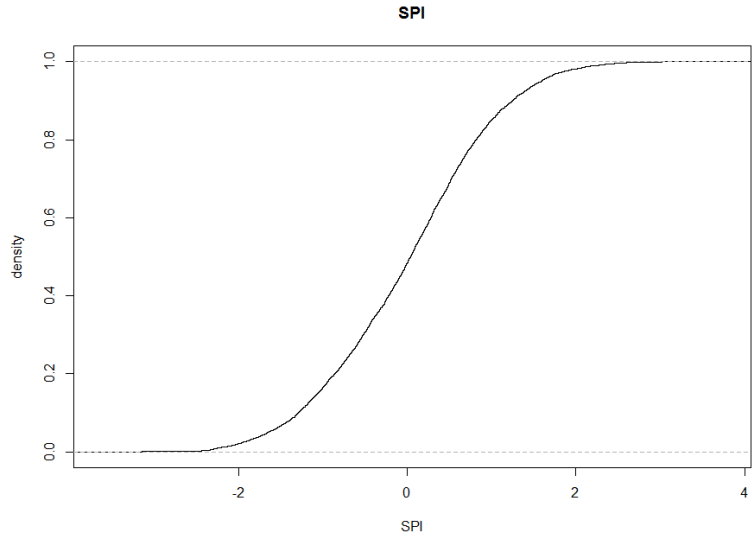

Fig. 18 Empirical distribution function of SPI, on x axis there is a range of SPI, on y axis is the density of SPI

On Tab. 8 the maximum value of Standardized Precipitation Evapotranspiration Index was the same on two meteorological stations: on 2940 (Bohumin) and 4530 (Letovice). The minimum value of SPEI was on 2400 (Děčin).

Tab. 8 Basic statistics for SPEI

DBC	*SPEI*						
	Min	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*MAX*	*sd*
1980	-2,3670	-0,7106	0,0305	-0,0052	0,7386	2,5320	0,9777
2110	-2,3680	-0,6837	0,0011	-0,0049	0,7353	2,4200	0,9774
2400	**-2,4940**	-0,6423	0,0040	-0,0005	0,6717	2,1860	0,9817
2940	-2,3370	-0,6906	-0,0238	-0,0027	0,6436	**2,5500**	0,9815
3450	-2,0850	-0,6758	0,0106	-0,0025	0,6795	2,3330	0,9776
3511	-2,2060	-0,7357	0,0271	0,0015	0,6519	2,3090	0,9794
3540	-2,2030	-0,7583	-0,0071	-0,0026	0,7533	1,9988	0,9783
3550	-2,1410	-0,7475	0,0074	-0,0018	0,6954	2,2450	0,9772
4215	-2,1780	-0,6982	-0,0393	-0,0019	0,7067	2,3530	0,9779
4410	-2,0510	-0,6436	-0,0098	-0,0091	0,6882	2,4230	0,9804
4530	-2,0820	-0,8344	0,0561	-0,0074	0,7633	**2,5500**	0,9807
4540	-2,0960	-0,7993	0,0221	-0,0073	0,6915	2,5050	0,9801
4650	-2,1760	-0,8063	0,0393	-0,0008	0,7669	2,0770	0,9816

In Fig. 19 is shown the histogram of SPEI, where shown the most frequently occurring value of SPI.

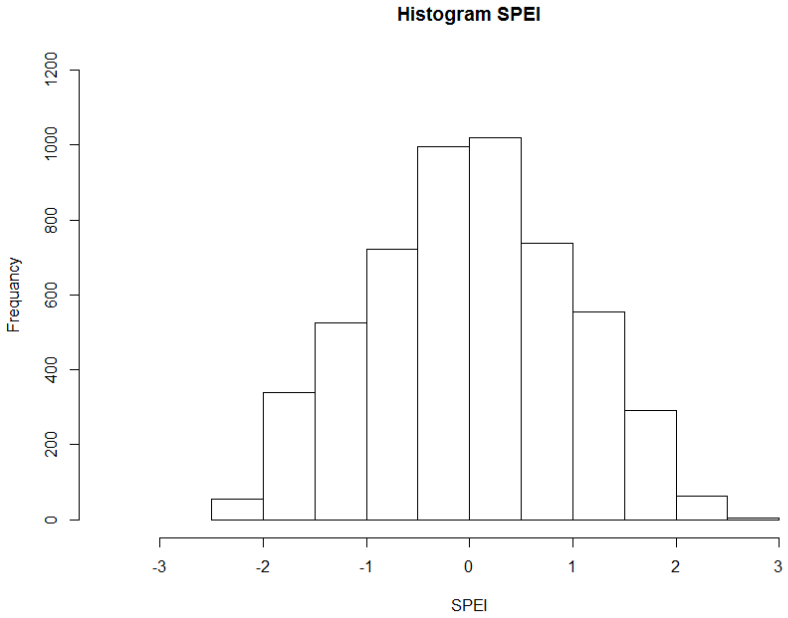

Fig. 19 Histogram of SPEI, on x axis there is a range of SPEI, on y axis is the frequency of SPEI

For SPEI empirical distribution is shown in Fig. 20.

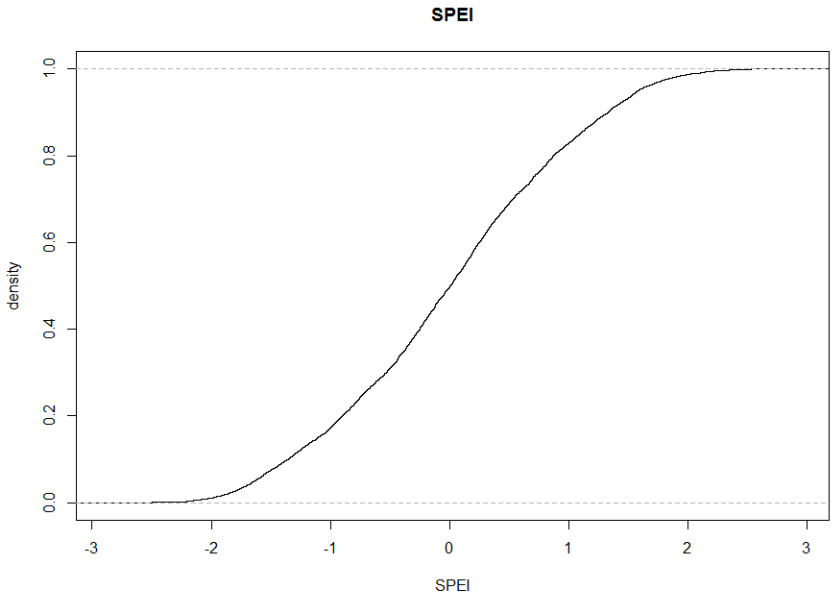

Fig. 20 Empirical distribution function of SPEI, on x axis there is a range of SPEI, on y axis the density of SPEI

As seen on Tab. 9, the maximum value of the Standardized Soil Moisture Index was on 2110 (Cihelny) and minimum was on 3511 (Šumperk).

Tab. 9 Basic statistics for SSI

DBC	*SSI*						
	Min	*1st Qu.*	*Median*	*Mean*	*3rd Qu.*	*MAX*	*sd*
1980	-2,7480	-0,6956	0,0407	-0,0018	0,7525	1,9010	0,9765
2110	-2,9910	-0,6070	0,1521	-0,0019	0,7837	**2,2800**	0,9965
2400	-2,5950	-0,7293	0,0861	0,0002	0,6717	2,1860	0,9730
2940	-3,4490	-0,6522	0,1123	-0,0044	0,7036	1,9370	1,0040
3450	-3,6120	-0,2961	0,1763	-0,0032	0,7360	1,4940	1,0517
3511	**-3,9670**	-0,4850	0,1157	-0,0035	0,7738	1,6310	1,0595
3540	-2,8200	-0,6071	0,1341	-0,0019	0,7804	1,5510	0,9916
3550	-3,0350	-0,5663	0,0953	-0,0018	0,8572	1,6540	0,9982
4215	-2,9920	-0,6538	0,0699	-0,0018	0,7856	1,8430	0,9855
4410	-3,3330	-0,5579	0,1487	-0,0014	0,7695	1,7660	1,0057
4530	-2,7260	-0,6907	0,1651	-0,0017	0,8561	1,8910	0,9880
4540	-2,6940	-0,7128	0,1570	-0,0021	0,8161	1,9350	0,9923
4650	-2,2960	-0,7552	0,0998	0,0003	0,6912	1,9560	0,9658

In Fig. 10 is represented histogram of SSI, where represented the most frequently occurring value of SSI.

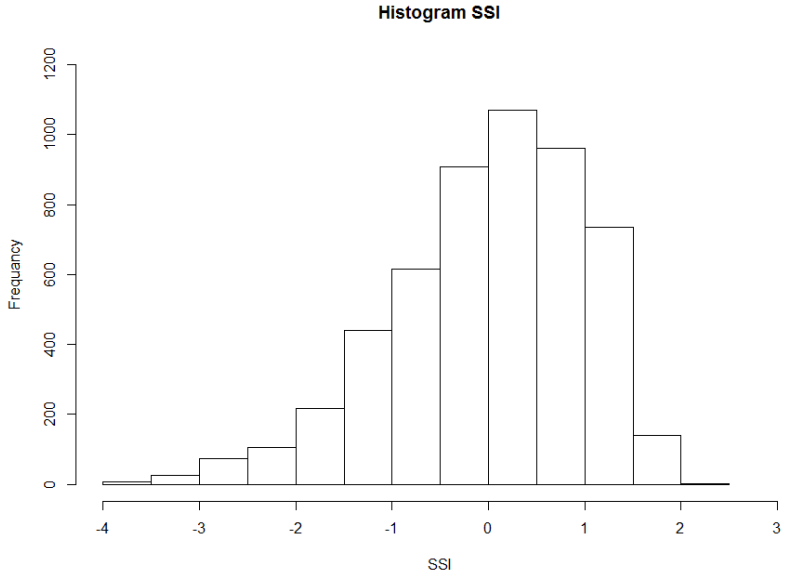

Fig. 10 Histogram of SSI, on x axis there is a range of SSI, on y axis is the frequency of SSI

For SSI empirical distribution is shown in Fig. 11.

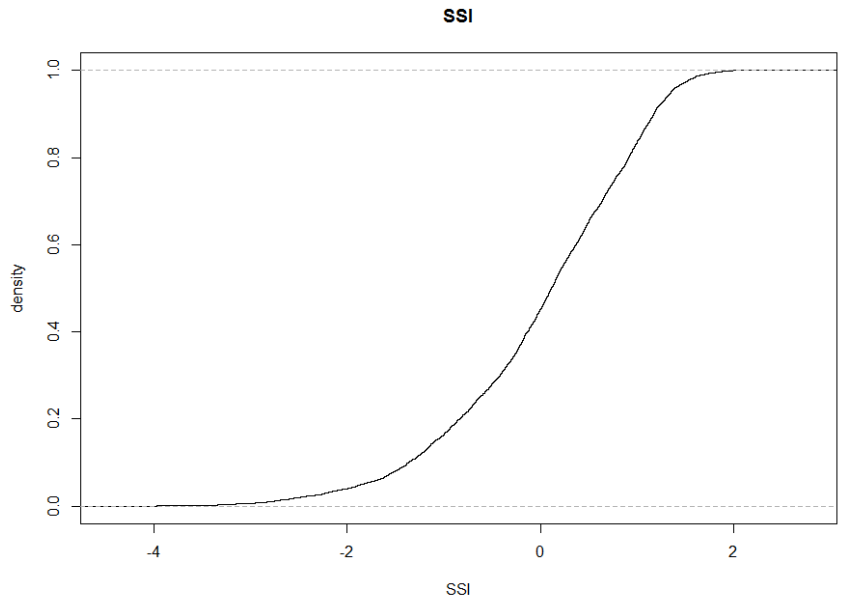

Fig. 11 Empirical distribution function of SSI, on x axis there is a range of SSI, on y axis is the density of SSI

4.1.3 Correlation

Correlation is a means of defining the strength of the relationship between two variables (Gordon et al., 2004). On presented below Tab. 12 - 14 shown correlation for every drought index between every meteorological station.

On Tab. 12 presented correlation between every meteorological between themselves. I created correlation with function "*cor*" in R studio. This function use to produce correlations. It shows which station correlated least of all or most of all respectively.

So, the best correlation was between two: 4530 Křetínka and 4540 Svitava. The worst correlation was between 2940 Odra and 2110 Teplá.

SPI	1980	2110	2400	2940	3450	3511	3540	3550	4215	4410	4530	4540	4650
1980	1.0000000	0.8728238	0.9375858	0.3960404	0.5137150	0.5287840	0.6442335	0.6059368	0.5092400	0.5480717	0.4669775	0.4734340	0.7108672
2110	0.8728238	1,0000000	0.8123412	0.2499089	0.4610982	0.4623020	0.5876634	0.5380100	0.3919514	0.4876570	0.3650726	0.3661687	0.5562124
2400	0.9375858	0.8123412	1,0000000	0.5721255	0.6755776	0.6815247	0.7920660	0.7578538	0.6822235	0.7220234	0.6426313	0.6510615	0.8484320
2940	0.3960404	0.2499089	0.5721255	1,0000000	0.7981413	0.8510411	0.7055393	0.7855098	0.9159901	0.7428212	0.8036784	0.8417420	0.6162213
3450	0.5137150	0.4610982	0.6755776	0.7981413	1,0000000	0.9600741	0.9128380	0.9710352	0.8784856	0.8857412	0.8157133	0.8426761	0.6813828
3511	0.5287840	0.4623020	0.6815247	0.8510411	0.9600741	1,0000000	0.8793485	0.9526949	0.8859241	0.8523864	0.7920046	0.8271838	0.6956120
3540	0.6442335	0.5876634	0.7920660	0.7055393	0.9128380	0.8793485	1,0000000	0.9738079	0.8592510	0.9036221	0.8474664	0.8643942	0.7927169
3550	0.6059368	0.5380100	0.7578538	0.7855098	0.9710352	0.9526949	0.9738079	1,0000000	0.9016321	0.9005201	0.8512262	0.8741139	0.7553708
4215	0.5092400	0.3919514	0.6822235	0.9159901	0.8784856	0.8859241	0.8592510	0.9016321	1,0000000	0.8364865	0.8779059	0.9091073	0.7107785
4410	0.5480717	0.4876570	0.7220234	0.7428212	0.8857412	0.8523864	0.9036221	0.9005201	0.8364865	1,0000000	0.8629777	0.8899930	0.8039669
4530	0.4669775	0.3650726	0.6426313	0.8036784	0.8157133	0.7920046	0.8474664	0.8512262	0.8779059	0.8629777	1,0000000	0.9889975	0.7400209
4540	0.4734340	0.3661687	0.6510615	0.8417420	0.8426761	0.8271838	0.8643942	0.8741139	0.9091073	0.8899930	0.9889975	1,0000000	0.7422361
4650	0.7108672	0.5562124	0.8484320	0.6162213	0.6813828	0.6956120	0.7927169	0.7553708	0.7107785	0.8039669	0.7400209	0.7422361	1,0000000

Tab. 12 Correlation between every meteorological station for SPI

SPEI	1980	2110	2400	2940	3450	3511	3540	3550	4215	4410	4530	4540	4650
1980	1,0000000	0.8570078	0.9484637	0.5110974	0.5650604	0.5694894	0.6658586	0.6335750	0.5805557	0.6173511	0.5647843	0.5676920	0.7583200
2110	0.8570078	1,0000000	0.8020349	0.8020349	0.3470008	0.5300419	0.5294301	0.6021576	0.5710847	0.4459268	0.5454741	**0.4355611**	0.5954598
2400	0.9484637	0.8020349	1,0000000	0.6637150	0.7071501	0.6954472	0.8004950	0.7724295	0.7397798	0.7595886	0.7214635	0.7230899	0.8702612
2940	0.5110974	0.3470008	0.6637150	1,0000000	0.8194729	0.8321298	0.7578355	0.8024214	0.9243832	0.8275237	0.8596001	0.8867395	0.7062046
3450	0.5650604	0.5300419	0.7071501	0.8194729	1,0000000	0.9659541	0.9171517	0.9650331	0.8779648	0.8944881	0.8312689	0.8510077	0.7181675
3511	0.5694894	0.5294301	0.6954472	0.8321298	0.9659541	1,0000000	0.8926061	0.9493465	0.8667433	0.8643041	0.8118889	0.8313309	0.7289256
3540	0.6658586	0.6021576	0.8004950	0.7578355	0.9171517	0.8926061	1,0000000	0.9806322	0.8869468	0.8908308	0.8683731	0.8799351	0.8047525
3550	0.6335750	0.5710847	0.7724295	0.8024214	0.9650331	0.9493465	0.9806322	1,0000000	0.9080040	0.8909364	0.8648854	0.8795497	0.7814645
4215	0.5805557	0.4459268	0.7397798	0.9243832	0.8779648	0.8667433	0.8869468	0.9080040	1,0000000	0.8824468	0.9160335	0.9371208	0.7808398
4410	0.6173511	0.5454741	0.7595886	0.8275237	0.8944881	0.8643041	0.8908308	0.8909364	0.8824468	1,0000000	0.8968086	0.9160811	0.8229426
4530	0.5647843	0.4355611	0.7214635	0.8596001	0.8312689	0.8118889	0.8683731	0.8648854	0.9160335	0.8968086	1,0000000	0.9910685	0.8058816
4540	0.5676920	0.4346248	0.7230899	0.8867395	0.8510077	0.8313309	0.8799351	0.8795497	0.9371208	0.9160811	**0.9910685**	1,0000000	0.7979794
4650	0.7583200	0.5954598	0.8702612	0.7062046	0.7181675	0.7289256	0.8047525	0.7814645	0.7808398	0.8229426	0.8058816	0.7979794	1,0000000

Tab. 13 Correlation between every meteorological station for SPEI

SSI	1980	2110	2400	2940	3450	3511	3540	3550	4215	4410	4530	4540	4650
1980	1,00000000	0.7992789	0.9401559	0.5614409	0.5791229	0.5192501	0.6692899	0.6314326	0.6115355	0.6994082	0.5976261	0.6223244	0.7122266
2110	0,79927890	1,00000000	0.7517926	0.3689268	0.4356029	0.3778477	0.5171956	0.4676631	0.4469007	0.5661465	0.5149367	0.5322844	0.6074354
2400	0.9401559	0.7517926	1,00000000	0.7031231	0.7258986	0.6608905	0.8176956	0.7851808	0.7709317	0.8211116	0.7303493	0.7555599	0.8488799
2940	0.5614409	0.3689268	0.7031231	1,00000000	0.7870941	0.7879118	0.7656412	0.7739040	0.8737551	0.8430494	0.8177489	0.8443992	0.6943587
3450	0.5791229	0.4356029	0.7258986	0.7870941	1,00000000	0.9411812	0.9210170	0.9503902	0.8443927	0.8926297	0.7547143	0.7877488	0.7563751
3511	0.5192501	0.3778477	0.6608905	0.7879118	0.9411812	1,00000000	0.8920939	0.9382115	0.8547776	0.8424598	0.7761524	0.7973329	0.7526104
3540	0.6692899	0.5171956	0.8176956	0.7656412	0.9210170	0.8920939	1,00000000	0.9845790	0.9033048	0.9115780	0.8533182	0.8757793	0.8409045
3550	0.6314326	0.4676631	0.7851808	0.7739040	0.9503902	0.9382115	0.9845790	1,00000000	0.9074059	0.8933791	0.8258587	0.8479345	0.8267518
4215	0.6115355	0.4469007	0.7709317	0.8737551	0.8443927	0.8547776	0.9033048	0.9074059	1,00000000	0.8823712	0.8990791	0.9162061	0.8140824
4410	0.6994082	0.5661465	0.8211116	0.8430494	0.8926297	0.8424598	0.9115780	0.8933791	0.8823712	1,00000000	0.8755764	0.9022519	0.8150983
4530	0.5976261	0.5149367	0.7303493	0.8177489	0.7547143	0.7761524	0.8533182	0.8258587	0.8990791	0.8755764	1,00000000	0.9926909	0.7821424
4540	0.6223244	0.5322844	0.7555599	0.8443992	0.7877488	0.7973329	0.8757793	0.8479345	0.9162061	0.9022519	0.9926909	1,00000000	0.7908612
4650	0.7122266	0.6074354	0.8488799	0.6943587	0.7563751	0.7526104	0.8409045	0.8267518	0.8140824	0.8150983	0.7821424	0.7908612	1,00000000

Tab. 14 Correlation between every meteorological station for SSI

On Fig. 12 – 14 presented correlations between every drought indices, where was the best correlation between SPEI and SPI.

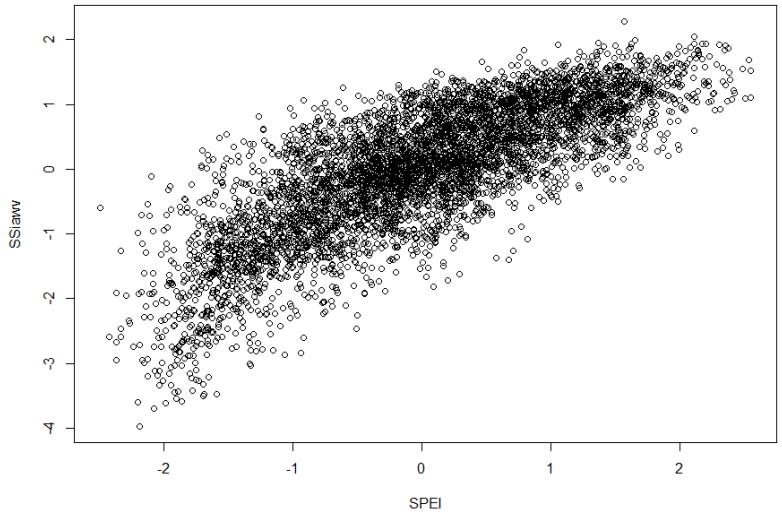

Fig. 12 Correlation between SSI and SPEI, on x axis is SPEI index, on y axis is SSI index

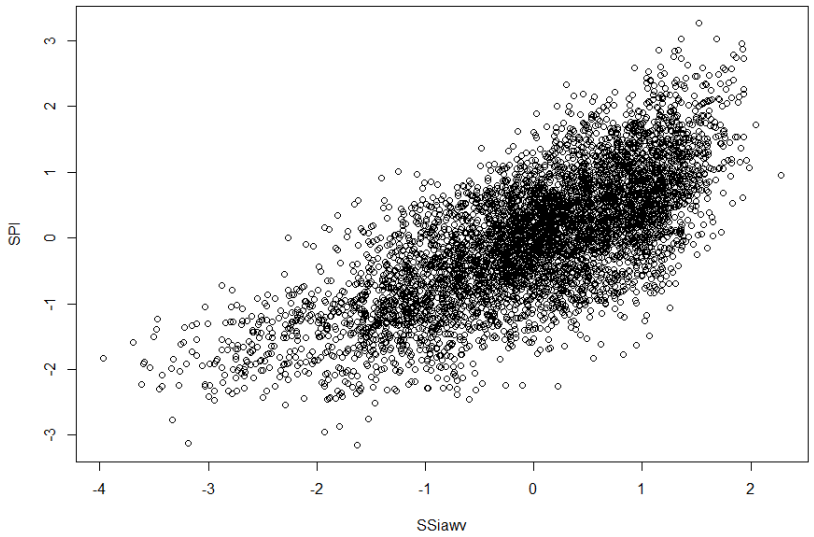

Fig. 13 Correlation between SPI and SSI, on x axis is SSI index, on y axis is SPI index

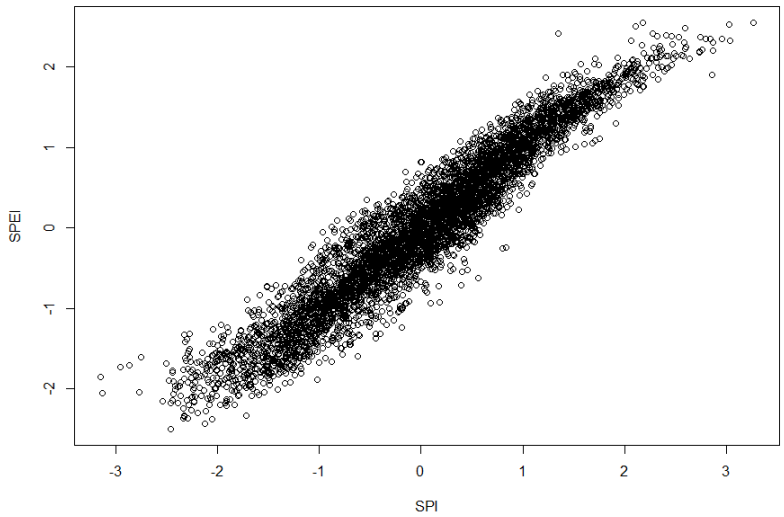

Fig. 14 Correlation between SPEI and SPI, on x axis is SPI index, on y axis is SPEI index

4.2 Neural Network Models

In my research, I used MLP with two hidden layers and one output layer (Fig. 15). MLP is one of the most popular ANN architecture. It consists of weighted connections and neurons arranged in layers. Each neuron collects the values from all of its input connections and produces a single output passing through an activation function. In practical usage, the MLPs are known for their ability to approximate non-linear relations. (Mareš T., 2012; Morid S., 2006; Savaci et al., 2005).

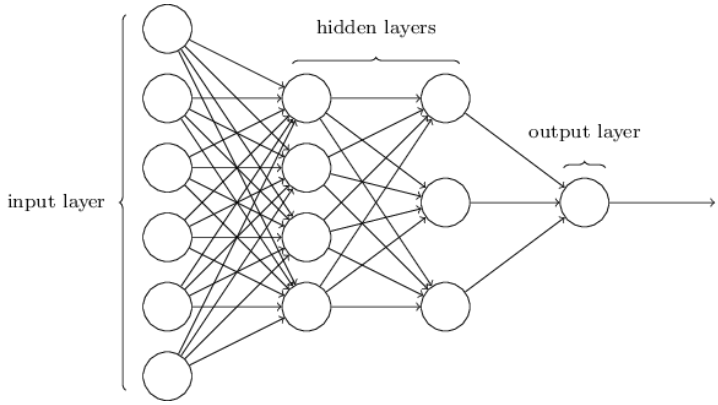

Fig. 15 Multilayer perceptron with two hidden layers (Nielsen M. et al., 2017)

For training, designing and using a neural network to classify I used the R package "AMORE". There is other way for creating a simple feedforward neural network with using package "neural". Comparing with the package "AMORE", the second one is rather slow (Wee-Jin Goh, 2006).

First, I had to prepare the input data. It means that I had to normalize the input data, so that all the inputs are at a comparable range. It is very important step, because it may lead to useless results or to a very difficult training process (Michy A., 2015).

Then I began design and train a Neural Network. Neural network has 10 iterations, due to initialization of weights. Number of iterations shows the number of times data passed through the algorithm, in case of ANN it means the "*forward pass*" and "*backward pass*". But there is a different between an epoch and an iteration. One epoch is a one forward or backward pass of all training examples. For 0,0001 global learning rate there are 5,000 tested training epochs, where best training epoch is 2, 252 (Kim et al, 2003; Morid el at., 2007).

Error criterium argument specifies the way you will determine, at each iteration, how close the network is to predicting its target. As the best option, I used LMS (for least mean squares) because it works well in many cases.

As an activation function of the hidden layer neurons I used a hyperbolic tangent sigmoid transfer function "*tansig*". For output layer, I used a transfer function "*purelin*".

Adaptive gradient descent, with the momentum term included ("*ADAPTgdwm*"), specifies solution strategy for converging on weights within the network (Wilson et al., 2001).

Before training the neural network, I had to divide data on two parts: first was for calibration, second for validation. Calibration is the process of estimating model parameters by comparing model predictions for a given observed dataset in same conditions. Once the training, optimisation, faze is completed, the performance of the trained network has to be validated. It is important to know, that the validation dataset not have been used as part of the training process. Validation involves running a model using input parameters measured or determined during the calibration process (Maier et al., 1999; Morasi et al., 2006).

Then I began training NN with backpropagation algorithm. Basically, this algorithm is based on error-correction learning rule. This error-propagation process consists of two passes through the different layers of the network. In the first pass, an input vector is applied to the neurons and its effect propagates through the network, layer by layer. A set of output is produced as the actual response of the network (Kim et al, 2003; Maier et al., 1998; Tsoukalas et al., 1997).

The process of optimising the connection weights is training algorithm. The weights are different in the hidden and output layers, and their values can be changed during the process of network training (Kim. et al., 2003).

After, I had to simulate the calibration and validation data with trained neural network. All the results of forecasting every drought indices are shown on Fig. 16 - 54.

Forecasting SPI, SPEI and SSI was based on past observed data of these drought indices for every meteorological station using MLP.

4.2.1 Forecasting SPI

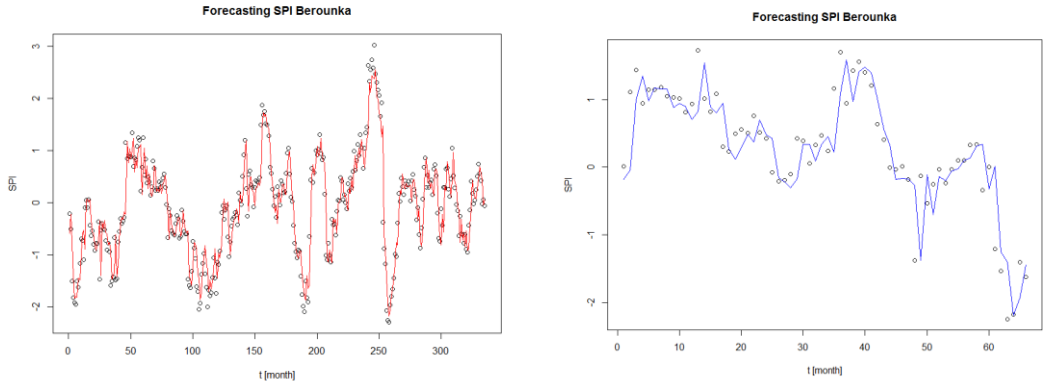

Fig. 16 Result of SPI forecasting using MLP, Berounka. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

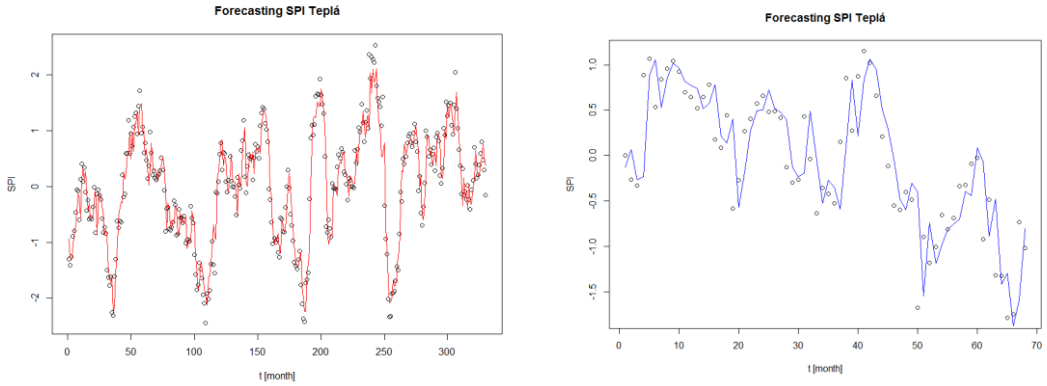

Fig. 17 Result of SPI forecasting using MLP, Teplá. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

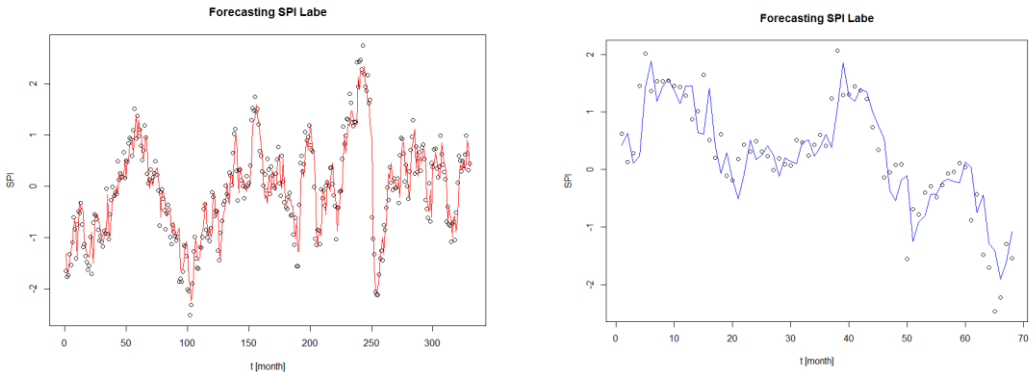

Fig. 18 Result of SPI forecasting using MLP, Labe. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

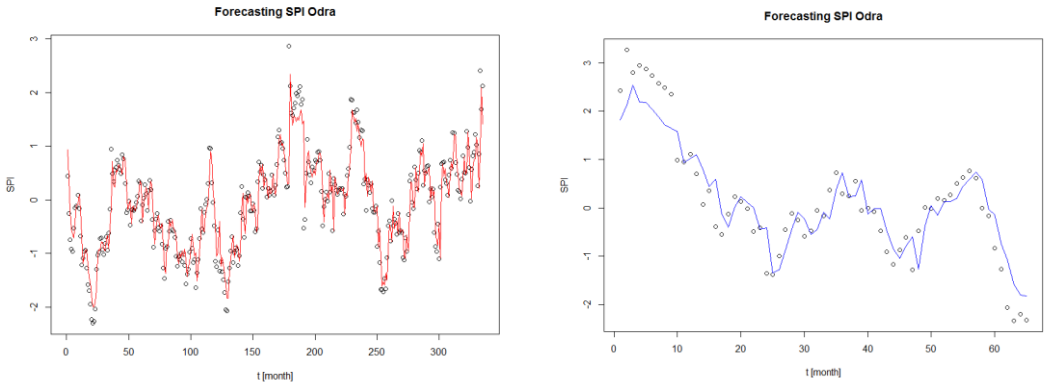

Fig. 19 Result of SPI forecasting using MLP, Odra. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

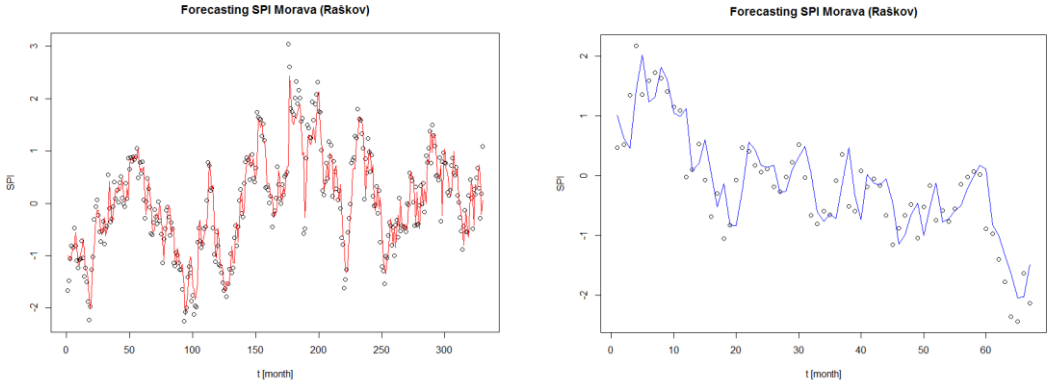

Fig. 20 Result of SPI forecasting using MLP, Morava (Raškov). On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

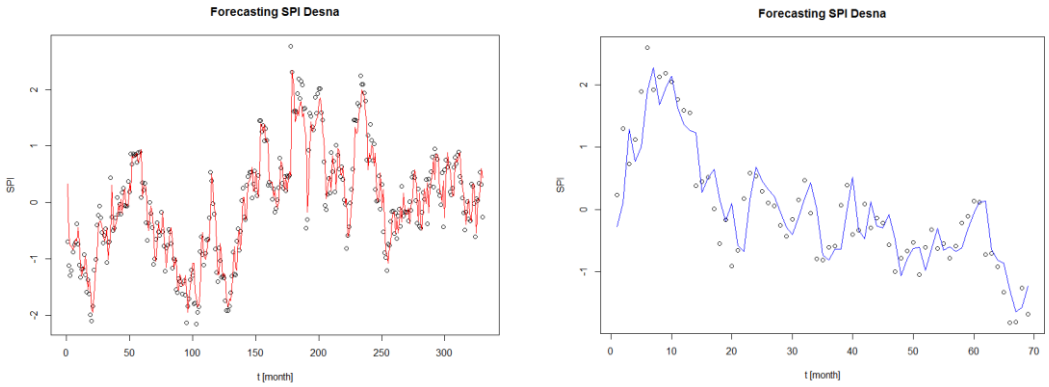

Fig. 21 Result of SPI forecasting using MLP, Desna. On x axis is time in month (396 month), on y axis is a range of SPI index (-2.14,2.76). The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

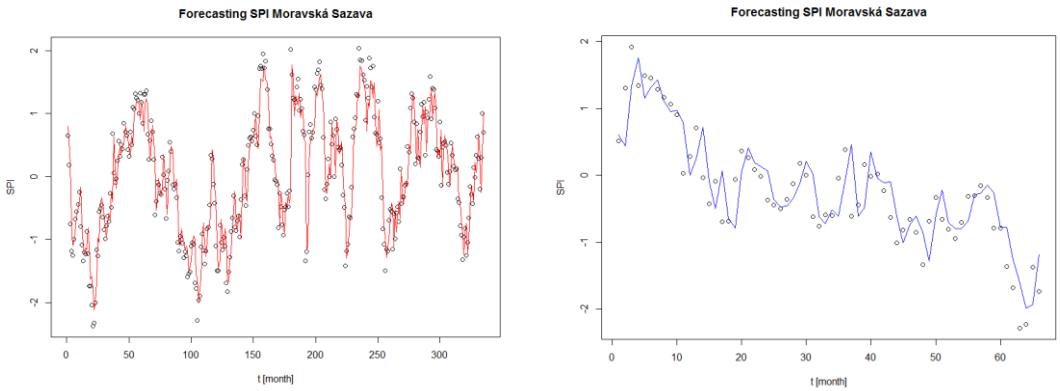

Fig. 22 Result of SPI forecasting using MLP, Moravská Sázava. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

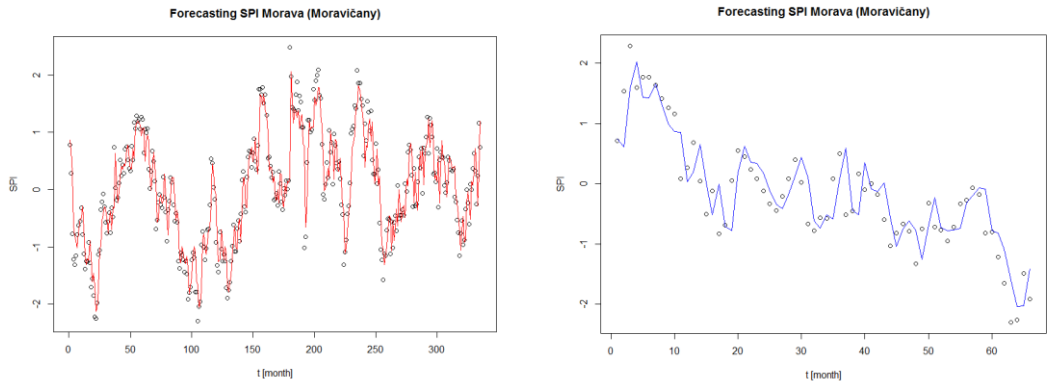

Fig. 23 Result of SPI forecasting using MLP, Morava (Moravičany). On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

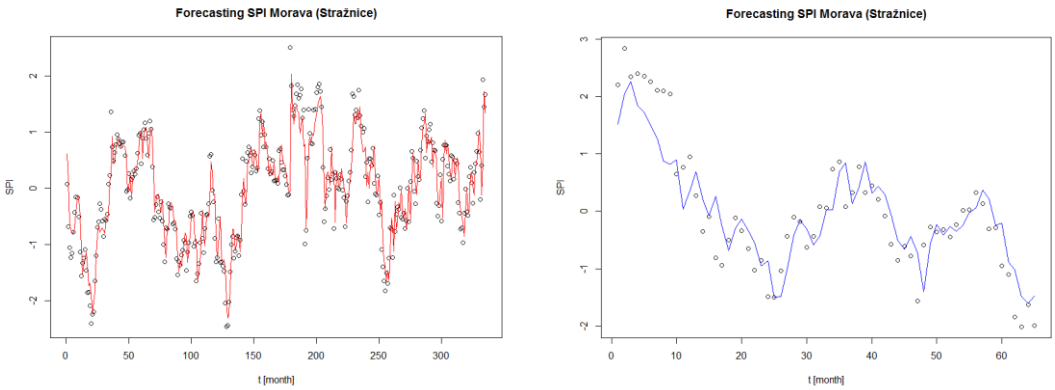

Fig. 24 Result of SPI forecasting using MLP, Morava (Stražnice). On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

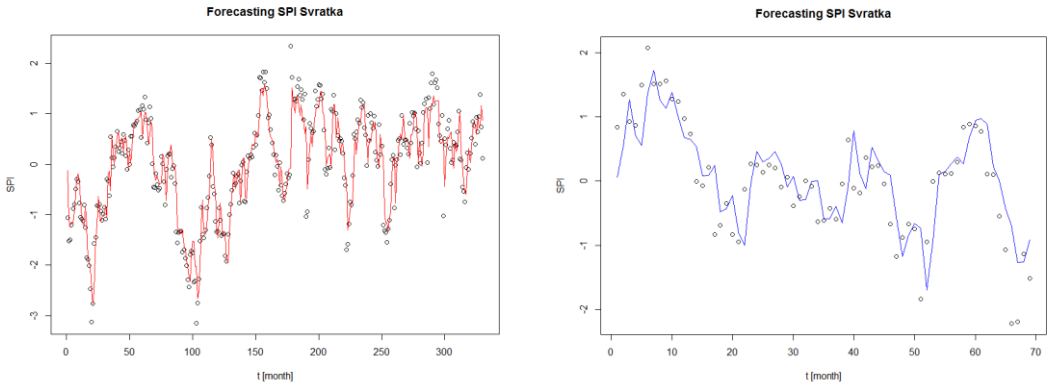

Fig. 25 Result of SPI forecasting using MLP, Svratka. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

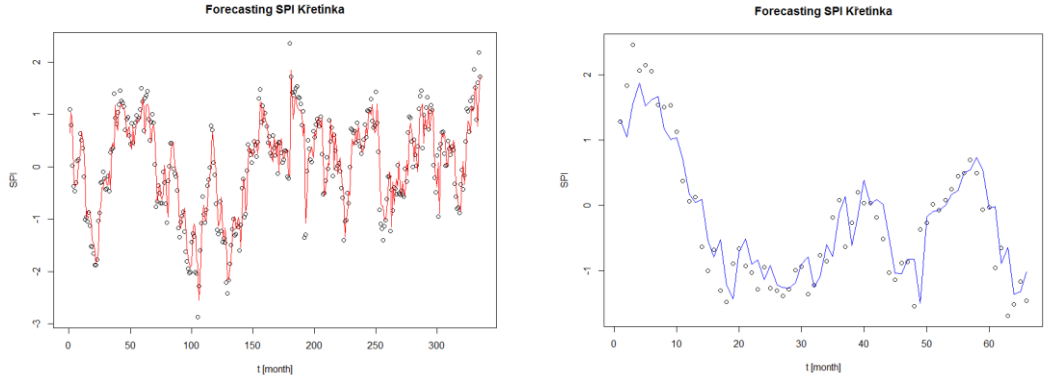

Fig. 26 Result of SPI forecasting using MLP, Křetínka. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

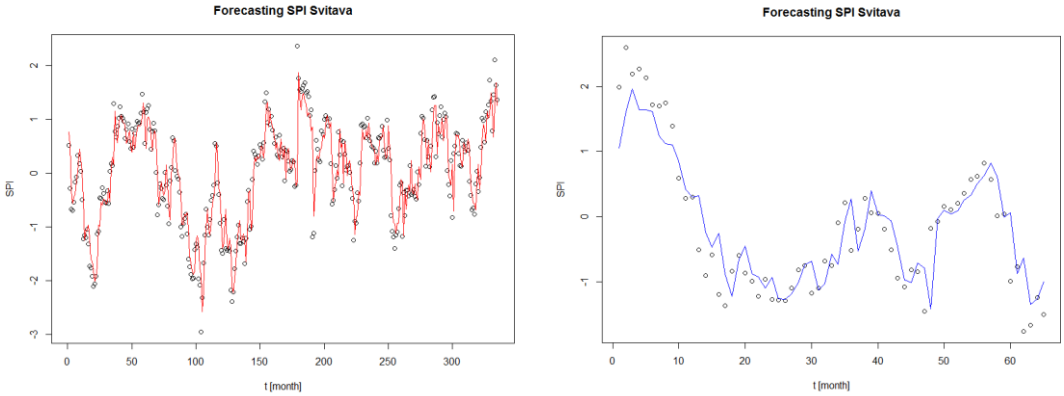

Fig. 27 Result of SPI forecasting using MLP, Svitava. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

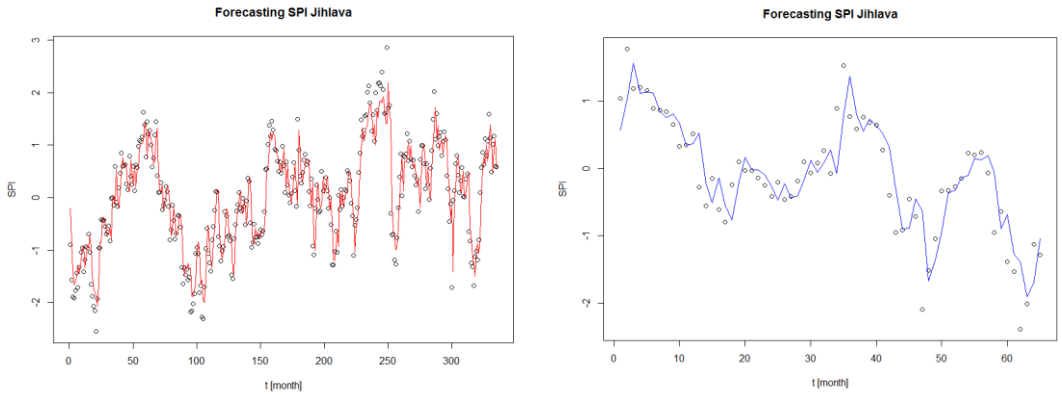

Fig. 28 Result of SPI forecasting using MLP, Jihlava. On x axis is time in month (396 month), on y axis is a range of SPI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

4.2.2 Forecasting SPEI

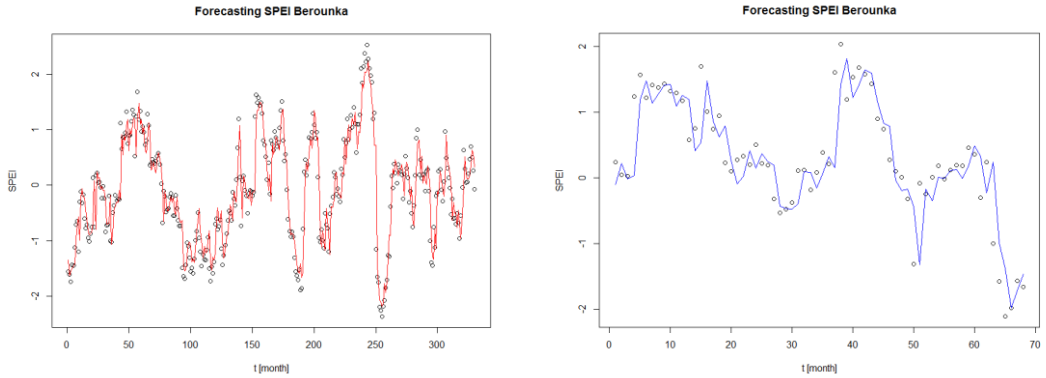

Fig. 29 Result of SPEI forecasting using MLP, Berounka. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

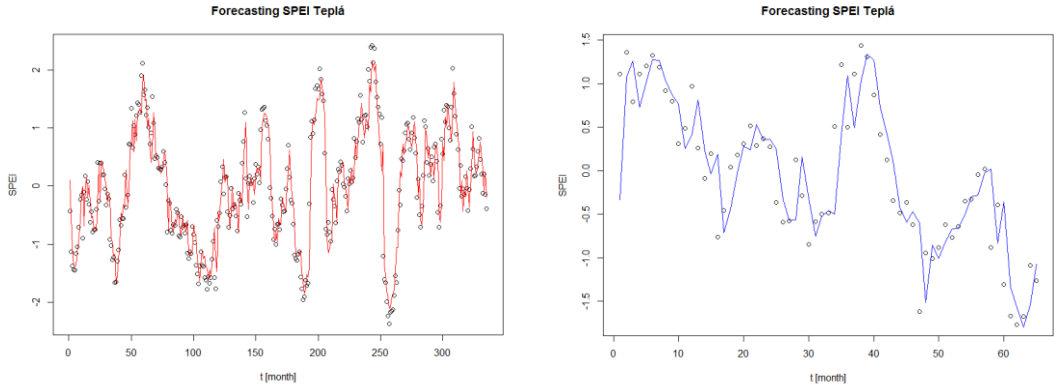

Fig. 30 Result of SPEI forecasting using MLP, Teplá. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

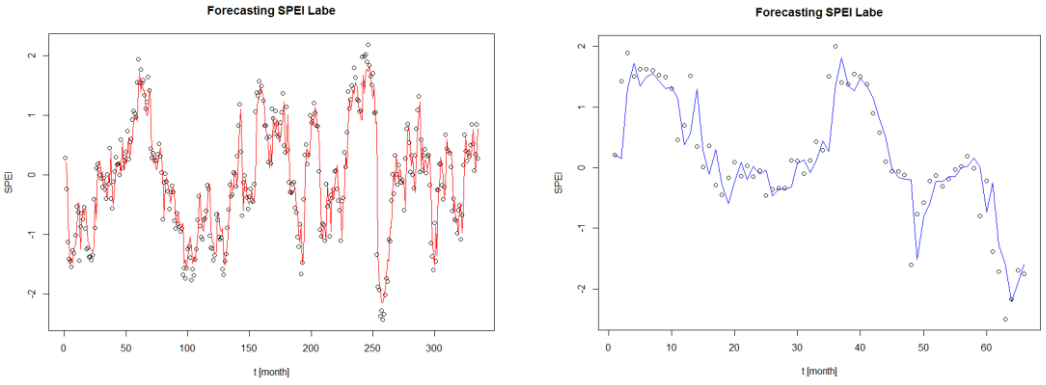

Fig. 31 Result of SPEI forecasting using MLP, Labe. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

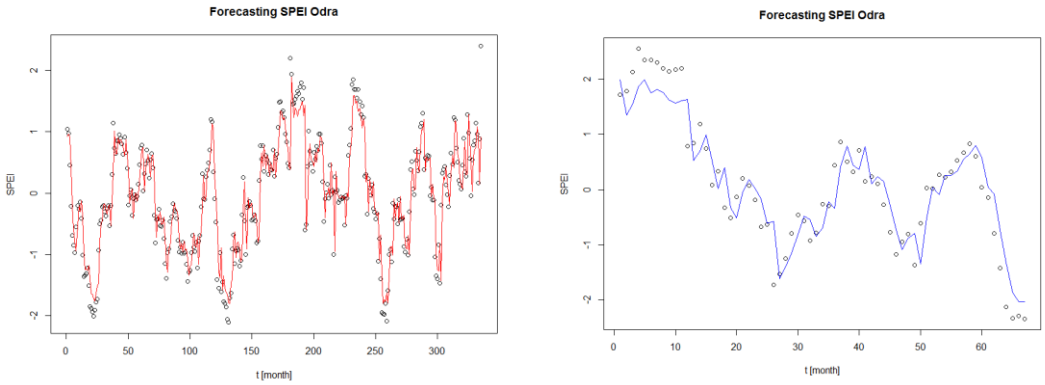

Fig. 32 Result of SPEI forecasting using MLP, Odra. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

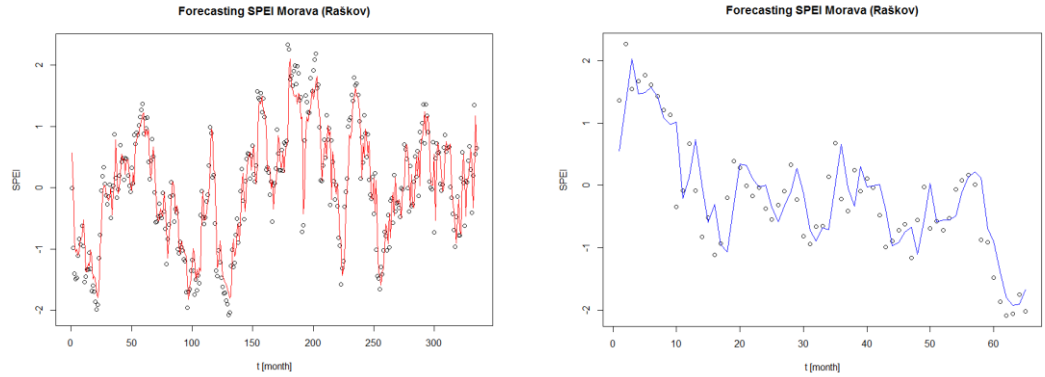

Fig. 33 Result of SPEI forecasting using MLP, Morava (Raškov). On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

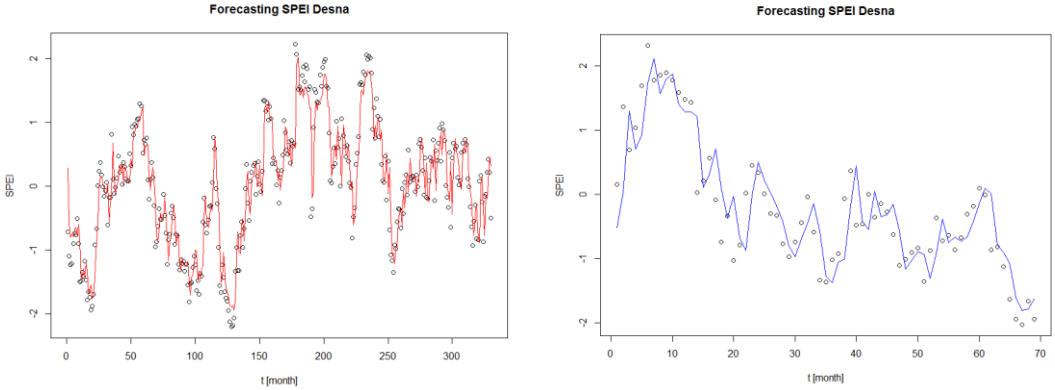

Fig. 34 Result of SPEI forecasting using MLP, Desna. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

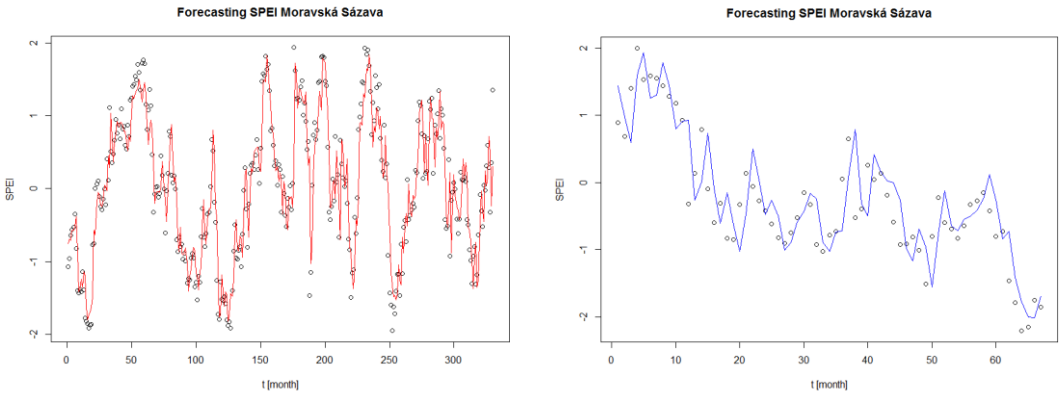

Fig. 35 Result of SPEI forecasting using MLP, Moravská Sázava. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

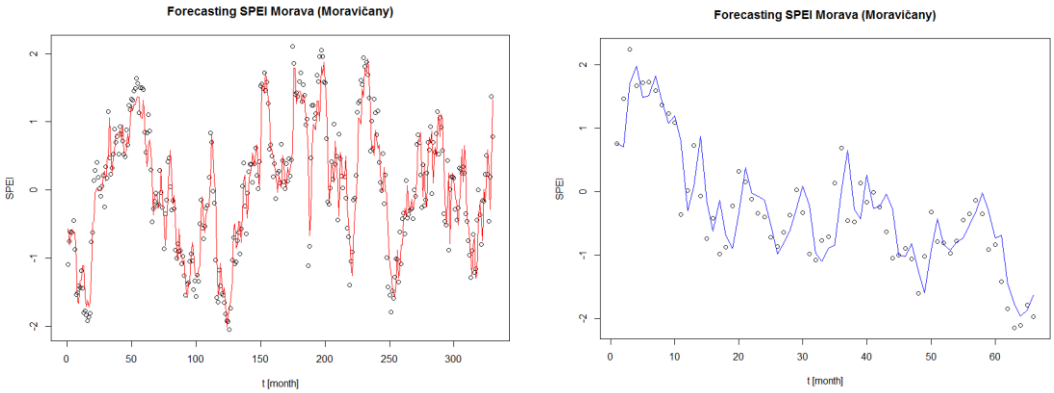

Fig. 36 Result of SPEI forecasting using MLP, Morava (Moravičany). On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

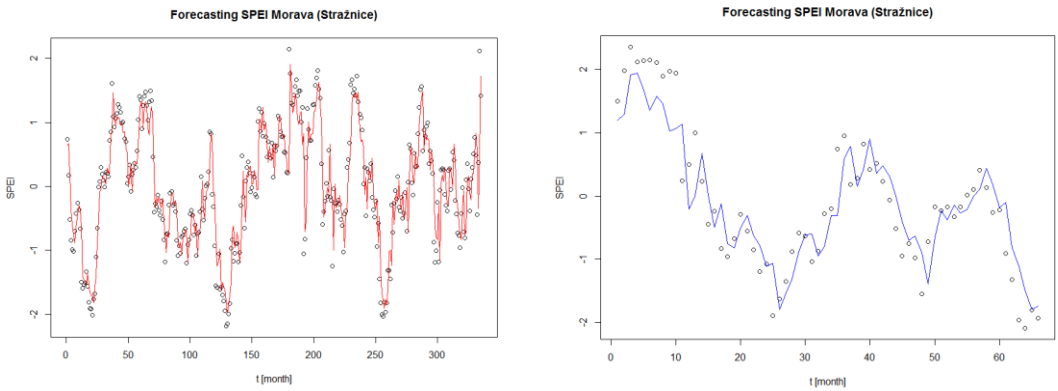

Fig. 37 Result of SPEI forecasting using MLP, Morava (Stražnice). On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

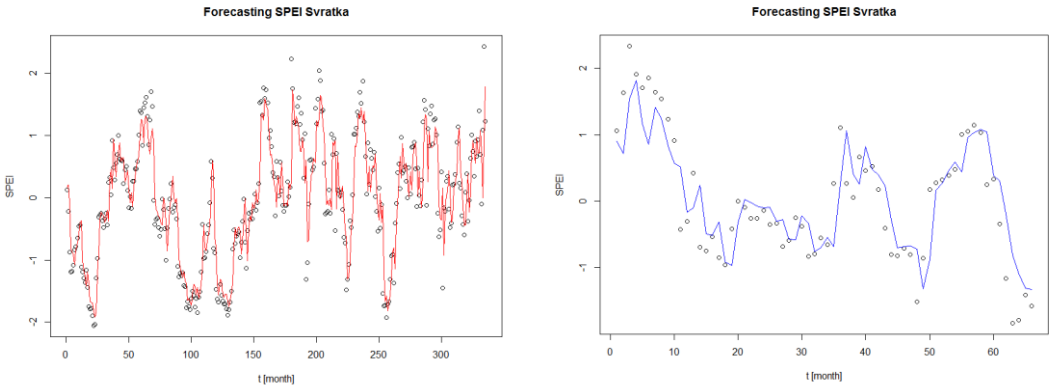

Fig. 38 Result of SPEI forecasting using MLP, Svratka. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

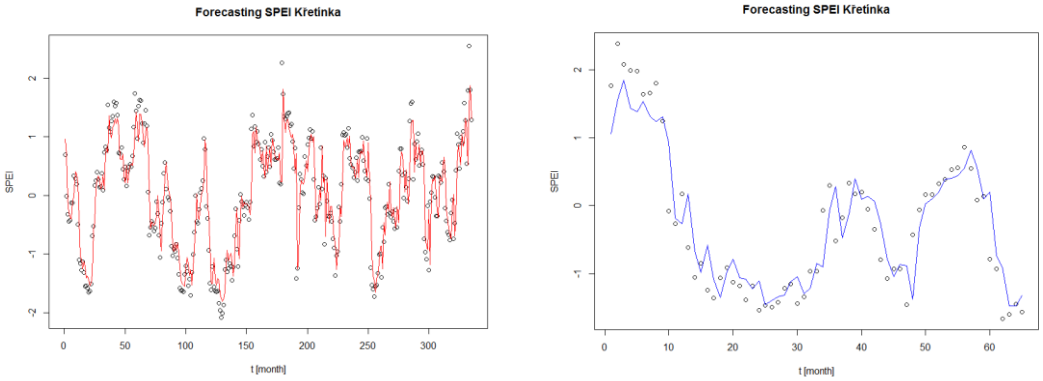

Fig. 39 Result of SPEI forecasting using MLP, Křetínka. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

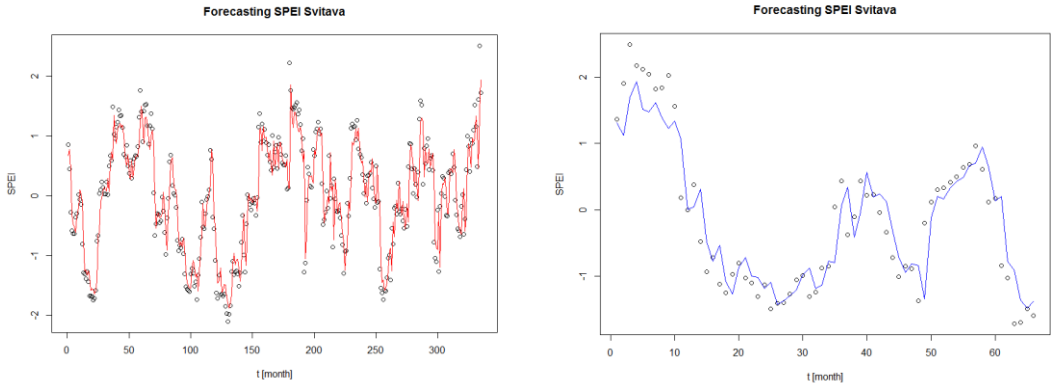

Fig. 40 Result of SPEI forecasting using MLP, Svitava. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

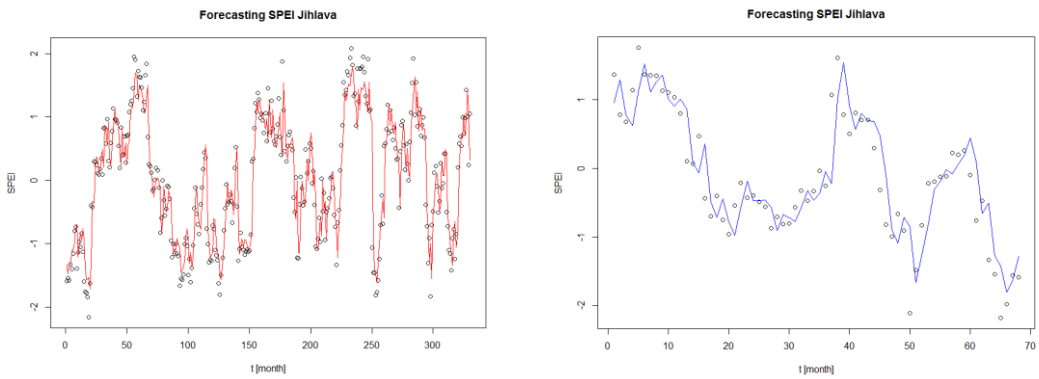

Fig. 41 Result of SPEI forecasting using MLP, Jihlava. On x axis is time in month (396 month), on y axis is a range of SPEI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

4.2.3 Forecasting SSI

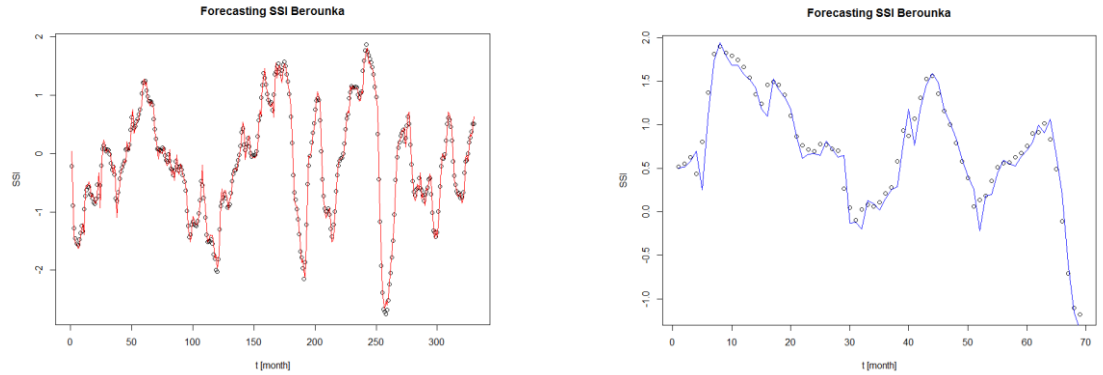

Fig.42 Result of SSI forecasting using MLP, Berounka. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

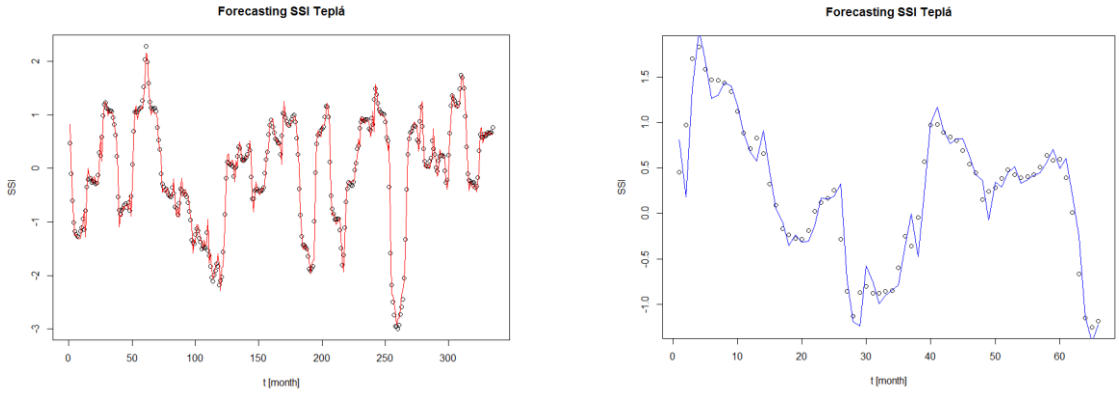

Fig. 43 Result of SSI forecasting using MLP, Teplá. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

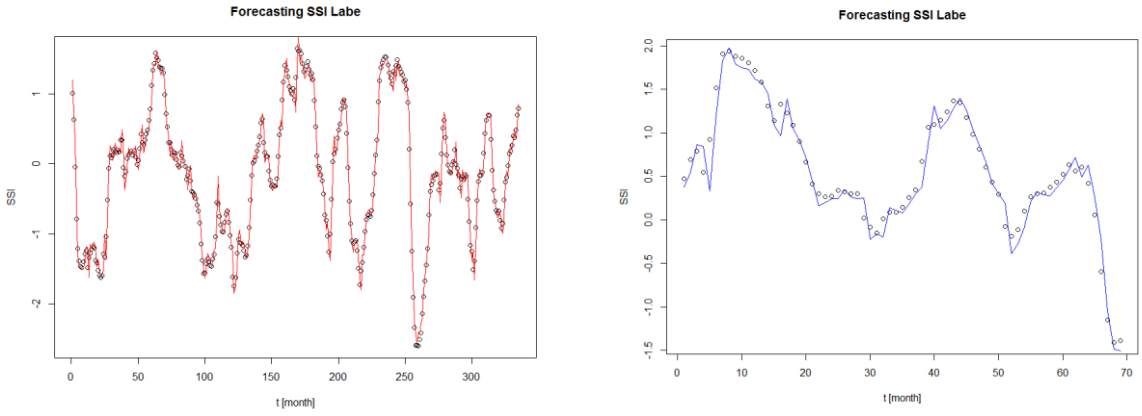

Fig. 44 Result of SSI forecasting using MLP, Labe. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

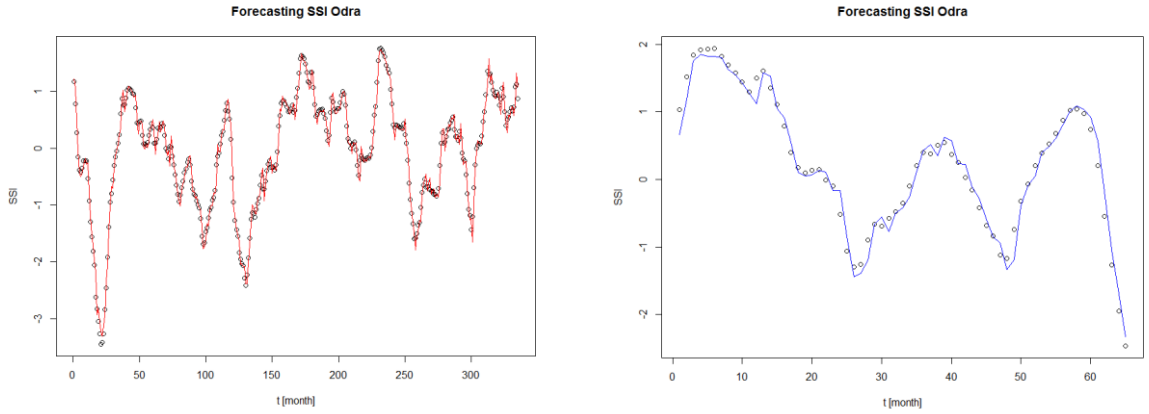

Fig. 45 Result of SSI forecasting using MLP, Odra. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

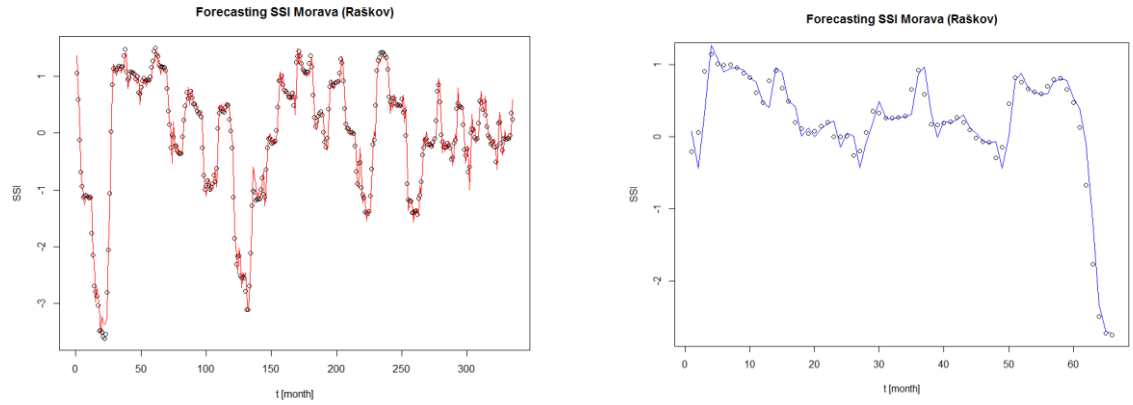

Fig. 46 Result of SSI forecasting using MLP, Morava (Raškov). On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

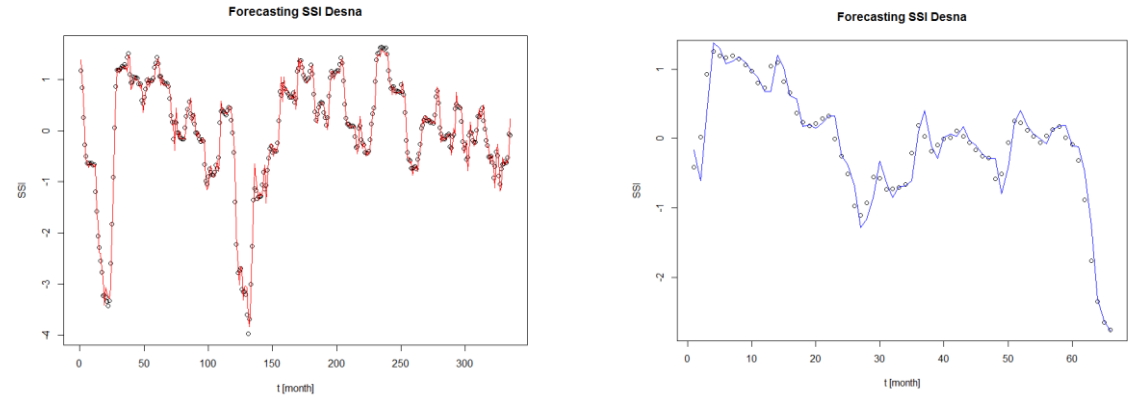

Fig. 47 Result of SSI forecasting using MLP, Desna. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

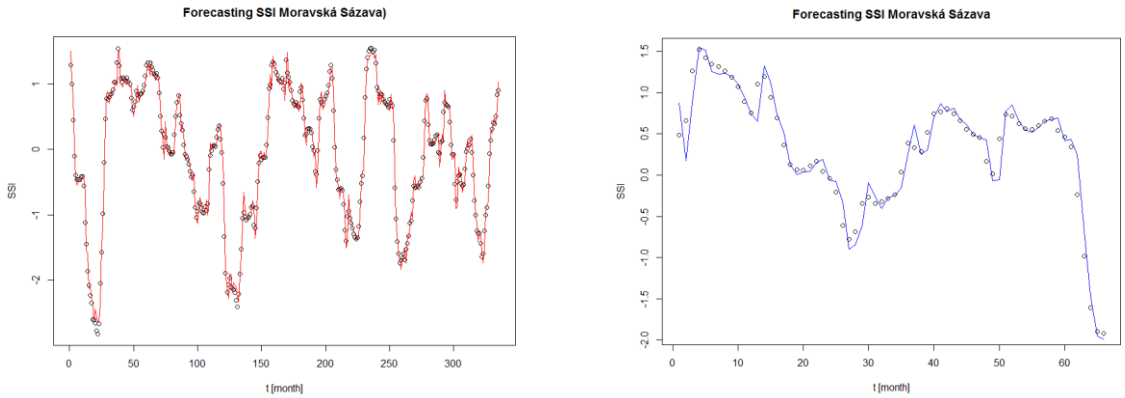

Fig. 48 Result of SSI forecasting using MLP, Moravská Sázava. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

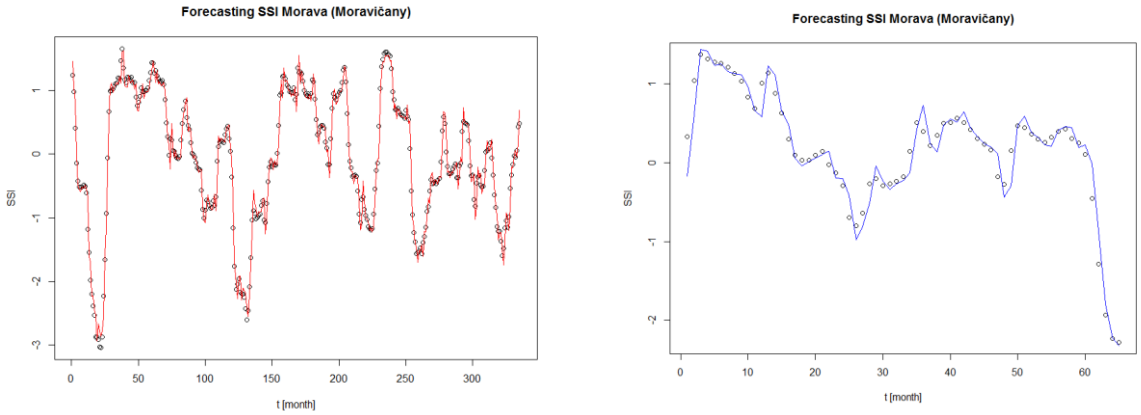

Fig. 49 Result of SSI forecasting using MLP, Morava (Moravičany). On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

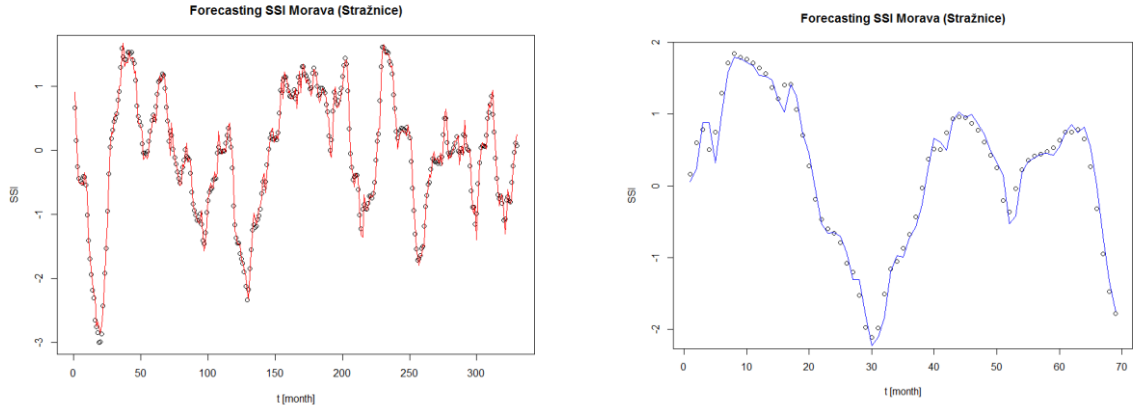

Fig. 50 Result of SSI forecasting using MLP, Morava (Stražnice). On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

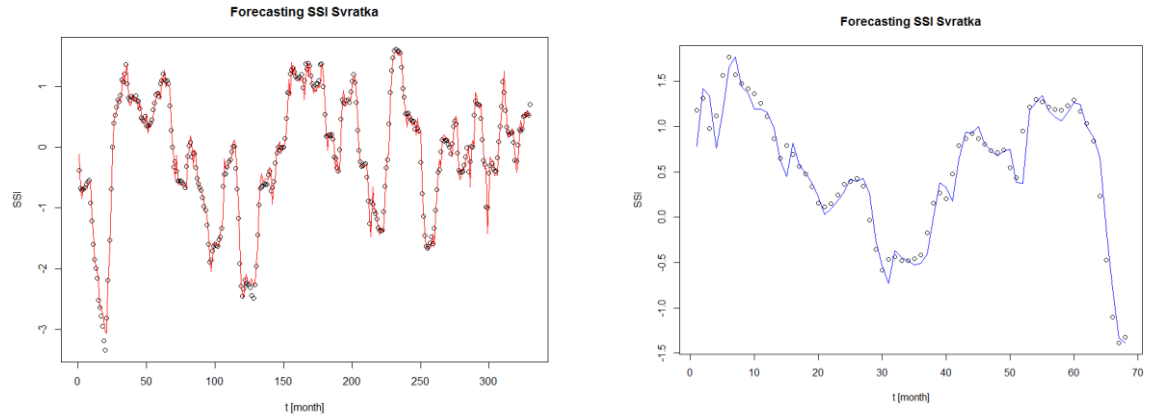

Fig. 51 Result of SSI forecasting using MLP, Svratka. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

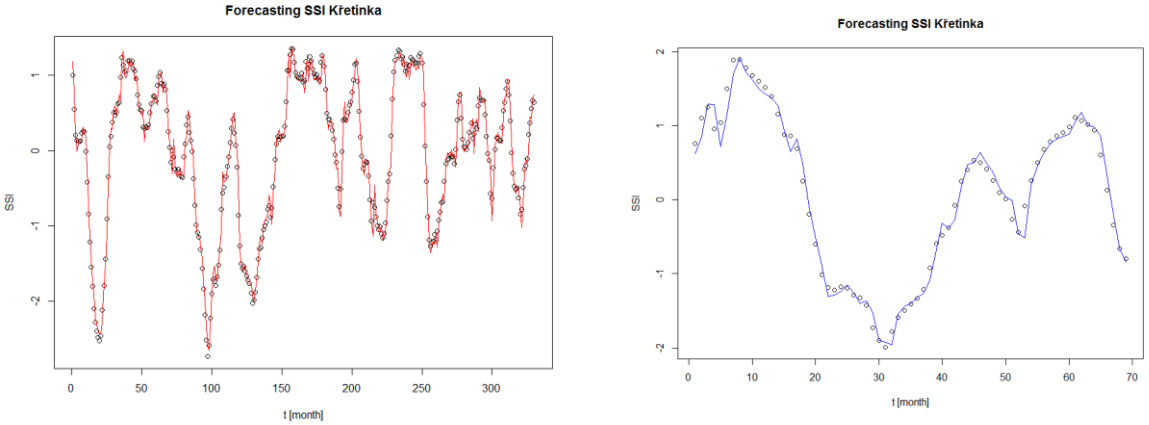

Fig. 52 Result of SSI forecasting using MLP, Křetínka. On x axis is time in month (396 month), on y axis is a range of SSI index.The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

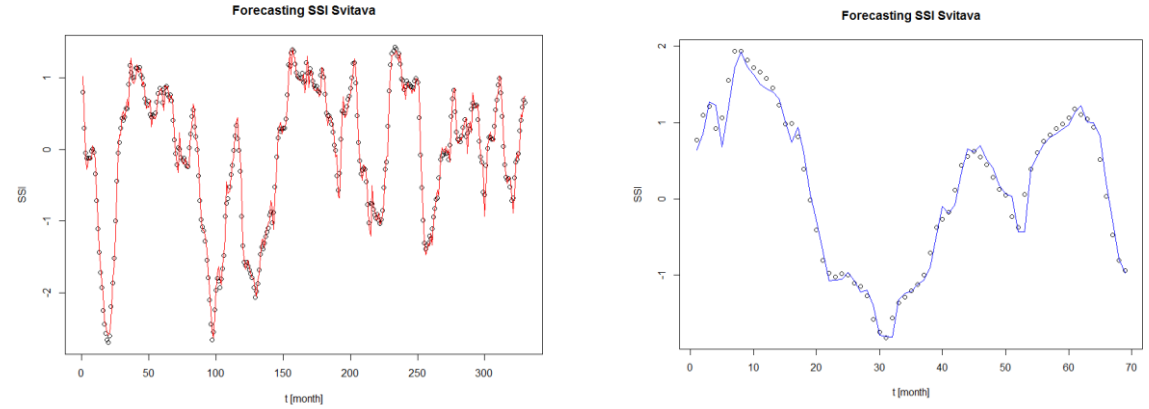

Fig. 53 Result of SSI forecasting using MLP, Svitava. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

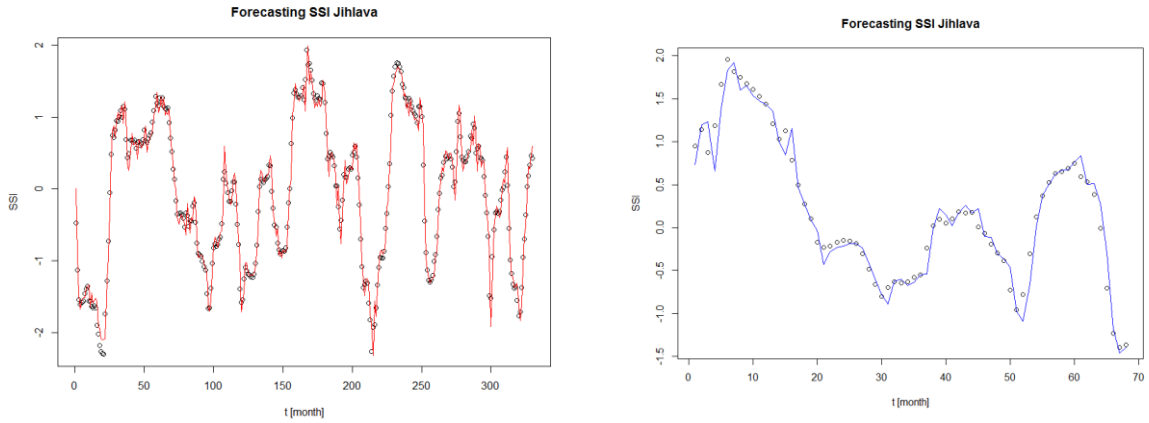

Fig. 54 Result of SSI forecasting using MLP, Jihlava. On x axis is time in month (396 month), on y axis is a range of SSI index. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

4.3 The Performance of ANN models

Then I must evaluate models with evaluation statistics. For calculating additional statistical analyses, I used "hydroGOF" package. Goodness-of-fit function for numerical and graphical comparison of simulated and observes time series, mainly focused on hydrological modelling: Mean Error (ME), Percent Bias (PBIAS), Root Mean Square to Standard Deviation (RSR), Nash-Sutcliffe Efficiency (NSE), Mean Absolute Error (MAE), Ratio of Standard Deviations (rSD), Modified Nash-Sutcliffe efficiency (mNSE), Relative Nash-Sutcliffe efficiency (rNSE), Index of Agreement (d), Modified Index of Agreement (md), Relative Index of Agreement (rd), Coefficient of persistence (cp), Kling-Gupta Efficiency (KGE), Volumetric Efficiency (VE), Coefficient of Determination (R^2).

Coefficient of determination (R2)

The coefficient of determination R^2 describe the degree of collinearity between simulated and measured data. In other words, it describes the proportion of the variance in measured data explained by the model. It ranges from -∞ to *1*. Typically, when the values are greater than *0,5*, it considered as acceptable (Maca P., 2015; Moriasi et al., 2006).

$$R^2 = 1 - \frac{\sum_{i=1}^{N}(Q_{obs} - Q_{sim})^2}{\sum_{i=1}^{N}(Q_{obs} - \bar{Q})^2} \quad (6)$$

$$\text{where } \bar{Q} = \frac{1}{N}\sum_{i=1}^{N} Q_{obs}$$

where Q_{obs} is the observation,

Q_{sim} is simulated value,

$\bar{Q}$ is the mean of observed data for the constituent,

N is a total amount of observations.

Mean Error (ME)

Mean error estimates the systematic errors. Usually when the values are close to *0*, it shows a perfect fit of model (Maca P. et al, 2015).

$$ME = \frac{1}{N}\sum_{i=1}^{N}(Q_{obs} - Q_{sim}) \quad (7)$$

Q_{sim} is simulated value,

Q_{obs} is the observation,

N is a total amount of observations.

Mean Absolute Error (MAE), Root Mean Squared Error (RMSE)

They are also called error indices, which usually used in model evaluation. It includes three indices: The Mean Absolute Error (MAE) and The Root Mean Squared Error (RMSE). They range from *0* to ∞, values of *0* indicates a perfect fit. MSE is a network performance function, it measures the performance of the network according to the mean squared errors. From the MSE we can create square root, the result is the RMSE. The Mean Absolute Error is used to measure how close forecast values are to the observed values. It is the average of the absolute errors (Belayneh A., 2013; Maca P. et al., 2015; Moriasi et al., 2006).

$$MAE = \frac{1}{N}\sum_{i=1}^{N} Q_{obs} - Q_{sim} \quad (8)$$

The Root Mean Squared Error (Belayneh A., 2013):

$$RMSE = \sqrt{\frac{SSE}{N}} \quad (9)$$

Where *SSE* is the sum of squared errors and *N* is the number of samples used, *SSE* is given by (Belayneh A., 2013):

$$SSE = \sum_{i=1}^{N}(Q_{obs} - Q_{sim})^2 \quad (10)$$

where Q_{obs} is the observation,

Q_{sim} is simulated value,

$\bar{Q}$ is the mean of observed data for the constituent,

N is a total amount of observations.

For better understanding, in Tab. 15 represented summary of selected functions which are usually using in hydrological modelling.

Tab. 15 Summary of selected functions

Criteria	*Range*	*The best*
ME	[-∞;+∞]	0
MAE	[0; +∞]	0
RMSE	[0; +∞]	0
R^2	[-∞; 1]	1

4.4 Analysis of models performance

I was tested out autoregressive model, which had a different number of length. LAG is effectively the lead time of the forecast, which varied from 1 to 12 months. Because it is a medium-range and the long-range forecasts that are critical for drought preparedness (Morid et al., 2007).

All the model evaluation statistics, described above, are represented in Tab. 16 - 18. On tab. 16. At calibration of SPI forecasting the highest coefficient of determination was in 2110 Teplá (Cihelny), the lowest was in 6 meteorological stations: 2940 Odra (Bohumín), 4215 Morava (Strážnice), 4410 Svratka (Borovnice), 4530 Křetínka (Letovice), 4540 Svitava (Letovice), 4650 Jihlava (Dvorce). At validation of SPI forecasting the highest coefficient of determination was in two meteorological stations: 2940 Odra (Bohumín) and 4540 Svitava (Letovice), the lowest was in 3: 1980 Berounka (Beroun), 3540 Moravská Sázava (Lupěné), 3450 Morava (Raškov).

Tab. 16 Results of SPI forecasting

DBC		Calibration				Validation			
	LAG [month]	ME →0	MAE →0	RMSE →0	R^2 →1	ME →0	MAE →0	RMSE →0	R^2 →1
1980	7	0,00	0,29	0,38	0,85	-0,02	0,32	0,44	0,73
2110	10	0,00	0,28	0,36	**0,88**	0,03	0,35	0,45	0,63
2400	10	0,00	0,31	0,40	0,82	-0,01	0,35	0,48	0,77
2940	8	0,00	0,28	0,39	0,81	0,00	0,43	0,54	**0,83**
3450	11	0,00	0,32	0,41	0,82	0,07	0,38	0,48	0,73
3511	9	0,00	0,29	0,39	0,84	0,04	0,34	0,43	0,81
3540	12	0,00	0,30	0,40	0,84	0,03	0,36	0,44	0,73
3550	4	0,00	0,29	0,39	0,84	0,04	0,37	0,45	0,77
4215	8	0,00	0,30	0,41	0,81	-0,04	0,43	0,55	0,78
4410	9	0,00	0,33	0,45	0,81	0,09	0,38	0,50	0,67
4530	7	0,00	0,31	0,42	0,81	-0,02	0,36	0,48	0,80
4540	8	0,00	0,31	0,42	0,81	0,03	0,34	0,45	**0,83**
4650	8	0,00	0,35	0,44	0,81	0,05	0,32	0,43	0,74

At calibration of SPEI forecasting the highest coefficient of determination was in 2110 Teplá (Cihelny), the lowest in four stations: 3450 Morava (Raškov), 4215 Morava (Strážnice), 4530 Křetínka (Letovice) and 4540 Svitava (Letovice). At validation of SPEI forecasting the highest coefficient of determination was in two meteorological stations: 4530 Křetínka (Letovice) and 2940 Odra (Bohumín), the lowest was in 2110 Teplá (Cihelny).

Tab. 17 Results of SPEI forecasting

DBC		Calibration				Validation			
	LAG [month]	ME →0	MAE →0	RMSE →0	R^2 →1	ME →0	MAE →0	RMSE →0	R^2 →1
1980	10	0,00	0,34	0,44	0,81	0,05	0,32	0,43	0,75
2110	8	0,00	0,31	0,40	**0,84**	0,01	0,34	0,45	0,71
2400	7	0,00	0,31	0,41	0,82	-0,02	0,33	0,47	0,78
2940	6	0,00	0,29	0,40	0,82	0.01	0,38	0,45	**0,86**
3450	8	0,00	0,35	0,44	0,80	0,05	0,37	0,47	0,75
3511	9	0,00	0,31	0,40	0,82	0,05	0,35	0,44	0,81
3540	11	0,00	0,32	0,42	0,82	0,05	0,35	0,44	0,77
3550	12	0,00	0,33	0,42	0,81	0,01	0,33	0,42	0,81
4215	7	0,00	0,31	0,42	0,80	-0,01	0,40	0,48	0,82
4410	7	0,00	0,34	0,46	0,78	0,03	0,38	0,49	0,73
4530	8	0,00	0.31	0,43	0,80	0,04	0,31	0,40	**0,86**
4540	7	0,00	0.31	0.42	0,80	0,01	0,33	0,45	0,85
4650	10	0,00	0,33	0,42	0,82	0,05	0,31	0,41	0,80

It is clear, that all statistics are better of SSI index. On Tab. 18 at calibration of SSI forecasting the highest coefficient of determination was in 7 stations, the lowest was just in 3550 Morava (Morvičany). At validation, the highest coefficient of determination was in two stations: 4530 Křetínka (Letovice) and 4540 Svitava (Letovice), the lowest at validation was in 2110 Teplá (Cihelny).

Tab. 18 Results of SSI forecasting

DBC	*LAG [month]*	*Calibration*				*Validation*			
		ME →0	*MAE →0*	*RMSE →0*	*R^2 →1*	*ME →0*	*MAE →0*	*RMSE →0*	*R^2 →1*
1980	9	0,00	0,10	0,13	**0,98**	-0,03	0,11	0,16	0,94
2110	7	0,00	0,13	0,18	0,97	-0,01	0,15	0,21	0,93
2400	9	0,00	0,11	0,14	**0,98**	-0,02	0,11	0,15	0,96
2940	8	0,00	0,10	0,13	**0,98**	0,00	0,13	0,17	0,97
3450	7	0,00	0,15	0,20	0,97	0,01	0,14	0,21	0,93
3511	7	0,00	0,13	0,18	0,97	0,01	0,14	0,20	0,94
3540	7	0,00	0,12	0,16	**0,98**	0,00	0,12	0,18	0,94
3550	8	0,00	0,12	0,18	0,94	0,00	0,12	0,18	0,94
4215	9	0,00	0,12	0,16	**0,98**	0,01	0,13	0,18	0,94
4410	10	0,00	0,13	0,17	0,97	-0,01	0,12	0,17	0,94
4530	9	0,00	0,10	0,14	**0,98**	0,00	0,10	0,14	**0,98**
4540	9	0,00	0,10	0,14	**0,98**	-0.01	0,02	0,15	**0,98**
4650	10	0,00	0,00	0,17	0,97	-0.01	0,11	0,16	0,96

4.5 The best ANN models

Monthly rainfall recorded at 13 meteorological stations had been used as the input data to train and test the neural network. The rainfall data recorded over the time 1982 – 2015.

For calculation SPI, SPEI and SSI was used the "SPEI" R package. For calculation SSI were used data with water content in layer 0 – 100 cm. The probability distribution was expressed using the three-parameters log - logistic probability distribution, the SPI probability distribution was calculated using the Gamma distribution. For SSI, the probabilities are obtained empirically using the empirical Gringorten approach.

In this thesis, drought conditions were predicted by attempting to forecast future SPI, SPEI and SSI values. Through testing various network and learning methodologies, the Feedforward Neural Network with standard back propagation algorithm was found to be the most suitable. Those values were calculated for a range of time lags from 1 to 12 months. A drought lag time parameter was introduced to quantify the time between the start of a moisture anomaly and the onset of drought.

The evaluation of the accuracy of model forecasts was carried out using Root Mean Square Error, Mean Average Error, Mean Error and Coefficient of determination.

So, I had to choose the best model according to LAG. For those 13 meteorological stations that I had, the best models for all drought indices are represented on Tab. 19. Coefficient of determination at SPI forecasting was the best at calibration when LAG equals 10 month.

Coefficient of determination at SPEI forecasting was the best at validation when the LAG equals 6 months.

SSI shows the best results of forecasting. All statistical analysis show very good simulation. When the LAG equals 9 two meteorological stations show at calibration and validation the closest to the perfect fit results

Tab. 19 The best models of SPI forecasting

			R^2	
	DBC	*LAG*	*Calibration*	*Validation*
SPI	2110	10	0,88	0,63
	2940	8	0,81	0,83
	4540	8	0,81	0,83
SPEI	2940	6	0,82	0,86
SSI	4530	9	0,98	0,98
	4540	9	0,98	0,98

The results show the mathematical similarity of values of SPI and SPEI drought indices. It can be seen, that SPI and SPEI produce very close values. Small differences

between two Standardised Indices of drought show the fact that the temperature trends were not visible.

The results indicate that soil moisture exhibits higher persistence that precipitation.

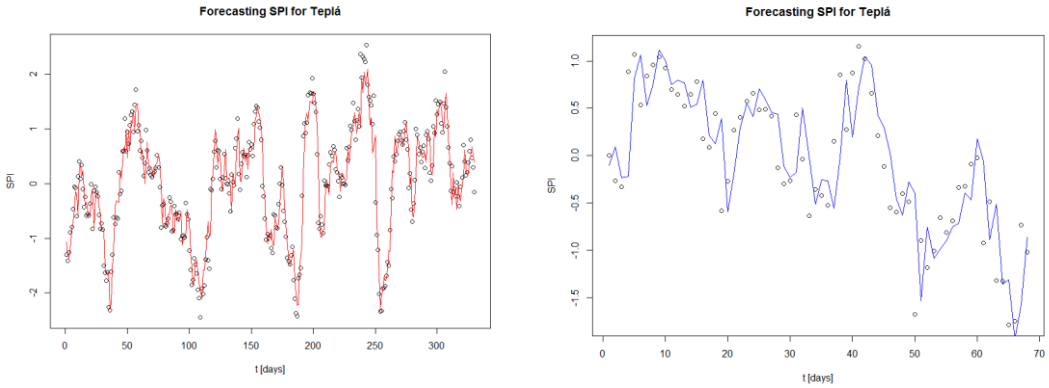

Fig. 55 Calibration and validation results of SPI forecasting for the best model according to LAG, Teplá. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

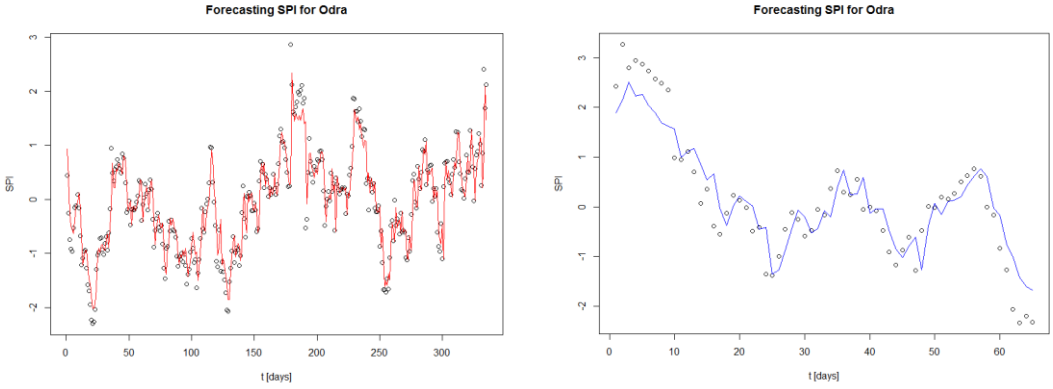

Fig. 56 Calibration and validation results of SPI forecasting for the best model according to LAG, Odra. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

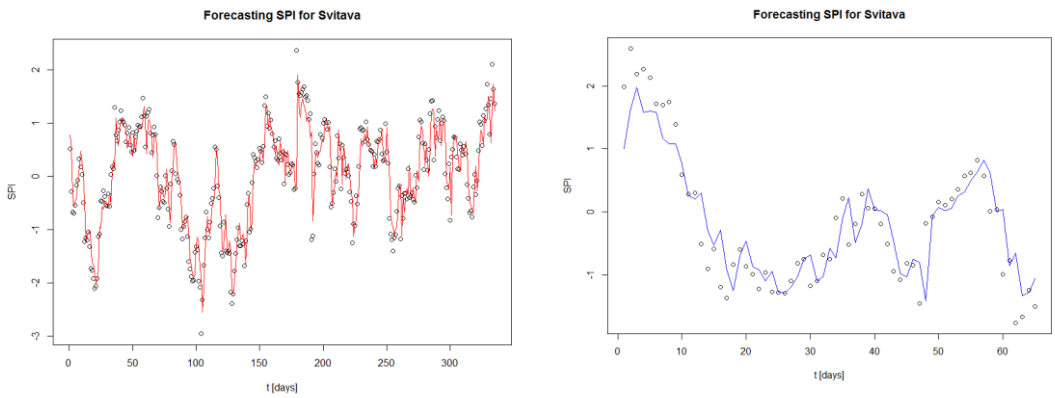

Fig. 57 Calibration and validation results of SPI forecasting for the best model according to LAG, Svitava. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

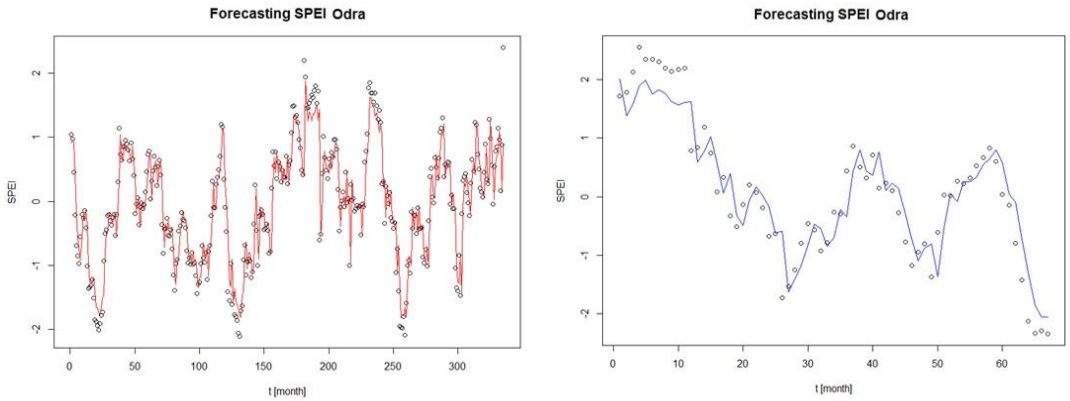

Fig. 58 Calibration and validation results of SPEI forecasting for the best model according to LAG, Odra. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

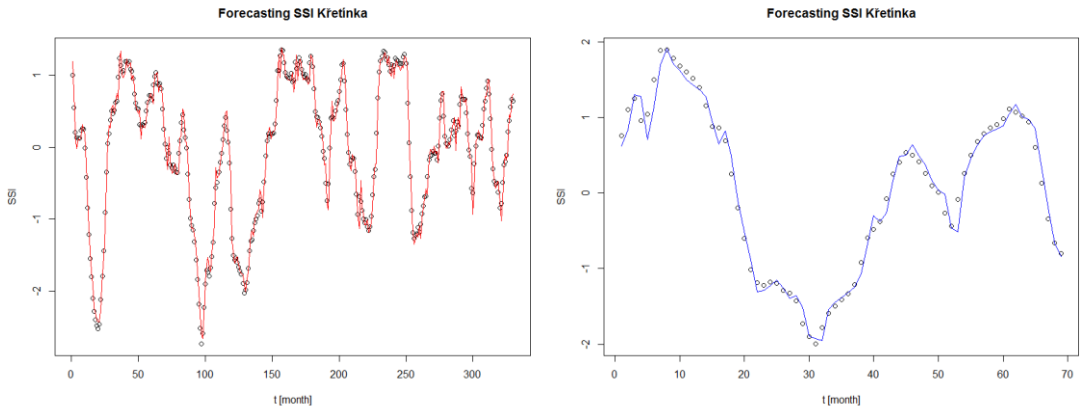

Fig. 59 Calibration and validation results of SSI forecasting for the best model according to LAG, Křetínka. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

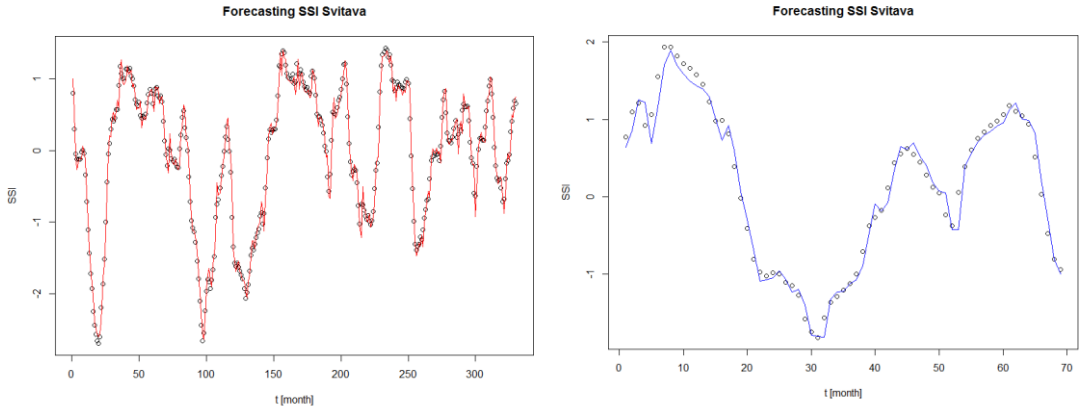

Fig. 60 Calibration and validation results of SSI forecasting for the best model according to LAG, Svitava. The red line: simulated calibration data from MLP model, the blue line: simulated validation data from MLP, points: observations

5 Discussion

Drought is hard to detect and describe. Most of the recent drought prediction models which are based on statistical or artificial neural network. Application of statistical models has a long history in drought forecasting. First who used Markov and regressions models for drought forecasting were Gabriel and Neumann (1962) and Torranin (1976) (Morid et al., 2007).

But nowadays new statistical technique called Artificial Neural Network show the superior performance for forecasting in many areas. A lot of studies presented their good results with using ANN. For example, Morid et al. (2007) used ANN for drought forecasting and time series of drought indices in the Tehran Province of Iran. In their study, they used two drought indices: The Standard Precipitation Index and The Effective Drought Index (EDI). As input data, they used monthly and daily rainfall data from meteorological stations from January 1970 to December 2000. Both of their best models have the coefficient of determination in range of values 0.66 – 0.79 for a lead time of 6 month. The EDI show better performance. In my case, the coefficient of determination was 0.98 for SSI index, which shows the superior results.

Sonnadara and Illeperuma (2009) used ANN to forecast drought in Sri Lanka. Predictions were made using the Standardized Precipitation Index as the drought monitoring index. In their study the highest correlation coefficient was found 0.94 for 3-month time window. As input data, they used the rainfall data from 13 climatological stations from 1870 to 1980. The results of his work show that neural network models trained on SPI can be used to forecast water scarcity.

Other example using artificial neural network approach for modelling rainfall-runoff due to typhoon by Che. Wang and Tsou (2013). They used rainfall data from a river basin of 27 typhoons between the years 2005 and 2009 in Taiwan. The feed forward back propagation network and conventional regression analysis were employed their performances were presented in their study. From the statistical evaluation, the coefficient of determination was 0.969 for feed forward back propagation network and 0.284 conventional regression analysis.

Zulifgar Ali et al., (2017) used ANN for forecasting SPEI. They used monthly time series data of Standardized Precipitation Evapotranspiration Index for 17 climatological stations: (Balakot, Kotli, Chirat, Chilas, Islamabad, Gupis, Pechawar, Saidu Shareef, Muzafarabad, Bunji, Di Khan, Drosh Gari Dubata, Dir, Gligit and Kakul) located in Northern Area and Pakistan including capital territory from 1975 to 2012. Time series data o observed SPEI with 1,3, 6 and 12-month time scales. The result of their work shows that ANNs have the power to capture the variation in selected drought indices with one month time scale.

There are a lot of other examples with using ANN for forecasting natural hazards. Kim and Valdes (2003) forecasted drought using dyadic wavelet transforms and neural network. Mishra et al. (2007) used SPI to compare the forecasting performance of ANN and linear stochastic model in the Kansabati River basin, India. Bacamli et al. (2009) investigated SPI and used Adaptive Neuro-Fuzzy Inference System for drought forecasting.

I'm sure that in near future using ANN in forecasting drought or other areas will develop more. People will create other indices for calculating drought. For example, in March 2017, Minister for Primary Industries Nathan Guy has welcomed a new tool to monitor drought in New Zealand's regions, The New Zealand Drought Index (NZDI). It combines four commonly-used drought indices: The Standardized Precipitation Index, Soil Moisture Deficit, Soil Moisture Deficit Anomaly and Potential Evapotranspiration Deficit. NZDI was created for showing levels of dryness and time when it turns into drought conditions (Guy N., 2017).

6 Conclusion

In recent years, a lot of studies were created to analyse the spatial patterns of drought risk and a big attention was drawing to this problem.

Regularly numerous number of natural indicators have to be controlled and monitored to determine the onset, ending and spatial characteristics of drought. Some of them must be valuated on frequent time steps (Wilhite D. et al., 2009).

Along with a drought effective drought early warning systems must integrate others climatic parameters such as snow pack, lake levels, soil moisture into a comprehensive assessment of current and future drought and water supply conditions (Wilhite D. et al., 2009).

In the first part of this thesis I describe the drought, its classifications and impacts of drought. Then I analysed three indices, which were developed to identify a drought. First of them, is The Standardized Precipitation Index, which based only on precipitation data. Second of them is The Standardized Precipitation Evapotranspiration Index. It is based on precipitation and potential evapotranspiration. And the last one was The Standardized Soil Moisture Index. Soil moisture is a main indicator for agricultural drought monitoring.

Then I describe Artificial Neural network and the procedure for drought forecasting using the MLP and its application in Czech Republic. Three drought indices – the SPEI, the SPI and SSI – have been used as the predictands. The SPI, SPEI and SSI neural network forecast was based on data obtained from January of 1982 to December of 2015 from 13 meteorological station situated in different parts of Czech Republic.

Tested Artificial Neural Network was Multilayer Perceptron with two hidden layers and was trained using Back-propagation algorithm. After evaluating the ANN models performance, the results of all three model performance indices were superior. The best results show The Standardized Soil Moisture Index.

Two of indices SPEI and SPI are using the precipitation data have some similarity, but just because they mathematically the same, just the SPEI include PET.

7 Bibliography

Ali Z., Hussian I., Faisal M., Nazir H.M., Hussian T., Shad M.Y., Shoukry A.M., Gani S.H., 2017: Forecasting Drought using Multi-Layer Perceptron Artificial Neural Network Model.

AghaKouchak A., 2014: A Baseline Probabilistic Drought Forecasting Framework Using Standardized Soil Moisture Index: Application to the 2012 United States Drought, Hydrology and Earth System Sciences.

Allen R.G., Periera L.S., Raes D., Smith M., 1998: Crop Evapotranspiration: Guidelines for Computing Crop Water Requirements Irrigation and Drainage Paper.

American Meteorological Society (AMS), 2004. Statement on Meteorological Drought. Bull. Am. Meteorol. Soc. 85, 771–773.

Asrari E., Masoudi M., 2014: A New Methodology for Drought Vulnerability Assessment Using SPI (Standardized Precipitation Index). 9 pp., 425-432.

Bacanli U.G., Firat M., Dikbas F., 2009: Adaptive Neuro-Fuzzy Inference System for Drought Forecastinh: Stachastic Environmental Research and Risk Assessment. 1143-1154

Beguería, S., and Vicente-Serrano, S. M. 2009: A Multiscalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index (SPEI), International Journal of Climatology 23.

Belayneh A., Adamowski J., 2012: Standard Precipitation Index Drought Forecasting Using Neural Networks, Wavelet Neural Networks, And Support Vector Regression.

Belayneh A., Adamowski J., Khalil B., Ozga-Zielinsky B., 2013: Long-term SPI Drought Forecasting in The Awash River Basin In Ethiopia Using Wavelet Neural Network And Wavelet Support Vector Regression Models.

Bishop C.M., 1995: Neural Networks for Pattern Recognition. Oxford University Press. ISNB 0198538642. 482 s.

Cancelliere A., G. Di Mauro, B. Bonaccarso, Rossi G., 2007: Drought Forecasting Using the Standardized Precipitation Index. Water Resource Manage 21:801-819.

Choi M., Jacobs J.M, Anderson M.C., Bosch D.D., 2012: Evaluation Of Drought Indices Via Remotely Sensed Data With Hydrological Variability, Journal of Hydrology.

Chen S.M., Wang Y.M., Tsou I., 2013: Using Artificial Neural Network Approach for Modelling Rainfall-Runoff Due to Typhoon.

Dai A., 2011: Drought Under Global Warming: A Review. Advanced Review. Volume 2.

Gabriel K.R., Neumann J., 1962: A Markov Chain Model for Daily Rainfall Occurrences at Tel Aviv. Quarterly Journal of The Royal Meteorological Society.

Gommes R., Petrassi F., 1994: Rainfall Variability and Drought in Sub-Saharan AfricaSince 1960. Agro-meteorology Series Working Paper 9.

Gordon N.D., McMahon T.A., Finlayson B.L., Stream Hydrology: An Introduction for Ecologists

Gringorten, I. I., 1963: A Plotting Rule for Extreme Probability Paper, J. Geophys. Res., 68, 813–814.

Guttman, N.B., 1998: Comparing the Palmer Drought Index and the Standardized Precipitation Index. Journal of the American Water Resources Association, 34:113–121.

Guttman, N.B., 1999: Accepting the Standardized Precipitation Index: A Calculation Algorithm. Journal of the American Water Resources Association, 35:311–322.

Hayes, M., M. Svoboda, Wall N., Widhalm M., 2011: The Lincoln Declaration On Drought Indices: Universal Meteorological Drought Index Recommended. Bulletin of the American Meteorological Society, 92(4):485–488.

Haykin S., 2008: Neural Network and Learning Machines. Third Edition.

Hollinger S.E., Kumar A., 1993: A New Soil Moisture Drought Index for Predicting Crop Yields In: Preprints, Eight Conf. on Applied Climatology.

Jacobi J., Perrone D., Duncan L. L., Hornberger G., 2013: A Tool for Calculating The Palmer Drought Indices, Water Resources Reserch, Vol 49, 6086-6089.

Juana J.S., Makepe P.M., Mangadi K.T., Naryana N., 2014: The Socio-Economic Impact Of Drought In Botswana, Volume 11, PP. 43-60.

Kim T.W., Valdes J. B., 2003: Nonlinear Model for Drought Forecasting Based On A Conjunction Of Wavelet Transforms And Neural Networks. Journal of Hydrologic Engineering, vol. 8.

Maca P., Pech P., 2015: Forecasting SPEI And SPI Drought Indices Using the Integrated Artificial Neural Networks.

Maca P., Pech P., Pavlasek J., 2014: Comparing the Selected Transfer Functions and Local Optimization Methods for Neural Network Flood Runoff Forecast, Hindawi.

Maier H. R., Dandy G. C., 1998: The Effect of Internal Parameters and Geometry On The Performance Of Back-Propagation Neural Networks: An Empirical Study. Environmental Modelling & Software 13: 193–209.

Maier H. R., Dandy G. C., 2000: Neural Networks for The Prediction and Forecasting of Water Resources Variables: A Review Of Modelling Issues And Application. Environmental Modelling & Software 15: 101–124.

Mares T., 2012: Artificial Neural Networks in Calibration of nonlinear models. 12 pp.

Mangiafico S.S., 2016: Summary and Analysis of Extension Program Evaluation In R., Descriptive Statistics.

McKee, T.B., N.J. Doesken and J. Kleist, 1993: The Relationship of Drought Frequency and Duration to Time Scales. Proceedings of the 8th Conference on Applied Climatology, 17–22 January 1993, Anaheim, CA. Boston, MA, American Meteorological Society.

Mishra, A. K., Singh, V. P., 2011: Drought modeling – A review. Journal of Hydrology. JUN 6 2011, 403, 1-2, s. 157–175.

Mishra A. K., Singh V. P., A., 2010: A Review of Drought Concepts, Journal of Hydrology 391 202 – 216.

Mishra A.K., Desai V., Singh V.P., 2007: Drought Forecasting Using a Hybrid Stochasting And Neural Network Model. Journal of hydrology, 626 – 638.

Mishra A.K., Desai V.R. 2006: Drought Forecasting Using Feed-Forward Recursive Neural Network. Ecologic Modelling 198.

Monteith J.L., 1965: Evaporation and Environment. In the State and Movement Of Water In Living Organism. 205-234.

Morasi D.N., Arnold J.G., Van Liew M.W., Binger R.L., Harmel R.D., Veith T.L., 2006: Model Evaluation Guidelines for Systematic Quantification of Accuracy in Watershed Simulations.

Morid S., Smakhtin V., Bagherzadeh K., 2007: Drought Forecasting Using Artificial Neural Networks and Time Series of Drought Indices, International Journal Of Climatology, Vol. 27, No. 15. 2103–2111, 2007.

Nicholson S.E., Davenport M.L., Malo A.R., 1990: A Comparison of The Vegetation Response to Rainfall in The Sahel And East Africa, Using Normalized Difference Vegetation Index From NOAA-AVHRR. Clim Change 17(2-3).

Obasi G.O.P., 1994: WMO ‘s Role in The International Decade For Natural Disaster Reduction. Bull. Am. Meteorol. Soc. 75 (1). 655-1661.

Palmer W.C., 1965: Meteorological Drought. US Department of Commerce, Weather Bureau.

Palmer W.C., 1968: Keeping Track of Crop Moisture Conditions, Nationwide: The New Crop Moisture Index.

Penman, H.L., 1948: Natural Evaporation from Open Water, Bare Soil and Grass. Proceeding of the Royal Society of London. Series A., Mathematical and Physical Scences, Vol 193, p. 120-145.

Peters E., Torfs, P.J.J.F., van Lanen H.A., Bier G., 2003: Propagation Of Drought In Groundwater In Semiarid And Sub-Humid Climatic Regimes, In: Hydrology In Mediterranean And Semiarid Regions: International Conference, Montpellier, France.

Pickup G., 1998: Desertification and Climate Change the Australian Perspective. Clim Res 11:51-63.

Prieto A., 1991: Artificial Neural Networks, Lecture Notes in Computer Science (540).

Rebetez M., Mayer H., Dupont O., Schidler D., 2006: Heat and Drought 2003 In Europe: A Climate Synthesis.

Rumelhart D.E., Hinton G.E., Williams R.J., 1986: Learning Internal Representations by Error Propagation.

Savaci A., 2005: Artificial Intelligence and Neural Networks, 14th Turkish Symposium, TAINN 2005 Izmir, Turkey, June 16-17, 2005, Revised selected Papers.

Sonnadara U., Illeperuna G., 2009: Forecasting Drought using Artificial Neural Networks. National Symposium on Disaster Risk Reduction and Climate Change Adaptation.

Stahl K., 2001: Hydrological Drought – a Study across Europe, Ph.D. thesis. Albert-Ludwigs Universitat.

Svoboda M., Fuchs B.A., World Meteorological Organization (WMO) and Global Water Partnership (GWP), 2016: Handbook of Drought Indicators and Indices. Integrated Drought Management Programme (IDMP), Integrated Drought Management Tools and Guidelines Series 2. Geneva.

Tallaksen, L. M., Van Lanen, H. A. J., 2004: Hydrological Drought: Processes and Estimation Methods For Streamflow And Groundwater, Developments In Water Science; 48, The Netherlands, Elsevier Science B.V., Amsterdam, The Netherlands.

Thornwaite C., 1948: An Approach Toward a Rational Classification of Climate. Geographical Review 38: 55 – 94.

Torranin P. 1976: Unpredictably of hydrological drought. In Proceeding of the Second International Symposium in Hydrology.

Tsoukalas L.H., Uhrig R. E., 1997: Fuzzy and Neural Approaches in Engineering.

United Nations. Economic and Social Commission for Western Asia (UNESCWA) 2005: Vulnerability of the region to socio-economic drought.

Vicente-Serrano S.M., Begueria S., Lopez-Moreno J.I., 2010: A Multi-Scalar Drought Index Sensitive to Global Warming: The Standardized Precipitation Evapotranspiration Index -SPEI.

Van Rooy M.P., 1965: A Rainfall Anomaly Index Independent of time and space. Notos 14.

Wee-Jin Goh, 2006: Creating A Simple Neural Network With AMORE.

Werbos P.J., 1974: Beyond Regression: New Tools for Prediction and Analysis In The Behavioural Sciences. PhD thesis, Harvard University.

Wilhite D., Glanty M., 2009: Understanding the Drought Phenomenon: The Ride of Definitions. Water International 10(1985) 111-120.

Wilson D.R., Martinez T.R., 2001: The Need For Small Learning Rates On Large Problems.

Zotarelli L., Dukes M.D., Romero C.C., Migliaccio K.W., Morgan K.T., 2010: Step by Step Calculation Of The Penman-Monteith Evapotranspiration (FAO-56 Method), IFAS Extension, University of Florida.

8 Internet resources

CHMI, Czech Hydrometeorological Institute, ©2015: Drought in the Czech Republic in 2015. A preliminary summary. Available from <http://portal.chmi.cz/files/portal/docs/meteo/ok/SUCHO/zpravy/en_summary_drought2015_ol.pdf>

Chrislb, 2005: Single Layer Neural Network. Available from: <https://commons.wikimedia.org/wiki/File:SingleLayerNeuralNetwork.png>

DeepLearning 0.1 Documentation, 2017: Multi Layer Perceptron. Available from <http://deeplearning.net/tutorial/mlp.html#mlp>

Guy Nathan, 2017: New drought measurement index launched. Available from <http://business.scoop.co.nz/2017/03/24/new-drought-measurement-index-launched/>

Leverington D., 2009: A Basic Introduction to Feedforward Backpropagation Neural Network. Available from <http://www.webpages.ttu.edu/dleverin/neural_network/neural_networks.html>

Michy A., 2015: Fitting a neural network in R. Available from: <https://www.r-bloggers.com/fitting-a-neural-network-in-r-neuralnet-package/>

NCAR, National Centre for atmospheric research, Climate Data Guide, 2017: Standardized Precipitation Evapotranspiration Index (SPEI). Available from <http:1url.cz/PtVgt>

Nielsen M., 2017: Neural Network and Deep Learning. Available from <http://neuralnetworksanddeeplearning.com/chap1.html>

NOAA, National Centres For environmental information, 2016: Potential Evapotranspiration. Available from: <https://www.ncdc.noaa.gov/monitoring-references/dyk/potential-evapotranspiration >.

UCAR, 2017: The Standardized Precipitation Evapotranspiration Index (SPEI). Available from <https://climatedataguide.ucar.edu/climate-data/standardized-precipitation-evapotranspiration-index-spei>

UNISEF, ©2008: Drought. Czech Republic. Available from <http://www.un.org/esa/agenda21/natlinfo/countr/czech/drought.pdf>