# 3.5 Influence of Measurement errors

Every measurement includes measurement errors. The measurement errors can be of different magnitude. Until now the training data that were looked at did not include measurement errors. In this section noise was included in the training data.

The training samples for this test result out of a back calculation of (Eqn. 3.1-11). The effectiveness was first calculated then the outlet temperature of the water in the outlet flow was calculated with (Eqn. 3.1-12). The measurement error (Fig. 3.5-1) was artificially included here. An random error in the range of plus or minus *0.25K* was included for half of the samples the other half of the samples included a random error in the range of plus or minus *0.5K*. This is a distribution of error that can come close to an error distribution that can be observed in reality. Samples for which the difference between the temperature of the air and the water smaller than *2K* were skipped from the data set. It was assumed that these samples were taken at conditions that are non-operating and therefore not interesting. This assumption had to be preformed because the error of maximal 0.5K in the outlet temperature can cause the calculated effectiveness to yield results that are not reasonable.

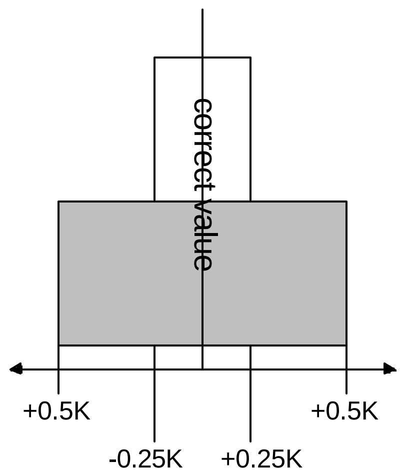


*Fig. 3.5-1 Error distribution*

In the following the given number of samples still include the trainng samples that were later skipped. The number of skipped samples for 150 trainng samples is about 15. The number of skipped samples for 500 training samples is 48.

In (Fig. 3.5-2) the effectiveness was calculated for 150 unequally spaced training samples including noise as described in (Fig. 3.5-1). The smoothness parameter resulted again out of the wiggle-method. One more inflection was allowed in the search for the smoothness parameter in the wiggle method. Measurement errors make it harder to find a value for the smoothness parameter that accommodates for smoothness but does not let the precision become bad. The precision of the prediction is worse than for the previous examples including no noise.

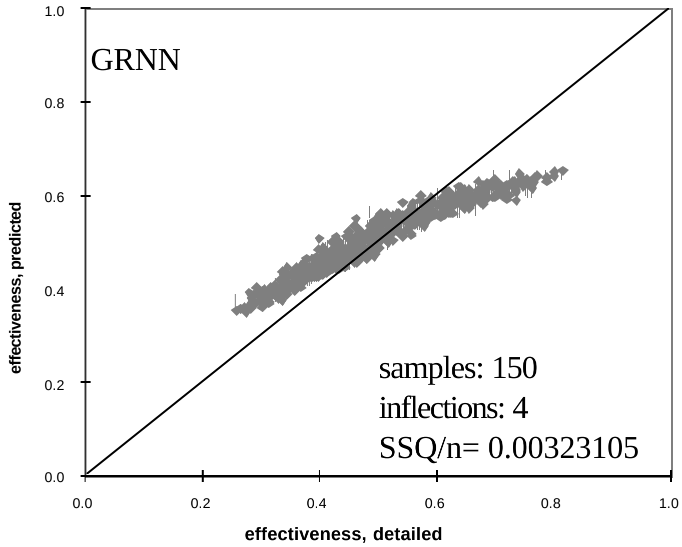


*Fig. 3.5-2 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 150 unequally spaced training samples including noise, ignoring knowledge about simple model*

In (Fig. 3.5-3) the simple model was used as an underlying function. The precision is better than for the example without using the underlying function but not very much. The smoothness parameter was selected using the wiggle method again. One more inflections was allowed for the algorithm in the wiggle method to count for the noise in the data.

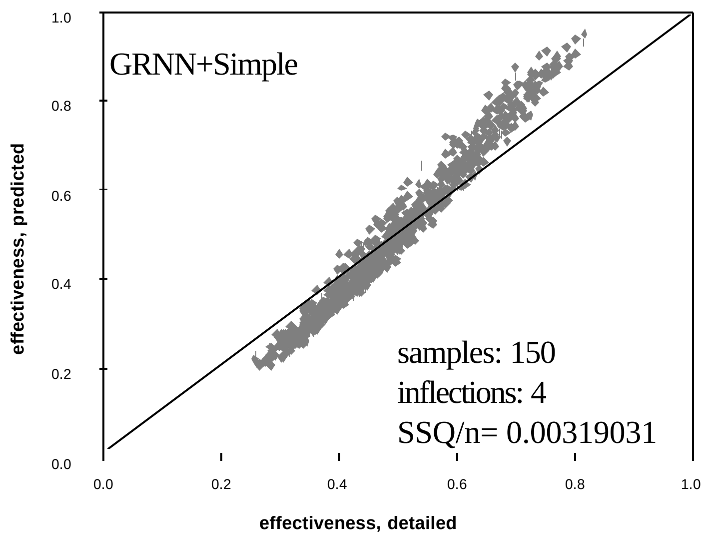


*Fig. 3.5-3 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 150 unequally spaced training samples including noise, using the approach to correct simple model*

In (Fig. 3.5-4, Fig. 3.5-5) predictions were performed using the smoothness parameter from the previous predictions. The training samples were the same as for the prediction in (Fig. 3.5-2, Fig. 3.5-3). These plots include the real effectiveness calculated with the same model the training samples were calculated. The two predictions using the simple model or not are plotted in each plot. The results for the simple model are included in these plots too.

The shape of the curve representing the real effectiveness and the prediction using no simple model is more pronounced for these examples including noise than for the previous examples that did not include noise. The predictions using the simple model perform a lot better than the approach using no simple model. The shape of the real effectiveness and the predicted effectiveness are a lot closer. The reason for this is the fact that the simple model approximates the real effectiveness (Fig. 3.5-4, Fig. 3.5-5).

The results for the predictions deviate significantly from the real effectiveness, especially the prediction of the effectiveness using no simple model as an underlying function. The simple model again supports the prediction.

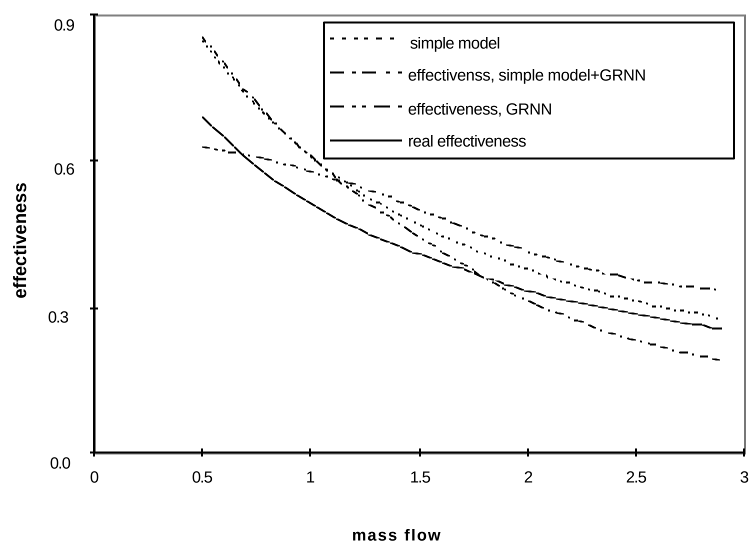


*Fig. 3.5-4 Comparison of prediction using simple model or not. 150 unequally spaced training samples including noise, $T_{i,water}$ =360K, $T_{air}$=300K, velocity=12.5m/s*

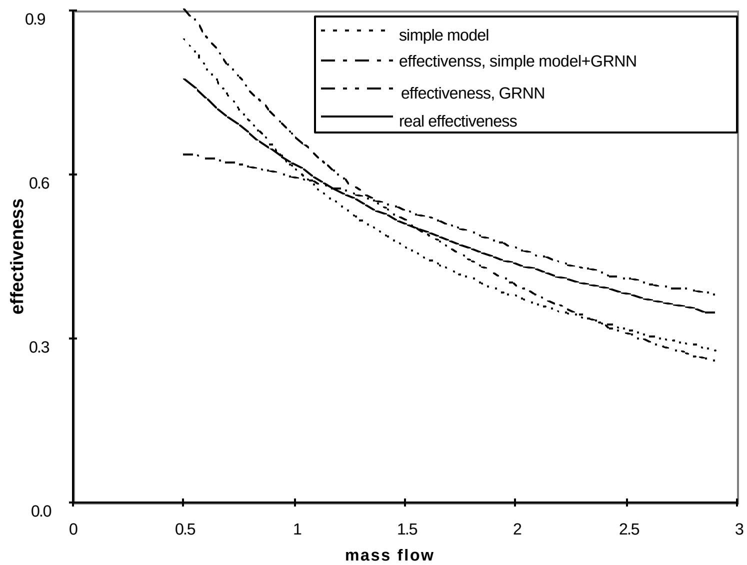


*Fig. 3.5-5 Comparison of prediction using simple model or not. 150 unequally spaced training samples including noise, $T_{in,water}$ =360K, $T_{air}$=300K, velocity=25m/s*

150 samples are not very many samples. The same analysis is repeated for 500 training samples that are unequally spaced and include noise like the 150 training samples. Again samples were skipped for which the difference between the air temperature and the water inlet temperature was closer than *2K*. The precision of the prediction is better for the 500 training samples than for only 150 training samples (Fig. 3.5-6, Fig. 3.5-7). The difference in precision is not very big between the example using 150 training samples or 500 training samples. The distribution is not as wide for the prediction using 500 training samples.

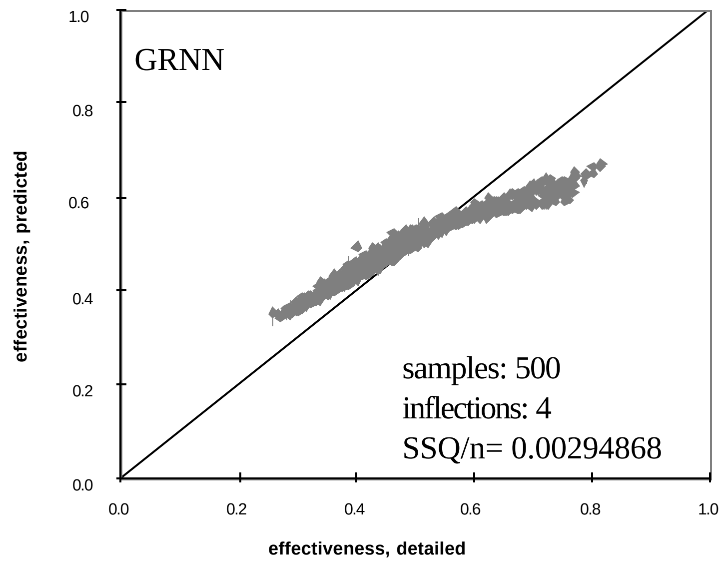


*Fig. 3.5-6 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 500 unequally spaced training samples including noise, ignoring knowledge about simple model*

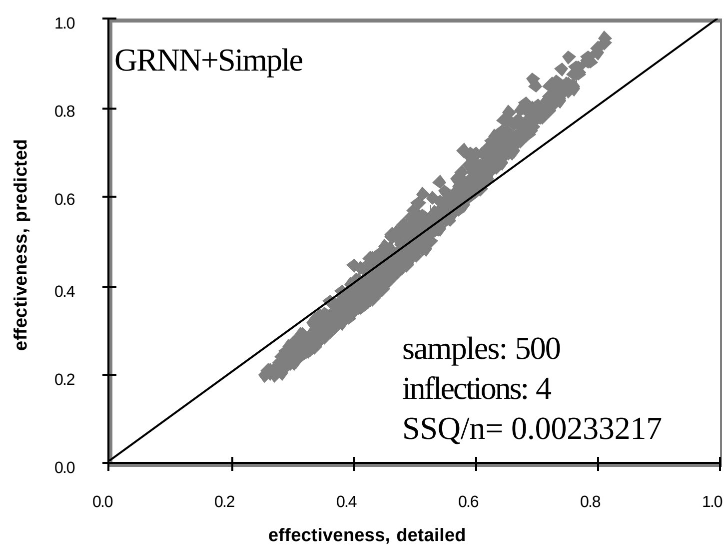


*Fig. 3.5-7 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 500 unequally spaced training samples including noise, using the approach to correct simple model*

In (Fig. 3.5-8, Fig. 3.5-9) predictions are shown again for 500 training samples for air temperature of *300K*, Water inlet temperature of *360K* and for a velocity of air for (Fig. 3.5-8) of *12.5m/s* and for (Fig. 3.5-9) the velocity of the air was *25m/s*. As it was shown already in (Fig. 3.5-6) and (Fig. 3.5-7) the precision improved a little, but not as much as it was hoped for.

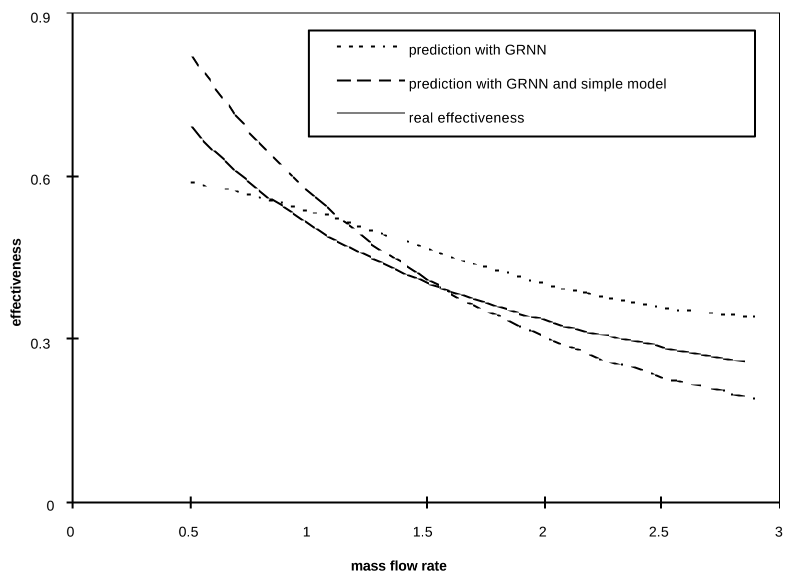


*Fig. 3.5-8 Comparison of prediction using simple model or not. 500 unequally spaced training samples including noise, $T_{in,water}$ =360K, $T_{air}$=300K, velocity=12.5m/s*

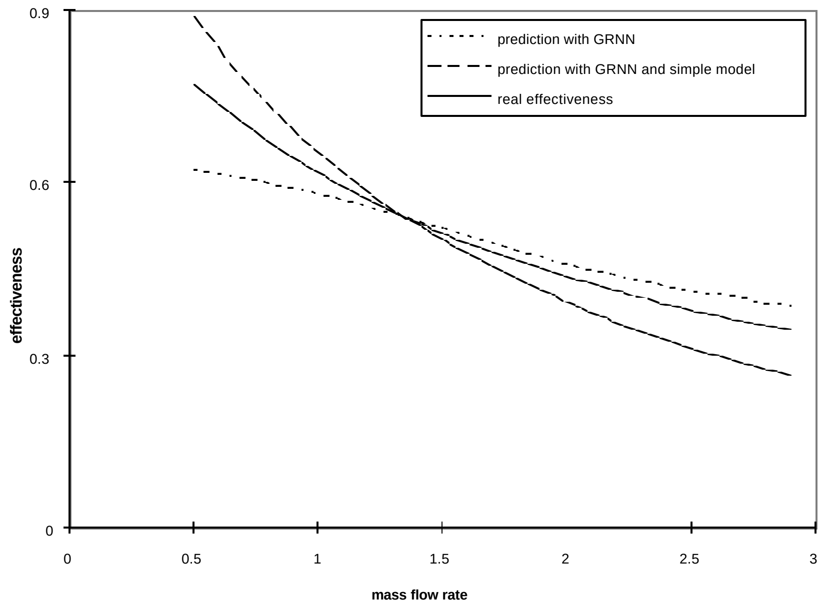


*Fig. 3.5-9 Comparison of prediction using simple model or not. 500 unequally spaced training samples including noise, $T_{in,water}$ =360K, $T_{air}$=300K, velocity=25m/s*

As it was previously discussed, the influence of the number of inflections that are allowed in the wiggle method is big. The influence of increasing the number of the allowable inflections was tested again for unequally spaced data including noise. The allowable number of inflections was increased to eight. The same predictions with a different smoothness parameter are performed again, that were previously performed already.

The results are shown in (Fig. 3.5-10) and (Fig. 3.5-11) for 150 unequally spaced training samples that are including noise. Again the real effectiveness and the predicted effectiveness using the simple model or not are plotted for the conditions that were already used before. The amount of wiggles that the curve now include, (Fig. 3.5-10, Fig. 3.5-11) especially the curve using no underlying function, is drastic. The shape of the curve for the use of the simple model is not that drastically bad. The simple model supports the prediction very much because the simple model itself is not bad. The correction has only a smaller influence on the shape and the precision than for the approach using no simple model.

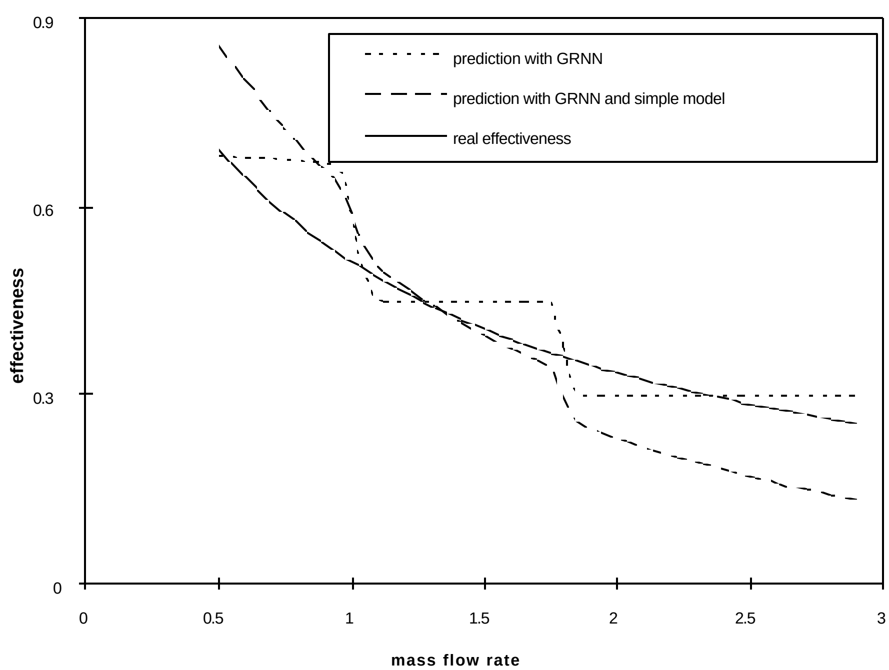


*Fig. 3.5-10 Comparison of prediction using simple model or not. 150 unequally spaced training samples including noise, allowing more inflections, $T_{in,water}$ =360K, $T_{air}$=300K, velocity=12.5m/s*

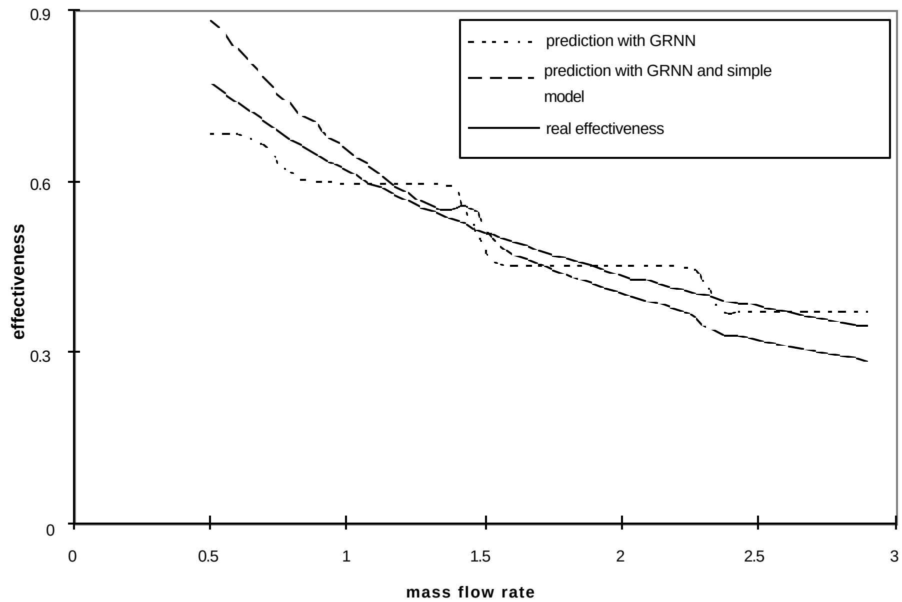


*Fig. 3.5-11 Comparison of prediction using simple model or not. 150 unequally spaced training samples including noise, allowing more inflections, $T_{in,water}$ =360K, $T_{air}$=300K, velocity=25m/s*

The results for the plot of predicted effectiveness versus real effectiveness are shown in (Fig. 3.5-12, Fig. 3.5-13). The results for both approaches are scattered.

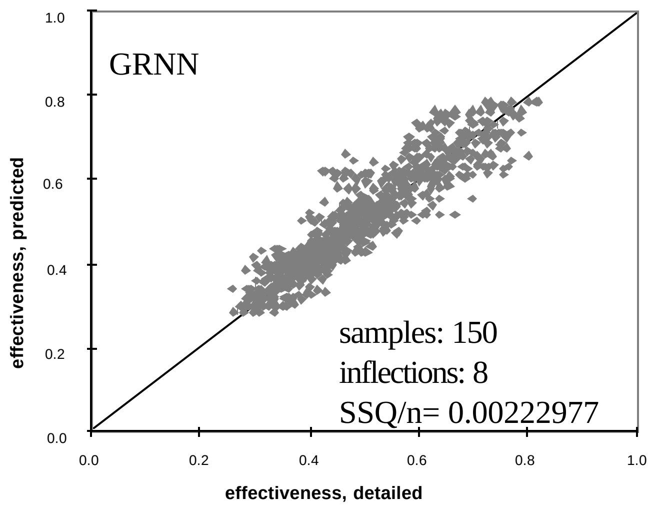


*Fig. 3.5-12 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 150 unequally spaced training samples including noise, allowing more inflections, ignoring knowledge about simple model*

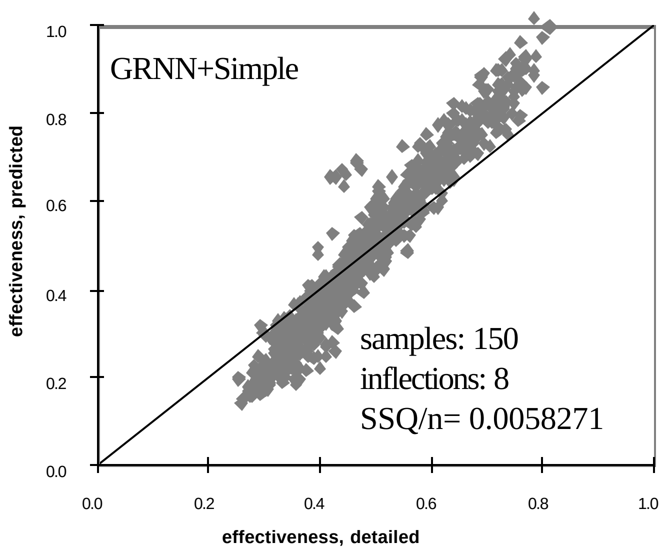


*Fig. 3.5-13 Predicted effectiveness versus effectiveness, calculated with the detailed model for 999 randomly picked points for 150 unequally spaced training samples including noise, allowing more inflections, using simple model*

Allowing five more inflections in the wiggle method does not increase the precision of the prediction by a lot. The sum of Squares devided by the number of predictions for the increase in the number of allowable inflections for the problematic cases of unequally spaced data and data including noise decreased by about 25% to 50%. The precision got a little better but the predcition includes more wiggles now. Again the tradeoff between the precision and the smoothness of a prediction has to be made.

# 3.6 Chapter Summary

In order to judge the quality of a prediction it is not only important to look at the precision of a prediction. It is as well important to look at the smoothness of the prediction. In (Table 3.6-1) a summary of the results is given. The sum of squares (SSQ) divided by the number of compared points resulted out of the plots from effectiveness of the prediction versus the real effectiveness.

| ***Number of samples*** | ***spacing*** | ***including simple model or not*** | ***number if inflections in wiggle method*** | ***SSQ/n*** |
|---|---|---|---|---|
| -- | -- | simple model | --- | 0.004103 |
| 81 | equal | no | 2 | 0.001443 |
| 81 | equal | yes | 2 | 0.000694 |
| 81 | equal | no | 6 | 0.003699 |
| 81 | equal | yes | 6 | 0.000735 |
| 150 | unequal | no | 3 | 0.00614425 |
| 150 | unequal | yes | 3 | 0.00090529 |
| 500 | unequal | no | 3 | 0.00297433 |

| | | | | |
|---|---|---|---|---|
| 500 | unequal | yes | 3 | 0.00061012 |
| 9+9 | equal | no | 2+2 | 0.0035317 |
| 9+9 | equal | yes | 2+2 | 0.00140332 |
| 150 | unequal | no | 8 | 0.00201946 |
| 150 | unequal, noise | no | 4 | 0.00323105 |
| 150 | unequal, noise | yes | 4 | 0.00319031 |
| 500 | unequal, noise | no | 4 | 0.00294868 |
| 500 | unequal, noise | yes | 4 | 0.00233217 |
| 150 | unequal, noise | no | 8 | 0.00222977 |
| 150 | unequal, noise | yes | 8 | 0.0058271 |

*Table 3.6-1 Sum of squares depending on the conditions used for the prediction*

Each of the methods discussed performs more or less well. Using an underlying function certainly helps a lot to accommodate for the individual needs. The use of GRNN in addition to the simple model definitely has some potential. It is for example not necessary to built up a very complicated model to perform simulations. A simpler model is enough if there are data available for GRNN to learn.

If well designed experiments are possible it is desirable to use this kind of information using equally spaced data versus a larger number of uneqally spaced data. Smaller sets of training samples allow the procedure to perform a lot faster than for bigger training sets. In (Table 3.6-1) it can be seen that the quality of the prediction for 81 equally spaced training samples is about as good as for 500 unequally spaced training samples.

If enough knowledge is available to come up with proof that certain variables are sufficiently independent, then the training samples can be reduced even more and the procedure is getting faster. This is only possible if experiments under very well designed conditions are possible.