Examensarbete
TVVR 12/5013

Long term variability of Swedish river discharge as represented by EC-Earth in the past and future climates

David Corral Díaz
Olena Siergieieva

Division of Water Resources Engineering
Department of Building and Environmental Technology
Lund University

Avdelningen för Teknisk Vattenresurslära
TVVR-YY/5013
ISSN-1101-9824

Long term variability of Swedish river discharge as represented by EC-Earth in the past and future climates

David Corral Díaz
Olena Siergieieva

ACKNOWLEDGEMENT

This study was conducted at the Division of the Water Resources Engineering at Lund University, and is a part of the project Hydroimpacts 2.0, SMHI.

This thesis would not have been possible without guidance and help of several people who in different ways made their own valuable contribution into this work.

We would like to express our gratitude to our supervisor Cintia Bertacchi Uvo for her passion and wisdom in guiding us through this work, for her support and unique ability to find necessary words and give us further inspiration and enjoyment from the work. Special thankfulness we address to our co-supervisor Kean Foster for his help, necessary assistance, positive thinking and ability to make complex problems unbelievably simple. We would like to thank Grigory Nikulin for providing us with data, helpful explanations and suggestions within our work.

We are very grateful to our families, who were always supporting and encouraging us with their wise advices and warm wishes.

ABSTRACT

The aim of this study is to identify changes in pattern of long-term variability of seasonal discharge for the period 1851-2100 among four selected Swedish rivers based on output data from EC-Earth climate model.

In order to achieve the main goal and be able to analyze and compare changes in patterns of variability of water discharge for the past time period 1851-2005 and future time-period 2006-2100, discharge peaks should be represented as oscillations of frequency. This was done by means Fourier and wavelet analysis. Since data, obtained from the climate model are large-scale atmospheric variables with different units and magnitudes, they are incomparable with small-scale river discharge data. As consequence data need to be downscaled and standardized, it was done by using Singular Value Decomposition (SVD).

Results imply that natural variability becomes weaker in the future and is affected by radiative forcing. According to the model results, long-term variability is expected to increase in the future, whereas short-term variability tends to disappear and becomes weaker with higher RCP. For the past period model simulations showed that long-term variability tends to decrease in frequency towards lower latitudes, whereas frequency of short-term variability is increasing in the South compared to the North of Sweden.

TABLE OF CONTENT

TABLE OF FIGURES

INTRODUCTION

Nowadays climate change and global warming are strongly discussed topics. The fact that the climate has always changed and will continue to change due to natural factors is well known. These changes exist on a short term so we have seasons on the Earth and on long term that causes stadials (periods of low temperatures) and interstadials (warm periods). However current scientific discussions around the world about climate change refer these facts to human activity. Humans produce too much greenhouse gases, such as methane and nitrous oxide, together with sulfate aerosol emissions and changes of the land use causes changes in the Earth's atmosphere temperature. It has been evaluated by many scientists that CO_2 level could be doubled by 2100 compared to 1990 if the rate of emissions to the atmosphere is not reduced (Loaiciga et al., 1996). Therefore estimation of causes and results of possible climate changes currently became a dominant climate research field.

For instance Hay and McCabe (2010) in their research made general estimation of the monthly water balance for the Yukon River Basin, where rise of the annual discharge was predicted. This was caused mainly by an increment of the precipitation rate and by comparatively small effects of changes in temperature. Houghton et al. (1990; 1992; 1995) as cited in Xu (1999) conducted another research, where an assessment of the raise in temperature for the Northern Hemisphere and changes in precipitation rate by 15% under doubled CO_2 conditions is given.

One of the important aspects in this problem is water resources, or to be more accurate: effect of climate change on water availability in the future.

Rise of temperature will change water cycle and as a result higher losses of water through evaporation and increase of atmospheric water vapor content will occur, changes in the rate of precipitation and discharge, reduction of snow cover etc., which is valid only for temperature and sub-arctic climate (Bates et al., 2008). This in turn will significantly affect flow rates in rivers. Flood periods will move from spring to early spring/late winter. Such roughly predictable changes in the river flows have big influence on water distribution, energy use, drainage systems, irrigation, fishing and wildlife management. Extremely high importance of water for society and for wildlife does not allow delaying in predictions of possible consequences. These effects are widely known theoretically based on the previous knowledge and basic principles of water cycle, and they have been investigated experimentally for the short-time scale to a certain extend.

For instance Foster and Uvo (2010) studied seasonal and inter-annual stream-flow predictability for several rivers in Norway. Research was conducted by means of multi-model downscaling approach for producing forecast four to five months in advance.

However, there is still a high need to investigate experimentally possible consequences of the climate change also in a long-term scale.

Global climate models can be used to simulate present climate conditions and predict future changes in a long-term scale. In this project data from the global climate model EC-Earth are used. In comparison with more traditional coupled atmosphere-ocean general circulation models (AOGCMs) the EC-Earth model includes additional climate components such as ocean biochemistry, dynamic vegetation, atmospheric chemistry, carbon cycle components, and dynamic ice sheet. These additions will most likely improve sensitivity of the model and will expand accurate knowledge of the main climate feedbacks (Bintanja, 2011).

It is well-known that future climate is very dependent on the past climate. Therefore, to make predictions of EC-Earth model more reliable, it is necessary to consider climate conditions from previous years.

Output of GCMs can be processed through different methods to use their output at lower scales, within this study statistical method Singular Value Decomposition (SVD) was chosen to downscale data. Due to SVD maximizes square covariance accounted by each mode; this parameter was used as criterion for the selection of the most representative large-scale atmospheric variables and their combinations. Besides statistical parameters, selection of the variables was based on knowledge of the physical processes occurring on the study area. Relating statistics with physical background provide more reliable forecasts.

In addition, SVD is a good choice to downscale the model output, since it makes possible to compare two multi-dimensional fields, such as predictor – represented by the model data – and predictand – represented by the observed discharge data.

It is considered that model predictions for the future can be improved by simulations that start from the observed conditions. Using large number of the measurements of the past changes will partly increase confidence in the model and will give more accurate representation of the physical conditions and processes (Rossby Centre, 2010). This will provide an important and fundamental understanding of changes of the present conditions in the future on the decadal time-scale. Therefore long-term historical data from the model were involved in the SVD training. Historical data were tested against climatology (averages of seasonal river discharge data) to achieve better quality of the forecast.

Downscaled time-series of the historical run were compared with three different future scenarios by applying Fourier and wavelet analyses, in order to identify changes in frequency of the time-series. Three future scenarios were simulated for different representative concentration pathways (RCP) of greenhouse gasses, which depend on the economic, demographic and technological states, to mention some. Various combinations

of the distinct states give different radiative forcing pathways. Therefore three time-series of possible scenarios were compared to the historical time-series to find out any changes in peaks of discharge rates in terms of frequency. Obtained results of the comparisons were analyzed and described based on the knowledge of the fundamental physical principles and current conditions.

1. AIM

This thesis forms part of the integral project Hydroimpacts 2.0 carried on by the Swedish Meteorological and Hydrological Institute (SMHI), which focuses on the assessment of diverse impacts of the climate change on the hydrological processes at different temporal and spatial scales within Sweden (Olsson 2010).

The present study is concentrated in identifying the long-term seasonal variability of rivers discharge, at decadal scale, between the past years (1851-2005) and three different future scenarios (2006-2100) of concentration of greenhouse gases for the different climatologic conditions of Sweden by applying statistical downscaling to EC-Earth model output.

2. CLIMATE AND RIVERS DISCHARGE WITHIN THE STUDY AREA

Figure 1 Geographical map of the region

The study area is shown in Figure 1, it covers the whole Sweden (latitude ranging between 55° and 67° N and longitude between 11° and 24° E). The eastern side of the country is represented by a long coastal line, whereas the western part is represented by highlands and Scandinavian mountain chain, with the highest point in the North-West Kebnekaise (2,104 m). Besides the western part, terrain within Sweden is mostly flat with lowlands. A large part of Sweden is covered by forest and woodland, which becomes denser northward. The Southern area is represented by agricultural lands, which constitute about 8% from the whole territory of the country.

Climate in the South is temperate with cold winters, but mainly without snow, summers are cool and rainy; whereas in the North, climate is subarctic, which is represented by very cold winters with big amount of snow and short summers. Climate in Sweden is mainly affected by westerly and south-westerly winds, that come from the Atlantic Ocean and bring frequent storms and moist, which turns into precipitation. Main rivers of the country flow in the direction from North-West to South-East and belong to Baltic Sea. Within this project four rivers: Kalix and Torne in the North, Vindel in the middle part and Helge in the South were considered. Selected river

stations are shown in Figure 2. They were chosen in accordance to the representativeness of the different climatic conditions of Sweden, among other criteria that will be discussed further in section Data.

Combination of such factors as temperature, radiation, humidity, wind, precipitation and snow cover determines climatic conditions of a particular region, affecting runoff and river discharge. In this section all factors mentioned above are described. All numbers stated below represent mean values of the respective parameter for the period 1961-1990 according to the National Atlas of Sweden (NAS) (Raab and Vedin, 1995).

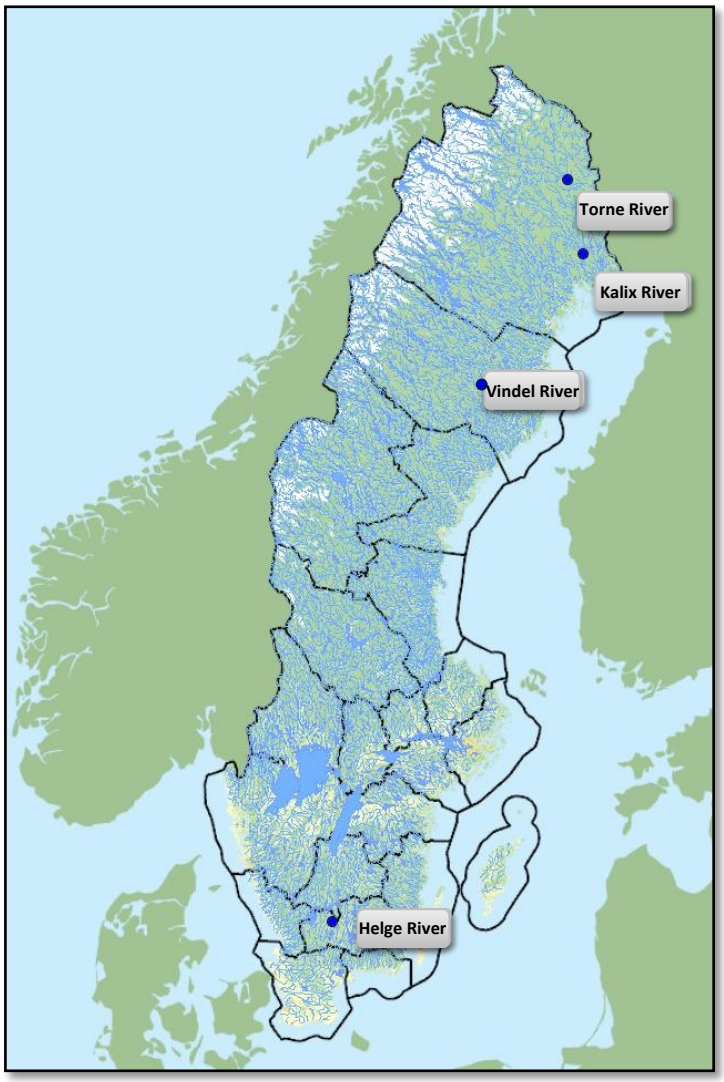

Figure 2. Selected stations for the rivers, which represent different climatic conditions of Sweden

Solar radiation balance is one of the factors that influences climate, thus it should be considered in the issues of climate change. Two of the mainly recorded parameters of the solar radiation are global radiation, which is represented by shortwave radiation on the horizontal surface, and sunshine duration. Within the given study area the global radiation changes in the limit of about 15% from year to year. However the normal value of the global radiation varies in the range of 800-1,100 kWh/m^2. Sunshine duration is dependent on the clouds conditions and seasons. Southern highlands and mountains closer to the North of Sweden get the least amount of the sunshine. However due to the long duration of the days during summer in the North, the total amount of sunshine is almost equally distributed along the whole coastline of Sweden.

Solar radiation is one of the main factors that regulate air temperature. Air temperature changes with altitude, thus all weather stations measure the air temperature at the same height, which is about 1.5 m above the ground surface. In summer the temperature in Sweden is approximately the same within the whole country, about 15-16°C in the North and 16-17°C in the South during July. However quite big difference can be observed during winter, which is about -15 to -17°C in the North and 0 to -1°C in the South in January-February.

The air pressure changes between 935 hPa in deep depressions and 1,065 hPa in high altitudes, thus the average pressure at sea level is defined as 1,013.2 hPa. In the North of Sweden the mean air pressure in January is determined as 1,009 hPa, whereas in the South is about 1,013 hPa. In July these numbers are equal to 1,010 hPa and 1,013 hPa, respectively. The maximum mean wind speed, that was measured ten meters above ground surface, is about 35-40 m/s along the coast line and in the high mountains.

Water vapor is a very important gas among the greenhouse gases and its pressure rate has a relevant role on natural water cycle. Mean values of the vapor pressure vary within seasons and differs in different parts of Sweden. During January this number is equal to 3 hPa and 5 hPa in the North and South, respectively; whereas in January the mean value reaches 8-12 hPa and 13-15 hPa in the North and South, respectively. Water vapor is a result of the evaporation (including evapotranspiration). The rate of annual evaporation within Sweden is about 400-500 mm in the South and 100-300 mm in the North.

As part of the continuous water cycle, water vapor rises into the air and condenses into clouds. In Sweden cloud cover is determined as follow: January – 60-70% in the North and 75-80% in the South, whereas in July – 65-75% in the North and 55-65% in the South.

In Sweden precipitation is very dependent on the westerly winds, which bring rain, as a result of collision between warm and cold air. The highest rate of precipitation is present at high mountains (close to Norway) and Southern Swedish Highlands (west side of the country). In Southern parts of the country the average annual precipitation rate is about 1,100 and 1,200 mm. In the North, apart from the obvious precipitation peaks in the mountainous area, the annual precipitation is 600-700 mm along the east coast.

Combination of precipitation and low temperatures determines the spread of the snow cover, which explains the fact that high-lying areas usually have deep snowpack. Distribution of the snow within Sweden is unequally spread in time. Southern Sweden has its deepest snowpack in January-February (if there is snow), while in North-West Lapland snow still continue to accumulate in April-May. Mean values of the snow depth within the period between 1961 and 1990 are unequally distributed as follows: January - 20 cm of snow cover with 60% of years with snow in the South, while in the North is about 50-80 cm with 99% of years with snow cover; April – 20 cm of snow cover with 1-20% of years with snow in the South, while in the North this value reaches 60-110 cm of snow cover with 99% of years with snow.

The process of accumulation and snow melting determines to certain extent the runoff in rivers. Furthermore runoff is to a large extent determined by the relation between evapotranspiration, precipitation and storage. The North-west part of Sweden contributes with the largest amount of runoff, about 1300 mm, while the lowest rates occur at the South-east region, about 100 mm. In summer, evapotranspiration is high and almost none of precipitation contributes to runoff, groundwater storage acts as source to feed watercourses. Due to this reason the groundwater table decreases during summer. In autumn runoff is produced by precipitation and groundwater table is replenished with

autumn rains. Groundwater also plays an important role in feeding watercourses at the North regions during winter, since precipitation is stored as snowpack. From observations during previously mentioned period the yearly mean value for the runoff corresponds to about 380 mm.

River discharge varies considerably throughout the year. In the North and middle part of Sweden a spring flood period due to melting of accumulated snow during winter, usually occurs during May-July. For the Southern part the flood period comes in March-May. At the North and middle of Sweden this variability is substantially higher than in the South. For instance, the mean discharge measured at Kallio and Vindel station (Figure 3) during 1961-1990 varies in the order of one hundred between spring flood and winter discharge. In contrast to this in the South of Sweden, discharge is rather constant and is mainly affected just by rains (Figure 4).

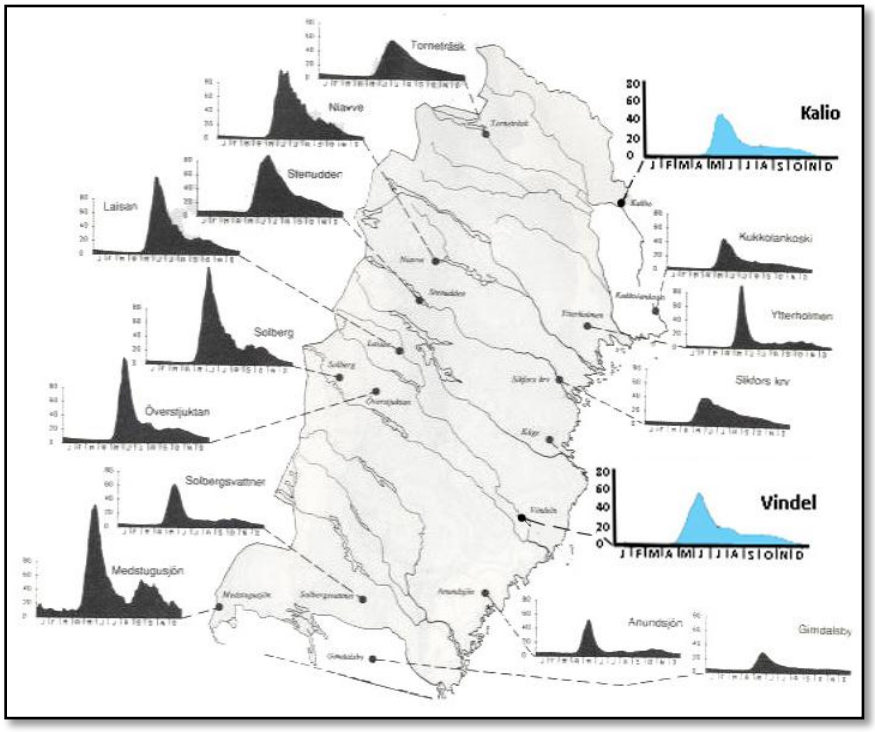

Figure 3. Hydrographs represent variation of the yearly discharge on different river stations in the North of Sweden for the period 1961-1990. Discharge marked with blue shows river stations, which were considered in this study. Discharge in l/(s·km^2).

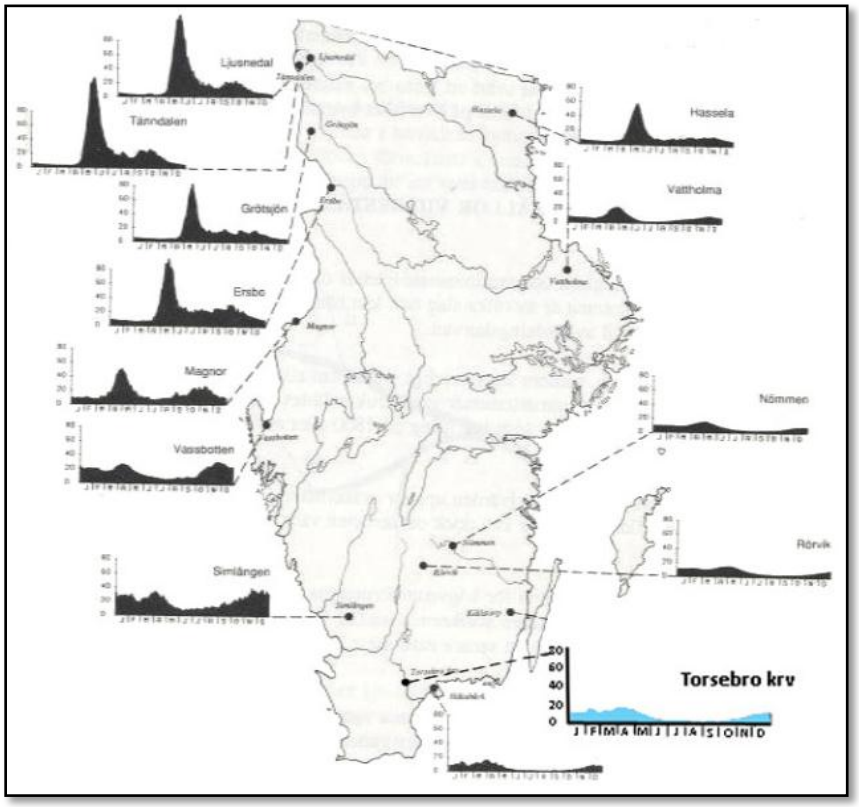

Figure 4. Hydrographs represent variation of the yearly discharge on different river stations in the South of Sweden for the period 1961-1990. Discharge marked with blue shows river stations, which were considered in this study. Discharge in l/(s·km^2).

Temperature variations make possible water mixing and circulation in lakes and oceans, this is an important consideration for climate models. Water temperature is dependent on sun rays, wind and radiation. At the beginning of spring water is cold since ice layers in rivers and lakes just melted. Water is warmed due to higher air temperature at the surface, diffuse radiation and sun effects. After water reach its higher temperature in watercourses in summer, about 15 ◦C in the North and 20 ◦C in the South, it is cooled by heat radiation into the atmosphere and chilly wind. The process of heat exchange between water and atmosphere decreases after water is covered with ice and snow in winter. During winter, temperature in watercourses decreases almost to 0°C in the South and frozen point in the North.

3. THEORETICAL BACKGROUND OF THE EC-EARTH MODEL

General Circulation Models or Global Climate Models (GCMs) are used to develop the climate projections based on mathematical expressions of the atmospheric and/or oceanic circulation. Through those projections different specific needs and requirements at national and international scales are satisfied taking into account previous knowledge as well as new information regarding climatic conditions. In addition, they can serve as starting background for research activities and development of new modules of the climate model, which will improve their skill in forecasting at different time-scales (Hurk, 2009).

Climate models simulate communication between atmosphere, ocean, land surface, living organisms and ice, and its effects on climate changes. The goal of these models is not only to understand connections of all climate components, but also to forecast their interactions (Walsh, 2008). In order to achieve these purposes, climate models receive information from the incoming and outgoing energy of the sun. This amount of energy is reflected or absorbed unequally on the surface of the Earth due to different factors, such as the composition of the atmosphere, affecting significantly Earth temperature gradients. GCMs process received information based on principles and laws of fluid-dynamics and thermodynamics, containing parameterization for the processes involving atmospheric and ocean components (Slute, Clement and Lohmann, 2001). Most of the climate models are based on a similar methodology, but differ by its complexity and varies in weighting of the particular variables or in the forcing (Walsh, 2008).

The application of the GCMs to predict hydrological effects at local scale was analyzed in the past (Kuhl and Miller, 1992; Miller and Russell, 1992 cited in Xu, 1999). In those studies it was observed that GCMs are not reliable regarding hydrological forecasts. Although GCMs could give a good representation of vertical transfer of water through evapotranspiration, precipitation and ground water storage, in most of the cases their hydrologic cycle representation is not complete, lateral transport of water is not included. Therefore, discharge is not taken into account (Xu, 1999). Hydrological forecast could be considerably enhanced if GCMs are coupled with hydrological models at regional level; it gives possibility to receive continuous feedback of all stages of the water cycle; however this involves scale differences between the models. The EC-Earth model belongs to the group of GCMs and in its last release (2.0) a closed hydrological cycle was implemented. It gives more physical and plausible representation to achieve a better model feedback.

Data from EC-Earth model of the last release of the experiments of the Coupled Model Intercomparison Project (CMIP5) was used in the given study. EC-Earth is a new generation model, which is a common coupled atmosphere-ocean model (AOGCM), developed by national weather services and institutes of 11 countries (Hurk, 2009).

EC-Earth is built according to the concept of “seamless prediction”, which unites weather forecasting and climate change researches into one framework. The objective of the EC-Earth model is to develop a global Earth System model (ESM), which combines and takes into account relation between biochemical and human processes with physical climate system (Hazeleger, 2010).

EC-Earth model includes the following components: IFS - module, which represents the atmosphere, NEMO - oceanic module, LIM – module for the sea-ice, HTESSEL – represents land surface and vegetation, TMS – module for the atmospheric chemistry. Components are coupled through OASIS coupler to the IFS module, which is already well-tested component (EC-Earth, 2010).

The EC-Earth model takes into account initial ocean conditions which are obtained from the previous long pre-industrial control simulation and for external forcing including changes in greenhouse gases and aerosols emissions, solar radiation, land-use and volcanoes (EC-Earth, 2010). Initial state of the coupler determines three modules, which together create initial conditions of the EC-Earth model: atmosphere, ocean and sea-ice. For the atmosphere component these conditions are included in the basic configuration as a T159L62, whereas restart files (updated in 2009) were used for ocean and sea-ice modules NEMO/ORCA1 configuration and they were forced with DF3 climatological data at the surface. The model in standard configuration uses reduced N80 Gaussian grid, which corresponds to a 1.125 degrees spacing (125 km) and a spectral resolution of T159, which is about 1.1°x1.1° and is the highest among all other models within CMIP5. EC-Earth model is running with time step of 1 hour (Brandt, 2010), however within this project monthly data were processed.

An important remark is that climate models in general (including EC-Earth) does not represent concept of time; it is impossible to obtain knowledge about when a particular event or climatological conditions occurs. Model represents power spectrum and frequency of occurrence of diverse states of natural variability, but not natural variability in temporal aspect. This is one of the reasons why it is impossible simply to compare observations with outcome from the model. To do so it is necessary to conduct further additional evaluation or downscaling (Rossby Centre, 2009).

4. DATA

The present study involves three following categories of data sets. Two types of monthly time series of atmospheric variables from the EC-Earth global model: historical run and three future runs. The third dataset is represented by discharge observations for different stations across Sweden, provided by the SMHI.

Large-scale atmospheric variables that are involved in this study come from the historical and three future climate scenario simulations of the EC-Earth model. Taken variables derived mostly from the atmospheric module, besides the variable sea surface temperature that was taken from the ocean module.

All model variables represent predictors or independent variables that were used to forecast the discharge tendency throughout Sweden. Model runs considered in this work are climatological runs, which means that model reproduce real climate in a statistical way. Due to this reason there will be no precise correspondence with observation data in any specific year (Brandt, 2010).

Since regional climate depends on synoptical climate, time series considered for the development of this project lie within the longitudes 75°W to 75°E and latitudes 20°N to 80°N and not just in coordinates near Sweden.

Historical time series cover the period from January 1851 to December 2005 and are represented by monthly data. Those time series come from the 16 variables listed in Table 1.

Table 1. EC-Earth historical atmospheric variables (Christensen, Gutowski and Nikulin, 2012).

Variable	Name	Unit
hfls	Surface Upward Latent Heat Flux	W m-2
hfss	Surface Upward Sensible Heat Flux	W m-2
hus	Near Surface Specific Humidity (850)	1
pr	Precipitation	Kg m-2 s-1
psl	Sea Level Pressure	Pa
ta	Air Temperature (850)	K
tas	Near-Surface Air Temperature	K
tauu	Surface Downward Eastward Wind Stress	Pa
tauv	Surface Downward Northward Wind Stress	Pa
tos	Sea Surface Temperature	K
ua	Eastward Wind (850)	m s-1
uas	Eastward Near-Surface Wind Velocity	m s-1
va	Northward Wind (850)	m s-1
vas	Northward Near-Surface Wind Velocity	m s-1
wap	Lagranian Tendency of air	Pa s-1
zg	Geopotential Height (850)	m

Some historical variables were given in a range of different pressure levels, namely: near surface specific humidity, air temperature, eastward wind and Northward wind. However, for this study they were extracted for the 850mb pressure level, which is referred to the elevation about 1000 m.

To assess the future long term discharge variability, three scenarios of representative concentration pathways (RCP), for the period from January 2006 to December 2100 were considered: RCP 2.6, RCP 4.5 and RCP 8.5. Each scenario is the result of different combination of economic, technological, demographic policy and institutional states that lead to different radiative forcing pathways, as indicated in Table 2 according to (Moss et al., 2010). From all these scenarios time series of relevant atmospheric variables were taken. Future scenarios differ in radiative forcing and have the same initial conditions.

Table 2. Radiative forcing at 2100 for RCP2.6, RCP4.5 and RCP8.5

Scenario	Radiative forcing
RCP2.6	Peak at 3 W/m^2 before 2100 and then declines
RCP4.5	4.5 W/m^2 stabilization after 2100
RCP8.5	8.5 W/m^2 reaches in 2100

All large scale atmospheric variables derived from the model were originally in the netcdf-format and model output was represented by three-dimensional matrices conformed by longitude-latitude-time.

In order to test SVD, observations of the discharge rate were used. Those discharge observations were selected from different river stations considering regulation state, observation length and geographical position. As it was mentioned before, in order to cover a wider spectrum of different climate conditions and therefore variability of the discharge in the whole Sweden, rivers were chosen from the Northern, middle and Southern part of Sweden (Figure 2). As a representation of the Northern region the unregulated rivers Kalix and Tornä, with stations Räktfors (#17) and Kallio 2 (#2395) respectively, were considered. Those rivers were processed together when SVD was applied for a 68-year period, corresponding to the observation length (1937-2005). For the middle part of Sweden Vindel river and the station of the same name (#50023) was chosen. Available data for this station has an observation period of 93 years (1912-2005). For the Southern region the best river, considering the above mentioned conditions, was Helge with station Torsebro KRV (#2191) and time period of 96 years (1909-2005). Observed discharge at mentioned stations can be seen in Figures 5 to 8.

The discharge data used for this study are seasonal means coming from daily measurements at different stations. Due to spring floods give the extreme river flow rate, and therefore the strongest signal for the model, seasonal spring flood data were used within this project for the three mentioned regions. From historical observations of the SMHI it was considered that flood season starts on May 1 and ends on July 31 for the North and middle of the country while at the South the discharge is evenly distributed through the year, therefore the selected period has no influence on the streamflow values and the considered period corresponds to March 1 to May 31.

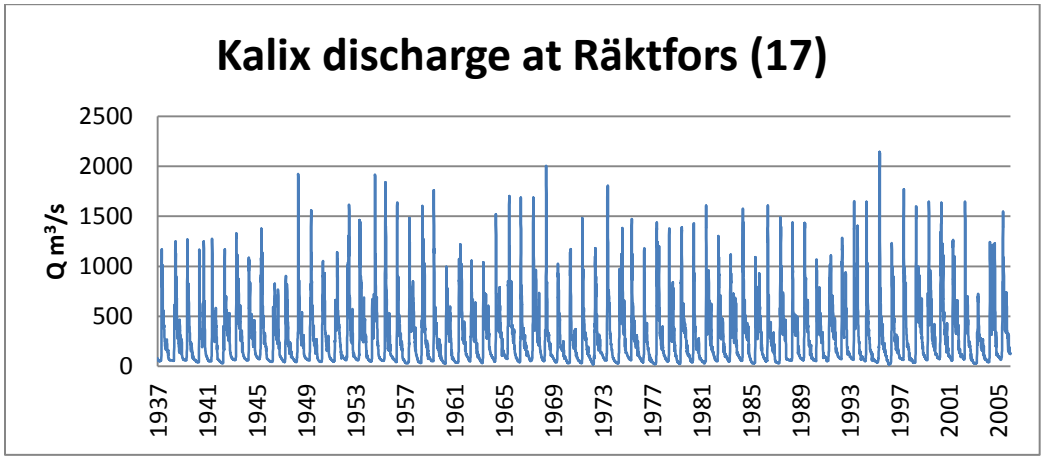

Figure 5. Historical discharge (1937-2005) for Kalix at Räktfors station

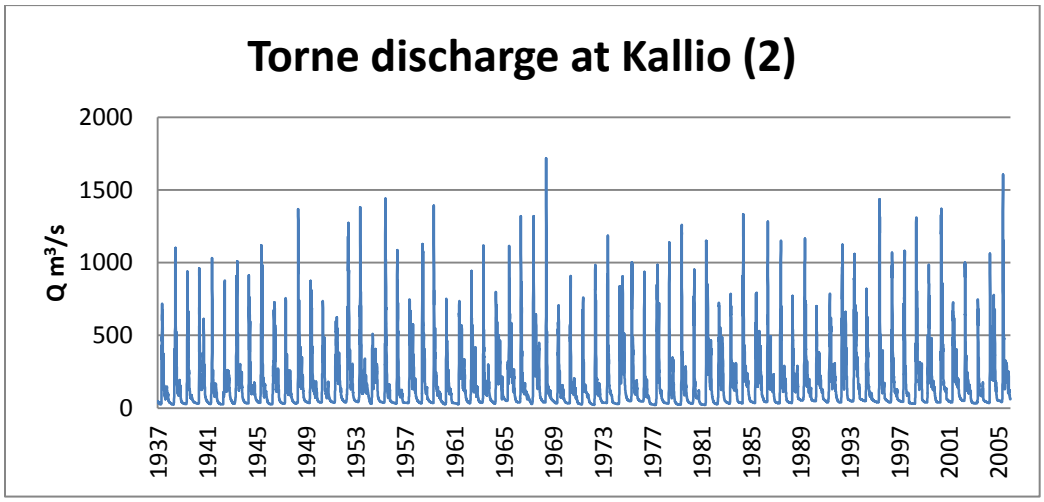

Figure 6. Historical discharge (1937-2005) for Torne river at Kallio station

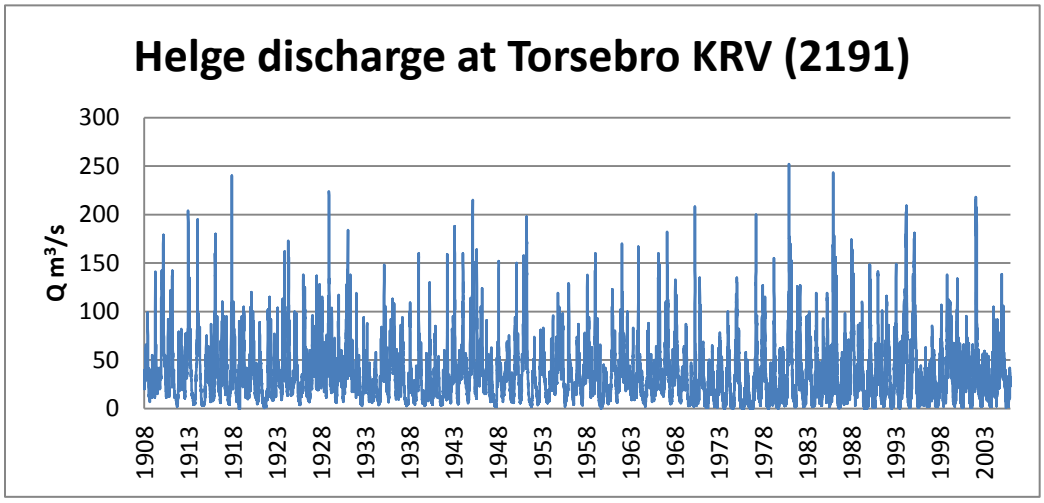

Figure 7. Historical discharge (1908-2005) for Helge at Torsebro KRV station

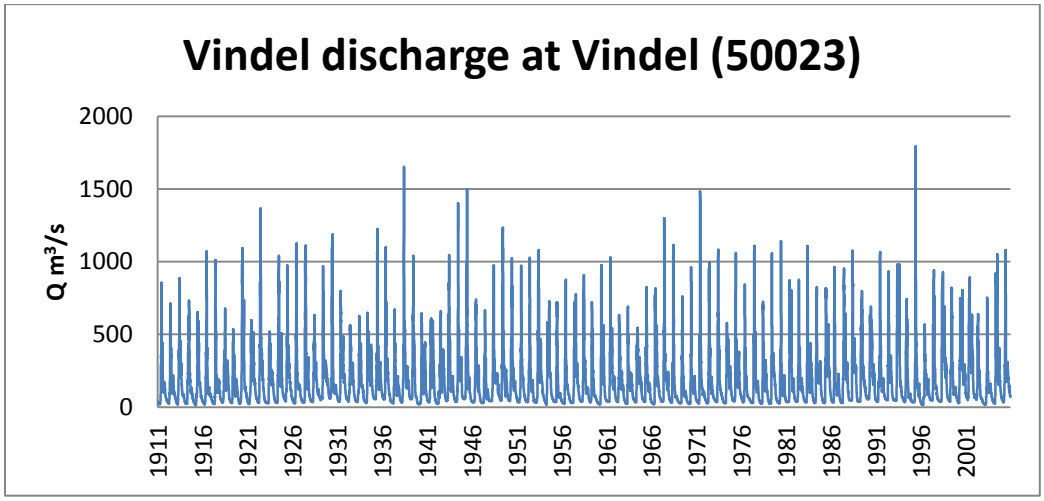

Figure 8. Historical discharge (1911-2005) for Vindel at Vindel station

5. METHODOLOGY

In order to reach the main goal of the study, to be able to analyze and compare seasonal patterns of variability of the water discharge among four rivers in two different periods of time (past and future), discharge peaks should be represented as frequency oscillations. This was done through the Fourier and Wavelet analysis. However, data obtained from the global climate model are large-scale atmospheric variables which are incomparable with small-scale river discharge data; moreover all data have different units and magnitude. Due to this reason data should be downscaled and standardized. As a downscaling technic Singular Value Decomposition (SVD) was used. Detailed methodology is described below by means of flowchart, see Figure 14.

A. USED SOFTWARE

Within this study for the data preparation and for the application of the main statistical methods Matlab 2010b was used.

Matlab is a programming environment which combines data analysis, advanced graphics and visualization, powerful programing language and numeric computations. An important feature is that it uses matrix as its basic unit instead of scalar values, this ease work with data of multiple dimensions. In addition, Matlab gives possibility to process technical problems faster than with traditional programming languages.

The main requirements for the software within the given study were the possibility of solving mathematical expressions involving large amount of data within a reasonable time frame,

processing and analyzing data in netcdf format, since this is the original data format from the EC-Earth, which is very common format for the climatological data.

Appendix A, B and C include the principal codes that were used in Matlab during this study.

B. DATA PREPARATION

Raw data as output from the model contains information of the variables at each grid point and each time step. Obtained data were in the netcdf format and represented as 3D arrays conformed by longitude, latitude and time as shown in Figure 9. To apply SVD it is required to get the information in 2D arrays. Thus the original 3D matrices were rearranged to get information of the location (longitude and latitude) as columns and time as rows (Figure 10).

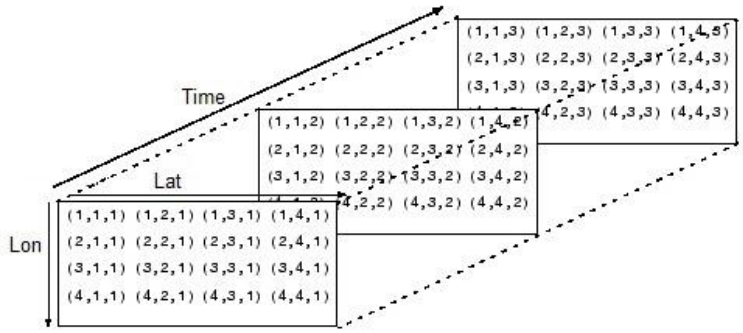

Figure 9. 3D array, obtained as an original data from the EC-Earth model

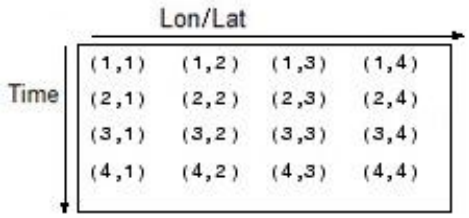

Figure 10. 2D array, obtained through the rearrangement of the original 3D matrix

From the original monthly data obtained from EC-Earth model, seasonal averages for each large-scale atmospheric variable for December-January-February (DJF), January-February-March (JFM) and March-April-May (MAM) were created. These seasons were chosen in order to identify the period in which the large atmospheric variables have the biggest influence on the river discharge. For the observation data the flood period is different for the diverse regions. Based on the data from the SMHI flood periods for the rivers were defined as follow: May-June-July (MJJ) for the Northern rivers and middle part of Sweden and March-April-May (MAM) for the Southern region. Seasonal river data of the discharge for the flooding period were created from the daily observations at the stations presented in Figure 2.

C. DOWNSCALING

GCMs use low resolution which is not suitable for providing information about climate and hydrologic response at local scale (Wilby and WIgley, 2000). Nevertheless it is known that regional climate is dependent on synoptic-scale processes (Hewitson and Crane, 2006). This dependency makes possible to apply downscaling techniques to use outputs from GCMs at local levels.

In order to make downscaling process more efficient it was necessary to select preliminarily the best predictors. Usually atmospheric and circulation data come from GCM, whereas the local models mainly provide precipitation or discharge. However every model is different and the principal aspect in the selection of predictors is the degree of representation of the physical processes (Hewitson and Crane, 2006). This is verified using the grade of correlation between observed value and model output for each predictor. The selection was done by running SVD with the mean seasonal value for each predictor and predicand observations to obtain the normalized square covariance (nsc), statistical parameter which indicates the degree of predictand-predictors dependency. It is important to mention that although the correlation between the predictor and the predictand is high, there may be some degree of uncertainty in the model, which can lead to misconceptions, for instance external and internal forces that tend to change climate, such as greenhouse gas emissions, land cover change and weather variability to name a few.

According to Xu (1999) climate downscaling could be classified into dynamic and empirical downscaling. The first one focuses on explicitly solving physical processes i.e. by using GCM with finer grid on the area of interest or by doing simulations from regional-scale limited-area models with GCM outputs as boundary conditions. Although the grid scale decreases considerably the computational requirements are still very high. The empirical downscaling approach consists of statistically relating data from global scale (predictors) to local meteorological series (predictands). Once this relationship is defined, validation of this interaction between predictor and predictand should be done using independent time series, which gives possibility to forecast predictands at local scale. Singular value decomposition (SVD) and canonical correlation analysis (CCA) are examples of this approach.

I. SVD

Due to the low computational requirements and the possibility to conduct multivariate analysis, SVD was chosen as downscaling method. In the present research SVD may be referred as a multivariate statistical analysis which relates large scale atmospheric variables with the local observed discharge, maximizing the normalize square covariance (nsc). In this way, the signal of a set of multiple variables as function of another set of several variables can be calculated. The theory behind SVD presented in this paper is based in previous studies by Uvo et al. (2001) and Bretherton, Smith and Wallace (1992), where a wider explanation of this method can be found.

SVD is a method which factorizes a given matrix (A) into the eigenvalues matrix (s) and the left and right eigenvectors (g and h respectively), being k the mode of the eigenvector as stated in eq. 1.

$$A = g_k s h'_k \qquad \text{eq. 1}$$

Given two data sets as it is shown on the Figure 11, the predictor $Y_{nt,ny}$ and the predictand $Z_{nt,nz}$ where nt is the time steps and ny and nz are the space points/variables, the spatial patterns (g and h), which are a linear combination from Y and Z and explains most of the temporal covariance between the two fields (predictor-predictand), can be found by applying SVD to the cross-covariance matrix of Y and Z (C_{YZ}), C_{YZ} is given by eq. 2. Therefore time series are linearly related to the eigenvectors taking the form shown in eq. 3 and eq. 4. From the new time series U and V, the firsts modes which explain most of the covariance between the two fields can be used to avoid noise in future analysis, being the modes orthogonal among them. Due to the different units and magnitudes of the related variables, it is important to mention that data must be zero mean and standardized (eq. 5) in order to obtain C_{YZ}.

$$C_{YZ} = \frac{Y'Z}{nt-1} \qquad \text{eq. 2}$$

$$U_{nt,k} = Y_{nt,ny} * g_{ny,k} \qquad \text{eq. 3}$$

$$V_{nt,k} = Z_{nt,nz} * h_{nz,k} \qquad \text{eq. 4}$$

$$\frac{Variable - Variable\ mean}{standard\ deviation} \qquad \text{eq. 5}$$

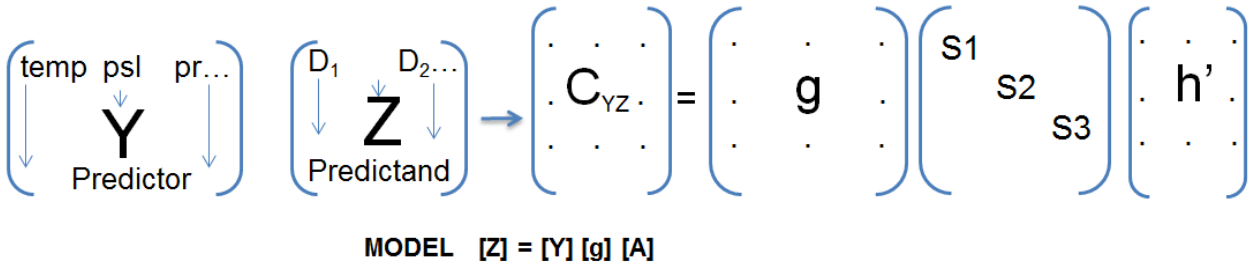

Figure 11. Factorization of the covariance matrix done by SVD

Once that the SVD for C_{YZ} has been done, it is required to quantify the strength of the relationship between the predictors and the predictands in the modes obtained from different SVD expansions. This is done by calculating the normalized squared covariance (nsc) according to eq. 6. var_i and var_j are the variances, i^{th} is the grid point in the left field and j^{th} in the right field and $var_i * var_j = 1$ since data are normalized. Depending on the spatial relationship between the two fields nsc varies between zero and one; the closer nsc to one, the higher the correlation between the two fields. In addition, this parameter is useful to compare SVDs from different fields, as it was done for the variable selection within this study.

$$nsc_k = \left(\frac{s_k^2}{\Sigma_i \Sigma_j var_i var_j}\right)^{1/2} \quad \text{eq. 6}$$

The final step is creating the model. This is done by relating values of the predictor time series (U) to the predictand field (Z) as result the regression matrix (A) is created see eq. 7. Then the time series can be extrapolated in time according to eq. 8, were $\hat{Z}$ is the forecasted field and is referred as zhat in the appendix B, where the full process to develop the model is presented.

$$A = \frac{U'Z}{U'U} \quad \text{eq. 7}$$

$$\hat{Z} = UA = YGA \quad \text{eq. 8}$$

D. FOURIER ANALYSIS

Discrete Fourier Analysis or Fourier Transform was chosen as the best option in the given case, due to climate models are not representing the concept of time correctly and within this study the main interest is to identify changes in frequency of the flow rates.

Fourier Transform is a very common technique for extracting spectrum content of the time-series data. The main idea of this method is to transform time series from the time domain into the frequency domain (Uvo C.B., 2012a).

As a basis Fourier Transform uses sine and cosine functions (Figure 12). Through this method time-series data are decomposed to the sum of sine and cosine waves, which represents frequency of the original data. This sum of waves creates a continuous sinusoidal function according to eq. 9.

$$Y_t = \bar{Y} + \sum_{k=1}^{n/2}\left(A_k \cos\left(\frac{2\pi kt}{n}\right) + B_k \sin\left(\frac{2\pi kt}{n}\right)\right) \quad \text{eq. 9}$$

Where Y_t is a given data series; A_k, B_k are Fourier coefficients, which are a function of the frequency; n is the length of the time-series (number of observations); k is the harmonic number (number of sinusoids, which are harmonically related); t is time.

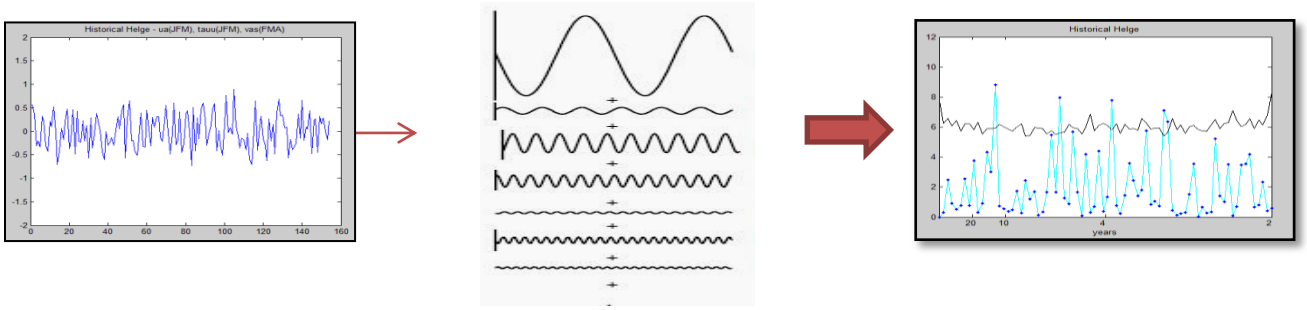

Figure 12. Fourier transform of the time-series from time-domain into frequency-domain, using sinus and cosines functions.

Fourier transform does not change the data, but it represents them in another way. It should be noticed that this transformation does not preserve time information. The concept of time is lost and data are represented as frequencies (ω_k), see eq. 10.

$$\omega_k = \frac{2\pi k}{n} \qquad \text{eq. 10}$$

It is significant to mention that the lowest possible frequency that can be represented by Fourier transform is 1/n, whereas the maximum is n/2. It means that it is impossible to get higher frequency than the half of the available observations (Wilks, 2011).

The coefficients A_k and B_k for the particular time series Y_t can be found through the generalized eq. 11 and eq. 12.

$$A_k = \frac{2}{n}\sum_{t=1}^{n} y_t cos\left(\frac{2\pi\, kt}{n}\right) \text{ eq. 11}$$

$$B_k = \frac{2}{n}\sum_{t=1}^{n} y_t sin\left(\frac{2\pi\, kt}{n}\right) \text{ eq. 12}$$

However it should be noticed that eq. 11 and eq. 12 can be applied only in case when data is equally spaced in time and contains no missing values (Wilks, 2011). The amplitude-phase for the harmonic k can be calculated through the eq. 13.

$$C_k = [A_k^2 + B_k^2]^{1/2} \qquad \text{eq. 13}$$

Time series transformed by Fourier method are analyzed through graphical representations, plotting spectrums, which show squared amplitude C_k^2 as a function of frequency ω_k. Spectrum represents the proportion of variations in the time series through the oscillations of the harmonic frequencies (Wilks, 2011). A step by step description of Fourier transformation applied in this study is presented in Appendix C.

The main advantage of the Fourier Transform is that it gives possibility to see which variables (processes) have the biggest influence on variance of the data series. This is done through determining the highest peaks of frequencies and amplitudes, which are statistically significant. Fourier method makes possible to get better information of the strongest signal of the variables without unnecessary noise (Uvo, 2012b).

Analysis of the statistical significance at 5% was done within this study through the comparison of the spectrum of the time-series with red-noise spectrum.

Fourier transform as well as most of the methods has its conditions. As it is known, pattern of the time-series data can be characterized by trend, seasonality and whether data are stationary or not. One of the requirements of the Fourier transform method is that data are supposed to be stationary for the accurate tracking of changes. Climate data that were processed within this project are stationary. Data are stationary in the case when statistical properties, such as frequency and amplitude, do not change in time; they should be time independent (Uvo C.B., 2012b).

E. WAVELET ANALYSIS

Another method to identify variability of time-series is wavelet transform. This analysis was done in order to have stronger confidence of the analysis of the discharge variability.

In contrast to the Fourier analysis, where the signal transformation is obtained by addition of the continuous sinus and cosines functions of different frequencies, wavelet analysis uses signals of quick local variation. The transformation is done by comparing waveforms of limited duration to the signal and calculating their correlation coefficient. Subsequent shifting of the same wavelet through the full signal period and repeating this process with different scale of the wavelet in order to analyze different frequencies is shown on Figure 13 (Uvo, 2012a). This process is described by eq. 14 for the case of continues wavelet transform and eq. 15 for continues wavelet transform of discrete data series. It is important to mention that the wavelet should be normalized to have unit energy for the further comparison of the wavelet coefficients at different scales and from other time series.

$$C(scale, position) = \int_{-\infty}^{\infty} f(t)\Psi(scale, position, t)dt \qquad \text{eq. 14}$$

Where f(t) is the signal, Ψ (scale, position, t) is the wavelet and C (scale, position) is the wavelet coefficient.

$$W_n(S) = \sum_{n'=0}^{N-1} X_{n'}\Psi\frac{(n'-n)dt}{s} \qquad \text{eq. 15}$$

Where s is the scale, Xn is the data series and Wn(S) is the wavelet coefficient (Torrence and Compo,1998).

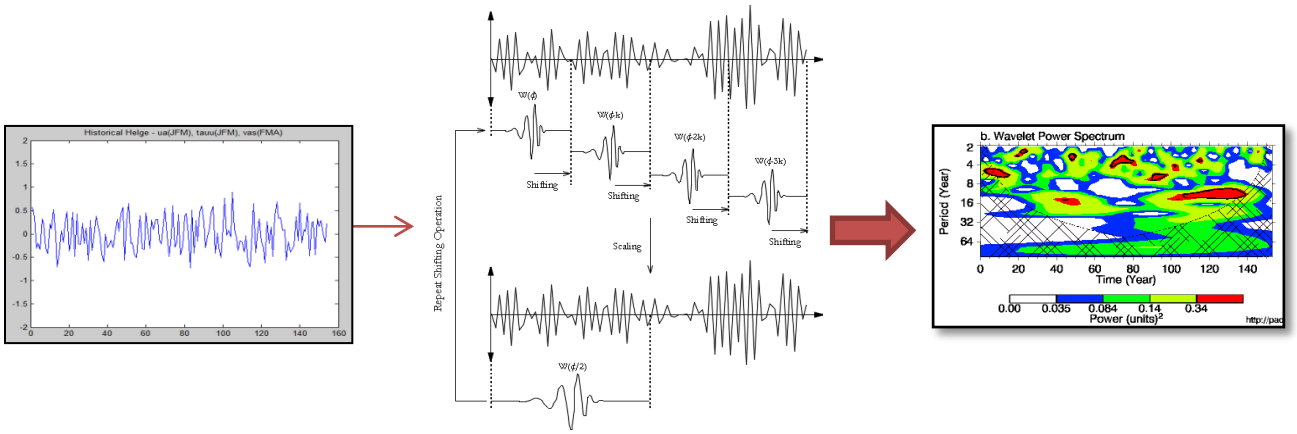

Figure 13. Transformation of the time-series from time-domain into frequency-domain by wavelet analysis

Wavelet analysis, in addition to represent time-series through frequency and power, shows time. For this study the information about time was not considered since the EC-EARTH, as well as other climatic models, does not represent time accurately (as it was mentioned before). The results of the wavelet transform were used to compare with Fourier analysis output and get a better representation of the discharge variability.

Statistical significance analysis at 5% (as well as Fourier transform) was done through the comparison of the wavelet spectrum of the time-series with red-noise spectrum. As the type of the base function a non-orthogonal, complex wavelet function Morlet was chosen. Complex functions return information about amplitude and phase better than real functions adjusted for the tracking of the oscillation behavior, which means that it contains more oscillations in comparison with other functions. Compared to such function as Mexican hat, wavelet Morlet has higher frequency resolution, however worse time resolution. During the wavelet transform in the given study time-series was padded with zeros and cone of influence was applied. This gave possibility to eliminate the edge effect, since errors take place at the beginning and the end of the power spectrum in the finite time-series (Torrence and Compo,1998).

Within this study wavelet transform was done using an interactive wavelet plot, which is part of the work of Torrence and Compo (1998). This tool is available at (http://paos.colorado.edu/research/wavelets/) and is able to perform wavelet analysis based on different types of wavelet base functions, scales and gives possibility to choose confidence interval against white or red noise among other features.

F. METHODOLOGY FLOWCHART

For an easier and better representation of the project methodology, all performed steps within this study are summarized in the flowchart in Figure 14.

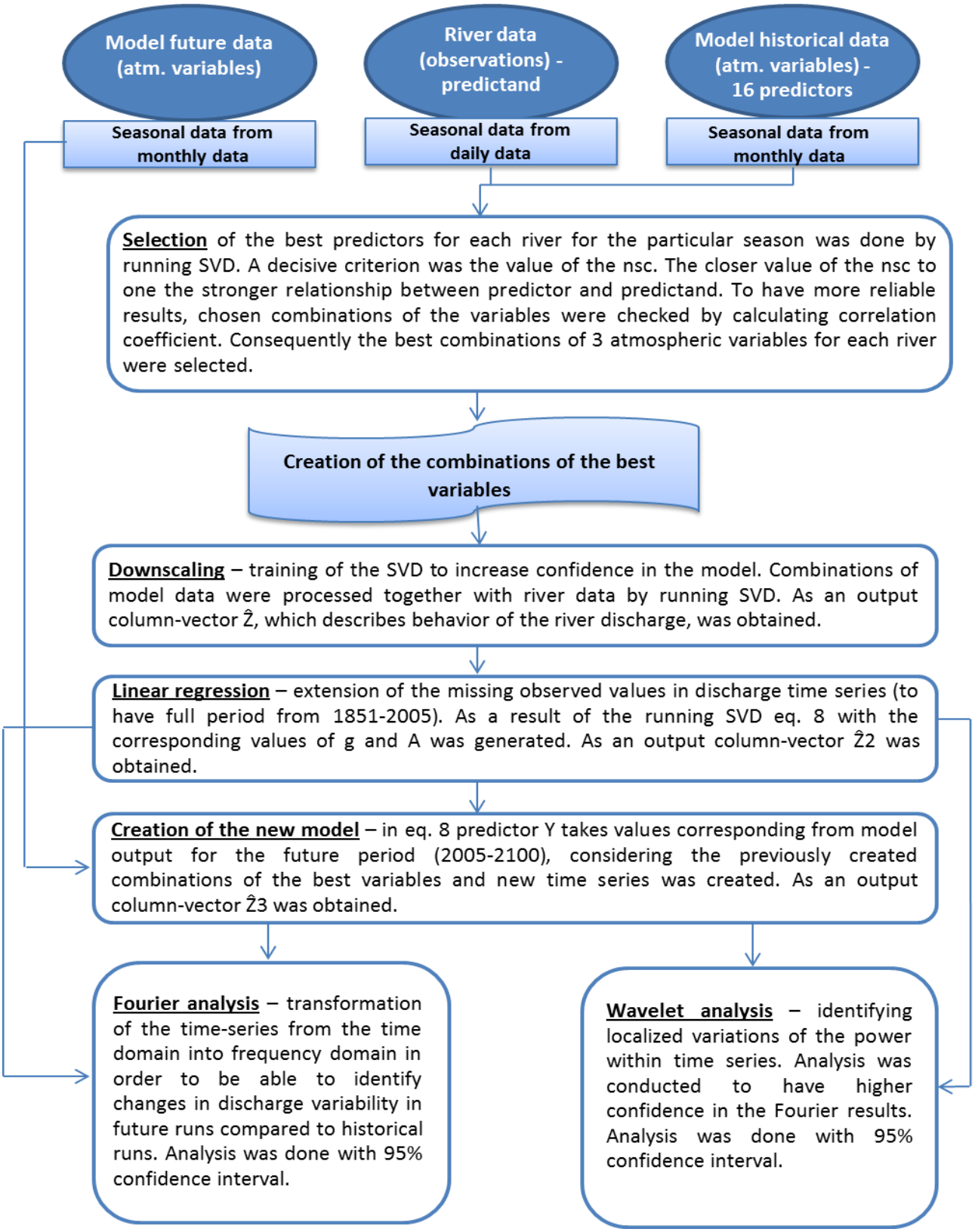

Figure 14. Methodology flowchart

6. RESULTS

In this section the final results of the Fourier and wavelet analysis are presented. In addition, the interim results of the application of SVD for the selection of the predictors and identifying best combinations of them are briefly described.

A. SVD

As starting point 16 atmospheric variables were obtained from the EC-Earth model. Only a few of them were selected as predictors by running SVD since this method maximizes nsc (appendix A). The closer the value of the nsc to 1 the higher the statistical dependency between predictor and predictand. When dealing with climatological data, it is expected to get relatively low values for nsc (between 0.1 and 0.3), those values may be acceptable since phenomena and processes involved in climate are very complex. Explaining 10 - 30% is enough to get an overview of the parameters that have larger influence on certain phenomena. Tables 3 and 4 present nsc for the variables that have the highest value of nsc for each particular river at the different considered seasons. For Kalix and Torne the predictors selection give basically the same results, as it can be seen in Table 3. Due to their geographical proximity they share similar climatological conditions. The statistical relation explained by the nsc was supported by analyzing the physical interaction between atmospheric variables and river discharge in different regions.

Table 3. Best predictors for Kalix and Torne rivers in the North of Sweden according to their nsc

Kalix river				Torne river			
Variable	**DJF**	**JFM**	**FMA**	**Variable**	**DJF**	**JFM**	**FMA**
hfss	0.1221	-	-	hfss	-	0.1219	-
tas	0.1247	-	-	tas	0.1247	-	-
pr	-	-	0.1271	pr	-	-	0.1271
ta	0.1274	-	-	ta	0.1274	-	-

Table 4. Best predictors for Vindel river in the middle part of Sweden and Helge river in the South of Sweden according to their nsc

Vindel river				Helge river			
Variable	**DJF**	**JFM**	**FMA**	**Variable**	**DJF**	**JFM**	**FMA**
psl	-	-	0.1528	ua	-	0.1334	-
tas	-	-	0.1329	tauu	-	0.1275	-
ua	-	-	0.1389	vas	-	-	0.1252
zg	-	-	0.1465	zg	-	0.1419	-
uas	-	-	0.1315	psl	-	0.1389	-

Representative combinations of the selected predictors were created and tested by applying SVD to generate different prediction models. The accuracy of these predictor models was measured considering their correlation coefficients, which were performed at

99% confidence level; the best combination of variables and their respective season for each river are presented in Figure 15. Correlation coefficient represents linear relation between two independent time series and can vary between 1 and -1, both extreme values represent 100% of positive or negative correlation respectively, whereas correlation coefficient of value 0 means that there is no linear correlation between the time series.

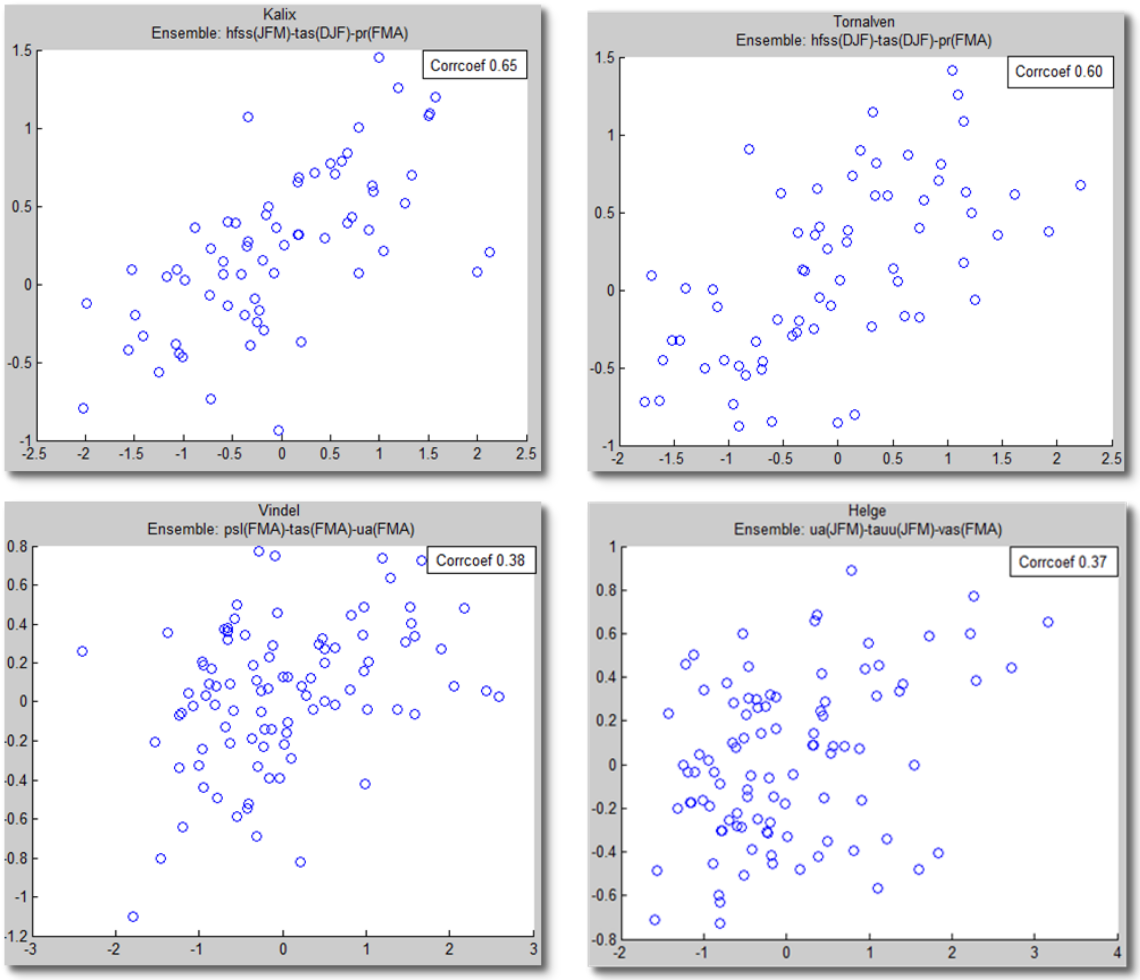

Figure 15. Correlation coefficient at 99% confidence level for the combinations of the best variables with their respective seasons for the 4 selected rivers

B. FOURIER AND WAVELET ANALYSIS

Fourier and wavelet analysis were applied to compare downscaled historical time-series with downscaled time-series of three different RCP scenarios, they were created based on eq. 8. The following results represent time-series in the frequency domain for the four selected rivers. The results are organized in the following way: all graphs are structured by scenarios to simplify the comparison process: first RCP2.6 is presented, followed by RCP4.5 and RCP8.5. Each scenario includes all rivers with their respective six graphs, two charts on the top represent the index time series for historical (to the left) and future discharge (to the right) based on standardized fields, output of the Fourier analysis for the historical time-series and future period are shown at the middle where Fourier demonstrates frequency of the time-series which is represented with the blue line and tested within 95% statistical confidence interval performed by bootstrap analysis (black solid line), whereas the two charts at the bottom show results for the wavelet analysis. Areas which have a black contour in wavelet graphs are significant at 5%.

Charts of Fourier analysis below represent time-series by peaks of frequency of natural variability on the x-axis and amplitude of oscillation on the y-axis. Wavelet analysis represents the same information, however frequency is represented on y-axis, power of oscillations is shown by color, from white (weak) to red (strong) and in addition time information is provided in x-axis.

I. SCENARIO RCP2.6

Figures 16 and 17 represent standardized discharge and results of the frequency for the historical time series and for the RCP2.6 scenario for the Northern rivers of Sweden: Kalix and Torne respectively.

The two first charts of each figure presents standardized discharge with a slightly visible upward trend in data for the future period in both Northern rivers, since data were not detrend.

As it can be seen from the Fourier results for the historical time-series model represent long term variability about 18, 30 and 65 years for the unregulated Kalix river, whereas only 18-year variability is significant in the wavelet power spectrum, all others peaks are in the zone padded with zeros, which makes this information statistically less reliable. By comparing these results with frequency for the future time-series it can be seen that only one statistically significant peak about 65 years is present according to the Fourier charts. However wavelet spectrum shows this peak in the zone padded with zeros.

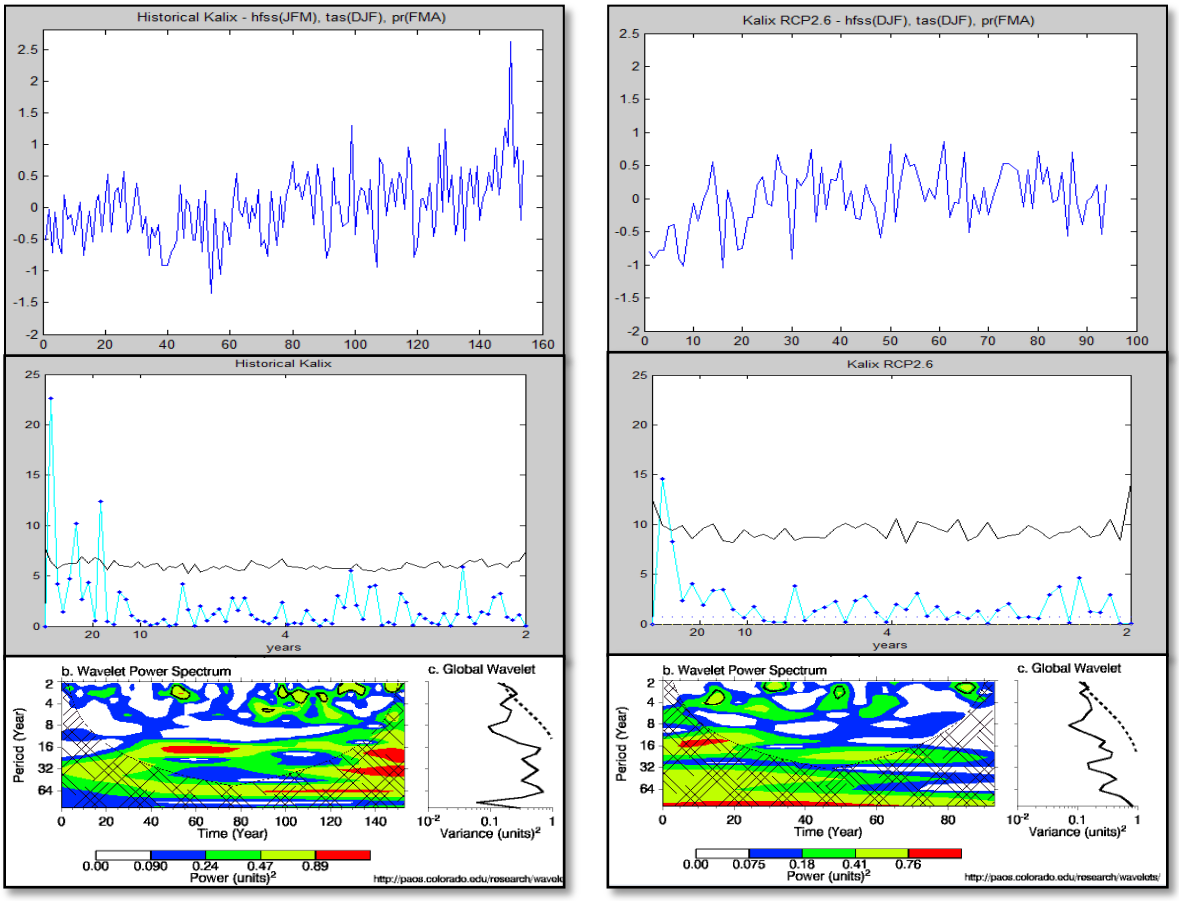

Figure 16. Fourier and wavelet analysis for discharge variability between historic period and RCP 2.6 for Kalix river

Discharge variability for historical time-series of the Torne river shows short-term variability of 3 and 5 years and long-term peaks about 50 years. Short-term variability is significant in wavelet spectrum, whereas long-term variability is in the zone padded with zeros, which makes this event debatable. For the future scenario model represent long-term variability about 30 years, which is confirmed by both methods, however short-term variability disappeared.

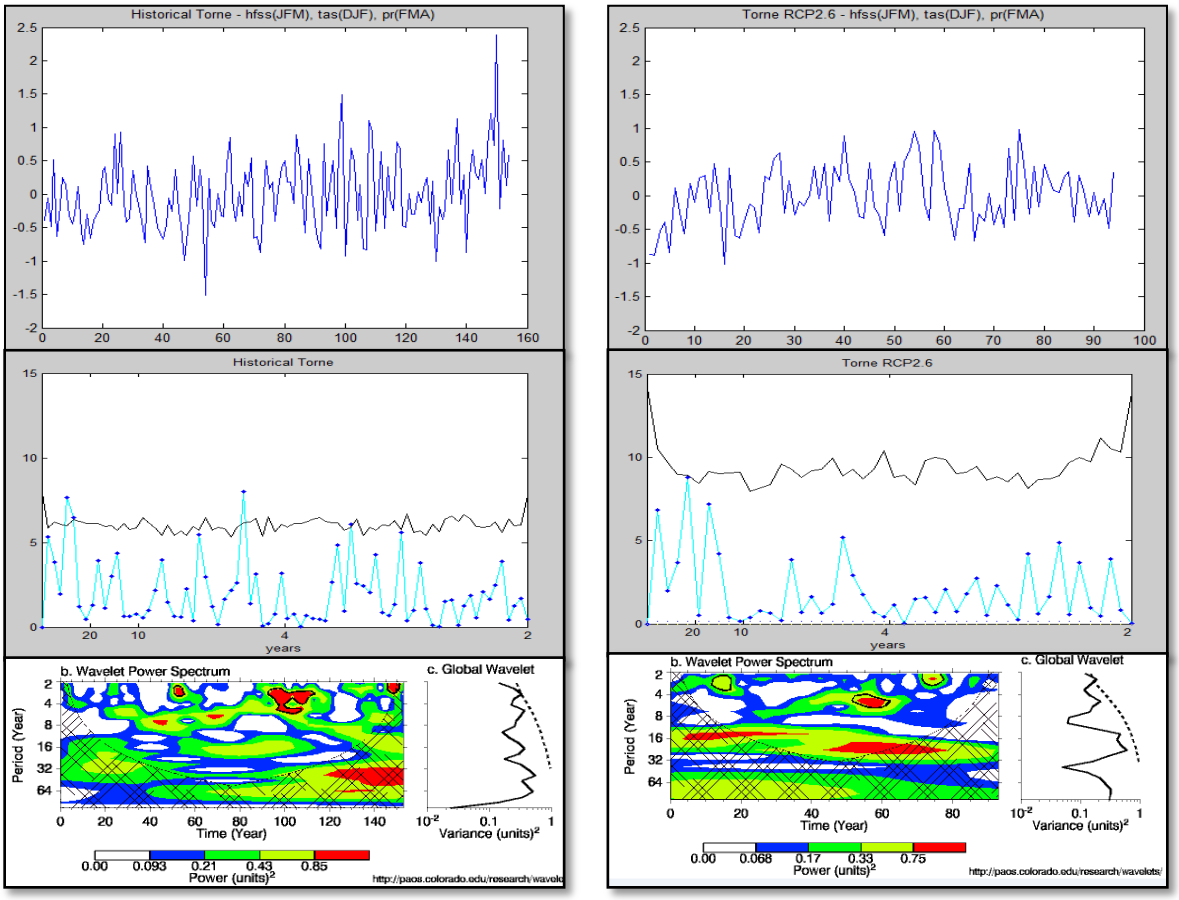

Figure 17. Fourier and wavelet analysis for discharge variability between historic period and RCP2.6 for Torne river

Figures 18 and 19 represent frequency variations in the discharge of Vindel and Helge rivers respectively for the historical time-series and future scenario RCP2.6.

Trend, which was slightly visible in the Northern rivers for the future period, is not represented in Vindel and Helge rivers, as can be seen from the two first graphs of standardized discharge.

Discharge rate variability for the historical data of Vindel river is clearly significant at 5% for the short-term (between 3 and 8 years) and long-term (about 30 years), that is proved by Fourier and wavelet analysis. However in the future for the scenario RCP2.6 this situation changes: 30-year variability is not significant and short-term peaks are represented mainly by 3 years variability according to Fourier and 3 and 10 years according wavelet spectrum.

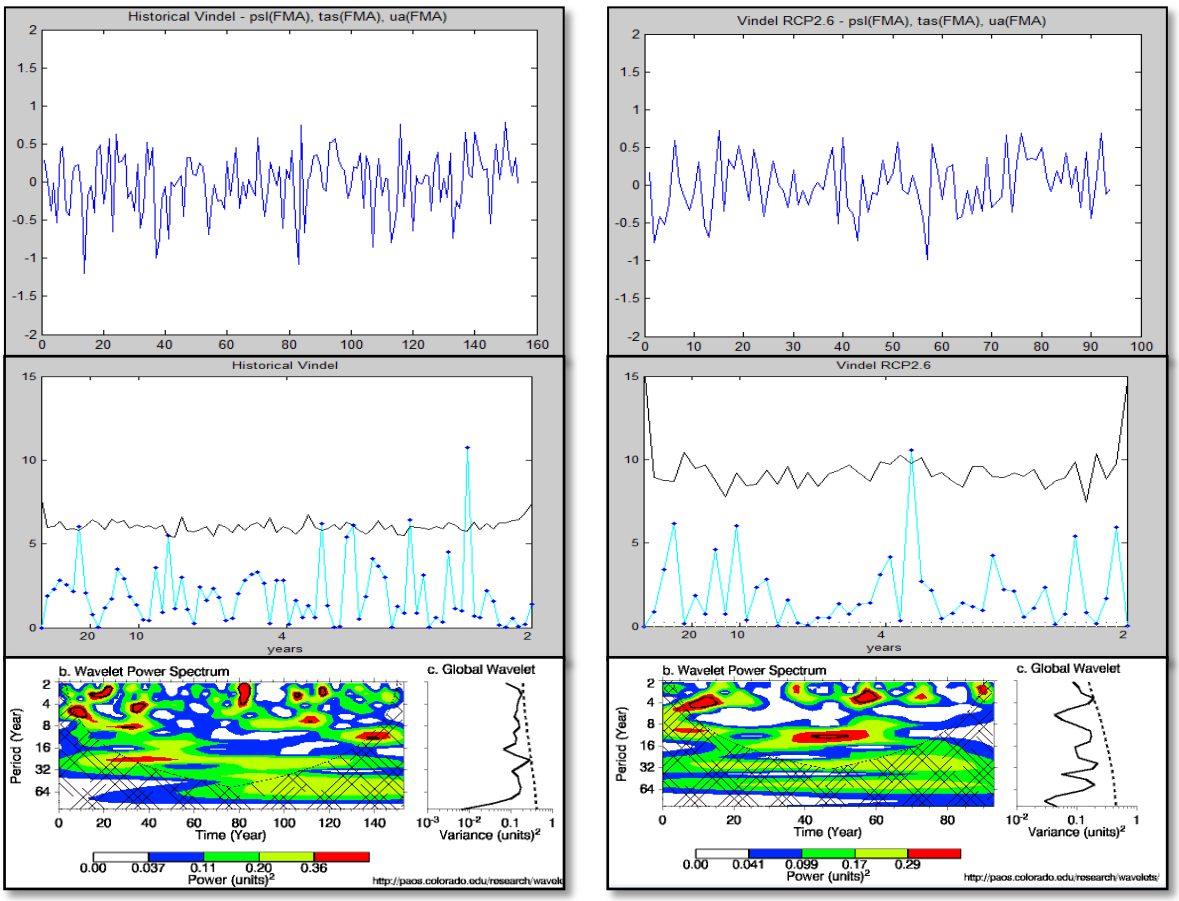

Figure 18. Fourier and wavelet analysis of discharge variability between historic period and RCP 2.6 for Vindel river

For historical period of the Southern river Helge, model shows variability at more time scales, compared to the other rivers, variability in the discharge rate: short-term is present at 3, 4 and 6 years and for long-term it is present at 14 years. These peaks coincide in the outputs for both methods. However, this situation changes for the future scenario. Fourier transform does not provide any variability within 95% confidence interval, whereas wavelet power spectrum shows variation between 8 and 16 years and short-term oscillation between 2 and 4 years that are significant at 5%.

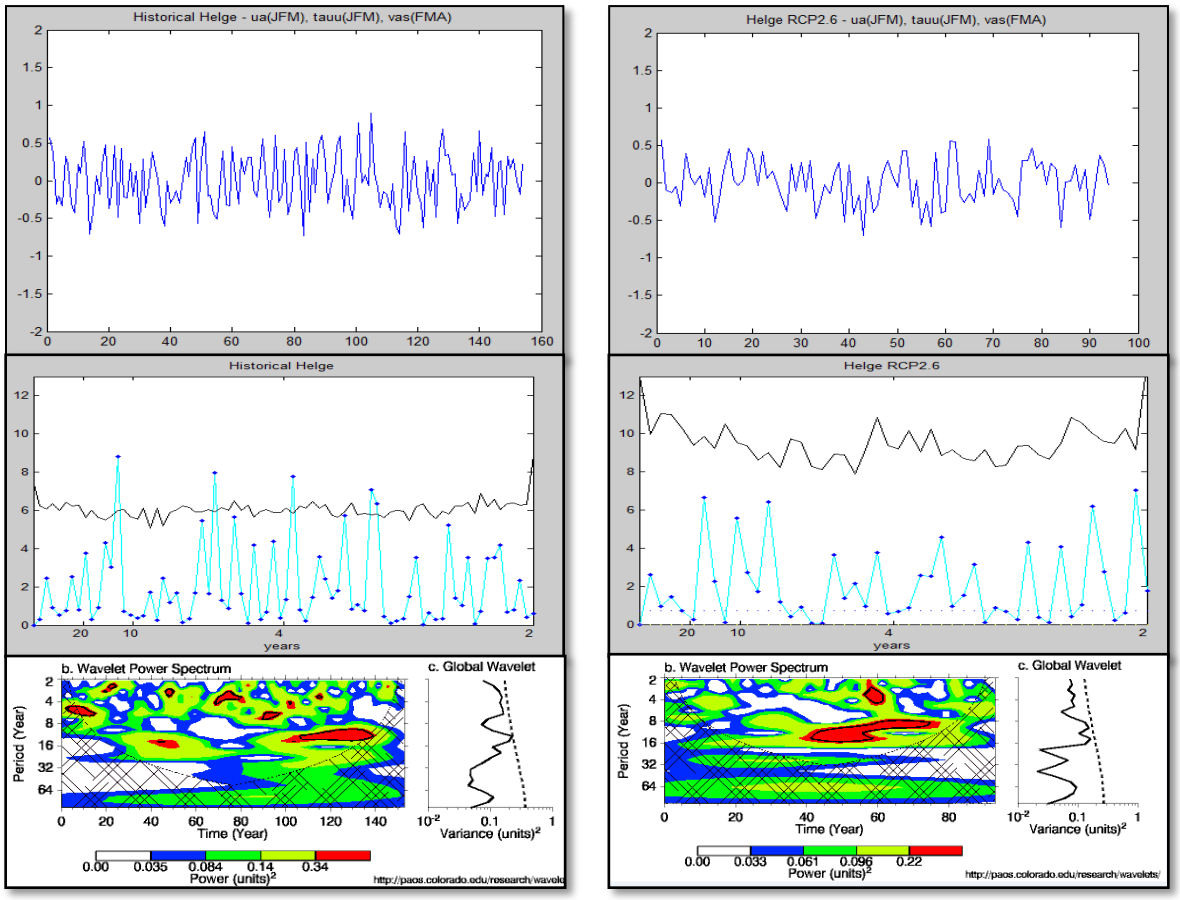

Figure 19. Fourier and wavelet analysis of discharge variability between historic period and RCP 2.6 for Helge river

II. SCENARIO RCP4.5

Figure 20 and Figure 21 represent frequency of the discharge variability of Kalix and Torne rivers in the historical period and future for the RCP4.5.

As it can be seen from the following figures, upward trend in data for Kalix and Torne rivers become more evident compared to the RCP2.6 scenario.

Discharge variability for Kalix river in the historical run of the model is represented by low power short-term variability between 2 and 4 years and long-term variability about 65 years, this is confirmed by both methods. In contrast with the historical run, prediction of the future situation is considerably different between both analyses. Fourier shows power of amplitude 37 for the long-term variability of 65 years, whereas wavelet analysis still indicates that there is low power and short-term variability between two and four years. Long term variability with a frequency of 70 years falls into the padded with zeros region.

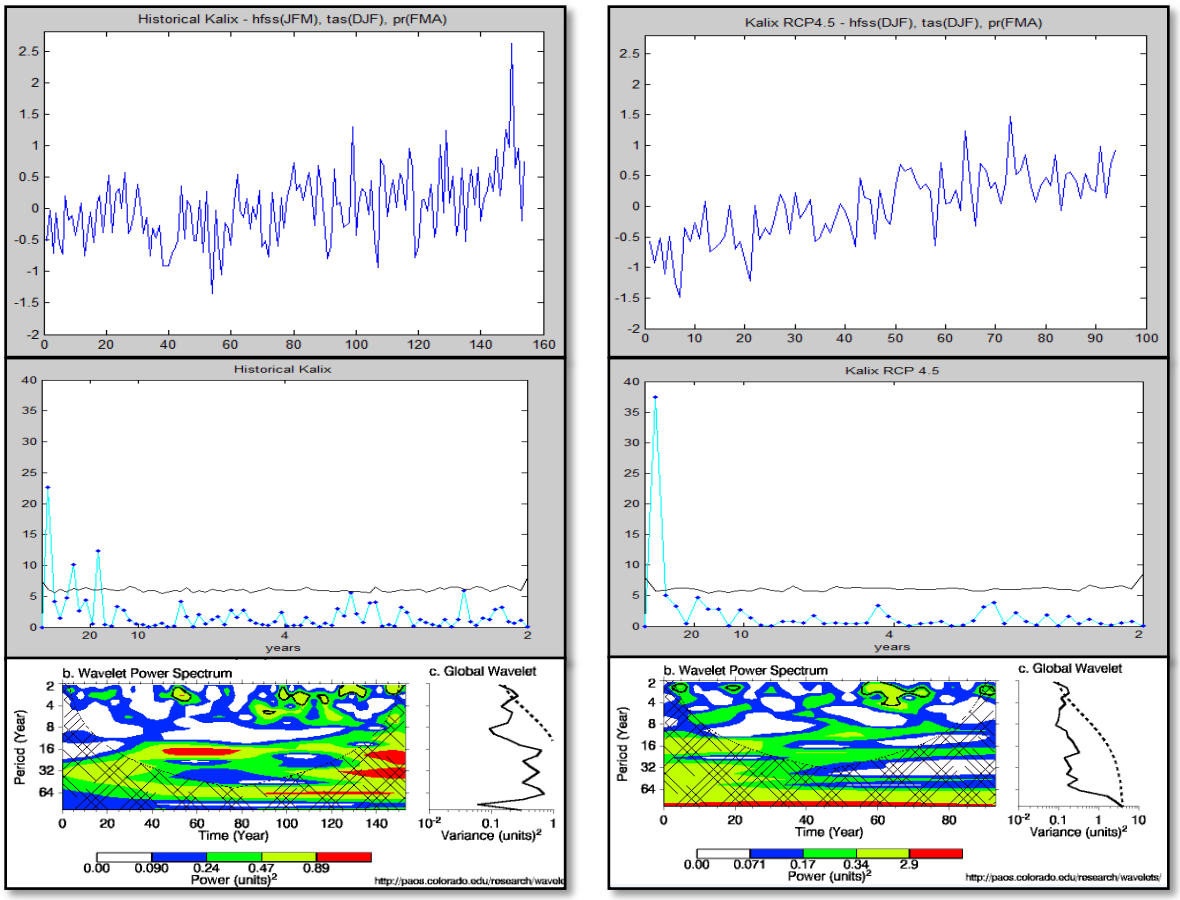

Figure 20. Fourier and wavelet analysis of discharge variability between historic period and RCP 4.5 for Kalix river

Coming to the Torne river it can be seen from Figure 21 that historical data shows short-term variability between three and five years, which is identified by both approaches, significant at 5%. Long-term variability is shown in the Fourier spectrum within 30-50 years, while this peak is powerful but falls into the region padded with zeros in wavelet spectrum. For the future scenario RCP4.5 short-term variability is still preserved for three years. However long-term variability is clearly shown by powerful peak (amplitude 25) at frequency of 65 years in the Fourier spectrum, whereas concerning wavelet analysis this long-term variability falls into the padded with zeros region.

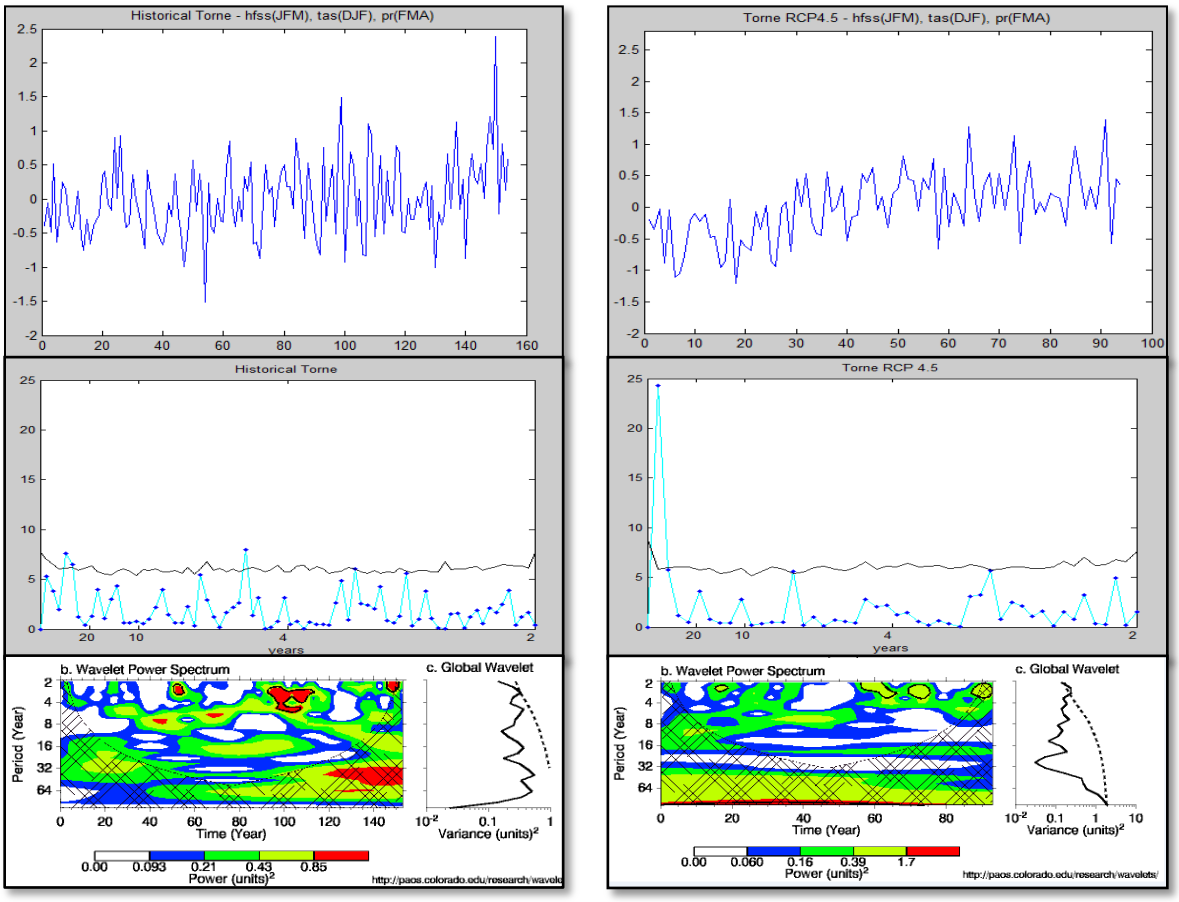

Figure 21. Fourier and wavelet analysis of discharge variability between historic period and RCP 4.5 for Torne river

Figures 22 and 23 present the frequency of the discharge variability for the Vindel and Helge rivers respectively for the historical and future period for the RCP4.5.

These rivers do not represent any visible trend in the discharge time-series, as it was already noticed in the previous case, scenario RCP2.6.

Vindel shows significant short-term variability in period between two and four years according to both methods during historical period, however wavelet spectrum indicates as well that there is variability of eight years. Long-term variability is significant in both approaches about 30 years. For the this future scenario short-term variability of 3 years is still present and powerful. However 30 years variability is no longer present, instead there is 60 years discharge variability which is only detected by Fourier analysis.

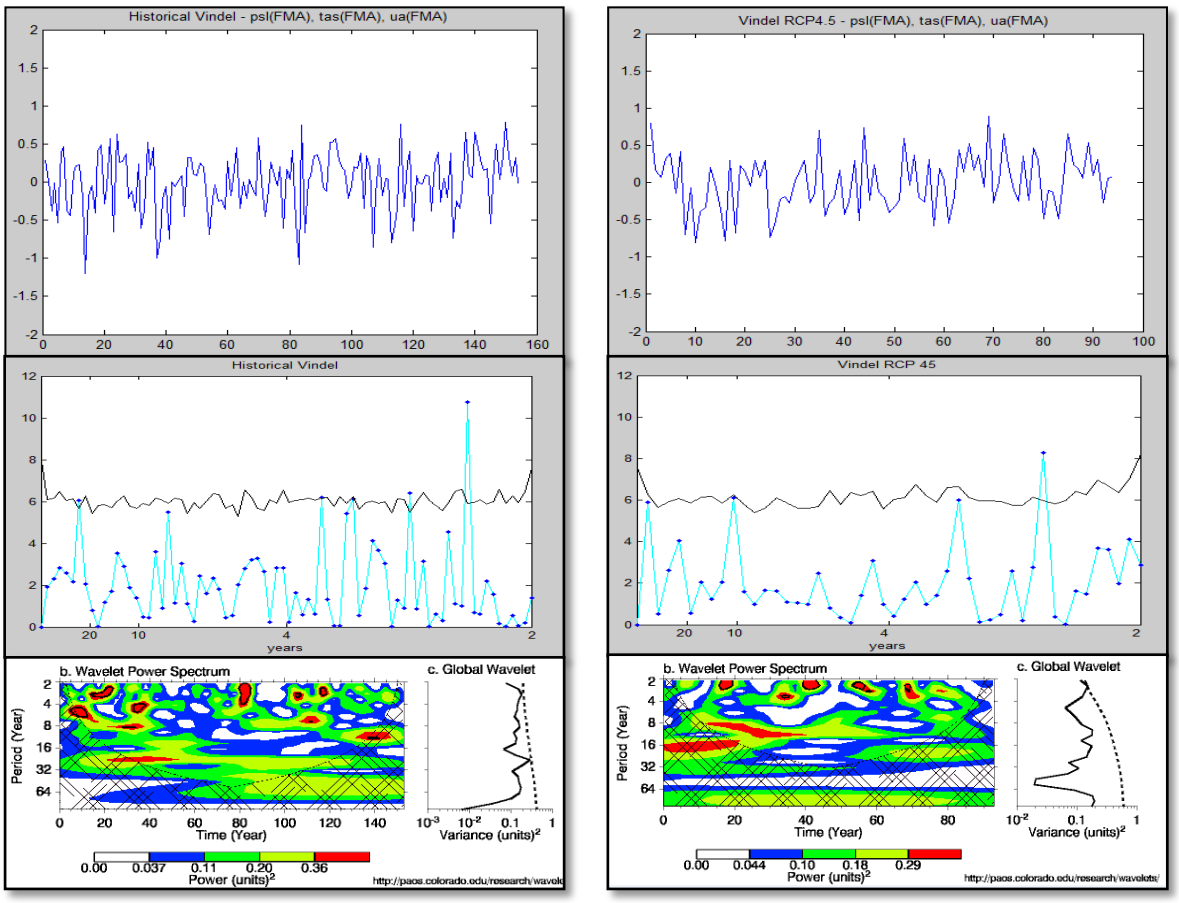

Figure 22. Fourier and wavelet analysis of discharge variability between historic period and RCP 4.5 for Vindel river

Helge river, as in RCP2.6, represent variability at many frequencies in the historical period (Figure 23). Short-term significant oscillations of 3, 4 and 6 years and long-term variability of 18 years are represented on the charts of both approaches. In the future scenario short-term variability is present mainly at 3 years, and long-term variability appears approximately in the same manner as oscillation every 18 years. However it should be noticed that previously mentioned long-term variability is not that powerful in the future scenario as in the historical period and on the wavelet spectrum a considerable part of it falls into the region padded with zeros.

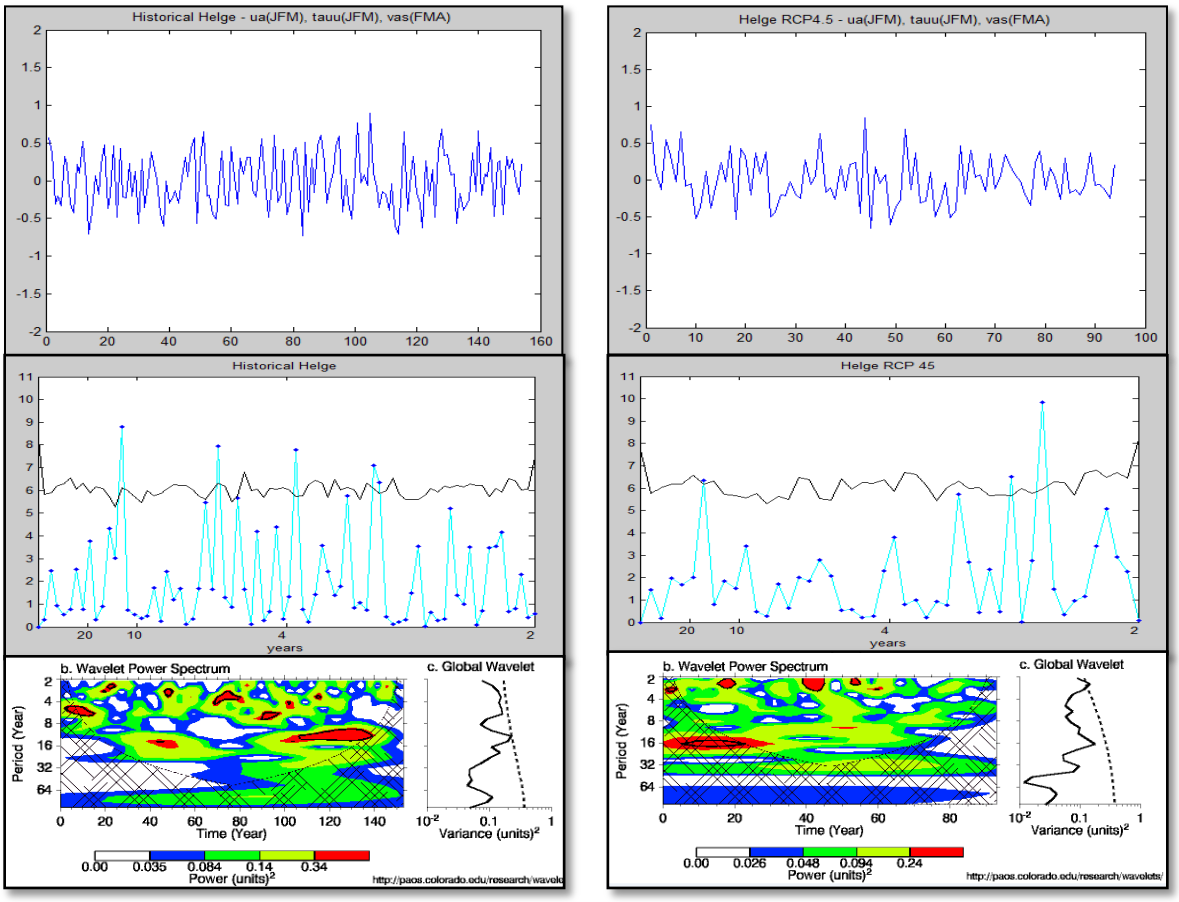

Figure 23. Fourier and wavelet analysis of discharge variability between historic period and RCP 4.5 for Helge river

III. SCENARIO RCP8.5

Figures 24 and 25 represent variability of the discharge of the Northern rivers, Kalix and Torne rivers respectively, for the historical time-series and the future scenario RCP8.5.

RCP scenario has a noticeable influence on the behavior of trends in data. RCP8.5 scenario shows significant upward trend for the Northern rivers, which is more explicit than it was perceived in previous scenarios.

Historical data of Kalix river represent a low-power oscillations between 2 and 4 years on short-term variability. Long-term variability of 65 years is clearly shown on the Fourier spectrum, but is located into the zone padded with in wavelet power spectrum. In the future scenario this situation changes: no significant short-term peaks are present, but 65-year powerful oscillation is clearly presented in the Fourier spectrum, however wavelet analysis displays this variation in the zeros region, this is inconsistent with Fourier analysis therefore it is difficult to get reliable conclusions for this frequency.

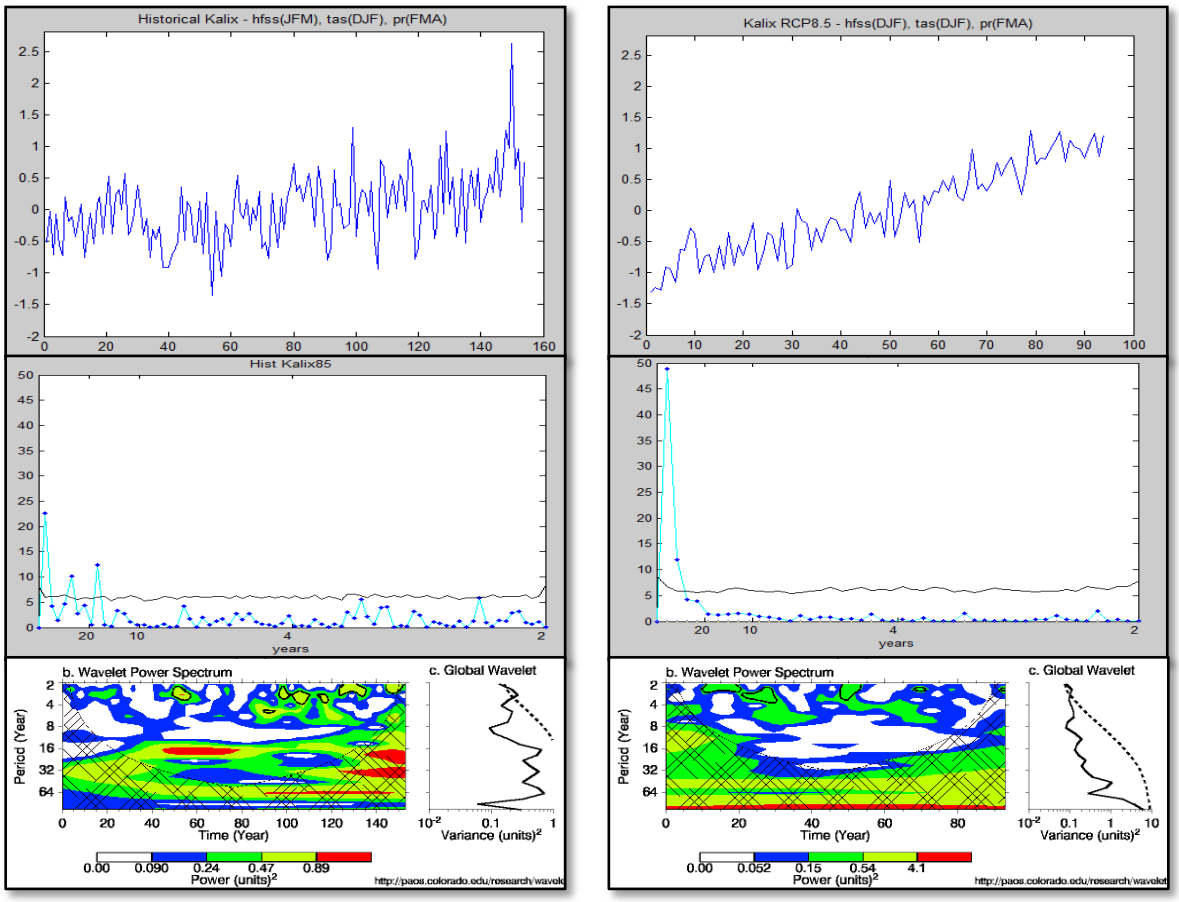

Figure 24. Fourier and wavelet analysis of discharge variability between historic period and RCP 8.5 for Kalix river

For the Torne river short-term variability between 3 and 5 years in the historical period is found. Long-term variability is present at 50 years according to the Fourier analysis and is in the padded with zeros area on the wavelet spectrum. Based on the scenario RCP 8.5 in the future period short-term variability disappears and only powerful (amplitude 40) significant oscillation of 65 years occurs according to the Fourier and in the zeros zone in the wavelet spectrum.

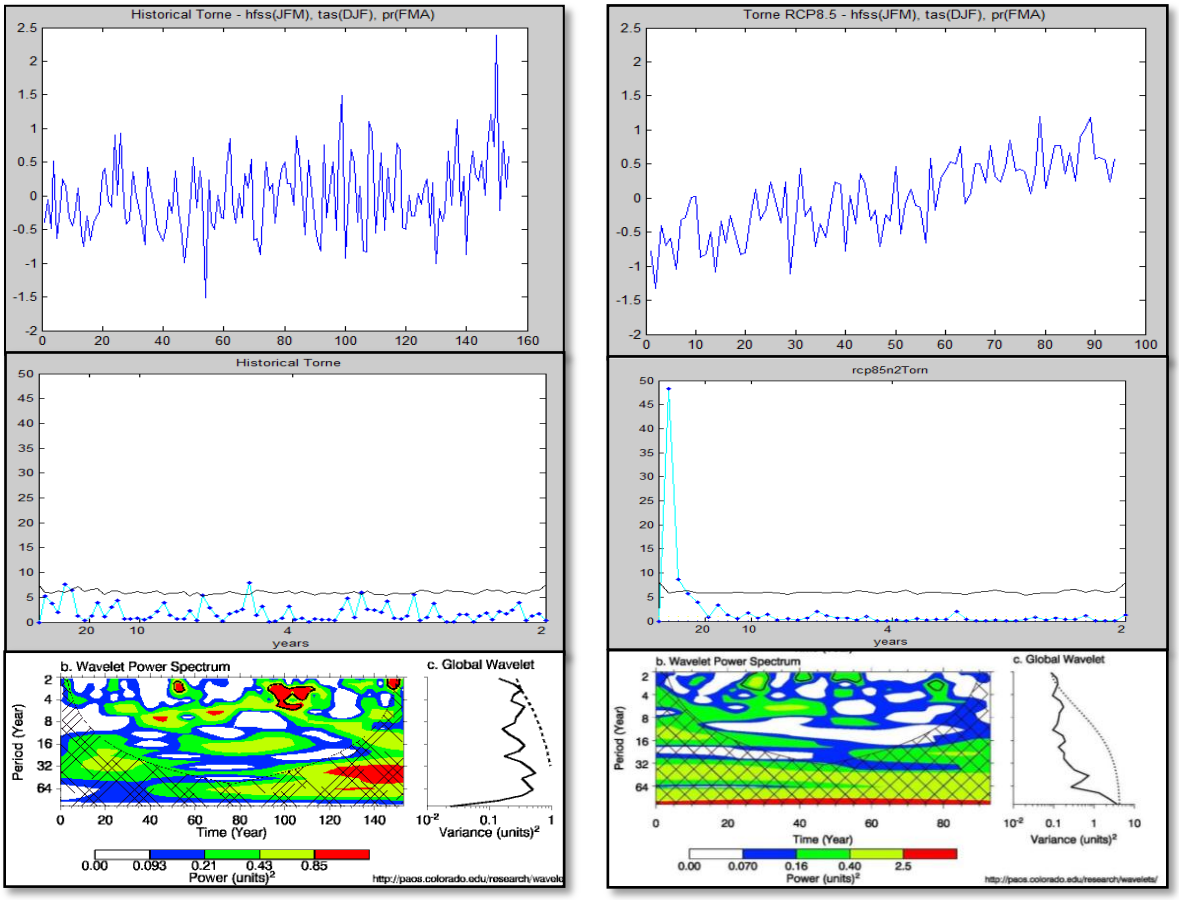

Figure 25. Fourier and wavelet analysis of discharge variability between historic period and RCP 8.5 for Torne river

Figures 26 and 27 show frequency variability of the discharge in Vindel and Helge rivers respectively, for the historical time-series and for the future scenario RCP8.5.

For the historical period of the Vindel river model results represent short-term variability of 3 and 8 years and long-term variability of 30 years that are 5% significant according to both methods. For the future period short-term peaks present lower power and are no longer significant. However Fourier transform show significant peak for the long term at 65 years, which falls into the padded with zeros zone of the wavelet analysis.

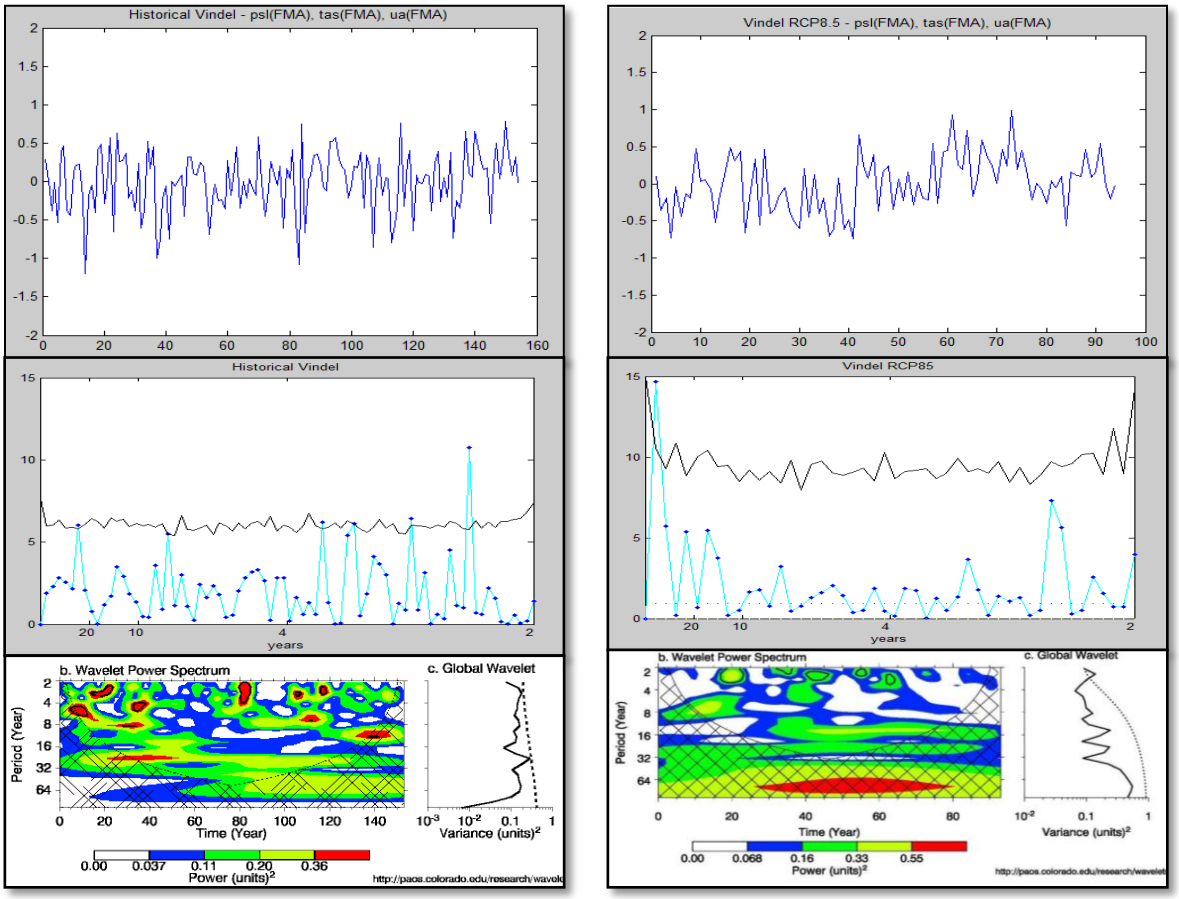

Figure 26. Fourier and wavelet analysis of discharge variability between historic period and RCP 8.5 for Vindel river

In Figure 27 historical period of Helge river presents variability of 3, 4 and 6 years and significant powerful long-term oscillation of 14 years according to the both approaches. For the future scenario model does not show any short-term variability according Fourier analysis. However the wavelet power spectrum presents variations between 2 and 4 years, which are not very powerful but still significant at 5%. Long-term variability about 50 years is as well significant at 5% based on the Fourier transform, whereas wavelet power spectrum represents this variability on the padded with zeros region.

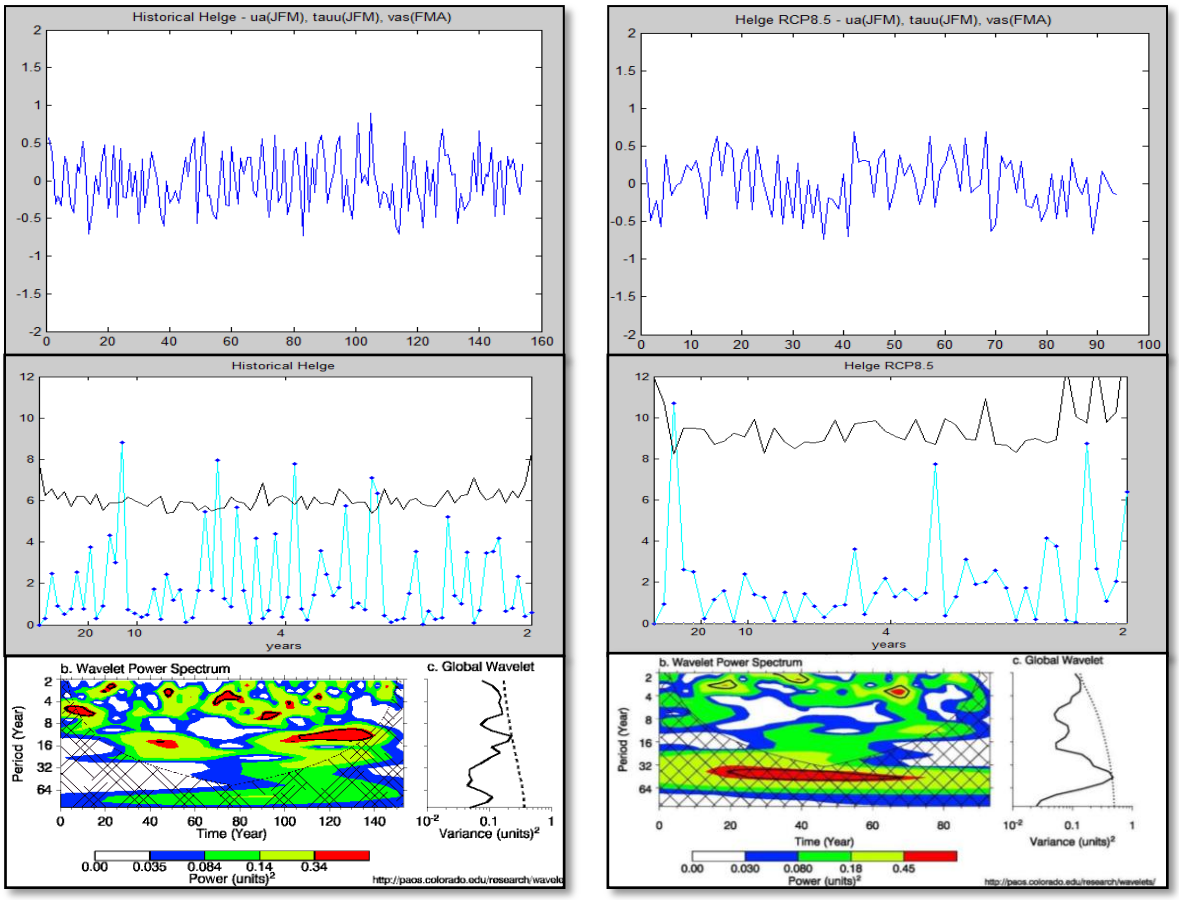

Figure 27. Fourier and wavelet analysis of discharge variability between historic period and RCP 8.5 for Helge river

7. DISCUSSION

The main goal of the present study is to identify the long-term seasonal variability of rivers discharge for past and future climates based on data obtained from the EC-Earth model. To achieve this goal data were downscaled using SVD, based on the previously identified variables that have the highest influence on the rivers discharge for each particular studied region. Variables were identified based on the nsc parameter for the further identification of the best combinations of variables based on the correlation coefficient.

The higher the nsc coefficients shown in the tables 3 and 4, the stronger the physical relation between each variable and river discharge. According to model results several relations can be found. Variables for Kalix and Torne rivers show very similar nsc values, which can be explained by their geographical proximity and therefore similar climatic conditions. The most representative season for Kalix river is DJF, in which all variables showed the highest nsc. These results can be connected to natural physical processes, in the Northern Sweden this season is characterized by the most intensive snowpack accumulation. Based on results another dependence was found. Atmospheric variables which have stronger relation to the river discharge in the North are mainly associated with temperature; such variables are surface upward sensible heat flux (hfss), near-surface air temperature (tas) and air temperature (ta). However the most representative variables in the South of Sweden are related to wind; such variables are surface downward eastward wind stress (tauu), eastward wind (ua), northward near-surface wind velocity (vas) and geopotential height (zg). This is consistent with previous study of Foster and Uvo (2010) carried at different basins in Norway. In their research it was found that climate in the South of the country is strongly influenced by the westerly winds coming from the Atlantic Ocean. Westerly and south-westerly winds bring frequent storms and moist, which turns into precipitation and to a greater extend affect the Southern part of Sweden in comparison to the Northern. In addition, the South of Sweden receives Northward winds from the Baltic. The driven forces in the middle part result from a combination of both: temperature and wind.

Correlation coefficients between observed discharge and historical discharge obtained through the SVD model are considerably different between the Northern, middle and Southern parts of Sweden. In the North correlation coefficient is equal to about 0.65, whereas in the middle part of the country and in the South it is about 0.40. This may be caused due to the fact that Northern rivers have stronger and more constant signal of the flood season in comparison with the middle and Southern parts, since in the North the period of snow accumulation is much longer and more intensive than it is in the middle and in the South.

Based on nsc and correlation coefficient the best variables were chosen. The best combinations were running through the SVD in order to create statistical model for producing a new time series of the discharge for the future period, which together with historical time-series served as input for the Fourier and Wavelet analysis.

Table 5 represent a summary of the obtained model results of the Fourier and wavelet analysis, where three different scenarios are divided according to the long- or short-term variability and four selected rivers with respective amplitude and frequency. Arrows and signs (+/-) show behavior of the frequency and amplitude of oscillations. Numbers are organized in the way that values before dash represent historical period and after dash future period. It should be noticed that summarized results just take into account the longest frequency regarding long-term variability.

Table 5. Summary of the results

		RCP2.6		RCP4.5		RCP8.5	
		Long-term	Short-term	Long-term	Short-term	Long-term	Short-term
KALIX	Frequency	= 65-65	↓ 3-0	= 65-65	↓ 3-0	= 65-65	↓ 3-0
	Amplitude	(-) 23-15	(-) 6-0	(+) 23-37	(-) 6-0	(+) 23-50	(-) 6-0
TORNE	Frequency	↓ 50-30	↓ 5,3-0	↑ 50-65	↓ 5,3-3	↑ 50-65	↓ 5,3-0
	Amplitude	(+) 8-9	(-) 7,6-0,0	(+) 8-25	(-) 7,6-6,6	(+) 8-50	(-) 7,6-0,0
VINDEL	Frequency	↓ 30-0	↓ 8,3-0,3	↑ 30-60	↓ 8,3-3,0	↑ 30-65	↓ 8,3-0
	Amplitude	(-) 6-0	(-/+) 10,6-0,10	(=) 6=6	(-)10,6-8,0	(+) 6-15	(-) 10,6-0,0
HELGE	Frequency	↓ 14-0	↓ 6,4,3-0	↑ 14-18	↓ 6,4,3-0,3,3	↑ 14-50	↓ 6,4,3-0
	Amplitude	(-) 9-0	(-) 8,7-0,0	(-) 9-6	(+/-) 8,7-9,6	(+) 9-11	(-) 8,7-0,0

From table 5 it can be seen that RCP4.5 and RCP8.5 scenarios show the same behavior regarding variability: long-term variability increasing in the future and short-term variability tends to disappear, whereas scenario RCP2.6 shows opposite situation for the long-term variability. This is quite debating situation, since it was expected that this scenario will show the same behavior, but perhaps with lower magnitudes. Further analysis of the long-term variability is more concentrated on the RCP4.5 and RCP8.5.

As model presents, short-term variability in the historical period becomes more relevant toward lower latitudes. In the South, short-term variability shows more significant peaks compared to Northern rivers. Such behavior can be explained by the effect of westerly winds, which affects climatic conditions in the whole Sweden, but predominantly in the South.

Another patron is that short-term variability tends to disappear in the future and it decreases with higher RCP. Usually short-term variability in Sweden can be explained by the climatic phenomenon North Atlantic Oscillation (NAO). This teleconnection pattern occurs in the North Atlantic Ocean due to difference in the Icelandic low and Azores high sea level pressure. In positive phase of NAO Sweden receives stronger westerly winds, which bring frequent storms, moist and warmth. However in negative phase of NAO, Sweden gets dry climate conditions with extreme temperatures and low rates of precipitation. Frequency of occurrence of this phenomenon is about 4-6 years (Climate Predictor Center, 2012). According to the model results short-term variability tends to disappear in the future, this may indicate some changes in the behavior of NAO. Several scenarios can be assumed. For instance, NAO will change its pathway, so it will not affect Sweden in the future; its

effect will become weaker compared to the past or it will change its frequency of occurrence, e.g. from 4-6 years to about 10 years, so it will not has influence on the short term variability. However to make any conclusion concerning such changes further researches should be done.

Results suggest that long-term variability tends to decrease towards lower latitudes in the past and to less extend in the future. In the table it is clearly shown that frequency of the long-term variability is much lower in the South compared to the North. Such results might be affected by some weaknesses of the model simulations. However, since EC-Earth is a recent model, not that many studies have been conducted in order to make conclusions about representativeness of the reality of the model. It is common that GCM present bias at extreme latitudes, this could be one reason for such differences of the long term variability in the North and South of Sweden. In addition, although Kalix and Torne rivers are closely located, there are noticeable difference in the power and frequency of oscillations for the historical period. This variance can be explained by differences in terrain, for instance in highlands changes such as accumulation and melting of snow, occur faster than in flat areas. However deeper understanding of the model is required and particular attention with high criticism should be given to the Northern regions, in order to understand if obtained results are reliable or they are highly influenced by limitations of model simulation at high latitudes.

Model results show another relation, long-term variability increases in frequency in the future compared to the past and tends to increase in magnitude with higher RCPs, as it can be noticed from the results RCP4.5 and RCP8.5 scenarios. Usually long-term variability can be explained by the climatic phenomenon Pacific Decadal Oscillation (PDO). This phenomenon occurs with periodicity of twenty to thirty years and is manifested through changes in sea surface temperature at the Northern Pacific, affecting climate in a large scale (Joint Institute for the Study of the Atmosphere and Ocean, 2000). Taking into consideration the model results, it can be seen that frequency of oscillations will become larger up to 65 years. Such results might indicate about possible changes in frequency of occurrence of PDO. Further researches should be conducted to confirm this hypothesis.

8. UNCERTAINTIES

It should be noticed that results may include some degree of uncertainty. The SVD model was created based on the historical data, therefore for the future period, conditions are assumed to be stationary, the same as it in historical period. Due to this reason many factors that evolve with time may have different pattern of extrapolation for the future.

Another point which can lead to some degree of uncertainty is that the length of periods of observation are different, therefore observations were extended by linear regression in

order to encompass the same period as the model data (1851-2005). These actions give some uncertainty during the creation of the statistical model for the future time-series and in the wavelet analysis. Since data-series with shorter period of time will be more sensitive in relation to the area padded with zeros.

Trend in data, mainly represented in the time-series for the Northern rivers, might affect power of the long-term variability. To obtain higher accuracy of the magnitude of oscillations data should be detrended. But the main purpose of the given study was to identify changes in frequency of the decadal variability, therefore data were not detrended.

Mentioned uncertainties have certain influence on the obtained results, but they are not relevant constrains to achieve the main goal of this study.

9. CONCLUSIONS

The main purpose of this study was to identify changes in the long-term variability of discharge in Swedish rivers by comparing past and future climates using wavelet analysis and Fourier transform. Data were downscaled through the SVD method using variables, which have the biggest influence on the rivers discharge in each of the studied regions.

Variables, which showed high values of the statistical parameters nsc and correlation coefficient have strong physical meaning which is related to the following factors: geographical position, period of the intensive snow-pack accumulation and effect of westerly winds.

Overall model results imply that natural variability might become weaker in the future and it is considerably affected by the radiative forcing level. As model shows, frequency of the long-term variability is expected to increase in the future compared to the past, whereas short-term variability tends to disappear and decreases with higher RCP scenario. Frequency of long-term variability will become shorter toward lower latitudes, whereas frequency of short-term variability becomes considerably higher in the in South in comparison to the North for the past period.

There are no similar studies to compare with for Swedish conditions. This study represents a good starting point for further researches. Following suggestion can be made for the future studies:

- Remove trend from time-series, which is more clearly represented in the discharge-data for the Northern rivers, since it might affect amplitude magnitude of the long-term variability. One of the options can be subtraction of low order polynomial fit to data. This will give more reliable results in relation of amplitude of the oscillations.
- Extend research through the addition of several family members or/and add more GCM to have more confidence in the obtained results.

REFERENCES

Bates, B.C., Kundzewicz, Z.W, Wu, S. and Palutikof J.P. 2008. Climate Change and Water. Technical Paper of the Intergovernmental Panel on Climate Change, IPCC Secretariat, Geneva, p.210.

Bintanja,W. 2011. *EC-EARTH: goals, developments and scientific perspectives.* [online] KNMI. Available at: <http://www.knmi.nl> [Accessed 07 April 2012].

Brandt M. 2010. EC-EARTH: the European Community Earth system model. [online] IS-ENES. Available at: <https://verc.enes.org> [Accessed 07 April 2012].

Bretherton, C. S., Smith, C. and Wallace, J. M., 1992. Intercomparison of Methods for Finding Coupled Patterns in Climate Data. *Journal of Climate*, 5, pp. 541-560

Christensen O.B., Gutowski B. and Nikulin G., 2012. *CORDEX Archive Design, version 26/1/2012*. [online] Available at: <http://cordex.dmi.dk/joomla/images/CORDEX/cordex_archive_specifications_120126.pdf> [Accessed 10 April 2012].

Climate Predictor Center, 2012. *North Atlantic Oscillation*. [online] Available at: <http://www.cpc.ncep.noaa.gov/data/teledoc/nao.shtml> [Accessed 06 June 2012].

Hazeleger, W. et al. 2010. *EC-Earth: A Seamless Earth-System Prediction Approach in Action.* [online]KNMI. Available at: <http://www.knmi.nl> [Accessed 16 April 2012].

Hay, L.E. and McCabe, G.J. 2010. Hydrologic Effects of Climate Change on the Yukon River Basin: Climatic Change. *Springer,* v. 100, no. 3-4, pp.509-523.

Hewitson, B.C. and Crane, R. G., 2006. Consensus between GCM climate change projections with empirical downscaling: Precipitation downscaling over South Africa, International Journal of Climatology, 26, pp. 1315-1337.

Houghton, J.T., Callander, B.A. and Varney, S.K. 1992. *Climate change 1992: The supplementary report to the IPCC assessment.* Cambridge: Cambridge University Press.

Houghton, J.T., Jenkins, G.J. and Ephraums, J.J. 1990. *Climate change. IPCC assessment.* Cambridge: Cambridge University Press.

Houghton, J.T., Meira Filho, L.G., Callander, B.A., Harris, N., Kattenberg, A. and Maskell, K. 1995. *Climate change 1995: The science of climate change: contribution of Working*

*group I to the second assessment report of the Intergovernmental Panel on Climate Change. Ca*mbridge: Cambridge University Press.

Hurk, B. van den et al., 2009. *The EC-Earth modelling challenge.* [online]KNMI. Available at: <http://www.knmi.nl> [Accessed 15 April 2012].

EC-Earth, 2010. *About EC-Earth*. [online] Available at: < http://ecearth.knmi.nl > [Accessed 10 April 2012].

Foster, K. and Uvo, C.B. 2010. Seasonal streamflow forecast: a GCM multi-model downscaling approach. *Hydrology Research* 41.6, pp.503-507.

Loaiciga, H.A., Valdes, J.B., Vogel, R., Garvey, J. and Schwarz, H. 1996. Global warming and the hydrologic cycle. *Journal of Hydrology* 174, pp.83–127.

Moss, R.H. et al., 2010. The next generation of scenarios for climate change research and assessment. *Nature*, 463, p.747-756.

Olsson, J. 2010. *Hydroimpacts2.0: Strategic research on the hydrological climate change impact.* [online] SMHI. Available at: < http://www.smhi.se > [Accessed 06 April 2012].

Rossby Centre, 2010. *EC-Earth and CMIP5*. [online] SMHI. Available at: <http://www.smhi.se > [Accessed 06 April 2012].

Rossby Centre, 2009. *Rossby Centre Newsletter*. [online] SMHI. Available at: <http://www.smhi.se> [Accessed 30 April 2012].

Slute, M., Clement, A. and Lohmann, G. 2001. Global climate models: Past, present and future. *Proceedings National Academy of Sciences, USA,98,* Issue 19, pp.10529-10530.

SMHI, 1994. Svenskt Vattenarkiv: *Vattenföring i Sverige. Del 4. Vattendrag till Västerhavet*. Vattenföringsserier tom 1990. Svalöv: Grafik Studion.

Torrence C. and Compo G.P., 1998. A practical guide to Wavelet Analysis. *Bulletin of the American Meteorological Society*, 79 (1), pp.61-78.

Uvo C.B., 2012a. Fourier and Wavelets Transforms, *WR001F Linear and non-linear data analysis – Part 1: One-dimensional Analysis.* Lund University. Unpublished.

Uvo C.B., 2012b. Spectral Analysis, *WR001F Linear and non-linear data analysis – Part 1: One-dimensional Analysis.* Lund University. Unpublished.

Uvo C.B. et al., 2001. Statistical atmospheric downscaling for rainfall estimation in Kyushu Island, Japan. *Hydrology and Earth System Sciences* 5(2), pp.259-271.

Walsh, J., 2008. *Encyclopedia of Global Warming and Climate Change:* Climate Models. [online] SAGE. Available at: <http://sage-ereference.com.ludwig.lub.lu.se>.

Wilby, R.L. and Wigley, T.M.L., 2000. Precipitation predictors for downscaling: Observed and general circulation model relationships. *International journal of climatology*, 20, pp.641-661

Wilks, D.S., 2011. *Statistical Methods in the Atmospheric Sciences*. 3rd ed. New York: Academic Press.

Xu, C.Y. 1999. From GCMs to river flow: A review of downscaling methods and hydrologic modelling approaches. *Progress in Physical Geography*, 23, pp.229–249.

APENDICES

APPENDIX A: Routine to obtain nsc through SVD

The following routine, created by Uvo (2001), performs SVD between 2 fields y and z, where y is the left field (predictor) and z is the right field (predictand) as defined by Bretherton et al. (1992). The nsc between the two fields is calculated.

```
% load your data files

y = <predictor>;  % Variable containing mean seasonal values from model output
z = <predictand>; % Variable containing river observations

[nt,ny] = size(y);
[nt,nz] = size(z);

% Standardize your data

% extract mean on y
y = y - ones(nt,1)*mean(y);
% divide by std
y = y ./(ones(nt,1)*std(y));

% extract mean on z
z = z - ones(nt,1)*mean(z);
% divide by std
z = z ./(ones(nt,1)*std(z));

% preparing for the SVD calculation
% calculating cross covariance matrix

disp('Starts SVD calculation')
ccov = y' * z / (nt-1);
sumsq=sum(sum(ccov.^2));
[g,s,h] = svd(ccov,0);

% calculation of nsc (normalized squared covariance)

sigma=diag(s,0);
sigma2=sigma.^2;
ngrade=ny;
nestac=nz;
somat=ngrade*nestac;
nsc=sqrt(sigma2/somat);
fprintf('nsc'); nsc(1:1)
fid = fopen('nsc.dat','w');
fprintf(fid,'%5.3f \n',nsc);
fclose(fid);

clear fid ngrade nestac somat nsc eiv explvar sigma sigma2 s
clear sumsq
```

APPENDIX B: Routine to generate a SVD model

The following routine, developed by Uvo (2001), performs SVD between 2 fields y and z, where y is the left field (predictor) and z is the right field (predictand) as defined by Bretherton et al. (1992). This model is used to forecast z field using regression matrix A and eigenvector g.

```
% load your data files

y = <predictor>; % Variable containing mean seasonal values from the model output
z = <predictand>; % Variable containing river observations

[nty,ny] = size(y);
[nt,nz] = size(z);

% Standardize your data

% extract mean on y
y = y - ones(nty,1)*mean(y);
% divide by std
y = y ./(ones(nty,1)*std(y));

% extract mean on z
z = z - ones(nt,1)*mean(z);
% divide by std
z = z ./(ones(nt,1)*std(z));

y1=y(<observations length>,:); %

[nt,ny] = size(y1);
[nt,nz] = size(z);

% preparing for the SVD calculation
% calculation cross covariance matrix

disp('Starts SVD calculation')
ccov = y1' * z / (nt-1);
sumsq=sum(sum(ccov.^2));
[g,s,h] = svd(ccov,0);

% Calculating time series

u = y1 * g; % y time series - dimension time x mode
v = z * h;  % z time series - dimension time x mode

% Construction of the model. Calculation A such as zhat = u*A
% using only few (choose yourself) modes of the eigenvector u.

u3 = u(:,1:1);
A = inv(u3'*u3) * (u3' * z);

% for cross-val u3 becomes Y(left out year)*g(1:3)
```

```
%zhat = Y * A;

zhat = y1 *g* A;
zhat2= y*g*A;

% Correlations between predicted and observed:

for i=1:size(z,2)
    [a,p]=corrcoef([z(:,i) zhat(:,i)]);
    r1(i,:)=[a(2,1) p(2,1)];
    bb(1,i)=a(2,1);
    cc(1,i)=p(2,1);
end

aa=[1 2];
stats(1,1:2)=aa;
stats(2,1:2)=bb;
stats(3,1:2)=cc;
clear a p

save vin2_SVD2.mat stats zhat zhat2 A g
clear r

% Plotting observed discharge vs model estimated discharge
plot(z(:,1),'Color','red','DisplayName','z(:,1)','YDataSource','z(:,1)');figure(g
cf)
hold on
plot(zhat(:,1),'DisplayName','zhat(:,1)','YDataSource','zhat(:,1)');figure(gcf)
%save vin2_SVD2.fig
hold off
```

APPENDIX C: Routine to apply Fourier transformation

The following routine applies Fourier transformation to river discharge signals. It was developed by Uvo (n.d.) and is based on Torrence and Compo (1998):

```
% Insert your data here

riv = <data> % river discharge calculated with SVD
Fs = 1;      % sample rate number of observations per time unit (year)
N = length(riv);
f =(0:N/2)*Fs/N;    % frequencies
per = 1./f(2:end);  % periods
Y = fft(riv);
pow = abs(Y(1:N/2+1)).^2/(N/2*std(riv)^2);  % power ^2 normalized by
N/2variance(riv)

% statistical significance test using bootstrap 95% confidence interval

rnb = bootstrp(1000,@rednoisefft,riv); % bootstrap 1000 red noise SSA
rnb = sort(rnb);
powrn = rnb(950,:); % the 950 highest value of the 1000 spectrum at each freq

% plotting

figure
plot(f',pow,'c-', f,pow,'b.')
axis([0 .5 0 12])
hold on
plot(f,powrn,'k-')
axis([0 .5 0 12])
xlabel('years')
set(gca,'XTick',[0.05 0.1 0.25 0.5 1 2])
set(gca,'XTickLabel',[1/0.05;1/0.1;1/0.25;1/0.5;1;1/2])
```