

Modeling commodity prices for valuation and hedging of mining projects subjected to volatile markets

Mathieu Sauvageau

Doctor of Philosophy

Mining and Materials Engineering

McGill University

Montreal, Quebec

July 2017

A thesis submitted to McGill University in partial fulfillment of the
requirements of the degree of Doctor of Philosophy

©Mathieu Sauvageau, 2017

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisor Dr. Mustafa Kumral for giving me the opportunity to perform research in a field which I am passionate about, finance. I hope my contribution will help the Mineral economics, mine reliability and asset management lab to gain a greater exposure in the research community.

I would also like to thank my lab colleagues Salvatore, Martha, Asli, Tarek, Burak, Aldin and Faruk for sharing your thoughts on my research, but more importantly, your extremely rich and diversified backgrounds. A special mention goes to Salvatore for cooking me pasta and showing me how to make my own pizza Margherita.

I would also like to thank my life partner, Joanie, for accepting to put all kinds of projects on hold while I am still a student. Believe me, I will prove to you that it was well worth the wait! I would also like to thank my parents for encouraging me to pursue my studies. My mother always said when you like what you are doing you will never have to work. I pretty much apply this concept on an every day basis and I can't stress enough how much it is true. My father has always been there for me, giving me advice when needed. I am glad to have a very supportive family which values education and hard work.

CONTRIBUTION OF AUTHORS

The author of this thesis is the primary author for all manuscripts contained within. Professor Mustafa Kumral is the supervisor of the candidate Ph.D., and is included as a second author for each of these manuscripts.

Chapter 3 - Sauvageau, Mathieu, and Mustafa Kumral. “Analysis of Mining Engineering Data Using Robust Estimators in the Presence of Outliers.” *Natural Resources Research* 24.3 (2015): 305-316.

Chapter 4 - Sauvageau, Mathieu, and Mustafa Kumral. (under review) “A Kalman filtering approach to model Net Present Value of an iron mine.” *Natural Resources Research*.

Chapter 5 - Sauvageau, Mathieu, and Mustafa Kumral. “Genetic algorithms for the optimization of the Schwartz-Smith two-factor model: A case study on a copper deposit.” *International Journal of Mining, Reclamation and Environment*. <http://dx.doi.org/10.1080/17480930.2016.1260858>

Chapter 6 - Sauvageau, Mathieu, and Mustafa Kumral. Under revision “Cash-Flow at Risk valuation of mining project using Monte-Carlo simulations calibrated on historical data”, *The Engineering Economist*.

ABSTRACT

Duration between discovery of a mineral deposit and delivery of the material (e.g. metal or concentrate) yielded from this deposit to the market can take several years. After the discovery, a feasibility study is conducted to see if the mining operation on this deposit is economically viable. This feasibility study is exposed to two types of uncertainties, which add to risks to the mining project. These are (1) technical risks arising from sparse data (e.g. grade, recovery and geotechnical characterization) and (2) financial risks arising from unknown future events (e.g. commodity price, discount rate and exchange rate). Among the others, commodity price is a significant concern for the executives of mining enterprise. Given that mining products and their derivatives are traded in commodity, stock and future markets, market dynamics are very complex. Furthermore, it is very sensitive to politics of global world and very open to speculation and manipulation as well as demand and supply. In the past, the mining industry witnessed that many mining operations were suspended or ceased due to unresponsiveness to price fluctuations. Project valuation based on long-term price is a quite naive approach at present day. This can jeopardize the financial resources of the investor company. Therefore, the risks associated with commodity price are assessed, quantified, mitigated, diversified or managed. The analysis of commodity prices starts with the study of historical transactions in financial markets. To facilitate the analysis, it is often necessary to convert commodity prices into returns. Then, the next task is to model the distribution of returns using a statistical distribution. One of the main characteristics of the distribution of commodity price returns is that it tends to have excess kurtosis. This can be explained either by a stochastic volatility or jump component in the diffusion equation describing

the evolution of prices. For this reason, it is necessary to consider other models than the Geometric Brownian Motion and Mean-Reverting price models when modeling the dynamics of commodity prices. The objective of this thesis is to construct a robust workflow capable of reproducing the observed price dynamics in the commodity markets. With such calibrated models, it is possible to value mining projects or estimate their exposure to market risk. In the first case, the valuation process is made in a risk-neutral framework using a Real Options approach. In the second case, real world probabilities are used to simulate commodity price paths and assess how a mining project may be exposed to market price fluctuations. Following an introduction and a Literature review, the thesis is divided in four additional parts, corresponding to four different publications. In the first publication, the use of robust estimators for the detection and mitigation of outliers is investigated. The paper starts with an overview of multiple linear regression and assess how the model assumptions can be violated. The second part of the paper deals with detecting outliers using studentized residuals or the Mahalanobis distance. Then robust regression is used to diminish the effects of outliers in mining engineering data including price. The second paper investigates how the dynamics of iron ore future can be modeled with a dynamic linear model. Traditionally, iron ore futures have been traded using long-term commitment contracts. This paper investigates how relatively recent financial instruments such as futures on iron ore can affect the NPV profile of an iron ore project. The third paper deals with the optimization of the parameters in a commodity model using a genetic algorithm. With correctly calibrated parameters, Monte Carlo simulations of commodity spot and futures are performed and an active trading strategy is implemented in an NPV valuation framework. The last publication deals with

the choice of the stochastic process when measuring market risk of a mining project. Several stochastic processes are calibrated on historical data and used to calculate the cash flow at risk of a mining project. The models are calibrated using an hybrid metaheuristic calibration strategy. Particle swarm optimization is first used to find a solution close to the global minimum. Then, a gradient based routine is used to find the optimal solution.

ABRÉGÉ

Plusieurs années se produisent habituellement entre la découverte d'un nouveau gisement minéral et la livraison des matériaux produits sur les marchés financiers. Une étude de faisabilité est généralement effectuée pour évaluer la viabilité du projet minier. L'une des tâches les plus importantes est d'évaluer la valeur actuelle nette du projet minier. Cette variable est la somme des flux de trésorerie actualisés que le projet minier générera au cours de sa durée de vie. Les flux de trésorerie sont actualisés à un niveau approprié pour refléter le risque du projet. Quantifier le risque des projets miniers peut être une tâche encombrante. En effet, les projets miniers sont soumis à de multiples sources d'incertitudes qui doivent être estimées à l'aide de modèles stochastiques. L'une de ces inconnues est le prix des produits de base, qui doit être projeté sur toute la durée du projet minier. Pour faciliter l'analyse, il est souvent nécessaire de convertir les prix des matières premières en rendements. Ensuite, la tâche suivante consiste à modéliser la distribution des rendements à l'aide d'une distribution statistique. L'une des principales caractéristiques de la distribution des rendements des prix des produits de base est qu'il tend à avoir un excès de kurtosis. Cela peut être expliqué soit par une volatilité stochastique ou une composante de saut dans l'équation de diffusion décrivant l'évolution des prix. Pour cette raison, il est nécessaire d'envisager d'autres modèles que les modèles de mouvement géométrique brownien et de réversion moyenne lors de la modélisation de la dynamique des prix des produits de base. L'objectif de cette thèse est de construire un workflow robuste capable de reproduire la dynamique des prix observée sur les marchés des matières premières. Avec ces modèles calibrés, il est possible d'évaluer les projets miniers ou d'estimer leur exposition au risque de marché. Dans le premier cas, le processus d'évaluation

est effectué dans un cadre risque-neutre en utilisant une approche axée sur les options réelles. Dans le second cas, les probabilités réelles sont utilisées pour simuler les scénarios de prix des produits de base et pour évaluer comment un projet minier peut être exposé aux fluctuations des prix du marché. La thèse est divisée en quatre parties, correspondant à quatre publications différentes. Dans la première publication, l'utilisation d'estimateurs robustes pour la détection et l'atténuation des valeurs aberrantes est étudiée. L'article commence par un aperçu de la régression linéaire multiple et évalue comment les hypothèses du modèle peuvent être violées. La deuxième partie de l'article traite de la détection des valeurs aberrantes en utilisant des résidus studentisés ou la distance de Mahalanobis. Puis, une régression robuste est utilisée pour diminuer les effets des valeurs aberrantes dans les données d'ingénierie minière. Le second article étudie comment la dynamique du minerai de fer peut être modélisée avec un modèle linéaire dynamique. Traditionnellement, les contrats à terme de minerai de fer ont été négociés à l'aide de contrats d'engagement à long-terme. Cet article étudie comment des instruments financiers relativement récents comme les contrats à terme sur le minerai de fer peuvent affecter le profil de la VAN d'un projet de minerai de fer. Le troisième article traite de l'optimisation des paramètres dans un modèle de marchandise utilisant l'algorithme génétique. Avec des paramètres correctement étalonnés, des simulations Monte Carlo de spot et de contrats à terme sur matières premières sont effectuées et une stratégie de négociation active est mise en œuvre dans un cadre d'évaluation VAN. La dernière publication traite du choix du processus stochastique lors de la mesure du risque de marché d'un projet minier. Plusieurs processus stochastiques sont étalonnés sur des données historiques et utilisés pour calculer le flux de trésorerie à risque d'un projet minier. Les modèles sont calibrés en utilisant une stratégie de calibration métaheuristique

hybride. L'optimisation des essaims de particules est d'abord utilisée pour trouver une solution proche du minimum global. Ensuite, une routine basée sur le gradient est utilisée pour trouver la solution optimale.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
CONTRIBUTION OF AUTHORS	iii
ABSTRACT	iv
ABRÉGÉ	vii
LIST OF TABLES	xiii
LIST OF FIGURES	xv
1 Introduction	1
1.1 Overview	1
1.2 Original Contributions	4
2 Literature Review	7
2.1 Overview	7
2.2 Market risk regarding commodity markets	9
2.3 Robust regression	12
2.4 Stochastic processes	14
2.5 State space representation of stochastic processes	18
2.6 Genetic algorithms	21
2.7 Options pricing theory in financial markets	23
2.8 Introduction to Real Options valuation	25
3 Analysis of mining engineering data using robust estimators in the presence of outliers	29
3.1 Abstract	29
3.2 Introduction	30
3.3 Outliers detection	34
3.4 Robust regression	35
3.4.1 The least absolute regression model (L_1 estimator)	37
3.4.2 M-estimator	38
3.4.3 LTS estimator	39
3.4.4 LMS estimator	39
3.4.5 MM-estimator	40
3.5 Case Studies	41

	3.5.1 Mineral processing costs recovery relation	41
	3.5.2 Iron and alumina grade relation	45
3.6	Conclusion	48
3.7	Chapter Conclusion	49
4	A Kalman filtering approach to model net present value of an iron mine	50
4.1	Abstract	50
4.2	Introduction	51
4.3	Methodology	56
	4.3.1 Model Input Data	56
	4.3.2 Maximum Likelihood Optimization	61
	4.3.3 Schwartz Two-Factor Model	62
4.4	Results	65
	4.4.1 Time Series Analysis	65
	4.4.2 Kalman Filtering	69
4.5	Conclusion	78
4.6	Appendix	80
4.7	Chapter Conclusion	81
5	Genetic algorithms for the optimization of the Schwartz-Smith two-factor model: A case study on a copper deposit	82
5.1	Abstract	82
5.2	Introduction	83
5.3	Methodology	88
	5.3.1 Schwartz-Smith two-factor model	88
	5.3.2 Genetic algorithms	92
	5.3.3 NPV Incorporating an active trading strategy	94
5.4	Case study	96
	5.4.1 Data	96
	5.4.2 Two-factor model calibrated on artificial data	101
	5.4.3 Two-factor model calibrated on crude oil futures prices	103
	5.4.4 Two-factor model calibrated on copper prices	106
	5.4.5 NPV valuation of a copper mine	109
5.5	Discussion	112
5.6	Conclusion	115
5.7	Chapter Conclusion	117
6	Cash-flow at risk valuation of mining project using Monte-Carlo simulations calibrated on historical data	118
6.1	Abstract	118
6.2	Introduction	118
6.3	Literature review	120
6.4	Methodology	123

6.4.1	Merton jump-diffusion models	123
6.4.2	Stochastic volatility models	124
6.4.3	Estimation of CFaR	127
6.5	Case study	128
6.5.1	Data	128
6.5.2	Model calibration	134
6.5.3	CFaR analysis	137
6.6	Discussion	139
6.7	Conclusion	141
7	Conclusion	145

LIST OF TABLES

<u>Table</u>	<u>page</u>
3-1 Summary of the OLS Performed on the Original Processing Costs and Recovery Dataset as Well as on the Cleaned Dataset Where Observation 14 was Dropped.	43
3-3 Summary of OLS-, L1-, M-, LTS-, LMS-, and MM-Estimation Intercept and Coefficient	45
3-5 Summary of OLS-, L1-, M-, LTS-, LMS-, and MM-Estimation Intercept and Coefficient	47
4-1 Summary statistics for the training and prediction sets of the iron ore time series.	65
4-2 Initial and optimized parameters with the MLE algorithm after 300 iterations	69
5-1 Parameters used in simulations of the SSTF model. These parameters are taken from the original publication of Schwartz and Smith (2000a).	97
5-2 Minimum, maximum and the simulated values of the parameters in the SSTF model simulations. Simulated values are the parameters the GA will try to retrieve and the range between minimum and maximum value is the search space for the optimal solution.	101
5-3 Optimal parameters after 25 generations of the GA for different period lengths T.	103
5-4 Optimal solution of the GA after 25 generations	104
5-5 Mean error, standard deviation of errors and mean absolute deviation of the fitted maturities for the crude oil futures dataset.	106
5-6 Parameters of the original Goodwin (2013) paper and parameters optimized with the GA for two different sub-periods affected by different market dynamics.	107
5-7 Mean error, standard deviation of errors and mean absolute deviation of the fitted maturities for the copper futures dataset.	108

5–8	cash flows of the mining project. The project lasts 12 years with an initial outlay of \$175 millions. Mining costs are the costs to extract the materials from the mine and processing costs are the costs to extract the metal from the ore.	110
6–1	Descriptive statistics for the gold, silver and platinum spot price returns as well as the crude oil futures contracts returns. The period 1 covers the weeks between 8 February 1993 and 27 December 2004. The second period covers the weeks between 10 January 2005 and 5 December 2016. The total period covers the whole time range.	131
6–2	Jarque-Bera statistics for normality on the 99% level of confidence for the four different commodities spot price returns. .	134
6–3	Calibrated parameters of the GBM, Heston SV and MJD on gold, silver platinum and crude oil between 15 February 1993 and 5 December 2016.	135
6–4	In-sample and out-of-sample RMSE for the different commodity datasets. The KF is calibrated using 80% of the first observations. The out-of-sample RMSE uses the calibrated parameters on the remaining 20% of the dataset.	137
6–5	Characteristics of the open-pit gold mine. These parameters are used for the CFaR calculations.	137
6–6	CFaR of the gold-mining project for quarterly, monthly, weekly and daily time intervals	138

LIST OF FIGURES

<u>Figure</u>	<u>page</u>
3-1 On the left, processing costs as function of recovery. A line fitted with OLS estimation is superimposed. On the right, corresponding studentized residuals of the OLS regression.	43
3-2 Comparison of robust estimators with OLS for the processing costs and recovery case study.	45
3-3 On the left, iron concentration in function of alumina concentration. A line fitted with OLS estimation is superimposed. On the right, corresponding studentized residuals.	46
3-4 Comparison of robust estimators with OLS for the iron and alumina concentration case study.	47
4-1 The solid black line represents closest to maturity daily Iron Ore 62% Fe, CFR China contracts, which can be interpreted as a proxy for TSI index level (or spot price). For each trading day, forward curve is extracted for contract maturing in the next year. The graphic is generated using the Schwartz 97 R package (Erb et al., 2014)	58
4-2 Iron Ore 62% Fe, CFR China contracts, TSI index level	60
4-3 Training (above) and prediction (below) partitions for the iron ore time series. The histogram represents the distribution of log returns. The blue line is a kernel density estimation (kde) plot using a Gaussian kernel and the dashed green line represents a normal distribution with the theoretical mean and variance presented in Table 4-1.	66
4-4 Aggregated volume from all contracts for the prediction partition of the iron ore futures time series.	67
4-5 The top plot represents a heat map of traded volumes in function of maturity and date. Red color indicates highest traded volumes. The bottom plot represents the sum of volume of all contracts for a given maturity. Nearest maturities have the greatest cumulative volumes.	68

4-6	Left plot shows nearest actual contracts (black line) as well as future curves for each trading day in the studied period of time. The right plot shows the fitted corresponding variables estimated with Kalman filter.	70
4-7	Simulated paths (in gray color) with parameters derived from the MLE algorithm. The continuous red line represents the mean of all the simulations and the red dashed lines represent the 95% confidence interval. The continuous green line is the actual price.	71
4-8	Kernel density plot of residuals for the first five maturing contracts.	72
4-9	Histograms of NPV for the three different scenarios. The neutral scenario is in blue, while the bullish and bearish scenarios are respectively in green and in red.	73
4-10	NPV with a varying proportion of production sold in futures markets. The analysis is performed for the optimistic, neutral and pessimistic price profiles in futures markets.	75
5-1	Simulated time series containing 259 observations. Plot a) shows the long-term price index as well as the sum of the long-term price and short-term deviations. Plot b) shows simulated commodity contracts with maturities of 1, 5, 9, 12 and 17 months.	98
5-2	Time to expiration of crude oil contracts aggregated from Quandl API. The time series ranges between 1990-01-02 and 1995-02-17. Contracts are expiring up to 17 months in the future.	100
5-3	Crude oil futures contracts aggregated from the Quandl API. For each trading week (time t), contracts are expiring at 1, 5, 9, 12 and 17 months into the future.	101
5-4	Convergence of the GA. The blue line represents the average fitness of the population at each generation. The red line represents the solution within the population having the highest fitness.	103
5-5	Plot a) represents the t+1 month crude oil contract (in blue) as well as the KF estimate of the t+1 future contract (in green). The long-term price estimate of the KF is represented in red. Plot b) represents residuals (in %) of the KF estimates for each maturing contract.	105

5-6	Plot a) of the Figure represents the $t+1$ month copper contract (in blue) as well as the KF estimate of the $t+1$ future contract (in green). The long-term price estimate of the KF is represented in red. Plot b) represents residuals (in %) of the KF estimates for each maturing contract.	108
5-7	Plot a) of the Figure represents the KF estimate of the long-term trend and the estimated spot price. Plot b) represents simulation of copper contracts for the calibrated SSTF model. Plot c) represents the size of the stockpile, conditional on difference between estimated spot price and the long-term trend.	111
5-8	Value added (alpha) of the active trading strategy compared to the basic NPV investment. Alpha is calculated by differencing NPV using an active strategy to passive NPV for each simulation outcome.	112
6-1	Price levels for gold, silver, platinum (in \$USD/ounce) and crude oil futures (in \$USD/barrel) during 8 February 1993 and 8 November 2016.	129
6-2	Logarithm of returns of gold spot prices during 8 February 1993 and 8 November 2016. The first period ranges between 8 February 1993 and 27 December 2004. The second period ranges between 29 December 2004 and 8 November 2016. . .	130
6-3	Boxplots for gold, silver, platinum and crude oil futures price returns. The points are considered as outliers when the distribution is assumed to be normal.	132
6-4	Histogram of spot gold price returns. The dashed PDF corresponds to a normal distribution with the same mean and variance as observed in the empirical dataset. The continuous PDF is calculated with a kernel density estimator.	133

CHAPTER 1

Introduction

1.1 Overview

The valuation of a mining project requires assessing risk such as market, operational, geological, political and environmental risk. To assess the geological risk, workflows involving stochastic modeling of an ore body followed by its production schedule optimization to determine future production rates have proven to be very promising. However, when converting forecasted production rates into cash flows, forecasted spot prices of a commodity also have to be taken into account.

The Net Present Value (NPV) approach is widely used to assess the present economic value of a mining project over its entire life. The approach consists in projecting future cash flows and then discounting them with a rate reflecting the risk of the investment. Discounted cash flows are then summed and compared to the initial expenditures required to undergo the mining project, and if the difference is positive, the project should be accepted. However, mining projects can often last more than ten years, a period where cycles, jumps and spikes will influence commodity prices. These changes in prices will greatly influence forecasted cash flows, affecting the value of a mining project. Moreover, choosing the right discount rate to calculate NPV may be very subjective. A too small discount rate may make an unviable project look profitable while a too big discount rate may underestimate cash outflows needed for the completion phase of a mining project. Finally, managerial flexibility is not addressed when performing NPV valuation. For these

reasons, the NPV valuation technique may be too rigid to reflect the true complexity of a mining project.

A promising and increasingly popular approach for valuing mining projects is the Real Options valuation (ROV). This method of valuation is similar to NPV valuation in the sense that present and future cash flows are discounted and summed to value an investment. The difference is that the management team of a mining project are given options, for example, to mothball, close, open, accelerate or pause a project. Also, instead of affecting a project directly, managers are given options to hedge their production using options or future markets. This increase in flexibility adds some value to a mining project but can be hard to put in practice because of the costs associated with changes in the project. However, when commodity markets are subjected to extreme events such as spikes or strong cycles, ROV may be necessary as a complementary approach to NPV.

Workflows based on ROV are usually implemented in a Monte Carlo (MC) simulation framework. The general MC framework requires modeling different scenarios of spot prices using a stochastic process. The parameters of the stochastic process may be calibrated on historical data or inferred from the knowledge of the modeler. If the historical approach is chosen for the calibration, the stochastic process parameters are adjusted to reflect price dynamics in the recent past. The Kalman filter has proven to be a very useful tool for the calibration of stochastic processes on historical data. When using the Kalman filter, the stochastic process is written in a state-space representation which decomposes the process into two parts: the transition equation and the measurement equations. The transition equation is the true process while the measurement equation is the measured process. The measurements are usually noisy. A maximum likelihood routine can be used to adjust the parameters

governing the transition equation to better fit historical data. The procedure consists of applying the Kalman filter on historical observations with a vector of parameters and computing the logarithm of likelihood estimate. Then, an optimization routine is used to adjust the vector of parameters to maximize the likelihood. In practice, this procedure is converted to a minimization problem by taking the negative of the likelihood.

The optimization of the parameters of the Kalman filter is a major task in this thesis. Usually, the optimization routines used to perform the calibration are gradient based and can be stuck in a local minimum. To overcome this problem, several different starting values can be tested to ensure the convergence to a global minimum. However, a better approach is to use a different class of optimization routines that can handle non-linear and multidimensional cost functions. Namely, in this thesis, the use of a genetic algorithm and a particle swarm optimization are tested for the calibration of historical parameters. This class of algorithm mimics the behavior of nature to solve problems. In the case of the genetic algorithm, the theory of evolution is implemented in the algorithm. The particle swarm optimization mimics how swarm of individuals in a herd collaborate to survive. In any case, several solutions are tested at each iteration and the likelihood to get stuck in a local extremum thus is diminished.

Once the parameters of the stochastic process have been calibrated, they can be used to perform simulations. An active trading strategy using the Kalman filter and the Schwartz-Smith two-factor (SSTF) model is tested. The SSTF makes the assumption that the price of a commodity can be decomposed into two different terms. The long-term price is expected to evolve following a (geometric Brownian motion) GBM while the short-term price is mean-reverting. The application of the active trading strategy is implemented using

both price components in the SSTF. When the short-term price drives the spot price lower than implied by its long-term trend, a percentage of the production of a mining project is stockpiled for later use. When the short-term price drives the spot price higher, a part of the production that was previously stockpiled is sent to the market. This mean-reversion strategy can increase the NPV of a mining project by 5% to 10%.

Another use of the Kalman filter calibration workflow is to assess the market risk of a mining project. The commodity returns are generally better fitted with a fat-tail distribution. This is due to the fact that prices can jump or be affected by volatility clusters. These features are not well captured by the GBM. On another hand, the Merton jump (MJ) and Heston stochastic volatility (SV) model are able to reproduce behavior observed in the markets. The use of the MJ and SV models for the calculation of cash-flow at risk (CFaR) are investigated. These stochastic processes are better suited to reproduce fat-tails and provide a value of CFaR which is more conservative than the GBM.

The next section lists the different publications associated with this thesis.

1.2 Original Contributions

The first part of the project is to apply robust regression algorithms on a dataset containing outliers. One must be very careful before deleting outliers from a dataset since they can be good data points that may have resulted from an extreme case of the sample generating process, or perhaps a secondary process. Several tools exist to detect and calculate the influence and leverage of outlier points. Then, using robust regression estimators, parameters from the regression can be estimated without any bias. The work done in robust regression has led to the publication of an article called Analysis of mining engineering data using robust estimators in presence of outliers, in the Natural

Resources Research Journal. The original contribution of this paper is the development of a workflow capable of detecting outliers and downsizing them to reduce the effect on regression models. The tools developed in this research are later used to detect abnormal jumps in the distribution of commodity returns.

The second publication uses a new approach using long-term commitment contracts (LTC) and futures contracts to value iron ore projects. This approach supposes that the management team of the iron mine has the ability to use two types of financial instruments to sell their production. The first instrument is the long-term commitment contracts and the second instrument is the futures contract which can be traded in financial markets. The long-term instrument is considered as a risk-free instrument, while the futures contract is actively traded and fluctuates. The Kalman filter is used to estimate the dynamics of futures contracts, iteratively predicting unknown observations, then updating the parameters to improve the next prediction. The Kalman filter is the ideal estimator when the transition equation is linear and when the residuals are following a normal distribution. However, iron ore markets are subject to extreme events, such as jumps, which cannot be considered linear. These jumps are generally followed by an increase of volatility in markets. The first objective of this paper is to improve the Kalman filter estimates by scaling the observation uncertainties with the volume of traded contracts. Then, simulations of iron ore prices are performed with correctly calibrated parameters. Results show that this workflow can be used to adjust the proportion of the productions sold using LTC or futures contracts according to the risk appetite of the company. Results show that the calibration and NPV simulation workflow can be effectively used to assess the profitability of a mining project, accounting for the exposure in futures markets.

The third publication uses a genetic algorithm (GA) to calibrate the parameters of the Schwartz-Smith two-factor (SWTF) model on oil and copper futures contracts. Usually, the optimization is done using a deterministic optimization called the Expectation-Maximization (EM) algorithm which can get stuck in a local maximum. Although they are slower than EM algorithms because they use random number generators to search for the optimal solution, genetic algorithms optimize a population of solutions instead of working on only one solution at the time. Moreover, a constraint on the range parameter can be applied to ensure the parameter has a sound economic meaning. To demonstrate the performance of the proposed approach, a case study was conducted on a copper deposit. The simulations were based on the SWTF model whose parameters are determined by GA. An active management strategy of the stockpile, dependent on discrepancies in commodity futures prices is tested. Results show that the active management strategy produces positive returns over the passive investment approach.

The last publication compares several models used for the estimation of the CFaR of a mining project. The model parameters are calibrated using the Kalman filter. Then, Monte-Carlo simulations using the calibrated parameters are performed to compare the stochastic volatility model, the Merton jump diffusion process, and the geometric Brownian motion when assessing market risk exposure of a mining project. The contribution of this research is to provide a robust workflow to calibrate commodity price dynamics models to historical data using the Kalman filter. The parameters of the model are fitted using a particle swarm optimization routine combined with a gradient-based search method.

CHAPTER 2

Literature Review

2.1 Overview

In this section, an extensive literature survey on the concepts tackled in the thesis is presented. The literature survey revolves around the concept of options pricing in financial markets and its extension to a Real Option framework. Other concepts, such as genetic algorithms and the Kalman filter are introduced because they can be used to simulate or derive unknown parameters in stochastic process models.

The first section discusses risk related to commodity markets and strategies to protect a mining company from that risk. First of all, some basic concepts about the mining industry are covered. After, some basic management concepts on mining projects are covered. These management strategies involves option and futures contracts and are designed to reduce volatility in quarterly earnings of mining companies. Then, a review of how volatility forecasting models have been used to forecast volatility of commodity prices is made. These commodities include petroleum and minerals, but also agricultural commodities such as soybeans, and wheat.

The second section introduces robust regression as a tool for outlier detection and mitigation. It is well documented that financial markets are subjected to crashes. These extreme movements cannot be modeled with the normal distribution and will generally make the conventional statistical estimates biased. For this reason, another class of robust statistics models was introduced. One of the main tasks when treating for outliers in a dataset is to detect them. A

review of outlier detection methods is made as well as a review of robust estimation techniques. These tools are used later in this thesis to accommodate for outliers in commodity market returns.

The next section introduces some stochastic processes used to generate simulations in stock, bonds, commodities and futures markets. Stochastic processes are very useful to simulate possible price paths and assess how a mining project can be affected by the market fluctuations. Another popular use of these stochastic processes is to assess market risk of mining projects. Then, it is possible to estimate how a hedging program can help to mitigate this market risk. A variety of stochastic processes exists in the literature. Some, like the geometric Brownian motion, are more parsimonious in terms of the number of parameters to estimate. Other models have more parameters constraining the stochastic process. Each stochastic process has its own advantages and inconveniences and they should be chosen in function of the problem to be solved.

The fourth section introduces the state space representation of stochastic processes and the Kalman filter. The Kalman filter is a very powerful tool used for the simulation of stochastic processes as well as for parameter estimation. Once a stochastic process has been represented in its state space form, it is possible to estimate its transition equation using the Kalman filter. In this thesis, the Kalman filter is a central topic because of its flexibility. However, it can be difficult to calibrate the Kalman filter with historical observations because the optimization routine can yield to infeasible parameter estimates. For this reason, genetic algorithms were used for the calibration of the Kalman filter.

The next section introduces genetic algorithms. The genetic algorithm is a metaheuristic approach which is particularly suitable to solve non-linear

problems. When the objective (or cost) function of an optimization problem contains multiple dependent parameters and the landscape of this function is irregular with multiple local minimums, the genetic algorithm proves to be superior to the steepest descent approach for optimization. This is due to the fact that genetic algorithms work with a population of points to estimate instead of working on a single optimization point at a time. A tradeoff of using the genetic algorithm is computation efficiency.

The sixth section introduces pioneering work as well as recent work on option-pricing theory. With a calibrated stochastic process (using the Kalman filter and genetic algorithm workflow), it is possible to perform simulation and to estimate the value of derivatives contracts dependent on price path simulations of commodities, stocks or bonds. Then, an extension from the valuation option-pricing theory to Real Options valuation is presented. Real options are useful to determine the value of managerial flexibility in a capital budgeting projects. For example, the real option to close a mining project when the commodity price is unfavorable may add value over the alternative status quo (or doing nothing). A parallel can be made between real options and financial options. For example, closing a mine project can be seen as a financial put option: the assets are sold at a given price and the project is terminated. The last section presents how Real Options valuation can be applied to the valuation of mining projects.

2.2 Market risk regarding commodity markets

The mining industry can be separated into three principal spheres: exploration, mining and metallurgy (Svetlana, 2010). Each sphere is affected by mineral market fluctuations. For example, in a mineral boom, exploration companies are getting a lot of liquidity to rapidly find the next ore deposit in order to bring minerals on the market while mining companies may try

to acquire smaller mines or boost their already existing production. For this reason, valuing mining companies is difficult due to the impact of mineral and economical cycles (Svetlana, 2010). The first cyclical effect refers to the fact that the value of the commodity company is not only affected by the price of the commodity but also by the expected volatility in that price, leading to higher volatility in earnings than other sectors in the market. The second cyclical effect refers to the high fixed operating costs of mineral extraction projects; because of these high fixed costs, mining companies need to keep their operations active even in economic downturns. Another important aspect of mining companies is that ore deposits have a finite geometry meaning that if mining companies do not invest in exploration campaigns or acquire other mining properties, they are condemned to disappear (Svetlana, 2010).

Once the fact that mining companies are subject to some cycles has been established, several questions need to be answered. First of all, how are such types of companies going to be managed? The basic management problem can be stated very simply. The prices of commodity metals fluctuate by large amounts from one period to another, whereas costs are relatively more stable, resulting in large profit instability (Adams, 1991). Financial markets tend to give more valuation to companies with stable earnings since they are predictable. What tools can mining company's managers use to stabilize earnings throughout the life of a (or a portfolio of) mining project(s)? In his research, Adams (1991) denotes three principal management strategies to get through downward mining cycles: Capital repayment reduction, restructuring net operating costs and risk sharing. The first strategy is to reduce the fixed repayment schedule by offering metal loans (i.e. a gold loan) or bonds. The second strategy involves varying volumes of production and margins during mining cycles in order to stabilize the earnings; hence increasing stability at

the cost of profitability. The last strategy involves hedging potential losses by offering futures contracts to buyers of the produced commodity. In all cases, such strategies need to be implemented before downward cycles start.

Commodity prices have always been very volatile and hard to predict. Kroner et al. (1995) were among the first researchers to use econometric models to perform long horizon (up to 225 days) forecasts of volatility on commodity markets. To do so, they forecasted volatility using expectations derived from options prices (market expectation or presumed volatility forecasts), forecasts using time series modeling, and forecasts with a combination of market expectation and time-series methods. Kroner et al. (1995) showed that the workflow using a combination of implied and realized volatility gives the best results over the two other workflows. The reason why the combination method is better than the others is that options markets are inefficient. Econometric models were later used to forecast volatility in agricultural commodities such as corn, wheat and soybeans (Manfredo and Leuthold, 1998). According to the authors, value at risk (VaR) and GARCH forecasting appears to be a promising management tool for the agricultural industry, since it can quantify exposure to market risk of cultures and provide incentives to develop hedging strategies. Giot and Laurent (2003) used econometric models such as RiskMetric, skewed Student APARCH and skewed Student ARCH to predict value at risk of long and short positions in commodities markets. APARCH models can account for the tails asymmetrical distribution, allowing more weight on past negative returns than past positive returns. Also, since the authors use a Student distribution, they can account for fat tails, thus, he is able to better reproduce strongly negative returns. The authors showed that the three methods could be used to do very short-term predictions (1 day predictions). In most cases, RiskMetric failed to predict VaR 1 while skewed Student APARCH gave the

best predictions. However skewed Student ARCH is a good compromise since there is no need to perform non-linear optimizations in the procedure. It is not sure if econometric models bring an added value over forecasts of supply and demand for long-run predictions, but the methodology could definitely be implemented to perform midterm predictions (Giot and Laurent, 2003). Later, Sadorsky (2006) used econometric principles to forecast volatility in petroleum markets. The data used for his study were observations on the daily closing futures price returns on West Texas Intermediate crude oil, heating oil #2, unleaded gasoline and natural gas from the 1988-2003 period. The time period of the study makes the evaluation of econometric models very robust since oil prices were affected by large fluctuations during that period of time. The conclusion of this study was that a TGARCH model fits well for heating oil and natural gas volatility and a GARCH model fits well for crude oil and unleaded gasoline volatility. Then, non-parametric value at risk measures based on volatility forecasts of oil and gas markets can provide a useful way to measure risk exposure of an investment in oil and gas commodities (Sadorsky, 2006).

Since datasets can be contaminated with outliers, it is important to detect them and when possible, treat them with robust statistical methods. The next section introduces robust regression.

2.3 Robust regression

Before introducing stochastic processes, it is important to understand that any dataset can be contaminated with outliers. Outliers are different from bad data in the sense that they may be extreme values of a generation process or they can perhaps be generated from a secondary process (Ortiz et al., 2006). Several methods can be used to detect outliers. These methods are often based on distance measures or analysis of clustering. There is also

a distinction between the univariate and multivariate case. In the first case, there is no covariance term to calculate between the different variables. As a result, a simple method like the studentized residual method can be used to detect outliers (Thompson, 1935). In the multivariate case, it is often more useful to use a distance measure to detect outliers. For example, Filzmoser et al. (2005) used the Mahalanobis distance to assess which data points can be considered as outliers in a regression plot. This method is suitable for a multivariate case because it does not assume the distribution of the data.

Once the outliers are detected, they can be corrected using robust regression. Huber (1973) introduced the M-regression based on maximum likelihood. This estimator has the advantage to be very efficient. Rousseeuw (1984) introduced the least-median square (LMS) which is less affected by the presence of outliers. The LMS regression method is a particular case of quantile regression (Powell, 1986). A least-trimmed square estimator is useful when the dataset is contaminated with a substantial amount of outliers (Leroy and Rousseeuw, 1987). Yohai (1987) proposed a high breakdown point and highly efficient robust estimator called the MM-estimator. The MM-estimator has the highest breakout point possible of 0.5 meaning that up to 50% of the dataset can be contaminated by outliers before the robust estimator breaks.

Outliers will often appear in financial time series so it is important to understand how they can affect different estimates. Martin and Simin (2003) showed how outliers could affect estimates of betas used in the Capital Asset Pricing Model (CAPM). Authors propose an alternative approach than OLS to calculate betas. This complementary approach can also be coupled with the traditional calculation of beta to detect outliers. Robust methods can also be used to model operational risk in financial institutions and help to characterize how they can be affected by outlier events Chernobai and Rachev

(2006). The aim of using robust regression is not to discard extreme events but rather to partition a dataset in an outlier-free and outlier contaminated sets. This results in a much more meaningful analysis of time series since the estimates are not biased by a few but extremely influent outlier points. Robust statistics in finance can also be used to construct portfolios resistant to outliers (Welsch and Zhou, 2007). Results show that robust portfolios have a lower asset turnover and an increased risk adjusted return.

Some stochastic processes are better suited to deal with outliers. The following paragraphs, a review the different stochastic processes.

2.4 Stochastic processes

Samuelson (1965) introduced the first stochastic process models used to estimate returns in financial markets. The author used geometric Brownian motion (GBM) to describe how financial returns could evolve between two points in time using a drift and a volatility parameter. It was shown that in GBM diffusion models, financial returns are independent from previous returns (Samuelson, 1965). Stochastic interest rates tend to grow indefinitely when applying simple GBM to model them (Vasicek, 1977). Instead, an Ornstein-Uhlenbeck process is introduced in the diffusion process to describe the short-term rate dynamics.

Stochastic volatility (SV) methods were later developed to overcome the strong constant volatility assumption in previously developed diffusion model. Andersen and Shephard (2009) published an exhaustive literature review on stochastic volatility. They made a clear distinction between SV methods and ARCH based methods (often confused for SV methods). In stochastic volatility methods, the conditional likelihood function is not known as opposed to one-step-ahead predictions such as ARCH. In SV, the asset returns are instead approximated by a mixture distribution where the level of activity of new

arrivals is taken into account. The Heston SV (Heston, 1993) model links these biases within the Black-Scholes (BS) model to the dynamics of the spot price and the distribution of spot returns. The author shows that even if the BS model produces option prices similar to the stochastic volatility models for at-the-money options, SV models are better suited to explain skewness of option spreads by modeling the correlation between volatility and the spot price. Merton (1976) used a different approach and introduced an independent jump component to the GBM. As a result, the stochastic process is able to reproduce outliers that can't be modeled with the GBM. Later the work of Bates (1996) extended the Heston model and the Merton jump by combining both stochastic volatility and jump components. Bates (1996) shows how these jumps are important for modeling options prices simultaneously across the strike and time-to-maturity spectrum.

The convenience yield has to be taken into account when modeling diffusion processes governing different commodities. Unlike stock markets, commodity markets are accompanied with forward and future markets which can provide insights of future prices (Geman, 2005). Because of the high level of similarities between forward curves in commodities and yield curve in bond markets, theories developed in the fixed-income field in finance have been applied to model commodities (Geman, 2005). Schwartz (1997a) compares three models that take into account mean reversion. In the first model, a one-factor model following a mean-reverting process governs the diffusion process. The second model is a two-factor diffusion model taking the stochastic convenience yield of a commodity as a factor. The third model is a three-factor model that also includes randomly varying interest rates. Casassus and Collin Dufresne (2005) showed how the convenience yield could be modeled using the Schwartz three-factor model.

There is a fundamental difficulty in historical volatility forecasting: volatility is not directly observable (Danielsson, 2011). In other words, to detect volatility increases, one must first observe changes in price ranges to volatility changes to occur. For this reason, volatility is called a latent variable, which means there is a delay between the present time and volatility. However, there is an exception to this observation. For example, futures on volatility indices such as the VIX from the Chicago Board Option Exchange are based on the spread on future options, which means volatility forecasting is based on future values of implied volatility (Alexander, 2008). For this reason, such an approach is called implied volatility forecasting (Christensen and Prabhala, 1998). Day and Lewis (1988) observed that when approaching the expiration date for stock index options, volatility tends to increase significantly. Studies on implied volatility forecasting are based on the idea that option spread should reflect realized volatility given the fact that option markets are efficient.

The very first method used for historical volatility forecasting was the simple moving average (SMA). Volatility forecasts are based on returns on past days of trading (Danielsson, 2011). Since volatility tends to cluster, it is important to have a volatility forecast method which puts more weight on recent observations. The SMA volatility forecast tend to underestimate volatility because it applies equal weight to each trading day. For this reason, Morgan (1996) released a publicly available volatility forecasting model called the Exponential Weight Moving Average (EWMA), also called RiskMetric. The difference between SMA and EWMA is that the latter method applies an exponentially decreasing weight from the most recent days of trading to the preceding days. One can choose the decay factor of the EWMA calculation which is a number between 0 and 1. Morgan (1996) suggests to use 0.94 as

a decay factor for all types of investment securities, which is a subjective but suitable value of decay.

The volatility forecasting models discussed in the previous paragraphs were univariate. Extending the theory to take into account several assets at the same time is not always an easy task, since the covariance between different assets needs to be calculated (Danielsson, 2011). One of the firsts multivariate model developed was the BEKK (Engle and Kroner, 1995) where the matrix of conditional covariances is a function of the product of lagged returns and lagged conditional covariances, constrained by a parameter matrix. Another approach to model multivariate data is to assume that the observed data can be linearly transformed into a set of uncorrelated components using an inversion matrix. Parameters of the inversion matrix depend on conditional information which can be calculated directly from time series.

Another family of volatility forecasting methods is the Stochastic Volatility (SV) (Danielsson, 2011). This family of methods uses random number generation to generate equiprobable scenarios of volatility constrained by an input model. Also, SV models differ from ARCH based models where the conditional variance of returns is given by past returns observed, while in the SV approach the predictive distribution of returns is specified indirectly, via the structure of the model (Andersen and Shephard, 2009). In SV models, price changes are due to exogenous shocks to the stochastic distribution of returns. The stochastic features of time series have been observed for a long time in the literature. Mandelbrot (1963) describes a succession of returns with the random walk theory, where returns have a Brownian motion.

This family of forecasting methods was developed to overcome the strong normal distribution assumption in the BS option pricing model (Heston, 1993). The Heston SV model links these biases within the BS model to the dynamics

of the spot price and the distribution of spot returns. The author shows that even if the BS model produces option prices similar to the stochastic volatility models for at-the-money options, SV models are better suited to explain skewness of option spreads by modeling the correlation between volatility and the spot price. Later Bates (1996) extended the Heston model to allow for jumps in volatility. Bates (1996) shows how these jumps are important for modeling option prices simultaneously across the strike and time-to-maturity spectrum.

Brennan and Schwartz (1985) were the first to introduce a model that infers a relationship between commodity spot prices and futures. To link these two aspects, they introduced the convenience yield which is a continuously compounded benefit that the owner of a forward or futures contract renounce to when buying the contract. The greater the convenience yield is, the lower the fair value of the contract is. In Brennan and Schwartz (1985), the spot price moves according to a geometric Brownian motion (GBM) with deterministic convenience yield and risk-free interest rates. Later, Gibson and Schwartz (1990) used the the previous models and tested a two-factor model for pricing weekly oil futures contracts using spot price and the instantaneous convenience yield as variables.

A very useful algorithm for calibrating stochastic processes is the Kalman filter (KF). The next section introduces the KF and its diverse applications for solving engineering problems.

2.5 State space representation of stochastic processes

The KF was introduced in 1960 and have found numerous applications in solving engineering problems (Kalman et al., 1960). One of the first important uses of the KF was to implement the recursive algorithm responsible of tracking the Apollo spacecraft and its lunar module (Willems, 1978). The KF separates

a signal into two different partitions. The measurement partition contains measurement errors and can be used to estimate the transition equations, which describe the phenomenon to be modeled. For example, in the case of a radar measurement, the position of a vessel can be measured with a certain amount of noise, but the instantaneous velocity needs to be inferred from different subsequent measurements of position (Brookner, 1998). In this case, the transition equations will describe how velocity is related to position and the measurements will be used in the KF to estimate the correct velocity.

A wide range of engineering problems can be solved using the KF. Moreover, the KF can be used to solve engineering problems containing multiple independent variables or in multiple dimensions. For example, the KF can be used to track objects in 3 dimensions using the Global Positioning System (GPS) (Sasiadek et al., 2000). More recently, the KF have been widely used to track Unmanned Aerial Vehicles (UAV), also known as drones (Benini et al., 2013). The previous examples showed how the KF could be used for tracking objects in multiple dimensions. Another important use of the KF is for the interpretation of sensor measurements used in automated processes. An extension of the KF, the Ensemble Kalman filter (EnKF), can be used to solve engineering problems related to geology and geography (Evensen, 2003). Bendorf et al. (2014) used the EnKF for the real-time online characterization of properties of coal moved on conveyor belts. The workflow enables the reconciliation of the geostatistical model and actual measurements of the material properties. The extended Kalman filter (eKF) was introduced to solve problems where the transition equations are non-linear (McElhoe, 1966). The method uses Taylor series expansions to linearize the problem. Other approaches can be used to describe non-linear problems. The particle filter (PF) was introduced to overcome the linearity assumptions of the KF (Del Moral, 1996). The

PF is more accurate than the KF when the problem is non-linear but since it relies on the estimation of discrete particles instead of a normal distribution assumption of the error term, it is more computationally intensive.

The KF and its derivatives methods can be used to infer hidden state variables from commodity futures contracts observations. Schwartz (1997a) introduced the state space representation of stochastic processes in the context of commodity futures estimation. The author compared three models that took into account mean reversion in the convenience yield. In the first model, a one-factor model following a mean-reverting process governs the diffusion process. The second model is the Gibson and Schwartz (1990) model. The third model is a three-factor model that also includes randomly varying interest rates. In order to correctly estimate parameters, Schwartz (1997a) uses KF techniques to continuously and iteratively correct the covariance matrix. Ribeiro and Hodges (2004) further extended the long-term Schwartz two-factor model by replacing the Ornstein-Uhlenbeck by a Cox-Ingersoll-Ross to model the convenience yield. This ensures that the model is arbitrage free. Furthermore, they examined the spot price volatility as proportional to the square root of the convenience yield level.

Both Ribeiro and Hodges (2004) and Schwartz (1997a) models are adept at pricing short-term contracts, but fail when pricing longer term structures. Schwartz and Smith (2000a) eliminated the concept of convenience yield and modeled the log of the spot price as the sum of a short and a long-term component. The model allows mean reversion in the short-term and uncertainty in the long-term price. Movements in prices for long maturity futures contracts describe the long-term equilibrium level and the difference between short-term contracts and long-term describes how spot prices mean revert in

the short-term. This specification is better at pricing longer term futures contracts than previous models. Aiube and Samanez (2014) reconducted analysis in oil futures markets using longer maturity contracts that were not available in previous studies. Authors showed that the use of these multifactor models tend to underestimate the risk premium observed in futures prices.

The KF is very useful to estimate commodity markets dynamics. However, calibrating the parameters can be challenging. In this thesis, the use of the genetic algorithm for the calibration of the Schwartz-Smith model (Schwartz and Smith, 2000a) is investigated. The next section presents a brief literature survey on genetic algorithms.

2.6 Genetic algorithms

Genetic algorithms (GAs) appeared in the mid 70's with the pioneering work of Holland (1975). In the first formulation of the GA, a set of parameters to be optimized were encoded using a binary representation. GAs can ignore the gradient of the optimized function since they rely on an objective function to evaluate the transition from a solution to another. This enables the algorithm to consider a population of points instead of focusing on the gradient optimization of a single point which can be stuck in a local maximum (or minimum) (Goldberg and Holland, 1988). This makes the GA particularly suitable to optimize highly non-linear functions. As a result, GAs are employed in a wide range of engineering optimization problems.

Some researchers have used the GAs to literally simulate the evolution of virtual creatures (Sims, 1994). The authors have created a 3D virtual environment subject to physical constraints and used the GAs to simulate the evolution of randomly generated creatures. The objective function represented tasks the virtual creatures were required to perform such as jumping, swimming or walking. The GAs helped to design complex morphologies that were

optimal for performing the tasks. Later, GAs were used for the optimization of the shape of the wing of an aircraft (Obayashi et al., 2000). In this case, the shape of the wing could be modified iteratively to have better aerodynamic characteristics.

In the previous examples, GAs were used to optimize the shape of computer generated objects. GAs are also suitable for solving abstract engineering problems. Mahfoud and Mani (1996) used the GA to forecast individual financial stock performance. The authors used the GAs to model how the combination of 15 attributes can be used to forecast favorable market circumstances. Later, Shin and Lee (2002) used the GA to predict bankruptcies of Korean financial institutions. The GA can have multiple objectives to optimize at the same time. Tapia and Coello (2007) denoted five types of problems that are solved using multi-objective GA optimization. The five types of application are: investment portfolio optimization, financial time series modeling, stock ranking and screening, risk-Return analysis and economic modeling.

Other algorithms use behavior observed in nature to solve problems. Eberhart and Kennedy (1995) introduced particle swarm optimization (PSO). This algorithm mimics how a herd of individuals cooperate to better survive. Each individual is allowed to move independently at a given speed. The swarm collectively knows which solution is the best, but individual member of the swarm move randomly to try to find a better solution. Maniezzo (1992) introduced ant colony optimization (ACO), which mimics the behavior of an ant colony which tries to optimize the route between their colony and a source of food.

Any of the presented metaheuristics approaches can be used to optimize the KF parameters. Ting et al. (2014) used the GAs to optimize the parameter of the KF in a battery management system. Authors show the methodology is suitable because it produces low root mean-square (RMS) error values. Ortmanns et al. (2015) used a combination of GA and unscented Kalman filter (UKF) for the estimation of parameters used in a sigma-delta modulator analog-to-digital converter. Luis Enrique Coronado (2015) used a combination of extended KF and GAs for the optimization of neuro-mechanical parameters responsible to prevent human from falling. The workflow is better suited than actual trial and error methods for finding a solution to the problem.

The next section introduces options pricing theory. This will be useful to better understand real options, a central concept in this thesis.

2.7 Options pricing theory in financial markets

An extensive literature survey exists in the valuation of options traded in financial markets. Black and Scholes (1973) developed a widely used formula to price European-style options. The BS formula takes a stock price, the option exercise price and its time to expiration, the risk-free interest rate and the volatility as an input and computes the price an investor should pay for that option. In the BS formula, stock price returns obey a geometric Brownian motion, and they follow a log normal distribution. Moreover, in BS formulation, the volatility is constant. It is well known that these assumptions are not true because volatility tends to be clustered in periods of high and low volatility (Fama, 1965). If a stock price can be modeled as a GBM affected by shocks such as important news or political decisions, volatility should also be affected by these shocks (Fama, 1965).

Later, Merton (1976) improved the BS model by allowing for stochastic jumps in the stock price returns while considering the volatility constant over

time. As a result, a premium has to be added to options where the stock prices are exhibiting strong changes in magnitude. Schwartz (1977) solved partial differential equations in the BS and Merton models to take account for dividend-paying stocks. A numerical example based on the pricing of American Telephone and Telegraph (ATT ticker) options comparing Schwartz model with BS without dividends and BS with dividends, show a discrepancy between real market prices of options and Schwartz model. According to Schwartz (1977), the discrepancy could be due to a bias in the model, a bias in the market value or a discrepancy in the historical variance derived to price the option using the model and the variance predominant in the market valued options. Cox et al. (1979) developed a numerical pricing option based on a binomial lattice. Their discrete-time approach uses simple mathematics and the no-arbitrage assumption in derivative markets. The binomial lattice method converges to the BS formulation when discrete time steps converge to zero. For this reason, the binomial lattice algorithm can be seen as a numerical implementation of the BS model.

Boyle (1977) later introduced a Monte-Carlo simulation approach in options pricing theory. The approach uses computational power to generate a number of realizations of stock price return paths and has an advantage over analytic methods in cases where jump processes are present in the underlying stock price returns. Longstaff and Schwartz (2001) developed options valuation with a Monte Carlo-based approach using least square regression to estimate the expected payoff of an option contract holder. The model is called the least square Monte-Carlo (LSMC) regression. Their method can calculate American type options, which is a significant advantage over methods that can only calculate European type options. Longstaff and Schwartz (2001) recognized that using methods such as generalized methods of moments (GMM)

when estimating the conditional expectation function could lead to better results when residuals from the regression are heteroscedastic. Jonen (2011) improved the LSMC algorithm for when the conditional expectation function contains outliers by replacing the least-squares estimator with a robust estimator. Jonen (2011) showed that the robust Monte-Carlo regression algorithm significantly reduces the bias without increasing the variance when using robust loss functions such as Huber (Huber, 1973) in presence of outliers. Even if Monte-Carlo regression based techniques can tackle more complex problems by allowing for multiple simultaneous assets, stochastic volatility cannot be easily implemented because it would require solving a non-linear regression at each time step. In any case, analytical methods presented in this section are based on the assumption that volatility is constant over time, or non-stochastic.

The next section introduces real options, which is an extension of options pricing theory to the pricing of options on real assets.

2.8 Introduction to Real Options valuation

Because of the high level of similarities between management of a mine and the management of options traded in financial markets, Real Options valuation was first introduced by Myers (1977) who made an analogy between an irreversible investment and a call option (or option to buy) in financial markets. An American call option can be exercised at any time before the expiration date. As an analogy, the possibility to undergo an irreversible investment such as opening a mine can happen whenever after the completion of the feasibility study whenever the financial resources are available. At the project expiration date, financial actors may decide to invest their money elsewhere and the option to undergo the mining project expires. The added optionality using options pricing theory is referred in the literature as managerial flexibility (Trigeorgis, 1993). In natural resource investment projects, management

teams have the possibility to defer or accelerate development, alter operating scale and abandon or temporarily close a mine (Trigeorgis, 1993).

Because Real Options valuation theory is based on an analogy with options traded in financial markets, practitioners decided to adapt option-pricing theory in a Real Options framework. Copeland et al. (2001) made a comprehensive book on applying Real Options valuation to real world problems. The authors are principally using decision tree analysis (DCA) with an approach similar to Cox et al. (1979) to value real financial assets as opposed to securities. Luehrman (1998) considered characteristics of a real world investment problem as input variables into the BS model. For instance, present value of a project operating assets can be seen as the stock price, expenditures required to acquire the project assets are the call option exercise price, the period the decision can be postponed is the time to expiration, the time value of money corresponds to the risk-free interest rate, mean-reverting and the risk of the project corresponds to the variance of returns on stocks.

In the recent literature, several attempts have been made to apply Real Options to the valuation of mining projects. Even if the approach seems promising compared to the DCF approach, practitioners still have a lot of difficulties applying ROV to mining projects. One of the reasons is that ROV models are oversimplified and do not reflect the true complexity of a mining project investment (Dimitrakopoulos and Abdel Sabour, 2007). According to the authors, one of the major simplifications that render ROV impracticable is that production rates in a mine are considered to be constant. To overcome this difficulty, Dimitrakopoulos and Abdel Sabour (2007) focus on implementing a workflow integrating mine production sequence in the ROV framework. Another problem that is not currently well addressed in ROV

for mining projects is the variability of foreign exchange rates (Dimitrakopoulos and Abdel Sabour, 2007). The authors address the problem by modeling foreign-exchange rates as a mean reverting process taking into account the correlation with other stochastic variables. However, the authors did not explain exactly how foreign exchange rates were simulated.

Brennan and Schwartz (1985) explained how classical DCF valuation techniques were not appropriate to the valuation of resource projects. The stochastic aspect of commodity prices and the cyclical aspect of mining industry require developing models specific to the resource extraction industry since commodity prices can encounter price swings of 25%-40% per year. Samis et al. (2005) compared conventional DCF analysis to ROV. They observed that ROV valuation tends to add value to a project because of the added managerial flexibility. The authors emphasize that ROV is not a hidden value detection tool, but rather a planning tool. In other words, one cannot predict the future of commodity prices, but can plan strategies if different scenarios occur. Authors used a binomial lattice method such as in the work of Cox et al. (1979) to model the discrete time value of Real Options. Bertisen and Davis (2008) produced an econometric study that showed production costs are systematically undervalued in mining projects. According to the authors, this is due to the fact that feasibility studies are realized by third-party firms hired by mining companies. Another reason is the scarcity of financing resources compared to the quantity of mining projects. Finally, authors showed that the effect of inflation is not negligible when projects take a longer time to start.

Because of the known limitations of DCF analysis, ROV is increasingly used in the industry. In fact, it is accepted in NI43-101 reports (Samis et al., 2011). The difficulty to implement ROV in feasibility studies is the need of

having a qualified person (QP) to perform the analysis. Such QP is very rare. Nevertheless, Samis et al. (2011) presented a NI43-101 report comparing the valuation of the Oyu Tolgoi project operated by Ivanhoe and Entree Gold using ROV and DCF. Authors showed that ROV can better integrate risk related to commodity prices even in the early stages of the feasibility studies. Finally, Samis and Davis (2014) showed how ROV valuation could be used to plan a hedging program for a mining project. Authors use a MC discounted cash flow method to analyze financing proposed for an African gold project. Authors showed the risk mitigation plan proposed by the financial institution was transferring too much wealth from the mining company to the financial institution. The authors argued that even if mining management teams are not financial engineers, using ROV to evaluate the risk mitigation program offered from a financial institution can help to grasp what are the trade-offs of using such hedging plans.

In the next chapters are presented the papers accompanying this thesis.

CHAPTER 3

Analysis of mining engineering data using robust estimators in the presence of outliers

3.1 Abstract

Ordinary least squares (OLS) regression is an estimation technique widely used in mining research to model relationship among a dependent variable and a set of explanatory variables. However, OLS is based on strong assumption such as normality of the error term, exogeneity of explanatory variables, linearity between regression coefficients, and homoscedasticity. When one or more of the assumption are not held, OLS can lead to biased and inefficient estimates of the true population parameters. Using biased estimates of the population parameters may result in serious problems in mining decision making, especially if key decisions are based on the biased parameter estimates. One of the main reasons that can lead OLS to provide biased estimates is the presence of outlier data points in the sampled population. In this paper, we focus on alternatives to the OLS estimator, which are more robust (or resistant) to the presence of outliers. Two case studies were conducted to compare OLS and robust regression approaches namely L_1 -estimation, M-estimation, least trimmed squares, least median squares, and MM-estimation. The case studies showed that inference based on OLS in the presence of outliers could lead to bad decisions if regression coefficients are not interpreted correctly. Robust regression approach can provide estimates useful even if a dataset is contaminated with outliers.

3.2 Introduction

In mining engineering research, regression analysis based on the ordinary least squares (OLS) approach is widely used to model relationships between variables. For example, O’Hara (1980) established regression models relating equipment sizing, daily ore-waste production, and cost for open pit mines. Taylor (1986) introduced a relationship between production rate and reserves. In sampling broken ores, K , which is a constant factor being the product of the mineralogical factor, the shape factor, the granulometric factor, and the liberation factor, and α , which is the cube of the nominal top-size of the fragments in the sample, are predicted by linear regression modeling (Francois-Bongarcon and Gy, 2002; Minnitt et al., 2007). Karacan (2008) researched the relative effects of different coal bed parameters on the migration of methane by multiple regression analysis. Jorjani et al. (2009) predicted the combustible value and combustible recovery of coal flotation concentrate by regression and artificial neural network based on proximate and group macerals analysis. Demirel (2011) fitted an equation relating geological strength index and average hourly stripping amount. Gray (1996) established a regression equation relating conductivity to sulfate concentration in both acid mine drainage and contaminated surface waters. Rajesh Kumar et al. (2013) used regression analysis to estimate rock characteristics from sound levels generated throughout drilling.

A linear model estimates the relationship between a dependent variable and a set of independent observations. The multiple regression analysis can be represented as

$$y_i = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} + \varepsilon_i \quad (3.1)$$

In equation 3.1, y_i is the predicted variable, x_{1i}, \dots, x_{ni} are the n explanatory variables, β_0, \dots, β_n are the n coefficients that can be adjusted to

minimize the difference between the predicted variable and explanatory variables and ε_i is the error term. Since observations are taken from a sample of a population, the coefficients will always be an approximation of the true population coefficients. To isolate the error term from the predicted and explanatory variables, Equation 1 can be re-written

$$\varepsilon_i = y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni}) \quad (3.2)$$

The optimal OLS coefficients to minimize the difference between observed and explanatory variables are found by minimizing the sum of squared difference between measured value of the predicted variable y_i and its OLS estimate:

$$\text{minimize } \sum_{i=1}^k (y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni}))^2 \quad (3.3)$$

One should be very cautious when deciding which explanatory variable to include inside the regression model. A good model is based on a natural process. There are five critical assumptions behind OLS regression (Greene, 2003). First of all, as expressed in Equation 3.2, there must be a linear relationship (w.r.t. β) between predicted variable and all explanatory variables. If this assumption does not hold, it may be still possible to find an optimal solution using non-linear regression, but this approach would not be OLS. The second assumption is that all explanatory variables must be independent, ensuring that there is a unique solution to the minimization problem. The vectorial space spanned by the different explanatory variables can be seen as a projection space where the estimates of the model can be generated. If two or more columns are linearly dependent, there is no unique solution to the minimization problem. The third assumption is that all explanatory variables must not be correlated with the error term (exogeneity of the error term). The

fourth assumption is that variance has to be the same for all disturbances (homoscedasticity) and there must be no correlation between error terms. When variance is non-stationary (or heteroskedastic), there are workarounds like using the General Regression Model, but again, it's a different approach than OLS. Finally, OLS requires the error term to be normally distributed. Failing to meet the five assumptions will often lead to biased or inefficient estimates when performing OLS regression.

From our observations, in many engineering applications, all the assumptions of the OLS are rarely met. Outliers can either be extreme values of a distribution or values generated from a secondary process that does not taken into account in the regression model (Filzmoser et al., 2005). For example, residuals often exhibit an increasing spread in time, which means that there is a violation of the homoscedasticity assumption. Since OLS weights each observation equally, the presence of heteroskedasticity can affect coefficients estimates by giving too much weight to observations that shows a higher level of variance. On a cross plot of predicted variable in function of explanatory variables in presence of heteroskedasticity, high variance points will appear as outliers (Greene, 2003). Another example of outlier generation is the short regression (Greene, 2003). When an explanatory variable in the multivariate regression equation (or the model) is neglected to be included, the variation of the relevant explanatory variable must still be taken into account and it will be implicitly included in the error term. As a result, residuals will exhibit a higher level of variance. In mining research, collecting samples can be very expensive. This often causes engineers to work with sparse datasets. If there is an outlier among sparse sample, using OLS can lead to very large residuals. Using OLS in presence of outlier data in mining may result in serious financial, safety or environmental consequences and losses in mine design and valuation

processes. As such, robust (resistant) techniques are required as an alternative to the OLS method (Leroy and Rousseeuw, 1987).

In OLS, each element of vector parameter estimates is seen as a weighted average of the vector of independent variables (Davidson and MacKinnon, 2004). Therefore, outlier points can have high influence on OLS regression. OLS error term can be affected by the presence of a point (x, y) which is significantly apart from the rest on other y values (“outlier point”), other x values (“leverage point”) or on both y and x values simultaneously (“outlier–leverage point”) (Ortiz et al., 2006). Outlier-leverage points will have the most dramatic effect on OLS regression since it will have a tendency to “pull” on the regression line and inflate the error term at the same time. Robust regression techniques have been designed to be resistant against outliers, leverage or outlier-leverage points at the same time.

In this paper, the focus will be on alternatives to the OLS estimator, which are more robust (or resistant) to the presence of outliers when modeling cross-sectional data. Main contribution of this paper is to introduce robust regression techniques to mining practitioners, engineer, and decision makers. Given that the regression based on least squares fitting in the presence of outlier can lead to destructive consequences, robust regression techniques can generate more reasonable fittings. In addition to this, the paper provides R codes of various robust regression techniques. Thus, a practitioner can directly use it. We will first discuss the definition and detection of outlier points. Then, alternatives to the OLS estimation will be proposed, namely least absolute deviation or L_1 -estimator, M-estimator, LTS-estimator, LMS-estimator and LTS-estimator. Then, two case studies consisting of cross-sectional data estimated with different robust techniques will be presented.

3.3 Outliers detection

In this section, the available methods to detect outliers in datasets will be presented. First of all, it is important to distinguish between cross-sectional data and time series. In a mining context, cross-sections correspond to datasets that do not vary in time. For example, a relationship between copper and arsenic concentration of ore can be viewed as cross-sectional data since for each pair of observed points values cannot change over time. On the other hand, time series are highly correlated in time and violate the fourth assumption (correlation in time) of the OLS. In both cross-sectional and time series, outliers often arise for similar reasons, and the failure to account outliers in estimation process may lead to biased estimates.

Detecting outliers in cross-sectional data is a qualitative task. Outliers are different from bad data in the sense that they may be extreme values of a generation process or they can be generated from a secondary process (Ortiz et al., 2006). A high quantity of outlier points in a dataset is often a warning that the model is not well defined. Bad data are generated by an erroneous input, which can be due to an error with the manipulation of a measure instrument, or errors during the transcription of the result. Detecting outliers implies that a critical analysis of the model must be done while eliminating bad data points is generally done directly during the process of quality assessment of samples.

A common method to detect outliers is to evaluate extreme percentiles for a given distribution. In the univariate case, this can be achieved by fitting a distribution to a set of observations represented in a histogram. The second step is to calculate the cumulative distribution function (CDF), and determine a percentile threshold to reject a sample. When one or more suspect data points have been generated by a different population than the observed data,

studentized residuals approach can be used to detect outliers (Greene, 2003). The approach consists of calculating the regression coefficients and variance omitting each observation subsequently, and then standardizing the modified residuals:

$$e(i) = \frac{\frac{e_i}{(1-h_{ii})}}{\sqrt{\frac{e'e - e_i^2/(1-h_{ii})}{n-1-K}}} \quad (3.4)$$

Where e_i is the residual for the regression with i outlier point dropped, e is residual for the whole sample, and h_{ij} is the influence measure according to (Belsey et al., 1980). Studentized residuals should have a t distribution with $n - 1 - K$ degrees of freedom, and observations that have studentized residuals greater than 2 in absolute value should be considered as outliers (Greene, 2003).

For the multivariate case, one of multiple approaches is to use a measure of distance like the Mahalanobis distance (MD) (Filzmoser et al., 2005):

$$MD_i = \sqrt{(x_i - t)^T C^{-1} (x_i - t)} \quad (3.5)$$

where x_i is one observation, t is the estimated multivariate location, and C is the estimated covariance matrix. It is common to use as t the centroid of a multivariate distribution. The MD follows a Chi-squared distribution of p degrees of liberty, where p is the number of variables. Given it is possible to estimate the covariance matrix, we can find the resulting Chi-squared distribution and determine a threshold for outliers in order to be considered as outliers for the given Chi-squared distribution.

3.4 Robust regression

There are two main approaches to reduce the effect of outliers. The first approach diagnoses outliers and then corrects, removes, or downsizes them.

OLS estimation is then used on the cleaned set of data. When outliers are not caused by errors in the sampling method but rather generated by rare occurring events, it may be preferable to take them into account in the model. To do so, it becomes necessary to use robust estimators, which are not affected by the presence of outliers, rather than using OLS estimators. This approach refers to the inclusion approach where outliers are kept but robust estimators replace OLS estimator.

There are two measures for robustness (Ortiz et al., 2006):

- Breakdown point measures how good an estimate can resist outlier data before it fails. The finite sample breakdown point of an estimator at a sample set is the smallest fraction of contamination. Only one outlier is adequate to carry the OLS over all bound. For a sample of size n , the breakdown point is $1/n$. That is, it is apt to 0% for large number of sample. The maximum value of breakout point that can be obtained is 50%.
- Influence function is a local measure of how much a single outlier affects the estimate. It has local characteristics and defines the effect of an infinitesimal contamination in the sample on the estimator. The methods are classified in the basis of whether the influence function is bounded on the x and/or y axes, depending on leverage and/or influence of outliers.

Not one robust estimator is universally better than any other in terms of breakdown point, efficiency, or influence function. It is often preferable to use more than one robust estimator in the presence of outliers in a sample, simply because there is always a tradeoff between breakdown point measure, efficiency, and influence. Efficiency refers to the tendency of an estimator to give the best unbiased estimate for a particular set of observations (Greene, 2003). For each robust estimator presented in this paper, the breakdown point

and influence will be discussed qualitatively so the reader can decide which estimator to use depending on how outliers are bound to a different axis.

The next paragraphs present the robust estimators, which are a good alternative to OLS. First, L_1 -estimation is introduced. Second, M-estimation is presented, followed by quantile regression estimators such as least trimmed squares (LTS) and LMS. Finally, S-estimator is introduced as a first step of MM-estimation.

3.4.1 The least absolute regression model (L_1 estimator)

We introduce the least absolute errors regression (L_1), which minimizes the absolute value of the error term instead of its square:

$$\text{minimize } \sum_{i=1}^k |y_i - (\beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni})| \quad (3.6)$$

Since the error term is not squared in Equation 3.6, the effect of outliers will be diminished. However, large residuals will be still problem. Using L_1 regression may seem simple at first, but it requires using iterative algorithms to solve (Barrodale and Roberts, 1973). Moreover, to use L1 regression over OLS, the statistician needs to make an assumption on the distribution of the error term (Dasgupta and Mishra, 2004). In order to be efficient, the regression estimator needs to converge to the maximum likelihood estimator (MLE), which is the value that is most likely to result from the population that is modeled. However, depending on the distribution of the errors, either the OLS or the L1 can be closer to the MLE. When the distribution of errors is normal, the OLS is the best choice. But when the distribution of errors is Laplacian, or double exponential, the best choice is the L1 regression Hunt et al. (1974). If the assumption that the error term are Laplacian, it is possible to solve L1 regression with maximum likelihood method. However, it is important to be careful when using MLE with a small data set since it can be affected by local

maximums during the optimization (Greene 2003). MLE is discussed in the next section. The breakdown point of the L1 regression is 0% (Ortiz et al., 2006).

3.4.2 M-estimator

Another class of estimator is the M-estimation. Maximum likelihood-type regression stands for the class of regression methods, which aims to minimize or maximize a sum of functions representing each observation (Huber, 1973). Huber M-estimator is solved by the minimization of a function

$$f(y_1, \dots, y_n | \theta) = \prod_{i=1}^n f(y_i | \theta) = L(\theta | y) \quad (3.7)$$

If we want to find the maximum value of the pdf, which corresponds to the most likely value of the sample mean, we need to derive Equation 3.7 w.r.t θ and find the zero. However, it may be difficult to derive Equation 3.7 since it will be necessary to use derivation chain rule for the whole set of functions representing each observation. It is more convenient to use the logarithm of $L(\theta | y)$:

$$\text{minimize } \sum_{i=1}^k \rho(r_i) \quad (3.8)$$

where q is the Huber q function, a symmetric positive-definite function with a unique minimum at zero. Huber q function is less increasing than square residuals in OLS regression and is defined as (Huber, 1973)

$$\rho(x) = \left\{ \begin{array}{ll} \frac{1}{2}x^2 & \text{for } |x| < c \\ c|x| - \frac{1}{2}c^2 & \text{for } |x| > c \end{array} \right\} \quad (3.9)$$

The c tuning constant may be adjusted to control asymptotic efficiency of the M-estimator, q is symmetric and continuously differentiable, and $q(0) = 0$.

There exists $c > 0$ such that q is strictly increasing on $[0, c]$ and constant on $[c, \infty]$.

To solve Equation 3.8, an iterative re-weighted algorithm may be used to subsequently optimize weight of each observation. Huber M-estimation has the advantage of being more efficient than L1-estimation but has a breakdown point of $1/n$ (Anderson and Schumacker, 2003).

3.4.3 LTS estimator

Least trimmed squares (LTS-estimator) are obtained by minimizing the h quantile of squared residuals (Ortiz et al., 2006; Leroy and Rousseeuw, 1987):

$$\min_{a,b} \left\{ \sum_{i=1}^k (y_i - (a + bx_i))^2 \right\}_{i:n} = \min_{a,b} \left\{ \sum_{i=1}^k r^2 \right\}_{i:n} \quad (3.10)$$

where $(r^2)_{1:n} \leq (r^2)_{2:n} \leq \dots \leq (r^2)_{n:n}$ are the ordered squared residuals.

LTS is similar to OLS but largest squared residuals do not appear in the distribution. For instance, LMS regression, which is presented in next section is a particular case of LTS estimator, but where the quantile to trim is 50%. When index k of the sum in the objective functions attains a value of approximately $n/2$, breakout point of LTS estimation is 50% (Ortiz et al., 2006).

3.4.4 LMS estimator

Instead of minimizing the absolute value of errors, a method was introduced to minimize the median of errors (Rousseeuw, 1984): A simple way to compare LMS and OLS is that OLS will fit a line that minimize the sum of distances between residuals and fitted line, while LMS tries to find the narrowest strip covering half of the points. This method has a very high breakdown point ($\epsilon^* = 50\%$) but has a very low efficiency ($n^{1/3}$ convergence). The value of the breakdown point is not a valid condition to evaluate if a regression method is good or not. It can instead be viewed as an necessary condition (Rousseeuw, 1984). In the LMS, even if 50% of the data is outlier at maximum,

the accuracy of curve fitting will not be affected because median residual is only considered. Therefore, the breakdown point of the LMS is 50%. It is to be noted that LMS-estimation is practically never used because of its low efficiency (Anderson and Schumacker, 2003). However, it is presented in the case studies in the next section as a comparison to OLS and other robust estimators.

3.4.5 MM-estimator

A class of high breakdown point and high efficiency estimators has been introduced by Yohai (1987). These estimators are computed in three stages. The first stage is to compute a high breakdown point but not necessarily efficient estimate of the regression. In doing so, S-estimation can be used to give the scale estimate (or estimate of the spread of the dependent variable) (Anderson and Schumacker, 2003). The scale estimate is the value of s , which is the solution of

$$\frac{1}{n} \sum_{i=1}^n \rho\left(\frac{u_i}{s}\right) = b \quad (3.11)$$

where n is the size of the sample, $u_i = y_i - x'_i\beta$ or the error term in vectorial annotation, ρ is a minimization function, and b may be defined by $E_\varphi[\rho(u)] = b$, where $E[\cdot]$ stands for expectation function and φ stands for standard normal distribution. The second stage uses the scaled estimates of the first S-estimation as an input in an M-estimation to find a close-by but more efficient solution (Anderson and Schumacker, 2003). The third step is to find the final MM-estimates (Yohai, 1987). Let $\rho_1(u)$ another minimizing function such that

$$\rho_1(u) \leq \rho(u) \quad (3.12)$$

Let $\psi_1 = \rho'_1$ then, the MM-estimate is defined as any solution of

$$\sum_{i=1}^n \psi_1 \left(\frac{r_i(\theta)}{s_n} \right) x_i \quad (3.13)$$

which verifies

$$\sum_{i=1}^n \rho_1 \left(\frac{r_i(\theta)}{s_n} \right) \leq \sum_{i=1}^n \rho \left(\frac{r_i(\theta)}{s_n} \right) \quad (3.14)$$

Assuming that the error term is normal, MM regression will both be efficient and consistent. With the addition of a high breakdown point, MM-estimator is truly a good choice when dataset is in the presence of outliers.

3.5 Case Studies

To demonstrate the difference between linear regression based on OLS and robust regression estimators, two case studies are provided. The first dataset consists in pairs of processing costs and recovery values for ore samples sent at the smelter. The second dataset consists in iron and alumina concentrations in ore samples.

3.5.1 Mineral processing costs recovery relation

The original mineral processing costs and recovery dataset contains fourteen observations obtained by a series laboratory tests. OLS regression was performed on processing costs and recovery dataset, using built-in *lm* function in R programming language. The R code used to perform regression as well as figures presented in this article is in the appendix section. Since the two case studies presented use the same code, only mineral processing costs and recovery relation R code is presented in appendix. Regression was performed on the original as well as on a cleaned version of the dataset, where the outlier point was manually removed. Parameters of the OLS regression are summarized in Table 1. Recovery coefficient is positive; meaning that aiming to obtain higher percentages of recovery will lead to an increase in processing costs. Moreover, p-value is 1.76×10^{-6} meaning the coefficient is significantly different from 0,

even if the sample contains very few data points. As shown in Figure 3-1, observation number 14 looks far from other observation in the cross-plot of processing costs in function of recovery. Using the studentized residual approach, observation 14 is flagged because its calculated (with Equation 3.4) value is greater than 2. When a dataset is very scarce and contains few outlier points, the studentized residuals approach can be used to detect outlier data points that may have been generated by a different process. If there are reasons that the outlier point is caused by an error in the sampling process, it can be removed and the OLS can be performed on the cleaned dataset. In this case, the recovery coefficient decreases from 44.726 to 39.599 when observation 14 is removed. An increase from 0.861 to 0.925 can also be observed when OLS is performed with outlier point removed. However, removing points is not always a good practice since they can contain relevant information. In the processing costs and recovery case study, it may be that given a certain recovery threshold, processing costs are increasing dramatically. In this situation, removing the outlier point may lead to an underestimation of processing costs for high percentage of recovery. In the next section robust regression techniques, which can resist to the presence of outliers will be presented.

Table 3–1: Summary of the OLS Performed on the Original Processing Costs and Recovery Dataset as Well as on the Cleaned Dataset Where Observation 14 was Dropped.

	OLS	Cleaned OLS
(Intercept)	-6.417 (3.63)	-3.994 (2.35)
Recovery	44.276 -5.143	39.599 -3.397
p-values	1.76×10^{-6}	1.57×10^{-7}
R ²	0.861	0.925
Adjusted R ²	0.849	0.918
N	14	13
Standard errors in parentheses		

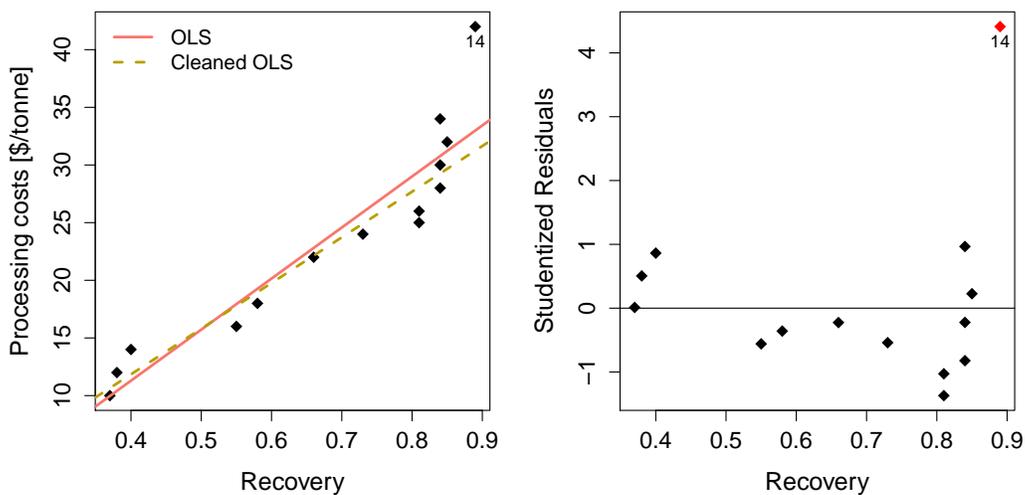


Figure 3–1: On the left, processing costs as function of recovery. A line fitted with OLS estimation is superimposed. On the right, corresponding studentized residuals of the OLS regression.

In Figure 3–2 are presented different robust regression fit. As expected, OLS is the most affected regression method in presence of outliers while LMS is the least affected because it has a breakout point of 50%. Other methods coefficients fall between OLS and LMS. L1 regression was performed in R using *quantreg* package, using a tau value of 0.5 (Koenker, 2013). LMS and M-regression was performed using *MASS* package in R using iterated re-weighted least squares (Venables and Ripley, 2002). Huber’s weight function was used in M-estimation because it corresponds to a convex optimization problem and gives a unique solution. LTS-estimation and MM-estimation were performed using *robustbase* R package (Todorov and Filzmoser, 2009). Coefficients calculated by each regression methods are presented in Table 3–3. The use of one or another estimate coefficient in mine planning depends on the degree of risk managers are willing to accept. For example, OLS overestimates processing costs for recovery values ranging from 0.4 to 0.8 while it underestimates values for recovery greater than 0.8. If a mine manager wants to use a conservative cost coefficient and aim mid-level recovery values, OLS will give the most conservative estimates. However, these estimates may be too conservative and robust estimates may be a more suitable choice. For this particular example, it seems there is a structural break for values over 80% recovery, meaning processing costs and recovery pairs are generated by a different population. This can be due to a change in processing methods to obtain higher percentage of recovery. If a mining project manager plans to obtain higher recovery percentages, a better practice would be to split the dataset in two classes and fit a regression line for each class to take into account the observed structural break.

Table 3-3: Summary of OLS-, L1-, M-, LTS-, LMS-, and MM-Estimation Intercept and Coefficient

	OLS	L1	M	LTS	LMS	MM
Intercept	-6.417	-5.310	-4.831	-2.864	1.130	-3.687
Coefficient	44.276	41.379	41.193	36.364	30.435	38.927

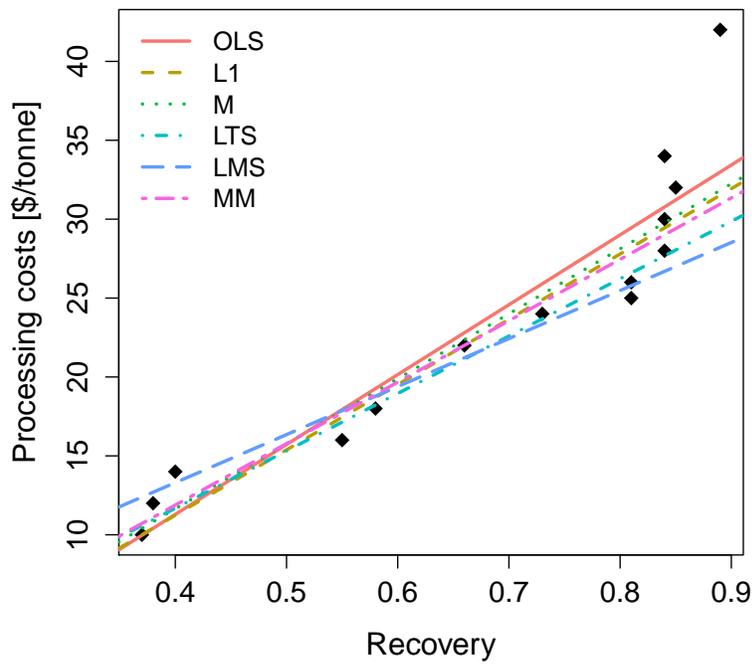


Figure 3-2: Comparison of robust estimators with OLS for the processing costs and recovery case study.

3.5.2 Iron and alumina grade relation

As shown in Figure 3-3, the iron and alumina dataset contains 172 observations with a large amount of outlier points. It seems there are two populations sampled in the same dataset. Moreover, performing OLS regression will lead to erroneous estimates since the fitted line will be adjusted for the two samples at the same time, resulting in a bad fit for both of the samples.

OLS fitted line will be pulled slightly towards lower values for low level of alumina concentrations. Moreover, studentized residuals approach flag a large proportion of outliers having a value less than -2, using Equation 3.4. In this case study, the best practice would be to split the dataset in two categories and perform OLS separately for both sampled population.

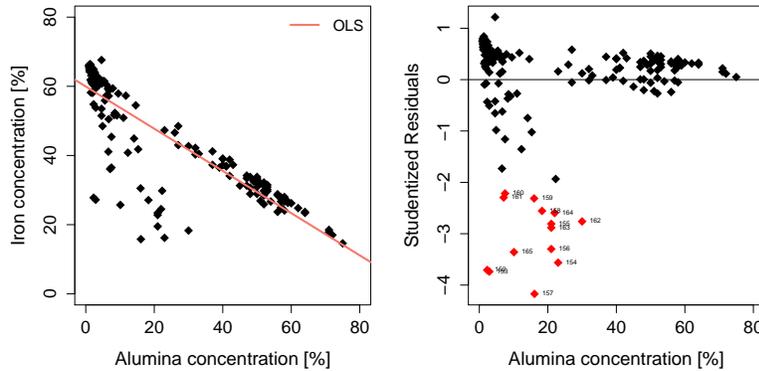


Figure 3-3: On the left, iron concentration in function of alumina concentration. A line fitted with OLS estimation is superimposed. On the right, corresponding studentized residuals.

Supposing the objective of a process engineer is to fit a line to the family that has the most samples, ignoring outliers, Figure 3-4 shows how robust estimators performs compared to OLS. Again, LMS is the least affected by outliers although all robust estimators have similar coefficients as shown in Table 3-5. Again, it is clear that using OLS without taking into account data contamination may lead to biased estimates. The difference between OLS coefficient and robust estimates varies between 4-5% in the alumina and iron case study.

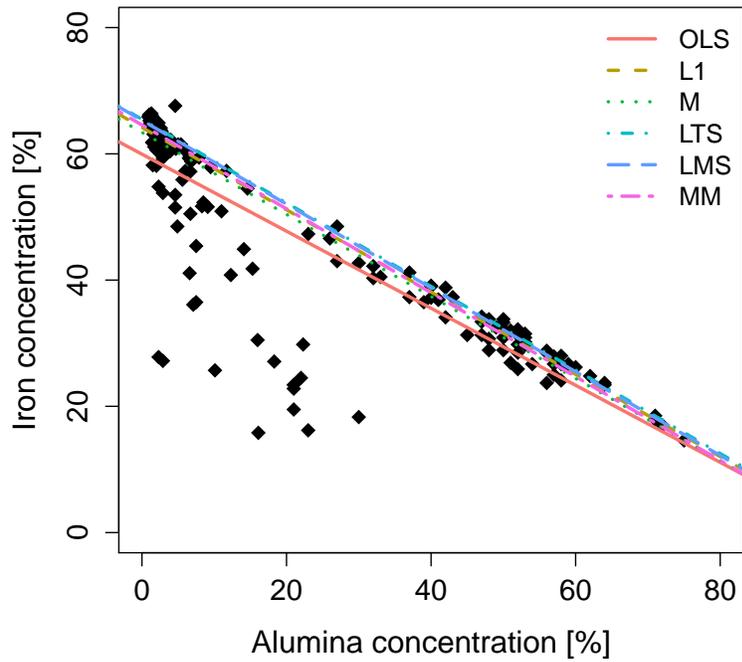


Figure 3–4: Comparison of robust estimators with OLS for the iron and alumina concentration case study.

Table 3–5: Summary of OLS-, L1-, M-, LTS-, LMS-, and MM-Estimation Intercept and Coefficient

	OLS	L1	M	LTS	LMS	MM
Intercept	59.950	64.172	63.407	65.478	65.322	64.574
Coefficient	-0.610	-0.652	-0.650	-0.662	-0.664	-0.664

3.6 Conclusion

In this paper, several robust estimators were compared to OLS in two simple case studies having each one predicted variable and one explanatory variable. OLS and robust estimation can be extended to the multiple regression case where multiple explanatory variables are used to predict a variable. Nevertheless, case studies showed that a mineral industry practitioner should be very careful when fitting a line to data. Even though OLS is well known and used widely, it may not always be an appropriate choice. If there are outliers, a robust estimator should be selected depending upon data characteristics. Robustness of regression estimators is always a compromise between efficiency and consistency. Some robust estimators (such as LMS and LTS) may not be entirely satisfactory even if they have high breakdown points because they can be inefficient. The efficient M-estimators are not robust in the explanatory variables and have breakdown points of zero. S-estimators can be highly inefficient, and MM-estimators require an iterative approach to be solved. When deciding which estimator to choose, one should put great attention in the definition of the regression model or non-linear trends in a sample. A decision maker should know difference between outliers and bad data. Outliers are good data that may have resulted from an extreme case of the sample generating process or can be generated by a secondary process. Another way to deal with outliers is to use data transformation to downsize their importance. For example, a logarithmic transformation can decrease the importance of very low values in a sample. It is to be noted that when performing data transformation, it is also necessary to change the interpretation of the transformed data.

3.7 Chapter Conclusion

This paper serves as a foundation for the rest of the thesis. Regression models are used in every subsequent chapters, and it is very important to understand the limitation of these models when a dataset is subjected to outlier data. In the next chapters, regression models will be used to analyze time series on metal futures contracts. The regression model will be implemented in a Kalman filtering framework, where an iterative algorithm is applied on a time series to calibrate its parameters on historical observations. As shown in this paper, there are several alternatives to the OLS regression approach. The next chapters will also show different approaches to the Kalman filter algorithm which can also be influenced by outliers.

The next chapter introduces the Kalman filter in the context of capital budgeting of an iron mine. In this paper, we use a Kalman filtering framework to calibrate the parameters of a stochastic model on historical observations of iron ore futures contracts. The Kalman filter is used, in conjunction with a Maximum Likelihood calibration approach to calibrate a geometric Brownian motion process on iron ore prices. Results shows that even if iron futures are more volatile, they have a higher expected return than long-term commitments, the privileged instrument in the industry to sell the iron ore production. As a result, a mining company can adjust its risk profile by gaining exposure to iron ore futures market. To better assess all the possible scenarios when investing in iron ore futures, a simulation approach with calibrated parameters can be used.

CHAPTER 4

A Kalman filtering approach to model net present value of an iron mine

4.1 Abstract

Iron ore was traditionally traded using long-term commitment (LTC) contracts. In the last decade, with the surging demand from China, a futures market was created for iron ore. In this paper, using historical information from this futures market, we focus on modeling market dynamics of Iron Fine 62% Fe - CFR Tianjin Port (China) futures contracts to determine optimal parameter values of the Schwartz (1997) two-factor model. A new approach using LTC and futures contracts is proposed to assess the Net Present Value (NPV) of an iron ore mining project. We apply Kalman filtering techniques to calibrate the two-factor commodity model to iron ore futures for the January 2014 to November 2016 period. The Kalman filter is useful to infer unobservable variables from noisy measurements. In the Schwartz (1997) two-factor model, the unobservable spot price and convenience yield are inferred from futures contracts transactions. Model parameters are fitted using maximum likelihood optimization. Using parameters derived from the Kalman filtering and the maximum likelihood approach, spot price simulations for the next seven years are made for three scenarios. The NPV of a mining project is calculated for each scenario. Then, both LTC and futures markets are treated separately and the mining company can choose which proportion of its production to sell in each market. Results show that the calibration and NPV simulation workflow can be effectively used to assess the profitability of a mining project, accounting for the exposure to futures markets.

4.2 Introduction

Evaluating the feasibility of a mining project requires a forecast of the expected future value of the commodity to be produced. Mining project developers usually rely on consensus forecasts of iron ore prices to predict the tendency of iron ore price movements over a horizon of less than two years. However, such forecasts are deterministic in nature and do not take into account the volatility of iron ore prices. Another valuation approach consists of simulating price paths and valuing the economic viability of a mining project for a broad range of scenarios. Traditionally, it was straightforward to perform such an analysis for iron ore projects, because the industry relied on long-term supply contracts ranging from 5 to 25 years (Rogers and Robertson, 1987). These contracts were renegotiated annually based on benchmark prices. Moreover, the long-term commitment (LTC) contract was traditionally used as a collateral to obtain capital for mining project funding. For an iron ore producer, relying on LTC contracts is a method to hedge price risk. The producer is on the short side of the hedge, whereas the buyer of iron ore is on the long side.

In recent years, China has become a dominant importer of iron ore to meet the growing demand for steel (Wilson, 2012). The demand for iron ore of China grew from 70 to nearly 630 megatonnes per year between 2000 and 2009. Demand is expected to increase even more through the 21st century (Patino Douce, 2016). In 2014, China imported more than 930 megatonnes of iron ore, confirming the upward trend (marketrealist.com, 2015). As a result of this increased demand, major mining companies (e.g., Vale, BHP Billiton and Rio Tinto) provided some of their iron ore production at spot price (Astier, 2015), incurring an upward pressure on prices because of the limited seaborne supply of the resource compared to the demand. An issue with short-term

surges in prices is the usually concomitant increases in price volatility. Higher prices are not always sustainable and are subject to drastic changes. The upward pressure on iron ore prices also encourages higher cost producers to enter the industry, leading to an increase in supply and a subsequent decrease in prices (Hurst, 2015), assuming demand does not exceed supply.

Since iron ore mining companies have the choice to sell their production with LTC contacts or through futures markets, they may decide which risk profile they prefer. Selling 100% of the production using LTC contacts can be considered risk-free, while selling 100% of the production in futures markets is very risky. The proportion of the commodity sold in LTC contacts or futures markets will directly affect the Net Present Value (NPV) of the project. Mining company managers tend to underestimate the price risk when assessing the feasibility of mining projects (Bertisen and Davis, 2008). As a result, errors are made in timing and capacity of mining projects, and the realized NPV is much lower than the NPV projected in the feasibility study of the project.

To provide reliable and independent spot prices for iron ore, The Steel Index (TSI), created in 2008, consists of daily iron ore spot market transactions supplied from more than 600 sources in the iron supply chain such as mining companies, steel makers and speculators. The dataset is cleaned and treated to account for outliers. Then, one spot price is calculated through a volume-weighted average, and the resulting iron ore spot price is redistributed to the index subscribers. The existence of a reliable and impartial index led to the creation of several financial instruments to offer mining companies and steel makers the possibility to hedge risk in price volatility. One of these instruments is the Iron Ore 62% Fe, China (TSI) Futures Settlements created by the Chicago Mercantile Exchange (CME) group (CME, 2015).

A wide range of studies have focused on the relationship between spot prices and futures. In general, the methodology consists of describing price dynamics using a stochastic differential equation (SDE) and applying Ito's lemma to solve the SDE. Brennan and Schwartz (1985) were the first to introduce a model that infers a relationship between commodity spot prices and futures. To link these two aspects, they introduced the convenience yield, which is a continuously compounded benefit that the owner of a forward or futures contract renounces when buying the contract. The greater the convenience yield, the lower the fair value of the contract. In Brennan and Schwartz (1985), the spot price moves according to the geometric Brownian motion (GBM) with deterministic convenience yield and risk-free interest rates.

In reality, the convenience yield should vary according to fluctuations in the inventory levels of the commodity and changes in expectation in supply and demand for different term structures. Moreover, macro-economic factors should influence the risk-free interest rates and hence the NPV through the discount rate. Gibson and Schwartz (1990) improved the previous models and tested a two-factor model for pricing weekly oil futures contract prices using the spot price and instantaneous convenience yield as variables. In their model, both random variables followed a joint correlated stochastic process, and the use of a stochastic convenience yield helped to better price short-term contracts. Zhang et al. (2014) used the mean reverting model to value Real Options (RO) on a mining gold project, and they showed that the flexibility value of the RO is maximized when the average mining cost equals the spot price. Schwartz (1997a) compared three models that took into account mean reversion. In the first model, a one-factor model following a mean reverting process governs the diffusion. The second model is the Gibson and Schwartz (1990) model. The third model is a three-factor model that also

includes randomly varying interest rates. In order to correctly estimate parameters, Schwartz (1997a) used Kalman filtering techniques to continuously and iteratively correct the covariance matrix.

The Kalman filter is widely used in finance for calibrating stochastic processes, but it is also used in geosciences to characterize orebodies (Nejadi et al., 2015). Ribeiro and Hodges (2004) further extended Schwartz's long-term two-factor model by replacing the Ornstein-Uhlenbeck with a Cox-Ingersoll-Ross to model the convenience yield. This ensures that the model is arbitrage free. Furthermore, they examined the spot price volatility as proportional to the square root of the convenience yield level. The multifactor models can be used to value mining projects in a RO valuation framework. Hedging strategies using RO can even be implemented and tested (Haque et al., 2016). Other types of RO models use production parameters such as cutoff grades and consider stockpiling to optimize the NPV of mining projects (Zhang and Kleit, 2016). Result show that such models significantly alter the cutoff grade used in the project since they consider the processing of the stockpiled material.

Both the Ribeiro and Hodges (2004) and Schwartz (1997a) models are adept at pricing short-term contracts, but fail when pricing longer term structures. Schwartz and Smith (2000a) eliminated the concept of convenience yield and modeled the log of the spot price as the sum of a short- and a long-term component. The model allows mean reversion in the short-term price and uncertainty in the long-term price. Movements in prices for long maturity futures contracts describe the long-term equilibrium level and the difference between short-and long-term contracts describes how spot prices mean revert in the short-term. This specification is better at pricing longer term futures contracts than previous models. Aiube and Samanez (2014) analyzed oil futures markets using longer maturity contracts that were not available in previous

studies, and they showed that the use of these multifactor models tends to underestimate the risk premium observed in futures prices.

Mining companies situated outside of the United States can also be affected by exchange rates because commodities are mainly traded in US Dollars (USD) (Rudenko, 1998). If the local currency of a mining company is positively correlated with a commodity, a downward pressure on commodity prices will be attenuated by an appreciation of the USD. However, if the local currency and the commodity are negatively correlated, they will move in the opposite direction. Thus, when the commodity is subjected to a downward pressure, the local currency will be stronger, resulting in lower cash flows. It is possible to hedge or completely eliminate currency markets risk using forward contracts.

In this paper, LTC and futures contracts are considered as two financial instruments upon which a mining company can rely to sell their production. LTC contracts expected futures prices are known with certainty, since they are fixed at the beginning of an iron ore mining project. However, futures markets can have higher expected returns than LTC contracts, but are typically volatile. To model the futures markets, we apply the Schwartz (1997) model to iron ore futures. The parameters are estimated with maximum likelihood optimization in a Kalman filtering framework. With an estimate of the parameters generating the spot and future prices, Monte-Carlo simulations are then applied to generate price paths calibrated to observed futures prices. A NPV valuation is performed with different iron ore futures price profiles corresponding to an optimistic, a neutral and a pessimistic scenario. Simulation outcomes are then combined with LTC to adjust expected return and volatility to the desired risk profile of a mining company. Results show that considering iron ore mining projects in a portfolio framework helps to mitigate

the risk of undertaking a project with overestimated NPV, because the production may be hedged with LTC. Although exchange rate variations can also significantly affect the NPV of mining projects, they are excluded from this paper. Instead, the focus here is on the application of an integrated framework to combine price behavior in two iron ore markets: LTC and futures markets. To the best of our knowledge, this is the first study applying Kalman filtering techniques to iron ore futures. The Kalman filter is used to calibrate the model to futures contracts using maximum likelihood optimization, and parameters are then used as an input in futures prices simulations to model the volatility inherent in futures markets.

4.3 Methodology

The analysis is focused on creating the iron ore time series. Then, state space models are introduced and used to predict future paths of iron ore prices. Finally, simulated iron ore prices are used in a NPV valuation framework to estimate the viability of an iron ore mining project.

4.3.1 Model Input Data

An iron ore 62% Fe, CFR China (TSI) futures contract fixes the prices today for one metric tonne of iron ore containing 62% iron, as well as other components such as moisture, alumina and silica to be delivered at a specific time in the future. The difference between an iron ore forward contract and a futures contract is that the latter is backed by a clearinghouse such as the CME group, thus eliminating counterparty credit default risk. The mechanism to reduce risk of default of hedgers and speculators is to mark-to-market gains and losses on a daily basis, depending on the movement of the spot price. For this reason, the position on any futures contract will pick up the volatility of the underlying commodity, regardless of the time to maturity. Futures contract investments are risky and volatile regardless of their time horizon.

There are two mechanisms to terminate a long or short position in a futures contract. The first is obviously to wait for expiry and obtain a cash settlement or delivery of the underlying commodity. The second is to offset the first transaction by an opposite transaction of the same quantity. In this situation, the clearinghouse recognizes that the net position of the investor is zero. These measures favor market liquidity since it is always possible to rollover or cancel futures contracts to avoid delivery. Because of their liquidity, metals are generally traded exclusively on futures markets and closest maturing contracts are considered as a proxy for spot prices. Iron ore futures are an exception to this tendency, because the industry still relies heavily upon LTC. This implies that the spot price and convenience yield relationship will not be fully factored into the analysis, because not all of the information is observable.

The iron futures prices are based on TSI (Platts, 2015). A script was developed in the R programming language (R Core Team, 2015) to download from Barchart daily Iron Ore 62% Fe, CFR China swap futures prices for each trading day in the last two years. For each trading day, future prices are available for contracts expiring on a monthly basis for the next 23 months. Contracts having the same maturities are then gathered and trading days are sorted in chronological order. Finally, forward curves are extracted (Fig. 1) and are produced using the Schwartz (1997) package of the R programming language (Erb et al., 2014).

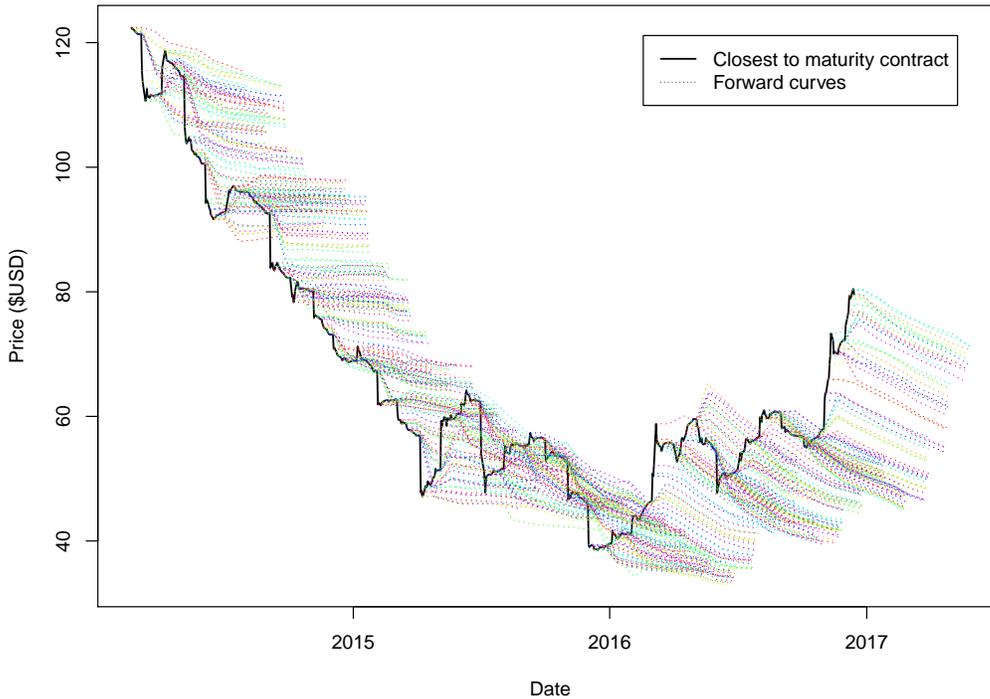


Figure 4–1: The solid black line represents closest to maturity daily Iron Ore 62% Fe, CFR China contracts, which can be interpreted as a proxy for TSI index level (or spot price). For each trading day, forward curve is extracted for contract maturing in the next year. The graphic is generated using the Schwartz 97 R package (Erb et al., 2014)

In Figure 4–1, the colored dashed lines are forward curves extracted for each trading day. They are constructed with contracts expiring up to 14 months into the future. Forward curves indicate where market participants locked in iron ore prices in the future for a given trading day. Figure 4–1 show only three years of history. Iron ore derivatives are a relatively new market compared to other futures markets, thus the calibration period is shorter than in other studies (e.g., Schwartz (1997)), where study periods ranged from five to ten years for oil and gold futures, respectively.

The solid black line in Figure 4–1 is constructed with the closest to maturity contracts and can be thought of as a proxy for the spot price of iron

ore. Because the closest to maturity contracts are stitched together to form a continuous contract, the continuous future curve can appear discontinuous, especially near expiration dates because of contango and backwardation effects (Masteika and Alexander, 2012). One way of obtaining a smoother, continuous future curve than the one presented in Figure 4–1 is to smooth prices on a time window surrounding maturity dates. However, in our case, we can measure the spot price directly because Iron Ore 62% Fe, CFR China is a derivative whose underlying asset is TSI (Platts, 2015).

The spot prices are obtained from the TSI (Fig. 4–2) from the Wall Street Journal using the Quandl API (Raymond McTaggart et al., 2015). The Wall Street Journal stopped publishing daily iron ore data at the end of 2015. Thus, the time series includes 18 February 2014 and 2 November 2015. Prices in the TSI time series have been converted to the logarithm of returns. To ensure the time series exhibit similar features no matter where it is analyzed in time, it has been partitioned in two samples. The training set comprises 350 daily observation of log returns between 18 February 2014 and 8 July 2015. The prediction set comprises 85 observations between 9 July 2015 and 2 November 2015.

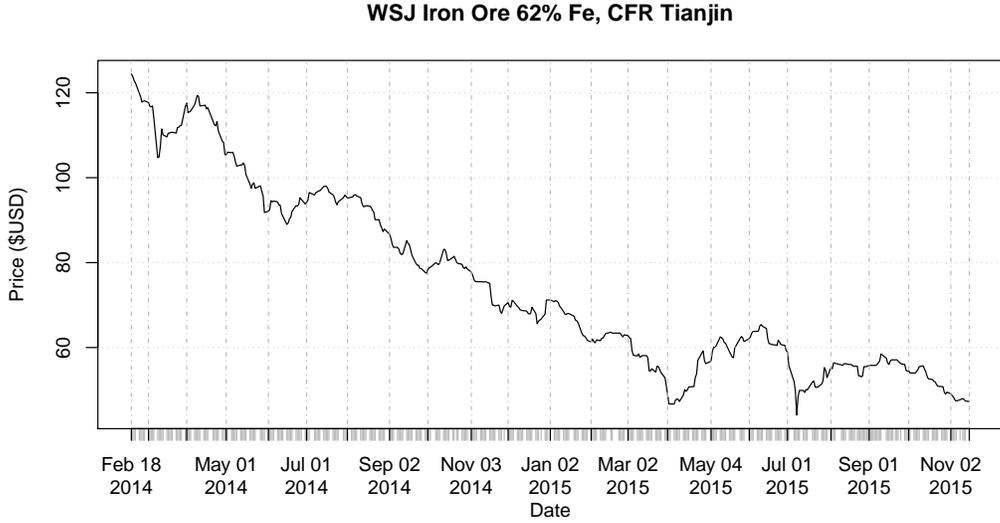


Figure 4–2: Iron Ore 62% Fe, CFR China contracts, TSI index level

Time series and future curve changes over time and it is convenient to model them with the Kalman filter. Any SDEs describing dynamics of spot or future markets can be represented in a state space form (Harvey, 1990). The measurement equation is given by:

$$y_t = Z_t \alpha_t + d_t + \varepsilon_t, \quad t = 1, \dots, T \quad (4.1)$$

where y_t is a multivariate time series containing N elements, Z_t is a $N \times m$ matrix, α_t is a $m \times 1$ vector and is called the state vector, d_t is a $N \times 1$ vector and ε_t is a $N \times 1$ vector of identically independently distributed disturbances with $E(\varepsilon_t) = 0$ and $Var(\varepsilon_t) = H_t$. The elements of α_t are not observed directly and are generated by a first-order Markov process having the following transition equation:

$$\alpha_t = T_t \alpha_{t-1} + c_t + R_t \eta_t, \quad t = 1, \dots, T \quad (4.2)$$

where T_t is a $m \times m$ matrix, c_t is a $m \times 1$ vector, R_t is a $m \times g$ matrix and η_t is a $g \times 1$ vector where $E(\eta_t) = 0$ and $var(\eta_t) = Q_t$.

The Kalman filter is a recursive estimator composed of a prediction and a measurement step. The prediction step occurs when predicting the variables at the next time step and consists of a motion of the probability density distribution. Since the prediction process obeys the total probability rule and the motion process has an error term with its own Gaussian distribution, information is lost. The linear algebra equations describing the prediction process are:

$$\alpha_{t|t-1} = T_t \alpha_{t-1} + c_t \quad (4.3)$$

$$P_{t|t-1} = T_t P_t T_t^T + R_t Q_t R_t^T \quad (4.4)$$

where $\alpha_{t|t-1}$ is the prediction, T_t is the state transition matrix, c_t is the motion vector, P_t is the transition covariance, R_t is the measurement noise and Q_t is the measurement covariance.

The measurement step is based on Bayes' theorem and consists of multiplying the prior density with a likelihood density distribution in order to get the best estimate of the mean and variance of a variable. In matrix notation, the measurement step can be described with the following equations:

$$a_t = a_{t|t-1} + P_{t|t-1} Z_t^T F_t^{-1} (y_t - Z_t a_{t|t-1} - d_t) \quad (4.5)$$

$$P_t = P_{t|t-1} - P_{t|t-1} Z_t^T F_t^{-1} Z_t P_{t|t-1} \quad (4.6)$$

where $F_t = Z_t P_{t|t-1} Z_t^T + H_t$. Each prediction-measurement cycle is an iteration of the Kalman filter where the new covariance matrix is estimated from all the available information at $t - 1$.

4.3.2 Maximum Likelihood Optimization

The Kalman filter offers a very robust method to adjust the covariance matrix of the transition equation to fit the actual observations. However,

parameters governing the transition equations are static and must be estimated prior to using the Kalman filter. The likelihood function of the Kalman filter is presented in Appendix 1 and can be used to minimize the residuals spread between predicted and actual state variables.

The Erb et al. (2014) R package has a maximum likelihood estimation (MLE) function that enables choosing which parameters are to be guessed and which are held fixed in the MLE optimization algorithm. The MLE optimization may lead to unrealistic parameters selection since the optimization problem is non-linear. For this reason, MLE is conducted multiple times by varying the initial parameters. This method reduces the chances of being stuck in a local maximum.

4.3.3 Schwartz Two-Factor Model

Since iron ore is a commodity that can be stockpiled, several factors such as the short-term interest rate, cost of storage and convenience yield will affect its forward curve. These factors can be used to model the theoretical relationship between spot and future prices. Some factors (e.g., interest rates on capital expended to acquire equipment used to produce the commodity or storage costs) will increase the cost of owning a commodity. This implies that the theoretical price of a futures contract, locking the price of a commodity today, should include these costs. For commodities whose convenience yield is lower than storage costs added to the risk-free rate, the futures price will be greater than the spot price so the market will be in contango. In reality, a downward sloping curve is often encountered, meaning that the market is in backwardation. This can be explained by the benefit of owning the commodity being greater than the costs of storing it plus the risk-free rate (Brennan and Schwartz, 1985). This non-monetary benefit is accrued to the owner of the commodity instead of the owner of the futures contract. For example, a mining

company can benefit from short-term surges in metal prices due to unexpected changes in market equilibrium. The profit is made rushing the schedule to put metals on the market when such price surges occur. Other factors such as hedging pressure can affect the term structure of the future curve. Iron ore markets are generally in normal backwardation. This implies that an excess of producers want to hedge price risk by taking a short position in the contract. They lock the selling price to a known level, because they own the asset, and they fear prices of the commodity may go down. When they lock the selling price to a known level, there is an excess of short positions in the market. As a result, speculators and long hedgers demand a higher risk premium to take the opposite position, exerting a downward pressure on future prices.

Spot and future prices of a commodity are related with the following equation (Geman, 2005):

$$f^T(t) = S(t)e^{(r+A-y)(T-t)} \quad (4.7)$$

where $f^T(t)$ is the future price of the commodity expiring at time T measured at time t , $S(t)$ is the spot price at time t , r is the risk-free interest rate, A is the storage cost, y is the convenience yield and $(T - t)$ denotes the remaining term of the contract.

The Schwartz (1997) two-factor model is used to describe the relationship between iron ore spot prices and instantaneous convenience yield:

$$dS_t = (\mu - \delta_t)S_t dt + \sigma_S S_t dW_S \quad (4.8)$$

$$d\delta_t = \kappa(\alpha - \delta_t)dt + \sigma_E dW_\epsilon \quad (4.9)$$

where the GBM term in spot price of the commodity S_t and the instantaneous convenience yield δ_t are correlated following $dW_S dW_\epsilon = \rho dt$. In equation 4.8,

spot price is a GBM term, whereas the instantaneous convenience yield is an Ornstein-Uhlenbeck process (Schwartz, 1997a). κ is the mean reversion parameter of the Ornstein-Uhlenbeck process and controls how fast the convenience yield time series reverts to its long-term mean α . Volatility of the spot price and convenience yield are modeled with σ_S and σ_E parameters, respectively. μ is the long-term drift parameter of the spot price and will be positive if prices are increasing in the long run.

The state space representation of the dynamics of $y_t = [\ln F(T_i)]$ is given by the measurement equation, where N is the number of futures contracts vectors:

$$y_t = d_t + Z_t[X_t, \delta_t]' + \varepsilon_t \quad i = 1, \dots, NT \quad (4.10)$$

with:

$$d_t = A(T_i), \quad i = 1, \dots, N, \quad N \times 1 \quad (4.11)$$

$$Z_t = \left[1, \frac{1 - e^{-\kappa T_i}}{\kappa}\right], \quad i = 1, \dots, N, \quad N \times 1 \quad (4.12)$$

with $E[\varepsilon_t] = 0$ and $Var[\varepsilon_t] = H$

Components of the transition equation described in Equation 4.2 for this model can be written as:

$$[X_t, \delta_t]' = c_t + Qt[X_{t-1}, \delta_{t-1}]' + \eta_t \quad i = 1, \dots, NT \quad (4.13)$$

where

$$c_t = [(\mu - 1/2\sigma_1^2)\Delta t, \kappa\alpha\Delta t] \quad (4.14)$$

$$Q_t = \begin{bmatrix} 1 & -\Delta t \\ 0 & 1 - \kappa\Delta t \end{bmatrix} \quad (4.15)$$

$$\text{and } E[\eta_t] = 0 \text{ and } \text{Var}[\eta_t] = \begin{bmatrix} \sigma_1^2\Delta t & \rho\sigma_1\sigma_2\Delta t \\ \rho\sigma_1\sigma_2\Delta t & \sigma_2^2\Delta t \end{bmatrix}$$

4.4 Results

4.4.1 Time Series Analysis

Descriptive statistics for the training and the prediction sets are presented in Table 4–1. For both samples, the means are negative. However, the in-sample partition is negatively skewed whereas the prediction partition is positively skewed. Moreover, kurtosis is higher in the training partition.

	n	Mean	SD	Min	Max	Skew	Kurtosis
Training set log(ret)	350	-0.002	0.018	-0.113	0.095	-0.461	8.365
Prediction set log(ret)	85	-0.001	0.014	-0.041	0.059	1.211	4.763

Table 4–1: Summary statistics for the training and prediction sets of the iron ore time series.

The density plots of the training and prediction sets show that the logarithm of returns in both sets are better modeled using fat-tailed distributions (Fig. 3). The training set contains more outliers and has more observations concentrated around the mean.

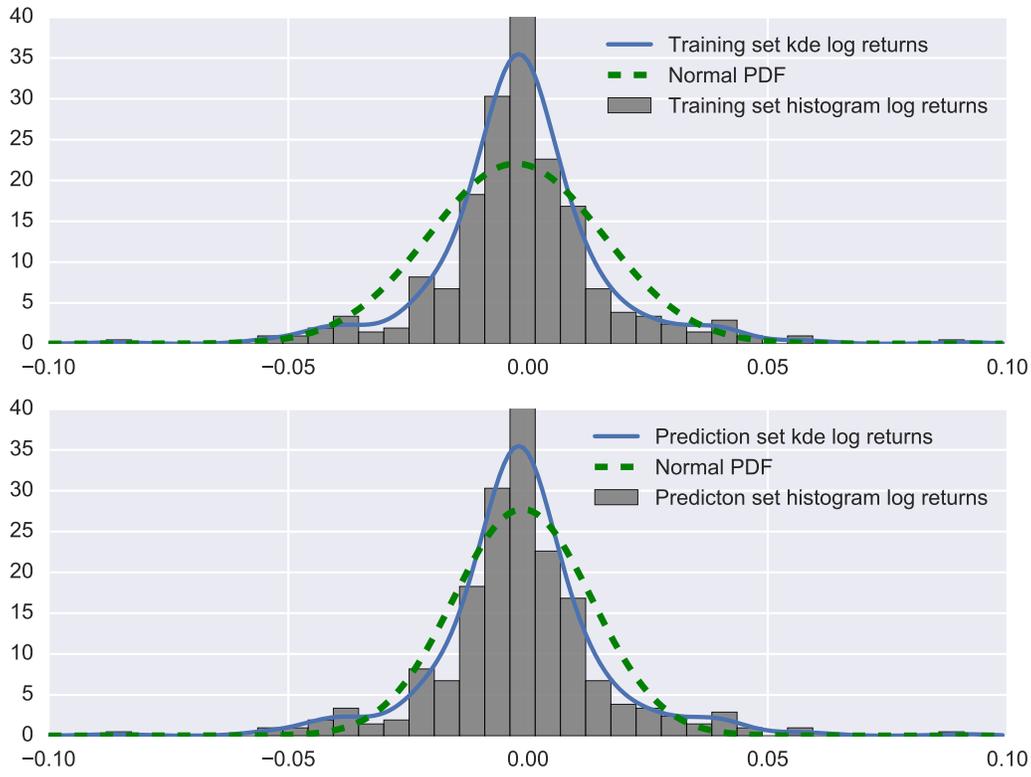


Figure 4–3: Training (above) and prediction (below) partitions for the iron ore time series. The histogram represents the distribution of log returns. The blue line is a kernel density estimation (kde) plot using a Gaussian kernel and the dashed green line represents a normal distribution with the theoretical mean and variance presented in Table 4–1.

The next plots analyze the iron ore futures contracts based on TSI. Because futures contracts are not traded uniformly, an analysis of the volume activity was performed. Figure 4–4 shows how the aggregated volume from each contract varies in time. From the inception date to March 2015, traded volumes in all contracts were quite low. There is an increase in volume from May 2015 until November 2015, then the volume then decreases. Since the dataset is very small, no observation was discarded based on volume.

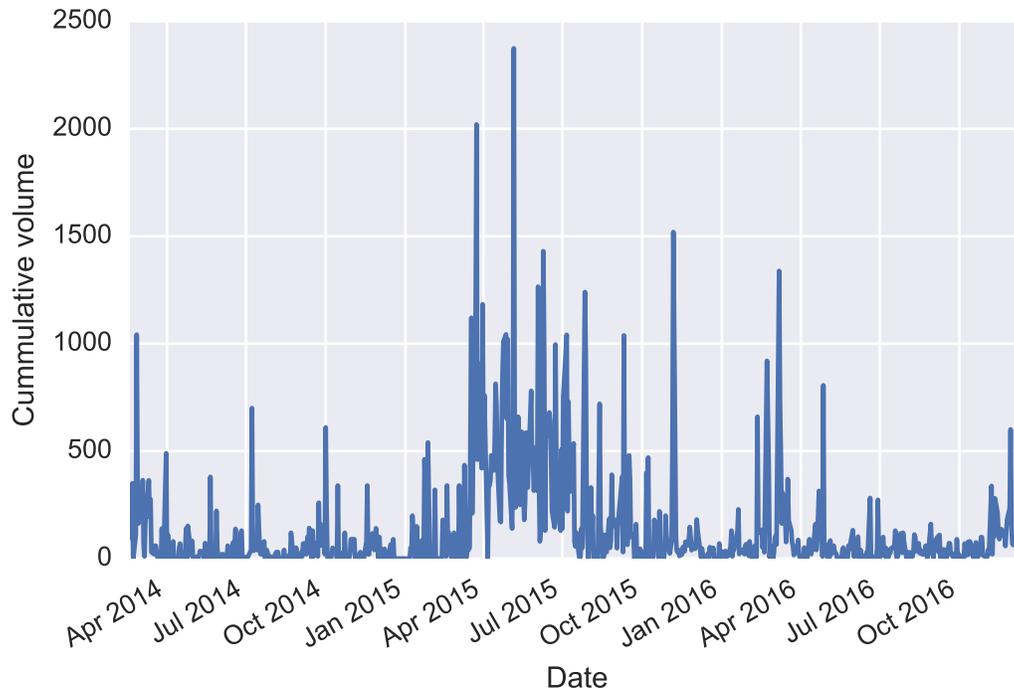


Figure 4–4: Aggregated volume from all contracts for the prediction partition of the iron ore futures time series.

In addition, volumes in nearest maturity are greater than in those contracts that are far from expiration (Fig. 5). Geman (2005) also found a similar relationship in the Brent WTI, a much more traded security. Nearest maturity contracts can have ten times the volume of longer maturities ones. In the Kalman filtering algorithm, measurement error standard deviations are set proportional to the average traded volumes, giving more weight to contracts that have more transactions. In this fashion, there is no need to discard long-maturity contracts since the weight will be scaled down automatically.

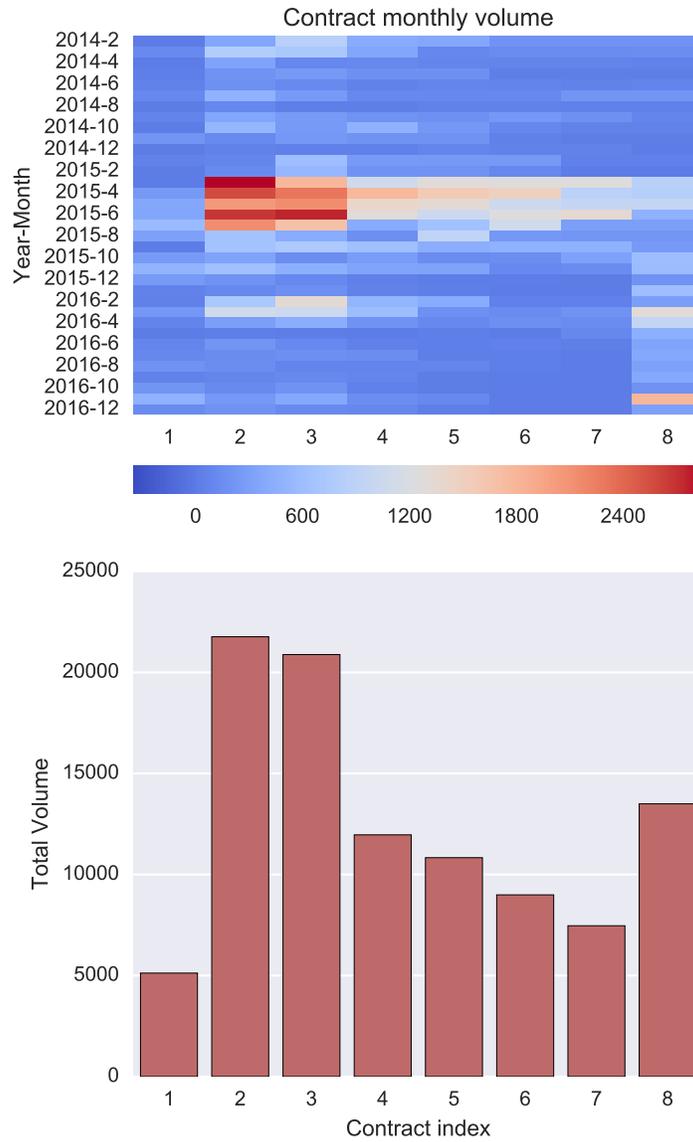


Figure 4–5: The top plot represents a heat map of traded volumes in function of maturity and date. Red color indicates highest traded volumes. The bottom plot represents the sum of volume of all contracts for a given maturity. Nearest maturities have the greatest cumulative volumes.

4.4.2 Kalman Filtering

As a first step, every parameter was unrestricted to let the MLE algorithm find the best fit. This led to unrealistic parameters such as a negative α . Therefore, α was set to 0.01 and the MLE optimization was performed on the remaining unrestricted parameters. The optimized parameters are shown on Table 4–2 and results are shown on Figure 4–6. The standard deviation of the parameters describes how the parameters changed between iterations during the MLE optimization process. It can vary widely for different parameters and can be used to assess how easily parameters converged during optimization. It is important to note that α and r were fixed, so their standard deviation are zero.

Parameter	Initial Estimate	Optimized Estimate	Standard Deviation
Drift (μ)	0.1	−0.12	0.13
Volatility of the spot price (σ_S)	0.8	0.39	0.15
Speed of mean reversion (κ)	2.0	3.47	0.75
Mean reverting level (α)	0.01	0.01	0
Volatility of the convenience yield (σ_E)	0.7	0.81	0.44
Correlation coefficient (ρ)	0.7	0.63	0.03
Risk-free rate (r)	0.03	0.03	0

Table 4–2: Initial and optimized parameters with the MLE algorithm after 300 iterations

In Table 4–2, μ represents the drift or the periodic return of spot price. The returns are calculated on an annual basis. σ_S is the volatility of the spot price. κ represents the speed of mean reversion of the convenience yield. The higher κ is, the higher the convenience yield will mean revert to its mean level α . σ_E represents the volatility of the convenience yield and ρ is the correlation between spot price returns and periodic changes in the convenience yield. r is the risk-free rate.

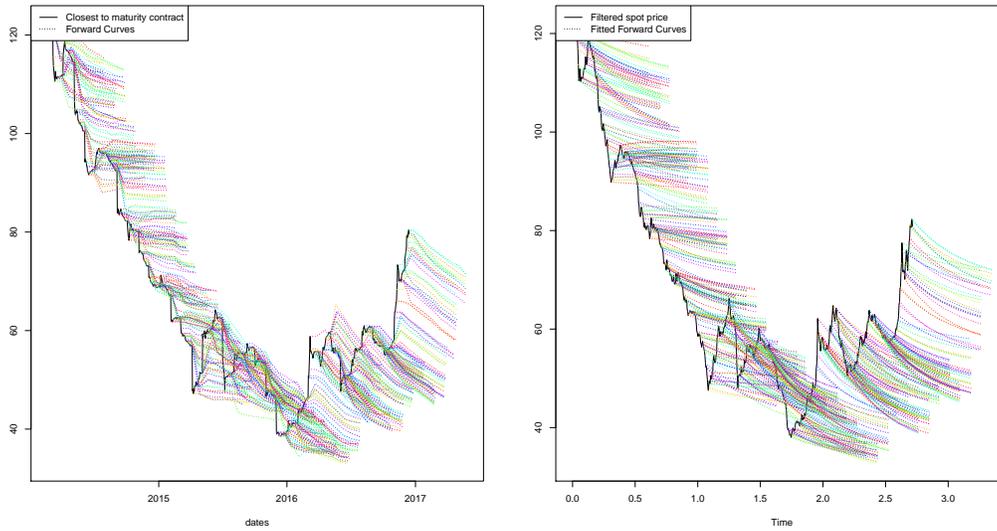


Figure 4–6: Left plot shows nearest actual contracts (black line) as well as future curves for each trading day in the studied period of time. The right plot shows the fitted corresponding variables estimated with Kalman filter.

Although the simulated forward curves are smoother than the actual values, they are able to correctly reproduce the slope of the different term structures. The nearest contract, represented by the black line in both plots, is also well fitted.

Figure 5.12 shows price and convenience yield simulation using the calibrated parameters shown in Table 4–2. The actual price falls between the bounds of the 95% confidence interval for the studied period of time most of the time, except at the end.

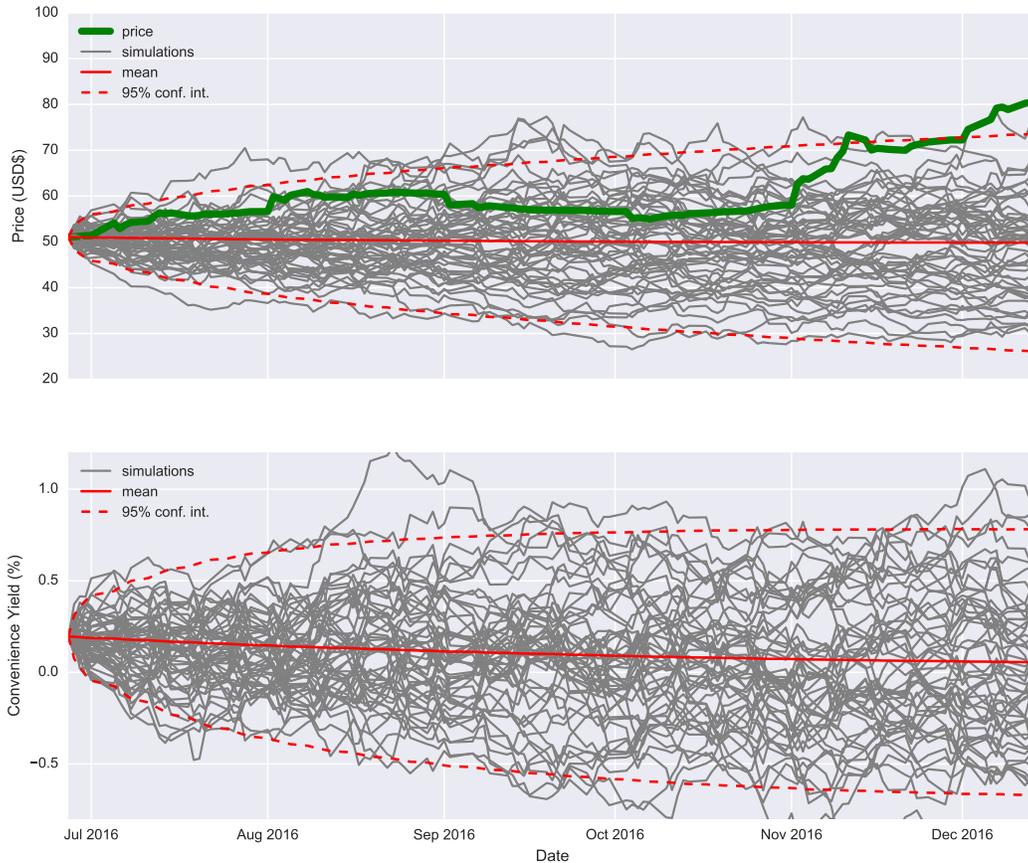


Figure 4–7: Simulated paths (in gray color) with parameters derived from the MLE algorithm. The continuous red line represents the mean of all the simulations and the red dashed lines represent the 95% confidence interval. The continuous green line is the actual price.

Figure 4–8 shows the residuals between the Kalman filter and actual observation of spot price. The distribution of the residuals is approximately normal with a mean dispersed around 0 for the first five maturing contracts. One of the assumptions of the Kalman filtering technique is that random disturbances affecting state space variables are normally distributed. Figure 4–8 confirms this assumption holds for the analysis.

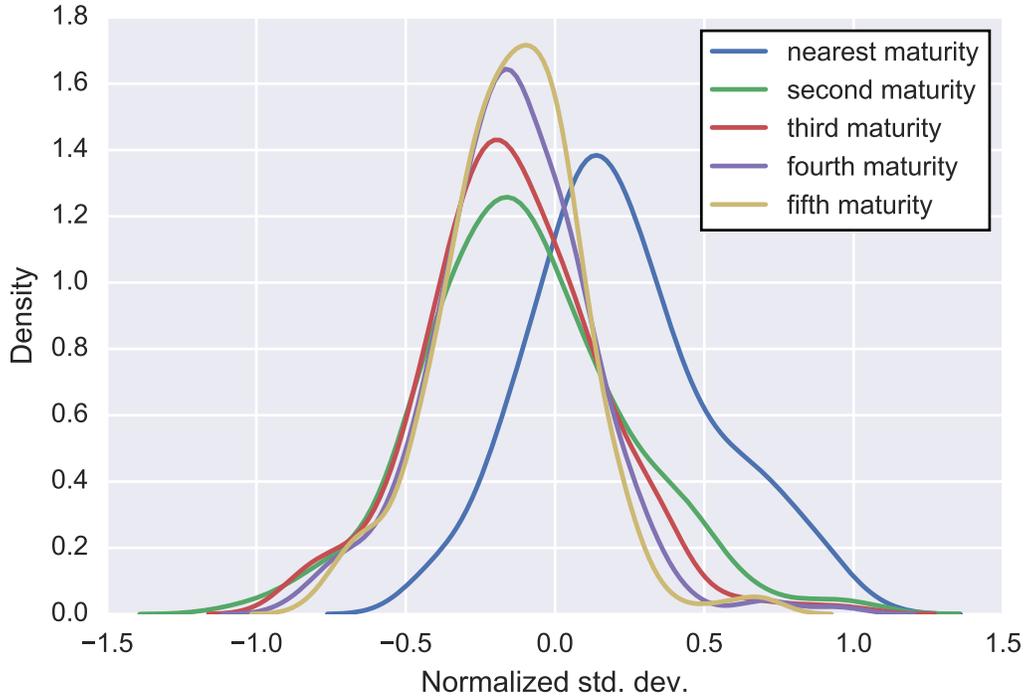


Figure 4–8: Kernel density plot of residuals for the first five maturing contracts.

Mining companies can decide upon the degree at which they will rely on LTC contracts and futures markets to sell their production. Let w_l be the weight of production sold in form of LTC contracts and w_f be the weight sold in future markets. $E(R_l)$ and $E(R_f)$ are expected returns using LTC contracts and futures markets respectively. Since LTC have zero volatility, the expected return and volatility on a portfolio of two assets is described with the following equations:

$$E(R_{portfolio}) = w_l E(R_l) + w_f E(R_f) \quad (4.16)$$

$$\sigma_{portfolio} = w_f \sigma_f \quad (4.17)$$

According to Equation 4.19, the volatility of iron ore prices returns in the portfolio is determined entirely by the weight of the production sold in

futures markets. If the w_f is 100%, the volatility of returns will be the one observed in futures markets. On another hand, if w_f is 0%, the volatility of returns will be zero. This relationship can be observed in Figure 4–10 where the NPV analysis is performed with different proportion of production sold on futures markets. When the proportion is zero, NPV has the same value, since it depends entirely on LTC prices. When the proportion sold in futures markets increases, the NPV of the optimistic scenario increases linearly, but it also decreases in both neutral and pessimistic scenarios.

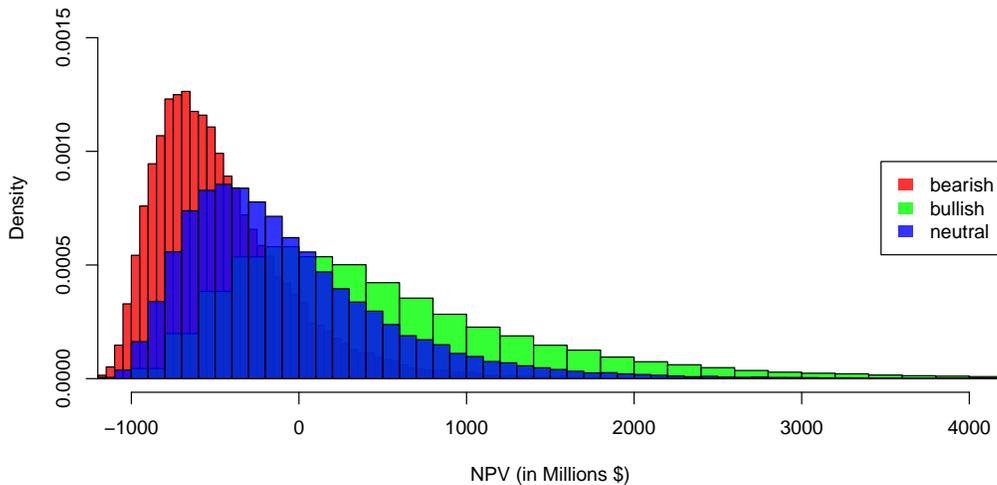


Figure 4–9: Histograms of NPV for the three different scenarios. The neutral scenario is in blue, while the bullish and bearish scenarios are respectively in green and in red.

Mining companies can decide the degree at which they will rely on LTC and futures markets to sell their production. Let w_l the weight of production sold in form of LTC and w_f the weight sold in future markets. $E(R_l)$ and $E(R_f)$ are expected returns using LTC and futures markets respectively. Since LTC have zero volatility, the expected return and volatility on a portfolio of two assets is described with the following equations:

$$E(R_{portfolio}) = w_l E(R_l) + w_f E(R_f) \quad (4.18)$$

$$\sigma_{portfolio} = w_f \sigma_f \quad (4.19)$$

According to Equation 4.19, the volatility of iron ore prices returns in the portfolio is determined entirely by the weight of the production sold in futures markets. If the w_f is 100%, the volatility of returns will be the one observed in futures markets. On another hand, if w_f is 0%, the volatility of returns will be zero. This relationship can be observed in Figure 4–10 where the NPV analysis is performed with different proportion of production sold on futures markets. When the proportion is zero, NPV has the same value since it depends entirely on LTC prices. When the proportion sold in futures markets increases, the NPV of the optimistic scenario increases linearly, but it also decreases in both neutral and pessimistic scenarios. The volatility is reflected in the proportion of production sold into futures markets.

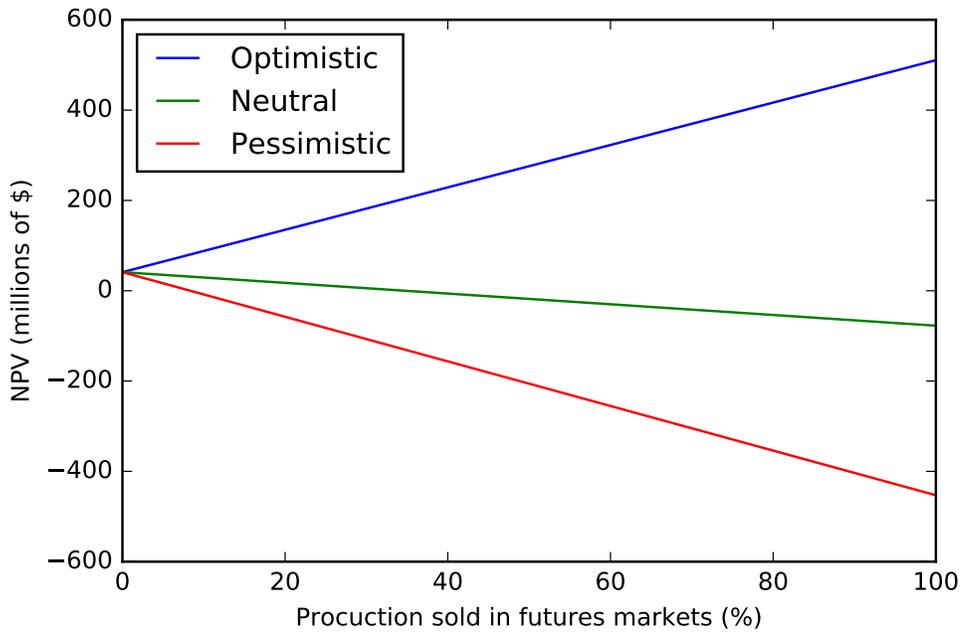


Figure 4–10: NPV with a varying proportion of production sold in futures markets. The analysis is performed for the optimistic, neutral and pessimistic price profiles in futures markets.

We modeled LTC and futures markets price movements to value the NPV of a mining project. When an iron ore company solely relies on futures markets to assess the NPV of a project, it may underestimate the market risk that it may face, since futures prices can be very volatile. As a result, markets can change drastically and using short-term futures prices, projects may appear profitable based on the NPV criterion. A better approach to assess the profitability of a project is to consider prices from LTC and futures markets separately. LTC contracts reflect required prices for both steel mills and iron ore mining companies to be sustainable. Then, depending on the willingness of a company to assume risk to increase expected returns, the proportion of production sold in futures markets should be adjusted accordingly. For example, a company may require that, on average, the simulated NPV is never negative. In the case of this study, this would imply finding where the

NPV profile of the pessimistic scenario intersects the w_f axis in Figure 4–10. In this example, the risk profile of the mining company has increased, but the potential upside return (in form of a higher NPV) has also increased. Since the pessimistic, or worst case scenario has a zero NPV, the sustainability of the mining project is not jeopardized. However, it is not generating any value for shareholders.

Discount rates are usually calculated with the Capital Asset Pricing Model (CAPM) (Rudenno, 1998). In the CAPM the risk is separated into two categories. The unsystematic risk can be diversified away and investors do not demand a premium for bearing this risk. On another hand, the systematic risk is inherent to the project and cannot be diversified away. In the CAPM, the systematic risk is measured using the relationship between the company's share price and an index composed of companies of similar sizes, in the same industry. If the project is similar to other projects the company invests in, then, this measure of systematic risk is appropriate for the calculation of the CAPM discount rate. Mining companies have the tendency to use relatively high discount factors due to the systematic volatility in metal markets or economic factors (Rudenno, 1998). Using a high discount rate does not guarantee that the NPV of a mining project has been correctly accounted for, especially when prices have increased drastically, as they have for iron ore in recent years. Moreover, using high discount rates may lead to an underestimation of the closing costs of a mine. Bertisen and Davis (2008) showed that capital costs required to open a mine are systematically biased downwards. With the methodology presented in this paper, it is possible to model the portfolio diversification effect of selling a percentage of production using LTC and the rest of the production in futures markets. This results in an overall better

NPV estimation framework, because it takes into account the risk associated with higher expected values of futures markets.

The calibration set is used to determine the best parameters to describe the time series. The Kalman filter is run until the optimal parameters are found for the given set, using MLE techniques. It is important to note that the optimization can converge to meaningless parameters, and it is thus advisable to revise and adjust accordingly. Then, these parameters are used as inputs to price and convenience yield simulations using Equations 4.8 and 4.9. Each of these equations has a stochastic and a deterministic component. The stochastic component is characterized by a GBM, meaning the volatility will increase by a factor of \sqrt{t} . This is why the confidence interval shown in Figure 5.12 keeps increasing with time. The deterministic part of the equation determines how the trend will move in time. The long-term drift coefficient μ is slightly negative. Only the spot and convenience yields are random in the Schwartz two-factor model. Thus, the hedging pressure risk premium is held constant. However, it is very likely that such risk premium will also vary in time. For example, if there is a risk of an oversupply in the iron ore market and there is an imbalance between hedgers and speculators willing to take the hedging risk, the market risk premium should change. For actively traded commodities, the US Commodity Future Trading Commission (CFTC, 2015) publishes the commitment of traders each Friday, which indicates the spreading between long and short participants in the commodity futures market. This information can be used to assess market risk premium. The risk premium should also be correlated with volatility, since higher volatility should induce higher uncertainty for speculators (Basu and Miffre, 2013). However, volatility is held constant in the Schwartz (1997) model. Therefore, it is impossible to capture this relationship.

Kalman filters require that the SDE describing the evolution of the state variables are well-defined. For example, a Kalman filter designed to track an object moving at a constant speed will not work properly if there is an acceleration component in the real process generating the object movement. The same can occur when describing evolution of future markets. In the Schwartz (1997) model, it is assumed that the SDE describing the relationship between the spot price and the instantaneous convenience yield fully reflects the dynamics of the futures market. Since the stochastic part of the process is governed by GBM, the error term is identically independently distributed, meaning that there is no serial correlation in the error term. However it is clear that other factors are not accounted for in the model. For instance, the price movement is constantly decreasing from its peak because of macro-economic factors, meaning the returns should be slightly skewed to the left. Moreover, only a tiny fraction of iron ore transactions are performed at the spot prices, meaning the effect of LTC contracts will not be accounted for in the Schwartz two-factor model. These factors will implicitly affect the error term in the Kalman filter, inducing serial correlation, heteroscedasticity in the residual or outlier points.

4.5 Conclusion

In this paper, we combined LTC and futures markets in a NPV valuation framework using concepts of portfolio optimization. To model futures markets, we applied Kalman filtering techniques combined with MLE optimization to fit the Schwartz two-factor model on iron ore futures (Erb et al., 2014). The data consisted of daily observations of a vector of futures contracts maturing at different months. Then, calibrated parameters were used to simulate joint random paths of iron ore spot prices and convenience yield. Even if the iron ore time series exhibited extreme movements for the studied period, the Kalman

filter and MLE procedure was able to adequately fit a model describing the spot price and forward curve relationship.

This approach can be used in the iron ore mining industry to assess risks in financial markets and correctly determine the sustainable NPV of a mining project. The implication for the mining industry is to consider the diversification effect in their NPV valuation approach. Another important aspect that should be accounted for in the analysis is the effect of the depreciation of the currency where the mining project is operating. Haque et al. (2015) showed that the Australian and US dollar are co-integrated. This implies iron ore prices and the Australian dollar move in the same direction in the long term. There should be a similar relationship between the Canadian dollar and metal prices. Since Canadian companies are selling their production in USD, the depreciation of the Canadian currency has the effect of smoothing the loss of iron companies in periods of crisis. In future work, the depletion of the resource is also a factor that should be considered in the analysis (Othman and Jafari, 2012).

MLE optimization can yield unrealistic parameter estimation because the optimization algorithm maximizes the fit between predictions and actual data, regardless of the real physical meaning of the parameters being optimized. As future work, it would be interesting to apply genetic algorithms to optimize the Kalman filter parameters, since it is possible to constrain each parameters within a range. Ranges such as a drift parameter between -0.1 and 0.1, or volatility in spot prices between 20% and 80%, could be chosen to be realistic. As a result, the genetic algorithm would lead to the best solution within a constrained range of possible scenarios. Another possible improvement of this work could be to extend the analysis to the case of a mine producing multiple

metals. The diversification effect of producing multiple commodity types could be addressed.

4.6 Appendix

Let y_1, \dots, y_T be a set of independent and identically distributed observations. The joint density function is given by (Harvey, 1990):

$$L(y; \psi) = \prod_{t=1}^T p(y_t | Y_{t-1}) \quad (4.20)$$

where $p(y_t | Y_{t-1})$ is the conditional distribution of y_t knowing all available information at time $t - 1$. The measurement equation of the Kalman filter can be written as:

$$y_t = Z_t a_{t|t-1} + Z_t(\alpha_t - a_{t|t-1}) + d_t + \varepsilon_t \quad (4.21)$$

The conditional distribution of y_t is normal with a mean:

$$E(y_t) = Z_t a_{t|t-1} + d_t \quad (4.22)$$

and a covariance matrix given by:

$$F_t = Z_t P_{t|t-1} Z_t' + H_t \quad (4.23)$$

For a gaussian model, the likelihood function can be written as:

$$\text{Log}L = -\frac{NT}{2} \log(2\pi) - \frac{1}{2} \sum_{t=1}^T \log |F_t| - \frac{1}{2} \sum_{t=1}^T v_t' F_t^{-1} v_t \quad (4.24)$$

4.7 Chapter Conclusion

This chapter introduced the Kalman filter and the Schwartz 1997 two-factor model. This iterative calibration approach can prove to be very useful to adjust the risk profile of a mining project. In this paper, it was shown that the calibration process using Maximum Likelihood estimation can be used to calibrate a stochastic process to iron ore futures. To perform the calibration, it was necessary to try using multiple initial conditions to ensure the algorithm does not get stuck in a local minimum. In the next paper, another calibration approach is used. Instead of transitioning from one solution to another using only one point estimate, the genetic algorithm uses a population of randomly selected values for the transition. As a result, it limits the chance the algorithm gets stuck in a local minimum.

The genetic algorithm works by mimicking the principle of the survival of the fittest in the nature. At each iteration, the best solutions are combined together to generate new solutions that should be closer to the optimal solution. In the next chapter, the algorithm is applied on the Schwartz-Smith two-factor model to assess the profitability of a copper mining project. Results shows that the solution from the genetic algorithm is more robust than the solution from the Maximum Likelihood estimator. The results are then used to simulate an active stockpile management strategy where the mean reversion of copper prices is taken into account. The strategy is able to produce significant alpha over the passive production strategy. The next chapter presents the methodology and results in details.

CHAPTER 5

Genetic algorithms for the optimization of the Schwartz-Smith two-factor model: A case study on a copper deposit

5.1 Abstract

Mining companies typically seek ways to hedge risks affecting their production. One useful instrument to mitigate the financial risk is the futures contracts on commodity prices. Information from the transactions in futures markets is publicly available and can be analyzed with the Schwartz-Smith two-factor model. However, finding the parameters governing this model can be very challenging. This step is done using a deterministic optimization approach called the Expectation-Maximization algorithm (EM). The starting values of the model will have a significant effect on the convergence of the EM. To ensure the solution does not get stuck in a local maximum, the EM algorithm is performed multiple times with different starting values. This paper assesses the value of genetic algorithms to optimize the parameters of the SWTF model. Although they are slower than EM algorithms because they use random number generators to search for the optimal solution, genetic algorithms (GA) optimize a population of solutions instead of working on only one solution at the time. Moreover, a constraint on the range parameter can be applied to ensure the parameter has a sound economic meaning. Once the SWTF parameters have been calibrated on the observation of futures contracts, the model can be used for the simulation of spot and futures prices. To demonstrate the performance of the proposed approach, a case study was conducted on a copper deposit. The simulations based on the SWTF model whose parameters are determined by GA are used. An active management

strategy of the stockpile, dependent on discrepancies in commodity futures prices is tested. Results show that the active management strategy produces positive returns over the passive investment approach.

5.2 Introduction

Commodities play a major role in finance. They have been a very good choice to diversify risk in portfolios due to low levels of correlation with stock markets during the 80s and the 90s (Gorton and Rouwenhorst, 2006). Commodities can also be traded using futures contracts. These contracts are standardized and represent an agreement between two parties to trade a pre-determined standardized quantity of the commodity at a given price in the future. They are traded at high volume in exchange markets such as the Commodity Exchange, Inc. (COMEX), the New York Mercantile Exchange (NYMEX) and the Chicago Board Options Exchange (CBOE). As a result, information related to spot and future prices, volumes and open interest are readily available at little cost. Several models have been developed to describe dynamics in the commodity futures markets. Gibson and Schwartz (1990) introduced a two-factor stochastic model where the first factor is the commodity spot price and the second factor is the convenience yield. The first factor is a Geometric Brownian Motion (GBM) and the second factor is an Ornstein–Uhlenbeck (OU) process. Both factors follow a joint stochastic process and they are correlated. Gibson and Schwartz (1990) showed that the two-factor model is appropriate at pricing futures contracts with short-term maturities. Schwartz (1997a) later introduced a three-factor model where the third stochastic variable is interest rates. When compared to the two-factor model, the use of a stochastic interest rate adds one layer of complexity, without improving much the fitness of the model. Similar to the two-factor model, the three-factor model of the Schwartz (1997a) model does not predict

longer term maturing contracts very well. Ribeiro and Hodges (2004) extended the Schwartz (1997) two-factor model so that the OU process becomes a Cox-Ingersoll-Ross (CIR) process. Also, to account for the link between the volatility and the convenience yield (and indirectly, the stock level of the commodity), the volatility is proportional with the square root of the convenience yield. As a result, volatility is high when stocks for a given commodity are low. Later, Schwartz and Smith (2000a) introduced the Schwartz-Smith two-factor model (SSTF) model. The SSTF model has been widely used to model commodity prices levels because of the simplicity to interpret its results. It divides the observed spot price into two components. The first component is a short-term price deviation which is not expected to persist. This unobservable variable is linked with the convenience yield as described in the Gibson and Schwartz (1990) paper. The second component is the long-term price path which is affected by shifts in the supply of and demand for the commodity. The notion of convenience yield is completely ignored. The SSTF model is a mean-reverting level but the long-term price level to which the model reverts is stochastic.

In this paper, the SSTF model is used to model the price dynamics of crude oil and copper futures contracts. The SSTF is implemented in a Kalman filtering framework to fit spot price and futures contracts. However, this requires the use of optimization techniques to correctly determine the parameters of the Kalman filter (KF). Expectation Maximization (EM) algorithms can be used to iteratively maximize the fitness using the Maximum Likelihood (ML) criteria but it cannot be guaranteed to converge to the global maximum. In the EM workflow, only one vector is optimized at a time. For this reason, it is important to recompute the workflow several times with different initial estimates to ensure the optimization does not get stuck in a local maximum.

Moreover, since the parameters are not constrained, the use of EM algorithms may lead to unrealistic parameters determination (Lüthi et al., 2014). Therefore, the use of Genetic Algorithms (GA) is investigated to address issues with conventional optimization methods. GAs search for a population of points instead of a single point so the initial estimate is not a major issue. Moreover, GAs transition rule from a solution to another is not based on gradient methods, but is rather probabilistic (Goldberg and Holland, 1988). This enables more flexibility in the choice of the objective function to maximize.

One of the most heavily traded commodities, with futures contracts expiring up to ten years in the future, is crude oil (Geman, 2005). Hedgers, arbitrageurs and speculators are very active in the crude oil market, making it very efficient and thus, ideal to test the GA workflow. Although it is not always possible, it is convenient to test the GA on a time series which is covariance stationary and mean-reverting. The time period studied by Schwartz and Smith (2000a) ranges between 2 January 1990 and 17 February 1995 and tends to mean-revert to a stationary level. The workflow is then applied to copper futures to calibrate parameters and infer the profitability of a copper-mining project. Once the optimal parameters of the KF have been found, the hidden state variables can be used to evaluate the profitability of a mining project using an active trading strategy. Several countries have adopted a standardized procedure to report the estimated profitability of mining projects. In Canada, the NI43-101 Disclosure Standards (Canadian Securities, 2012) is used and every public company is required to disclose an estimate of the profitability of a mining project using the Net Present Value (NPV) framework. NPV is the sum of all periodic discounted cash flows a project will generate. For mining projects, these cash flows depend on several parameters such as the discount rate, commodity prices, the production schedule, tax and royalties, the grade

of the ore, extraction costs, recovery and many other parameters. To assess the impact of these parameters on NPV, scenario analyses are performed by changing a vector of parameters to reflect a base case, a higher and a lower uncertainty scenario. To project commodity prices in the future, a moving average is applied to spot prices to forecast prices in the long-term. However, the projection of a moving average completely ignores the term structure of future prices. The use of the SSTF model provides a better framework for estimating spot prices as well as future prices.

NPV analyses fail to recognize value when a project exhibits a high level of uncertainty (Copeland et al., 2001). In fact, NPV analyses consider the management team of a project has no control on the parameters generating the cash flows. For example, if the price of the commodity a mine produces falls drastically, NPV analysis assumes that the management cannot scale down the production to cut losses. A promising and increasingly popular approach for valuing mining projects is the Real Options valuation (ROV) method. This method of valuation is similar to NPV in the sense that present and future cash flows are discounted and summed to value an operation. The difference is that the management team of a mining project is given options, for example, to mothball, close, open, accelerate or pause a project. Luehrman (1998) took the concept of financial options on stocks and applied it to value options contingent on real assets. He used the Black-Scholes model parameters to describe the project's parameters. Samis et al. (2005) showed that the traditional NPV of discounted cash flows analysis fails to recognize value in the operation of a copper mine project. This is because the flexibility to abandon (temporarily or permanently) to cut losses, expand or switch between projects is not captured by the NPV approach. Authors used ROV to find out the NPV decision rule would have led the mining company to forego a profitable investment. Lemelin

et al. (2006) used the Schwartz (1997) single factor model to simulate gold and nickel price paths to value the Ragland mining project. ROV using the least-square Monte Carlo method was compared to the conventional NPV analysis. With this approach, the investment decision can be compared to an American call option where the option is exercisable at any time before expiration.

Abdel Sabour and Poulin (2010) used ROV to value the option to expand or abandon the production of a copper mine. Authors have modeled the relationship between the uncertainty level (volatility) and these two Real Options. With low uncertainty levels, the threshold price for abandonment is high while the threshold price for expansion is low. When the volatility increases, the threshold of both options respectively increases and decreases. Haque et al. (2014) showed that as financial options, the volatility parameter σ is the major driver of the value of ROV over the NPV approach. For example, a mining company cannot easily suspend a mine production when the commodity prices are low. Later, Haque et al. (2016) tested a hedging strategy based on the fact that commodity prices are mean-reverting. In this paper, the management has the option, but not the obligation, to use an active trading strategy to manage the proportion of production sent to a stockpile or to the smelter. This option is inexpensive to put in place when compared to conventional ROV options.

The first part of this paper focuses on generating an artificial commodity spot and future price dataset with known parameters to test the robustness of GA for the optimization of the KF parameters. Then, the workflow is applied to the same crude oil futures dataset presented in Schwartz and Smith (2000a). In the third part, the workflow is applied on copper prices, comparing KF parameters estimated with the GA to gradient search methods (Goodwin, 2013). Finally, a mining feasibility study is undertaken using the

NPV approach and an active trading strategy based on the difference between short-term and long-term prices derived with the KF. The active trading strategy is compared to the conventional NPV analysis to assess the added-value of managing the stockpile. The originality of this paper rests on the use of a new optimization approach to find the optimal parameters of the SSTF model. The use of GA requires less user intervention than the conventional EM approach because they study a population of initial parameters vectors instead of optimizing one point at a time.

5.3 Methodology

In this section the background material to perform the analysis is presented. The first part explains how to derive the SSTF in the state space form and use matrix algebra to generate an artificial dataset. The second part details the GA workflow. The last part discusses how to use the SSTF calibrated on the observed market term structure of futures contracts to perform simulations of the state-space variables. These simulations are then used to value a mining project using an active trading strategy.

5.3.1 Schwartz-Smith two-factor model

In the SSTF model, the logarithm of the price of a commodity ($\ln S_t$) is separated in a short-term (χ_t) and a long-term (ξ_t) component (Schwartz and Smith, 2000a):

$$\ln S_t = \chi_t + \xi_t \tag{5.1}$$

The short-term price deviations emerge from unexpected shocks on the demand for the commodity and are not expected to persist. The long-term prices are related to macroeconomic factors that induce persistent changes in the equilibrium prices. The long-term trend is a Brownian Motion while the short-term deviation is mean-reverting to the long-term price level. The rates of change for both short-term and long-term prices are respectively:

$$d\chi_t = -(\kappa\chi_t - \lambda_\chi)dt + \sigma_\chi dz_\chi \quad (5.2)$$

$$d\xi_t = (\mu_\xi - \lambda_\xi)dt + \sigma_\xi dz_\xi \quad (5.3)$$

Where κ is the mean reversion rate, λ_χ is the risk of borrowing money over the short-term, σ_χ is the volatility of the short-term price deviations, μ_ξ is the drift rate of long-term prices, λ_ξ is the risk of borrowing money over the long-term and σ_ξ is the volatility of long-term prices. dz_χ and dz_ξ are increments of the standard Brownian motion and are correlated with the following relation:

$$dz_\chi dz_\xi = \rho_{\chi\xi} dt \quad (5.4)$$

Equation 5.3 describes price movements in a context where investors are risk-averse, meaning they have to be compensated for the risk they take by investing in futures markets. λ_χ and λ_ξ represents the short-term and long-term risk premiums respectively and can be considered as constant rates subtracted from the drift parameters. The risk-free notation is mostly useful for pricing options on futures contracts or for valuing capital budgeting projects in a ROV framework.

Using these equations, it is possible to generate an artificial dataset with known parameters. This dataset is used to test the robustness of the GA. The first step is to convert the continuous form stochastic differential equations in a discrete model. Using the notation of Fu and Tai (2009):

$$x_t = c + Gx_{t-1} + w_t \quad (5.5)$$

$$y_t = d + F'x_t + v_t \quad (5.6)$$

Where

$$x_t = \begin{bmatrix} \chi_t \\ \xi_t \end{bmatrix}, c = \begin{bmatrix} 0 \\ \mu_e \end{bmatrix}, G = \begin{bmatrix} e^{-\kappa\Delta t} & 0 \\ 0 & 1 \end{bmatrix} \quad (5.7)$$

$$y_t = \begin{bmatrix} \ln E\{S_{T_1}\} \\ \ln E\{S_{T_2}\} \\ \vdots \\ \ln E\{S_{T_n}\} \end{bmatrix}, d = \begin{bmatrix} a(T_1) \\ a(T_2) \\ \vdots \\ a(T_n) \end{bmatrix}, F' = \begin{bmatrix} e^{-\kappa T_1} & 1 \\ e^{-\kappa T_2} & 1 \\ \vdots & 1 \\ e^{-\kappa T_n} & 1 \end{bmatrix} \quad (5.8)$$

Where

$$a(T) = (\mu_\xi - \lambda_\xi)T - (1 - e^{-\kappa T}) \frac{\lambda_\chi}{\kappa} + (1 - e^{-2\kappa T}) \frac{\sigma_\chi^2}{4\kappa} + (1 - e^{-\kappa T}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} + \frac{\sigma_\xi^2}{2} T \quad (5.9)$$

w_t and v_t are zero mean Gaussian noises due respectively, to unexplained variability in spot prices and errors in the reporting of future prices. T_n stands for the maturity (in years) of futures contracts. The covariance matrix of w_t and v_t are

$$Cov[w_t] = Cov[(\chi_{\Delta t}, \xi_{\Delta t})] = \begin{bmatrix} (1 - e^{-2\kappa\Delta t}) \frac{\sigma_\chi^2}{2\kappa} & (1 - e^{-\kappa\Delta t}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} \\ (1 - e^{-\kappa\Delta t}) \frac{\rho_{\chi\xi} \sigma_\chi \sigma_\xi}{\kappa} & \sigma_\xi^2 \Delta t \end{bmatrix} \quad (5.10)$$

And

$$Cov\{v_t\} = \begin{bmatrix} s_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & s_n^2 \end{bmatrix} \quad (5.11)$$

Equations 5.5 and 5.6 can be used in a KF framework (Schwartz and Smith, 2000a; Fu and Tai, 2009). The KF works in two steps:

Prediction step:

$$\hat{x}_{t|t-1} = c + G\hat{x}_{t-1} \quad (5.12)$$

$$\hat{y}_t = d + F^T \hat{x}_{t|t-1} \quad (5.13)$$

Updating step:

$$R_t = GC_{t-1}G^T + Cov\{w_t\} \quad (5.14)$$

$$Q_t = F^T R_t F + Cov\{v_t\} \quad (5.15)$$

$$A_t = R_t F Q_t^{-1} \quad (5.16)$$

$$\hat{x} = \hat{x}_{t|t-1} A_t (y_t - \hat{y}_t) \quad (5.17)$$

$$C_t = R_t - A_t Q_t A_t^T \quad (5.18)$$

In the prediction step, observation ranging from the beginning of the time series to time $T = t - 1$ is used to predict the value of the states variables (\hat{x}) at time t . The update step uses the difference between the predicted value and the observation to readjust covariance matrices and further improve the next prediction.

The prediction and transition equations can be used to generate an artificial dataset constrained by known parameters. First of all, a vector containing a set of parameters to plug in the discrete model can be defined as:

$$\theta = \{\kappa, \sigma_\chi, \lambda_\chi, \mu_\xi, \sigma_\xi, \rho_{\chi\xi}, \chi_0, \xi_0, s_1^2, \dots, s_n^2\} \quad (5.19)$$

χ_0 and ξ_0 are the starting values of the short-term and long-term simulated prices and s_n represents the diagonal elements of the covariance matrix v_t .

To generate an artificial dataset, the following workflow can be used, using the multivariate normal distribution:

1. Initialize the parameter vector θ and the total number N of increments of Δt to simulate.
2. Draw a random sample from a multivariate normal distribution with covariance w_t as described in Equation 10.
3. Generate an observation of x_t using the parameter vector θ and Equation 5.5.
4. Draw a random sample from a multivariate normal distribution with covariance v_t as described in Equation 11.
5. Generate an observation of y_t using the parameter vector θ and Equation 5.6
6. Repeat steps 2 to 5 until increment N is reached.

Simulating the SSTF model in the state space form gives simulations of the whole term structure of futures contracts as well as generated known values for the χ_t and ξ_t variables.

5.3.2 Genetic algorithms

In this research, the Distributed Evolutionary Algorithms in Python (DEAP) library was used to optimize parameters of the SSTF model. GA is a meta-heuristic algorithm that mimics the natural selection process to solve optimization problems. In GA, any member of the population is defined by a set of parameters (chromosomes). When an individual mates with another, information is exchanged through the crossover operator and chromosomes of both

individuals are randomly selected to create a new individual. The mutation mechanism also ensures the newly created individual has the probability to have a random change in its chromosomes. Finally, appropriateness of every individual created is valued by solving an objective function. The individuals that are more apt at maximizing the objective function are selected to mate and produce a new generation. GAs are very different from deterministic optimization approaches. They can completely ignore the gradient of the optimized function since they rely on an objective function to evaluate the transition from a solution to another. This enables the algorithm to consider a population of points instead of focusing on the gradient optimization of a single point which can be stuck in a local maximum (or minimum) (Goldberg and Holland, 1988). Another advantage of this approach is that it does not depend on linear hypothesis. The workflow of the GA used in this paper is based on (Fortin et al., 2012; Bäck et al., 2000):

```

initialize P(t) from parameters vector  $\theta$ ;
while number of generations <  $N$  do
    |  $t = t + 1$ ;
    | Mate new population from  $P(t - 1)$ ;
    | Crossover;
    | Mutation;
    | Select best individuals from tournament;
    |  $P(t) =$  best individuals
end

```

Algorithm 1: Workflow of the genetic algorithm used in this paper.

The initialization process consists of generating a population from the parameter vector θ . Each parameter is randomly drawn from a uniform distribution with lower and upper bounds determined to have a sound economic meaning.

For example, correlation between long-term and short-term deviations can be set to lie between -1 and 1. This ensures the search space for the optimization procedure will yield realistic results. At each iteration of the GA, a new population is generated.

Two mechanisms are used to ensure the newly generated population is different from the parent. The crossover consists of exchanging information between two parents, generally by swapping one or more sets of parameters. The other mechanism, mutation, ensures a random mutation unrelated with the parents has a chance of occurring (Bäck et al., 2000). This is performed by generating a new parameter from a normal distribution with a predetermined standard deviation. The higher the standard deviation is, the more extreme the mutation has the chance to be.

The last step of the algorithm is to rank each newly generated individual in the population based on a fitness function. The randomly generated parameters are used as inputs in the SSTF model, and the corresponding implied future curves are compared to the actual observed futures contracts. The optimization of the objective function is performed using maximum likelihood. The individuals fitting the best futures contracts are selected based on a tournament, which allows for negative fitness values (Fortin et al., 2012).

5.3.3 NPV Incorporating an active trading strategy

The true power of the workflow is to be able to calibrate parameters of the SSTF model to the actual term structure of futures contracts to perform simulations of hidden state variables and futures contracts. With such a simulated dataset, it is possible to test how an active strategy can be implemented to generate active returns. Reeve and Vigfusson (2011) have found that although futures prices have not historically outperformed random walk forecasts by a high margin, they are performing well when there is a big difference between

spot and future prices. In the case of SSTF models, a big discrepancy between spot and futures prices can be observed when the long-term mean and short-term deviation diverge. When the short-term variation is well above the long-term mean, the market is said to be in strong backwardation. This market imbalance is unsustainable and is expected to mean revert to the long-term price. On the other hand, when the short-term deviation is well below the long-term mean, the market is in strong contango and spot price is undervalued. The active trading strategy consists of tracking the discrepancy between the short-term variation and the long-term trend.

NPV analysis is a cash flow valuation tool widely used to assess the value of a mining project. NPV analysis generates the after-tax cash flows that a project will create and discounts them using an appropriate discount rate. The discount rate can be seen as the hurdle rate required to satisfy debt holders and shareholders (investors) of the company owning the mining project. With unlimited available funds to invest, the decision rule is to accept projects with positive NPV. This analysis is usually performed using a passive investing strategy, where the management has no control over the project. In this paper, a basic NPV valuation is compared to the valuation using an active strategy. For the basic NPV valuation, copper prices are simulated with fixed production and milling rates. Free cash flows are then calculated on an after-tax basis and discounted at the appropriate rate. The periodic production is sold at a stochastic spot price and no material is stockpiled.

For the incorporation of the active strategy in the NPV valuation, three situations can prevail. In the first scenario, the estimated spot price is close to the long-term mean. In this case, the copper ore is mined and sent directly to the mill just as in the passive investment case. In the second scenario, the spot price is below the long-term trend by a given threshold. In this

case, the same quantity of ore is mined, but a fraction is sent to the mill and the other part is sent to a stockpile. As a result, mining costs are incurred immediately, but milling costs and sale proceeds are deferred to the future. Since the short-term discrepancy between the estimated spot price and the long-term mean is expected to mean-revert to zero, the goal is to withdraw material from the stockpile in the future when the spot price rises above the long-term mean. Withdrawal of the material from the stockpile is considered in the third scenario, where the estimated spot price is above the long-term trend by a given threshold.

5.4 Case study

5.4.1 Data

The first part of this paper focuses on the implementation of a workflow using GA to optimize the parameter used in the SSTF model. To assess the validity of the workflow, three datasets were generated. The first one, with 259 weekly observations, is the same length as in the original Schwartz and Smith (2000a) paper. The two other datasets have respectively 500 and 1,000 weekly observations. In the early iterations of the KF, the state may be in a transient phase, reducing the fitness between KF estimates and actual data. The use of samples of 500 and 1,000 observations will help to assess if the GA performs better with a greater sample size. Another important aspect of using a bigger sample size is to verify how the GA is efficient in terms of calculation time.

To increase the comparability of the study, the input parameters used to generate the dataset were taken from Schwartz and Smith (2000a) and are summarized in Table 5-1:

Parameter	Value
κ	1.49
σ_χ	0.286
λ_χ	0.157
μ_ξ	-0.0125
σ_ξ	0.145
μ_{ξ^*}	0.0115
$\rho_{\chi\xi}$	0.3
χ_0	0.117
ξ_0	3.01

Table 5–1: Parameters used in simulations of the SSTF model. These parameters are taken from the original publication of Schwartz and Smith (2000a).

A simulated time series containing 259 weekly observations is presented in Figure 5–1. The first part of the simulated dataset represents the state variables, also defined as the long-term price dynamics and short-term deviation.

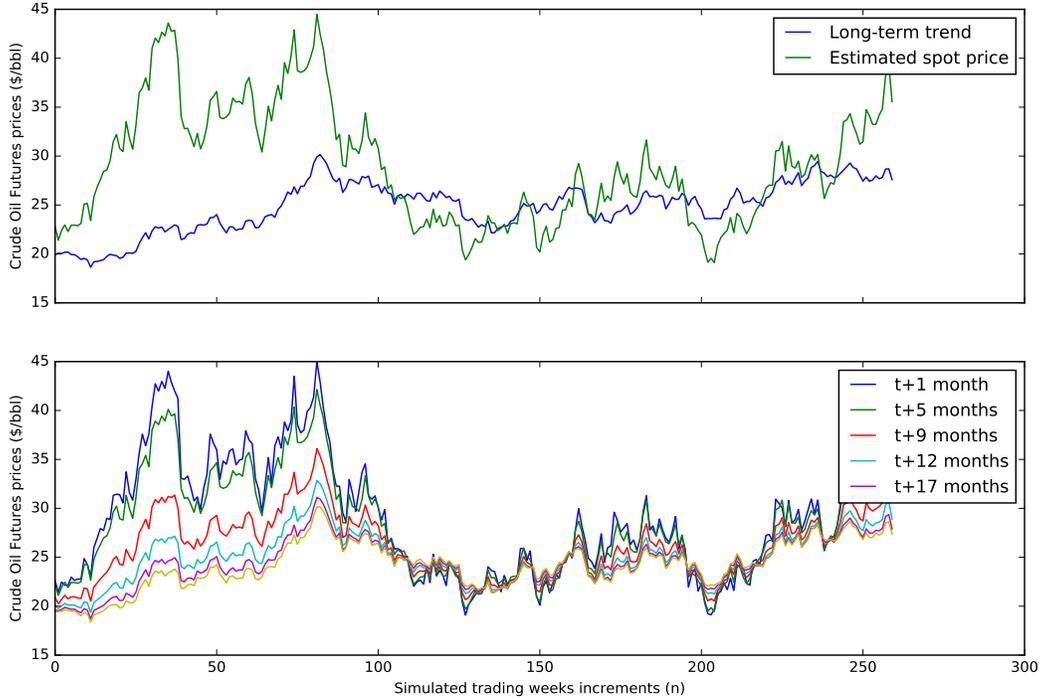


Figure 5–1: Simulated time series containing 259 observations. Plot a) shows the long-term price index as well as the sum of the long-term price and short-term deviations. Plot b) shows simulated commodity contracts with maturities of 1, 5, 9, 12 and 17 months.

In Figure 5–1 a), the red line represents the sum of the short-term deviations and the long-term mean which can be interpreted as the estimated spot price (which is not generally easily observable) through the KF. The green line represents the long-term price trend of the commodity, which can be modeled as a random walk. Short-term deviations are not expected to persist and the time-series should on average mean-revert to the long-term mean. This characteristic is replicated on Figure 5–1, where the spot price has a tendency to follow the long-term trend in the long run.

Figure 5–1 b) represents simulations of futures contracts expiring at different maturities. At each point in time, the expiration dates are on average 1, 5, 9, 12 and 17 months in the future. These futures contracts are used as an input in the GA and replicate the real problem where hidden state variables

are unobservable and need to be estimated using actual transactions occurring in commodity futures markets.

To assess the robustness of the GA on a real dataset, crude oil futures contracts expiring between 1983 and 2021 were obtained from Quandl API (<https://www.quandl.com/collections/futures/cme-wti-crude-oil-futures>). The contracts were then aggregated by their average time to expiration using Pandas (McKinney, 2015) library of the Python programming language. Then, they were averaged on a weekly basis and trimmed to match the sampling period of January 1990 to February 1995 of the Schwartz and Smith (2000a) dataset. Figure 5–2 shows time to expiration for each contract during the time period. There are a total of 17 contracts with a mean maturity ranging from 10 to 345 trading days. With 240 trading days per year, the contract expiring in 345 days corresponds to month 17. Contracts are simply stitched together and no roll methodology, such as applying a weighted average, was performed. The workflow to aggregate futures contracts from the Quandl API has been modified to be used on all commodities available on Quandl servers. This way, the SSTF model can be computed in real-time, or as soon as new data are available.

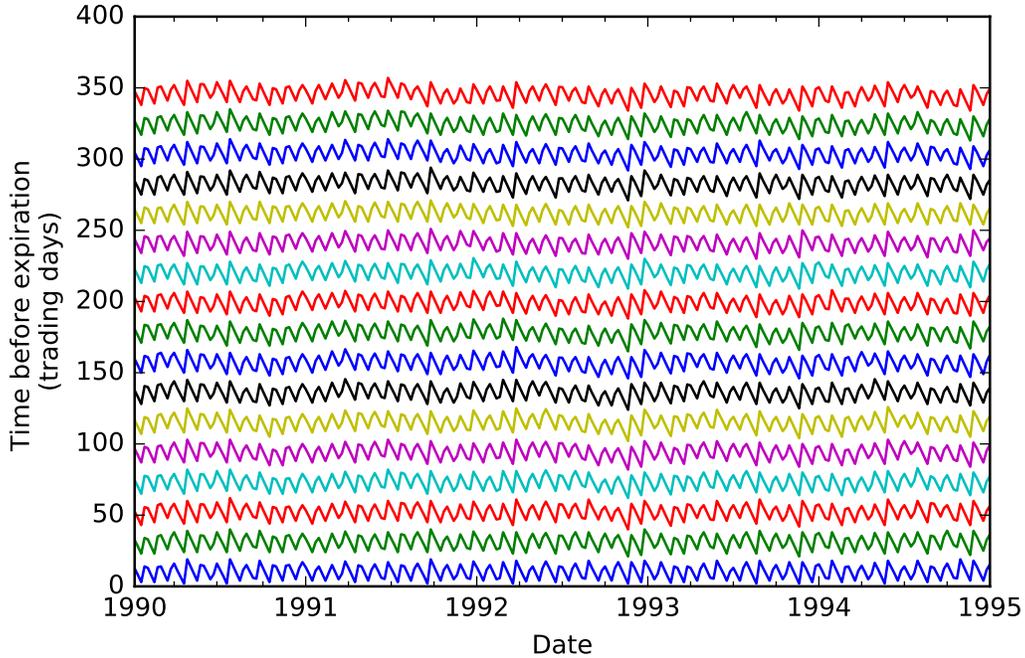


Figure 5–2: Time to expiration of crude oil contracts aggregated from Quandl API. The time series ranges between 1990-01-02 and 1995-02-17. Contracts are expiring up to 17 months in the future.

The sample of the dataset is presented in Figure 5–3 and consists of a continuously priced index with five different maturities. Although the source of futures contracts in this paper is different from the original paper of Schwartz and Smith (2000a), the aggregated futures for corresponding maturities are very similar. The authors have attributed the spike in the price to the Gulf War in 1990. It is also important to note that the artificial dataset has the same underlying parameters as the real dataset presented in this section. This means both datasets should exhibit similar features, and should revert to their long-term trends in the same amount of time. Moreover, the volatility is identical in both cases as is the drift parameter. It is also worth noticing that in periods of high volatility, the spread between the different futures contracts increases. This feature is observable for both artificial and real datasets.

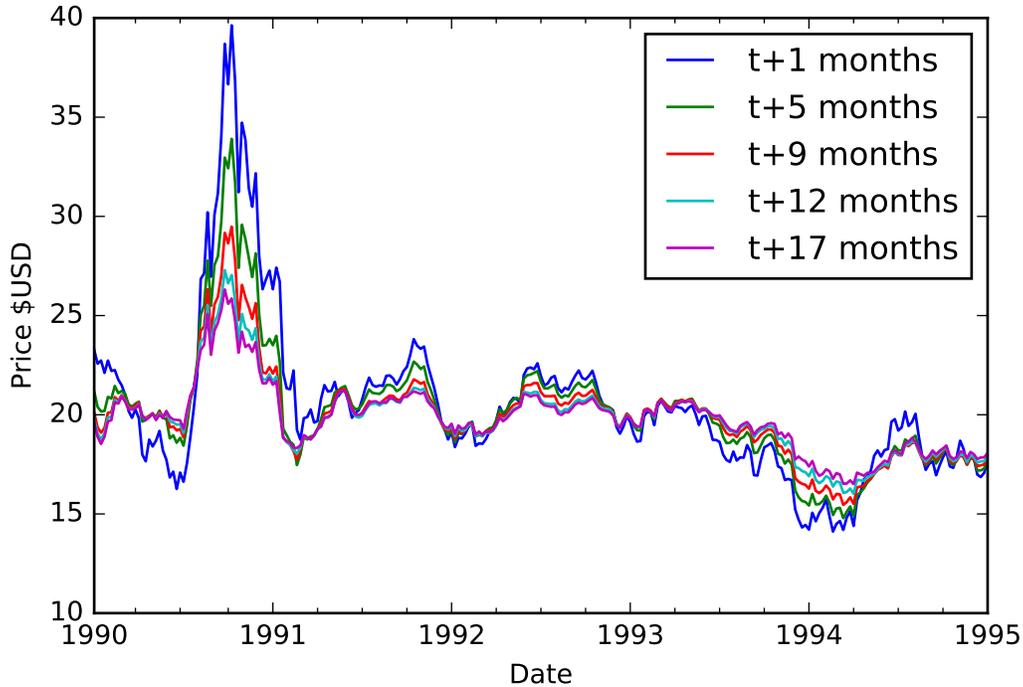


Figure 5–3: Crude oil futures contracts aggregated from the Quandl API. For each trading week (time t), contracts are expiring at 1, 5, 9, 12 and 17 months into the future.

5.4.2 Two-factor model calibrated on artificial data

A GA formulation was used to optimize the parameters of the KF on the artificial dataset created. The input vector consisting of a range for each parameter is presented in the following table:

Parameter	Minimum Value	Simulated Value	Maximum Value
κ	0.10	1.49	4.00
σ_χ	0.050	0.286	0.500
λ_χ	0.020	0.157	0.300
μ_ξ	-0.2000	-0.0125	0.3000
σ_ξ	0.020	0.145	0.200
μ_{ξ^*}	0.0020	0.0115	0.2000
$\rho_{\chi\xi}$	0.15	0.30	0.75

Table 5–2: Minimum, maximum and the simulated values of the parameters in the SSTF model simulations. Simulated values are the parameters the GA will try to retrieve and the range between minimum and maximum value is the search space for the optimal solution.

The GA uses a random uniform distribution bounded by the ranges of the parameters presented in Table 5–2 to generate each individual of the population. A penalty function is added to the fitness function to ensure that the solution lies in the defined bounded region. This procedure provides an advantage when compared to the classical EM algorithm because parameters will stay consistent with the *a priori* knowledge of the model. On the other hand, this requires an adjustment to the range of parameters according to the uncertainty. If the search space is too restrictive, the optimal solution could lie outside of the defined range and the solution could be infeasible.

The GA was performed with an initial population of size $n = 200$. Rather than using a convergence criteria, the algorithm was simply run on 25 generations. The crossover and mutation probabilities are respectively 25% and 20% and when a mutation occurs, the standard deviation of the mutated parameter is 0.25. Figure 5–4 shows the maximum fitness as well as the average fitness for each iteration. The interest lies in finding the solution with the maximum fitness. For this reason, the convergence of the maximum fitness solution is the predominant criteria to assess if the GA performs satisfactorily. The average fitness line seems to have several peaks and troughs, but it is only due to the strong effect that some parameters can have on the fitness function when a mutation occurs. After 25 generations, the optimal parameters are summarized in Table 5–3. The GA ran for 393 seconds.

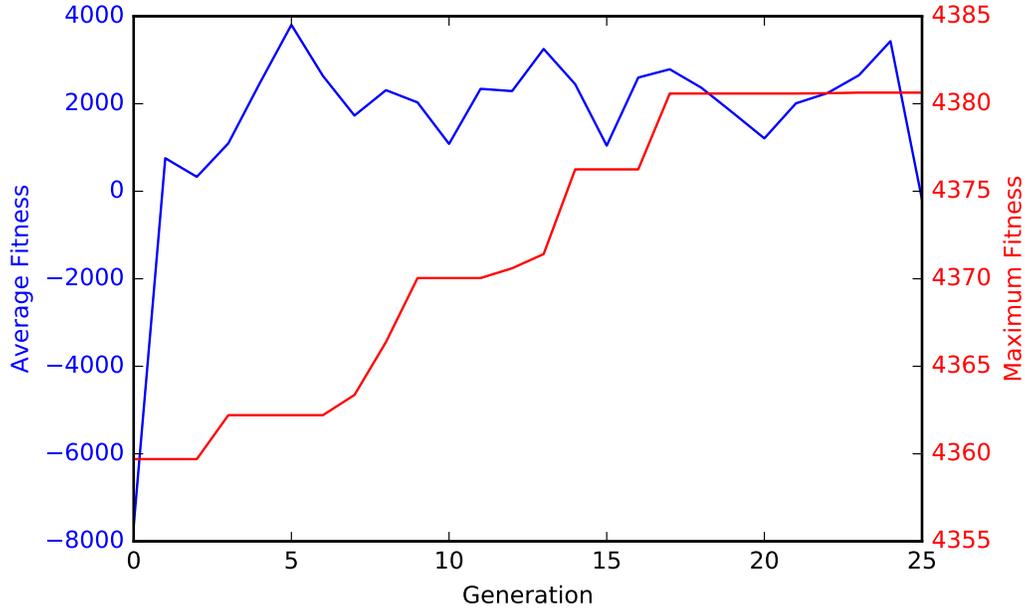


Figure 5–4: Convergence of the GA. The blue line represents the average fitness of the population at each generation. The red line represents the solution within the population having the highest fitness.

Parameter	True Value	Estimated $T = 200$	Estimated $T = 1,000$	Estimated $T = 2,000$
κ	1.490	1.462	1.505	1.509
σ_χ	0.286	0.269	0.272	0.273
λ_χ	0.157	0.125	0.142	0.147
μ_ξ	-0.0125	-0.0912	-0.0124	-0.0182
σ_ξ	0.145	0.124	0.152	0.173
μ_{ξ^*}	0.0115	0.0094	0.0150	0.0152
$\rho_{\chi\xi}$	0.300	0.443	0.382	0.321

Table 5–3: Optimal parameters after 25 generations of the GA for different period lengths T .

5.4.3 Two-factor model calibrated on crude oil futures prices

The GA was applied on the same dataset used in Schwartz and Smith (2000a). The calibration procedure took 25 generations. The optimal solution is presented in Table 5–4. The results are slightly different from the original paper with respect to parameters μ_ξ , μ_{ξ^*} and $\rho_{\chi\xi}$. The fitness function using

Schwartz and Smith (2000a) gives a value of 3,401 while the optimized parameter selection yields a fitness of 3,431. The difference is not attributed to a lack of convergence but rather to having found a solution with higher levels of fitness. Another reason for explaining the discrepancy may be differences in the dataset used. Schwartz and Smith (2000a) used crude oil futures from Knight-Ridder Financials and this research was performed from the Quandl API.

Parameter	Estimated (Schwartz and Smith, 2000a)	Estimated With GA
κ	1.490	1.590
σ_χ	0.286	0.275
λ_χ	0.157	0.156
μ_ξ	-0.0125	0.0441
σ_ξ	0.145	0.145
μ_{ξ^*}	0.0115	0.0031
$\rho_{\chi\xi}$	0.300	0.492

Table 5-4: Optimal solution of the GA after 25 generations

Figure 5-5 a) represents original observations of the t+1 months crude oil contracts (in blue) as well as the estimated t+1 price using the KF (in green). The long-term equilibrium level is in red. Plot 5-5 b) represents residuals between observed contracts prices and prices estimated for each maturity. The residual term exhibits a level of heteroscedasticity for the near maturing contract (1 month). This characteristic is consistent with Schwartz and Smith (2000a) results, where the near maturing contract was the most uncertain. Also, there is a discrepancy in the first 25 iterations of the KF. This may be caused by the fact that the KF is still in a transient state. The objective function was modified to take into account this transient state when valuing the fitness. However, discarding the first 10% of observations to be valued with the fitness function did not seem to improve noticeably the GA methodology.

Table 5–5 presents mean error, standard deviation of the error term and the mean absolute deviation (MAD). All three measures of error are higher with the near-maturing contract. As Schwartz and Smith (2000a) remarked, these error terms are mainly imposed by the diagonal terms of the covariance matrix used in the SSTF model. In this case, the nearest maturing contract had the highest variance value (0.042). When imposing the covariance matrix of the KF, it is important to choose the diagonal terms not to overfit observations of the model.

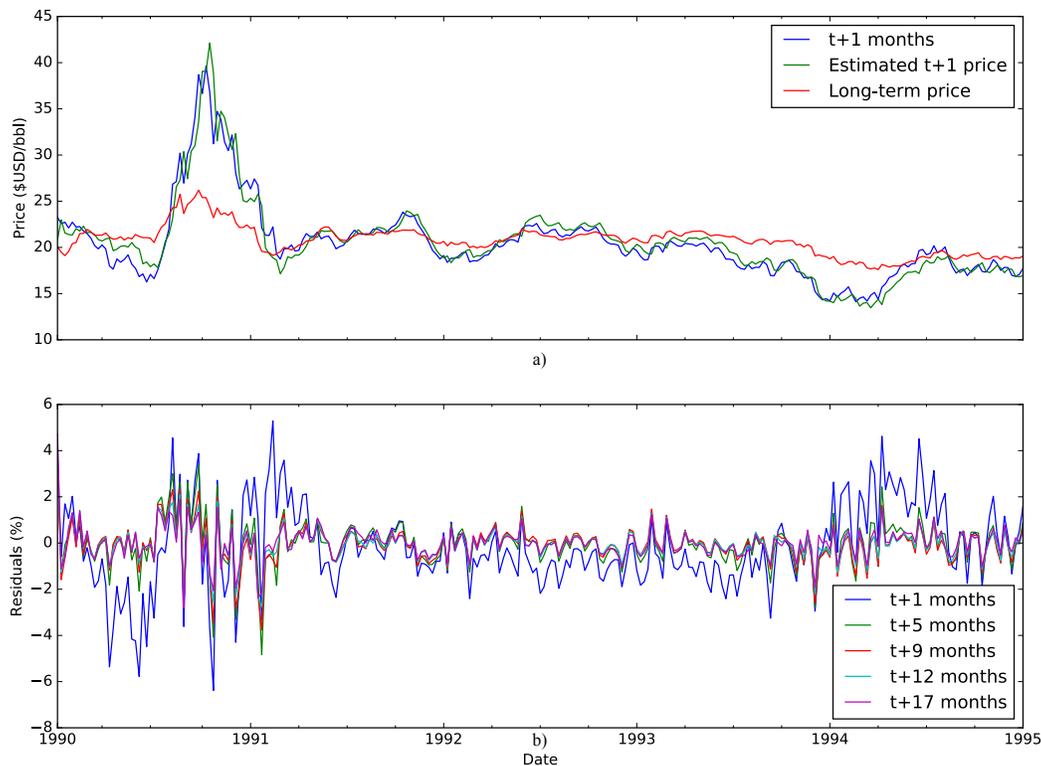


Figure 5–5: Plot a) represents the t+1 month crude oil contract (in blue) as well as the KF estimate of the t+1 future contract (in green). The long-term price estimate of the KF is represented in red. Plot b) represents residuals (in %) of the KF estimates for each maturing contract.

Maturity	mean error	st. dev.	MAD
t+1	0.0527	0.00700	0.0415
t+5	0.0285	0.00014	0.0198
t+7	0.0235	0.00060	0.0160
t+9	0.0205	0.00023	0.0138
t+17	0.0195	0.00021	0.0135

Table 5–5: Mean error, standard deviation of errors and mean absolute deviation of the fitted maturities for the crude oil futures dataset.

5.4.4 Two-factor model calibrated on copper prices

The first part of this research aimed at developing and testing a workflow to calibrate parameters of the KF in the SSTF model. Now, the workflow is applied on copper futures obtained from the Quandl Python API. Copper is an industrial metal heavily used in the electric equipment and electronic component industries. Since its production is dependent on industry cycles, an increase in copper prices attracts new players in the copper-mining industry, resulting in a higher supply of the commodity. For this reason, the price of copper has a tendency to mean-revert to a long-term level. This feature of the copper market was confirmed by performing Dickey-Fuller unit root test solved on a long period (Laughton and Jacoby, 1993).

Copper is also very sensitive to recessions and is particularly affected by drops in the US GDP (Chevallier and Ielpo, 2013). Drops in copper futures prices can thus be of high magnitude, as observed in the 2008 financial crisis. Goodwin (2013) has identified a structural break in the price dynamics of the copper market. For this reason, he separated the dataset in two subsamples, each one affected with different market dynamics. The first period is between 21 July 1993 and 27 September 2002 and the second sub-period is between 30 September 2002 and 5 March 2013 and includes the spectacular drop associated with the 2008 financial crisis. Both datasets consist of daily observations of the spot price and the 3, 15 and 27-month contracts. The GA workflow was

applied on this dataset and yielded similar parameter estimates, as reported in Table 5–6.

Param.	sub-period 1 (Goodwin, 2013)	sub-period 1 with GA	sub-period 2 (Goodwin, 2013)	sub-period 2 with GA
κ	0.690	0.477	0.338	0.367
σ_χ	0.152	0.212	0.204	0.201
λ_χ	-0.015	0.073	-0.101	-0.105
μ_ξ	-0.016	-0.020	0.183	0.1066
σ_ξ	0.152	0.091	0.228	0.237
μ_{ξ^*}	0.060	0.023	-0.085	-0.084
$\rho_{\chi\xi}$	-0.133	-0.051	-0.209	-0.192

Table 5–6: Parameters of the original Goodwin (2013) paper and parameters optimized with the GA for two different sub-periods affected by different market dynamics.

In sub-period 2, parameters are very close to the ones calculated as discussed in Goodwin (2013). The parameter μ_ξ is slightly different, but as noted by the authors, this parameter is estimated with very low precision since it has no effect on the maximum likelihood estimate.

The workflow was applied on an updated version of Sub-period 2, ranging from 30 September 2002 to 7 July 2016. In this subset, copper futures mean-reverted to a lower long-term level. The fitted dataset is presented in Figure 5–6.

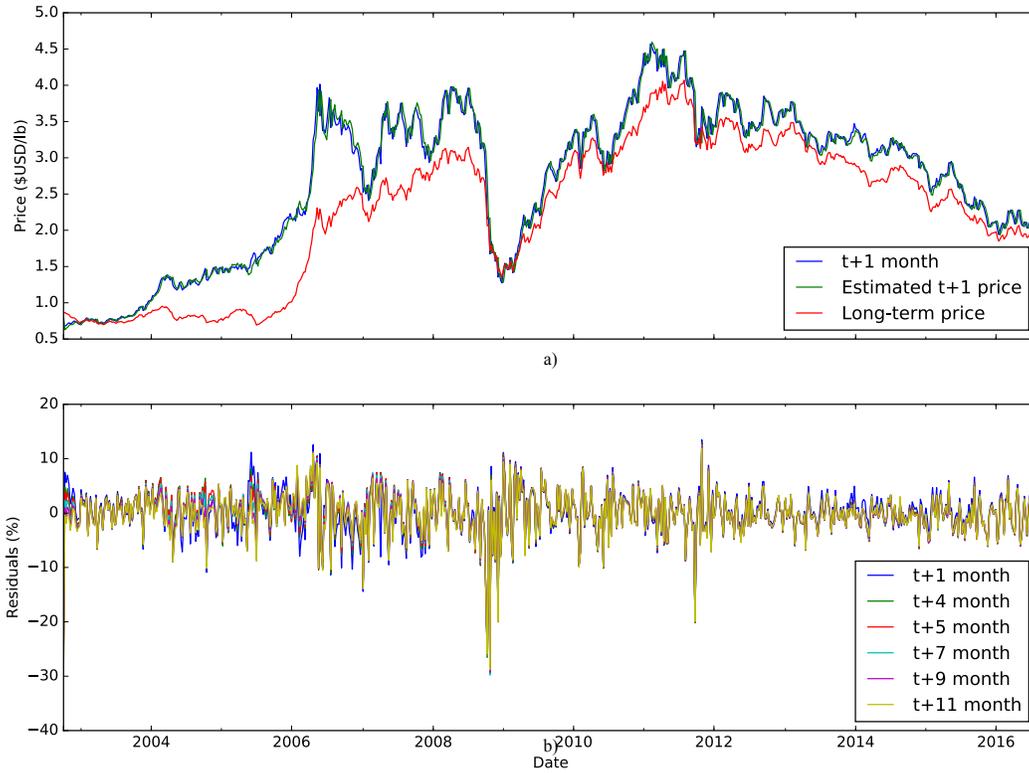


Figure 5–6: Plot a) of the Figure represents the t+1 month copper contract (in blue) as well as the KF estimate of the t+1 future contract (in green). The long-term price estimate of the KF is represented in red. Plot b) represents residuals (in %) of the KF estimates for each maturing contract.

Maturity	mean error	st. deviation of errors	MAD
1 month	0.0195	0.00021	0.0135
4 months	0.0285	0.00014	0.0198
5 months	0.0285	0.00014	0.0198
7 months	0.0235	0.00060	0.0160
9 months	0.0205	0.00023	0.0138
11 months	0.0527	0.00700	0.0415

Table 5–7: Mean error, standard deviation of errors and mean absolute deviation of the fitted maturities for the copper futures dataset.

As Goodwin (2013) noted, the equilibrium price in recent economic times varies significantly and tends to move with the spot price. Copper future prices

are more or less following a random walk, with a mean-reversion parameter κ less than 1. The half-life for the mean reversion is thus $-\ln(0.5)/0.367 = 1.88$ years. The GA performed very well to find optimal parameters in copper markets. The mean error, standard deviation of errors and mean absolute error are presented in Table 5-7.

5.4.5 NPV valuation of a copper mine

In this section the NPV of a copper mine is calculated using an active trading strategy and then being compared to the passive NPV valuation. The NPV analysis is performed on a simple mining project where the only commodity produced is copper. The project is financed using 60% equity (w_e) and 40% debt (w_d). The required rate of return is 10% for debt holders (r_d) and 16% for shareholders (r_e). The tax rate (T) is 35%. On an after-tax basis, the weighted average cost of capital ($WACC$) for the project is thus 12.2% ($WACC = w_d * r_d (1 - T) + w_e * r_e$ or $WACC = 40\% * 10\% (1 - 35\%) + 60\% * 16\%$). The project requires an initial fixed capital investment of \$175 millions that will be depreciated on a straight line basis and that will be salvaged for \$100 millions at the end of the project. The project requires an additional \$25 millions of working capital investment that will be withdrawn entirely at the end of the project as shown in Table 5-8. Cash flows are calculated on an after-tax basis. The tax shield provided by the depreciation method is added back to the after-tax net sales. To calculate the NPV, all after-tax inflows and outflows are summed and discounted at 12.2%.

Year	0	1-4	5-11	12
Fixed capital investment (FCInv)	(175)			
Salvage value of FCInv				100
Working capital investment	(25)			25
Total ore production (Mt/year)		5.5	9	
Grade (%)		0.5	0.5	
Mining costs (\$/tonne)		17	17	
Processing costs (\$/tonne)		15	15	

Table 5–8: cash flows of the mining project. The project lasts 12 years with an initial outlay of \$175 millions. Mining costs are the costs to extract the materials from the mine and processing costs are the costs to extract the metal from the ore.

The first step to perform the valuation is to simulate a population of futures time-series with parameters calibrated from the GA optimization. Figure 5–7 represents one simulation of the SSTF model. The upper section of the figure represents the simulation of the state variables. The green line represents the long-term mean and the red line represents the sum of the long-term mean and the short-term deviation. This last time series oscillates around the long-term mean because it is mean-reverting. On the middle graph, copper spot and future prices are simulated on a weekly basis. Simulated futures contracts include the t+1, t+4, t+5, t+7, t+9 and t+12 months. The last graph represents the stockpile when using an active trading strategy. For any given week, when the estimated spot price (the sum of the long-term mean and short-term deviation) is below the long-term mean by a threshold of 10%, the mining company can stockpile 10% of its weekly production. When the estimated spot price is over the long-term mean by 10%, 10% of the weekly production is withdrawn from the stockpile and sent to the smelter. If the estimated spot price is between the threshold limits for stockpiling or withdrawing from the stockpile, the stockpile remains untouched.

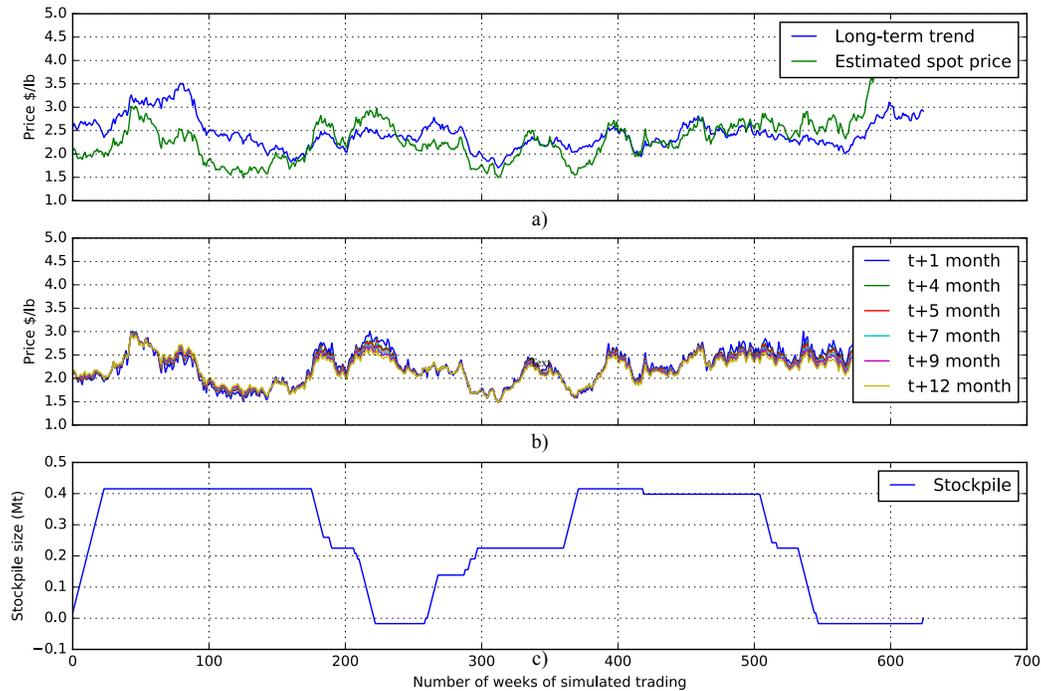


Figure 5–7: Plot a) of the Figure represents the KF estimate of the long-term trend and the estimated spot price. Plot b) represents simulation of copper contracts for the calibrated SSTF model. Plot c) represents the size of the stockpile, conditional on difference between estimated spot price and the long-term trend.

The value added from the active trading strategy is obtained by comparing it to the passive NPV valuation. The comparison is performed on a population of simulated time series. For each simulation, the NPV is obtained using both an active and passive trading strategy. For each simulation, the alpha (value added from the active strategy) is calculated as the difference between the NPV using an active trading strategy and the NPV without stockpiling. An histogram of alpha for 1,000 simulations is presented in Figure 5–8. The mean alpha for 1,000 simulations is \$2.039 millions with a standard deviation of \$5.451 millions.

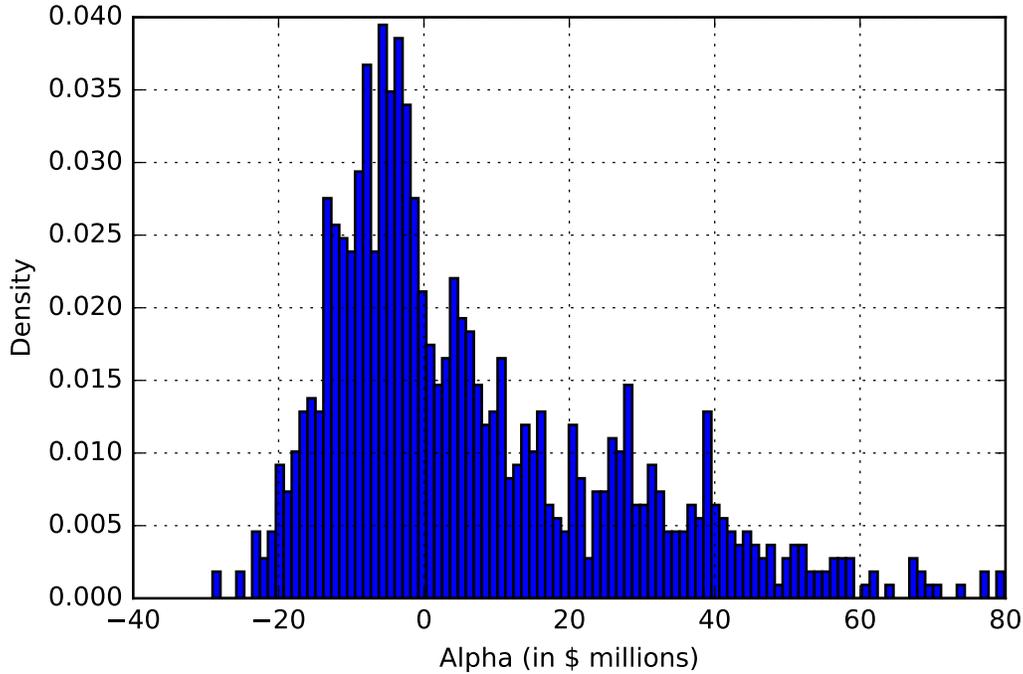


Figure 5–8: Value added (alpha) of the active trading strategy compared to the basic NPV investment. Alpha is calculated by differencing NPV using an active strategy to passive NPV for each simulation outcome.

5.5 Discussion

In this paper, a framework to calibrate the SSTF model parameters using GA is presented. Although it needs long computing time, it proves to be very useful when the objective function is noticeably non-linear with multiple dimensions. Moreover, the GA is more flexible since it can incorporate a penalty function when a parameter falls outside a feasible range. Using EM algorithms, it is not possible to apply such a constraint. EM algorithms are heavily dependent on the initial values of the model parameters to converge. Schwartz and Smith (2000a) reported that they optimized the parameters with several different starting values in order not to get caught in a local maximum. Goodwin (2013) reported that different initial parameters could yield very different optimized parameters. In the GA workflow, the process of starting with different initial parameter estimates is embedded in the algorithm, eliminating

the need to perform the workflow several times before obtaining convergence of the parameters.

With the GA, the optimality of the solution cannot be guaranteed. The algorithm is designed such that on average, at each iteration, there is a chance a newly generated individual in the population will prove to be a better candidate in terms of fitness. It is important to run the algorithm several times by varying parameters such as mutation probability and amplitude, population size and crossover probability. When the initial population is very large, there is a high probability an individual or a combination of individuals will have parameters close to the solution. In such a case, the algorithm will perform well using crossover operations since best candidates will be mated together and their siblings will have a combination of better parameters than their parents. A major drawback of having a large initial population is that the algorithm has more solutions to evaluate, rendering the process cumbersome.

On the other hand, if the initial population is small, it is important to set a high level of mutation probability (around 20% in the case of this project) so the new generation has higher levels of population diversity. However, if mutation probability is too high, the GA can diverge and give completely unreliable results even with the inclusion of a penalty function.

The fitness function used in this paper attributed equal weight to each maturity being fitted. If the modeler is more interested in the fitness of near maturity contracts, the fitness function could be modified to attribute more weight to these contracts. Also, the covariance matrix of the observed contract can be adjusted to yield KF estimates that better match any given maturity. It is important not to overfit the observed contracts.

The real strength of the workflow is to calibrate parameters of the SSTF model to historical data to retrieve parameters of the generating process. With

these parameters, it is possible to perform simulations of state-space variables and futures contracts calibrated with the observed term structure in the market. These simulations were exploited to perform an NPV valuation using an active trading strategy. The active trading strategy was compared to a passive trading strategy where the weekly production was sold at spot price without any variation to the stockpile. As shown on the histogram of differences between both strategies (Figure 5–8) this difference can be positive or negative. For example, if the long-term price is constantly increasing and the short-term deviation is below it, a large percentage of the production will be stockpiled and sold only at the end of the project. As the value of money decreases with time, proceeds from sending the stockpile to the smelter will be lower if they are received later. For 1,000 simulations, the expected value added using an active over a passive strategy is about \$2 millions in this case. The thresholds for sending ore to or get ore from the stockpile and the size of the stockpile have not been optimized. Moreover, the cost of stockpiling the ore has not been taken into account.

One of the governing parameters adding value from the active trading strategy is the speed of mean-reversion κ . When κ is low, it takes more time for the estimated spot price to mean-revert to the long-term trend. In this paper, the κ parameter calibrated on crude oil contracts was much higher than κ from copper futures (1.59 for crude oil vs 0.367 for copper in period 2). The assumption for a proper active trading strategy is that on average, the difference between the short-term deviation and the long-term mean will mean-revert rapidly to 0 when κ is high. This implies proceeds from stockpiled ore sent to the smelter will not be deferred long into the future. Parameters can be highly variable when the calibrated period of the time-series change. Moreover, even if it is calibrated on the term-structure of futures contracts,

the analysis is performed on historical data. It is assumed that the parameters derived from the GA will not change (the past will repeat itself). It is very unlikely that the model parameters will remain static.

Real Options workflows involving Kalman Filters generally rely on financial theory to solve the value of real assets contingent on commodity prices in a risk-neutral form. These workflows use binomial tree models or recursive algorithms to solve for the value of an American option on simulated risk-neutral price paths Abdel Sabour and Poulin (2006). The RO involved in these studies generally involves making significant changes to projects such as terminating or expanding them. As a result, the difference between ROV and NPV is generally significant. In this research, the active trading strategy can be performed without impacting the project significantly. In this project, the difference of \$2 millions between the ROV and NPV may seem to be very small when compared to previous ROV studies, but it does not involve additional capital investments for the expansion or closure of the project.

5.6 Conclusion

The aim of this paper was to show the robustness of GA to fit parameters of the SSTF model and use simulations of futures contracts to test an active trading strategy. An artificial dataset with known parameters and hidden state variables was created to test the algorithm. The DEAP library was used to assess the use of this optimization method. Since DEAP is implemented in Python, it might not be as fast as other genetic algorithms written in say, C (Fortin et al., 2012). Another possible improvement in this research could be to parallelize the GA to improve the execution speed of the algorithm. GAs are well suited to be parallelized and already implemented in the DEAP library. Also, several extensions of the Kalman Filter used in this research exist. These extensions can take into account nonlinearity in the transition

equations or relax the assumptions of non-Gaussianity in the residual term. For example, the Extended Kalman Filter or the Particle Filter could yield to better parameter estimation without having to rely on GAs. For future works, Particle Swarm Optimization (PSO), another family of algorithms similar to GA in the sense that they work on a population of points at the same iteration, could be tested. These algorithms have the tendency to be faster than GA while yielding similar results (James et al., 2013). On the assumption that the GA captures the true parameters generating the futures contracts term structure, the simulations of spot and future prices were made to assess the NPV of mining projects. Since the discrepancy between the estimated spot price and the long-term trend is expected to mean revert to 0, the active strategy consisting of stockpiling the mined ore when prices are low and processing it when prices are higher proves to be effective. The proposed approach can be used to perform NPV with a simple active trading strategy based on the SSTF model.

Acknowledgements

The authors gratefully thank NSERC for the support (Project No: 236482).

5.7 Chapter Conclusion

In this chapter, the genetic algorithm for the calibration of the Schwartz-Smith two-factor model was presented. Since the transition from a solution to another is done using a statistical method, the optimality of the solution is not guaranteed. However, the main advantage of the algorithm is that it reduces the chance the optimization process gets stuck in a local minimum.

The algorithm was used to calibrate a stochastic process on historical observations of gold prices. Using the calibrated parameters, an active trading strategy using mean-reversion in the short term and the Geometric Brownian Motion in the long run was used to implement an active stockpile management strategy. However, such model does not account for jump components or stochastic volatility sometimes observed in commodity markets. The next chapter presents another set of stochastic processes that are able to cope with outliers or stochastic volatility. Namely these stochastic processes are the Heston stochastic volatility process and the Merton jump process.

To be able to calibrate these processes in a Kalman filtering framework, it is necessary to use another version of the Kalman filter called the Unscented Kalman filter. The Kalman filter relies on the normality assumption for the distribution of the error term while the unscented Kalman filter uses an approximation. This improves the accuracy of the filter, especially when the dataset contains outliers. The next chapter will introduce the Unscented Kalman filter as well as the particle swarm optimizer to calculate the Cash-Flow at risk of a gold mining project.

CHAPTER 6

Cash-flow at risk valuation of mining project using Monte-Carlo simulations calibrated on historical data

6.1 Abstract

Mining projects are subject to multiple sources of market uncertainties such as metal price, exchange rates and their volatilities. Assessing a mining project exposure to market risk usually requires Monte-Carlo simulations to capture a range of probable outcomes. The probability of a major loss is extracted from the probability density function of simulated prices at a given time into the future. This paper proposes an approach to calibrate the stochastic process to be used in Monte-Carlo simulations. The simulations are then used for measuring the cash-flow at risk of a mining project. To assess the performance of the proposed approach, a case study is conducted on a mining project. The results show that the calibration approach is robust and apt at fitting various stochastic processes to historical observations.

6.2 Introduction

Mining projects are exposed to significant financial risks. Since the produced commodity is traded in markets that are affected by supply and demand dynamics, mining companies are subject to market risk. Commodity price and its volatility are among the most important sources of risk. For example, gold price, which was approximately 1,800 \$/ounce in 2011, is approximately 1,150 \$/ounce at the end of 2016. The price mechanisms of commodities are complex because many commodities and their derivatives are traded in spot and futures markets. They are also indirectly related to capital markets because mining companies usually use equity markets for financing purposes.

Because commodities are generally traded in US dollars, exchange rates fluctuations also affects mining projects operating outside of the US. The joint effect of price and exchange rates dynamics can affect the net present value (NPV) of a mining project (?). With the assumption that financial markets are efficient, exogenous information constantly changes the expectations of parties involved in the transactions. These market participants can be short hedgers such as mining companies who seek to sell their products in advance to prevent losses from sudden drop in the price of the metal they produce. The other party of the transaction can be a long hedger who seeks to fix a ceiling on its production inputs. Other stakeholders such as arbitrageurs and speculators try to profit from market inefficiencies. When the market is unfavourable for short hedgers, they need to pay a premium (sell their production at a higher discount). Because this can be very expensive, mining companies cannot hedge the entirety of their production. They need to correctly measure their risk exposure to market fluctuations and hedge the right amount of their production depending on the risk they are willing to take.

In the recent literature, the KF has been used to value real options in engineering projects. Real options generally add value to engineering projects but since they are complex to evaluate and interpret, they are not widely used (Block, 2007). Real options are generally used before the project actually starts, to verify that the added flexibility increases the NPV of a project. This paper proposes a different use of the Kalman filter (KF) calibration workflow for market risk management. Three different stochastic processes are used to model dynamics of gold, silver, platinum and crude oil prices. These processes are the geometric Brownian motion (GBM), the Merton jump diffusion (MJD) and the Heston Stochastic volatility (SV) models. The GBM

and MJD processes are calibrated using a maximum likelihood estimation routine. The Heston SV is calibrated using the Unscented Kalman filter. To find the optimal parameters of the Heston SV model, a sub-optimal solution is first found using particle swarm optimization metaheuristics approach (Eberhart and Kennedy, 1995), from which the optimal solution is found using a gradient-based method. Using calibrated parameters from historical observations, simulations of gold prices are realized over a three-month period to calculate the Cash-Flow at Risk (CFaR) of a gold-mining project. Results show that the MJD and Heston SV models are better suited to reproduce tail behaviour of commodity returns.

The originality of this paper lies on the derivation of a robust calibration framework using a hybrid metaheuristics optimization approach and the KF to perform CFaR analysis of a mining project. In this research, price volatility in Heston SV model and jump component in MJD model are calibrated through a combined approach using particle swarm optimization and Kalman filter. The most significant advantage of the paper is that unlike real options valuation, the proposed methodology can be used at any stage of a mining project to assess how cash flows are exposed to market risk in the near future. This gives the opportunity to better assess the exposure to excessive cash-flows drawdown and thus, the ability to implement hedging strategies when markets are stable and prices of derivatives relatively low.

6.3 Literature review

The Value at Risk has been used by risk managers over the last decade (Jorion, 2001). The approach consists of determining the distribution of profit and losses for an investment over a given time horizon. Then, the lower tail of the distribution is analyzed. The threshold for the lower-tail can be set, for example, to 5%. The value corresponding to the 5th percentile of the profit

and losses distribution and can be interpreted as the minimum loss that an investment will suffer 5% of the time. The Value at Risk approach is useful to infer risk measures on a portfolio of assets. In the case of mining investments, another approach is to use the cash-flow at risk measure of risk (Alesii, 2003). Because the fixed capital costs of mining projects are sunk and therefore very illiquid, the CFaR provides a more useful metric for valuing the exposure to market risk. This is because mining companies suffer when cash flows become negative after a sudden drop in commodity prices. This approach can also be implemented in a Real Options valuation framework. The MC simulations approach can be used to calculate the CFaR. This approach is based on the simulation of multiple price paths using a stochastic process such as the geometric Brownian motion. GBM is the most widely used stochastic process to model price dynamics of stocks, commodities or options contracts. Because of its simplicity, it is generally used as a reference to benchmark other stochastic processes.

The logarithm of returns on commodity futures are often positively skewed with a significant amount of excess kurtosis (Gorton and Rouwenhorst, 2006). This implies that the distribution of commodity logarithm of returns can be better described with a fat-tailed distribution. Schöne (2014) described two mechanisms contributing to the fat-tailed behaviour of commodity returns. The first mechanism is SV and the second is the inclusion of jumps in prices. The Heston SV (Heston, 1993) model can be used to describe behaviour of volatility clusters in periods of high and low volatility. The MJD model (Merton, 1976) adds a jump component to the stochastic process. Schöne (2014) studied how the choice of the stochastic process in a real options valuation framework can influence the NPV of a mining project. The author calibrated the GBM, SV and MJD models to commodity spot prices by minimizing the

square difference between the probability density function (PDF) of models and the kernel density estimated PDF of the observations. Hammond and Bickel (2013) studied how the choice of the stochastic process can affect the NPV rankings of mining projects with embedded real options.

One of the main tasks of stochastic volatility modeling is to calibrate the model with actual market observations. Since they were primarily developed for the pricing of options on financial assets, such models are calibrated on observed option prices (Gatheral, 2011). In this framework, the goal is to adjust parameters of the model to fit the actual implied volatility surface of all traded options for a given security. However, when modeling stochastic volatility in commodity markets, historical options on commodity spot or futures may be hard to find. Another class of models instead uses historical future contracts observations with the Kalman filter to calibrate the model parameters. For example, Schwartz (1997b) used the KF to describe the evolution of commodities spot price subject to a stochastic convenience yield. A two-factor commodity model consisting of a short-term price deviation and a long-term trend was later developed (Schwartz and Smith, 2000b). The authors used the KF to adjust model parameters on historical observations of commodity future prices. Hammond and Bickel (2013) showed how the choice of the stochastic process can affect the NPV ranking of oil investments. The authors consider a simple case with no options and the case where real options are embedded in the NPV analysis.

To use the KF, the stochastic process is first derived in a state-space form. One of the main assumptions of the KF is that the equations describing the transition between different observations are linear. Javaheri et al. (2003) showed that the Unscented Kalman filter and the particle filter are better suited when the transition equation of a stochastic process such as the Heston

SV model is non-linear. Schwartz and Smith (2000b) used a quasi maximum-likelihood routine to find the optimal parameters of the model. The routine uses a gradient-based optimization method to derive the optimal parameters of the model. When the objective (or cost) function have several parameters to adjust, gradient-based optimization may quickly get stuck in a local maxima (Sauvageau and Kumral, 2016). Alternate initial value vectors can converge onto different optimized parameters. Schwartz and Smith (2000b) observed a similar issue in the two-factor model. The authors used several different initial values to prevent the optimization routine from not getting stuck in a local minimum. A better approach is to use metaheuristic algorithms to ensure a robust convergence when problems are non-linear (Subulan et al., 0).

6.4 Methodology

6.4.1 Merton jump-diffusion models

Merton (1976) extended the GBM to account for random price discontinuities or jumps. The frequency of such jumps is modeled with a Poisson process which is independent from the GBM diffusion process. Using the notation of Remillard (2013), the MJD can be represented with the following equation where S_0 is the spot price at initiation, and $S(t)$ is the expected price at any given point in time:

$$S(t) = S_0 e^{X(t)} \tag{6.1}$$

with

$$X(t) = \left(\mu - \lambda\kappa - \frac{\sigma^2}{2}\right)t + \sigma W(t) + \sum_{j=1}^{N(t)} \xi_j \tag{6.2}$$

where μ is the drift parameter, λ is the intensity of the Poisson process, κ is the expected value of the jump, σ is the annualized standard deviation, $W(t)$

is a standard Brownian motion and $\sum_{j=1}^{N(t)} \xi_j$ is the compound Poisson process with $\xi_j \sim N(\gamma, \delta^2)$. When the jump process is Gaussian, the κ parameter is:

$$\kappa = e^{\gamma + \frac{\delta^2}{2}} - 1 \quad (6.3)$$

These equations can be used to simulate the MJD model. However, in the valuation or risk management workflows, the model needs to be calibrated with historical observations. The maximum likelihood (ML) method can be used to perform such calibration (Remillard, 2013). The probability density function (PDF) of price returns can be represented with the following equation:

$$f_{R_i}(r) = \sum_{k=0}^{\infty} e^{-\lambda h} \frac{(\lambda h)^k}{k!} \frac{e^{-\frac{1}{2} \frac{(r-a-k\gamma)^2}{\sigma^2 h + k\delta^2}}}{\sqrt{2\pi(\sigma^2 h + k\delta^2)}} \quad (6.4)$$

To perform the ML optimization of the parameters, the experimental PDF is approximated with a kernel density estimator. Then, the parameter set $\theta = (\mu, \sigma, \lambda, \gamma, \delta)$ is optimized through a gradient based minimization routine in MATLAB. For example, the `fminunc` (find minimum of unconstrained multivariable function) in MATLAB can be used.

6.4.2 Stochastic volatility models

Another class of stochastic models capable of reproducing fat tails is the Heston SV model. The evolution of spot price in this model can be represented with the following set of equations (Gatheral, 2011):

$$\begin{aligned} dS(t) &= \mu S(t)dt + \sqrt{V(t)}S(t)dW_1 \\ dV(t) &= \kappa(\theta - V(t))dt + \eta\sqrt{V(t)}dW_2 \\ E[dW_1 dW_2] &= \rho dt \end{aligned} \quad (6.5)$$

Where μ is the drift, κ is the speed of mean reversion, θ is the long-term mean of volatility, η is the volatility of the volatility parameter and ρ is the correlation between the two standard Brownian motions dW_1 and dW_2 .

The calibration of the SV model to historical observations is a non-trivial task. Javaheri et al. (2003) proposed the use of the UKF to perform the calibration. This algorithm is suitable when the transition equation is non-linear as in the SV model. A Gaussian approximation is made using sigma points. The weight assigned to each sigma point is defined as:

$$W_0^{(m)} = \frac{\lambda}{n_a + \lambda} \quad (6.6)$$

and

$$W_0^{(c)} = \frac{\lambda}{n_a + \lambda} + (1 - \alpha^2 + \beta) \quad (6.7)$$

and for $i = 1 \dots 2n_a$

$$W_i^{(m)} = W_i^{(c)} = \frac{\lambda}{2(n_a + \lambda)} \quad (6.8)$$

where λ , α , β are tuning parameters for the sigma points and n_a is the dimension of the augmented state. In the algorithm, $\lambda = \alpha^2 n_a - n_a$. In this paper, the parameters are $\alpha = 0.001$, $\beta = 2$ which are optimal for a Gaussian approximation. Then, the sigma points are passed through the non-linear transition equation of the KF and a Gaussian distribution is approximated through the transformed points. The rest of the UKF algorithm is presented in the Appendix.

The parameters of the UKF $\alpha = (\mu, \kappa, \theta, \eta, \rho)$ are estimated through the quasi maximum likelihood (QML) method. The QML estimate is given by:

$$L_{1:N} = \sum_{k=1}^N \ln(P_{z_k z_k}) + \frac{z_k - m_{z_k}}{P_{z_k z_k}} \quad (6.9)$$

where z_k and $P_{z_k z_k}$ are defined in the Appendix.

To calibrate the Heston SV model, a first sub-optimal solution is found using the PSO routine (Eberhart and Kennedy, 1995). Instead of working with a single vector of initial parameters, the PSO tests a population of vectors. Each vector corresponds to a single particle which is allowed to move on the search space at a given speed. Then, at each iteration, the algorithm searches for a better solution without using the partial derivatives of the cost function. The particles are allowed to move at a given velocity given by:

$$v_{k+1}^i = \omega_k v_k^i + c_1 r_1 (P_k^i - x_k^i) + c_2 r_2 (P_k^{Global} - x_k^i) \quad (6.10)$$

where ω is the inertia weight of the particles, r_1 and r_2 are random draws from the uniform distribution, and (c_1, c_2) are the acceleration coefficients of the particles. The algorithm is stopped after a number of iterations selected for convergence criteria and the sub-optimal solution is used as the starting values in the gradient-based optimization routine. The routine is presented in Algorithm 2. Then, as with the MJD model, the set of parameters is optimized through the gradient-based algorithm in MATLAB.

Initialize a particle swarm with random values positions and velocities
 from D dimensions in the search space;

```

while number of iterations  $k < N$  do
  | for Each particle  $i$  do
  | | Evaluate objective function  $f(x_k^i, y_k^i)$ ;
  | | Update the particle best solution  $P_k^i$ ;
  | | Update the swarm best solution  $P_k^{Global}$ ;
  | | Update the velocity with Equation 6.10;
  | | Update the position of all particles using  $x_{k+1}^i = x_k^i + v_{k+1}^i$ ;
  | end
end

```

Algorithm 2: PSO algorithm

6.4.3 Estimation of CFaR

Unlike financial institutions, where VaR measure is the norm to infer market risk, mining companies own physical assets and are dependent on commodity prices. In general, their investment cannot be liquidated without any major loss. For this reason, the risk metric studied in this paper is the CFaR. It can be interpreted as the minimum cash-flow from operations (CFO) that the mining project can earn for a given period. When the risk of negative cash flows is high, mining companies seek ways to hedge the market risk by using commodity derivatives such as futures contracts.

Different approaches can be used to calculate CFaR. The historical approach consists of fitting a statistical distribution to historical observations and determining the 5th percentile of the cumulative density function (CDF). The drawback of this approach is that it is unlikely the past will repeat itself and the true risk could be underestimated. Another approach supposes that

the distribution of price returns can be approximated with a normal assumption. The CFaR is extracted from a PDF fitted to the distribution. However, if the production is hedged with derivatives contracts, the CFaR distribution becomes non-linear and becomes very difficult to fit with an analytical method. In this paper, MC simulations are used to calculate the CFaR. The parameters inferred from calibrated stochastic processes are used to perform MC simulations of price paths. The simulation parameters are calibrated on historical data but they can be modified to account for econometric predictions or can be stress-tested to estimate the error on the CFaR predicted value. CFaR calculations are performed in a real-world P probability framework. The stochastic processes of interest are the GBM, the MJD and the Heston SV.

To calculate the CFO, the following function is used, with stochastic gold prices from the GBM, MJD and Heston SV models:

$$\begin{aligned}
 CFO = & [throughput * price * grade * recovery] - \\
 & [mining costs * (throughput + waste production)] - \\
 & [processing costs * throughput] - Fixed costs
 \end{aligned} \tag{6.11}$$

Then, the different CFO are aggregated on daily, weekly, monthly and quarterly periods, and ranked. The CFO corresponding to the 5th percentile is taken as the calculated CFaR.

6.5 Case study

6.5.1 Data

The dataset used to perform the analysis consists of gold, silver and platinum spot prices as well as a continuous series of the CME group crude oil nearest maturing contracts. Historical observations of gold and silver spot prices are from the London Bullion Market Association (LBMA). The LBMA fixes

gold prices twice a day by matching bid and ask prices of buyers and sellers of the commodity. The spot platinum dataset is from The Johnson Matthey Base Prices dataset consisting of the company’s quoted selling prices. The crude oil contracts are taken from the Chicago Mercantile Exchange (CME) group database and consists of raw assembled contracts without adjustment. Figure 6–1 presents the time series for the four datasets ranging between 8 February 1993 and 8 November 2016. The time period includes the 2009 financial crisis. The dataset was split in two different periods. The first period corresponds to the pre-crisis era while the second corresponds to the price dynamics after the crisis.

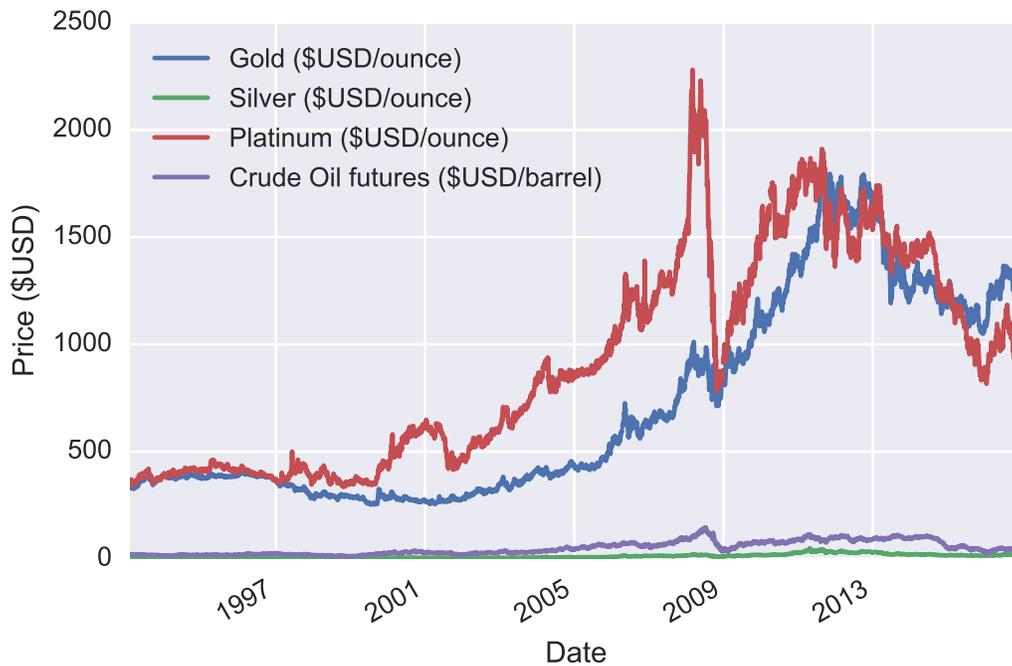


Figure 6–1: Price levels for gold, silver, platinum (in \$USD/ounce) and crude oil futures (in \$USD/barrel) during 8 February 1993 and 8 November 2016.

To calibrate the models, price needs to be converted to logarithm of returns. Figure 6–2 represents respectively log returns and spot price of gold in two different periods. The first period ranges between 8 February 1993 and 27 December 2004. The second period ranges between 29 December 2004 and 8

November 2016. It is worth noticing how volatility tends to cluster on the first plot of Figure 6–2. This characteristic can be observed in both periods, but the volatility tends to be higher in the second period. The same phenomenon can be observed for the rest of the dataset.

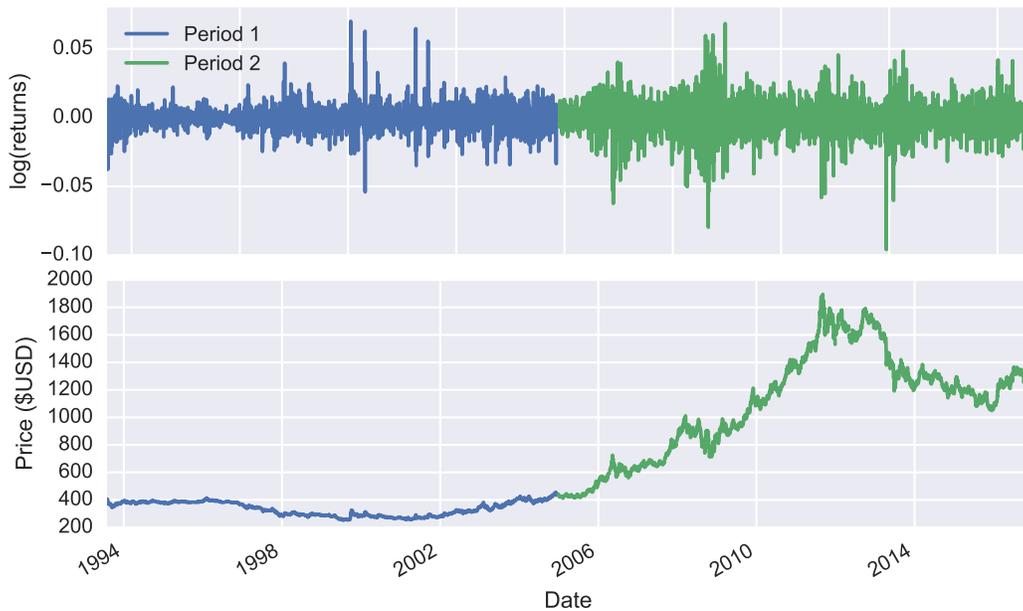


Figure 6–2: Logarithm of returns of gold spot prices during 8 February 1993 and 8 November 2016. The first period ranges between 8 February 1993 and 27 December 2004. The second period ranges between 29 December 2004 and 8 November 2016.

In Table 6–1 descriptive statistics of the distribution of commodity returns in the three different periods are presented. The returns are calculated using daily observations but the mean and variance are annualized so they are easier to interpret. In all cases, the distribution of commodities price returns are characterized by an excess kurtosis which is a characteristic of fat-tailed distributions. The positive skewness of gold and platinum spot returns in Period 1 implies that there is more weight in the right tail than the left tail. However, in Period 2, this characteristic is inverted and both distributions

become negatively skewed. Finally, the minimum and maximum values are generally higher in the second period.

Another way to detect extreme values or analyze the dispersion of the distribution is to use a box plot. Figure 6–3 shows box plots for the four commodities in the two different periods. The central part of the box plot is limited by the median and the 25th and 75th percentiles. The whiskers extend to the 10th and 90th percentiles. The data points that are outside the whisker range are considered as outliers. Both plots share the same y-axis. It is easy to see that both periods contain a considerable amount of outlier points and that Period 2 has more outliers than Period 1. Again, Figure 6–3 shows that the silver spot price returns dataset contains the most extensive outliers.

Table 6–1: Descriptive statistics for the gold, silver and platinum spot price returns as well as the crude oil futures contracts returns. The period 1 covers the weeks between 8 February 1993 and 27 December 2004. The second period covers the weeks between 10 January 2005 and 5 December 2016. The total period covers the whole time range.

		counts	ann. mean	ann. std	skewness	kurtosis	min	max
Period 1	gold	2913	0.0191	0.1271	0.6277	8.1926	-0.054	0.0701
	silver	2945	0.0376	0.2536	-0.0829	4.6821	-0.112	0.0991
	platinum	3026	0.0511	0.1952	0.4949	9.011	-0.0741	0.1393
	crude oil	2895	0.0687	0.3579	-0.3326	3.7277	-0.1654	0.1423
Period 2	gold	2907	0.0389	0.1929	-0.4234	4.8751	-0.096	0.0684
	silver	2940	0.0417	0.3609	-0.5557	8.8629	-0.1869	0.1828
	platinum	2878	-0.0161	0.2435	-0.7157	9.9826	-0.1554	0.1264
	crude oil	2928	0.0256	0.3944	0.0669	5.2571	-0.1576	0.1641
Total period	gold	5820	0.029	0.1633	-0.1959	6.7039	-0.096	0.0701
	silver	5885	0.0396	0.3118	-0.4522	9.3112	-0.1869	0.1828
	platinum	5904	0.0184	0.22	-0.3018	10.3185	-0.1554	0.1393
	crude oil	5823	0.0471	0.3767	-0.1042	4.7122	-0.1654	0.1641

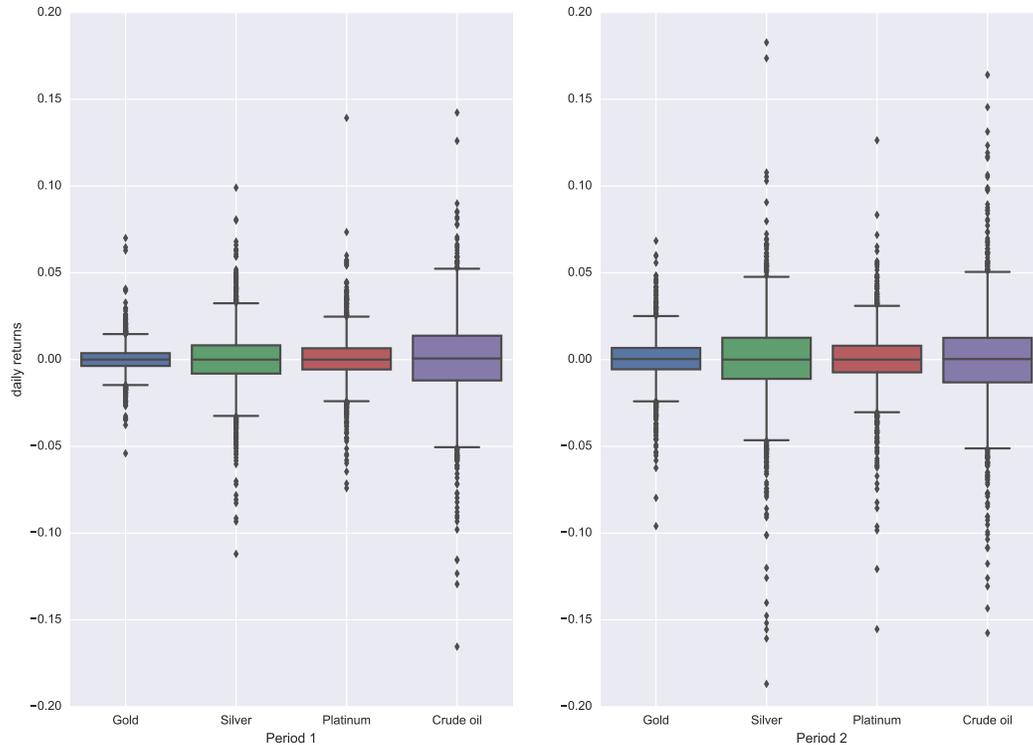


Figure 6-3: Boxplots for gold, silver, platinum and crude oil futures price returns. The points are considered as outliers when the distribution is assumed to be normal.

Figure 6-4 shows the normalized histograms of silver price returns in both periods. The dashed line corresponds to the normal distribution with the same mean and standard deviation presented in Table 6-1. The continuous blue line shows a probability density function estimated with a kernel density estimator (KDE). The KDE is very useful to fit any kind of distribution nothing that the Gaussian kernel has only a bandwidth parameter to adjust; the normal distribution was the appropriate function to describe silver spot returns, both lines would overlay. In Period 2, silver spot returns are also negatively skewed when compared to Period 1. The normal distribution is symmetric and should have a skewness of 0. Finally, the Jarque-Bera (Jarque and Bera, 1980) statistic test was performed on the commodity spot price returns (Table 6-2). This test statistic follows a Chi-square distribution with two degrees of freedom

and is reliable when the sample contains at least 2,000 observations. The null hypothesis is that the observations are normally distributed. This hypothesis cannot be rejected when the test statistics are below the critical value. The results presented in Table 6–2 shows sub-optimal that the normality hypothesis is rejected at the 99% level of confidence.

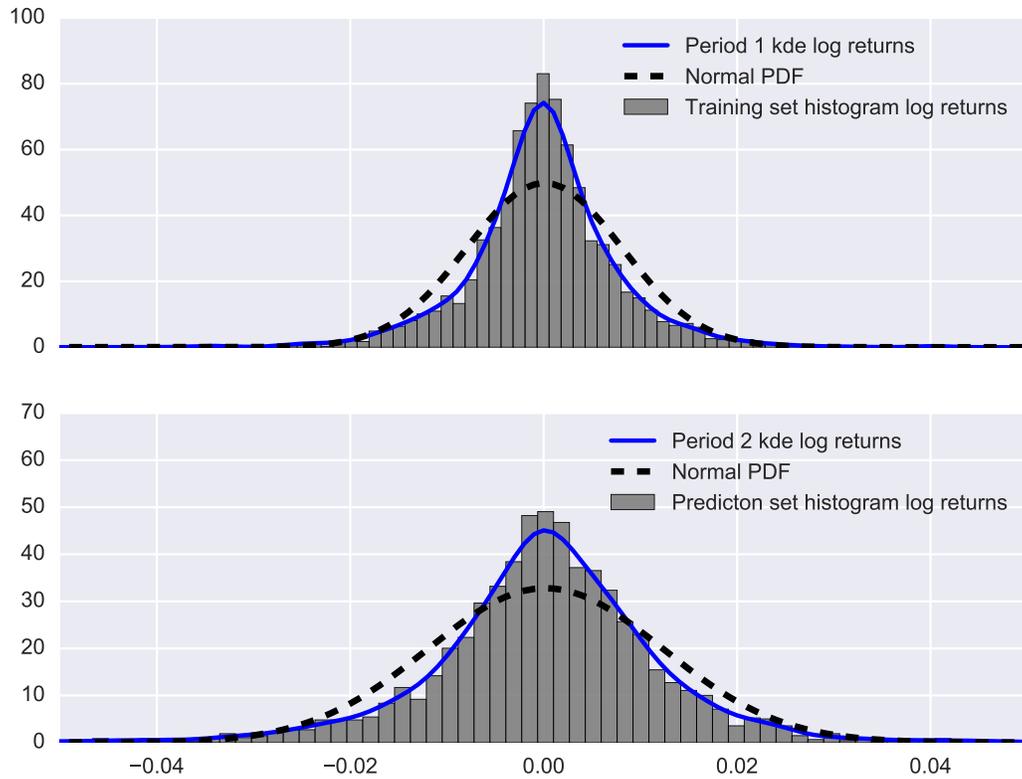


Figure 6–4: Histogram of spot gold price returns. The dashed PDF corresponds to a normal distribution with the same mean and variance as observed in the empirical dataset. The continuous PDF is calculated with a kernel density estimator.

Table 6–2: Jarque-Bera statistics for normality on the 99% level of confidence for the four different commodities spot price returns.

Commodity spot price returns	Test statistic	Critical value (99% conf. level)
Gold	8,440.23	9.21
Silver	2,697.13	9.21
Platinum	10,353.91	9.21
Crude oil futures	1,573.46	9.21

6.5.2 Model calibration

The calibrated parameters and the estimation of their error for the GBM, the SV and the MJD models are presented in Table 6–3. To reduce the likelihood of getting stuck in a local minimum, several starting parameters have been tested using PSO. After 15 iterations, PSO is stopped and the sub-optimal best solution in the population becomes the starting values for a gradient-based optimization approach. Also, to ensure the calibrated parameters are realistic, a change of variable was performed in the minimization routine. For example, volatility parameters need to be positive and correlation needs to be between -1 and 1. For volatility parameters, the logarithm of volatility is parsed in the minimization routine and converted back taking its exponential before evaluating the objective function. This ensures that the search space will always contain positive values for volatility. A similar approach is used with the correlation parameter using a tangent and its inverse function. Then, the numerical Jacobian method is used to calculate the error on calibrated parameters (Remillard, 2013). This ensures the Fisher information matrix is always estimated with a positive-definite Hessian matrix. Finally, the Feller condition ($2\kappa\theta > \eta$) is applied on the SV model to ensure the process never reaches 0 (Albrecher et al., 2006). The results presented in Table 6–3.

Table 6–3: Calibrated parameters of the GBM, Heston SV and MJD on gold, silver platinum and crude oil between 15 February 1993 and 5 December 2016.

Param	Gold		Silver		Platinum		Crude oil		
	est.	std	est.	std	est.	std	est.	std	
GBM	μ	0.0540	0.0079	0.0729	0.0065	0.0475	0.0046	0.0790	0.0079
	σ	0.1656	0.0020	0.3138	0.0012	0.2223	0.0008	0.3831	0.0020
Heston SV	μ	0.1412	0.0332	0.0563	0.0574	0.0783	0.0361	0.0509	0.0632
	κ	1.5008	0.1609	1.7396	0.0245	1.467	0.0787	1.435	0.1584
	θ	0.031	0.0120	0.0947	0.0100	0.0506	0.0141	0.1452	0.0224
	η	0.3184	0.0529	0.3248	0.0707	0.4665	0.0424	0.3729	0.0361
	ρ	0.2075	0.0350	0.2791	0.0520	0.0431	0.0200	-0.4232	0.0346
Merton Jump	μ	0.0603	0.0659	0.0957	0.0629	0.059	0.0445	0.1151	0.0789
	σ	0.0764	0.0025	0.2052	0.0040	0.0028	0.0076	0.2793	0.0055
	λ	155.87	11.51	56.45	5.6626	59.04	5.6782	45.33	5.97
	γ	-0.0002	0.0002	-0.0018	0.0011	-0.0006	0.0007	-0.0026	0.0016
	δ	0.0114	0.0003	0.0305	0.0012	0.0215	0.0008	0.0386	0.0019

The GBM is the simplest model having only two parameters to adjust. The drift μ can be interpreted as the annual continuously compound return when investing in the commodity. All commodities exhibit positive drift parameters. Crude oil futures exhibit the highest level of volatility while gold prices are less volatile. The volatility of the MJD model is lower than in the GBM but the real volatility of the process is the sum of the GBM and the Poisson process. The jump mean η is negative for the four commodities, but very close to zero. This means that the jump processes are centered around zero and do not significantly affect the skewness of the returns distribution. The λ affects the kurtosis of the distribution by increasing the occurrence of returns in the tails of the distribution.

In the Heston SV model, the η parameter determines the volatility of the stochastic volatility process while the long-term trend θ represents its mean reverting level. When the volatility level is above the long-term mean, it is expected to drift to lower values with a speed of mean reversion κ . The initial

volatility plays a crucial role on the outcome of simulated results. When this value is far from the mean reversion level, simulated paths will exhibit a transition from the initial volatility to the mean-reverting level. This value was estimated using a rolling standard deviation window on the whole sample and taking the standard deviation corresponding to the average of the first 10% of observations. The correlation determines the relationship between price and volatility levels. When the correlation coefficient is positive, a higher level of volatility will cause price returns to move in the same direction. Correlation is positive for gold and silver, neutral for platinum and strongly negative for crude oil futures.

To ensure the calibrated parameters fit the past observations and are also useful for out-of-sample predictions, the root mean square errors have been calculated for the GBM and Heston SV models. The values are presented in Table 6–4. The dataset was split into two different partitions. The first partition is used for the calibration of the KF parameters and consists of 80% of the observations. The second partition is out-of-sample and consists of the remaining 20% observations. Out-of-sample values are systematically lower than in-sample values. This is due to the nature of the data, which contains many more outliers and volatility clusters in the in-sample period. The root-mean-square error (RMSE) is almost identical for the GBM and Heston SV models. The UKF performs a linear correction on a non-linear problem. In this case study, it does not improve the RMSE when compared to the GBM.

Table 6–4: In-sample and out-of-sample RMSE for the different commodity datasets. The KF is calibrated using 80% of the first observations. The out-of-sample RMSE uses the calibrated parameters on the remaining 20% of the dataset.

	GBM		Heston SV	
	in-sample	out-of-sample	in-sample	out-of-sample
Gold	0.0104	0.0105	0.0104	0.0105
Silver	0.0203	0.0169	0.0204	0.0169
Platinum	0.0144	0.0124	0.0144	0.0124
Crude oil	0.0246	0.0223	0.0245	0.0224

6.5.3 CFaR analysis

This paper assesses the CFaR of an open-pit gold mine. The weekly production of the mine is presented in Table 6–5. To obtain the CFaR value, it is necessary to first calculate the cash flows from operations (CFO) for each trading period. Since the objective of this paper is to assess the effect of the stochastic process on CFaR calculations, the only random variable is the gold price.

Table 6–5: Characteristics of the open-pit gold mine. These parameters are used for the CFaR calculations.

Parameter	Value
Mine type:	Open pit
Processing method:	Grinding and flotation
Processing capacity:	180,000 tonne/week
Throughput:	140,000 tonne/week
Waste production:	180,000 tonne/week
Average grade:	2.6 g/tonne
Mining cost:	3.2 \$/tonne
Processing cost:	12.5 \$/tonne
Gold price:	Simulated \$/gr
Gold recovery:	84%
Fixed costs:	1,500,00 \$/week

Using the parameters of Table 6–5 and Equation 6.11, the CFO with a gold price of 42\$/g is:

$$\begin{aligned}
CFO &= 140,000 * 42 * 2.6 * 0.84 - \\
& [3.2 * (140,000 + 180,000)] - \\
& [12.5 * 140,000] - 1,500,000 = \$8,567,920
\end{aligned}
\tag{6.12}$$

The daily simulations of gold prices are performed using the GBM, the MJD and the Heston SV models calibrated with historical data. The CFaRs are calculated for a daily, a weekly, a monthly and a quarterly interval. First, 5,000 simulations of gold prices are performed using each model, and then the prices are used as an input in Equation 6.11. The calculated CFaR for each period are presented in Table 6–6. The results show the simulated cash-flows corresponding to the 5th percentile of the simulated cash-flow distribution. The first thing to notice is that the GBM model systematically undervalues CFaR for each time period. Also, the longer the studied timeframe is, the bigger the difference between CFaR calculated with each different methodology will be. For the quarterly CFaR calculation, the difference between the MJD and the GBM can vary between 6% and 50.5% when calculated on a quarterly basis while the differences range between 1.9% and 6.1% on a daily basis. Although CFaR values are all positive, they may not meet the hurdle rate required to satisfy shareholders of the company operating the project. If these cash flows were negative, the company might even need to dilute its equity with a secondary offering at an unfavorable market price of shares to prevent a default payment.

Table 6–6: CFaR of the gold-mining project for quarterly, monthly, weekly and daily time intervals

	Quarterly CFaR	Monthly CFaR	Weekly CFaR	Daily CFaR
GBM	\$26,465,238	\$10,247,064	\$2,189,013	\$544,234
MJD	\$17,577,417	\$ 8,671,698	\$2,070,152	\$512,755
Heston SV	\$24,739,856	\$9,819,090	\$2,152,277	\$535,176

6.6 Discussion

Most mining companies use derivative instruments on commodities or exchange rates to reduce their exposure to market risk (Armstrong et al., 2009). To correctly assess the exposure to market risk, stochastic models are used to describe the price dynamics of commodities. Model risk emerges when the model used to describe commodity price dynamics is misused. This can be due to a violation of the assumptions of the model or a bad calibration. The GBM assumes that commodity price returns are normally distributed. It is clear that commodity price returns are leptokurtic and the model used to simulate price paths calibrated on historical observations should take into account stochastic volatility or jump components. To reduce model risk, different stochastic processes have been calibrated on historical data. Using the KF workflow, the models are calibrated in-sample and out-of-sample. It is important to test the model on an out-of-sample set to be sure the model is not overfitting past observations. In the KF framework, or using ML, the model parameters are calibrated on a proportion of the sample and tested out-of-sample on the rest of the observations.

The MJD and Heston SV models are more complex and are harder to calibrate than the GBM. In a robust valuation workflow, different kinds of models should be tested and benchmarked against the GBM. Different strategies may be used to calibrate models, depending on their complexity. The MJD model is a GBM with an independent jump process. The PDF of such model is straightforward and can be calibrated using ML. Another good calibration strategy could have been to model a GBM with outlier correction and then model the outliers in the jump component of the model. In the SV model, the diffusion process of both price and volatility are correlated. The model has five parameters to calibrate. The topology of the cost function is highly

non-linear and the use of a metaheuristic approach for the optimization helps to find a first good approximation of the solution (Donkor and Duffey, 2013). Then, the sub-optimal solution of the PSO is used as an input in the gradient-based optimization, the initial value is very close to the global minimum. The Jacobian and Hessian matrices may not exist in the entire domain of the cost function, but it can be approximated in the neighborhood of the sub-optimal solution. As a result, the gradient-based optimization requires fewer iterations to find the optimal solution. The main advantage of the gradient-based optimization is that it yields an estimate of the parameters error, which can be used to construct confidence intervals or assess if the parameters are reliable.

Because of the diversity of stochastic processes that can be used to model commodity price dynamics, there is an inherent lack of generality in the parameter calibration process. To solve this problem, Schöne (2014) developed a general calibration method based on the minimization of a theoretical and a kernel based PDF. In this paper, a different approach is used, and the criteria to assess how well each model compared to each other are in-sample and out-of-sample RMSE measures. This procedure ensures that no matter how the stochastic process is calibrated, it is effective at fitting past observations and is also useful for making predictions. Historical calibration converges onto the estimation of parameters that may not be forward-looking. If the calibration is performed in the recent past, the calibrated parameters will reflect market dynamics of the past period. However, it is very unlikely that the past will repeat itself indefinitely and the estimates should be readjusted to reflect the expectation of future events. This is particularly important when valuing market risk exposure.

The CFaR was calculated considering only one stochastic parameter. In general, mining projects will be affected by more than one risk factor. For

example, CFaR could be calculated with a stochastic exchange rate level, or with varying production parameters such as the grade of the ore. The CFaR of a project incorporating all of these uncertainties would be lower than the individual CFaR. For example, lower commodity prices could be countered by a more advantageous exchange rate level. Another important task would be to verify that the calculated CFaR actually reflects past observations. If the CFaR is calculated using the historical method, the CFaR will be exactly the one corresponding to historical observations. On another hand, the CFaR calculated using the workflow presented in this paper relies on MC simulations and should be back tested. One drawback of the CFaR methodology is that it yields an upper limit of the minimum cash-flow that a project will produce at a given time. In reality, the actual minimum cash-flow could be much lower. Another approach to better quantify the actual loss is the Conditional Value at Risk (CVaR) (Acerbi and Tasche, 2002). The risk metric produced with this approach can be interpreted as the expected loss assuming that the VaR threshold has been triggered. This can be extended to CFaR calculations.

6.7 Conclusion

Stochastic processes are widely used to analyze risks associated with commodity prices through MC simulations. The important question is how to find or calibrate appropriate values for the parameters of stochastic processes. This paper proposed a new hybrid approach to calibrate the parameters to be used in Heston SV and MJD models. For calibration, KF were used. Since KF requires an initial sub-optimal solution, PSO was used to generate an initial solution. This paper investigates how the choice of the stochastic process can affect the CFaR of a mining project. The risk analysis is implemented using real world probabilities. Depending on the complexity of the model, different

calibration strategies may be used to fit historical observations. The KF calibration workflow proves to be a very robust. This paper uses the UKF to model the SV in a filtering framework. The model parameters are calibrated using a hybrid metaheuristics approach using PSO and a gradient based optimization method. This approach linearizes a non-linear problem using a Gaussian approximation. As future work, another algorithm called the particle filter (PF) could be tested. However, since the PF relies on individual particles that are each transformed using the state-space transition equation, it is much more computationally intensive. The choice of the stochastic process affects the calculated CFaR risk measure. This is because more complex models such as the MJD and the Heston SV models are better suited for reproducing fat tails observed in empirical observations.

Acknowledgements

The authors gratefully thank NSERC for the support (Project No: 236482).

Appendix

The UKF algorithm is taken from Javaheri et al. (2003). First, the mean and covariance are initiated:

$$\begin{aligned}\hat{x}_0 &= E[x_0] \\ P_0 &= E[(x_0 - \hat{x}_0)(x_0 - \hat{x}_0)^T]\end{aligned}\tag{6.13}$$

The state vector is concatenated with the system noise and the observation noise:

$$x_{k-1}^a = \begin{bmatrix} x_{k-1} \\ w_{k-1} \\ u_{k-1} \end{bmatrix}\tag{6.14}$$

Therefore:

$$\hat{x}_{k-1}^a = E[x_{k-1}^a | z_k] = \begin{bmatrix} \hat{x}_{k-1} \\ 0 \\ 0 \end{bmatrix} \quad (6.15)$$

and

$$P_{k-1}^a = \begin{bmatrix} P_{k-1} & P_{xw}(k-1|k-1) & 0 \\ P_{xw}(k-1|k-1) & P_{ww}(k-1|k-1) & 0 \\ 0 & 0 & P_{uu}(k-1|k-1) \end{bmatrix} \quad (6.16)$$

For $i = 1 \dots n_a$, the sigma points are calculated:

$$\chi_{k-1}^a(i) = \hat{x}_{k-1}^a + \left(\sqrt{(n_a + \lambda) P_{k-1}^a} \right)_i \quad (6.17)$$

And for $i = n_a + 1 \dots 2n_a$:

$$\chi_{k-1}^a(i) = \hat{x}_{k-1}^a - \left(\sqrt{(n_a + \lambda) P_{k-1}^a} \right)_{i-n_a} \quad (6.18)$$

$$\chi_{k|k-1}(i) = f(\chi_{k-1}^x(i), \chi_{k-1}^w(i)) \quad (6.19)$$

for $i = 0 \dots 2n_a + 1$ so:

$$\hat{x}_k^- = \sum_{i=0}^{2n_a} W_i^{(m)} \chi_{k|k-1}(i) \quad (6.20)$$

and

$$P_k^- = \sum_{i=0}^{2n_a} W_i^{(c)} (\chi_{k|k-1}(i) - \hat{x}_k^-) (\chi_{k|k-1}(i) - \hat{x}_k^-)^T \quad (6.21)$$

The innovation is:

$$z_{k|k-1}(i) = h(\chi_{k-1}(i), \chi_{k-1}^u(i)) \quad (6.22)$$

then:

$$\hat{z}_k^- = \sum_{i=0}^{2n_a} W_i^{(m)} z_{k|k-1}(i) \quad (6.23)$$

and:

$$v_k = z_k - \hat{z}_k^- \quad (6.24)$$

For the measurement update:

$$P_{z_k z_k} = \sum_{i=0}^{2n_a} W_i^{(c)} (z_{k|k-1}(i) - \hat{z}_k^-)(z_{k|k-1}(i) - \hat{z}_k^-)^T \quad (6.25)$$

and

$$P_{x_k z_k} = \sum_{i=0}^{2n_a} W_i^{(c)} (\chi_{k|k-1}(i) - \hat{x}_k^-)(\chi_{k|k-1}(i) - \hat{x}_k^-)^T \quad (6.26)$$

The Kalman gain is:

$$K_k = P_{x_k z_k} P_{z_k z_k}^{-1} \quad (6.27)$$

The next estimate is:

$$\hat{x}_k = \hat{x}_k^- + K_k v_k \quad (6.28)$$

and the covariance is:

$$P_k = P_k^- - K_k P_{z_k z_k} K_k^T \quad (6.29)$$

CHAPTER 7

Conclusion

In the last decade, several researchers predicted that Real Options valuation would become as widely used as the NPV valuation approach for mining projects. In reality, it is not always clear if the classical Real Options valuation workflow is useful for mining projects. Once the investment decision is made, the capital expenditures necessary to bring a mineral field to production are very high. Even if the price of the commodity drops substantially, the mine cannot simply close and reopen later when the prices are higher. As a result, the NPV approach is almost always used for the valuation of a mining project. Moreover, the real option will always bring additional value compared to the base case NPV scenario. If a project is already profitable using the NPV criteria, the real option does not change the investment decision.

In this thesis, a different set of tools aimed at helping mining companies to increase their value or reduce their market risk exposure in their daily activities are developed. The tools are borrowed from the financial engineering field of research and are applied in the context of a mining project. Unlike the classical Real Options paradigm, these new methods do not require significant alterations in the nature of the mining operation. The contributions of this research are to provide a robust workflow to calibrate stochastic processes to historical observations. The contributions are discussed in the next paragraphs.

The first contribution was to consider the mining operation in a portfolio valuation framework. In this project, stochastic processes are calibrated on historical iron ore futures, and the production can be sold into two different

segregated markets. The first market is considered risk-free and consists of long-term commitment contracts; the second market is the spot market and is subject to price fluctuations. The riskier market has a higher expected rate of return. A mining company can decide the proportion of their production to be sold in any market depending on the risk they are willing to take. For future work, it would be interesting to study a multi-metal project. The two-factor model could be expanded to include a correlated stochastic process. Then, depending on the correlation between the different stochastic processes and a risk-free asset, a portfolio could be constructed and optimized.

The second contribution consisted of managing the stockpile of a mining project based on the differences between the short-term and long-term prices. When the weekly selling price is lower than the expected long-term level, a percentage of the production is set apart in a stockpile to be sold when prices are more advantageous. This approach is based on the mean reversion of the short-term disturbance of the price level. On average, the active trading strategy adds value to the mining project. In this project, only one active trading strategy was tested. The parameters of the strategy could be optimized to take the best decision. In financial engineering and portfolio management, multiple trading strategies exist and it would be interesting to assess if they can be implemented to value mining projects. For example, an arbitrage strategy exploiting market inefficiencies could be tested using this methodology.

The third contribution can also be used in a mining project daily activities and is particularly useful to measure the market risk of a mining project subject to extreme variations in commodity prices. The approach uses the GBM, the Merton and the Heston stochastic models to calculate the CFaR of a mining operation. Results show that the CFaR value obtained using either the Merton or the Heston models are more conservative than the GBM. The

model parameters were optimized with a hybrid metaheuristics approach. Results show that this approach is apt at finding the global minimum of the cost function. The CFaR methodology is very useful to assess the market risk of a mining project. The method can be used even when the dataset is contaminated with outliers since the Merton or Heston models can accommodate for fat-tailed distributions.

The stochastic processes relied on a Kalman filtering workflow to be calibrated on historical data. The parameters of the stochastic processes were calibrated using metaheuristics algorithms such as the genetic algorithm of particle swarm optimization. These algorithms are well suited to optimize highly nonlinear and multidimensional problems. Also, they are apt at dealing with cost functions having an irregular or discontinuous topology.

The workflow presented in this thesis requires calibration of stochastic processes using historical observations. This methodology implies the past is likely to repeat itself. However, it is always possible markets will be affected by an unprecedented event that cannot be captured by the historical simulation approach. Mining companies managers should always stay informed with major events that will likely affect commodity markets. The historical calibration workflow presented in this paper should not be seen as a substitute for sound economic analysis, but rather as a tool to assess how a mining project can be affected by different market dynamics.

When pricing options contracts, sell-side financial institutions are more interested in actual prices observed in the market. They use the term structure of futures contracts or actual option prices at different maturities to infer the properties of the stochastic model they use. The approach is often called implied volatility. And is very useful to grasp the instantaneous market sentiment regarding a financial product. However, obtaining such information

can be quite expensive and formatting the data to be interpreted in a Kalman filter workflow may prove to be a difficult task. Another method used to make the analysis more forward-looking is to incorporate information from different sources in a Bayesian framework. Using this methodology, estimates from historical observations could be used as the *a priori* that can be updated with a likelihood function. The Bayesian framework proves to be very useful to incorporate different types of information in a unified model.

References

- Abdel Sabour, S. A., Poulin, R., 2006. Valuing real capital investments using the least-squares Monte Carlo method. *The Engineering Economist* 51 (2), 141–160.
- Abdel Sabour, S. A., Poulin, R., Dec. 2010. Mine expansion decisions under uncertainty. *International Journal of Mining, Reclamation and Environment* 24 (4), 340–349.
- Acerbi, C., Tasche, D., 2002. Expected shortfall: A natural coherent alternative to value at risk. *Economic Notes* 31 (2), 379–388.
URL <http://dx.doi.org/10.1111/1468-0300.00091>
- Adams, R. G., 1991. Managing cyclical businesses. *Resources Policy* 17 (2), 100–113.
- Aiube, F. A. L., Samanez, C. P., Jul. 2014. On the comparison of Schwartz and Smith’s two- and three-factor models on commodity prices. *Applied Economics* 46 (30), 3736–3749.
- Albrecher, H., Mayer, P., Schoutens, W., Tistaert, J., 2006. The little heston trap.
- Alesii, G., 2003. Controlling cfar with real options. a univariate case study. Tech. rep., Working Paper.
- Alexander, C., 2008. Moving average models for volatility and correlation, and covariance matrices. *Handbook of Finance*.
- Andersen, T. G., Shephard, N., 2009. Stochastic volatility: Origins and overview. *Handbook of Financial Time Series* 233, 254.

- Anderson, C., Schumacker, R. E., 2003. A comparison of five robust regression methods with ordinary least squares regression: Relative efficiency, bias, and test of the null hypothesis. *Understanding Statistics: Statistical Issues in Psychology, Education, and the Social Sciences* 2 (2), 79–103.
- Armstrong, M., Galli, A., Lautier, D., et al., 2009. A reality check of hedging practices in the mining industry. Tech. rep., Paris Dauphine University.
- Astier, J., May 2015. Evolution of iron ore prices. *Mineral Economics* 28 (1-2), 3–9.
- Bäck, T., Fogel, D. B., Michalewicz, Z., 2000. *Evolutionary computation 1: Basic algorithms and operators*.
- Barrodale, I., Roberts, F. D., 1973. An improved algorithm for discrete L1 linear approximation. *SIAM Journal on Numerical Analysis* 10 (5), 839–848.
- Basu, D., Miffre, J., 2013. Capturing the risk premium of commodity futures: The role of hedging pressure. *Journal of Banking & Finance* 37 (7), 2652–2664.
- Bates, D. S., 1996. Jumps and stochastic volatility: exchange rate processes implicit in deutsche mark options. *Review of Financial Studies*.
- Belsey, D. A., Kuh, E., Welsch, R. E., 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley.
- Benini, A., Mancini, A., Longhi, S., 2013. An imu/uwb/vision-based extended kalman filter for mini-uav localization in indoor environment using 802.15.4a wireless sensor network. *Journal of Intelligent & Robotic Systems*, 1–16.
- Benndorf, J., Buxton, M., Shishvan, M., 2014. Sensor-based real-time resource model reconciliation for improved mine production control: a conceptual framework. In: *Proceedings of the SMP Symposium on Orebody Modelling and Strategic Mine Planning: Integrated Mineral Investment and Supply*

- Chain Optimisation.
- Bertisen, J., Davis, G. A., 2008. Bias and error in mine project capital cost estimation. *The Engineering Economist* 53 (2), 118–139.
- Black, F., Scholes, M., 1973. The pricing of options and corporate liabilities. *The journal of political economy*, 637–654.
- Block, S., 2007. Are real options actually used in the real world? *The Engineering Economist* 52 (3), 255–267.
- Boyle, P. P., 1977. Options: A monte carlo approach. *Journal of Financial Economics* 4 (3), 323–338.
- Brennan, M. J., Schwartz, E. S., 1985. Evaluating natural resource investments. *Journal of Business* 58 (2), 135–157.
- Brookner, E., 1998. Tracking and Kalman filtering made easy. A Wiley-Interscience publication. Wiley.
URL <https://books.google.ca/books?id=Ph5TAAAAMAAJ>
- Canadian Securities, 2012. National instrument 43-101. Standards of Disclosure for Mineral Projects, 43–101.
- Casassus, J., Collin Dufresne, P., 2005. Stochastic Convenience Yield Implied from Commodity Futures and Interest Rates.
URL <http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.2005.00799.x/full>
- CFTC, 2015. Weekly commitment of traders reports. Tech. rep., U.S. Commodity Futures Trading Commission, 1155 21st Street, NW, Washington, D.C.
URL www.cftc.gov
- Chernobai, A., Rachev, S., 2006. Applying robust methods to operational risk modeling. *Journal of Operational Risk* 1 (1), 27–41.
- Chevallier, J., Ielpo, F., 2013. *The Economics of Commodity Markets*.

- Christensen, B. J., Prabhala, N. R., 1998. The relation between implied and realized volatility. *Journal of Financial Economics*.
- CME, 10 2015. Futures and options trading for risk management.
URL www.cmegroup.com
- Copeland, T. E., Antikarov, V., Copeland, T. E., 2001. Real options: a practitioner's guide. Texere New York.
- Cox, J. C., Ross, S. A., Rubinstein, M., 1979. Option pricing: A simplified approach. *Journal of Financial Economics* 7 (3), 229–263.
- Danielsson, J., Apr. 2011. Financial Risk Forecasting. The Theory and Practice of Forecasting Market Risk with Implementation in R and Matlab. John Wiley & Sons.
- Dasgupta, M., Mishra, S. K., 2004. Least absolute deviation estimation of linear econometric models: A literature review. Available at SSRN 552502.
- Davidson, R., MacKinnon, J. G., 2004. Econometric theory and methods. Vol. 5. Oxford University Press New York.
- Day, T. E., Lewis, C. M., Oct. 1988. The behavior of the volatility implicit in the prices of stock index options. *Journal of Financial Economics* 22 (1), 103–122.
- Del Moral, P., 1996. Non-linear filtering: interacting particle resolution. *Markov processes and related fields* 2 (4), 555–581.
- Demirel, N., Dec. 2011. Effects of the rock mass parameters on the dragline excavation performance. *Journal of Mining Science* 47 (4), 441–449.
- Dimitrakopoulos, R. G., Abdel Sabour, S. A., Sep. 2007. Evaluating mine plans under uncertainty: Can the real options make a difference? *Resources Policy* 32 (3), 116–125.
- Donkor, E. A., Duffey, M., 2013. Optimal Capital Structure and Financial Risk of Project Finance Investments: A Simulation Optimization Model

- With Chance Constraints. *The Engineering Economist* 58 (1), 19–34.
- Eberhart, R. C., Kennedy, J., 1995. A new optimizer using particle swarm theory. In: *Proceedings of the sixth international symposium on micro machine and human science*. Vol. 1. New York, NY, pp. 39–43.
- Engle, R. F., Kroner, K. F., 1995. Multivariate simultaneous generalized ARCH. *Econometric theory* 11 (01), 122–150.
- Erb, P., Luethi, D., Hinz, J., Oetziger, S., 2014. *schwartz97*: A package on the Schwartz two-factor commodity model. R package version 0.0.6.
URL <http://CRAN.R-project.org/package=schwartz97>
- Evensen, G., 2003. The ensemble kalman filter: Theoretical formulation and practical implementation. *Ocean dynamics* 53 (4), 343–367.
- Fama, E. F., 1965. The behavior of stock-market prices. *The Journal of Business* 38 (1), 34–105.
- Filzmoser, P., Garrett, R. G., Reimann, C., 2005. Multivariate outlier detection in exploration geochemistry. *Computers & geosciences* 31 (5), 579–587.
- Fortin, F.-A., De Rainville, F.-M., Gardner, M.-A., Parizeau, M., Gagné, C., jul 2012. DEAP: Evolutionary algorithms made easy. *Journal of Machine Learning Research* 13, 2171–2175.
- Francois-Bongarcon, D., Gy, P., 2002. The most common error in applying Gy’s Formula’in the theory of mineral sampling, and the history of the liberation factor. *Journal of the South African Institute of Mining and Metallurgy* 102 (8), 475–479.
- Fu, M., Tai, X., 2009. On parameter estimation of the Schwartz-Smith short-term/long-term model. *Asian Control Conference. IEEE, Proceedings of IEEE ICASSP*.
- Gatheral, J., 2011. *The volatility surface: a practitioner’s guide*. Vol. 357. John Wiley & Sons.

- Geman, H., 2005. Commodities and commodity derivatives. Modeling and Pricing for Agriculturals.
- Gibson, R., Schwartz, E. S., 1990. Stochastic convenience yield and the pricing of oil contingent claims. *The Journal of Finance* 45 (3), 959–976.
- Giot, P., Laurent, S., 2003. Market risk in commodity markets: a VaR approach. *Energy Economics*.
- Goldberg, D. E., Holland, J. H., 1988. Genetic algorithms and machine learning. *Machine Learning* 3 (2), 95–99.
- Goodwin, D., 2013. Schwartz-Smith Two-Factor Model in the Copper Market: before and after the New Market Dynamics.
- Gorton, G., Rouwenhorst, K. G., 2006. Facts and fantasies about commodity futures. *Financial Analysts Journal* 62, 47–68.
- Gray, N. F., 1996. Field assessment of acid mine drainage contamination in surface and ground water. *Environmental Geology* 27 (4), 358–361.
- Greene, W. H., 2003. *Econometric Analysis—International Edition*. New York University.
- Hammond, R. K., Bickel, J. E., 2013. On the Decision Relevance of Stochastic Oil Price Models: A Case Study. *The Engineering Economist* 58 (3), 209–230.
- Haque, M. A., Topal, E., Lilford, E., 2014. A numerical study for a mining project using real options valuation under commodity price uncertainty. *Resources Policy* 39, 115–123.
- Haque, M. A., Topal, E., Lilford, E., Jun. 2015. Iron ore prices and the value of the Australian dollar. *Mining Technology* 124 (2), 107–120.
- Haque, M. A., Topal, E., Lilford, E., Mar. 2016. Estimation of Mining Project Values Through Real Option Valuation Using a Combination of Hedging

- Strategy and a Mean Reversion Commodity Price. *Natural Resources Research* 25 (4), 459–471.
- Harvey, A. C., 1990. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.
- Heston, S. L., 1993. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies* 6 (2), 327–343.
- Holland, J. H., 1975. *Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence*. MIT press.
- Huber, P. J., Sep. 1973. Robust Regression: Asymptotics, Conjectures and Monte Carlo. *The Annals of Statistics* 1 (5), 799–821.
- Hunt, J. G., Dowling, J. M., Glahe, F. R., Jan. 1974. L1 estimation in small samples with laplace error distributions. *Decision Sciences* 5 (1), 22–29.
- Hurst, L., Sep. 2015. Resources Policy. *Resources Policy* 45 (C), 247–254.
- James, G., Witten, D., Hastie, T., Tibshirani, R., 2013. *An introduction to statistical learning*. Vol. 6. Springer.
- Jarque, C. M., Bera, A. K., 1980. Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics letters* 6 (3), 255–259.
- Javaheri, A., Lautier, D., Galli, A., 2003. *Filtering in finance*. Wilmott.
- Jonen, C., 2011. *Efficient Pricing of High Dimensional American Style Derivatives*.
- Jorion, P., 2001. *Value at risk: the new benchmark for managing financial risk*. NY: McGraw-Hill Professional.
- Jorjani, E., Poorali, H. A., Sam, A., Chelgani, S. C., Mesroghli, S., Shayestehfar, M. R., 2009. Prediction of coal response to froth flotation based on coal

- analysis using regression and artificial neural network. *Minerals Engineering* 22 (11), 970–976.
- Kalman, R. E., et al., 1960. A new approach to linear filtering and prediction problems. *Journal of basic Engineering* 82 (1), 35–45.
- Karacan, C. Ö., 2008. Evaluation of the relative importance of coalbed reservoir parameters for prediction of methane inflow rates during mining of longwall development entries. *Computers & geosciences* 34 (9), 1093–1114.
- Koenker, R., 2013. *quantreg: Quantile Regression*.
- Kroner, K. F., Kroner, K. F., Kneafsey, K. P., Kneafsey, K. P., Claessens, S., 1995. Forecasting volatility in commodity markets - Kroner - 2006 - *Journal of Forecasting* - Wiley Online Library. *Journal of Forecasting*.
- Laughton, D. G., Jacoby, H. D., 1993. Reversion, timing options, and long-term decision-making. *Financial Management* 22 (3), 225–240.
- Lemelin, B., Abdel Sabour, S. A., Poulin, R., Mar. 2006. Valuing mine 2 at raglan using real options. *International Journal of Mining, Reclamation and Environment* 20 (1), 46–56.
- Leroy, A. M., Rousseeuw, P. J., 1987. *Robust regression and outlier detection*. J. Wiley&Sons, New York.
- Longstaff, F. A., Schwartz, E. S., 2001. Valuing American options by simulation: a simple least-squares approach. *Review of Financial Studies*.
- Luehrman, T. A., Jun. 1998. Investment opportunities as real options: getting started on the numbers. *Harvard business review* 76, 1–16.
- Luis Enrique Coronado, R. C. R. M. M. A. C. D. P., Aug. 2015. Combining Genetic Algorithms and Extended Kalman Filter to Estimate Ankle’s Muscle-Tendon Parameters, 1–10.
- Lüthi, D., Erb, P., Otziger, S., 2014. Using the schwartz97 package.

- Mahfoud, S., Mani, G., Dec. 1996. Financial forecasting using genetic algorithms. *Applied Artificial Intelligence* 10 (6), 543–566.
- Mandelbrot, B., Sep. 1963. The Variation of Certain Speculative Prices. *The Journal of Business* 36 (4), 394–419.
- Manfredo, M. R., Leuthold, R. M., 1998. IDEALS @ Illinois: Agricultural Applications of Value-at-Risk Analysis: A Perspective. . . . : Working paper of the University of
- Maniezzo, A. C. M. D. V., 1992. Distributed optimization by ant colonies. In: *Toward a Practice of Autonomous Systems: Proceedings of the First European Conference on Artificial Life*. Mit Press, p. 134.
- marketrealist.com, 2015. China’s steel production is greatly influencing iron ore prices.
URL <http://marketrealist.com/2015/07/big-chinese-appetite-iron-ore-imports/>
- Martin, R. D., Simin, T. T., 2003. Outlier-resistant estimates of beta. *Financial Analysts Journal*.
- Masteika, Saulius, A. V. R., Alexander, J. A., 2012. Continuous futures data series for back testing and technical analysis. *Conference Proceedings, 3rd International Conference on Financial Theory and Engineering*. 29, 265–269.
- McElhoe, B. A., 1966. An assessment of the navigation and course corrections for a manned flyby of mars or venus. *IEEE Transactions on Aerospace and Electronic Systems* (4), 613–623.
- McKinney, W., 2015. pandas: a python data analysis library. see <http://pandas.pydata.org>.
- Merton, R. C., 1976. Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics* 3 (1), 125–144.
- Minnitt, R., Rice, P. M., Spangenberg, C., 2007. Part 2: Experimental calibration of sampling parameters K and alpha for Gy’s formula by the sampling

- tree method. *Journal of the South African Institute of Mining and Metallurgy* 107 (8), 513–518.
- Morgan, J. P., 1996. Riskmetrics: technical document.
- Myers, S. C., 1977. Determinants of corporate borrowing. *Journal of Financial Economics*.
- Nejadi, S., Trivedi, J., Leung, J. Y., 2015. Estimation of facies boundaries using categorical indicators with p-field simulation and ensemble kalman filter (enkf). *Natural Resources Research* 24 (2), 121–138.
URL <http://dx.doi.org/10.1007/s11053-014-9233-0>
- Obayashi, S., Sasaki, D., Takeguchi, Y., 2000. Multiobjective evolutionary computation for supersonic wing-shape optimization. *IEEE transactions on ...* 4 (2), 182–187.
- O’Hara, T. A., 1980. Quick guide to the evaluation of ore bodies. *CIM Bulletin* 88, 34–43.
- Ortiz, M. C., Sarabia, L. A., Herrero, A., 2006. Robust regression techniques: A useful alternative for the detection of outlier data in chemical analysis. *Talanta* 70 (3), 499–512.
URL <http://www.sciencedirect.com/science/article/pii/S0039914006000270>
- Ortmanns, M., Becker, J., Lorenz, M., Aug. 2015. Hybrid of unscented Kalman filter and genetic algorithm for state and parameter estimation in sigma-delta modulators. *Electronics Letters* 51 (17), 1318–1320.
- Othman, J., Jafari, Y., 2012. Accounting for depletion of oil and gas resources in malaysia. *Natural Resources Research* 21 (4), 483–494.
URL <http://dx.doi.org/10.1007/s11053-012-9192-2>
- Patino Douce, A. E., 2016. Metallic mineral resources in the twenty-first century: Ii. constraints on future supply. *Natural Resources Research* 25 (1), 97–124.

- URL <http://dx.doi.org/10.1007/s11053-015-9265-0>
- Platts, Oct. 2015. The steel index.
- URL <https://www.thesteelindex.com>
- Powell, J. L., 1986. Censored regression quantiles. *Journal of Econometrics* 32 (1), 143–155.
- R Core Team, 2015. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.
- URL <https://www.R-project.org/>
- Rajesh Kumar, B., Vardhan, H., Govindaraj, M., Vijay, G. S., 2013. Regression analysis and ANN models to predict rock properties from sound levels produced during drilling. *International Journal of Rock Mechanics and Mining Sciences* 58, 61–72.
- Raymond McTaggart, Gergely Daroczi, Clement Leung, 2015. Quandl: API Wrapper for Quandl.com. R package version 2.7.0.
- URL <http://CRAN.R-project.org/package=Quandl>
- Reeve, T. A., Vigfusson, R. J., 2011. Evaluating the Forecasting Performance of Commodity Futures Prices. FRB International Finance Discussion Paper 1025.
- Remillard, B., 2013. *Statistical Methods for Financial Engineering*. CRC Press.
- Ribeiro, D. R., Hodges, S. D., 2004. A two-factor model for commodity prices and futures valuation. EFMA 2004 Basel Meetings Paper.
- Rogers, C. D., Robertson, K., 1987. Long term contracts and market stability: The case of iron ore. *Resources Policy* 13 (1), 3–18.
- Rousseeuw, P. J., 1984. Least median of squares regression. *Journal of the American Statistical Association* 79 (388), 871–880.
- Rudenno, V., 1998. *The mining valuation handbook*. Wrightbooks.

- Sadorsky, P., 2006. Modeling and forecasting petroleum futures volatility. *Energy Economics*.
- Samis, M., Davis, G. A., 2014. Using Monte Carlo simulation with DCF and real options risk pricing techniques to analyse a mine financing proposal. *International Journal of Financial Engineering and Risk Management* 1 (3), 264–281.
- Samis, M., Davis, G. A., Laughton, D., Poulin, R., 2005. Valuing uncertain asset cash flows when there are no options: a real options approach. *Resources Policy* 30 (4), 285–298.
- Samis, M., Eng, P., Martinez, L., Davis, G. A., Whyte, J. B., Branch, C. F., Sep. 2011. Using Dynamic DCF and Real Option Methods for Economic Analysis in NI43-101 Technical Reports.
- Samuelson, P. A., 1965. Rational theory of warrant pricing. *Industrial Management Review* 6, 13–31.
- Sasiadek, J., Wang, Q., Zeremba, M., 2000. Fuzzy adaptive kalman filtering for ins/gps data fusion. In: *Intelligent Control, 2000. Proceedings of the 2000 IEEE International Symposium on*. IEEE, pp. 181–186.
- Sauvageau, M., Kumral, M., 2016. Genetic algorithms for the optimisation of the schwartz–smith two-factor model: a case study on a copper deposit. *International Journal of Mining, Reclamation and Environment* 0 (0), 1–19. URL <http://dx.doi.org/10.1080/17480930.2016.1260858>
- Schöne, M., 2014. *Real Options Valuation: The Importance of Stochastic Process Choice in Commodity Price Modelling*.
- Schwartz, E., Smith, J. E., 2000a. Short-term variations and long-term dynamics in commodity prices. *Management Science* 46 (7), 893–911.
- Schwartz, E., Smith, J. E., 2000b. Short-term variations and long-term dynamics in commodity prices. *Management Science* 46 (7), 893–911.

- Schwartz, E. S., 1977. The valuation of warrants: implementing a new approach. *Journal of Financial Economics* 4 (1), 79–93.
- Schwartz, E. S., 1997a. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance* 52 (3), 923–973.
- Schwartz, E. S., 1997b. The stochastic behavior of commodity prices: Implications for valuation and hedging. *The Journal of Finance*.
- Shin, K. S., Lee, Y. J., 2002. A genetic algorithm application in bankruptcy prediction modeling. *Expert Systems with Applications* 23 (3), 321–328.
- Sims, K., Jul. 1994. *Evolving virtual creatures*. ACM, New York, New York, USA.
- Subulan, K., Baykasoğlu, A., Akyol, D. E., Yildiz, G., 0. Metaheuristic-based simulation optimization approach to network revenue management with an improved self-adjusting bid price function. *The Engineering Economist* 0 (0), 1–30.
- URL <http://dx.doi.org/10.1080/0013791X.2016.1174323>
- Svetlana, B., 2010. *Valuation of Metals and Mining Companies*. collaboration with the University of Zürich.
- Tapia, M., Coello, C., 2007. Applications of multi-objective evolutionary algorithms in economics and finance: A survey. ... 2007 CEC 2007 IEEE Congress on, 532–539.
- Taylor, H. K., 1986. Rates of working of mines—a simple rule of thumb. *Institution of Mining and Metallurgy Transactions. Section A. Mining Industry* 95.
- Thompson, W. R., 1935. On a criterion for the rejection of observations and the distribution of the ratio of deviation to sample standard deviation. *The annals of mathematical statistics*.

- Ting, T. O., Man, K. L., Lim, E. G., Leach, M., 2014. Tuning of Kalman Filter Parameters via Genetic Algorithm for State-of-Charge Estimation in Battery Management System. *The Scientific World Journal* 2014 (2), 1–11.
- Todorov, V., Filzmoser, P., 2009. An Object-Oriented Framework for Robust Multivariate Analysis. *Journal of Statistical Software* 32 (3), 1–47.
- Trigeorgis, L., 1993. Real options and interactions with financial flexibility. *Financial Management*, 202–224.
- Vasicek, O., Nov. 1977. An equilibrium characterization of the term structure. *Journal of Financial Economics* 5 (2), 177–188.
- Venables, W. N., Ripley, B. D., 2002. *Modern Applied Statistics with S*.
- Welsch, R. E., Zhou, X., 2007. Application of robust statistics to asset allocation models. *REVSTAT–Statistical Journal* 5 (1), 97–114.
- Willems, J. C., 1978. Recursive filtering. *Statistica Neerlandica* 32 (1), 1–39.
- Wilson, J. D., 2012. Chinese resource security policies and the restructuring of the asia-pacific iron ore market. *Resources Policy* 37 (3), 331–339.
- Yohai, V. J., 1987. High breakdown-point and high efficiency robust estimates for regression. *The Annals of Statistics* 15 (2), 642–656.
- Zhang, K., Kleit, A. N., Mar. 2016. Mining rate optimization considering the stockpiling - A theoretical economics and real option model. *Resources Policy* 47 (C), 87–94.
- Zhang, K., Nieto, A., Kleit, A. N., Jul. 2014. The real option value of mining operations using mean-reverting commodity prices. *Mineral Economics* 28, 11–22.