RECENT EXPERIENCES IN THE APPLICATION OF PLS IN THE STEEL INDUSTRY

Barry Lennox*, Ognjen Marjanovic^ and Erik Sandberg+

* School of Engineering, University of Manchester, Manchester, UK
^ Perceptive Engineering Ltd., Lymm, Cheshire, UK (http://www.perceptive-engineering.co.uk)
+ MEFOS, Lulea, Sweden
Barry.Lennox@man.ac.uk

Abstract
This paper provides results from a series of projects investigating the application of PLS to provide process engineering solutions in the steel industry. These applications involve the prediction of final product quality in a *LD converter*, the prediction of *NOx* in a reheating furnace and an analysis of the effects that changing scrap types have on final product quality in an electric arc furnace. These applications have been undertaken by the University of Manchester and Perceptive Engineering Ltd. in partnership with several industrial partners, including Swedish Steel and MEFOS. The applications are each very different and together they provide an interesting insight into the benefits that PLS can offer in the manufacturing of steel.

1. Introduction

Since it was first proposed by Wold (1966) in the 1960's, partial least squares (PLS) has been applied to a broad range of problems. In the process industries the algorithm has been used primarily as a regression tool, able to cope with the highly correlated data routinely encountered in this field. More recently, application studies have demonstrated the ability of PLS to provide a convenient and robust method for detection and isolating fault conditions on industrial plants (Kourti *et al*, 1995).

In this paper the application of PLS to generate soft-sensing and predictive models for three case-studies within the steel industry is described. These case studies are each very different and involve the use of PLS to provide:

- optimisation capabilities in a reheating furnace;
- a prediction of the effects that varying scrap types have on final product quality in an electric arc furnace;
- a prediction of the end-point in a batch operated LD converter.

The following three sections provide an overview of each of the above case studies. This is followed by a brief list of conclusions that can be made from this work.

2. Case Study 1: Estimation of NOx in a Reheating Furnace

Nitrogen oxides (NOx) are produced in a reheating furnace in a hot-strip rolling mill as a result of burning fuel. In the present climate of highly politicised environmental issues it is of paramount importance to ensure that emissions of these gases are minimised with minimal impact on productivity. The objective for this study was to develop a model able to describe the causal relationships within the reheating furnace. If successful this model could then to be used to provide advisory information, indicating how operating conditions could be changed to reduce NOx emissions. Whilst the primary use for the model is to provide optimisation capabilities, in terms of minimising NOx emissions, it can also be used as a soft-sensor, reducing the cost of expensive hardware-based analysers in the ever more competitive manufacturing industry.

<u>Reheating Furnace</u>
The process that was used for the development of the NOx prediction model was chosen to be a reheating furnace in the hot-strip rolling mill at the SSAB (Swedish Steel) site in Borlange, Sweden. Reheating furnaces are the first component of hot-strip rolling mills, intended to reheat the steel slabs from nominally ambient temperature to around 1300 degrees C. The source of energy is the burning (oxidation) of volatile chemical liquids (oil or LPG). As a by-product of oxidation, nitrogen oxides are produced and discharged into the atmosphere through a stack.

<u>Model Development</u>
To develop an accurate prediction model a series of experiments were conducted on the furnace. In particular, the flow rates of air and fuel into the burners preheating zones were varied to sufficiently excite the plant. In addition to these primary cause variables several other process variables from the furnace, such as zone temperatures, combustion air temperatures and slab distribution in each of the preheating zones have been included in the model. To capture the dynamic nature of the process, a FIR (finite impulse response) structure was declared for the model. Figure 1 illustrates the prediction accuracy of the model over an unseen data set. The accuracy of this model was considered to be suitable for this application.

Online Implementation
The above PLS estimator has been implemented online at SSAB site in Borlange. In addition to the prediction model, additional cause-validation PCA monitors and bias adaptor have been implemented to improve the robustness of the overall system.

Cause-validation PCA monitors have been implemented to ensure integrity of the NOx estimation in the presence of thermocouple failures. Hence, if one or more thermocouples, used in the NOx model, does fail then the corresponding temperature measurement is obtained using the single component regression technique (Nelson *et al*, 1996). Therefore despite sensor failure, the on-line system remains operational.

In order to ensure zero-mean prediction error, the bias of the NOx prediction is estimated by passing the prediction error (whenever a NOx measurement is available) through a low-pass filter, and adding this value to the original raw NOx prediction.

The schematic of the on-line system is illustrated in figure 2. This schematic provides an estimation of the NOx value and also indicates which of the sensor measurements are considered to be 'bad'. These sensor measurements, displayed in red on the left hand side of the schematic are inferred using the PCA model.

3. Case Study 2: Quality Estimation in an Electric Arc Furnace

Electric arc furnaces (EAF) are used to melt solid metallic materials. It can be described as a chemical batch reactor that utilises electricity to transform solid metal to molten metal. Raw materials are loaded in baskets (normally two baskets per batch, or melt) which are charged to the top of the furnace. When all the steel in the furnace is melted, a taphole is opened in the bottom of the furnace and the steel is tapped into a ladle. In most meltshops (except stainless and speciality steel producers) a 'hot heel' is left in the furnace to avoid any slag being sucked into the vortex forming above the taphole.

Electric arc furnaces are most commonly used to melt steel scrap and approximately 1/3rd of the world's production of steel is made through the electric arc furnace root. The main source of energy in the EAF is of course electricity (electric arcs are formed between the electrodes and the scrap/molten steel), but in many furnaces a significant part of the total energy input comes from various types of chemical fuel, such as natural gas, oil, coal, non-noble metals and metallurgical dust. These fuels may be utilised in different ways, such as combustion in the furnace, scrap preheating and melt injection, which each affect the efficiency of the applied energy. Figures 3a and 3b illustrate this with a schematic of the material and energy flow in an EAF.

A lack of knowledge regarding the properties of the scrap types used in the EAF is today limiting the precision by which the electric arc furnace process can be predicted, modelled and controlled. The objective of this project was to provide a prediction of the final steel analysis, based upon the amounts and types of scrap material used in the particular melt. The predictions are of course useful for the operator when controlling the process, but through coefficient analysis they can also be used to estimate the mean and standard deviation of scrap properties such as chemical analysis, specific melting energy and metal content.

This project is still in its infancy, however, during the project it is expected that independent models will be made for several different steelworks and the results will be collected into a database of scrap properties. The results presented here refer to one such steelworks.

To identify a predictive model for this furnace, data was analysed from over 3000 melts. The effect variables for this model were the various composition measurements taken from the final steel analysis, such as tin and copper content. The cause variables were the types and weights of scrap sent to the furnace, energy inputted to the furnace, the final steel temperature and the analysis from the previous melt (which provides the composition of the 'hot heel'). The co-efficients for the model were then identified using the PLS algorithm.

Figure 4 illustrates the accuracy of this model when used to predict the tin content of the final melt. The accuracy of this model is considered to be acceptable by plant operators and engineers and similar results have been produced for predictions of other components within the steel, such as copper and sulphur. The accuracy of the model developed here has been compared with that which is obtained if more traditional identification algorithms, such as multiple linear regression (MLR) or Principal Component Regression (PCR) are used. The PLS model produced similar results to the PCR model, but with fewer components and significantly outperformed the MLR model, which was unable to cope with the high levels of co-linearity between the cause variables.

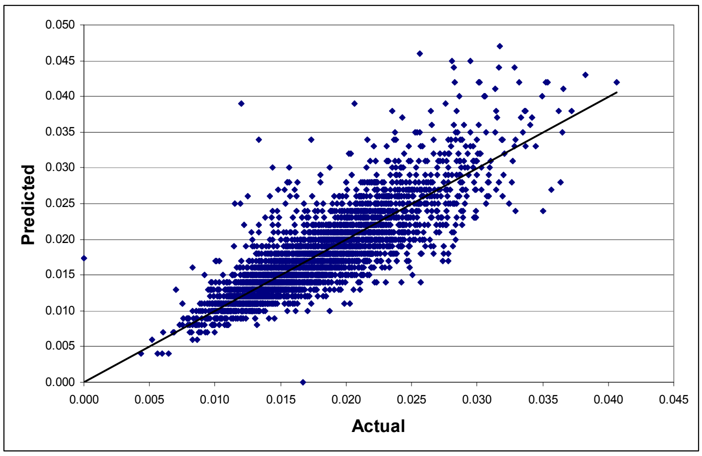

Figure 4: Prediction of steel

4. Case Study 3: Final Quality Prediction in an LD Converter

This application involves a batch operated process that forms part of a SSAB Oxelösund (Swedish Steel) steel-making plant. The process, termed a LD converter reduces the carbon levels in iron by blowing oxygen into a bath to form low carbon steel. Two inert gases, Ar and N_2, are blown from the furnace bottom to mix the bath. The blowing process takes approximately 20 minutes per batch and is split into three phases. During the first phase, between 3-6 minutes, Si, Mn and Fe are oxidised and lime is dissolved to form slag. The second phase takes approximately 15 minutes and commences after the Si has been oxidised. The decarburisation rate is almost constant during this stage, and decreases rapidly during the brief third phase due to carbon depletion in the metal, which results in heavy Fe oxidation. The briefness of the third phase makes it difficult to stop the oxygen blowing at the correct time. More accurate timing of the oxygen shut-off would allow the carbon target to be achieved more closely and lead to a reduction in average blowing time and an increase in product consistency. Prediction of FeO content in the slag, based on data from the end of recorded batches, can be used to define the batch endpoint and thus indicate when to cease oxygen blowing. The objective for this case study was therefore to provide an on-line prediction of FeO content in the slag.

Batch processes tend to exhibit strong non-linearities as the conditions at the end of the batch are very different to those at the start. Unfortunately, PLS is a linear tool, which limits its effectiveness when applied to such processes. However, for the analysis of batch data, studies have found that the major non-linearities in the data can be removed by transforming the data prior to application of the MSPC tools. The most common form of data transformation, termed multi-way MSPC, was proposed by Nomikos and MacGregor (1994). Since then other researchers have adopted the approach and applied it to a variety of processes. For example, Gallagher *et al* (1996) applied the technique to monitor nuclear waste storage vessels and Lennox *et al* (2001) investigated the detection of faults in fed-batch fermentation processes.

There are two important factors which must now be considered regarding the multi-way approach described above. The first is that a condition of the unfolding technique described above is that the run length of each batch is consistent. This is not the case with the LD Converter where run lengths may vary by an order of magnitude of three. Although more complicated methods exist, such as dynamic time warping of data (Kassidas *et al*, 1998), the standard approach to coping with variable run lengths is to simply determine the shortest run length and then ignore the data collected after this time in all batches. For the LD Converter this approach is unsuitable because it is the data collected towards the end of the batch, the third phase of operation that is of particular concern. The method that was adopted in this study was to ignore the data collected at the start of the batch and to use the data collected during the latter stages of each batch. It was found that if the final 50 samples were used from each batch then this encompassed the third phase of operation for each batch.

The second factor is that it is relatively straight forward to develop the PCA or PLS model. However, the approach suffers from the fact that if the model is to be used on-line then it requires the final 50 measurements made from the batch. Such information will only be available once the batch has been stopped, which is clearly unacceptable. Therefore, for the purposes of on-line FeO prediction, the model will be run over a moving window of fifty samples, providing a rolling estimate of the iron oxide concentration. It is envisaged that the endpoint would be declared, and oxygen blowing ceased, when the predicted FeO concentration reached a pre-defined level.

As mentioned earlier, the iron oxide content of the slag was taken as the variable to be predicted to define batch endpoint. Static measurements of FeO were available from the end of each batch. In all, 35 static variables were included and 5 dynamic variables, with a window length of 50 and a sample interval of 6 seconds, giving a total of 285 variables in the unfolded matrix. After initial evaluation, records of 119 batches were included in the analysis. The first 80 batches were used for PLS modelling and the remaining batches used for evaluation. The prediction accuracy of the approach is illustrated in figure 5. The PLS model that is used to generate these predictions contained 5 latent variables. In this figure the first 80 measurements were used to develop the PLS model, with the remainder being used to evaluate it. It is evident from figure 5 that the model is far from perfect, however, it is capable of identifying the major trends in the FeO measurement. The quality of prediction is viewed very favourably by SSAB who at present have no comparable method of identifying FeO content on-line.

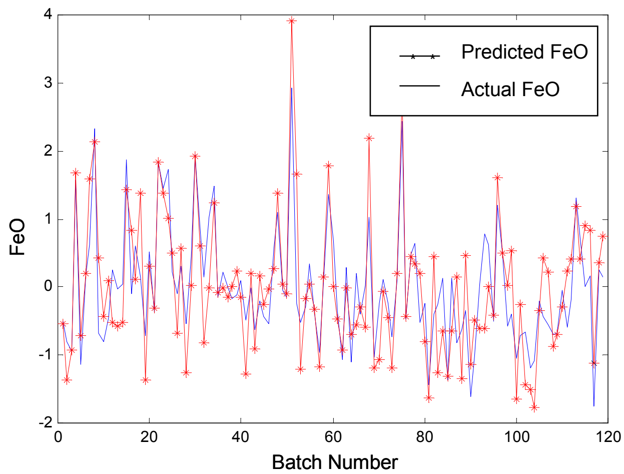

Figure 5: PLS Prediction of FeO

Figure 6 shows a simulated run of the on-line predictor using a typical batch from the test set. It can be seen that the initial predictions are poor, because the model was developed on data from the end of batches and thus cannot reproduce the behaviour of the first stage of the blowing process. As the batch progresses, the predictions improve and tend towards the endpoint value. The PLS model, reducing the prediction problem from 285 variables to five latent variables thus provides an accurate on-line indicator of FeO concentration from which the endpoint of the batch may be more accurately predicted.

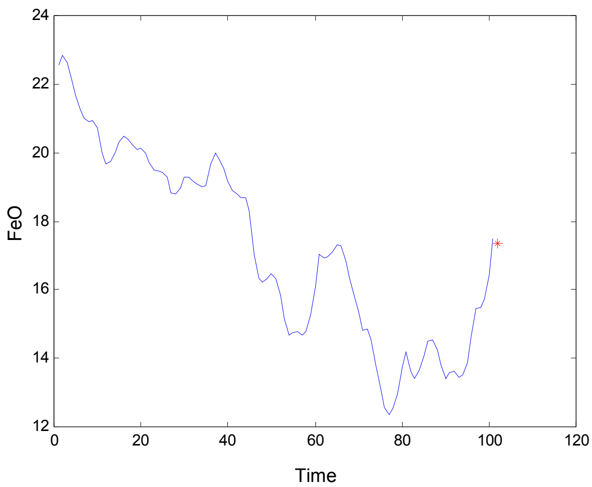

Figure 6: On-line FeO Prediction

5. Conclusions

This paper has provided an overview describing the results from three case studies which have focused on realising the benefits that multivariate statistics, and in particular PLS, has in the steel industry. The application studies that have been reported show that there are many benefits to be gained through the use of PLS, particularly as a tool for identifying predictive models for process plant that contain high levels of colinearity.

Each of the studies reported have been well received by the industrial engineers and operators working on this project. The case studies that have been investigated each involve the analysis of a process that is either poorly understood or there is very little instrumentation available to help plant operators and engineers make informed decisions. In each application PLS has provided useful information regarding plant conditions which can be used to improve process efficiency.

Acknowledgments

The authors would like to thank Jernkontoret for sponsoring certain elements of this work and the various engineers who have helped in this study. In particular thanks go to Per-Olof Norberg, Magnus Norberg, Lennart Klarnas, Risto Ponkala, Daniel Widlund and the staff of Perceptive Engineering Ltd., all of whom have spent valuable time helping with the projects described in this paper.

References:

Gallagher, N.B., Wise, B.M. and Stewart, C.W. (1996), “Application of multiway principal components analysis to nuclear waste storage tank monitoring”, *Comp. Chem. Eng.*, **20S**, 739-744

Kassidas, A., MacGregor, J.F. and Taylor, P.A., 1998, ‘Synchronisation of batch trajectories using dynamic time warping’, *AIChE Journal*, 3-25

Kourti T, Nomikos P. and MacGregor J.F. “Analysis, Monitoring and Fault Diagnosis of Batch Processes Using Multiblock and Multiway PLS”, *Journal of Process Control*, Vol. 5, No. 4, pp. 277 – 284, (1995)

Lennox, B., Montague, G.A., Hiden, H.G., Kornfeld, G., Goulding, P.R. (2001), ‘Process monitoring of an industrial fed-batch fermentation’, *Biotechnology and Bioengineering,* **74** (2), 125-135

Nelson, P.R.C., Taylor, P.A. and MacGregor, J.F. (1996), “Missing data methods in PCA and PLS: score calculations with incomplete observations”, *Chemometrics and Intelligent Laboratory Systems*, **35**, 45-65

Nomikos, P. and MacGregor, J.F. (1994), “Monitoring batch processes using multi-way principal component analysis”, *AIChE Journal*, **40** (8), 1361-1375

Wold H., Estimation of Principals Components and related Models by Iterative Least Squares, in Multivariate Analysis, ed. By P. Krishnaiah, pp. 391 – 420, Academic Press, New York

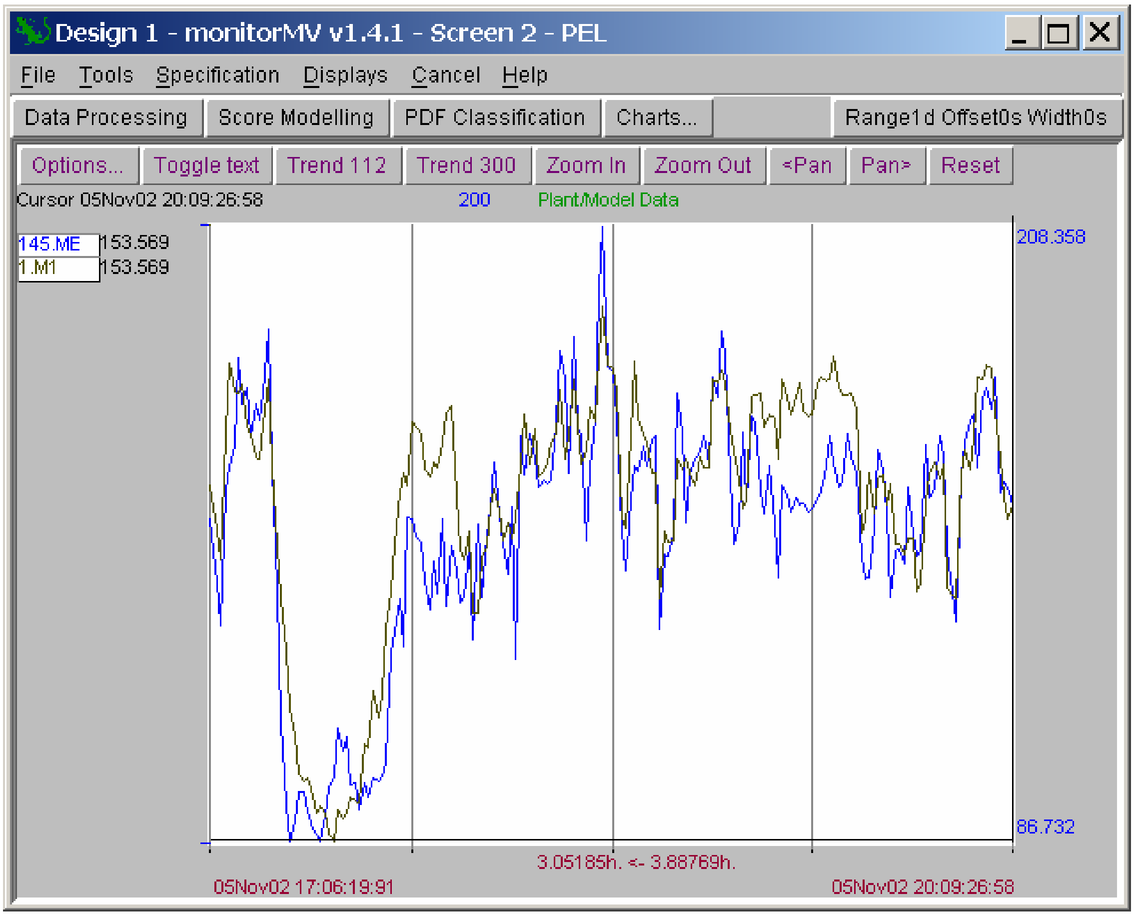

Figure 1: Prediction of NOx by PLS model

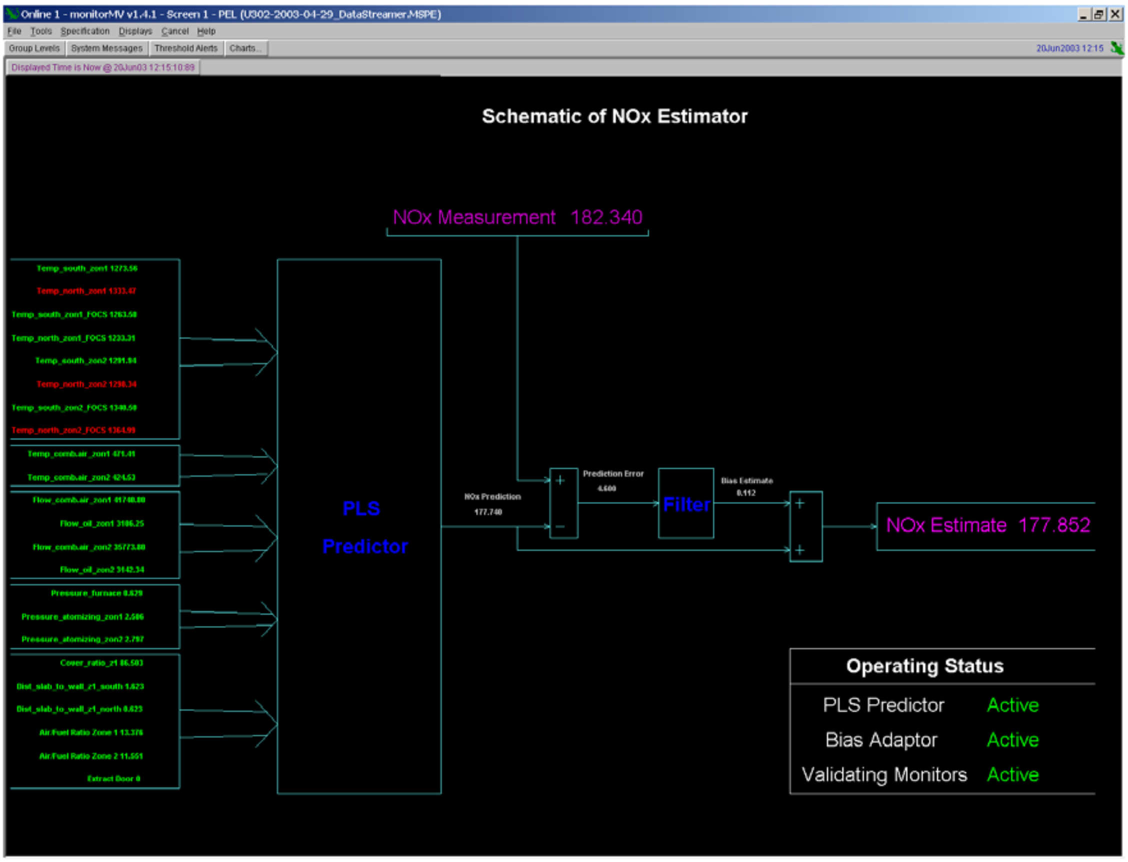

Figure 2: NOx on-line prediction schematic

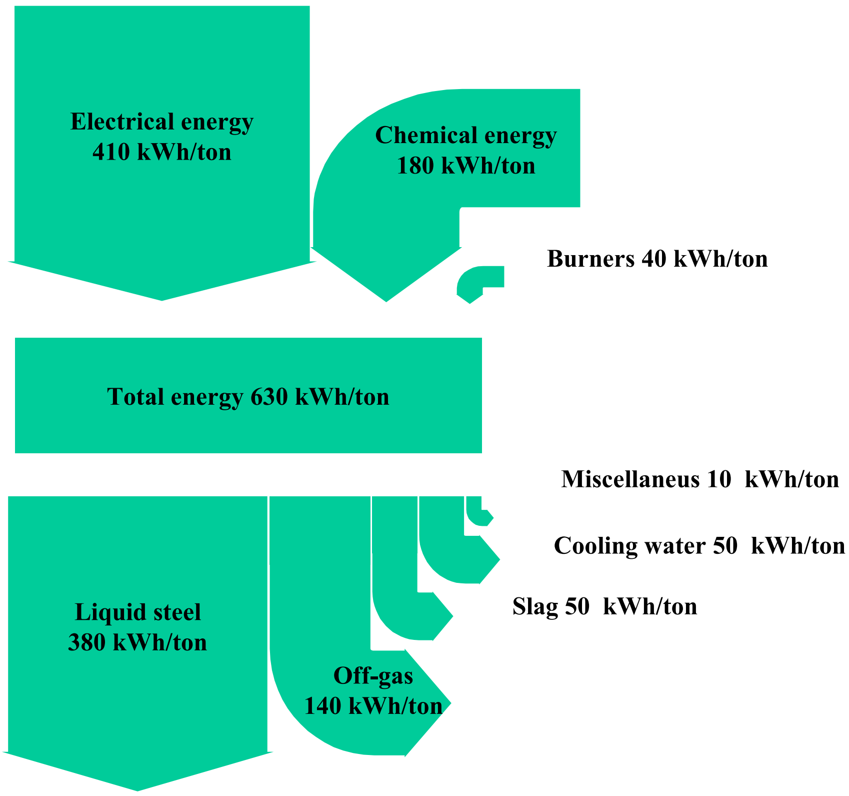

Figure 3a: Energy Flow in EAF

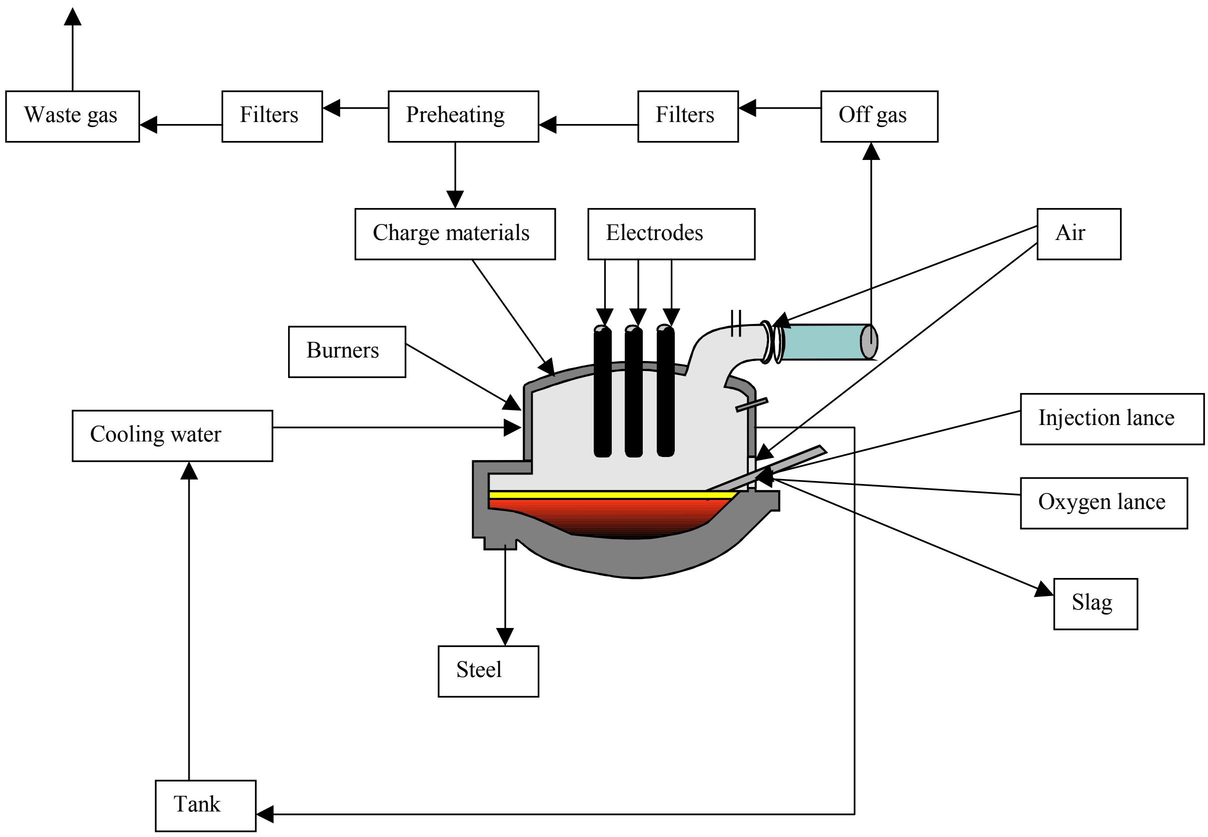

Figure 3b: Material Flow in an EAF