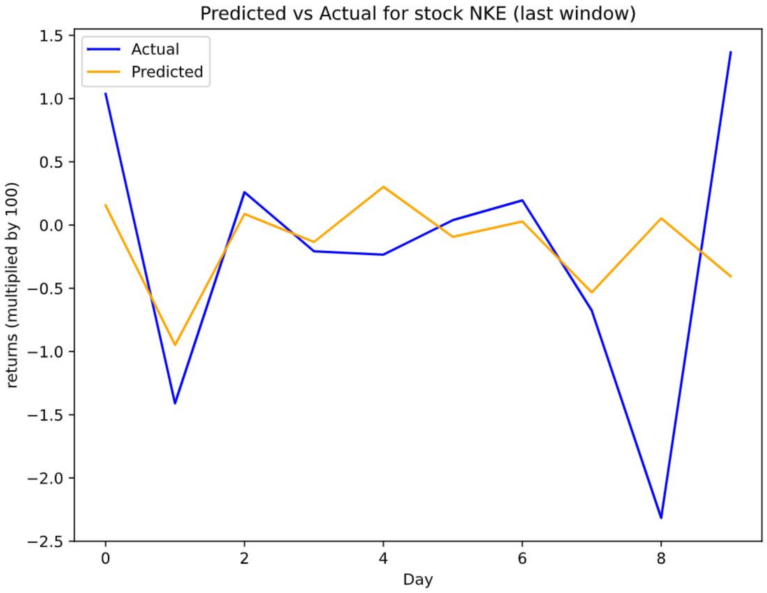
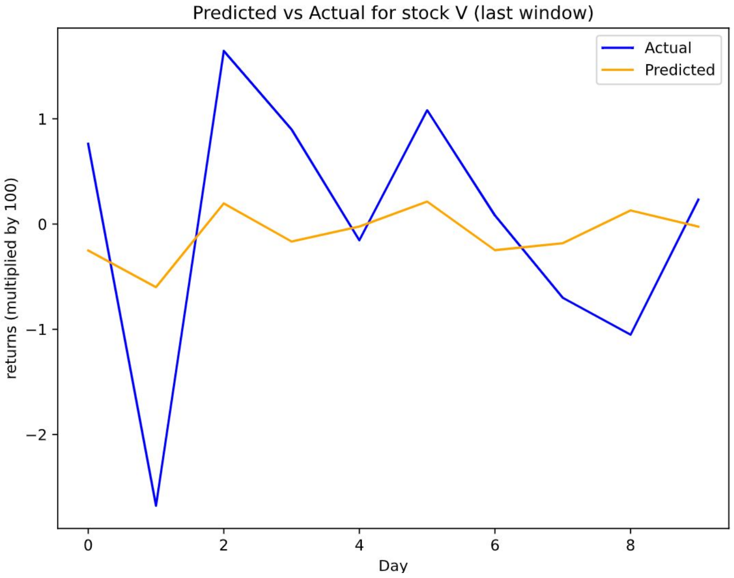
CSCD94 Week 2 Update

Dominik Luszczynski

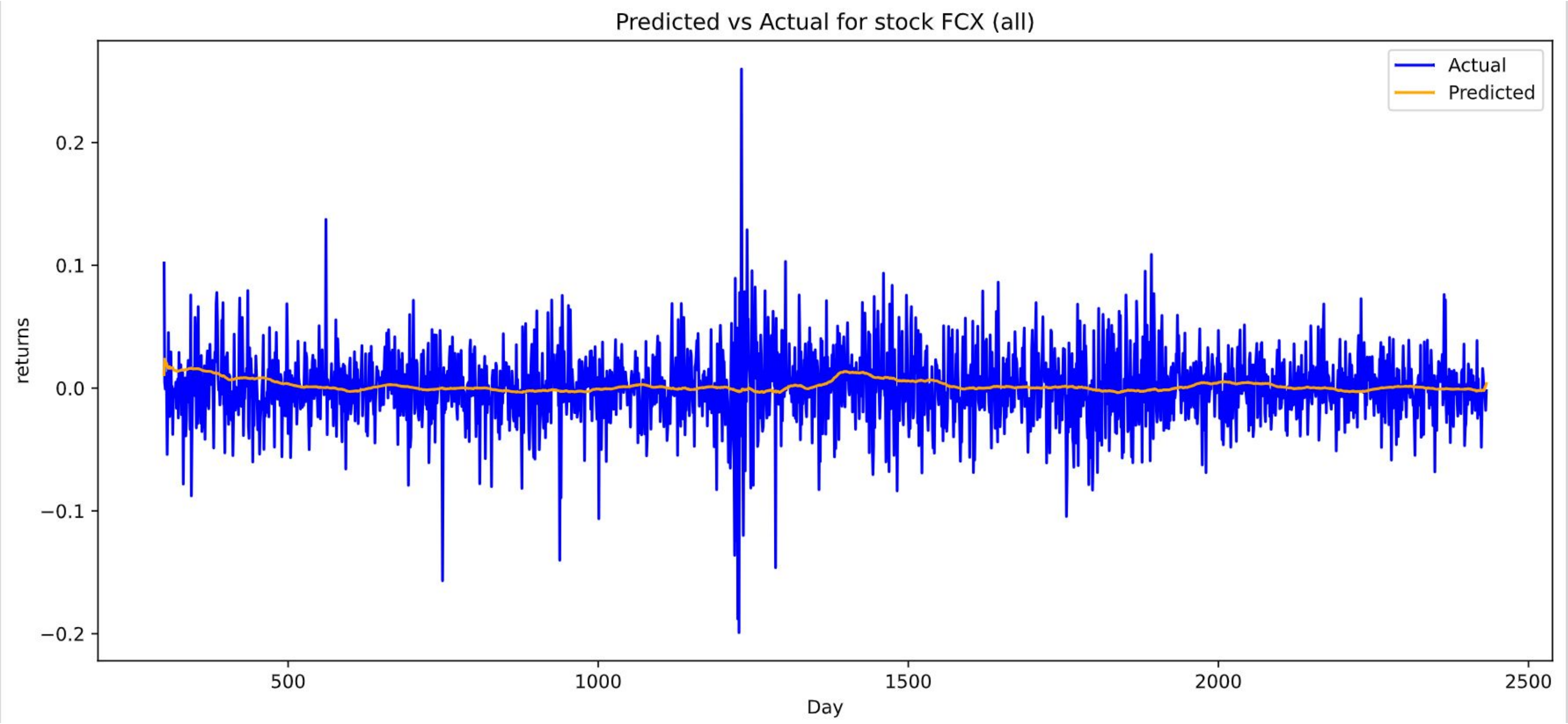
Week 2 Progress

- Training to predict returns





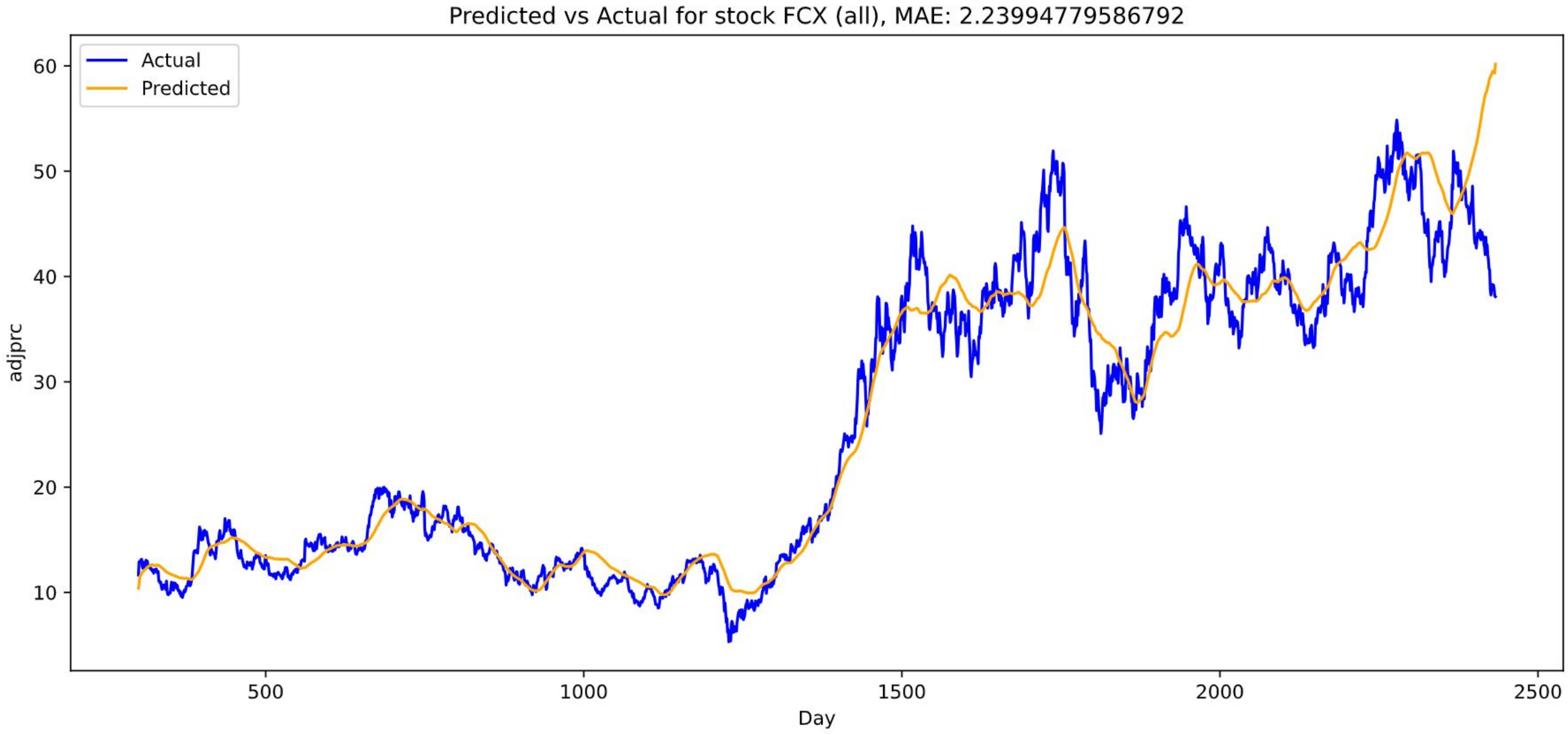
Full Recording - ret



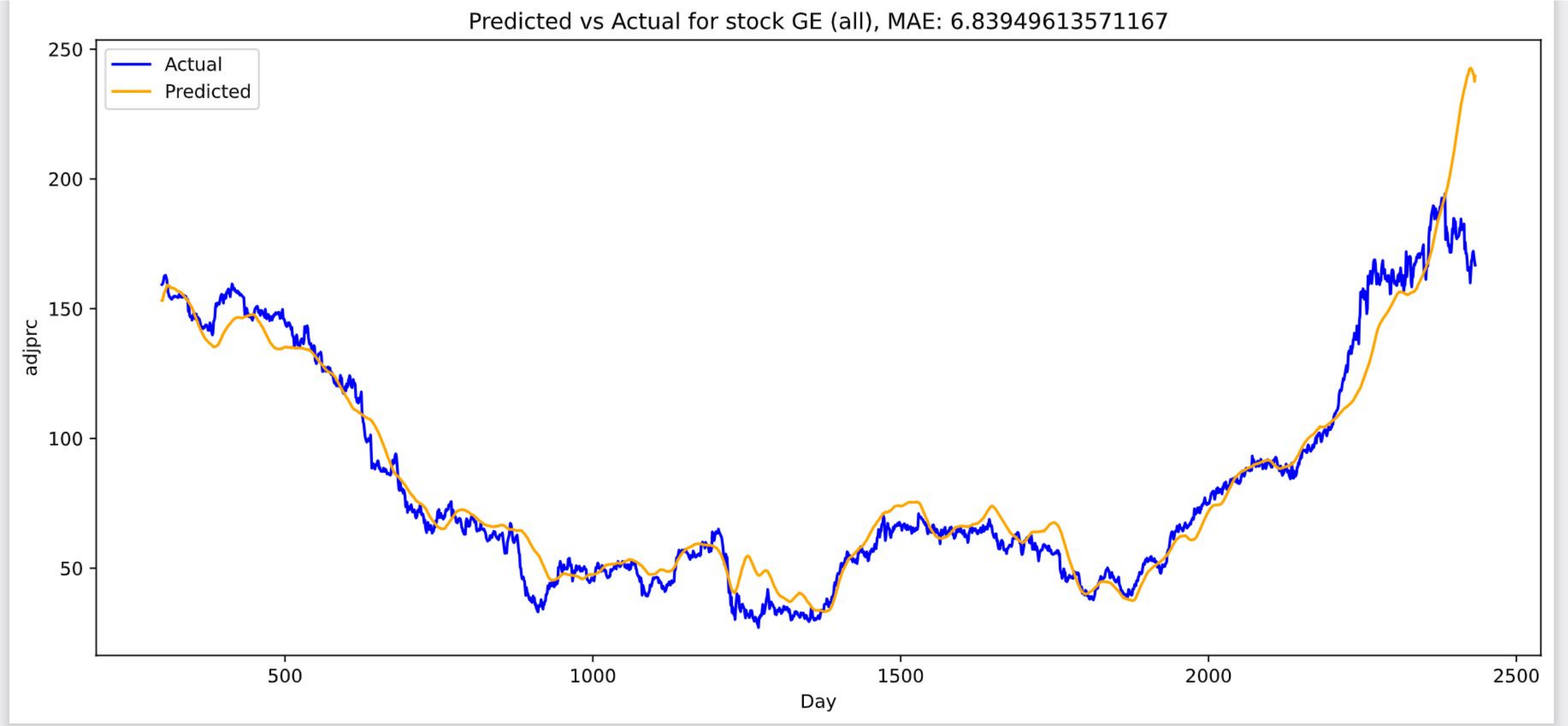
Week 2 Progress

- Focused on adjprc again
  - Experimented with different loss functions (MSE, MASE (Mean Absolute Scaled Error), SMAPE (Symmetric Mean Absolute Percentage Error))
    - Best loss: Quantile loss (QuantileLoss(quantiles=[0.001, 0.01, 0.05, 0.5, 0.95, 0.99, 0.999]))
  - Extended the max encoder length (max lookback time) to 300, and now we predict 50 days in the future.
    - Filtered data to remove companies with less than 1000 days of data
  - Derived new features from adjprc
    - Log returns (log(adjprc_t/adjprc_(t-1)), Exponential Moving Averages, Rate of Change (percent change from the previous 5 days)
    - Removed the rolling means and stdev over 3 days
  - Used a Pytorch Lightning Tuner to find the optimal learning rate.

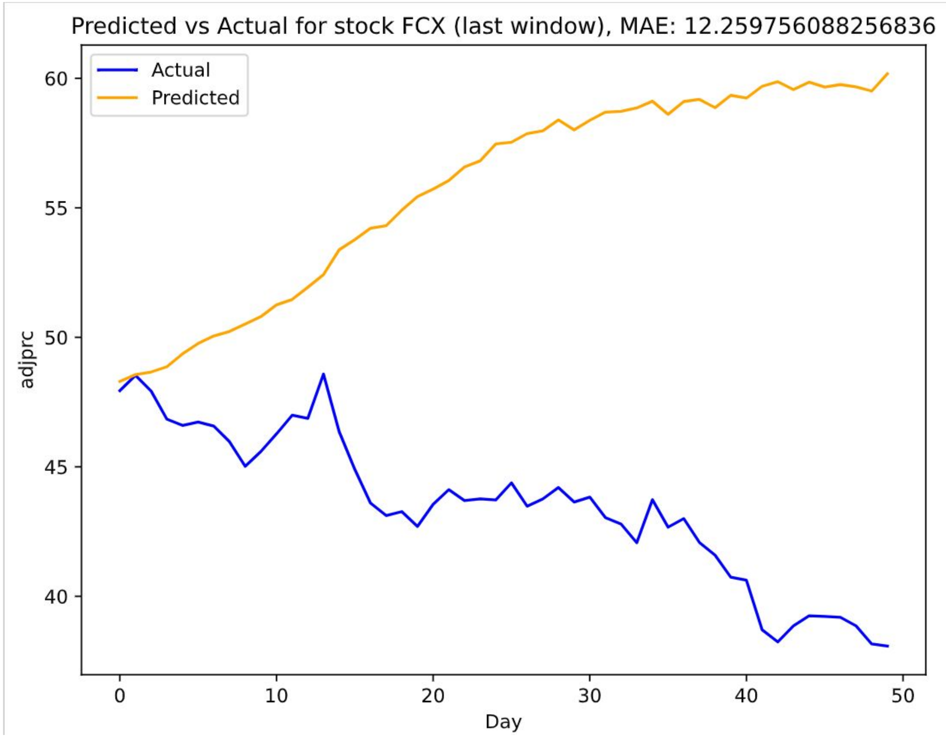
Sample forecasts

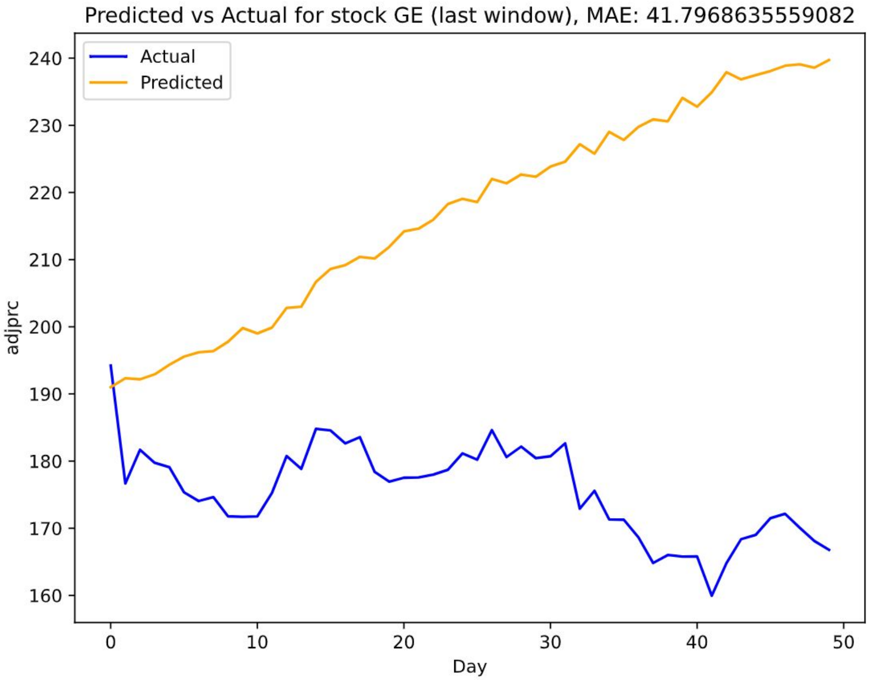


Sample Forecasts



Forecasts on Last Window



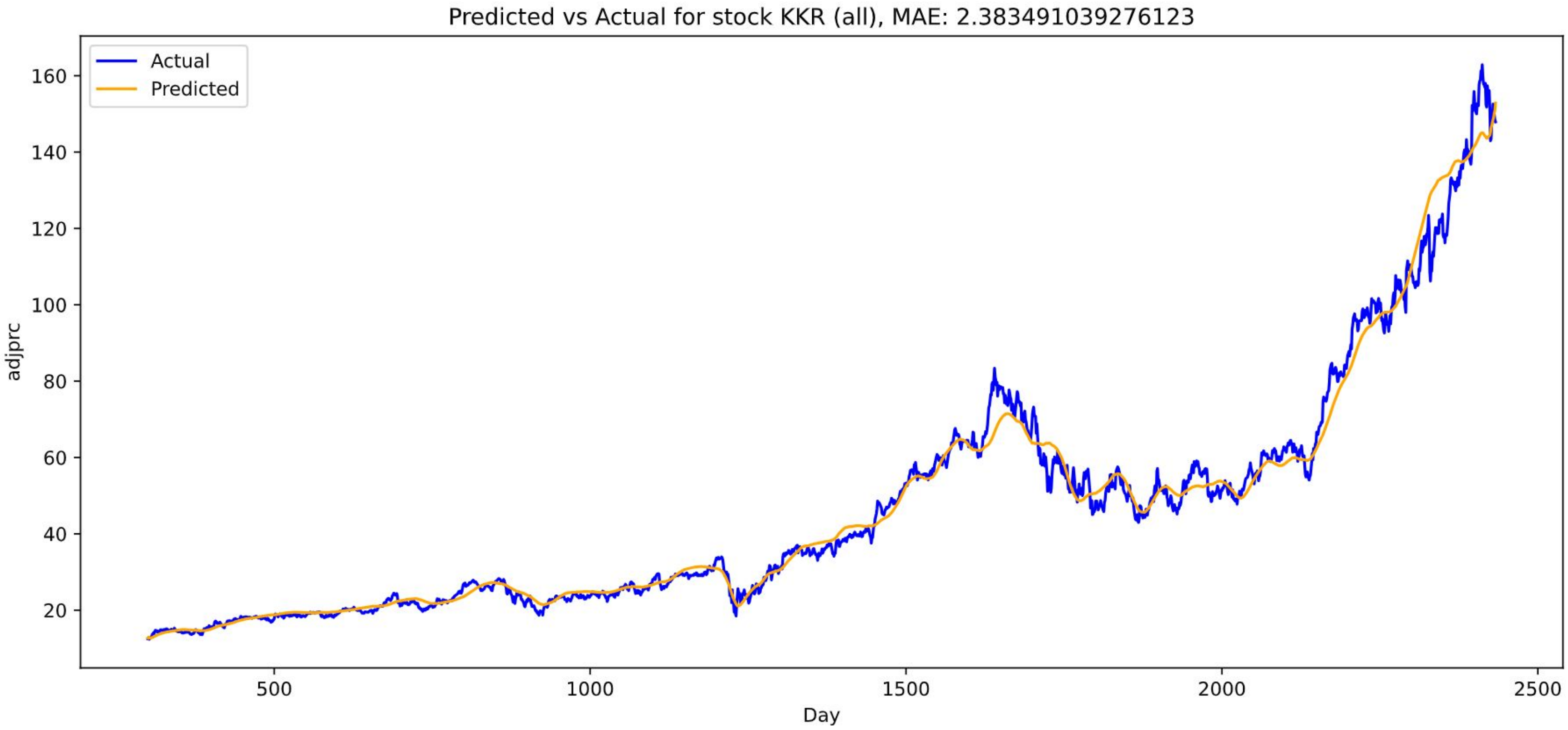


MAE Over Full Recording vs Last Window (Test Set)

- Full Recording: 7.21
- Last Window: 19.27

- Reason:
    - NHiTs performs predictions in rolling windows of the max encoder length (look-back period). As a result, the full recording predictions are averaged over each window allowing for smoother and more accurate forecasts.

Would it be an issue? - I don't think so



Next Steps…

- Now that the forecasting model outputs good results I can:
    - Implement existing gradient methods to get a baseline attack level.
    - Develop a working GAN model.