Predicting Hydropower Generation from Generation Flow

By Rafael Almeria

Engineering Data Analysis and Modeling
Talk Outline

• Objective: What are we trying to do?
• Significance: Why are we doing it?
• Methodology: How did we solve this?
• Results: What did the analysis yield?
• Conclusion: What did we learn?
Objective

• To estimate the daily amount of power generated from a hydroelectric dam using only the daily amount of generation flow.
Significance

• Current power generation data is not a publicly known quantity.

• Energy company X stands to gain from advance knowledge of the supply of power available to competitors.

• If company Y is going to have a surplus of power then it may lower prices.

• Company X would like to know as soon as possible what Y’s supply is so that it can lower prices first and thus prevent the loss of customers to company Y.
Methodology

- Training set contains 20 years of daily data.
- 2 test sets each contain one year of data.
- $N \gg P$
- Single input (generation flow).
- Single output (power generation).
- Used a linear least squares model.
Methodology (continued)

• Used $R^2$ (coef. of multiple determination) to measure model fit.
• Used 99% confidence limits to measure the prediction success.
Results

• $R^2 = 0.89312$ which means that the model fits the data reasonably well since $R^2$ is much closer to 1 than it is to 0.

• Predictions for 2001 and 2002 are 100% successful as all of the test data fall within the 99% confidence limits.

• Plots of predictions on 2001 and 2002 data follow.
QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.
QuickTime™ and a TIFF (LZW) decompressor are needed to see this picture.
Conclusion

• The model provided good results given the measure of success and the limitations of a single available input variable.

• Unfortunately, the predictions may not be precise enough (to power marketers) since the variance of the power generation data is significant (about 5K Megawatt Hours).

• More inputs such as the number of online generators (if they were available) could possibly yield more precise predictions.