Universität St.Gallen

Prediction of BTC energy consumption

Aleksandra Bakhareva (21-604-475)

Mario Blauensteiner (20-602-744)

Paul Kilian Kreutzer (20-600-599)

Tim Ludwig Leonard Matheis (21-603-907)

Smart Data Analytics

University of St.Gallen

https://github.com/QuantLet/SDA_2021_St_Gallen/tree/master/SDA_2021_St_Gallen_Prediction_of_BTC_energy_consumption

Team

Aleksandra Bakhareva
Master of Science in
Quantitative Economics

Mario Blauensteiner
Master of Arts in
Banking and Finance

Paul Kilian Kreutzer
Master of Arts in
Banking and Finance

Tim Ludwig Leonard Matheis
Master of Arts in Banking
and Finance

Outline

1. **Introduction**
2. **Data collection**
3. **Exploratory analysis**
    - Plotting historical BTC network power demand and covariates
    - Descriptive statistics
    - Stationarity test
4. **Historical BTC network power demand prediction**
    - ARIMA, VAR, LSTM
5. **Conclusion**

BTC-related energy consumption

- Hashing activity on BTC blockchain increased 10-billion-fold (2010–2020)
- Total energy consumption related to BTC increased 10-million-fold (2010–2020)
- BTC network cost relative to the volume of transactions stable at ~ 1% since 2010 (Song & Aste, 2020)

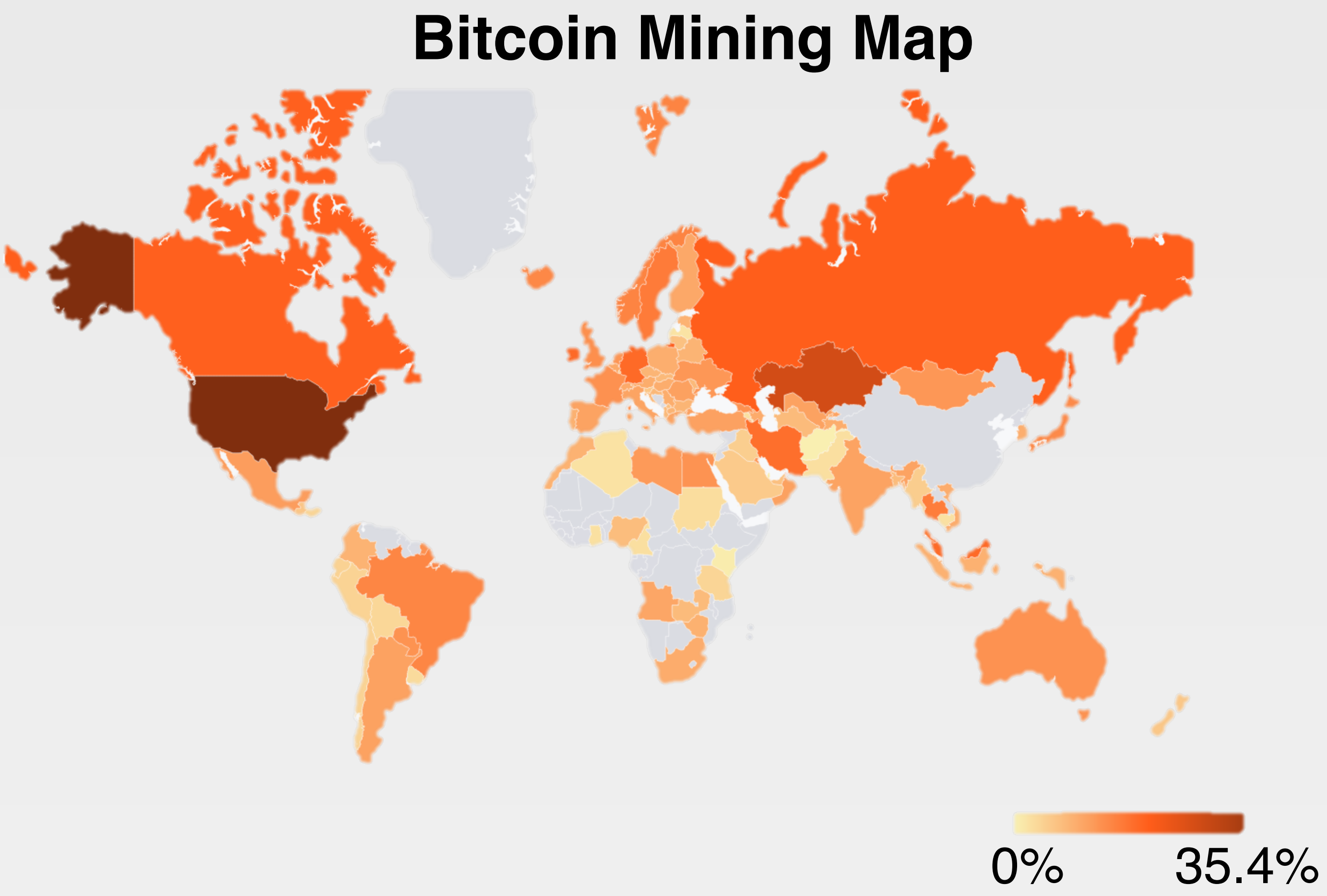
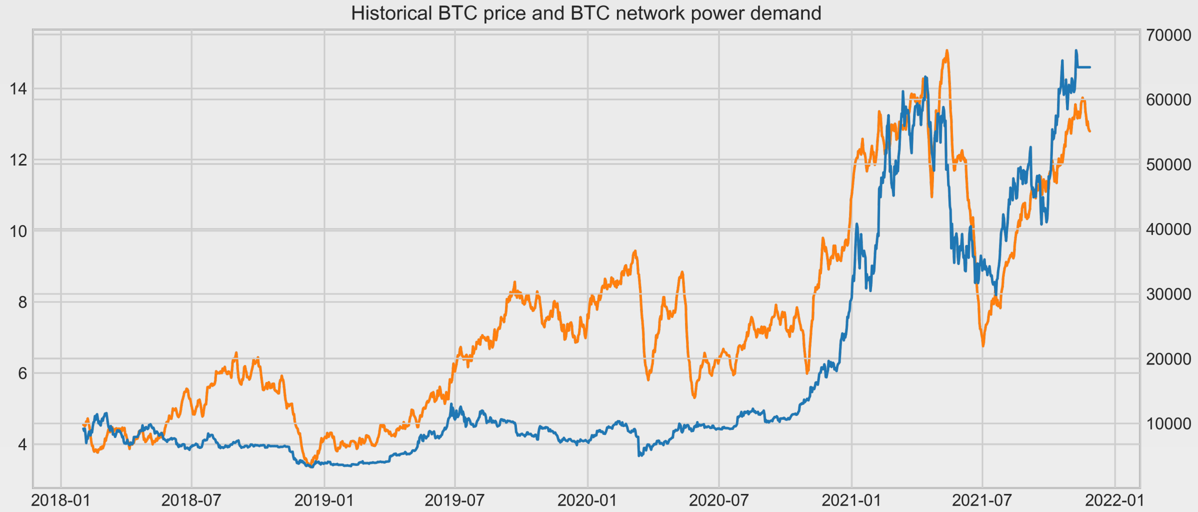


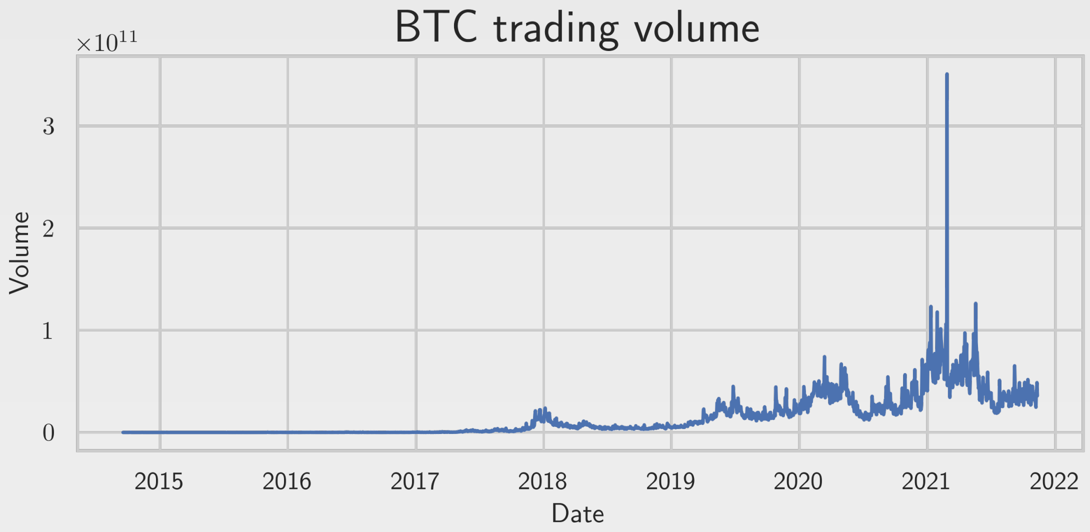
Image source: Cambridge Bitcoin Energy Consumption Index (https://ccaf.io/cbeci/mining_map)

Goal and method

- Using sentiment and historical data to predict future energy consumption related to BTC network
  - Methods used: ARIMA, VAR, and LSTM for forecasting future energy demand, enhanced by sentiment analysis

Data collection (1/4)





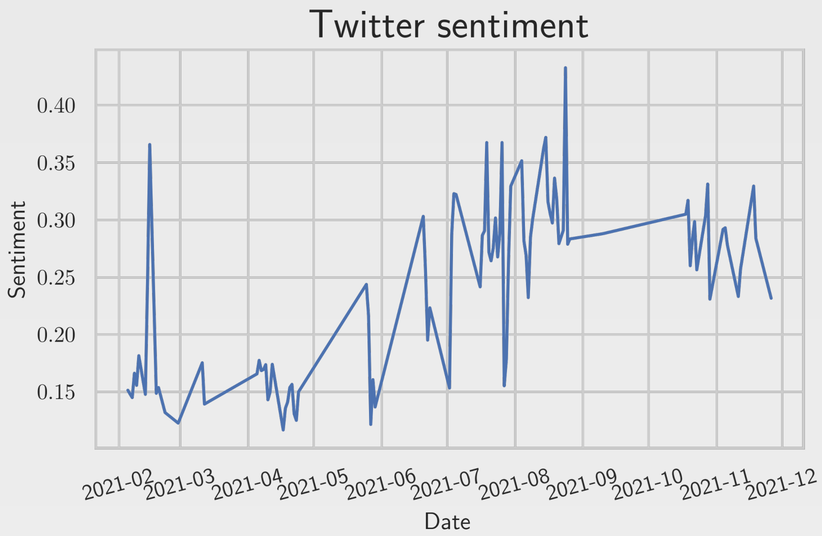
- Sharp increase in BTC price and transaction volume
- Concomitant rise in power demand
- High correlation of BTC **price** and **network energy consumption**

Data sources: coinmarketcap.com, Cambridge Bitcoin Energy Consumption Index (https://ccaf.io/cbeci/mining_map)

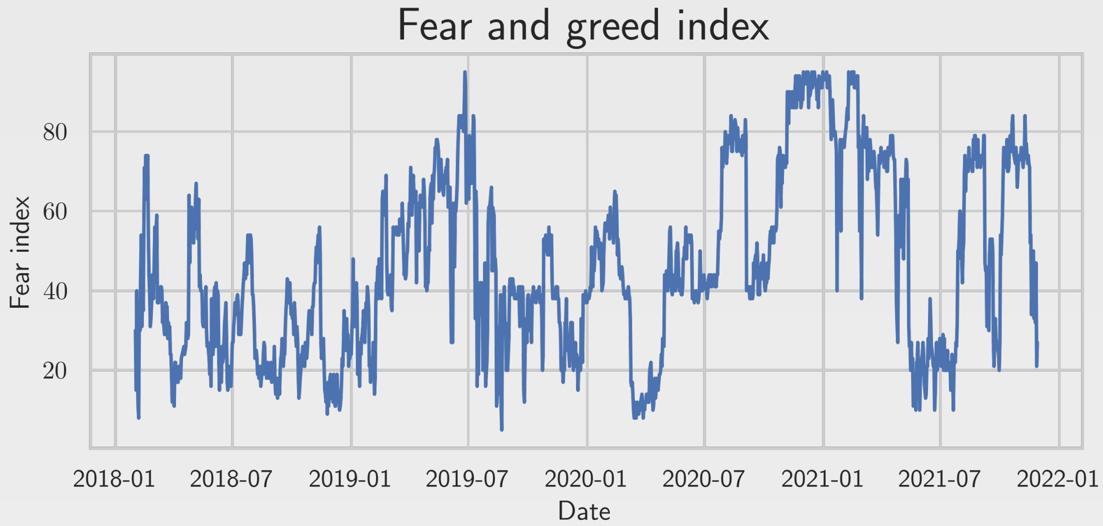
Data collection (2/4)

- Sentiment analysis
  - Crawl data from Twitter using an API (tweepy)
  - Compute sentiment with Vader, a pretrained model
  - Alternatively compute sentiment manually with a formula such as $\frac{\#positive\ words\ -\ \#negative\ words}{\#positive\ words\ +\ \#negative\ words}$ and a list of positive and negative words (provided in data folder)

Data collection (3/4)



from -1 to 1, -1 = terrible sentiment, 1 = great sentiment

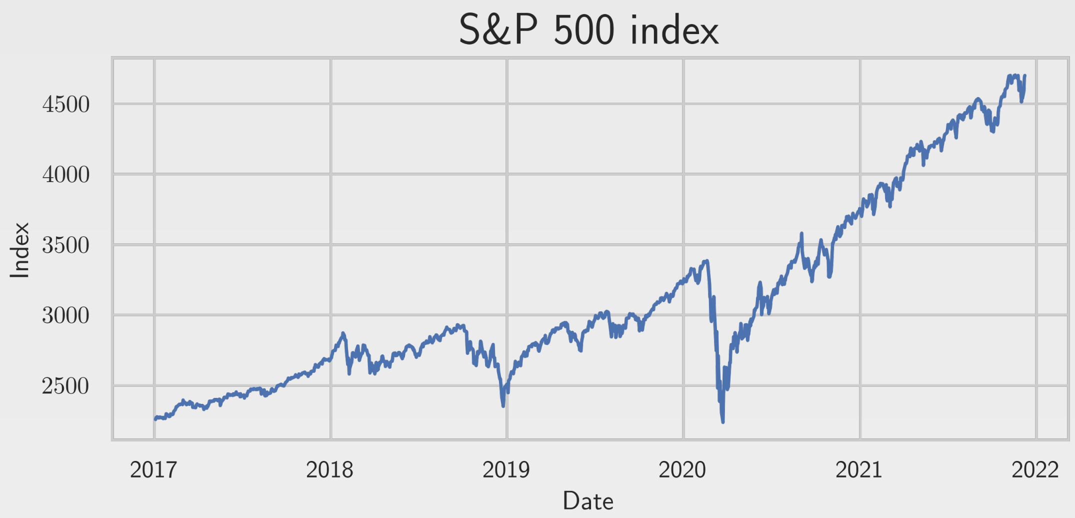


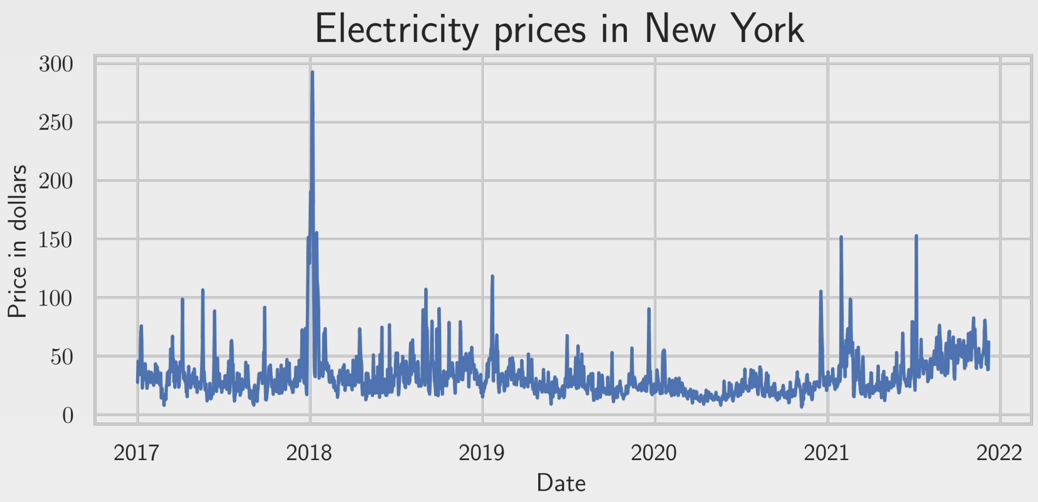
from 0 to 100, 0 = extreme fear, 100 = extreme greed

- Extreme fluctuations in Twitter sentiment on BTC and “Fear and greed index” is also very volatile

Data sources: Twitter, alternative.me

Data collection (4/4)



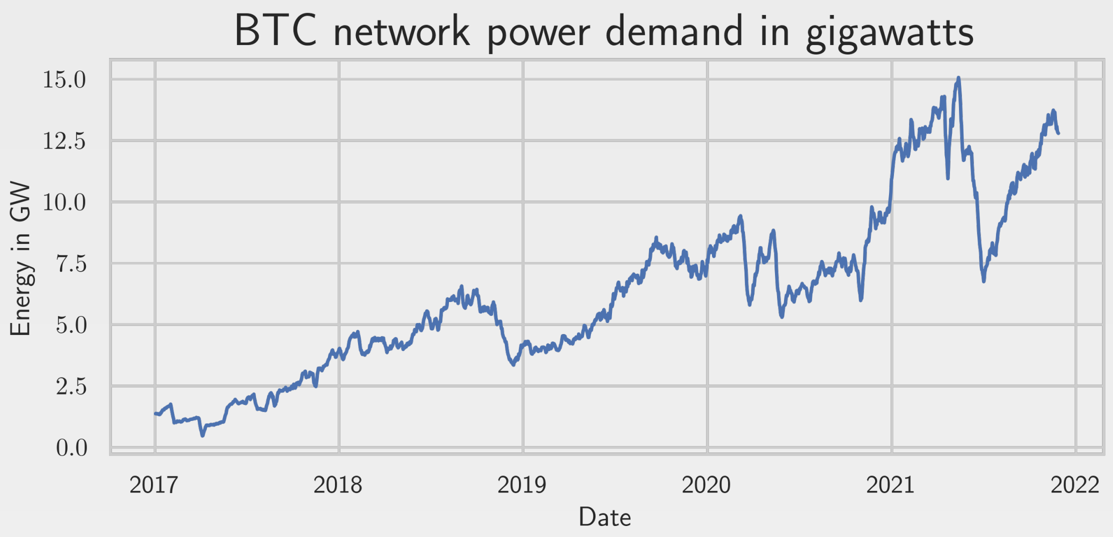


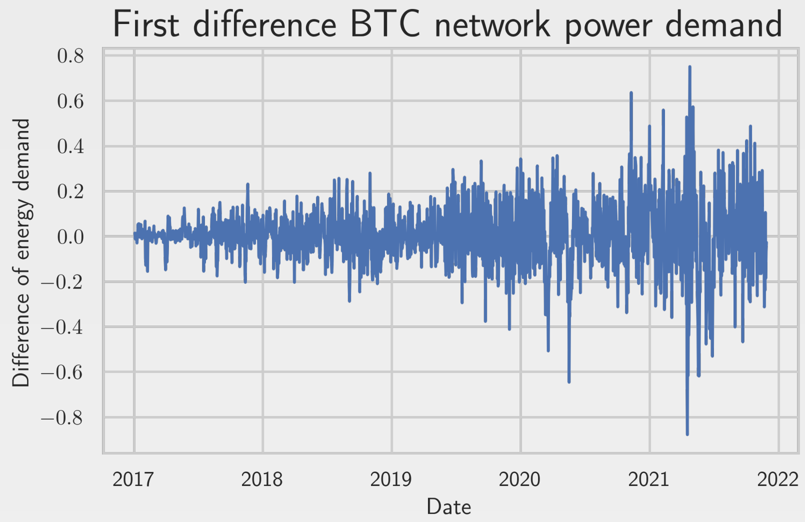
- S&P 500 index as a return benchmark
- High electricity prices potentially make mining less attractive
- High electricity prices could lead to new prime locations for mining

Data sources: MarketWatch, ENGIE Resources

Exploratory analysis (1/6)

- Original network power demand series does not look stationary
- Differentiate the original series to make it stationary
- Differentiated time series looks stationary; tested via Dickey-Fuller test (p-value 0.000)

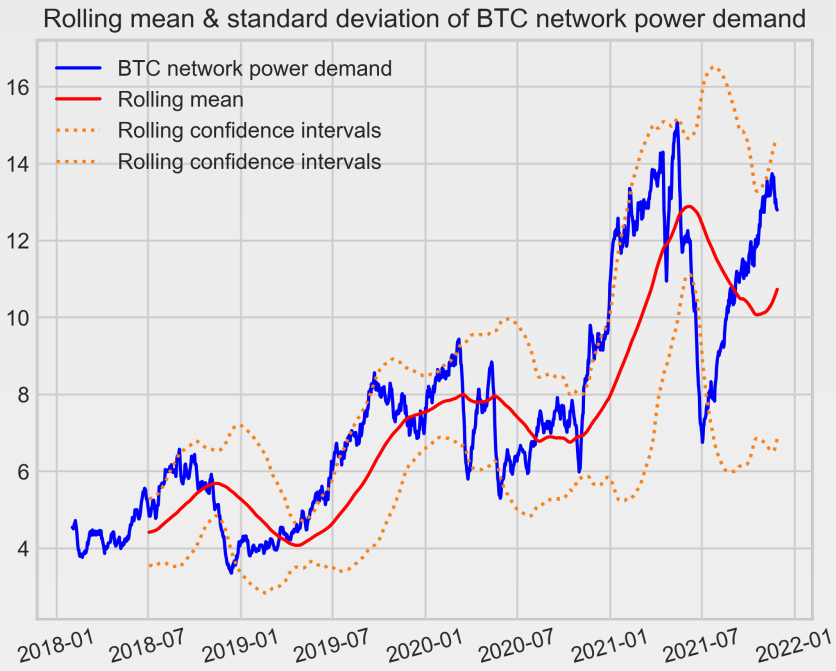


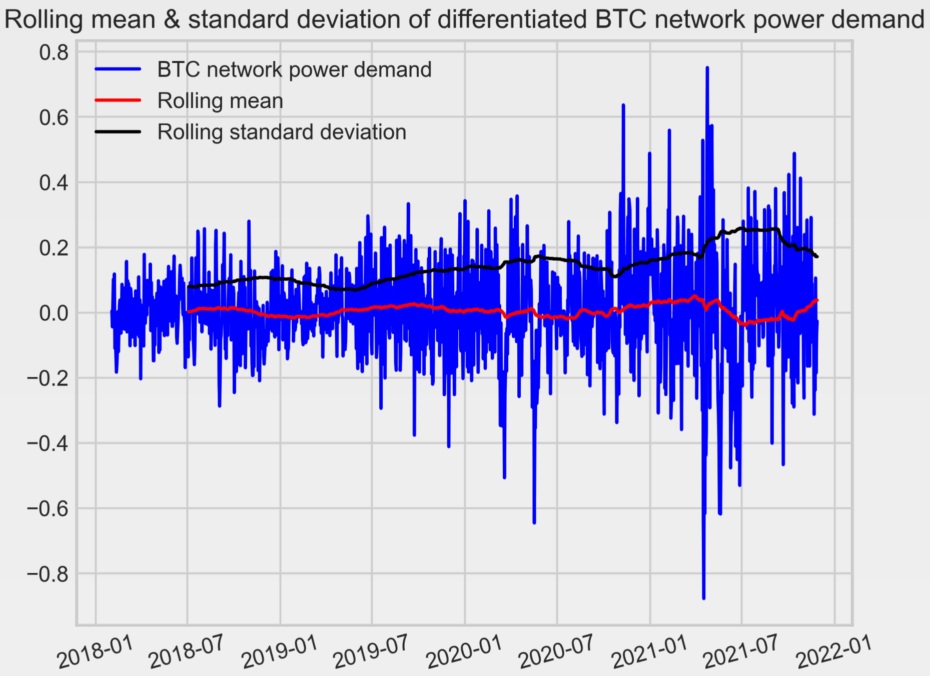


Data source: Cambridge Bitcoin Energy Consumption Index (https://ccaf.io/cbeci/mining_map)

Exploratory analysis (2/6)

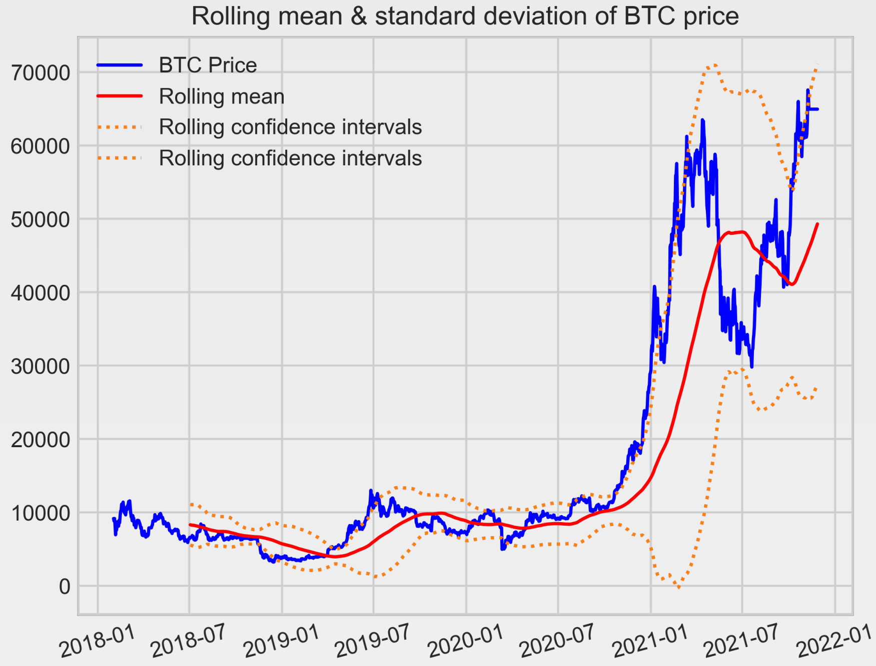
- Calculating rolling mean and standard deviation of BTC network power demand using window of 5 months (150 days)

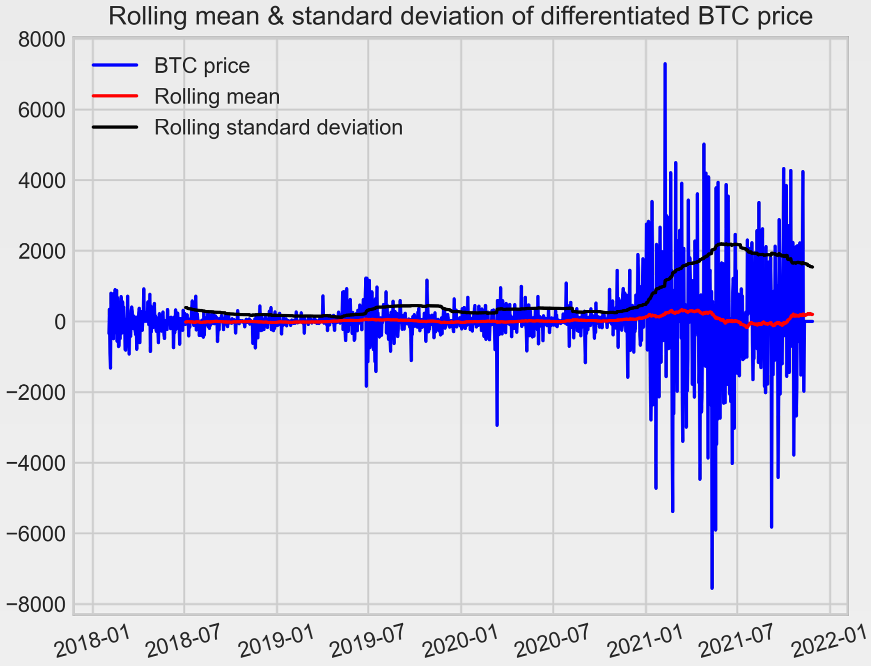




Exploratory analysis (3/6)

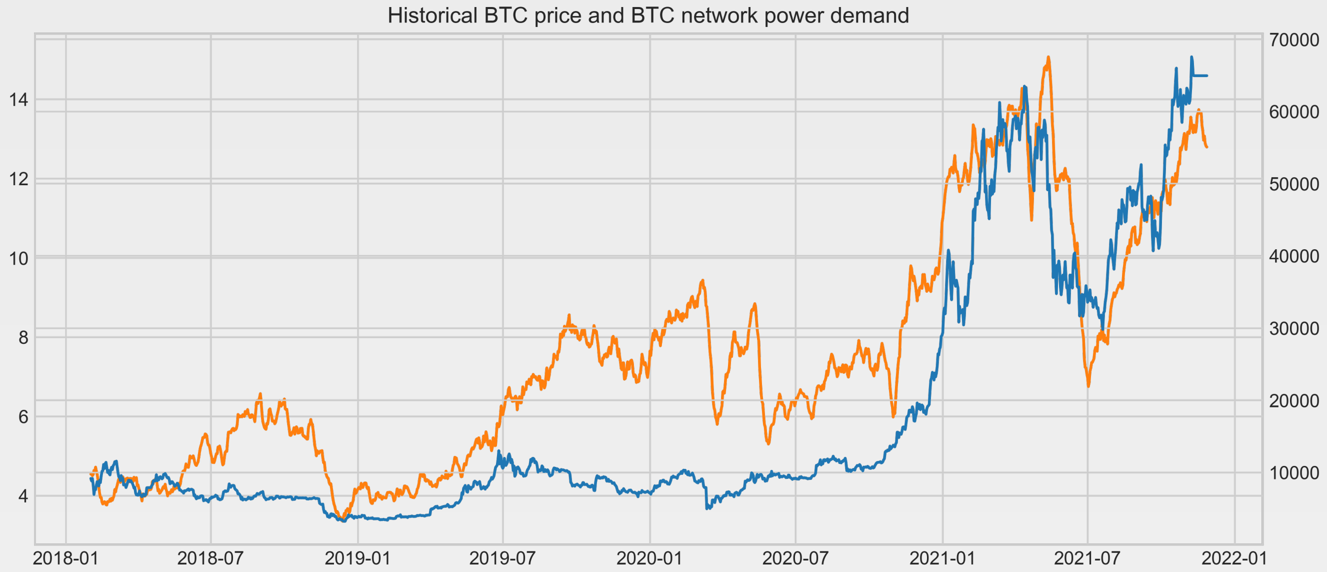
- Calculating rolling mean and standard deviation of BTC price using window of 5 months (150 days)

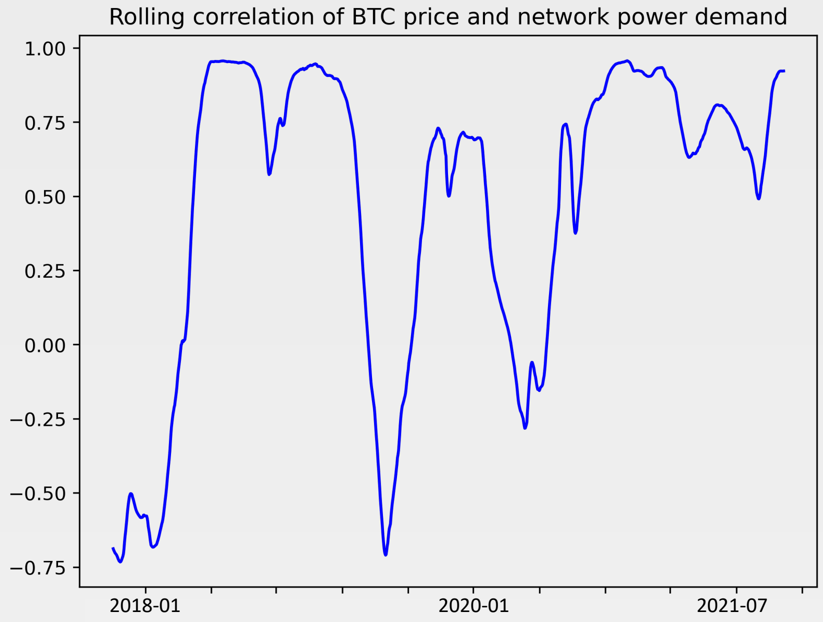




Exploratory analysis (4/6)

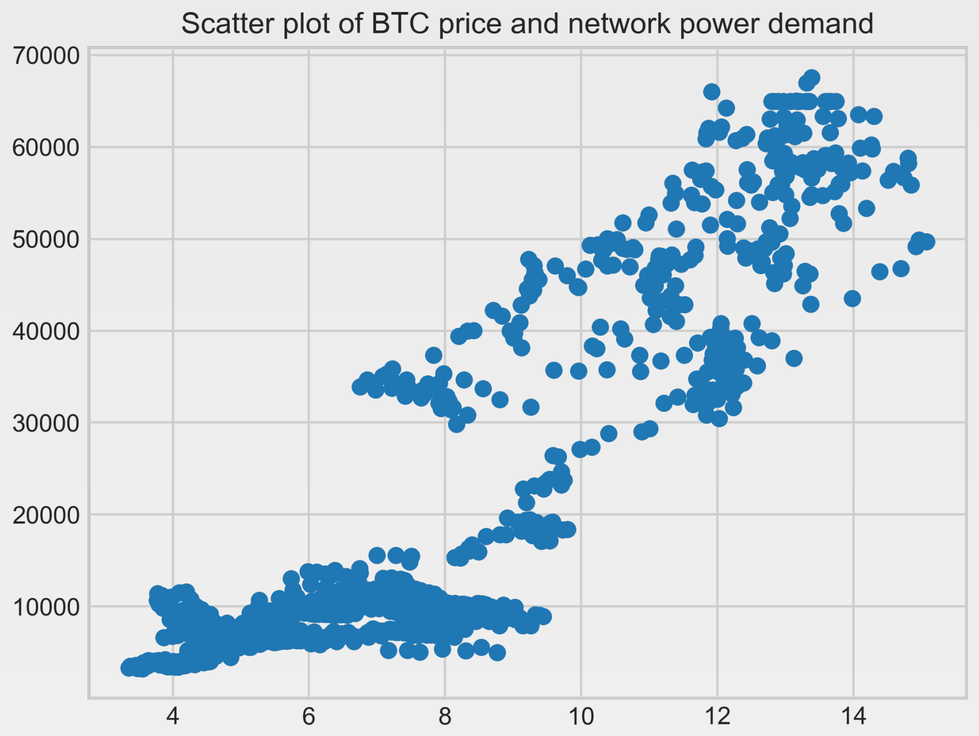
- Calculating rolling correlation between BTC network power demand and BTC price
- Using window of 5 months (150 days)
- Plotting rolling correlation

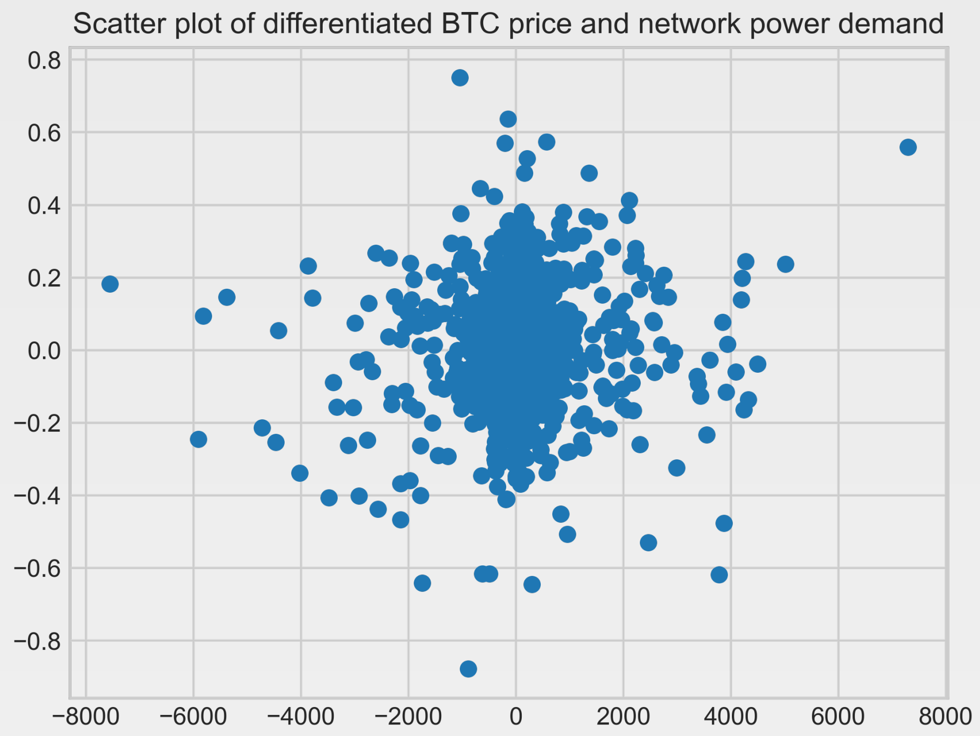




Exploratory analysis (5/6)

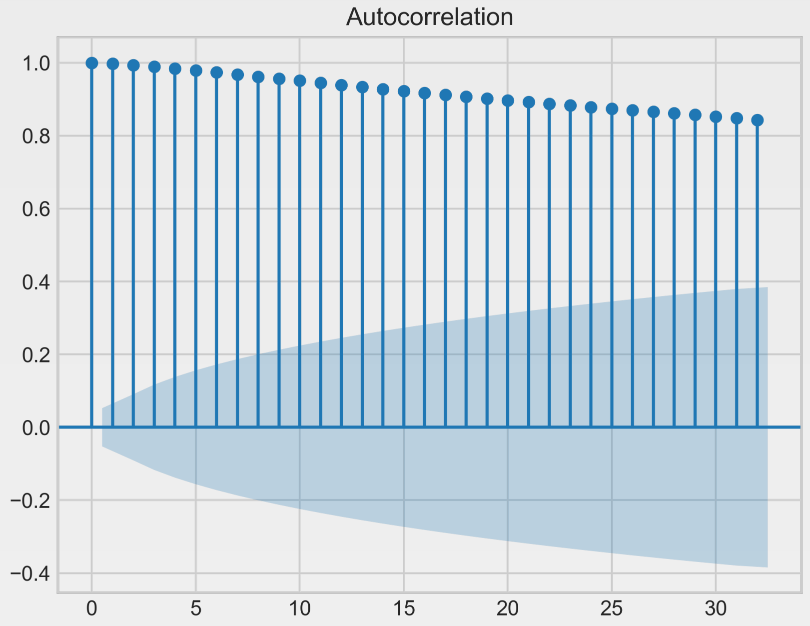
- Scatter plots for BTC price and network power demand and price

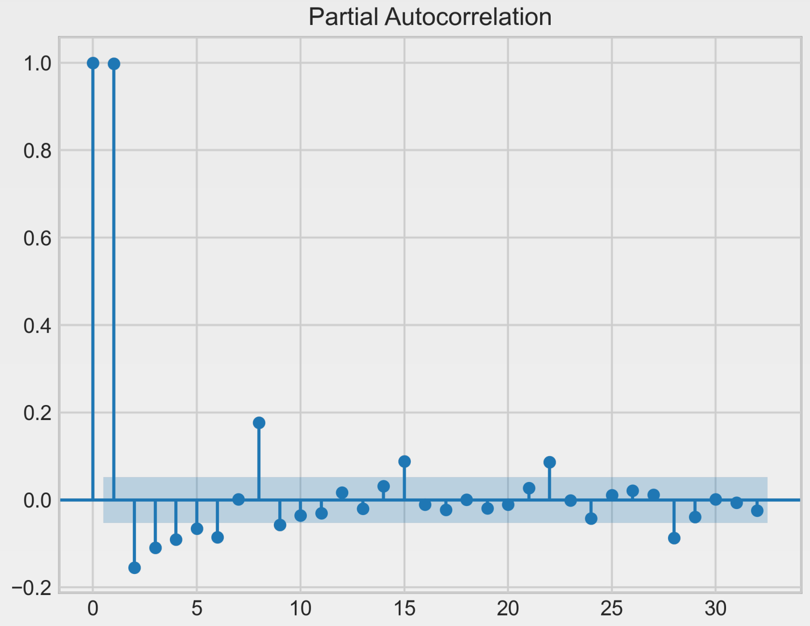




Exploratory analysis (6/6)

- Autocorrelation of historical Bitcoin network power demand
- Decay in autocorrelation and around 5 significant lags in the partial autocorrelation plot
- Assumption: ARMA(5,0)





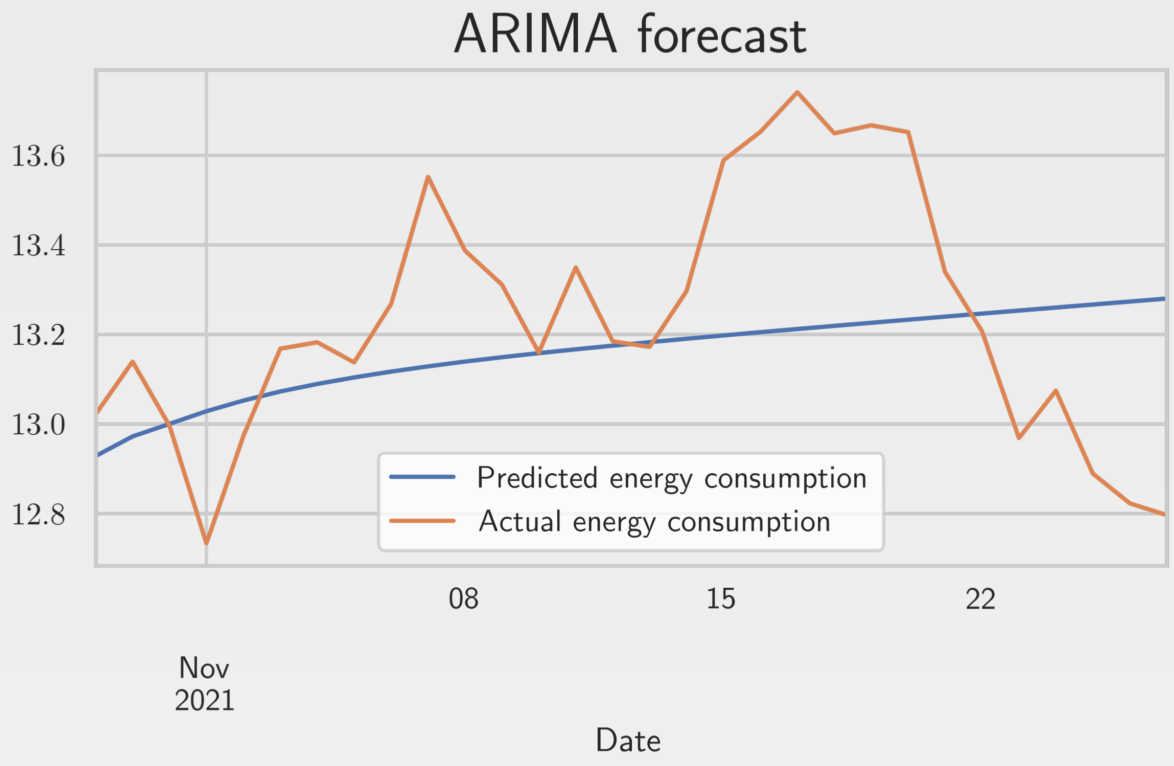
Bitcoin network power demand prediction (1/5)

- Model: ARIMA (5, 1, 0)

ARIMA Model Results

Dep. Variable:	D.GUESS GW	**No. Observations:**	1760
Model:	ARIMA(5, 1, 0)	**Log Likelihood**	1120.211
Method:	css-mle	**S.D. of innovations**	0.128
Date:	Sun, 28 Nov 2021	**AIC**	-2226.423
Time:	21:15:35	**BIC**	-2188.111
Sample:	01-03-2017	**HQIC**	-2212.264
	- 10-28-2021		

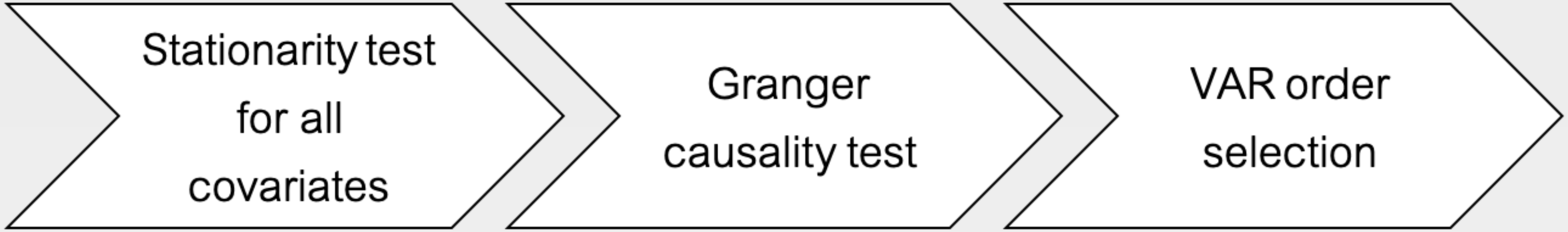
	coef	std err	z	P>|z|	[0.025	0.975]
const	0.0066	0.006	1.049	0.294	-0.006	0.019
ar.L1.D.GUESS GW	0.1842	0.024	7.742	0.000	0.138	0.231
ar.L2.D.GUESS GW	0.1100	0.024	4.572	0.000	0.063	0.157
ar.L3.D.GUESS GW	0.0662	0.024	2.738	0.006	0.019	0.114
ar.L4.D.GUESS GW	0.1024	0.024	4.244	0.000	0.055	0.150
ar.L5.D.GUESS GW	0.0559	0.024	2.345	0.019	0.009	0.103



Bitcoin network power demand prediction (2/5)

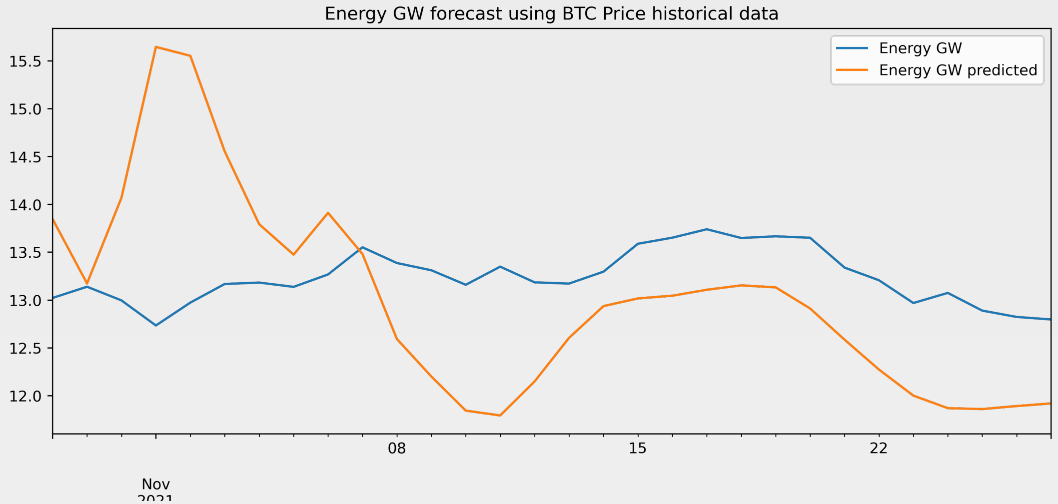
- Vector Autoregression (VAR) model equation

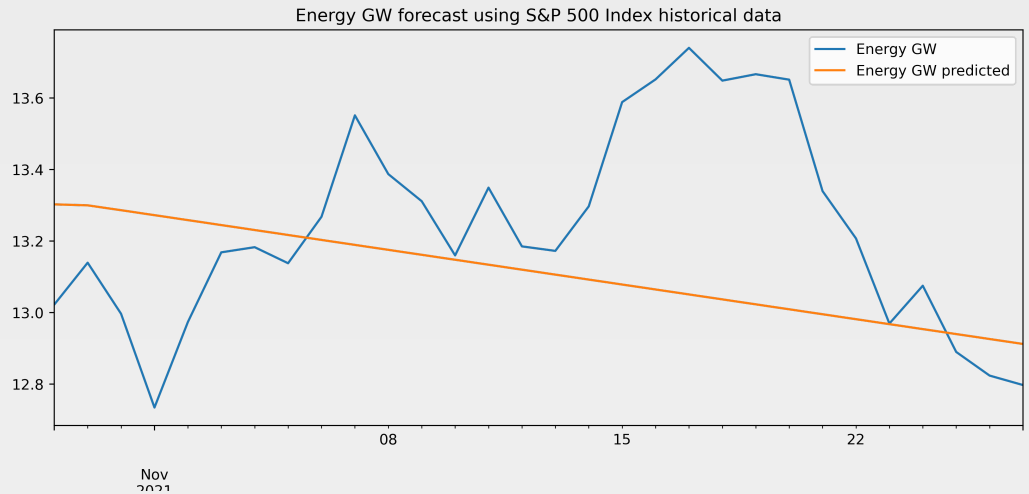
$$Y_{1,t} = \alpha_1 + \beta_{11,1} Y_{1,t-1} + \beta_{12,1} Y_{2,t-1} + \epsilon_{1,t}$$
$$Y_{2,t} = \alpha_2 + \beta_{21,1} Y_{1,t-1} + \beta_{22,1} Y_{1,t-1} + \epsilon_{2,t}$$



Bitcoin network power demand prediction (3/5)

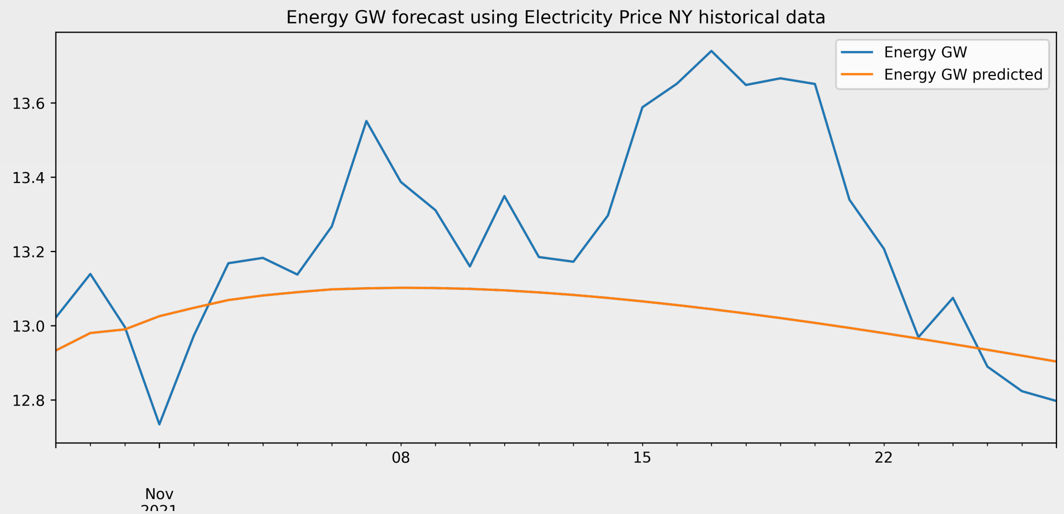
- VAR(11) model using BTC network power demand and BTC price
- VAR(1) model using BTC network power demand and S&P 500 index
- RMSE of 1.09 and 0.33 respectively

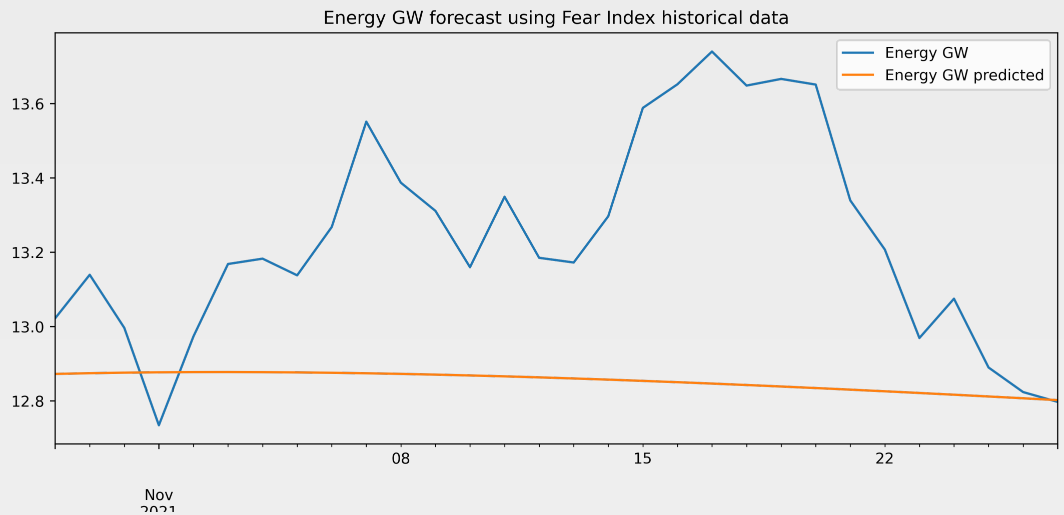




Bitcoin network power demand prediction (4/5)

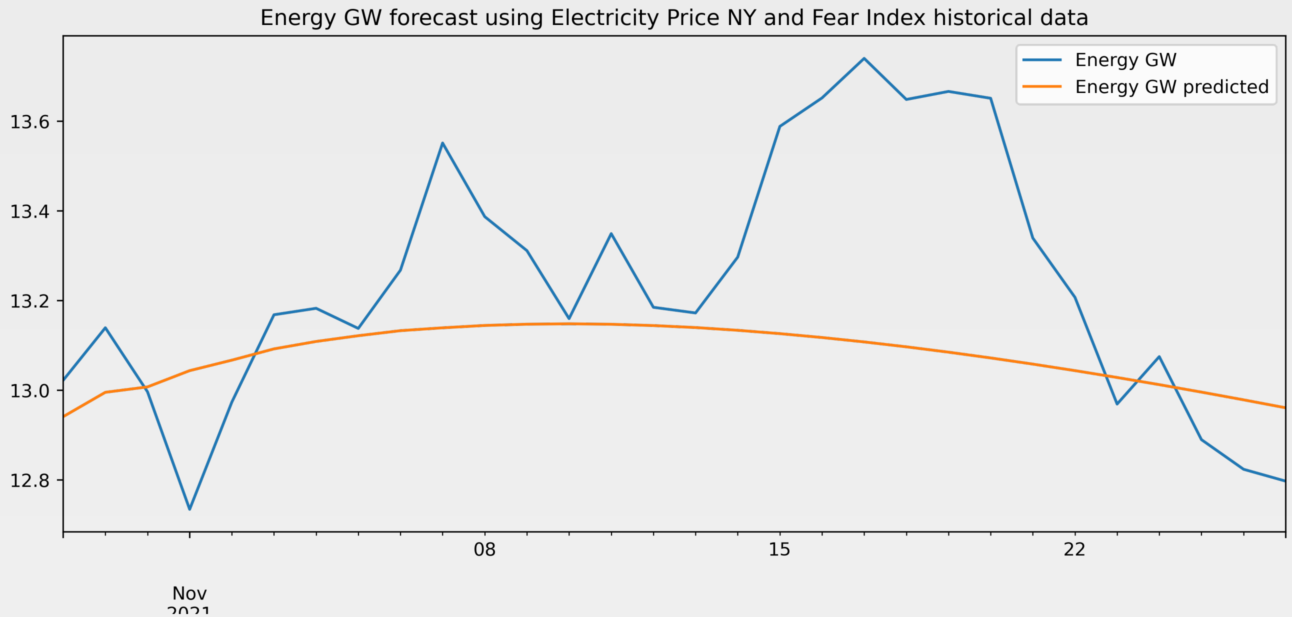
- VAR(7) model using BTC network power demand and electricity price
- VAR(1) model using BTC network power demand and fear index
- RMSE of 0.32 and 0.47 respectively





Bitcoin network power demand prediction (5/5)

- Including 3 variables in VAR model
- VAR(7) model using BTC network power demand, electricity price, and fear index
- RMSE of 0.29



LSTM (= Long Short-Term Memory)

- Recurrent Neural Network (RNN)
- LSTM rectifies a huge issue -> short memory
- LSTM model manages to keep, forget or ignore data points based on a probabilistic model

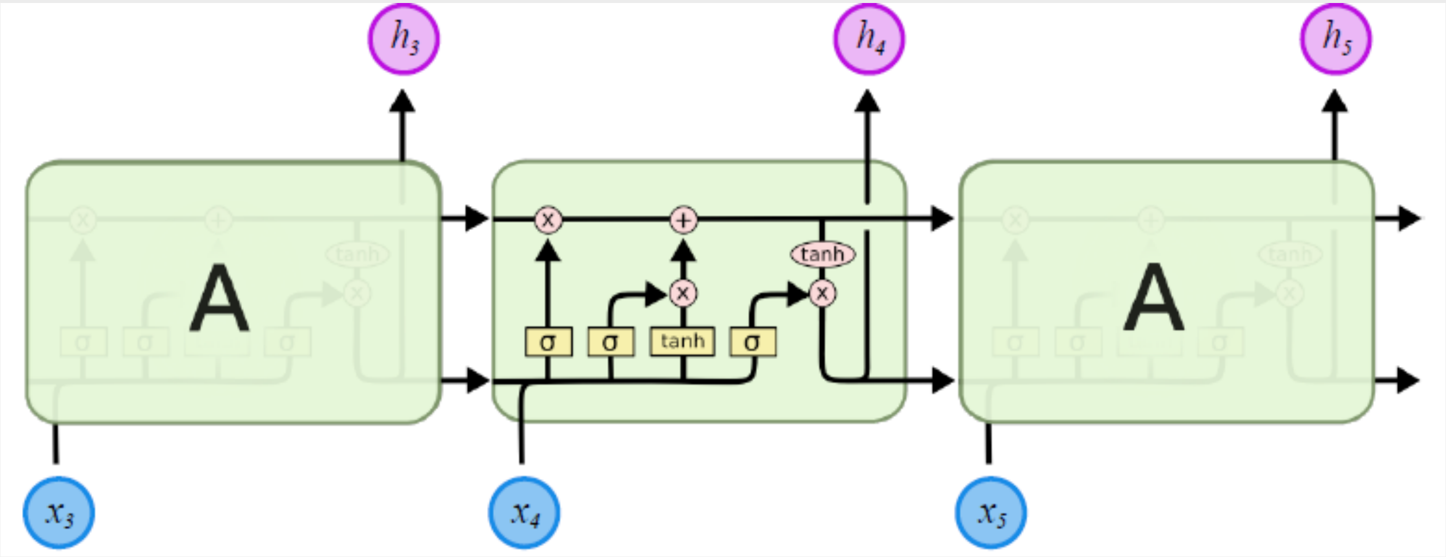
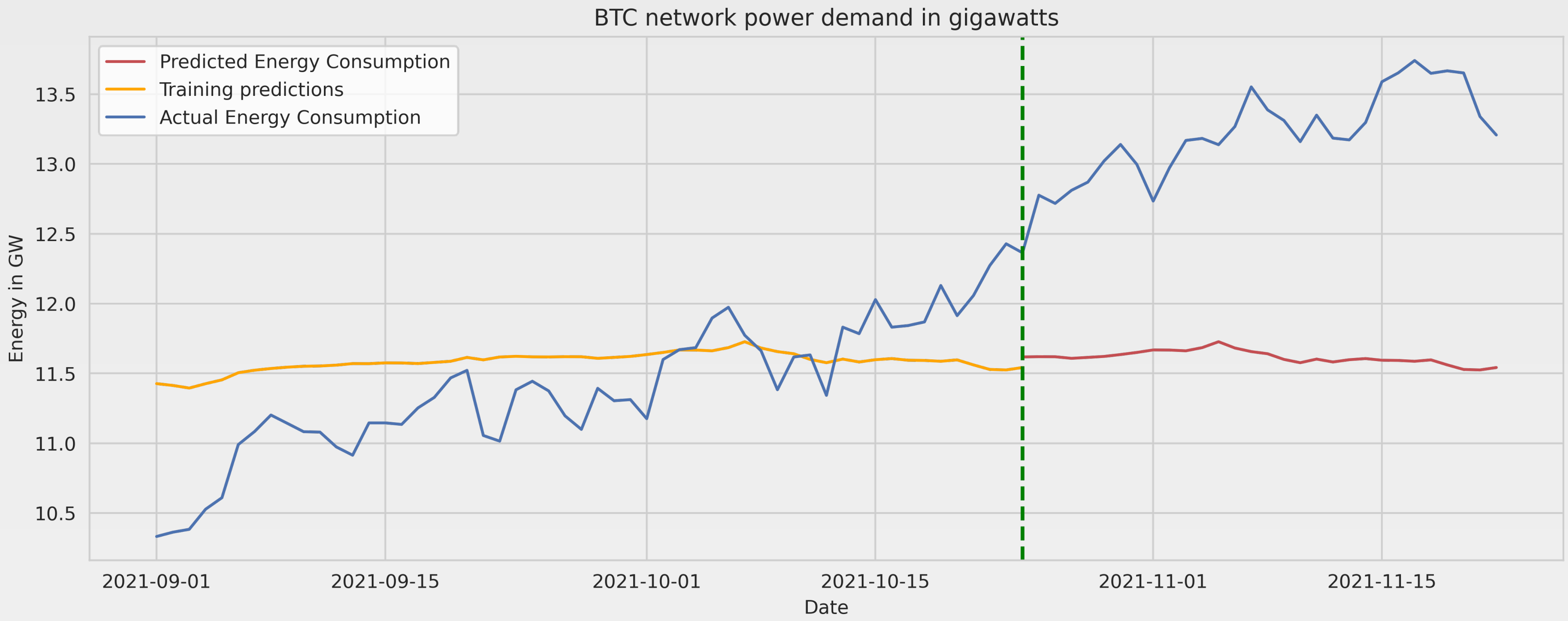


Image source: Universität St. Gallen - Financial Technology Course Lesson 8, Fall 2020, Andrea Barbon

Bitcoin network power demand prediction

- LSTM model using as features BTC network power demand, BTC Price, BTC volume, fear index, S&P 500 and electricity price



Comparison of predictions

	ARIMA (5, 1, 0)	VAR (7)	LSTM
RSME	0.28	0.29	0.90
Data	Requires only energy consumption data	Additional data for covariates is required	In theory, the more data, the better
Advantages	Easy to apply, relatively robust especially when generating short-term forecast	Easy to estimate, good forecasting capabilities, and no need to specify endogeneous and exogeneous variables	Capable of learning long-term dependencies, more robust to the problem of short memory
Drawbacks	Poor at predicting turning points, models are "backward looking"	Many parameters to be estimated	High computing complexity and required memory

Conclusion

- We used daily historical data of Bitcoin network power demand, Bitcoin price, Bitcoin trading volume, fear index, S&P 500 index, and electricity price starting from 01.02.2018 till 27.11.2021 to predict Bitcoin network power demand for the next 30 days
- Our prediction using ARIMA (5, 1, 0) model and historical Bitcoin network power demand daily data has RMSE of 0.28 and perfectly catches the trend
- We applied VAR to predict Bitcoin network power demand using historical daily data of this variable and one additional variable, which lead us to 4 outputs where the VAR(7) model involving historical daily data of electricity price shows the lowest RMSE of 0.32
- Our prediction using VAR(7) model with 2 additional covariates (electricity price and fear index) showed the best result among all VAR models with RMSE of 0.29
- We applied LSTM model to predict Bitcoin network power demand using historical daily data of all other covariates, which showed the result of RMSE equals to 0.90
- Inclusion of additional variables may overload the model and decrease Root Mean Square Error, which is corroborated by the fact that our forecast using ARIMA model showed the lowest RMSE comparing with VAR and LSTM
- GARCH model could lead to improvement of the results compared to ARIMA since it takes volatility into consideration

Data sources

- BTC price data: obtained with web-scraping from finance.yahoo.com; 10 February 2017 – 10 November 2021
- Energy consumption: downloaded from https://ccaf.io/cbeci/index; daily data points 10 February 2017 – 10 November 2021
- Twitter sentiment: obtained with a crawler from twitter.com
- Fear and greed index: downloaded from https://alternative.me/crypto/fear-and-greed-index/; daily data points 1 February 2018 – 10 November 2021
- S&P 500 data: https://www.marketwatch.com/investing/index/spx/download-data?startDate=1/1/2017&endDate=12/8/2021 /; daily data points 1 February 2018 – 10 November 2021
- Electricity prices in New York: downloaded from https://www.engieresources.com/historical-data#reports_anchor /; daily data points 1 February 2018 – 10 November 2021

Sources

- Song, Y. D., & Aste, T. (2020). The cost of Bitcoin mining has never really increased. Frontiers in Blockchain, 3, 44.
- Bentour, El Mostafa. (2015). A ranking of VAR and structural models in forecasting. MPRA Paper 61502. University Library of Munich, Germany.

Prediction of BTC energy consumption

Aleksandra Bakhareva (21-604-475)

Mario Blauensteiner (20-602-744)

Paul Kilian Kreutzer (20-600-599)

Tim Ludwig Leonard Matheis (21-603-907)

Smart Data Analytics

University of St.Gallen

https://github.com/QuantLet/SDA_2021_St_Gallen/tree/master/SDA_2021_St_Gallen_Prediction_of_BTC_energy_consumption