

Time series interpolation with Volatility Bridge

Dr. Ventsislav Nikolov, Snezhina Yanakieva, Aleksandar Krastev

Technical University of Varna, Computer Science and Engineering Department, 9010, 1 Studentska Street, Varna, Bulgaria

When working with market data it is not always easy to rely on price series being consistent. Some of the series may contain gaps causing calculation errors. One way to approximate the data in these gaps is to perform a deterministic interpolation. However, deterministic interpolation techniques fail to capture the correct probability distribution at these intermediate times. Using Monte Carlo simulation the Volatility Bridge interpolation manages to capture the correct joint distribution of the data.

Интерполация на времеви серии с Volatility Bridge (доц. Венцислав Николов, Снежина Янакиева, Александър Кръстев). Работейки с пазарни данни не може винаги да се разчита на консистентност на предоставените ценови серии. В някои серии има липсващи данни, които водят до грешки при изчисления. Един от начините, по които се изчисляват липсващите стойности, е детерминистична интерполация. И все пак детерминистичната интерполация не успява да улови правилната вероятност за разпределение в тези междинни моменти. Използвайки Monte Carlo симулация Volatility Bridge интерполацията успява да изчисли правилното разпределение на мединните данни.

Introduction

The financial calculation algorithms achieve better performance if the time series these calculations are working on have consistent distribution. The two techniques that are usually used to consistently rescale market observed time series data are interpolation and standardization. Data interpolation provides reorganization of the observable information to appear similar across all records and fields.

Today big data networking and machine learning tools are widely used for market analysis and predictions. All these market analysis tools usually work with large volumes of observed market prices series. However, the data observed at the market is not always consistent. The gaps in the observed series may cause calculation errors or inaccuracies. Fig. 1 shows an example of observed on the market data series for aluminum.

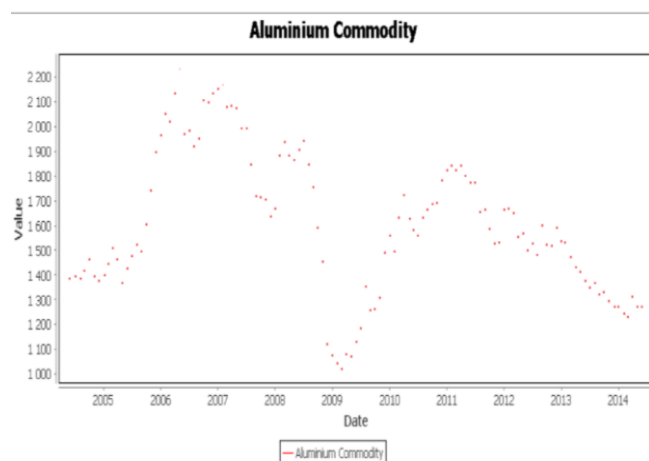


Fig.1. Actual market series of aluminum commodity prices

Data interpolation is used to organize data to appear similar across all records and fields. There are two main big

groups of interpolation techniques: deterministic and stochastic. Deterministic interpolation techniques give a way to create surfaces from observed market points, based on the behavior of the series. Stochastic interpolation techniques utilize the statistical changes of the measured data. Each method generates dates grid according a given frequency, start and end dates.

After an analysis of the Principal Component solution of Factor analysis model and Roweis' missing value solution in that context based on the Expectation minimization (EM) algorithm [2] and its modified version implementation [1] we propose an approach based on Monte Carlo simulation and Volatility Bridge algorithm. The purpose of the Volatility Bridge is generating smaller time interval values between the bigger time interval values of a time series using a Monte Carlo simulation.

Deterministic interpolation techniques

Deterministic interpolation techniques are used to create surfaces from sample market observable points using mathematical functions. Here the following types are analyzed:

- Use last value
- Linear interpolation

Use last value

This method adjusts a series to the generated grid. The adjust approach is: if there is no quote for the (grid) date the last available (original series) quote is considered as valid. The implementation of this method applied on the example data series for aluminum is shown at Fig. 2.

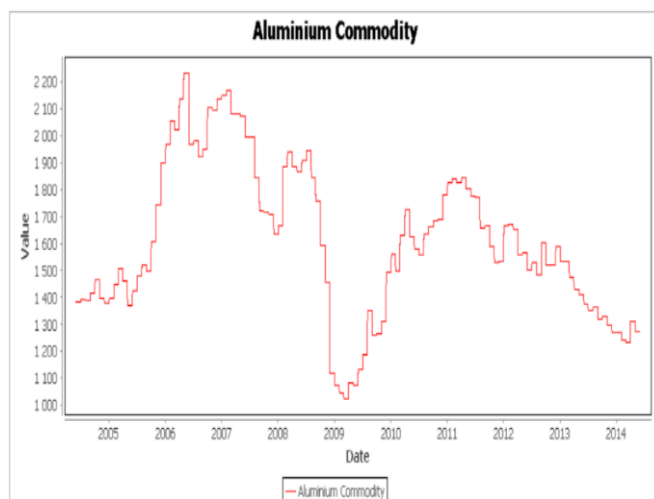


Fig.2. Use last value interpolation of aluminum commodity prices

Linear interpolation

This method adjusts a series to the generated grid. The adjust approach is: if there is no quote for the date (grid), use linear interpolation between the (original series) quotes before and after the date (grid) to fill it. Linear interpolation applied to the example data is shown at Fig. 3.

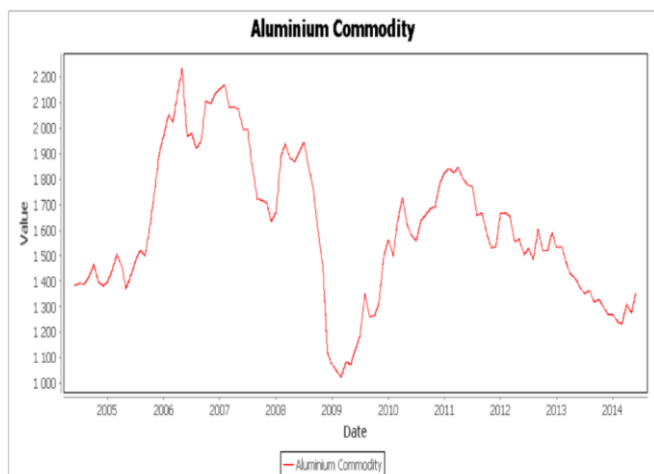


Fig.3. Linear interpolation of aluminum commodity prices

Monte Carlo Simulation

The Monte Carlo Simulation can create scenarios for the market development changes in the future and can prepare the series to be analyzed successively, by using scenarios. The resulting outcome series represent a price distribution, indicating the risks with confidence. Fig. 4 shows the simulated by Monte Carlo algorithm scenarios for a market observable data series.

The systematic evaluation of the Monte Carlo Simulation is based on several steps:

- 1) Volatility vector and correlation matrix for a set of the market factors.
- 2) Covariance matrix is obtained by a matrix multiplication of the volatility vector and correlation matrix.
- 3) A Cholesky matrix is constructed from the covariance matrix.

4) Matrix Perturbation or Spectral Decomposition need to adjust the covariance matrix to be positive definite, in case the original matrix was not positive definite, as a result of the risk factors.

5) A random generator produces independent sample data for each market factor, using standard normal distribution (0, 1) that get stored into deviation vectors.

6) Deviation vector is being multiplied by the Cholesky matrix, taking into account the covariance between every market risk factor pair. This produces delta vectors of normally distributed and at the same time correlated deviations.

7) Vectors are applied to actual market factor values.

8) Steps 6 is repeated a number of times (5000...10000). The large scenario set affects the calculation of price distributions of pricing expression. The resulted price distribution is calculated by counting the appearance of the values within many close ranges.

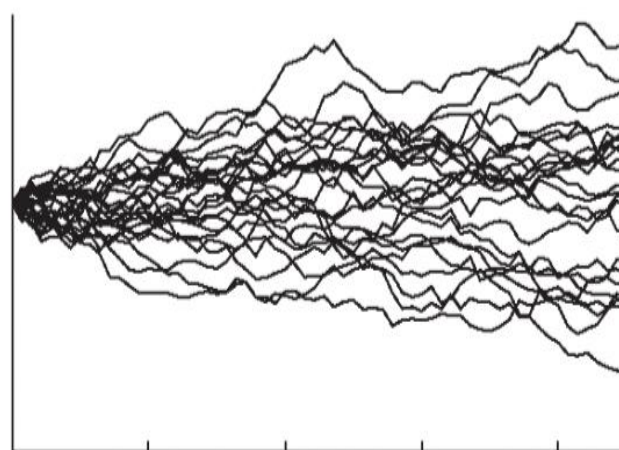


Fig.4. Generated scenarios

Volatility Bridge algorithm

Input data

The algorithm works with the original time series of an observed on the market factor. Some series are not eligible for volatility bridge normalization. The series not eligible for volatility bridge interpolation match at least one of the following criteria:

- Have at least one invalid quote (returns and volatility cannot be calculated).
- Have BIENNIAL frequency (requires history longer than 200 years for representative sample). The series frequency is assumed to be the frequency whose periods occupy most days within the series. For example:
Series with 2 BIENNIAL quotes and 30 DAILY quotes will have BIENNIAL frequency.
Series with 300 DAILY and 5 MONTHLY quotes will have DAILY frequency.
- Have volatility less than 0.01 % (constant series) or greater than 100% (out of models domain).

Output data

After the volatility bridge method is applied, the time series has additional values amongst the bigger time interval values.

Algorithm

The volatility bridge generates scenarios for arbitrary horizons in between the original time series values.

The volatility bridge process generates values with a normal distribution in between two time points, where at each time point the value is known. Price scenarios, generated by Volatility Bridge, are normally distributed.

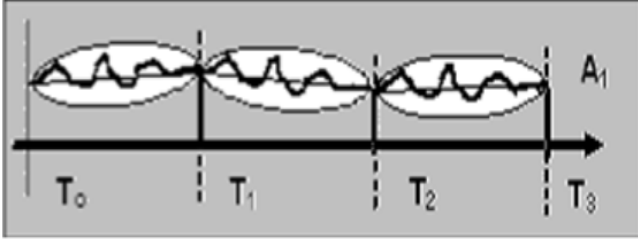


Fig.5. Generating values between several time points

Source: Adapted from [2]

The original time series values represent level 1 series points, while Volatility Bridge generated values are at level 2 series points. Fig. 5 represents example generated values between two time series points. The Volatility Bridge simulation performs a stochastic interpolation between the original points, using the volatility bridge. The following properties are taken into account during this process.

- Expected values between the control points are linearly interpolated.
- Daily fluctuations are generated by a Monte Carlo simulation, which generates the changes derived from the historical volatility and correlation.

The simulation of scenarios is independent of the time of simulated planned items. An example simulation is shown at Fig. 6.

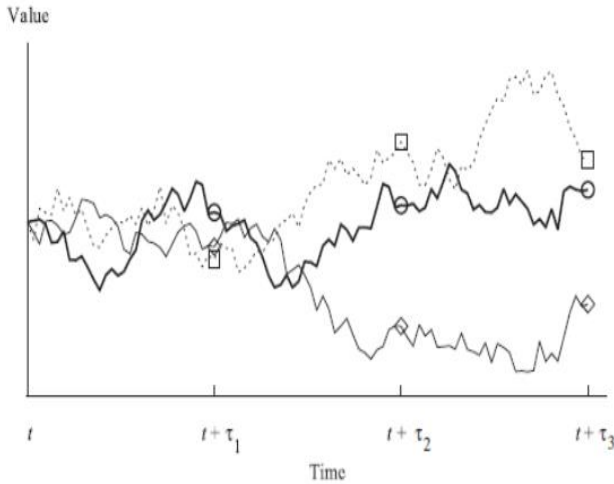


Fig.6. Simulated price paths with their level 1 and level 2 points

Source: Adapted from [2]

For $b=1, 2, \dots, \tau_1, \tau_1 + 1, \dots, \tau_R$ the daily prices are generated with the following properties:

Between level 1 dates, for $\tau_{r-1} < b < \tau_r$, level 2 scenarios are distributed according to the assumptions of daily dynamics of d , conditional on d passing through $S_{t+\tau_{r-1}}^q$ and $S_{t+\tau_r}^q$, where $S_{t+\tau_r}^q$ are level 1 prices.

In order to simulate the volatility bridge process for d ,

the following algorithm has to be used:

1) Independent normally distributed random variables u_l^q for $l = 1, \dots, \tau_r$ are generated with mean 0 and variance σ^2 .

2) At level 1 dates (that is, for $b = \tau_r, r = 1, \dots, R$), set $d_{t+\tau_r}^q = S_{t+\tau_r}^q$

3) For all other dates set

$$(1) \quad d_{t+b}^q = \frac{\tau_r - b}{\tau_r - \tau_{r-1}} S_{t+\tau_{r-1}}^q + \frac{b - \tau_{r-1}}{\tau_r - \tau_{r-1}} S_{t+\tau_r}^q + \sum_{l=\tau_{r-1}+1}^b u_l^q - \frac{b - \tau_{r-1}}{\tau_r - \tau_{r-1}} \sum_{l=\tau_{r-1}+1}^{\tau_r} u_l^q$$

In general, the u_l^q can be generated according to an arbitrary distribution. This allows the accounting for any special characteristics of movements, while maintaining the covariance structure of the volatility bridge.

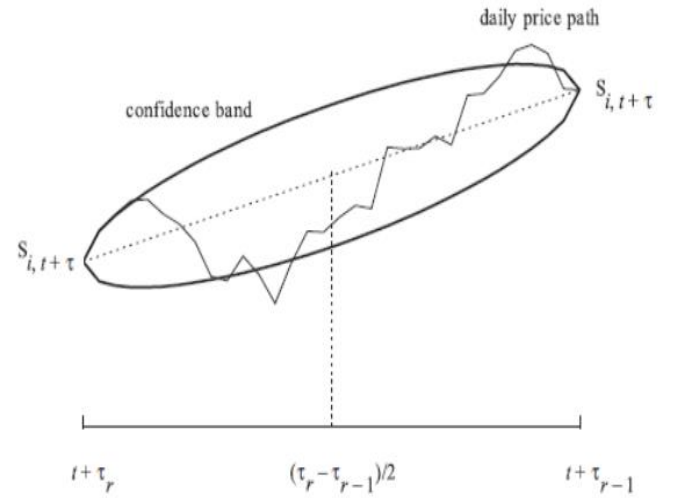


Fig.7. Volatility bridge algorithm

Source: Adapted from [2]

Results

The results from the implementation of the algorithm on the example aluminium series are shown at Fig. 8.

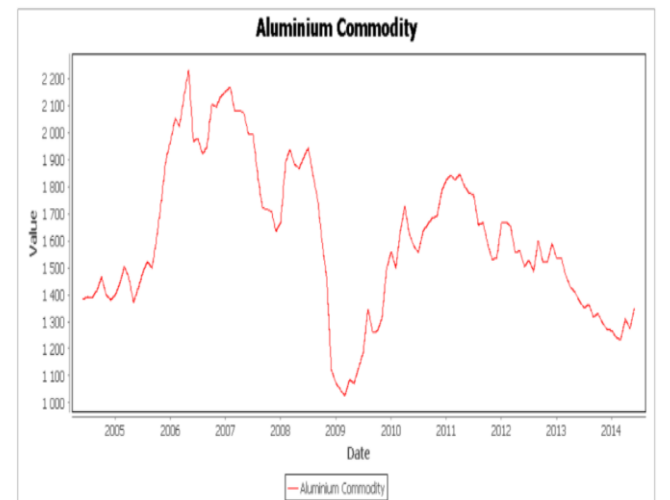


Fig.8. Volatility Bridge interpolation of aluminum commodity prices

Conclusion

This paper presents a comparison of stochastic and deterministic interpolation methods for market factors interpolation before these factors can be used in a variety of big data and artificial neural networks based calculations and analyses.

Practical implementation results showing observed on the market series of factors were used to compare the interpolation methods. The results show that the Volatility Bridge method chosen as preferred stochastic interpolation method captures the correct joint distribution by sampling from a Gaussian distribution, which helps it to better represent the stochastic nature of the series and to generate values that are closer to the actual data observed on the market.

The presented Volatility Bridge interpolation algorithm has been implemented as a web based library used in insurance, banking and trading tools for market analysis and predictions.

REFERENCES

- [1] Nikolova, N., Toneva-Zheynova, D., Naydenov, D., Tenekedjiev, T. Imputing Missing Values of Environmental Multi-Dimensional Vectors Using a Modified Roweis Algorithm, 2012.
- [2] Roweis, Sam T, .EM Algorithms for PCA and SPCA. NIPS , 1997.
- [3] Eugene Stern, Risk management for non-financial corporations, RiskMetrics Group, 1999.
- [4] Lee, A. Corporate metrics, RiskMetrics Group, 1999.
- [5] Eshel, G. The Yule Walker Equations for the AR Coefficients, Technical report, Bard College at Simon's Rock, 2009.
- [6] Hamilton, J. Time Series Analysis. Princeton University Press, ISBN: 0-691-04289-6, 1994.
- [7] Chatfield, C. The analysis of time series. An introduction. Fifth edition. Chapman & Hall/CRC, ISBN-10: 0412716402, ISBN-13: 978-0412716409, 1996.
- [8] Box G. E. P., G. M. Jenkins. Time Series Analysis: Forecasting and Control, San Francisco, Holden-Day, 1970.
- [9] Cameron, C. Trivedi, P. Regression Analysis of Count Data. Cambridge university press, 1998.
- [10] Draper, N. Smith, H. Applied Regression Analysis. Wiley Series in Probability and Statistics, 1998.
- [11] Balakrishnan, N., Nevzorov, V. A primer on statistical distributions. John Wiley & Sons Inc., 2003.
- [12] Bury, K. Statistical Distributions in engineering. Cambridge University Press, 1999.
- [13] Krishnamoorthy, K. Handbook of statistical distributions with applications. Chapman & Hall, 2006.
- [14] Mun, J. Modeling Risk. Applying Monte Carlo Simulation, Real Options Analysis, Forecasting, and Optimization Techniques. Wiley, 2006.
- [15] Ricci, V. Fitting Distributions with R. 2005.
- [16] Robert, C., Casella, G. Monte Carlo statistical methods. Springer-Verlag, 1999.
- [17] Vallentin, M. Probability and Statistics Cookbook, 2011.
- [18] Vose, D. Fitting distributions to data and why you are probably doing it wrong. 2010.