

Fast Estimation of the Relation between Aggregated Train Power System Information and the Power and Energy Converted

Lars Abrahamsson and Lennart Söder
Electric Power Systems
Royal Institute of Technology
Stockholm, Sweden

Abstract—Transports on rail are increasing and major investments in the railway infrastructure, including the Railway Power Supply System (RPSS), are expected. The future railway power demands are naturally not known for certain. The more remote the uncertain future, the greater the number of scenarios that have to be considered. Large numbers of scenarios make time demanding simulations unattractive.

The aim of this paper is to present a fast approximator that uses aggregated RPSS information. Since the electrical and mechanical relations governing an RPSS are quite intricate, an approximator based on Neural Networks (NN), is applied. This paper presents a design suggestion for an NN estimating the power and energy flows through each converter station, given RPSS data and levels of train traffic. Even if the future usage of the NN is investment planning, the modeling of such an approximator has a value in itself concerning the understanding of the relations between RPSS and train traffic.

I. INTRODUCTION

For environmental and economic reasons, in Sweden and the rest of Europe, both personal and goods transports on rail are increasing. Therefore major railway infrastructure investments are expected. An important part of this infrastructure is the RPSS. One phase low frequency AC RPSSs are normally connected to the ordinary 50 Hz power system through frequency converters, see Figure 1.

The converters may be merely used as power sources, as in Sweden for example, whereas in other countries the railway administrations produce electricity on their own. The railway overhead contact line from which the locomotive extracts electric power is called the catenary. The original type of catenary, Booster Transformer (BT), has a relatively high impedance. The RPSS can be strengthened by replacing BT catenaries with Auto Transformer (AT) ones, placing the converter stations closer to each other, or connecting a High Voltage (HV) transmission line in parallel to the catenary system.

RPSSs are changing all the time. When a train moves, the impedances between it, other trains, and the feeding points changes. Both the active and reactive power demands of the locomotives may vary with the slopes of the railway, the train weights, the desired train velocity, etc.

When considering future possible RPSS investments, for each possible expansion alternative, many different situations of railway operation causing different loading on the RPSS

have to be studied. This is a kind of transmission expansion planning. Similar approaches can be found in the references [1]–[3]. In the case of the railway, however, the locations of feeding points are up to the RPSS administrator, and not to the actors on the market as in the case with the public grid.

For each possible investment alternative, the power grid calculations have to be performed fast. There are a number of simulators available, e.g. TPSS (Train Power System Simulator) [4] and also commercial software such as TracFeed Simulation [5]. These are however not fast enough, though, when several thousands of cases have to be studied. The intention of this paper is to propose a solution of how to design an approximator that rapidly estimates some of the properties of an RPSS which are considered important in the studies of future expansion alternatives. This approximator uses aggregated results from many load flow calculations in a fast way.

The two most important consequences of train traffic related to the state of the RPSS, having the railway operation costs in mind, are:

- The maximal power consumption and the energy consumption of the grid, preferably divided up by individual converter stations, which is studied in this paper.
- The impact of catenary voltages on the minimal traveling times, studied in [4], [6], [7].

Since the electrical and mechanical relations of an RPSS are quite intricate, an approximator of the black box kind is used. Therefore, NNs (Neural Networks), which basically are a kind of nonlinear predictors, were selected as a suitable approximator type. An NN with too many inputs and outputs may need a tremendous number of training cases in order to become reliable and general [8]. Therefore TPSA just uses a few aggregated inputs and outputs.

II. THE MODEL OF THE APPROXIMATOR

Since a detailed description of RPSS modeling in general, and the modeling of TPSS in particular can be found in [4] no such details will be presented in this paper. Moreover, in [4], [6], [7] detailed backgrounds to the need for fast approximators as well as thorough descriptions of the first part of the approximator, TPSA-V (Train Power System Approximator Velocity) are present.

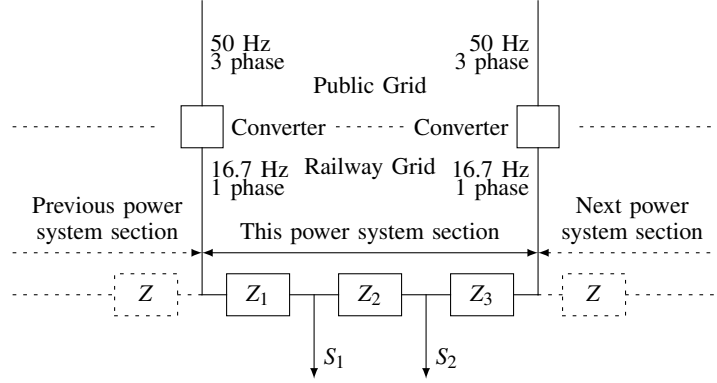


Figure 1: A section of the RPSS, illustrated as an electric circuit.

TPSA-V tells, in aggregated form, how fast an extra train on an RPSS section can go on average – given the RPSS technology used on the section and the already existing traffic levels there. In this paper, the second part of TPSA, which is called TPSA-E (Train Power System Approximator Energy), is presented. In TPSA-E, energy consumption and peak apparent power usage are estimated. In the following section, the choices of aggregated inputs and outputs are described and motivated.

A. Choices of Approximator Inputs and Outputs

1) *The Aggregation of Inputs:* The TPSA-E inputs can be classified as either parameters giving rise to power consumption, informing of the RPSS impedances, related to the relative locations of the loads, or combinations of those.

In the TPSS simulations, which results are used for the aggregation of training data for the inputs and outputs to the NNs building up TPSA, the trains are allowed to stop only at the locations of converter stations. In addition to that, the simulations are restricted to one type of trains traveling in one direction, and finally, all the trains try to maintain the same velocity. These are all facts that reduces the number of possibly needed input variables. The exact details of the simulation setups can be found in the full report [4].

The intention is that TPSA should be able to work as a traffic constraint with respect to the RPSS for software that is planning for the future train traffic. This also restricts the choices of possible inputs to the TPSA. For example, in a train time table, only positions and locations are known, therefore only the average velocity can be calculated. For more details about these ideas, see [4], [6], [7].

Two binary TPSA inputs, $\rho_{catenary}$ and η , describes the RPSS technologies used and also which of the four NNs in Figure 2 to use. When $\rho_{catenary}$ has the value 1 one of the NNs representing AT catenaries are active, and if it equals 0 one of the two BT catenary NNs are active. The second binary variable, η , tells whether or not an NN representing an RPSS with an HV line should be used. The remaining six TPSA-E inputs are NN inputs.

The reason not to let inputs that cannot be modeled as continuous variables be inputs to the NNs have been

explained by [8]. Shortly, the kind of NNs used in this paper, backpropagation ones [9], [10], can be very good approximators of continuous functions but are comparatively bad at coping with discrete inputs of a more classifying kind. In the following paragraphs, the NN inputs are listed, described and motivated.

a) *The Length of the Power System Section:* is the distance between the pair of converters defining the borders of the RPSS section. This kind of division of the power grid is motivated in [4]. The distance gives information of the RPSS impedance. And in this paper it is also always the distance traveled for all trains, not necessarily within the time window of the studied case though, see Section II-B.

b) *The Average Inclination:* of the RPSS section. The average inclination gives information of the net potential energy consumed by the trains traveling through the section.

c) *The Standard Deviation of the Inclination:* of the RPSS section. The average inclination is not expected to always be good enough. It would for example equal zero for both flat ground as well as for a rail section with 20 per mille uphill half the section and 20 per mille downhill the remaining half. The standard deviation is a measure of how much the inclinations fluctuate, which will influence the consumed electric power of the trains.

d) *The Average Number of Trains:* on the section is also calculated. The number of trains is important because the more trains, the greater the need for electricity. The average number of trains is calculated as total train traffic time divided by the length of the time window. In this study, all time windows have the same size of 17 minutes, for details see Section II-B.

e) *The Average Velocity of the Trains while Driving:* on the studied RPSS section, within the time window is also measured. This parameter is used because higher train speeds means greater power consumption. The average velocity is measured only when trains are in service.

f) *The Relative Average Locations:* of the trains in traffic on the section within the studied time window, $\rho_{location}$. This parameter is supposed to be helpful for the approximator to determine how much of the by the trains consumed electric power are taken from each of the two converter stations. A $\rho_{location}$ valued less than 0.5 indicates that the train centroid is

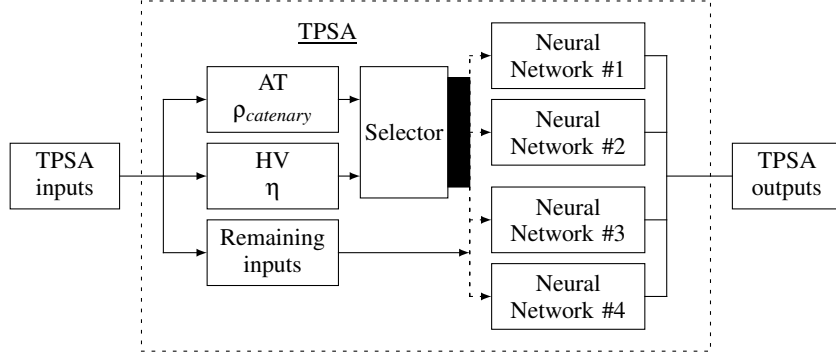


Figure 2: There are four available NNs, AT and BT catenaries, which can be either connected to or without an HV line, are allowed.

closer to the left converter station, and conversely, a parameter value between 0.5 and 1 means that the trains most of time is located closer to the right-hand side converter station.

2) *The Aggregation of Output:* There are six different outputs studied in this paper. The list below consists of only three items, but since there are two converter stations and it is desirable to be able to predict which of the converter stations that are loaded at a certain level, this number must be doubled.

a) *The Peak Apparent Power:* flow through each converter station within the time window is considered to be important for dimensioning reasons. The higher the capacity needed of the RPSS, the higher the investment cost.

b) *The Peak Six-minute-average Apparent Power:* flow through each converter station within the time window is another common dimensioning measure.

c) *The Consumed Electric Energy:* by each converter station is of importance because the more energy that is consumed, the more the railway administrator has to pay to the power companies.

B. The Time Window Size

First of all, TPSA-E does, unlike TPSA-V, consider all trains equally important. Therefore, the sizes of the time windows used for aggregation of TPSA inputs and outputs cannot be defined by the action of some specific train, they have to be predefined and preferably equal in size.

In theory, when having the future use of TPSA as a tool for calculating operation costs of the railway, the time windows could be pretty big, orders of magnitude like days. In reality, however, the aggregated data is created from simulation results of more detailed kind. As indicated earlier in this paper, the NNs, that TPSA is made of, has to be trained with aggregated data as inputs and outputs. Creating a variety of simulated cases is a necessity if the approximator should be reliable for all sorts of inputs. The longer the cases are, the more time demanding they will be to simulate also. Therefore, the time windows cannot be too long. Moreover, even more time is saved if already existing simulations can be used. So the time windows of this study is determined by the durations of the shortest traffic studies in the 400 simulations made for the studies of TPSA-V presented in [4], [6], [7]. The shortest of

these simulations represented 17 minutes of RPSS operation. Here, the RPSS is considered to be in operation for all time steps where power flows through the converter stations in the system. Sometimes the power flows are zero in the first time step, this happens if all trains are assumed to be standing still in the beginning. In such cases, time step one is still considered as one when the RPSS is in operation.

TPSS determines train positions and velocities in time step $t + 1$ by usage of the load flow results from time step t . The load flow calculations gives the accelerations, that in turn gives the velocities and the positions for the following time step.

Many of the existing 400 simulations also represent RPSSs in operation for much longer than so – there are cases enduring up to 99 minutes. Since it would be a waste just using 17 minutes of each simulation for the creation of the training cases for the four different NNs, each simulation is divided into $\lfloor \frac{\tau}{w} \rfloor$ different time windows, where τ is the simulated time, and w is the length of the time window. That means that the remaining pieces of the simulations that are shorter than 17 minutes are not used in this paper for training set creation. The above described chopping up of simulations into equally sized time windows results in 283 cases for the pure BT system, 293 cases for the pure AT system, 292 cases for the BT system with parallel HV line, and 290 cases for AT with HV line. These numbers can be explained by the fact that the trains go faster the stronger the RPSSs are, so the simulations are longer for weaker systems. In the pure BT case however, which is an exception, nine out of 100 simulations did not converge [4] so the cases are a bit fewer.

C. The Neural Network Design

As in the previous descriptions of TPSA-V [4], [6], [7], also TPSA-E uses one hidden neural layer with tansig (tanh) transfer functions. The NN is also here trained with the algorithm `trainbr` [11], with an error goal of 10^{-5} and the limit in number of allowed epochs (iterations) is set to 1000. Before training, the inputs and outputs are normalized to span from -1 to 1. The output layer has linear transfer functions.

In order to determine a suitable number of neurons in the hidden layer, for each of the four different RPSS types, the aggregated data was divided into two subsets. The biggest

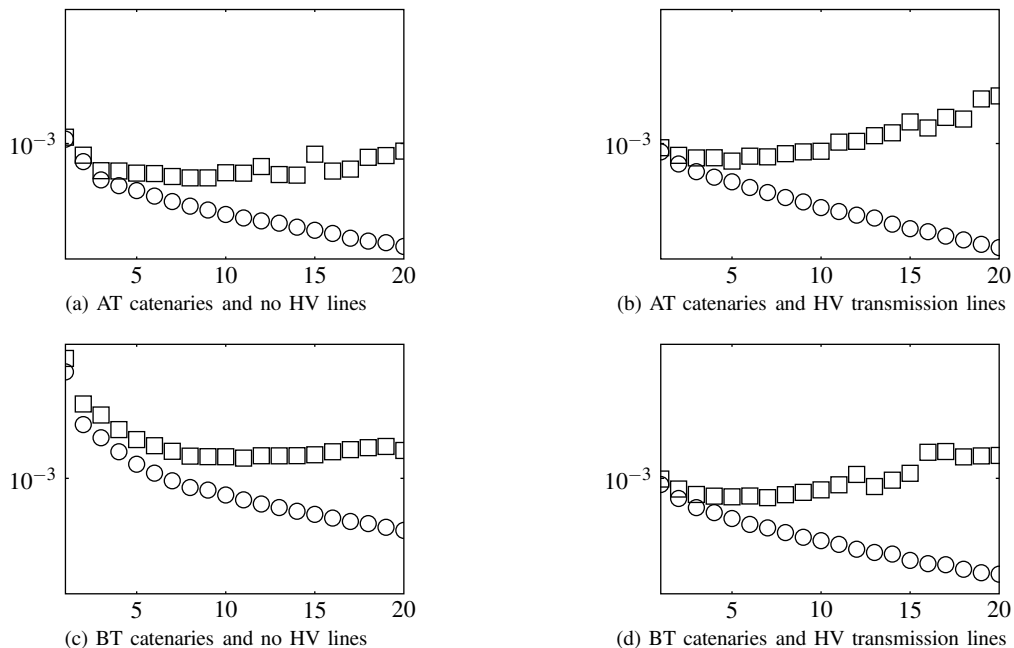


Figure 3: A study of the average errors on the training and testing sets as a function of the number of hidden neurons. Circles denote the training set and squares denote the test set. The scale of the vertical axis is logarithmic.

subset, two thirds of the cases, constitutes the training set against which the parameters of the NN are estimated. The remaining third of the cases are used as a testing set. The NN performs better against the training set than against the testing set. However, since both sets are coming from the same kind of study, one can expect that if the NN catches the trends of the training set, also the properties of the testing set will be caught. This is true as long as the NN has a limited enough number of degrees of freedom. If the NN can adapt too good to the training set, its adaptation capability against the testing set will decrease. That is in turn a sign of an NN that is not good enough at generalization, i.e. an overlearned NN. A reasonable strategy is to choose the number of neurons that gives the smallest approximation Mean Square Error (MSE) for the training set. In Figure 3 one can see an illustration of how the mean of the MSEs for the training and testing sets vary for different numbers of neurons in the hidden layer. The mean is taken over all six inputs, but also over 20 different random choices of training and testing sets in order to even out the noise and making it possible seeing the trend.

It is not easy to see in the graph, but for the pure BT case in Figure 3c, the MSEs keeps on decreasing until there are 11 neurons in the hidden layer. The difference between 8 and 11 neurons are however so small that it is considered safer to choose 8. In the pure AT case, Figure 3a, it is actually for 8 neurons the smallest MSEs for the testing set are produced. When the BT system is connected to the HV lines one can in Figure 3d see that the testing set MSEs decrease for an increasing number of neurons, up to 5. Noteworthy is however, that the MSE is at the smallest for 7 hidden neurons. In the

case of AT catenaries combined with HV lines, it is clearly visible in Figure 3b that the MSE of the testing set is the smallest for 5 neurons.

The above presented results ends up in a choice for a TPSA-E model where for the NNs representing RPSSs without the additional parallel HV line, the number of neurons in the hidden layer are 8. For the NNs representing RPSSs with the HV line, on the other hand, the number of neurons in the hidden layer are chosen to be 5.

It is noteworthy that different numbers of neurons are needed for the RPSSs with HV line present compared to the RPSSs without the HV line. Possibly this is because with an HV line present, the power does not have to be produced locally, which is the case with a catenary as the only transmitter of electric power.

III. NUMERICAL RESULTS

A. The Impact of the Relative Location of the Loads

In this part of the paper, all available data are used for the training of the NNs. The studies are focused on the possibility to allocate the correct amounts of power and energy to the two converter stations, in particular how big the gain of the input variable $\rho_{location}$ is.

First, the correlation coefficients were calculated for the input variable $\rho_{location}$ against the six outputs. The peak apparent power inflow, is in the following denoted \hat{S}^1 for the left hand side converter station, and \hat{S}^2 for the right one. Furthermore, the total energy consumed within a time window is denoted E^1 for the left hand side converter station, and

Table I: The correlations between the input $\rho_{location}$ and the six different TPSA-E outputs for the four different RPSS types.

Power Supply	The correlation coefficient, $\frac{\text{cov}(\rho_{location}, X)}{D(\rho_{location})D(X)}$, where ...					
Technology	$X = \widehat{S}^1$	$X = \widehat{S}^2$	$X = E^1$	$X = E^2$	$X = \widehat{S}_{6m}^1$	\widehat{S}_{6m}^2
BT	$1.58 \cdot 10^{-1}$	$2.95 \cdot 10^{-1}$	$1.51 \cdot 10^{-1}$	$2.91 \cdot 10^{-1}$	$1.45 \cdot 10^{-1}$	$3.08 \cdot 10^{-1}$
AT	$2.44 \cdot 10^{-1}$	$3.03 \cdot 10^{-1}$	$2.45 \cdot 10^{-1}$	$2.84 \cdot 10^{-1}$	$2.36 \cdot 10^{-1}$	$3.02 \cdot 10^{-1}$
BT+HV	$2.44 \cdot 10^{-1}$	$2.68 \cdot 10^{-1}$	$2.40 \cdot 10^{-1}$	$2.49 \cdot 10^{-1}$	$2.39 \cdot 10^{-1}$	$2.64 \cdot 10^{-1}$
AT+HV	$2.79 \cdot 10^{-1}$	$2.79 \cdot 10^{-1}$	$2.67 \cdot 10^{-1}$	$2.68 \cdot 10^{-1}$	$2.75 \cdot 10^{-1}$	$2.74 \cdot 10^{-1}$

E^2 for the right one. Finally, the maximal six-minute-average apparent power flows are denoted \widehat{S}_{6m}^1 and \widehat{S}_{6m}^2 , respectively.

One can in Table I see that the stronger the RPSS, the more equal the correlations are. The correlations are quite weak, between 0.15 and 0.30. One peculiar thing is that it seems like the power consumption always increases if the traffic moves to the right. One explanation is probably that the system is not linear enough to be explained with correlations. Another explanation could be that the slopes are heavier at the end of the track [4]. Also a too small set of training data might be an explanation.

We must investigate the differences in NN performance with and without the $\rho_{location}$ input. One could expect that it might still be a non-linear dependence there. In Figure 4 the results of the nonlinear investigations are shown. The average MSEs are compared for the six outputs and the four RPSS types both with and without the $\rho_{location}$ input. In order to even out the results, the MSE plots of Figure 4 are averaged for 20 different training situations, where the ordering of the input and output data sets are randomly changed for each of the situations.

A trend all over Figure 4, except for the case with AT catenaries and HV line parallel to it, is that it seems to be harder to predict the power and energy flows through the right hand side converter station than through the left one. No obvious explanation exists, and deeper investigations as well as greater data sets for training are needed for making such. However, one possible contribution is that there are probably less training sets where the right hand side station is heavily loaded. This is because in the simulations used, the trains depart from the left station, and arrive at the right one. Naturally, trains consume more power while accelerating compared to when braking, but that does not explain it all since accelerating takes only about three to four kilometers, and braking is done even faster [4]. A greater contribution is believed to be that many ends of the simulations are not used for aggregation into the data sets used. If for example a simulation represents 50 minutes of RPSS operation, then only the first 34 minutes will be used for aggregating two data sets of inputs and outputs, and the last 16 minutes will simply be wasted. Therefore, the risk is that less cases with high consumption on the right side will be registered. It is also possible that these differences in MSEs are just coincidences, one should bear in mind that the MSEs are still quite small.

This study indicates that it is only for RPSSs with high impedances, i.e. the pure BT RPSSs, that the information about the relative locations of the loads are needed.

B. The Performance of TPSA-E for Cases it is Not Trained for

This section exist, not primarily to benchmark or evaluate TPSA-E, but more in order to give a visual description of how it can predict the outputs both for cases simulated and interpolate and extrapolate for cases not simulated.

There are infinitely many possibilities of what to show. However, in this paper the choice on fell on showing the energy consumption of the two converter stations during a time window. All parameters but the velocity are being held constant. The RPSS simulated in Figure 5 is a pure BT system with 114 km between the converter stations, and in this type of study the inclinations are uniquely given by the inter-converter distance [4]. Since the median mean velocity over all cases on this distance and RPSS type is about 108 km/h, the reference point is chosen as this median sample from the simulations. Moreover, the number of average trains on the section are 4.50, and their average relative location is 0.473. This median-velocity case is plotted as a star in Figure 5.

The graphs are not completely realistic far from the training data points. The most obvious example is that in Figure 5b, the consumed average energy is negative for really low train speeds. This does not happen if no trains are braking regeneratively. More generally, however, the approximation is reasonable because one can expect that the higher the speeds, the higher the loads.

For the 400 simulations made for the NN training set, the trains normally are located close to the left-hand-side converter station when driving slow. This is because all trains start there. That can be one explanation to why the energy goes down so much faster for low train speeds for the right-hand-side converter station than for the left one in TPSA-E. One should also bear in mind that in the simulated material it is rare that the trains drive slowly, this is another reason to why TPSA has a harder time approximating well for low speeds.

IV. CONCLUSIONS

It has been shown that an approximator based upon NNs can estimate power and energy flows through RPSS converter stations in a fast and reliable way given the RPSS data and the intensity level of the train traffic. Moreover, adding information to the NN about the relative average locations of the loads makes it easier predicting the locations of the power and energy flows. However, it is probably only worth doing so for RPSS technologies like pure BT where the impedance is quite high and power needs to be consumed locally.

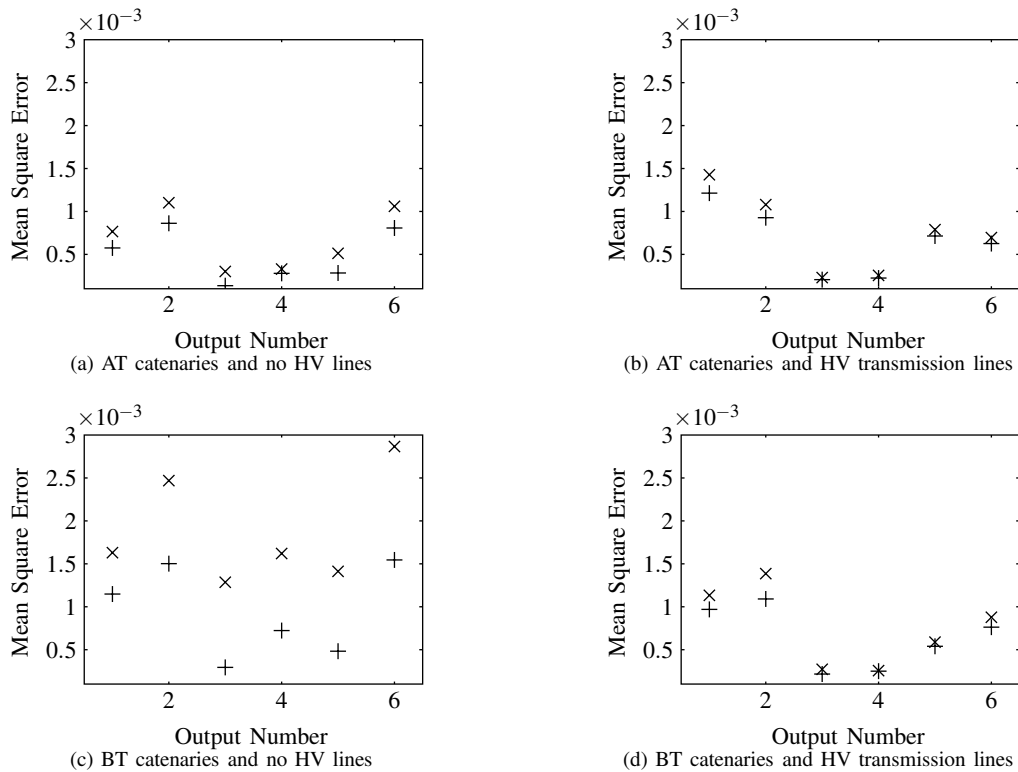


Figure 4: The average MSEs are compared for an NN given the average relative location of the loads, and for another NN that is not given it. The prior are marked \times and the latter $+$.

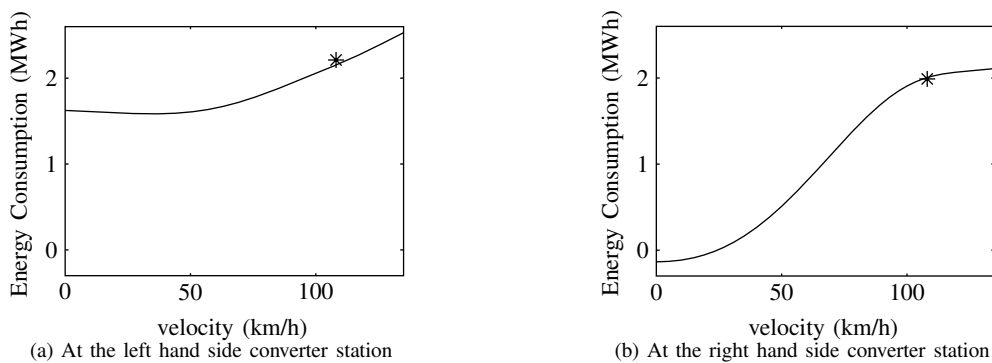


Figure 5: The consumed electric energy at both of the converter stations, the only input that is varied is the velocity.

REFERENCES

- [1] R. D. Cruz-Rodriguez and G. Latorre-Bayona, "HIPER: Interactive Tool for Mid-Term Transmission Planning in a Deregulated Environment," *Power Engineering Review, IEEE*, vol. 20, pp. 61–62, Nov. 2000.
- [2] M. O. Buygi, G. Balzer, H. M. Shanechi, and M. Shahidehpour, "Market-Based Transmission Expansion Planning," *Power Systems, IEEE Transactions on*, vol. 19, pp. 2060–2067, Nov. 2004.
- [3] B. Chennapragada, C. Radhakrishna, and R. Vallampati, "Risk-based approach for Transmission Expansion Planning in Deregulated Environment," in *Probabilistic Methods Applied to Power Systems, 2006. PMAPS 2006. International Conference on*, (Stockholm, Sweden), 1–4.
- [4] L. Abrahamsson, "Railway Power Supply Models and Methods for Long-term Investment Analysis," tech. rep., Royal Institute of Technology (KTH), Stockholm, Sweden, 2008. Licentiate Thesis.
- [5] J. Stern, "TracFeed Simulation, Reference Manual," Tech. Rep. BBSE951112-BNA, Balfour Beatty Rail, Mar. 2006.
- [6] L. Abrahamsson and L. Söder, "Fast Estimation of Aggregated Results of Many Load Flow Solutions in Electric Traction Systems," in *to be presented at Comprail 2008*, 2008.
- [7] L. Abrahamsson and L. Söder, "Fast Estimation of Relations between Aggregated Train Power System Data and Traffic Performance," submitted to: *IEEE Journal of Vehicular Technology*, July 2008.
- [8] Communication with Björn Levin, "SICS - Swedish Institute of Computer Science," Sept. 2007. Kista, Sweden.
- [9] K. Gurney, *An Introduction to Neural Networks*. CRC Press, 2003.
- [10] S. Haykin, *Neural Networks*. Englewood Cliffs: Prentice Hall, 1999.
- [11] "Neural network toolbox™," July 2008. About the "trainbr" training algorithm.