Basis Risk Quantification Using SAS®

Allen Bryson, Kevin Kindall, and Xianqiao Chen

Abstract

The natural gas industry continues to be a dynamic industry. With recent high prices and the renewed interest in LNG, the business activity within the industry is expected to continue undiminished for some time to come. Clearly, understanding risk is important to decision makers. Although there are many risks to consider, basis risk is one of the primary risks embedded in many natural gas portfolios. While many of the standard Value at Risk models do not explicitly model basis risk, we show how to model basis risk directly, and develop a SAS-based Monte Carlo process to estimate exposure metrics. This paper is divided into three sections: an extended introduction and overview of the gas market, mathematical background and risk measures, and a SAS-based solution.

Gas Market Overview

Natural gas is an energy commodity that is used for industrial uses, as a heating fuel, and for power generation. It is primarily composed of methane. Based on data from the Energy Information Agency, industry and residential/commercial uses individually account for just less than 40% of demand. With the large residential component, demand is seasonal with a peak in the winter. To meet winter demand, gas is injected into storage during the summer months. The book "Trading Natural Gas" by Fletcher Sturm (1997) is a good reference on the commercial gas business.

Market Participants

Participants in the gas market include producers, marketers, end users, and speculators. A market needs diverse participants with different objectives to be successful. Over the last several years, the market has experienced changes triggered in part by the bankruptcy of Enron and the credit downgrades and regulatory problems of many other marketing companies. Large producers and now a number of large financial institutions have been able to fill the vacuum left by the marketers.

Infrastructure

The structure of the North American gas market consists of oil and gas fields, gathering systems, intrastate and interstate pipelines, storage facilities, gas processing plants, local distribution companies and end users. Once produced, gas needs to be transported from the well to the interstate pipeline network. The pipes that connect the wells to the main pipelines are collectively called the gathering system. All the major pipelines have quality requirements that gas must meet to be accepted by the pipeline. Natural gas plants take gas from the gathering systems, separate out the methane from the heavier molecules, and remove impurities to get the gas to pipeline specifications. Heavier hydrocarbon molecules generally command a higher price.

Producing regions are connected to consuming areas through an interconnected network of pipelines. The network is shown in Figure 1. Gas is traded at many different locations, and prices can be materially different across the network as they are driven by local supply and demand factors.

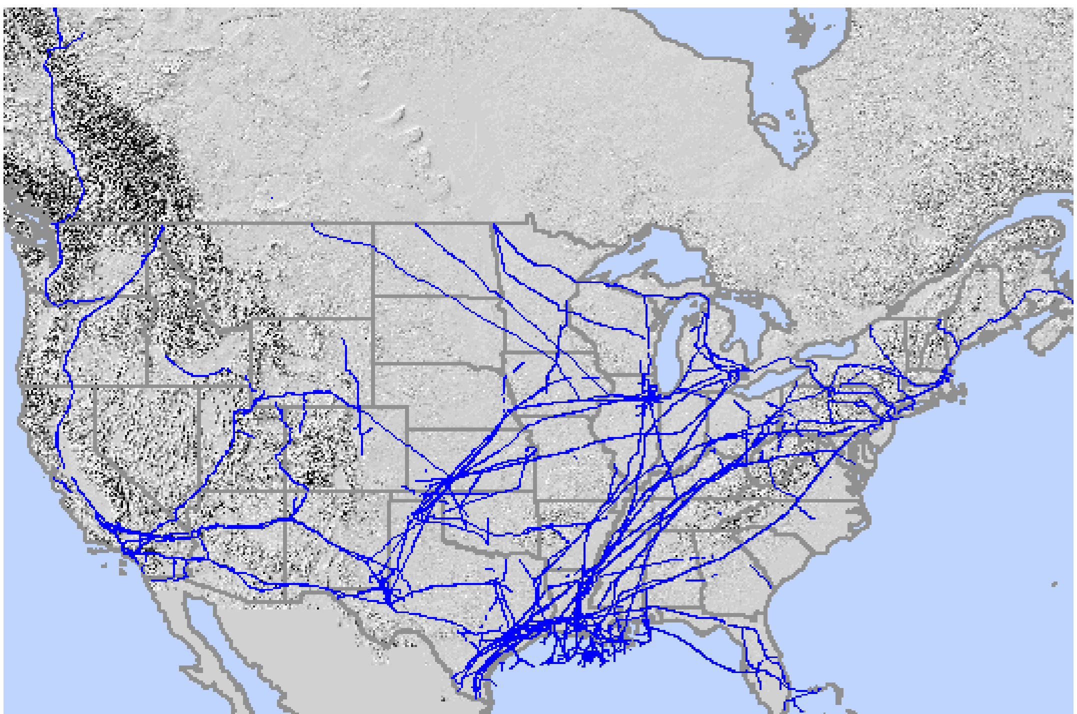

Figure 1. North American natural gas pipeline network (24 inch and larger).[1]

[1] Energy Velocity, Velocity Suite.

Pricing

Common methods of pricing gas include the use of price indices, a differential to the New York Mercantile Exchange (Nymex) futures contract (known by the term “basis” within the industry), and the use of a specified or fixed price. Index prices are published by several industry publications such as Platts (Inside FERC and Gas Daily), Intelligence Press (Natural Gas Intelligence), and Natural Gas Week. The Natural Gas Exchange (NGX) publishes prices for Canadian gas including both physical and financial instruments. The Canadian Gas Price Reporter publishes the Canadian indices based on NGX data.

Indices are available for both monthly and daily transactions. The monthly U.S. indices are based upon fixed price and basis transactions made during the last week of the month for natural gas deliveries for the upcoming calendar month. This period is called “bid week” as most transactions for the coming month get made during this time interval. Basis trades are only used in the setting of indices for points east of the Rocky Mountains (these are made during the first three days of bid week before the Nymex contract expires). Basis trades are excluded from the index setting process in the west.

The surveys for the daily indices are intended to collect fixed price transactions made during morning trading the day in advance of the flow date (the date of physical delivery). The survey for weekends and holidays also picks up the flow date that immediately follows the holiday or weekend. The various industry publications define which pipeline delivery points each index represents.

Types of instruments traded

Within the gas industry there are two broad categories of instruments (or contracts) that are transacted (traded) between parties. The first is the forward contract that is a contract for physical delivery of a specific amount of gas to another party at a specified location and time period. The second type of contract is known as a derivative. These are (usually) financial contracts that are almost always cash settled. Derivatives can be traded on an exchange or traded directly between counterparties, sometimes with the assistance of a broker (over-the-counter, or OTC contracts). A derivative is a contract whose value is determined based on the price or value of an observable but related entity. The entity is frequently referred to as the underlier. For gas, the underlier may be the Henry Hub price for natural gas, or an industry-specific index. Some contracts call for one observable while others have two observable values.

One form of derivative is the futures contract. The standard Nymex futures contract for natural gas is based on a volume of 10,000 MMBtus of gas for delivery in a specified contract month. Delivery of the gas is at Henry Hub Louisiana, which is located at a convergence of several pipelines. Approximately 72 different delivery months are trading at any one time.

Other derivatives are traded over-the-counter (OTC). The most common form of an OTC transaction is the swap. A swap involves cash settlement based on the underlying index(s) specified in the contract. Another form of derivative contract commonly used in the natural gas business is the monthly index basis swap. This transaction involves one party paying a monthly index price and the other party paying a price tied to the Nymex gas futures contract plus or minus a fixed differential. The traded basis differential is generally negative in the producing regions and positive in the consuming regions of the country. This reflects local supply and demand along with the cost of transporting the gas to the consuming regions.

Having thus concluded the general market overview, we now turn our attention to the development and estimation of certain risk measures.

Mathematical Background and Risk Measures

Various events over the past few years have highlighted the need for comprehensive risk management systems. While risk management is itself a large topic, risk managers tend to focus primarily on *market risk*: the potential change in the value of a trade resulting from (primarily) changes in prices. Virtually all energy trading firms estimate a particular metric known as Value-at-Risk (VaR). VaR is defined as the maximum potential change in value of a portfolio of financial instruments with a given probability over a given time horizon[2]. Given a distribution of the change in portfolio values, the VaR would correspond to a particular percentile. The art of market risk management is to develop this distribution of portfolio changes in a meaningful and intuitive way that also accurately reflects the price dynamics observed in the market, and to do it quickly enough to be useful for decision making. If constructed well, the risk exposures as well as the key drivers of the profit and loss become apparent. A detailed discussion of VaR may be found in the RiskMetrics Technical Document.

There are three main ways to calculate VaR: analytical, monte carlo, and historical. The historical approach is largely nonparametric as it relies primarily on historical prices. Although this is the preferred way to calculate VaR, issues concerning data quality coupled with the difficulty in continuously obtaining historical prices each day for various illiquid commodities have limited its use in the energy industry. Details on historical VaR may be found in the literature[3], but we will restrict our discussion to the other two techniques.

Monte Carlo VaR

The monte carlo approach is quite common, and borrows much from the option pricing literature. The idea behind the approach is quite simple. Prices for a given future time horizon are generated using a stochastic process. The portfolio is then revalued using the set of simulated prices. Given the set of simulated portfolio values, the distribution of the

[2] RiskMetrics-Technical Document, Fourth Edition, page 6.
[3] See, for example, Beyond Value at Risk, pages 193-196.

changes in the market value of the portfolio can then be estimated by subtracting the current portfolio value from the set of simulated portfolio values. The VaR will correspond to the percentile of choice, usually the 1^{st} or 5^{th} percentile. One of the advantages of the monte carlo approach is its ability to handle nonlinear instruments, such as options, that are in the portfolio. We provide some details below.

Perhaps the most common stochastic process in use today is geometric brownian motion (GBM). The formula for GBM relates the change in price of a security to a number of other parameters, and is given in Equation 1.

$$dS = \mu S dt + \sigma S dz \qquad (1)$$

Equation 1 is a stochastic differential equation. Among the terms shown here, it is the "*dz*" that deserves special attention. This term not only introduces randomness as *dz* is a normally distributed random variable, but also requires a different calculus. Embedded within Equation 1 are a number of continuity and independence that, for our purposes here, are only mentioned in passing, but are discussed at length within the literature. Equation 1 can be directly integrated to yield the following:

$$S_{t+1} = S_o e^{\left(r - \frac{1}{2}\sigma^2\right)t + \sigma\sqrt{t}dz} . \qquad (2)$$

Furthermore, one can use Ito's lemma to derive the differential equation that the change in the value of a derivative (*f*), such as an option written on the security, must satisfy.

$$df = \left(\frac{\partial f}{\partial S}\mu S + \frac{\partial f}{\partial t} + \frac{1}{2}\frac{\partial^2 f}{\partial S^2}\sigma^2 S^2 \right) dt + \frac{\partial f}{\partial S}\sigma S dz . \qquad (3)$$

As the change in the value of the derivative and the change in value of the underlier are driven by the same source of randomness, one can construct a portfolio consisting of the underlier and the derivative in such a way as to eliminate the randomness altogether[4]. A portfolio that is constructed in this way is riskless, and thus must grow at the risk free rate. Following this line of reasoning, one may derive the famous Black-Scholes equation.

$$\frac{\partial f}{\partial t} + rS\frac{\partial f}{\partial S} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 f}{\partial S^2} = rf . \qquad (4)$$

Given a particular payoff diagram, the Black-Scholes equation may be solved any number of ways. Some insight may be gained by realizing that the Black-Scholes equation is a heat equation in nonstandard form.

---

[4] Options, Futures, and Other Derivatives, Third Edition, page 237.

In practice, many firms use a reduced form of the general GBM process in their risk management algorithms.

$$S_{t+1} = S_o e^{\sigma \sqrt{t} dz} . \tag{5}$$

Equation 5 is indeed quite similar to the Equation 2 except that the "t" term has been removed. For small forecast horizons, the difference is quite small. However, from a practical perspective, Equation 5 is much easier to implement and computationally faster.

Equation 5 may also be written in a slightly different form

$$\ln\left(\frac{S_{t+1}}{S_0}\right) = r = \sigma \sqrt{t} dz . \tag{6}$$

This form gives some insight into the distribution assumptions alluded to earlier. As dz is distributed N(0,1), the geometric return, r, is distributed as

$$r \sim N\left(0, \sigma^2 t\right). \tag{7}$$

This result is also useful for parameter estimation and validation.

The monte carlo process is easily demonstrated for a single instrument. Each row in the table below corresponds to one possible outcome. The change in portfolio value is simply the difference between the simulated price and the starting price, which in this case is $60. For each iteration, dz is selected (as a sample from a standard normal distribution), and inserted into Equation 5.

Iteration	dz	St+1	Change in Portfolio Value
1	1.33386	$61.02	$1.02
2	-0.497751	$59.62	-$0.38
3	-1.073086	$59.19	-$0.81
4	-1.913004	$58.57	-$1.43
⋮	⋮	⋮	⋮

Table 1. Monte carlo simulation detail.

The VaR at 95% confidence corresponds to the fifth percentile of the change in portfolio value distribution (also known as change in mark to market). This is found empirically by sorting the changes in portfolio value from least to greatest, and then reading off the 5^{th} percentile. For 500 iterations, the resulting distribution function is shown in Figure 2.

The VaR in this example is estimated to be $1.27. Supposing for the moment that the portfolio consists of one share of stock, the expected one in twenty day loss is $1.27. If the VaR is too large relative to the risk tolerance of the investor, various hedging actions can now be conducted to reduce the risk profile.

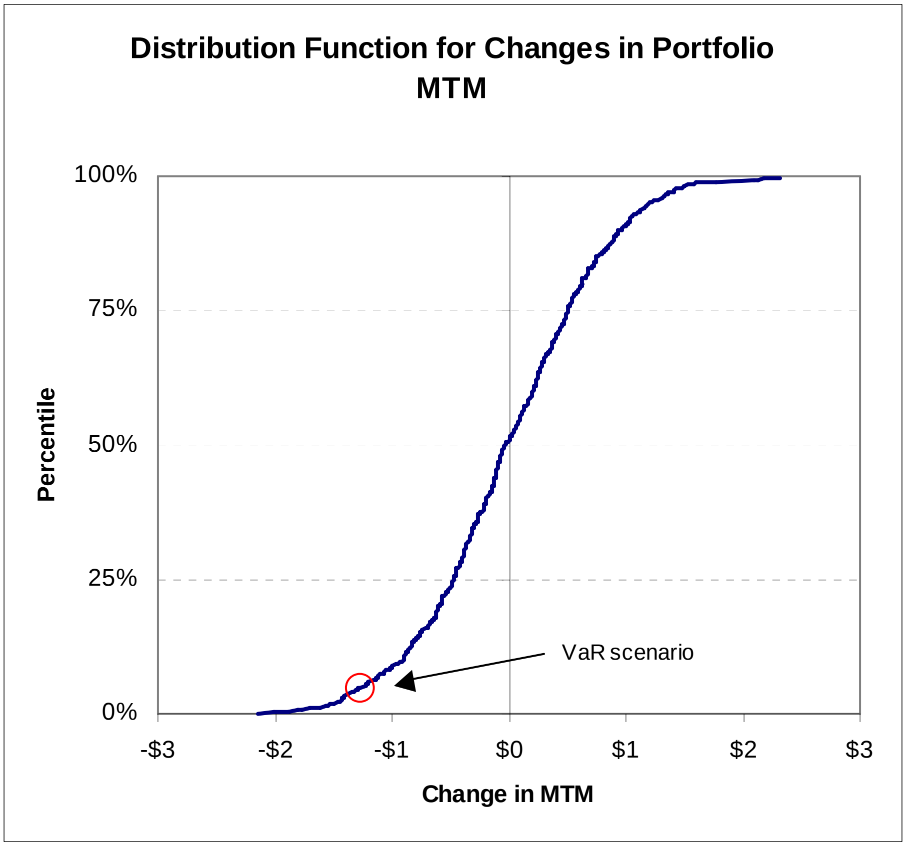

Figure 2. Monte Carlo estimate of the change in MTM distribution function.

Analytic VaR

VaR may also be estimated analytically. For linear instruments and short forecast horizons, the difference between an analytical VaR and a monte carlo VaR is quite small for reasons that we shall see in a moment.

In most cases, the analytical VaR is a multiple of the standard deviation of the portfolio value. This approximation works particularly well if the distribution is normal. As the normal distribution is characterized by its first and second moments, the VaR can be found by multiplying the standard deviation estimate by the appropriate multiplier, which in turn is determined by the confidence level of interest. Consistent with the discussion earlier, the VaR at 95% confidence would use a multiplier of approximately –1.65. The formula for a typical analytical VaR is given in Equation 8.

$$VaR = -1.65 * S_0 * \sigma\sqrt{t} \ . \qquad (8)$$

As mentioned above, a monte carlo based VaR estimate, and an analytical VaR estimate for a linear portfolio are very close. This is, of course, not coincidental.
Consider the formula for the reduced form GBM.

$$S_{t+1} = S_o e^{\sigma\sqrt{t}dz}$$

For a single linear instrument, the VaR scenario will correspond to the case where dz = -1.65 so that

$$VaR = S_{VaR} - S_0 = S_0 e^{\sigma\sqrt{t}(-1.65)} - S_0 \ . \qquad (9)$$

Expanding the exponent in a Taylor series (and ignoring higher order terms), one finds

$$VaR = S_0\left(1 - 1.65\sigma\sqrt{t}\right) - S_0 \ . \qquad (10)$$

Rearranging terms, we derive the formula for an analytical VaR.

$$VaR = -1.65 * S_0 \sigma\sqrt{t} \ . \qquad (11)$$

According to Equation 11, the distribution of changes in the portfolio value is normal, and has a mean of zero. One advantage of Equation 11 is its simplicity, ease of calculation, and, as shown above, for linear instruments produces a VaR estimate that is virtually identical to a VaR estimate derived by the monte carlo approach. However, in general, the presence of nonlinear portfolio instruments prohibits the analytical VaR to be a fully general solution. We now turn our attention to the multivariate case.

Multivariate VaR

As almost all non-trivial portfolios consist of more than one instrument, it is necessary to generalize the above approach. In the monte carlo framework, this can be accomplished by modeling each instrument as a correlated GBM. The correlation reflects how the two instruments are expected to move together. Given 40 instruments, one would then have 40 correlated GBMs. The next few paragraphs describe the process for generating correlated GBMs.

Fundamentally, the multivariate monte carlo approach is based upon the idea of sampling from multivariate normal distribution. If the dz's are chosen to reflect the correlation structure, then they can be inserted in Equation 5 (one such equation for each instrument) and the process described earlier repeated. Indeed, we shall see that many of the formulas shown above become matrix equations.

There are a number of different ways to sample from a multivariate normal distribution. At risk of compromising rigor, some insight can be gained by considering the following line of thinking. If the distribution of an individual return is given by Equation 7, then the multivariate distribution of portfolio returns will be given by

$$R \sim N(0,\Sigma) = \sqrt{\Sigma} N(0,I), \tag{12}$$

where Σ is the correlation matrix, and I is the identity matrix. $N(0,I)$ is simply the standard normal multivariate distribution. The $\sqrt{\Sigma}$ term may be interpreted in the light of Equation 13

$$\Sigma = A^T A, \tag{13}$$

where A is an upper triangular matrix. This process of decomposing the correlation matrix in this way is known as a Cholesky decomposition.[5]

However, certain conditions must be satisfied in order for such a decomposition to be possible—one of which is positive definiteness. Although the correlation matrix is both real and symmetric, in practice correlation matrices tend to not to be positive definite, so certain perturbative procedures have been developed to ensure that the correlation matrices used in risk management algorithms are positive definite. The real world consequence is that a positive definite correlation matrix ensures that the set of all possible portfolios has a positive variance.

The equations for analytical VaR generalize nicely to the multivariate case. As before, the VaR is a multiple of the standard deviation of the portfolio value. For a two asset case, Equation 8 becomes

$$VaR = -1.65\sigma_{std}, \tag{14}$$

where σ_{std} is given by

$$\sigma_{std} = \sqrt{\begin{pmatrix} SV\sigma & SV\sigma \end{pmatrix} \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix} \begin{pmatrix} SV\sigma \\ SV\sigma \end{pmatrix}}. \tag{15}$$

Such simplicity allows for quick estimates of VaR to be calculated using the matrix functions in Excel.

Although not discussed here, model calibration and validation are very important, and at times can become quite involved. Details can be found elsewhere.[6]

---

[5] RiskMetrics Technical Document, Fourth Edition, page 253.
[6] See, for example, the discussion regarding backtesting in VaR, Understanding and Applying Value-at-Risk, pages 52-54.

Basis Risk

As mentioned in the first section, basis risk is one of the key risks embedded in many gas portfolios. While the discussion above is limited to generating prices, traders are more concerned with basis (or price spreads). Thus, not only should the mark-to-market reports clearly show exposure to basis, but also the risk reports need to be adapted to this paradigm.

All the mathematics that were discussed earlier dealt with prices, not price spreads. Although one can simulate two prices and subtract them to synthetically create a basis, empirical studies of basis behavior indicate that there are dynamics in the basis that may not be accurately reflected by synthetically created basis. While these errors are small for small time horizons, they become important once the risk horizon is expanded. Furthermore, one cannot use a GBM process for the basis itself, as a basis may be negative.

One approach to this problem is to model the Nymex price using a GBM process, and modeling the basis separately. As a first step, one can assume that the changes in the basis follow some known distribution. Once the parameters of the distribution are appropriately estimated, the monte carlo process is generalized to include the following term:

$$S_{t+1} = S_0 e^{\sigma\sqrt{t}dz} + B_{t+1} . \qquad (16)$$

As the basis is modeled as a separate entity, questions regarding basis exposure and basis VaR may be addressed, while still preserving the ability to calculate a total VaR as before. Care must be taken when constructing the correlation matrix as the correlation matrix must reflect the behavior of geometric returns and changes in the basis.

This approach does, however, open the door to a wide range of models for the basis. The basis could be modeled using a mean reverting process, or with one of a large number of econometric models. For example, one could use the following mean reverting model for the basis:

$$B_{t+1} = B_t + \kappa(\mu - B_t) + \sigma\varepsilon_t . \qquad (17)$$

Here μ represents the long run mean estimate, while κ represents the speed of mean reversion. In the monte carlo process, one would generate samples for the error term, $\sigma\varepsilon_t$, in such a way as to reflect the correlation of the geometric returns with the residuals of Equation 17.

Model Basis Risk Using SAS

Commodity trading firms typically handle large amounts of data for prices and positions. Most firms maintain a central data warehouse as the official system of record that stores data and produces mark-to-market and risk reports. If a SAS user is familiar with the data structure of the corporate IT system, it is straightforward to use PROC SQL to query information out of the system for processing in SAS. We introduce this section by providing a brief overview of the SAS solution in the following two paragraphs.

In our code example, we assume that a trader has some open positions in Nymex futures contracts as well as the Chicago basis and wants to compute a VaR estimate on these positions. Prices are stored in the data warehouse. The positions, measured in mmbtu, are given below.

Month	Nymex	Chicago
8/1/2005	5,000	10,000
9/1/2005	10,000	20,000
10/1/2005	20,000	10,000
11/1/2005	10,000	15,000
12/1/2005	20,000	20,000
1/1/2006	-10,000	-10,000
2/1/2006	25,000	5,000
3/1/2006	5,000	10,000
4/1/2006	5,000	-20,000

Table 2. Trade detail.

We use the DATA step to access a text file produced by the system of record that has both the current prices and 75 days of historical prices for Chicago and Nymex—the historical prices being necessary for parameter estimation. Geometric returns for Nymex and basis changes for Chicago are then calculated. A separate DATA step uses SAS RANNOR function to generate 100,000 random numbers for each Chicago and Nymex pricing point. The output datasets are then used in a PROC IML module where the bulk of VaR calculation takes place, and includes the estimation of volatilities and correlations, correlating the random samples, price simulation, portfolio revaluations, and finally the estimation of the VaR. Figure 3 depicts the entire process.

As mentioned previously, Nymex prices are to be modeled using Equation 5. As Equation 5 contains a volatility parameter, we need to obtain a time series for geometric returns of Nymex prices to estimate volatilities and correlations—the correlations used in generating the "dz" term.

Chicago basis prices, on the other hand, are to be modeled using the following equation

$$B_{t+1} - B_t = \sigma\sqrt{t}dz \qquad (18)$$

where $dz \sim N(0,1)$. Note, however, that the mean reversion model in Equation 17 can be readily used once the parameters have been estimated for the long run mean and the speed of mean reversion.

To calibrate the model in Equation 18, we need to obtain the historical time series for absolute basis price changes for Chicago to estimate both volatilities and correlations. The first few data steps and proc steps in the example generate a 75×18 matrix which represents 18 instruments (location and contract month combinations, such as Chicago_Mar06) and 75 historical observations for each instrument. In addition, the mean is calculated for these instruments and placed into a 1×18 vector for use in the calculation of volatilities and correlations in the PROC IML module. Current prices are also placed into a 1×18 vector.

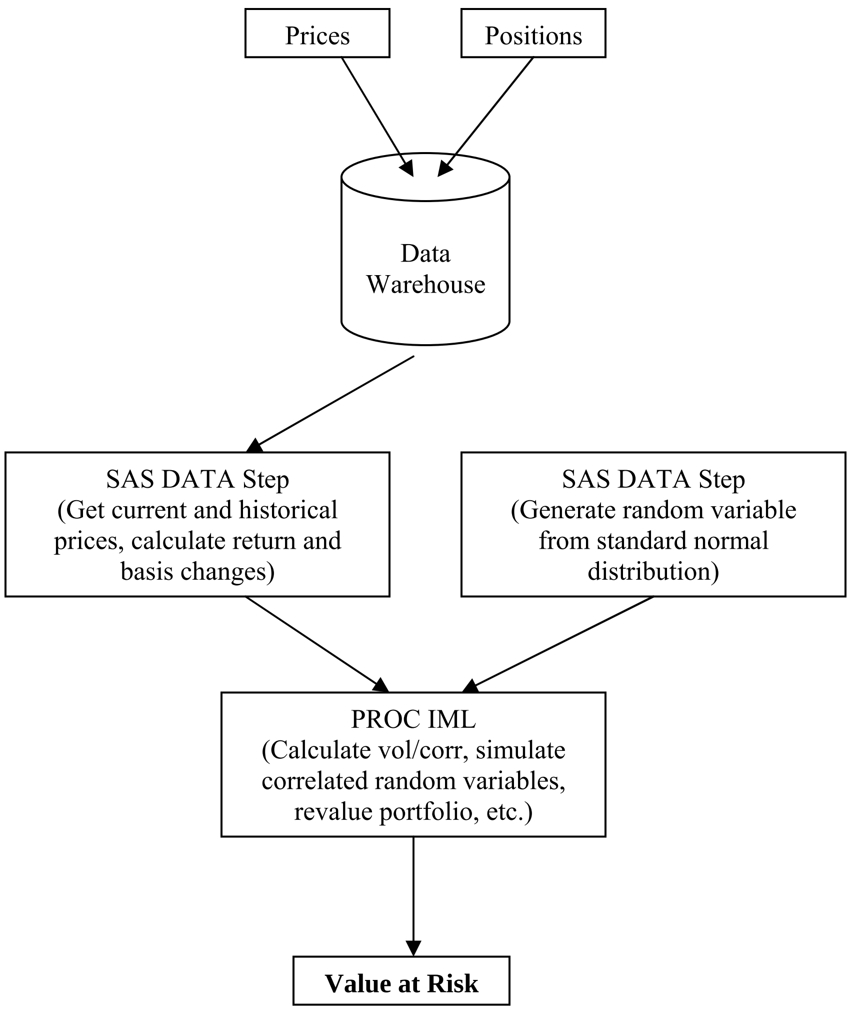

Figure 3. SAS monte carlo VaR flow diagram.

Code

```
data OriginalPriceHistory;
* Import text file;
 infile '\\RISKSERVER\ROOT$\Price_file_history\price050729.rpt'
DLM=',';
 input name:$ type $ Nearby Date:mmddyy10. Currency $ Price;
 format date mmddyy10.;
run;
data FilteredPriceHistory;
* Query prices;
 set OriginalPriceHistory;
 where nearby<>10 and nearby <>0 and (name='NGI_CHIC' or
name='NYMEX_NG');
 key=nearby || name;
 drop currency name type nearby;
run;
proc sort data=FilteredPriceHistory;
 by date;
run;
proc transpose data=FilteredPriceHistory
out=FilteredPriceHistory;
 var price;
 by date;
 id key;
run;
data FilteredPH;
 set FilteredPriceHistory;
 array F{*}_numeric_;
 array Nymex(9);
 array fp_basis(9);
 do i=1 to 9;
  Nymex(i)=F(i+10);
  fp_basis(i)=F(i+1); * +1 to exclude the date column;
 end;
 array basis(9);
 do i=1 to 9;
      basis(i)=fp_basis(i)-Nymex(i);
 end;
 array l_nymex(9);
 array r_nymex(9);
 do i=1 to 9;
      l_nymex(i)=lag1(nymex(i));
      r_nymex(i)=log(nymex(i)/l_nymex(i));
 end;
 array l_basis(9);
 array c_basis(9);
 do i=1 to 9;
      l_basis(i)=lag1(basis(i));
      c_basis(i)=basis(i)-l_basis(i);
 end;
run;
Data FPH;
 set filteredPH(firstobs=2);
 keep r_nymex1-r_nymex9 c_basis1-c_basis9;
run;
```

```
Data CurrentPrice;
 set filteredPH(firstobs=76);
 keep nymex1-nymex9 basis1-basis9;
run;
proc means data=FPH noprint;
 output out=Mn(Drop=_:) mean=;
run;
```

Another DATA step is used to generate the dz. This produces a $100{,}000 \times 18$ matrix which represents 18 instruments and 100,000 random variables for each instrument. Note, however, that the set of dz's are random drawings from a standard normal distribution and are not correlated across instruments. They will be "forced" to correlate across instruments in the PROC IML module.

<u>Code</u>

```
Data R_number (drop=i j);
 array Y(18);
 do i=1 to 100000;
  do j=1 to 18;
      Y(j)=rannor(631280);
  end;
  output;
 end;
run;
```

The output datasets from preceding steps are then read into the PROC IML module. The first step in this subsection is to calculate volatilities and correlations using an exponentially weighted moving average approach (EWMA). This approach is different from the standard approach of calculating standard deviation for a given time series in that it gives the latest observations the highest weight in the volatility estimate. This approach has three important advantages over the equally weighted model. First, the volatility estimate reacts faster to shocks in the market as recent data carry more weight than data in the distant past. Second, following a shock (a large return), the volatility declines exponentially as the weight of the shock observation becomes smaller over time. Third, the estimator naturally selects its sample size.[7] The formula for calculating volatility is given in Equation 19:

$$\sigma = \sqrt{(1-\lambda)\sum_{t=1}^{T} \lambda^{t-1}(r_t - \bar{r})^2}\,, \tag{19}$$

where λ ($0<\lambda<1$) is the decay factor and has a value of 0.94 in our example. The formula for calculating correlation is given in Equation 20:

$$\rho_{1,2} = \frac{\sigma_{1,2}{}^2}{\sigma_1 * \sigma_2} \tag{20}$$

[7] RiskMetrics-Technical Document, Fourth Edition, page 78.

Code

```
%let decay=0.94;
proc iml;
 use FPH;
 read all into ReturnMtx;
 use mn;
 read all into MeanVector;
 CorrMtx=j(18,18,0);
 VolVec=j(1,18,0);
 * Calculate correlation first;
 do i=1 to 18;
  do j=1 to 18;
      tempTotal=0;
      vol1=0;
      vol2=0;
      do k=1 to 75;
            temptotal=temptotal+&decay**(75-k)*(ReturnMtx[k,i]-
            MeanVector[1,i])*(ReturnMtx[k,j]-MeanVector[1,j]);
            vol1=vol1+&decay**(75-k)*(ReturnMtx[k,i]-
            MeanVector[1,i])**2;
            vol2=vol2+&decay**(75-k)*(returnmtx[k,j]-
            meanvector[1,j])**2;
      end;
      vol1=sqrt((1-&decay)*vol1);
      vol2=sqrt((1-&decay)*vol2);
      CorrMtx[i,j]=(1-&decay)*temptotal/(vol1*vol2);
  end;
 end;
 * then calculate volatility;
 do i=1 to 18;
      vol=0;
      do k=1 to 75;
            vol=vol+&decay**(75-k)*(ReturnMtx[k,i]-0)**2;
      end;
      volVec[1,i]=sqrt((1-&decay)*vol);
 end;
```

The preceding code produces an 18×18 correlation matrix and a 1×18 volatility vector. As mentioned previously, the correlation matrix is essential in generating a set of correlated random variables for price simulation. The correlation matrix is to be decomposed into an upper triangular matrix using Cholesky decomposition. However, we must first make sure that the correlation is positive definite. Stated differently, we need to check to see if the smallest eigenvalue of the matrix is greater than zero. If not, we need to perturb the matrix to make it positive definite.[8]

[8] LongRun Technical Document, page 151.

Code

```
*Check to see if correlation matrix is positive definite and
perturb it if necessary;
 lambda=min(eigval(CorrMtx));
 x=-lambda/(1-lambda)+0.0001;
 NewCorrMtx=CorrMtx+(x*(I(18)-CorrMtx));
 *Check to see if new correlation matrix is positive definite;
 y=min(eigval(NewCorrMtx));
 print lambda x y;
```

The resulting correlation matrix is "guaranteed" to be positive definite, and we then use the SAS ROOT function to perform Cholesky decomposition on this new correlation matrix to derive the upper triangular matrix A in Equation 13. The next step is to import the $100{,}000 \times 18$ dz matrix and perform the matrix multiplication of this matrix with the upper triangular matrix. The result is a $100{,}000 \times 18$ correlated dz matrix.

Code

```
*Get Cholesky root and correlated random variables;
 c_CorrMtx=root(NewCorrMtx);
 use R_number;
 read all into Rnumber;
 c_Rnumber=Rnumber*c_CorrMtx;
```

With a set of correlated random variables and a volatility vector, we can generate 100,000 simulated prices for Nymex future and Chicago basis. The formulas for simulating prices for the next day for both Nymex and basis are provided here again for convenience:

$$S_{t+1} = S_o e^{\sigma dz}$$
$$B_{t+1} = B_t + \sigma dz$$

Code

```
*Get simulated price;
 use CurrentPrice;
 read all into Price;
 s_price=j(100000,18,0);
 do i=1 to 18;
  do j=1 to 100000;
      if i<10 then
s_price[j,i]=price[1,i]*exp(volvec[1,i]*c_Rnumber[j,i]);
      else s_price[j,i]=price[1,i]+volvec[1,i]*c_Rnumber[j,i];
  end;
 end;
```

The preceding code produces a $100{,}000 \times 18$ simulated price matrix. Now we are just steps away from getting a VaR. We create a 1×18 current price vector and an 18×1 position vector, and multiply the two to get the current portfolio value. Next, each element in the simulated price matrix is multiplied by the corresponding position in the position vector and added across instruments to produce a $100{,}000 \times 1$ simulated portfolio value vector. Subtraction of the current portfolio value from each element of this vector creates a $100{,}000 \times 1$ vector that represents change in portfolio value. This vector is then sorted in ascending order. Observation #5000 is the 5^{th} percentile and also our VaR estimate.

Code

```
*Create some positions, calculate current MTM, simulated MTM and
PNL;
 position={5000,10000,20000,10000,20000,-10000,25000,5000,5000,
                  10000,20000,10000,15000,20000,-
                  10000,5000,10000,-20000};
 c_mtm=price * position;
 s_mtm=j(100000,18,0);
 s_mtmtotal=j(100000,1,0);
 do i=1 to 100000;
  do j=1 to 18;
      s_mtm[i,j]=position[j,1]*s_price[i,j];
      s_mtmtotal[i,1]=s_mtmtotal[i,1]+s_mtm[i,j];
  end;
 end;
 s_pnl=j(100000,1,0);
 do i=1 to 100000;
  s_pnl[i,1]=s_mtmtotal[i,1]-c_mtm;
 end;
 call sort(s_pnl,1);
 VaR=s_pnl[5000,1];
 print VaR;
quit;
```

It is easy to calculate a VaR on the basis positions only by setting the Nymex position to be identical to zero in the position vector.

Summary

We have shown how to produce a Value-at-Risk estimate for the basis risk in a portfolio. Clearly, there are distinct advantages to modeling the basis separately. Although there are more complex basis models than the one used here, the mechanics of the simulation will largely be the same.

The authors may be contacted at the following emails:
Xianqiao.Chen@conocophillips.com
Allen.Bryson@conocophillips.com
Kevin.G.Kindall@conocophillips.com

References

Crnkovic, Cedomir and Drachman, Jordan. “Quality Control.” In VaR: Understanding and Applying Value-at-Risk. London: KPMG and RISK Publications, 1997.

Hull, John. Options, Futures, and Other Derivatives. Upper Saddle River, NJ: Prentice Hall, 1997.

Jorion, Philippe. Value at Risk: The New Benchmark for Controlling Market Risk. New York: McGraw-Hill, 1997.

Kim, Jongwoo, Allen Malz, and Jorge Mina. LongRun Technical Document. 1999.

RiskMetrics-Technical Document. New York, 1996.

Sturm, Fletcher. Trading Natural Gas: Cash, Futures, Options and Swaps. Tulsa, OK: PennWell Publishing Company, 1997.