PARAMETER SENSITIVITY AND OPTIMIZATION OF A CATCHMENT-SCALE HYDROLOGIC MODEL ACROSS SWEDEN

ROYA MEIDANI

Master’s thesis
Physical Geography and Quaternary Geology, 60 Credits

NKA 67
2012
Preface

This Master’s thesis is Roya Meidani’s degree project in Physical Geography and Quaternary Geology at the Department of Physical Geography and Quaternary Geology, Stockholm University. The Master’s thesis comprises 60 credits (two terms of full-time studies).

Supervisor has been Steve Lyon at the Department of Physical Geography and Quaternary Geology, Stockholm University. Examiner has been Jerker Jarsjö at the Department of Physical Geography and Quaternary Geology, Stockholm University.

The author is responsible for the contents of this thesis.

Stockholm, 20 December 2012

Lars-Ove Westerberg
Director of studies
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

Abstract

Understanding of the timing and spatial patterns of the fluxes of water is a vital part of any hydrologic-nutrient transport model. To investigate hydrological and nutrient transport interactions, the Baltic Nest Institute (BNI) has implemented the catchment simulation model (CSIM) in the Baltic Sea drainage basin (BSDB). This study focuses on the application of the CSIM model across Sweden as a part of the Baltic Sea drainage basin with specific focus on parameter sensitivity and optimization/calibration. To this end, the spatiotemporal hydrology parameter sensitivity of the CSIM model is explored. In addition, potential improvement over the existing base model parameter calibration is considered using a Genetic Algorithm (GA) optimization method.

The CSIM model parameters show remarkable spatial and temporal (seasonal scale) variation in Sweden. Several regions and watersheds are rather insensitive to the parameters of the CSIM model. Further, the spatial parameter sensitivity emphasizes the relevance of site-specific calibration of the model. The temporal sensitivity analysis clearly shows distinct sensitivity variations in the northern-boreal basins relative to the southern-temperate basins. The annual and seasonal optimization results show that the GA could improve model performance particularly when considering season-specific values for the most sensitive parameters in the CSIM model. Moreover, the optimization outcomes highlighted that the CSIM model, like any hydrological model, faces parameter uncertainties and the concept of equifinality should be considered in the optimization process. This consideration can help to assess the uncertainties that the CSIM model derives from either formulation (structure) or data (calibration).

Key words: CSIM, parameter sensitivity, spatial, temporal, optimization, Genetic algorithm (GA), Baltic Sea drainage basin (BSDB), Sweden, uncertainty, equifinality.
Contents

1. Introduction ......................................................................................................................... 1

2. The CSIM model configuration ......................................................................................... 3
   2.1. CSIM model theory .................................................................................................... 4
   2.2. CSIM model data requirements .................................................................................. 7

3. Background on sensitivity and calibration for GWLF and CSIM ..................................... 8

4. Study Area ......................................................................................................................... 10

5. Materials and Methods ..................................................................................................... 11
   5.1. Sensitivity analysis .................................................................................................... 12
   5.2. Optimization process, Genetic algorithms (GA) ......................................................... 15

6. Results ................................................................................................................................ 19
   6.1. Sensitivity analysis .................................................................................................... 19
      6.1.1 Annual sensitivity analysis .................................................................................... 19
      6.1.2 Seasonal sensitivity analysis ................................................................................. 29
   6.2. The Genetic Algorithm Optimization analysis .......................................................... 34
      6.2.1 Annual GA optimization ....................................................................................... 34
      6.2.2 Seasonal GA optimization .................................................................................... 36
      6.2.3 Model improvement through GA optimization ....................................................... 39

7. Discussion ........................................................................................................................... 42
   7.1. Annual sensitivity analysis ......................................................................................... 42
   7.2. Seasonal sensitivity analysis ...................................................................................... 43
   7.3. Optimization process .................................................................................................. 45

8. Conclusion ............................................................................................................................ 48

Acknowledgement ................................................................................................................ 49

Reference .................................................................................................................................. 50

Appendix A .............................................................................................................................. 54
Abbreviations and Symbols

\begin{align*}
A_t & \quad \text{The antecedent moisture of the previous five days} \\
BB & \quad \text{Bothnia Bay} \\
BNI & \quad \text{Baltic Nest Institute} \\
BP & \quad \text{Baltic Proper} \\
BS & \quad \text{Bothnia Sea} \\
BSDB & \quad \text{Baltic Sea drainage basin} \\
CN_{kt} & \quad \text{Curve number for land-type k} \\
CSIM & \quad \text{Catchment Simulation model} \\
CV_t, \ cv & \quad \text{Cover coefficient for evapotranspiration estimation} \\
DNA & \quad \text{Deoxyribonucleic acid} \\
DP_t & \quad \text{Deep Seepage from the lower saturated zone into the deep saturated zone on day t} \\
DS & \quad \text{Danish sounds} \\
DS_{kt} & \quad \text{Detention parameter for land-type k} \\
E_t & \quad \text{Evapotranspiration on day t} \\
e_t & \quad \text{Saturated water vapor pressure on day t} \\
EU-JRC & \quad \text{European Union- Joint research center} \\
GA & \quad \text{Genetic algorithm} \\
GIS & \quad \text{Geographic information system} \\
Gr_{I_t} & \quad \text{Groundwater discharge from the upper saturated zone into stream on day t} \\
Gr_{II_t} & \quad \text{Groundwater discharge from the lower saturated zone into stream on day t} \\
grcoef & \quad \text{Seepage coefficient to box 2} \\
GW & \quad \text{Ground water} \\
GWLF & \quad \text{Generalized Watershed Loading Function} \\
H_t & \quad \text{Average number of daylight hours in a day for the specific month} \\
IS & \quad \text{Initial snow}
\end{align*}
KT  Kattegat

$M_i$  Modeled values

$M_t$  Snow melt on day t

$N_{bit}$  Number of bits in each variable

$N_{gene}$  Number of genes in chromosomes

$N_{pop}$  Initial population

$N_{var}$  Number of variables

$O_i$  Observed values

$PC_t$  Percolation into the upper saturated zone on day t

$PE_t$  potential evapotranspiration on day t

$Q_t$  Runoff on day t

$R_t$  Rain on day t

$recesscoef2, rec2$  Ground water recession coefficient-Box 2

$recesscoef1, rec1$  Ground water recession coefficient-Box 1

$RMSE$  Root mean square error

$S_{I, satstor1}$  Water in the upper saturated storage zone (Soil water-Box I) on day t

$S_{II, satstor2}$  Water in the lower saturated storage zone (Deep ground water-Box II) on day t

$seepcoef$  Deep seepage coefficient (loss)

$SP_t$  Seepage from the upper saturated zone into the lower saturated zone on day t

$T_t$  Average daily temperature on day t

$temptresh$  Temperature threshold

$tempcoef$  Temperature coefficient

$U_{unsatstor}$  Water in unsaturated storage zone on day t

$X_{rate}$  Selection rate

$%Artificial$  Percentage of artificial land

$%Bare$  Percentage of bare land
<table>
<thead>
<tr>
<th>Percentage Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>%Cultivated</td>
<td>Percentage of cultivated land</td>
</tr>
<tr>
<td>%Decid-For</td>
<td>Percentage of deciduous forest</td>
</tr>
<tr>
<td>%Mixed-For</td>
<td>Percentage of mixed forest</td>
</tr>
<tr>
<td>%Shrub-Herb</td>
<td>Percentage of shrub and herb lands</td>
</tr>
<tr>
<td>%Snow-Ice</td>
<td>Percentage of snow and ice covered lands</td>
</tr>
<tr>
<td>%Water</td>
<td>Percentage of open water coverage in land</td>
</tr>
<tr>
<td>%Wetland</td>
<td>Percentage of wetland</td>
</tr>
</tbody>
</table>
1. Introduction

Eutrophication, a major coastal problem which has been growing for decades in many regions, has attracted the attention of many water-science researchers. There is no doubt that achieving the best possible water resource management in coastal regions requires comprehensive understanding of the water balance within the region’s watershed area. This includes understanding of the timing and spatial patterns of the related fluxes of water such as precipitation, infiltration, soil moisture, discharge, and evapotranspiration along with good knowledge about how these factors interact. Considering the central role of water in nutrient transport, such hydrologic knowledge is vital for investigating the travel pathways of nutrients and to estimate solute loads in and influence on recipient water bodies. Changes in nutrient concentrations and resultant loads could be expected in coastal regions due to many factors such as land use change, soil erosion, population increase, and climate change. To investigate the implications of such changes on coastal regions, watershed models like the generalized watershed loading function (GWLF) create a good opportunity to develop and investigate potential future scenarios highlighting hydrological and nutrient transport interactions. GWLF (Haith et al., 1992) was originally developed for modeling stream-flow, sediment, and nutrient fluxes in the United States as a lumped hydrological model (Mörth et al., 2007). This makes it a good candidate for simulating large-scale systems and potentially suitable for addressing issues associated with excess nutrients and pollution.

Focusing on coastal regions, one can consider the well-known story of Baltic Sea eutrophication. Starting from the 1800s, the Baltic Sea has been changed from an oligotrophic clear-water sea into a eutrophic marine environment. The main reasons of eutrophication in the Baltic Sea are excessive nitrogen and phosphorus loads from land-based sources. Whereas about 75% of the nitrogen load and at least 95% of the phosphorus load are released into the Baltic Sea from rivers or as direct waterborne discharges, just about 25% of the nitrogen load comes as atmospheric deposition (HELCOM, 2006). To help manage these problems, and considering reasonable outcomes of GWLF in large-watershed modeling, the Baltic Nest Institute (BNI) decided to implement the base GWLF model in a manner compatible with conditions for the Baltic Sea drainage basin (BSDB) watersheds. The resulting model framework, called the catchment simulation (CSIM) model, has become the cornerstone of BNI’s whole-basin management approach to the Baltic Sea. Following the EU-JRC data for European watersheds, the modeling framework partitions the Baltic Sea drainage basin into 117 watersheds (Mörth et al., 2009). Up to now, the CSIM model has been applied to evaluate the annual flows and conveyed nutrient loads in these watersheds (Lyon et al., 2012). It is apparent, however, that CSIM (like most of the watershed models working at such large scales) has a number
of parameters which cannot be directly measured or determined at the appropriate spatial scale (i.e., that of a catchment). As such, the accuracy and reliability of the CSIM model (or any lumped catchment-scale hydrologic model) depends on a combination of model structure, input data availability/quality and parameter estimation \((Li \ et \ al., \ 2010)\).

From this perspective, the current study intends to investigate the application of the CSIM model across Sweden as a part of the Baltic Sea drainage basin with specific focus on parameter sensitivity and calibration. This region was selected due to its relatively high data quality and for its long gradient of hydro-climatic conditions spanning the temperature southern to the sub-arctic northern regions. Whereas hydrological modeling is a presentation of some phases of the hydrologic cycle (by a simplified system), physical principles of a given specific hydrologic process should be considered in model formulation. Sensitivity analysis can provide a better understanding of these physical processes and their underlying conceptual representations as they are implemented in a modeling environment. Pointing to the main three phases for modeling a physical system (i.e., formulation, calibration and verification), sensitivity analysis of model parameters can be considered as a vital part of most calibration strategies \((McCuen, \ 1973)\). In addition to parameter sensitivity, the calibration procedure considered is important as it can influence model performance. In recent years, automatic optimization techniques have been widely used in hydrological science for calibration of multivariate models. Genetic algorithms (GAs) as global optimization methods based on the biological evolution, for example, have become increasing popular optimization methods.

Taken together, this study investigates the CSIM model response to spatial and seasonal changes in Sweden (i.e., the spatiotemporal variability) with regards to parameter sensitivity and explores additional optimization/calibration over the existing base model. This follows the initial approach by \(Mörth \ et \ al. \ (2007)\) who established a BSDB-consistent data and modeling framework. To this end, the objective of this study is to determine the most sensitive parameters for discharge estimation within the CSIM model across Sweden. Further, once these parameters are identified, the study seeks to investigate the influence of the model calibration/optimization procedure for the most sensitive parameters within selected watersheds. This procedure inherently addresses the uncertainties in the CSIM model as its optimization is based on the comparison of observed and simulated discharges alone.
2. The CSIM model configuration

According to Mörth et al. (2007), the main goal of the CSIM model as a large-scale model is simulation of nutrient transport to the Baltic Sea. CSIM is a lumped hydrological model based on the GWLF model. The GWLF model was initially developed, tested and described for temperate zone watersheds in North America. GWLF was also originally developed for stream-flow, sediment and nutrient fluxes simulation at the watershed scale. In order to achieve a better description and simulation of hydrology and dissolved nutrient transport in streams across the Baltic Sea drainage basin (and in northern watersheds in particular), GWLFs implementation has been revised to make the CSIM model. The main modification made in creating the CSIM model was the addition of a third conceptual compartment (First ground water box in Figure 1) through which water can transit as it moves from the landscape to the stream. Thus the three conceptual components in CSIM model are as follow (Smedberg et al., 2006):

**Top Soil Water:** Rain and snow water that flushes the topsoil layer with no buffering pool. This water is especially active during spring and contributes instantaneously to the stream. Further, in CSIM snowmelt is simulated using a degree-day approach and the split between runoff and infiltration determined using a curve number approach.

**Soil Water:** Soil water representing near-surface subsurface flows with a shorter residence time than deeper groundwater which drains the soil below the near-surface. From this soil water conceptual pool (shallow ground water), the water is apportioned to either the stream or the deeper groundwater.

**Deep Groundwater:** Deeper groundwater here is responsible for maintaining catchment base flow. It represents water with relatively longest residence time of the three conceptual zones.

By dividing each watershed into a number of land use categories (Deciduous, Coniferous and Mixed forest, Herbaceous, Wetlands, Cultivated areas, Bare lands, Water, Snow and Ice and Artificial areas) and applying precipitation (observed) and evapotranspiration (estimated) to each land cover type, stream-flow is simulated in CSIM (Figure 1) by aggregation of total surface runoff from each land use and the two components of groundwater from the saturated zone (Mörth et al., 2007).
2.1. CSIM model theory

As presented in Figure 1, the upper soil box containing the unsaturated zone is conceptually connected to canopy (land use) and can create surface runoff. The magnitude of the surface runoff from this compartment is computed by the US. Soil Conservation Service’s Curve Number method (Rahm et al., 2005):

\[ Q_{kt} = \frac{(R_t + M_t - 0.2DS_{kt})^2}{R_t + M_t + 0.8DS_{kt}} \]  
(1)

Where \( R_t \) and \( M_t \) are rainfall and snowmelt respectively per day [cm]. \( DS_{kt} \) is called detention parameter [cm] which determines the fraction of precipitation and snowmelt that is routed through the surface runoff (Rahm et al., 2005):

\[ DS_{kt} = \frac{2540}{CN_{kt}} - 25.4 \]  
(2)

Where \( CN_{kt} \), the curve number, is determined as a function of antecedent soil moisture of the previous five days. As such, the curve number can range from dry to wet values based on antecedent
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

conditions. The base curve number ($CN_{2k}$) depends on land use, hydrologic condition and hydrologic soil group and values typically vary between 30 and 95 (Rahm et al., 2005; Haith et al., 1992). Curve number values are based on plot-scale studies of rainfall-runoff relations carried out over much of North America. From this base curve number, dry ($CN_{1k}$) and wet ($CN_{3k}$) relevant curve number values can be determined using the following (Hawkins, 1978):

\[
CN_{1k} = \frac{CN_{2k}}{2.34 - 0.01334CN_{2k}} \quad (3)
\]

And,

\[
CN_{3k} = \frac{CN_{2k}}{0.4036 - 0.0059CN_{2k}} \quad (4)
\]

The antecedent moisture of the previous five days is calculated by aggregation of rainfall ($R_n$) and snow melt ($M_n$):

\[
A_t = \sum_{n=t-5}^{t-1} (R_n + M_n) \quad (5)
\]

In CSIM, the curve numbers are selected as a function of $A_t$ and the relevant growing/dormant condition of the current month. For instance, for $A_t < 1.6$ cm in dormant seasons and $A_t < 3.6$ cm in wet seasons the CSIM model chooses the dry condition and consequently $CN_{1k}$, for $A_t > 2.8$ cm in dormant seasons and $A_t > 5.3$ cm in growing seasons $CN_{3k}$ is used similar to the snowmelt conditions. For $A_t$ values between defined constraints $CN_{2k}$ are used (Rahm et al., 2005; Haith et al., 1992).

Water that does not go directly to runoff is allowed to infiltrate and evapotranspire from an unsaturated storage. As mentioned by Haan (1972) the moisture holding capacity of this storage has two main sections. The first is a volume readily available for evapotranspiration, $M_1$, with the maximum capacity assumed to be $U^*$; and the second is a volume that is less readily available for evapotranspiration, $M_2$. Moreover, after precipitation is divided into infiltration and surface runoff, $M_1$ stores all of the infiltrated water until its entire capacity is filled. Any additional infiltrated water, thus, goes directly to $M_2$. In the condition that both of storages are full, the entire precipitation produces runoff. Evapotranspiration is equal to potential evapotranspiration while readily available water is accessible. Over time, water stored in $M_1$ and $M_2$ is also allowed to percolate into lower saturated storages.

Considering Haan’s (1972) soil moisture concept and using daily mass balance which occurs in the unsaturated, two shallow saturated and deep saturated zones the three governing equations in the CSIM model could be defined as:
And,

\[ U_{t+1} = U_t + R_t + M_t - Q_t - E_t - PC_t \quad (6) \]

And,

\[ S_{II}t+1 = S_{II}t + PC_t - Gr_{II}t - SP_t \quad (7) \]

And,

\[ S_{II}t+1 = S_{II}t + SP_t - Gr_{II}t - DP_t \quad (8) \]

With regards to Figure 1, the parameters in these three equations could be defined as following:

- \( U_t \): Water in unsaturated storage zone on day \( t \) in cm,
- \( R_t \): Rain in cm on day \( t \),
- \( M_t \): Snow melt in cm on day \( t \),
- \( Q_t \): Runoff in cm on day \( t \),
- \( E_t \): Evapotranspiration in cm on day \( t \) (Eq.10 &11),
- \( PC_t \): Percolation in cm into the upper saturated zone on day \( t \),
- \( S_{II}t \): Water in the upper saturated storage zone (Soil water-Box I) on day \( t \) in cm,
- \( S_{II}t \): Water in the lower saturated storage zone (Deep ground water-Box II) on day \( t \) in cm,
- \( Gr_{II}t \): Groundwater discharge from the upper saturated zone into stream on day \( t \) in cm (Eq.12),
- \( Gr_{II}t \): Groundwater discharge from the lower saturated zone into stream on day \( t \) in cm (Eq.13),
- \( SP_t \): Seepage from the upper saturated zone into the lower saturated zone in cm on day \( t \) (Eq.14)
and
- \( DP_t \): Deep Seepage from the lower saturated zone into the deep saturated zone in cm on day \( t \) (Eq.15).

Indeed, some of these parameters like evapotranspiration and percolation need special conditions to occur and have various physical constraints. By definition, percolation occurs when the maximum capacity of unsaturated soil is exceeded. Additionally, evapotranspiration cannot exceed the available moisture in the unsaturated zone (Haith et al., 1992):

\[ E_t = \text{Min}(CV_t \cdot PE_t, U_t + R_t + M_t - Q_t) \quad (9) \]

Where \( CV_t \) is cover coefficient and \( PE_t \) is daily potential evapotranspiration in cm by Hamon (1961) method (Haith et al., 1992):

\[ PE_t = \frac{0.021H_t^2e_t}{T_t + 273} \quad (10) \]
Where $H_t$ refers to the average number of daylight hours in a day for the specific month, $T_t$ is the average daily temperature on day $t$ and $e_t$ is the saturated water vapor pressure in millibar which can be approximated for all the $T_t > 0$ by Bosen (1960) formula (Haith et al., 1992):

$$e_t = 33.8639 [(0.007387 T_t + 0.8072)^8 - 0.000019 (1.87 T_t + 48) + 0.001316]$$ (11)

With regard to Haan (1972), ground water discharges ($Gr_I$ and $Gr_{II}$, respectively) to the stream channel, seepage and deep seepage from each boxes of shallow saturated zone to the deeper saturated zones are modeled by a simple linear reservoir:

$$Gr_I_t = recesscoef_1 * S_I_t$$ (12)

And,

$$Gr_{II}_t = recesscoef_2 * S_{II}_t$$ (13)

And,

$$SP_t = grcoef * S_I_t$$ (14)

And,

$$DP_t = seepcoef * S_{II}_t$$ (15)

Where recesscoef1 and recesscoef2 are groundwater recession coefficients in the upper and lower boxes of shallow saturated zone, respectively. These values can be estimated from the hydrograph separation procedure (Chow, 1964) but can also be treated as purely calibrated parameters. Further, grcoef and seepcoef are the constant rates of seepage and deep seepage from the two groundwater boxes (Figure 1). It is worth noting that the seepage from the lower saturated zone into the deep aquifer has no influence on the stream-flow discharge and therefore on nutrient loads (Haith et al., 1992). As such, these four parameters (recesscoef1, grcoef, seepcoef and recesscoef2) control the timing of water movement in CSIM through the subsurface domain.

2.2. CSIM model data requirements

Daily temperature and precipitation are the main drivers of the CSIM model. For the Baltic Sea drainage basin within the Baltic NEST framework, these are extracted from 700 to 800 synoptic stations cover the whole Baltic drainage basin. These data are available in an interpolated resolution of 1x1 degree from 1979 to present. Beside the meteorological database, each watershed includes land use, soil type, erosion and sediment, nutrient concentration in runoff and hydrology databases. These data together were used to calibrate the currently available version of the CSIM model (e.g., Mörth et al., 2007). This calibrated model forms the basis for this current study. The hydrology
compartment of CSIM which absorbs the main focus of this study consists of different parameters such as recession coefficients, ground water coefficient, seepage coefficient, temperature threshold and coefficient, initial unsaturated and saturated storages, initial snow and cover coefficient for estimating the evapotranspiration (Rahm et al., 2005).

3. Background on sensitivity and calibration for GWLF and CSIM

Some previous sensitivity analysis and calibration approaches have been carried out for the GWLF model. These include the sensitivity analysis and comparison performed by Staley et al. (2006) looking at modeling channel erosion at the watershed scale. Li et al. (2010) considered GWLF watershed model calibration using multi-objective optimization and multi-site averaging that looks specifically into three groups of parameters in GWLF according to the Lee et al. (2000) approach. This classification was performed due to influences of the parameters on the model response: group one which refers to essential parameters, group two with parameters that can improve the model performance and group three which contains parameters that can be set to default values. As such, they calibrated only the hydrology and nutrient parameters in groups one and two by excluding the curve number (CN) and soil erosion parameters.

In the mentioned approach by Lee et al. (2000), they evaluated the predictive ability of GWLF with regards to water flux, nitrogen and phosphorus export in the Choptank basin (within two gauged areas). Accordingly, as stated earlier, the model parameters were categorized in three groups with regards to their effect on the model response. Group one or “essential parameters” should be locally adjusted to achieve the better model performance. The curve number (CN) which is related to land use and soil drainage, sediment delivery ratio, the groundwater recession coefficient, the seepage coefficient, unsaturated available water capacity (UAWC), dissolve and particular nutrient concentration in each land use, point sources and population are the parameters in this group. The second group indicated a class of parameter which can improve the model performance. These were called “useful” parameters and included the crop coefficient for evapotranspiration. The third group or “nonessential” parameters could be set to default values due to their influence on just a few months of the model output. The influence of this group could be avoided by running the model for a year prior to period of interest with regard to recommendation by Haith et al. (1992). These nonessential parameters included unsaturated storage, saturated storage, initial snow, and 5-day antecedent precipitation. Lee et al. (2000) concluded that from all of the hydrological parameters in GWLF, the ground water recession coefficient is the most important. Further, they emphasized that the model parameters are very site dependant and for the most accurate application, they should be calibrated and verified locally.
Considering the CSIM model at the scale of the Baltic Sea drainage basin, Mört et al. (2007) presented some base application of the CSIM model to describe the annual stream-flow and total N loads to the Baltic Sea. This application performed fairly well on an annual basis with stream-flow, for example, only overestimated with 4% across the entire Baltic Sea drainage basin by the model. In addition they presented that the inter-annual correlation coefficients of individual watersheds show temporal discrepancies in some cases. Meanwhile on average, the inter-annual correlation coefficient per watershed was 0.69 for stream-flow where this value for the major watersheds, ranged between 0.13 and 0.75.

However, there has been no formal sensitivity analysis or detailed/guided parameter calibration performed for the CSIM model in its application to the Baltic Sea drainage basin. Further, a detailed investigation into the spatiotemporal performance at a sub-annual scale for this model is lacking. This forms the main focus of this current study.
4. Study Area

The Baltic Sea drainage basin covers an area about 2,000,000 km$^2$. The northern region, with boreal forests and wetlands, drains into the Gulf of Bothnia and the southern region which is dominated by cultivated areas and deciduous forests drains into the rest of the Baltic Sea (Figure 2).

Following the EU-Joint Research Center the entire Baltic Sea drainage basin can be divided to 117 watersheds (Mörth et al., 2009) with both direct drainage and diffuse drainage into the Baltic Sea. The Baltic Sea itself is characterized as seven main basin districts (Bothnia Bay, Bothnia Sea, Baltic Proper, Gulf of Riga, Gulf of Finland, Kattegat and Danish Sounds). These basins in the Baltic Sea can be used to group the various drainage areas across the entire Baltic Sea drainage basin.

This study focuses on Sweden as a collection of the western Baltic Sea drainage watersheds, which extends 1600 km from South to North (55° to 69°) and 500 km from West to East (Figure 2). This area has been divided traditionally into 3 regions: Norrland in the north with vast mountains and forests, Svealand in center with lowlands in the east and highlands in the west, and Götaland in the south. The highest peaks of Sweden are located in Norrland and are north of the Arctic Circle at about 2089 meters and 2011 meters above sea level, respectively, in the Scandinavian Mountains chain containing numerous glaciers. Most of the surface level of eastern Svealand and...
northern Götaland is below the sea level due to a dominance of lakes, fragmented bedrocks, fertile clayey plains and sandy ridges by glaciers. Table 1 shows the land class areas of the each considered Baltic Sea main basin districts in Sweden.

Table 1. Land class areas (ha) of each Baltic Sea main basin districts in Sweden: Bothnia Bay (BB), Baltic Proper (BP), Bothnia Sea (BS), Danish Sounds (DS), Kattegat (KT), Shapefile source Mörth et al. (2007) (Baltic NEST institute).

<table>
<thead>
<tr>
<th>Basin</th>
<th>%Bare</th>
<th>%Artificial</th>
<th>%Shrub-</th>
<th>%Cultivated</th>
<th>%Mixed-</th>
<th>%Conif-</th>
<th>%Decid-</th>
<th>%Wetland</th>
<th>%Snow</th>
<th>%Water</th>
<th>Basin area (ha)</th>
<th>Contains N watersheds</th>
</tr>
</thead>
<tbody>
<tr>
<td>KT</td>
<td>0.16</td>
<td>2.04</td>
<td>6.48</td>
<td>15.25</td>
<td>3.85</td>
<td>52.62</td>
<td>3.01</td>
<td>2.69</td>
<td>0</td>
<td>13.90</td>
<td>7345444</td>
<td>9 (6)</td>
</tr>
<tr>
<td>DS</td>
<td>0</td>
<td>9.92</td>
<td>0.73</td>
<td>76.96</td>
<td>0.09</td>
<td>5.81</td>
<td>5.11</td>
<td>0.08</td>
<td>0</td>
<td>1.30</td>
<td>262271</td>
<td>1 (0)</td>
</tr>
<tr>
<td>BP</td>
<td>0.29</td>
<td>3.15</td>
<td>4.27</td>
<td>22.76</td>
<td>3.47</td>
<td>53.01</td>
<td>2.83</td>
<td>0.84</td>
<td>0</td>
<td>9.38</td>
<td>8379146</td>
<td>12 (9)</td>
</tr>
<tr>
<td>BS</td>
<td>1.14</td>
<td>0.58</td>
<td>22.88</td>
<td>2.83</td>
<td>4.39</td>
<td>49.82</td>
<td>4.22</td>
<td>7.13</td>
<td>0</td>
<td>7.01</td>
<td>17994816</td>
<td>13 (12)</td>
</tr>
<tr>
<td>BB</td>
<td>6.82</td>
<td>0.38</td>
<td>23.91</td>
<td>1.07</td>
<td>5.93</td>
<td>36.74</td>
<td>6.62</td>
<td>11.72</td>
<td>0.31</td>
<td>6.51</td>
<td>13105114</td>
<td>10 (9)</td>
</tr>
<tr>
<td>Sum (ha)</td>
<td>1135515</td>
<td>595046</td>
<td>8084868</td>
<td>3877679</td>
<td>2140832</td>
<td>22101373</td>
<td>2097785</td>
<td>3086764</td>
<td>41374</td>
<td>3925555</td>
<td>47086791</td>
<td>45 (36)</td>
</tr>
</tbody>
</table>

*The number in the parentheses shows the number of gauged watersheds and the other out of that shows the total number of watersheds (gauged and un-gauged(coastal)).

Following the Köppen climatic classification (Köppen-Geiger, 2011) the north and major part of central Sweden are located in Continental Subarctic Climate; while the southern part is classified in Humid Continental zone. Accordingly, it is not surprising that the growing season varies from 240 days in south to 120 days in north and the cultivated area is less than one-tenth of Sweden’s area with most arable lands in the south. In Sweden, the temperature changes from -40°C in winter, in the northern area, to 17°C in summer in the southern part. Annual precipitation is about 600 mm which falls throughout the year but mostly in late summer and autumn (Encyclopædia, 2012). Indeed, each of these “Characteristics” according to their significance has an influence on the watersheds hydrological and hydro-chemical behavior.

5. Materials and Methods

The CSIM model has been written in Visual Basic 6.0 with an Access database for data storing and retrieving. Its framework handles input data and results, besides keeping track of the history of simulation and files. This current study adopts the base model calibrated by Mörth et al. (2007) as a starting point of the analysis. The CSIM model initial calibration was performed using the SDK solver module (www.solver.com). The solver module used long term monthly data of 1995 to 2000 to minimize the RMSE for finding the best fit between modeled and observed values. The climatic and discharge model inputs in this study are daily data from January of 1990 to the end of December 2000 for the sensitivity analysis and from January 1995 to end of December 2000 for the Genetic
Algorithm optimization. These data were used in the CSIM model parameter sensitivity analysis for 36 of the total 45 Swedish watersheds draining into the Baltic Sea (eliminating the 9 diffuse coastal watersheds). For a better clarification of the region, the watersheds were divided to the “main” Baltic Sea drainage basins which are Bothnia Bay (BB), Baltic Proper (BP), Bothnia Sea (BS), and Kattegat (KT) (Figure 2).

For the CSIM parameter sensitivity analysis (Section 5.1) and further the model optimization (Section 5.2) carried out in this current study, the CSIM model code was manipulated using a systematic changing of the hydrological model parameters. The following sections provide methodological theory on the approaches considered and descriptions of how they were implemented in this study. Results were characterized and compared using a combination of MATLAB (2011a) and ARCGIS.10.

5.1. Sensitivity analysis

Enhancing model calibration efficiency implies concentrating effort on the most sensitive parameters of the model (Beven, 2008). By definition, sensitivity is the rate of change in one parameter with regard to change in another factor. Considering the Taylor series expansion, the general definition of sensitivity can be presented mathematically as (McCuen, 1973):

\[ F_0 = x(F_1, F_2, ..., F_n) \] (16)

Changes in factor \( F_i \) would affect factor \( F_0 \) as following form:

\[ x(F_i + \Delta F_i, F_j | j \neq i) = F_0 + \frac{\partial F_0}{\partial F_i} \Delta F_i + \frac{1}{2!} \frac{\partial^2 F_0}{\partial F_i^2} + \cdots \] (17)

If the nonlinear terms were considered negligible with respect to the linear term, Eq.17 reduces to:

\[ x(F_i + \Delta F_i, F_j | j \neq i) = F_0 + \frac{\partial F_0}{\partial F_i} \Delta F_i \] (18)

And further to:

\[ \Delta F_0 = x(F_i + \Delta F_i, F_j | j \neq i) - F_0 = \frac{\partial F_0}{\partial F_i} \Delta F_i \] (19)

Consequently, the general definition of sensitivity is:

\[ S = \frac{\partial F_0}{\partial F_i} \] (20)
Furthermore, model parameter sensitivity refers to the relationship between changes in input parameters and resulting differences in model response. Taken together, one method to conduct model parameter sensitivity analysis is to take the local gradient of the response surface defined by some performance measure along the parameter axis of the parameter being considered. In essence, this makes the normalized sensitivity index of the following form (Beven, 2008):

$$S_i = \frac{dZ/dx_i}{x_i}$$  \hspace{1cm} (21)

Where $S_i$ is the sensitivity index of the parameter $i$ with value $x_i$ and $Z$ is performance measure. In this study, root mean square error (RMSE) was selected for the performance measure of the CSIM model at the considered point in the parameter domain (see following section).

To assess the sensitivity index in this study, the percentage of change in an individual model parameter was compared to the root mean square error (RMSE) (Eq.22) between modeled and observed values (corresponding to changed parameter) and the sensitivity index was calculated as the slope of this comparison. The RMSE, which has been used widely in hydrology to measure the goodness of fit between modeled and observed values, is defined as (Hall, 2001):

$$RMSE = \sqrt{\frac{1}{n} \sum^n_i (O_i - M_i)^2}$$ \hspace{1cm} (22)

Where $O_i$ and $M_i$ are the observed and modeled values (stream-flow in this study), respectively, and $n$ is the number of observations. It is clear that the lowest RMSE presents the better fit between simulated and observed values and a zero value indicates a perfect fit. The RMSE was calculated to evaluate the sensitivity of the model parameters (see previous section).

With regards to the Lee et al. (2000) approach and the CSIM model configuration, the hydrological parameters of the CSIM model were all presented in Table 2.
<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Unit</th>
<th>Variable</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Curve number</td>
<td>For computing the surface runoff</td>
<td>-</td>
<td>land use</td>
<td></td>
</tr>
<tr>
<td>CV</td>
<td>Cover coefficient for evapotranspiration estimation</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Base-flow recession 1</td>
<td>GW recession coefficient-Box 1</td>
<td>day^-1</td>
<td>recesscoef1</td>
<td>**</td>
</tr>
<tr>
<td>Base-flow recession 2</td>
<td>GW recession coefficient-Box 2</td>
<td>day^-1</td>
<td>recesscoef2</td>
<td>*</td>
</tr>
<tr>
<td>Groundwater coefficient</td>
<td>Seepage coefficient to box 2</td>
<td>day^-1</td>
<td>grcoef</td>
<td>**</td>
</tr>
<tr>
<td>Seepage coefficient</td>
<td>Deep seepage coefficient (loss)</td>
<td>day^-1</td>
<td>seepcoef</td>
<td></td>
</tr>
<tr>
<td>Temperature coefficient</td>
<td>For Snow melt calculation (degree-day method)</td>
<td>-</td>
<td>tempcoef</td>
<td>*</td>
</tr>
<tr>
<td>Temperature threshold</td>
<td>For Snow melt calculation (degree-day method)</td>
<td>°C</td>
<td>temptresh</td>
<td>*</td>
</tr>
<tr>
<td>U</td>
<td>unsaturated storage</td>
<td>Cm</td>
<td>unsatstor</td>
<td>***</td>
</tr>
<tr>
<td>S_I</td>
<td>saturated storage1</td>
<td>Cm</td>
<td>satstor1</td>
<td>***</td>
</tr>
<tr>
<td>S_II</td>
<td>saturated storage2</td>
<td>Cm</td>
<td>satstor2</td>
<td>*</td>
</tr>
<tr>
<td>IS</td>
<td>Initial snow</td>
<td>Cm</td>
<td>snow</td>
<td>***</td>
</tr>
</tbody>
</table>

* In result of differences in number of groundwater boxes and snow-melt process, it was not mentioned by Haith et al. (1992).

** In result of differences in number of groundwater boxes, it does not have the same name as Haith et al. (1992) paper.

*** With regards to Haith et al. (1992), this parameter only affects first few months.

Considering the water residence time as an important issue in Hydrological-Nutrient modeling, this study specifically looks into the two groundwater compartments of the CSIM model (saturated zone). Therefore, the parameters which influence the over-land and unsaturated zones were excluded from this approach (e.g. curve number, temperature coefficient, cover coefficient and etc.), although, their influence on the hydrological process of the saturated zone could not be neglected. The italic-gray cells in Table 2 represent the parameters which are considered in the sensitivity analysis.

To find the most sensitive parameters, the parameters were individually changed from their initial calibrated values from -100% to +100% of that value by increments of 10% for each of the 36 Swedish watersheds. By considering 36 watersheds, 20 steps for 10 % increments of changing and 6 hydrologic parameters, the CSIM model was run 4320 times by adding a loop script to the model code. Further, the model results for each of these adjusted parameter values were compared to the annual/seasonal observed values by calculating the RMSE. The annual sensitivity graphs (e.g., plots of performance against parameter adjustment) were presented in two viewpoints. First, the sensitivity graphs were grouped to compare the parameter spatial sensitivity of the main basin districts draining across Sweden. Second, for each main basin district, the sensitivities of the four most sensitive parameters were plotted for relative ranking.

Considering Eq.21, the parameter sensitivity index, $S_i$, was calculated as the slope of the sensitivity graph over each 20% domain of the parameter change. Indeed, for finding the most sensitive parameters or evaluating the spatial sensitivity of one parameter, the magnitude of the sensitivity...
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

index is more important than its direction. To this end, the absolute value of the RMSE slope (regardless of its sign) was considered for rating the sensitivity index. Thus, the derivatives of the annual-RMSE curves (by increments of 20%) for all of the 36 Swedish watersheds were calculated. Further, absolute of the calculated values (of all watersheds) were mapped to show the spatial sensitivity of the considered parameters along Sweden. Again, these absolute values have been used to make an average over 36 Swedish watersheds and consequently to rank the average annual-sensitivity of the 4 parameters.

As the other objective of this study is to investigate temporal sensitivity of the CSIM model parameters, the model results (due to several runs and parameters changing) were regulated seasonally over the study period (1990-2000). Seasons in this study were divided to Winter, Spring, Summer and Autumn; of which Winter consists of December, January and February, Spring starts from March and ends in May, Summer represents Jun, July and August values and finally Autumn comprises September, October and November. Further, the seasonal sensitivity was evaluated by seasonal RMSE calculation for the modeled and observed discharges. The sensitivity graphs of each parameter were mapped seasonally for comparing the behavior of the main basin districts.

5.2. Optimization process, Genetic algorithms (GA)

Simply, optimization refers to the processes which make something better or search for the “Best solution” by adjusting the inputs. The term “best” implies that the problem faces more than one solution with different values for acceptance. The definition of “best” is tightly connected to the considered problem, its solution method, tolerances of convergence and of course, the person who formulates the problem. Some problems may have exact answers (e.g., roots) and some may have various minimum and maximum solutions which known as extrema (e.g., optimal points). The Genetic Algorithm, by John Holland (1975), as a global optimization method presents the process in nature which is remarkably successful in optimizing of natural phenomena (Haupt and Haupt, 2004).

Some advantages of Genetic Algorithms are:

- There is no need to use derivatives.
- It can deal with the large number of variables.
- Variables with extremely complex cost surfaces can be optimized by jumping out of local minimums.
- Instead of one single solution, it provides a list of optimums.

Accordingly, in the current study the genetic algorithm was used for the annual and seasonal model optimization. As such, a MATLAB version of GA which is a modified code by Haupt and Haupt (2004)
was revised to be compatible with our problem in hand (the CSIM model) and the cost function (RMSE). The four hydrological parameters which have been evaluated in the sensitivity analysis were calibrated and optimized with the GA. The RMSEs between the observed and the modeled discharges, as the cost function, were minimized in the optimization process for four selected watersheds: Mörrumsån (93), Dalälven (25), Ångermanälven (33) and Torne älv (10) (see Appendix A, Figure A1 for their locations). The 6 years data, 1995 to 2000, were used for the CSIM model in the GA optimization process. This period has been considered for making the comparison possible between the first calibrated values which have been done by Mörh et.al (2007) and the GA’s optimized values. For avoiding the influence of “nonessential” parameters (See section 3 for the definition), the CSIM model was run a year prior to 1995 with regard to recommendation by Haith et al. (1992) as a spin up period. As such, the CSIM model was run from 1994 to 2000 while the optimization process just considered the 1995 to 2000 values.

Like any optimization techniques, the GA starts with defining the optimization variables, the cost function and the cost; it ends, likewise, by testing for convergence. However, the difference lies with the methods between these points. Considering a gene as the basic unit of heredity, an organism’s genes are carried on one of pair of chromosomes in the form of DNA (deoxyribonucleic acid). DNA in the form of double helix carries the genetic code of the organisms. The GA considers optimization input variables as chromosomes and by using a set of chromosomes generates an output for the cost function. First, it starts by defining chromosomes (an array of variable values) which are considered to be optimized. By having N variables for the optimization process which is defined by $P_1, P_2...$ and $P_{N_{\text{var}}}$ then the chromosome is:

$$chromosome = [P_1, P_2, P_3, ..., P_{N_{\text{var}}}] \quad (23)$$

In this study, number of variables is 4 and the variables are defined by the sensitivity analysis. It is clear that each chromosome offers a cost by evaluating the cost function ($f$), here RMSE:

$$cost = f(chromosome) = f(P_1, P_2, P_3, ..., P_{N_{\text{var}}}) \quad (24)$$

It should be noted here that the unconstraint optimization puts no limit in variable domain, thus variables could take any values. In contrast, constrained optimization limits the variable domain by incorporating a penalty function to the cost function. This prevents finding unreasonable chromosomes during the optimization process (Karamouz and Kerachian, 2003). In this approach the upper limit of parameters was set to twice as their initial calibrated value by Mörh et al. (2007) and the lower limit was set to 0. It is worth noting that there is no penalty function and no pre-defined range for the considered parameters in this study.
Genetic Algorithm works with binary encoding, 0 and 1; it converts the variable values to binaries and further because the cost functions requires continuous variables starts decoding. In encoding process number of bits in each variable is considered as number of genes in chromosomes, an example of an encoded chromosome with 10 bits and \( N_{\text{var}} \) is:

\[
\text{chromosome} = [1111001001001100111111 \ldots 00001011011]
\]

In this case each gene can experience \( 2^{10} \) possible values for one variable (Haupt and Haupt, 2004). In this study number of bits was set to 8 for the input variables, while in one special case this was changed to 12 for two parameters to consider a lower weight assignment to these least sensitive parameters.

These briefly presented principles of GA can be better appreciated in Figure 3 which shows the Binary Genetic Algorithm process.

![Figure 3. A binary Genetic Algorithm flow chart (Haupt and Haupt, 2004)](image-url)
As presented, the GA starts with generating an initial population. This population is generated as an \( N_{\text{pop}} \times (N_{\text{bits}}=N_{\text{gene}} \times N_{\text{var}}) \) matrix which filled by randomly produced zeros and ones. In this study the initial population is set to 32.

After calculating of the cost for \( N_{\text{pop}} \), the costs and correspondent chromosomes are ranked due to the value of their cost (in this case lower to higher). Further, only the best survive to continue and generate new offspring, while the rest are deleted. The selection rate is the fraction of population that is selected for the next step of mating; \( X_{\text{rate}} \) was set to 50% in this study with regards to Haupt and Haupt’s (2004) recommendation, except in one special case for the Mörrumsån watershed. After selection of good chromosomes, two chromosomes should be selected from the mating pool to produce two new offspring. The mating continues until new offspring replace the deleted chromosomes. Figure 4 shows a most common mating form of two parents that produce two offspring. For further information about parent selection and mating see Practical Genetic Algorithm by Haupt and Haupt (2004).

![Figure 4. Two parents are mating to generate two offspring (Haupt and Haupt, 2004)](image)

Alternating a certain percentage of the bits, from 0 to 1 or vice versa, in the list of chromosomes is mutation. Mutation is an initiative way in GA for exploring the cost surface for preventing to be trapped by popular solution convergence. The first chromosome (due to elitism) and the last iteration are free from mutation. In the current approach, the initial and final mutation rates were set to 0.05 and 0.005 respectively (except in a special case in Mörrumsån watershed). After the mutation process, the cost of mutated chromosomes and offspring are calculated and the iteration process continues until the chromosomes and associated costs convergence. The number of generation (iterations) in this study varied from 20 to 40 for the annual model optimization and 30 to 40 for the seasonal model optimization.
6. Results

6.1. Sensitivity analysis

6.1.1 Annual sensitivity analysis

The CSIM model was run for the period January 1990 to December 2000 with sensitivity measured by changing the hydrologic parameters and comparing modeled and observed runoff differences. This was done by changing of the hydrologic parameters from -100% to 100% of their initial calibrated values in every model run. The study focuses on the 4 most sensitive parameters control the movement of water through the subsurface since, by comparison, the other parameters showed no signs of being sensitive and were primarily associated with the initialization period.

The 36 considered Swedish watersheds were categorized to the “main” Baltic Sea drainage basins (Figure 2). This was done because presenting the model sensitivity results for a large number of watersheds can make interpreting of the spatial correlation difficult. Following this, Figure 5 (a-d) shows the sensitivity curves of the four most sensitive parameters. These were recesscoef1, recesscoef2, grcoef and seepcoef for the main basin districts along Sweden. In these curves, the curve steepness indicates general parameter sensitivity.
Figure 5. Sensitivity curves for a) recesscoef1, b) recesscoef2, c) grcoef and d) seepcoef for the considered main basin districts BB, BS, BP and KT in Sweden, RMSE was calculated for modeled and observed runoffs over 1990-2000.

As shown in Figure 5, the RMSE, in most of the cases, tends to increase as the parameter value changes from its initially calibrated value (i.e., moving left or right from 0 along the X axis). This could help to find the effective range of the parameters in the calibration process. At a glance, by looking to the slope of the RMSE curves it is apparent that the parameters seepcoef and recesscoef2 tend to be less sensitive to variations in parameter values as their curves remains mainly steady even in large changes. The recesscoef2 only shows a steep curve when at the extreme case of -100% changes which force the parameter back to a 0 value. The RMSE curves of the parameter recesscoef1 obtain a steep-parabolic shape which shows a rapid variation due to parameter changing. The parameter grcoef seems to have gentler parabolic shape compared to recesscoef1 curve while it shows more rapid variation to change of parameter compared to the curves of parameter recesscoef2 and seepcoef. For instance, Figure 5a shows roughly 16 mm change in RMSE by decreasing the recesscoef1 from its initial value to 0 (decreasing 100%) in the KT basin. In BP, however, over the same range this amounts in a RMSE change is only 6 mm. Figure 5d shows increasing the seepcoef to twice as its initial calibrated value in KT basin changes the RMSE less than 0.1 mm (from 9.4 to 9.5) and also less than 1 mm for decreasing the parameter to 0. In addition, this reveals that BP has the lowest sensitivity compared to the other main basin districts, as its sensitivity
curves have gradual behavior even with large variation in the considered parameters. The KT and BB basins behavior could be considered as having higher sensitivity in Figure 5a compared to the BB basin. However, these basins’ sensitivity is hard to determine in Figures 5b, c and d because there is no evidence of large difference.

In addition, the model performance in the main basin districts (for their initial calibrated value) is comparable in Figure 5 (a-d). For instance, the KT basin shows a RMSE around 13 mm over 11-years of simulated discharge which presents the lowest model performance among the other main basin districts. In contrast, its insensitive neighboring basin, BP, shows a lowest RMSE (around 7 mm) among the considered basins; which indicates the best model performance for this basin.

For a clearer presentation of spatial sensitivity for the 4 most sensitive parameters Figure 6 (a-d) shows the sensitivity curves of Figure 5 in a different way. In these figures the sensitivity curves of the four parameters are presented for each of the main basin districts. These graphs help to find likely differences of the parameter sensitivities due to their spatial variation.
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

a) BB

b) BS
Figure 6. Sensitivity curves comparison for recesscoef1, recesscoef2, grcoef and seepcoef in the considered main basin districts in Sweden a)BB, b)BS, c)BP and d)KT, RMSE was calculated for modeled and observed runoffs over 1990-2000.

By Looking to Figure 6a, it is clear that parameters recesscoef1 and grcoef are the most sensitive. The parameter recesscoef1 shows a higher sensitivity than grcoef in the domain of 0 to 40% increase. There is higher sensitivity, however, for grcoef between 40% and 60% rise when the
recesscoef1 curve goes to its minimum value and the grcoef curve is a steep straight line. Further, the recesscoef1 sensitivity curve tends to be steeper when increasing its initial value from 40% to 100%. Similarly, in the left half (decreasing area) of Figure 6a recesscoef1 is the most sensitive parameter. However, in the changing domain of -80% to -100% recesscoef2 is the most sensitive parameter. This change of recesscoef2 over recesscoef1 is visible in the BS basin (Figure 6b). Besides, the parameter recesscoef2 is more sensitive in the BB and BS basins compared to grcoef. In the BS basin (Figure 6b) the parameter recesscoef1 and grcoef have higher sensitivity compared to recesscoef2 and seepcoef. With respect to parabolic shape of recesscoef1, when its curve goes to the minimal point by changing the parameter from 40% to 60% of its initial value, the other parameters grcoef1, recesscoef2 and seepcoef appear to have higher sensitivity. Further, from 60% to 80% increases the parameter recesscoef1 is the most sensitive parameter. In the BP basin (Figure 6c) the parameter recesscoef1 is also the most sensitive parameter. The parameter grcoef has the higher sensitivity compared to recesscoef2 and seepcoef in the KT basin (Figure 6d), except in decreasing domains -80% to -100 in which recesscoef2 has the higher sensitivity. The parameter recesscoef1, however, behaves as the most sensitive parameter over most conditions.

We can approximate a sensitivity index by calculating the absolute slope of the sensitivity curves for the parameters considered for each of the 36 Swedish watersheds. The spatial variability of the parameters’ sensitivity can be assessed spatially by mapping these indices in Figure 7(a-d). The slope calculation was limited from -40% to +40% of the changes in parameters initial value to avoid the influence of extreme changes; therefore, each map categorizes the results in 4 domains.
Figure 7. Sensitivity map for a) recesscoef1, b) recesscoef2, c) grcoef and d) seepcoef in 36 Swedish watersheds, data from 1990-2000.
The sensitivity map of recesscoef1 shows remarkable sensitivity index variation for the 36 Swedish watersheds. As presented in Figure 7a, the sensitivity of recesscoef1 in the northern watersheds is higher than in the central watersheds. This tendency of higher sensitivity can also be found in small southern watersheds. In addition, by increasing the parameter recesscoef1 the sensitivity index decreases for the majority of watersheds. The parameter recesscoef2, however, has a lower range of variation for the sensitivity index compared to recesscoef1 (Figure 7b). Again, the northern and small southern watersheds have higher sensitivity indices. Besides, the central watersheds (except Norrström which shows no response to the parameter changes) become more sensitive when decreasing the value of recesscoef2. When grcoef is changed, however, the sensitivity index does not display any spatial patterns similar to the previously considered parameters. As shown in Figure 7c, northern watersheds behave differently; a collection of high-sensitive and insensitive watersheds can be seen in northern part of Sweden. The very northern watersheds with higher sensitivity become more sensitive as grcoef decreases, while those with lower sensitivity maintain a constant sensitivity index. The small southern watersheds are sensitive with similar patterns as displayed for the other parameters. In addition, the range in which the sensitivity index estimated for grcoef varies is larger than that of recesscoef2 and lower than that of recesscoef1. According to the presented results in Figure 6(a-d) there is no surprise that the sensitivity index of seepcoef shows only a slight variation. In Figure 7d the majority of watersheds show very low values of sensitivity index. In this case, the range of the sensitivity index variation is the lowest among the 3 parameters.

To characterize the sensitivity of each parameter the sensitivity index (the absolute slope of the sensitivity curves in each 20% domain) of all 36 Swedish watersheds were averaged. This is presented in Figure 8 for recesscoef1, recesscoef2, grcoef and seepcoef.
Again, the parameter recesscoef1 has the highest averaged-sensitivity index compared to the other parameters. The sensitivity appears to increase as the values decrease. The second most sensitive parameter seems to be grcoef while there is no evidence of large variation in its sensitivity by changing the parameter. The parameter recesscoef2 is situated at the third place. Its averaged sensitivity index seems to increase slightly by decreasing the parameter values. The lowest averaged sensitivity index is for seepcoef which has an averaged sensitivity index very close to zero without any obvious change in the considered domain.

### 6.1.2 Seasonal sensitivity analysis

Following the annual CSIM model sensitivity analysis, the model results and parameter sensitivities were considered from a seasonal perspective to evaluate temporal variations in sensitivity of the CSIM model in the main Swedish basins over years 1990 to 2000. Figures 9 (a-d) to 12 (a-d) present the temporal sensitivity (seasonal scale) of the four hydrologic parameters in the four main basin districts considered.
The results in Figure 9 present variation in sensitivity of recesscoef1 at the seasonal scale. As presented, the KT basin shows a steeper sensitivity curve during winter and afterward its sensitivity decreases during spring. It seems that the sensitivity of recesscoef1 tends to be increased slightly in the summer and this increasing continues to the end of autumn. In this case the parabolic shape of the KT basin’s curve remains constant while its slope under decreasing values of recesscoef1 (left side of plots) changes more than under increasing values. The basins BB and BS have relatively similar behavior during winter and spring. Their lowest sensitivity in winter is replaced by the highest spring-sensitivity among the other basins. However, in summer the behavior of the BB and BS is different. BS keeps its shape with a lower sensitivity than spring and the BB curve tends to have smooth shape. The winter trend, again, comes back in autumn. This, in general, leads to the lowest relative sensitivity for these basins among those considered here. The watershed BP displays a sensitive behavior in winter with a declining trend of sensitivity during spring and summer, while in summer loses its sensitivity.
Figure 10. recesscoef2 seasonal sensitivity graph for the main Swedish basin districts over 1990-2000, Top-left) Winter, Top-right) Spring, Bottom-left) Summer and Bottom-right) Autumn.

The graphs in Figure 10 comprise a lower sensitivity variation for the parameter recesscoef2 in different seasons. In winter the KT basin shows the highest sensitivity compared to the other basins. In spring, because of the quite insensitive behavior, it is difficult to rank the basin’s relative sensitivity to this parameter. In summer, the basin BB has the highest sensitivity and thereafter, the BB and BS basin sensitivity becomes rather similar. Among the four main basin districts, in the case of recesscoef2 seasonal sensitivity evaluation, BP has the lowest rank of sensitivity, as its sensitivity curve tends to be straight in all seasons.
The seasonal variation in the sensitivity of grcoef is drawn in Figure 11. As presented in these graphs, the southern basin, KT, in winter shows slightly more sensitivity compared to the other basins. In spring, the BS basin tends to show more sensitivity than the other basins. In summer, the straight shape of the BS sensitivity curve is changed to parabolic and generally, there is more sensitivity across all basins. In autumn, all of the basins seem to behave insensitive compared to spring or summer.
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

Figure 12. Seepcoef seasonal sensitivity graph for the main Swedish basin districts over 1990-2000, Top-left) Winter, Top-right) Spring, Bottom-left) Summer and Bottom-right) Autumn.

The seasonal sensitivity of the parameter seepcoef, as shown in Figure 12, seems to remain rather constant in different seasons and in different basins. By zooming-in to the seasonal graphs, however, the watersheds BB and BS show slightly different behavior compared to the quite insensitive basins BP and KT.

The other noticeable factor in the presented figures is variation of the model performances in different seasons (i.e., the magnitude of each curve in the figures). By considering the initial calibrated values (e.g., x=0), different model performances for the considered basins could be seen in four seasons. Table 3, presents the seasonal model performance for the four considered main basin districts at their initial calibrated value.

<table>
<thead>
<tr>
<th></th>
<th>BB</th>
<th>BS</th>
<th>BP</th>
<th>KT</th>
</tr>
</thead>
<tbody>
<tr>
<td>Winter-RMSE</td>
<td>6.56</td>
<td>5.97</td>
<td>9.52</td>
<td>11.11</td>
</tr>
<tr>
<td>Spring-RMSE</td>
<td>13.38</td>
<td>28.74</td>
<td>11.87</td>
<td>9.78</td>
</tr>
<tr>
<td>Summer-RMSE</td>
<td>19.63</td>
<td>11.10</td>
<td>8.22</td>
<td>5.01</td>
</tr>
<tr>
<td>Autumn-RMSE</td>
<td>9.89</td>
<td>5.71</td>
<td>8.22</td>
<td>11.24</td>
</tr>
</tbody>
</table>

Table 3 shows that in winter KT has the highest model error and BS has the highest model performance. This rank is repeated in autumn. In spring, however, BS has the lowest performance while KT has the best model fit. KT keeps its position with best model performance in summer while
BS is replaced by BB as the worst modeled basin. BB and BS have their worst model performance in summer and spring, respectively. BP has its best model performance in summer and its worst in spring. Similarly, KT has its best model fit in summer and its worst in autumn.

6.2. The Genetic Algorithm Optimization analysis

6.2.1 Annual GA optimization

The parameters recesscoef1, grcoef, recesscoef2 and seepcoef (from the sensitivity analysis) were considered for GA optimization in four selected watersheds. This optimization includes annual (and seasonal – see following section) searches for the optimum set of parameters in the Mörrumsån, Dalälven, Ångermanälven and Torne älv watersheds. These watersheds we selected as they span the whole of Sweden.

A first step in the GA optimization was to explore potential algorithm setups and determine an adequate approach to restrict the search algorithm. As such, three scenarios were defined to compare the optimized results (see Appendix A for all results). The first scenario used the defined limits of parameter changes from -100% to +100% of their initial values; this is called here after coef_1 scenario. For the second scenario, coef_2, the initial population of the parameters recesscoef1 and grcoef were replaced by the results from the first optimization scenario. This means that GA looks for the optimized set with regards to recesscoef2 and seepcoef while it has the optimized set for the recesscoef1 and grcoef from the prior detection (i.e., the coef_1 scenario). The third scenario (coef_3), however, is not similar for the four selected watersheds. For Torne älv, the parameters recesscoef2 and seepcoef remained constant and GA only searched for the optimum sets of recesscoef1 and grcoef. For the rest of selected watersheds different bits were defined for the model parameters to prepare a wider range of search for the less sensitive parameters. For the Mörrumsån watershed in the scenario coef_3 alongside assigning different bits to the parameters the selection rate, X_r, was set to 0.6 and the initial and final mutation rates were set to 0.03 and 0.003, respectively. Table A2 in Appendix A summarizes the characteristics of three different scenarios while Figures A4-A5 compare the calculated RMSEs of three different optimization scenarios in the selected watersheds.

Once the appropriate algorithm setup was determined, it was possible to optimize the best set of parameters for each watershed (Table 4). Table 4 presents the optimized set of parameters and the corresponding RMSEs of the selected watersheds. For comparison, the RMSEs from the parameters by Mörth et al. (2007) determined using the SDK solver calibration are shown; this serves as a base
to compare the optimization results. Also, in Table 4, the multiplication factor is the value by which the originally calibrated parameter values needs to be scaled to obtain GA optimization.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Ångermanälven</th>
<th>Torne älv</th>
<th>Dalälven</th>
<th>Mörrumsån</th>
</tr>
</thead>
<tbody>
<tr>
<td>recesscoef1</td>
<td>1.83</td>
<td>1.55</td>
<td>1</td>
<td>1.29</td>
</tr>
<tr>
<td>Grcoef</td>
<td>0.75</td>
<td>1.7</td>
<td>1.99</td>
<td>0.73</td>
</tr>
<tr>
<td>recesscoef2</td>
<td>1.2</td>
<td>0.64</td>
<td>2</td>
<td>1.25</td>
</tr>
<tr>
<td>Seepcoef</td>
<td>1.91</td>
<td>0.66</td>
<td>0.002</td>
<td>0.87</td>
</tr>
<tr>
<td>RMSE-SDK solver</td>
<td>14.89 mm</td>
<td>9.1 mm</td>
<td>10.68 mm</td>
<td>8.51 mm</td>
</tr>
<tr>
<td>RMSE-GA</td>
<td>14.63 mm</td>
<td>6.67 mm</td>
<td>9.25 mm</td>
<td>6.96 mm</td>
</tr>
</tbody>
</table>

For the Ångermanälven watershed, there is only a minor difference between the RMSE of the model using the initially calibrated parameters and those from the GA optimization. Basically, the GA optimization was able to enhance the Ångermanälven model discharge prediction by just 0.26 mm. What is clearly different, however, is the actual values for each parameter used to achieve these optimized results (i.e., the numbers in Table 4 with a value of 1 meaning the parameter does not change from the initial value). Under the GA optimization, for example in Ångermanälven, the parameter recesscoef1 should be multiplied by 1.83, grcoef with 0.75, recesscoef2 by 1.2 and seepcoef by 1.91 of their original values.

For the watershed Torne älv, the best GA optimization results can be compared with the initially calibrated parameters in the CSIM model that produce a RMSE of 9.1 mm. In the optimization process of Torne älv watershed the GA could increase the CSIM model performance seen through a reduction in RMSE to 6.67 mm. Again, the parameter set identified to achieve this reduction in RMSE was quite different than the original parameter set.

For the Dalälven watershed, the initial calibrated parameter values from Mörrth et al. (2007) give a RMSE value of 10.68 mm. For the parameter recesscoef1 the multiplication factor is 1 and for grcoef and recesscoef2 the multiplication factor is around 2. The multiplication factor for the parameter seepcoef is close to zero in the Dalälven watershed.

In the southern watershed, Mörrumsån, the RMSE of the first calibration with SDK solver by Mörrth et al. (2007) was 8.51 mm. With regards to the best set of parameters presented in Table 4, the RMSE reduces to 6.96 mm which indicates the model performance improvement compared with
SDK solver results. Similarly to the other watershed, there is a large variation in multiplication factors between the defined parameters in the GA optimization and the original parameters.

### 6.2.2 Seasonal GA optimization

Based on the CSIM model performance variation and parameter sensitivity observed in different seasons, seasonal optimization of the CSIM model was carried out. These seasonal GA optimizations create uniquely optimized parameter sets for each watershed for each season. The multiplication factors for each of these seasonally optimized parameter sets can be compared to the multiplication factors obtained from the annual GA optimization. Figures 13 to 16 present the outcomes of the GA seasonal optimization for the Ångermanälven, Torne älv, Dalälven, and Mörrumsån watersheds, respectively.

As presented in Figure 13, the annual optimized multiplication factor for the parameter recesscoef1 (1.83) is closest to the seasonal multiplication factor in summer (1.93). The other seasons have quite different values compared to the annual factor. The annual multiplication factor of 0.75 for the parameter grcoef is lower than all of the seasonal multiplication factors (which are over 1). Similar for the parameter recesscoef2, it is rather hard to find the closest season to the annual value. The
highest multiplication factor for the parameter seepcoef is situated in spring, 1.86, and it is closest to the annual value when considering the annual multiplication factor equal to 1.91.

Figure 14 compares the seasonal optimization results of the GA with the annual values for the Torne älv watershed.

![Figure 14. Comparison of the seasonal and annual GA optimization results for the watershed #10](image)

With regards to the GA outcomes, the parameter recesscoef1 should be increased in three seasons of a year (winter, summer and spring) and decrease around 10% in autumn to achieve better model performance. For the parameters recesscoef2 and seepcoef, however, this is changed to three seasons where the parameter decreases and one season where it increases. For the parameter grcoef, the annual multiplication factor (1.70) would be nearest to the autumn and summer multiplication factors. The multiplication factor for the parameter recesscoef2 is smaller than one for all seasons except spring; thus the annual optimized value, 0.64, is most similar to winter and autumn. For the parameter seepcoef, with seasonal multiplication factor near to zero in winter, summer and spring and 1.09 in autumn, finding the best seasonal fit with the annual value is impractical.
The comparison of seasonal and annual optimization results of the Dalälven watershed are presented in Figure 15.

For the Dalälven watershed, it is not difficult to find the closest seasonal values to the annual values. The parameters grcoef and recesscoef2 are constant in all seasons (1.99); the annual optimized multiplication factor for both of the parameters is 1.99 which fits with the optimized value of all seasons. For the parameter seepcoef which has a multiplication factor very close to zero (0.05) in whole year, the annual multiplication factor (roughly 0.002) could be considered coincident with all values. The annual multiplication factor for the parameter recesscoef1 is equal to 1. Among the four seasons, winter and summer are assigned the multiplication factor quite close to one.

Figure 16 presents the seasonal optimization results of the GA beside the annual optimized values for the Mörrumsån watershed.
Figure 16. Comparison of the seasonal and annual GA optimization results for the watershed #93

The variation of multiplication factor in different seasons is quite noticeable for the parameters grcoef, recesscoef2 and seepcoef. For instance the multiplication factor for the parameter seepcoef varies from 0.05 in summer, winter and autumn to 1.99 in spring. The parameter recesscoef1 has an increasing optimized factor for all seasons. The closest seasonal value to the annual factor (1.29) lies between winter (1.14) and spring (1.38). The parameter grcoef experiences the multiplication factors of 0.16 in spring to 1.99 in winter. The closest value to the annual multiplication factor (0.73) could be found in summer around 0.86. Similarly, the annual value of the parameter recesscoef2 (1.25) has the closest seasonal value in summer (1.19).

6.2.3 Model improvement through GA optimization

The percentage of reduction in RMSE (both annual and seasonal) for the CSIM model by using the GA optimization is presented in Table 5 for the selected watersheds. It should be noted, by considering the negligible difference between RMSEs in different annual (optimization) scenarios, the seasonal optimization process was performed with no weighting criteria.
Table 5. The annual and seasonal CSIM model reduction in RMSE by the GA compared to SDK solver

<table>
<thead>
<tr>
<th></th>
<th>Annual</th>
<th>Winter</th>
<th>Spring</th>
<th>Summer</th>
<th>Autumn</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ångermanälven</td>
<td>1.70%</td>
<td>0.90%</td>
<td>3.72%</td>
<td>9.49%</td>
<td>9.06%</td>
</tr>
<tr>
<td>Torne älv</td>
<td>26.66%</td>
<td>2.34%</td>
<td>54.27%</td>
<td>25.24%</td>
<td>35.83%</td>
</tr>
<tr>
<td>Dalälven</td>
<td>13.41%</td>
<td>12.49%</td>
<td>10.52%</td>
<td>5.92%</td>
<td>25.01%</td>
</tr>
<tr>
<td>Mörrumsån</td>
<td>18.25%</td>
<td>4.69%</td>
<td>28.37%</td>
<td>32.60%</td>
<td>2.92%</td>
</tr>
</tbody>
</table>

With regards to Table 5, the Ångermanälven watershed faces to the worst CSIM model performance in summer, however, the GA optimization was able to decrease the RMSE by 9.5%. In winter which is the season with the highest model performance, the CSIM model RMSE could be reduced by only 0.9% if the initial calibrated values were replaced by GA optimization parameter set.

In the Torne älv watershed, the GA was able to decrease the RMSE in the worst model performance season (summer with RMSE equal to 13.46 mm) by 25.24%. Further, in spring, the GA improves the ability of the CSIM model in discharge prediction in Torne älv by more than 54%. In winter, however, the model predicts the discharge much better than the other seasons with the RMSE equal to 3.15 mm such that, in this case, GA could not decrease the RMSE more than 2%.

With respect to Dalälven, the GA results would improve the CSIM model prediction ability of the two worst seasons, summer and spring, by 5.92% and 10.52%, respectively. For the seasons with best discharge prediction efficiency, the RMSE is reduced by 25% in autumn and 12.49% in winter.

By knowing that the Mörrumsån watershed is located in the BP basin, as stated earlier in Table 3, the seasonal variation of the CSIM model performance is quite low. The best seasonal model performance occurs in summer when the SDK solver RMSE is 5.31 mm. The GA optimized set of parameter can decrease the model RMSE by 32.6% in summer. In winter (in which the SDK solver gives the highest RMSE close to 9.24 mm) the GA can decrease the model error to 8.81 mm (less than 5%). In spring the optimization process can improve the model discharge prediction ability by 28.4% while the GA potency of RMSE reduction decreases to less than 3% in autumn for the Mörrumsån watershed.

In summary, the results show the highest annual and seasonal model RMSE reduction for the pristine watershed, Torne älv, close to 27% reduction annually and 54% in spring (Table 5). However, for the heavily dam regulated watershed, Ångermanälven, the GA could not decrease the annual model RMSE more that 1.7%. Following the GA optimization results, to reduce the seasonal RMSE of a watershed like Ångermanälven in summer by less than 1 percent there is remarkable change in the model parameters required.
As several different parameter sets are achieved from the seasonal optimization process for the hydrological parameters, the role of these best seasonal parameters could be tested on the annual model prediction performance. As such, the 4 seasonally obtained parameter sets were used in the CSIM model (instead of applying only one annually derived optimum set of parameters) for the selected watersheds to evaluate the annual model response. Table 6 compares the annual CSIM model responses using different sets of optimized values (specifically, the SDK solver set, GA-annual set and GA-4 seasonal parameters sets).

<table>
<thead>
<tr>
<th></th>
<th>Ångermanälven</th>
<th>Torn älv</th>
<th>Dalälven</th>
<th>Mörrumsån</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDK solver- annual RMSE</td>
<td>14.89 (0%)</td>
<td>9.1(0%)</td>
<td>10.68(0%)</td>
<td>8.51(0%)</td>
</tr>
<tr>
<td>GA-annual RMSE</td>
<td>14.63 (1.8%)</td>
<td>6.67(26.7%)</td>
<td>9.25(13.4%)</td>
<td>6.96(18.2%)</td>
</tr>
<tr>
<td>GA-Annual RMSE using seasonal optimized values</td>
<td>13.75 (7.7%)</td>
<td>6.09(33.1%)</td>
<td>8.73(18.3%)</td>
<td>6.32(25.7%)</td>
</tr>
</tbody>
</table>

It is clear, from the results in Table 6 that the annual CSIM model performance is enhanced by using seasonal sets of parameters obtained from GA seasonal optimization. Therefore, it could be claimed that the application of 4-seasonal sets of optimized parameters in the CSIM model is a twofold beneficial method for enhancing both annual and inter-annual CSIM model performances.
7. Discussion

7.1. Annual sensitivity analysis

The annual sensitivity results show remarkable spatial variability in Sweden. Several regions and watersheds come up as being rather insensitive to the parameters of the CSIM model. Considering sensitivity curves of the main basin districts presented in Figure 5 (a-d) and watershed sensitivity maps shown in Figure 7 (a-d), the BP basin (especially Norrström and Motalastraöm watersheds) has the tendency to be less sensitive to the parameters in CSIM. This could potentially be due to large amount of lake area in that region, roughly 10% (Table 1). Mälaren and Vättern lakes, the two largest lakes in Sweden with water residence times on the order of years to decades, are located in this region. These lakes could dampen the potential influence of groundwater (as it is characterized in the CSIM model) to make BP function more closely to a flow-regulated basin. A second factor could also be the high rate of annual surface runoff production in BP by Norrström and Motalastraöm rivers (Humborg et al., 2008). Again, this would limit the impact of groundwater parameters in the CSIM model performance relative to other process representations. Thus, in Norrström, where the average runoff ratio between 1985 and 2010 is 3.1, the infiltration into the groundwater boxes in CSIM might be relatively minor compared with surface runoff. Therefore, the calculated RMSE of the BP basin shows less variability to changes of groundwater parameters except in some southern small watersheds.

Looking at KT basin and considering the areas occupied by lakes (roughly 14%, Table 1) coupled with the high rate of surface runoff from, for example, Göta älv, the same behavior as BP would be expected. However, its behavior in Figures 5a and 5c shows more sensitivity to changes in parameter values. This could be interpreted as an amplified influence of groundwater in this area. This would be consistent with, for example, the existence of Sweden’s largest aquifer, Alnarpsströmmen. The groundwater effect on the model performance in KT basin is, accordingly, stronger than that in the BP basin. It could be explained by the differences of precipitation in south-west and south-east of Sweden (roughly twice as south-east). As such, the high rate of precipitation in south-west of Sweden causes a high rate of groundwater accumulation (Swedish Environmental Protection Agency, 2000). Figure 6d shows sensitivity curves of different parameters in the KT basin with the steeper curves of recesscoef1 and grcoef presenting engagement of the two groundwater boxes in this basin.

In the CSIM model, the runoff is simulated by coarse representation of surface runoff and macropore flow. Conceptually, the CSIM model omits shallow groundwater discharge with transit time longer than a day. Indeed, in dam-regulated watersheds such as those in the BS basin (Ume älv,
Parameter sensitivity and optimization of a catchment-scale hydrologic model across Sweden

Ångermanälven, Indalsälven, Ljungan, Ljusnan, Dalälven (Humborg et al., 2008)) river conversion into lakes increases the water residence time. Therefore, the modest sensitivity slope of BS (compared to BB) to the parameter recesscoef1 could be interpreted as the excess precipitation accumulated in dam reservoirs that infiltrates to lower soil layers later (Figure 5a). In contrast, in the drainage area to the BB basin which is almost an unperturbed mountainous region the sensitivity of recesscoef1 is rather high (Figures 3a and 5a). By considering perturbed conditions against unperturbed, the two dam-regulated basins Skellefte älv and Lule älv show different behavior than rest of the BB basin (Figure 7a).

Summary of all parameters sensitivity is presented in Figure 8 by calculating the average RMSE absolute slope over the 36 Swedish watersheds. In agreement with Lee et al. (2000) approach, the figure clearly shows that the parameter recesscoef1 is the most sensitive parameter in the CSIM model. The sensitivity of recesscoef1 increases sharply when decreasing the parameter while its sensitivity decreases by increasing the parameter from its initial calibrated value. The increasing trend might be due to upper soil-box saturation be caused recesscoef1 reduction. Under such condition, if there is no vertical way for water escape, the model becomes more sensitive to the value of recesscoef1 that regulation flow out of this upper box. The second most sensitive parameter, grcoef, controls the loss of near surface water into the deeper layer. The third most sensitive parameter is recesscoef2 and seepcoef is the least sensitive. The reason of this ranking is likely linked to the response of the CSIM model to the two different groundwater flows. As stated earlier, the stream-flow consists of the aggregated surface runoff from individual land uses and groundwater flow from two compartments in the saturated zone. The upper box of groundwater discharge has a relatively quick response to precipitation while the lower box responds slowly to climate with no substantial effect on seasonal stream discharge. In the boreal forested hillslope watersheds (like those found across Scandinavia) because of the high hydraulic conductivity of the upper layer, the main groundwater flow is a lateral flux in the upper soil horizon. This process-based difference makes the upper layer likely more sensitive compared to the lower groundwater box in the current CSIM model representation.

7.2. Seasonal sensitivity analysis

The seasonal sensitivity results indicate temporal variability of the CSIM parameter sensitivity in Sweden. The seasonal sensitivity graphs (Figures 9, 10 and 11) indicate that the northern Swedish basins which drain into Bothnian Bay and Bothnian Sea (BB and BS) display more sensitivity in spring and/or summer to changes of the parameters recesscoef1, recesscoef2 and grcoef compared to the southern basins. This can be explained by their sub-polar climate (Voss et al., 2011), land cover
domination mainly by boreal forest (49%-59%, Table 2) and wetland coverage (7%-12%, Table 2) and, of course, soil characteristics. As stressed by Bledring et al. (2000) till soil, which is deposited by glacier ice, is the main surface deposit in the Nordic countries (e.g. Sweden). This is poorly sorted sediment with a wide range of particle size. The till soil dominated landscape mainly has a thin soil cover (about 5-10 m according to Smedberg et al. (2009)) with a high saturated hydraulic conductivity. The thin and unconsolidated structure of till soils causes the short water residence time of these boreal regions. The high specific discharge and low temperature (due to the high latitude) which brings rather low evapotranspiration (10%-30% of precipitation) can further shorten the water residence time (Voss et al., 2011). Although, there is no certain estimation about the water residence time in boreal watersheds, Laudon et al. (2007) claim that for the small boreal watersheds the likely water residence time varied from weeks to a few months whereas, Burgman et al. (1987) contend for the larger main [northern] Swedish basins (e.g. BB and BS) the water residence time is in average 11 months (from 2 to 27 months). With regards to high hydraulic conductivity of near surface layer in till dominated areas and therefore, the groundwater level rising into soil horizons, more rapid lateral flow by mobilizing of the pre-event water stored in the hill-slope to stream is anticipated (Laudon et al., 2007). Thus, the groundwater outflow from saturated and unsaturated discharge areas may respond rapidly to snowmelt and rainfall (Bledring et al., 2000). The higher sensitivity of the northern Swedish watersheds to the parameter recesscoef1 would be supported, as well, by the findings of Voss et al. (2011) who state that roughly 50% of the runoff in boreal and sub-arctic rivers is generated within a few weeks in May and June flushing specially the rich organic matter of the top soil.

The high saturated hydraulic conductivity of the boreal watersheds, however, decreases rapidly with soil depth on bedrock of comparatively low hydraulic conductivity (Lind and Lundin, 1990). This is coincident with longer water residence time in deeper soil layers which is called groundwater box-II in this study.

With regards to Laudon et al. (2007), who studied 15 nested boreal streams in northern Sweden, the soil-water isotope in the wetland dominated watershed displays two hydrological pathways for water flow. First, is a shallow pathway close to the surface and second, so called preferential pathway at 200 to 250 cm depth. This preferential pathway, presumably, is caused by a continuous frozen melt-water which is infiltrated to the peat. The melt-water, probably, originates by infiltrated water from the forest (Table 2) surrounds the wetland or from the wetland perimeter (flowing from the layers with higher hydraulic conductivity) (Sirin et al., 1998). A thick frozen soil layer which originates due to wet conditions and high groundwater levels in autumn (before the cold winter) is a
common event in wetlands in boreal areas (Laudon et al., 2007). This impermeable frozen soil layer does not let the water infiltrate into the soil. Hence, the water finds its way to overland or through deeper preferential pathway. Therefore, the spring hydrograph in wetland dominated watersheds has similar proportion of event and [deep] pre-event water (Laudon et al., 2007). Consequently, it can be inferred that the seasonal sensitivity of the BB and BS basins caused by change of parameters participating into and from the deeper groundwater box, grcoef and recesscoef2, may be related to their occupied area by wetlands. According to Table 2, wetland area is roughly 12% in the BB basin and 7% in the BS basin. Whereas, the effect of damming in the BB basin that lengthens the water residence time could also be considered as well.

The southern main basin district, e.g., KT, however, displays earlier and gentler seasonal sensitivity to the parameters recesscoef1, recesscoef2 and grcoef shift compared to northern basins. This happens in KT for the parameter recesscoef1 in winter and thereafter again in autumn. For the parameters grcoef and recesscoef2 the highest sensitivity is shown in winter. Similar to the BB and BS basins the reason could be found in land and climate characteristics of this basin. With regard to Table 2, KT is dominated by forest around 60% of its area; besides, it has wetland area close to 3% while the lake occupation is roughly 14%. The climate is oceanic temperate with productive soil on glacial sediments support intensive agriculture (Voss et al., 2011). There is no doubt, hence, that the snow melt event in southern Swedish watersheds happens earlier compared to northern sub-arctic watersheds. Moreover, the higher temperature in this region causes more evapotranspiration from cultivated land, forest and specifically from lakes. As Dingman (1973) discussed the higher rate of evapotranspiration brings low recession constants. Hence, the more moderate sensitivity behavior of KT (compared to northern basins) may be associated to these lower recession constants.

The BP, however, presents seasonal sensitivity just for the parameter recesscoef1 mostly in winter and spring and afterward in autumn. By warming up of the weather, summer drives the basin to be insensitive. It could be translated to engagement of only first groundwater box in the CSIM model.

7.3. Optimization process

Mörth et al. (2007) showed that the CSIM model inter-annual performance of the 105 studied watersheds in BSDB show temporal discrepancies in some cases; while, overall, the annual CSIM model performance for predicting the annual stream-flow and Tot-N loads is fairly well. They showed on average, the inter-annual correlation coefficient per watershed was 0.69 for stream-flow and for the major watersheds this ranged between 0.13 and 0.75. Accordingly, they suggest the seasonal flux calibration for all rivers to achieve better inter-annual correlation coefficient between simulated and measured data.
By using the CSIM model parameters in the annual and seasonal optimization processes, this study indicates the success of GA optimization, in both annual and seasonal hydrological events (Table 5). It could be claimed that the simultaneous annual and seasonal CSIM performance enhancement is in the hand of using four-seasonal sets of parameters, although, it may cause complexities in the model structure.

The quite low ability of the GA in increasing the CSIM model performance in the Ångermanälven watershed (annual and seasonal) which needs remarkable changes of the model parameters (Figure 13), however, draws a new perspective. This leads to think about the probability of there being more than one optimum set of parameters available for the CSIM model and seeing the problem through the lens of “Equifinality”. For the better understanding of equifinality in the CSIM model the two most sensitive parameters, recesscoef1 and grcoef, were selected to draw a response surface for the CSIM model (Figure 17).

![Figure 17. The CSIM model response surface (RMSE) for the parameters recesscoef1 and grcoef, the X and Y axis show the percentage of change in each parameter.](image)

While the dark blue area represents the minimum RMSE due to change of the parameters recesscoef1 and grcoef, it is apparent that there is more than one combination of these two parameters available in the “best fit” area. The acceptance domain of the parameter recesscoef1 changes between its initial calibrated values (0) and 5% more than that. However, the acceptable changing range for the parameter grcoef varies between -10% and 40%. Consequently, while the parameter grcoef shows less sensitivity compared to the parameter recesscoef1 the wider range of values can satisfy the best fit area via a low RMSE value. Similarly, the GA process starts to find the
best value for the most sensitive parameters (recesscoef1 and grcoef) and further, searches for the best set of the least sensitive ones (recesscoef2 and seepcoef); whereas, the wider range of those could give the best model fit. This is in agreement with the current sensitivity analysis of this study.

Nevertheless, it should be noted that, to date, no method could be found to optimize the structure of a model which is [presently] a subjective understanding of the hydrologic system by hydrologists (Hreiche et al., 2002). In relation to the concept of equifinality, according to Beven (1993), this (equifinality) carries two other concepts: first, the equal probability of a given solution and second, the “acceptability”. Although, the acceptability can be harnessed by limiting the model parameters’ domain, it is inevitably a subjective issue. Hreiche et al. (2002), however, noted that these two concepts and equifinality differ in their application. The equifinality, for example, had been neglected for a long time when there was no tendency to search for the best “set” of parameters as long as one set could give good results. To this end, Sorooshian and Gupta (1983) stressed three reasons that could violate the principle of uniqueness: (1) structure of the model, (2) the model inadequacy in representing reality and the data and (3) inherent errors.

Unluckily, despite the evidence from GA which reveals that the CSIM model (like any hydrological model) faces to equifinality, it could not address in which phase the model endures more uncertainties. For example, Li et al. (2010) stressed the prediction failure of the GWLF model in unusual high storm-flow conditions. They quoted from Lee et al. (2000) that this poor performance of the GWLF model could be due to model inadequacy in storm-flow prediction; while they believed that the storm-flow calibration does not necessarily lead to a good partitioning between base flow and stream flow. Hence, for an adopted model from GWLF, with non-realistic compartment of process representation, there is no surprise if the same inadequacies exist. This inadequacy could be seen in Table 3; while in spring the average CSIM model performance of the considered main basin districts over this season shows the worst model discharge prediction ability compared to other seasons. Taken together, as the parameter sensitivity (spatial and temporal) of the model could be justified by physical concepts, watershed characteristics and climatic events, this can also address the uncertainties in the CSIM model structure or/and observed values (e.g. in flood or/and drought periods).

Optimization, in general, assumes that the model and the observations are free from errors, although, it is clear at least for hydrological models that both the model and observations are not error-free. Importantly, even an improved calibration method is not capable to reduce the model structure’s errors which cause prediction discrepancies (Li et al., 2010). Besides, the concept of an optimum parameter set might be ill-founded in hydrological modeling in the reason of equifinality as
usually many parameters set could be found that give reasonable fit to the observations. When an optimum parameter set gives only a single prediction, the multiple parameter sets prepares a range for the prediction. Nevertheless, equifinality may be considered as an advantage for assessing the uncertainties in prediction and decision making (Beven, 2008).

8. Conclusion

The outcomes of the sensitivity analysis, as it was hypothesized, showed remarkable variations in both spatial and temporal parameters sensitivity in the CSIM model. In case of model calibration, identification of the most sensitive parameters (instead of using the whole hydrological parameters) in optimization process helped to accelerate the operation; as the greater sensitivity of the model response to a parameter leads to a faster parameter optimization process (McCuen, 1973). By considering the uncertainties that the CSIM model faces, the spatial sensitivity results of this study propose the question of the model regeneration capability: if one or more component of the CSIM model leads to a low change in the output response in a watershed or in a season the model complexity could be reduced/revised in that area or time, even though, the CSIM model response could be justified by existing concepts. In such case, McCuen (1973) offers feasibility evaluation of using “component” sensitivity. This helps to measure the importance of a specific sub-system in the model (e.g. revising the evapotranspiration process from different land covers and open water bodies). Moreover, for the most accurate application, the spatial parameter sensitivity emphasizes on site-calibration (local calibration) of the model parameters specifically.

It is found that the groundwater recession coefficient (recescoef1) is one of the most important parameters in the CSIM model. Hence, for the watersheds which show higher engagement (sensitivity) of this parameter in the CSIM operation, the hydrograph separation analysis (e.g. straight-line method and fix-base methods) for defining the parameter could be recommended (see Linsley et al. (1975), Dunne and Leopold (1978) and Chow et al. (1988)). This could be accomplished as well with regards to temporal sensitivity of the parameter.

Following the optimization results, it was presented that the GA could improve the CSIM model performance in both annual and seasonal scales. The CSIM seasonal optimization results, however, gave a better model predictive performance when each season was assigned a specific parameter set. This may point to the feasibility of using four-seasonal values for the most sensitive parameters in the CSIM model to achieve both annual and inter-annual predictive improvement since it increases the number of parameters needed in the modeling.

The importance of the equifinality concept was discussed, as the optimization process presents multiple sets of acceptable parameters rather than only one. This helps to assess the uncertainties
that the CSIM model, as a hydrologic model, carries in either its formulation or in the data considered for calibration.

Acknowledgments

I am truly indebted and heartily grateful to my advisor, Dr. Steve Lyon, for his kind support and guidance. Dear Steve, I am sure that this project would have not been possible without your help and encouragement. Also, I would like to show my gratitude to Dr. Carl-Magnus Mörth for giving me the opportunity of working with CSIM model and helping me to explore that. Besides, I would never forget smart inspiration of Dr. Ali Haghighi and Abbas Zangeneh in application of Genetic Algorithms in my project. I wish to thank my examiner, Dr. Jerker Jarsjö, for showing me the right way to start my thesis. Of course, I owe sincere and thankfulness to Maria Damberg, student counselor at the department of Physical Geography and Quaternary Geology, for her unconditioned support and help.

Finally, I would like to express my special and deepest thanks to my parents, Homa and Javad, my first and best teachers ever; to Mehrashk, my brother, who persuade me to think like an engineer; to my beloved, Abbas, for being in my life; and with all my love thank to my country, Iran, for growing me in her beauty, history and patience.
Reference


Encyclopædia Britannica online, 2012: SWEDEN.
   http://www.britannica.com/EBchecked/topic/576478/Sweden, 06/01/2012


Swedish Environmental Protection Agency. 2000. Environmental Quality criteria, Groundwater. ARALIA.


World Maps of Köppen-Geiger climate classification, 2011: Present climate
http://koeppen-geiger.vu-wien.ac.at/present.htm, 18/10/2011.
Appendix A

Table A1. The Name and WSID of the Swedish watersheds

<table>
<thead>
<tr>
<th>Id</th>
<th>WSID</th>
<th>Name</th>
<th>Id</th>
<th>WSID</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>Rickleån</td>
<td>19</td>
<td>91</td>
<td>Helge å</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>Skellefte älv</td>
<td>20</td>
<td>93</td>
<td>Mörrumsån</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>Pite Ålv</td>
<td>21</td>
<td>95</td>
<td>Lyckebyån</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>Alterälven</td>
<td>22</td>
<td>96</td>
<td>Ljungbyån</td>
</tr>
<tr>
<td>5</td>
<td>6</td>
<td>Lule älv</td>
<td>23</td>
<td>97</td>
<td>Emån</td>
</tr>
<tr>
<td>6</td>
<td>8</td>
<td>Kalix älv</td>
<td>24</td>
<td>98</td>
<td>Botorpströmmen</td>
</tr>
<tr>
<td>7</td>
<td>10</td>
<td>Torne älv</td>
<td>25</td>
<td>99</td>
<td>Motala ström</td>
</tr>
<tr>
<td>8</td>
<td>24</td>
<td>Forsmarksån</td>
<td>26</td>
<td>100</td>
<td>Nyköpingsån</td>
</tr>
<tr>
<td>9</td>
<td>25</td>
<td>Dalälven</td>
<td>27</td>
<td>101</td>
<td>Norrström</td>
</tr>
<tr>
<td>10</td>
<td>26</td>
<td>Gavleån</td>
<td>28</td>
<td>142</td>
<td>Rönne å</td>
</tr>
<tr>
<td>11</td>
<td>27</td>
<td>Ljusnan</td>
<td>29</td>
<td>143</td>
<td>Lagan</td>
</tr>
<tr>
<td>12</td>
<td>28</td>
<td>Delängersån</td>
<td>30</td>
<td>145</td>
<td>Nissan</td>
</tr>
<tr>
<td>13</td>
<td>29</td>
<td>Ljungan</td>
<td>31</td>
<td>147</td>
<td>Åtran</td>
</tr>
<tr>
<td>14</td>
<td>31</td>
<td>Indalsälven</td>
<td>32</td>
<td>149</td>
<td>Viskan</td>
</tr>
<tr>
<td>15</td>
<td>33</td>
<td>Ångermanälven</td>
<td>33</td>
<td>151</td>
<td>Göta älv</td>
</tr>
<tr>
<td>16</td>
<td>35</td>
<td>Ume älv</td>
<td>34</td>
<td>341</td>
<td>Gideälven</td>
</tr>
<tr>
<td>17</td>
<td>71</td>
<td>Råneälven</td>
<td>35</td>
<td>342</td>
<td>Lödäälven</td>
</tr>
<tr>
<td>18</td>
<td>72</td>
<td>Töreälven</td>
<td>36</td>
<td>343</td>
<td>Öreälven</td>
</tr>
</tbody>
</table>

Figure A1. Location of the Swedish watersheds, the catchments which starts with 90** are coastal areas.
Table A2. Characteristics of three different scenarios used in the GA optimization process

<table>
<thead>
<tr>
<th>Scenario</th>
<th>generation</th>
<th>bits</th>
<th>Watershed</th>
<th>Scenario</th>
<th>generation</th>
<th>bits</th>
<th>Watershed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef_1</td>
<td>20</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
<td>Coef_1</td>
<td>20</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
<tr>
<td>Coef_2</td>
<td>20</td>
<td>(rec1 &amp; grcoef) 8</td>
<td>#33</td>
<td>Coef_2</td>
<td>20</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
<tr>
<td>Coef_3</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
<td>Coef_3</td>
<td>20</td>
<td>not included in process</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 12</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
<tr>
<td>Scenario</td>
<td>generation</td>
<td>bits</td>
<td>Watershed</td>
<td>Scenario</td>
<td>generation</td>
<td>bits</td>
<td>Watershed</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>---------------</td>
<td>-----------</td>
<td>----------</td>
<td>------------</td>
<td>---------------</td>
<td>-----------</td>
</tr>
<tr>
<td>Coef_1</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
<td>Coef_1</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
<tr>
<td>Coef_2</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td>#25</td>
<td>Coef_2</td>
<td>40</td>
<td>(rec1 &amp; grcoef) 8</td>
<td>#93</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
<tr>
<td>Coef_3</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
<td>Coef_3</td>
<td>30</td>
<td>(rec1 &amp; grcoef) 8</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 12</td>
<td></td>
<td></td>
<td></td>
<td>(rec2 &amp; seepcoef) 8</td>
<td></td>
</tr>
</tbody>
</table>
Figure A2. The GA optimization results for the watershed #33 by considering 3 scenarios.

Figure A3. The GA optimization results for the watershed #10 by considering 3 scenarios.
Figure A4. The GA optimization results for the watershed #25 by considering 3 scenarios

Figure A5. The GA optimization results for the watershed #93 by considering 3 scenarios