LEARNING SPATIOTEMPORAL DYNAMICS IN WHOLESALE ENERGY MARKETS WITH DYNAMIC MODE DECOMPOSITION

Clay T. Elmore and Alexander W. Dowling*
Department of Chemical and Biomolecular Engineering
University of Notre Dame, Notre Dame, IN 46556

Abstract Overview

Energy markets are ubiquitous across the globe and offer significant revenue opportunities for emerging technologies including energy storage and flexible energy-intensive industrial consumers. Effectively exploiting these opportunities requires reliable and accuracy short-term forecasts. Both classical time-series analysis and machine learning techniques such as ARIMA, GARCH, neural networks, nearest neighbors, and hybrid learning have been applied to energy price forecasting with mixed success. In contrast, in this work, we quantify the ability of the emerging Dynamic Mode Decomposition (DMD) techniques to learn nonlinear dynamics that drive spatiotemporal price fluctuations and make predictions for optimal control formulations. We find that standard DMD is unable to reliably extract low-rank structures from California whole energy market data. This is due to a well-known limitation of singular value decomposition techniques to efficiently compress invariances in datasets. We then show that Augmented DMD (ADMD) overcomes this limitation, giving fast and accurate day-ahead market energy price forecasts. Preliminary benchmarks show that ADMD requires less tuning, less training data and less computational time than alternative forecasting techniques.

Keywords

Dynamic Mode Decomposition, Singular Value Decomposition, Time-Series Forecasting, Energy Markets

Introduction

The California Independent System Operator (CAISO) runs wholesale electricity markets that service over 39 million California residents, businesses, and industries. The total usage of electric energy in the CAISO in 2015 amounted to the usage of 261,000 GWh for a staggering 40 billion dollars in electric energy sales. With such a vast, complex market comes economic opportunity for energy buyers and sellers. Time variant prices in the CAISO present a basic desire to employ "buy low, sell high" trading strategies in order to either minimize energy costs or maximize profit from energy sales. Using historical market data and signals, we have previous analyzed optimal market participation strategies for several energy conversation and storage technologies including combined heat and power (CHP) generators, concentrated solar thermal (CSP) generations with thermal storage, grid-scale batteries, and energy intensive industry systems. (Dowling et al, 2017, 2018; Sorourifar et al, 2018) For each of these technologies, we consistently found systematic multi-timescale multi-product market participation can boost revenues from between 1.5 to 10 times. These and many other technoeconomic analyses assume perfect information of future energy prices, which means the results are an upper bound on achievable revenue.

Calculating realistic expected revenues requires accurate forecasts of energy prices. Classical time-series analysis models such as (Seasonal) Autoregressive (Integrated) Moving Average, i.e., (S)AR(I)MA, and Generalized Autoregressive Conditional Heteroskedasticity, i.e., GARCH, have been employed with varying success to energy price forecasting. Recently, machine learning techniques as neural networks, nearest neighbors, and hybrid learning have also been considered for energy price forecasting. All of these methods seek to discovery and exploit statistical trends to make forecasts. In this work, we instead explore the ability of emerging Dynamic Mode Decomposition (DMD) techniques to learn nonlinear dynamics that drive spatiotemporal price fluctuations and make predictions for optimal control formulations.

* To whom all correspondence should be addressed. adowling@nd.edu

Dynamic Mode Decomposition

DMD is a so-called equation free dimensionality reduction technique that has been used to extract complex, nonlinear, periodic behavior from computational fluid dynamic simulations and other spatiotemporal datasets. (Kutz et al, 2016) These learned dynamics are typically used for diagnostics, short-term predictions, and/or control. Recently, DMD has been successfully analyze and forecast stock market data, resulting in new economic insights and high-performance trading algorithms. (Cui & Long, 2016; Mann & Kutz, 2016; Hua et al, 2016) In this paper, we explore the ability of DMD to extract spatiotemporal trends and forecast in CAISO day-ahead wholesale electricity markets.

We now highlight the key mathematical properties of DMD in the context of energy market forecasting using the notation from Kutz et al (2016). DMD starts with an input matrix $X_t \in \mathbb{R}^{m \times n}$ where each row corresponds to a spatial feature (nodes/locations in an energy market) and each column is a discrete timestep.

$$X_t = [x_1\, x_2 \dots x_n] \tag{1}$$

This matrix is partitioned into two additional matrices:

$$X = [x_1\, x_2 \dots x_{n-1}] \tag{2}$$
$$X' = [x_2\, x_2 \dots x_n] \tag{3}$$

The overall goal of DMD is to produce a best fit linear operator $A \in \mathbb{R}^{m \times n}$ which advances the energy price dynamics by one hour (timestep):

$$X' \approx A\,X \tag{4}$$

Thus A is a finite dimensional approximation to the infinite dimensional Koopman operator. A key property of the Koopman operator is that it can map any nonlinear state function exactly. The DMD result can be interpreted as the best fit finite dimensional linear system that mimics the infinite dimension Koopman operator. It is important to highlight that DMD is not built from a local linearization. We can compute a low-rank linear operator via singular value decomposition (SVD):

$$X = U\,\Sigma\,V^* \tag{5}$$

Retaining the r most dominant singular values and vectors compresses the data matrix:

$$X \approx U_r\,\Sigma_r\,V_r^* \tag{6}$$

The low-rank operator $\tilde{A}$ is obtained via projection with U_r:

$$\tilde{\boldsymbol{A}} = U_r^*\,\mathrm{A}\,U_r = U_r^*\,X'\,V_r\,\Sigma_r^{-1} \tag{7}$$

This allows us to avoid directly computing $\boldsymbol{A}$, which can be high-dimensional (e.g., millions by millions for CFD simulations). Because $\tilde{A}$ is a linear operator, the price trajectories can be calculated analytically using an eigendecomposition:

$$x(t) = \Phi\ D(t)\,b \tag{8}$$

where $\boldsymbol{\Phi}$ are the so-called DMD modes which arise from the eigenvectors of $\tilde{A}$ and encode spatial patterns. Likewise, $\boldsymbol{D(t)}$ are the temporal modes, arise from the eigenvalues, and encode temporal patterns. Thus, DMD give us an analytic prediction for any continuous time, making it well suited for computationally fast forecasting.

Results

We performed systematic analysis of DMD applied to hourly timesteps (8760 in total) for 6587 nodes in CAISO for calendar year 2015. In particular, we explore the following questions:

1. How does the choice of SVD truncation level r impact reconstruction error?
2. How does the window size m and n impact reconstruction error?
3. Do the best r, m and n recommendations for reconstruction translate to low forecasting error?

Due to space limitations, we only summarize the key results. We find that truncation error is highly sensitive to the choice in r, which can change dramatically for each training dataset. We also find that DMD performs best using all nodes (largest m) and the trainset dataset should include at least as many hours as you wish to forecast. We also find that the optimal truncation level r for reconstruction is near optimal for forecasting. Overall, we found median (mean squared) forecasting error of 16%. This mediocre performance is explained by the fact that DMD often failed to extract low-rank structures in the energy price dataset.

Further investigation revealed the poor performance of "vanilla" DMD is due to the well-known "standing wave" problem; in other words, SVD methods such as DMD and Principle Component Analysis (PCA) are unable to handle invariances in the data. To overcome this limitation, we consider Augmented DMD (ADMD), which simply involves stacking time-shifted copies of the data to form a larger, augmented data matrix. Figure 1 compares DMD and ADMD, highlighting the improved performance. Overall, ADMD results is approximately 10% forecasting error. Surprisingly, we find ADMD performs nearly as well considering either only one node or multiple nodes together.

This suggests ADMD is not apply to identify meaningful spatial patterns to improve forecasting.

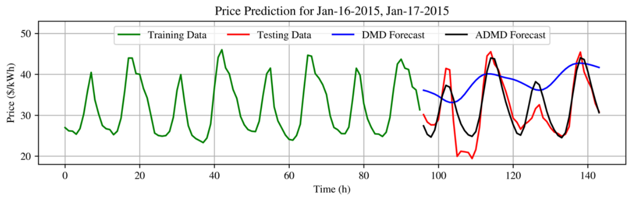

Figure 1. Comparison of DMD (blue) and Augmented DMD (black) forecasts.

We the compare DMD, ADMD, and backcasting forecasting in a rolling horizon optimal control framework for a generalized energy storage system. We consider all 6587 nodes for all hours in calendar year 2015. We find these methods capture 80 to 84% of the maximum achievable revenue (calculated with perfect information). The geography distribution of revenues are shown in Fig. 2.

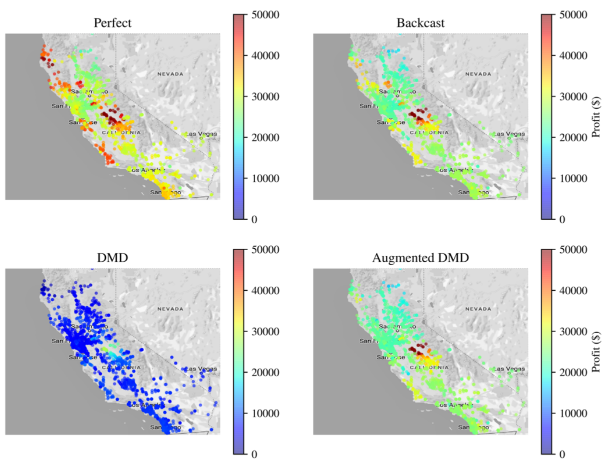

Figure 2. Comparison of market revenue realized from optimal control of an energy storage system with perfect information, backcasting, DMD, and Augmented DMD forecasts.

Conclusions

We present one of the first published applications of DMD to extract spatiotemporal dynamics in wholesale energy markets. We find, somewhat surprisingly, the success of DMD in stock markets does not directly translate to energy markets. Standard DMD is unable to identify low-rank dynamics and gives mediocre electricity price forecasts. Augmented DMD overcomes these limitations and achieves much more accurate forecasts. However, additional spatial information does not significantly improve forecasting accuracy, suggesting spatial dynamics are less importance than temporal dynamics. Although DMD does not immediately offer new economic insights, it is a promising forecasting tool. Augmented DMD is robust to SVD truncation level choice, meaning it requires less tuning that ARIMA and other classically time-series approaches. Likewise, ADMD works well with only few of data, which is orders of magnitude less than machine learning techniques. Finally, DMD and ADMD gives analytic forecasts, which are much faster than classical time-series analysis and machine learning approaches in our comparative benchmarks. As future work, we plan to explore multiresolution DMD and other extensions to further improve forecasting accuracy.

Acknowledgments

C.T.E. thanks Profs. Jeffery Kantor and Patrick Wensing for insightful suggestions and serving on his senior thesis committee.

References

Cui, L.-X., Long, W. (2016). Trading strategies based on dynamic mode decomposition: Tested in Chinese stock market. *Physica A*, 461, 498-508.

Dowling, A. W., Kumar, R., & Zavala, V. M. (2017). A multi-scale optimization framework for electricity market participation. *Applied energy*, 190, 147-164.

Dowling, A. W., & Zavala, V. M. (2018). Economic opportunities for industrial systems from frequency regulation markets. *Computers & Chemical Engineering*, 114, 254-264.

Dowling, A. W., Zheng, T., & Zavala, V. M. (2018). A decomposition algorithm for simultaneous scheduling and control of CSP systems. *AIChE J*., 64 (8), 2480-2417.

Hua, J.-C., Roy, S., McCauley, J.L., Gunaratne. (2016). Using dynamic mode decomposition to extract cyclic behavior in the stock market. *Physica A*, 448, 172-180.

Kutz, J.N., Brunton, S.L., Brunton, B.W., Proctor, J.L. (2016). Dynamic Mode Decomposition: Data-Driven Modeling of Complex Systems, SIAM.

Mann, J., Kutz, J.N. (2016). Dynamic mode decomposition for financial trading strategies. *Quantitative Finance*, 16:11, 1643-1655.

Sorourifar, F., Zavala, V.M, & Dowling, A.W. (2018). Integrated Multiscale Design, Market Participation, and Replacement Strategies for Battery Energy Storage Systems, IEEE Trans. Sustainable Energy, early access.