

Western US Stream Flow Metric Dataset: Modeled Flow Metrics for Stream Segments in Selected Basins of the Western United States under Historical Conditions and Forecast Climate Change Scenarios

Dataset Description

25 January, 2011; updated 27 June 2011

http://www.fs.fed.us/rm/boise/AWAE/projects/modeled_stream_flow_metrics.shtml

Overview

The flow regime is of fundamental importance in determining the physical and ecological characteristics of a river or stream, but actual flow measurements are only available for a small minority of stream segments, mostly on large rivers. Flows for all other streams must be extrapolated or modeled. Modeling is also necessary to estimate flow regimes under future climate conditions. To date there are few databases of modeled stream flows that are broad in coverage, fine in resolution, and available for both historical and future climate conditions. Here we present such a database.

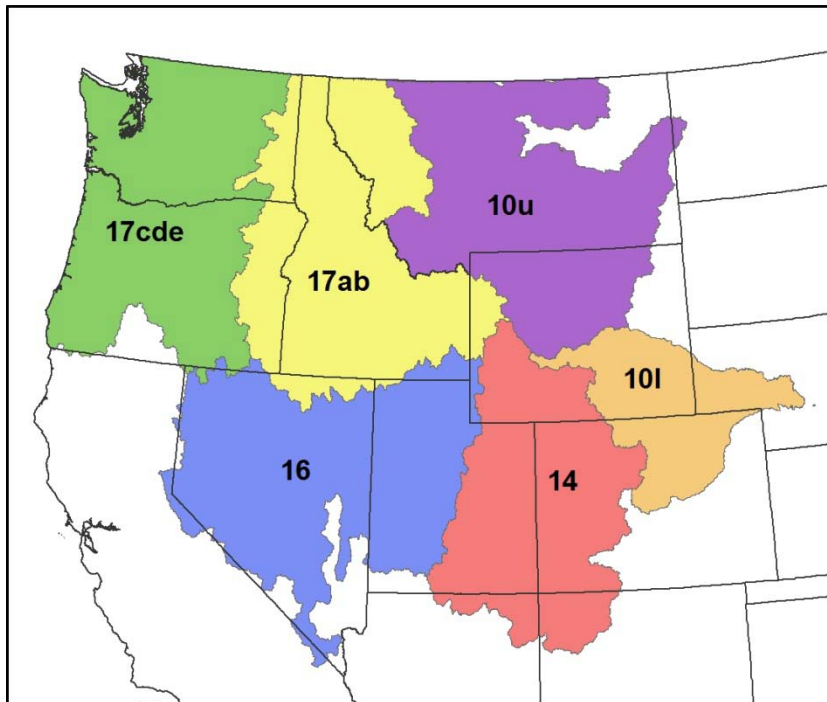
The University of Washington Climate Impacts Group, Trout Unlimited, and the US Forest Service Rocky Mountain Research Station (RMRS) used the Variable Infiltration Capacity (VIC) macroscale hydrologic model to estimate stream flows at a daily timestep under historical and forecasted future climate conditions across major river basins of the western United States. Forecasts were for the A1B emissions scenario for the 2040s and 2080s, using (1) an ensemble of 10 global climate models (GCMs), (2) MIROC3.2, a model that predicts a large amount of warming in the region, and (3) PCM1, a model that predicts a low amount of warming. The individual models (MIROC and PCM) were used to bracket the range of uncertainty in warming and precipitation, while the ensemble mean represented a best-guess model consensus. For the historical scenario and each of the six future scenarios, we estimated a hydrograph for every individual stream segment in the 1:100,000 scale National Hydrography Dataset ([NHD](#)), excluding large rivers. We then calculated a set of summary flow metrics to describe key attributes of the flow regime for each segment. These may be attached to the NHD stream segments for visualization, or used for various kinds of analyses.

Data Coverage

These files cover portions of four major river basins in the Western US: the Columbia (plus coastal drainages), the Upper Colorado, the Great Basin, and selected portions of the Upper Missouri (Figure 1). The coverage area was defined to encompass the historical range of inland cutthroat trout, but for practical reasons the Rio Grande and the Arkansas basins were excluded. Rivers with a drainage area larger than about 250 km² are not covered in the dataset. This is because our method for flow routing (Wenger et al. 2010a) does not include travel time in the river channel. This produces reasonable results

for streams of small to moderate drainage, in which flow responds to precipitation within hours rather than days, but it is not realistic for larger rivers. In addition, routed flows for large rivers in some regions (particularly the Pacific Northwest), are already available from the Climate Impacts Group.

Figure 1. Coverage of modeled flow metric dataset. The data are divided into six groups, labeled by NHD regions/production units.



Methods

VIC is a fully-distributed and largely physically-based model that balances surface energy and water fluxes. Infiltration, runoff, and baseflow processes are based on empirically derived relationships (Liang et al. 1994) and characterize the average conditions over each grid cell. In this analysis, grid cells were 1/16 degree (~5km) on a side, except in the Great Basin, where 1/8 degree cells were used. For the historical simulations, meteorological forcing data were produced using interpolated weather station data, corrected for biases in temperature and precipitation using PRISM products (Daly et al. 1994, Hamlet and Lettenmaier 2005, Maurer et al. 2002). Simulations were performed on a daily time step from the period 1915 through 2006. The VIC modeling methods used here are described in more detail in Elsner et al., 2010.

For the projected climate scenarios, we used meteorological data from GCMs for the 2040s and 2080s associated with the A1B greenhouse gas emissions trajectory (IPCC 2007). The A1B is a middle-of-the road scenario in terms of its assumptions for the accumulation of atmospheric greenhouse gases (IPCC

2007). For each of the future time periods we used projections from three models: (1) MIROC 3.2, which tends to project greater warming and less summer precipitation in the study region than other GCMs, (2) PCM1, which tends to project less warming and more summer precipitation, and (3) a mean of the 10 IPCC models with the lowest bias in simulating observed climate across the region of interest (Littell et al. 2010). The MIROC and PCM models served to bracket the range of possible outcomes; the former projected a mean increase of 5.51 C in mean summer temperature by the 2080s for the study region, whereas the latter projected a mean increase of 2.49 C for the same period. GCM model simulations were downscaled using a spatially explicit delta method (Littell et al. 2010). The traditional delta method involves perturbing the historical time series of meteorological data with spatially uniform monthly changes in temperature and precipitation derived from GCMs. This has the advantage of preserving natural interannual variability in precipitation, which tends to be underestimated by GCMs (Elsner et al., 2010), and allows straightforward comparisons between current and future conditions. The spatially explicit delta method incorporates spatial variability in temperature and precipitation trends from the GCM projections, making the method more appropriate for broad-scale analysis.

To create stream hydrographs, we combined the baseflow and runoff values from the VIC modeling to yield an estimated daily stream flow produced by each cell. We then used a simplified routing procedure to translate cell-based flows into flows for stream segments in the 1:100K NHD stream network, based on stream drainage area and a lag to simulate the time required for flows to exit a cell (Wenger et al. 2010a). As noted above, rivers of greater than about 250 km² drainage area were excluded from the processing because the simplified routing method is intended for small streams, and would overpredict peak flow magnitude in larger streams.

The result of this was a set of seven files for each stream segment: one for the historical scenario, and one for each of the six forecasts. Each file, about 300 kb uncompressed, included a single column of 33,602 numbers representing daily stream flow values (i.e., a 92-year daily hydrograph). From these hydrographs we extracted a set of 12 flow metrics to represent selected attributes of the flow regime. Our emphasis was on flow metrics that had a hypothesized relationship to trout distributions in the region. We validated the model performance by comparing the estimated metrics with observed metrics at a subset of sites with USGS gaging stations and relatively unmodified flows (Wenger et al. 2010a). We describe the flow metrics and the results of the validation in the section “Flow Metrics,” below.

File Naming and Organization

Files are organized in folders corresponding to NHD regions (2-digit hydrologic unit codes) or a subdivision of regions based on NHDplus “production units” (<http://www.horizon-systems.com/NHDPlus/>). Production units are designated by letters appended to the region code, such as “17a” (a part of the Pacific Northwest region that covers the Snake River Basin). The coverage of each of the files is shown in Figure 1 and summarized below:

10u: Units 10g and 10h of the Missouri in Montana and western Wyoming. Because Region 10 (the Missouri Basin) is so large, it is subdivided into an upper group (with units 10e, 10f, 10g and 10h) and lower group (with units 10a, 10b, 10c, 10d and 10i). ["Upper Missouri"]

10l: Unit 10c of the lower section of the Missouri basin ["Platte"].

14: Upper Colorado (units 14a and 14b) ["Upper Colorado"].

16: Great Basin (units 16a and 16b) ["Great Basin"].

17ab: Eastern Columbia Basin (units 17a, 17b) ["Interior Columbia"].

17cde: Western Columbia Basin and coastal watersheds in the Pacific Northwest (units 17c, 17d and 17e) ["Costal PNW"].

Within these folders, filenames begin with "fm" for "flow metric", followed by the folder number, followed by a suffix for the scenario. The suffixes are defined below.

hist: historical

2040c: 2040s A1B scenario with the composite (ensemble) of 10 models

2040m: 2040s A1B scenario with the MIROC model

2040p: 2040s A1B scenario with the PCM model

2080c: 2080s A1B scenario with the composite (ensemble) of 10 models

2080m: 2080s A1B scenario with the MIROC model

2080p: 2080s A1B scenario with the PCM model

Each file has 13 fields. The first is the "ComID," which provides a unique identifier for each NHD stream segment. This field is the key for merging (joining) the file with NHD stream tables. The remaining twelve fields are flow metrics, which are described in the next section.

Flow Metrics

The 12 flow metrics are described below. For the purposes of these flow metrics, winter was defined as December 1 through February 28. The start of summer was calculated individually for each stream segment and each year as the first day after June 1 when flows fell below the mean annual value; this ensured that summer started after the subsidence of the snowmelt flood. Summer was assumed to end on September 30, regardless of the starting date.

Daily Mean (DM). The mean daily flow, averaged over a year. Also sometimes called mean annual flow. Units: cubic feet per second (cfs).

Winter2yr (W2). The probability of a 2-year flow event occurring in the winter. Units: probability (0-1).

Winter 1.5yr (W1.5). The probability of a 1.5-year flow event occurring in the winter. Units: probability (0-1).

Winter99 (W99). The number of days in the winter in which flows are among the highest 1% for the year. Units: frequency (0+).

Winter 95 (W95). The number of days in the winter in which flows are among the highest 5% for the year. Units: frequency (0+). Of the winter high flow metrics, we found this one best correlates to bull trout occurrence. Bull trout occur fairly commonly in streams with values of less than 0.5 for this metric, uncommonly at values of 0.5-3, and are rare at values above 3.

Channelflow (Q1.5). The 1.5-year flow, sometimes considered the channel-forming flow. Units: cfs.

CtrFlowMass (CFM). Timing of the center of the mass of flow. Also called center of timing or CT. The day of the water year at which 50% of the year's flow has passed. Units: day number of the water year.

Summer95 (S95). The number of days in the summer in which flows are among the highest 5% for the year. Summer here is defined as Based on our validation this is usually NOT accurate and is not recommended for use.

Summer20 (S20). The number of days in the summer in which flows are 20% of the daily mean. Based on our validation this is usually NOT accurate and is not recommended for use.

MeanSummer (MS). The mean flow during the summer. Units: cfs.

Highlow (HL). The ratio of high flow magnitude to low flow magnitude. Based on our validation this is usually NOT accurate and is not recommended for use.

Flow7q10 (7Q10). The 7-day low flow with a 10-year return interval. Units: cfs.

Performance statistics for selected flow metrics, based on comparison of predicted vs. observed values for 55 gaging stations in the Pacific Northwest, are shown in the table below. Data are from Wenger et al. 2010a. “MAPE” is the Mean Absolute Percent error; “Bias” is the mean bias.

Flow metric	MAPE	Bias
MA	18%	-12%
W2	32%	4%
W1.5	29%	9%
W99	27%	-7%
W95	22%	-3%
S95	245%	181%
S20	83%	-29%
MS	32%	-10%
7Q10	57%	-10%

Using the Files

The flow metric files may be joined with NHD-plus stream segments using the ComID field for analyses and presentation in a GIS. The files are in the format of comma separated values (.csv), which is an efficient format for most purposes, but is slow when used in ESRI ArcGIS, which works fastest with files in “.dbf” format. Conversion to .dbf can be done with certain spreadsheet and statistical analysis software. Note that newer versions of Excel do not support writing to .dbf, and older versions often introduced errors when doing so. We recommend using the free statistical software *R*. Code for converting is simple:

```
library (foreign)
fmtable <- read.csv("C:/flowmetrics/fm10u_2040c.csv")
write.dbf(fmtable, "C:/flowmetrics/fm10u_2040c.dbf")
```

In the code above, the file pathways in quotes should be changed to the appropriate location of the source .csv file and the desired location for the resulting .dbf file. Note that in *R*, a forward slash is used in file pathways rather than a backward slash.

For some purposes it can be useful to summarize flow metrics values at the scale of watersheds or subwatersheds. This can be done based on mean, median, weighted mean, or the value at the watershed outlet, depending on the objective. To facilitate this, a file that relates ComID to the 12-digit hydrologic unit code (“comid-huc12.csv”) is available on the Western US Stream Flow Metric Dataset website.

Limitations

As noted above, previous validations (Wenger et al. 2010a) found that some of the flow metrics were not modeled accurately. It is also important to recognize that the VIC model used to generate these data

was calibrated and validated only in the Columbia Basin and the Pacific Northwest, and there have been no formal published validations of the performance of this dataset in the Colorado Basin, Missouri Basin or Great Basin. We conducted some limited-scale validations at sites in the Upper Missouri and Upper Colorado (Wenger et al 2010b), and found that while some aspects of the flow regime were faithfully reproduced, there was a consistent bias to the prediction of the timing of the spring snowmelt runoff. In all cases we examined, spring snowmelt was predicted to occur earlier than was observed. This appears to be the same phenomenon we observed in the Pacific Northwest validation (Wenger et al. 2010a), in which rain-dominated sites were biased late and snow-dominated sites were biased early in the predictions of snowmelt runoff timing (CFM). A simple regression of predicted vs. observed CFM at the 55 validation sites produced the following relationship, which can serve as a correction factor:

$$\text{Actual.CFM} = (1.25 * \text{Predicted. CFM}) - 44$$

Without this correction, the projected rate of change in CFM due to climate warming will be underestimated. We suggest that a correction such as this be employed until the underlying cause of the bias in the VIC modeling can be identified and corrected.

Other potential issues with the data are discussed in Wenger et al. 2010a and other publications on VIC modeling. We strongly recommend that users carefully consider potential errors, biases and limitations before drawing inferences from analyses based on these data.

Additional Data and Information

The number of flow metrics included in this dataset was limited for practical reasons. Although additional flow metrics can be derived from the full stream hydrographs, we cannot distribute the hydrographs themselves via the Internet because the full database is in excess of a terabyte in size. We will do our best to accommodate requests for the full hydrographs on an ad hoc basis. Direct data requests and any questions or comments to:

Seth Wenger, Trout Unlimited. swenger@tu.org

Charlie Luce, RMRS. cluce@fs.fed.us

References

Daly, C., R.P. Neilson, and D.L. Phillips, 1994. A statistical topographic model for mapping climatological precipitation over mountainous terrain, *Journal of Applied Meteorology* 33: 140-158.

Elsner M., L. Cuo L, N. Voisin, J. Deems, A. Hamlet, J. Vano, K. Mickelson, S. Lee, and D. Lettenmaier, 2010. Implications of 21st century climate change for the hydrology of Washington State. *Climatic Change*: 102, 225-260.

Hamlet, A.F. and D.P. Lettenmaier, 2005. Production of temporally consistent gridded precipitation and temperature fields for the continental United States, *Journal of Hydrometeorology* 6: 330-336.

IPCC, 2007. *Climate Change 2007: Working Group 2: Impacts, Adaptation and Vulnerability*. Intergovernmental Panel on Climate Change, Geneva, Switzerland.

Liang, X., D.P. Lettenmaier, E.F. Wood, and S.J. Burges, 1994. A simple hydrologically based model of land-surface water and energy fluxes for general-circulation models. *Journal of Geophysical Research-Atmospheres* 99: 14415-14428.

Littell J.S., M.M. Elsner, G.S. Mauger, E. Lutz, A.F. Hamlet, and E. Salathe, 2010. *Regional Climate and Hydrologic Change in the Northern US Rockies and Pacific Northwest: Internally Consistent Projections of Future Climate for Resource Management*. Available online: http://cses.washington.edu/picea/USFS/pub/Littell_etal_2010.

Maurer, E.P., A.W. Wood, J.C. Adam, D.P. Lettenmaier, and B. Nijssen, 2002. A long-term hydrologically based dataset of land surface fluxes and states for the conterminous United States, *Journal of Climate* 15: 3237-3251.

Wenger S.J., C.H. Luce, A.F. Hamlet, D.J. Isaak, and H.M. Neville, 2010a. Macroscale hydrologic modeling of ecologically relevant flow metrics. *Water Resources Research* 46: W09513, doi 10.1029/2009WR008839.

Wenger S.J., C.H. Luce and D.J. Isaak, 2010b. *Comparison of hydrologic predictions from the Variable Infiltration Capacity (VIC) model and the MC1 model to observed gage data in the region around the Shoshone National Forest*. Trout Unlimited/USFS Rocky Mountain Research Station, Boise, ID. Available online: http://www.fs.fed.us/rm/boise/AWAE/projects/modeled_stream_flow_metrics.shtml.

Acknowledgments

This dataset was developed with funding from grant 2008-0087-000 from the National Fish and Wildlife Foundation's Keystone Initiative for Freshwater Fishes. VIC modeling was performed by Alan Hamlet and Marketa McGuire Elsner of the Climate Impacts Group, University of Washington. Development of the flow metric database was by Seth Wenger (Trout Unlimited), Charlie Luce (RMRS), Dan Isaak (RMRS), and Helen Neville (Trout Unlimited). Bruce Rieman (RMRS), Jason Dunham (US Geological Survey) and Kurt Fausch (Colorado State University) helped to select appropriate hydrologic metrics.

Correction: The third line of R code on page 6 was incorrect in the original version of this document. It was corrected on 27 June, 2011.