Learning the Optimal Energy Supply Plan with Online Convex Optimization

Spiros Chadoulos and Iordanis Koutsopoulos

Department of Informatics, Athens University of Economics and Business, Greece
Email: {spiroscha, jordan}@aueb.gr

Abstract—**In this paper, we propose an online learning approach, utilizing the framework of Online Convex Optimization (OCO) to tackle the problem of learning the optimal energy supply plan, in terms of total incurred cost, from the perspective of an electricity retailer that also owns power generators and renewable energy sources. The retailer does not have prior knowledge about key dynamic processes that affect the problem, such as future demand, renewable energy supply, and wholesale market prices. The retailer sequentially learns how to adjust the power generation plan so as to minimize the power generation cost plus the cost of buying additional power from the wholesale market, in case the planned generated amount is not enough to cover the demand. We use Online Mirror Descent (OMD) and Online Gradient Descent (OGD), and we verify that they both achieve sublinear static and dynamic regret, which compare the cumulative cost of each algorithm against that of the optimal offline static and dynamic solution respectively. In particular, dynamic regret appears to be well aligned with the considered setting since it captures the fact that a power generator can dynamically change its power output between consecutive time slots based on only small adjustments and not in an arbitrary fashion, due to ramp constraints. Our model can capture different settings of electricity markets. Simulations with real data verify that OMD precisely learns the optimal dynamic supply policy for small power adjustments between consecutive time slots.**

I. Introduction

Electricity retailers are responsible for selling electricity to end-consumers, charging them fixed or dynamic prices based on their contracts. In some cases, a retailer may also have power generation infrastructure in its possession, as well as renewable energy sources. The retailer may want to utilize these assets to directly fulfill its customers' demand, instead of selling the generated electricity through the wholesale market, given that the market structure and regulation allow it. In case the company resources do not manage to cover the customer demand, additional energy shall be purchased from the wholesale electricity market.

Hence, the retailer has two main alternatives to support the supply. One is through power generation, and the other is through the wholesale market, where various interested entities such as generators and resellers trade electricity. One of the most popular market designs consists of the day-ahead market where these entities offer and request energy for the next day with an hourly granularity, and the real-time spot market where the actual economic dispatch takes place for every 5-minute slot during the current day, taking into account the outcome of the day-ahead market, real-time bids, and all the constraints necessary for the delivery of electricity to end-consumers [1].

In order to enhance energy efficiency, it is crucial for electricity retailers to plan the future supply to their customers, and thus manage their power generation resources appropriately. The power generation at any given moment is affected by several dynamic processes such as consumer power demand, power supply from renewable sources, and energy market prices. These processes are often non-stationary in the sense that their statistics are time-varying, hence difficult to predict and characterize statistically, while they jointly affect the power generation decision, and the choice between the two power supply alternatives. Thus, Online Convex Optimization (OCO) is an ideal framework to model such non-stationary dynamic processes since it constitutes a distribution-free online approach [2], contrary to supervised learning models that can capture the dynamics of non-stationary processes only with the utilization of complex deep learning algorithms which require a significant amount of data, often difficult to collect [3].

Various machine learning models have been widely utilized in smart grid related applications like Demand Response (DR), load forecasting, renewable supply prediction, energy disaggregation, electricity pattern recognition, and load peak detection, as well as load profiling and management [4]. Supervised learning models regarding the prediction of renewable energy generation, energy market prices and consumer electricity demand can be utilized by retailers in order to accumulate prior knowledge, and adjust the power generation accordingly. However, these types of models need quality training data with multi-dimensional feature vectors including exogenous variables like weather conditions. Collecting such data might sometimes be difficult and costly, especially in cases where most of the houses do not have smart meters. Thus, supervised learning is often not the best tool when there are little or inaccurate data available, and when the dynamic processes of the model are non-stationary.

Contrary to supervised learning techniques that require availability of the entire dataset, an online model can make decisions without full knowledge of the dataset, by learning sequentially from data streams arriving in real time. An online learner makes a decision x_t at time t, and then nature chooses

w_t (which is an adversary process). The learner experiences a cost $f(x_t, w_t)$ based on a cost function $f(\cdot,\cdot)$. The learner adjusts its decisions in consecutive time slots based on the incurred cost and hopes to converge to the best decision as time goes to infinity [2]. An online learning model can be directly deployed by a retailer, by conducting a "warm-up" phase on a limited set of historical data and perform accurately from the very first moment in production.

In this paper we use the framework of OCO to address and tackle the problem of learning the optimal (in terms of cost efficiency) energy supply plan in the presence of non-stationary dynamic processes such as consumer demand, renewable supply and wholesale market prices. We consider an electricity retailer that faces the problem of how much power to generate at each time slot, in order to minimize the total incurred cost. There exists a cost of power generation which depends on the amount of power to generate, as well as exogenous factors that affect the generation cost. There also exists a time-varying monetary cost associated with buying the power needed to satisfy the demand from the wholesale market. The solution involves the planning decision of how much power to generate.

A. Related work

OCO has found application in several domains, and among them in the smart grid ecosystem. In [5] the authors propose a demand response framework with real-time pricing using OCO to tackle demand uncertainty for the centralized power load scheduling problem, from the distribution network operator perspective, which is decomposed into separate scheduling sub-problems for each customer.

In [6] the authors propose a distributed energy management system for networked microgrids, and they use the online alternating direction method of multipliers along with OCO. The problem objective is to minimize the microturbine power generation cost along with the costs of transactions with other microgrids. The decision variables include the active and reactive microturbine power generation and demand, the amount of power transacted with other microgrids, and the microgrid voltages.

The authors of [7] utilize OCO to tackle the problem of scheduling power generation in smart microgrids with wind turbines. The objective is to minimize the microgrid power generation cost along with the spot market transaction cost. Specifically, the model assumes a standard electricity generation forecast decided in the previous day through offline demand prediction, and the OCO algorithm learns the cost received from the difference between actual and forecasted demand, and the unpredictability of wind turbines. The authors include demand forecasting in the model, but demand might not always be predicted accurately and deviates from the setting where the learner has no prior knowledge about the future.

In [8] an algorithm called Averaging Fixed Horizon Control (AFHC) is studied for OCO with ramp constraints. This approach assumes that the learner has a perfect lookahead for a specific number of future cost functions, which might not always be realistic because of the effects of non-stationary dynamic processes. The algorithm is validated with simulations for a simplified version of the economic dispatch problem, where the objective is to minimize the grid power generation cost, while satisfying the total demand and the generator ramp constraints. Transactions with the wholesale market and other parameters are not taken into account since the focus of the paper is on ramp constraints applied to OCO.

In [9] the authors tackle the problem of online optimal grid power flow with renewable energy sources, from the perspective of a network operator. They use OCO to minimize the cumulative bus generation and spot market transaction cost, while taking into account transmission network and operational constraints, so that the cost at each bus of the grid is minimized. Differently from [9], we tackle the problem of optimal power generation from the perspective of an electricity retailer with generation capabilities, and study the performance of the algorithms using the dynamic regret formulation which captures the effect of generator ramp constraints.

B. Contributions

The contributions of our work are the following:

- We address and formulate the basic planning problem of retailer cost minimization which entails the decision for energy supply planning, given the instantaneous values of customer power demand, supply from renewable energy sources and wholesale market prices.
- We propose an OCO approach using the Online Mirror Descent (OMD) algorithm to solve the problem. This approach learns jointly all three evolving dynamic processes and adjusts the power generation to minimize the expenses made by a retailer.
- We use dynamic regret to compare our online solution to that of an offline agent that has complete prior information for all the processes and applies an optimal dynamic policy. Hence, this offline agent used for comparison is the equivalent of a perfect forecasting model for all the dynamic processes, which has superior performance than the state-of-the-art in time-series prediction. Furthermore, dynamic regret is closely aligned with our problem, since by definition it captures the fact that the power generators have inherent limitations in different time slots due to ramp constraints.
- We show through simulations on real data that OMD converges much faster than OGD, since it learns the power generation pattern within some days, whereas OGD would need 4 months. This is particularly important since the model operates in the absence of prior information about market prices, renewable source supply, and consumer demand.

The rest of the paper is organised as follows. In section II we formulate the problem of learning the optimal retailer energy supply plan with static and dynamic regret, and describe how we utilized OGD and OMD. In section III we present numerical results and discuss the performance of the algorithms, and in section IV we conclude our work.

II. Learning the optimal energy supply plan

A. Model

We consider an electricity retail company with power generation capabilities that needs to plan the amount of power P_t to generate at each time slot $t = 1, 2, \ldots, T$ (i.e. at the beginning of time slot t), where T is the data time horizon. Each time slot is in the order of several minutes or even several hours, since power generators suffer from ramp constraints and cannot adjust their output in an arbitrary manner.

The retail company owns a set of power generators, with a total power generation cost function $G(P_t)$ for a power generation of P_t, and a set of renewable energy sources with a power generation output of r_t at time t. We also define the cumulative power demand of the retailer's customers as d_t.

If the power from the generators and the renewable sources does not eventually suffice to cover the demand, the retailer has to acquire additional power to satisfy the demand. This is achieved through the wholesale energy transaction market, where the retailer can buy an additional amount of electricity needed at time t at a price q_t ($/MWh). This price is determined based on the structure and regulation of the wholesale market, for example through bidding. We assume that the retail company has no prior knowledge about future electricity demand, renewable power supply and market prices, while trying to decide on the amount of power to be generated in the future. At the beginning of time slot t the generators produce the amount of power P_t that has been planned. When time slot t ends, the actual values of these three unknown processes are revealed and the retailer suffers the total cost $c_t(P_t)$ for time slot t. The arbitrary processes that show the evolution of demand, market prices and renewable supply respectively, are unknown a priori to the retailer, while their non-stationary nature implies that their statistics cannot be easily characterized.

B. Problem formulation

The single time step problem is to minimize the total cost for each time slot t, which consists of:

(a) The power generation cost $G(\cdot)$, which is piece-wise linear and convex, since the differential cost of power generation increases as the demand increases. This means that generating additional units of power becomes more expensive as the demand grows, because the activation of more expensive power plants is needed to cover increasing demand [10].

(b) The monetary cost of buying additional power at price q_t from the wholesale market only if the produced amount P_t plus the generated power from renewable sources r_t, does not cover consumer demand d_t.

The formulation of the problem for each time slot t (i.e. if d_t, r_t, q_t are given), where the total cost is minimized and the decision variable is the amount of power P_t to be generated at time t, is as follows:

$$\begin{aligned} &\min_{P_t} c_t(P_t) = G(P_t) + q_t \max\{0, d_t - r_t - P_t\}, \\ &\forall t \in \{1, \ldots, T\} \end{aligned} \tag{1}$$

and the overall cost function $c_t(\cdot)$ is convex as the sum of two convex functions. The single time slot problem cannot be directly solved since it includes unknown parameters, hence the generation decisions need to be taken sequentially as data is revealed, making the OCO framework an appropriate solution.

Remark. Our model and problem formulation are stated in a general form so that they can be applied on real electricity market models with few modifications depending on each market's regulation. For example, in a day-ahead market where the retailer would have to declare its planned generation for the next 24 hours, our model would remain the same but instead of being scalars, P_t, q_t, d_t, and r_t would be vectors, each with 24 elements, where each element would account for a specific 1-hour interval during the next day. However, the structure of electricity markets can vary significantly based on regulations, hence we study the generic case of hourly supply planning.

C. Online Convex Optimization

According to [11], the OCO framework can be regarded as a game where an online player chooses $x_t \in \mathcal{K}$ at iteration t, where $\mathcal{K}$ is the convex feasible region. Consider $\mathcal{F} : \mathcal{K} \to \mathbb{R}$ to be a family of bounded convex cost functions, from which an adversary assigns a cost function $f_t(\cdot) \in \mathcal{F}$ to the player at each iteration. One metric to evaluate the performance of an online learning algorithm is the so called static regret. The static regret for an online control policy is defined as the difference between the total costs incurred by an online sequential policy and that of the optimal *static* policy that minimizes the total cost over the time horizon.

In our problem, a dynamic policy π is a sequence of power generations $\{P_1, P_2, \ldots, P_T\}$. We are interested in policies for which the regret is sublinear in T and therefore the model will learn the optimal policy in the long run. The static regret of an OCO algorithm for the problem of learning the power generation plan using $c_t(\cdot)$ as the cost function at each time slot becomes:

$$SR(\pi) = \sum_{t=1}^{T} c_t(P_t) - \sum_{t=1}^{T} c_t(P^*), \text{with} \tag{2}$$

$$P^* = \arg\min_{P \in \mathbb{R}^+} \sum_{t=1}^{T} c_t(P) \tag{3}$$

where the first term refers to the cumulative cost of the online learning algorithm, and the second term represents the cumulative cost of the optimal *static* offline policy that has all available information. If the regret $SR(\pi)$ is sublinear in terms of T, this means that the average regret $\frac{SR(\pi)}{T}$ goes to zero as t increases and that in the long run, the dynamic policy performs on average very close to the optimal one [11].

D. Dynamic Regret

A regret definition that fits better the setting of our problem is that of dynamic regret [12], where the learning algorithm is compared against a dynamic offline strategy instead of a static one. The dynamic regret formulation is more appropriate for our setting, due to the nature of the power generation adaptation, where the generated power can be adjusted on an hourly basis in order to react to demand, renewable supply and price patterns. However, most power generators suffer from ramp constraints and do not have the ability to arbitrarily adjust the amount of generated power in a short amount of time. Consequently, the dynamic regret captures the real-life constraints above. In other words we allow limited adjustment of the offline policy between successive time slots, by defining the path length of a sequence/policy $\{P_1, P_2, \ldots, P_T\}$: $\sum_{t=1}^{T-1} d(P_t, P_{t+1})$, where $d(\cdot,\cdot)$ is the difference between two consecutive values. The definition of the dynamic regret of an OCO algorithm for the problem of learning the power generation pattern, using $c_t(\cdot)$ as the cost function at each time slot is:

$$DR(\pi, L) = \sum_{t=1}^{T} c_t(P_t) - \min_{A' \in \mathcal{A}(T,L)} \sum_{t=1}^{T} c_t(A'_t) \tag{4}$$

where $\mathcal{A}(T, L)$ is the set of sequences with T elements and a maximum allowed path length of L. As with static regret, the goal is to minimize dynamic regret and achieve sublinearity, so that the model will learn the optimal dynamic policy in the long run. We also define the average allowed adjustment between consecutive slots as $\frac{L}{T}$, which in our problem captures the average allowable amount of change of the generation between consecutive time slots.

Remark. Parameter L reflects the total amount of change of the optimal decision between successive slots, over a time horizon. Large values of L do not apply any limitations to the amount of power to be generated in two consecutive slots, thus leading to an optimal unconstrained offline dynamic policy. However, smaller L values might be more appropriate due to generation limitations. This is mainly because power generators need time to adjust their power output level, and the bigger the adjustment is, the more time is needed.

Furthermore, a general bound for dynamic regret cannot be derived, but bounds in terms of maximum path length L can be obtained [13]. Hence, if L takes large values, dynamic regret is not guaranteed to achieve sublinear bounds. Consequently, it is only feasible to achieve sublinear dynamic regret when the maximum path length is small [14]. For the aforementioned reasons, the maximum path length L should be assigned small values.

E. Algorithms

Online Gradient Descent (OGD) is one of the first proposed algorithms for online learning and for solving OCO problems [11], hence we use it as a baseline. The OGD algorithm [2], [11] is described below as Algorithm 1:

Algorithm 1 Online Gradient Descent (OGD)

Input: learning rate: $\theta > 0$
initialize: $P_1 = 0$
for $t = 1, 2, \ldots, T$ **do**
  update rule: $P_{t+1} = P_t - \theta \nabla c_t(P_t)$
end for

where $\{c_1(\cdot), c_2(\cdot), ..., c_T(\cdot)\}$ is the sequence of convex cost functions and $\{P_1, P_2, ..., P_T\}$ is the sequence of power generation values the algorithm chooses. It has been shown that, subject to certain conditions, OGD achieves a sublinear static regret bound of $\mathcal{O}(\sqrt{T})$, and a dynamic regret bound of $\mathcal{O}(\sqrt{T}(1 + L))$ [12], [13], [15].

In Online Mirror Descent (OMD) the update is performed in a "dual" space defined by the choice of a regularization function $R(\cdot)$ that affects the update rule [11]. It is described below as Algorithm 2:

Algorithm 2 Online Mirror Descent (OMD)

Input: learning rate: $\theta > 0$
initialize: $P_1 = 0$
for $t = 1, 2, \ldots, T$ **do**
  update rule: $P_{t+1} = \arg\min_{P \in \mathbb{R}^+} \theta \langle \nabla c_t(P_t), P \rangle + D_R(P, P_t)$
end for

In the above, $D_R(\cdot,\cdot)$ is the Bregman Divergence for a convex regularization function $R(\cdot)$ which is defined as:

$$D_R(P, P_t) = R(P) - R(P_t) - \langle \nabla R(P_t), P - P_t \rangle \tag{5}$$

where $\langle \cdot,\cdot \rangle$ stands for the inner product of two vectors. OMD is a generalization of OGD, since for $R(P) = \frac{1}{2}P^2$ the OGD algorithm is derived from OMD. It can be shown that OMD achieves a static regret bound of $\mathcal{O}(\sqrt{T})$ [2]. In addition, OMD achieves sublinear dynamic regret bounds only for small values of maximum path length L, hence if the average allowed adjustment between consecutive slots $\frac{L}{T}$ is small, then OMD is guaranteed to have sublinear dynamic regret [14]. This makes OMD an ideal algorithm for learning the optimal power generation plan.

III. Numerical Results

In this section, we present the results of the simulations conducted with real smart grid data, to demonstrate that the proposed model can quickly learn the optimal power generation planning policy.

A. Dataset and model parameters

The dataset we used originates from the Independent Electricity System Operator (IESO) of Ontario, Canada [16], and it includes 6 months of hourly values regarding demand (MW), intra-day energy prices ($/MWh) and power generation by

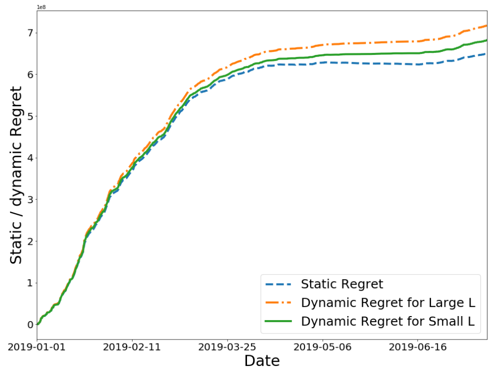

Fig. 1. Comparison between static regret, dynamic regret with $L = 5 \cdot 10^5$ MW and dynamic regret with $L = 2 \cdot 10^6$ MW for OGD over a time horizon of $T = 4,727$ hours.

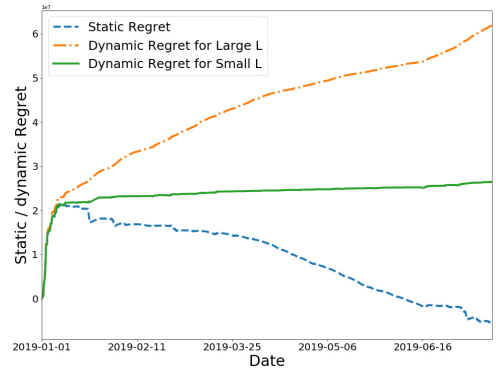

Fig. 2. Comparison between static regret, dynamic regret with $L = 5 \cdot 10^5$ MW and dynamic regret with $L = 2 \cdot 10^6$ MW for OMD over a time horizon of $T = 4,727$ hours.

fuel type. Since real datasets from retail companies are not easy to obtain, we assume that all of Ontario's consumers are customers of a single retailer. The features we use are: wind and solar as the only renewable energy sources (r_t), total power demand (d_t), and intra-day energy transaction prices (q_t). Furthermore, we consider the following generation cost function: $G(P_t) = 0.0001P_t^2$ which captures a realistic generation cost as described in the literature [9], [7], [10]. A different cost function can be utilized for different generators based on their characteristics and placement in the grid.

B. Simulations

The length of a time slot in the dataset is 1 hour, and we ran the simulations for $T = 4,727$ hours corresponding to 6 months of hourly data using a learning rate of $\theta = 0.2$ for OGD and OMD. In our experiments we used absolute difference as the distance measure between consecutive generation values for (4).

First, we compared OGD's performance with the optimal offline static solution, which in our case was 21,017 MW. The static regret for this setting was sublinear as depicted in Fig. 1. However, the retailer can also perform adjustments between consecutive time slots in order to lower the cost. For this reason, OGD's predictions were also compared to the best offline dynamic solution with a maximum path length set to $L = 2 \cdot 10^6$ MW, because for $L \geq 2 \cdot 10^6$ MW there is no restriction regarding power generation changes between consecutive time slots. This happens because $\frac{L}{T} = 423.1$ MW, which means that the retailer can make adjustments of up to 423.1 MW on average from one time slot to the next. The dynamic regret (4) of this setting was found to be sublinear as depicted in Fig. 1. We also tested the case where the maximum path length has a smaller value of $L = 5 \cdot 10^5$ MW, i.e. $\frac{L}{T} = 105.7$ MW. This translates to a more realistic scenario, where the power generators can only apply small generation adjustments between consecutive time slots because of physical limitations in their means of generating power. The dynamic regret was also found to be sublinear in this case, as can be observed in Fig. 1. It is also evident from Fig. 1 that OGD achieved a lower static regret compared to the two dynamic regret cases, which makes sense since the static offline solution is easier to be learned than the dynamic one, especially if the latter involves large values of L.

The same series of simulations were also conducted for the OMD algorithm using the following regularization function for (5): $R(P) = \frac{1}{64}P^2$ so that OMD can perform larger generation adjustments between consecutive time slots compared to OGD, and resemble as much as possible the optimal generation pattern. As we can observe from Fig. 2, the OMD algorithm also achieved a sublinear regret in all three cases. From the results of Fig. 1 and Fig. 2, it is clear that OMD:

(a) Converges much faster than OGD, i.e. in some days, compared to 4 months.
(b) Achieves superior performance compared to the optimal offline static solution, since the static regret goes to zero.
(c) Learns the optimal offline dynamic solution for small L.

Fig. 3 illustrates the cost for the cases of OMD, OGD and the optimal offline dynamic solution with $L = 2 \cdot 10^6$ MW for a specific time window consisting of 8 days. It shows that if the retailer applies OMD, then in most cases a much lower cost compared to OGD is achieved, while in some cases the cost is almost equal to the optimal offline solution for large values of L ($\frac{L}{T} = 423.1$ MW). One interesting outcome derived from Fig. 3 is that during night time, OMD has a slightly higher cost than OGD. This happens most probably due to its ”aggressive” nature, but is insignificant compared to the cost reduction during daytime.

Finally, in Fig. 4 we observe OMD and OGD perform similar to what they would in a real setting, trying to learn the optimal hourly amount of power generation, compared to the offline solutions. OMD performs much better than OGD, and it achieves faster convergence, while closely resembling the power generation pattern of the optimal offline dynamic solution for small values of L. OMD converges to the optimal offline dynamic power generation policy much faster than OGD, i.e. in a matter of some days instead of several months. This translates to significant cost reduction for the retailer that can utilize OMD to achieve a near-optimal performance

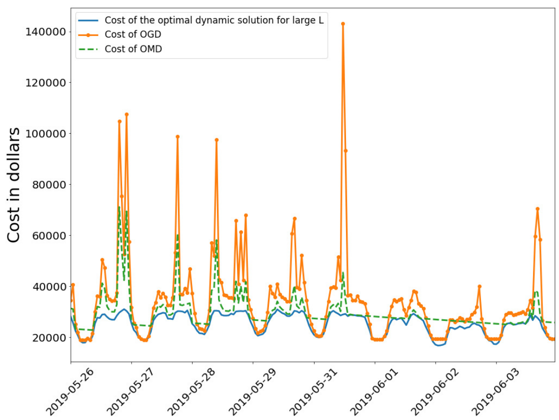

Fig. 3. Cost comparison between OGD, OMD and the optimal dynamic policy for $L = 2 \cdot 10^6$ MW over a time horizon of $T = 4,727$ hours.

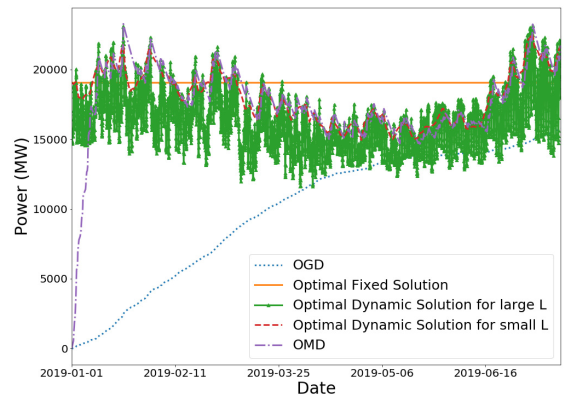

Fig. 4. Online power generation prediction of OGD and OMD compared to optimal fixed generation, optimal dynamic generation for $L = 5 \cdot 10^5$ and optimal dynamic generation for $L = 2 \cdot 10^6$ MW over a time horizon of $T = 4,727$ hours.

in terms of total cost in a short amount of time, and in the absence of prior knowledge about the unknown dynamic processes.

IV. Conclusion

In this paper the framework of OCO is utilized for the problem of learning the optimal energy supply plan for an electricity retailer. Our model learns the power generation policy in an online manner without prior knowledge of customer demand, renewable supply and market prices. In addition, we use dynamic regret to characterize the online learning policy, which is realistic since power generators can perform small output adjustments between consecutive time slots due to ramp constraints. Simulations with real smart grid data showed that OMD can learn the optimal dynamic power generation policy for small values of maximum path length much faster that OGD, while also achieving lower cost for the retail company.

We identify several future research directions. Our model is amenable to adaptation to different market structures, and it would be interesting to investigate how it will change in the presence of both day-ahead and intra-day markets. The cost function can be enhanced by including a cost for realizing DR. If the retailer is also an aggregator, then the cost would amount to the DR incentives provided to consumers, otherwise this cost would be that of remuneration of the aggregator to perform DR on its behalf [17]. Energy trading could also be incorporated in the cost, by allowing the retailer to sell excess energy back to the grid at a possibly time-varying price.

Acknowledgement

This work was supported by project InterConnect (Interoperable Solutions Connecting Smart Homes, Buildings and Grids), which has received funding from European Union's Horizon 2020 Research and Innovation Programme under Grant 857237.

References

[1] P. Cramton, "Electricity market design," *Oxford Review of Economic Policy*, vol. 33, no. 4, pp. 589–612, 2017.

[2] S. Shalev-Shwartz, "Online learning and online convex optimization," *Foundations and Trends in Machine Learning*, vol. 4, no. 2, pp. 107–194, 2012.

[3] A. S. Bedi, A. Koppel, K. Rajawat, and B. M. Sadler, "Nonstationary nonparametric online learning: Balancing dynamic regret and model parsimony," *arXiv preprint arXiv:1909.05442*, 2019.

[4] Y. Wang, Q. Chen, T. Hong, and C. Kang, "Review of smart meter data analytics: Applications, methodologies, and challenges," *IEEE Transactions on Smart Grid*, vol. 10, no. 3, pp. 3125–3148, 2018.

[5] S. Bahrami, Y. C. Chen, and V. W. Wong, "An autonomous demand response algorithm based on online convex optimization," in *2018 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm)*, 2018.

[6] W.-J. Ma, J. Wang, V. Gupta, and C. Chen, "Distributed energy management for networked microgrids using online admm with regret," *IEEE Transactions on Smart Grid*, vol. 9, no. 2, pp. 847–856, 2016.

[7] B. Narayanaswamy, V. K. Garg, and T. Jayram, "Online optimization for the smart (micro) grid," in *Proceedings of the 3rd international conference on future energy systems: where energy, computing and communication meet*, 2012.

[8] M. Badiei, N. Li, and A. Wierman, "Online convex optimization with ramp constraints," in *2015 54th IEEE Conference on Decision and Control (CDC)*. IEEE, 2015, pp. 6730–6736.

[9] S.-J. Kim, G. B. Giannakis, and K. Y. Lee, "Online optimal power flow with renewables," in *2014 48th Asilomar Conference on Signals, Systems and Computers*, 2014.

[10] I. Koutsopoulos and L. Tassiulas, "Optimal control policies for power demand scheduling in the smart grid," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 6, pp. 1049–1060, 2012.

[11] E. Hazan, "Introduction to online convex optimization," *Foundations and Trends in Optimization, vol. 2, no. 3-4, pp. 157-325*, 2015.

[12] M. Zinkevich, "Online convex programming and generalized infinitesimal gradient ascent," in *20th International Conference on Machine Learning*, 2003.

[13] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, "Online optimization: Competing with dynamic comparators," in *Artificial Intelligence and Statistics*, 2015, pp. 398–406.

[14] E. C. Hall and R. M. Willett, "Online convex optimization in dynamic environments," *IEEE Journal of Selected Topics in Signal Processing*, vol. 9, no. 4, pp. 647–662, 2015.

[15] L. Zhang, S. Lu, and Z.-H. Zhou, "Adaptive online learning in dynamic environments," in *Advances in neural information processing systems*, 2018, pp. 1323–1333.

[16] Independent electricity system operator power data. [Online]. Available: http://www.ieso.ca/power-data

[17] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis, "The role of aggregators in smart grid demand response markets," *IEEE Journal on selected areas in communications*, vol. 31, no. 7, pp. 1247–1257, 2013.