Machine Learning the Carbon Footprint of Bitcoin Mining

15th Risks International Forum 2022,
'Climate Footprint of Bitcoin'

H. Calvo-Pardo, T. Mancini and J. Olmo
(Economics, FSS, U. of Southampton)

22 Mar 2022
CCI-RIdF-Paris

Question and Motivation

- *Does Bitcoin mining contribute to climate change?*
  - Bitcoin blockchain validation process requires specialized hardware and vast amounts of electricity, translating into a significant carbon footprint.
  - e.g. the 2017 carbon footprint of Bitcoin reached 69 million metric tons of CO_2-equivalent ($MtCO_2e$), forecasting a violation of the Paris COP21 UNFCCC Agreement by 2040 *due to* Bitcoin's cumulative emissions alone (Mora et al., 2017);
  - But controversial: subsequent estimates heavily revise downwards Mora et al.'s (2017) projections...
  - Why? Because miners are globally geo-located, facing very different energy costs, and employ hardware with unknown energy intensities $\Longrightarrow$ Difficulty in measuring the Bitcoin mining network power consumption;
  - And, *the recent five-fold increase in Bitcoin prices (€35,561.49 as of 14/03/2022) has heightened public concern, despite of China's recent ban on Bitcoin mining, Elon Musk's rejection of Bitcoin as payments for Tesla cars, etc.*

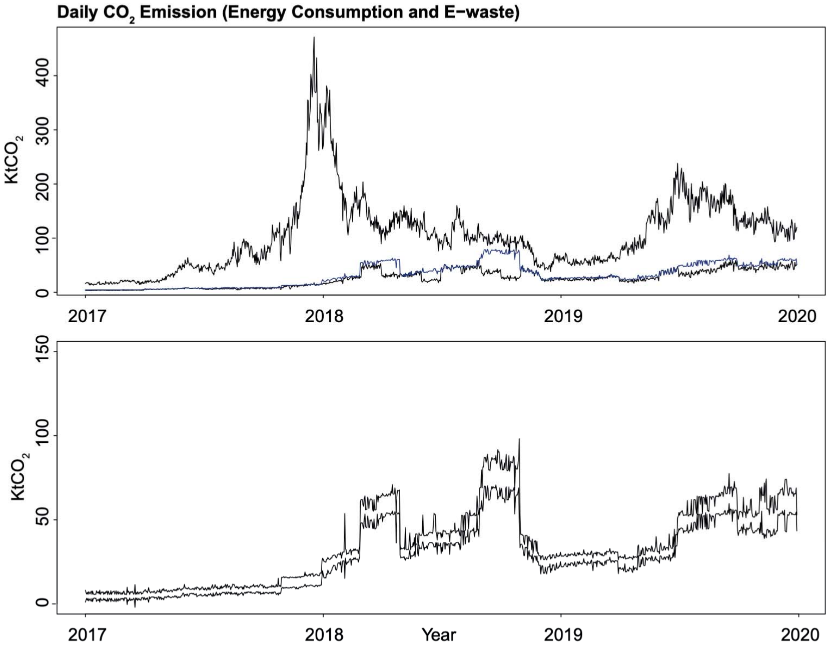

Market Summary > Bitcoin
35,561.49 EUR
+35,254.47 (11,482.78%) today
14 Mar, 12:05 UTC · Disclaimer
1D
5D
1M
6M
YTD
1Y
5Y
Max
60,000
50,000
40,000
30,000
20,000
10,000
0
2018
2020
7,213.89 10 Jan 2020
1
BTC
35561.49
EUR

What do We Do

- We deploy statistical/machine learning (ML) methods (*feedforward deep neural networks*, DNNs) to measure the carbon footprint of Bitcoin mining ('target/output') and the associated uncertainty (prediction intervals, PIs), which:
  - Are absent from the ongoing debate, frames it, and are crucial to inform policies**;**
  - *Nest* existing techno-economic approaches, and are superior to statistical methods, because
  - Bitcoin miners geo-location, actual sources of energy and hardware intensities used are *unobserved* $\Longrightarrow$ model misspecification mistakes compromise reliable statistical/causal inference;
  - ML automates model selection and estimation, replacing exact functional form specification and covariate selection by '*high quality approximation*' ('universal approximation theorems') of the unknown input-output relationship $\Longrightarrow$ reliable statistical inference;
  - Inputs ($P = 42$): (i) predictors of the Bitcoin price level (e.g. monetary economics); (ii) factors driving investors' interest in/attention to the cryptocurrency; (iii) exchange rates with other currencies; or (iv) supply-side factors for the costs incurred by Bitcoin and ASIC mining chips producers, related to rational for-profit mining decisions.

What do We Find I

1. Substantial *uncertainty reduction* around the estimated CO_2 emissions, relative to the economic upper and lower bounds (Figure 1, upper panel) when compared to the associated 95% PIs (Figure 1, lower panel):



Daily CO_2 Emission (Energy Consumption and E−waste)
$KtCO_2$
0
100
200
300
400
2017
2018
2019
2020
$KtCO_2$
0
50
100
150
2017
2018
Year
2019
2020

What do We Find II

2. Substantially *lower* estimated CO_2 annual emissions, from our novel bottom-up approach:

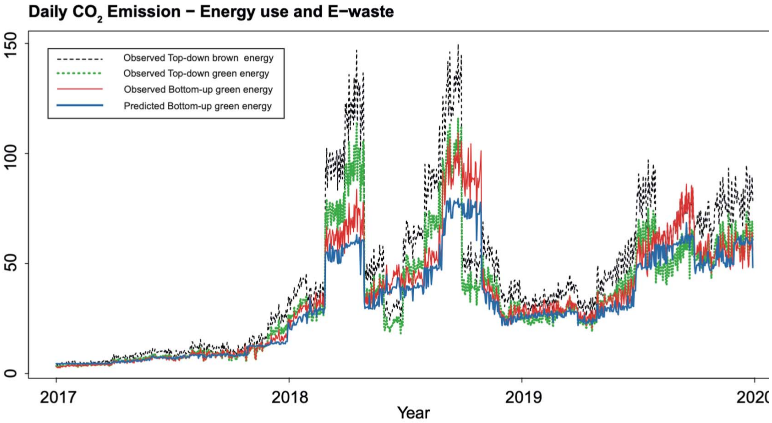
$$\underset{[\text{ktCO}_2 \text{ per day, per TH/s}]}{CO_2^r} = E^r \cdot I + CO_2^{rw} = PUE \cdot e^r \cdot H \cdot I \times 10^{-9} + CO_2^{rw}$$

$$\underset{[\text{ktCO}_2 \text{ per day, per TH/s}]}{CO_2^{er}} = E^r \cdot I^e + CO_2^{rw} = 1.05 \cdot e^r \cdot H \cdot I^e \times 10^{-9} + CO_2^{rw}$$

$$\underset{[\text{ktCO}_2 \text{ per day, per TH/s}]}{CO_2^{BU}} = 1.05 \cdot \sum_{m=1}^{M} e_m^r \cdot \sum_{c=1}^{C} s_{cm}^{ASIC} \cdot I_c^e \cdot H_c \times 10^{-9} + CO_2^{rw}$$

Optimal ReLu DNN Target/Year	2017	2018	2019
CO_2^{BU} ($MtCO_2e$) [95% PI]	2.77 [1.98,3.56]	16.08 [14.19,17.97]	14.99 [13.25,16.73]
CO_2^{re} ($MtCO_2e$) [95% PI]	2.98 [0.42,6.70]	18.11 [16.34,19.88]	17.45 [15.76,19.14]
CO_2^{r} ($MtCO_2e$) [95% PI]	3.72 [2.90,4.54]	23.98 [22.46,25.51]	20.06 [18.53,21.59]

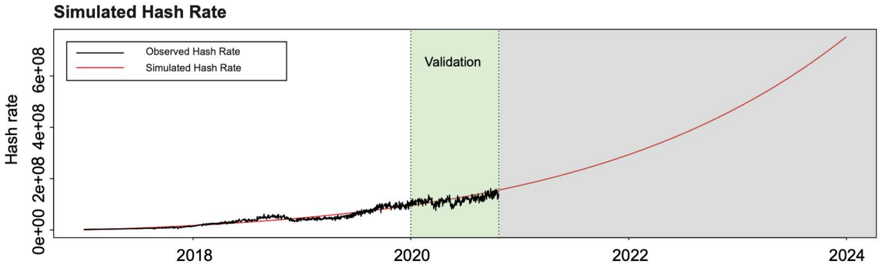
Figure: 'Brown' (energy intensity I, CO_2^r in black) and 'green/clean' (energy intensity I^e, CO_2^{re} in green), 'green' bottom-up approach (CO_2^{BU} in red). ReLu DNN point estimates of $\hat{CO}_2^{BU}$ in blue:

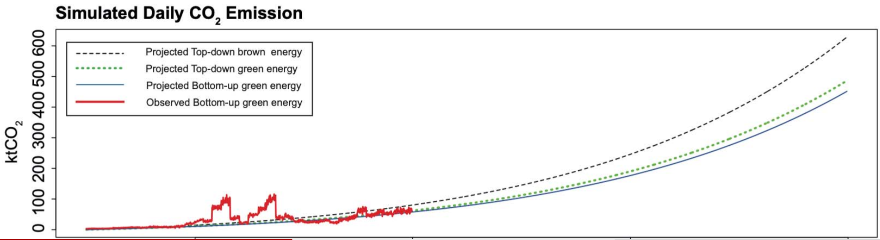


What do We Find III

3. Yet, raising alarming inmediate concerns about Bitcoin mining GHG annual emission levels, forecasted to increase to 29.05 by the end of 2021, to 50.46 by 2022, and to 83.41 by 2023, to reach an alarming 132.01 by the end of 2024, all in $MtCO_2$-e, *similar to the combined annualized 2019 GHG of Belgium (100 $MtCO_2e$) and Denmark (32 $MtCO_2e$)*:

Figure: Top panel: observed (black) and simulated (red) Bitcoin network hashrate $\widehat{H}_t = (3.52 \times 10^4 \cdot t) \cdot exp\{8.31 \times 10^4 \cdot t\}$. Bottom panel: Projected CO_2 emissions for CO_2^{BU} (blue), CO_2^r (black), and CO_2^{re} (green); observed CO_2 (red).





Conclusion

- We deploy supervised ML methods (Optimal ReLu DNN) to better, more reliably and *timely* assess concerns with the carbon footprint of Bitcoin mining, that:
  1. Improve upon existing (techno-economic and statistical) approaches, based on an economic model of rational Bitcoin mining;
  2. Encompass available estimates, framing the debate, and yet
  3. Raise immediate concerns, calling for urgent policy action while offering a novel methodology to track and evaluate alternative policies;
- Next: Policy evaluation calls for causal/counter-factual analysis with ML tools, e.g. Athey and Wager (2021, *Econometrica*) for Random Forests, or Farrell et al. (2021, *Econometrica*) for DNNs.
- Open access publication: *J. Risk Financial Manag.* 2022, 15(2), 71.

Methods I

- Why ML methods, and within those, DNNs?

Method	Target Output	MAE	MSE	RMSE
Optimal ReLu DNN	CO_2^r	8.29	123.97	11.13
Optimal ReLu DNN	CO_2^{re}	6.17	58.76	7.67
Optimal ReLu DNN	CO_2^{BU}	**4.50**	**33.59**	**5.80**
Optimal ReLu DNN, no inputs	CO_2^{BU}	18.37	363.56	19.07
Cross-validated ReLu DNN	CO_2^{BU}	5.35	48.48	6.96
Random Forest	CO_2^{BU}	7.17	82.62	9.09

- Pairwise model comparison test statistic of the difference in out-of-sample MSE of our optimal ReLu DNN against:
  - (rf) random forest: 3.77 (p-value < 0.0001), and
  - (cv) equally-sized cross-validated ReLu DNN: 1.93 (p-value of 0.0269)

Methods II

- What are Optimal ReLu DNNs?
  - In Calvo-Pardo, H. F., Mancini, T. and Olmo, J. (2020) "Optimal Deep Neural Networks by Maximization of the Approximation Power", SSRN, we show that recent advances in combinatorial optimization software (RStudio) can be exploited to optimally allocate hidden units ($\{Z_l\}_{l=1}^{L}$) within ('width') and across layers in deep architectures of a given size $Z = \sum_{l=1}^{L} Z_l$. Adopting the lower bound on the maximal number of linear regions that ReLu DNNs can approximate as maximization criterion (Montufar, Pascanu, Cho and Bengio, 2014), $LB(L, \{Z_l\}_{l=1}^{L-1}; P) \equiv \left(\prod_{l=1}^{L-1} \left\lfloor \frac{Z_l}{P} \right\rfloor^{P} \right) \sum_{r=0}^{P} \binom{Z - \sum_{l=1}^{L-1} Z_l}{r}$, the optimal depth $\widehat{L}$ and width $\{\widehat{Z}_l\}_{l=1}^{\widehat{L}}$ of a DNN obtains from:

$$\text{(OPT)} \quad (\widehat{L}, \{\widehat{Z}_l\}_{l=1}^{\widehat{L}}) \in \arg \max_{(L, \{Z_l\}_{l=1}^{L-1})} LB(L, \{Z_l\}_{l=1}^{L-1}; P)$$

  - Since (OPT) finds the optimal depth and width (layer-wise) *given the network architecture size*, $Z = \sum_{l=1}^{L} Z_l$, bigger and more complex datasets $\{y_i, \mathbf{X}_i\}_{i=1}^{N}$ naturally summon architectures with more hidden units, Z.