A new SPC tool in the steelshop at ArcelorMittal Gent designed to increase productivity

Rodrigo Rezende Amaral

A dissertation submitted to the Faculty of Engineering and Architecture of Ghent University in partial fulfillment of the requirements for the degree of Master of Science in Industrial Management

Advisor: Prof. Dr. Peter Ottoy
Supervisor: Ir. Luc Van De Putte

May 2012
Abstract

This work describes the development and application of a new tool for Statistical Process Control (SPC) in the steelshop at ArcelorMittal Gent. Among other features, this tool allows the Process Technology and Operations staff to identify not only operational problems (as does classical SPC) but also opportunities to increase the productivity of the plant.

The central feature of this new tool is the possibility to explore underlying data when using Shewhart Control Charts. In the classical approach, control charts are commonly used to detect abnormal variations of a process variable. The tool presented in this work helps making the link between those variations and other process variables, allowing a faster diagnosis of the root causes of the process variability. Additionally, through defining which variables account for more or less problems in the long run, it boosts the knowledge generation within the organization, leading to higher productivity.

In order to have an overview of the tool and its applicability in the department, the processes that take place in the steelshop are briefly presented, followed by comments on the importance of SPC for the business. The existing tools and the previous use of SPC in the company are subsequently described, demonstrating the possibilities and needs that could be fulfilled by the new tool.

The technical details about the development of this tool are then presented and examples of real applications are given. These examples show how productivity gains can be achieved by applying it. The work is concluded with comments on the lessons learned during the development and utilization of the tool and suggestions of other possibilities to apply it.

Keywords:

Statistical Process Control (SPC)
Performance Management
Continuous improvement
Deming Cycle (PDCA)
Shewhart Control Charts
# Table of Contents

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Introduction</td>
<td>4</td>
</tr>
<tr>
<td>1.1 The Company</td>
<td>6</td>
</tr>
<tr>
<td>1.2 The Steelshop</td>
<td>7</td>
</tr>
<tr>
<td>1.3 How process performance impacts production costs</td>
<td>11</td>
</tr>
<tr>
<td>1.4 Statistical Process Control Basics</td>
<td>14</td>
</tr>
<tr>
<td>1.5 The Purpose of This Work</td>
<td>18</td>
</tr>
<tr>
<td>2. Development</td>
<td>20</td>
</tr>
<tr>
<td>2.1 The Existing Tools</td>
<td>21</td>
</tr>
<tr>
<td>2.1.1 Real-time process control, data acquisition and data storage</td>
<td>21</td>
</tr>
<tr>
<td>2.1.2 Online knowledge application and decision support systems</td>
<td>25</td>
</tr>
<tr>
<td>2.1.3 Data analysis and reporting</td>
<td>27</td>
</tr>
<tr>
<td>2.2 The New Tool</td>
<td>28</td>
</tr>
<tr>
<td>2.2.1 Overview</td>
<td>29</td>
</tr>
<tr>
<td>2.2.2 Main Input</td>
<td>29</td>
</tr>
<tr>
<td>2.2.3 Main Output</td>
<td>34</td>
</tr>
<tr>
<td>3. Applications</td>
<td>49</td>
</tr>
<tr>
<td>3.1 Increasing the heat size</td>
<td>49</td>
</tr>
<tr>
<td>3.1.1 The economic gains of a increased heat size</td>
<td>50</td>
</tr>
<tr>
<td>3.1.2 Defining the target heat size</td>
<td>52</td>
</tr>
<tr>
<td>3.1.3 Using the spc tool to increase heat size</td>
<td>56</td>
</tr>
<tr>
<td>3.2 Reducing alloying costs</td>
<td>59</td>
</tr>
<tr>
<td>3.2.1 The costs of prediction errors and the determination of the target</td>
<td>60</td>
</tr>
<tr>
<td>3.2.2 Pursuing and predicting the target chemical composition of a heat</td>
<td>61</td>
</tr>
<tr>
<td>3.2.3 Using the spc tool to reduce prediction errors of the chemical composition</td>
<td>62</td>
</tr>
<tr>
<td>3.3 Finding best practices to avoid material losses</td>
<td>64</td>
</tr>
<tr>
<td>3.3.1 Reasons and methods for slag removal operations</td>
<td>65</td>
</tr>
<tr>
<td>3.3.2 Losses concerning slag removal operations</td>
<td>67</td>
</tr>
<tr>
<td>3.3.3 Using the spc tool to spot opportunities to avoid material losses</td>
<td>69</td>
</tr>
<tr>
<td>4. Conclusion</td>
<td>71</td>
</tr>
<tr>
<td>4.1 Lessons Learned</td>
<td>72</td>
</tr>
<tr>
<td>4.2 A Late Example</td>
<td>73</td>
</tr>
<tr>
<td>4.3 Further development</td>
<td>76</td>
</tr>
<tr>
<td>References</td>
<td>77</td>
</tr>
<tr>
<td>List of Figures</td>
<td>79</td>
</tr>
<tr>
<td>List of Tables</td>
<td>81</td>
</tr>
<tr>
<td>List of Abbreviations</td>
<td>82</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction

The steel industry is a capital intensive business and, as such, requires high value investments in assets and involves a high level of variable and fixed costs. The profitability of the business depends to a great extent on the stability of the production processes and its productivity level, i.e. the relation between output and input [Asian Efficiency, 2011].

\[
\text{Productivity} = \frac{\text{Output}}{\text{Input}}
\]

There are different forms of input and output in a production process. Raw material, equipment and manpower are the most common inputs, while finished products are the most common outputs. There are other forms of input and output which are more abstract and difficult to measure, as knowledge and customer satisfaction. Moreover, there are forms of input and output which are not wanted, as defective raw material, waste and pollution. Since these forms have different units of measurement – or are simply impossible to measure directly – input and output are normally expressed in a currency unit, representing respectively the total cost and the total income of the production process in a time interval.

In periods when the market demand for steel is very high, it is desired to keep the production level close to the capacity. In this situation, the total cost of production rises, but the profits compensate the increase.

\[
\text{Productivity} \uparrow = \frac{\text{Output} \uparrow}{\text{Input} \uparrow}
\]

On the other hand, in periods of economic difficulties, the production level has to match a low demand. To sustain a good productivity level, it is essential to reduce the costs.

\[
\text{Productivity} \uparrow = \frac{\text{Output} \downarrow}{\text{Input} \downarrow}
\]

Reducing costs is also an objective in good periods, but that is not always the main focus. For example: as it will be described later, a steel plant produces steel mainly from hot metal and metal scrap. Depending on the market price of those raw materials and considering technical constraints, there is a mix which yields an optimal production cost and another mix that yields maximal production volume. The level of demand determines the mix to be chosen: the former is usually chosen during periods of economic downturn and the latter during periods of economic growth.

Productivity can be increased by improving the efficiency and/or the effectiveness of the business. Effectiveness is often defined as doing the right things and efficiency as doing the things right. Those definitions could be rewritten respectively as doing what the market wants and doing it with the lowest cost. For instance, when the production is increased to match a higher demand, the business becomes more effective. When the mix of hot metal and metal scrap is chosen to minimize the production costs, it becomes more efficient. Being both effective and efficient is the ideal to be pursued. In the example, this would mean to always fulfill the market demand at a minimum cost for the necessary production level.
Fulfilling the market demand should be understood not only from the end-customers’ point of view, which includes volume, quality and on-time delivery, but also the other stakeholders’ (shareholders, community and government, among others) expectations should be taken into account. In this broader view, issues like long-term strategy, environmental concerns and social responsibility have to be considered.

Figure 1.1 shows the four ways in which productivity improvements contribute to a higher profitability of an organization. To simplify the demonstration, the variable costs and income were considered directly proportional to the number of products manufactured and sold, respectively. The profitability gains demonstrated in the figure are achieved by a combination of efficiency and effectiveness improvements. For example, production capacity can be increased by using the installations in a smarter way (lower left graph), which leads to a lower relative production cost (total cost per number of products sold), or, in other words, an efficiency improvement. As a consequence, this allows the fulfillment of a higher number of orders, which is an effectiveness improvement.
In order to achieve high productivity levels in production, it is first of all necessary to know the processes and the business well, being able to connect the process variables to the business. Second, it is necessary to be able to measure, monitor, predict and control those variables. This work intends to give an overview of how a new tool for Statistical Process Control is helping the steelshop at ArcelorMittal Gent to accomplish these goals.

1.1 The Company

ArcelorMittal Gent is part of the ArcelorMittal Group, the largest steel producer in the world and the market leader for automotive, construction, household appliances and packaging segments. The group was formed in 2006 by the merger of Arcelor and Mittal Steel, both groups being formed as results of previous mergers. In 2011, ArcelorMittal had revenues of US$94 billion and crude steel production of 91.9 million tons, representing approximately 6% of the world’s steel output. [ArcelorMittal Website, 2012]

The production site in Ghent is an integrated steel plant located on the right bank of the Ghent-Terneuzen Canal, about 20 km from the centre of the city of Ghent. It is the largest private employer of the province of East Flanders, employing approximately 5,000 people. In 2011, the company shipped around 4.4 million tons of flat carbon steel. [ArcelorMittal Gent Intranet, 2012]

In general, an integrated steel plant combines several production stages on one site. The main production departments can be observed in Figure 1.2. The steelshop is the production department which is the focus of this work. Its processes are explained into more detail in the next section.

Figure 1.2 – ArcelorMittal Gent [ArcelorMittal Gent Intranet, 2012]
1.2 The Steelshop

The blast furnaces produce hot metal using sinter and coke as main raw materials. Sinter is produced from iron ore in the sinter plant and coke from metallurgical coal in the coke plant. These materials, filled into the top part of the blast furnace, are continuously melted by a hot blast which is blown into its bottom part. The hot metal is periodically cast into units called torpedo cars at temperatures of around 1250°C [This is how ArcelorMittal Gent produces steel, 2010]. Most of the torpedo cars in ArcelorMittal Gent have a nominal capacity of 150ton. The blast furnaces at ArcelorMittal Gent can be seen in Figure 1.3 and the hot metal production process is demonstrated in Figure 1.4.

![Figure 1.3 – Blast furnaces at ArcelorMittal Gent](image)

![Figure 1.4 – Hot metal production at the Blast Furnaces](image)

The chemical properties of the hot metal are not suitable for most of the steel applications. Table 1.1 shows the main differences in the composition of those materials.

<table>
<thead>
<tr>
<th>Chemical element</th>
<th>Presence in hot metal</th>
<th>Presence in steel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fe</td>
<td>94%</td>
<td>99.4%</td>
</tr>
<tr>
<td>C</td>
<td>4.1 – 4.5%</td>
<td>0.02 – 0.2%</td>
</tr>
<tr>
<td>Mn</td>
<td>0.30 – 0.35%</td>
<td>0.08 – 1.2%</td>
</tr>
<tr>
<td>Si</td>
<td>0.35 – 0.40%</td>
<td>≤ 0.6%</td>
</tr>
<tr>
<td>P</td>
<td>0.11%</td>
<td>≤ 0.015%</td>
</tr>
<tr>
<td>S</td>
<td>0.021%</td>
<td>≤ 0.015%</td>
</tr>
</tbody>
</table>

Table 1.1– Chemical composition of hot metal and steel [Staalbereding, 2008]

The transformation of hot metal into steel happens in a plant called steelshop. The first step is the desulphurization process, which occurs while the hot metal is still in the torpedo car. Using calcium carbide, this process decreases the sulphur concentration in the hot metal from values between 100ppm and 400ppm, depending on the operational conditions of the blast furnaces, to values between 10ppm and 90ppm, depending mainly on the desired steel grade. After the desulphurization, the hot metal is cast from the torpedo car into a ladle and the slag generated during the process removed. The hot metal ladles have a capacity for around 260ton, what means that a hot metal ladle is filled by different torpedo cars and a
Liquid steel is produced by blowing oxygen in a mixture of hot metal and metal scrap. Thus, parallel to the hot metal preparation, the converter, which is the installation where the oxygen blowing process takes place, is charged with some metal scrap. The hot metal ladle is then cast into it, melting the scrap. Oxygen is finally blown by means of a water-cooled lance in order to decrease the carbon level in the bath, which reaches temperatures around 1650°C in the end of this process. The converter charging and the oxygen blowing process are shown in Figure 1.7. The casting of a hot metal ladle can be seen in Figure 1.8.
The fraction of hot metal and metal scrap in this mixture depends on many factors: the price, quality and availability of these raw materials, the desired output volume and grade, the required temperature for the blowing process, among others. The choice of the mix has impacts on the amount of oxygen to be blown, on the duration of this process and on properties of the steel which are of great importance for the next production steps, such as chemical composition and temperature.

This temperature plays an important role for the logistics and costs of the steel plant. Ideal temperatures are needed for alloying and casting the steel. Thus, if the temperature after oxygen blowing is too low, the heat has to be reblown. When this happens, more iron and other elements become further oxidized and are lost in the slag, more aluminum has to be used to deoxidize the steel afterwards, more oxygen and more refractory are consumed, and the blowing of the next heat has to be delayed, causing more temperature and energy losses [ArcelorMittal Gent Know-how book: CVT - Kost herbazen, 2011].

On the other hand, aiming a higher temperature at all the times also causes more costs: longer blowing duration, higher average consumption of oxygen, refractory and cooling scraps etc. Besides these losses, there is another important limitation: very high temperatures (about 1740°C) cannot be supported by the ladles without risks of breakthroughs, which can lead not only to huge economic losses but also to risks to the safety of employees and installations.

After the oxygen blowing process, the liquid steel is cast into a steel ladle and taken by a crane to the secondary metallurgy installation, where final adjustments on its chemical properties are made. During this process the slag resulting out of the oxygen blowing process is removed, the excess of oxygen in the bath is cleared away by means of aluminum alloys and other alloying elements, such as chrome, manganese and nickel are added. The tapping operation and the secondary treatment are illustrated in Figure 1.9.

![Figure 1.9 – Converter Tapping and Secondary Treatment](This is how ArcelorMittal Gent produces steel)

Again the temperature at the end of the process is a relevant concern. If it is too high, it increases the risk of breakthroughs during the continuous casting, the next production step. To avoid that, high quality cooling scrap should be added during the secondary treatment.

The average net weight of a full steel ladle is currently around 300ton. A higher average could increase the productivity of the plant since the fixed costs (wages, terrain etc.), maintenance and “heat-related” losses (refractory, sampling equipment etc.) would be spread in a relatively higher output volume and, in case of strong demand, the department would profit from an increased production capacity. However, the weight has to be limited due to the lifting capacity of the cranes and the available volume in the converter and the ladle, which varies according to the number of heats after a refractory relining.

At the end of the secondary treatment the steel ladles are transported by cranes to one of the two continuous casting machines where the steel is solidified into slabs. The ladle is first placed into a turret, which can be seen in Figure 1.10, and then protected with a ladle cover to maintain the temperature. When the previous ladle is completely cast, the turret turns 180° and the full ladle can be cast. The liquid steel flows from the ladle into an 80ton capacity tundish and from this to the mould where it starts becoming...
solid. The tundish works as a buffer of liquid steel which allows the continuity of the casting while the turret is turning.

![Continuous casting machine (Turret) (This is how ArcelorMittal Gent produces steel)](image1)

After leaving the mould, the steel is supported by a series of rolls, through which the slab is guided. Large quantities of cooling water are then sprayed upon the slab between the rolls in order to solidify it. The slab is cast vertically, but is steadily bent horizontally. Each casting machine has two casting lines. The slabs leaving the casting machine are cut to length by natural gas and high-pressure oxygen cutting machines. The process is demonstrated in Figure 1.11.

![Continuous Casting (This is how ArcelorMittal Gent produces steel)](image2)

The casting width lies between 1310mm and 2630mm at casting machine number 1 and between 950mm and 2000mm at casting machine number 2. Figure 1.12 shows the liquid steel solidifying into slabs.

![Liquid steel solidifying into slabs](image3)

The steel slabs are the main product of the steelshop. They are further processed in the hot strip mill, where they are used for the production of hot rolled coils. Besides the hot metal, scrap and alloys, the steelshop needs various other services and raw materials. Electrical and mechanical maintenance, refractory care, cooling water treatment, by-products (gas, slag etc.) handling and administrative services are examples of it.
1.3 How process performance impacts production costs

As mentioned in the previous section, some process variables play essential roles in the cost of steelmaking. The temperature after oxygen blowing of a heat, for instance, determines if it has to be reblown, the energy and refractory spent in the converter process, the amount of cooling scrap which needs to be added in the secondary metallurgy and, in the worst case, the possibility of a breakthrough of the steel ladle. In a simplified view, reblowing represents a relatively high cost; the energy, refractory and cooling scrap represent a medium cost which increases with the difference between achieved and ideal temperatures; and the breakthrough of a ladle represents a very high cost.

In the current operational standards, it is not possible to predict and drive the temperature after oxygen blowing exactly to a predefined value. As in almost all production processes, there are levels of accuracy and precision, concepts explained in Figure 1.13.

![Accuracy and Precision](Figure 1.13 – Accuracy and Precision [Wikipedia: Accuracy and Precision, 2012])

In practice, what happens is that a target temperature is defined and actions to achieve it are planned based on practical and empirical knowledge of the processes and measurements of other variables. For different reasons, the achieved value is sometimes higher and other times lower than the target. Some of these reasons are:

- Insufficient (or insufficiently implemented) knowledge of the process
- Inadequate measurement of process variables (inaccuracy or imprecision)
- Unexpected machine behavior (lack of maintenance, improper use, inappropriate design)
- Non-conforming raw materials
- Incorrect execution of the planned actions (human error)
- Influence of the external environment
- ...

Knowing that the achieved temperature will not always equal the target, but instead be distributed around it, one of the decisions which have to be made in order to minimize the production costs is to define the target temperature, taking into account the above mentioned costs and the probability distribution of the achieved temperature given the target.

To demonstrate how this could be done, consider the sketch of the probability density function shown in Figure 1.14. The bias (difference between the target and the average temperature) is not being considered in order to facilitate the analysis. The global reblowing cost is proportional to the fraction of the heats for which the achieved temperature is lower than the reblow threshold; the global cost of energy, refractory and cooling scrap is proportional to the average temperature; and the global breakthrough cost is proportional to the sum of the product of the probability of achieving a temperature and the risk of
breakthrough at this temperature, for all possible temperatures. These three costs are represented as reblowing costs, wasted energy costs and breakthrough costs respectively.

If the target temperature is set as a higher value, the configuration of the costs will change, as shown in Figure 1.15.

This perception was used by Dr. Genichi Taguchi to create a function which relates the performance of a variable compared to the target and the cost of the process [iSixSigma, 2012]. This function, known as Taguchi’s Loss Function, is represented in Figure 1.16. As it can be observed, there is a value for the target temperature for which the sum of the global costs is optimal.
A second optimization exercise that can be done is to decrease the variability of the process. When this is done, the configuration of the costs also changes, what can be observed in Figure 1.17.

For the new value of the process variance, it should also be possible to determine a new optimal cost, which will be lower than the previous. This leads to the conclusion that there is a function which relates variance and cost, as shown in Figure 1.18.
The variability of a process can be reduced by mitigating the factors listed in the beginning of this section. All the same, it also incurs costs. Buying better equipment, reinforce maintenance or studying the processes further require investments. The relation between the cost reduction and the investments to be made has to be considered.

For instance, if the same investment is necessary to reduce the variance from point 1 to point 2 and from point 2 to point 3 on the previous graph, the first one will have a shorter payback time. Depending on the impact of the cost of the process on the whole business, it could be that the first investment would be welcome and the second not.

The reasoning demonstrated in this example can be used when studying many other process variables in the steelshop. The average net weight of a full steel ladle, for instance, is one of them. As mentioned in the previous section, the productivity of the plant could be increased by a higher average weight, but the lifting capacity of the cranes and the available volume in the converter and ladles limit the weight of each heat. The degree in which these constraints are respected and consequently the average weight depend on the variability of difference between aimed and achieved weights.

Another example is the use of alloying material and the achievement of the chemical specifications for the steel in the secondary treatment. Each steel grade defines limits for the mass fraction of relevant chemical elements, usually expressed in parts-per-million (ppm). The fraction of the heats which fulfill the requirements depends on the variability of the difference between aimed and achieved concentrations of the different elements, while the amount of alloying material used depends on the average target. Not fulfilling the requirements or using more than necessary alloying material are costs which have to be minimized.

More examples for which the approach could be used: duration of operations (torpedo casting, ladle transportation, converter tapping etc.); number of equipment breakdowns per period or per produced volume; amount of defective machine parts replaced during inspections; percentage of cooling water that is recycled, among many others. Since all these process variables have a direct impact on the effectiveness and efficiency of the steelshop, the need for methodologies and tools to define, achieve and maintain acceptable levels of variability in the processes or, in other words, to implement Statistical Process Control (SPC), is very clear.

### 1.4 Statistical Process Control Basics

In the implementation of SPC, two phases can be distinguished: achieving and improving control (engineering phase), and maintaining control (operational phase) [Peter Ottoy, 2011]. During the first phase, the causes of process variability should be studied. After an analysis of the possible gains and necessary investments, actions should be taken to eliminate permanent (or common) causes and mechanisms to detect and eliminate transitory (or special) causes of variability should be designed. These mechanisms should be summarized in an Out of Control Action Plan (OCAP). The engineering phase is validated when the variability of the process yields the desired economic results.

During the operational phase, the performance of the process should be regularly evaluated. While its variability arouses only by unknown, unavoidable or tolerated (too expensive to be mitigated) common causes, the process is considered to be in control. If the inaccuracy or imprecision increase, a special cause should be determined and the OCAP executed. In case the special cause cannot be found or the OCAP is not sufficient to bring the process back to control, the engineering phase should be reviewed.

The behavior of process variables is the means used by SPC to define process variability, accuracy and precision. In the whole implementation of SPC, attention has to be given to the measurement and interpretation of those variables. Therefore statistical tools are employed throughout all the phases.
The PDCA (plan-do-check-act) cycle, made popular by Dr. W. Edwards Deming, is one of the methodologies used nowadays in industry to implement SPC. The first two stages of the cycle (plan and do) correspond to the engineering phase. The plan stage comprehends the study of the causes of variability (both common and special) and the possibilities to eliminate them. This also includes the conception of the OCAP.

The actions to eliminate common causes are performed in the do stage. If it is feasible to force some special causes, the OCAP can also be tested. In case the actions do not lead to the expected results, the plan stage should be restarted. Otherwise, the engineering phase is concluded.

The two following stages (check and act) correspond to the operational phase. The regular evaluation of the process is the main purpose of the first one and the detection of variations induced by special causes triggers the transition to the second, when the OCAP should be executed. If the process can be brought back to control, the check stage is resumed. Otherwise, the whole cycle has to be restarted. Figure 1.19 shows some management tools which can be used in each stage of the PDCA cycle.

The Control Charts, originally developed by Walter A. Shewhart and also known as Shewhart Charts, are a well spread tool to determine whether a process is in statistical control. In short, a control chart consists of points plotted in chronological sequence representing the individual measurements of a process variable or a statistic calculated by grouping those measurements into so-called rational subgroups. The statistic can be, for instance, the mean, the range or the standard deviation of the subgroups.

Besides the points, three horizontal lines are also calculated and plotted. They are called centerline (CL), upper control limit (UCL) and lower control limit (LCL). The centerline is set equal to the mean or the expected value of the plotted points. The control limits are usually set equal to the centerline plus and minus three theoretical standard deviations of the plotted values.

The plotted statistic determines the control chart type. The most used chart types are summarized in Figure 1.20. Not all of them were included in the original set proposed by Shewhart. The Cumulative Sum Chart...
(CUSUM), for instance, was developed by E. S. Page, and the Exponentially-weighted Moving Average Chart (EWMA) by S. W. Roberts.

The choice of the chart type and the control limits depends on the type of data, the expected behavior of the process, the kind of process variation which it is supposed to detect, among others. These design issues are discussed into more detail in [James Benneyan, 1998]. [Peter Ottoy, 2011] explains how the points and limits should be calculated for each chart type.

<table>
<thead>
<tr>
<th>Chart type</th>
<th>Plotted statistic</th>
</tr>
</thead>
<tbody>
<tr>
<td>X-chart (or i-chart)</td>
<td>individual measurements</td>
</tr>
<tr>
<td>mR-chart</td>
<td>moving range of the individual measurements</td>
</tr>
<tr>
<td>mA-chart</td>
<td>moving average of the individual measurements</td>
</tr>
<tr>
<td>(X\bar{b})ar</td>
<td>subgroup means</td>
</tr>
<tr>
<td>R-chart</td>
<td>subgroup ranges</td>
</tr>
<tr>
<td>s-chart</td>
<td>subgroup standard deviations</td>
</tr>
<tr>
<td>EWMA-chart</td>
<td>exponentially weighted moving average</td>
</tr>
<tr>
<td>CUSUM-chart</td>
<td>cumulative sum of individual measurements or subgroup averages</td>
</tr>
<tr>
<td>c-chart</td>
<td>count of defects in subgroup</td>
</tr>
<tr>
<td>u-chart</td>
<td>defects per unit in subgroup</td>
</tr>
<tr>
<td>np-chart</td>
<td>number of defective units in subgroup</td>
</tr>
<tr>
<td>p-chart</td>
<td>proportion of defective units in subgroup</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Type of data</th>
<th>Subgroup size</th>
<th>Chart type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Continuous variable</td>
<td>Individual</td>
<td>(X) (or (i)) - mR - mA - CUSUM</td>
</tr>
<tr>
<td>measurements</td>
<td>Fixed size</td>
<td>(X\bar{b} - R - z) - EWMA - CUSUM</td>
</tr>
<tr>
<td></td>
<td>Variable size</td>
<td></td>
</tr>
<tr>
<td>Atribute data</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Count of nonconformities</td>
<td>Fixed size</td>
<td>c-chart</td>
</tr>
<tr>
<td>Nonconforming units</td>
<td>Variable size</td>
<td>u-chart</td>
</tr>
<tr>
<td>(OK / NOK data)</td>
<td></td>
<td>np-chart</td>
</tr>
<tr>
<td></td>
<td>Variable size</td>
<td>p-chart</td>
</tr>
</tbody>
</table>

Figure 1.20 – Most used control chart types

An example of an X-chart can be observed in Figure 1.21. The points on this chart could represent, for instance, the difference between the actual and the expected duration of a certain operation for each batch in seconds (e. g. transport of a ladle, tapping of converter etc.). A high value means that the operation took much longer than expected.

Figure 1.21 – X-chart example
The general rule states that a point out of the control limits is an indication that the process might be out-of-control and should be studied. This is because, assuming the points follow a normal distribution, only about 0.27% of them are expected to lie outside the limits when the process is under statistical control. Thus, whenever it happens, the existence of special causes for variations should be investigated. The last mentioned percentage will be false-alarms.

Some authors and organizations have created extra rules to detect trends or systematic patterns on the charts. In [Western Electric, 1956] the quality staff of the Western Electric Company defined a set of rules known nowadays as Western Electric Rules. Instead of having only one upper and one lower control limit, they first divided the chart into three zones, as described in Table 1.2.

| Zone A          | Within ±2σ of the centerline and the control limit (±3σ) |
| Zone B          | Within ±1σ and ±2σ of the centerline                   |
| Zone C          | Within ±1σ of the centerline                           |

Table 1.2 – Chart zones for Western Electric Rules

Next they considered that an out-of-control alarm should be given whenever one of these situations occurs:

1. Any single point falls out of zones A, B or C (out of the control limits);
2. Two out of three consecutive points fall in or beyond zone A on the same side of the centerline;
3. Four out of five consecutive points fall in or beyond zone B on the same side of the centerline;
4. Eight consecutive point fall on the same side of the centerline.

Figure 1.22 shows how these situations could appear on an Xbar-chart.

Western Electric has also suggested other rules depending on the type of data and chart, probability distribution and the kind of variation that should be detected. A similar set of rules was also suggested by Lloyd S. Nelson in 1984 and is currently known as Nelson Rules. It suggests, for instance, that out-of-control alarms should be given when six consecutive points are continually increasing or decreasing, or when fourteen consecutive points alternate in direction, increasing then decreasing [Nelson, 1984].

Obviously, the introduction of more control rules increases the rate of false alarms, since the described situations can occur even when the process is under statistical control.
1.5 The Purpose of This Work

The values of a process variable can often be associated to other variables or properties. Process batches, for instance, can be from different classes or grades, made by different equipment or operators, under different machine configurations, produced with raw material from different types or suppliers etc. It is often useful to know the influence of each of those properties on the variation of a process variable measured for each batch.

Nevertheless, as it could be observed in the previous section, the classic Control Charts are made using information about only one variable. So, when using the charts in the traditional way, the influence of the different properties cannot be easily detected. Three reasons are:

1. **The subgroups have data from different categories**

   Given all the possible combinations of properties, it is usually unfeasible to create charts where all subgroups have exactly the same attributes.

2. **The points are auto-correlated in some charts**

   In some charts, like the Moving Average (mA) or the Page’s Cumulative Sum (CUSUM), the value of a point depends on the values of the previous points. Therefore, even when the subgroup size is one (individual data) a point cannot be fully linked to the properties of its corresponding batch.

3. **The control rules associate points next to each other**

   Rules like the ones suggested by Western Electric and Nelson are often defined based on situations which can occur with consecutive points. However, as batches with similar properties are not always made in succession, the patterns which these rules aim to detect can exist for determined properties and still not be found on the charts.

These problems could be solved by filtering the data beforehand and making a chart for each category. To illustrate that, consider again the use of a X-chart to monitor the difference between the actual and the expected duration of a certain operation, as the one in Figure 1.21. Consider also that the operation can be executed on different classes of products, in different installations, by different operators. Figure 1.23 shows the same chart, but the points that represent operations executed by a given operator are marked.

![Figure 1.23](image)

*Figure 1.23 – An out-of-control situation that could be detected in advance if data is filtered beforehand*
It can be easily observed that the concerned operator takes in average longer than the others to accomplish the operation. Still, applying the Western Electric Rules for the whole data on the chart, only the last operation executed by him is marked as out-of-control. The use of other chart types, such as EWMA or CUSUM would not help detecting it either because the impact of the variable operator would be shared by the different points on the charts.

However, if a control chart for each operator is made separately, an alarm is given much earlier, when the situation eight consecutive on the same side of the centerline occurs. And even without applying the Western Electric Rules, an alarm would still appear earlier if the above-mentioned charts are used.

The idea of making a chart for each category entry in every chart has not been applied in the company before this work and constitutes its first objective. In other words, it means developing and applying a tool that enables the easy creation of control charts based on underlying properties of the production batches in order to detect variations of a process variable in advance when compared to the traditional tools.

An expected disadvantage of this approach is that too many charts would be generated for one variable. In the previous example, it means that different charts would have to be made for each operator, each machine, each product class etc. Knowing that charts produce false alarms, checking every out-of-control situation in all of them manually would unnecessarily increase the workload in the department, what should surely be avoided. Thus, the second objective consists in designing mechanisms to interpret and filter the alarms, drawing conclusions about which properties of the batches account for less or more process variation.

The third – and presumably the most important – objective of this work is to demonstrate how the use of the new tool could increase the productivity of the department. The achievement of this objective is particularly valuable because it opens the path for spreading the employed techniques to other departments and production sites, generating even greater economic gains.

It is not the objective of this work to replace the existing tools and practices for SPC used in the company, but to extend it.
Chapter 2

Development

ArcelorMittal Gent and especially the steelshop department have always been considered a reference for the ArcelorMittal Group and for other steel companies in relation to its automation and informatization level. Data availability and reliability is relatively high and the staff is well prepared to store, read and evaluate it.

Moreover, ArcelorMittal Gent has also been a reference for knowledge about the steelmaking processes. Those two facts are not independent: according to Knowledge Management concepts, data and information are the basis for knowledge generation, as illustrated in Figure 2.1, and this has since long been an objective pursued and emphasized by the management of the company.

Nevertheless, becoming a reference in those fields requires much effort. Process data needs to be consistently captured and stored. Then it has to be carefully grouped, filtered and analyzed to be transformed into information which should be displayed in the right way to the right people. When patterns are found within the information and their implications are understood, knowledge is generated. But knowledge by itself does not generate economic gains: it has to be applied in the processes. Applying it in a structured way through management systems and techniques, such as Statistical Process Control, reflects the wisdom of the organization.

The SPC tool which is the core of this work can be considered an extension of the tools previously used in the steelshop for achieving these objectives. For this reason, the first section of this chapter intends to give an overview of these previous tools, clarifying the context in which the new tool is implemented.

In the second section, the development of this tool is described. This description demonstrates how the implemented functionalities allow the achievement of the objectives stated in the introduction of this work.
2.1 The Existing Tools

The process data management tools in the steelshop are presented in this section divided into three groups:

1. Real-time process control, data acquisition and data storage
2. Online knowledge application and decision support systems
3. Data analysis and reporting

2.1.1 Real-time process control, data acquisition and data storage

Most of the operations in the steelshop are supported by programmable logic controllers (PLCs), which are digital computers used for automation of electromechanical processes. Unlike general-purpose computers, PLCs are designed for multiple inputs and output arrangements, extended temperature ranges, immunity to electrical noise, and resistance to vibration and impact [Wikipedia: Programmable logic controller, 2012]. They execute programs to control machine operation in real time and therefore must be highly reliable and have fast response time. The most common PLC used in the steelshop is the model SIMATIC S7-400 produced by Siemens and is shown in Figure 2.2.

![Figure 2.2 – Siemens PLC (Model SIMATIC S7-400)](image)

The PLCs receive data from sensors present in the plant, such as thermometers and pressure gauges, and, after performing calculations, send signals to actuators, such as valves and motors. The calculations are made not only using data from the sensors: there are computers connected to the PLCs which provide the operators with human-machine interfaces (HMIs), allowing them to visualize the processes and give low level commands as the opening of a valve or the start of a motor. Figure 2.3 shows a HMI used to control and monitor the slag removal operation before the secondary metallurgy.

Besides these HMIs, the PLCs communicate with so-called Process Computers, which are linked to databases, decision supporting systems and other computers. They also have interfaces which enable inputs from the operators and the visualization of the processes, allowing higher level commands, as starting a process, readapting the production planning or downgrading a non-conform product. Figure 2.4 shows an interface of the Process Computer used for the follow-up of the hot metal flow.

In the steelshop at ArcelorMittal Gent, the PLCs and the process computers are grouped by production installation. There is a group of PLCs and Process Computers for the hot metal stands and the converters, a group for the secondary metallurgy and a group for the continuous casting machines. Figure 2.5 gives an overview of the whole system.
Figure 2.3 – Human Machine Interface (HMI) for the slag removal operation

Figure 2.4 – Process Computer interface for the follow-up of the hot metal flow
Figure 2.5 – Overview of the computer system used for process control in the steelshop.
There are different kinds of databases. The production databases are built and used to allow the real-time operations. They are designed to store data in a high consistency level and to communicate efficiently with PLCs, HMIs and process computers. Therefore their tables and fields are strongly normalized and queries which are frequently performed become automatically optimized. A description of database normalization can be found in [Christopher Date, 1999].

These characteristics of production databases make their direct use for reporting and research inefficient. For this reason, their data are automatically selected, filtered and stored in other databases. The mainframe database, called DB2, is the main of them and is updated daily. The tables in these databases are mostly product or equipment oriented, i.e. each row contains data of one batch (heat or slab) or of one tool or installation (torpedo car, steel ladle etc.).

To store continuous data from processes which comes from the PLCs, ArcelorMittal Gent has developed a tool called Process Data acquisition Centre (PDC). It is shown in Figure 2.6. It replaces some functionalities of Plant Information Management Systems (PIMS), which are available on the market but were not purchased by the company. The best known PIMS products are the Aspen PIMS, developed by AspenTech and the PI, developed by OSIsoft.

![Figure 2.6 – Process Data acquisition Centre (PDC)](image)
2.1.2 Online knowledge application and decision support systems

The complexity of the processes, i.e. the amount and the interdependency of variables, requires computational tools to evaluate the variables automatically and suggest or implement actions. For instance, choosing the best production route given a production planning or determining the types and quantities of cooling scrap which should be added when the temperature of the liquid steel is too high are tasks which should be assisted by computers in order to achieve efficiency and reliability.

One of the tools in the steelshop designed for these goals is the Knowledge Base System (KBS). The KBS contains the models developed by the process specialists. These models require input variables from the processes and return output variables by applying logical rules and performing calculations. Techniques as linear programming and neural networks can be implemented in the models.

Generally speaking, the operators use a process computer to collect the required variables from the PLCs, HMIs and other computers, send a calculation request to the KBS and wait for its output. This is then evaluated, adapted if necessary and forwarded to the PLCs, which will drive the operations accordingly.

The determination of the types and quantities of cooling scrap to be added in the secondary metallurgy is an example of it. After the first phase of alloying, the operators take a sample of the temperature and chemical composition of the steel. This command is given in the process computer interface and sent to the PLCs. The temperature is immediately measured and the sample is sent by pneumatic tubes to a nearby chemical laboratory, where the chemical analysis is made.

These measurements are sent back to the process computer and the operators call a routine from the KBS which calculates the recommended quantities and types of cooling scrap based on the results of the analysis, the actual and desired temperature, the desired steel grade, the cost of each type of scrap etc. The process computer receives this output and the operators can evaluate, adapt and send it to the PLCs. Figure 2.7 shows a part of the model used by the KBS for this calculation.

![Figure 2.7 – The Knowledge Base System (KBS)](image)

In many cases, the KBS makes the inverse calculation when the process is finished: it calculates the expected output based on the actions that were actually made. For instance, it calculates which temperature would be achieved according to the model considering the quantities and types of cooling scrap that were actually added. The difference between this recalculated temperature and the actually achieved, which is measured after the process, is the error of the model. This can be used to adapt the model automatically, as a closed-loop control system, as shown in Figure 2.8. Nevertheless, the effect of this correction is limited and having a good model of the process is essential to achieve satisfactory results. Therefore, the data used in every calculation of the KBS is also stored in a database for further analysis.
Another important decision support tool in the steelshop which is closely related to this work is the Online Statistical Process Control (OSPC). It receives data from PLCs and Process Computers and makes control charts online, allowing the operators to immediately detect out-of-control situations of process variables. These variables can represent calibration of sensors, prediction error of KBS models, fulfillment of requirements for the chemical composition in each step of the production process, among others. Figure 2.9 is an example of a chart made by OSPC. It is an I-chart of the model error on the prediction of the temperature of the steel after the oxygen blowing process at the converter.

Besides the control charts, OSPC calculates process statistics and capabilities, allows the configuration and visualization of OCAPs and displays underlying data to accelerate investigation when an out-of-control situation is found. It is a very powerful tool and it is not the intention of the new tool described in this work to suppress its use.

On the other hand, adding new variables to OSPC cannot be done straightaway by the Process Technology staff since it depends on data from PLCs and Process Computers. It is not possible, for instance, to create a chart based on data from a spreadsheet or from a simple database query. This lack of flexibility motivated the creation of a new tool instead of extending the OSPC functionalities.

There are many other systems to support the steelshop operations, such as the Coordination System (KOORDI), which adapts the production planning in real-time according to operational circumstances, and the ArcelorMittal Integrated Global Optimization (AMIGO), which objective is to optimize material management. These systems, however, have a less direct relation to the new SPC tool and will not be explained in detail.
2.1.3 Data analysis and reporting

The steelshop has decided to use MATLAB as the main tool for data analysis and reporting. MATLAB is a product developed and distributed by MathWorks which consists of a high-level technical computing language and interactive environment for algorithm development, data visualization, data analysis and numerical computation [MATLAB, 2012].

MATLAB can be used for a wide range of applications, including signal and image processing, communications, control design, test and measurement, financial analysis and computational biology. Add-on toolboxes (collections of special-purpose programs) extend the MATLAB environment to solve particular classes of problems in these application areas. The steelshop has acquired the Statistics Toolbox from MathWorks and is developing its own toolbox (the AMG Toolbox), which allows reading data from its systems (mainframe and process computer databases, KBS, PDC etc.) and facilitates the use of the most common MATLAB functions, including the ones from the Statistics Toolbox. The tools of the AMG Toolbox are periodically compiled in executable formats (.EXE) and made available for everyone in the plant.

One of the main tools developed using MATLAB which is daily used by the Process Staff is called dex (Data Exploration Tool) and is shown in Figure 2.10. dex is part of the AMG Toolbox and centralizes its main data exploration functions, allowing users to load data from the mentioned systems or from other sources such as ordinary Microsoft Excel spreadsheets, make different kinds of graphs and apply statistical techniques such as multiple regression and hypothesis tests, all in a very quick, flexible and user-friendly way.

Figure 2.10 – dex

Based on studies carried out with dex, the Process Technology staff can improve the processes by adapting the working methods, refining the models, selecting better raw materials, adjusting machines etc. Furthermore, the graphs created by dex can be easily included in automatically generated reports to follow-up Key Performance Indicators (KPIs), which are available on the company’s intranet and daily updated. Finally, the studies can also be used to justify investments.
2.2 The New Tool

The new tool can be classified in the third group, i.e. *data analysis and reporting*, presented in the previous section. It uses SPC techniques (Control Charts) to enhance data analysis and reporting, but, differently from OSPC, it is not designed to support operational decisions online.

Following the trend of using MATLAB for these purposes, the new tool is created as part of the AMG Toolbox. It benefits from the previously implemented data reading features and adopts a layout similar to the one used in *dex*, which the Process Technology staff is already used to. The interaction between the two tools is smooth: it is easy and quick to create control charts with the new tool using a variable chosen for making a graph in *dex* and vice-versa. The new tool was named *spc* in the AMG Toolbox and it will be referred as such in this work from now on.

Apart from the AMG Toolbox, the Statistics Toolbox from MathWorks comprises a wide range of possibilities for Statistical Process Control, including functions for the creation of Control Charts. However, its original form has a low level of user-friendliness: after collecting and filtering the data, the user has to type all parameters for the chart in the MATLAB command line. These parameters include the chart type, the target, the variance, the control limits, the control rules, among many others.

*spc* provides a user-interface which combines the data reading features from the AMG Toolbox and the Statistical Process Control functions from the Statistics Toolbox. This interface can be compiled into a stand-alone application and executed in any computer, not requiring the installation of MATLAB.

This section will describe how the tool works from the user’s point of view. It is divided in three subsections:

1) **Overview**
   A. Start the program
   B. Program Layout

2) **Main Input**
   A. Load data
   B. Select a variable
   C. Choose a Reading Method
   D. Choose the Chart Parameters

3) **Main Output**
   A. Charts
   B. Process and Subgroups Statistics (*stats*)
   C. Plotted Data (*plotdata*)
   D. Histogram (*histfit*)
   E. Capability
   F. CUSUM
   G. Root Cause Advisor (*RCA*)
   H. Creating reports
2.2.1 Overview

A. Start the program

`spc` can be started up directly from MATLAB by the command `spc` or by its compiled executable file (`spc.EXE`) which is a stand-alone application. In the first case, the user must have MATLAB and the AMG Toolbox installed on his computer. In the second case, only the MATLAB Runtime Compiler is needed, which ArcelorMittal Gent is allowed to distribute.

Another way of starting `spc` is by a shortcut available in `dex`.

B. Program Layout

The `spc` interface is shown in Figure 2.11. It is divided in two blocks: Main Input and Main Output. These two blocks are described more into details in the next subsections.

![Figure 2.11 – The spc interface](image)

2.2.2 Main Input

A. Load data

In accordance with the other tools in the AMG Toolbox, `spc` works with data structures called *datasets* in MATLAB. In short, a MATLAB *dataset* is a table where each column has a name and the elements represent a variable. Similarly to `dex`, `spc` can call the routines from the AMG Toolbox to create datasets by reading data from the computer systems or ordinary files. Figure 2.12 shows how a dataset could be created in `spc` from a MS Excel spreadsheet.
B. Select a variable

Once the dataset is loaded, the data for the charts can be selected. Therefore, three fields are foreseen: Variable, Labels and Select. The first is mandatory and the other are optional. These fields are visualized in Figure 2.13.

**Variable** should be an expression that results in the process variable which statistic will be plotted.

**Labels** should be a dataset column containing the names which will be given to the rational subgroups. It can be, for instance, the product lot or the production date. If the column indicating labels is not unique per row, it is used to effectively create the subgroups (see next topic).

**Select** should be a logical expression to define which rows from the dataset should be considered for the control chart. It can be used, for instance, to consider only the operations executed in determined production installation or by determined shift team.
C. Choose a Reading Method

Depending on the type of the data, the way it is stored in the dataset and the desired chart type, a reading method should be defined, as shown in Figure 2.14.

The following reading methods are foreseen:

1. **Variable** is a matrix of measurements\(^1\) of a continuous variable and each row is a group in time order. (Xbar)
2. **Variable** is a vector of individual measurements of a continuous variable. (I-chart)
3. **Variable** is a vector of measurements of a continuous variable and **Labels** should be used for grouping. (Xbar)
4. **Variable** is a vector of individual measurements of a continuous variable that should be grouped each \(N\) elements. (Xbar)
5. **Variable** is a vector of proportions. A scalar or a variable from the dataset given in the field **Units** will indicate the number of inspected units per row.\(^2\) (np-chart)
6. **Variable** is a vector of binary variables and they should be grouped each \(N\) elements to form a vector of proportions. (p-chart)
7. **Variable** is a vector of binary variables and they should be grouped by **Labels** to form a vector of proportions. (p-chart)
8. **Variable** is a vector of counts of defects. A variable or scalar given in **Units** will indicate the number of inspected units per row. (u-chart)

Examples of variables, chart types and respective recommended reading methods are shown in Table 2.1.

---

\(^{1}\) In a MATLAB, a dataset column can contain a numeric matrix with several columns. Besides that, the user can choose more than one dataset column in **Variable** to create a matrix containing these columns.

\(^{2}\) **Units** and \(N\) are fields which appear when required by the reading method.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Chart type</th>
<th>Plotted statistic</th>
<th>Sample size</th>
<th>Reading method</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average duration of operation per N executions</td>
<td>Xbar</td>
<td>Subgroup average</td>
<td>Fixed</td>
<td>1* or 4</td>
</tr>
<tr>
<td>Model error for the calculation of the temperature at end blowing for each heat.</td>
<td>I-chart</td>
<td>Individual measurement</td>
<td>Individual</td>
<td>2 (Labels = heat number)</td>
</tr>
<tr>
<td>Average duration of operation per day</td>
<td>Xbar</td>
<td>Subgroup average</td>
<td>Variable</td>
<td>3 (Labels = date)</td>
</tr>
<tr>
<td>Fail rate (e.g. reblow) each N operations</td>
<td>np-chart</td>
<td>Proportion (fraction)</td>
<td>Fixed</td>
<td>5** or 6***</td>
</tr>
<tr>
<td>Fail rate (e.g. reblow) per day</td>
<td>p-chart</td>
<td>Proportion (fraction)</td>
<td>Fixed</td>
<td>5** or 7***</td>
</tr>
<tr>
<td>Count of defective machine parts replaced during inspection</td>
<td>u-chart</td>
<td>Defects per unit</td>
<td>Fixed or Variable</td>
<td>8</td>
</tr>
</tbody>
</table>

* Only if Variable is a matrix with more than one column
** Only if Variable is a vector of proportions (numbers from 0 to 1)
*** Only if Variable is a vector of binary data (0 = ok, 1 = fail)

Table 2.1 – Process variables, chart types and reading methods

The tool provides a “guessing algorithm” which evaluates the input fields and automatically detects the type of the variable and suggests the reading method and the chart type.

The combination of the presented data loading, selection and reading features allows the accomplishment of the first objective stated in Chapter 1 by “enabling the easy creation of control charts based on underlying properties of the production batches”.

D. Choose the Chart Parameters

Once the input variable is selected and the reading method is defined, the parameters for creating the control charts can be chosen. Figure 2.15 shows how it can be done using the spc interface and Table 2.2 gives a description of each parameter group.

![Figure 2.15 – Panel to select the Chart Parameters](image-url)
**Parameter** | **Description**
---|---
CHART TYPE | The desired control chart types to be plotted. There are four sets of compatible types: XBAR, S, R, and EWMA; I, MR, and MA; P and NP; U and C.
TARGET MEAN | Value for the process mean. This is the $p$ parameter for p- and np-charts, the mean defects per unit for u- and c-charts, and the normal $\mu$ parameter for other charts.
SIGMA | Either a value for sigma ($\sigma$), or a method of estimating it chosen from among std (the default) to use the average within-subgroup standard deviation, range to use the average subgroup range, and variance to use the square root of the pooled variance. When creating I, mR, or mA charts for data not in subgroups, the estimate is always based on a moving range.
CONTROL LIMITS | Two options:
1. $n$-sigma: the number of sigma multiples from the center line (target mean) to a control limit (default is 3)
2. user-defined values for each control line (lower control limit, center line and upper control limit).
SPECIFICATION LIMITS | Lower and upper specification limits (LSL and USL), i.e. the customer requirements or expectations. It can be, for instance, the limits for the concentration of a chemical element in a finished slab. These limits are used to calculate the process capability. Since specification limits typically apply to individual measurements, this parameter is primarily suitable for I-charts. They are not plotted on R, S, or mR charts.
CONTROL RULES | Western Electric or Nelson Rules. The codes for the control rules are shown in Figure 2.16
LAMBDA | A parameter between 0 and 1 controlling how much the current prediction is influenced by past observations in an EWMA plot. Higher values of LAMBDA give less weight to past observations and more weight to the current observation. The default is 0.4.
WIDTH | The width of the window used for computing the moving ranges and averages in mR and mA charts, and for computing the sigma estimate in I, mR, and mA charts. Default is 5.

Table 2.2 – Control Chart parameters
2.2.3 Main Output

After defining all input parameters, the user should click the Plot button. If no error is found, i.e. the expressions for Variable, Labels or Select are all valid and compatible to Chart type and Reading Method, and all Chart Parameters are valid, the Main Output is calculated and indicated by a menu of buttons, as shown in Figure 2.17.

![Plot and Main Output buttons](chart)

In order to make the topics of this subsection clearer, the output of two examples of input are presented. Table 2.3 summarizes the input parameters chosen for these examples.

<table>
<thead>
<tr>
<th>Example nr.</th>
<th>Variable</th>
<th>Input parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Variable: f.mn_afg_werk - f.mn_afg_her (difference between actual and expected)</td>
<td>Labels: f.lading (lading = heat) Select: f.lading &gt; max(f.lading)-3000 Reading Method: 2 Target mean: 0 Sigma: std Control limits: n-sigma = 3 Specification Limits: ± 1000 ppm Control rules: we2, we3, we6, we7</td>
</tr>
<tr>
<td>2</td>
<td>Variable: b.herblazen (herblazen = reblow; 0 or 1)</td>
<td>Labels: b.lading Select: b.cv == 2 &amp; b.lading &gt; max(b.lading) - 4000 Reading Method: 6 , N = 25 Chart type: p Target mean: auto Sigma: std Control limits: LCL: 0 , CL: 0.15 , UCL: 0.30 Specification Limits: none Control rules: none</td>
</tr>
</tbody>
</table>

Table 2.3 – Examples of Main Input for spc
A. Charts

By choosing the chart option, the control charts are displayed. Figure 2.18 and Figure 2.19 present the charts for the given examples. When a point is clicked, a tooltip shows the label of the subgroup and, if applicable, the control rule which classifies the point as being out-of-control.

![Figure 2.18 – Control Chart for example nr. 1](image)

![Figure 2.19 – Control Chart for example nr. 2](image)

B. Process and Subgroups Statistics (stats)

Depending on the chart type, the program calculates process and subgroup statistics. The process statistics are:

- **mu**: Estimated (or given) process mean
- **sigma**: Estimated (or given) process standard deviation
- **p**: Estimated (or given) proportion defective
- **m**: Estimated mean (or given) defects per unit
The subgroup statistics are:

- mean: Subgroup means
- std: Subgroup standard deviations
- range: Subgroup ranges
- n: Subgroup size, or total inspection size or area
- i: Individual data values
- ma: Moving averages
- mr: Moving ranges
- count: Count of defects or defective items

Figure 2.20 and Figure 2.21 demonstrate how these statistics are displayed.

C. Plotted Data (plotdata)

The lines plotted in the charts can be represented numerical columns, which is done by the plotdata option. It facilitates exporting and recreating the charts in other programs (e.g., Microsoft Excel). The columns and their meanings used for this representation are:

- pts: Plotted point values
- cl: Center line
- lcl: Lower control limit
- ucl: Upper control limit
- se: Standard error of plotted point
- n: Subgroup size
- ooc: Logical that is true for points that are out-of-control

Figure 2.22 and Figure 2.23 show how this representation is presented to the user.
D. Histogram (histfit)

The plotted statistic from Xbar, I, p, np, u, and c charts can be visualized in histograms. It is possible to estimate parameters for known probability density functions, including the Normal, the Exponential, and the Poisson distributions.

**Figure 2.24 – Histogram for example nr. 1**

**Figure 2.25 – Histogram for example nr. 2**
E. Capability

When specification limits are given and the selected chart types include an Xbar, I or p chart, the program calculates the capability indices for the process. This calculation is based on the assumption that the points follow a normal distribution. The estimated parameters and the calculated indices are:

- $\mu$: Estimated process mean or proportion defective
- $\sigma$: Estimated process standard deviation
- $P$: Estimated probability of being within limits
- $Pl$: Estimated probability of being below $L$
- $Pu$: Estimated probability of being above $U$
- $Cp$: $\frac{(USL-LSL)}{(6*\sigma)}$
- $Cpl$: $\frac{(\mu-LSL)}{(3*\sigma)}$
- $Cpu$: $\frac{(USL-\mu)}{(3*\sigma)}$
- $Cpk$: $\min(Cpl, Cpu)$

More information about process capability can be found in [Peter Ottoy, 2011].

Figure 2.26 shows the capability plot and indices for example nr. 1. They cannot be calculated for example nr. 2 because the specification limits were not defined in the input.

F. CUSUM

The CUSUM-chart is not included in the Statistics Toolbox from MathWorks and was implemented as part of the AMG toolbox. Some modifications in the original proposition made by E. S. Page were made.

The objective of a CUSUM-chart is to small detect shifts of the process output or persistent deviations from the target. To achieve this objective, the variables $C^+$ and $C^-$ are created according to the following equations:

\[
C_i^+ = \max[0, x_i - (T+K) + C_{i-1}^+] \quad (C_i^+ = 0)
\]
\[
C_i^- = \max[0, (T-K) - x_i + C_{i-1}^-] \quad (C_i^- = 0)
\]
\( X \) is the vector of containing the values of the plotted statistic of an Xbar, I or \( p \) control chart, while \( T \) and \( K \) are respectively the target mean and the threshold parameter of the CUSUM-chart.

\( C^+ \) increases when the difference between \( T \) and \( X \) is greater than \( K \), and decreases otherwise, but not below zero. The slope of increasing or decreasing is \( x_i - (T+K) \). Analogously, \( C^- \) increases when the difference between \( T \) and \( X \) is smaller than \( -K \).

If the process is stable, both \( C^+ \) and \( C^- \) tend to become zero after eventual increases. However, if the process mean is shifted, one of them will indefinitely increase (\( C^+ \) in case of a positive shift and \( C^- \) in case of a negative shift). The control limit \( h \), the second parameter of the CUSUM-chart defines the maximum value that \( C^+ \) or \( C^- \) can reach before an out-of-control alarm.\(^1\)

In the implementation of the chart in the AMG toolbox, two changes were made. First, the equations for \( C^+ \) and \( C^- \) were modified in order to avoid these variables from reaching extreme values:

\[
\begin{align*}
C_i^+ &= \max \left[ 0, \ x_i - (T+K) + D^+ \right] \quad (C_i^+ = 0) \\
C_i^- &= \max \left[ 0, (T-K) - x_i + D^- \right] \quad (C_i^- = 0)
\end{align*}
\]

\[
D^+ = \begin{cases} \ C_{i-1}^+ & \text{if } C_{i-1}^+ < h \\ 0.5*h & \text{otherwise}\end{cases}^2
\]

\[
D^- = \begin{cases} \ C_{i-1}^- & \text{if } C_{i-1}^- < h \\ 0.5*h & \text{otherwise}\end{cases}
\]

According to this modification, \( C^+ \) or \( C^- \) are virtually reset to \( 0.5*h \) for the calculation of subsequent values when the previous value exceeds the control limit \( h \). As a consequence of this change, it is not possible to define if the process is out-of-control only by observing a point on the chart and checking if it is above or below the limit. After a (small) positive shift, for instance, a point subsequent to the first one above \( h \) will probably be below \( h \).

However, if the deviation from the target persists, \( C^+ \) is unlikely to become zero again. Thus, as a second change, the out-of-control alarm does not stop until \( C^+ \) become zero. Obviously, the same is valid for \( C^- \).

These changes allow to detect not only a shift of the process variable or statistic, but also the return to normality. Only the last point of the chart must be checked to determine if the process is in control or not. Figure 2.27 illustrates this for an artificially created vector where the first and the last 20 elements follow the standard normal distribution, while the 60 points in the middle follow the same distribution with the average shifted by +1 unit.

---

\(^1\) The threshold and the control limit parameters (\( K \) and \( h \)) for the CUSUM-chart can be different for detecting positive or negative trends. Four different parameters are defined in this case: \( K^+ \), \( K^- \), \( h^+ \) and \( h^- \). To simplify the demonstration, it was considered that \( K^+ = K^- = K \) and \( h^+ = h^- = h \).

\(^2\) The factor 0.5 can be given as a parameter as well.
Figure 2.28 shows the CUSUM-chart for example nr. 2. The p-chart shown in Figure 2.19 is replicated in the upper portion to permit a clear understanding of the plotted values in the CUSUM-chart.
It is interesting to observe that the CUSUM-chart gives an alarm when the reblow rate remains low for a sustained period, which is not detected by the p-chart. This alarm disappears when the process becomes normal again.

G. Root Cause Advisor (RCA)

The Root Cause Advisor (RCA) is the main feature of spc in the context of this work. In the RCA panel, the user can select variables from the original dataset to create charts for each entry of these variables and compare the results of each of these entries. The variables can be the product class or grade, the shift team, the production installation etc. Figure 2.29 shows how this selection is made.
After the selection of the charts and the variables which represent the categories to be analyzed, the RCA can be performed. The output of the RCA for the example nr. 1 is shown in Figure 2.30. The steel grade (nuance) is the category chosen for the analysis.

Figure 2.30 – RCA for example nr. 1: influence of steel grade
When the output is ready, the user can select a category entry by clicking on its name on the table. The button *Make chart* becomes enabled, and clicking on it will plot the chart only for the selected entry, as shown in Figure 2.31.

![Control Chart created for a category entry](image)

**Figure 2.31 – Control Chart created for a category entry**

Additionally, when one or more categories are selected on the table, the buttons *Boxplot* and *Trend* become enabled, allowing the comparison between the selected categories and the rest by means of a boxplot or a trend graph. These possibilities are shown in Figure 2.32 and Figure 2.33 respectively.

![Boxplot integrated with RCA](image)

**Figure 2.32 – Boxplot integrated with RCA**
The objective, however, is not to let the user check each chart individually, but let the computer determine, based on criteria defined beforehand, the categories which deserve special attention. Therefore, a series of hypothesis tests are performed, as described in Table 2.4. All the tests are described in [W & W, 1990]. The usual parameters for the tests (required confidence level, which default value is 95%; minimum sample size etc.) can be chosen by the user.

<table>
<thead>
<tr>
<th>Test name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>z-test</strong> (two-proportions z-test)</td>
<td>Indicates if the proportion of out-of-control alarms in the chart of one category entry is significantly higher than the proportion by the other entries. If the test detects a significant difference for a category entry, it is repeated for all other entries leaving the data from the “bad” one out.</td>
</tr>
<tr>
<td><strong>t-test</strong> (one-sample t-test)</td>
<td>Indicates if the mean value of the plotted statistic is significantly different from the expected value.</td>
</tr>
<tr>
<td><strong>vartest</strong> (Chi-square variance test)</td>
<td>Indicates if the variance of the plotted statistic is significantly higher than the predefined sigma (given as input) or the calculated sigma in stats.</td>
</tr>
<tr>
<td><strong>runstest</strong> (runs test)</td>
<td>Indicates if the signs of (plotted statistic – expected value) come in a random order.</td>
</tr>
<tr>
<td><strong>CUSUM ooc</strong></td>
<td>Based on the adapted CUSUM-chart (see previous topic), it indicates if the last plotted statistic values deviate from the target value systematically.</td>
</tr>
</tbody>
</table>

Table 2.4 – Hypothesis tests in RCA
The results of the tests can then be used to define levels of alarm for each category entry. A possible combination of results and alarm levels is shown in Table 2.5. Choosing the parameters for the tests and how the combinations of results generate alarms are the criteria to be defined.

<table>
<thead>
<tr>
<th>Color</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>RED</td>
<td>positive z-test AND any other positive test</td>
</tr>
<tr>
<td>ORANGE</td>
<td>positive z-test but no other positive test OR any positive test but negative z-test</td>
</tr>
<tr>
<td>GREEN</td>
<td>no positive test</td>
</tr>
</tbody>
</table>

Table 2.5 – Results of tests and RCA alarms

The described approach allows the accomplishment of the second objective stated in Chapter 1 by “interpreting and filtering the alarms, drawing conclusions about which properties of the batches account for less or more process variation”.

Important remarks on RCA:

1. Recommended chart types

   Although RCA can be used for creating separately charts from any type (Xbar, R, I, mA, p etc.), the tests to determine the alarm level for the categories assume that the values of plotted statistic are statistically independent. As a consequence, it is not applicable for charts which points are correlated, as EWMA and mR. RCA is particularly recommended for Xbar, I, p and u charts.

2. Allow more Type I errors in order to have less Type II errors

   There are two types of erroneous outcomes when a single Control Chart is analyzed: the process can be considered in control when it is out-of-control (Type I error) or vice-versa (Type II error). This is described in Table 2.6.

<table>
<thead>
<tr>
<th>Chart does not give an alarm</th>
<th>Process is actually in control</th>
<th>Process is actually out-of-control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Correct outcome</td>
<td></td>
<td>Type II error</td>
</tr>
<tr>
<td>Type I error (False alarm)</td>
<td>Correct outcome</td>
<td>Correct alarm</td>
</tr>
</tbody>
</table>

   Table 2.6 – Types of errors in Control Charts

   There is a relation between the rates of the two types of error. The rate of Type II errors can be reduced by strengthening the control limits and adding control rules, but it would increase the rate of Type I errors (false alarms). Contrarily, the rate of Type I errors can be reduced by relaxing the limits and rules, but it would increase the rate of Type II errors.

   In the traditional approach for designing control charts, a balance between these rates has to be planned considering the costs of checking and acting upon false alarms and the costs of not detecting an out-of-control situation. An important design parameter of a classical control chart is therefore the in-control average run length (ARL).

   However, the alarms given by the charts are not individually checked in the RCA algorithm and no costs can be directly associated to it. Furthermore, if the process is stable and only false alarms are given, it should happen homogeneously to all categories and the z-test would not indicate a significant difference between them. This is indeed how RCA filters false alarms.

   As a consequence, accepting a higher rate of false alarms in comparison to the traditional approach is recommended since it would also increase the rate of correct alarms. Nevertheless the former should remain moderate (< 30%), otherwise it would mask the correct alarms. Using $n$-sigma = 2 or control rules $we_2$, $we_3$, $we_6$ and $we_7$ (Figure 2.16) is advisable.
3. RCA when category entries have different control parameters

It is possible to use the RCA algorithm when the groups have different target means, sigma values, control limits or specification limits. In this case, the variables TARGET_MEAN, SIGMA_VALUE, LCL, CL, UCL, LSL and/or USL should be included as columns in the original dataset and should have unique values for each category entry. Other possibility is to let RCA calculate these values automatically. This can be done by choosing the option auto for target mean and/or sigma and n-sigma for control limits (Main Input).

In both cases, the option Use the same control limits, target mean and process standard deviation whenever possible, which appears when the button Options is clicked (see Figure 2.30), should then be unmarked before running the RCA algorithm.

It is important to remark that when one of these possibilities are chosen, some of the suggested tests become meaningless. For instance, if the sigma value is different for each category entry, the vartest should not be considered.

4. Prior analysis of the original charts and correct interpretation of RCA results

In the industrial environment, it is common to program the production of batches with similar properties sequentially or concentrated in periods. This non-random assignment of categories can mislead the conclusions of the RCA algorithm. For instance, batches with a determined property (or category entry) can be predominantly produced in a period when the process as a whole is (by coincidence and because of other reasons) out-of-control, affecting the assessment made by RCA.

To illustrate this, consider the production of 2000 batches from two different categories (A and B) in two distinct periods (1 and 2). In each period, 1000 batches are produced. In the period 1, 60% of the batches are from category A and 40% from category B. In the period 2, 25% of the batches are from category A and 75% from category B. Due to changes in the production process, the rate of out-of-control alarms decreased from 20% in period 1 to 10% in period 2. The situation is illustrated in Figure 2.34.

![Figure 2.34 – Example of how changes in process can affect RCA results](image-url)

**Global rate of out-of-control alarms**

\[
\text{A} = \frac{(120+25)}{(480+120+225+25)} = 17.1\%
\]

\[
\text{B} = \frac{(80+75)}{(320+80+675+75)} = 13.5\%
\]

Figure 2.34 – Example of how changes in process can affect RCA results
The rate of alarms in each period does not depend on the categories, but if all data is put together, the rates are 17.1% for batches from category A and 13.5% for batches from category B. Given the number of batches in each period, the z-test from the RCA algorithm would appoint a significant difference in the proportion of out-of-control alarms for each category.

To avoid a misinterpretation of the results, the original Control Charts plotted before the execution of the RCA algorithm should be first observed, providing an overview of the evolution of the process as a whole. In fact, if the original charts are observed on time, the change in the process can generally be detected before the non-random assignment of categories affects the RCA assessment.

Other way to check if the RCA results are consistent is by observing the trend curves for the “bad” categories and the others, as shown in Figure 2.33. If the points corresponding to batches from the “bad” categories present the same behavior as the other points but are concentrated in “bad” periods, the assessment is probably not correct.

5. Exclude data after major process changes for category assessment

RCA uses all input data to perform the tests on the categories. For this reason, after implementing a change in the process (adjusting a machine, change the raw material requirements, adapting the model etc.), the preceding data should be excluded (using the Select field in the Main Input) before executing the RCA algorithm, as described in the previous item. This contrasts with the traditional use of the charts only for online control.

On the other hand, the creation of separate charts by RCA for each category entry without filtering the data is an interesting way of evaluating the impact of a change implemented for a specific entry. For instance, after adjusting a machine which output was systematically deviating from the target, the performance before and after the can be compared using its separate CUSUM-chart.

H. Creating reports

The MATLAB code executed by spc to read the input data, make the charts and perform the RCA algorithm can be automatically generated and copied to a common MATLAB script using the option shown in Figure 2.35.

MATLAB offers the possibility to publish the results of a script in various formats, including HyperText Markup Language (.HTML), Extensible Markup Language (.XML), Portable Document Format (.PDF), Microsoft Word (.DOC) or Microsoft PowerPoint (.PPT) and LaTeX formats. A published script can be visualized in any computer which can open the files from the selected format, even if it does not have MATLAB software installed.

In the steelshop at ArcelorMittal Gent, it is possible to select scripts which are automatically published every day by a server computer. The resulting files are made available on the company’s intranet. Using the code generated by spc, scripts can be written and selected to be automatically published, enabling the creation of daily SPC reports. An example of a HTML report is presented in Figure 2.36.
Mn: actual - calculated (tapping)

Difference between actual and calculated Mn after converter tapping.

Contents
- Description
- Boxplot
- Charts per category
- Conclusion

Description

VARIABLE: f.mn_afq_werk - f.mn_afq_her  -> werk = actual ; her = calculated
LABELS:  f.lading
SELECT:  f.lading > max(f.lading)-3000  -> last 3000 heats
CHARTS:  T-chart,CUSUM-chart (basis for RCA), using Control Rules: w=2, w=3, w=6, w=7
DATE:  06-Apr-2012 20:48:15

Boxplot

Figure 2.36 – Report created using spc
Chapter 3

Applications

The techniques implemented in the spc tool are applicable to any process for which Control Charts can be used, especially when the process variable in question can be linked to other known properties which can have impacts on the process performance. It can be used, for instance, to monitor and control the calibration of measurement systems (sensors), to evaluate the performance of different operators or shift teams and define best practices, or to gradually improve the mathematical models for the processes.

Three examples of applications of the tool are described in this chapter. The first application is related to increasing the heat size, the second to reducing alloying costs and the third to defining best practices to avoid material losses.

3.1 Increasing the heat size

As mentioned in Chapter 1, the productivity of the steelshop could be increased if the average net weight could be safely increased. Two reasons for this were presented. The first reason is that some costs (fixed, maintenance and “heat-related”) would be spread in a higher output volume. The second reason is that it would increase the total production capacity from the plant by reducing the production lead time, since part of it is constituted by “heat-related” operations (charging and transport operations, sample taking etc.). This would allow the company to fulfill more orders in periods of higher demand.

Recalling the concepts around productivity given in the introduction of this work, the first reason represents an efficiency improvement (reduce relative production costs) and the second an effectiveness improvement (increase the number of fulfilled orders). Both lead to a productivity increase, as shown in Figure 3.1.

![Figure 3.1 – How productivity gains can be achieved by increasing the heat size](image)

In this section, the ways in which increasing the heat size would impact the economic results of the steelshop are first explained, then the method used to determine the target heat size is described and finally, the applicability of the spc tool to achieve an higher average heat size is demonstrated and the actual gains calculated.
3.1.1 The economic gains of an increased heat size

The number of fulfilled orders is increased and the fixed costs per produced ton are decreased only if the extra capacity generated by an increased heat size is actually used. This decision is not made by the steelshop but at higher management levels in the company. However, “heat-related” costs are always reduced by increasing the heat size. Only the reduction of these costs will be considered in this section, although the actual gains can be even higher.

“Heat-related” costs are a special kind of variable costs which arise each time a converter heat is made. Differently from “tonnage-related” costs, such as hot metal and metal scrap, “heat-related” costs do not depend directly on the total weight which is produced, but are proportional to the number of heats which are made. If the average heat size is increased, less heats must be made to reach the same production level and this is the reason why these costs would be reduced. The main “heat-related” costs are summarized in Table 3.1 and are subsequently described.

<table>
<thead>
<tr>
<th>Description</th>
<th>€/heat</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Steel losses during the production processes</td>
<td>1,445.00</td>
</tr>
<tr>
<td>2. Hot metal losses during converter inter-campaign</td>
<td>465.00</td>
</tr>
<tr>
<td>3. Refractory for converters and ladles</td>
<td>560.00</td>
</tr>
<tr>
<td>4. Other (heat-related maintenance, sampling lances and other consumables)</td>
<td>120.00</td>
</tr>
<tr>
<td>TOTAL</td>
<td>2,590.00</td>
</tr>
</tbody>
</table>

Table 3.1 – Heat-Related costs

1. **Steel losses during the production processes**

During the production steps, part of the steel is lost. For instance, during the converter tapping and slag removal operations, some liquid steel is disposed along with the slag, and during the continuous casting process, not all content from the ladle can be cast depending on the format of the refractory layer of its bottom. Some of these losses can be recovered and used as scrap in subsequent heats, but it involves energy losses and reprocessing costs.

In general, the steel losses per heat are independent from the net weight of the ladle, which means that an increase of the heat size would decrease the value of the losses per produced ton.

2. **Hot metal losses during converter inter-campaign**

During the relining of the refractory layer of one of the two converters, which happens around every 140 days and takes about 13 days, all the production of the steelshop goes through the remaining converter. In this period, called *inter-campaign*, the overall capacity of the steelshop is reduced and the hot metal produced by the blast furnaces cannot be completely processed in its liquid form. Part of this material is thus solidified, processed by an external company and used as scrap in future heats.

Although it is not exactly a direct “heat-related” cost, it should be included in the present context since an increase of the heat size would increase the capacity during the inter-campaign, allowing more liquid hot metal to be processed and avoiding energy losses and unnecessary processing costs.

3. **Refractory for converters and ladles**

The refractory of converters and ladles is usually relined in function of the number of heats for which they were used. The costs related to it would be decreased if less heats have to be made for a desired production level. Since the relining is mainly executed by own personnel, the labour costs are considered to be part of the fixed costs and are not included in the calculation. Only the material costs were considered.
4. **Other heat-related costs**

Some material, equipment and actions required for the production of steel are related to the number of heats which are produced, regardless of their net weight. It includes disposable equipment for sampling (lances), laboratory processing and analysis, control and cleaning of installations, covering powders for thermal isolation of the liquid steel bath during the secondary treatment and continuous casting etc.

The sum of all “heat-related” costs per year can be calculated as follows:

\[
\text{Total “heat-related” costs (€/year) = } \frac{\text{Production level (ton/year) \times “heat-related” costs for 1 heat (€/heat)}}{\text{Average heat size (ton/heat)}}
\]

A histogram of the achieved net weight for all heats in March 2012 is presented in Figure 3.2. The average weight is 300t.

**Histogram: achieved net weight**

![Histogram: average achieved net weight of full ladles (March 2012)](image)

Considering a steel production of 4.7Mton/year, which is the goal set for the steelshop for 2012, and “heat-related” costs for 1 heat of 2590 €/ton mentioned in Table 3.1, the total “heat-related” costs per year is calculated as:

\[
\text{Total “heat-related” costs (€/year)} = \frac{4.7 \cdot 10^6}{300} \times 2590 = 40,576,667.00 \text{ €/year}
\]

Supposing the average heat size could be increased to 303t (1% more), for instance, the total “heat-related” costs would be reduced to:

\[
\text{Total “heat-related” costs (€/year)} = \frac{4.7 \cdot 10^6}{303} \times 2590 = 40,174,917.00 \text{ €/year}
\]

In other words, an average heat size increase of 1% would lead to a cost reduction of 401,750.00 €/year. It should be stressed that this value is the minimum gain, calculated in a conservative approach and not considering the use of the extra capacity for fulfilling more order nor an eventual reduction of workforce due to the diminution of “heat-related” activities (refractory relining, installation control etc.) when keeping the same production level.
3.1.2 Defining the target heat size

A target net weight for a heat is defined before the oxygen blowing process, when the quantities of metal scrap and hot metal should be calculated. Among other things, this calculation has also to take into account the estimated weight losses during the slag removal and gains during alloying and eventually cooling. Some corrections can eventually be made on the weight during further processes, but setting the target and striving for it since the selection of the converter charge is the cheapest way of achieving the desired weight.

The main constraints for defining the target net weight are:

- The crane lifting capacity
- The converter volume capacity
- The ladle volume capacity (including the freeboard requirement)

According to their specifications, the cranes cannot lift more than 430ton of gross weight for more than 2% of the transport operations and never more than 450ton. Since the weight of the empty ladles (i.e. the tare weight) is measured with satisfactory accuracy and precision levels before each heat, it is possible to know the maximum steel weight with which the ladle can be filled according to this constraint.

The volume capacity of the ladles and of the converter constitutes the other constraints. In both cases, the available volume depends on the thickness of the refractory layer, which decreases according to the number of heats made after a refractory relining. In the past, the Process Technology staff used to make predictions on this volume based on regressions and other statistical techniques. Currently, it is measured by a laser scanning system before each heat, which output is presented as in Figure 3.3. Because the steel density is known, it is also possible to determine the maximum steel weight which the ladle can be filled with according to the volume constraints.

![Figure 3.3 – Laser scanning system for measurement of the available volume in converters and ladles](image-url)
In the case of the ladles, a freeboard, i.e. the distance between the top of the ladle and the steel bath level, of 50cm is required. This measure is taken to avoid that the steel would flow over during transport operations, risking the safety of the people and equipment in the plant. If this freeboard limit is not respected, the transportation has to be done very carefully, much slower and with possible evacuation procedures. The laser scanning system takes the freeboard into account when calculating the available volume, as show in Figure 3.4.

Each constraint determines a maximum allowed weight, which absolute value depends on operational circumstances such as the tare of the assigned ladle or the thickness of the refractory layer of the converter. If the probability distribution of the achieved net weight given a target is known or estimated, it should be possible to determine the maximum target weight for each heat in a way that all constraints (and respective tolerances), are respected in the long run.

As an example, consider that the tare of a ladle assigned for a given heat is 115ton. Since the tolerated risk of having a gross weight higher than 430ton is 2% according to the crane lifting capacity constraint, the maximum target weight should be defined in a way that the probability of achieving a net weight of more than 315ton (= 430ton – 115ton) is equal or less than 2%, or \( Pr(\text{achieved net weight} < 315\text{ton}) \leq 2\% \). Using the previous heats to estimate the probability density function of the achieved net weight given a target, it is done as illustrated in Figure 3.5. A similar approach should be used to the other constraints and the smallest of the maximum values should be chosen as the effective target.

---

1 In theory, a value higher than 2% could be allowed in this case since the weight will be limited by the other constraints for some of the heats, but this would make the analysis more complex.
Since the absolute value of the effective target depends on how the operational circumstances affect the constraints, an alternative way to represent the presented method which provides a better view of the performance of the process is shown in Figure 3.6. In this representation, all the curves are shifted by subtracting the effective target weight. This allows the inclusion of a histogram of the difference between the achieved and target net weight calculated for each heat.

Number of heats

The data shown in the figure represents all heats from March 2012. The probability density function (solid red line) was estimated based on these values, assuming they follow a normal distribution. The parameters calculated for this estimation are $\mu = -1.8$ ton (bias) and $\sigma = 4.7$ ton. This line is the same plotted in Figure 3.5, but in Figure 3.6 it is rescaled to fit the histogram.
In both figures, it is easy to visualize how a smaller variance would allow the target, and consequently the average achieved net weight, to be set closer to the limits defined by the constraints, as demonstrated by the dark grey dashed line.

Many reasons can be listed for the variation. Temperature model errors and logistic issues can influence the amount of cooling scrap which is actually added during the processes and consequently the achieved net weight. For instance, if the achieved temperature after the oxygen blowing process at the converter is too high, more cooling scrap has to be added in the secondary treatment and the achieved weight becomes higher than expected. On the other hand, if a full ladle has to wait too long before it is taken to the secondary metallurgy installation due to unexpected operational delays, less cooling scrap is added and the achieved weight becomes lower than the expected.

The weight model itself, which is implemented in the Knowledge Base System and executed when each heat is planned, is also a source of variation. Taking the steel grade specifications, the above-mentioned constraints and other process variables (raw material availability, production planning etc.) into account, this model estimates gain and losses in each process and defines a target weight and the material input (hot metal, metal scrap, alloying additions, cooling scrap etc.) for each heat. The achieved net weight depends on the accuracy and precision of the estimations.

Errors from other models, such as the alloying model discussed further in the work (Section 3.2), measurement deviations and unforeseen changes in the production planning are other examples of sources of variability. To assure this variability is in control, the difference between achieved and target net weight is calculated after each heat and followed by OSPC, as shown in Figure 3.7. An I-chart is used to monitor the evolution of this variable.

Figure 3.7 – The difference between achieved and target net weight is followed by OSPC

The use of OSPC helps maintaining the variance stable and, as a consequence, the average achieved weight and the productivity at the current levels. However, no extra information is given which could allow a further reduction of the variance and an increase of the average weight and of the productivity level.
3.1.3 Using the spc tool to increase heat size

The variables *achieved* and *target net weight* are also available in the AMG toolbox and can be used in the *spc* tool. Figure 3.8 shows how the chart made in OSPC (Figure 3.7) would be presented in Matlab.

![Figure 3.8 – The same chart made in OSPC can be made with the spc tool](image)

Many other process variables can be loaded and selected as categories for the RCA algorithm. Using data from March 2012, the RCA algorithm was executed for three different categories: converter, shift team and steel grade group. The control rules *we2*, *we3*, *we6* and *we7* (Figure 2.16) were used and the RCA alarms were defined as described in Table 3.2.

<table>
<thead>
<tr>
<th>RED</th>
<th>positive z-test AND positive t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ORANGE</td>
<td>(negative t-test AND positive z-test) OR (positive t-test AND negative z-test)</td>
</tr>
<tr>
<td>GREEN</td>
<td>no positive test</td>
</tr>
</tbody>
</table>

Table 3.2 – RCA alarms for heat size evaluation

For the first two categories (converter and shift team), orange alarms were given to all category entries because of positive t-tests, indicating that the average (bias) for every group was significantly different from the target (zero), but the z-tests were all negative, indicating that the proportion of out-of-control situations was not significantly different between the groups. For the third category, however, red alarms were given to two of the steel grade groups, as presented in Figure 3.9.

![Figure 3.9 – Influence of steel grade group (nuancegroep) on the weight calculation is detected by spc](image)
The different behavior of the two groups (T and TB) and the other can be observed in the boxplot made with the spc tool and shown in Figure 3.10.

![Boxplot](image)

Figure 3.10 – Different behavior between groups with and without RCA alarm for the weight

The steel grade group determines process parameters and production routes which permit the achievement of determined steel grades. The steel grade itself determines the ranges for the concentration of each chemical element and consequently the addition of alloying and other materials.

Generally speaking, the bias of the process indicates the average of unexpected weight loss or gain. A significantly different bias detected within a group is a hint for the physical reason for this. Using this hint, the model – or maybe the process as whole – can be improved. This would require some effort and investment, but the first step, i.e. the acknowledgement that some groups behave differently, is taken. This shows how the spc tool can boost the knowledge generation within the company.

Meanwhile, and as a temporary solution, the calculation of the weight losses for the steel grade groups with red alarms can be offset by the difference between the biases of these groups and the other. As demonstrated in Figure 3.11, the difference should be 2.9ton. By doing that, a higher target would automatically be set by the model for the “bad” groups.

![Table and Diagram](image)

Figure 3.11 – Target net weight can be offset by 2.9ton for heats from steel grades T and TB
Two types of gains are expected from this change:

1. Direct gains

For all heats from steel grades T and TB, the average net weight would be increased by 2.9ton. As shown in Figure 3.12, these heats represent around 25% of the steelshop production. This would result in a 0.7ton (= 2.9 · 25%) increase of the overall average net weight.

![Figure 3.12 – Pareto analysis: number of heats per steel grade group](image)

2. Indirect gains

After implementing the change, the overall variance is expected to be reduced by 0.2ton. This value is found by adding 2.9ton to the variable \( \text{achieved} - \text{target net weight} \) for the heats from steel grades T and TB and using the standard deviation to estimate the new variance. This variance reduction would allow a 0.2ton increase of the target and, as a consequence, of the average net weight for all heats.

The expected increase of the average net weight is thus 0.9ton (= 0.7+0.2). The Total “heat-related” costs per year would be reduced to:

\[
\text{Total “heat-related” costs (€/year)}_{\text{avg. heat size}=300.9\text{ton}} = (4.7 \cdot 10^6 / 300.9) \cdot 2590 = 40,455,301.00 \text{ €/year}
\]

In other words, the change, which had no implementation costs, has led to savings of:

\[
40,576,667.00 - 40,455,301.00 = 121,366.00 \text{ €/year}
\]
3.2 Reducing alloying costs

The chemical composition of the steel is a determining factor for properties such as hardness, ductility, and tensile strength. The desired levels of these properties depend on the final application for which the steel is designed. Different levels are required by the automotive, the construction, the household appliances and the packaging sectors. The so-called steel grades were created to standardize these requirements through the definition of ranges for the presence of the chemical elements in the steel.

Many of the chemical elements included in the specifications, such as chromium and molybdenum, are not present in sufficient proportions in the hot metal or in the metal scrap used to produce liquid steel. For this reason, alloying materials are added during the secondary treatment. Examples of alloys used in this process include ferromanganese (FeMn), ferrovanadium (FeV), ferroniobium (FeNb), ferroboron (FeB), ferromolybdenum (FeMo) and ferrochromium (FeCr).

Oppositely, some elements, such as oxygen and silicon, are present in the liquid steel after the converter process in proportions which are often higher than specified. Other materials are then used during the secondary treatment in order to remove them. For instance, aluminium blocks are used to remove the oxygen by forming aluminium oxides, which go to the slag on the top of the steel bath and are not cast during the continuous casting process.

The precision and accuracy levels in which the desired specifications are achieved depend to a great extent on the ability to predict the composition of the steel after the processes. For this reason, this ability plays an important role for the quality of the final products and for the cost of raw materials. In other words, the effectiveness and efficiency (and thus the productivity) of the business, are highly affected by the prediction error. Reducing the error would allow the steelshop to decrease the fraction of products which do not fulfill the specifications and, at the same time, reduce the raw material costs, leading to higher productivity levels, as illustrated in Figure 3.13.

![Figure 3.13 – How productivity gains can be achieved by reducing the prediction error](image)

In the next subsections, the costs caused by prediction errors on the chemical composition of the steel are discussed and the approach used to define the target concentration for each element is presented. Then the methods applied to pursue the targets are described and finally an example of how the spc tool can be used to reduce the errors is given.
3.2.1 The costs of prediction errors and the determination of the target

The costs of prediction errors can be divided in two main types: *downgrade costs* and *exceeding raw material costs*. The first type incurs when the chemical composition of a heat does not fulfill the specification requirements. In this case, the product is downgraded, i.e., considered to be from a poorer grade or quality. Possible consequences of this for the business are:

- Customers pay less for the products (decreased added value)
- Customers wait longer for the right product (increased lead time)
- Customers change from supplier (decreased market share)
- Higher levels of internally rejected products are reached (increased variable costs)
- Higher levels of safety stocks of “good” products are needed (increased fixed costs)
- ...

The second type of costs, *exceeding raw material costs*, is the difference between the costs of the minimum necessary raw material to produce the steel with the required specifications and the amount of raw material which is actually used.

Figure 3.14 illustrates how these two types of costs are related to the probability density function of the concentration of an element for heats from a determined steel grade. Following the same line of thought used in the example given in Section 1.3, it should be possible to determine a value for the target which would minimize Taguchi’s Loss Function given the variance of the variable (*achieved – target concentration*). \(^1\)

![Figure 3.14 – Costs caused by prediction errors on the expected concentration of an element](image)

---

\(^1\) In some cases, the additions used in the secondary metallurgy affect the concentration of more than one element and the minimization of the separate Taguchi’s Loss Functions can lead to infeasible or suboptimal combinations of targets. Therefore, Linear Programming techniques are used to define the targets in such a way that the sum of all costs are minimized, considering the effects and prices of each addition. This, however, does not impact the analysis made in this work.
3.2.2 Pursuing and predicting the target chemical composition of a heat

In order to determine the necessary additions to achieve the steel grade specifications during the secondary treatment, an alloying model is implemented in the Knowledge Base System. The main input parameters for this model are:

- The specified ranges for the concentration of the chemical elements (steel grade specifications)
- The results of chemical analysis of liquid steel samples, which are taken before each phase of the process and sent by pneumatic tubes from the steelshop to a nearby chemical laboratory
- The estimated variance of the variable \((\text{achieved} - \text{target concentration})\) from previous heats to determine the target
- The costs of the additions and their expected effect on the chemical composition
- The estimated downgrade costs
- Process and operational variables such as steel mass, temperature, production route and planning etc.

Using these parameters, the model determines an optimal target for each chemical element and calculates the amount of each type of addition needed to achieve the desired composition. The results are sent to the process computers, evaluated (and eventually adapted) by the operators and sent to the PLCs.

At the end of the process, a final sample is taken to check if the steel grade requirements were achieved. The difference between the results of this sample and the expected value by the model for each element is the prediction (or model) error, i.e. the variable \((\text{achieved} - \text{target concentration})\). This variable is followed up in OSPC and is also available in the AMG Toolbox. The I-chart presented in Figure 3.15 is made using the spc tool.

![Mn concentration: prediction error per heat](image)

*Figure 3.15 – Prediction error on the Mn concentration per heat*
3.2.3 Using the spc tool to reduce prediction errors of the chemical composition

The steel grade of a heat has direct and indirect effects on the model input. The direct effects are linked to the specification limits and the indirect effects are linked to the other process and operational variables. As mentioned in Section 3.1, heats from different steel grades go through different production routes (e.g. production installations) and respect different process parameters (e.g. ideal temperature for alloying and casting). They are therefore a possible source of prediction errors and were chosen as the category for the RCA algorithm in the spc tool.

Using data between May 2011 and April 2012, the prediction error of various elements were evaluated. The same control rules and alarm criteria described in Section 3.1 were applied. As an example, the output for the error on the Mn concentration is shown in Figure 3.16.

<table>
<thead>
<tr>
<th>nuance</th>
<th>count</th>
<th>Fraction_occ</th>
<th>mean</th>
<th>std</th>
<th>ztest</th>
<th>ttest</th>
<th>vertest</th>
<th>runtest</th>
<th>cusum_occ</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK80</td>
<td>31</td>
<td>0.8710</td>
<td>136.7742</td>
<td>67.6677</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>KK50</td>
<td>31</td>
<td>0.6774</td>
<td>65.0452</td>
<td>45.4060</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K450</td>
<td>37</td>
<td>0.2703</td>
<td>61.6757</td>
<td>47.8046</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>KH50</td>
<td>37</td>
<td>0.1351</td>
<td>-33.7569</td>
<td>94.4144</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K340</td>
<td>16</td>
<td>0.1250</td>
<td>-59.6875</td>
<td>71.8273</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>KM40</td>
<td>56</td>
<td>0.0714</td>
<td>-26.3038</td>
<td>69.0870</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>KM50</td>
<td>98</td>
<td>0.0408</td>
<td>-33.8980</td>
<td>65.6783</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K350</td>
<td>72</td>
<td>0.0278</td>
<td>-26.6389</td>
<td>60.9942</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K570</td>
<td>37</td>
<td>0.0270</td>
<td>22.8108</td>
<td>52.7853</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KT10</td>
<td>24</td>
<td>0</td>
<td>22.2917</td>
<td>46.6481</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KM60</td>
<td>23</td>
<td>0</td>
<td>23.4783</td>
<td>33.2714</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KM30</td>
<td>48</td>
<td>0</td>
<td>15.5208</td>
<td>42.4349</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K390</td>
<td>16</td>
<td>0</td>
<td>15.6250</td>
<td>16.5040</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K120</td>
<td>75</td>
<td>0</td>
<td>19.1067</td>
<td>33.0653</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K140</td>
<td>23</td>
<td>0.1739</td>
<td>-31.7523</td>
<td>67.5305</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KT10</td>
<td>10</td>
<td>0.1000</td>
<td>-9.5000</td>
<td>84.9382</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KT30</td>
<td>21</td>
<td>0.0952</td>
<td>6.6190</td>
<td>96.7460</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KT70</td>
<td>34</td>
<td>0.0882</td>
<td>17.8235</td>
<td>70.7537</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K140</td>
<td>30</td>
<td>0.0687</td>
<td>14.6667</td>
<td>53.2658</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K120</td>
<td>67</td>
<td>0.0597</td>
<td>-0.7164</td>
<td>69.9052</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>KI70</td>
<td>62</td>
<td>0.0484</td>
<td>8.3065</td>
<td>58.0290</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K170</td>
<td>28</td>
<td>0.0357</td>
<td>11</td>
<td>41.4845</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K320</td>
<td>298</td>
<td>0.0362</td>
<td>0.4128</td>
<td>54.8241</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K130</td>
<td>166</td>
<td>0.0241</td>
<td>12.4035</td>
<td>47.5037</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K210</td>
<td>209</td>
<td>0.0239</td>
<td>-12.5407</td>
<td>55.8300</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K370</td>
<td>396</td>
<td>0.0227</td>
<td>-6.4590</td>
<td>50.4600</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>KM80</td>
<td>58</td>
<td>0.0172</td>
<td>-3.8103</td>
<td>61.1488</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>K320</td>
<td>273</td>
<td>0.0073</td>
<td>-7.0596</td>
<td>45.9730</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>K580</td>
<td>15</td>
<td>0</td>
<td>51.6000</td>
<td>35.9261</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3.16 – Influence of steel grade group (nuance) on the prediction error is detected by spc

As observed, the RCA algorithm detected that the prediction error for the steel grades KK80, KK50 and K450 is significantly higher than for the other grades and that a significantly higher number of alarms is given for these grades when the control charts are made separately. The boxplot made with the spc tool shown in Figure 3.17 allows the comparison between the groups with and without the red alarm.
The actual *downgrade costs* caused by the Mn concentration for the mentioned steel grades during the analyzed period were not relevant. No heat from these groups was downgraded because the Mn concentration did not fulfill the specifications.

Nevertheless, the difference between the expected and the actual *exceeding raw material costs* can be calculated. As the Mn concentration is always higher than the target for the steel grades in question, it can be concluded that too much alloying material was added. This can be considered as the costs of the prediction error.

The cheapest ferroalloy which raises the Mn concentration is the Standard Ferromanganese. This alloy is cheaper because it also raises the concentration of other elements such as carbon and silicon, and cannot be used in high quantities for every heat, especially when the concentration of these elements is a limiting factor. In any case, this alloy can be used as a reference to determine the minimum avoidable costs.

Standard Ferromanganese prices are around 950€/ton (0.95€/kg). It is considered that the Mn concentration of an average-sized heat increases approximately 2.6ppm with the addition of 1kg of this alloy. The costs which could be avoided with a reduction of the prediction error for the steel grades identified by the *spc* tool are calculated as shown in Table 3.3.

<table>
<thead>
<tr>
<th>Steel grade</th>
<th>Average error (ppm)</th>
<th>Wasted Standard Ferromanganese per heat (kg/heat)</th>
<th>Heats per year(^1)</th>
<th>Avoidable costs (€/year)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KK80</td>
<td>136</td>
<td>((136 / 2.6) = 52.3)</td>
<td>73</td>
<td>((52.3 \cdot 73 \cdot 0.95) = € 3,627.00)</td>
</tr>
<tr>
<td>KK50</td>
<td>86</td>
<td>((86 / 2.6) = 33.1)</td>
<td>129</td>
<td>((33.1 \cdot 129 \cdot 0.95) = € 4,056.40)</td>
</tr>
<tr>
<td>K450</td>
<td>62</td>
<td>((62 / 2.6) = 23.8)</td>
<td>94</td>
<td>((23.8 \cdot 94 \cdot 0.95) = € 2,125.34)</td>
</tr>
</tbody>
</table>

Table 3.3 – Calculation of costs which can be avoided by reducing the prediction error

Reducing the prediction error for these steel grades would also reduce the overall variance and allow a further reduction of the costs for the other grades. Repeating this exercise for the other elements (and other categories), substantial gains can be achieved.

---

\(^1\) The number of heats shown in Table 3.3 is based on the company’s records for the period between May 2011 and April 2012 (same period used for the RCA), but it is much higher than the count shown in Figure 3.16 because the prediction error could not be found for all the heats.
3.3 Finding best practices to avoid material losses

Material losses constitute a highly concerning issue for the steelshop. Some examples of losses were given in the Subsection 3.1.1. They include the liquid steel which is disposed during slag removal operations and the content of the steel ladle which cannot be cast in the continuous casting process.

It is estimated that 15.5 ton of liquid steel is lost per heat during the processes. Part of the steel is recovered after solidifying, processed by an external company and used as metal scrap in future heats. Nevertheless, the wasted raw material, energy and processing costs account for economic losses of approximately 95 €/ton. Obviously, it is vital for the productivity of the steelshop to control and preferably reduce these losses.

Figure 3.18 shows a daily report available in the company’s intranet used to monitor the steel losses.

![Staalverlies per lading](image)

Figure 3.18 – Report on the company’s intranet shows the liquid steel losses per heat

In this section the reasons and methods for slag removal operations are explained, information concerning the losses during these operations is presented and an application of the spc tool to reduce the losses is demonstrated.
3.3.1 Reasons and methods for slag removal operations

Slag is formed during different steps of the steelmaking. In general, it consists mostly from elements which are not wanted in the final product. During the desulphurization process, for instance, calcium carbide is used to decrease the sulphur concentration from the hot metal. When these substances react, slag rich in sulphur is formed and goes to the top of the bath. This slag must be removed before the hot metal is cast into the converter, otherwise the sulphur will be present in the steel.

During the oxygen blowing process in the converter, slag is also formed, consisting mostly of calcium, phosphorus, silicon, magnesium, manganese and iron oxides. This slag must be removed before the secondary treatment for an adequate alloying process.

In order to minimize the carry over of slag to the steel ladle, the content of the converter is tapped in two steps: first into the ladle and then to a slag pot, as shown in Figure 3.19. A special camera detects the moment when the slag begins to be cast and sends a signal to the PLCs to turn the converter to the opposite direction. The slag is then poured in a smaller ladle (slag pot).

Because the movement of the converter is relatively slow, this procedure does not prevent some slag from being cast into the steel ladle. There are devices available in the market designed to avoid the carry over of the slag. The so-called slag stoppers are an example of these devices. They are electromechanical arms coupled to the converter which block the tapping hole in fractions of seconds when the signal from the camera is received.

Naturally, the use of such devices involves considerable implementation and operational costs. ArcelorMittal Gent has not made these investments and tried to develop a technique called slag free tapping internally. It consists in placing a slag pot close to the steel ladle in the transfer car, which moves forward when slag is detected by the camera. The slag free tapping technique is represented in Figure 3.20.
The movement of the transfer car is faster than the converter, but not enough to avoid the carry over of the slag completely. The proposed solution was to move the transfer car in advance, allowing some steel to go to the slag pot on it. Not all slag is cast into this pot: at the same time the transfer car moves, the converter turns to the opposite side (as usually) and the rest of the slag is cast into another pot. The original idea included using the content of the first slag pot as a hot input for the subsequent converter heat. This would lower the material and energy losses, reduce the processing costs and eliminate the traditional slag removal operation before the secondary treatment explained hereafter.

The implementation of the slag free tapping was however not as successful as expected. Using the content of the first slag pot turned out to be infeasible due to logistic issues, the steel losses were not reduced significantly and the increase of the operational costs (e.g. refractory relining of two slag pots instead of one) was considerable. The technique is currently used for about 41% of the heats, as shown in Figure 3.21.

![Slag Free Tapping (April 2012)](image)

Figure 3.21 – Number of heats for which the slag free tapping technique was planned

For the other heats, the traditional slag removal operation takes place in the deslagging station, where a device called slag skimmer is available. Generally speaking, the slag has a lower density than the steel and floats on top of the bath. An electromechanical structure allows the slag skimmer to travel along the surface and pull the slag out of the ladle, which is slightly inclined. The slag removal operation is illustrated in Figure 3.22. The movements and the depth of the slag skimmer and the inclination of the ladle are manually controlled by the operators. The objective of the operation is to remove the slag (almost) completely in an adequate time, disposing minimum quantities of steel.

![Slag removal operation](image)

Figure 3.22 – Slag removal operation
3.3.2 Losses concerning slag removal operations

The main losses related to slag removal operations are caused by inadequate slag removal and excessive steel disposal. The first cause has a negative impact on the secondary treatment. The yield of the alloying elements and other addition is reduced and more material has to be used in order to achieve the steel grade specifications. Furthermore, the steel is contaminated by the unwanted elements present in the slag and the risk of other quality problems increases.

The disposal of steel along with the slag constitutes the second main cause of losses. As mentioned in the beginning of this section, these losses are estimated in 95€/ton.

The two causes of losses are somehow connected: it is easier to remove more slag by disposing more steel along with it. However, this relation is not so straightforward; a carefully executed operation can remove the slag adequately and dispose a relative small quantity of steel, while a poorly executed operation disposes more steel while removing an unsatisfactory amount of slag. The losses should thus be evaluated separately.

The second cause of losses, i.e. the amount of steel disposed during slag removal operations, is the focus of the analysis presented in this section. The quantities of steel and slag in the ladle before the slag removal operation cannot be measured apart from each other. Instead, the weight of the slag is estimated based on its depth, which can be measured using a lance and a ruler. Since the dimensions of the ladle and the average density of the slag are known, it is possible to estimate its weight. Assuming that the remaining slag weight after the operation is negligible, the weight of the disposed steel can be calculated according to the equation represented in Figure 3.23.

![Figure 3.23 – Calculation of the disposed steel weight](image)

The total disposed weight (slag and steel) is followed up in OSCP, as shown in Figure 3.24. The estimated disposed steel weight was not chosen as control variable to avoid the error on the estimation of the slag weight to be appointed as cause of deviations. Based on operational experience, the Process Staff considers that slag removal operations can be satisfactorily executed keeping an average of 4ton for the total disposed weight and that it leads to an average disposed steel weight of 3ton per heat.

![Figure 3.24 – The amount of steel disposed during slag removal operations is followed up in OSCP](image)
Figure 3.25 presents a histogram of the disposed steel weight per heat during April 2012. The negative values are found due to measurement and estimation errors, but considering that these errors are random and occur in both directions (positive and negative), they do not impact the analysis made in this section. The average disposed steel weight per heat in the mentioned period was 3.5ton, which is 0.5ton higher than the target, which has previously been achieved in the department.

**Histogram: disposed steel weight per heat**

![Histogram: disposed steel weight per heat](image)

Using the heat size of 300ton/heat and the production target of 4.7Mton/year for 2012 as reference (values mentioned in Section 3.1), and considering that it should be possible to avoid the disposal of liquid steel which exceeds the target (i.e., avoid disposing 0.5ton/heat), the following calculations can be made:

**Number of heats per year which go through the slag removal operation:**

\[
59\% \cdot 4.7 \cdot 10^6 \text{ Mton/year} / 300 \text{ ton/heat} = 9,243 \text{ heats}
\]

**Avoidable liquid steel losses during slag removal operations per year:**

\[
9,243 \text{ heats} \cdot 0.5 \text{ ton/heat} = 4,622 \text{ ton}
\]

**Avoidable economic losses:**

\[
95.00 \text{ €/ton} \cdot 4,622 \text{ ton} = 439,058.00 \text{ €}
\]
3.3.3 Using the spc tool to spot opportunities to avoid material losses

The slag removal operation involves the ability of the operators to control and synchronize the movements of the slag skimmer and the inclination of the steel ladle. Productivity improvements can be achieved by identifying the best operators, from who the best practices could be learned, and the worst, who should be alerted and receive training. The spc tool was used for this purpose and the results are presented in this subsection.

The steelshop does not keep record of the operator responsible for each operation, but the shift team which was in charge is stored and it was chosen as a category for the RCA in the spc tool. Data from April 2012 was used for the analysis. The I-chart was selected, the parameter n-sigma for the calculation of the control limits was defined as 2, no special control rule (Western Electric or Nelson) was considered and the criteria for the RCA alarms were the same as the examples given in Sections 3.1 and 3.2. The output of the program is shown in Figure 3.26.

<table>
<thead>
<tr>
<th>ploeg</th>
<th>count</th>
<th>fraction_occ</th>
<th>mean</th>
<th>std</th>
<th>stest</th>
<th>ttest</th>
<th>var</th>
<th>runtest</th>
<th>cusumocc</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>150</td>
<td>0.1600</td>
<td>4.3547</td>
<td>2.8855</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>D</td>
<td>151</td>
<td>0.0662</td>
<td>3.7980</td>
<td>2.3225</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>205</td>
<td>0.0732</td>
<td>3.4412</td>
<td>2.1733</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>A</td>
<td>195</td>
<td>0.0154</td>
<td>2.8044</td>
<td>1.8988</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>

Figure 3.26 – Significant differences in the performance of the shift teams is detected by the spc tool

No alarms were given for the shift teams A and B, which indicates that the average steel weight disposed during slag removal operations performed by these groups is not significantly different from the target (3ton) and that the proportion of out-of-control situations detected when the control charts are plotted separately is not significantly higher than for the other groups.

An orange alarm was given for the shift team D because of a positive t-test, which indicates that the average disposed steel weight is significantly higher than the target for this group.

A red alarm was give for the shift team C since both t-test and z-test were positive. Not only the average disposed steel weight is higher than the target, but the proportion of out-of-control situations detected by the control chart for this group is also significantly higher when compared to the other groups.

The different performance of the shift teams can also be observed in the boxplot presented in Figure 3.27.
The control charts used by the algorithm for the shift teams A and B are shown in Figure 3.28. The different behavior of the plotted variable can be easily perceived.

![Shift team A chart](image)

![Shift team C chart](image)

Figure 3.28 – Separate charts for shift teams A and C

Sharing and standardizing the techniques used by the shift team A can allow the achievement of the target of 3ton for the average disposed steel weight – and even make it smaller.
Chapter 4

Conclusion

Most management systems used for process, quality or project control and follow-up have Continuous Improvement as one of the goals. Modern management strategies, especially the ones based on Kaizen or Six Sigma principles, suggest the implementation of small improvements and standardization of processes, continually aligned to the organization’s purposes and guided by Statistical Process Control techniques, in order to achieve this goal. In this context, the features of new tool presented in this work reinforce Continuous Improvement in the steelshop by expanding the possibilities of the traditional Control Charts.

In terms of Knowledge Management, the traditional control charts generate information through detecting out-of-control situations using raw process data as input. The approach implemented in the new tool goes further, allowing the identification of patterns in the behavior of process variables by detecting properties which significantly account for a less or more out-of-control situations. This idea is represented in Figure 4.1.

While the traditional approach helps maintaining the processes and the productivity levels stable, the new tool generates knowledge which allows the reduction of processes variability and leads to productivity improvements.

In this Chapter, the experiences and lessons learned during the development and application of the tool are shared, an example to illustrate how the company should use these lessons is presented and suggestions for further development are given.
4.1 Lessons Learned

During the application of the tool in the steelshop at ArcelorMittal Gent, it could be observed that, although the Process Staff is skilled and takes decisions based on a statistical mindset, they do not make a frequent use of more complex control charts types such as EWMA, CUSUM, mA charts, while I-charts are overused.

The examples given in Chapter 3 intended to emphasize the RCA algorithm included in the new tool, a feature which was not available in the previous software. In this context, the same charts designed and used for years in OSPC were remade with the spc tool and the new possibilities were explored. Good results were achieved and it is possible to conclude that the objectives listed in Chapter 1 are accomplished.

However, as it may have been noticed, the choice has always been for I-charts, despite the fact that both tools allow the creation of many other charts types, which could give better results depending on the situation. The application shown in Section 3.3 is an example of this. The probability distribution of the estimated disposed steel weight per heat is clearly not normal, as observed in Figure 3.25, and Xbar-Charts would be more advisable. These charts, made using the same data as for Figure 3.28, are presented in Figure 4.2. Not many points are necessary to realize the discrepancy in the performance of the different shift teams, which is not the case when the values per heat are plotted (I-charts). If the correct chart would have been chosen, the RCA algorithm would possibly be unnecessary to detect the divergence.

Figure 4.2 – Xbar-Chart for the disposed steel weight (example given in Section 3.3)
The general perception is that control charts which require grouping data from more batches (e.g. p-chart for reblow rate as described in Example nr. 2 of Subsection 2.2.3) are hardly used. A possible reason for this is the strong emphasis on the fast detection and solution of problems. This leads to the idea that online charts should display data from every single heat.

As a matter of fact, only major problems can be detected this way, i.e. without grouping data. Minor problems – which, when summed, account for most of the variability and costs in the long-term – can only be detected by checking the evolution and correlation of variables and their statistics in representative time intervals.

Although the Process Staff of the steelshop at ArcelorMittal is aware of this and uses tools such as dex (presented in Subsection 2.1.3) for coherent statistical analyses, the referred minor problems are only detected and studied when their consequences become more serious or visible. Then, besides the losses already incurred, considerable time and money are spent on trying to gather information from the past, which is frequently no longer available or has dubious precision and accuracy.

The control is thus not always implemented as consistently and systematically as recommended by the SPC guidelines. It may be important to stress that the above-mentioned minor problems are special causes of variability – though they can hardly be distinguished from common causes if not clearly identified – and it should be possible to design control charts (with grouped data) to detect them.

Summing this up, there is room to improve the check stage of the PDCA methodology. The new tool and techniques explained in this work can aid this, but, as before, the level of success of the implementation depends on the familiarity of the users with general SPC concepts and on the maturity with which they can design and employ control charts, evaluate capability indices etc.

### 4.2 A Late Example

An example to illustrate the message of the last section is the follow-up of the cooling coefficients of the additions used during the steelmaking processes. As mentioned in Chapter 1, when the temperature of the liquid steel is too high for the subsequent production steps, it can be reduced by adding cooling material. Different types of additions can be used and the amount and type of each addition which is actually added depend on their expected effects on the temperature and on the chemical composition of the steel.

The effect on the temperature is described by cooling coefficients, usually expressed in °C/(kg of added material/ton of steel). For instance, if 1200kg of determined addition which cooling coefficient is 2°C/(kg/ton) is added to a 300ton heat, the temperature is expected to decrease approximately 8°C (= 2 \cdot 1200 / 300). As explained in Section 1.3, the performance of the steelshop depends on the precision and accuracy in which the aimed temperatures are achieved, and the correctness of the cooling coefficient estimations is crucially relevant for this.

The physical properties of the materials determine their cooling coefficients. Studying these properties is the first step to estimate the coefficients of each addition. However, the additions are constituted by different mixes of materials, in which the proportion and the properties of each material is not always precisely known. Statistical techniques, such as linear regression, are used to adjust the first estimation.

The accuracy of the estimation for the cooling coefficient for an addition can be evaluated by checking the linear coefficient of the regression of the Prediction Error on Temperature (°C) (i.e. achieved minus expected temperature) on Amount of Addition (kg) for previous heats. If the cooling coefficient is correctly estimated, the expected value for the linear regression coefficient is zero. In other words, if the estimated cooling coefficient corresponds to the real coefficient, the Prediction Error on Temperature should not depend on the Amount of Addition.
Two graphs representing the *Prediction Error on Temperature versus Amount of Addition* for an addition internally known as *slibbri* or *briketten uit slib* are shown in Figure 4.3. The graphs were made using data from November 2011 and April 2012 respectively.

**Prediction Error on Temperature versus "Slibbri"**

**November 2011**

\[ Y = -0.924869 + (3.86411 \times 10^{-5}) \times X \]

**Prediction Error on Temperature versus "Slibbri"**

**April 2012**

\[ Y = -2.69174 + (0.00133869) \times X \]

*Figure 4.3 – Prediction Error on Temperature vs. Amount of Addition (slibbri)*
At a 95% confidence level, the linear coefficient of the regression for the first graph is not significantly different from zero, while for the second it is. The regression coefficients and their corresponded p-values are shown in Table 4.1.

<table>
<thead>
<tr>
<th>Graph</th>
<th>Linear Regression Coefficient</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>November 2011</td>
<td>-3.86 $\times 10^{-5}$</td>
<td>0.9140</td>
</tr>
<tr>
<td>April 2012</td>
<td>-1.33 $\times 10^{-3}$</td>
<td>&lt; 0.0002</td>
</tr>
</tbody>
</table>

Table 4.1 – Regression coefficients and p-values

Surprisingly, the cooling coefficient used by the model was not adapted between November 2011 and April 2012. The real cooling coefficient was actually increased due to modifications on the constitution of the addition (i.e. a problem at the raw material supplier), but the effect of this change was unnoticed by the Process Staff and not registered on the model. As a consequence, the achieved temperatures began to be lower than expected, especially when high quantities of slib bri were added, resulting in higher process variability, higher reblow rate and higher costs.

Reasons why following the evolution of the prediction error on temperature for each heat using an I-chart as traditionally done in the steelshop was not sufficient to detect the a change include:

- Many other variables influence the prediction error
- The introduction of other major and minor process changes is frequent
- There is a considerable spread in the amount of slib bri which is added to each heat and the effects of the inaccurate estimation depend on this amount.

The estimated cooling coefficients of the additions are only reevaluated when the rising costs of the process call the attention of management. But the deterioration of the temperature prediction due to the change on the addition could have been detected much earlier and the related losses avoided by using a control chart to follow the evolution of the regression coefficients, calculated periodically as demonstrated in Figure 4.4. Each point on the graph represents the linear regression coefficient calculated for a group of 100 heats. The moment in which the process has changed can be easily visualized.

![Linear regression coefficients: Prediction Error on Temperature vs "Slibbri" (100 heats per regression)](image-url)
This approach is based on the fact that regression coefficients of variables from a stable process are random variables. If they begin to systematically deviate from the expected value, there is a high probability that the process has changed.

It should be remarked that, although an I-chart was chosen to follow the evolution of the regression coefficients, each point on the graph accounts for a period of 100 heats, which is representative for the process.

4.3 Further development

The examples given in Chapter 3 and the observations added in the present chapter were intended to accomplish the final objective of this work stated in Chapter 1, i.e. “demonstrate how the use of the new tool could increase the productivity of the department”. Considering this objective was achieved, the next step would be spreading the employed techniques and use the new tool in other departments and production sites.

In the software perspective, many opportunities of expanding the tool can be listed. User-friendliness can be improved (e.g. let user choose fonts and colors, create keyboard shortcuts, etc.), code can be optimized for faster performance, new possibilities to import data from other systems can be foreseen, setup assistants (“wizard”) for the creation of HTML reports can be provided, among others. As software developers often say, “a computer program is never finished”.

From the statistical analysis point of view, the approach used in the example given in the last section, i.e. the periodic evaluation of regression coefficients, could be implemented as an add-on for the RCA algorithm. This could be applied when the user wants to check the impact of continuous variables on a variable chosen for an I-chart. For example: the user would first use the tool to create an I-chart (e.g. Prediction Error on the Temperature per heat) and then choose other continuous variables (e.g. Amount of Slibbri added) for the RCA. Based on pre-defined parameters (e.g. number of points per regression and expected linear coefficient), the program would make the chart for the regression coefficients as the one shown in Figure 4.4.

In the business standpoint, the potential gains which can be generated by using the tool and the described techniques are immeasurable, as are the gains of applying SPC or PDCA methodologies. The actual gains, however, depend on the users’ knowledge about the production processes, ability to evaluate data and understanding of the concepts implemented by the tool. These factors are raised by experience and thus using the tool is actually the best way to improve its results.
References

[Asian Efficiency, 2011]
Thanh Pham. What is Productivity?
Retrieved from http://www.asianefficiency.com/productivity/what-is-productivity/

[ArcelorMittal Website, 2012]
ArcelorMittal Company. Public website
Retrieved from http://www.arcelormittal.com

[ArcelorMittal Gent Intranet, 2012]
ArcelorMittal Gent. Local intranet

[This is how ArcelorMittal Gent produces steel, 2010]
ArcelorMittal Gent. This is how ArcelorMittal Gent produces steel

[Staalbereding, 2008]
Frederik Engels, Philip Henrion and Marc Ladon. Cursus Siderurgie - 2de jaar


[Wikipedia: Accuracy and Precision, 2012]
Wikipedia. Accuracy and precision

[iSixSigma, 2012]
iSixSigma. Using Taguchi's Loss Function to Estimate Benefits

[Peter Ottoy, 2011]
Peter Ottoy. Lecture notes on Quality Engineering and Industrial Statistics (Ghent University)

[James Benneyan, 1998]
James C Benneyan. Use and interpretation of statistical quality control charts

[Western Electric, 1956]
Western Electric Company. Statistical Quality Control handbook
[Nelson, 1984]
Lloyd S. Nelson. *Technical Aids*
Published in the Journal of Quality Technology 16, no. 4 (October 1984), 238-239

[Gene Bellinger, 2004]
Gene Bellinger. *Knowledge Management – Emerging Perspectives*

[Wikipedia: Programmable logic controller, 2012]
Wikipedia. *Programmable logic controller*

[Christopher Date, 1999]
Christopher Date. *An Introduction to Database Systems*

[MATLAB, 2012]
MATLAB. *Product Overview*

[W & W, 1990]
List of Figures

Figure 1.1 – Productivity and Profitability .......................................................................................................... 5
Figure 1.2 – ArcelorMittal Gent [ArcelorMittal Gent Intranet, 2012] ................................................................. 6
Figure 1.3 – Blast furnaces at ArcelorMittal Gent [This is how ArcelorMittal Gent produces steel] .............. 7
Figure 1.4 – Hot metal production at the Blast Furnaces [This is how ArcelorMittal Gent produces steel] ....... 7
Figure 1.5 – Desulphurization and casting [This is how ArcelorMittal Gent produces steel] ..................... 8
Figure 1.6 – Casting of a torpedo car [This is how ArcelorMittal Gent produces steel] ........................................ 8
Figure 1.7 – Converter charging and oxygen blowing [This is how ArcelorMittal Gent produces steel] ........ 8
Figure 1.8 – Casting of a hot metal ladle [This is how ArcelorMittal Gent produces steel] .......................... 8
Figure 1.9 – Converter Tapping and Secondary Treatment [This is how ArcelorMittal Gent produces steel] .. 9
Figure 1.10 – Continuous casting machine (Turret) [This is how ArcelorMittal Gent produces steel] .......... 10
Figure 1.11 – Continuous Casting [This is how ArcelorMittal Gent produces steel] ...................................... 10
Figure 1.12 – Liquid steel solidifying into slabs ................................................................................................ 10
Figure 1.13 – Accuracy and Precision [Wikipedia: Accuracy and Precision, 2012] ........................................ 11
Figure 1.14 – Costs and probability density function of the temperature after oxygen blowing .................. 12
Figure 1.15 – Cost configuration with a higher target temperature ............................................................ 12
Figure 1.16 – Costs in function of target temperature .................................................................................. 13
Figure 1.17 – Cost configuration with a smaller variance in relation to the target temperature ................... 13
Figure 1.18 – Relation between process variance and optimal cost .......................................................... 13
Figure 1.19 – The PDCA cycle ....................................................................................................................... 15
Figure 1.20 – Most used control chart types ................................................................................................. 16
Figure 1.21 – X-chart example ....................................................................................................................... 16
Figure 1.22 – Western Electric Rules ............................................................................................................ 17
Figure 1.23 – An out-of-control situation that could be detected in advance if data is filtered beforehand . . . 18

Figure 2.1 – Knowledge generation [Gene Bellinger, 2004] ......................................................................... 20
Figure 2.2 – Siemens PLC (Model SIMATIC S7-400) ..................................................................................... 21
Figure 2.3 – Human Machine Interface (HMI) for the slag removal operation ............................................. 22
Figure 2.4 – Process Computer interface for the follow-up of the hot metal flow ....................................... 22
Figure 2.5 – Overview of the computer system used for process control in the steelshop ............................ 23
Figure 2.6 – Process Data acquisition Centre (PDC) .................................................................................... 24
Figure 2.7 – The Knowledge Base System (KBS) ......................................................................................... 25
Figure 2.8 – KBS control loop ....................................................................................................................... 26
Figure 2.9 – Online Statistical Process Control (OSPC) ............................................................................. 26
Figure 2.10 – \texttt{dex} ..................................................................................................................................... 27
Figure 2.11 – The \texttt{spc} interface ................................................................................................................. 29
Figure 2.12 – Loading a dataset from a MS Excel spreadsheet .................................................................... 30
Figure 2.13 – Selecting a variable ................................................................................................................ 30
Figure 2.14 – Reading Methods possibilities ............................................................................................... 31
Figure 2.15 – Panel to select the Chart Parameters ...................................................................................... 32
List of Tables

Table 1.1– Chemical composition of hot metal and steel [Staalbereding, 2008] ....................................................7
Table 1.2 – Chart zones for Western Electric Rules .................................................................................................17

Table 2.1 – Process variables, chart types and reading methods ............................................................................32
Table 2.2 – Control Chart parameters ....................................................................................................................33
Table 2.3 – Examples of Main Input for spc ...........................................................................................................34
Table 2.4 – Hypothesis tests in RCA ......................................................................................................................44
Table 2.5 – Results of tests and RCA alarms .........................................................................................................45
Table 2.6 – Types of errors in Control Charts .......................................................................................................45

Table 3.1 – Heat-Related costs ...............................................................................................................................50
Table 3.2 – RCA alarms for heat size evaluation ....................................................................................................56
Table 3.3 – Calculation of costs which can be avoided by reducing the prediction error ....................................63

Table 4.1 – Regression coefficients and p-values ..................................................................................................75
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AMG Toolbox:</td>
<td>ArcelorMittal Gent MATLAB toolbox</td>
</tr>
<tr>
<td>AMIGO:</td>
<td>ArcelorMittal Integrated Global Optimization</td>
</tr>
<tr>
<td>c-chart:</td>
<td>Count of defects in subgroup (Control Chart)</td>
</tr>
<tr>
<td>CL:</td>
<td>Centerline</td>
</tr>
<tr>
<td>CUSUM-chart:</td>
<td>Cumulative sum of individual measurements or subgroup averages (Control Chart)</td>
</tr>
<tr>
<td>dex:</td>
<td>Data Exploration tool (AMG Toolbox)</td>
</tr>
<tr>
<td>EWMA-chart:</td>
<td>Exponentially weighted moving average (Control Chart)</td>
</tr>
<tr>
<td>HMI:</td>
<td>Human-Machine Interface</td>
</tr>
<tr>
<td>I-chart:</td>
<td>Individual measurements (also referred as X-chart) (Control Chart)</td>
</tr>
<tr>
<td>KBS:</td>
<td>Knowledge Base System</td>
</tr>
<tr>
<td>KOORDI:</td>
<td>Steelshop Coordination System</td>
</tr>
<tr>
<td>KPI:</td>
<td>Key Performance Indicator</td>
</tr>
<tr>
<td>LCL:</td>
<td>Lower Control Limit</td>
</tr>
<tr>
<td>LSL:</td>
<td>Lower Specification Limit</td>
</tr>
<tr>
<td>mA-chart:</td>
<td>Moving average of the individual measurements (Control Chart)</td>
</tr>
<tr>
<td>mR-chart:</td>
<td>Moving range of the individual measurements (Control Chart)</td>
</tr>
<tr>
<td>n1, n2, ...</td>
<td>Nelson Rules</td>
</tr>
<tr>
<td>np-chart:</td>
<td>Number of defective units in subgroup (Control Chart)</td>
</tr>
<tr>
<td>OCAP:</td>
<td>Out of Control Action Plan</td>
</tr>
<tr>
<td>OOC:</td>
<td>Out-of-control</td>
</tr>
<tr>
<td>OSPC:</td>
<td>Online Statistical Process Control</td>
</tr>
<tr>
<td>p-chart:</td>
<td>Proportion of defective units in subgroup (Control Chart)</td>
</tr>
<tr>
<td>PDC:</td>
<td>Process Data acquisition Centre</td>
</tr>
<tr>
<td>PDCA:</td>
<td>plan-do-check-act (also known as Deming’s cycle)</td>
</tr>
<tr>
<td>PIMS:</td>
<td>Plant Information Management Systems</td>
</tr>
<tr>
<td>PLC:</td>
<td>Programmable Logic Controller</td>
</tr>
<tr>
<td>RCA:</td>
<td>Root Cause Advisor (implemented in the spc tool)</td>
</tr>
<tr>
<td>R-chart:</td>
<td>Subgroup ranges (Control Chart)</td>
</tr>
<tr>
<td>s-chart:</td>
<td>Subgroup standard deviations (Control Chart)</td>
</tr>
<tr>
<td>spc:</td>
<td>Statistical Process Control tool (AMG Toolbox)</td>
</tr>
<tr>
<td>SPC:</td>
<td>Statistical Process Control</td>
</tr>
<tr>
<td>u-chart:</td>
<td>Defects per unit in subgroup (Control Chart)</td>
</tr>
<tr>
<td>UCL:</td>
<td>Upper Control Limit</td>
</tr>
<tr>
<td>USL:</td>
<td>Upper Specification Limit</td>
</tr>
<tr>
<td>we1, we2, ...</td>
<td>Western Electric Rules</td>
</tr>
<tr>
<td>Xbar-chart:</td>
<td>Subgroup means (Control Chart)</td>
</tr>
<tr>
<td>X-chart:</td>
<td>Individual measurements (also referred as I-chart)</td>
</tr>
</tbody>
</table>