Operational Real-time Flood Forecasting under Climate Change Impacts:  
The COFFEE Model for Coastal Storm Dominated Watersheds in British Columbia  
Charles Luo, Ph.D., P.Eng.  
BC River Forecast Centre  
E-mail: Charles.Luo@gov.bc.ca  

Abstract  

From the perspective of climatology and hydrology, climate change for British Columbia implies warmer average annual temperatures, increasing annual precipitation, the earlier onset of freshet, and more frequent heavy rainfall events. These changes have fundamentally impacted operational real-time flood forecasting in British Columbia. In order to meet the challenge of climate change impacts, highly time-efficient real-time flood forecasting models with relatively high confidence of accuracy become necessary. The Coastal Fall Flood Ensemble Estimation (COFFEE) Model is such a model for real-time flood forecasting for the coastal storm dominated watersheds in British Columbia during the fall-winter season. 

The COFFEE Model is basically a unit hydrograph model forecasting instantaneous peaks at a daily time step. Coastal storm driven floods have significantly high levels of uncertainty induced by errors included in the observed and forecast climate data and model approximation. In order to account for this uncertainty without expanding computer power, the concept of “ensemble forecasting” is borrowed into the model. In this study, “ensemble” does not mean that the model generates a set of forecasts but rather produces an average and the maximum and minimum by using the historical statistics. The historical rainfall events are used to construct the unit hydrograph for each of the watersheds. 

The model can be run at any time of the year when a storm system is building up over the coastal regions of British Columbia. One run of the model can finish within a minute on a personal computer and producing a five-day forecast of instantaneous peaks for all of the 94 watersheds takes about 10 minutes.  

Keywords: Climate change, real-time flood forecasting, rainfall dominated watershed.  

1. Introduction  

According to the Climate Change 2014: Synthesis Report by Intergovernmental Panel on Climate Change (IPCC, 2014), since 1950s, many of the observed changes to the global climate system over the past decades are unprecedented. These observed changes include that the atmosphere and ocean have warmed, the amounts of snow and ice have diminished, the sea level has risen, the number of extreme precipitation events have increased in a number of regions, and the number of cold days and nights has
decreased and the number of warm days and nights has increased on the global scale. In respect of extreme precipitation events, there are likely more land regions where the number of heavy precipitation events has increased than where it has decreased. The increasing trend in extreme precipitation, which usually incurs extreme discharges in some watersheds, implies greater risks of flooding at a regional scale. Meanwhile, impacts of the climate-related extremes, such as heat waves, droughts, floods, cyclones and wildfires, may endanger some ecosystems and many human communities (IPCC, 2014).

British Columbia (BC), which is located on the west coast of the Pacific Ocean, is impacted by climate changes as a part of the global climate system. In order to address the climate changes on a local scale for BC, an analysis of historical data conducted by BC Ministry of Environment (2016) indicates that many properties of climate have changed during the 20th and early 21st centuries, affecting the marine, freshwater, and terrestrial ecosystems in BC. These include (a) the average annual temperature has warmed by 1.4°C per century across the province; (b) the northern regions of BC has warmed more than the provincial average; (c) night-time temperatures have increased across BC in all seasons; (d) night-time minimum average temperature in winter has increased by 3.1°C per century; (e) annual precipitation has been increasing across the province overall; (f) lakes and rivers become free of ice earlier in the spring; (g) the bulk of river flow occurs earlier in the year; (h) the average sea level has risen along most of the BC coast; (i) sea surface temperatures have increased along the BC coast; (j) summertime temperatures in the Fraser River are warmer; and (k) more heat energy is available for plant and insect growth. In addition to the above findings, a different study by Groisman et al. (1999) found that the probability of daily precipitation exceeding 25.4 mm (1 inch) in northern countries including Canada has increased by about 20%, which is nearly four times the increase of the mean. Summarily, climate change for BC from the perspective of climatology and hydrology implies warmer average annual temperatures, increasing annual precipitation, the earlier onset of freshet, and more frequent heavy rainfall events.

These changes, especially the early onset of freshet and more frequent heavy rainfall events, have fundamentally impacted operational real-time flood forecasting in BC. This requires that operational real-time flood forecasting models are able to run all year round, encompassing not only the freshet season but also the periods before the traditional freshet and the fall-winter season. Such models were traditionally unavailable for BC. In order to meet modeling capability requirements for more frequent heavy rainfall events, highly time-efficient models with relatively high confidence of accuracy become necessary. Moreover, these models must also be able to accommodate themselves to the Canadian and/or BC data system, and the current stuffing and computing resources. As far as the data system is concerned, most climate stations in BC provide only daily maximum and minimum temperatures and daily precipitation. And the current stuffing and computing resources do not favor models that are labour and computationally intensive. Thus, the Coastal Fall Flood Ensemble Estimation (COFFEE) Model
was developed to meet the challenges posed by climate change impacts for operational real-time flood forecasting for the coastal storm dominated watersheds in BC during the fall-winter season. The study areas and climate data uncertainty is described in the next section, and then the methodology, model application, and discussion are detailed in two successive sections.

2. Study areas and climate data uncertainty

2.1 Study areas – coastal BC watersheds

BC is situated on the Pacific coast of Canada between latitudes 49° and 60° north and has a total land area of about 947,900 km². Details of BC’s geomorphologic and climatologic characteristics can be found in the Technical Reference for the CLEVER Model (Luo, 2015). Coastal BC is influenced greatly by the Pacific Ocean and prevailing westerly winds and is therefore mild and wet in the winter and warm and dry in the summer. As the consequence of the geographic location, most watersheds located in coastal BC are dominated by the fall-winter seasonal storms.

Hydrometrically, coastal BC is comprised of the following basins: the Northwest, Stikine, Skeena-Nass, Haida Gwaii (Islands), Central Coast, South Coast, Lower Fraser and Vancouver Island (Figure 1). Among the above hydrometric basins, the Stikine, Skeena and Nass are three of the seven major watersheds in BC. In the current stage, the COFFEE Model covers the majority of these basins comprising a total of 94 watersheds with a total area of about 225,000 km². These watersheds are critical to the local communities. Water Survey of Canada (WSC) hydrometric stations are available for all these 94 watersheds, for which climate stations are either located in or close to the watersheds. This study used Environment and Climate Change Canada climate stations and automated snow weather stations and fire weather stations managed by the Province of British Columbia. Figure 2 (a) and (b) shows the study watersheds modeled by the current COFFEE Model.

2.2 Climate data uncertainty

Forecasting of BC coastal storm driven floods has significantly high levels of uncertainty due to the uncertainty of the climate factors, especially the uncertainty in rainfall amounts and spatial and/or temporal distributions of rainfall. This uncertainty becomes a great challenge faced by real-time flood forecasting modellers. Significant model errors could be induced by the climate data errors because of the uncertainty. These climate data errors can be classified into two categories: 1) from the observation and/or forecast data itself, and 2) by spatial interpolation and/or extrapolation when the climate stations are not located in the watersheds. In this study, the second category of data errors is prominent because that the number of climate stations is limited and that many watersheds are so small that no climate stations are located in or even close to them. The Carnation Creek watershed on south Vancouver Island is a typical example of the problem of climate data interpolation.
Figure 1. Map of coastal BC (color filled areas)
Figure 2. COFFEE Model watersheds – color-filled patches with four digit watershed IDs (a) North
Figure 2. COFFEE Model watersheds – color-filled patches with four digit watershed IDs (b) Central and South

Figure 3 shows the geographic locations of the Carnation Creek watershed (CRNT), the adjacent Sarita River watershed (SART) and related climate stations. The area of the Carnation Creek watershed is only 11 km². The closest climate station is a fire weather station, Summit (SMM), which is 22 km away, and the second closest climate station is also a fire weather station, TS Effingham (TEF), which is 38 km away.

On January 28, 2018, both stations SMM and TEF recorded an incredible amount of 24-hour precipitation, 168 and 214 mm, respectively. However, WSC hydrometric station Carnation Creek at the Mouth (08HB048) recorded a negligible response to this event. Figure 4 (a) shows 24-hour precipitation recorded at the two climate stations and the instantaneous hydrograph at an interval of 5 minutes recorded at the Carnation Creek at the Mouth (08HB048) from January 26 to 30, 2018. Figure 4 (b) shows the plot of readings at the WSC station located in the adjacent watershed, Sarita River near Bamfield (08HB014), for comparison.
The 1.01-year return period discharge is plotted in Figure 4 (a) and both 1.01- and 2-year return period discharges are plotted in Figure 4 (b) so that the hydrometric responses can be seen more clearly. Comparing the two responses in Figure 4 (a) and (b), the response at the Carnation Creek at the Mouth (08HB048) is obviously negligible. This phenomenon demonstrated that the January 28, 2018 rainfall event did not really occur over the small Carnation Creek watershed even though both of the two closest climate stations recorded an incredible amount of rainfall. On the other hand, a completely opposite situation may also be possible for a small watershed such as the Carnation Creek watershed; that is that heavy rainfall may be pouring down onto the small watershed while no adjacent climate stations record any rainfall. Either case induces significant uncertainties and challenges to the flood forecasting in the coastal storm dominated watersheds in BC.
Figure 4. Hydrometric responses of two WSC stations to January 28, 2018 rainfall event (hydrographs from January 26 to 30, 2018)
3. Methodology

3.1 Developing discrete unit hydrograph for each watershed

The unit hydrograph is the unit pulse response function of a linear hydrologic system. The unit hydrograph of a watershed is a direct runoff hydrograph resulted from a unit of excess rainfall (e.g. 1 mm) which is generated uniformly over the drainage area at a constant rate for an effective duration, such as 24 hours. The unit hydrograph is a simple linear model that can be used to produce the hydrograph that results from any amount of excess rainfall input (Chow et al., 1988). Given a series of excess rainfall input (pulses) $P_m$ and the unit hydrograph $U_{n-m+1}$ at time $n$, the discrete response equation of runoff at time step $n$ can be written as:

$$Q_n = \sum_{m=1}^{n=M} P_m U_{n-m+1}$$  \hfill (1)

Equations (1) can be rewritten into a matrix equation as:

$$[P][U] = [Q]$$  \hfill (2)

Equation (2) can be solved by linear regression method:

$$[U] = [Z]^{-1}[P]^T [Q]$$  \hfill (3)


In this study, the time step is one day or 24 hours, and thus the excess rainfall $P_m$ is the total volume of rainfall in a period of 24 hours or a day (one time step), $Q_n$ is the daily average discharge, and the unit response $U_{n-m+1}$ is a dimensionless variable.

3.2 Generating forecast of daily average discharge

The net water input at each time step, 24 hour or a day in this study, to each watershed is calculated by the following water balance equation:

$$W = R + M + G - E - I$$  \hfill (4)

in which $W \geq 0$ is the net water input or excess rainfall to the watershed and has the unit of mm/day, and this unit is used for all the terms on the right-hand side of the equation, $R$ is the rainfall, $M$ is the snowmelt, $G$ is the groundwater seepage which is the base flow in this study, $E$ is the evapotranspiration, and $I$ is the infiltration to the unsaturated soil and recharge to the groundwater. In this study, snowmelt $M$ is calculated using the temperature index method, where $G$, $E$ and $I$ are assumed to be constants subject to calibration. The relationship between $P_m$ and $W$ is given by:

$$P_m = W A / \Delta t$$  \hfill (5)

where $A$ is the watershed area and $\Delta t$ is the time step which is 24 hours or a day in this study.
The daily average discharge can be estimated by substituting Equation (5) into Equation (1). However, forecasting daily average discharges is not the objective of this study. The objective of developing the COFFEE Model is to produce forecasts of instantaneous peaks for the coastal-storm dominated watersheds in BC. Therefore, an efficient method must be developed to convert the daily discharges into the instantaneous flows. This method must also be able to tackle the issues of climate data uncertainty discussed in subsection 2.2.

3.3 Ensemble-analogue forecasts of instantaneous peaks

In order to account for the uncertainty in forecasting coastal storm-driven floods in BC, especially the uncertainty induced by the climate data errors, the concept of “ensemble forecasting” is borrowed into the COFFEE Model in this study.

Ensemble forecasting was first used in numerical weather prediction (NWP) and is known as ensemble prediction systems (EPS). Instead of making a single deterministic forecast of the most likely weather, a set (or ensemble) of forecasts is produced to give a range of possible future states represent the total uncertainty in modeling. Recently, operational and research flood forecasting systems around the world are increasingly moving towards ensemble forecasting using EPS as data inputs to drive their flood forecasting systems to produce river discharge predictions (Cloke and Pappenberger, 2009).

Besides using ensemble (multiple) NWP/EPS data inputs to the hydrologic models, ensemble forecasts can also result from multiple models with multiple parameter sets and from probabilistic predictions. However, ensemble forecasting is also faced with a number of challenges, such as difficulties in understanding the full range of uncertainties and interactions between uncertainties in the forecast systems, requiring great efforts in hydrologic data assimilation, and demanding huge computer power (Cloke and Pappenberger, 2009).

Without sacrificing time efficiency and without expanding computer power in this study, “ensemble” does not really mean that the COFFEE Model generates a set of forecasts, either by using NWP/EPS data as model input, or using multiple parameter sets for the model, or by carrying out probabilistic predictions; instead, the model produces an average, maximum (upper bound) and minimum (lower bound) flows by using historical statistics. By doing so, “ensemble” really means “ensemble-analogue” in this study. WSC’s historical statistics of hydrometric data include annual instantaneous peaks and annual peaks of daily average discharges. Ratios ($R_n$) of the instantaneous peaks ($O_n$) to the daily average peaks ($D_n$) for year $n$ can be found by:

$$R_n = \frac{O_n}{D_n}$$

(6)

And
\[
\begin{align*}
R_{ave} &= \frac{\sum_{n=1}^{N} R_n}{N} \\
R_{max} &= \max(R_n) \\
R_{min} &= \min(R_n)
\end{align*}
\]

(7)

in which \(R_{ave}\) is the average, \(R_{max}\) is the maximum and \(R_{min}\) is the minimum in a total of \(N\) years of data, in which historical statistics of annual instantaneous peaks and annual peaks of daily discharges are both available.

The ensemble-analogue forecasts of instantaneous peaks are obtained by timing the relevant ratios to the forecast daily discharges:

\[
\begin{align*}
Q_{ave,n} &= R_{ave} Q_n \\
Q_{max,n} &= R_{max} Q_n \\
Q_{min,n} &= R_{min} Q_n
\end{align*}
\]

(8)

in which \(Q_{ave,n}, Q_{max,n}, Q_{min,n}\) are the average, maximum and minimum instantaneous peaks at time step \(n\), respectively.

Some WSC hydrometric stations have only water level data. In this case, a rating cure is used to convert the forecast discharges into water levels.

3.4 Adding up upstream inflows

There is no open channel routing in the COFFEE Model. However, a larger watershed is always split into a number of smaller sub-basins when the watershed is too large to simplify into one node. For example, the Skeena River watershed (area = 54,432 km\(^2\)) is split into 13 sub-basins. The total outflow at the watershed outlet is simply calculated simply by finding the sum of the staggered hydrographs of the outflows from all the upstream sub-basins. A constant time lag (\(\Delta t_j\)) is first determined by the distance from an upstream sub-basin to the watershed outlet. Then, a staggered hydrograph can be obtained by shifting the data series by \(\Delta t_j\) time steps ahead. The total outflow at the watershed outlet (\(Q_{total,n}\)) at time step \(n\) is given by:

\[
Q_{total,n} = \sum_{j=1}^{J} Q_{n-\Delta t_j}
\]

(9)

in which \(Q_{n-\Delta t_j}\) is the staggered outflow from the \(j\)-th upstream sub-basin.
4. Model application and discussion

4.1 Model establishment

In this study, the 94 coastal storm dominated watersheds (as shown in Figure 2) are incorporated into a single model. For each of the 94 watersheds, the following two steps must be finished before the model can run: (i) constructing the unit hydrograph using historical rainfall events and Equation (3), and (ii) carrying out statistical analysis to estimate the ratios in Equations (6) and (7). For step (i), a series of historical rainfall events may exist and one of the events can be used to generate a unit hydrograph. The final hydrograph for a watershed is the arithmetic average of all the unit hydrographs. There may not be enough historical records available for some of the watersheds to complete step (ii). In this case, a set of initial ratios are given and these ratios are then subject to calibration.

4.2 Model calibration

Parameters subjected to calibration include: the factor applied to precipitation ($F_p$), snowmelt rate ($M$), the base flow ($G$), evapotranspiration rates ($E$), and infiltration rate ($I$). The ratios in Equation (7) ($R_{ave}$, $R_{max}$ and $R_{min}$) are also calibrated for the reasons discussed in the above subsection, and so that the forecast discharge is able to accommodate as many historical peaks as possible. The model calibration includes two stages, advance calibration and operational calibration. The advance calibration is carried out before the model is put into operation for real-time forecasting, and the operational calibration is performed when the model is running for real-time forecasting.

Statistical methods are not employed for the calibration due to the high levels of uncertainty in the coastal storm flood forecasting in this study. Basically, the model calibration was performed visually in this study by comparing plots of the observed discharges to the estimated discharges. The advance calibration includes comparing the observed to the estimated daily average discharges and comparing the observed instantaneous discharges to the forecast instantaneous peaks for the average, maximum (upper bound), and minimum (lower bound). The major criterion for the advance calibration is that the forecast upper and lower bounds accommodate as many observed instantaneous peaks as possible. The operational calibration only involves comparing the observed and forecast daily average discharges for the first ten days of the total time steps that the model runs for.

Figure 5 shows an example of model calibration for the Kitimat River below Hirsch Creek (08FF001) for October 24-25, 2017 rainfall event. Figure 5 (a) is the comparison for the daily average discharges and Figure 5 (b) is that for the instantaneous peaks. It can be seen from Figure 5 that the model underestimates the larger peak in the calibration of daily average discharges but the observed instantaneous peak falls within the estimated upper bound and average in the calibration of instantaneous peaks.
(a) Calibration of daily average discharges

(b) Calibration of instantaneous peaks

Figure 5. Model advance calibration for Kitimat River below Hirsch Creek (08FF001) for October 24-25, 2017 rainfall event
4.3 Model forecast

The COFFEE Model runs for a total time steps of 15 days, with 10 days of observation and 5 days of forecast for a total of 94 watersheds and sub-basins. The estimated hydrograph starts from the initial flow or the base flow, which is the smallest flow over the 10 days of observation. The forecast discharge is the estimated discharge for the last 5 days without any adjustment. One run of the model on a desktop person computer with a CPU clock of 3.20GHz can finish in 30 seconds to one minute, depending on the number of rainfall events and rainfall intensities. It takes less than 10 minutes to produce a set of output of five-day forecast, including a Google map with color coded stations, tables of numerical forecasts, and a detailed chart of forecast for each of the 94 stations.

The US National Oceanic and Atmospheric Administration (NOAA) Advanced Hydrologic Prediction Service (AHPS) is a robust and efficient system which provides hydrologic forecasts through 13 River Forecast Centers covering most watersheds around the US (McEnery et al., 2005). The Columbia River watershed, which has a watershed area of 668,000 km² which is triple that covered by the COFFEE Model, is one of the huge watersheds that covered by the AHPS. Checking the public website of North West River Forecast Center, one can see that it takes the AHPS one to two hours to update a 10-day hydrometric forecast for the Columbia River watershed. It is clear that the COFFEE Model is also very time-efficient comparing to the AHPS.

Figure 6 is an example of the Google map output from the COFFEE Model for October 22, 2017. The color codes reflect the return periods of the forecast maximum in the next five days. In this example, a heavy rainfall event was moving over BC central coast and the forecast maximum peaks over the next five days are floods between the 50 to 100-year return periods.

The numerical forecasts are also compiled in separate tables for all the areas of coastal BC, listing all the stations modeled in the COFFEE Model and the values of forecast average, maximum and minimum instantaneous peaks.

Figure 7 is an example of the detailed forecast chart for the Kitimat River below Hirsch Creek (08FF001) from the COFFEE Model for October 22, 2017. This detailed chart includes a bar chart of the observed and predicted rainfall intensity over a 15-day period, a hydrograph including 10 days of observation and 5 days of forecast average, maximum and minimum peaks, the numerical forecast over the next five days for this station, and other information for user references.
Figure 6. Google map output from the model on October 22, 2017.
4.4 Discussion - forecast errors

The accuracy of model forecasts was not evaluated using statistical methods due to the high level of uncertainties and the complexity induced by the ensemble-analogue forecasting. However, the forecast errors are straightforward and easily seen by comparing the forecast to the observed instantaneous peaks when they become available.

Forecast errors in the COFFEE Model mainly stem from: (i) errors in the forecast climate data especially the forecast rainfall data, (ii) errors from the spatial interpolation and/or extrapolation of climate data, and (iii) model errors including errors from model approximation and errors in the unit hydropaths. The forecast errors, especially those from the third source – model errors, may well subside when the ensemble-analogue forecasting is adopted and if the observed peaks falling within the range between the forecast upper and lower bounds is considered accurate (such as the example shown in Figure 7).

*Figure 7. A detailed chart for Kitimat River below Hirsch Creek (08FF001) output from the model forecast for October 22, 2017*
However, in some extreme cases, the forecast errors from the first and second sources – errors from the forecast climate data, either coming with the forecast data itself or induced by spatial interpolation and/or extrapolation, may not subside even if the ensemble-analogue forecasting is carried out. Figure 8 shows an example of this case. The forecast 24-hour rainfall for the Roberts Creek watershed (34 km²) was 100, 117 and 20 mm for January 28, 29 and 30, 2018, respectively. Therefore, the forecast maximum peak (upper bound) at station Roberts Creek at Roberts Creek (08GA047) for these three days was greater than the 200-year return period flood (Figure 8 (a)). However, the actual rainfall for these three days was only 50, 30 and 2 mm, respectively, and the observed peak was only 17.5 m³/s (Figure 8 (b)), which was well below the forecast lower bound (48.9 m³/s) shown in Figure 8 (a).

Figure 9 shows the model “forecast” for the Roberts Creek watershed by using the observed rainfall as the model input for January 28, 29 and 30, 2018. The “forecast” was actually reconstructed on January 31, 2018 when the rainfall event was over and the observed rainfall data became available. One can see from this figure that the observed peak on January 29, 2018 is very close to the forecast lower bound. This forecast error, which is much smaller than that shown in Figure 8, reflects the model errors.

5. Conclusions

After the above description, derivation and analyses, it can be concluded that: (a) the COFFEE Model is a very time-efficient model; (b) the accuracy of the ensemble-analogue forecasts of average, maximum and minimum instantaneous peaks is acceptable; and thus (c) the COFFEE Model is able to meet the challenges of climate change impacts for operational real-time flood forecasting for the coastal storm dominated watersheds in BC.

6. Acknowledgements

The author thanks all staff members of BC River Forecast Centre for their full support in all aspects during the model development and tests.
Figure 8. An example of forecast rainfall errors and forecast and observed discharges at Roberts Creek at Roberts Creek (08GA047) for January 28 to 30, 2018.
Figure 9. Reconstructed model forecasts using the observed rainfall as model input for Roberts Creek at Roberts Creek (08GA047) for January 28 to 30, 2018

References


