Volume 34 (1), pp. 65–81
http://orion.journals.ac.za

ORiON
ISSN 0259–191X (print)
ISSN 2224–0004 (online)

Simulation of a coal stacking process using an online X-Ray Fluorescence analyser

RF Rossouw*† RLJ Coetzer‡ NJ le Roux§

Received: 2 May 2017; Revised: 9 December 2017; Accepted: 13 March 2018

Abstract

The Sasol Coal Value Chain is a complex system consisting of blending, stacking and reclaiming of no fewer than six different coal sources with vastly different coal qualities. The amount and quality of the gas produced from coal depend crucially on the quality of the coal reclaimed from the coal stacking yards. In this paper the development of a real time coal quality simulation model using information from an online X-Ray Fluorescence analyser, integrated with various data sources from the Coal Supply Facility, is presented. The integration of different data sources is discussed to create a centralised and standardised data framework for input to the simulation model. The simulation of a heap profile of the coal quality for each heap stacked, together with the quality of the reclaimed coal, is discussed in detail. It is shown how the generated information from the model is utilised in the development of a reclaiming strategy.

Key words: Coal properties, coal quality profile, coal stacking process, coal value chain, stacking simulation model, statistical simulation model, X-Ray Fluorescence analyser.

*Sasol Group Technology R&T, Sasol, Private bag 1, Sasolburg, 1947, Republic of South Africa, email: ruan.rossouw@sasol.com

†Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, 7602, Matieland, South Africa, email: ruan.rossouw@sasol.com

‡Sasol Group Technology R&T, Sasol, Private bag 1, Sasolburg, 1947, Republic of South Africa, email: roelof.coetzer@sasol.com

§Department of Statistics and Actuarial Science, University of Stellenbosch, Private Bag X1, 7602 Matieland, South Africa, email: njlr@sun.ac.za

http://dx.doi.org/10.5784/34-1-575

1 Introduction

Sasol Mining, South Africa, produces no fewer than 40 million tons of coal annually [11]. The coal is delivered to the *Sasol Synfuels Coal-to-Liquids* (CTL) factory at Secunda from six surrounding collieries. The CTL complex in Secunda is the largest coal to syngas production facility of its kind in the world. A total of 84 Sasol® FBDB™ gasifiers have a combined production of no less than $5 \times 10^6 m^3 n/h$ pure synthesis gas. The coal from the collieries goes through a number of processing and preparation steps, including transport on overland conveyors, transfers, and stacking and reclaiming, before it reaches the gasification plant. The coal handling and preparation facilities are collectively called the *Coal Value Chain* (CVC). A high level overview of the system is shown in Figure 1 [6].

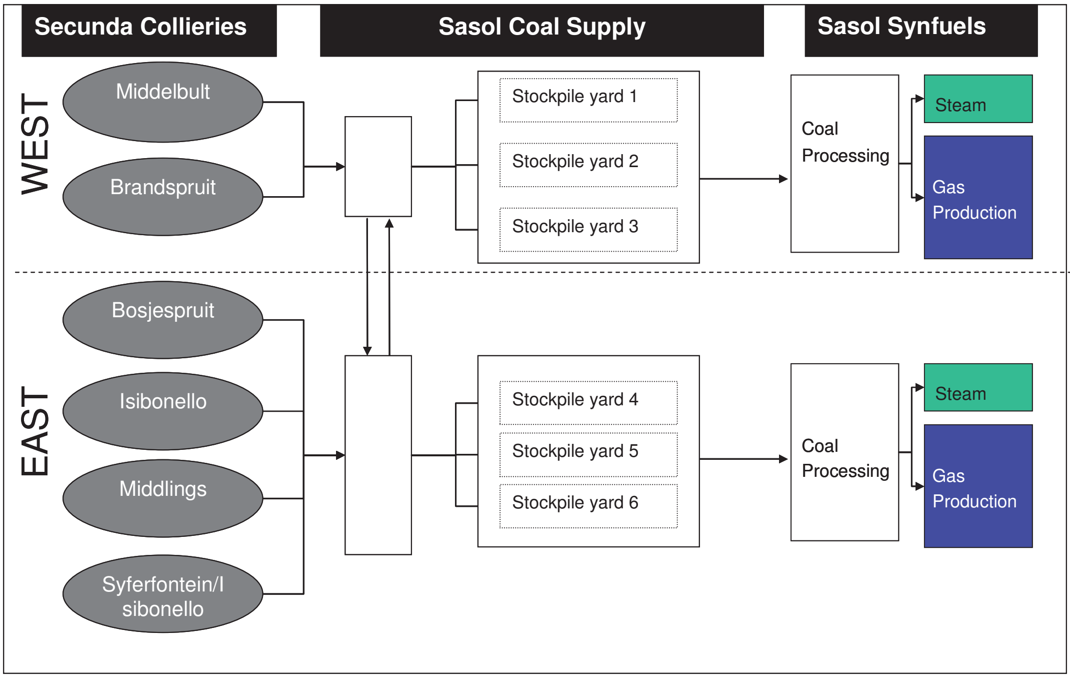

Figure 1: Overview of Secunda Coal Value Chain (CVC).

Two of the mines are located on the Western side of the factory and four mines on the Eastern side. Overland conveyors are used to transfer the coal from the different mines to two identical stockpile facilities at the Western and Eastern factory, respectively. Coal from the mines on the Western side is mainly transported to the Western stockpile facility, and coal from the mines on the Eastern side is mainly transported to the Eastern stockpile facility. However, coal is also transferred via heavy duty conveyor belts between the Eastern and Western stockpile facilities when required.

Each stockpile facility consists of three stockpile yards. The length of the stockpile is about 600 meters, which can contain up to six coal heaps [6]. The coal heaps are built by a stacker, and reclaimed by a reclaimer. The stacker and reclaimer operations will be discussed in more detail in Section 4. The stockpiles fulfil two distinct purposes. First, they act as a supply buffer between the mining operations and the gasification operations. Mining is shift based, and coal is not produced continuously. In contrast, gasification (and the downstream operations) is a continuous process. During mining production excess coal is deposited on the stockpiles, and is then utilised during non-production periods to ensure continuous feed to the gasification process.

The qualities of the coal vary significantly between the different coal sources and the *run-of-mine* (ROM) coal is stacked and, in effect, blended on the stockpiles as it is received. Therefore, the second purpose of the stockpiles is to serve as a homogenisation step for improved stability in the coal qualities. Great effort is made to reduce the variability in the coal qualities on the heaps through planning and blending. However, the blending is constrained by the capacity of the stockpile yards and the rate of production at the different mines. It was shown previously that the quality of the coal feedstock, and the stability thereof, has a significant effect on the amount of gas produced [1, 2, 3]. Therefore, to ensure stable feed at desired qualities being delivered to the factory an efficient coal stacking process is required. A detailed discussion of the scheduling and blending process and challenges thereof can be found in [4] and [11].

Currently, the coal sources blends are planned and scheduled based on coal qualities that are not available in real time. Specifically, the laboratory analysis for ash content takes approximately two days, and historic data are therefore used in the blend plan *i.e.*, there is no real time data available for efficient planning and optimisation of the blends. Given that ash content is the main coal property being monitored, having real time information available during stacking and reclaiming would significantly improve the planning and blending of the coal sources for delivering coal with desired quality and stability to the factory. Furthermore, the turnaround time for the stockpiles (from stacking to reclaiming) is about two to three days, and the implication thereof is that a stockpile could be reclaimed before the quality of the coal on the stockpile is known. To alleviate this, a project was initiated to install instrumentation for on-line measurement of the coal qualities. Specifically, an on-line *X-Ray Fluorescence* (XRF) analyser was installed on one stacker conveyor belt to provide real time information on the ash content and ash composition of the coal.

In this paper the development of a statistical simulation model is presented for the coal stacking process. The model utilises the on-line XRF data to calculate and to provide real time information on the ash content of the coal stacked across the length of the heap, and to calculate the ash content of the coal blends on the heaps as it is reclaimed. In addition, the model provides an improved understanding of the effect of the blending and stacking parameters on the coal properties for purposes of minimising the standard deviation in the ash content subject to a specified mean ash content. The stacking simulation model is an enabler for the online XRF data to improve the coal sources blending, and to deliver stable feed with desired quality to the factory.

The paper is outlined as follows. In Section 2 a brief overview of the X-Ray Fluorescence analyser is provided. This is followed in Section 3 by a description of the various coal data sources utilised in this study. The main contribution of the paper, the development and utilisation of a stacking simulation model, is detailed in Section 4. Finally, some conclusions are highlighted in Section 5.

2 X-Ray Fluorescence analyser (XRF)

The theoretical chemistry of X-Ray fluorescence is outside the scope of this paper. However, a short explanation of the measurements from the XRF analyser will be given here [9]. In principle, X-Ray fluorescence is the emission of characteristic "secondary" (or fluorescent) X-rays from a material that has been excited by bombarding with high-energy X-rays or gamma rays. Primary X-rays are bombarded onto a sample. X-rays are either absorbed by the atom or scattered by the material. During this process, if the primary X-ray has sufficient energy, electrons are ejected from the inner shells, creating vacancies. These vacancies present an unstable condition for the atom. As the atom returns to its stable condition, electrons from the outer shells are transferred to the inner shells and in the process yield a characteristic X-ray whose energy is the difference between the two binding energies of the corresponding shells. Each element has a unique set of energy levels and therefore produces X-rays at a unique set of energies.

Figure 2: An image of an X-Ray Fluorescence analyser (XRF).

One XRF analyser has been installed on Stacker 4 at the Eastern stockpile yard. The XRF runs on a sled on top of the coal passing on the coal conveyor (see Figure 2). This ensures proximity between the X-Ray tube and the coal on the conveyor. The calibration of the XRF analyser is outside the scope of this paper as it is performed by the supplier. Note the calibration is performed on run-of-mine coal samples which ensures that the results are robust.

3 Coal data sources

Developing a stacker simulation model requires many different sources of data including, online coal quality data from the XRF analyser, laboratory coal quality analyses, material movement data on the coal sources and information on each coal stockpile. The data from the different sources come in different formats and are located on different network servers. Therefore, the development of a standardised and integrated data management framework was required for the capturing and integration of the different data sources, and for development of a stacking simulation model and a real time monitoring system. The different data sources are now discussed briefly.

- *Online XRF coal quality data.* One XRF analyser is installed at Stacker 4 on the coal stack yard 4 (see Figure 1). The data from the analyser are captured on a local PC attached to the analyser by the Monaco software from J&C Bachmann [5]. The data are retrieved, cleaned and stored in a local MySQL database.
- *Laboratory coal quality analyses.* A set of Excel [8] spreadsheets containing laboratory coal quality analyses is distributed as attachments on an email every morning. The spreadsheets contain data on ash content, particle size distribution and inherent moisture for each coal source. The data from the email attachments are captured automatically by triggering Visual Basic for Applications (VBA) macros once the emails arrive in the Microsoft Outlook inbox.
- *Material Movement Data.* The material movement information is generated in the SCS control room when the different mines are selected, and tons are specified, for transport on the conveyors from the mines. The data are captured on a Microsoft SQL Server [7] database. The data from the SQL Server database are downloaded with the R software [10] and exported to a local MySQL database. The material movement data record the movement of coal on the conveyors from the different mines to the stackers, and from the different heaps reclaimed to the gasification plant. This information is invaluable, since in combination with the online XRF data, it is used to build a profile of the coal qualities for each mine. In addition, this information is used to predict the properties of the coal for each individual coal heap. This is discussed in more detail in Section 4.
- *Stockpile Information.* Information about the stockpiles is made available on Excel files two times a day (06:00 and 18:00) by the mining department. These files contain various important information for example, the length of the heap (in meters), start and end position of the heap, planned tons and actual tons. The start and end positions are defined on a scale +300 meters to -300 meters relative to a centre point of zero for the stack yard. The positive values indicate that the heap is built on the factory side of the stack yard, and the negative values indicate that the heap is built on the mining side of the stack yard. To capture these data in a convenient format a combination of R scripts and Excel VBA macros is utilised. The relevant

information is captured and exported to the same MySQL database as the other data sources.

Although all the data discussed above are available on various servers, individually only limited insight can be generated from the different sources. Maximum intelligence is generated through the integration of all the different sources into one data framework, which is then readily available for model development and real time performance monitoring.

4 Stacker simulation model

4.1 Preliminaries

The objective of the stacking simulation model is to provide information on the ash content of the coal stacked across the length of the heap, and to provide approximations of the ash content of the coal blends on the heaps as it is reclaimed. The online XRF analyser provides the unique opportunity to specify a real time profile of the coal qualities from the different sources measured in real time. In this section the development of the stacking simulation model will be discussed. The real time coal profile information from the XRF analyser on Stacker 4 is utilised, in combination with the material movement data to develop the statistical simulation model for the stacking process. Note, in this study it is assumed that the coal properties for each coal source going to Stacker 4 are representative of the coal properties for the same source going to the other stackers.

Huge variation in coal qualities exists within and between the different mines. To homogenise coal properties on the reclaimer belts different coal sources are blended via a stacking procedure. Two different processes of blending takes place. First, each heap in the stack yard consists of more than one feed source, and second, the sources are stacked on the heap in several layers.

Chevron-strata stacking is described here, where the coal from the different mines is longitudinally stacked in layers on top of each other. For example, if the target length of the heap is 120 meters, the stacker will throw coal from position 0 to position 120 and then immediately from position 120 to position 0.

In contrast with stacking, reclaiming starts from one end of the heap, and reclaims vertical segments of the heap. Although the reclaimers can move in both directions, they can only reclaim a heap from one direction until the heap is finished. Typically, more than one heap will be reclaimed at the same time, but at different speeds *i.e.*, tons/hour. Therefore, if the profile of the properties of the coal on the heap can be accurately simulated, and the reclaimed coal from the heaps can be predicted over the length of the heaps, the operator can ensure that a heap with less desirable coal qualities is reclaimed at lower volumes. In addition, a heap with desirable coal qualities can be reclaimed simultaneously with a heap with poorer qualities to mitigate the effect of the lower qualities on the processing plant.

The combination of stacking horizontally and reclaiming vertically leads to homogenisation of the coal. The actual combination of mines and the number of layers and sequence of the mines on the heaps will however impact the homogeneity of the resulting reclaimed coal.

Given that the stacker moves at a constant speed there are three levers to manipulate the layers of the feed sources (blending) on the heap:

1. The length of the heap.
2. The tons/meter of coal on the conveyor.
3. The sequence of feed sources.

The length of the heap will impact the layers in that a longer heap will have fewer layers of any specific mine for each segment of the heap. A shorter heap will have the opposite effect. Demand constraints, as well as equipment constraints do however limit the length of the heaps that can be build at any given time. Loading fewer tons per hour on the conveyors from each feed source will have the effect of thinner layers and therefore more layers of each source on the heaps. This will improve homogenisation of the coal qualities. There are however demand constraints from the factory on the tons per hour.

The sequence of stacking the supply from the mines may be used as an additional lever to modify the heap quality profile and improve stability in coal qualities. Due to variability within the feed sources it could be beneficial to have more layers of sources inter-dispersed on the heaps. One cause of variability in coal qualities within sources is that each source consists of different coal seams. Consequently, the coal properties can differ significantly between the seams.

In this paper a statistical methodology is proposed for accurate and efficient simulation for heap profiling. Before model development can commence, the input data must be prepared *i.e.*, capturing of the data on the run-of-mine coal from the different conveyors feeding into the stacker conveyor. The material movement data available is illustrated in Table 1. From the table it can be observed that the resolution of the data is not appropriate for a stacking simulation model. Specifically, coal movement on the conveyor is a continuous process with variation in coal qualities and tons. For example, Table 1 indicates that coal source D was deposited first on the heap with specified start end times of the stacker. However, the time intervals are too long which yield tons and coal qualities that are not on a continuous scale for the particular coal source. It is possible to approximate a continuous process by breaking it up into very small discrete time buckets. Choosing the appropriate time buckets for the simulation is a trade-off between efficiency and accuracy. Smaller time buckets increase the accuracy of the model, but also increase simulation time as well as memory use.

Consider time buckets of one second for the simulation. For example, assume the maximum amount of coal passing on the conveyor is 2000 tons/hour. This translates to approximately 555 kg per second, which is a substantial amount of coal. The stacker speed is $\frac{1}{7}$ m/s and a heap can be as short as 80 meters. A time bucket greater than seven seconds will therefore lead to blocks of more than a meter in the simulation, and less than 100 actual points in the resulting stack. From in–house experimentation, it was found that using one second time buckets is sufficient for providing an accurate simulation of the coal qualities profile on the heaps. In addition, using one second time buckets reduces the complexity of the time format for the simulation model. Therefore, in this paper, one second time buckets is used for the statistical simulation model.

Variability needs to be introduced with the material movement data for the simulation of

Row	SCS_ID (1)	Source (2)	Start Time (3)	End Time (4)	Mass (ton) (5)
1	15374	Source D	1409242890	1409242950	68.87
2	15374	Source D	1409243072	1409243550	254.14
3	15374	Source D	1409243640	1409245441	970.00
4	15374	Source D	1409245440	1409245441	105.73
5	15374	Source D	1409245470	1409248051	1300.77
6	15374	Source A	1409248710	1409260410	5876.08
7	15374	Source A	1409261220	1409263144	962.36
8	15374	Source A	1409263140	1409263144	75.46
9	15374	Source A	1409263175	1409271032	3940.58
10	15374	Source C	1409271540	1409272170	288.12
11	15374	Source C	1409272470	1409272623	42.14
12	15374	Source C	1409274360	1409281470	3356.50
13	15374	Source B	1409288911	1409290953	979.54
14	15374	Source B	1409291250	1409300460	4438.46
15	15374	Source B	1409303700	1409304720	473.86
16	15374	Source B	1409304720	1409304720	89.44
17	15374	Source B	1409304750	1409306190	508.26
18	15374	Source B	1409306160	1409306190	53.32
19	15374	Source B	1409306220	1409306911	18.92
20	15374	Source E	1409308830	1409310420	244.02
21	15374	Source B	1409310990	1409315164	1960.80
22	15374	Source B	1409315610	1409315670	2.58
23	15374	Source E	1409316120	1409316870	120.54
24	15374	Source E	1409317653	1409321731	144.06
25	15374	Source B	1409318520	1409318640	30.10
26	15374	Source B	1409318732	1409321731	1402.66
27	15374	Source B	1409324790	1409325111	125.56
28	15374	Source B	1409453460	1409458590	2453.58
29	15374	Source B	1409458981	1409459521	239.08
30	15374	Source E	1409510580	1409516156	1382.78
31	15374	Source B	1409515950	1409516156	67.94

Table 1: A material movement file for one heap.

the coal qualities on the heap. If the data are used as is, and just broken up into smaller time buckets, the coal tons will be constant for long periods, which will lead to much less variability in the simulated reclaimed coal properties compared to the actual reclaimed coal. This will not reflect reality. The actual distribution of the coal tons is unknown. However, after some investigation it was determined that the coal is distributed uniformly over the belt due to the natural packing of the coal transported over long distances on the conveyors from the feed sources. Therefore, the uniform distribution is employed to simulate the coal tons. To conform to the actual mass balance the tons are simulated as a composition *i.e.*, the total of the buckets for a time period is equal to the tons in the material movement file for the same time period.

Algorithm 1: Generating random compositional data.

- Draw independent random $u_i \sim \text{Uniform}(0,1), \quad i = 1, \dots, n-1$.
- Sort u_i from smallest to largest *i.e.*,

$$\boldsymbol{u}^T = (u_{(1)}, u_{(2)}, \dots, u_{(n-1)}).$$

- Specify $\boldsymbol{v}^T = (0, \boldsymbol{u}^T, 1)$, thus $\boldsymbol{v}$ is of dimension $(n+1) \times 1$.
- Let $d_i = v_{i+1} - v_i, \quad i = 1, \dots, n$. Thus $d_i \sim \text{Uniform}(0,1)$ and $\sum_i d_i = 1$.
- Therefore, any quantity x can be represented by n uniformly distributed variables d_i through the identity

$$\sum_i d_i x = x. \tag{1}$$

4.2 Simulation model

The simulation process will now be discussed and illustrated for one heap. Specifically, it will be illustrated how the heap is tracked from the material movement files to a completed heap on the stockpile yard. In the following discussion one replication will be considered. However, since the process involves random distributions multiple replications are performed in the commercial application of the methodology.

Consider the material movement file in Table 1 for one heap. Each row contains the tons of coal (Mass) from a specific mine that passed over the conveyor between Start Date and End Date. The time stamps are in the POSIX time format denoting number of seconds from 1 January 1970. Note that although the material movement file captures the coal movement on the conveyor, the coal is stacked on a specific heap, and the information is therefore directly applicable to the specific heap. Let x_{nj} denote the tons of coal for time period $n = 1, \dots, N$ for mine $j = 1, \dots, J$. For example, from Table 1, $N = 31$ and $J = 5$ for the mines transferred on the conveyor for the heap.

Simulation of the coal tons is as follows:

- Applying Algorithm 1, the tons of coal on the conveyor for each second z in each interval of length k_{nj} can be simulated as $c_{znj} = d_{znj} x_{nj}, \quad z = 1, \dots, k_{nj}$ and $d_{znj} \sim \text{Uniform}(0,1)$ with $\sum_{z=1}^{k_{nj}} d_{znj} x_{nj} = \sum_{z=1}^{k_{nj}} c_{znj} = x_{nj}$.
- Note that $\boldsymbol{c}_{nj}$ is therefore a $k_{nj} \times 1$ vector of simulated coal tons for row n and mine j for each one second interval on the conveyor.
- Let $T_b = t_{1b}$ and $T_e = t_{Ne}$ and $M = T_e - T_b + 1$, where t_{1b} and t_{Ne} denote the first and last time stamps for the heap respectively.
- An $M \times 1$ vector $\boldsymbol{m}$ consisting of all the one second buckets of simulated coal tons on the conveyor, is defined for the heap.
- Each $\boldsymbol{c}_{nj}$ will occupy $\boldsymbol{m}_{(t_{nb} - T_b + 1)}$ to $\boldsymbol{m}_{(t_{ne} - T_b + 1)}$.
- Note that the $\boldsymbol{m}_{(t_{ne} - T_b + 2)}$ to $\boldsymbol{m}_{(t_{(n+1)b} - T_b)}$ values for $n = 1, \dots, N-1$ will be occupied by zero values.

For example, for Table 1, $T_b = 1409242890$ and $T_e = 1409516156$. Therefore, $M = T_e - T_b + 1 = 273267$. For $n = 1$, $(t_{1b} - T_b + 1) = 1409242890 - 1409242890 + 1 = 1$, and $(t_{1e} - T_b + 1) = 1409242950 - 1409242890 + 1 = 61$, and $m_{(1)}$ to $m_{(61)}$ will therefore populate $\boldsymbol{c}_{1j}$. Similarly, for $n = 1, (t_{ne} - T_b + 2) = 1409242950 - 1409242890 + 2 = 62$ and $(t_{(1+1)b} - T_b) = 1409243072 - 1409242890 = 182$, therefore $\boldsymbol{m}_{(62)}$ to $\boldsymbol{m}_{(182)} = 0$.

The total tons of coal is calculated by $\sum_n x_{nj}$ and for the specific heap, $\sum_n x_{nj} = 31976$ compared to 32000 planned from Table 1. The total tons for each mine j is calculated as

$$T_j = \sum_n \begin{cases} x_{nj}, & \text{if } j_n = j \\ 0, & \text{otherwise} \end{cases}$$

and the mine percentages as

$$P_j = \frac{T_j}{\sum_j T_j} \times 100. \tag{2}$$

The histogram of calculated mine percentages is shown in Figure 3 for the heap.

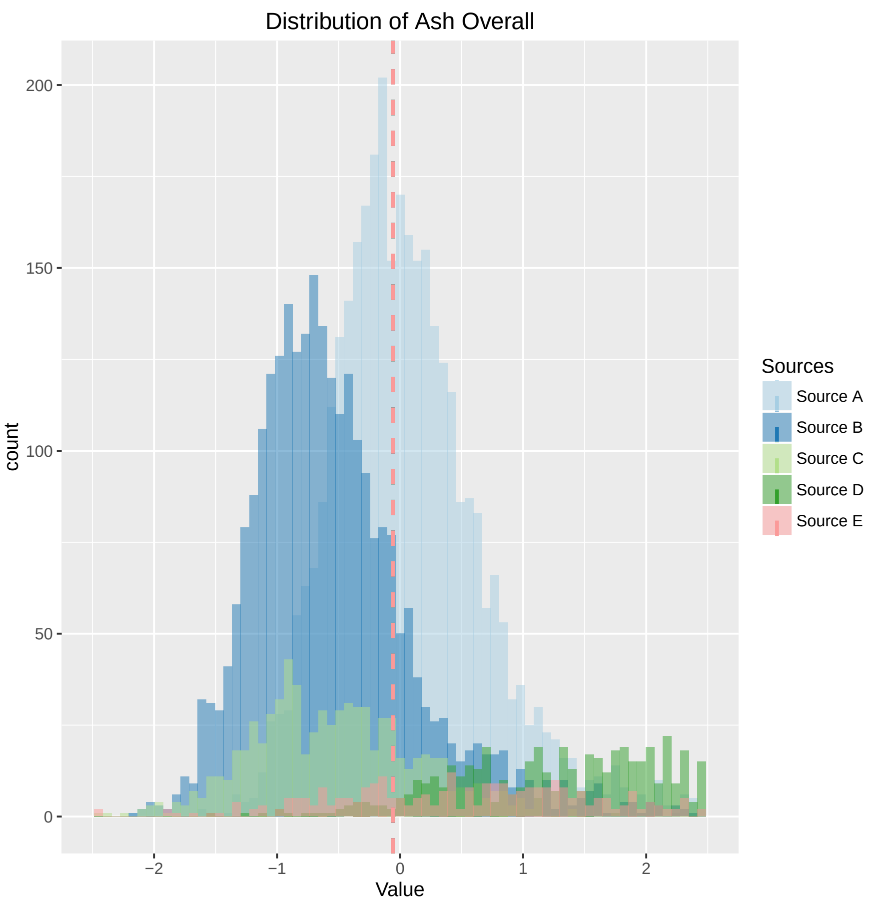

Figure 3: Histograms of overall ash per source.

Only the ash content of the coal will be discussed here, as the calculation of the other properties is similar. The XRF data are captured on a time stamp and an ash value is recorded every 90 seconds when coal is traveling on the conveyor to the stacker. Using the material movement data, the mine data is merged with the online XRF data on the appropriate time stamps. This yields a table with a time stamp, an ash value, and a mine ID in each row.

For utilisation in the stacking simulation, the online XRF data must be relevant to the specific heap, and have an appropriate resolution. The 90 seconds resolution of the XRF data, although high resolution, is not appropriate for the stacking model if inferences need to be made about the distribution of ash over the heap length. As mentioned above, the stacker moves at $\frac{1}{7}$ m/s, and 90 seconds will therefore represent a constant value for 12.86 meters. Therefore, the same approach used above for simulating the distribution of the coal tons on the conveyor is applied for simulating the distribution for the ash content.

First, it is required to extract the appropriate range of data from the measured and stored online XRF data. Specifically, the data selected must be as close as possible in time to the time stamps in the material movement file. A default of two working days from the current start time in the material movement file is used to initiate the search for the time stamps. A check is performed if any data are available for the specific mine, and whether enough data points are available. If not, the start date is moved back a further 24 hours into history, and the search is repeated. This is repeated until the appropriate number of data points (in this case 30) is obtained.

To sample from the available data, it is assumed that the data are normally distributed. The assumption of normality was considered to be sensible given the amounts (tons) of coal passing underneath the online instrument. Alternatives, such as bootstrap sampling, were also investigated but it did not improve the results or the practical significance thereof. Therefore, the appropriate number of points are randomly sampled from the normal distribution to complement the one second material movement data. Histograms of the ash data for the different mines are provided in Figure 3 for a 90 day period, and it is clear from the different histograms that the XRF analyser on Stacker 4 is capable of distinguishing ash values between the individual mines.

The normal distribution is specified by two parameters; the population mean μ and the population variance σ^2 (*i.e.*, $\mathcal{N}(\mu, \sigma^2)$). The population mean and variance for each mine j can be estimated by the sample mean ($\bar{S}_{jq}$) and variance $\hat{\sigma}^2_{S_{jq}}$ from the XRF data as shown in Table 2. The goal is to simulate k_{nj} ash values for each row n in Table 1. Let v_{znj} be a sample of k_{nj} values from $\mathcal{N}(\bar{S}_{jq}, \hat{\sigma}^2_{S_{jq}})$. Therefore, $\boldsymbol{v}_{nj}$ will be a $k_{nj} \times 1$ vector of simulated ash values.

An $M \times 1$ vector $\boldsymbol{a}$ consisting of all the one second buckets of simulated ash tons on the conveyor going to one heap is defined by multiplying elementwise the ash percentages in $\boldsymbol{v}_{nj}$ with the coal tons in $\boldsymbol{c}_{nj}$. Each $\boldsymbol{v}_{nj} \cdot \boldsymbol{c}_{nj}$ will occupy $\boldsymbol{a}_{(t_{nb}-T_b+1)}$ to $\boldsymbol{a}_{(t_{ne}-T_b+1)}$. Note that the $\boldsymbol{a}_{(t_{ne}-T_b+2)}$ to $\boldsymbol{a}_{(t_{(n+1)b}-T_b)}$ entries are occupied by zero values for $n = 1, \ldots, N-1$.

As before, $\boldsymbol{m}$ is a vector of coal tons, and $\boldsymbol{a}$ the vector of ash tons passing on the conveyor in one second. If the length of the heap is defined as L (from Table 3 $L = 160m$ for the heap) meters, and the length of $\boldsymbol{m}$ as M ($M = 273267$ seconds for the same heap), the

Sources	Tons	Mean Ash	SD Ash	% in Blend	Data From	Data To
Source A	10854	29.68	2.39	33.95	29-Aug-14 22:16	30-Aug-14 23:59
Source D	2700	36.00	4.00	8.44	28-Aug-14 18:28	28-Aug-14 19:44
Source C	3687	27.19	3.68	11.53	30-Aug-14 02:24	30-Aug-14 04:39
Source E	1891	37.20	4.69	5.92	30-Aug-14 08:56	31-Aug-14 21:43
Source B	12844	27.97	3.01	40.17	30-Aug-14 05:33	31-Aug-14 06:15

Table 2: Mine properties for one heap.

Heap Summary		
% Ash Summary	Average = 29.7%	Standard Error = 0.12%
Total Tons	Planned = 32000	Actual = 31976
Stack Date	Start = 28-Aug-14 18:21	End = 31-Aug-14 22:15
Total Heap Length	160 Meters	200 Tons/Meter

Table 3: Summary statistics for one heap.

number of layers B on the heap can be specified as

$$B = \lceil M/(\frac{L}{s}) \rceil \tag{3}$$

where $\lceil x \rceil$ is defined as ceiling(x) (the smallest integer not less than x) and s is defined as the stacker speed in m/s. Let $l = \frac{L}{s}$, then the simulated heap can be defined as the $B \times l$ matrix $\boldsymbol{H}$. This is an obvious oversimplification as coal particles will not stack perfectly, and in practice the endpoints of the heap will not be square, but for the sake of simplicity the assumption of a rectangular heap will suffice. In addition, given the amounts of coal on a heap (*e.g.* about 32 000 tons), assuming a rectangular heap will not significantly affect the predicted mean ash content of the heap.

As the stacker moves forward and backwards building the heap, the stacking can be envisioned as a folding of the coal on the conveyor into multiple layers. The length of each fold will be l, and the number of folds will be B. As B was rounded up it is necessary to add $B \times l - M$ empty one second buckets to the front of vectors $\boldsymbol{m}$ and $\boldsymbol{a}$. Each row u of $\boldsymbol{H}$ will be populated in the following equation,

$$\boldsymbol{H}_{u(1,\ldots,l)} = \begin{cases} \boldsymbol{m}_{(u\times l)+1,\ldots,(u\times l)+l}, & \text{if } u \text{ is odd} \\ \boldsymbol{m}_{(u\times l)+l,\ldots,(u\times l)+1}, & \text{if } u \text{ is even.} \end{cases}$$

A $B \times l$ matrix of ash tons $\boldsymbol{A}$ is defined utilising $\boldsymbol{a}$ similar to $\boldsymbol{H}$. The ash percentage over the heap length can then be calculated by first calculating the tons ash ($\boldsymbol{ta}$) and ton coal ($\boldsymbol{th}$) for each row in $\boldsymbol{A}$ and $\boldsymbol{H}$ as seen in the following two equations,

$$ta_i = \sum_{j=1}^{B} \boldsymbol{A}_{(ij)} \quad , i = 1, \ldots, l \tag{4}$$

$$th_i = \sum_{j=1}^{B} \boldsymbol{H}_{(ij)} \quad ,i = 1, \ldots, l. \tag{5}$$

The percentage of ash over the length of the heap ($\boldsymbol{ap}$) is therefore calculated as

$$ap_i = \frac{ta_i}{th_i} \quad ,i = 1, \ldots, l. \tag{6}$$

The resulting output of the simulation for the heap is shown in Figure 4. Figure 4 provides valuable information on the percentage ash over the length of the heap. This information is utilised to prepare a reclaiming strategy to stabilise the feed to the factory (Section 4.3).

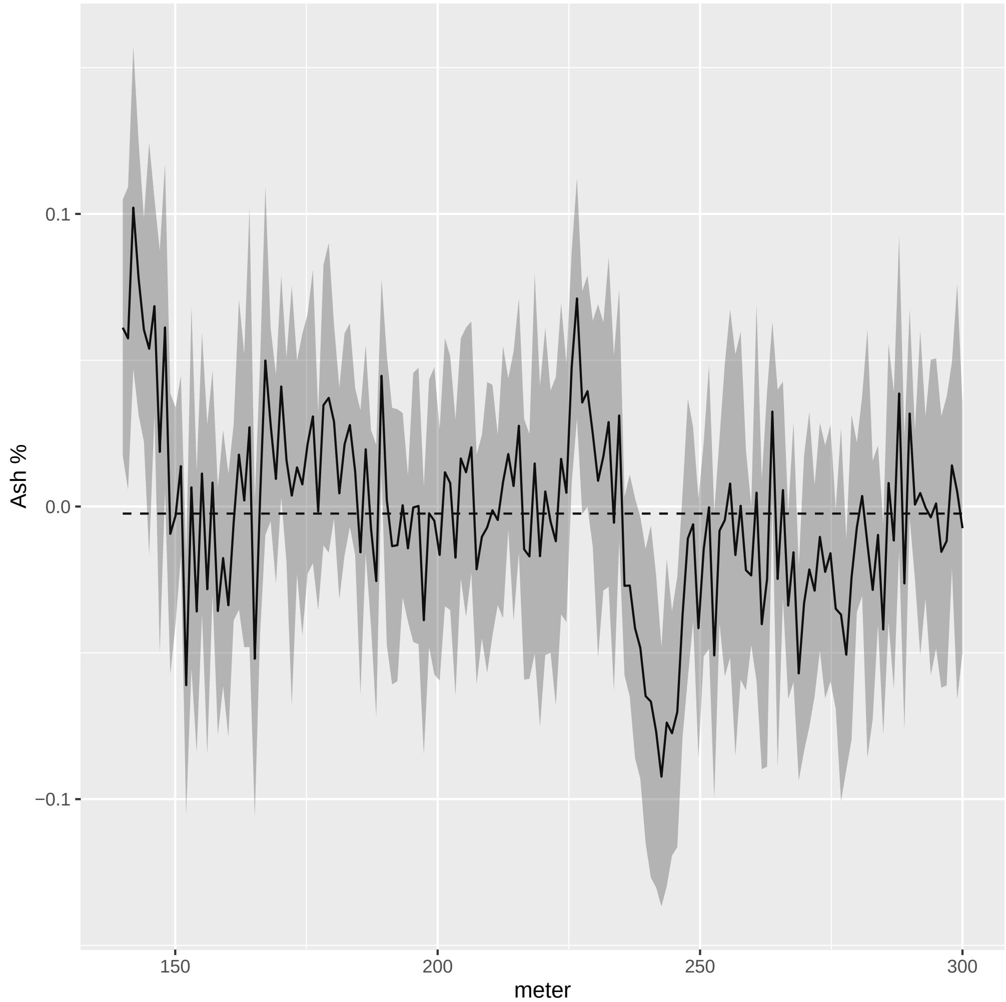

Figure 4: Ash percent profile over the length for one heap.

Table 2 contains the XRF information for the heap. The overall average ash and standard error of the average ash are calculated as follows:

- Let P_j be the proportion of mine j on the heap, x_j the tons of coal of mine j, $\bar{w}_j$ the average of the ash from the XRF analyser for mine j, $\hat{\sigma}_j$ the standard deviation of the ash from the XRF analyser for mine j, and n_j the number of XRF measurements for each mine j.
- The overall mean ash percent is calculated as a proportional mean as

$$\bar{W} = \sum_j \bar{w}_j \times P_j. \tag{7}$$

- The standard error of the mean ash content for the heap is calculated as

$$SE_w = \sqrt{\sum_j P_j^2 \times \frac{\hat{\sigma}_j^2}{n_j}}. \tag{8}$$

- The standard deviation of the ash for the heap is calculated as

$$S_w = \sqrt{\frac{\sum_j x_j^2 \times \hat{\sigma}_j^2}{(\sum_j x_j)^2}}. \tag{9}$$

The above summary statistics are shown in Table 3 for the heap. Included are the planned and actual tons, the start date and end date for the heap, the length of the heap and tons per meter. Some additional information that is obtained from the material movement information is the layering of the mines on the heap. Refer to Figure 5. The horizontal axis depicts the length in meters of the heap. The vertical axis depicts the layers over time for the heap. The mines are depicted by the colours in the legend. Note that the white layers depict the times when the conveyor was empty and the stacker was still moving. It is important to note that the layers depict the movement of the stacker over time and not the actual positions of the coal on the heap, as it does not include the tons on the layers. This graph provides important information to management concerning the homogeneity of the coal properties on the heap.

4.3 Application of the stacker simulation model

Since the amount and quality of the gas produced from coal depend crucially on the quality of the coal reclaimed from the coal stacking yards, it is critically important to have information on the quality of the coal being sent to the processing plant in real time. Therefore, the information generated from the simulation model can now be used to predict the quality of the coal reclaimed.

In the petrochemical industry as well as most other process driven industries large volumes of data are generated by various on-line instruments and analysers. These data are captured and stored in a *Distributed Control System* (DCS). At Sasol the predominant DCS system for the production facilities is the OSIsoft PI system. The tons reclaimed by each reclaimer are available from the PI DCS system.

The output from the stacker simulation model is combined with the data from the material movement files, and the reclaimer data from the PI DCS system to calculate the ash

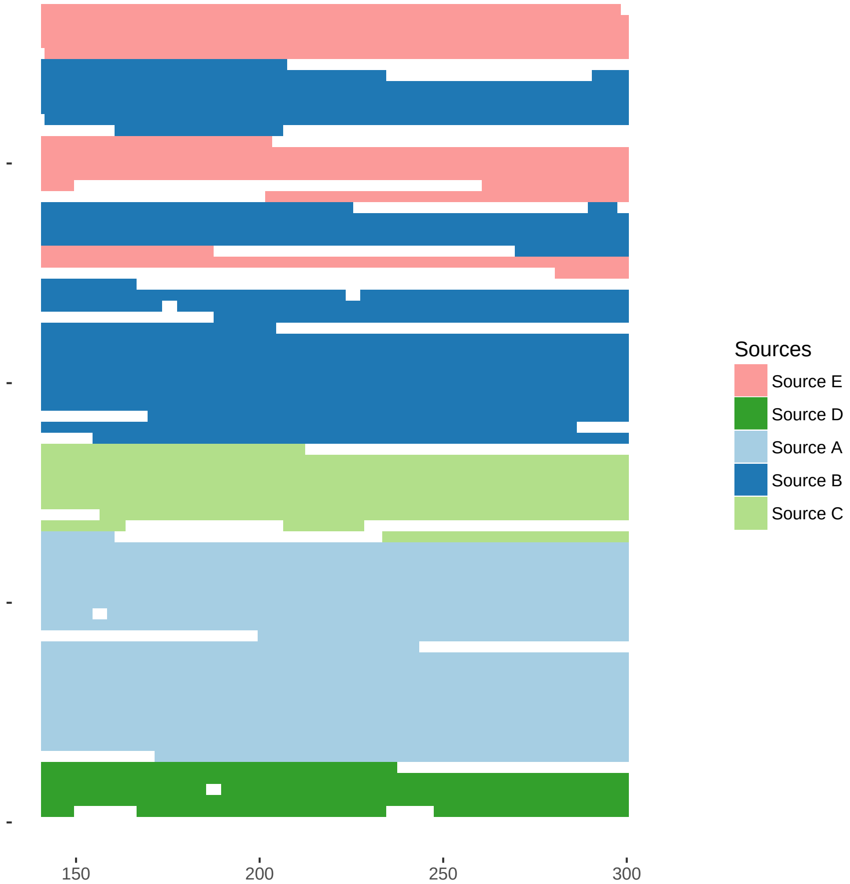

Figure 5: Source layers over the length for one heap.

percent. Figure 6 shows a profile of ash content reclaimed across the length of the heap. This information is now used to develop a reclaiming strategy to minimise the impact of the heaps with unfavourable ash properties on the gasification factory.

Specifically, at any given time coal is reclaimed from two to three heaps simultaneously, which is sent to the processing plant via overland conveyors. Note about 4000 tons of coal are reclaimed and processed per hour. Therefore, having information available in real time on the quality of the coal on each heap, the plant can select which heaps need to be reclaimed together. Furthermore, the tons per hour to be reclaimed from each heap can be specified for minimising the variability in the coal quality reclaimed, subject to the minimum tons of coal required per hour. Therefore, the data extracted from the various data sources at the SCS facilities have been converted into valuable information through an operational simulation model, which can be used for informed decision making.

5 Conclusions

A stacker simulation model was developed in-house and implemented on the commercial plant to predict the properties of the heaps on all the stack yards. The model utilises the

Figure 6: Predicted reclaimed coal ash percentage.

online XRF data in combination with the data from the SGS laboratories, the material movement files and the stockpile information files. The stacker simulation model allows for the prediction of coal properties (including but not exclusive to ash) over the length of the heap, as well as the average ash percentage and standard error of the ash percentage for the heap. This information is used to do blend planning for the week, and is compared to the laboratory analysis of the data for validation purposes.

As a direct consequence of the developed modules and implementation discussed in this paper, Sasol has ordered seven additional XRF analysers. These analysers will be installed on the remaining five stackers, as well as on the conveyors feeding the reclaimed coal to the Sasol Gasification plant (one each for the Western and Eastern factories). In addition, a study is in progress to evaluate various technologies to reliably monitor the position of the stackers and reclaimers. As discussed in Section 4, knowing the position of the stacker will improve the accuracy of the simulation model.

In conclusion, evidence is provided in this paper of the successful implementation of a combined operations research and data science approach to real-time multivariate process monitoring.

References

[1] Coetzer R & Keyser M, 2002, *Experimental design and statistical evaluation of a full-scale gasification project*, Fuel Processing Technology, **1632**, pp. 1–16.

[2] Coetzer R & Keyser M, 2004, *Robustness studies on coal gasification process variables*, ORiON, **20(2)**, pp. 89–108.

[3] Coetzer R, Rossouw R & Lin D, 2008, *Dual response surface optimization with hard-to-control variables for sustainable gasifier performance*, Applied Statistics, **57**, pp. 567–587.

[4] Conradie DG, 2007, *Scheduling coal handling facilities using metaheuristics*, Master's Thesis, University of Pretoria.

[5] J&C Bachmann, 2015, *Monaco*, J&C Bachmann GmbH, Im Stöckmädle 23/1, 76307, Karlsbad, Available from https://jcbachmann.com/en/startpage

[6] Mabuza N, 2011, (Unpublished) Technical Report, Sasol Technology - Research and Development.

[7] Microsoft Corporation, 2005, *Microsoft SQL Server 2005*, Microsoft Corporation USA, Redmond (WA), Available from https://www.microsoft.com/en-za/

[8] Microsoft Corporation, 2010, *Microsoft Excel 2010*, Microsoft Corporation USA, Redmond (WA), Available from https://www.microsoft.com/en-za/

[9] Moitsheki L, Coetzer R, Koekemoer A, Rossouw R, Grobler D & Raijmakers N, 2014, (Unpublished) Technical Report, Sasol Techonology - Research and Development.

[10] R Core Team, 2016, *R: A language and environment for statistical computing*, R Foundation for Statistical Computing, Vienna, Available from https://www.R-project.org/

[11] Swart M, 2005, *Scheduling model for a coal handling facility*, Master's Thesis, University of Pretoria.