Investigating sales and production volumes of Gold, PGM, Iron Ore and Manganese using a Bayesian change point detection approach

Michael Ingham[1], Sonali Das[2], Sarma Yadavalli[3]

[1]*Department of Industrial Engineering University of Pretoria*
Mike.j.ingham@gmail.com

[2]*Advanced Mathematical Modelling, CSIR Modelling and Digital Science, P.O. Box 395, Pretoria 0001, South Africa; School of Statistics and Actuarial Science, University of Witwatersrand, Johannesburg, Private Bag 3, Wits 2050, South Africa; Department of Statistics, PO Box 77000, Nelson Mandela Metropolitan University, Port Elizabeth 6031, South Africa; DST-NRF Centre of Excellence in Mathematics and Statistical Sciences (CoE-MaSS), South Africa*
sdas@csir.co.za

[3] *Department of Industrial Engineering University of Pretoria*
Sarma.Yadavalli@up.ac.za

Abstract:
This paper identified change points in the production volumes and sales of various Industrial Metals in South Africa from the period 2003 – 2016. The socio-economic environment in which the mining sector operates is complicated, with countless factors influencing the performance of industrial metals. This complexity makes it difficult for mining stakeholders to accurately forecast the production and sales for specific metals. A descriptive model that links the identified change points to causative events or factors was constructed. This model provides mining stakeholders with information on the events or factors that have the greatest impact on the performance of the various metals. This information allows mining stakeholders to focus forecasting efforts on the identified factors. Combining the focused forecasts with the impact that similar events had in the past helps the mining stakeholders to alter production levels or schedule investments before forecasted events take place, minimising the potential negative impact of said event. In this study, the monthly production volumes and sales of Gold, Platinum Group Metals (PGMs), Iron Ore and Manganese were analysed. The data spanned from 2003 – 2016 and was supplied by StatsSA. The data was analysed using the Bayesian Change Point Analysis (BH) (Barry and Hartigan, 1993). The results reveal that production drops were caused predominantly by mining strikes and increases in production costs, while sales were influenced by changes in the exchange rates and rand value of the commodity. Future papers will use sales volumes instead of Actual Rand values in order to identify changes that can be attributed to shifts in the demand of each metal. Future papers will also focus on performing a multivariate analysis on similar data to determine whether certain factors influence numerous metals.

***Keywords*:** Statistics, Mining, Metals, South africa

INTRODUCTION

According to statistics released by the Chamber of Mines, mining provided one million jobs and contributed 7.7 % of the GDP in 2015 (Chamber of Mines South Africa, 2015). These figures show that the performance of South Africa's industrial metals has a significant impact on the performance of South Africa's economy. Given the importance of the mining sector in South Africa, it is imperative for mining stakeholders to have the ability to forecast the performance of various metals in order to create effective business plans.

In this paper, a Bayesian change point approach was applied to the production volumes and sales of PGM, Gold, Iron Ore and Manganese. This was done in order to identify change points in the data. The identified change points were used to identify the factors or events that have the most significant impact on the performance of the abovementioned metals. The objective of which was to create a descriptive model that provides mining stakeholders with information on the events or factors that have the greatest impact on the performance of the various metals. This information allows mining stakeholders to focus forecasting efforts on the identified factors in order to predict the future performance of the metals in question. Mining stakeholders refers to management of mines as well as any potential investors.

(Adams and MacKay, 2007) describes change points as locations where abrupt variations in the generative parameters of a sequence of data occur. Change point detection involves ascertaining whether or not a change (or several) in the parameters of a sequence has occurred and identifying the times of these change(s). Change points may occur when the process that generates the data changes or if the method of data collection is altered. Change point detection has applications in areas such as finance, biometrics and robotics.

METHODOLOGY

The Bayesian Change Point Analysis approach was used to identify change points in this paper. The overview is detailed below.

Bayesian Change Point Analysis (BH)

The Bayesian approach (Barry and Hartigan, 1993) states that the sequence of data has an underlying sequence of generative parameters that are divided into contiguous blocks. The blocks are divided so that observations in each block have equal parameter values. The beginning of each block will therefore represent a change point.

BH assumes that the observations X_i are independent and that the probability of a change at a position i is p. Each X_i has a density that is dependent on $\theta_i, i = 1,...,n$. The number of blocks is unknown with partitions $\rho = \{i_0,i_1,...,i_b\}$ such that $0 = i_0 < i_1 < ... < i_n$ and $\theta_i = \theta_{ib}$ when ir−1 < i 6 ir .

(Barry and Hartigan, 1993) defined X_{ij} as the sequence of observations from X_{i+1} to X_j, where X_{i+1} is the observation at time point i + 1. The density of x_{ij} given θ_j (when $\theta_{i+1} = \theta_{i+2} = ... = \theta_j$) is defined as $f_{ij}(x_{ij}|\theta_j)$.

The parameter values are approximated by Markov sampling techniques. The assumption is that the parameter sequence θ_i forms a Markov Chain. Given θ_i, θ_{i+1} equals θ_i with probability $1 - p_i$ or has a density $f(\theta_{i+1}|\theta_i)$ with probability p_i.

The probability of a partition $\rho = (i_1,i_2,...,i_b)$ is given by $f(p) = Kc_{i0}i_1c_{i1}i_2 \cdot\cdot\cdot c_{ib-1}i_b$, where c_{ij} are prior cohesions for each possible block ij. Cohesions are of the form

$$c_{ij} = (j-i)^{-3} \text{ for } 0 < i < j < n,$$
$$c_{ij} = (j-i)^{-2} \text{ for } i = 0 \text{ or } j = n,$$

and

$$c_{0n} = n^{-1}$$

Independent priors are specified for each parameter: p,μ_0,σ^2 and $w = \sigma^2/(\sigma^2_0 + \sigma^2)$.

$$\begin{aligned} f(\mu_0) &= 1, & -\infty \leqslant \mu_o \leqslant \infty, \\ f(\sigma^2) &= 1/\sigma^2, & 0 \leqslant \sigma^2 \leqslant \infty, \\ f(p) &= 1/p_0, & 0 \leqslant p \leqslant p_0, \\ f(w) &= 1/w_0, & 0 \leqslant w \leqslant w_0. \end{aligned}$$

p_0 and w_0 are prespecified numbers in [0,1]. (Barry and Hartigan, 1993) recommend using $p_0 = 0.2$ and $w_0 = 0.2$. BH estimates the probability of a change point at each location. This is a distinct advantage over BP, that merely estimate the locations of the change points. BH provide a more informative summary reflecting the degree of uncertainty in the change points. For a full breakdown of the approach refer to (Barry and Hartigan, 1993).

RESULTS

Bayesian Change Point Analysis (Barry and Hartigan, 1993)

Included below are tables indicating the most probable change points of the production volumes and sales of Gold, PGMs, Iron Ore and Manganese. Each table shows the date, probability of a change point occurring at the identified point and the posterior estimated mean of the specific point.
BCP: Bayesian Change Point Analysis
Date: Date of identified change point
Prob.: Posterior Probability of a Change Point at that specific location X1:
Posterior estimate of mean

Table 1:BCP Analysis for Gold Production

	Date	Prob.	X1
12	12/2003	0.704	192.94
36	12/2005	0.742	154.83
60	12/2007	0.736	129.81
117	09/2012	0.814	88.70

Table 2: BCP Analysis for Gold Sales

	Date	Prob.	X1
41	05/2006	0.658	2617.50
104	08/2011	0.996	4953.19
153	09/2015	0.702	5183.62

Table 1 identified change points at the end of the calendar year for 2003, 2005 and 2007. These changes were due to the new production plan implemented at the beginning of each year. This production plan takes into consideration the increased production costs due to rises in labour, electricity and fuel as well as the fact that the gold near the surface has been mined, forcing the miners to dig deeper. The 09/2012 change point was caused by strikes in various Gold mines.

Since the Sales values were in Actual Rands, the change points identified were linked directly to the Rand weakening against the Dollar. In order to be able to obtain meaningful results, sales volumes will be used in future papers.

Table 3: BCP Analysis for PGM Production

	Date	Prob.	X1
60	12/2007	0.732	105.62
108	12/2011	0.738	107.11
133	01/2014	0.624	82.20
146	02/2015	0.708	79.68
156	12/2015	0.626	99.82

Table 4: BCP Analysis for PGM Sales

	Date	Prob.	X1
41	05/2006	0.934	4068.95
62	02/2008	0.924	6755.37
70	10/2008	0.668	6882.54
108	12/2011	0.890	6913.25
134	02/2014	0.694	7128.10
146	02/2015	0.930	6320.43
156	12/2015	0.870	8025.27
160	04/2016	1.000	6351.71
162	06/2016	1.000	11194.93

In the case of PGMs, it is important to note that South Africa supplies the vast majority of the world's PGMs. This means that a decrease in production of South Africa's PGMs decreases the world's PGMs supply and therefore drives up the price and vice versa. The change points of 12/2011 and 01/2014 were both caused by miners strikes, which in-turn decreased supply and increased the price of PGMs and increased the sales for the corresponding period.

Table 5: BCP Analysis for Iron Ore Production

	Date	Prob.	X1
65	05/2008	0.960	72.74
96	12/2010	0.984	99.86
97	01/2011	0.742	71.87
112	04/2012	0.794	109.14
117	09/2012	1.000	127.60
118	10/2012	0.994	86.99
125	05/2013	0.646	125.12
126	06/2013	0.992	143.75
131	11/2013	0.992	118.85
132	12/2013	0.992	160.54
136	04/2014	1.000	122.60
156	12/2015	0.988	116.10
161	05/2016	1.000	96.53

Table 6: BCP Analysis for Iron Ore Sales

	Date	Prob.	X1
68	08/2008	0.988	1538.15
86	02/2010	0.996	1988.99
92	08/2010	0.642	3195.04
98	02/2011	0.750	4484.71
101	05/2011	0.850	5093.30
103	07/2011	1.000	5969.63
104	08/2011	1.000	3581.64
108	12/2011	0.658	5192.57
110	02/2012	0.632	4221.45
115	07/2012	0.948	4943.90
119	11/2012	0.856	3749.76
136	04/2014	0.692	5673.46
139	07/2014	0.620	4596.12
150	06/2015	0.640	3421.22
158	02/2016	0.896	2863.32
166	10/2016	0.910	3527.82

In Table 5,the change points up to 2014 were fluctuations in production, with the base level steadily increasing. The fluctuations were caused by mining related issues. In 2014 the Iron Ore market became over saturated and the price of Iron Ore plummeted. The main reason for the oversupply of Iron Ore, was that the Chinese Steel Industry cut back on their production. The plummeting Iron Ore process resulted in mines cutting back in their production volumes.

Due to the apparent volatility of Iron Ore Production Volumes, quarterly productions volumes will be used in future studies.

Table 7:BCP Analysis for Manganese Production

	Date	Prob.	X1
72	12/2008	0.906	91.27
99	03/2011	0.906	104.95
124	04/2013	0.704	136.22
133	01/2014	0.664	149.16
141	09/2014	0.804	197.75
154	10/2015	1.000	229.19
161	05/2016	0.908	177.11

Table 8: BCP Analysis for Manganese Sales

	Date	Prob.	X1
63	03/2008	1.000	761.59
67	07/2008	0.996	1567.81
69	09/2008	1.000	2440.71
73	01/2009	1.000	1030.50
146	02/2015	0.804	1407.18
159	03/2016	0.912	883.33
162	06/2016	0.972	1839.08
166	10/2016	1.000	1337.40
167	11/2016	1.000	2854.20
168	12/2016	1.000	3929.91

Table 7 shows various change points, the most significant being the change points in 12/2008 and 10/2015. The former change point was caused by the economic crisis of 2008 and the latter being attributed to the dropping Chinese Steel production levels. Since Chinese Steel production is the main user of steel products and 95% of Manganese is used in steel production, the Chinese Steel Industry should be studied in order to gain an understanding of the performance of Manganese in the South African context.

CONCLUSION

This paper used the Bayesian Change Point Approach (Barry and Hartigan, 1993) to identify change points in the export volumes and total sales of Gold, PGMs, Iron Ore and Manganese. The identified changes were linked to causative events and/or factors in order to create a descriptive model of the factors that have significant impacts on the supply and/or demand of these industrial metals.

The descriptive model will give mining stakeholders information on the most significant factors that influence the production and/or sales of each metal. This information will help the stakeholders to focus forecasting efforts on the identified factors and events. These forecasts will in turn allow mines to more effectively plan production levels.

REFERENCES

Adams, R. P. and MacKay, D. J. 2007. Bayesian online changepoint detection.

Bai, J. and Perron, P. 2003. Computation and analysis of multiple structural change models. *Journal of Applied Econometrics*, 18(1):1–22.

Barry, D. and Hartigan, J. A. 1993. A bayesian analysis for change point problems. *Journal of the American Statistical Association*, 88(421):309–319.

Chamber of Mines South Africa. *Facts & Figures 2010* (pp 32). http://www.chamberofmines.org.za/industry-news/publications/facts-and-figures

Chamber of Mines South Africa. *Mining in South Africa: The Challenges and the Opportunities* (pp 20). http://http://www.chamberofmines.org.za/industry-news/publications/presentations/send/72015/269mining-in-south-africa-the-challenges-and-opportunities

Erdman, C. and Emerson, J. W. 2007. bcp: An r package for performing a bayesian analysis of change point problems. *Journal of Statistical Software*, 23(3).

Lee P.M. 1989. *Bayesian Statistics: An Introduction*, 1: 36

Naylor, T.and Finger,J.M. 1967. Verification of Computer Simulation Models. *Management Science*, 2:B92–B101.

Olshen, A. B., Venkatraman, E. S., Lucito, R., and Wigler, M. 2004. Circular binary segmentation for the analysis of array-based dna copy number data. *Biostatistics*, 5(4):557–572.

Renganathan, V. 2016. Overview of frequentist and bayesian approach to survival analysis. *Applied Medical Informatics* , 38(1):25.

Sen, A. and Srivastava, M. S. 1975. On tests for detecting change in mean. *The Annals of Statistics*, 3(1):98–108.