# Calibration Diagnostic Plots: A Comprehensive Guide

## Introduction

Calibrating a hydrological model involves adjusting its parameters to make its outputs match observed data as closely as possible. This process can be complex, with multiple parameters interacting in non-linear ways. The following diagnostic plots help us understand how well our calibration process worked, which parameters are most important, and how different aspects of model performance relate to each other.

The use of diagnostic plots in hydrological modelling has been advocated by numerous researchers to gain insights into model behaviour and performance beyond what single summary statistics can provide (Gupta et al., 2008; Wagener & Gupta, 2005).

## 1. Objective Evolution Plot (objective\_evolution.png)

### What it shows

This plot displays how different measures of model performance (called "objective functions" or "metrics") change as the calibration progresses. We typically use metrics like RMSE (Root Mean Square Error), KGE (Kling-Gupta Efficiency), NSE (Nash-Sutcliffe Efficiency), and MAE (Mean Absolute Error).

### How to read it

- Each line represents a different metric.

- The x-axis shows the iteration number (how far along we are in the calibration process).

- The y-axis shows the value of each metric.

- For RMSE and MAE, lower values are better. For KGE, KGEp, and NSE, higher values are better.

### What to look for

- Do the metrics improve (get better) over time?

- Do they eventually level off, suggesting the calibration has converged?

- Are there any sudden jumps or drops?

- Do some metrics improve while others get worse?

### Why it's useful

This plot helps us understand if our calibration is working and when it might be finished. It can also reveal if there are trade-offs between different aspects of model performance.

### Scientific context

The concept of multi-objective calibration and the potential trade-offs between different performance metrics has been extensively discussed in the hydrological literature (Efstratiadis & Koutsoyiannis, 2010). For example, Kollat et al. (2012) used similar plots to visualize the evolution of multiple objectives in hydrologic model calibration.

## 2. Parameter Convergence Plot (parameter\_convergence.png)

### What it shows

This plot shows how the values of each model parameter change throughout the calibration process.

### How to read it

- Each line represents a different parameter.

- The x-axis shows the iteration number.

- The y-axis shows the parameter value.

- Dotted lines often represent the allowable range for each parameter.

### What to look for

- Do parameter values stabilize (converge) over time?

- Are there parameters that keep changing even in later iterations?

- Do any parameters hit their upper or lower limits?

### Why it's useful

This plot helps us identify which parameters the calibration process is able to determine confidently, and which ones might be less important or more uncertain.

### Scientific context

The convergence behavior of parameters during calibration can provide insights into parameter sensitivity and identifiability (Vrugt & Sadegh, 2013). Pokhrel et al. (2008) used similar plots to analyze parameter convergence in the calibration of a rainfall-runoff model.

## 3. Parallel Coordinates Plot (parallel\_coordinates.html)

### What it shows

This plot represents each parameter set as a line going through several vertical axes, one for each parameter and one for the performance metric.

### How to read it

- Each vertical axis represents a parameter or the performance metric.

- Each line crossing all axes represents one parameter set.

- Lines are often colored based on how well that parameter set performed.

### What to look for

- Are there clear patterns in the parameter values that lead to good performance?

- Are there parameters where the best-performing sets all have similar values?

- Are there parameters where good performance occurs across a wide range of values?

### Why it's useful

This plot can reveal relationships between parameters and performance that might not be obvious when looking at parameters individually.

### Scientific context

Parallel coordinate plots have been used in hydrological modeling to visualize high-dimensional parameter spaces and identify parameter interactions (Khu & Madsen, 2005). Inselberg (2009) provides a comprehensive treatment of parallel coordinates and their applications in various fields.

## 4. Pairwise Scatter Plot (pairwise\_scatter.png)

### What it shows

This plot creates a grid of scatter plots, showing how each parameter relates to every other parameter and to the performance metric.

### How to read it

- Each dot represents one parameter set tested during calibration.

- Dots are often colored based on the performance of that parameter set.

- The diagonal shows the distribution of values for each parameter.

### What to look for

- Are there clear relationships between pairs of parameters?

- Are there parameters that seem to have a strong influence on performance?

- Are there parameters that don't seem to affect performance much?

### Why it's useful

This plot helps us understand how parameters interact with each other and with model performance. It can reveal complex relationships that might not be obvious from other plots.

### Scientific context

Pairwise scatter plots, also known as scatterplot matrices, are widely used in hydrology for exploring parameter interactions and their relationships with model performance. Kingston et al. (2005) used such plots to ensure physically plausible parameterizations in hydrological modeling.

## 5. Best vs. Worst Plot (best\_vs\_worst.png)

### What it shows

This plot compares the parameter values that gave the best model performance with those that gave the worst performance.

### How to read it

- Each pair of bars represents one parameter.

- One bar shows the value for the best-performing parameter set, the other for the worst.

- Dotted lines often show the allowable range for each parameter.

### What to look for

- Which parameters have very different values in the best and worst cases?

- Are there parameters where the best and worst cases have similar values?

### Why it's useful

This plot can quickly highlight which parameters might be most important for getting good model performance.

### Scientific context

Comparing best and worst parameter sets is a simple yet effective way to identify influential parameters. Wagener & Gupta (2005) discuss various approaches to parameter identification in hydrological modeling, including methods that compare high and low performing parameter sets.

## 6. Parameter Correlation Heatmap (parameter\_correlation.png)

### What it shows

This plot shows how strongly each parameter is related to every other parameter and to the performance metric.

### How to read it

- Each cell in the heatmap represents the correlation between two parameters or between a parameter and the performance metric.

- Colors often range from blue (strong negative correlation) through white (no correlation) to red (strong positive correlation).

### What to look for

- Are there parameters that are strongly correlated with each other?

- Which parameters are most strongly correlated with model performance?

- Are there parameters that don't seem to correlate with anything?

### Why it's useful

This plot can reveal redundancies in our model (parameters that essentially do the same thing) and can highlight which parameters are most important for performance.

### Scientific context

Correlation analysis is a fundamental tool in understanding parameter interactions and their relationships with model outputs. Hill & Tiedeman (2007) provide a comprehensive treatment of sensitivity analysis in environmental models, including the use of correlation measures.

## 7. Objective Surface Plot (objective\_surface.png)

### What it shows

This plot gives a visual representation of how model performance changes as we vary the two most influential parameters.

### How to read it

- The x and y axes represent the values of two parameters.

- The color represents the value of the performance metric.

- Contour lines connect points of equal performance.

### What to look for

- Is there a clear "best" region?

- Is the surface smooth or rough?

- Are there multiple peaks or just one?

### Why it's useful

This plot helps us understand the nature of our calibration problem. A smooth surface with a single peak suggests an easy calibration problem, while a rough surface with multiple peaks suggests a more difficult one.

### Scientific context

Response surfaces have been widely used in hydrology to visualize the parameter space and understand the nature of the calibration problem. Duan et al. (1992) used such plots to illustrate the challenges in calibrating conceptual rainfall-runoff models, leading to the development of the SCE-UA algorithm.

## Conclusion

Together, these plots provide a comprehensive view of the calibration process. They can help us understand which parameters are most important, how parameters interact with each other, and whether our calibration process is working well. By carefully interpreting these plots, we can gain insights that help us improve our model and increase our confidence in its results.

Remember, the goal of calibration isn't just to find the "best" parameter values, but to understand our model better. These diagnostic plots are tools to help us build that understanding.

The use of multiple, complementary diagnostic plots aligns with the concept of diagnostic model evaluation advocated by Gupta et al. (2008), who argue for a more comprehensive approach to model assessment beyond simple performance metrics.

## References

Duan, Q., Sorooshian, S., & Gupta, V. (1992). Effective and efficient global optimization for conceptual rainfall-runoff models. Water Resources Research, 28(4), 1015-1031.

Efstratiadis, A., & Koutsoyiannis, D. (2010). One decade of multi-objective calibration approaches in hydrological modelling: A review. Hydrological Sciences Journal, 55(1), 58-78.

Gupta, H. V., Wagener, T., & Liu, Y. (2008). Reconciling theory with observations: Elements of a diagnostic approach to model evaluation. Hydrological Processes, 22(18), 3802-3813.

Hill, M. C., & Tiedeman, C. R. (2007). Effective groundwater model calibration: With analysis of data, sensitivities, predictions, and uncertainty. John Wiley & Sons.

Inselberg, A. (2009). Parallel coordinates: Visual multidimensional geometry and its applications. Springer.

Khu, S. T., & Madsen, H. (2005). Multiobjective calibration with Pareto preference ordering: An application to rainfall-runoff model calibration. Water Resources Research, 41(3).

Kingston, G. B., Maier, H. R., & Lambert, M. F. (2005). Calibration and validation of neural networks to ensure physically plausible hydrological modeling. Journal of Hydrology, 314(1-4), 158-176.

Kollat, J. B., Reed, P. M., & Wagener, T. (2012). When are multiobjective calibration trade-offs in hydrologic models meaningful?. Water Resources Research, 48(3).

Pokhrel, P., Yilmaz, K. K., & Gupta, H. V. (2008). Multiple-criteria calibration of a distributed watershed model using spatial regularization and response signatures. Journal of Hydrology, 418, 49-60.

Vrugt, J. A., & Sadegh, M. (2013). Toward diagnostic model calibration and evaluation: Approximate Bayesian computation. Water Resources Research, 49(7), 4335-4345.

Wagener, T., & Gupta, H. V. (2005). Model identification for hydrological forecasting under uncertainty. Stochastic Environmental Research and Risk Assessment, 19(6), 378-387.