ON CLIMATE PREDICTION : PERFORMANCE EVALUATION OF REGIONAL CLIMATE MODELS (RCMs)

SHILPY MADAN

Master's Thesis, 2012

LUND UNIVERSITY

Department of Mathematical Statistics

Faculty of Science

Abstract

In a changing climate scenario it is important to study how different components of a climate system may change. For many impact sectors, particularly relating to flooding and water resource management techniques, variation in precipitation intensity and amount are of much importance. Role of different climate model in studying these impacts is far beyond doubt. Climate modeling has evolved many folds since their introduction in 1960s to incorporate different aspects of earth system into modeling and improving their performance. This study is aimed at evaluating performance of different Regional Climate Models (RCMs) over Europe derived from Global Circulation Models (GCMs) by means of statistics of precipitation over Goteborg, Sweden.

Performance evaluation in prediction of precipitation is done using RCMs driven by two GCMs namely HadCM3Q0 and ECHAM5r3 with that to the observed values over Goteborg. Three datasets were used for the same purpose. The first set incorporated the daily mean precipitation data from 01 January 1961 to 31 December 2009. The second set is derived from this by computing total precipitation in each month for each year, i.e., from January 1961 to December 2009. Lastly, the third set accounts for annual precipitation values, i.e., from 1961 to 2009. We have assessed the statistics of data obtained from these models and compared them with those of the observed data to comprehend if these models can be used for significantly predicting future climate change and their application in hydrology.

Principle Component Analysis is used to analyze which of RCMs best track the observed precipitation statistics and can be relied upon for climatic predictions of 21st century. It was further supported by other techniques including comparison in regard to maximum daily, two-day, three-day and seven day maximum. We have used 99th percentile of daily precipitation data as our threshold value and computed the frequency of exceedances. These values were then used to fit a Poisson distribution. Annual maxima of precipitation values were used to fit Generalized Extreme Value distribution. We have also performed Fourier transform on the monthly precipitation data. Man Kendall trend analysis test was used to detect any significant trend in data. Lastly, we have compared the inter-annual variability of our sampled data using relative change and coefficient of variation approach.

The results obtained point towards the usefulness of these very high resolution regional climate models in studying various observed trends. PROMES from HadCM3Q0 comes close to the frequency spectrum of the observed data. It also follows inter annual relative change percentage of Goteborg precipitation closely and came closest to the two-day, three-day, seven-day-maximum averages of the observed data. Mann Kendall test indicated no significant trend both in observed data and model simulations. The number of exceedances above the threshold value accepted the Poisson distribution hypothesis and the mean value of PROMES data was in accordance with the observed. PCA also indicated the fact that PROMES came closest to the observed and all these results point towards achieving the RCM-PROMES that best describes the observed data statistics among the models used in study. Results also indicate that there is no one model which could be relied on correct predictions but significant improvement have taken place by incorporating different aspects of earth systems in modern day climate models.

Acknowledgements

I would like to extend my deepest regard and thanks to my supervisor Arun Rana for his valuable guidance and great support during this project. He inspired me greatly and aroused my interest in the field of climate change and urban hydrology.

I would also like to thank Ole Bøssing Christensen and Martin Drews at the Danish Meteorological Institute for their invaluable assistance in data collection. I greatly acknowledge the ENSEMBLES project, funded by the European Commission's 6th Framework Programme through contract GOCE-CT-2003-505539.

I would also like to extend my sincere gratitude and thanks to Magnus Wiktorsson for his invaluable inputs, suggestions and assistance throughout this project. I would also like to thank Cintia Uvo, Lars Bengtsson, Hanna Modin, Kean Foster, Andreas Jakobsson and Fabio Farias for their constant guidance and support. I would like to extend my most sincere regards to all members of Department of Mathematical Statistics and Department of Water Resource Engineering for their constant support and co-operation.

Finally, I would like to express my deepest regard, many thanks and appreciation to all my family and friends, who have always willingly stood by my side with all their support and encouragement.

Table of Contents

Chapter 1 : Introduction

1.1 Climate Change

Climate change can be defined as a significant and lasting change in the distribution of weather patterns over time frames ranging from decades to centuries to millions of years. Climate change may be confined to a specific geographic boundary or may occur across the whole globe. It is having profound effects on various environmental factors including rainfall in many regions across the globe. Precipitation is one of the most important climate variables for ecosystem, hydrology and agriculture services. It is of highest interest in the field of hydrology to assess and predict the possible climate change scenario. Decision makers in government, non-governmental organizations, and various industries and in general the public need detailed information on future climate to assess its impacts on various supporting services. It helps perceive environmental risks due to various factors, including the most sought after – anthropogenic emissions of greenhouse gases. Various projects have been undertaken by the European Union to build the climate scenarios by model designing. The PRUDENCE project ended in 2004 while the next major EU-financed regional climate modeling project ENSEMBLES ended in 2009. Immediately a new project, named CORDEX which is an international project, was launched.

The ENSEMBLES project gives the climate scenario over Europe in 25km and 50km resolution grid covering the transient periods 1951-2100 or 1951-2050 according to the A1B green house gas emission scenario. This addresses the characteristic uncertainties on a regional scale that were present in projections of future climate change that already existed. We no longer need to rely on projections from simple climate models or coarse resolution data from Atmospheric-Ocean General Circulation Models (AOGCMs) to study the potential impacts of climate change. Simple models only include the physical representation of the climate system while coarse resolution precludes simulating extreme events and detailed spatial patterns of variable like temperature and precipitation over heterogeneous surfaces like Mediterranean or the Scandinavia. Since GCMs have a relatively low resolution ~ 200-500 km, they have obvious limitations in providing fine regional scale information. In particular, this limitation is felt while studying indices related to extremes which are most relevant in studying climate change (Kunkel et al, 1999). Therefore, statistical downscaling methods are used to fill this gap. In this study we focus on statistics of output generated from Regional Climate Models (RCMs) using dynamic downscaling for comparing proximity to the observed climate on fine spatial scale. There are 19 Regional Climate Models data that is accessible through ENSEMBLES project. These models have been run by different institutes across the European Union by using output from various GCM simulations, defining their own boundary conditions. These models are based on similar physical relations as GCMs but work on a high resolution of about 25-50km. RCMs add two different type of small scale information to the GCMs results. Firstly, they add information about the local conditions at specific locations. This is important when there exists a large horizontal topographic gradient or the coastline. Secondly, they add information about some small scale processes which are not specific to the location but important in study, for example frontal systems, small scale convective precipitation and other meso-scale phenomenon (Lenderink et al., 2007). Readers are suggested to refer to http://ensemblesrt3.dmi.dk/ for more information on ENSEMBLES project.

Urbanization, industrialization and human activities are leading us towards climate change - which have bearing consequences towards temperature and precipitation variability. Some of the early works proved influence of urbanization on atmospheric properties; of them noteworthy were those undertaken by Oke (1973, 1979, 1982), Karaca et al, (1995) Zhou et al (2004). These factors have also been considered in the ENSEMBLES project and hence one might not consider the simulations of various RCMs as predictions but scenario. This is because human activity and effects of urbanization cannot be predicted and thence, one cannot model these factors with confidence.

Various studies suggest an already observable climate change over the course of century while some studies point towards expected changes in future. It is in regard of the climate change that the estimated global land precipitation increased by approximately 2% over the course of the 20th century (New, et al (2001), Smith, et al (2006)) The Science Daily issue of Dec'2002 also projected increase in rainfall variability across the world resulting from changes in global climate. Study by Dore and Mohammed (2005) highlighted broad implications for future global precipitation variance, suggesting that several regional precipitation trends can already be detected and are likely to increase in the future due to climate change. The effects of global warming can already be seen. Wet regions are experiencing higher levels of precipitation, and arid areas are witnessing reduced levels and becoming drier. Changes in precipitation are one of the most noteworthy and expected impacts of climate change. This study concluded that the changes are already observable – and are likely to intensify with additional warming.

Moreover, looking inward the climate change is expected to have a strong impact on Swedish hydrology. Sweden's geographic location in the world map puts it in a danger zone of experiencing extreme weather situations. Climate projections indicate higher temperatures, especially in the north and during winter. More precipitation with higher intensity is likely. The sea level is also expected to rise (Olsson, 2011). The Commission on Climate and Vulnerability was appointed by the Swedish Government in June 2005 to assess regional and local impacts of global climate change on the Swedish society. In the study it was resolved that "Sweden will become warmer and wetter" - Precipitation is likely to increase in most parts of the country mainly during the autumn, winter and spring time. In summer-time the climate will be warmer and drier, particularly in southern Sweden. The number of days of heavy precipitation is likely to increase. There will be significant increase in extreme rainfall events as suggested in the study. This may lead to greater risk of flooding and landslides. There are a few studies, which have examined precipitation variability in Sweden - C.C. Wallen (1958) studied that Sweden has much lower variability than should be expected from the world's conditions (See Fig 1. for reference). In studies of variability and fluctuations of precipitation in Sweden carried out by Angstrom (1941), it was concluded that overall trends of increasing precipitation exist in most parts of Sweden and especially in the north.

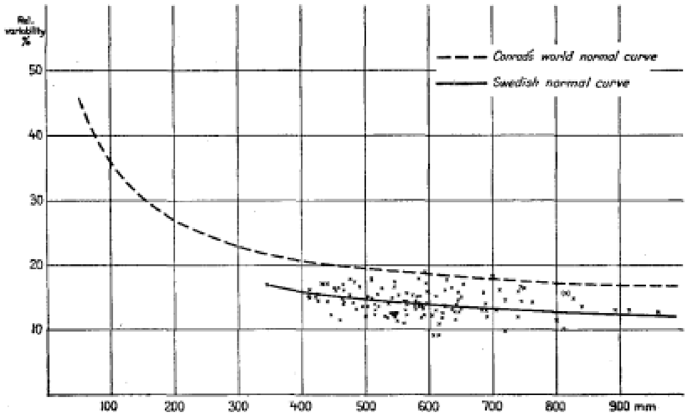


Fig 1: Isoanomals of relative variability of annual precipitation in Sweden. Anomalies with regard to world normal distribution of relative variability of precipitation according to CONRAD (1941)

Although precipitation is not easy to summarize as an average for the whole country, it has been indicated that precipitation increased significantly in Sweden during the 20th century. The lower levels of precipitation before 1920 was may be due to the fact that precipitation gauges used were different then, and their location was often more exposed to the wind. However, the upswing after the 1970s is undisputable. It was also found out that there was a positive trend in the annual precipitation in Sweden over the 1890–1990 time periods due to an upward shift around 1924 (Busuioc et al.,2001). The atmospheric circulation is the main factor for inducing regional climate variability, which exhibits variability not only on year-to-year time domain but also on decadal time domain. An important source of inter-annual variability in the atmospheric circulation includes North Atlantic Oscillation among others. Walker and Bliss (1932) established correlation between a NAO index and precipitation at a number of stations in Europe, while Hurrell (1995) provided a systematic study on the relationship between the NAO and surface temperatures and precipitation for various European regions. According to the IPCC, an increase in the average global temperature is very likely to lead to changes in precipitation and atmospheric moisture because of changes in atmospheric circulation and increases in evaporation and water vapor. Nigel W. Arnell (1999) explored the effect of climate change on hydrological regimes in Europe at the continental scale. He used a macro-scale hydrological model, with four climate change scenarios-A1, A2, B1, B2, and looked at changes in average annual runoff, seasonal and monthly runoff, and monthly low flow extremes. The main conclusions were: The four scenarios consistently pointed towards an increase in precipitation in northern Europe, and a decrease in southern Europe, with local differences.

It is thus desirable that any study on the potential consequences and changes of future climate should use the best available information about projected climates. These projections are most commonly retrieved from global climate model (GCM) simulations of the climate response to a given scenario of future greenhouse gas and aerosol emissions into the atmosphere. Future emissions and levels of greenhouse gases depend on human activities and are consequently unpredictable. In addition, the effects of changing emissions on climate are not yet thoroughly understood and hence cannot be modeled with certainty. For these reasons, model projections of climate change are described as scenarios rather than predictions (Mearns et al, 2001; Carter et al, 2001). However, with advancement of technology and knowledge, the limitation of coarse scale GCM simulations was felt. Better climate predictions were need of the hour, and hence statistical downscaling methods were used to equip decision makers with high resolution climate scenarios.

In our study, we wish to seek a model which captures the present day climate in the best possible way. Various RCMs have been studied before, assessments have been made, scenarios predicted and impacts studied. Here we compare five RCMs with the historic data (1961-2009) of Goteborg city of Sweden using various statistical measures. We intend to find a model that can be used for studying climatic scenario, whose statistics can give an insight into what we may expect in our future climate. The statistical significance of various parametric and non-parametric tests would be checked in order to conclude if any such model exists whose simulations can be relied upon for predicting a climatic scenario that awaits us.

1.2 Literature Review

This is the age of drastically observable climate change, where each country in the world is making policy changes to account for emission cuts of harmful greenhouse gases and more allocation of funds to study fuel alternatives and climate scenarios. Never before has so much emphasis been given on reducing Carbon Footprint of individuals, of organizations, and of countries. It is only imperative that climate prediction has become necessary and requires state of the art technology. To the aid of decision makers are now various GCMs and RCMs simulations which give climate scenarios until 2100. Various researchers have assessed the usefulness of these models.

Some of the earliest studies of the potential impacts of global warming in Europe were based on idealized GCM simulations. Some studies used results from only one model to illustrate potential impacts (e.g. Emanuel et al, 1985 for global natural vegetation zones) and some used a range of models for impact studies to ensure consistency (e.g., the IIASA/UNEP studies of impacts on agriculture in selected countries — Parry et al, 1988) (Cited in Mitchell, et al, 2004). Other studies recognized inter-model uncertainties and adopted outputs from several GCMs (e.g. Santer, 1985 for impacts on biomass potential in western Europe; Rotmans et al, 1994 for impacts on various sectors in Europe using the ESCAPE model). Many researchers attempted to predict future ecosystem responses to climatic change using output generated from general circulation models (GCMs) and regional climate models (RCMs) (Boer et al., 1992 and Caya et al., 1995). P.D. Jones and P.A. Reid (2001) studied the plausible increase in heaviest precipitation over Britain using RCM integrations. There were some more studies that compared the effectiveness of various RCM simulations with that of the historic data. Noteworthy among them were those conducted by Jeong, et al (2011) where they studied the diurnal cycle of precipitation amount and frequency in Sweden. RCM-RCA3 model simulations were compared with the historic observed data and it was concluded that the model well captured the spatial pattern of the phase of diurnal cycle with minor discrepancies in capturing the frequency and intensity of occurrences. Climate scenario for Netherlands was generated by Attema and Lenerink (2011) in their paper mean precipitation changes and coastal effects in the Ensembles regional climate model simulations. They studied mean precipitation response and precipitation patterns from simulations from all 19 RCMs and compared it with the historic data. Some models gave expected results with minor discrepancies while some gave absolute erroneous results. In general, the models were found to give a positive bias in mean precipitation for Netherlands however, they were consistent with the Dutch change patterns. Climate scenarios of temperature and precipitation for Netherlands were generated using regional climate models (Lenderink, et al.; 2007). Future climate extreme events in the Mediterranean region were studied by comprehending the similarities in statistics obtained from RCM-PROMES with the present day climate data. Model simulations were further used to assess possible intensity and frequency of extreme events (Sanchez, et al, 2004). An inter comparison of various RCMs was done using mean climate response and inter annual variability. Models like HIRAM and RACMO were close to the mean precipitation of present day data, ARPEGE and REMO gave a positive bias while PROMES and CHRM were quite dry. Inter annual variability was in good agreement for all models except CHRM which was consistently much drier than others. Urban hydrological assessment of Kalmar city of Sweden was done by applying climate model precipitation scenario (J.Olsson, et al.; 2008). RCM-RCA3 well represented the already observed high intensities but overestimated frequency of less rainfall intensities. Relative changes in rainfall characteristics were noted using Delta Method. Model simulations have also been used in civil engineering to study attic design using future climate projections (Vahid Nik, 2010) using relative changes in humidity and data distribution. Some more studies were done on inter comparison of regional climate models and compared with model performance with present day climate (Daniela Jacob et al, 2007; Timothy D.Mitchell et al, 2004; C. Archberger, 2003;).

Extreme events are the primary signals of climate change. In our study, extreme event simulation forms an integral part for model assessment. Occurrence of extreme events is directly related to higher temperatures, intensified hydrological cycle and vigorous atmospheric motions. The magnitude and frequency of heavy precipitation events are two important parameters for any impact and planning sectors. Extreme events can push any designed system of urban drainage, road surfaces, sewer works to failure. GCM results indicate that frequency and intensity of heavy rainfall are expected to increase under enhanced greenhouse conditions, particularly in non-summer months (Hennessy et al. 1997; McGuffie et al, 1999) The ongoing global warming of environment is ought to result in an increase in events of intense short-term storms (Trenberth et al., 2003). In UK an increase in number of heavy precipitation days along with an increase in precipitation amount on these days has been indicated from the results published by Osborn et al. (2000). The simulations of climate models also point towards more intense rainfall in northwestern Europe in future (Raisanen and Joelsson, 2001). Semmler and Jacob (2004) found from simulations that levels of rains for a given return period as expected to increase by 100%. Again, an expected increase in extreme daily precipitation by almost 10-15% in large parts of Norway was studied (Skaugen et al.,2003 cited in Bengtsson and Milloti, 2010). Hernebring (2008, cited in Bengtsson and Milloti, 2010) found that there was no general trend in return level with respect to return period for a long term rainfall time series for Goteborg, Sweden. Regional precipitation statistics from 1979-1996 were compared with data from 1997-2005 for Denmark (Madsen et al, 2009 cited in Bengtsson and Milloti, 2010). The comparison showed a general increase in extreme short-term rainfall intensity of about 10%. Results from a RCM were compared with the observational data of Britain to establish credibility or uncertainties in model performance in simulating extreme events of precipitation using quantile method. The results indicated a dramatic increase in heavy precipitation over Britain in future (Jones and Reid; 2001). Thus, it is of interest to us that the RCM should abide by the magnitude and frequencies of extreme events, so that future events of certain return level can be dealt with the required infrastructure and warning systems.

Improved methods of climate interpolation will enhance our ability both to quantify effects of climate and climate variability on natural and managed ecosystems, including forests and wetlands and to forecast the possible impacts of climate change. Although there are undoubtedly many conceptual problems and practical limitations in using these high resolution climate model output for predicting ecosystem responses, however, in most cases, simulation of ecosystem responses to climate does not require an exact representation of reality, so interpolation from sparse or incomplete records is usually acceptable (David et al., 2009).

1.3 Study Area

Gothenburg (Swedish: Göteborg) is the second largest city in Sweden by population. It is inhabited by approximately 937,000 people. It is situated on the west coast of Sweden and lies by the North Sea at the mouth of the river Göta Älv. The Gota river valley divides the area into western and eastern parts. Due to the Gulf Stream the city has a mild climate and quite a lot of rain. The Gulf Stream, together with its northern extension towards Europe, the North Atlantic Drift, is a powerful, warm, and swift Atlantic ocean current that originates at the tip of Florida, and follows the eastern coastlines of the United States and Newfoundland before crossing the Atlantic Ocean. Gothenburg has an oceanic climate according to Köppen climate classification. This kind of climate is often dominated entire year by the polar front, leading to changeable, often overcast weather. Despite its high northern latitude, temperatures are quite mild throughout the year and much warmer than places in similar latitude, mainly because of the moderating influence of the warm Gulf Stream. Goteborg lies at 57°42′N, 11°55′E on the longitude-latitude grid. The Sky View Factor (SVF) varies between 0.22 to 0.96 (Bjorn et al, 2006). On an average, the warmest month is July and coolest being January with temperatures touching 20oC and -5oC respectively. November is the wettest month, receiving an average precipitation of 84mm while February is the driest with only 41mm of rainfall. (World weather and Climate Information, http://www.weather-and-climate.com)

Goteborg has been a region of interest for climate studies in Sweden. Sweden is one of the countries in the world where reliable and relatively long time series of climate data is easily available. The Regional Climate Group (RCG) is primarily studying the local climate and its impact on environment in cities of Goteborg and Shanghai. The research is done at regional to global scales using data generated from global or regional climate models (GCMs/RCMs) and paleoclimate proxy records. The data for daily precipitation in Goteborg from 1961-2009 was easy available and henceforth used in our study.

1.4 Motivation

There is significant amount of variability and probable increase in precipitation levels across the globe and also specifically in Sweden. Precipitation levels have a direct implication on the life style of people, living conditions, food security, land under agriculture and drainage system. It also affects the economy of the state, the import – export scenario and possible risks of drought or flood. Thus, the study of trend of precipitation becomes imperative.

There are however large uncertainties over how these changes will affect hydrological processes, mainly because of the following issues (Olsson, 2011)

1) Each future projection – given by a combination of global and regional climate models, model properties and emission scenarios - gives a different result.
2) Climate model results have some uncertainties and practical errors that may be insignificant in a climate context but may be magnified in hydrological applications.
3) Despite the high resolution of climate models, results may be of too inferior nature for direct assessment of hydrological consequences, for example in urban hydrology.

To tackle these issues, activities in the hydrological climate impact research include

1) Ensemble modeling, where up to 19 different climate projections are used to drive a hydrological model and the result is evaluated in probabilistic terms.
2) Application of a post-processing method called Distribution-Based Scaling (DBS), which modifies raw climate model output based on performance in the control period.
3) Development of statistical methods for downscaling precipitation to sub-grid scales, based on additional data from the climate model such as precipitation type and cloud cover.

In our study, we take the third probable solution to the issues listed above. There exist in total nineteen regional climate models derived from seven global climate models. It is imperative that these models should preserve the statistical properties of the observed rainfall records, such as empirical distributions, periodicity, trend if any and characteristics of extreme events (Burlando and Rosso 1996; Menabde et al. 1999, Olsson and Burlando 2002). We evaluate RCMs derived from two of the GCMs and perform statistical analysis to perceive if any of these models explain the historic data characteristics. This study will help us find the suitable model whose simulations can be relied upon for making reliable assumptions and predictions by various sectors of the society for coming future for the city of Göteborg. Not before have the models been compared with the precipitation records using so many statistical techniques. Our study is limited to Goteborg city of Sweden, the second largest city in Sweden for drawing logical results. The results reported here are also only limited to the possible analysis done on data available. It is important to mention here that the scope of this study is large and it is only unfortunate that a lot of data was not available for some models that were of interest. However, similar methodology may be applied to other models and other locations based upon availability of data.

Chapter 2 : Data Description and Methodology

2.1 About Climate Models

Predictions of future climate change require complex computer models of the climate system to represent the full range of processes and interactions that influence climate. ENSEMBLES is an integrated research project that ran from 2004 to 2009 and is coordinated by the Met Office Hadley Centre. The data archive is maintained at DMI and can be reached via http://ensemblesrt3.dmi.dk/. It has produced probabilistic projections of climate for Europe to help inform researchers, analysts, policy makers, decision makers, businesses and the public with climate information from the latest climate modeling and analysis tools. Climate models are primary tools for the study of climate, its sensitivity to external and internal forcing factors, climate variability and change. The principal tools used to model future climate change are General Circulation Models which are deterministic high resolution models of the global atmosphere-ocean system. Though there has been significant improvement in terms of model development and performance since they were first used in late 1970s but still there is lot of work being put into for their improvement. Recently, an ensemble of 19 regional climate integrations with an increased resolution of 25 km has become available through the ENSEMBLES project [Hewitt and Griggs (2004)]. These integrations are based on the A1B emission scenario and run up to year 2100 with the area of focus being Europe. RCMs work by increasing the resolution of the GCM in a small, limited area of interest. The climate (temperature, precipitation, wind etc.) calculated by the GCM is used as input at the edges of the RCM as boundary conditions. RCMs can resolve the local impacts given small scale information about orography (land height), land use etc., giving weather and climate information on a regional basis.

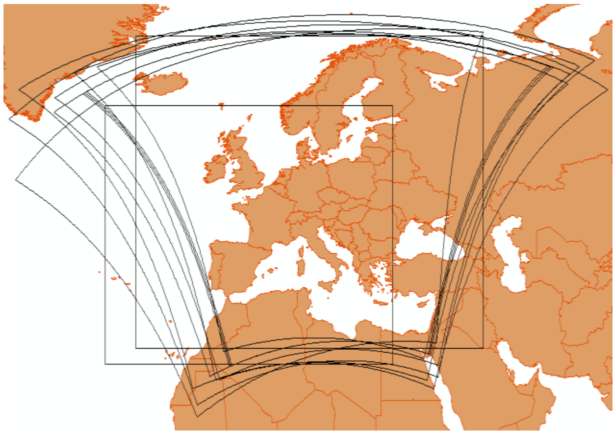


Fig 2: Spatial extent of the RCMs used in ENSEMBLES

Fig 3: Example for grid-points in a RCM

The ensemble contains 19 regional model integrations, driven by the following global climate models: HadCM3Q16 (2), HadCM3Q3 (2),HadCM3Q0 (5), ECHAM5 (6), ARPEGE (2), CGCM3 (1), and BCM (1). There are three versions of the GCM HadCM3 involved: the standard version HadCM3Q0, and two versions with adapted physical parameterizations, HadCM3Q16 being the one with high climate sensitivity and HadCM3Q3 being the one with low climate sensitivity. Different model integrations of ECHAM5 were available, but 5 out of the six regional model integrations use identical boundaries derived from ensemble member 3 (r3) of the ECHAM5 integrations using the A1B emission scenario. Thus, 7 different global model versions provide boundaries in ENSEMBLES, and among them 10 regional climate model integrations actually use boundaries from only two of these GCMs- HadCM3Q0 and ECHAMr3 (Attema and Lenderink, 2011). In addition, the climate change signal, such as the global temperature response and the response in atmospheric circulation over Europe, is also quite similar in these two integrations. Thus, we would be using the RCMs derived from these two GCMs in our study further. The factors considered in the model simulations include physical dynamic coupling, aerosol emission, radiation, horizontal-vertical atmospheric grid, microphysics, gravity wave drag, orograpy, hydrology, clouds, river routing and ocean.

HadCM3 (Hadley Centre Coupled Model, version 3) is a coupled atmosphere-ocean general circulation model (AOGCM) developed at the Hadley Centre in the United Kingdom. It was one of the main models used in the IPCC Third Assessment report in 2001. Unlike earlier AOGCMs at the Hadley Centre and elsewhere, HadCM3 does not need flux adjustment - additional artificial heat and freshwater fluxes at the ocean surface to produce a good simulation. The major factor responsible for this is higher ocean resolution of HadCM3; other factors include a good relation between the atmospheric and ocean components; and an improved ocean mixing scheme. It is composed of two components: the atmospheric model HadAM3 and the ocean model. We use the standard GCM-HadCM3Q0 driven RCMs. Simulations use a 360-day calendar, where each month is 30 days.

ECHAM5 is a coupled atmosphere-ocean GCM developed at DKRZ, the Deutsches Klimarechenzentrum GmbH and the Max-Planck Institute for Meteorology in Hamburg. This also does not require any flux adjustments between the atmosphere and ocean. The current simulation is one of the contributions to the IPCC AR4 work from the DKRZ and the Max-Planck Institute for Meteorology. It has been shown to perform well in terms of surface pressure patterns in west-central Europe indicating that the large scale circulation over Europe is realistic. RCMs driven from this GCM use a 365-day calendar for simulations.

In the selection of the RCMs the following criteria were considered:

1. Spatial extent of the RCMs cover the area of study, i.e., geographical location of Goteborg and its nearby areas.
2. Simulation results of the IPCC climate scenario A1B were readily available.

Model	Scenario	Global Model	Regional Model	Institute	Period	Legend
M1	A1	HadCM3Q0	CLM	ETHZ, Switzerland	150	⬛
M2	A1	HadCM3Q0	PROMES	UCLM, Spain	100	🟪
M3	A1	HadCM3Q0	HadRM3Q0	HC,United Kingdom	150	🟨
M4	A1	ECHAM5-r3	RACMO2	KNMI, Netherlands	150	🟩
M5	A1	ECHAM5-r3	REMO	MPI, Germany	150	🟥

Table 1: The members of the multi-model ensemble and the legend for most of the figures in this report

Regional Climate Model description:

a) RCM CLM: Climate version of the Lokal Modell (LM) is a non-hydrostatic regional climate model. It is developed by Swiss Institute of Technology, Switzerland, with a resolution of 25km grid size. A detailed description of LM is given in Jaeger et al. (2008)
b) RCM REMO: This model is based on the Europamodel from German Weather Service and is used by Max-Planck-Institute for Meteorology, Germany. It uses slightly modified physical parameterization schemes and has been tested in different climates (Semmler et al., 2004) with the focus of validation obtained from hydrological cycle (Frei et al., 2003; Lehmann et al., 2004). A detailed description of the model is given by Jacob (2001).
c) RCM PROMES: It is used by Universidad de Castilla La Mancha, Spain. It is a state of the art primitive equation model, hydrostatic and fully compressible. Output is available at a resolution of 25km in both directions; the full domain covers most of European region and Northern Africa. The PROMES model uses all principle physical processes (radiation, large-scale clouds, land surface processes, etc.) are included through parameterization. (Gallardo et al., 2001; Arribas et al., 2003)
d) RCM RACMO2: This model is used by The Royal Netherlands Meteorological Institute, Netherlands. It combines the land surface characteristics and the dynamical core of HIRLAM Numerical weather prediction system with the physical parameterization of the European Centre for Medium Range Weather Forecasting (ECMWF). Mainly the land surface scheme has been modified to increase the soil hydrological reservoir and reduce the sensitivity of canopy evaporation to drought conditions (Lenderink et al. 2003). More description on the model is given in van Meijgaard et al. (2008).
e) RCM HadRM3Q0: The Hadley Centre uses HadRM3Q0. The physical parameters incorporated in the model include calculation of large scale cloud and assumptions about the radiative effects convective clouds. However, the role of urban land surface has generally been ignored. Modifications were made to parameters in the precipitation scheme relating to precipitation efficiency to ensure reasonable vertical cloud profiles and radiation fields. More information on the model can be accessed from Collins et al. (2010)

2.2 Data Description

One of the most practiced methods while undertaking a comparative study is to take a long time series data and compare the statistics of real-time observed values with those of the predictions. For comparison with observed values of precipitation in Goteborg, three sets of data are used from the above mentioned models. The first set incorporated the daily mean precipitation data from 01-January-1961 to 31-December-2009. The second set is derived from this by computing total precipitation in each month for each year, i.e., from January 1961 to December 2009. Lastly, the third set accounts for annual precipitation values, i.e., from 1961 to 2009. The grid point data for these models was available in rotated latitude and longitude. These were converted to corresponding latitude and longitudinal decimal points. The data for grid point closest to Goteborg and eight adjoining points were extracted. The precipitation data was available in 'kg m-2 s-1' which was converted to 'm3 per day'. The average of these grid points were taken and were henceforth considered in various calculations. It can be figuratively represented as :

7 8 9

4 5 6

1 2 3

Fig 4: The centre point '5' represents grid point closest to Goteborg with the other adjoining points

2.3 Methodology

The RCMs on an average simulate some small amount of precipitation on regular daily steps. In order to standardize the output, precipitation less than 1 mm on any given day was equated to zero as a no event day. In comparison to smaller values like 0.1 or 0.5 mm, this choice of threshold makes the evaluation less sensitive to observational errors/measurement accuracy and also to the tendency of some models to give excessive occurrence of very weak precipitation (Frei, et al; 2002). Since the model simulations are not predictions but scenario, we assess the statistics obtained from these models and compare them with those of the observed data to comprehend if any of these models can be used for significantly predicting future climate change.

Statistical techniques like Principle Component Analysis is used to analyze which of RCMs best track the observed precipitation and can be relied upon for climatic predictions of 21st century. Then, we assess the empirical distribution of the observed data and model simulations. This being the primary step would help us visualize if the precipitation values over the given time scale of our different sets of data follow similar or varied distributions. We shall also see which models give similar distribution parameters as that of the observed rainfall records. Using the monthly precipitation dataset, we perform Fourier transform and check for periodicity in data. Presence of any trend in the data is analyzed using Mann Kendall Trend Analysis. In order to support the abovementioned findings we also see how the model statistics match with those of observed data with respect to average annual- daily maximum, two-day, three-day and seven day maximum. We use 99th percentile of daily precipitation data as our threshold value and compute the frequency of extreme events. These values are then used to fit Poisson distribution and assess the parameter estimate. Also, annual maxima are considered for fitting Generalized Extreme Value distribution using which we assess the Gumbel distribution hypothesis, return period and parameter estimates. Lastly, using the annual precipitation values we compare the inter-annual variability by computing the relative change percentage of our sampled data and coefficient of variation. The models that best describe the observed present day statistics can be henceforth used to analyze the changes that one may expect to see in coming future.

Chapter 3 : Results

3.1 Principle Component Analysis

Principal Component Analysis (PCA) is a multivariate statistical analysis, which attempts to simplify a complex set of interrelationships by creating one or few variables, with respect to those that allow a more convenient examination of the overall spatial relationship. Furthermore, it attempts to explain the overall variance in a data set by isolating a number of components with respect to newly defined axes, each of which corresponds to a variable (Richman, 1986; Preisendorfer, 1988; Graham, 1988). It helps us identify patterns in the data and express the data such that these similarities and differences are highlighted. Since patterns in higher dimension data can be hard to find, PCA is a powerful tool for data analysis. It can be more easily understood as a variable reduction procedure.

Hereby, our motive is to find which model simulation comes closest to the observed data set and check if our further analysis supports these findings. We use monthly precipitation set of data to carry out the analysis. In general, the number of components extracted is equal to the number of variables under consideration. Off these components one may decide which of them are worthy of being retained for further interpretation. We may expect that only first few components will account for explaining meaningful variance and later components may plausibly be ignored. Here we select four components as they explained more than 80% of the variance in data. It was found that first component explained 28.45% of variance in the data, while the second component accounted for 22.73% variance. Explained variance by each component is given in the table below.

Component	Explained Variance
1	28.45
2	22.73
3	17.98
4	12.18

Table 2: Explained variance by each component in PCA

It was found that all eigenvalues were positive. The coefficients of principle component were plotted.

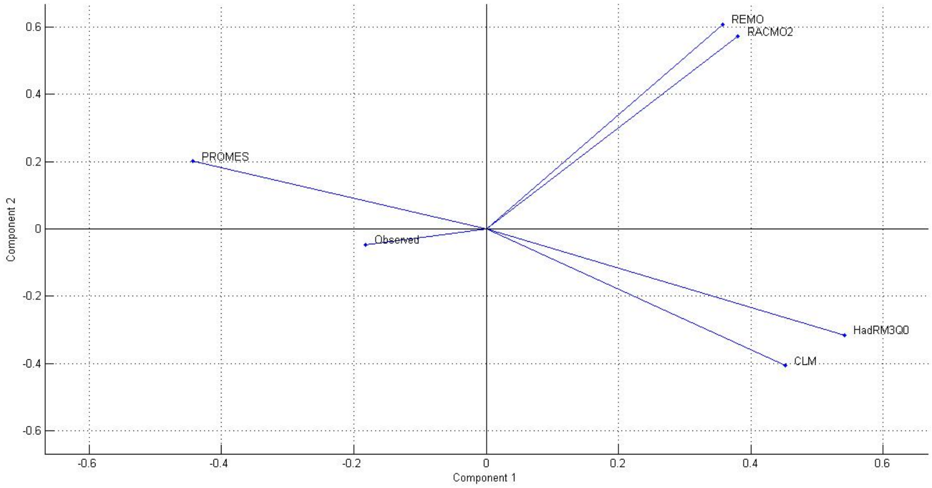


Fig 5: PCA using Component 1 and Component 2

The coefficients along principle component 1 and principle component 2 were plotted since they together accounted for more than 50% variability in the data. It can be seen that PROMES lied in the same phase as the Observed. However, Observed lied fairly close to the origin, suggesting that it is not greatly impacted by either of the two components. Since Observed is our reference point, it was important that the principle components we chose should significantly impact the Observed.

The first three principle components together explain approximately 70% variability in the data. Choosing the first 3 components and plotting the coefficients we get a 3D plot of PCA as follows:

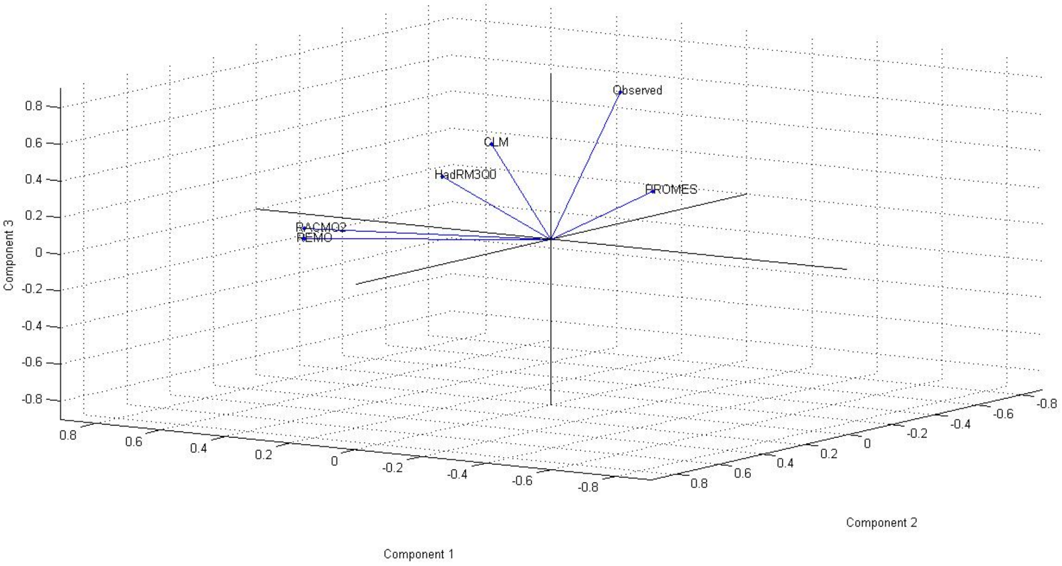


Fig 6: PCA using Component 1, Component 2 and Component 3

The 3D plot of PCA above suggests that Observed is most influenced by Component 3 among all principle components. Again PROMES lied in the same grid as the Observed while other models lied in different grids. As stated earlier, that since Observed is our reference point, it is important that we chose those components that most affect the Observed and thence carry out the comparison study. For simplicity we resort to 2D and chose principle component 1 and principle component 3. In Fig.7 it can be seen from the plot that PROMES and Observed lied in the same quadrant. This suggests that PROMES came closest to explaining the variability in Observed set. Both Observed and PROMES lie close to each other in the same quadrant while REMO, RACMO2, HadRM3Q0 and CLM lie in the adjoining quadrant.

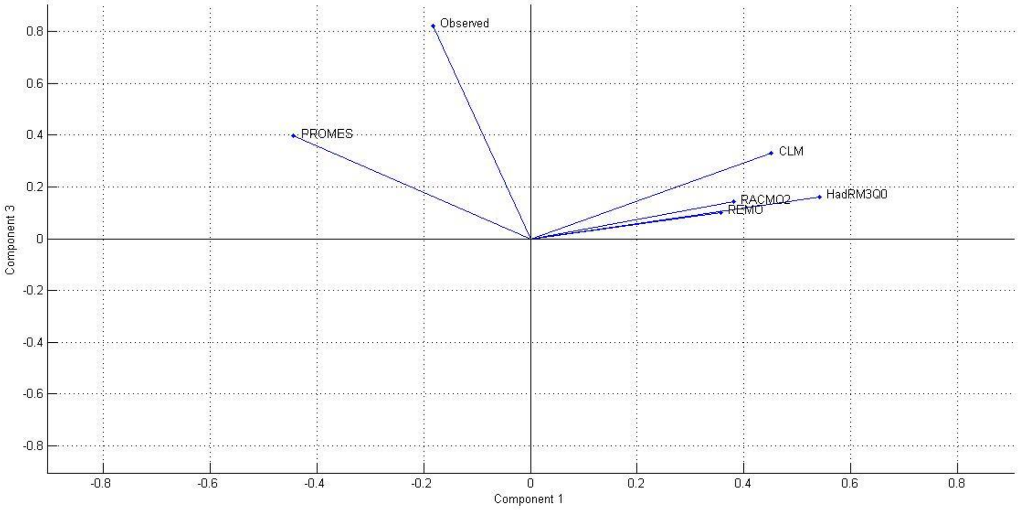


Fig 7: PCA using Component 1 and Component 3

This result suggests that simulations generated by RCM-PROMES are of similar nature as in reality. It would be henceforth studied if the statistics generated by this model also matches the statistics generated from the observed historic data.

3.2 Rainfall Distribution

The primary step before moving forward with our analysis is to check the distribution of our data. This is an important property and has been used extensively in climate study, for example Molnar and Burlando (2004) and Vorosmarty et al. (2004). Cumulative distribution $F(x)$ can be defined as the proportion of observations lying below a certain value x. The cumulative distribution for all models simulation runs is compared to observed data. It can be viewed that though all models generally fit a similar distribution as the observed, PROMES captures perfectly the end tail of the observed distribution. Also, unlike other models that give a steeper curve of the cumulative distribution function, PROMES behaves very similar to the observed distribution curve. From this plot one can say that PROMES gives a wetter climate than the actual observed while other models give a relatively drier climate. Historic data is missing for year 2004-2005 which explains the $F(0) \approx 0.05$ for the observed cdf.

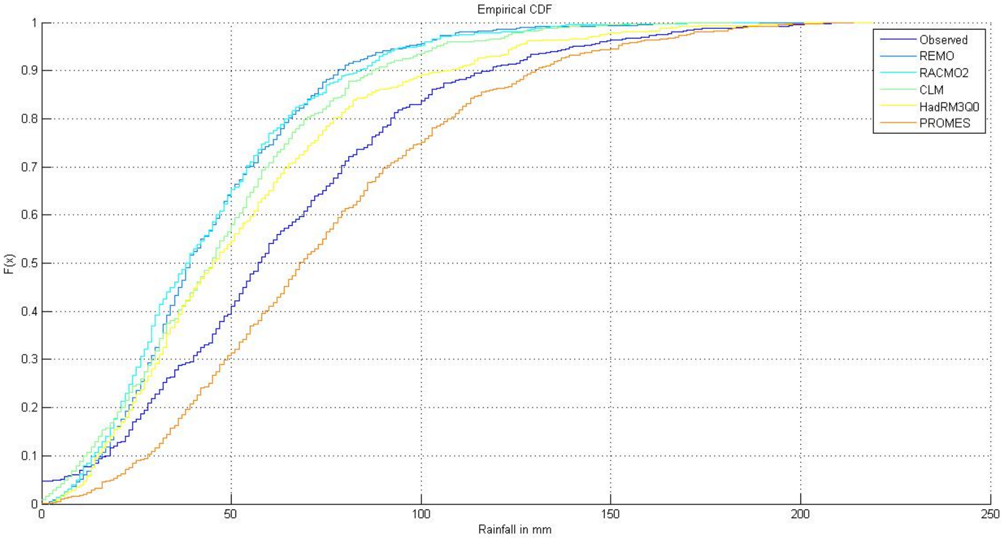


Fig 7: The cumulative distribution function of monthly rainfall amounts from observations and various model simulations.

We also check the parameter estimates obtained for each distribution fit of our models and observed data set. The first and second order moments of PROMES simulated data come closest to the moments of the observed set of data. The variability in recorded data is perfectly captured by this model simulation while there exists a minor positive bias in mean value of the distribution. Other models simulations give a negative bias in estimating the mean value of rainfall distribution and the spread of data is also not appropriately defined.

	Observed	REMO	RACMO2	CLM	HadRM3Q0	PROMES
Mean	63.33	45.12	44.32	47.98	53.76	74.60
Median	57	39	38.50	45	45	68
Std dev	40.51	27.23	28.30	30.61	36.46	40.38

Table 3: Distribution parameter estimates

3.3 Fourier Transform

The complexity of climate variability on all time scales requires the use of several redefined tools to unravel its primary dynamics from observations. One such method used most frequently is the Fourier Technique. The high resolution climate data was subjected to a Fourier transform to detect any annual or sub annual patterns.

If x_n denotes our time series data of length N, the convolution theorem allows us to do N convolutions simultaneously in Fourier space using the Discrete Fourier Transform (DFT). The DFT of x_n is given by:

$$\widehat{x_k} = \frac{1}{N}\sum_{n=0}^{N-1} x_n e^{-2\pi i k n/N}$$

where $k = 0,1,\dots,N-1$ is the frequency index.

To make it easier to compare different power spectra, it is desirable to find a common normalization index for the power spectra. Here the spectrum has been normalized by $N/2\sigma^2$ where N is the number of points and σ^2 is the variance of the time series. Thus, the normalized Fourier spectrum in Fig. is given by $(N \times |\widehat{x_k}|^2)/2\sigma^2$, where $\widehat{x_k}$ is give from the equation above. (Torrence and Compo, 1997)

The spectral analysis of the recorded observations revealed a periodic signal which appeared well above the uncertainty level in the Fourier spectrum: biennial; i.e. once every two years. This periodicity can be attributed to the 26-month Quasi-biennial oscillation or the El Nino Southern Oscillation. El Nino's effects on Europe are not entirely clear but there is evidence that El Nino may cause a wetter, cloudier winter in Nothern Europe.

It can be noted in plots below that among all models only PROMES was able to capture the periodicity in monthly precipitation very similar to the one as observed. None of the other models could do so as distinctly. Thus, backing the result of PCA, we again get the same model which best describes the real time data characteristics. Because of large amplitude, we scale our output on the y-axis to a log scale and can distinctly see the peaks generated by PROMES simulations and recorded observations at 26month periodicity.

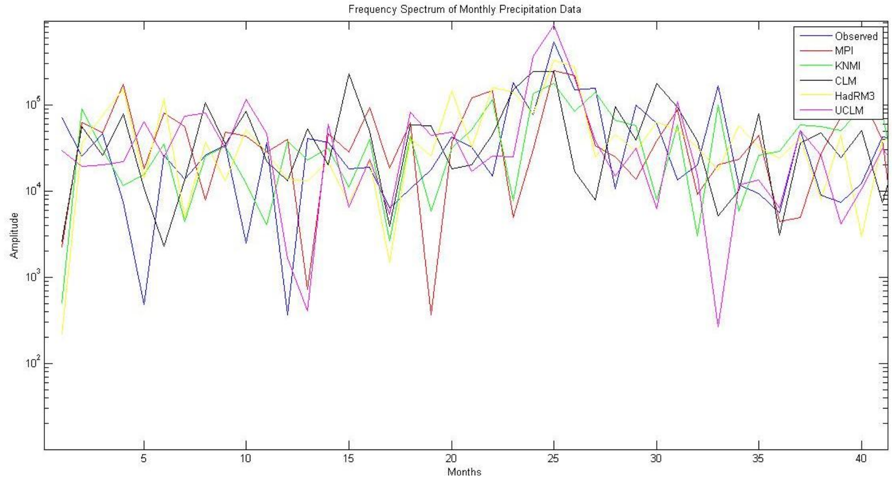


Fig 8: Frequency spectrum on the log scale

An enlarged view of the normalized Fourier spectra shows that there is a significant peak after a period of 26months for the historic data set. This frequency is captured by PROMES simulations perfectly but with minor bias in the amplitude. However, none of the other four models capture this periodicity in dataset significantly.

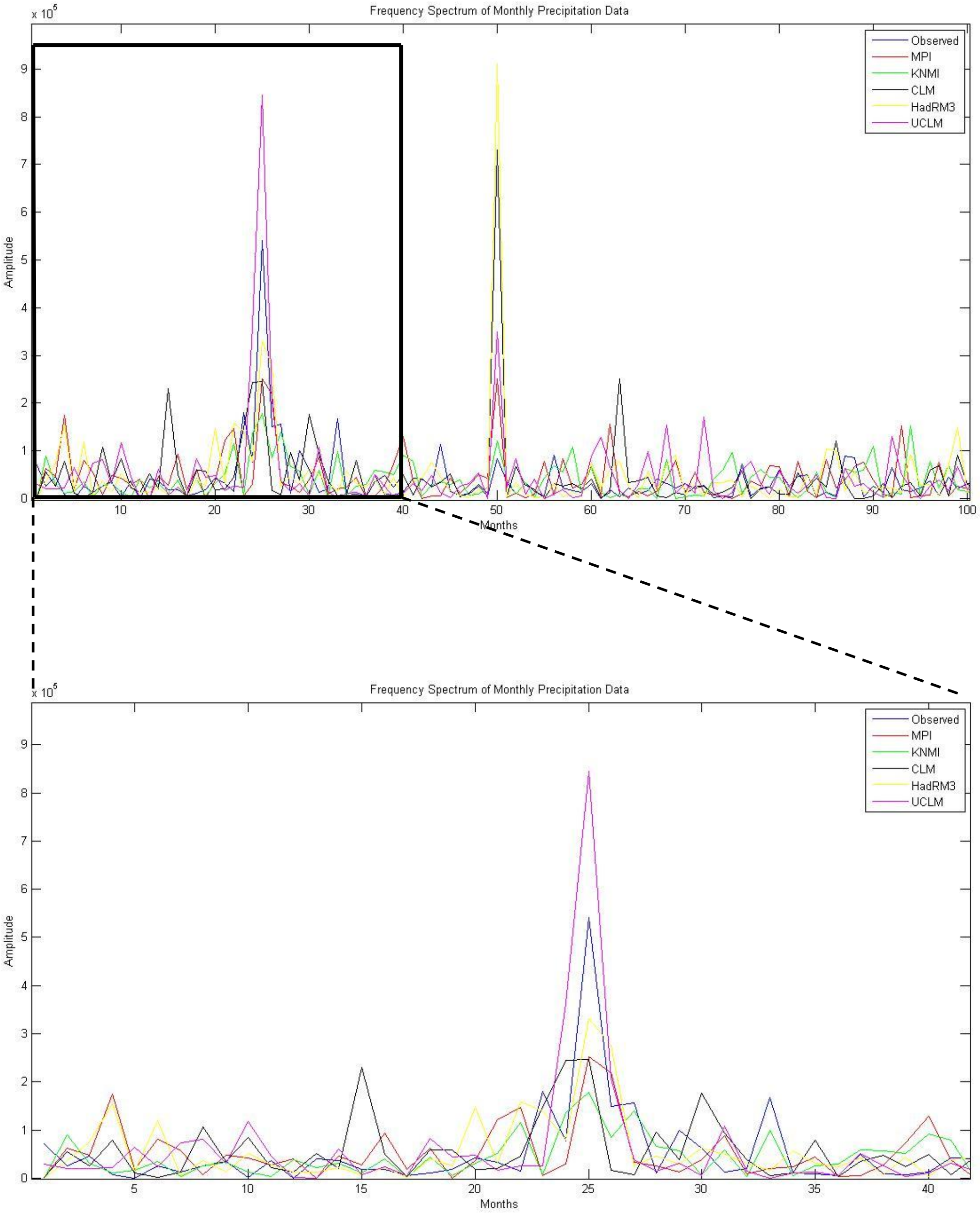


Fig 9: Normalized Frequency Spectrum of model simulated and observed data

3.4 Mann-Kendall Test

The existence of a trend in hydrological time series is detected by statistical tests. The purpose of trend testing is to determine if the values of a random variable under consideration significantly increase or decrease over some period of time (Helsel and Hirsch, 1992). Parametric or non-parametric statistical tests can be used to determine presence of any significant trend. The Mann-Kendall test is a popular statistical method used by contemporary climatologists (Kadioglu, 1997; Brunetti at al. 2000; Salinger and Griffiths 2001; Wibig and Glowicki 2002; Lu et al., 2004; Domroes and El-Tantawai, 2005 Gadgil and Dhorde 2005; Tomozeiu et al. 2006, and many others) as a significant test for checking the overall trend.

The major advantages of this test were highlighted by Rio, Penas and Fraile, 2005 as :

a) No assumptions about distribution are necessary
b) It is directly applicable to climate data for any given month or season. Also, it is not much affected by outliers because its statistic is based on the sign of differences and not directly on the values of the random variable.

The null hypothesis H0 is tested that there exists no trend versus the alternative hypothesis H1, that there is trend. The Mann-Kendall test is based on the statistic S. Each pair of observed values yi, yj (i > j) of the random variable is inspected to find whether $y_i > y_j$ or $y_i < y_j$. Let the number of former type of pairs be P, and the number of latter type of pairs be M. Then S is defined as

$$S = P - M$$

For large n, as in our case, the sampling distribution of S is as follows. Z follows the standard normal distribution where

$$Z = \begin{cases} \dfrac{S-1}{\sigma_S} & if\ S > 0 \\ 0 & if\ S = 0 \\ \dfrac{S+1}{\sigma_S} & if\ S < 0 \end{cases}$$

where $\sigma_S = \sqrt{\frac{n(n-1)(2n+5)}{18}}$

The null hypothesis that there is no trend is rejected when the computed Z is greater than $Z_{\frac{\alpha}{2}}$ in absolute value. The value of $Z_{\frac{\alpha}{2}}$ for 95% confidence level is 1.95. One may also use p value for hypothesis testing. We base our test on $\alpha = 0.05$ level of significance. We may reject H_0 if p value is less than α. We may define our hypothesis as:

H_0 : There is no trend in monthly precipiation data from $1961 - 2009$

H_1 : There exists a significant trend in data

It is of interest to us if there exists a significant trend in historic data and whether this trend is duly captured by any model simulation. Results of transient regional climate model integrations are compared with the monthly precipitation observational data set from January 1961 to December 2009. They have been tabulated as below:

Data	Trend	P value	Z value	S value
Observed	Absent	0.3178	0.9990	4755
REMO	Absent	0.9481	0.0651	311
RACMO2	Absent	0.4180	0.8099	3855
HadRM3Q0	Absent	0.6869	0.4030	1919
CLM	Absent	0.7445	0.3259	1552
PROMES	Absent	0.4122	0.8200	3903

Table 4: Mann Kendall Trend Analysis

It is observed that there does not exist any significant trend in the recorded real data. The null hypothesis is accepted at 0.05 level of significance. As in accordance with the observational data set none of the model simulations show the presence of any significant trend. It is interesting to note that the p vale, Z value and S counts of RCM-PROMES come closest to those of the observed. This analysis too lies in accordance with the PCA, empirical distribution and Fourier Analysis results.

3.5 Averages Using Daily Data

Daily maximum, two day maximum, three day maximum and seven day maximum were computed for each year and averaged over all years for each set of data. Here we focus on our results based on daily data simulations and visualize if any model gives relevant output on such a fine scale resolution.

A window of two day, three day and seven day was used to study the occurrences of extreme events and at the same time reduce the effect of outliers and substantiate our findings. It was found that all models gave fairly close value of these averages as were seen in recorded data. However, PROMES gave very similar results as the observed. It may be said that this model works well even on such fine scale resolution and captures these fractional changes well. The output generated from model simulations has been tabulated below:

(Average)	Observed	REMO	RACMO2	CLM	HadRM3Q0	PROMES
Day Max.	33.73	34.00	32.90	34.00	34.02	31.15
Two Day Max.	42.60	43.23	43.48	44.11	47.60	42.42
Three Day Max.	50.22	48.85	50.04	49.79	55.07	51.22
Seven Day Max.	74.17	62.00	64.95	65.92	76.56	78.05

Table 5: Averages

However, a more appropriate analytical tool for studying extreme events is using Generalized Extreme Value Distribution, Generalized Pareto Distribution-more commonly known as Peaks Over Threshold Method (POT) or Poisson Distribution. While GEV and GPD uses the magnitudes of extreme events, Poisson distribution is fit to the frequency of extreme events in a given time period. Extreme events act as a catalyst in climate change studies thus, any study on climate change modeling, impacts and trends is incomplete without going into fine details of extreme event analysis. We shall now study GEV and Poisson distribution in succeeding sections.

3.6 Generalized Extreme Value Distribution

Historically, precipitation extreme analysis constitutes the prime important example in environmental sciences where annual maxima based inference on return levels is the ultimate research goal. For this purpose generalized extreme value distribution is the basic tool.

One approach to working with extreme value data is to group the data into blocks of equal length and fit the data to the maximum value of each block. The type of extreme events considered here are the maximum of a sequence $E1 = \max\{Y_1, Y_2, \dots, Y_n\}$. We use annual maxima of daily precipitation amounts. The choice of block size is critical as too small block size can lead to bias and blocks that are too large generate inadequate block maxima while performing the fit. This may lead to large estimation variance. The block maxima approach is closely associated with the use of GEV family. We use the extremes package in R to fit our set of data to GEV family. All parameters are estimated by the method of Maximum Likelihood Estimation (MLE).

The Generalized Extreme Value distribution functions have the form:

$$G(y) = \exp\left(-\left(1 + \gamma \frac{y-\mu}{\sigma}\right)_+^{-\frac{1}{\gamma}}\right)$$

where $\sigma > 0, \mu \in R,\ \gamma \in R$ are the scale, location and shape parameters respectively.

We have three cases:

1. $\gamma \to 0\,; G(y) \to \Lambda(y)$
2. $\gamma > 0\,; G(y) \to \Phi(y)$
3. $\gamma < 0\,;\ G(y) \to \Psi(\mathrm{y})$

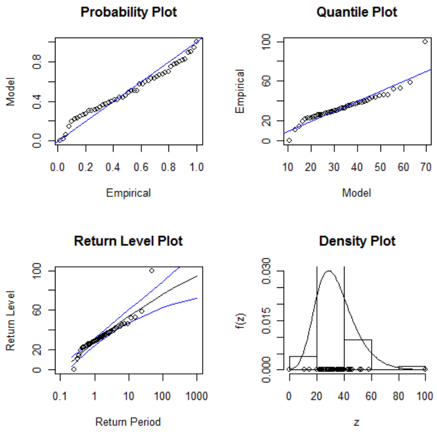
Where $\Lambda(.), \Phi(.)\ and\ \Psi(.)$ denote the Gumble distribution, Frechet distribution and Fisher distribution respectively.

It was found that all data sets fitted well to the GEV family. Except RACMO2, all data sets abided by the Gumbel distribution hypothesis. This implies that annual maximum precipitation fitted well to the Gumbel distribution from the GEV family of distributions. The parameter estimates and plots are shown below. It can be seen that CLM and PROMES data fit gave similar parameter estimates as the observed. The probability plot, quantile plot, return level plot and density plot of CLM and PROMES came closest to the ones generated by the observed data set. The quantile plot compares the model quantiles against the data quantile. Since the quantile plot does not deviate much from the straight line; it suggests that model assumptions are valid for the data plotted. The return level plot shows the return level against the return period, the blue lines being the 95% confidence interval. Here return level is the amount of precipitation in mm that is expected to exceed once in every n years. From the Observed return level plot one would expect precipitation in Goteborg to exceed 40mm once every 10 years, and the same is predicted by PROMES and CLM simulations.

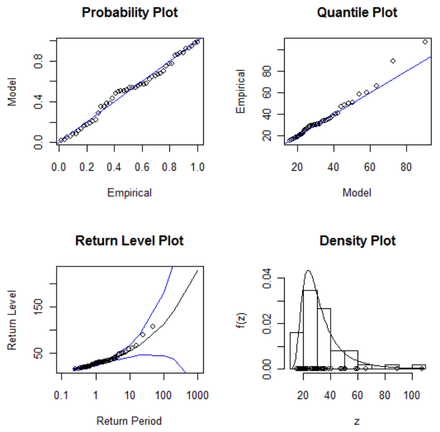
	Observed	**REMO**	**RACMO2**	**CLM**	**HadRM3Q0**	**PROMES**
Mu	27.98	25.55	28.50	29.14	30.42	28.60
Sigma	12.19	8.79	6.46	8.82	8.21	6.09
Gamma	-0.07	0.29	0.09	-0.01	-0.16	-0.19
Gumbel Hypothesis	Accept	Accept	Reject	Accept	Accept	Accept

Table 6: Maximum Likelihood Estimates of GEV parameters

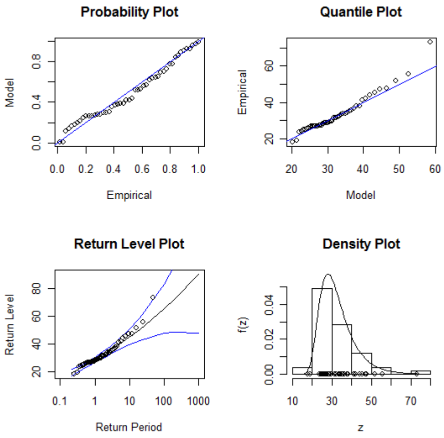
Observed



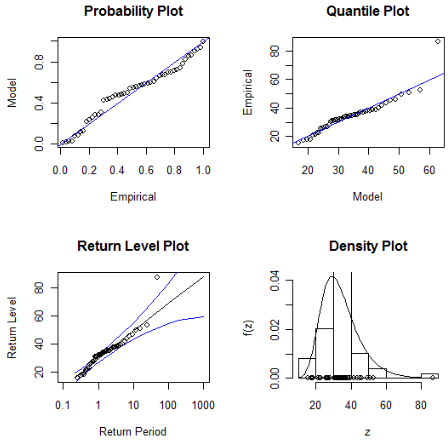
REMO



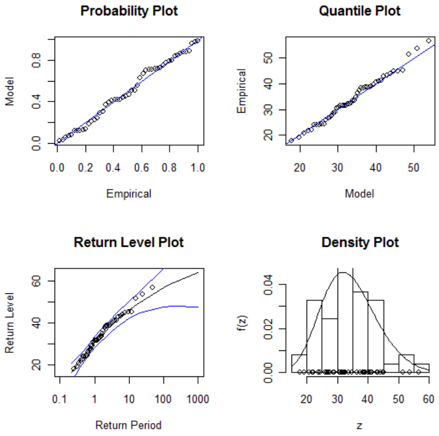
RACMO2



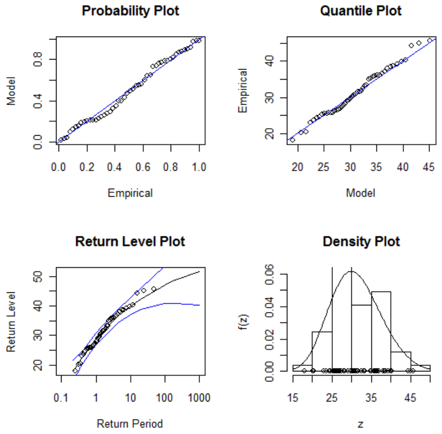
CLM



HadRM3Q0



PROMES



3.7 Poisson Distribution

Extreme events act as a catalyst over the concern if climate is changing. The study of extreme events in statistical theory has been used to demonstrate that the frequency of occurrences of such events is dependent on climate variability. Katz and Brown studied that climate models need to be designed such that they can detect the changes in climate variability. Poisson distribution is used to model the number of occurrences of an event within a given interval of time. It can be used for data that involves random sums of rare events. Here our interest lies in modeling the frequency of extremes rainfall events. We consider the extreme event of the type: exceedance from a threshold value, event $E1 = \{Y > c\}$, where constant c denotes the threshold. An extreme event was defined as the event exceeding 99th percentile of the data. From the observed daily data the threshold value was computed at 20mm of precipitation and the frequency of values exceeding this threshold were computed for each year. The fit to Poisson distribution was obtained as follows:

	Lambda	P value	Poisson hypothesis
Observed	4.65	0.09	Accept
REMO	2.18	0.06	Accept
RACMO2	2.79	0.53	Accept
CLM	3.02	0.52	Accept
HadRM3Q0	3.79	0.02	Reject
PROMES	4.24	0.65	Accept

Table 7: Poisson distribution fit to data set

It is observed that all models accept the Poisson distribution hypothesis for the frequency of exceedances except HadRM3Q0. Also, the parameter value for Poisson distribution of both observed and PROMES data set do not differ significantly. This implies that as suggested by the observed data, PROMES simulations also abide with, that on an average there would be 4-5 occurrences of such extreme events in any given year. The remaining models underestimate the frequency of such occurrences and on an average predict 2-3 events in year where daily precipitation would exceed 20mm.

3.8 Inter-Annual Varaibility

The frequency of intense precipitation and also the amount have increased substantially over the years. Since RCMs are not predictions but scenarios, it may be of interest to us if they capture the fractional changes as well. The changes in the large scale precipitation pattern over our full domain need to be analyzed. Herein we adopt two techniques. Firstly, we compute inter annual relative change percentage for all models and compare with that of the observed. Finally, we conclude our study by computing the mean coefficient of variation over our time scale. These two techniques have been used prior in studying climate change scenarios, for example, Giorgi et al (2004) and Attema and Lenderink (2011).

3.8.1 Relative Change Percentage:

We compute the total annual precipitation for each year from the recorded daily observations and also from each model simulation. Annual relative change is defined as follows:

$$RC = \frac{value\ in\ suceeding\ year - value\ in\ current\ year}{value\ in\ current\ year}\ X\ 100$$

The annual relative change was obtained for each model and historical data set and plotted for comparison:

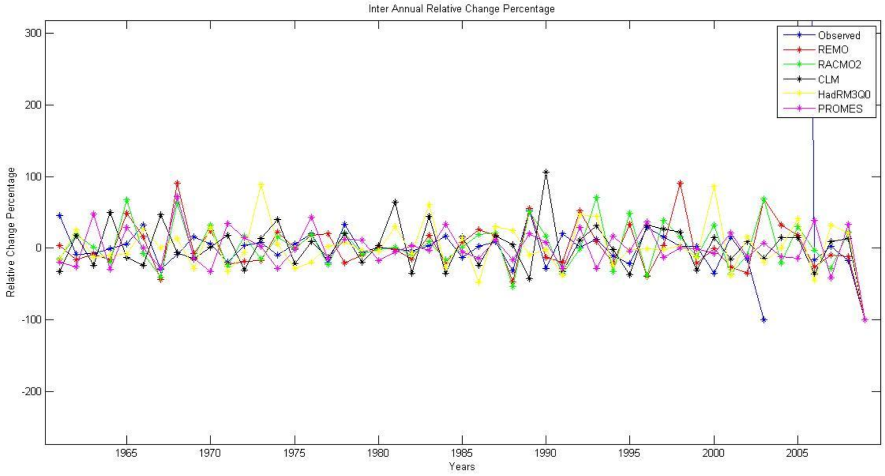


Fig 10: Inter annual relative change percentage

A visual interpretation suggests that the observed patterns of change in annual precipitation are captured well by PROMES and REMO while not so much by other models. The average relative change in precipitation in Goteborg over 1961-2009 has been 1.37%. PROMES and REMO give an average relative change over these years as 2.8% and 1.6% respectively.

3.8.2 Coefficient of Variation

Another method for giving a measure of inter annual variability is Coefficient of Variation index. We compute the coefficient of variation using monthly precipitation data from 1961-2009. The main advantage of using this method is that it removes the dependency of standard deviation on mean precipitation. The basic steps involved in calculating coefficient of variation are:

a) Filtering out effects of trend.
b) Compute the variability in data and its mean.

Coefficient of variation is given by the formula:

$$CV = \frac{\sigma}{\mu}$$

where μ is the mean of monthly rainfall for each year and σ denotes the standard deviation.

Annual coefficient of variation was computed for each model and recorded observations using monthly data for each year and then averaged over years.

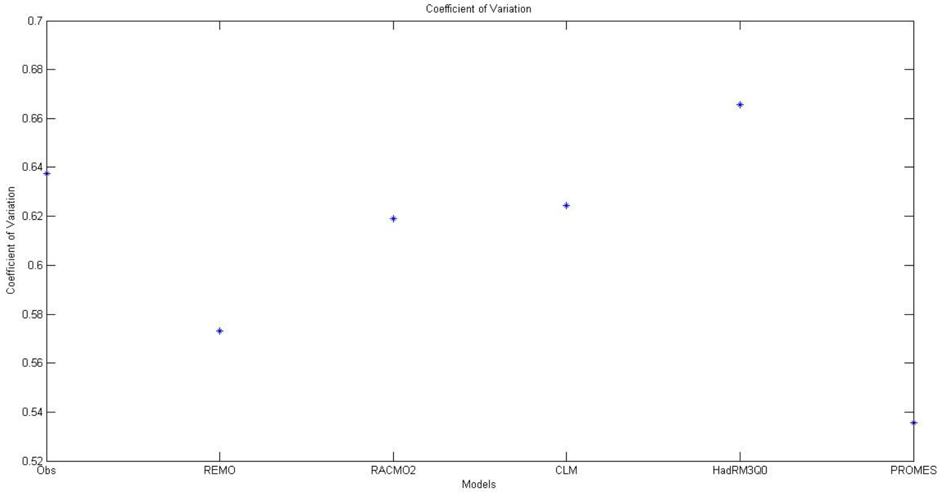


Fig 11: Coefficient of variation for each model

It can be seen from the above plot that CV is less than one for all data sets, implying that there is insignificant inter annual variation in precipitation in Goteborg. It can be seen from the plot above that REMO and PROMES gave slightly smaller index value of coefficient of variation in comparison to the observed. The coefficient of variation for observed data was found to be 0.63 while REMO and PROMES simulations gave the index value as 0.57 and 0.53 respectively. It has been studied before that wet RCMs usually underestimate the coefficient of variation (Giorgi et all, 2004). This could be one of the reasons behind an underestimated value of CV given by PROMES.

Chapter 4 : Summary and Conclusion

The findings of this could have large applicability in various hydrological sectors that specialize in design of storm water drainage, road surfaces, sewers, industries and agricultural practices. From a simple physical viewpoint a warmer atmosphere holds more moisture, enhancing the water cycle in the atmosphere leading to heavier precipitation events. Global warming and increase in green house gas emissions is a proven fact. There are assertions that this will lead us to a rapid and unprecedented rise in global temperature within the next fifty years. Thus, prospects of increase in intensity and frequency of precipitation cannot be neglected.

We are not accustomed to long-term forecasts of anything of such consequence. It is but natural question posed by one and all over the reliability of predictions. Various projects have been launched all across the world - Europe, USA to build future climate change scenarios. The need for such a high magnitude of study has never been felt more. For what is claimed, that if indeed these predictions are true and portray an accurate picture of the future, what choices are we left with and how should we prepare ourselves. Now to the aid of decision makers are the measurements simulated by various climate models. Given the potential implications of regional climate changes towards documenting national policies, and building infrastructure that meets climate related extreme events such as droughts, floods and hurricanes one needs to study how reliable are these projections of future change. One may ask if the uncertainties present in climate models are significantly large to ignore their projections. One needs to understand that neither climatological studies nor present climate models are sufficient to project how climate will change with certainty. And as it is usually said in the field of statistics, “There is no one good model.”

Our study aims at comparing various model projections in order to find that one fairly good model whose projections can be relied upon to estimate a future climate scenario. One method of doing such a study, which has been usually followed by various climatologists, is to take a long time series data of recorded observations of a station and compare it with the predictions. However, one cannot predict climate in the future with certainty mainly because of uncertainties that arise from a) natural variability of climate, b) inability to predict green-house gas and aerosol emissions, c) unpredictable potential factors such as volcanic eruptions or new factors that may stem from human activities and d) our still incomplete understanding of total climate system and inter-relationships between factors. Thus, we may assess the output from various regional climate models as scenarios and not predictions. It is thus not advisable to compare model simulations with the recorded dataset on a time scale. Statistical methods can be applied and results can be compared for drawing meaningful inference. We looked at RCMs driven from two GCMs namely, HadCM3Q0 and ECHAM5r3 and the results have been compiled below:

1. Principle Component Analysis :

   PCA was done on monthly precipitation dataset from 1961 to 2009. Our dataset comprises of historical data of precipitation of Goteborg and model simulations on same time domain. The observations were standardized. All eigenvalues were found to be positive. The first four components together explained 80% variability in data. We chose component 1 and 3 and the coefficients of principle components were plotted. It was observed that RCM-PROMES lied in the same phase as the Observed, whereas, all other models lied in different phase. It can be inferred that PROMES simulations come close to the real data off all other models. This finding is further backed by performing more tests on our dataset.

2. Rainfall Distribution:

   The empirical distribution of data $F(x)$ is defined as proportion of data values less than or equal to x. We overlay the distribution obtained from model simulations over the plot obtained from historical rainfall data distribution for comparison. It was observed that tail end of the observed distribution was captured perfectly by PROMES. There was some difference in the middle – PROMES accounts for more precipitation whereas other models account for less amount of precipitation when compared with the observed rainfall distribution. The same can be seen in that parameters obtained from the fitted distributions. PROMES showcases a small positive bias in mean precipitation while other models have a negative bias in comparison to the observed mean. The variability in data is perfectly captured by PROMES simulations. This result backs the findings of PCA and shows that PROMES is best among all models.

3. Fourier Transform

   We perform Fourier analysis of the dataset to find if there is any periodicity. It is expected that climate data should follow an annual or biennial cycle. The monthly precipitation values over years were Fourier transformed and a biennial periodicity was observed. This frequency was significantly captured only by RCM-PROMES while all other models did not yield a significant peak. This helps us conclude that PROMES simulations do follow the periodicity as observed in historic data which is an important property especially while dealing with climate data. None of the other RCMs simulations showcased a similar periodicity and thence, this finding proves again the inadequacy of other models in comparison to real data.

4. Mann Kendall Test

   Mann Kendall test for trend analysis is a non-parametric test for detecting the presence of any significant trend in data. We use this test mainly because of its two advantages. Firstly, it does not require any assumptions regarding the underlying distribution of the data and secondly, it is not affected by outliers in the data and hence can be easily used for climate studies. It was of initial interest to us if there existed a significant trend in historic data and whether it was adequately captured by any model. However, there was no significant trend in historic data and all models predicted the same. Again, the hypothesis for absence of trend is accepted and it was observed that the p-value, Z-value and the statistic value of the test- S-value of PROMES were closest to the values obtained from the historic data analysis.

5. Averages Using Daily Data

   After having studied periodicity and trend in data, it was of interest whether any model is good enough to capture fine scale fractional changes. To measure this we compute annual maxima, annual two days maximum, annual three days maximum and annual seven days maximum and average it over years. It was observed that mostly all models gave similar averages as the observed. RCM-PROMES gave all set of average values very similar to those of the observational data set. It is pointed out again that PROMES performs best in simulating a real climate scenario in comparison to all other models.

6. Generalized Extreme Value Distribution

   For any model to be redeemed good it is very important that it effectively models extreme events, their intensity and return period. This is of highest regard in climate studies. Extreme events act as catalysts in climate change and need to be studied intricately so that one is aware and prepared for expected extreme occurrences. We fit GEV to annual maxima. It was found that historic data fitted to Gumbel distribution from GEV family of distributions. All models except RACMO2 accepted the Gumbel distribution hypothesis. The observed data, RCM-PROMES and CLM give similar parameter estimates of location, scale and shape parameters of the distribution. Also, they estimate a similar return level of 40-45mm of rainfall at a return period of 10years. This test helps us infer that PROMES models extreme events perfectly in accordance with the observed.

7. Poisson Distribution

   The study of extreme events in statistical theory has been used to demonstrate that the frequency of occurrences of such events is dependent on climate variability. Here we used Poisson distribution to model the number of occurrences of extreme events, an extreme event being defined as the observation exceeding 99th percentile of data i.e. 20 mm of rainfall. Daily data was used and frequency of extreme events was computed for each year. The frequency of extreme events in all models accepted the Poisson distribution hypothesis except for RCM-HadRM3Q0. The parameter estimate, lambda of Poisson distribution for recorded real data was 4.65, i.e. on an average there would be approximately five occurrences when rainfall on any given day exceeds 20mm threshold level. All models except PROMES under-estimated this value and gave a parameter estimate in a range of 2-3. PROMES gave a parameter estimate of 4.24 which lies very close to the observed value.

8. Annual Variability

   Climate variability is an observable characteristic. One year may experience regular and high intensity rainfall while following year may face conditions of drought. Varied rainfall patterns have been seen amounting to different sum total rainfall in each year. Annual variability was measure using two techniques that are often used in climate change studies. Firstly, annual relative change percentage was computed between two successive years for entire time scale considered. A visual interpretation suggested that REMO and PROMES followed the annual variability pattern as observed in the historic data. The average relative change percentage over years was found to be 1.4% for the observed data while REMO and PROMES simulated data predicted the change in precipitation levels as 1.6% and 2.8% respectively. Other models simulated a higher percentage of relative change than the observed. Lastly, another measure for measuring annual variability is the coefficient of variation index. The coefficient of variation for observed data was found to be 0.63 while REMO and PROMES simulations gave the index value as 0.57 and 0.53 respectively. An underestimated value of CV could be due to a wetter climate scenario generated by PROMES. CV value less than 1 for all data sets indicated that there is insignificant inter annual variability over the time range considered. These tests suggest that among all models these two models best explain inter annual variability of precipitation levels over a given time period.

The differences between observed data and control integrations could be due to two factors. The model simulations are from a 25×25 km grid boxes rather than individual stations. Secondly, model data from nine points closest to Goteborg were taken and their values were averaged. The differences are however not as great as expected. Whatever minor discrepancies showcased in our analysis could be attributed to grid box measurements being averaged and compared. In general, one can finally conclude from this study that among all the five regional climate models considered RCM-PROMES simulation statistics are able to best define the periodicity, patterns and occurrence of events as seen in the recorded historic data. The various characteristics of PROMES simulated data match with the observed data and are as listed below:

a) Lies in the same phase as the observed in principle component analysis.
b) Simulates a wetter climate, but in general the rainfall data distribution is similar to the observed. Especially, the tail end of the distribution is in perfect alignment with the observed.
c) Captures the periodicity in data as observed in the historic data.
d) Does not showcase presence of any trend.
e) Gives averages of annual –one day maxima, two day maxima, three day maxima and seven day maxima very similar to the ones obtained from recorded data.
f) Accepts the Gumbel distribution hypothesis for generalized extreme value distribution and predicts a return level of 40mm in 10years.
g) Accepts the Poisson distribution hypothesis as a measure of frequency of extreme events. The parameter value of the distribution is in accordance with the observed. It predicts on an average five occurrences of precipitation more than 20mm on any given day in a year.
h) Follows the annual relative change pattern as observed in historic data. The coefficient of variation index does not deviate much from the one measure from real data.

The aim of this study was to find one model that best replicated the statistics obtained from real-time data. Various methods applied above indicate that RCM-PROMES derived from GCM-HadCM3Q0 and run by institute UCLM is the model whose simulations can be relied upon for making significant assumptions and predictions of climate scenario in coming future.

Reference

About Goteborg-Wikipedia; http://en.wikipedia.org/wiki/Gothenburg

Arnell, N., 1999. The effect of climate change on hydrological regimes in Europe: a continental perspective. Global Environmental Change, Vol. 9, Issue 1, 5-23

Attema, J.,and Lenderink, G., 2011. Mean precipitation changes and coastal effects in the ENSEMBLES regional climate model simulations, KNMI publication

Barron, E., 1995, Climate models: how reliable are their predictions?. Consequences, US global change information office, Vol. 1, Issue 3

BBC News, 2006. Concern over Europe 'snow crisis'-BBC News-2006-12-17

Bengtsson, L., and Milloti, S., 2010. Extreme storms in Malmo, Sweden. Hydrol. Process, Vol. 24, 3462–3475

Busuioc, A., Chen, D., and Hellstrom, C., 2001. Temporal and spatial variability of precipitation in Sweden and its link with the large-scale atmospheric circulation. Tellus, 53A, 348–367

C.C. Wallen, 1958. Variability of annual precipitation in Sweden. Theoretical and Applied Climatology, Vol. 9, Issue 2, 113-123

Climate Data, SMHI; http://www.smhi.se/klimatdata/

Climate Prediction; http://climateprediction.net/content/regional-climate-models

Deque, M., Rowell, D., Luthi, D., Giorgi, F., Christensen, J., Rockel, B., Jacob, D., Kjellstrom, E., M. de Castro, B. van den Hurk, 2007. An inter comparison of regional climate simulations for Europe: assessing uncertainties in model projections. Climate Change, 81:53-70

Dore and Mohammed, 2005. Climate Change and Changes in Global Precipitation Patterns: What Do We Know?. Environment International, Vol. 31, Issue 8, 1167-1181

Frei, C.,Christensen, J., Michel, D., Jacob, D., Jones, R.,and Vidale, P., 2003. Daily precipitation statistics in regional climate models:Evaluation and intercomparison for the European Alps. Journal of Geophysical Research, Vol. 108, Issue D3, 4124

Frei, Scholl, Fukutome, Schmidli, Vidale, 2006. Future change of precipitation extremes in Europe: Intercomparison of scenarios from regional climate models. Journal of Geophysical Research, Vol. 111

Giorgi, F., Bi, X., Pal, J., 2004. Mean, interannual variability and trends in a regional climate change experiment over Europe. II: climate change scenarios (2071–2100). Climate Dynamics, Vol 23, 839–858

Karaca, M., Tayanc, M., and Toros, H., 1995. Effects of urbanization on climate of Istanbul and Ankara. Atmospheric Environment, Vol. 29, Issue 23, 3411-3421

Katz and Brown, 1992. Extreme events in a changing climate: variability is more important than averages. Earth and Environmental Science, Vol 21, 289-302

Linacre, E., 1992. Climate Data and Resources A Reference and Guide. Nature

Map of Goteborg; http://edition.cnn.com/WORLD/9708/25/sweden.bomb/gothenburg.map.jpg

McCarthy, M., Harpham, Goodess, C., Jones, P., 2011. Simulating climate change in UK cities using a regional climate model, HadRM3. International Journal of Climatology

Moussavi Nik, V., 2010. Climate simulation of an attic using future weather sets-Statistical methods for data processing and analysis. Licentiate Thesis, Chalmers University of Technology

New, M., Todd, M., Hulme, M. and Jones, P., 2001. Review: Precipitation measurements and trends in the twentieth century. International Journal of Climatology, Vol. 21, Issue 15, 1889-1992

Olsson, J., 2011. Climate impacts - water in the future, SMHI Research, Hydrology, Accepted. http://www.smhi.se/en/Research/Research-departments/Hydrology/hydrological-impacts-of-climate-change-1.7897

Onoz, B., and Bayazit, M., 2003. The Power of Statistical Tests for Trend Detection. Turkish J. Eng. Env. Sci., 27, 247-251

Rodrı´guez-Fonseca, B., Sa´nchez, E., and Arribas, A., 2005. Winter climate variability changes over Europe and the Mediterranean region under increased greenhouse conditions. Geophysical Research Letters, Vol. 32

Science Daily, December 2002

Smith, T. M.; Yin, X.; Gruber, A., 2006. Variations in annual global precipitation (1979-2004), based on the Global Precipitation Climatology Project 2.5° analysis. Geophysical Research Letters 33: 10.1029/2005GL025393

Stathis, D., and Myronids, D., 2009. Principal Component Analysis of Precipitation in Thessaly Region (Central Greece). Global NEST Journal, Vol.11, 476-476

Sweden facing climate change– threats and opportunities. Final report from the Swedish Commission on Climate and Vulnerability, Swedish Government Official Reports, SOU 2007:60, Stockholm 2007

Torrence, C., and Compo, G., 1998, A Practical Guide to Wavelet Alanysis. Bull Amer. Meteror. Soc., Vol. 79, 61-78

Tsanis, I., Koutroulis, A., Daliakopoulos, I., and Jacob, D., 2011. Severe climate-induced water shortage and extremes in Crete. Climatic Change 106, 667–677

Yiou, Baert, E., Loutre, M., 1996. Spectral analysis of climate data. Surveys in Geophysics, Vol. 17, Issue 6, 619-663

Zhou, L., Dickinson, R., Tian, Y., Fang, J., Li, Q., Kaufmann, R., Tucker, C.,and Myneni, R., 2004. Evidence for a significant urbanization effect on climate in China. Proc Natl Acad Sci, 101(26), 9540-9544