THE BUSINESS INTELLIGENCE AS A SERVICE IN THE CLOUD

Victor Chang[1,2]

1. School of Computing and Creative Technologies, Leeds Metropolitan University, Leeds, UK.
2. School of Electronics and Computer Science, University of Southampton, Southampton, UK.
V.I.Chang@leedsmet.ac.uk

Abstract

Limitations imposed by the traditional practice in financial institutions of running risk analysis on the desktop mean many rely on models which assume a "normal" Gaussian distribution of events which can seriously underestimate the real risk. In this paper, we propose an alternative service which uses the elastic capacities of Cloud Computing to escape the limitations of the desktop and produce accurate results more rapidly.

The Business Intelligence as a Service (BIaaS) in the Cloud has a dual-service approach to compute risk and pricing for financial analysis. In the first type of BIaaS service uses three APIs to simulate the Heston Model to compute the risks and asset prices, and computes the volatility (unsystematic risks) and the implied volatility (systematic risks) which can be tracked down at any time. The second type of BIaaS service uses two APIs to provide business analytics for stock market analysis, and compute results in the visualised format, so that stake holders without prior knowledge can understand. A full case study with two sets of experiments is presented to support the validity and originality of BIaaS. Additional three examples are used to support accuracy of the predicted stock index movement as a result of the use of Heston Model and its associated APIs.

We describe the architecture of deployment, together with examples and results which show how our approach improves risk and investment analysis and maintaining accuracy and efficiency whilst improving performance over desktops.

Key Words

Heston Model Simulations; Heston Model; Business Intelligence as a Service (BIaaS); Calibration; APIs for stock index; Visualisation in the Cloud; SaaS in the private cloud.

1. Introduction

There are limitations to desktop-based financial software to run a large number of simulations. Often activities of large-scale software simulation to calculate risk and pricing take hours without simplifying the model [1]. Simplification of model was used in finance industry for some time prior 2008 crisis. It is known as "Gaussian copula" to model rates for mortgage, bonds and derivatives. His model could work on the desktop and results could be calculated for an acceptable amount of time. However, his model did not check risks properly and did not have any suggestions to counter risks when extreme events happened [1]. Here is a real story describing the background that lead to the development and widely adoption of "Gaussian copula" and our proposal to rectify problems caused by this model.

1.1 Background

David X Li improved concepts implemented by financial services and modelled 'credit derivatives' to calculate collateralised debt obligation (CDO), which is a type of structured asset-backed security (ABS) with multiple series issued by special purpose entities for debt obligations including bonds and loans. Li allowed his model to get yields of a corporation's bonds or the prices of the new credit swaps to model the 'survival time' of an individual corporation (the time until it defaults). Li resolved this issue by introducing 'copula function' that he learned in statistics. Copula function is commonly used in mortgage lending concept to calculate the impacts of defaults due to deaths of one of the spouses [2]. Copula function can calculate set of marginal functions (to specify the probability that the wife will die before a given age, and separate function that specifies that the husband will die at or before another age) to form the joint or 'multivariate' distribution function (to specify the probability that the wife will die before a given age AND the husband will die at or before another age). This is one of the fundamental errors that Li has implemented, since probability for both events to happen is different from one event to happen and results can be varied widely in extreme conditions.

Li then combined a popular financial software, CreditMetrics, and Copula function together to establish "Gaussian copula" model, which has been used by the finance industry since 1997 [1]. The "Gaussian copula" model is often called as CDO solution, since it is a term easily understood by the finance sector. By combining both approaches, finance sector could enjoy benefits from both sides – Gaussian's simplicity and familiarity and copula's unified and easy-to-use approach. Gaussian copula model allowed analysts to buy a pool of bonds or loans, raising their money to do so by selling investors' securities claimed on the cashflow generated from the pool. The drawback for this approach is the assumption that correlation stays low. If correlation goes too high, the holders of the highest investment were at risks, and there was no detailed way to compute the extent of volatility in such extreme conditions [1], since the model takes assumption on the 'bright side' of the trading due to simplification of "Gaussian copula" model to compute complex financial derivatives, prices and volatility.

Li then developed his final version of software called CDO Evaluator, which became more and more popular amongst quantitative developers and investment banks. There were a few reasons according to MacKenzie and Spears [1] and their interviews with experts and developers working in financial services. Firstly, developers needed not to think of many variables which are time-consuming to obtain. Secondly, they could possibly avoid using Monte Carlo simulations that took overnight and the weekend to run through all simulations and perform exhaustive testing of financial derivatives. Thirdly, it was easier to understand the problems, since there were fewer variables to know. Fourthly, it was also easier to communicate with other teams. Financial problems and derivatives were difficult to understand and even within teams with different skills and focus, communications were not easy or lengthy. The use of concepts of "Gaussian copula" model and the CDO Evaluator could ease the level of difficulties during communications. Fifthly, developers found it easy to reproduce Li's concepts due to the simplicity of the model and the problem. Sixthly, his software was backed by some leading quantitative developers at that time, and had widespread use in investment banks.

1.2 Business Intelligence as a Service (BIaaS) in the Cloud

There are some Grid based financial applications but they cannot be transferred to Cloud Computing easily [3]. Where applications can be moved, there are advantages to moving risk assessment calculations to Clouds including risk reduction, the opportunity to compare results and the capability to run less conventional models able to identify errors and uncertainties

which are currently missed. However, factors such as accuracy, speed, reliability and security of financial models and their attendant costs must be considered [4]. Public clouds are not suitable due to privacy and data ownership issues [5]. A hybrid cloud could be used but this requires implementation of security technologies which are not the focus of our research. Private clouds are the obvious choice for the financial sector and also relevant to our objective.

The term "Business Intelligence" refers to a set of methods, processes, architectures and technologies that can process and transform collected datasets into meaningful and useful information for business purposes, and often used in business-critical servers, applications and services. Advanced simulations and processing in financial computing such as derivatives, stock analytics and financial software are part of Business Intelligence (BI), which aims to simplify the complexity of datasets and presents them with information for IT strategies and operational activities [6]. Daily activities performed by investment banks can be used by BI systems as an alternative, and another additional advantage is that BI systems can integrate with other technologies such as Cloud Computing to offer more added values to organisations [6, 7, 8]. There are two examples here to illustrate the added values of using BI systems in the Cloud. Firstly, Xu [7] designed and developed Cloud BI systems for manufacturing, and demonstrated how Cloud BI can transform the way that manufacturing was used to work. Cloud BI allowed different machines to work collaboratively and efficiently. Secondly, Marston [8] explained the added values of Cloud BI systems for business perspective and they demonstrated examples on how organisations could get different levels of contributions by using Cloud services such as Cloud BI services. These examples acknowledge the benefits of adopting BI in the Cloud for improved technical and business perspectives, thus, a hybrid solution of integrating BI in the Cloud is another motivation for this paper.

We propose and describe Business Intelligence as a Service (BIaaS) which is a Cloud based service designed to improve the accuracy and quality of both pricing and risk analysis in financial markets, compared with traditional desktop technologies. BIaaS is a type of Software as a Service (SaaS) with the emphasis on how the application offers quality services in private cloud environments. This is important because incorrect analysis leads to excessive risk taking which may then lead to financial losses, damage to business credibility or destabilised markets. We illustrate its use with an example which shows price and risk assessments for investments such as stocks and shares or financial derivatives in the context of different levels of volatility, maturity and risk free rates.

BIaaS has the dual-service approach to address the following challenges:

1. Compute the risks and asset prices, and computes the volatility which can be tracked down at any time.

2. Performing a sufficiently high number of simulations in acceptable time.

The breakdown for this paper is as follows. Section 2 describes problems with the existing and popular model used by financial services and the proposal of BIaaS. Section 3 explains all formulas associated to Heston Model, the model behind BIaaS. Section 4 explains the system architecture, including their associated services and APIs. Section 5 presents the set up and deployment of a private Cloud. The two stages of BIaaS are described in Sections 6 and 7, which use a number of experiments to validate the results. Section 8 describes two sets of experiments to demonstrate the validity and originality of BIaaS. Section 9 compares BIaaS and other models running on desktop. Section 10 discusses the use of BIaaS and benefits of adopting BIaaS for risk modelling. Section 11 describes conclusion and future work.

2. Problems with existing Gaussian copula modelling and our proposal

This section is aimed at describing the problems caused by Gaussian copula modelling, a model widely adopted by financial services and investment banks to calculate the lending. It starts with backgrounds, the process of getting popularity and explanations about the problems associated with the model.

2.1 Simplicity at the expense of accuracy and performance

However, Li's contributions to the finance industry were known as "receipt to disaster" after the financial crisis since 2008. The model underestimated the probability of the risks and did not have any measure to counter the risks when they began to take an immediate effect. Several assumptions he made his model did not work in extreme conditions [1]. His model could work on the desktop and results could be calculated for an acceptable amount of time. By taking simplification of Li's model, financial services could run their services and get the results quickly. This can avoid the need to run days and hours of Monte Carlo simulations. However, risks are not properly checked as a consequence. His model took the simplicity and ease-of-use at the expense of the risk modelling and controls [1]. Li also admitted his weakness of the model and wrote: "The current copula framework gains its popularity owing to its simplicity....However, there is little theoretical justification of the current framework from financial economics....We essentially have a credit portfolio model without solid credit portfolio theory" [9]. However, we argue that it is unprofessional to take false assumption and encourage users to take simplicity due to convenience. It is important to run models properly without sacrificing the quality of results but also offer a prompt services without the need of running simulations for hours and days.

2.2 BIaaS Private Cloud is the solution other than desktop

Running simulations on desktop clearly need much longer time and a smart way of performing a high number of simulations with a short time is required. The use of Cloud Computing techniques and technologies can play an important role to optimise the speed and also allow more vigorous tests and simulations are performed. The BIaaS private cloud is designed and deployed to ensure all calculations of pricing and risks are as accurate as possible and also can be computed within seconds and minutes. This can help financial services to perform thorough testing without compromising to run simpler model which does not take risks (happened in extreme condition of probability of 2% and lower) seriously. Different models and scenarios of risks models will be presented.

2.3 BIaaS – requirements and the need to rectify errors left by Gaussian copula model

It becomes important for computer scientists to develop better models to improve such a situation. Better ways to calculate pricing and risks, rectify errors and perform accurate and fast simulations are highly desirable. Requirements for our BIaaS are thus as follows:

1. Based on the reputable models – Our BIaaS adopts reputable models including the Heston Model (which includes the Wiener process and the Stochastic Volatility) and the Visualisation APIs to compute the best pricing and risks for different scenarios.

2. Compute and track volatility – BIaaS can track the movement of volatility which can help investors to make better judgement for investment when prices are high and volatility is low (which happens but the period is short and is not always available).

3. Accuracy – our BIaaS can compute pricing and risk values to several decimal places and also calculate its mean, lower and upper range to get our results as accurate as possible.

4. BIaaS should not just limit its operations on desktop or a particular platform but on different types of Clouds and desktop.

Results or discussions in other sections of the paper will refer back to here to explain how other development can fulfil requirements for BIaaS.

3. Outline of BIaaS and Methods used by BIaaS

Gaussian copula models are used in financial modelling and many banks' mathematical models assume normal (Gaussian) distributions of events and may underestimate risks in real financial markets [10]. Moreover, these models make assumptions about market behaviour which may not always be true with the result that the models can fail to detect risks, as highlighted by the financial crisis in 2008. To address this alternative, non-Gaussian financial models are needed. Various studies conclude that modelling of financial markets needs to be addressed in two stages; one for pricing and another for risk analysis [1, 11]. This means a more suitable model is required for large-scale of financial analysis. BIaaS is the most commonly adopted and provides data for investors' decision-making amongst other models [12].

3.1 The Heston Model: fulfilling the first BIaaS requirement

This section explains the suitable model for BIaaS to fulfil the first requirement presented in Section 2.3. BIaaS is derived from mathematical Heston Model and additional APIs and is a computational technique used to calculate risk; the probability of an event or investment happening. BIaaS is based on probability distributions, so that uncertain variables can be described and simulated with controlled variables [12, 13]. BIaaS is suitable to generate data which investors can use when making decisions [10]. When volatility is known, put and call prices can be calculated [14, 15]. Moreover, BIaaS has specific techniques such as Wiener process to compute high-volume of simulations and track the movements of volatility for the computed data.

The Heston Model has a close relationship with Black-Scholes model, since it relaxes the constant volatility assumption in the classical Black-Scholes model by incorporating an instantaneous short term variance process [16]. This means the Heston Model can be used in a more flexible way and is not as theoretical-oriented as the classical Black-Scholes model does. In addition, there are both the Wiener process and the CIR process related to the Heston Model, and their explanation is as follows.

3.2 The Wiener process and the Heston Model

The Wiener process is a continuous-time stochastic process named after Norbert Wiener. It is a standard Brownian motion and can be in mathematics. In applied mathematics, the Wiener process is used to represent the integral of a Gaussian white noise process for electronics engineering, instrument filtering theory and control theory. It can be used effectively in the mathematical theory of finance, in particular the Black–Scholes option pricing model. There are three properties in the Wiener process Wt [15, 17, 18]

1. $W0 = 0$

2. The function $t \rightarrow Wt$ is almost surely everywhere continuous

3. $Wt-Ws \sim N(0, t-s)$ (for $0 \leq s < t$).

The Wiener process is a stochastic process with independent and stationary increments, which means the motion of a point whose consecutive displacements are independent and random with each other. The Wiener process has Lévy characterisation has continuous martingale with W0 = 0 and quadratic variation [Wt, Wt] = t. This implies that Wt2−t is a martingale.

Here is the description to explain the relationship between Wiener process and the Heston Model, which starts with the mathematical formula first.

The basic Heston model assumes that S_t, the price of the asset, is determined by a stochastic process [20]

$$dS_t = \mu S_t \, dt + \sqrt{\nu_t} S_t \, dW_t^S \tag{1}$$

where ν_t, the instantaneous variance, is a CIR process, which is a Markov process with continuous paths defined by the following stochastic differential equation (SDE):

$$d\nu_t = \kappa(\theta - \nu_t) \, dt + \xi \sqrt{\nu_t} \, dW_t^\nu \qquad (\nu_0 = \xi^2 \text{, which is} > \text{or} = 0) \tag{2}$$

and dW_t^S, dW_t^ν are Wiener process (i.e., random walks) with correlation ρ dt.

The parameters in the above equations represent the following:

- μ is the rate of return of the asset.
- θ is the long variance, or long run average price variance; as t tends to infinity, the expected value of ν_t tends to θ.
- κ is the rate at which ν_t reverts to θ.
- ξ is the volatility of the volatility; as the name suggests, this determines the variance of ν_t.

All these parameters can be used in calibration to determine their respective values, and is an important process used by other models to validate results in financial analysis. Details will be presented in Section 3.5 for the right type of parameters used, and in Section 8.1 for large-scale experiments for result validation.

If the parameters obey the following condition (known as the Feller condition) then the process ν_t is strictly positive [17, 18] $2\kappa_t\theta \geq \xi^2$.

3.3 The CIR process and the Heston Model

Heston model is initially derived from the CIR model of Cox, Ingersoll and Ross [18, 19] for interest rates. CIR process is a Markov process with continuous paths defined by the following stochastic differential equation (SDE) presented by formula (2). The CIR process is used to model stochastic volatility in the Heston model, which aims to resolve a shortcoming of the Black–Scholes model which corresponds to the fact that the implied volatility does tend to vary with respect to strike price and expiry. By assuming that the volatility of the underlying price is a stochastic process rather than a constant, stochastic volatility can make it possible to model derivatives more accurately.

The CIR paper also cites earlier and seminal work of Feller [20], including proper (Feller) boundary conditions, process non-negativity and the distribution for the general square-root diffusions.

3.4 Runge-Kutta method

The Runge–Kutta method (RKM) is a technique for the approximate numerical solution of a stochastic differential equation (SDE) [15, 18]. RKM can be used to generalise the ordinary differential equation to SDE. The related formulas are as follows.

the Itō diffusion X satisfying the following Itō stochastic differential equation [18, 21]

$$dX_t = a(X_t)\,dt + b(X_t)\,dW_t, \tag{3}$$

with initial condition $X_0 = x_0$, where W_t stands for the Wiener process, and suppose that we wish to solve this SDE on some interval of time $[0, T]$. Then the Runge–Kutta approximation to the true solution X is the Markov chain Y defined as follows:

- partition the interval $[0, T]$ into N equal subintervals of width $\delta = T/N > 0$:

$$0 = \tau_0 < \tau_1 < \ldots < \tau_N = T;$$

  - set $Y_0 = x_0$;
  - define Y_n for $1 \leq n \leq N$ by

$$Y_{n+1} = Y_n + a(Y_n)\delta + b(Y_n)\Delta W_n + \frac{1}{2}\left(b(\hat{\Upsilon}_n) - b(Y_n)\right)\left((\Delta W_n)^2 - \delta\right)\delta^{-1/2},$$

(4)

where $\Delta W_n = W_{\tau_{n+1}} - W_{\tau_n}$

and $\hat{\Upsilon}_n = Y_n + a(Y_n)\delta + b(Y_n)\delta^{1/2}$.

ΔW_n are independent and identically distributed normal random variables with expected value zero and variance δ. RKM can be used to calculate SDE, which is required by the Heston Model.

3.5 Root-mean square error (RMSE) – the first formulas for calibration

Calibration is used in Business Intelligence in a way that a known observation of the dependent variables is used to predict a corresponding explanatory variable. Systematic approaches are followed based on the adoption of right formulas. There are two associated formulas in regard to calibration. The first formula is the root-mean square error (RMSE) and the second is moving windows (MW) estimate [18, 21].

The root-mean square error (RMSE) is used to measure of the differences between values predicted by a model or an estimator and the values actually observed. RMSE also determines the goodness of fit of the Heston Model presented as follows [18, 21].

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}} \tag{5}$$

where n is the number of quoted options, X_{obs} is observed values and X_{model} is modelled values at time/place i. The parameters required for RMSE include (v_0, κ, θ, ξ, ρ) used for calibration and v_0 is the instantaneous variance at the starting point. Referring to formula (2), the rate of return of the asset can be calculated by multiplying κ and difference between θ and v_0.

Guillaume and Schoutens [22] explain that parameters in RMSE can be recalibrated every day to new market data, and optimal parameter set can vary significantly on a day-to-day basis. An important goal is to avoid sharp fluctuations of the model parameters. An alternative is to use or restore the time series or market quotes to calculate some of the model parameters. Guillaume and Schoutens [22] explain that parameters v_0 can be remained fixed before calculating the value of volatility index, where η (To avoid confusion with n, researchers use θ to differentiate) can be estimated by the following three ways:

1. Historical time series of the volatility: This is a common approach demonstrated by researchers [14, 17, 18, 21].

2. Market quote of the VIX options: VIX is known as the Chicago Board Options Exchange Market Volatility Index, which is a popular measure of the implied volatility of S&P 500 index options [22, 23]. The Moving Window estimate can be used for computation for forecasting of the market index.

3. Another model to deal with fluctuations: Another model and approach is required. A common approach is to use exponentially weighted moving average (EWMA) [22, 23].

In current BIaaS implementation, the first and second objective can be demonstrated to show that BIaaS can compute volatility and calculate the likely movement in the stock index. Referring to Section 2.3, both performances (tracking of volatility) and accuracy can be achieved by the combined approach in the first and two objectives above.

3.6 The Moving Window (MW) estimate and exponentially weighted moving average (EWMA) - the second and third formulas for calibration

As explained in Section 3.5, the Moving Window (MW) estimate is a suitable model in the use of VIX options. MV can be computed as the mean of variance of the stock price process over the time series window that moves forward in time. The formula becomes [18, 21, 22]

$$\eta^{MW} = \frac{1}{T^{VIX}} \int_{-T^{VIX}}^{0} \left(\frac{VIX(t)}{100}\right)^2 dt = \text{mean}_{-T^{VIX} \leq t \leq 0} \left(\frac{VIX(t)}{100}\right)^2 \tag{6}$$

For this paper, MW is used to compute the forecasted movement in the Heston Model.

On the other hand, exponential weighted moving average (EWMA) estimate the long run variance presented as

$$\eta^{EWMA} = (1-\alpha) \sum_{i=1}^{N} \alpha^{N-i} \left(\frac{VIX(t_i)}{100}\right)^2 \tag{7}$$

Where $\alpha \in (0,1)$, $t_i = t_0 - (N-i)\Delta t$ and where $N \rightarrow \infty$ is the number of data in the time series. EWMA is designed to facilitate the stock index that has fluctuations. Experiments with input values for output modelling will be presented in Section 8. The aim of using MV in BIaaS is to calculate the predicted movement behind the scene and present to the analysts the result in visual formats. Results and discussions will be presented in Section 8 and 10.

3.7 Average absolute percentage error (APE) of the mean price and aggregated relative percentage error (ARPE) – additional formulas for calibration

Two additional formulas are useful for validation of results. Average absolute percentage error (APE) of the mean price is a measure of accuracy of a method for constructing fitted time series values in statistics and financial computation [18, 21, 22].

$$\text{APE} = \frac{1}{\text{mean}_j \hat{P}_j} \sum_{j=1}^{N} \frac{\left| P_j - \hat{P}_j \right|}{N} \tag{8}$$

Calculations by APE may cause a problem. A few of the series with a very high APE might distort a comparison between the average APE of time series fitted with one method compared to the average APE when using another method. To improve on this situation, another model, aggregated relative percentage error (ARPE) is used.

$$\text{ARPE} = \frac{1}{N} \sum_{j=1}^{N} \frac{\left| P_j - \hat{P}_j \right|}{\hat{P}_j} \tag{9}$$

All the formulas in Section 3.5, 3.6 and 3.7 can utilise parameters (v_0, κ, θ, ξ, ρ) for calibration.

4. The APIs and System Architecture for BIaaS

Section 3 explains all the formulas associated with BIaaS. However, mathematical modelling is not easy to use and analysts need years of training to be sufficiently competent for financial computation. Additional work is then required, including the development of APIs that perform financial computations and only presents the results to the analysts. In order to understand the work behind the scene, this section explains the high-level functionality of BIaaS, the role and functionality of associated APIs and the system architecture in order to perform experiments for Section 5, 6 and 7.

4.1 Two major services to offer

For the demonstration for this paper, experiments are running at two sites starting from Southampton, and the processing take place mainly at ULCC in London. Figure 15 in the Appendix shows the deployment where the details of hardware implementations are described in Section 5. Two different Cloud services are developed and their descriptions are as follows:

- Heston Volatility and Pricing as a Service (HVPaaS): The request started and completed at Southampton clusters, including the processing of the HVPaaS. The objective was to track volatility and pricing simultaneously. The reason is that volatility and pricing can change significantly during the volatile period such as financial crisis in 2008 and 2009. A systematic approach is required to compute these volatile values in real-time.

- Business Analytics as a Service (BAaaS): After analysing the numerical computation of volatility and pricing, the next step is to compute them as a Business Analytic. This makes the analysis much easier and the stakeholders can understand. After the processing of HVPaaS completed in Southampton, results are sent to Greenwich and ULCC in London, where both sites can process BAaaS. ULCC has better platforms and can produce a large-scale of processing than Greenwich.

Details for each service will be presented in Section 6 and 7 respectively. We first present how each service works in their system architecture and then how the two services can work together in the collaborative Cloud environment.

4.2 The role of the Application Programming Interfaces (APIs) for BIaaS

Formulas and theories behind the Heston Model, the Wiener process, the CIR process and RKM can be complex and not easy to be interpreted correctly in a way appealing to stake holders and investors. It is important for Cloud services to be easy to use as usability is a key criteria for successful Cloud delivery [24, 25]. In this research, APIs are used to bridge the gap between the complexity of the mathematical models and the good delivery of Cloud services. The emphasis is on how these APIs can be used, the results and benefits offered by the private cloud development for financial services research. Section 4.4 and 4.5 will describe the functionality of five APIs. Section 6 and 7 will describe the detailed usage scenarios for these five APIs.

4.3 System Design and Architecture

The working System Design and Architecture is similar to Papazoglou and van den Heuvel's [26] and Chang's [27, 28, 29] cloud delivery model to streamline a structured process. A number of technologies mainly APIs are used for each step of the service. The major advantage being that each step has a preference for certain technologies or platforms due to its functionality. For example, a statistical computing service favours statistical languages or packages. Figure 1 shows System Design and Architecture, the structure of each service, how different services are connected to each other and how data is exchanged between different virtualised layers. The network infrastructure is based on the NGN concept proposed by Rings et al [30]. Our BIaaS demonstration is based on the development of two proposals – the virtualisation and architectures from Rings et al [30] and Papazoglou and van den Heuvel's [26].
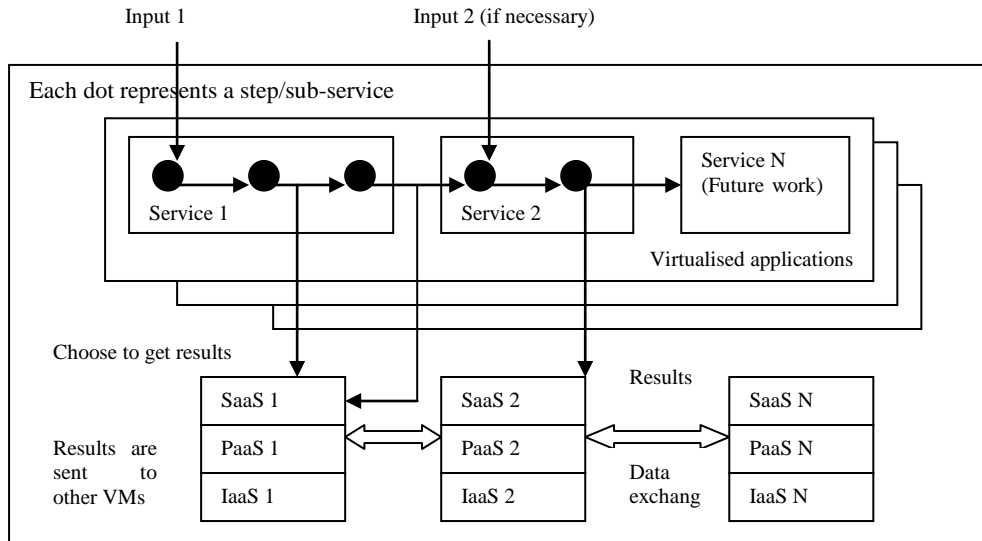
Figure 1: System Design and Architecture for BIaaS in Southampton and London Private Cloud sitting on top of NGN

4.4 Three APIs for the Heston Model

There are three APIs developed for the Heston Model to calculate the implied volatility. The Heston Model can use all of the Wiener process, the CIR process and Runge–Kutta method to achieve this as explained earlier. Each API can be used as a one-line command line. The APIs and functionality for each API are as follows.

1. **ItoProcess** API - this uses the Wiener process to launch stochastic process and calculate key variables for the Heston Model.

2. **BlockRandom** API – this is relevant to the Runge–Kutta method to compute stochastic differential equation for the Heston Model.

3. **CovarianceFunction** API – this is relevant to the Wiener process mainly and the CIR process to treat any process as a random process so that no input data is required. However, CovarianceFunction is used when key variable are known and the volatility is calculated. This means there is a higher level of confidence in those key values prior using CovarianceFunction.

There is an optional API, **Plot**, which computes all numerical results and displays them as 3D Visualisation. This allows stakeholders to understand it more easily.

Results are saved in numerical formats readable by each API service and then passed onto the next API. Service 2 is BAaaS which is itself comprised of two APIs. Results from the last API of first service in Southampton have two options. The first option is to report results in numerical formats, which can be interpreted by scientists and experienced data analysts. However, if such analysis is difficult to interpret, additional APIs are required to simplify the complexity for stakeholders. For the demonstration for this paper, BIaaS takes on the second option, which allows the result from "CovarianceFunction" API to pass onto the first API of BAaaS in London.

4.5 APIs for the Business Analytics Visualisation

APIs for visualisation are useful to present risks and asset prices in real-time. Visualisation can be used for purposes such as the review of the previous performance and review of the

current performance, particularly for the stock market analysis. There are two APIs available for the visualisation service and their function is described briefly as follows.

1. **FinancialData** API – this allows the BIaaS Cloud to obtain financial data from Google Finance and have all the major stock market data, particularly the US and UK stock exchange data.

2. **TradingChart** API – this allows the financial data to be presented in the trading chart format similar to the visualisation services offered by London Stock Exchange and Thomson Reuters. There are two additional functions, "SimpleMovingAverages" and "BollingerBands", which corresponds to the use of MW model to compute forecasted movement. "TradingChart" is the API to demonstrate both models (Heston and Financial data) can work together to deliver an integrated service.

Details for the analysis will be available in the Section 6 and 7.

4.6 Operation of BIaaS

Referring to Figure 1, results are saved in numerical formats readable by each API service and then passed onto the next API. This can ensure the entire service to operate as an integrated service. The dual-service approach of operating BIaaS includes:

1. Risk simulation service using BIaaS: This includes calculating the volatility and the asset prices, and their real-time numerical values. ItoProcess, BlockRandom, CovarianceFunction (based on the Heston Model) and Plot APIs are developed to support this function.

2. 3D Visualisation of pricing and volatility – The synchronised asset prices and volatility is computed and presented in 3D Visualisation and in a way that stake holders can understand without prior knowledge. FinancialData and TradingChart APIs are developed to support this function. FinancialData uses Google Finance and the Heston Model for financial computation.

3. Trading analytics: BIaaS offers the computation of the trading analytics. Although TradingChart is the main API for this functionality, it does require support from all other APIs to deliver this service.

5. Deployment for the experiments

This section describes the hardware, architecture and deployment scenarios used for performing experiments and simulations for BIaaS. Section 4.1 explains the hardware set up, section 4.2 describes the deployment scenario and section 4.3 presents the architecture.

5.1 Hardware used for experiments and benchmarks

This section describes a list of technology selections with their justification. The private cloud built at the University of London Computer (Data) Centre (ULCC) is based on a private-cloud SAN architecture made up of different clustering services, where each cluster is dedicated for one specific function. Various simulations and experiments for BIaaS have been performed using a high specification desktop environment, private and public clouds. The desktop machine has 2.67 GHz Intel Xeon Quad Core and 4 GB of memory (800 MHz). Figure 15 in the Appendix shows the architecture of the private cloud which involves six sites in total; two in London and two in Southampton. The University of Southampton resources are used for all experiments, and are also used to connect lead author's home cluster, Greenwich and London Data Centre. On Figure 15 in the Appendix, blue arrows indicate computational connections between internal networks.

Table 1: Selections of Technology Solutions

Technology selections	What is it used	Vendors involved	Focus or rationale	Benefits or impacts
Network Attached Storage (NAS)	To store data and perform automated and manual/personal backup.	Iomega/EMC Lacie Western Digital HP	They have a different focus and set up. HP is more robust but more time-consuming to configure. The rest is distributed between RAID 0, 1 and 5.	Each specific function is assigned with each NAS. There are 8 NAS at ULCC, including 2 for Archiving. Deployment Architecture is described in [23].
Infrastructure (networking and hosting solution)	Collaborator and in-house	ULCC University of Southampton Author's own resource	Some services need a more secure and reliable place. University of London Data Centre offers 24/7 services with around 500 servers in place, and is ideal for hosting solution.	Amount of work is reduced for maintenance of the entire infrastructure. It stores crucial data and used for archiving, which backup historical data and backup the most important data automatically and periodically.
Backup and computing applications	Third party and in-house	Open Source Oracle HP Vmware Symantec In-house development	There is a mixture of in-house development and third party solution. HP software is used for high availability and reliability. The rest is to support backup in between NAS. Vmware is used for virtual storage and backup.	Some applications are good in a particular service, and it is important to identify the most suitable application for particular services.
Virtualisation	Third party	VMware VSphere and Citrix	It consolidates IaaS and PaaS in private cloud deployment.	Resources can be virtualised and saves effort such as replication.
Security	Third party and in-house	ULCC University of Southampton Macafee Symantec F5	Security is based on the in-house solution and vendor solution is focused on secure firewall and anti-virus.	Remote access is given to a list of approved users.

London Data Centre has advanced Cloud and parallel computing infrastructure and network attached storage (NAS) service. In total it has CPUs totalling 30 GHz, 60 GB of RAM and 12 TB of disk space in place. Experiments performed in this environment get the best benefit of advanced optical fibre networking. There are two servers at London Greenwich, with a total of 9 GHz CPU and 20 GB RAM. The two servers at University of Southampton both have 6.0 GHz and 16 GB RAM. For the home cluster, the total hardware capability is 24.2 GHz CPU and 32 GB RAM.

Simulations and experiments on a desktop and two private clouds (one in Southampton and one in London) get the same results, and thus execution time to complete all simulation is the benchmark to differentiate their performance on different platforms. Apart from processing BIaaS applications at Southampton, most of advanced applications rely on the robust

infrastructure and platform at ULCC to provide reliable and accurate services. Design and Deployment is based on project requirements and their research focus. Selections of Technology Solutions are essential for Cloud Storage development as presented in Table 1.

5.2   Motivation for performing experiments for BIaaS

Business Intelligence as a Service (BIaaS) can also be used independently but it needs manual input of data for processing if that is the case. In this paper, the focus is to demonstrate that BIaaS in the Cloud can be delivered as an integrated service consisting of two different services. The first service is Heston Volatility and Pricing as a Service (HVPaaS) to process stochastic equations and present the calculated implied volatility and pricing. The second service is Business Analytics as a Service (BAaaS) to present financial data and analysis in a collaborative and easy-to-understand way. This makes integrations of two services useful for BIaaS development from both development and enterprise points of view. Technical developers can find that technical capabilities can allow them to compute two different types of business intelligence modelling once rather than twice. The business benefits for the enterprise is the reduction of time and cost to deliver services, which means the enterprise only pay the service once rather than twice to different service providers. Integrating both services requires the following:

- Results from the end of HVPaaS and the end of each API need to be saved as text (or numerical format) passed to the next step, allowing results from each API to be passed onto the next.

- Streamlining both services as a single process and ensure both are completed in one go rather than as two separate services.

The experiments are performed using a private cloud located in Southampton, distributed between two sites; two high performance servers with multiple VMs located at University of Southampton and two clusters of eight servers with VMs located at the lead author's home. All are connected to form a private cloud.

5.3   Preliminary set up to minimise risks

There are risks that can affect the performance of experiments. They mainly include synchronisation, network traffic control and library dependency. Risk-control rate is the rate to control risk of running experiments. The target is to maintain the incomplete API processing within 1%. Risk-control rate must be managed carefully and this can be achieved by the following steps.

- Synchronising all experiments: This can ensure experiments to London and Southampton start at the same time, and it makes the management of risk-control rate once rather than twice per experiment. Or all experiments are running in London, and all experiments start simultaneously.

- Network traffic and speed monitoring: The use of tool and scripts can measure the network traffic and upload speed time, which can be tracked and presented as a graph. The tool and scripts can report that the upload network speed between Greenwich and ULCC in London is as close as to each other. The expected upload bandwidth is 400 Mbps off-peak and 200 Mbps on-peak between Southampton and between Southampton and ULCC in London. The maximum difference of 4 Mbps for actual network upload time is within 99% of the confidence interval to the on-peak hours. If both network upload time has a difference of more than 4 Mbps, the entire BIaaS process will start again to ensure there is a consistency with network speed.

When the upload network speed reduces, checks will be carried out to see whether network speed is slow at all sites, particularly within ULCC. If there is more than one place having slower network speed and are within 4 Mbps difference to each other, experiments can continue. But if there is only one place with a slower network, then the entire experiment will halt until the network upload time is back to normal and within 4 Mbps of each other.

- Library or software dependency: It is useful to check with any existing API or tools that have library or software dependency. If there is, all updates will take place and the system reboot will ensure there is no any influence on the execution time and processing of APIs.

5.4 The execution time for running BIaaS dual-service in the local environment

This section describes the execution time for using BIaaS services, with the objective to demonstrate that BIaaS is efficient, quick and accurate to produce good-quality results. The first step is to test the execution time in each API in the local environment where risk concerning with performance of experiments is not a concern. It can be done on either server 2 at the University of Southampton or 1 of HPC servers at ULCC. Results are running one hundred times to get the average execution time. The standard deviation is always 0.10 and below and p value is less than 0.005 presented in Table 2.

Table 2: The execution time for each API or process in the local environment ($p < 0.005$)

API or process	Average execution time (seconds)	Standard deviation	Additional comments
ItoProcess	1.23	0.05	This process calculates key variables in the Heston Model.
BlockRandom	1.07	0.03	This process computes stochastic differential equation for the Heston Model.
CovarianceFunction	1.31	0.02	This process calculates implied volatility.
Plot	1.35	0.04	This process can plot all computed data and is completed in a short period of time.
FinancialData	2.04	0.10	It takes slightly more time to retrieve the data from Google.
TradingChart	1.11	0.03	The execution time to display the financial data in charts is short and almost available instantly after the command line.

5.5 The execution time for running BIaaS dual-service between Southampton clusters

There are two sites that can process BIaaS fully, and one site is located at the University of the Southampton (server 2) and one site is at ULCC (HPC servers) according to Figure 15. There are two additional experiments required. The first experiment is to make a request from server 1 to server 2 within the University of Southampton. The physical location between server 1 and 2 is about 100 meters and the network upload speed is 1 Gbps during the time experiments took place. The second experiment is to make a request in Southampton and process in ULCC in London and will be presented in the next section. The aim is to test the execution time while network speed becomes an influential factor. Results are running one hundred times to get the average execution time. The standard deviation is always 0.12 and below and p value is less than 0.005 presented in Table 3.

Table 3: The execution time for each API or process in the local environment ($p < 0.005$)

API or process	Average execution time (seconds)	Standard deviation	Additional comments
ItoProcess	1.30	0.07	This process calculates key variables in the Heston Model.
BlockRandom	1.16	0.03	This process computes stochastic differential equation for the Heston Model.
CovarianceFunction	1.38	0.04	This process calculates implied volatility.
Plot	1.42	0.04	This process can plot all computed data and is completed in a short period of time.
FinancialData	2.12	0.12	It takes slightly more time to retrieve the data from Google.
TradingChart	1.19	0.06	The execution time to display the financial data in charts is short and almost available instantly after the command line.

Although the average execution time is slightly higher than running in local environment, the processing of API and the delivery of services still short execution time.

5.6 The execution time for running BIaaS dual-service between Southampton and ULCC London clusters

This experiment is to make a request in Southampton and process in ULCC in London. The physical location between server in Southampton and ULCC is 100 miles and the network upload speed is 100 Mbps during the time experiments took place. Results are running one hundred times to get the average execution time. The standard deviation is always 0.12 and below and p value is less than 0.005. See Table 4.

Table 4: The execution time for each API or process in the local environment (p < 0.005)

API or process	Average execution time (seconds)	Standard deviation	Additional comments
ItoProcess	1.48	0.09	This process calculates key variables in the Heston Model.
BlockRandom	1.34	0.06	This process computes stochastic differential equation for the Heston Model.
CovarianceFunction	1.57	0.07	This process calculates implied volatility.
Plot	1.59	0.08	This process can plot all computed data and is completed in a short period of time.
FinancialData	2.31	0.18	It takes slightly more time to retrieve the data from Google.
TradingChart	1.37	0.08	The execution time to display the financial data in charts is short and almost available instantly after the command line.

Results show that despite of the network speed and physical distance difference, the difference in execution time is still small comparing execution time in Table 3. The APIs are designed not entirely to rely on network speed for service delivery and network speed is useful to send back results from server to the client. The emphasis of the API is designed to use mathematical formulas for computation.

Additional demonstrations and tests will be described in Section 6 and 7 based on the hardware infrastructure presented in this section.

6. The first service of BIaaS: Realisation of Heston's Stochastic Volatility Model

This section describes the first stage of BIaaS, focusing on Heston's Stochastic Volatility model. There are two ways of using this service. The first way is to calculate the volatility and compute them in 3D Visualisation. The focus is to investigate the movement of the volatility with the respect of time and percentage of profitability. This requires calculating the Stochastic Volatility to generate all possible cases and data points. The use of BIaaS then presents all these calculations as 3D Visualisation. There are two scenarios used to demonstrate this service.

6.1 Tracking prices and volatility simultaneously

BIaaS can calculate the implied volatility with the respect with time and profitability. Calculating the implied volatility is important because this risk factor varies with time and external conditions, and often it cannot be managed and controlled with tools and policies, as it acts as a systematic risk [14, 15]. According to Sharpe [31] in his Nobel prize speech, there are two types of risks. The first type is systematic risks which cannot be managed and controlled similar to the happening of earthquakes and epidemic. The second type is the

unsystematic risks, which can be managed and minimised. The implied volatility belongs to the first type of the risks. There are still ways to reduce the impacts caused by the implied volatility. This is an important aspect for risk engineering as innovative ways should be explored to detect the likelihood of extreme events to happen, and also recommendation to reduce the damaging impacts when they happen [1, 32].

Calculations of the implied volatility can be done by applying the Heston Model, which has the advantage over the Black Scholes Model (BSM) to calculate the implied volatility for normal and extreme conditions. The Heston Model uses the Stochastic Volatility (in the CIR process) to calculate the implied volatility as discussed in Section 3.3.

This section describes how to track prices and volatility simultaneously which has become more important for the finance sector [33]. As discussed in Section 4.4, an API is required for the Heston Model. We have an API known as "ItoProcess" to use the Wiener process to help to calculate key variables (x and t, see the IToProcess API) for the Heston Model, and later on can compute the implied volatility when values for key variables are known. The API can be used like a command-line as follows.

ItoProcess[sdeqns, expr, x, t, wdproc]

This command represents an Ito-process specified by a stochastic differential equation "sdeqns", output expression "expr", with state x and time t, driven by w following the process "dproc".

ItoProcess[..., {x, x0}, {t, t0}]
This is another command uses initial condition $x(t_0) = x_0$.

The implementation of this model requires the following procedures:

Step 1: Define the "IntoProcess" and use a correlated 2D Wiener process to define a Heston Model by SDEs. It looks like:

```
hm = ItoProcess[{
    ds[t] == μ s[t] dt + √ r[t] s[t] dwₛ[t],
    dr[t] == κ (θ - r[t]) dt + ξ √ r[t] dwᵥ[t]},
    {s[t], r[t]}, {{s, r}, {s₀, r₀}}, t, {wₛ, wᵥ} ≈ cW[ρ]];
```
                                                                            (a)

Step 2: Use a stochastic Runge-Kutta method to simulate the Heston model for Year 2012. It looks like:

```
td=BlockRandom[SeedRandom[2012];
RandomFunction[hm/.{μ→0,κ→2,θ→1,ξ→1/2,ρ→-
1/3,Subscript[s,0]→25,Subscript[r,0]→1.25},{0,1,0.005},6,Meth
od→"StochasticRungeKutta"]];
```
                                                                            (b)

These two steps calculate key values and then compute for the Heston Model.

Step 3 involves with the followings:

- Collect computed data in (b).

- Use a specific plotting function to plot the computed data (one for asset prices and one for volatility). Specify any additional input values for asset prices and volatility.

- Make both visualised results on the same row.

By doing so, the Heston Model can present both asset prices and volatility. The x-axis is the time for up to 1 year and is the same for both assets and volatility. The higher the volatility is, the less stable (higher risk) that the asset is liable to. The asset price is its market price with the respect with time. The Heston Model allows tracking both prices and volatility simultaneously for the investor's stock option in 2012. See Figure 2.
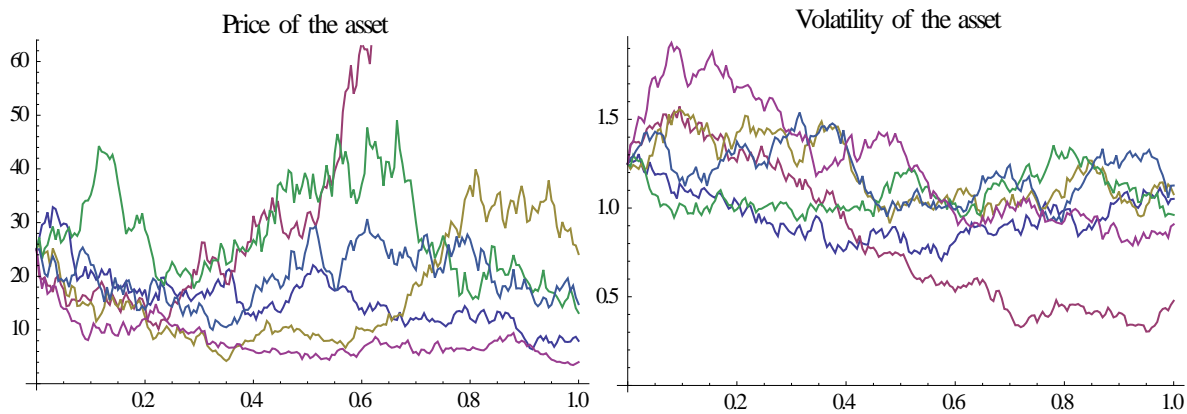

Figure 2: Tracking both prices and volatility simultaneously for a chosen stock in 2012

6.2 Calculations of the implied volatility and the API

As discussed in Section 4.4, an API is required for the Heston Model. We have an API known as "ItoProcess" to use the Wiener process to help to calculate key variables (x and t, see the ItoProcess API) for the Heston Model, and later on can compute the implied volatility when values for key variables are known. The API command-line is the same as described in Section 6.1, which already calculates the volatility (unsystematic risks, which can be reduced) and asset prices for a chosen stock. In this section, we aim to calculate implied volatility, which is classified as the systematic risks which cannot be controlled. The difference between implied volatility and volatility is that one is systematic and the later is the unsystematic risk. Apart from "ItoProcess", another command-line oriented API is "Covariance", and its usage is explained as follows.

Covariance[dist, i, j]
This command provides the covariance for the multivariate symbolic distribution dist.

The implementation of this model requires the following procedures:

Step 1: Define the "ItoProcess"

```
proc = ItoProcess[dx[t] == -x[t] dt + √(1 + x[t]^2) dw[t],
  x[t], {x, 1}, t, w ≈ WienerProcess[]]
```

Step 2: Use CovarianceFunction

```
CovarianceFunction[proc,s,t];
```

This step calculates the required values for key variables, s and t, and have generated computed data.

Step 3: Similar to Section 6.1, this step collects computed data for s and t and plots them. Calculations for the implied volatility by the Heston Model are in Figure 3, which can be rotated for 90 degrees as shown in Figure 4.
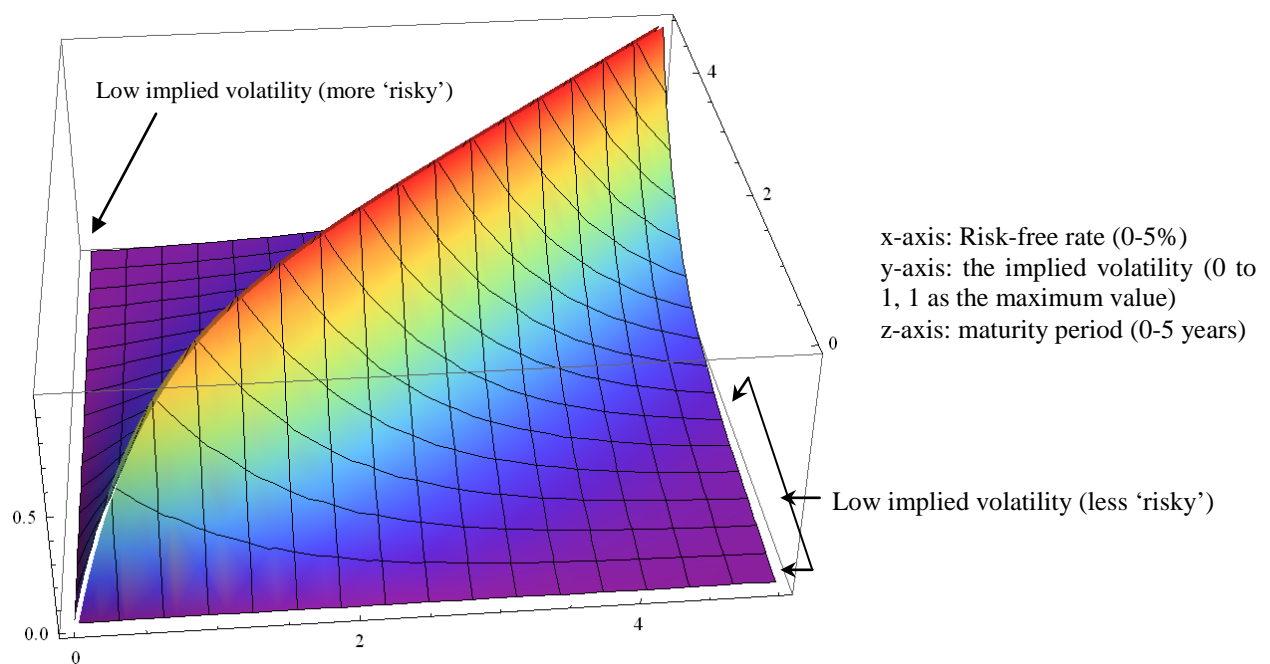

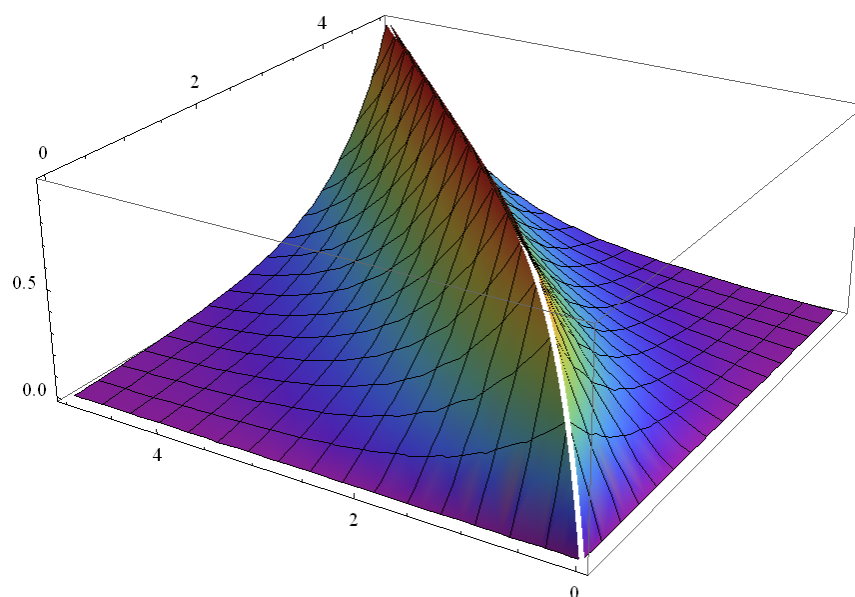Figure 3: Calculations of the implied volatility in the Heston Model (original)

Figure 4: Calculations of the implied volatility in the Heston Model (90 degrees)

Implications from results:

Both figures show that the implied volatility can reach its peak value at 1 (100%) at any period between the time of investment and four years from the investment, and also between risk-free rates between 0% and 5%. The shape of the modelling results look like a mountain centred between two plains – with higher levels of grounds towards the middle of Figure 3. This means that

- When the risk-free rate is low, it is very likely to get a high implied volatility at a short period of time. When the risk-free rate is improved to 4 and above, it allows the investment to hold on a longer period before the rise of the implied volatility.

- The implied volatility can be nearly at zero: (a) if the risk-free rate is high (5%) but the stock has to be sold within two years, or (b) the stock has to be sold at the end of four years with the risk-free rate is equal to zero. However, Figure 3 (original model) still confirms the option (a) is better as the option (b) needs a longer waiting period after going through a phase of high implied volatility.

Risk visualisation can help investors and analysts to identify when and how risks will be at the peak and the ideal period to avoid higher risks. This provides them useful information for their investment, such as the ideal times to sell or buy stocks due to the change of implied volatility, risk-free rate and time.

7. The second service of BIaaS: Business analytics of stock market performance

This section is aimed to describe the benefits the business analytics of stock market performance, which enabled by two APIs and belongs to the second type of services offered by BIaaS. Business analytics uses two APIs to enhance the visualisation of existing service and also allow analysts to obtain any stock market analysis. There are also demonstrations for how to use commands lines offered by these two APIs.

7.1 Two Visualisation APIs used as command lines

As explained in Section 4.5, two APIs, "FinancialData" and "TradingChart", are used to allow analysts and computer scientists to compute financial data to be presented in Visualisation. The focus for this paper is not to discuss what makes these two APIs but how they can be used to maximise the benefits of adopting BIaaS services.

- FinancialData – this command can capture financial data from Google Finance and make data to be understood by the private cloud to be ready for data computation and analysis. It follows the command-line approach which allows users to provide one sentence of line of code to obtain the data.

The usage scenario for the command line includes the followings:

FinancialData["name", {start, end, period}]

This command provides a list of dates and prices for the specified periods lying between start and end.

FinancialData["name", "prop", {start, end, ...}]

This command provides a list of dates and values of a property for a sequence of dates or periods.

- TradingChart – this command can compute the data obtained from "FinancialData" command and present them in a way similar to software provided for London Stock Exchange (LSE). Investors can see business analytics of their selected stock.

The usage scenario for the command line includes the followings:

TradingChart[{"name", daterange}]

This command makes a financial chart for the financial entity over a specific range of data.

TradingChart[{...}, {ind1, ind2, ...}]

This command makes a financial chart with indicators, which provides additional analytics functions such as plots and micro analysis.

There is a scenario in the next section to explain how all these commands can work to deliver a solution.

7.2 A scenario to explain how Visualisation works

This is a scenario to explain how to obtain the financial analytics computed by the private cloud. For example, an analyst would like to find out the stock market performance for Facebook (FB) for the past thirteen months after its initial public offerings (IPOs) between mid-May 2012 and the end of June, 2013. The command line is

data = FinancialData["FB", "OHLCV", {{2012, 5, 18}, {2013, 6, 25}}];

This command can obtain the data from Google Finance and get data ready for "TradingChart" computation. "OHLCV" means open, low, high, close and volume for the selected stock, which can be obtained from Google Finance to be presented for FinancialData.

TradingChart[data]

This command will present the Facebook stock data in a trading chart format. See Figure 5.

Figure 5: Trading chart for Facebook since its IPOs

The next step requires analytics of its recent stock market performance, which means the focus is within the most recent three to four months.

TradingChart["FB", {"Volume", "RelativeStrengthIndex"}]

This command can provide business analytics for FB, which includes

- displaying trading volume for FB;

- displaying Relative Strength Index for FB;

- offering R-squared values, which are used to determine how fit the data is to the overall plot, and is between -1 (rare) and 1. The higher the value, the closer the data is to the overall plot.

- calculating beta, which represents systematic risk value.

- providing an optional simplified plot to make analysis much easier for some investors.

See Figure 6 for details.

Figure 6: Trading chart for Facebook with business analytics

7.3 Tests for accuracy for a particular time period

Section 2 describes the problem faced by the finance industry and accuracy in business analytics is important to ensure results have a high quality and are very close to the actual values in asset prices and risks. There are two types of tests to be performed, and each test is designed for each type of BIaaS service. The second type of BIaaS services is easier to test since it retrieves data from Google Finance, and the output from "TradingChart" can compare results with London Stock Exchange, and also Yahoo UK Finance. Results are exactly the same and thus have 100% accuracy. The first type of BIaaS services can be tested with the description as follows.

- Calculate all values for asset prices, volatility, implied volatility and risk-free rate with respect to the time, and any dependencies between them. Three different stocks are chosen and all calculations are for each of three stock.

- Record down the value, and wait until the time lapse.

- Compare the computed results with actual results, and asset prices, volatility, time and risk-free rates (by the bank that offers the credit guarantee) are good indicators for comparison, particularly asset prices (all other three variables have the same values) since they are often used to indicate the values determined by the market.

- The difference between the computed and actual results should be as close as possible, and should be aimed for within 5% difference.

Three sets of results are recorded and compared as follows. See Table 5.

24

Table 5: Three sets of results (computed and actual results) to test accuracy

Sets of results	Computed results (Asset price as the main indicator)	Actual results (Asset price as the main indicator)	Difference (in numeric values and percentage)	Within 5% acceptance range?
1 (Stock 1)	**Asset price= 32.2;** volatility = 1.23; implied volatility = 0.45; time = 1.3	**Asset price= 32.4;** volatility = 1.23; implied volatility = 0.45; time = 1.3	0.2 (0.58%)	Yes
2 (Stock 2)	**Asset price= 17.4;** volatility = 0.78; implied volatility = 0.62; time = 0.7	**Asset price= 17.5;** volatility = 0.78; implied volatility = 0.62; time = 0.7	0.1 (0.57%)	Yes
3 (Stock 3)	**Asset price= 62.4;** volatility = 1.45; implied volatility = 0.77; time = 3.0	**Asset price= 63.1** volatility = 1.45; implied volatility = 0.77; time = 3.0	0.7 (1.11%)	Yes

Table 2 confirm that three sets of computed and actual results are close to each other and are always within 5% of difference. The difference is small when the time does not go beyond 2 years. Stock 3 is the one that has three years of investment period and has the largest difference of 1.11%. Results for a medium and long term investment may have greater differences but results are still within 5% of acceptance range. Additional supporting evidences will be demonstrated in Section 8.2 to test the validity and accuracy of BIaaS.

7.4 Summary of operating BIaaS

BIaaS allows two types of services. The first BIaaS service focuses on the Heston Model to calculate the asset price, volatility, maturity period, implied volatility and risk-free rate. The first service uses three APIs to deliver these services and detailed descriptions are given to demonstrate how APIs can be used and the benefits of using the first BIaaS service. The second BIaaS service focuses on using the open finance data and presenting them in visual formats similar to the LSE services, with the aid of another two APIs. Both services have usage scenarios to support how functionalities can meet the requirements for business analytics and provide analysts and investors high-quality of risk and pricing modelling.

Additional tests are carried out to test the performance and accuracy for BIaaS. Firstly, the execution time to run all these APIs is considerably fast. Secondly, the second BIaaS service has 100% accuracy as results are the same as LSE and Yahoo Finance. The first BIaaS service has three sets of computed and actual results (each set is a unique stock) and results are compared, where all three sets are less than 5% of acceptance range of difference.

8. A Full Case Study of using BIaaS service

The purpose of this case study is to demonstrate that experiments and results can be repeated in a similar environment and process. There are two experiments. The first set of experiment is to test the calibration of the Heston Model, where the formulas for the model is described between Section 3.2 and 3.4, and method and formula for calibration is described between Section 3.5 and 3.7. The second set of experiment is the comparisons between the actual stock index movement and predicted movement offered by "SimpleMovingAverage" and "BollingerBands". The first set of experiments is to validate historical results computed

BIaaS is accurate, where our contribution is the improved performance over the traditional practice of financial modelling on desktop. The second set of experiments is to validate current and forecasted results computed by BIaaS is close to real market values, which is unique from existing literature [16, 22, 23, 34, 35, 36] and contribute to current research.

8.1 The first set of experiments

VIX is a measure of the implied volatility of S&P 500 index options and it is used for the first set of experiment. The methodology is similar to calibration performed by other researchers [22, 23, 34, 35]. There are four periods of calibration with their associated parameters explained as follows.

- The first period: It is the low volatility regime period before financial crisis. The chosen date was 27th of October, 2006.

- The second period: It is when the financial crisis started, and the reaction of the market. The chosen date was 15th of September, 2008.

- The third period: It is when the financial crisis had its impacts on the Western and World economy. The chosen date was 15th of December, 2008.

- The forth period: It is when the financial crisis had calm down for the first time. The chosen date was 23rd of October, 2009.

The input parameters are calculated as a result of calibration, which then calculates output values for RMSE, APRE and APE test. These parameters are (v_0, κ, θ, ξ, ρ), which corresponds to key values in formula (2) and can be used by all formulas for calibration. Computed output values are compared with actual market values. Any differences are noted and discussed. Experiments for each calibration were performed three times to take the average execution time on the private cloud (Figure 15). Comparing to the traditional way of running simulations on desktop, private cloud helps improve performance. Standard deviation (SD) is always less than 0.1. Results are presented between Table 6 and Table 9.

Table 6: the low volatility period, 27th of October, 2006

	Calibration	MW ($T^{VIX} = 0.5$)	MW ($T^{VIX} = 3$)	MW ($T^{VIX} = 5$)	EWMA	Market-implied
RMSE	0.865433	1.705412	1.694425	1.119423	5.496902	1.149964
APRE	0.248758	0.322407	0.323417	0.259313	0.334489	0.245102
APE	0.09035	0.027005	0.022503	0.012487	0.050001	0.012846
(ν_0, κ, θ, ξ, ρ)	0.0106, 0.5671, 0.0394, 0.2114 and -0.8035	0.0125, 3.6447, 0.0205, 0.3837 and -0.9312	0.0125, 3.5789, 0.0228, 0.3832 and -0.9255	0.0125, 0.4598, 0.0404,0.1935 and -0.8825	0.0125, 3.3017, 0.0139, 0.2988 and -0.8813	0.0125, 1.4603, 0.0260, 0.2715 and -0.8341
Averaged execution time (sec)	18.44 (S.D: 0.07)	6.83 (SD: 0.03)	6.86 (SD: 0.03)	4.23 (SD: 0.03)	8.02 (SD: 0.04)	6.95 (SD: 0.04)

Table 7: the credit crisis period, 15th of September, 2008

	Calibration	MW ($T^{VIX} = 0.5$)	MW ($T^{VIX} = 3$)	MW ($T^{VIX} = 5$)	EWMA	Market-implied
RMSE	3.817004	5.388143	8.080046	8.388043	3.866422	3.849023
APRE	0.176188	0.376922	0.549589	0.563785	0.412274	0.356844
APE	0.032876	0.042756	0.065821	0.070101	0.032855	0.101516
(ν_0, κ, θ, ξ, ρ)	0.0694, 0.5671, 0.0394, 0.2114 and -0.8035	0.1108, 4.6795, 0.0514, 0.4714 and -1.0001	0.1101, 1.5322, 0.0321, 0.2859 and -0.9140	0.1101, 1.3207, 0.0285, 0.2729 and -0.9103	0.1101, 5.1093, 0.0631, 0.8034 and -0.9102	0.1101, 6.7475, 0.0602, 0.9000 and -0.8511
Averaged execution time (sec)	19.56 (SD: 0.09)	7.14 (SD: 0.04)	7.18 (SD: 0.04)	4.35 (SD: 0.03)	7.99 (SD: 0.04)	6.96 (SD: 0.04)

Table 8: The financial crisis had its impacts on the world economy, 15th of December, 2008

	Calibration	MW ($T^{VIX} = 0.5$)	MW ($T^{VIX} = 3$)	MW ($T^{VIX} = 5$)	EWMA	Market-implied
RMSE	5.005887	9.073945	6.377240	6.785792	14.720588	9.202151
APRE	0.246755	0.214123	0.630104	0.710044	9,3195442	0.214803
APE	0.027001	0.047682	0.040625	0.043455	0.084923	0.045879
(ν_0, κ, θ, ξ, ρ)	0.2405, 0.5525, 0.1274, 0.3746 and -0.9799	0.3108, 0.8397, 0.1948, 0.5713 and -0.9301	0.3108, 1.1903, 0.0580, 0.3715 and -0.9873	0.3108, 1.1604, 0.0428, 0.3168 and -0.9840	0.3108, 0.4118, 0.3761, 0.5563 and -0.9997	0.3108, 0.8401, 0.2006, 0.5809 and -0.9308
Averaged execution time (sec)	19.41 (SD: 0.08)	7.09 (SD: 0.04)	7.16 (SD: 0.04)	4.31 (SD: 0.03)	8.01 (SD: 0.04)	6.96 (SD: 0.04)

Table 9: The financial crisis had calm down for the first time, 23rd of October, 2009

	Calibration	MW ($T^{VIX} = 0.5$)	MW ($T^{VIX} = 3$)	MW ($T^{VIX} = 5$)	EWMA	Market-implied
RMSE	1.281240	1.504521	1.369142	3.035413	3.121217	1.718054
APRE	0.293871	0.257043	0.253085	0.290387	0.280411	0.307008
APE	0.011892	0.012545	0.011621	0.020458	0.020850	0.014056
(ν_0, κ, θ, ξ, ρ)	0.0435, 1.4974, 0.0920, 0.5231 and -0.8316	0.0465, 2.4484, 0.0779, 0.6185 and -0.8487	0.0465, 1.4718, 0.0893, 0.5120 and -0.8338	0.0465, 1.1012, 0.0606, 1.1523 and -0.9748	0.0465, 1.1718, 0.0601, 1.1848 and -0.9848	0.0465, 3.4359, 0.0735, 0.7072 and -0.08670
Averaged execution time (sec)	18.96 (SD: 0.08)	6.94 (SD: 0.03)	6.95 (SD: 0.03)	4.28 (SD: 0.03)	8.00 (SD: 0.04)	6.93 (SD: 0.04)

8.2 **The second set of experiments**

The second set of experiments is aimed top demonstrate that the use of Heston Model can help to forecast the stock market movement. Section 7.3 is a different test than this. Test of accuracy is provided in the way to cover more than 12 months instead of a particular time period (such as a particular day, or week). The validation is done by comparing the actual with predicted movement with the real stock options. Both "SimpleMovingAverage" and "BollingerBands" use the concept of Moving Window (MW) formula presented in Section 3.6. Facebook is used as the case study.

8.2.1 Facebook: Experiment by using "SimpleMovingAverage"

The first test is to use "SimpleMovingAverage" to compare the differences between the actual movement and forecast movement computed. BIaaS plots the average values based on the recent movements in the stock market performance by calibration. Figure 7 shows the result and blue line is the forecast movement. The green-red line is the actual movement, where the green indicates upward movement and red refers to downward movements. Both Moving Window (MW) and exponentially weighted moving average (EWMA) are used, where EWMA tracks the volatile movement of the previous time series and computes the next likely movement. The accuracy is 95% based on thousands of data point, where standard deviation is 2.5432. Some data points have achieved 100% but most of data points are within 95% of the Confidence Interval to the real data points. The command to compute the Facebook analytics is

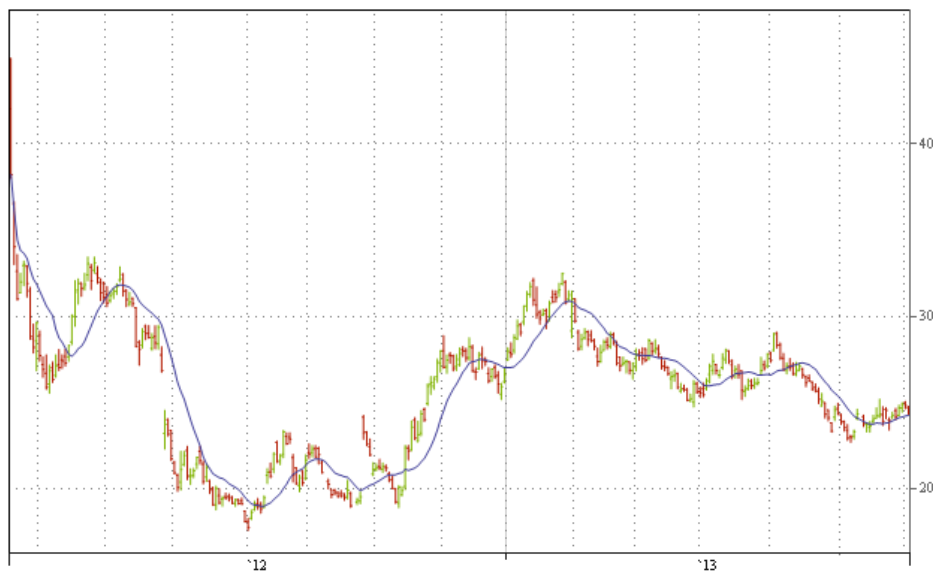*TradingChart["FB", {" SimpleMovingAverage"}]*

Figure 7: Trading chart for Facebook between actual and predicted movement offered by "SimpleMovingAverage" (between 18th of May, 2012 and 2nd of July, 2013)

8.2.2 Facebook: Experiment by using "BollingerBands"

Bollinger Bands is an analytical tool to measure mean and the upper and lower limits of the stock movement. "BollingerBands" is the function behind, and it uses the concept of Moving

Window (MW) and exponentially weighted moving average (EWMA) with 95% Confidence Interval (CI) to compute the predicted movements for the mean and lower and upper limit.

Bollinger Bands consist of the followings:

- an N-period moving average (MA) and the use of MW and EWMA

- an upper band at K times an N-period standard deviation above the moving average (MA + Kσ)

- a lower band at K times an N-period standard deviation below the moving average (MA − Kσ)

Figure 8 shows the results between the actual and predicted movements, where the middle blue line is the mean and the other two blue lines are the upper and lower limits respectively. In some volatile movements that are involved with the rapid fall and rise due to human speculation (such as some big investors speculate the market), either the upper or the lower limit fit. Comparing to the "SimpleMovingAverage", the actual prediction is slightly lower (but within 2%), however, the use of upper and lower limits can help the analysts to make a better judgement of the likely stock movement. The accuracy is 99.99% based on thousands of data points, where standard deviation is 3.6247. Almost all the data points are within the 95% of Confidence Interval to the real data points. The command to compute is

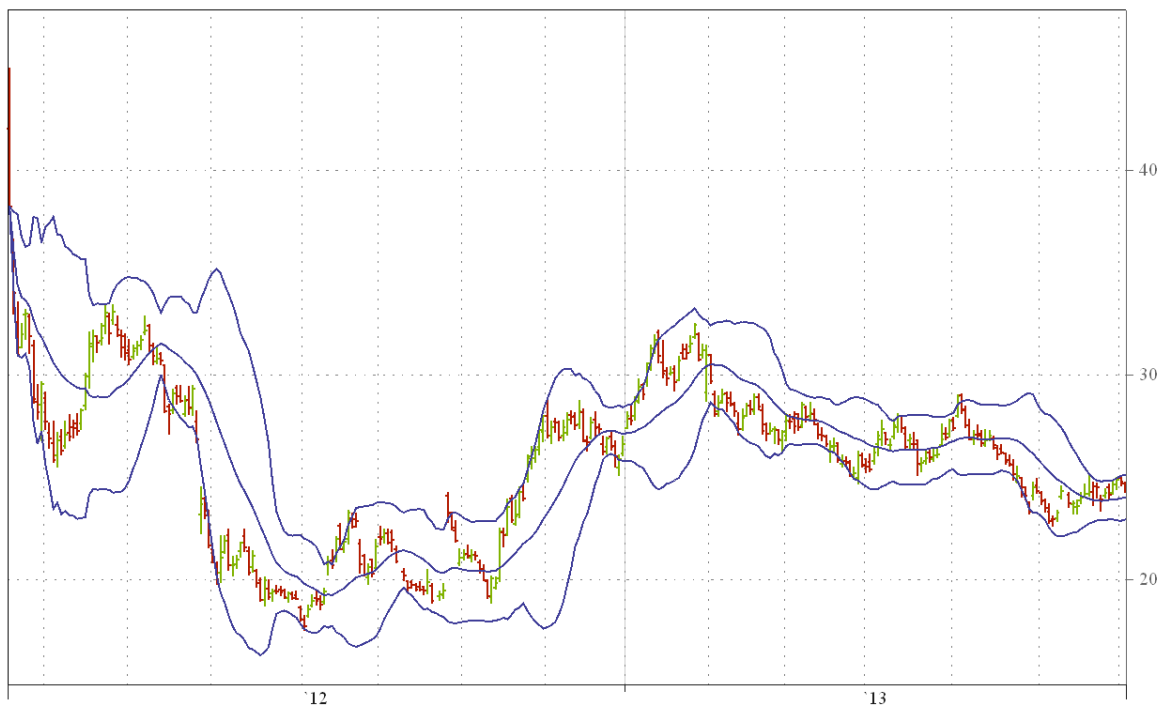***TradingChart["FB", {" BollingerBands"}]***

Figure 8: Trading chart for Facebook between actual and predicted movement offered by "BollingerBands" (between 18th of May, 2012 and 2nd of July, 2013)

The second set of experiments confirms that the actual and predicted movements are close to each other by the use of "SimpleMovingAverage" and "BollingerBands". Additional examples of using other stock options will be presented in Section 10.

9. Comparisons with other models

This section compares BIaaS with other platforms and other approaches to risk assessment and management. This is useful to understand the strengths and weaknesses of each model and understand how our model differs from others. See Table 10.

Table 10: Six elements for their relevance to Cloud adoption

Core elements	References
Usability	[24, 37, 38, 39]
Performance	[4, 14, 25, 40]
Security	[4, 38, 39, 40, 41]
Computational accuracy	[24, 25, 42, 43, 44]
Portability	[4, 14, 42, 43, 45, 46, 47, 48]
Scalability	[4, 24, 40, 48, 49]

9.1 Factors affecting Cloud adoption

These six elements are supported by literature (see Table 4) and a brief summary is described as follows:

- Relevance for usability for Cloud adoption is important with the support of real use cases [24, 37, 38, 39].

- Achieving good performance is an essential for Cloud adoption [4, 14, 25, 40].

- Security concern is a main reason for some organisations not to adopt Cloud Computing which researchers describe challenges and issues to be improved [38, 39, 40, 41].

- Computational accuracy is important to compute accurate results so that organisations have a higher trust and confidence for Cloud adoption [24, 25, 42, 43, 44].

- Service and data portability are highly relevant for Cloud adoption that researchers demonstrate their usefulness for adoption [4, 14, 42, 43, 45, 46, 47, 48].

- Scalability is a core characteristic for Cloud and the ability to scale up and down resources promptly for different demands is essential for Cloud adoption [4, 24, 40, 48, 49].

The next step is to use these six criteria to rate the platforms for BIaaS and desktops. See Table 2 for the explanation for the rating. The score rating is expanded on Hosono's work [50, 51] and our own scoring criteria as follows:

- A score between 1 and 3 is considered as a poor standard.

- A score between 4 and 6 is considered as average.

- A score of 7 is considered as satisfactory.

- A score of 8 to 10 is considered as excellent.

9.2 The expert review

Eleven experts in business intelligence and related areas are asked to rate the BIaaS and desktop applications by using the six criteria. They had extensive experience in using BIaaS

and desktop systems. They had been interviewed face-by-face or by telephone or by Skype, and were asked about their scores and supporting rationale for each criteria [52]. Their rationale and answers are summed up and explained as follows.

Business Intelligence as a Service (BIaaS), score is out of 10

- **Usability**: Most of BIaaS APIs are easy to use except one API requires further training. The overall score is 8 because at last 80% of the tools are easy to use and their manuals are self-explanatory. The other 20% of the functionalities require specialised knowledge about financial modelling to compute complex models.

- **Performance**: Performance on BIaaS is good. Computation takes a short time to get results. The score is 8.

- **Security:** BIaaS needs third party software and is not a model with a high level of security. Basic authentication and authorisation can still be achieved. As a result, the score is 4.

- **Computational accuracy:** Computational BIaaS results are accurate. Some banks have used BIaaS to calculate pricing and risks, and are close to the actual values. But BIaaS requires have accurate input values before getting the final results. This level of dependency is a limitation to prevent it to score 10. The overall score is 8.

- **Portability:** BIaaS is highly portable in most of the systems. All operating systems and computational devices can run BIaaS applications. The overall score is 9.

- **Scalability:** BIaaS tools are highly scalable. It can run on a single processor desktop, or clusters of high-end servers. Input variables can be highly adaptable to a wide range of values. Thus, the overall score is 9.

Desktops and desktop-based applications, score is out of 10

- **Usability**: There are several Excel or desktop-based tools which are available to financial services and are easy to use. There are enterprise systems such as Reuters and Bloomberg to offer similar services, although further training is required prior using these systems. The overall score is 8 out of 10.

- **Performance:** Performance on desktop-based system is good. Computation takes a short time to get results. The score is 8.

- **Security:** Desktop-based software need third party software and is not a model with a high level of security. Basic authentication and authorisation can still be achieved. As a result, the score is 4.

- **Computational accuracy:** This has a divided opinion among experts. Seven experts said that there is a high level of computational. The other three pointed out that the only limitation is that these applications are not designed to forecast risks. One example is that these applications could not forecast financial crisis happened in 2008 and could not calculate an alternative solution. However, applications in Bloomberg and Reuters have been market over 20 years and most of times calculations are accurate and reliable. Accuracy is high most of times and the average rating has a score of 8.

- **Portability:** Most of desktop applications work on Windows and need additional set ups. They have specific requirements for services. For example, if a digital certificate or licence key is not present, the entire application will halt. Not all applications can run on other platforms fully. This makes desktop applications less portable on some systems. However, some high-end applications have "apps" on mobile devices. The

percentage of easy-to-use application is still low in 2013 although situations may improve in the next two years. Based on all ratings, the overall score is 6.

- **Scalability:** Some desktop applications need specific set ups and input variables need to be defined prior computation. This makes desktop applications less scalable due to constrained options, although they can function well on desktops or clusters of high-end servers. The score is 4.

Based on the expert review, the overall rating for BIaaS and desktop-based tools are presented in Table 4.

Table 4: Using these six criteria to rate the platforms for BIaaS and desktops

Platforms Criteria	BIaaS	Desktops [54, 54]
Usability	Specially designed for finance sector. Good usability. Score: 8	Excel and desktop-based software has good usability. Score: 8
Performance	Good. Score: 8	Good. Score: 8
Security	Need third-party for security or make software to have authentication. Score: 4	Need third-party for security or make software to have authentication. Score: 4
Reliability	Reliable to get accurate calculations. Score: 9	Reliable to get accurate calculations. Score: 8
Portability	Can be used in different platforms and applications. Score: 9	Desktops and a few systems. Score: 6
Scalability	Can be scalable. Score: 9	Less flexible for scalability. Score: 4
Total score	47/60	38/60

9.3 Rationale: the score for Desktops

Business Experiments in Grid (BEInGRID) is an EU-funded project in 2009 to understand how businesses use Grid and Cloud computing using thirty-five case studies [43]. There are three financial modelling projects with desktop approaches; Financial Portfolio Management (FPM), Risk Management in Finance (RMF) and Anti-Money Laundering in Grid (AMONG).

Grid Service Company (GSC) has developed Financial Portfolio Management (FPM) software for finance industry which uses Grid Resource Broker technologies (GRB) to create "Grid of Grids". The objective is to hide all the complexity and enable customers to run complex applications as if they were using their desktops. FPM offers Excel-like services to allow customers to use financial functionalities and analysis with ease without any need for programming. They provide statistical analysis to review clients' financial performance. GSC claims that the use of the software allows their clients to analyse complex data whilst at the same time providing savings of time, resources and operational costs. FPM provides a basic level of risk modelling but not the comprehensive risk modelling offered by BIaaS.

The AXA Europe Group has developed Risk Management in Finance (RMF), which is based on American/Bermudian Monte Carlo approach. Their objective was to determine the price of the guarantee by taking into account the hidden option to the deferral period. They allow clients to choose and adapt the payoff and to have designed suitable algorithm to determine the frontier for the optimal exercise. They define a parallel algorithm to price and compute the embedded option on the Grid which includes the following steps:

- American/Bermudian style high dimension Monte Carlo methods techniques.
- State vector of key information and Option Value.
- Choice of a suitable algorithm to determine the optimal frontier of exercise.
- Pricing and sensitivities calculation for hedging.

This work is especially targeted at hedge fund managers and the software assists them to understand and analyse their hedging strategy. RMF software is easy to use but only provides 2D analysis aimed at desktop users. It does not offer the detailed risk modelling offered by BIAAS. Anti-Money Laundering in Grid (AMONG) is interested in an aspect of Grid and Cloud computing unrelated to the type of assessment carried out by BIaaS.

9.4 Summary of Comparisons

The features that set BIaaS apart from other related work are that it uses a structured approach which focuses on modelling of risks directly leading to an easily interpreted visual result. Taking this direct approach means that it need not consider business processes, accounting, brokering and so on. A second consequence is that its results are independent of the state of the economy, unlike other approaches which implicitly assume continuing favourable trading conditions.

10. Discussions

The key message of Cloud adoption and using BIaaS is that Cloud offers better performance, better efficiency, lower costs and better integration with other services. There are three topics of discussions to support the validity of BIaaS.

10.1 Summary of BIaaS work and key lessons learned

Section 3 explains all associated formulas related to the Heston Model and their APIs for stochastic differential equations and formulas for calibration. Section 4 explains the system architecture and operations of BIaaS in terms of how each API works and how they can deliver a joint service. Section 5 presents the technical sets ups and the execution time for processing each API. Descriptions for technology selections in building private cloud and minimisation of risks such as network latency are provided. All the execution time for running each API is between 1.07 seconds the lowest and 2.31 seconds the highest. Results also show that processing APIs on the Southampton and London private cloud are faster than the counter part on the desktop.

Section 6 and 7 demonstrate how each service in BIaaS can work. Section 6 explains the first part of BIaaS service in the use of each API and reinforces their corresponding relations to the selected formulas described in Section 3. Experiments are performed to show that risks can be calculated and presented in visualisation. A challenge is to track the implied volatility. The use of the Heston Model and its APIs can calculate the implied volatility and also interprets the implications from its 3D visualisation outputs. Section 7 presents the second part of BIaaS service in the use of API for visualisation. Results can offer the trading movement of the stock, including the trading volume, relative strength index, R-squared values and beta. Test of accuracy is provided in the way that when all the parameters, except the asset price is the same, comparisons between the actual and predicted movements are made. However, this test only focuses on each particular instance and not the whole year.

Section 8 presents a full case study for demonstrating BIaaS. The first set of experiments explain how to operate perform calibration, including the parameters for VIX model and validated results. Comparing to other research analysis, the calibration has a better

performance for achieving shorter execution time for calibration. The second set of experiments presents two other stock options that use "SimpleMovingAverage" and "BollingerBands". Results show that the predicted stock index movement is approximately 95% to the actual index and 99.99% to the actual index if the upper and lower limits of 95% of Confidence Interval are included.

Other researchers can follow the steps in Section 3 and 4 to design their own API or application, and Section 5 for the hardware infrastructure. They can use the methodology presented in Section 6 and 7 to develop their own applications that can work in similar environments. Descriptions in Section 8 can support validity and originality of BIaaS. The first set of experiments is used by researchers and the benefit of using BIaaS in the Cloud is the better performance in the form of shorter execution time. The second set of experiments confirms that results computed the Heston Model and their associated APIs can get a close estimation of 95% and 99.99% (including 95% C.I) to the actual stock index values.

The demonstration of BIaaS also fulfils the requirements described in Section 2.3 as follows. BIaaS is based on the Heston Model, which uses stochastic differential equations to model complex financial parameters and modelling. APIs developed by the Heston Model can compute risk and track volatility in the form of the Business Analytics as a Service (BAaaS). BIaaS has a good extent of accuracy in two ways. The first way is to calculate asset prices at a particular instance of time and has between 0.57% and 1.11% difference between the actual and computed results. The second way is to compute the entire stock index movements for Facebook, Google and Apple, and there is about 95% (actual) and 99.99% (95% C.I included) of accuracy between the actual and computed values. BIaaS can work on the private cloud which has relatively better performances than desktop. Each API takes between 1.34 and 2.31 seconds. All the complex calibration processes (before computing forecast index movements) take less than 19.56 seconds, where thousands of simulations have been running behind the scenes.

10.2 Other examples to compare actual and predicted movement

Section 8.2 presents the case study of Facebook to review its actual and predicted movement since its IPO in mi-May 2013. Both "SimpleMovingAverage" and "BollingerBands" offer close estimation to the real stock index and has the correct prediction rates of 95%. It is also important to use other stock options to support the originality and accuracy of predicted movements. Additional two examples are presented.

10.2.1 Google: Experiment by using "SimpleMovingAverage"

Google is chosen as an example as its stock index movement has improved from approximately 600 to 900 Dow Jones index in the past twelve months. It is a good example to test whether exponentially weighted moving average (EWMA) can accommodate the upward volatile movement (steep and steady rise) over a period of twelve months. Figure 9 shows the comparisons between the actual (green-red line) and predicted (blue line) stock index movement for Google.
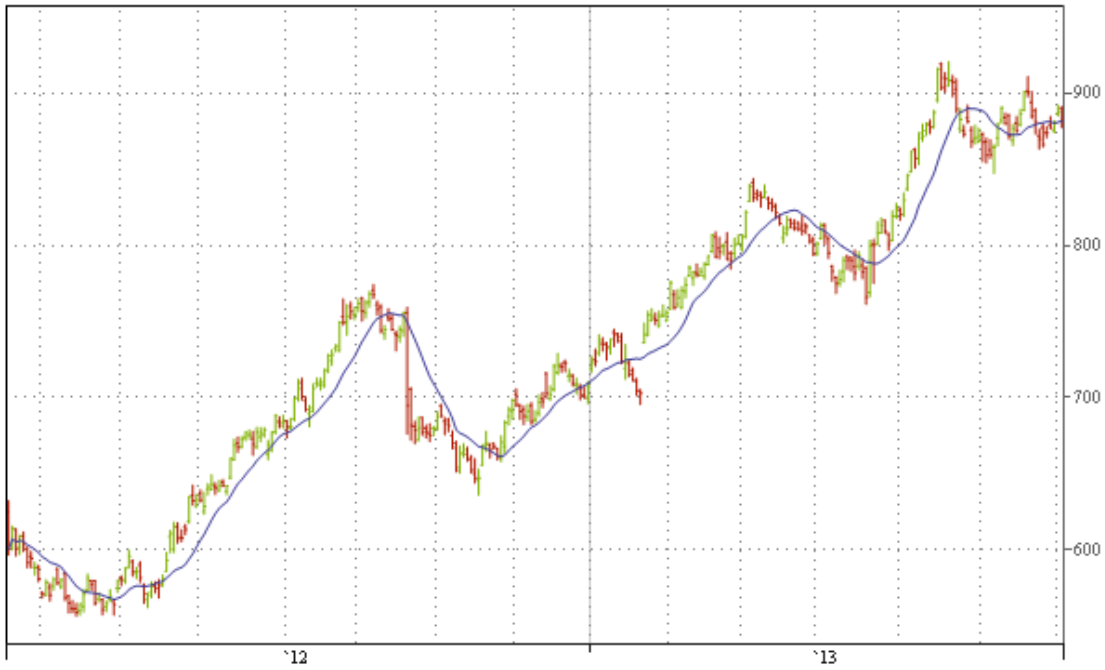
Figure 9: Trading chart for Google between actual and predicted movement offered by "SimpleMovingAverage" (between 18th of May, 2012 and 2nd of July, 2013)

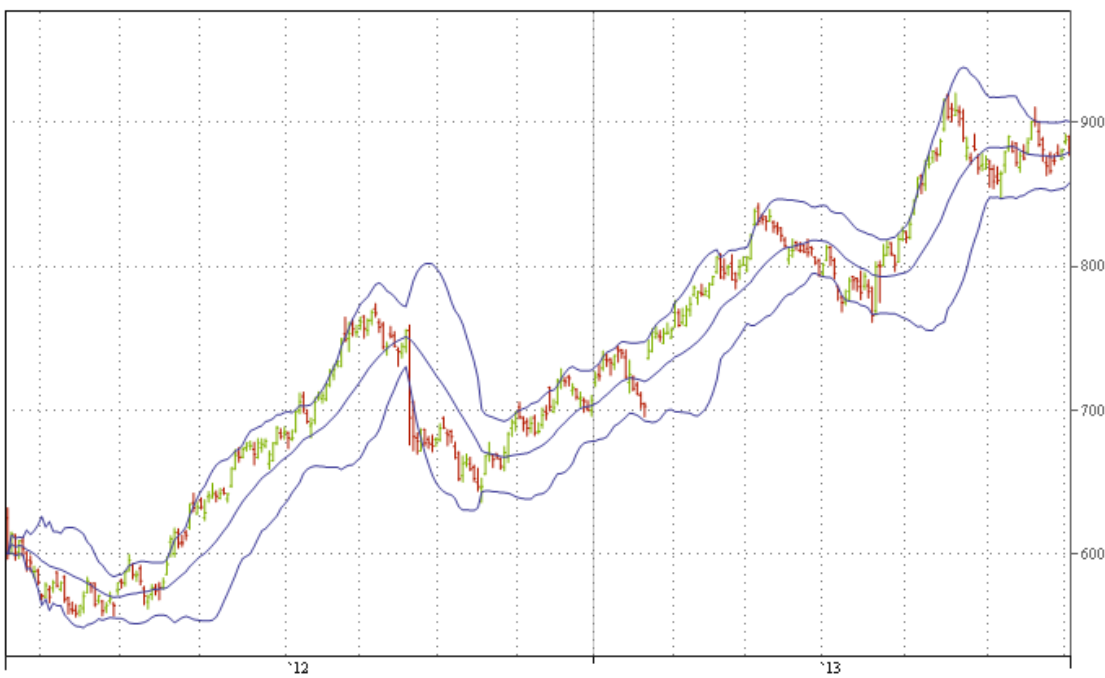## 10.2.2 Google: Experiment by using "BollingerBands"

Figure 10: Trading chart for Google between actual and predicted movement offered by "BollingerBands" (between 18th of May, 2012 and 2nd of July, 2013)

Figure 10 shows the comparisons between the actual and predicted stock index movement for Google while using "BollingerBands". A significant 99.99% of accuracy is observed when the upper and lower limits are included to calculate the predicted movements.

10.2.3 Apple: Experiment by using "SimpleMovingAverage"

Google is chosen as an example as it experienced both ups and downs. The rapid fall of its stock index fell in general since September 2012. It is a good example to test whether both MW and EWMA can follow both upward and then lengthier downward direction. Figure 11 shows the actual and predicted stock index movement for Apple. The accuracy of predicted movement is maintained at 95% when majority of data points fit in the overall movement.
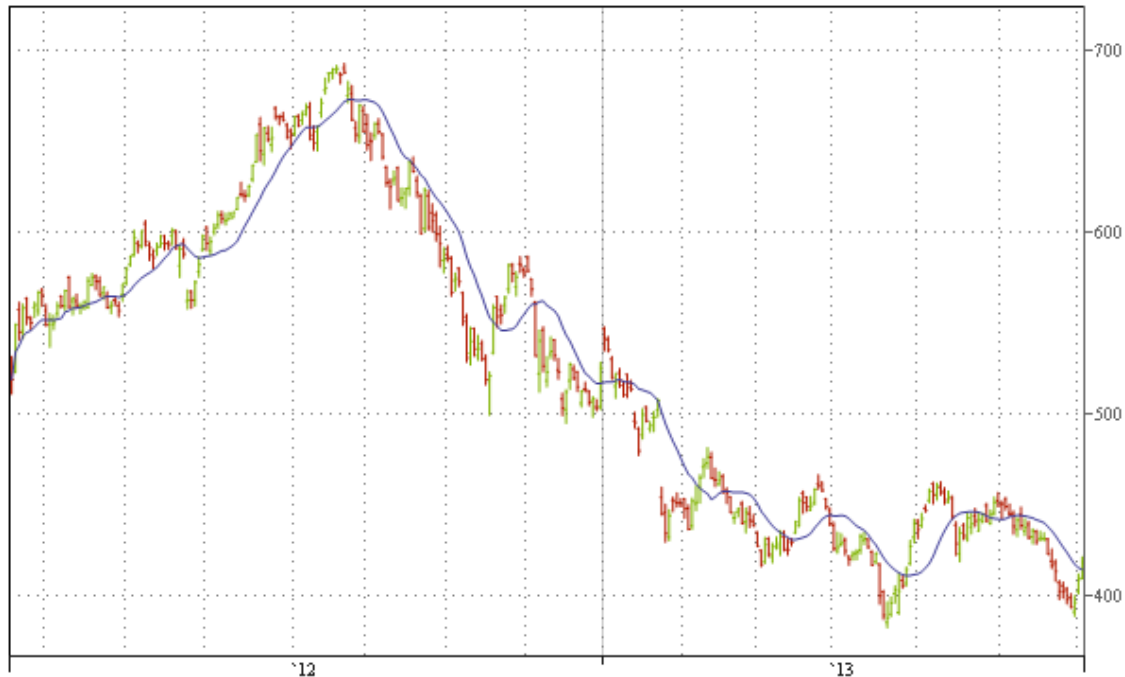

Figure 11: Trading chart for Apple between actual and predicted movement offered by "SimpleMovingAverage" (between 18th of May, 2012 and 2nd of July, 2013)

10.2.4 Apple: Experiment by using "BollingerBands"

Figure 12 shows the comparisons between the actual and predicted stock index movement for Apple while using "BollingerBands". Similar to Section 10.2.2, 99.99% of accuracy is observed when the upper and lower limits are included to calculate the predicted movements.
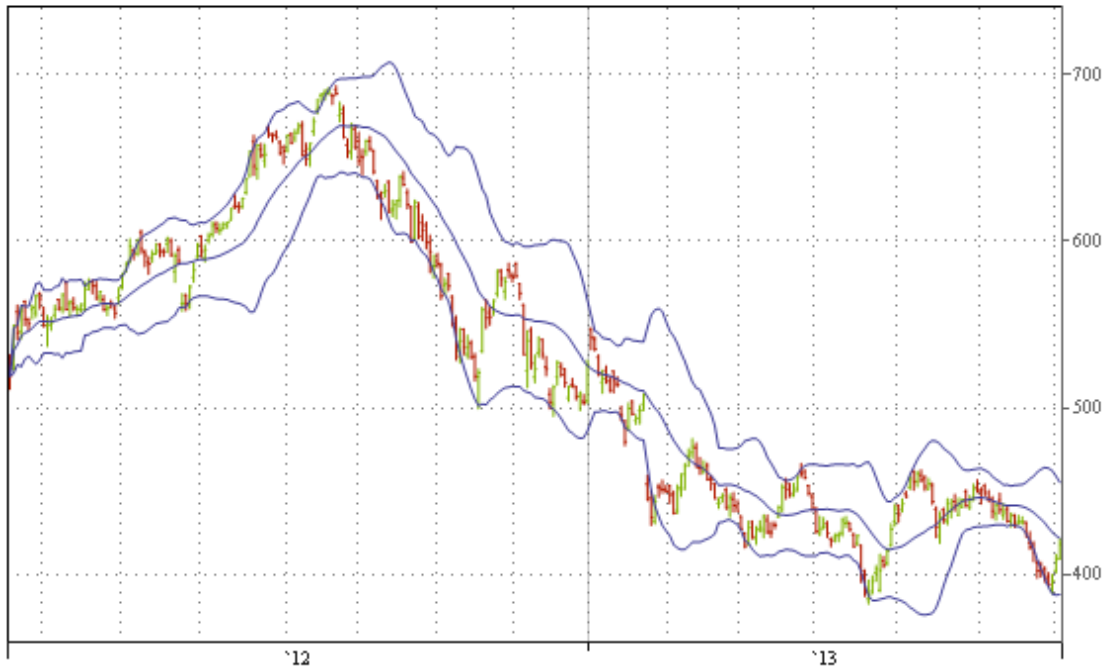
Figure 12: Trading chart for Apple between actual and predicted movement offered by "BollingerBands" (between 18th of May, 2012 and 2nd of July, 2013)

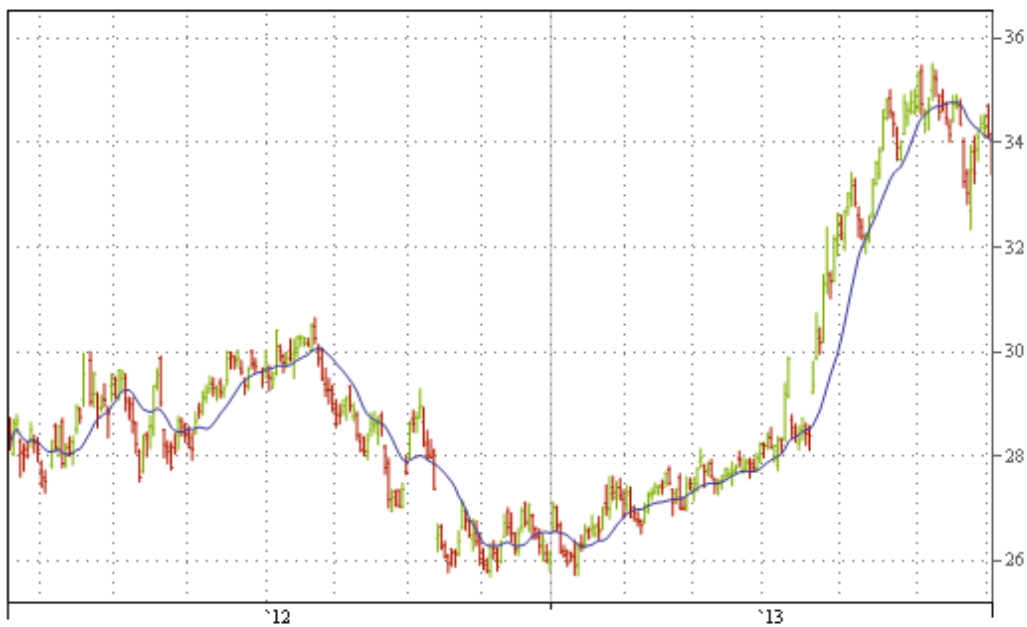## 10.2.5 Microsoft: Experiment by using "SimpleMovingAverage"

Figure 13: Trading chart for Microsoft between actual and predicted movement offered by "SimpleMovingAverage" (between 18th of May, 2012 and 2nd of July, 2013)

Microsoft is another example to support the validity and accuracy of BIaaS service. Figure 13 shows the actual and predicted stock index movement for Microsoft. Similar to previous examples, the accuracy of predicted movement is maintained at 95% when majority of data points fit in the overall movement.

10.2.6 Microsoft: Experiment by using "BollingerBands"

Similar to the previous examples, Figure 13 shows the comparisons between the actual and predicted stock index movement for Apple while using "BollingerBands". When the upper and lower limits are included to calculate the predicted movements, 99.99% of accuracy is observed.
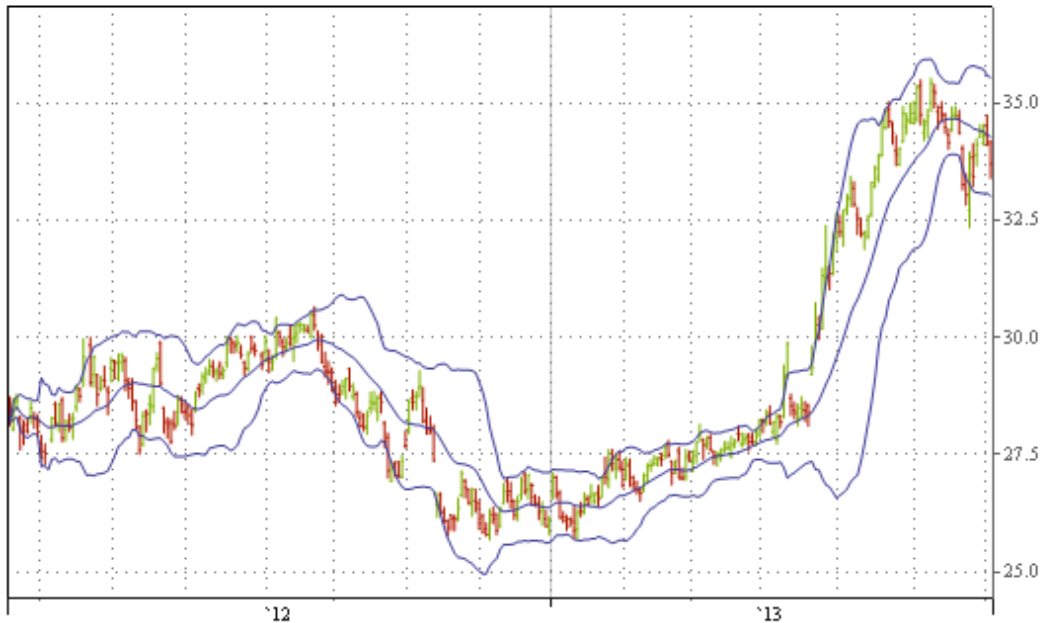

Figure 14: Trading chart for Microsoft between actual and predicted movement offered by "BollingerBands" (between 18th of May, 2012 and 2nd of July, 2013)

10.2.7 Summary of additional experiments

This section presents examples from Google, Apple and Microsoft to test the validity and accuracy of BIaaS. The "SimpleMovingAverage" is used to test that the predicted movement is 95% accurate comparing to the actual stock index movements. The "BollingerBands" is used to include both upper and lower limits of predicted movement and aim for 99.99% accuracy, where results of all the three examples agree and support with BIaaS.

10.3   Risk modelling can be performed better on Cloud than Gaussian copula

There are many types of risks that might affect businesses including market risk, credit risk, liquidity risk, legal/reputation risk and operational risk. Among these risks, operational risk is considered to be most directly related to the IT infrastructure as it might impact the business through fraud, workplace safety, damage to assets and business disruption.

Behavioural models of systems are often constructed to predict likely outcomes under different contexts and scenarios. Both analytical and simulations methodologies have been applied to these models to predict the likely outcomes, and our examples show some of these predictability features.

BIaaS is a cloud-based service which can compute pricing and risks with accurate results and good performance. Being a Cloud based system, risk data computed by different models such as Heston Model and its associated APIs could be simulated and shared within a secure platform (within the cloud) that offers anonymisation.

Following the 2008 financial crisis, many financial services businesses are scaling down their computing resources to save costs. Cloud computing allows them to do this but still offer similar levels of services by sharing resources more effectively. This paper demonstrates that financial clouds can enable users to benefit from modelling and simulations carried out on the Clouds which would be impractical otherwise. Compared with working on the desktop, users benefit from improved speed and accuracy: calculations can be completed in seconds and improved models provide a more accurate range of prices compared to traditional computation using normal distributions.

In this paper, BIaaS can compute the beta risk, which is an unsystematic risk (uncontrolled risk) that can adversely influence the future directions of many enterprises. A challenge is to calculate the accurate beta risk and new method needs to be in place to ensure the probability of calculating the beta risk is as close to the reality as possible. The use of R-squared values can determine how close the data is to the overall plot. A important task is focused on performing regression on all beta values, and computes results with both numerical analysis and visualisation presentations, so that both stakeholders and data analysts can all understand the implications offered by the beta risk. The use of Heston model can compute risks, both volatility and implied volatility, and the asset prices.

One key benefit offered by Cloud is cost. The implementation for all these systems and applications cost around 70% of the yearly licence for purchasing commercial products in our case. The private cloud development can save organisations costs in the long term.

11. Conclusion

Recent problems in the Financial Sector have arisen in part from inaccurate and inadequate assessment of risk arising. Because they have been constrained to operate in a desktop environment, financial institutions have used mathematical models of risk which assume a 'normal' or 'Gaussian' distribution of events which can dangerously underestimated the real risk. Using Business Intelligence as a Service in the Cloud (BIaaS) permits organisations to break the constraints of the desktop and use the power of Cloud Computing to adopt more complex models and improve accuracy in risk analysis and prediction.

BIaaS operates in two types of services:

1. The Heston Model to calculate asset prices, volatility, maturity period, implied volatility and risk-free rate with the use of calibration, which has been through stress test to validate results. The Heston Model track changes of prices and risks simultaneously in the form of numeric outputs.
2. Business analytics of stock market performance to allow investors to track or review the market performance of their chosen stocks. The Heston Model and its associated API can compute predicted stock index movement and offer a good extent of accuracy.

There are five different APIs developed to help BIaaS to deliver these two services. Their functionalities are explained and their usage scenarios are presented to demonstrate how BIaaS can achieve service delivery. BIaaS can calculate important key values for investments and allow investors to understand the change of risks in relation to changes in other variables. Often risk modelling cannot exploit the extent of growth in risks and BIaaS can provide recommendation for when to avoid possible investment and when is the right time for

investment through the use of risk visualisation. Risk visualisation allows analysts and investors to understand the extents of changes and they can make the corresponding actions.

Execution time is the de-facto benchmark for acceptable performance of this type of application and we have established that our technique achieves good accuracy and response times even when using a low-cost Cloud implementation. Existing work can perform calibration but either the execution time is prolong, or require expensive resources to support, or not being capable of computing predicted movement in real-time. Contributions to this research is that BIaaS can perform calibration quicker, is affordable and can offer around 95% of accuracy to the real stock index movement and 99.99% of accuracy while including 95% Confidence Interval.

Comparing to other approaches, BIaaS provides a structured way to deploy low cost, high quality risk assessment which avoids generating the unduly optimistic results which can arise from an implicit assumption in other techniques that favourable trading conditions will continue. The use of BIaaS offers better performance and better efficiency than using desktops and other systems for risk modelling, and collaborating organisations that adopt BIaaS also report the benefits over financial modelling on desktops and other systems.

Acknowledgement

Biography

He was an IT Manager/IT Lead in one of NHS Trusts, UK. He is working as a Senior Lecturer at Leeds Metropolitan University, UK since September 2012. He previously worked as an Assistant Professor and IT Architect at Business School, University of Greenwich, while doing and completing his part-time research with the University of Southampton. He is involved in enterprise research to help organisations in achieving good Cloud design, deployment and services. He received PhD (Computer Science) and PGCert (Higher Education) within 4 years part-time, whereby the distance between his work and research is hundred of miles away. He has 24 publications in his first PhD, and offers 100% effort for his full-time work. With 14 years of IT experience, he is one of the most active practitioners and researchers in Cloud in England.

References
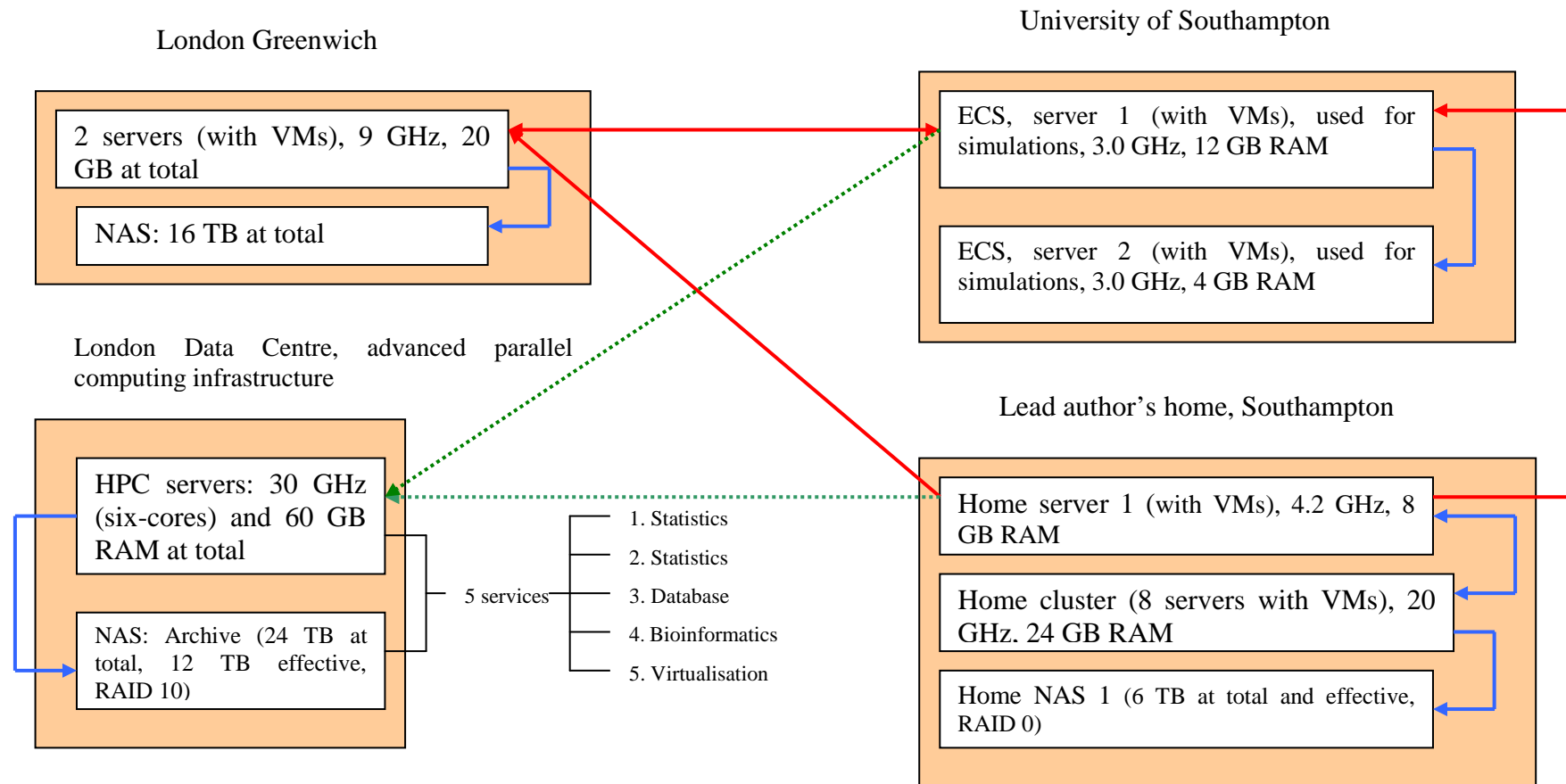
[1] D., MacKenzie, and T., Spears, The Formula That Killed Wall Street': The Gaussian Copula and the Material Cultures of Modelling, Working paper, School of Social and Political Science, the University of Edinburgh, June 2012.
[2] E. W., Frees, J., Carriere, and E., Valdez, Annuity Valuation with Dependent Mortality, Journal of Risk and Insurance 63:229-261, 1996.

[3] D., Kondo, B., Javadi, P., Malecot, F., Cappello, D.P., Anderson, Cost-Benefit Analysis of Cloud Computing versus Desktop Grids, in: 18th International Heterogeneity in Computing Workshop, Rome, Italy, 2009.

[4] E., Dixon, Cloud Computing - Benefits and Barriers for Financial Services organisations, in: Cobweb White Paper, 2011.

[5] M. Armbrust, A. Fox, R. Griffith, A.D. Joseph, R.H. Katz, A. Kownwinski, G. Lee, D. Patterson, A. Rabkin, I. Stoica, M. Zaharia, Above the Clouds: A Berkeley View of Cloud computing, in: Technical Report, UC Berkeley, Berkeley, 2009.

[6] Chang, J. F., Business Process Management Systems: Strategy and Implementation, Auerbach Publication, ISBN 0-8493-210-X, 2006.

[7] X., Xu, "From cloud computing to cloud manufacturing." Robotics and computer-integrated manufacturing 28, no. 1 (2012): 75-86.

[8] S., Marston, Z., Li, S., Bandyopadhyay, J., Zhang and A., Ghalsasi, Cloud computing—The business perspective. Decision Support Systems, 51(1), 176-189, 2011.

[9] G., Meissner, (editor), The Definitive Guide to CDOs, pg. 71, 2008, Risk Books.

[10] G. M., Gupton, C. C, Finger, and M., Bhatia, CreditMetrics: Technical Document, New York: J.P. Morgan, technical paper, 1997.

[11] E. W., Frees, J., Carriere, and E., Valdez, Annuity Valuation with Dependent Mortality, Journal of Risk and Insurance 63:229-261, 1996.

[12] J., Feiman, and D.W Cearly, Economics of the Cloud: Business Value Assessments, in: Core Research White Papers, Gartner RAS, 2009.

[13] J.C. Hull, Options, Futures, and Other Derivatives, Seventh Edition ed., Pearson, Prentice Hall, 2009.

[14] V., Chang, C.S., Li, D., De Roure, G., Wills, R.J., Walters, and C., Chee, The Financial Clouds Review, International Journal of Cloud Applications and Computing, 1 (2011) 46-63.

[15] J., Hull, and A., White, The Pricing of Options on Assets with Stochastic Volatilities, The Journal of Finance, Vol. 42 Issue 2, 1987.

[16] H., Albrecher, P., Mayer, W., Schoutens, and J., Tistaert, The Little Heston Trap, Technical paper, September 2006.

[17] R., Durrett, Probability: theory and examples,4th edition. Cambridge University Press, ISBN 0-521-76539-0, 2000.

[18] P., Wilmott, Paul Wilmott on quantitative finance (2nd ed.), ISBN 0470018704, 2006.

[19] J.C., Cox, J.E. Ingersoll, and S.A. Ross. (1985), A Theory of the Term Structure of Interest Rates, Econometrica 53: 385-408.

[20] W., Feller, Two Singular Diffusion Problems, Annals of Mathematics, 54 (1951), 173-182.

[21] P.E, Kloeden, and E., Platen, (1999). Numerical Solution of Stochastic Differential Equations. Berlin: Springer. ISBN 3-540-54062-8.

[22] F., Guillaume and W. Schoutens, Calibration risk: Illustrating the impact of calibration risk under the Heston model, Review of Derivatives Research, (2012) 15:57–79.

[23] F. Guillaume, W. Schoutens, Model and calibration risks, Eurandom report, techncial paper, 2010.

[24] A., Khajeh-Hosseini, D., Greenwood, and I., Sommerville, Cloud Migration: A Case Study of Migrating an Enterprise IT System to IaaS, 3rd IEEE International conference on Cloud Computing, Cloud 2010, 5-10 July 2010; Miami, USA.

[25] A., Khajeh-Hosseini, I., Sommerville, and I., Sriram, Research Challenges for Enterprise Cloud Computing, LSCITS Technical Report, 2010.

[26] M. P., Papazoglou, and W-J., van den Heuvel, Blueprinting the Cloud, IEEE Internet Computing, Vol. 15 No. 6, November/December 2011.

[27] V., Chang, G., Wills, G. and Walters, R., Towards Business Integration as a Service 2.0 (BIaaS 2.0), In IEEE International Conference on e-Business Engineering, The 3rd International Workshop on Cloud Services - Platform Accelerating e-Business , 19-21 October, 2011, Beijing, China.

[28] V., Chang, R. J., Walters, and G., Wills, (2012) Business Integration as a Service. International Journal of Cloud Applications and Computing, 2, (1), 16-40.

[29] V., Chang, (2013) Business Integration as a Service: Computational risk analysis for small and medium enterprises adopting SAP, International Journal of Next-Generation Computing, June 2013, 4, (3).

[30] T., Rings, G., Caryer, G., J., Gallop, J., Grabowski, T., Kovacikova, S., Schulz, and I., Stokes-Rees, Grid and Cloud Computing: Opportunities for Integration with the Next Generation Network, Journal of Grid Computing, DOI 10.1007/s10723-009-9132-5, August 2009.

[31] W.F., Sharpe, Capital Asset Prices with and without Negative Holdings, Nobel-Prize Economics Lecture, December 7, 1990.

[32] Financial Times, Interview with Lord Turner, Chair of Financial Services Authority, in: Financial Times, London, 2009.

[33] A.H., Grouard, S., Levy, and C., Lubochinsky, Stock market volatility: from empirical data to their interpretation, Research Report, Banque de France (FSR) Bank of France, June 2003.

[34] S., Hurn, K., Lindsay and A., McClelland, Estimating the Parameters of Stochastic Volatility Models using Option Price Data, Working paper, NCER Working paper series, October 2007.

[35] E., Benhamou, E., Gobet, M., Miri, SIAM Journal on Financial Mathematics 1, 1 (2010), page 289-325.

[36] X., Tang, K. Li, , G. Liao, K., Fang and F., Wu, A stochastic scheduling algorithm for precedence constrained tasks on Grid, Future Generation Computer Systems, Volume 27, Issue 8, October 2011, Pages 1083–1091.

[37] V., Chang, G., Wills, R., Walters, and W., Currie, (2012) Towards a structured Cloud ROI: The University of Southampton cost-saving and user satisfaction case studies. In, Sustainable ICTs and Management Systems for Green Computing. , IGI Global, 179-200, 2012.

[38] X., Chen, G. B., Wills, L., Gilbert, and D., Bacigalupo, Using Cloud for Research: A Technical Review, TesciRes Report for JISC, June 2010.

[39] T., Dillon, C., Wu, and E., Chang, Cloud Computing: Issues and Challenges, 2010 24th IEEE International Conference on Advanced Information Networking and Applications, pp.27-33, 20-23 April, 2010, Perth, Australia.

[40] R., Buyya, C. S., Yeo, S., Venugopal, J., Broberg, and I., Brandic, Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility, Journal of Future Generation Computer Systems, Volume 25, Issue 6, June 2009, Pages 559-616.

[41] M., Lockneed, Awareness, Trust and Security to Shape Government Cloud Adoption, LM Cyber Security Alliance and Market Connection White Paper, April 2010.

[42] A., Khajeh-Hosseini, I., Sommerville, J., Bogaerts, P., Teregowda, 2011. Decision Support Tools for Cloud Migration in the Enterprise. IEEE 4th Int. Conf. on Cloud Computing (CLOUD 2011), Washington DC, USA.

[43] A., Khajeh-Hosseini, D., Greenwood, J. W., Smith, and I., Sommerville, The Cloud Adoption Toolkit: Supporting Cloud Adoption Decisions in the Enterprise, Software: Practice and Experience, 2011.

[44] T. S., Behrend, E. N., Wiebe, J. E, London, and E. C., Johnson, Cloud computing adoption and usage in community colleges, Behaviour and Information Technology, 30:2, 231-240, 2011.

[45] Ahronovitz, M. et al. (about 70 authors), "Cloud Computing Use Cases White Paper", Version 4.0, National Institute of Standards and Technology, July 2010.

[46] A. A. , Friedman, and D. M.,, West, Privacy and Security in Cloud Computing, Issues in Technology Innovation, published by the Centre for Technology Innovation, Washington DC, No 3, October 2010.

[47] V., Chang, R. J., Walters, and G., Wills, (2012) Cloud Storage in a private cloud deployment: Lessons for Data Intensive research (Best student paper), the second international conference on Cloud Computing and Service Sciences (CLOSER 2012), Porto, Portugal.

[48] V., Chang, R. J., Walters, and G., Wills, (2013), The development that leads to the Cloud Computing Business Framework, International Journal of Information Management, June, 33, (3), 524-538.

[49] L., Youseff, M., Butrico, and D., Da Silva, Toward a Unified Ontology of Cloud Computing, Grid Computing Environments Workshop GCE '08, 12-16 Nov. 2008, Austin, TX, USA.

[50] S., Hosono, A., Kuno, M., Hasegawa, T., Hara, Y., Shimomura, and T., Arai, A Framework of Co-creating Business Values for IT Services, 2009 IEEE International Conference on Cloud Computing, September 21-25, 2009, Bangalore, India.

[51] S., Hosono, T., Hara, Y., Shimomura, and T., Arai, Prioritizing Service Functions with Non-Functional Requirements, CIRP Industrial Product-Service Systems Conference, pp.133-140, 14-15 April, 2010, Linkoping, Sweden.

[52] A., Assel, S., Chapman, L., C., Chee, Drapper, W., Fang, B., Hart, B., M., Khoe, D., Nelson., A., Ounnas, S.S., Yang and C., Wong, C., interviewees for expert review.

[53] J., Villasante, Approaching the Cloud: Better Business using Grid Solutions: Twenty-five successful case studies from BEINGrid, in, 2010.

[54] D., Allenotor, and R. K., Thulasiram, Grid resources pricing: A novel financial option based quality of service-profit quasi-static equilibrium model, in: The 9th IEEE/ACM International Conference on Grid Computing, IEEE, Washington, USA, 2008.

Appendix

London Greenwich

University of Southampton

2 servers (with VMs), 9 GHz, 20 GB at total

NAS: 16 TB at total

ECS, server 1 (with VMs), used for simulations, 3.0 GHz, 12 GB RAM

ECS, server 2 (with VMs), used for simulations, 3.0 GHz, 4 GB RAM

London Data Centre, advanced parallel computing infrastructure

Lead author's home, Southampton

HPC servers: 30 GHz (six-cores) and 60 GB RAM at total

NAS: Archive (24 TB at total, 12 TB effective, RAID 10)

5 services

1. Statistics
2. Statistics
3. Database
4. Bioinformatics
5. Virtualisation

Home server 1 (with VMs), 4.2 GHz, 8 GB RAM

Home cluster (8 servers with VMs), 20 GHz. 24 GB RAM

Home NAS 1 (6 TB at total and effective, RAID 0)

Red arrows: simulations and computational connections between different networks.

Blue arrows: simulations and computational connections between internal networks.

Green (dotted line): interactions between different sites which need to pay for access.

Figure 15: Private Cloud Deployment Architecture