**Price Forecasting Models**

**Corn, Soybeans, Wheat:**

**Hoffman, L.A., Irwin, S.H., & Toasa, J.I. (2007). Forecast Performance of Futures Price Models for Corn, Soybeans, and Wheat.**

本篇目的為使用月期貨價格(monthly futures price)、現貨價格(cash prices received)、基值(basis values, farm price received less futures)、月份額(marketing weight)來預測季平均價格(season-average farm price)，標的為美國玉米、黃豆、小麥。使用modified Diebold-Mariano test檢測是否本研究模型的預測(future model forecast)較World Agricultural Supply and Demand Estimates (WASDE)表現更好。

雖然由於期貨市場的有效性(efficient market hypothesis)，期貨價格是未來價格的不偏估計式(unbiased predictor)，但預測誤差的變異數(variance of forecast error)很大，而且會隨著期貨到期日增加，但期貨價格仍然是很有用的資訊。

在期貨到期的月分，以期貨價格減去基值(五年現貨移動平均減去五年期貨移動平均)作為該月的預測值，再根據USDA提供的生產量將月預測值做加權平均成季預測值。

玉米和小麥的期貨合約到期為十二月、三月、五月、七月、九月，玉米期貨價格使用#2 yellow corn futures contract traded on the Chicago Board of Trade，而小麥使用#1 hard red wheat contract traded on the Kansas City Board of Trade。黃豆期貨合約到期為十一月、一月、三月、七月、八月、九月，使用資料為#2 yellow soybean futures contracts traded on the Chicago Board of Trade。現貨價格的資料來源為National Agricultural Statistics Service (NASS)。

除了使用modified Diebold-Mariano (MDM) test檢定期貨模型(futures model forecasts)和WASDE price projection是否存在差異外，也使用傳統評估預測誤差的方式如mean error、mean absolute error (MAE)、mean absolute percentage error (MAPE)、mean squared error (MSE)、root mean squared percentage error (RMSPE)。

雖然futures model forecasts在許多月份的mean squared error都較WASDE小，但在MDM test中仍無法拒絕兩預測誤差具顯著差異的虛無假設。

H. Ahumada, M. Cornejo,Forecasting food prices: The case of corn, soybeans and wheat,International Journal of Forecasting,Volume 32, Issue 3,2016,Pages 838-848,ISSN 0169-2070,https://doi.org/10.1016/j.ijforecast.2016.01.002.

本篇目的在於考量變數間的相互相關性後，如何能夠提升預測的準確性，對象集中於玉米、黃豆、小麥。特別是將資料範圍設定相當不穩定的2008至2014年間，並且使用穩健法(robust approach)以及遞迴計算法(recursive scheme)，在多期預測時會同時使用迭代(iterate)及直接預測。

使用的資料為名目價格，品項分別有玉米、黃豆、小麥，資料頻率季資料，範圍是1994Q3至2014Q4，樣本外預測自2008Q1開始。除了預測的變數以外，在商品面，加入了商品產出(commodity production)以及存貨(inventories)。在需求面上，加入了實質GDP，包含的範圍有OECD、印度、中國。在貨幣與金融上加入了貨幣基數(US monetary base)、M2貨幣、資金流量(Fed’s flow of funds)、三個月國債券利率(3-month Treasury constant maturity rate)。最後加入美國消費者指數(US consumer price index)以及美國實質匯率(US real exchange rate)、美國乙醇產量(U.S. ethanol production)。

|  |  |  |
| --- | --- | --- |
| Variable | Colname | Website |
| US ethanol production | Table 3. Fuel ethanol supply and disappearance, by month | [USDA ERS - U.S. Bioenergy Statistics](https://www.ers.usda.gov/data-products/u-s-bioenergy-statistics/#:~:text=Significant%20biofuels%20include%20ethanol%20%28made%20mostly%20from%20corn%29%2C,biodiesel%2Frenewable%20diesel%20production%20totaled%20about%202.5%20billion%20gallons.) |
| Fed’s flow of funds | Money Market Funds; Total Financial Assets, Level | [Money Market Funds; Total Financial Assets, Level (MMMFFAQ027S) | FRED | St. Louis Fed (stlouisfed.org)](https://fred.stlouisfed.org/series/MMMFFAQ027S) |

首先將估計每個商品的均衡相關模型(equilibrium correlation model, EqCM)，EqCM可以同時包含長期以及短期行為。接著估計殘差交互相關(residual cross-correlation)，原因是機關常常根據經驗法則進行政策的調整，例如價格比值達一定值時執行。

單根(unit root)以及斷點(break)會造成序列的不平穩(non-stationarity)，面對斷點可以採取的方法是二階差分，結合EqCM即變成了二階差分向量均衡相關模型(double differenced vector equilibrium correlation model, DVEqCM)。

**Wheat:**

**Dias, J., Rocha, H. (2019). Forecasting Wheat Prices Based on Past Behavior: Comparison of Different Modelling Approaches. In, et al. Computational Science and Its Applications – ICCSA 2019. ICCSA 2019. Lecture Notes in Computer Science, vol 11621. Springer, Cham.** [**https://doi.org/10.1007/978-3-030-24302-9\_13**](https://doi.org/10.1007/978-3-030-24302-9_13)

比較五種模型用於預測小麥價格的效果，分別是ARIMA、Classification and Regression Tree (CART)、Random Forests (RF)、Support Vector Machines (SVM)、Multivariate Adaptive Regression Splines (MARS)。文中使用hard red winter wheat的出口價格，並且使用包含物價效果的名目價格，使用的資料頻率是月，從1992年2月至2017年2月。

使用Autocorrelation Function (ACF) 和Partial Autocorrelation Function (PACF) 圖表來觀察當期和落後期的相關性。隨著落後的期數越高，ACF下降的速度緩慢，大約24期後才沒有顯著相關。PACF是移除其他落後效果的自我相關性，PACF在第二期後便沒有顯著的相關。為了檢定是否需要做差分，文中使用unit root test中的Augmented Dickey-Fuller Test ( ADF)，檢定結果為需要做一階差分。

函式auto.arima輸出最佳的ARIMA模型為ARIMA(p=0, d=1, q=1)。Random Forests是決策樹組合，重複抽取部分樣本形成多棵決策樹，最後再整合所有結果。Support Vector Machines是監督式學習，找出hyperplane進行分類，可分為線性和非線性，非線性的情況需要將資料投影至更高維度。Multivariate Adaptive Regression Splines也可稱為flexible nonparametric regression modelling，是可自動計算非線性關係和互動效果的無母樹迴歸模型。

預測模型的訓練方式是樣本外預測，僅使用訓練資料自1992年2月至2012年2月來估計模型參數，模型訓練後會預測未來六個月的價格。接著將下一期的資料加入訓練資料重新估計模型參數，再預測接續六個月的價格，重複執行直到使用所有的樣本，因此本篇各模型都終將有六十個預測值。為了評估模型的成果，作者使用Mean relative absolute error (MAE)、Maximum relative absolute error (MaxAE)、Root Mean Square Error (RMSE)。

隨著預測的期數越高，模型的表現逐漸衰退，並沒有模型在所有預測期數下都表現最佳，在未來1至4期的預測中，SVM和MARS的表現較好。而在預測未來價格比當前高或低的情境中，SVM的準確性在任何條件下都為最高。

**Corn:**

**Musunuru, N., Yu, M., & Larson, A. (2016). Forecasting Volatility of Returns for Corn using GARCH Models. Texas Journal of Agriculture and Natural Resources, 26, 42-55.**

文章主要目的是建立GARCH模型來預測玉米市場的波動性。結果發現市場對於利多消息和利空消息的反應並不對稱，負面消息對於波動性有較大的影響。比較GARCH家族的模型後，發現EGARCH的效果最好。

首先使用隨機漫步模型(Random Walk)，本期的投資報酬率為平均報酬率加上誤差。但隨著時間進行，變異數不會保持不變，誤差變異樹為常數的假設無法被滿足，因此無法使用傳統計量方法如OLS進行估計，需要使用GARCH模型(Generalized Autoregressive Conditional Heteroskedasticity)。

GARCH模型同時考慮平均數和波動性，並且可以捕捉波動性聚類(volatility clustering)，即大波動會伴隨著大波動，小波動會伴隨小波動。GARCH模型假設未來變動的變異數預測值會和落後期的變異數相關。GARCH也假設市場對於消息的波動是對稱的，正面和負面消息的衝擊對於波動性有相同的效果。模型中的殘差變異數是由常數、前期殘差平方、前期變異所組成。

但在許多金融的文獻上說明，在相同的情況下，負面消息對波動性的影響比正面消息大。因此可以使用不對稱(asymmetric) GARCH模型家族中的TGARCH和EGARCH。TGARCH模型在殘差變異數中加入表示殘差正負號的指示函數，使其在正負號不同的情況下，會有不同的變異數。EGARCH的條件變異數為指數型態。指數型態提供正面及負面衝擊不同的變異數，也保證變異數為正。

作者使用的是CBOT的玉米價格資料，資料範圍為1995/1/3至2012/6/16。實證的資料分布為左偏(negative skewed)，峰度(kurtosis)為17，大於常態值3。Jarque-Bera統計值顯示實證分布不符合常態分佈。平均報酬率不顯著異於零，GARCH(1,1)和TGARCH(1,1)以及RGARCH(1,1)皆為顯著，顯示確實存在ARCH效果。而區分正負衝擊的係數也為顯著，顯示負面消息造成的波動較大。而在預測方面，在RMSE、MAE、TIC項目中，EGARCH(1,1)都為最佳的模型，而在MAPE中，較RW遜色。

Ge, Y., Wu, H. Prediction of corn price fluctuation based on multiple linear regression analysis model under big data. *Neural Comput & Applic* **32**, 16843–16855 (2020). <https://doi.org/10.1007/s00521-018-03970-4>

本篇使用單變量非線性迴歸以及多變量線性迴歸模型進行玉米價格的預測。使用單變量非線性迴歸的目的在於分析數列的趨勢，因為目前觀察到價格雖然逐年上升，但上升的趨勢和時間不呈線性關係。

資料的範圍是2005年至2016年，國際貿易的資料是來自於USDA，玉米價格的資料是來自IMF China。美國玉米市場年是9/1-8/31，而中國玉米市場年是1/1-12/31。

單變量迴歸式由兩個部分組成，第一個是描述價格趨勢的轉變，第二部分使用三角函數描述周期性的變動。求得參數解後，預測模型的R-squared值達到0.9883(over-fitting?)。而多變量線性迴歸模型中，自變項有玉米的產出、進口、出口、消費。

在ARMA(p, q)模型中，使用AIC criteria決定p和q。將自變數改為透過AR(p)以及MA(q)表示，及當期的數值由落後期組成，再帶入到原先的多元線性迴歸模型中。

乙醇價格的提升會增加農業的生產成本，因為乙醇是製造化學肥料的重要原料，連帶會導致玉米的價格上升。

**Soybean:**

**Xiaojie Xu, Yun Zhang, Soybean and Soybean Oil Price Forecasting through the Nonlinear Autoregressive Neural Network (NARNN) and NARNN with Exogenous Inputs (NARNN–X), Intelligent Systems with Applications, Volume 13, 2022, 200061, ISSN 2667-3053,** [**https://doi.org/10.1016/j.iswa.2022.200061**](https://doi.org/10.1016/j.iswa.2022.200061)**.**

本篇目的在於使用非線性自我迴歸神經網絡(nonlinear autoregressive neural network, NARNN)預測大豆價格，使用Levenberg–Marquardt algorithm求解。除了自我迴歸外，還在模型中加入外生輸入，使用的變數是大豆油。

資料更新的頻率為每日，使用的範圍超過五十年，自1968年至2021年。在Jarque-Bera test中拒絕分布為常態的假設。

NARNN模型可以將當期價格寫作過去價格的函數，函數由隱藏層(hidden layer)組成，而每個隱藏層裡有落後項的組合以及連接函數(transfer function)，連接函數多使用logistic sigmoid transfer function，特點是將所有數值映射到[0, 1]區間。NARNN-X在模型中加入外生變數，大豆和大豆油在對方模型中作為外生變數。

最終決定的大豆的NARNN模型包含了兩個隱藏層以及兩個落後項，大豆油的模型中有三個隱藏層以及三個落後項。而大豆的NARNN-X模型中有六個隱藏層以及五個落後項，大豆油的模型是四個隱藏層和三個落後項。

使用LM法求得以上需估計的模型參數，LM法可以求非線性最小化問題的局部解，結合並改善了牛頓法(Gauss-Newton approach)以及梯度下降法(steppest-descent algorithm)，因此相當適合用於緩慢收斂的問題。

將所有資料拆分成三筆，分別是訓練(train)、驗證(validation)、測試(test)，並且有不同的比例組合，對應不同的模型，最後計算出每個模型可能的RRMSE(relative root mean squared error)，選擇RRMSE低並且在三份資料集中數值都一致的模型。

最後使用Diebold-Mariano test比較各模型的表現，比較的對象是傳統AR模型，結果是四種模型的表現都顯著優於AR模型。未來建議可以嘗試其他可能的時間序列方法的組合或是使用機器學習的電腦圖論。

**Crude oil:**

**Lu, Q., Sun, S., Duan, H. *et al.* Analysis and forecasting of crude oil price based on the variable selection-LSTM integrated model. *Energy Inform* 4 (Suppl 2), 47 (2021).** [**https://doi.org/10.1186/s42162-021-00166-4**](https://doi.org/10.1186/s42162-021-00166-4)

由於原油市場和世界經濟與能源環境的相關性(linkage)，造成影響油價的組成相當複雜(complex)且發散(diverse)。地區事件、貿易制裁、OPEC政策等都會造成油價的波動(volatility)。

將可能的影響因子(influence factor)區分成三個主題，分別是供給與需求(supply and demand)、金融環境(finance factor)、技術(technology)。供給和需求是影響長期油價趨勢(trend)的基本因素(fundamental factor)。原油價格可以分解成原油供給衝擊(crude oil supply shock)、全部產業全球需求衝擊(the shocks to the global demand for all industrial commodities)、世界原油需求衝擊(the demand shock to the global crude oil market)。短期原油價格波動主要歸因於金融方面，受到市場預期與投機交易(speculative transactions)影響，如匯率等其他金融指標，連接至股票市場與貨幣市場。裂解價差(Crack spread)是原油和石油產品之間的價差，反映了原油市場和煉油市場(refined product market)的供需關係。

目前主要有四個預測方法應用於油價的建模與預測，分別是time series model、econometric models、qualitative methods、artificial intelligence techniques。時間序列方法主要使用autoregressive integrated moving average (ARIMA)、exponential smoothing (ETS)。計量方法有generalized autoregressive conditional heteroskedastic model (GARCH)、vector autoregression model (VAR)、state-space models、the threshold models。隨著資料量(data volume)變大以及影響因子複雜，機器學習方法(machine learning forecasting method)開始佔有優勢，特別是處理非線性問題(non-linear problem)以及短期預測，例如support vector machines (SVM)、artificial neural networks (ANNs)、genetic algorithms (GA)、wavelet analysis。

本文分別使用三個變數挑選(variable selection)法，分別是elstic-net regularized generalized linear model (GIMNET)、Bayesian model average (BMA)、spike-slab lasso method (SSL)，並且搭配機器學習方法long short-term memory (LSTM) model。

選用real monthly West Texas Intermediate (WTI) crude oil price作為依變數(<https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm>)，時間自2000年1月至2017年12月。GLMNET被認為在處理多元共線性(multicollinearity)和variable screening上最好的收縮法(contraction method)。BMA需要先假設每個模型的先驗機率(prior probability)，而後驗機率(posterior probability)可以用資料數據以及模型的接受度獲得。結合Markov chain Monte Carlo method (MCMC)可以加快速度、獲得全面的後驗分配、可以處理高相關BMA (high BMA correlation)

Long short-term memory (LSTM) neural networks是一種recurrent neural network (RNN)的特例，目的在於建立長期依屬關係(long-term dependencies)和最適落後期(optimal time lag)。RNN是一種深度神經網絡(deep neural network architecture)，常用於時間序列的建模。本期的預測值不只會受到自變數的影響，還會受到上一期預測值影響，因此稱為循環神經網絡。LSTM可以捕捉在RNN中消失的梯度(vanishing gradient)，因此可以觀察長期的趨勢。

將六種模型分別對應三種變數選取方法，便得到18個模型加上作為對照的RW和ARIMA。文中使用Diebold-Mariano (DM) test檢定預測準確性相同的虛無假設，Mean Squared Error (MSE)作為DM loss function，將所有模型和random walk比較。使用RMSE和MAPE測量模型的表現。最後用Pesaran-Timmermann (PT) test檢測directional symmetry (DS)。

經過檢定比較後發現，BMA變數降維法加上LSTM預測模型(BMA-LSTM integration model)的表現最好，不論是在預測誤差上或是預測方向上，注意的是BMA所選取的變數既非最多也非最少。

Dependent Variables:

WTI Spot Price <https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm>

Independent Variables:

1. Supply
   1. Crude Oil Production, Total OPEC <https://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_epc0_im0_mbblpd_m.htm>
   2. Crude Oil Production, Total Non-OPEC <https://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_epc0_im0_mbblpd_m.htm>
   3. Crude Oil Production, World <https://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_epc0_im0_mbblpd_m.htm>
   4. Henry Hub Natural Gas Spot Price <https://www.eia.gov/dnav/ng/ng_pri_fut_s1_m.htm>
   5. U.S. tight oil production <https://www.eia.gov/dnav/pet/pet_move_impcus_a2_nus_epc0_im0_mbblpd_m.htm>
2. Demand
   1. Petroleum Consumption, Total OPEC <https://www.eia.gov/dnav/pet/pet_cons_psup_dc_nus_mbblpd_m.htm>
   2. China oil import
   3. U.S. tight oil production
   4. Petroleum Consumption, Total OECD
   5. China oil import
   6. Fed fund effective
   7. Kilian Global economic index
   8. US CPI index seasonally adjusted
   9. US CPI energy seasonally adjusted
   10. US PPI seasonally adjusted
   11. US PPI energy seasonally adjusted
   12. EU 28 Countries PPI
   13. US PMI index
3. Inventory
   1. Petroleum inventory, Total OECD
   2. Crude Oil inventory, Total
   3. Crude Oil inventory, SPR
   4. Crude Oil inventory, Non-SPR
   5. Crude oil non-commercial net long ratio
   6. Real dollar index generalized
   7. The exchange rate of euro against US dollar
4. Stock market
   1. S&P 500 index
   2. Dow Jones Industrial Index
   3. NASDAQ index
5. Commodity market
   1. COMEX Gold Future closing price
   2. LME Copper Future closing price
   3. WTI-Brent spot price spread
   4. WTI crack spread actual value
   5. Brent crack spread actual value

(上述為文獻中提到的變數，多數已下載，原本下載連結也貼好了，淡旺季存檔，之後會補上)

**Ng’ang’a, F. and Oleche, M. (2022) Modelling and Forecasting of Crude Oil Price Volatility Comparative Analysis of Volatility Models. *Journal of Financial Risk Management*, 11, 154-187. doi:**[**10.4236/jfrm.2022.111008**](https://doi.org/10.4236/jfrm.2022.111008)**.**

比較不同的GARCH模型來估計風險價值(VaR, Value-at-Risk)，利用Kupiecs POC test分析VaR預測的準確性。

在能源市場中，價格具有以下四種特性：季節性(seasonality)、尖峰(spikes)、均值回歸(reversion)、不可儲存性(non-storability)。主要影響原油價格的因素有全球經濟環境、地區條件以及供給與需求的波動。

VaR (Value-at-Risk) 常被用於衡量風險，標示在給定的信賴區間下，某段時間內資產組合的最大損失。GARCH模型分析中使用不同的分布由助於估計VaR。

多數實證資料都顯示原油價格的時間數列(time series)，具有波動性類聚 (volatility clustering)、厚尾分布(fat tail distribution)、不對稱性(asymmetric)、均值迴歸(mean reversion)。

本篇研究步驟為：(1)使用Augmented Dickey-Fuller test以及Philips Perron test檢定時間序列是否具平穩性(stationarity)；(2)使用Lagrange multiplier test檢定是否存在ARCH( Autoregressive Integrated Moving-Averge)效果；(3)使用五個GARCH模型預測原油價格；(4)比較各模型預測VaR的結果。

Lagrange Multiplier (LM) test用於檢定變異數是否異質性，檢定相當於當期的平方項對所有落後期的平方項進行迴歸的F-statistic。Dickey-Fuller test將t期資料對t-1期進行迴歸，檢定係數是否為1，若為1則表示序列非平穩，對立假設為係數小於1，表示序列具平穩性。Augmented Dickey Fuller test拓展了DF test至更多落後期，適用於更大以及更複雜的模型。Philips-Perron test是ADF test的無母數版本，適用於自我迴歸出現間斷或政權轉移的情況。Jarque-Bera test是適合度檢定，對象為常態分配，使用樣本峰值(kurtosis)以及偏態(skewness)。

ARCH模型假設目前誤差項的變異數是過去誤差項的實現值的函數，並且和過去誤差項的平方相關。模型的特點是可以捕捉隨時間變動的波動性(time-varying volatility)以及波動性類聚(volatility clustering)。GARCH模型加入更多落後誤差實現值的平方項以及落後誤差項的變異數，進一步觀察長短期記憶性(long and short memory)以及不對稱槓桿效果(asymmetric leverage effects)。Exponential GARCH model允許在正面或負面衝擊下出現不對稱的效果，並且對於母數無限制。Integrated GARCH model改善GARCH model的弱平穩性問題。Threshold GARCH model同樣可以處裡槓桿效果，在函數中加入指標函數(indicator function)，可以將不同的誤差項放入不同的函數。Glosten-Jagannathan-Runkle GARCH model在模型中加入虛擬變數來捕捉不對稱效果。

VaR是在給定信心水準下，期間內最大的損失；換句話說，最大損失超過估計的VaR的機率是給定的誤差。使用Kupiec’s test和Cristoffersen’s test of independence評估VaR的估計值。

根據QQ-plot以及JB test都顯示資料並不呈現常態分布，在QQ-plot可以看到資料更接近student t分配。使用ADF test和PP test都顯示資料為平穩。計算所有ARIMA模型的AIC值後，發現ARIMA(2, 0, 4)的AIC值為最小。LM test證明ARCH效果的存在，說明使用GARCH model的必要性。

分別使用常態分配以及t分配估計模型參數，並且計算loglikehood。結果發現不論在何種分配以及資訊標準(information criteria)條件下，EGARCH model的表現都為最優。

在VaR backtesting中，計算實際損失超出估計VaR的次數是否符合信賴區間，在條件與非條件的L.R. statistics以及兩種分布的四情況下，都無法拒絕觀測失敗次數和預期誤差相同的虛無假設，四種GARCH模型都能預估正確的VaR在1%誤差水準下。在四個模型中IGARCH t model在估計VaR的表現最好。在報酬預測方面，比較了ME、RMSE、MAE後，發現IGARCH t model的預測誤差最小，而簡單GARCH model也有相當好的表現。

未來建議可以使用Markov Seitching Multifractal model，並且考慮skewed student t distribution和reparameterised Johnson distribution。

Khashman, A., & Nwulu, N. I. '' Intelligent prediction of crude oil price using Support Vector Machines''. In Applied Machine Intelligence and Informatics (SAMI), IEEE 9th International Symposium on pp . 165-169, 2011.

Siddhaling Urolagin, Nikhil Sharma, Tapan Kumar Datta, A combined architecture of multivariate LSTM with Mahalanobis and Z-Score transformations for oil price forecasting, Energy, Volume 231, 2021, 120963, ISSN 0360-5442, <https://doi.org/10.1016/j.energy.2021.120963>.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| S.NO | Feature Name | Abbreviation | Website | File Name |
| 1 | Crude Oil – West Texas Intermediate | WTI | <https://fred.stlouisfed.org/series/DCOILWTICO> | Crude Oil Prices - West Texas Intermediate (WTI) - Cushing, Oklahoma.xls |
| 2 | Gold Futures | GOLD | <https://www.nasdaq.com/market-activity/commodities/gc:cmx/historical>  <https://www.investing.com/commodities/gold-historical-data> | Gold (GC:CMX) Historical Data.csv  Gold Futures Historical Data.csv |
| 3 | S&P 500 Index | SP500 | <https://stooq.com/q/d/?s=%5Espx&c=0> | Historical data: S&P 500 - U.S. (^SPX).csv |
| 4 | US Dollar Index | USDINDEX | <https://finance.yahoo.com/quote/DX-Y.NYB/history?period1=31795200&period2=1677801600&interval=1d&filter=history&frequency=1d&includeAdjustedClose=true> | US Dollar(USDX) - Index - Cash (DX-Y NYB).csv |
| 5 | US 10 Year Bond Yield | US10B | <https://www.macrotrends.net/2016/10-year-treasury-bond-rate-yield-chart> | 10 Year Treasury Rate - 54 Year Historical Chart.csv |
| 6 | Dow Jones Utilities Average | DJU | <https://markets.businessinsider.com/index/dow-jones-utility-average> | Dow Jones Utility Average\_03\_02\_23-01\_03\_00.csv |

資料範圍是2000/1/4至2019/6/10，總共4947筆。

A VAR-SVM model for crude oil price forecasting, Lutao Zhao, Lei Cheng, Yongtao Wan, Hao Zhang and Zhigang Zhang. Published Online:May 20, 2015pp 126-144. <https://doi.org/10.1504/IJGEI.2015.069485>

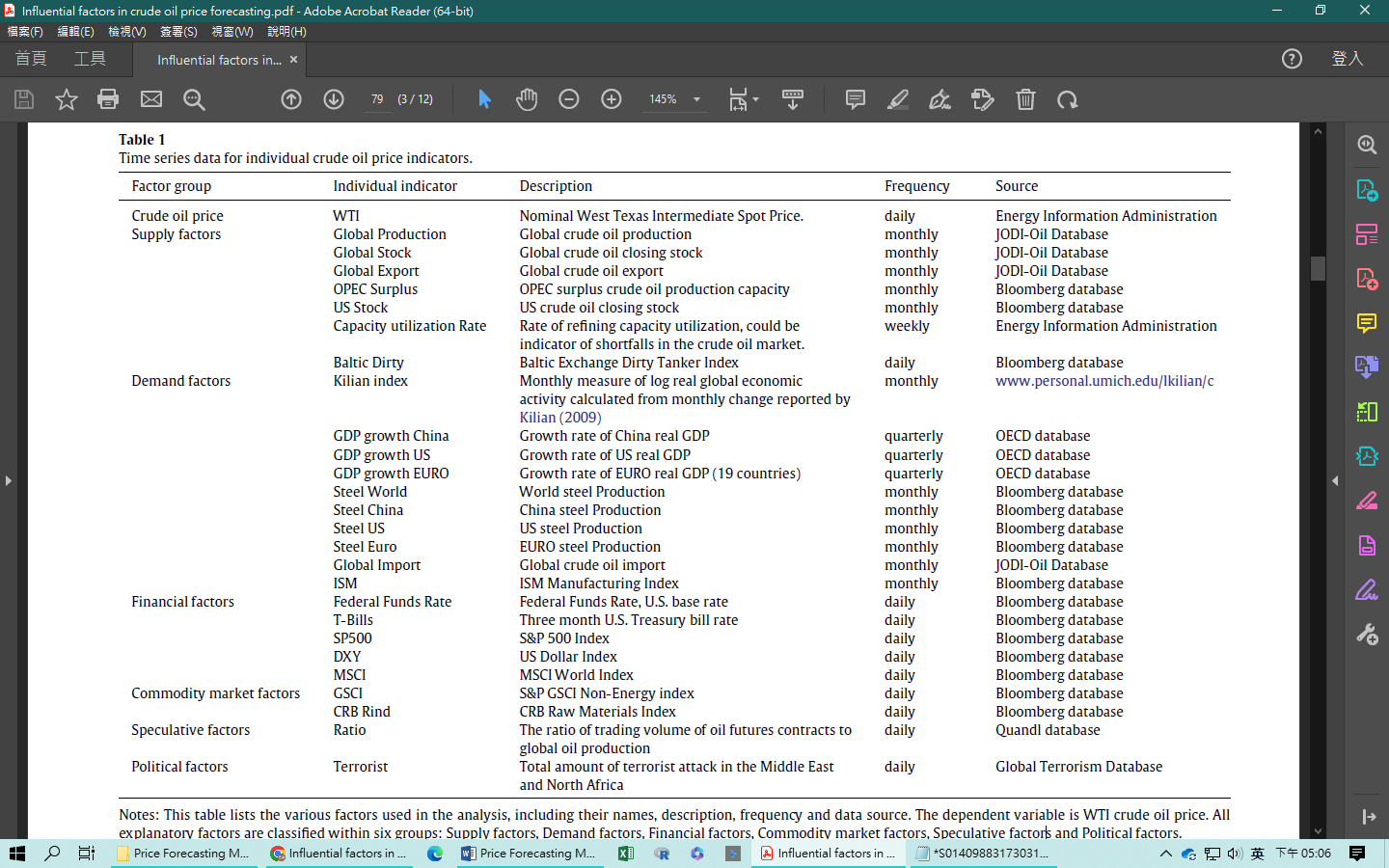
**$40.00**

[Mirmirani, S.](https://www.emerald.com/insight/search?q=Sam%20Mirmirani) and [Cheng Li, H.](https://www.emerald.com/insight/search?q=Hsi%20Cheng%20Li) (2004), "A COMPARISON OF VAR AND NEURAL NETWORKS WITH GENETIC ALGORITHM IN FORECASTING PRICE OF OIL", [Binner, J.M.](https://www.emerald.com/insight/search?q=Jane%20M.%20Binner), [Kendall, G.](https://www.emerald.com/insight/search?q=Graham%20Kendall) and [Chen, S.-H.](https://www.emerald.com/insight/search?q=Shu-Heng%20Chen) (Ed.) *Applications of Artificial Intelligence in Finance and Economics* (*Advances in Econometrics, Vol. 19*), Emerald Group Publishing Limited, Bingley, pp. 203-223. <https://doi.org/10.1016/S0731-9053(04)19008-7>

£23.00 (excl. tax)

The VAR-based forecast uses three endogenous variables: lagged oil price, lagged oil supply and lagged energy consumption.

Hong Miao, Sanjay Ramchander, Tianyang Wang, Dongxiao Yang,Influential factors in crude oil price forecasting,Energy Economics,Volume 68,2017,Pages 77-88,ISSN 0140-9883,https://doi.org/10.1016/j.eneco.2017.09.010.



在頁碼79，檔案的第3頁中table1 time series data for individual crude oil price indicators。將所有重要的變數分類，供給面、需求面、金融面、商品市場指數、投資、地域政治。本篇使用Least Absolute Shrinkage and Selection Operator (LASSO) regression，LASSO是一種篩選變數的方法，Stepwise method是同時使用backward elimination和forward selection的選取變數方法。使用資料範圍是2002/1/4至2015/9/25。比較的對象是現貨價格、期貨價格、以及全部變數放入的模型(full factor-based model)(VAR)、EIA的預測報告。

為了衡量各模型的表現，使用Mean Squared Prediction Error (MSPN)作為指標。並且使用Diebold-Mariano Test (DM test)。方法大致是將預測誤差取平方再相減，再將此差標準化，該統計量會符合標準常態分佈。Pesaran-Timmermann Test是用來評估預測方向的能力，該檢定屬於無母數方法，統計量符合常態檢定，虛無假設是模型無預測方向能力，即預測正確率並不比隨機猜測好。

Ramyar, S., Kianfar, F. Forecasting Crude Oil Prices: A Comparison Between Artificial Neural Networks and Vector Autoregressive Models. *Comput Econ* **53**, 743–761 (2019). <https://doi.org/10.1007/s10614-017-9764-7>

39,95 €

Price includes VAT (Taiwan (R.O.C.))

Fenghua Wen, Minzhi Zhang, Mi Deng, Yupei Zhao, Jian Ouyang,Exploring the dynamic effects of financial factors on oil prices based on a TVP-VAR model,Physica A: Statistical Mechanics and its Applications,Volume 532,2019,121881,ISSN 0378-4371, <https://doi.org/10.1016/j.physa.2019.121881>.

本文認為金融指標對於油價的影響會隨著時間變動，尤其在經歷2008年金融海嘯期間。因此使用Time-Varying Parameter Structural Vector Autoregression (TVP-VAR) model，特點在於可以表現隨時間改變的變數關聯。

本文使用1997/1至2016/6的月資料，使用的變數有WTI原油現貨價格、TWEXB作為美元匯率的指標、黃金價格、S&P 500、商業以及非商業的期貨交易行為。

Likun Lei, Yue Shang, Yongfei Chen, Yu Wei,Does the financial crisis change the economic risk perception of crude oil traders? A MIDAS quantile regression approach,Finance Research Letters,Volume 30,2019,Pages 341-351,ISSN 1544-6123,https://doi.org/10.1016/j.frl.2018.10.016.

本篇文章旨在討論2008金融危機前後原油交易者面對經濟風險的態度，使用經濟政策不確定(economic policy uncertainty, EPU)指數作為經濟危機的代理變數(proxy variable)。結果發現2008年前，EPU和報酬呈負相關，但2008金融危機後卻變成正相關。

使用的模型是MIDAS quantile regression model (Ghysels, 2016)。MIDAS Beta polynomial function的特色在於靈活性，可以包含多種不同的落後結構，並且模型中包含了兩種不同頻率的資料，分別是月資料及日資料。

使用的月資料是WTI原油現貨價格、原油期貨價格、美國生產指數(U.S. Industrial Production index)，資料範圍是1986年2月至2018年6月。而日資料包括經濟政策不確定性指數(U.S. EPU)、市場不確定性(U.S. Equity Market Uncertainty Index)、三個月期國庫券(3-Month T-bill)，範圍是1986/2/1至2018/6/30。

|  |  |  |
| --- | --- | --- |
| Variable | Full name | source |
| Spot oil | WTI crude oil spot prices | <https://www.eia.gov/dnav/pet/pet_pri_spt_s1_d.htm> |
| Oil futres | Oil futures price |  |
| EPU | Economic policy uncertainty index | <https://www.policyuncertainty.com/> |
| EUI | U.S. Equity Market Uncertainty Index | <https://www.policyuncertainty.com/EMV_monthly.html> |
| T-bill | 3-Month T-bill | <https://fred.stlouisfed.org/series/DTB3> |
| IP | U.S. Industrial Production index | <https://fred.stlouisfed.org/series/INDPRO> |

針對變數的處理，將EPU和EUI除以1000，其餘的轉換是取自然對數後進行一階差分，相當於報酬率。所有的數列都具有顯著的偏態及過高的峰值，Jarque-Bera test拒絕常態分佈的假設。Ljung-Box Q statistic檢測在5至20期後自我迴歸消失，顯示序列具有暫時的依賴性。ADF unit root test檢測所有序列都為平穩(stationary)。

Ani Shabri, Ruhaidah Samsudin, "Daily Crude Oil Price Forecasting Using Hybridizing Wavelet and Artificial Neural Network Model", *Mathematical Problems in Engineering*, vol. 2014, Article ID 201402, 10 pages, 2014. <https://doi.org/10.1155/2014/201402>

本篇目的在於使用小波轉換 (wavelet transform) 以及神經網絡模型來預測原油價格。小波轉換克服原本傅立葉轉換的問題，具有可調變的時頻窗口，窗口的寬度可以改變，並且可以具有有限的時常以及突變的頻率與震幅。在神經網絡的隱藏層中多數使用雙曲線正切函數 (hyperbolic tangent sigmoid, tanh)，tanh在原點附近有更大的梯度，可以解決時間序列以及梯度消失問題。最後在輸出層使用線性函數。許多模型都是使用反向傳播 (backpropagation) 進行訓練，但此演算法有慢收斂、不穩健等問題。LM 算法結合了最速下降法（Steepest Descent）和高斯-牛頓法（Gauss-Newton Method）的特點，能夠改善解決非線性問題出現的問題。