THE UNIVERSITY OF WESTERN ONTARIO
DEPARTMENT OF CIVIL AND ENVIRONMENTAL ENGINEERING

Water Resources Research Report

Selection of Calibration and Verification Data for the HEC-HMS Hydrologic Model
CFCAS Project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions

By:
Juraj Cunderlik
and
Slobodan P. Simonovic

Report No. 047
Date: January, 2004

ISSN: (print) 1913-3200; (online) 1913-3219;
ISBN: (print) 978-0-7714-2624-7; (online) 978-0-7714-2625-4;

Western

Selection of calibration and verification data for the HEC-HMS hydrologic model

CFCAS project:
Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions

Project Report II.

January 2004

Prepared by

Juraj M. Cunderlik

and

Slobodan P. Simonovic

CFCAS Project Team:

University of Western Ontario
Slobodan P. Simonovic
Gordon McBean
Juraj M. Cunderlik

University of Waterloo
Donald H. Burn
Linda Mortsch
Mohd Sharif

Upper Thames River Conservation Authority
Rick Goldt
Mark Helsten

Contents

List of Figures

List of Tables

I. INTRODUCTION

The main purpose of this report is to summarize the strategy applied for selecting the hydro-climatic data, which will be used for the calibration and verification of the US-ACE Hydrologic Engineering Center's Hydrologic Modeling System (HEC-HMS). The HEC-HMS model was chosen to be the most appropriate hydrologic modeling tool for achieving the goals set in the Canadian Foundation for Climatic and Atmospheric Sciences (CFCAS) funded project "Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions" ("project" hereafter), (*Cunderlik*, 2003).

Before selecting the calibration and verification data from available hydro-climatic records, it is necessary to become conversant with the streamflow and precipitation regimes in a given study area. Therefore, the first section of this report provides an insight into streamflow and precipitation regimes in the Upper Thames River basin (UTRb). The structure of the following text is then organized into two sections, one describing procedures used for selecting the hydro-climatic data for single-event hydrologic modeling, and the other one procedures for selecting the data for continuous hydrologic modeling. The last section summarizes the results, and formulates recommendations for subsequent project tasks. The terminology used in this report follows the terminology introduced and defined in *Cunderlik* (2003).

II. FLOW AND PRECIPITATION REGIMES IN THE UPPER THAMES RIVER BASIN

The knowledge of temporal streamflow and precipitation distributions, particularly the distributions of extremes, such as floods, droughts, and heavy thunderstorms, is crucial when selecting the data for the calibration and verification of hydrologic models. Therefore, this section provides a brief introduction to the regimes of hydro-climatic extremes in the Upper Thames River basin.

Seven streamflow and four precipitation gauges were selected for analyzing the seasonality of extremes in the study area. Figure 1 shows the location of the selected gauges together with a delineation of the Upper Thames River basin and its main subbasins. All selected gauges had to have continuous records during the 35-year long common observation period 1966-2000. The common period was chosen to be long enough to eliminate the effect of sampling variability, as well as to overlap most of the individual observation periods. The selected streamflow gauges represent small subbasins, with records not influenced by major dams or other constructions, which could negatively affect the results. Similarly to the streamflow gauges, the precipitation gauges were chosen to be spatially representative in terms of the precipitation regime in the UTRb. A list of the selected gauges is provided in Table 1.

Table 1. Number of POT events (NrPOT) in the selected streamflow (F) and precipitation (P) gauges.

F-ID	**Name**	**NrPOT**	**P-ID**	**Name**	**NrPOT**
02GD004	Middle Thames River at Thamesford	107	6142420	Foldens	109
02GD008	Medway River at London	107	6144475	London Airport	106
02GD011	Cedar Creek at Woodstock	112	6148105	Stratford MOE	105
02GD014	North Thames river near Mitchell	105	6149625	Woodstock	118
02GD018	Avon River below Stratford	112			
02GD019	Trout Creek	37			
02GD020	Waubuno Creek near Dorchester	117			

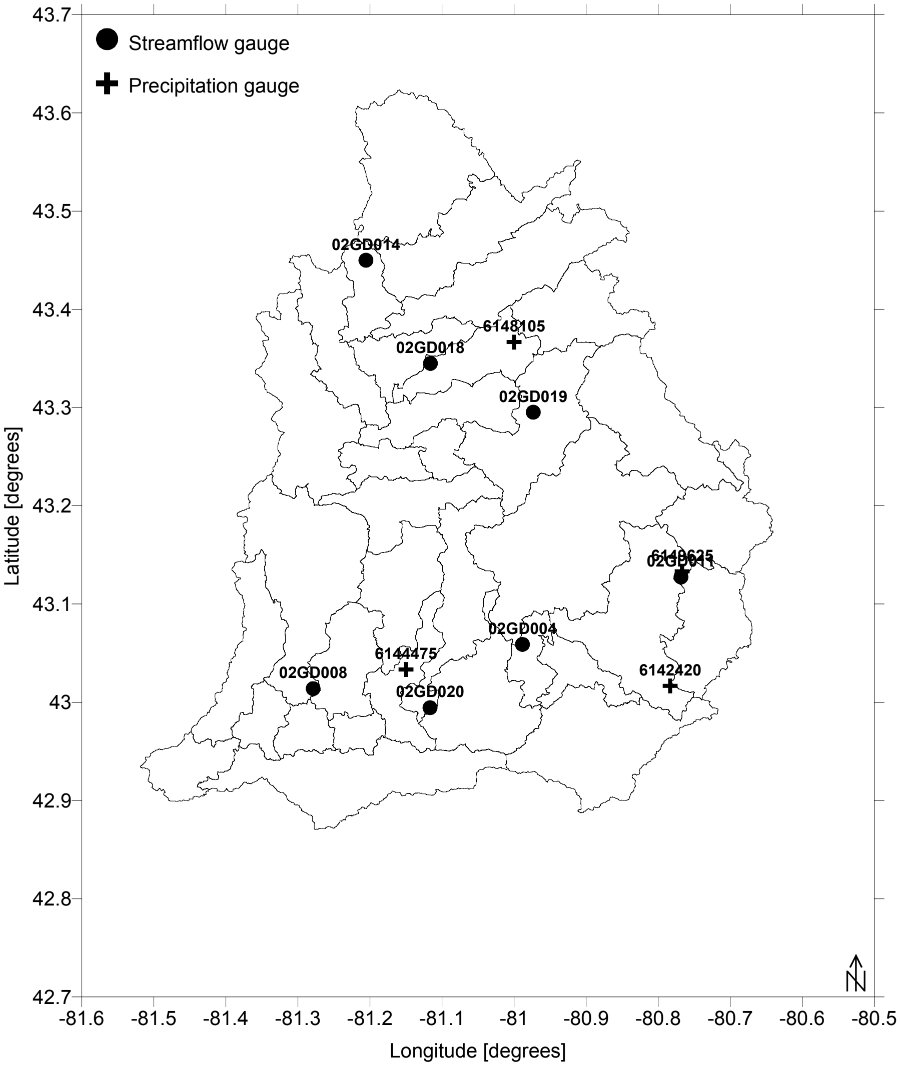

Figure 1. Location of the selected streamflow and precipitation gauges used for analyzing the seasonality of extremes in the Upper Thames River basin.

From the selected gauges, daily streamflow and precipitation records, corrected by Environment Canada (EC), were used for the analysis. The daily records were transformed into peaks-over-threshold (POT) records, containing on the average 3 maximum daily values (POT peaks) per year. The independence of POT peaks was assured by two criteria:

1. Two peaks had to be separated by at least three-times the average time to rise.
2. The minimum discharge in the trough between two peaks had to be less than two-thirds of the discharge of the first of the two peaks.

More information on POT data sampling and modeling can be found in *Lang et al.* (1999).

Table 1 lists the number of POT events for the selected streamflow and precipitation gauges from the common observation period 1966-2000. The average record length of streamflow POT records was 108 events (minimum 97 and maximum 117), and the average length of precipitation POT records was 109 (minimum 105 and maximum 118).

From the POT records, corrected relative frequencies of flood/precipitation occurrence, as well as the mean day of flood/precipitation and its variability of occurrence were estimated according to the guidelines given in *Cunderlik and Burn* (2002). Figure 2 shows the relative frequencies of POT flood occurrence for each of the seven selected streamflow gauges summarized in Table 1.

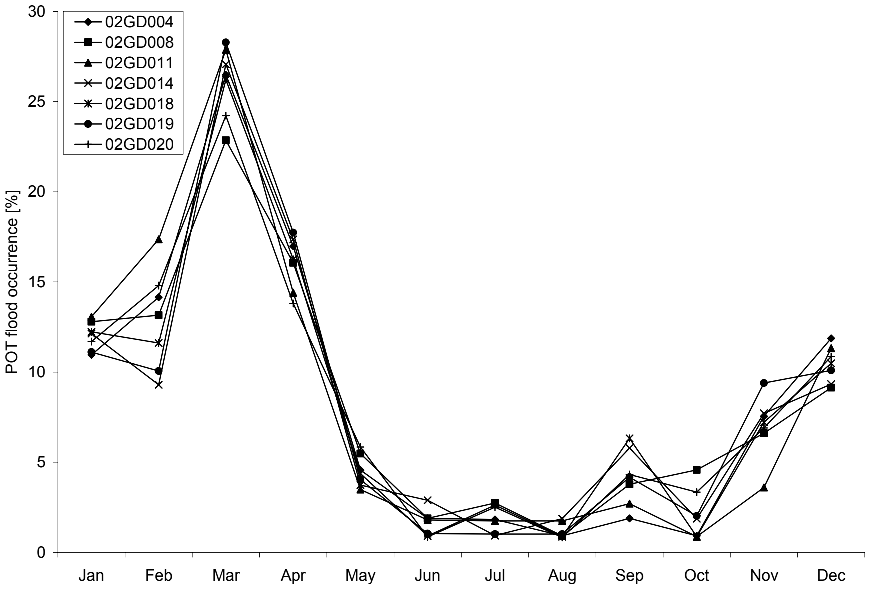

Figure 2. Relative frequencies of POT flood occurrence for the selected streamflow gauges in the Upper Thames River basin from the period 1966-2000.

Figure 2 indicates that the temporal distribution of POT floods from different parts of the basin has fairly uniform character. On the average, 25% of all floods in the UTRb occurred in

March, and more that 50% of all floods occurred in the period of February-April. Also, the months of November-January have higher probability of flood occurrence than the remaining summer months. In the period of November-April, flooding may arise from frontal precipitation, snowmelt, or from the combination of both. The period of March-April represents predominantly the period of catchment-wide, snowmelt-induced flood events. In November-December, floods are usually generated by frontal rainfall that falls on catchment saturated from previous rainfall events. Concurrent temperature records show that in December-April many floods result from sudden warming (snowmelt) accompanied by intensive frontal precipitation. Intensive summer storm floods are not so frequent and regular (and predictable) as snowmelt floods in the UTRb, but their peak discharges can exceed the peaks of snowmelt, or frontal rainfall floods. In fact, some of the highest floods in the analyzed streamflow records were generated from summer storms. Furthermore, a spatial analysis of several summer storm events showed that they can hit a substantial part of the UTRb.

Figure 3 depicts the relative frequencies of POT precipitation events for the four selected precipitation gauges summarized in Table 1. Similarly to flood distribution, the temporal distribution of precipitation across the basin has also quite uniform character. As Figure 3 suggests, heavy precipitation usually falls in the UTRb during the warm period of the year, the months of June-September account for more than 50% of all extreme precipitation events. An interesting feature in Figure 3 is a local decrease of the POT occurrence in August, apparent in all analyzed gauges. Also, there is a secondary POT probability increase in November resulting from enhanced frontal activity. The probability of extreme precipitation is lowest in December-February, and evenly increases towards the summer. The results presented in Figure 3 represent daily precipitation accumulations that are in good accordance with the runoff time concentrations found in the study area. Shorter time intervals will likely enhance the intensity of summer rainfall events, and such produce slightly different distributions as those showed in Figure 3.

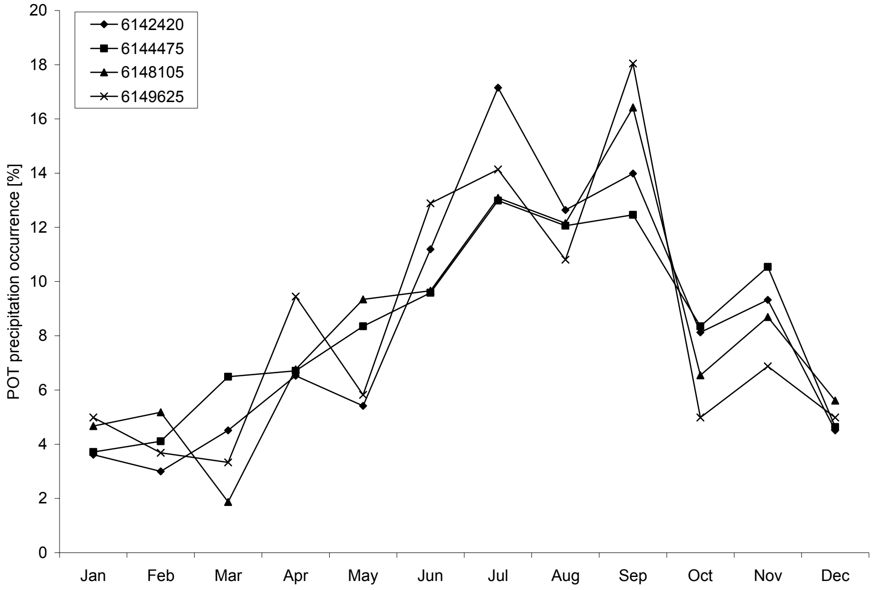

Figure 3. Relative frequencies of POT precipitation occurrence for the selected precipitation gauges in the Upper Thames River basin from the period 1966-2000.

Figure 4 shows the mean day of flood/precipitation plotted against the variability of flood/precipitation occurrence. The mean day of flood falls on the end of February, whereas the mean day of extreme daily precipitation falls on the end of August. Also, the average variability of the occurrence of extreme precipitation events is almost twice as high as the variability of flood occurrences. Relatively low variability of flood occurrences is caused by regular snow-induced flooding. The apparent lag between the mean occurrence of extreme precipitation and flood events confirms that not the precipitation, but the snowmelt is the main flood producing mechanism in the UTRb.

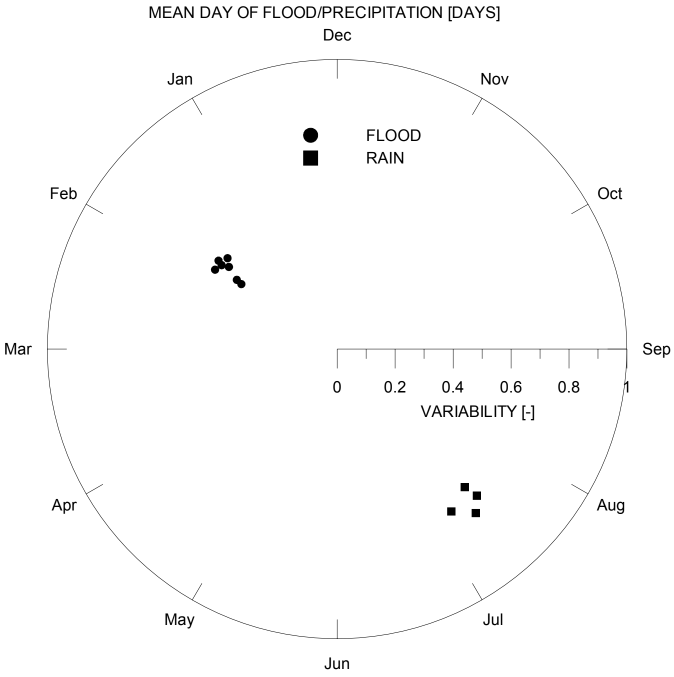

Figure 4. Mean day of flood/precipitation and the variability of flood/precipitation occurrence for the selected streamflow and precipitation gauges in the the Upper Thames River basin from the period 1966-2000 (Dec 31 = 360 degrees, 1 = maximum variability).

Droughts are the second type of hydrologic extremes important for the CFCAS project. Long-term mean monthly flows give a good picture of the temporal distribution of low flows. Standardized mean monthly flow values are depicted in Figure 5 for the selected streamflow gauges from the common observation period 1966-2000. As can be expected, the minimum flows occur in the UTRb at the end of summer, with the lowest values occurring in August. The August minimum flows may be connected with the local decrease in the POT precipitation occurrence, as was shown in Figure 3. The mean monthly flows in July and August are only 25-30% of the long term mean annual value. The pattern of intra-annual mean monthly streamflow distribution strongly resembles the pattern of the POT flood distribution showed in Figure 2. The rate of the mean monthly streamflow is highest in spring, in March up to 300% of the mean annual value.

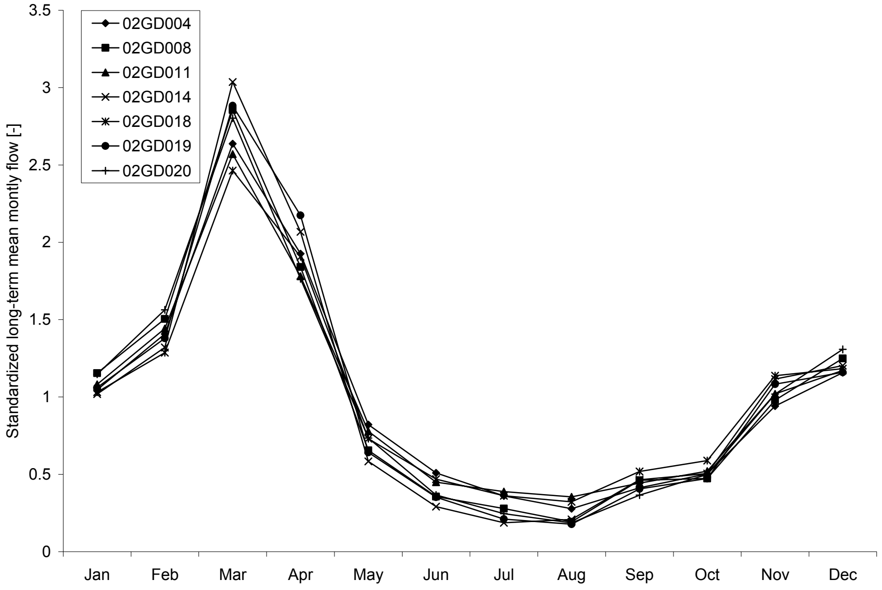

Figure 5. Standardized long-term mean monthly flows for the selected streamflow gauges in the Upper Thames River basin from the period 1966-2000.

In a summary, snowmelt is the main factor that produces flooding in the UTRb. Snowmelt-induced floods are most frequent in March. High flood events, with magnitudes sometimes exceeding the magnitudes of snowmelt-induced floods, are also generated by intensive summer storms. Although the size of the UTRb exceeds the physical limit of a single-cell storm, observations show that a storm can hit substantial portion of the basin. Nevertheless, the probability of catchment-wide (centered) storms is low. The frontal rainfall type of floods is frequent at the end of autumn. Severe flooding situations may arise when the frontal rainfall occurs in combination with snowmelt. This mixed flood events occur in December – April. Periods of low flow usually occur during the summer, and the risk of droughts is highest in the months of July and August. The above mentioned seasons represent the periods with the highest probability of the occurrence of critical hydrologic events. Therefore, the calibration and verification periods should cover these seasons, and the hydrologic modeling then, at the later stages of the project, used for simulating the identified critical events.

III. SINGLE-EVENT HYDROLOGIC MODELING

III.1 Selection criteria

The following criteria were applied for selecting individual rainfall-runoff events suitable for the calibration and verification of the HEC-HMS model:

- Maximum spatio-temporal data density of the observed hourly streamflow and rainfall records.
- Rainfall-runoff events generated by the same rainfall event.
- Streamflow peaks representing all runoff due to the selected rainfall event.
- Adequate spatial coverage of rainfall-runoff events, preferably covering the whole basin.
- The duration of rainfall events exceeding the time of concentration of the basin.
- The magnitude of rainfall events selected for calibration approximately equal the magnitude of rainfall events the model is intended to analyze.

Some of the above criteria were chosen according to the recommendations given in *US-ACE* (2000).

Spatio-temporal data density was examined by the distribution of hourly data during the 10-year period 1994-2003, calculated separately for all streamflow and precipitation gauges included in the HEC-HMS single-event basin model. There are 18 streamflow gauges and 16 precipitation gauges with hourly data in the UTRb that can be used for the calibration and verification of the HEC-HMS model. The location of these gauges is shown in Figure 6. Among them, several gauges were installed during the last ten years, and therefore the observation periods before 1994 were not analyzed. Table 2 provides a brief summary of the selected streamflow and precipitation gauges.

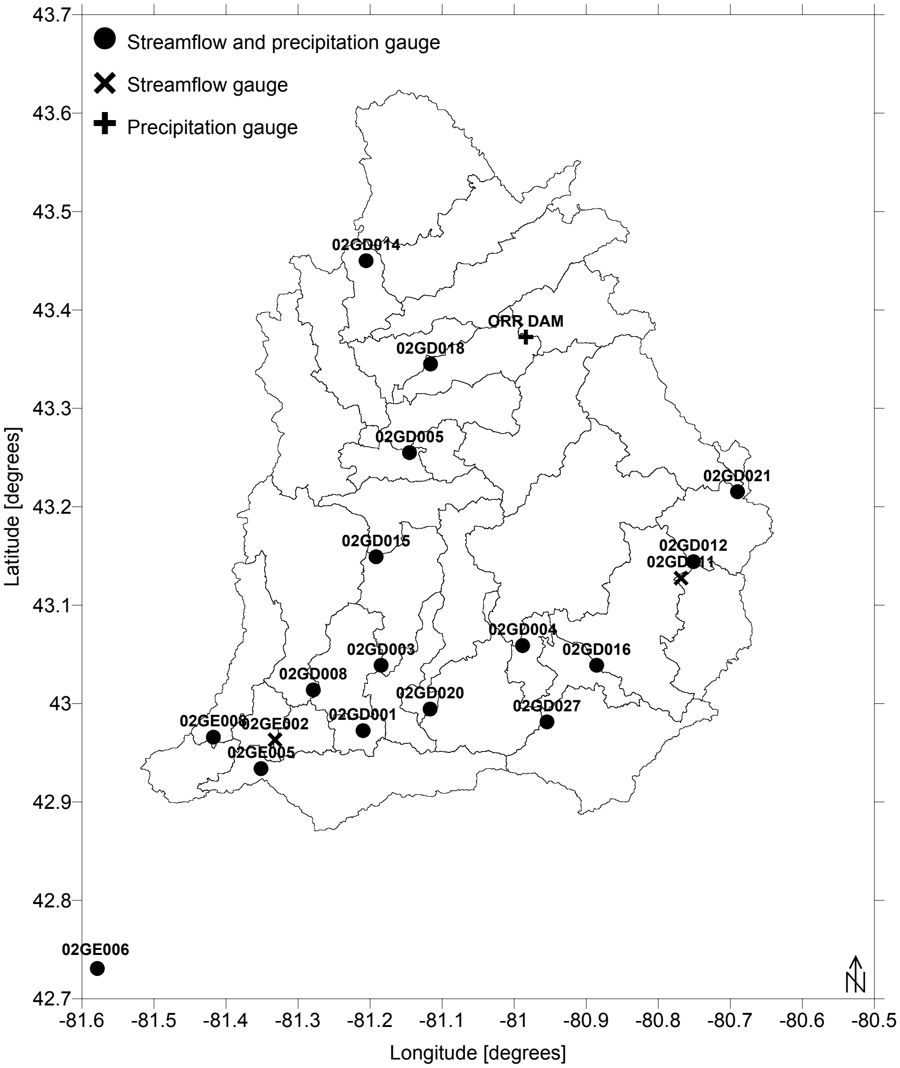

Figure 6. Location of the selected streamflow and precipitation gauges included in the HEC-HMS single-event basin model.

Table 2. Selected streamflow and precipitation gauges with hourly data records.

ID	**Name**	**Area [km^2]**	**Streamflow**	**Precipitation**
02GD001	Thames River near Ealing	1340	●	●
02GD003	North Thames River Below Fanshawe Dam	1450	●	●
02GD004	Middle Thames River at Thamesford	306	●	●
02GD005	North Thames River at St. Marys	1080	●	●
02GD008	Medway River at London	200	●	●
02GD011	Cedar Creek at Woodstock	93	●	
02GD012	Thames River at Woodstock	254	●	●
02GD014	North Thames river near Mitchell	319	●	●
02GD015	North Thames River near Thorndale	1340	●	●
02GD016	Thames River at Ingersoll	518	●	●
02GD018	Avon River below Stratford	144	●	●
02GD020	Waubuno Creek near Dorchester	108	●	●
02GD021	Thames River at Innerkip	149	●	●
02GD027	Reynolds	166	●	●
02GE002	Thames River at Byron	3110	●	
02GE005	Dingman Creek below Lambeth	146	●	●
02GE006	Thames river near Dutton	3760	●	●
02GE008	Oxbow Cr	89	●	●
---	Orr Dam*	91		●

* Precipitation gauge.

Figure 7 shows the spatio-temporal data density based on a monthly time step.

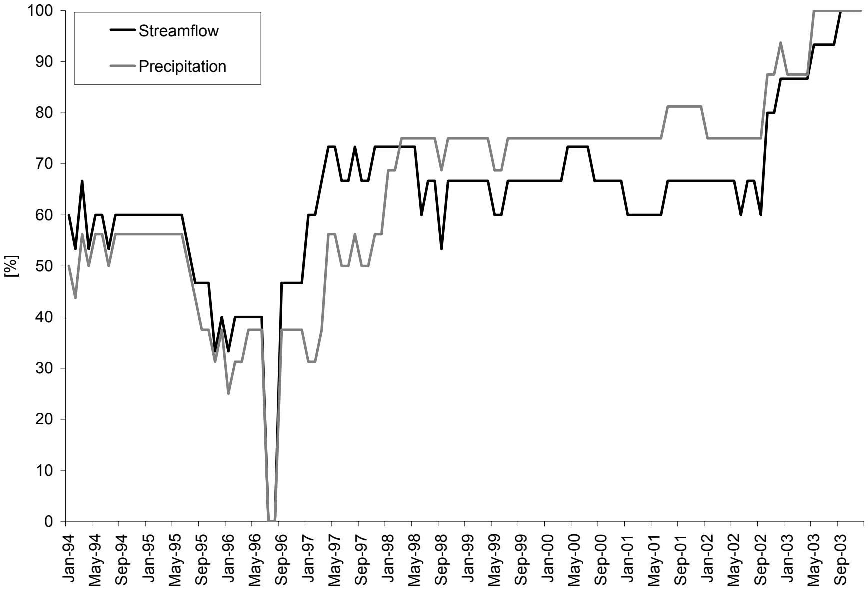

Figure 7. Spatio-temporal density of the hourly hydro-climatic data in the Upper Thames River basin (100% = all gauges included in the HEC-HMS model have hourly data in the given month).

The spatial data coverage before 1997 is rather poor, where only around 50% of all gauges included in the HEC-HMS model had hourly observations available. From 1997 the spatial data coverage improves to around 70%. In 2002 a couple of new gauges were installed in the UTRb (included in the HEC-HMS), and so the data density rises to 100%. Figure 7 suggests searching for suitable rainfall-runoff events preferably from the period 1997-2003.

III.2 Selected rainfall-runoff events

In the next step of the analysis, a number of rainfall-runoff events were identified according to the criteria given in the previous section. Only gauges with both streamflow and precipitation observations were used in this stage of the analysis. There were 15 such gauges included in the HEC-HMS model. The spatio-temporal density of these events is depicted in Figures 8 and 9.

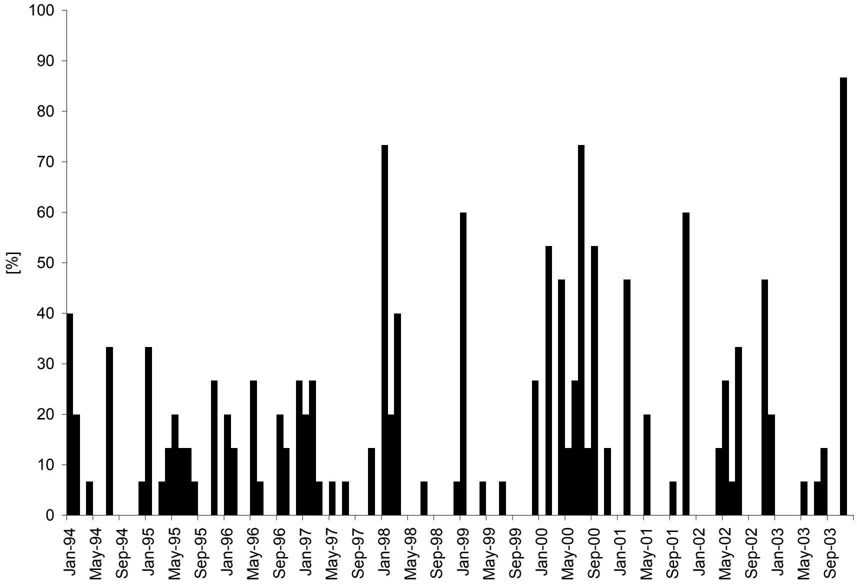

Figure 8. Spatio-temporal density of selected rainfall-runoff events (100% = all gauges included in the HEC-HMS model recorded the given rainfall-runoff event).

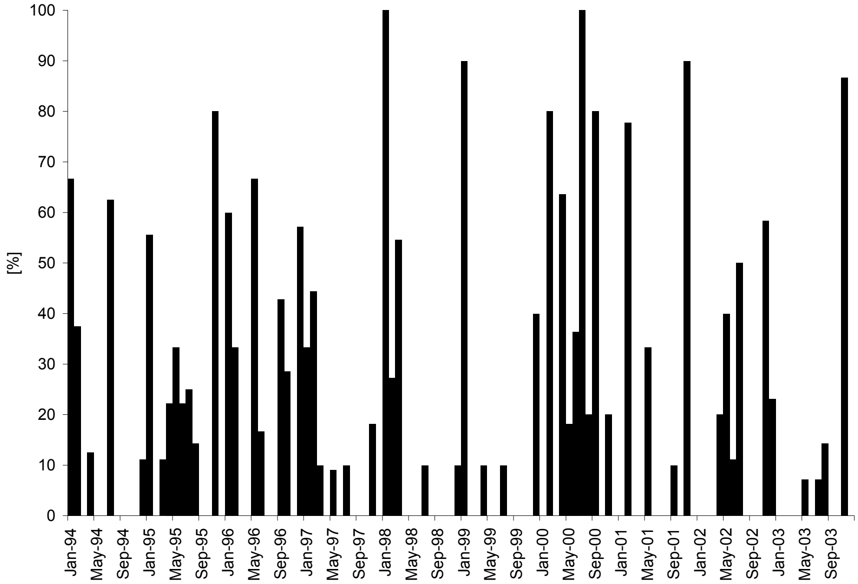

Figure 9. Spatio-temporal density of selected rainfall-runoff events (100% = all gauges with observations during the period of the given rainfall-runoff event, included in the HEC-HMS, recorded that event).

Figure 8 shows the ratio of the number of gauges, where a given rainfall-runoff event occurred, to the total number of gauges included in the HEC-HMS model; whereas Figure 9 shows the ratio of the number of gauges, where a given rainfall-runoff event occurred, to those gauges included in the HEC-HMS model, which had observations during the period of the given event. In other words, the latter figure represents the spatio-temporal data density relative to the data availability in a particular observation period.

Figure 8 indicates that the events, which occurred before 1997 have rather low spatial coverage, and that none of these events occurred in more than 50% of the gauges. Figure 9 then adds that among these events only one occurred in more than 70% of the gauges that had active observations in this period. Figure 8 further shows that during the whole analyzed period there are seven major flood events that occurred in more than 50% of the gauges: Jan 98, Jan 99, Feb 00, Jul 00, Sep 00, Nov 01, and Nov 03. Figure 9 shows again, that from these events only two – Jan 98 and Jul 00 occurred in all gauges included in the model. The remaining five events were captured by 80-90% of the gauges.

The flood of July 9-10 2000 is a well defined, storm-induced event that hit almost the entire basin. The spatial coverage of this event indicates that the storm cell had to be centered at the basin during the heavy outburst. Figure 10 depicts this event from the hourly observations at Thamesford@Thames (ID 02GD004). There was almost no rainfall activity before this event, and the streamflow was clearly at its minimum, baseflow level. The flood peak has a very simple, classical shape that makes this event excellent for the calibration. Furthermore, the magnitude of this event is noticeable, it is the maximum peak in the year 2000, and the 11-th largest annual maximum flood at Thamesford@Thames during the whole 63-year long observation period.

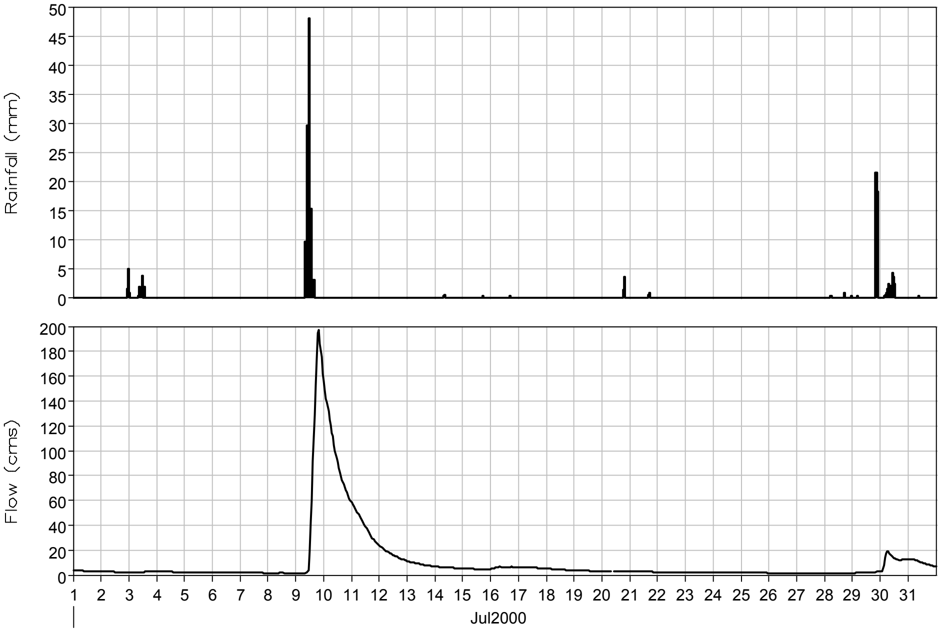

Figure 10. The storm and flood of July 9-10 2000 observed at Thamesford@Thames.

The Jan 98, Jan 99 and Feb 00 events are also well defined flood events, but the analysis of the corresponding temperature records revealed that they are likely to be a product of the combination of frontal rainfall fallen on an intensively melting snow cover (in Jan 98 the temperature rose to +10 degrees). The Sep 00 event lags behind the Jul 00 event with its magnitude and spatial extension. The remaining Nov 01 and 03 events have good spatial coverage, but represent more complex catchment responses on a series of rainfall events. In order to cover frontal rainfall driven flooding, the Nov 03 event, which has 100% spatial coverage in the UTRb, should be also used for the calibration. The Sep 00 and Nov 01 events represent both convective and frontal rainfall induced floods, and will serve as good verification events for the model.

IV. CONTINUOUS HYDROLOGIC MODELING

IV.1 Selection criteria

The main objective of the continuous hydrologic modeling in the CFCAS project is to simulate snow accumulation and melting processes, prolonged periods of low flows (droughts), and to address present and future long-term water budget/management requirements. There are three major criteria for selecting the data for continuous hydrologic modeling:

- Long sequences of concurrent daily streamflow, precipitation, and temperature data.
- Representative hydrologic variability (mean and extremes) in the selected sequences.
- Maximum spatio-temporal density of streamflow, precipitation, and temperature records.

Spatio-temporal data density was examined by the distribution of daily data during the period 1940-2002, calculated separately for all streamflow, precipitation and temperature gauges, included in the HEC-HMS continuous basin model. The time periods before 1940 have very low spatial density of gauges, and therefore were not analyzed. Also, the data for 2003 were not available at the time this report was being prepared. There are 18 streamflow, 10 precipitation, and 7 temperature gauges in the UTRb that can be used for the calibration and verification of the HEC-HMS continuous model. The location of these gauges is depicted in Figure 11. The 18 streamflow gauges are also listed in Table 2, and the corresponding EC climatic gauges in Table 3.

Table 3. Selected EC climatic gauges with daily data records.

EC-ID	Name	Precipitation	Temperature
6137362	St Thomas WPCP	●	●
6142066	Dorchester	●	
6149625	Woodstock	●	●
6148212	Tavistock	●	
6144475	London Airport	●	●
6148105	Stratford MOE	●	●
6143722	Ilderton Bear Creek	●	●
6142627	Fullarton	●	●
6142420	Foldens	●	●
6142295	Embro Innes	●	

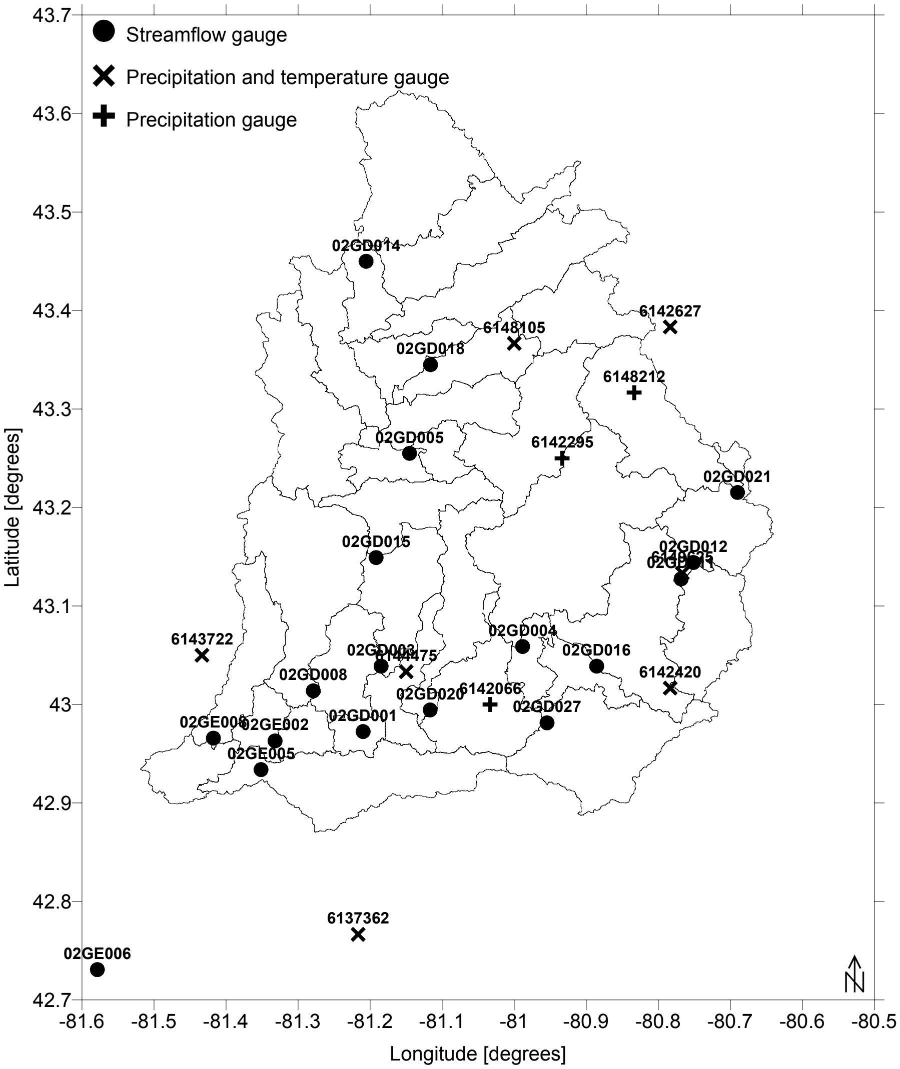

Figure 11. Location of the selected streamflow and precipitation gauges included in the HEC-HMS continuous basin model.

Figures 12 and 13 show individual and cumulative spatio-temporal data density based on a daily time step.

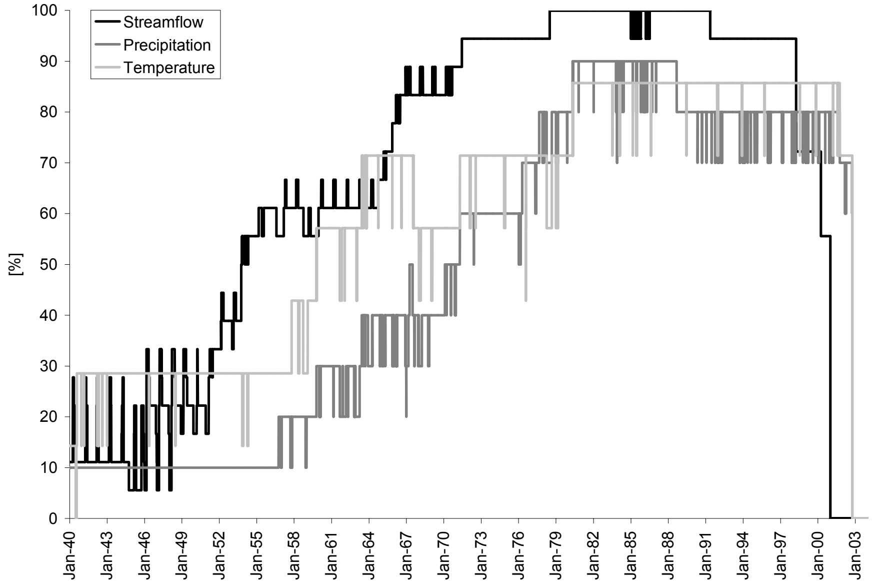

Figure 12. Spatio-temporal density of the daily hydro-climatic data in the Upper Thames River basin (100% = all gauges included in the HEC-HMS model recorded on the given day).

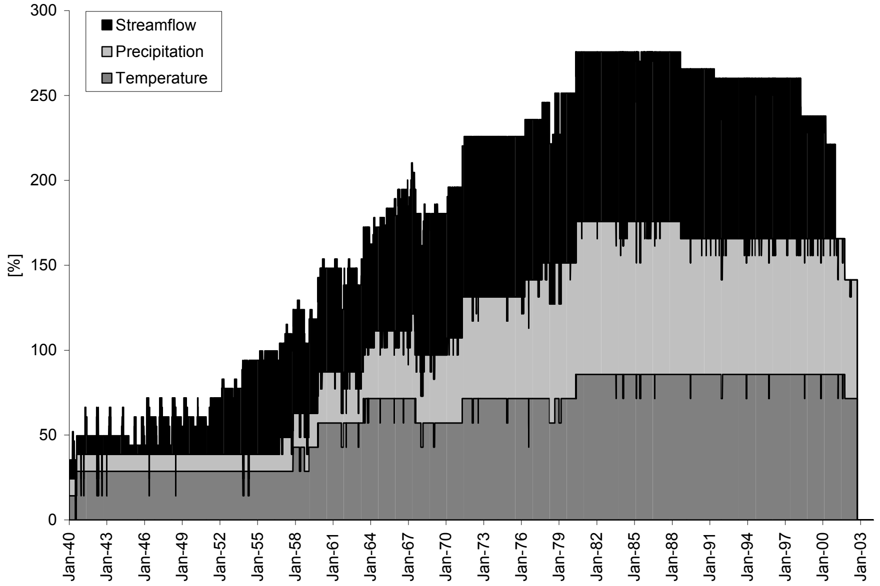

Figure 13. Cumulative spatio-temporal density of the daily hydro-climatic data in the Upper Thames River basin (300% = all hydro-climatic gauges included in the HEC-HMS model recorded on the given day).

Figures 12 and 13 show that before 1950, the spatial data coverage was very low, since only around 5-35% of all gauges included in the HEC-HMS model had daily observations available during this period. Streamflow data have high spatial density (95% and more) in the period of June 1971 – March 1998. The spatial density of precipitation records is high (70% and more) in the period of June 1977 – September 2002 (80% and more from November 1979 to August 1988). Finally, temperature records have high density (60-70%) from April 1971 to September 2002 (70-85% from March 1979 to September 2002). Both figures suggest using the period from March 1979 to March 1998 as a primary data source for the continuous hydrologic modeling.

IV.2 Selected observation periods

The high data-density period from March 1979 to March 1998 includes 18 complete hydrologic years that could be used for the model calibration and verification (hydrologic year was defined to start in November and to end in October). In order to address the hydrologic variability in the selected daily streamflow records, for each streamflow gauge, maximum and minimum values were found during this period. The temporal occurrence of these extreme values is depicted in Figure 14.

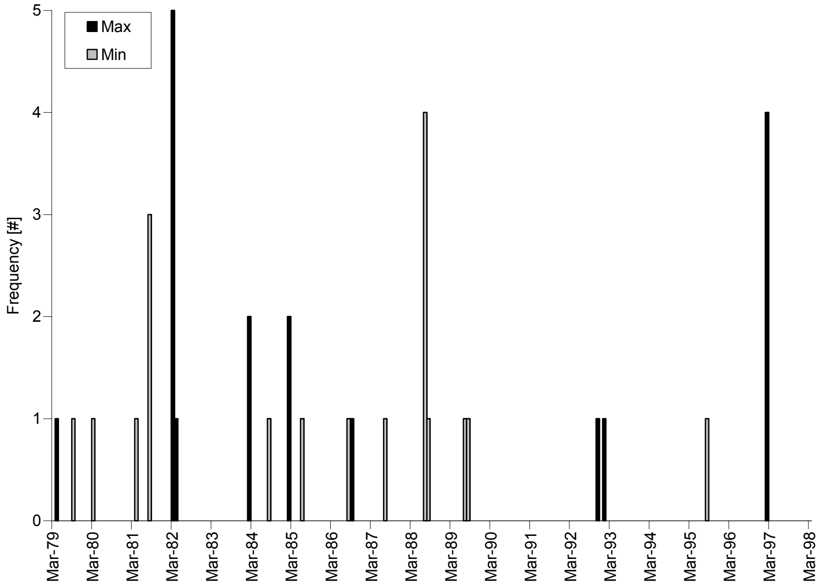

Figure 14. Frequency of maximum and minimum daily streamflow values from 18 streamflow gauges included in the HEC-HMS model for the high data density period March 1979 – March 1998.

Figure 14 clearly shows that the hydrologic variability is higher during the first half of the period March 1979 – March 1998. The period of November 1979 – October 1988 covers 10 complete hydrologic years, and includes minimum values from 85% of all gauges, and maximum values from 65% of all gauges. The length, spatio-temporal density, and hydrologic variability makes this period a convenient calibration sequence.

The remaining period from November 1988 to October 1997 includes 8 complete hydrologic years, and can be used for the verification of the model.

V. SUMMARY AND RECOMMENDATIONS

The main objective of this report was to provide the rationale for the selection of the most suitable hydro-climatic data for the calibration and verification of the HEC-HMS model. The analysis of the hydro-climatic extremes showed that there are three main flood-generating processes in the study area that produce four types of flooding. The most frequent type is a catchment-wide snowmelt-induced flooding, that occurs in the spring, with the highest probability in March. The second type of flooding is common in November-December, when floods are usually generated by frontal rainfall that falls on catchment surface saturated from previous rainfall events. The third type is generated by a combination of frontal rainfall with intensive snowmelt, which occurs in January-April. The last type represents intensive storm-induced floods that are most frequent in June-August. The analysis further showed that droughts have the highest probability of occurrence in the study area during the months of July and August.

A single-event hydrologic modeling should be used for simulating storm and frontal rainfall induced floods. Continuous modeling approach should be then employed for simulating snowmelt and mixed rainfall-snowmelt flooding, as well as for simulating the prolonged periods of summer low flows.

A set of criteria was defined for the selection of calibration and verification data for both single-event and continuous modeling approaches. According to these criteria, optimal hydro-climatic events and sequences were identified. The selected data cover all types of critical hydrologic events that occur in the study area. Table 4 lists the chosen events and data sequences.

Table 4. Selected events and data sequences for the calibration and verification of the single-event and continuous hydrologic models.

Hydrologic model	**Time step**	**Calibration**	**Verification**
Single-event	Hour	Jul 00, Nov 03	Sep 00, Nov 01
Continuous	Day	Nov 79-Oct 88	Nov 88-Oct 97

Two rainfall-runoff events were chosen for the calibration and verification of the single-event hydrologic model. They represent both convective and frontal rainfall driven floods. The calibration procedure should explore whether one set of model parameters can be used for simulating both types of events with acceptable accuracy.

Ten hydrologic years with high spatio-temporal density, and representative hydrologic variability were chosen for the calibration of the continuous hydrologic model, and eight complete years for the verification of the model. Data from these sequences will be used for simulating snow accumulation and melting processes, droughts, and long-term water balance studies.

For each modeling approach, a database in the HEC-DSS format was created, containing the selected data. These two databases will be used for hydrologic modeling in the later stages of the project.

VI. REFERENCES

[1] Cunderlik, J.M. (2003) Hydrologic model selection for the CFCAS project: Assessment of Water Resources Risk and Vulnerability to Changing Climatic Conditions. CFCAS Project Report I., 38 p.

[2] Cunderlik, J.M., Burn, D.H. (2003) Analysis of the linkage between rain and flood regime and its application to regional flood frequency estimation. *Journal of Hydrology*, 261, 115-131.

[3] Lang M., Ouarda T., Bobée B., 1999. Towards operational guidelines for over-threshold modeling. *Journal of Hydrology*, 225, 103-117.

[4] US-ACE (2000). *Hydrologic Modeling System HEC-HMS*. Technical Reference Manual. US Army Corps of Engineers, Hydrologic Engineering Center, 149 p.

VII. ABBREVIATIONS

CFCAS	Canadian Foundation for Climatic and Atmospheric Sciences
EC	Environment Canada
HEC	Hydrologic Engineering Center
HMS	Hydrologic Modeling System
US-ACE	United States Army Corps of Engineers
UTRb	Upper Thames River basin