COMMERCIAL IN CONFIDENCE

MCP Errors

Report No:	01327R00025
Issue No:	01
Date:	28 January 2005

MCP Errors

Report No:	01327R00025		
Issue No:	01		
Classification:	COMMERCIAL IN CONFIDENCE		
Prepared:	**Mike Anderson**	Signed Electronically	28 January 2005
Checked:	**Jeremy Bass**	Signed Electronically	28 January 2005
Approved:	**Mike Anderson**	Signed Electronically	28 January 2005

Revision History

Issue	Date	Nature And Location Of Change
01	28/01/05	First Created

CONTENTS

1.0 INTRODUCTION

A review of MCP techniques has been undertaken [Anderson, 2004] in which it was concluded that the Matrix Method should generally be used in preference to other existing methods. A limited analysis of the uncertainty in the predictions was discussed and it was tentatively suggested that the accuracy of the MCP process is typically ~5% on wind speed and ~9% on energy yield.

The purpose of this report is to further investigate the uncertainties in the MCP process with objective of attempting to quantify them more accurately. Three areas will be investigated:

- Confidence limits of the estimated model parameters. Where a least squares regression formulation is used, determining these is well documented and rather trivial. For the Matrix Method the standard approach is not possible and recourse will be made to the "The Bootstrap Method".
- The amount (length) of concurrent data.
- Seasonal variations. In Anderson [2004] all data was used and no account was taken of its seasonality.

2.0 ANALYSIS AND RESULTS

2.1 Parameter Estimation

2.1.1 The Bootstrap Method

A powerful statistical technique, commonly known as "The Bootstrap Method", [2] will be used as it has the advantage that knowledge of the underlying process is not required. The use of the term bootstrap derives from the phrase "to pull oneself up by one's bootstrap", widely thought to be based on one of the eighteen century "Adventures of Baron Munchausen", by R.E. Raspe. The Bootstrap Method uses the actual data set, with its N data points to generate any number of synthetic data sets, also with N data points. The procedure is simply to draw N data points at a time with replacement from the original data set. Because of the replacement, you don't simply get back your original data set each time but you get sets in which a random fraction of the original points, typically $\sim 1/e \approx 37\%$, are replaced by duplicated original points. These new sets are then subject to the same analysis, as was performed in Anderson [2004] to estimate the variation in the predicted mean wind speed and energy yield to infer the errors in the parameter estimation.

2.1.2 Results

For each of the 106 data sets, as described by Anderson [2004], 100 bootstrap realisations were created. From each realisation the predicted historical normalised bias wind/energy estimates were calculated and their standard deviation formed. The results of this analysis, presented as a function of the total length of data each set is shown in Figure 2.1. It should be noted that this error stems purely from fitting of the data and it is not indication of the total accuracy of the whole MCP process.

Presented in Figure 2.2 are the normalised errors of the energy yield plotted against the wind speed. Not surprisingly a strong linear relationship exists with a gradient approaching two, which is slightly higher than would have been expected.

Presented in Figure 2.3 is the normalised wind speed error plotted against the product moment correlation coefficient. Despite a rather poor overall correlation, the scatter suggests a weak tendency for the % error to decrease with increasing correlation.

2.2 Data Set Length and Seasonality

For each data set comprising *N* points , 100 blocks of *M* consecutive points were sampled and the normalised bias error of the predicted historical wind/energy error estimates calculated, subject to the constraint that $M < 0.8N$, for each block. For example a data set containing 20,000 points would have blocks *M* points ranging from 3,000 to 16,000 sampled from it. As in Anderson [2004] the normalised bias error for i^{th} block is defined as:

$$\mathrm{E}_{V,i} = \frac{\overline{V}_{h,pred} - \overline{V}_{h,act}}{\overline{V}_{h,act}} \quad [1]$$

where $\overline{V}_{h,pred}$ and $\overline{V}_{h,act}$ are the predicted and actual mean wind speeds for the historical period. The mean normalised bias error across all blocks of data was then calculated along with its standard deviation, (see Figure 2.4). The prediction error was then calculated by deriving the standard deviation of the normalised bias error for all data sets of *M* points in length taken from all 106 data sets. The results of this analysis are shown in Figure 2.5. From this figure can be observed that the wind speed prediction error for less than 1 year of data is greater than 4% but decreases asymptotically to about 2% for more than 2 years of data. It is also interesting to compare this Figure 2.5 with Figure 2.1 and note that the error resulting from the parameter estimation is substantially smaller. This suggests that the error in the prediction is mainly due to changes in the data between the concurrent and historical periods.

From same data the effect of seasonality can be investigated by taking the average of the standard deviation of the normalised bias error for all data sets of *M* points in length. The results of this analysis are shown in Figure 2.6 and it can be clearly seen that the error is a minimum at integer multiples of a year of data.

2.3 Total Wind Speed Prediction Error

Total wind speed prediction error will be a combination of the three errors discussed above. Shown in Figure 2.6 are the errors from parameter estimation and seasonality plotted against the error from data set length. This figure tends to indicate, not surprisingly that the errors are correlated and therefore we shall assume that the total is the arithmetic sum of the three components. see Figure 2.7. Also shown is an empirical fit to the data of the form:

$$Error(\%) = 375 \big/ \sqrt{N} \quad [1]$$

where *N* is the number of concurrent hours.

3.0 CONCLUSIONS

An analysis of the errors in the MCP process has been investigated. It is concluded that an empirical estimate of the total error in the wind speed estimate is used.

4.0 REFERENCES

[1] Anderson, M.B., 2004, “A Review of MCP Techniques”, RES Report 01327R00022, Issue 03.

[2] Davison, A.C and Hinkley D.V., 1997, “Bootstrap Methods and Their Applications (Cambridge Series in Statistical & Probabilistic Mathematics)”, Cambridge University Press.

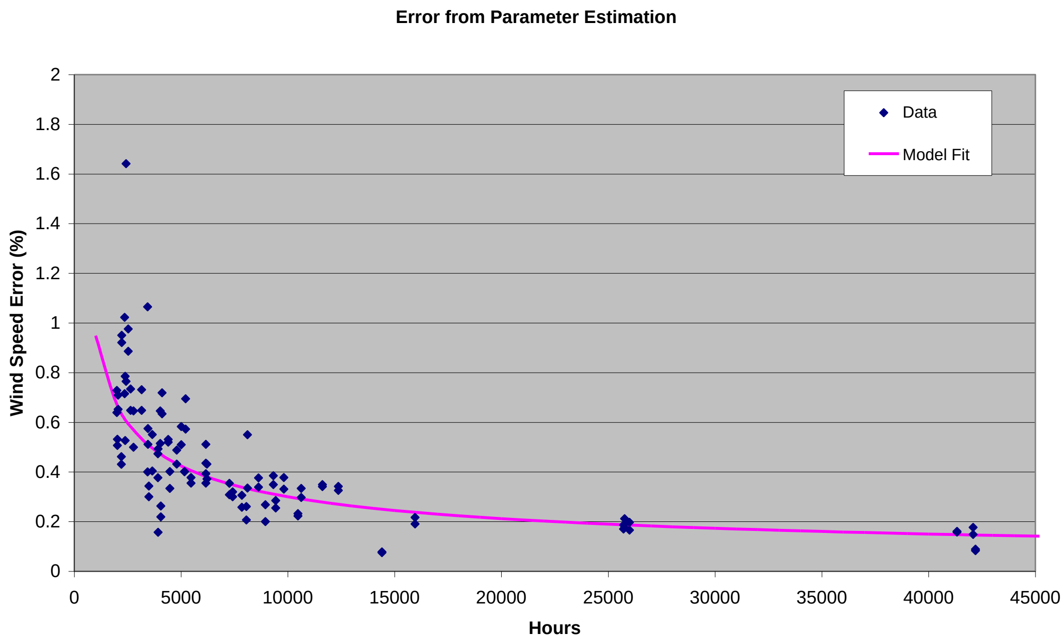

Figure 2.1 – Wind Speed Error from Parameter Estimation

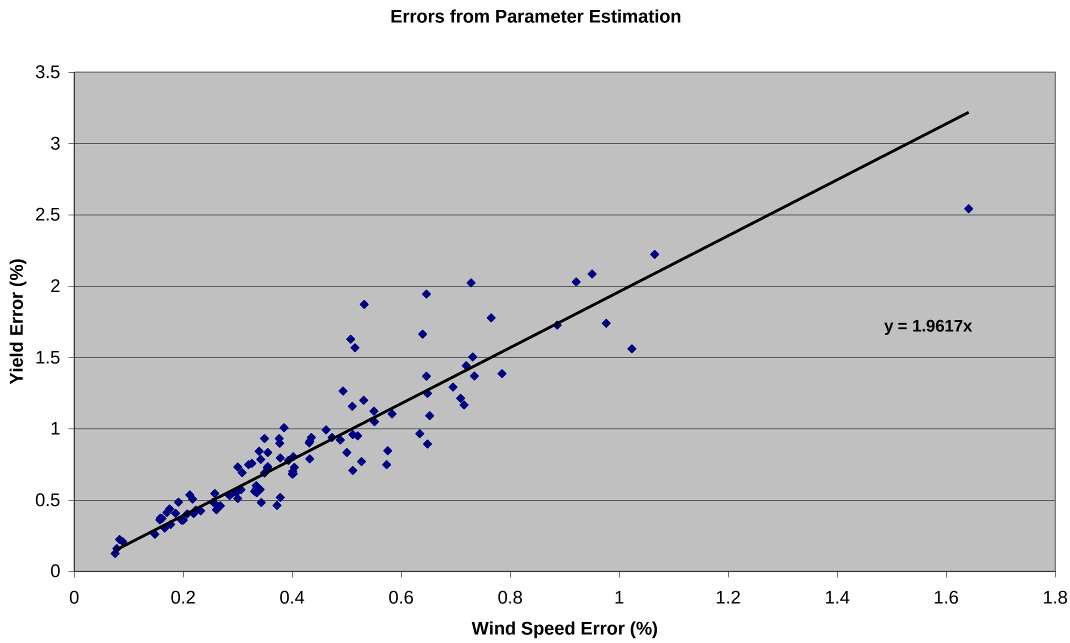

Figure 2.2 – Wind Speed and Energy Yield Errors from Parameter Estimation

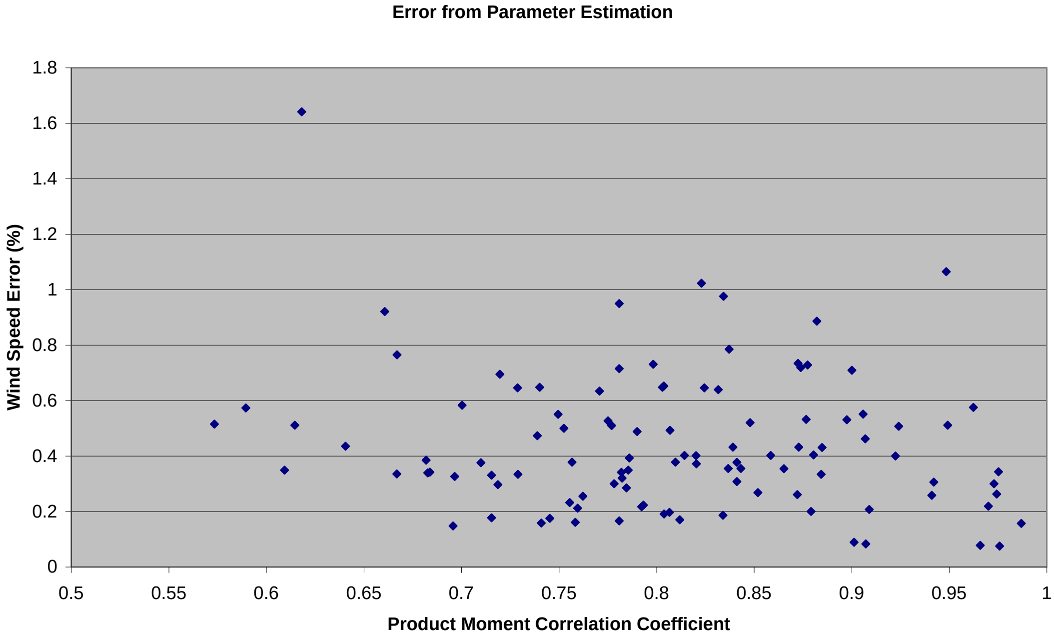

Figure 2.3 – Wind Speed and Product Moment Correlation Coefficient

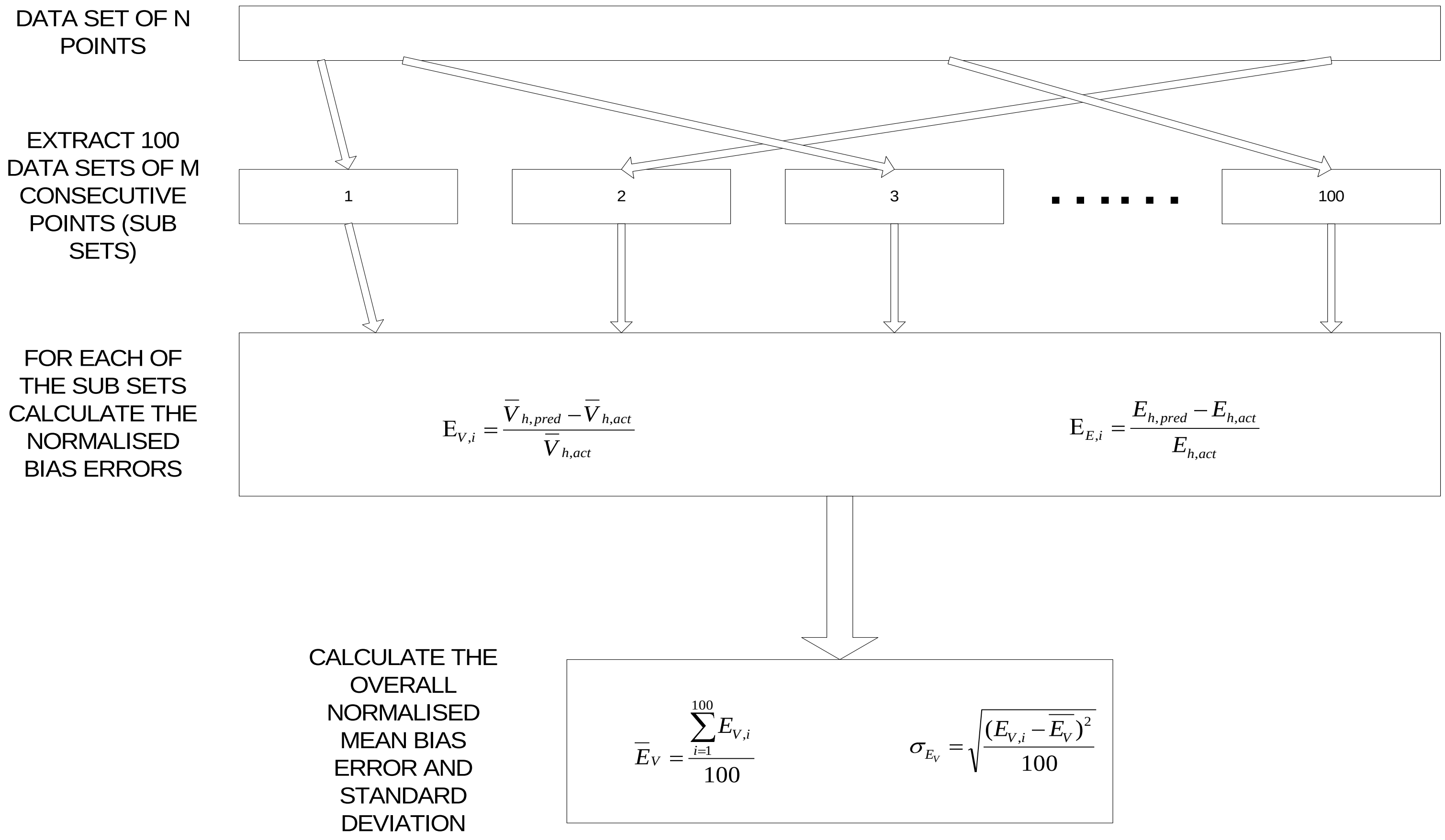

Figure 2.4 – Schematic Diagram of the Calculation Process

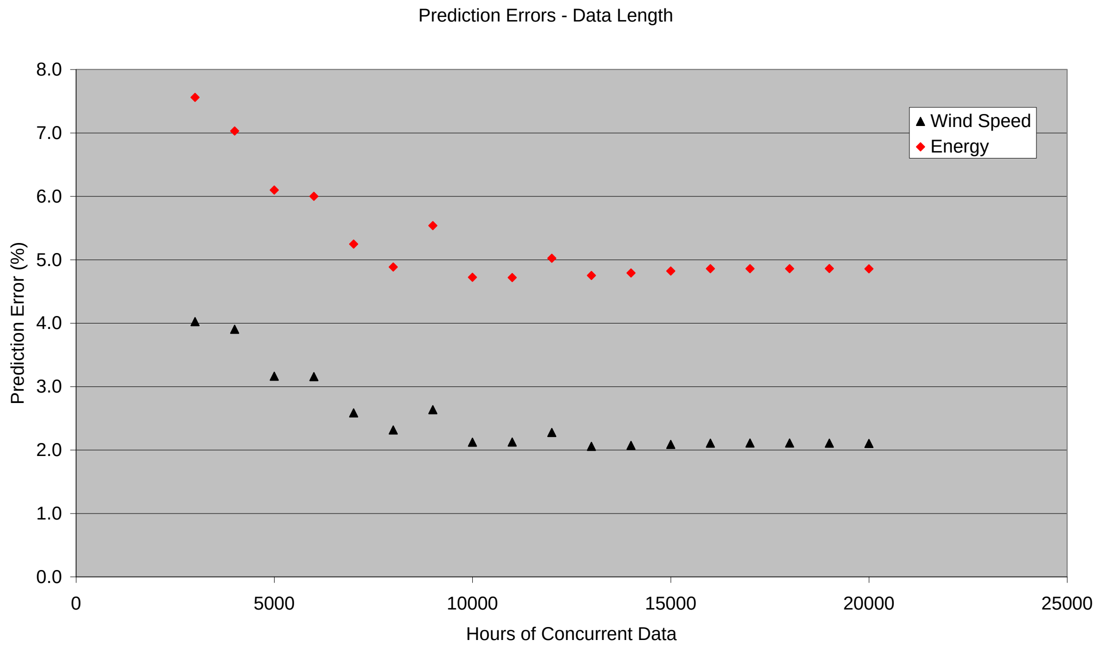

Figure 2.5 – Variation of Prediction Error with Data Length

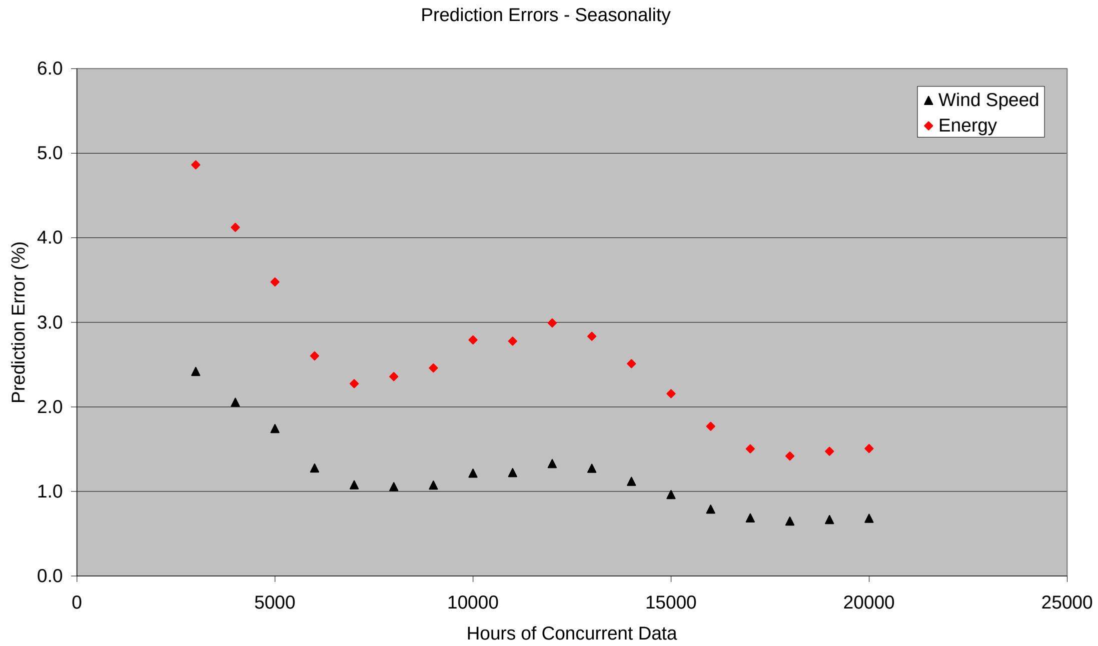

Figure 2.6 – Variation of Prediction Error with Data Length Due to Seasonality

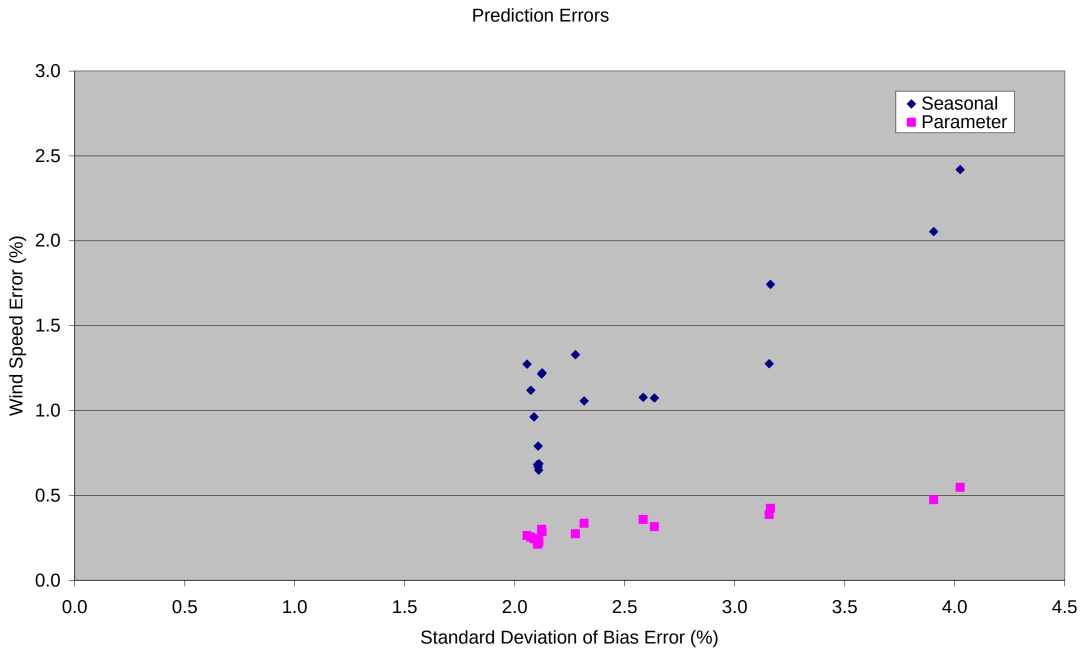

Figure 2.7 – Relationship Between Prediction Errors

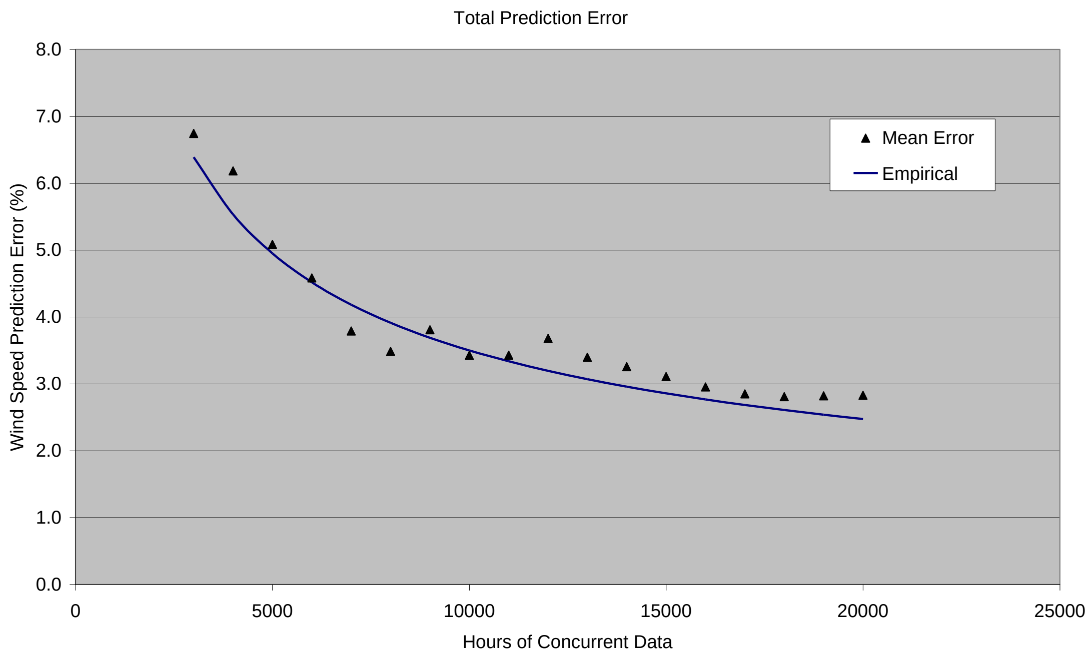

Figure 2.8 – Variation of Total Prediction Error with Data Length