If the stage discharge relationship is subject to change because of changes in the physical features that affect the gauge site, discharge is determined by the shifting-control method. In this method, correction factors based on individual discharge measurements and notes of personnel making the measurements are used when applying gauge heights to the rating tables. This method is also used if the stage discharge relationship is changed temporarily due to aquatic growth or debris on the control. Downstream flow obstructions may produce backwater effects that reach the gauge. Upstream obstructions may change the cross-sectional area.

Since there where many parts of the calibration fit in the objective function covering daily, monthly, annual and statistical features of flow, along with estimates of total evaporation, the overall problem with unmodeled hydrologic features is that the calibration will appear poor. If the long-term volumes match, the recharge and maximum saturated evapotranspiration estimates should be robust.

A very common measure of the performance of a hydrologic model is the Nash-Sutcliffe statistic (Moriasi et al. 2007). A Nash-Sutcliffe statistic equal to one is a perfect match between simulated and observed. A zero would mean that the average of the observations is a better model. Negative Nash-Sutcliffe values are possible, though they do not have a meaning. The Nash-Sutcliffe model performance categories are listed in Table 9-16. The spatial distribution of Nash-Sutcliffe values is show in Figure 9-26.

The calibration performance results for the watersheds are presented in Table 9-17. A total of 243 gauges within 50 HUC8 watersheds were used for calibration. Five HUC8 watersheds were ungauged and parameters were used from adjacent models to run them.

Figure 9-27 and Figure 9-28 compare measures of model performance against data quality. Note from the figures that the USGS has not identified any gauge as “Excellent” for 2009. The figures show that measures of model performance, like the Nash-Sutcliffe or percent bias, should not be the only way model performance is evaluated since these measures are also dependent on data quality. It can reasonably be asserted from Figure 9-27 and Figure 9-28 that the simulation is better than the data at representing the system because the simulation is constrained by mass balance, other gauges in the model and target evaporation values from literature.

Calibration plots and statistics are provided as appendices for all 243 gauges. These appendices are organized by model and named “Appendix T-XXXXXXXX” where “XXXXXXXX” is the HUC8 number of the model.

Table 9-16. Grading model calibration performance. Adapted from Moriasi et al. (2007)

Performance Rating	Percent Bias (Monthly)	Nash-Sutcliffe (Monthly)
Very good	< ±10	0.75 < NSE < 1.00
Good	±10 < PEM < ±15	0.65 < NSE < 0.75
Satisfactory	±15 < PEM < ±25	0.50 < NSE < 0.65
Unsatisfactory	> ±25	< 0.50