WEHY Validation

Mechele Pacheco

Abstract

The WEHY model is a watershed model that uses the ensemble-averaged forms of upscaled hydrologic conservation equations. The model can be used for ungauged watersheds because most of the model parameters can be estimated from the topography, soil, and land use/land cover characteristics of the watershed. However, some observed data is needed to both calibrate and validate the model parameters not estimated using these methods. If no observed data is available, the parameters not estimated using the topography, soil, and land use/land cover characteristics must be estimated using physical characteristics and tables in literature. In this paper, the estimated natural flows of Butte Creek from the Natural Flows Database will be used to validate the model using the default parameters for the time period 1980 to 1990. The validation was assessed using R^2, RSR, NSE, and PBIAS values. While using the default parameters, the model evaluation statistics are R^2 = 0.69, RSR = 12.08, NSE = 0.10, and PBIAS = 89 at the bottom of the watershed, R^2 = 0.39, RSR = 10.97, NSE = 0.20, and PBIAS = 64 at the middle of the watershed, and R^2 = 0.09, RSR = 11.08, NSE = 0.09, and PBIAS = -48 at the top of the watershed. While using the default parameters, the model evaluation statistics for the summer months are R^2 = 0.16, RSR = 6.13, NSE = 0.19, and PBIAS = 28 at the bottom of the watershed, R^2 = 0.09, RSR = 27.63, NSE = -10.60, and PBIAS = -195 at the middle of the watershed, and R^2 = 0.21, RSR = 39542.97, NSE = -18.48, and PBIAS = -2427378 at the top of the watershed. Therefore, while using the default conditions the model is unsatisfactory at all locations in the watershed.

Introduction

The DeSabla-Centerville Hydroelectric Project (D-C Project) diverts water from the West Branch Feather River (WB Feather River) into Butte Creek. These diversions transfer colder water from the WB Feather River into Butte Creek and are important for the salmon population. Since Butte Creek supports the largest population of spring-run Chinook salmon in California, it is important to understand the ecosystem of the region and how the D-C Project effects it. In order to understand the effects of the D-C Project operations on the spring-run Chinook salmon, the temperature in the streams must be modeled. The first step of modeling the temperature of the streams is modeling the rivers natural flows.

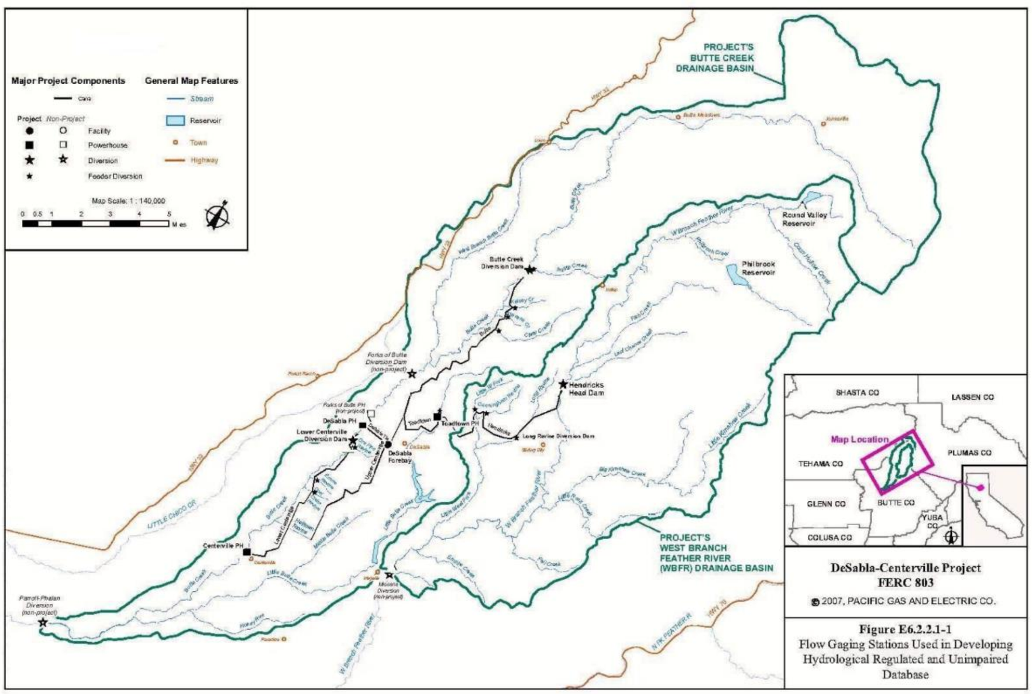

Figure 1: Map of the DeSabla-Centerville Project

The rivers natural flows will be modeled using the watershed environmental hydrology (WEHY) model. The WEHY model uses upscaled conservation equations that are the ensemble averaged forms of the point-location scale conservation equations. The parameters in the upscaled equations are areal averages and areal variances of the original point-scale parameters. Most of the model parameters can be estimated from the topography, soil, and land use/land cover characteristics of the watershed. The parameters not estimated from the topography, soil, and land use/land cover characteristics of the watershed can be calibrated using observed historical data. If no historical data is available, the parameters not estimated from the topography, soil, and land use/land cover characteristics of the watershed must be estimated using physical characteristics and tables in literature.

Objective

The main objective of this project is to validate the modeling of the river's natural flows by the WEHY model. In the WEHY model, default values will be used for the parameters not estimated from the topography, soil, and land use/land cover characteristics of the watershed. The model will be validated with the California Unimpaired Flows Database. The model will be considered satisfactory if $R^2 > 0.5$, $RSR < 0.70$, $NSE > 0.50$, and $PBIAS < \pm 25$. The model will be considered good if $RSR < 0.60$, $NSE > 0.65$, and $PBIAS < \pm 15$. The model will be considered very good if $0 < RSR < 0.50$, $0.75 < NSE < 1$, and $PBIAS < \pm 10$.

Hypothesis (optional)

Since default values are being used for the parameters not estimated using the topography, soil, and land use/land cover characteristics, it is expected that the model will meet the satisfactory conditions, but not the very good or good conditions.

Data Sources

The topography data was obtained from the National Elevation Database which is maintained by the U.S. Department of Agriculture (USDA) Natural Resources Conservation Services, National Cartography & Geospatial Center [8]. The soil data was obtained from the Soil Survey Geographic (SSURGO) Database which is maintained by the U.S. Department of Agriculture (USDA) Natural Resources Conservation Services [9]. The land use/land cover data was obtained from the FVEG database which is maintained by the Department of Forestry and Fire Protection [3]. The atmospheric data was obtained from the NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products [7]. The natural flow in Butte Creek was obtained from the California Unimpaired Flows Database v2.0.1 (Natural Flows Database) [10].

Methods and Assumption

1. Perform the watershed delineation for Butte Creek.
2. Estimate the model parameters using the topography, soil, and land use/land cover characteristics of the Butte Creek watershed. [8][9][3]
3. Obtain WRF outputs for the region. [7]
4. Run the snow model with the parameters estimated using the topography, soil, and land use/land cover characteristics of the Butte Creek watershed.
5. Run WEHY model for the time period 1980 to 1990 with default values for parameters not estimated using the topography, soil, and land use/land cover characteristics.
6. Convert WEHY model outputs from hourly values to monthly values.
7. Compare (R^2, RSR, NSE, PBIAS) WEHY model output to Natural Flows Database values [10]

The two models will be compared at three locations in the watershed (bottom, middle, and top). The bottom of the watershed will be compared in the WEHY reach 1. The middle of the watershed will be compared in the WEHY reach 6. The top of the watershed will be compared in the WEHY reach 16.

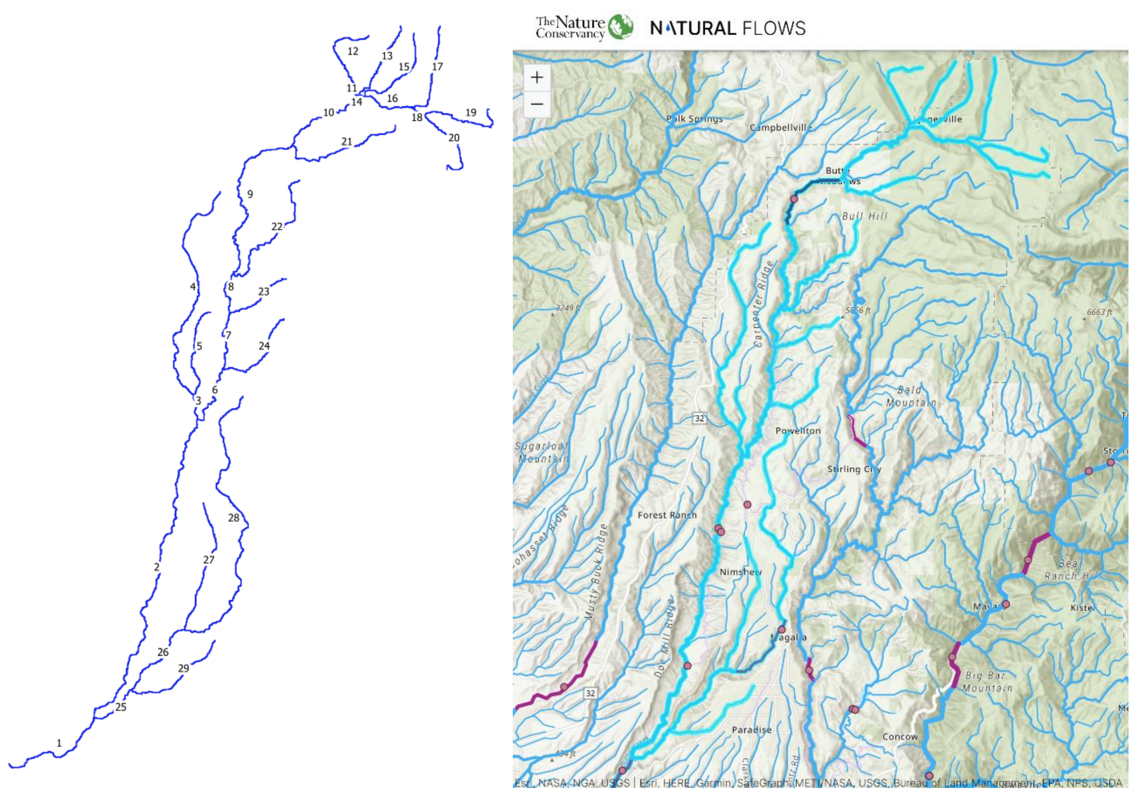

Figure 2: Comparison of WEHY reaches and Natural Flow Database Streams

Results

The monthly streamflow at the bottom of the watershed from WEHY and the Natural Flows Database (abbreviated NFD in the figure) can be seen in Figure 3. The time series data shows that the two models have similar seasonal patterns but the magnitude of the peak flows is much larger in the WEHY model. These peaks occur in the winter months of each water year. The peaks are usually preceded with a month of extremely low flow in the WEHY model. The same pattern can be seen in the middle and top of the watershed (*Figures 4 and 5*). Although, the magnitude of the difference between the WEHY model and the Natural Flows Database streamflow decreases further up the watershed. Both the WEHY model and the Natural Flows Database streamflow decreases further up the watershed, but the WEHY model decreases at a faster rate than the Natural Flows Database.

The statistical model evaluation techniques for each location in the watershed can be located in Table 1. The only value that meets the satisfactory criteria is the WEHY reach 1 R^2 value. Therefore, the model is unsatisfactory at all locations in the watershed. *In Table 2, the R^2 value is calculated for each year at the three locations.* In this table, the green background indicates the value is satisfactory and the orange background indicates the value is unsatisfactory. The water year type is not the cause of the difference between models since the beginning of the time period is wet and the end of the time period is dry and there does not seem to be a difference based on this pattern.

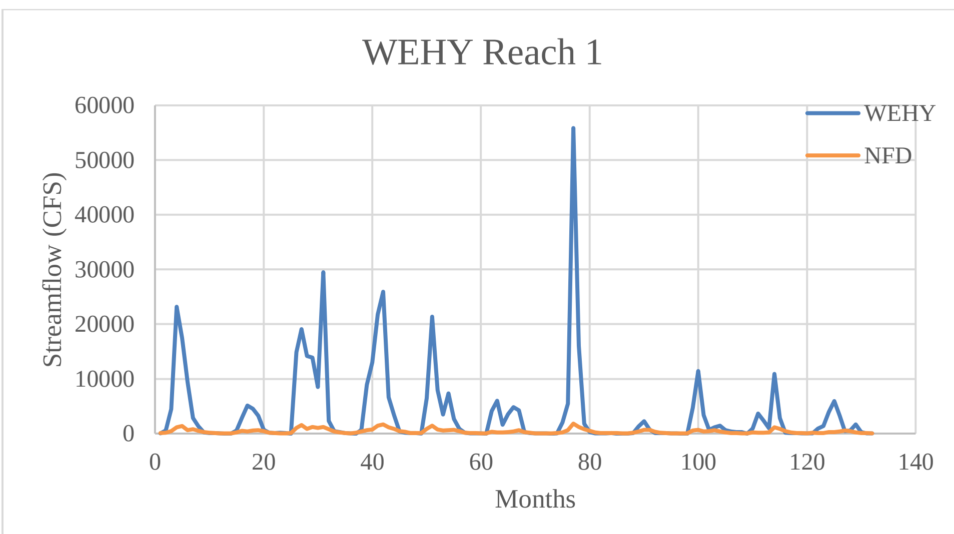

Figure 3: Monthly Streamflow at the Bottom of the Watershed

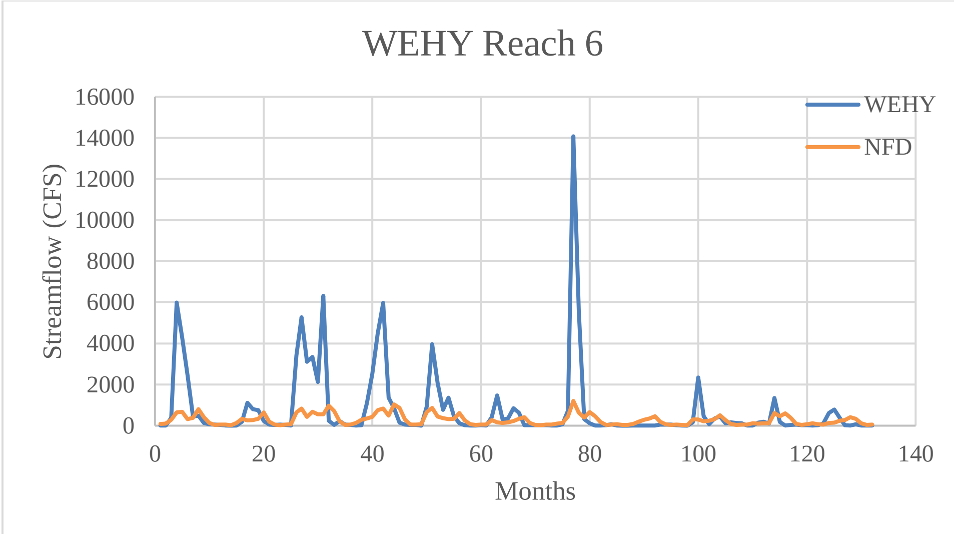

Figure 4: Monthly Streamflow at the Middle of the Watershed

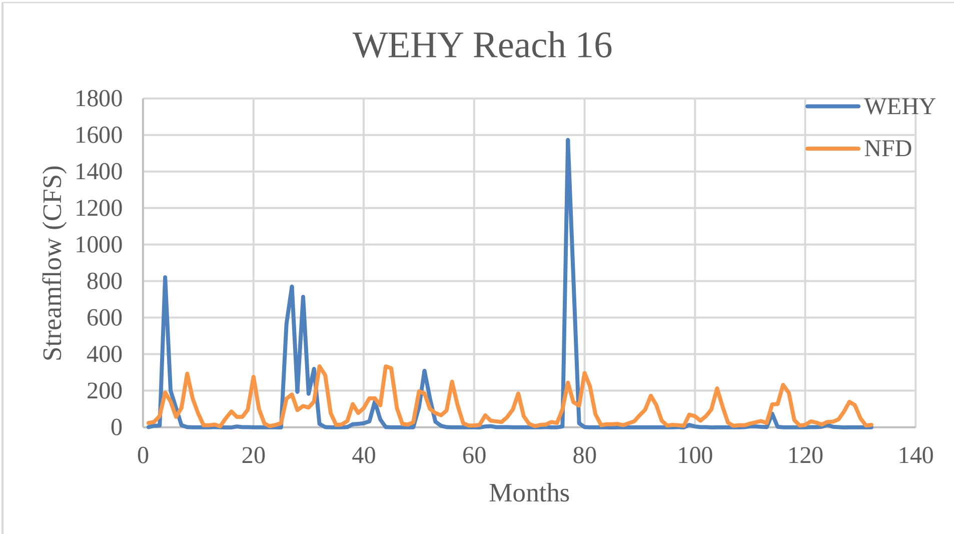

Figure 5: Monthly Streamflow at the Top of the Watershed

Table 1: Annual Statistics for 1980 to 1990

Annual			
	WEHY 1	**WEHY 6**	**WEHY 16**
R^2	0.690	0.387	0.091
RSR	12.081	10.965	11.084
NSE	0.102	0.204	0.093
PBIAS	89	64	-48

Table 2: Annual R^2 by Year

	R^2		
YEAR	**WEHY 1**	**WEHY 6**	**WEHY 16**
1980	0.976	0.949	0.861
1981	0.232	0.030	0.978
1982	0.959	0.011	0.946
1983	0.785	0.967	0.952
1984	0.866	0.736	0.976
1985	0.994	0.534	0.031
1986	0.187	0.267	0.973
1987	0.659	0.027	0.842
1988	0.953	0.090	0.999
1989	0.550	0.239	0.998
1990	0.928	0.977	0.981

The monthly streamflow for the summer months of June to September at the bottom of the watershed from WEHY and the Natural Flows Database (abbreviated NFD in the figure) can be seen in Figure 6. The time series data shows that the two models are similar until 1988 when the WEHY model has much higher streamflow than the Natural Flows Database in the months of June and July. These peaks in June and July are not visible in the WEHY model at the middle and top of the watershed (*Figures 7 and 8*). In fact, the Natural Flows Database has higher streamflow for both the middle and top of the watershed. The two models at the bottom of the watershed seem to follow the same pattern, but they become less similar further up the watershed. They WEHY model predicts almost no streamflow in the summer months at the top of the watershed.

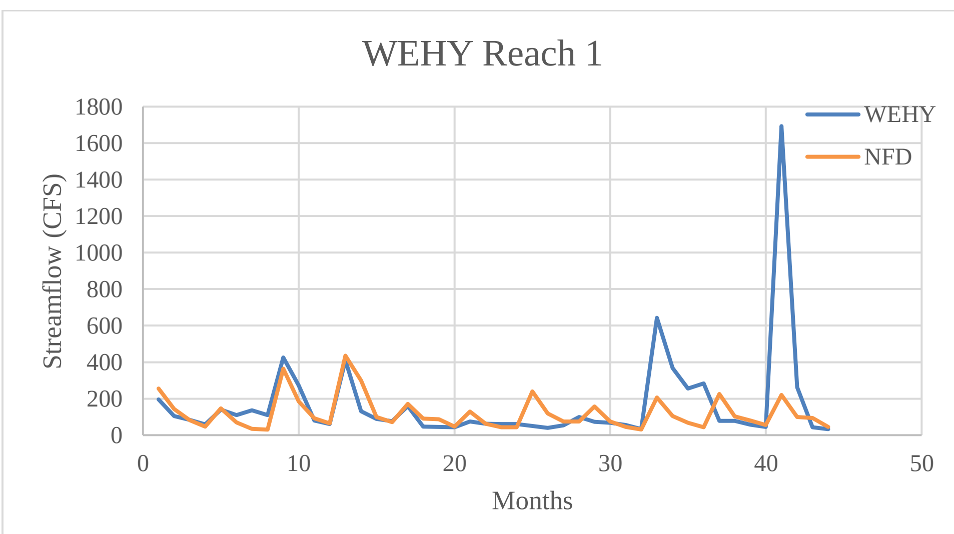

Figure 6: Monthly Streamflow for Summer Months at the Bottom of the Watershed

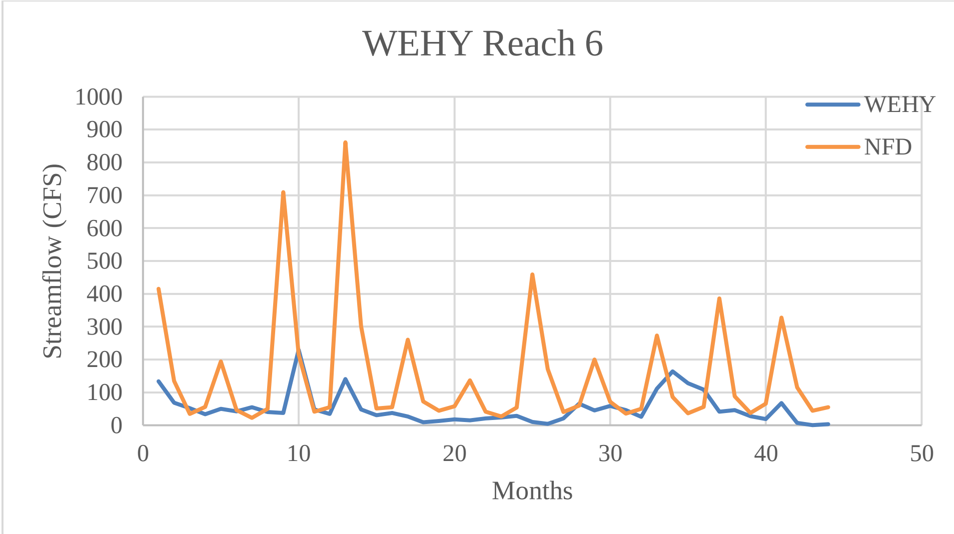

Figure 7: Monthly Streamflow for Summer Months at the Middle of the Watershed

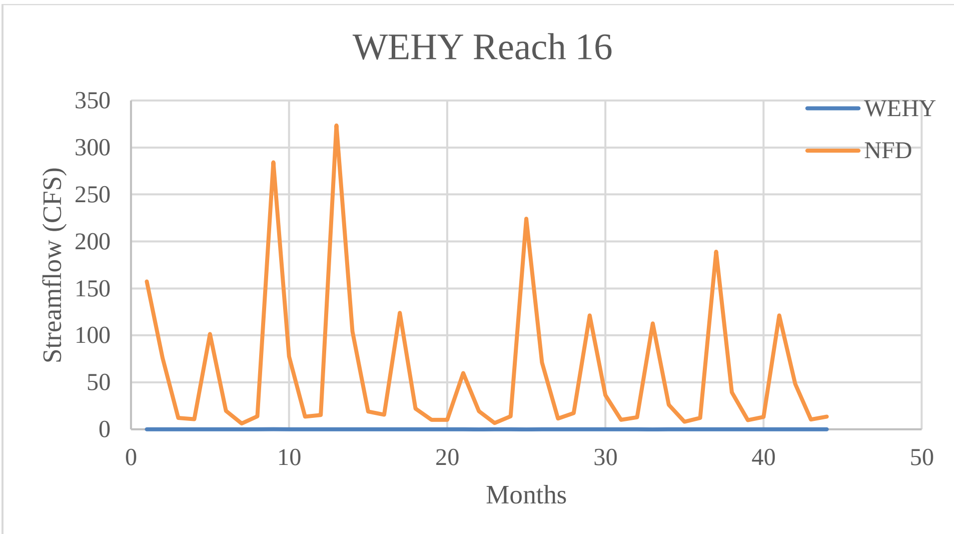

Figure 8: Monthly Streamflow for Summer Months at the Top of the Watershed

The statistical model evaluation techniques for each location in the watershed can be located in Table 3. None of the values meets the satisfactory criteria. Therefore, the model is unsatisfactory at all locations in the watershed. *In Table 4, the R^2 value is calculated for the summer months of each year at the three locations.* In this table, the green background indicates the value is satisfactory and the orange background indicates the value is unsatisfactory. The water year type is not the cause of the difference between models since the beginning of the time period is wet and the end of the time period is dry and there does not seem to be a difference based on this pattern. The table again shows that the WEHY model streamflow is most similar to the Natural Database streamflow at the bottom of the watershed as seen in Figures 6-8.

Table 3: Summer Months Statistics for 1980 to 1990

Summer (June-September)			
	WEHY 1	**WEHY 6**	**WEHY 16**
R^2	0.163	0.086	0.208
RSR	6.129	27.629	39542.976
NSE	0.187	-10.596	-18.484
PBIAS	28	-195	-2427378

Table 4: Summer Months R^2 by Year

R^2			
YEAR	**1**	**6**	**16**
1980	0.785	0.374	0.132
1981	0.665	0.191	0.002
1982	0.716	0.203	0.019
1983	0.813	0.278	0.003
1984	0.851	0.542	0.207
1985	0.484	0.075	0.007
1986	0.787	0.692	0.198
1987	0.652	0.054	0.043
1988	0.528	0.130	0.002
1989	0.669	0.310	0.043
1990	0.081	0.001	0.065

Conclusions

The hypothesis that the model would be satisfactory was incorrect. The model is unsatisfactory when analyzing both the full year and the summer months at all locations in the watershed. Therefore, the parameters not estimated from the topography, soil, and land use/land cover characteristics of the watershed need to be calibrated using observed historical streamflow data in Butte Creek. From the initial attempts at calibration, it was found that doubling the Chezy coefficient resulted only in a change in the thousandths place of the statistics. The most sensitive parameter in the calibration process is the soil depth adjusting factor. The roughness and albedo lengths in summer and winter are the next most sensitive parameters.

Recommendation/Limitations

The WEHY model and the Natural Flows Database may have disagreements that affect the validation process since the Natural Flows Database is a model itself. The WEHY model should first be calibrated and validated using observed historical data. The calibration and validation of the model is currently limited to the time period of 1980 to 1990 due to the WRF model outputs. There is available streamflow historical data from USGS for water year 1931 to present at the bottom of the watershed (WEHY reach 1) and water year 1961 to water year 1974 at the top of the watershed (WEHY reach 9). With these limitations the WEHY model can be currently calibrated and validated for water year 1980 to 1990 at the bottom of the watershed. Once this calibration and validation against the observed data is completed, the model can be validated again with the Natural Flows Database. In order to fully understand the region, the same process needs to be completed for the WB Feather River.

References

[1] California Environmental Flows Working Group (CEFWG). 2020. California Environmental Flows Framework. California Water Quality Monitoring Council Technical Report 37.

[2] Chen, Z. Q., et al. 2004. Watershed environmental hydrology (WEHY) Model: Model Application. J. Hydrol. Eng., 9(6), 480-490.

[3] FVEG. Department of Forestry and Fire Protection. https://frap.fire.ca.gov/mapping/gis-data/. March 18, 2021.

[4] Kavvas, M. L., et al. 2004. Watershed environmental hydrology (WEHY) model, based on upscaled conservation equations: Hydrologic module. J. Hydrol. Eng., 9(6), 450–464.

[5] Kavvas, M. L., et al. 2013. WEHY-HCM for modeling interactive atmospheric-hydrologic processes at watershed scale. I: Model description. J. Hydrol. Eng., 18(10), 1262-1271.

[6] Kure, S., et al. 2013. WEHY-HCM for modeling interactive atmospheric-hydrologic processes at watershed scale. II: Model application to ungauged and sparsely gauged watersheds. J. Hydrol. Eng., 19(10), 1272-1281.

[7] Saha, S., et al. 2010. NCEP Climate Forecast System Reanalysis (CFSR) 6-hourly Products, January 1979 to December 2010. Research Data Archive at the National Center for Atmospheric Research, Computational and Information Systems Laboratory. https://doi.org/10.5065/D69K487J.

[8] Soil Survey Staff. National Elevation Data 30 meter for California. United States Department of Agriculture, Natural Resources Conservation Service. https://gdg.sc.egov.usda.gov/. February 07, 2020 (FY2020 official release).

[9] Soil Survey Staff. Soil Survey Geographic (SSURGO) Database. United States Department of Agriculture, Natural Resources Conservation Service. https://sdmdataaccess.sc.egov.usda.gov. May 31, 2020.

[10] Zimmerman, J.K.H., et al. December 2020. California Unimpaired Flows Database v2.0.1. The Nature Conservancy. San Francisco CA. https://rivers.codefornature.org/. May 05, 2021.