Analyzing the Impact of Electricity Price Forecasting on Energy Cost-Aware Scheduling

Diarmuid Grimes, Georgiana Ifrim, Barry O'Sullivan, Helmut Simonis

{diarmuid.grimes, georgiana.ifrim, barry.osullivan, helmut.simonis}@insight-centre.org

INSIGHT Centre for Data Analytics,
University College Cork & University College Dublin, Ireland

Abstract

Energy cost-aware scheduling, i.e., scheduling that adapts to real-time energy price volatility, can save large energy consumers millions of dollars every year in electricity costs. Energy price forecasting coupled with energy price-aware scheduling, is a step towards this goal. In this work, we study cost-aware schedules and the effect of various price forecasting schemes on the end schedule-cost. We show that simply optimizing price forecasts based on classical regression error metrics (e.g., Mean Squared Error), does not work well for scheduling. Price forecasts that do result in significantly better schedules, optimize a combination of metrics, each having a different impact on the end-schedule-cost. For example, both price estimation and price ranking are important for scheduling, but they carry different weight. We consider day-ahead energy price forecasting using the Irish Single Electricity Market as a case-study, and test our price forecasts for two real-world scheduling applications: animal feed manufacturing and home energy management systems. We show that price forecasts that co-optimize price estimation and price ranking, result in significant energy-cost savings. We believe our results are relevant for many real-life scheduling applications that are currently plagued with very large energy bills.

Keywords:
Cost-aware-scheduling, Energy-price-forecasting, Smartgrid, Energy efficiency

1. Introduction

Energy cost-aware schedules are designed around a variable energy price tariff, and in their simplest form, aim to place energy intensive tasks at time slots with low energy price, therefore saving energy costs. In this paper, we analyze applications where the schedule is built day-by-day based on a day-ahead electricity price forecast, and evaluated using the true price. This corresponds to a scenario whereby a forecast price is available 24 hours in advance, but the price paid will be the actual market price. Since for many applications executing the schedule requires significant preparation work, we assume that we cannot continuously reschedule based on the current, actual, price.

In this paper we focus on two main aspects critical for understanding the value of designing cost-aware schedules. The first aspect regards how much cost can be saved by designing a cost-aware schedule based on a day-ahead price forecast. Electricity prices tend to fluctuate a lot during the trading day, for example in Ireland, an increased level of wind-generated energy has resulted in increased price volatility. Peak prices can be as much as 15 times higher than the average price (e.g., in 2010, max price of €766.35/MWh, average price €53.85/MWh), while the cheapest time periods can have negative prices, due to electricity exports (e.g., in 2010, min price of -

€88.12/MWh). We study a series of baselines for assessing the benefit of designing schedules that are aware of the price of energy at each time slot.

The second aspect focuses on the impact of price forecast properties on the end schedule-cost. In particular, we analyze how optimizing a price forecast for different error metrics affects scheduling-costs. Our previous work [1] has shown that even drastically improving price estimation based on regression error metrics (e.g., Mean Squared Error, in short, MSE), is not enough to guarantee schedule-cost savings. Here, we discuss how to optimize price forecasts such that the gains with respect to price forecasting metrics are observed in the end scheduling-cost. We show that both price estimation and ranking of time periods with respect to price (e.g., from cheapest to most expensive) should be co-optimized to reduce scheduling-cost.

In this work we focus on the Irish Electricity Market and analyze a range of price forecasts and their suitability for cost-aware scheduling. Currently most consumers in Ireland pay a fixed tariff (i.e., one electricity meter, one constant price no matter the time of day or season), day-night tariffs (i.e., two meters, cheaper rate between

23h-8h), or time-of-use tariffs[1] (smart meters, price depends on the time of day and season) [2]. Nevertheless, these tariff plans are not really linked to the real-time electricity market price (they are pre-computed and fixed well in advance), and the market price volatility is absorbed by the electricity supplier. Since electricity prices are increasing sharply[2], industrial consumers and electricity suppliers have an added incentive to move to real-time market prices [3]. Real-time pricing offers a huge potential benefit for large electricity consumers that can exploit the price volatility for their operations, but the risks of paying the market price are not well understood. In this paper, we consider the hypothetical scenario where consumers pay the real-time electricity price in the Irish market, and study the effect of building cost-aware schedules based on price forecasts. For a review of electricity price forecasting for other markets please consult [4].

In our previous study [1], we showed that the best price forecast with respect to price estimation (e.g., MSE), has the worst scheduling cost when employed in a manufacturing scheduling application. Here, we analyze the reasons behind this behaviour. We show that forecasts co-optimizing price over-estimation, under-estimation and Spearman rank correlation lead to significant scheduling-cost savings. Furthermore, these three metrics have different impact on the end schedule-cost, and we show how to learn metric importance-weights from a set of sample price forecasts and their scheduling cost.

Most of the forecasts investigated in this paper (other than very naive ones), result in scheduling cost not worse than 4% as compared to using the true price. This is lower than the mark-up required by suppliers for using fixed and time-of-use tariffs. Our best price forecasts further reduce this margin by 0.5-2% for the scheduling applications analyzed. The end metric for measuring success for the price forecasts analyzed, is how much cost-saving a particular price forecast leads to, when plugged into a cost-aware scheduling application.

2. Related Work

Data centres and manufacturing plants are large energy consumers, in critical need to reduce their electricity costs, which account for a big part of their operational cost (electricity accounts for 30-50% of operating costs) [5, 6, 7]. Understanding how to optimize price forecasts to deliver good cost-aware schedules is critical for this goal. A typical data centre can consume as much energy as a medium-sized town[3].

A single pellet press machine in the manufacturing plant used as a case-study in this paper (a real plant, COSYTEC, UK) has an average power rating of about 375kW, and there are 2-6 such machines. Electricity prices are also increasing: the cost for industrial consumers in Europe[4] has increased by almost 50%, from €0.06/kWh in 2005 to €0.094/kWh in 2013. Therefore, a considerable amount of recent work has focused on reducing the overall energy consumption, e.g., by designing energy-volume-aware operation schedules.

The main goal in this line of work is to reduce the overall amount of energy consumed, while minimizing impact on service quality. Examples of such work include [8, 9, 10] that propose technologies such as speed scalable processors, dynamic power-down and power-up mechanisms, new cooling technologies, replacing and consolidating applications to decrease the number of running servers, and multi-core servers. A review of energy-efficient algorithmic developments can be found in [11].

In contrast to energy volume-aware scheduling, energy cost-aware scheduling does not focus on reducing the amount of energy consumed per unit of work, but rather on reducing the cost for doing the work [12]. A number of recent papers focus on reducing both power usage and power cost by taking real-time energy price into consideration, or by considering various price tariffs such as day-night and time-of-use tariffs [13, 14, 15]. This work has mostly shown that allowing the schedule to consider time-variable prices can lead to significant savings as compared to a price-unaware schedule, but has assumed that knowledge of the actual price is available (or other simple baselines), rather than analyzing the effect of price forecasting model properties on reducing schedule-cost.

In our previous work [1] we started to study the problem of building schedules based on a day-ahead price forecast and assessing potential schedule-cost savings. We have pointed out that directly optimizing price forecasts for minimizing price estimation error is misleading, and can lead to increased schedule-cost. We established that different types of error in the price forecast (e.g., over-estimation vs under-estimation) have different effects on the schedule, but we stopped short of recommending ways to account for this (or other metrics) in order to build better schedules. In this paper we present strategies to build price forecasts directly aimed at reducing scheduling-cost. We show that accounting for price estimation error is not enough, and price ranking should also be considered. The price forecasts that co-optimize these two types of error metrics have a better scheduling-cost than previously achieved. In summary, we present an analysis of the Irish electricity market, build scheduling-driven day-ahead elec-

tricity price forecasts for this market, and show the positive effect of employing these forecasts for two scheduling applications (animal feed plant and home energy management system).

3. Electricity Price Forecasting for the Irish Market

The Irish electricity market is an auction-based market, with spot prices being computed every half-hour of a trading day. Under EU initiatives Ireland has an obligation[5] to supply at least 20% of its primary energy consumption[6] from renewable sources by 2020 [16]. The Irish government has set ambitious targets in 2007, for its energy usage: no oil in electricity generation by 2020, 15% of electricity from renewable resources by 2010, and 33-40% by 2020. Wind is the most abundant renewable energy source available in Ireland [16, 17]. However, introducing such renewable energy sources increases volatility in the market, making energy price prediction and cost-efficient planning considerably more challenging.

The methodology for calculating the price in the Irish all-island market, up to the end of 2011, was as follows: every half-hour of the trading day, the Single Electricity Market Operator (SEMO)[7] calculates the System Marginal Price (SMP). The SMP has two components: the *Shadow Price* representing the marginal cost per 1MW of power necessary to meet demand in a particular half-hour trading period, within an unconstrained schedule, e.g., no power transmission congestions; and the *Uplift* component, added to the Shadow Price in order to ensure the generators recover their total costs, e.g., start-up and no-load costs [18].

One day ahead of the trade-day the generators have to submit technical and commercial offer data: incremental price-quantity bids and technical specifications such as generator start-up costs, maximum capacity, etc. Generator units are scheduled in merit order according to their bids to meet the existing load. The SMP is bounded by a Market Price Cap (€1000/MWh) and a Market Price Floor (€-100/MWh), which are set by the regulatory authorities.

Two runs of the Market Scheduling and Pricing Software are particularly relevant for our work. The Ex-Ante (EA) run is carried out one day prior to the trade date which is being scheduled and as such uses entirely forecast wind and load data. A schedule of half-hourly forecasted SMP, shadow price, load and wind generation is produced by the market operator (SEMO) for the coming trade-day.

The Ex-Post Initial (EP2) run is carried out four days after the trade date which is being scheduled, and utilizes actual wind and load data. The system marginal prices produced in the EP2 run are used for weekly invoicing and the SMP determined in the EP2 run for a given half hour trading period is the price applicable to both generators and suppliers active in such a trading period.

In this section we discuss day-ahead price forecasting models for the Irish electricity market, in particular, the factors influencing the price and data collection.

3.1. Data Collection and Analysis

We begin our study of the Irish electricity market by analyzing the price and load (i.e., demand) profile from January 2009 to June 2011. Figure 1 shows the actual[8] half-hourly price (top frame, black) and demand (bottom frame, gray). We note that the load profile is fairly similar over time, showing clear periodicity, with higher load in winter. The price is much more volatile, with high variations during both cold and warm months. Furthermore, price volatility and magnitude increased considerably from 2009 to 2011.

Table 1 shows statistics about the price in this period. We observe an increasing median and average price, as well as increased price volatility (standard deviation) over time. This could be explained by increasing fuel prices and the ramp-up of wind-generated power, as well as other factors such as a higher percentage of unscheduled generator outages in 2011.

Table 1: Statistics of the Irish SMP (€/MWh) for 2009 to mid-2011.

Year	Min	Median	Mean	Stdev	Max
2009	4.12	38.47	43.53	24.48	580.53
2010	-88.12	46.40	53.85	35.49	766.35
2011	0	54.45	63.18	35.79	649.48

Figure 2 offers a closer look at the price versus load pattern in 2011, for the first week of the year and the week with the maximum price up to mid-2011. The daily periodicity of the load, and the volatility of the price can be seen more clearly here.

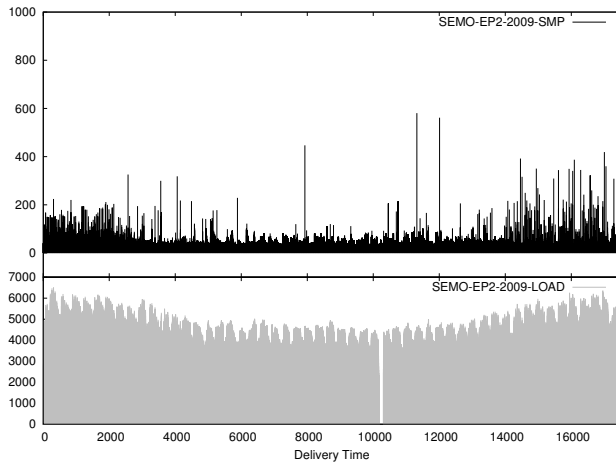## 3.2. Methodology for Building Price Forecasts

SEMO provides a web interface for public access to the historical Single Market Price (SMP), Shadow Price and load, back to January 2008. In November 2009, SEMO started providing day-ahead half-hourly actual (historical) values and forecasts for load, SMP, Shadow Price and
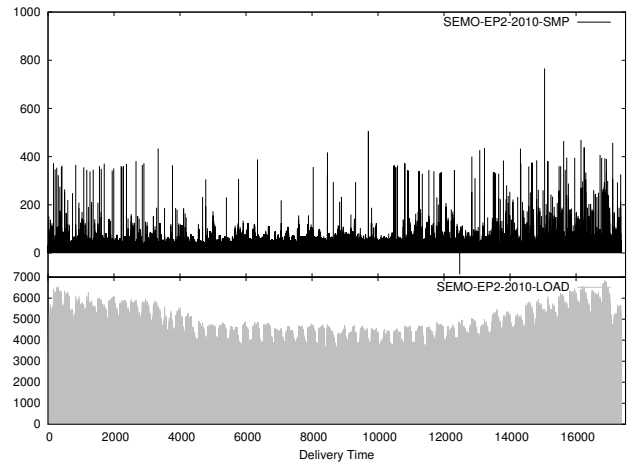
---

[5]http://www.eirgrid.com/media/
EirGridAnnualRenewableReport2013.pdf

[6]http://europa.eu/rapid/press-release_IP-08-80_en.htm

[7]http://www.sem-o.com

[8]http://www.sem-o.com/marketdata/
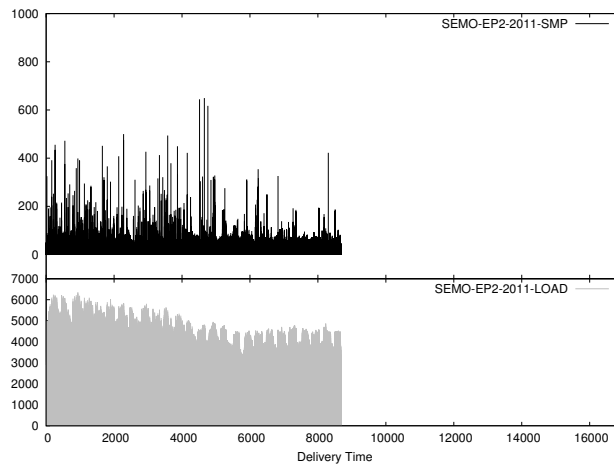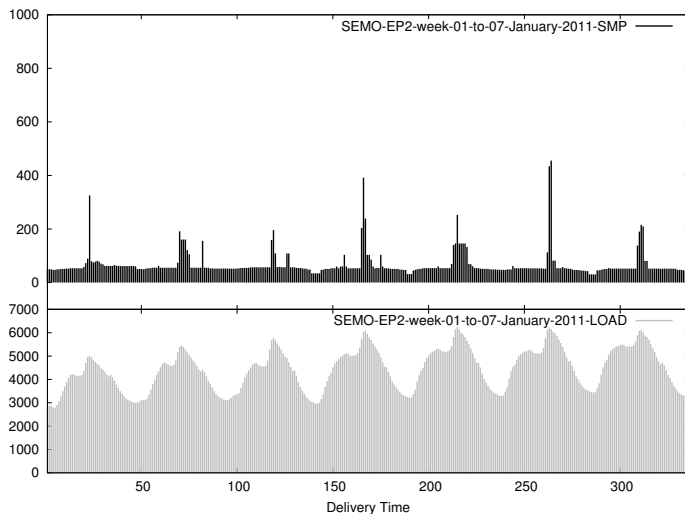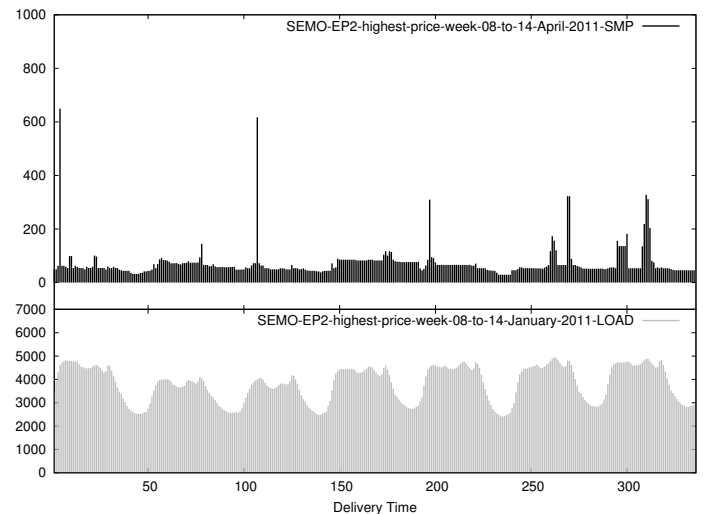
Figure 1: Half-hourly price (top-black) and demand (bottom-gray) from January-2009 to June-2011. The X axis represents the delivery time (every half-hour of a trading day). The Y axis for the top-black plot represents the SMP in €/MWh and for the bottom-gray plot, the Load in MWh. Missing data in the Load plot is missing from the SEMO tables.

Figure 2: Half-hourly price (top-black) and demand (bottom-gray) for two weeks in 2011.

wind-supply. Due to the later availability of more complete data (e.g., wind forecasts beginning 2010), we use data starting January-2010 to June-2011 for training and evaluating price forecasting models. We use the year 2010 for training our price forecasting models. Three months of 2011 (January, March and May) are used for validation (i.e., calibrating model parameters), and another three months of 2011 (February, April and June) for testing. The latter constitutes 88 test days.

The choice of training, validation and test is made in order to respect the time dependency in which we train on historical data of the past and forecast prices into the future. Months from different seasons were chosen for the test data, to avoid the bias of forecasting prices for summer or winter months exclusively (since prices in the winter tend to be more volatile than prices in the summer). Due attention was paid to the fact that in the Irish market, the actual values for SMP, load, etc., are made available only four days after the tradeday, thus for prediction we can only use historical data with a gap of four days back into the past from the current day. All evaluations of our models and comparisons to the SEMO price forecasts are done on the three test months that are not otherwise used in any way during training or validation.

We began our analysis with simple models, using only the historical SMP (i.e., the actual price values in the past) for predicting the SMP of the next tradeday. We then gradually introduced information about the Shadow Price, load and wind-generation and studied the effect of each of these new variables on the prediction quality. In order to estimate the expected supply, we have extracted information on the daily generator bids and planned generator outages available from SEMO and Eirgrid[9]. Information about demand and supply is important since price peaks are typically an effect of the mismatch between load and supply. Our data integrity checks revealed missing days/hours in the original SEMO data. We have filled in the missing half-hours by taking the data of the closest half-hour.

The data collected was available in different granularity (e.g., wind-supply obtained from Eirgrid was sampled every 15 mins) and units (Eirgrid wind-supply was in MW vs SEMO in MWh); we aggregated it to half-hourly granularity and converted the data to the same unit (MWh). Since we rely on SEMO forecasts for building our models, we estimated the SEMO forecast quality for each of the variables involved: load, wind, shadow price, and SMP. Our evaluation of SEMO's forecasts showed that the load forecast is most reliable, followed by shadow price, SMP and wind (we compared forecasts using the normalized MSE). In our models we use local forecast-quality-estimates as

additional features. Table 2 summarizes the features we used for building our day-ahead price forecasting models.

Table 2: Features used for building price forecasts.

Single Market Price (SMP)
Shadow Price (SHP)
Load (i.e., Demand)
Supply: Wind, Generator Bids, Planned Outages
Estimated quality of SEMO forecasts

3.3. Price Forecasting Models

We use machine learning to build price forecasting models and present here two of the models that worked best in our experiments (based on our prior work [1]).

The first model, *FM1*, aims to directly predict the price using the available features. This approach follows the classical line of price prediction in international electricity markets where the main idea is to use historical data, e.g., past prices, load, and supply, for training a price model for the next trade-day. From the time series data, we extract regression vectors as follows. For each half-hour of a tradeday, we take the actual SMP as a prediction target and use historical data for the same half-hour in the past as features. For example, if the SMP on 1st of January 2010, 7 AM, is €31.04/MWh, we take this as a learning target and the SMP at 7AM of D past days as features (in this case the most recent historical data is from 27 December 2009, due to the 4 days gap). The number of historical days D is a parameter of the model and is calibrated using the validation dataset.

Since we also have access to day-ahead forecasts of SMP, shadow price, load, wind and other-supply, we study those as additional features. We have additionally investigated weather and calendar information (e.g., weekend, bank or school holidays) as features, but these have not increased the quality of the model. This may happen since calendar and weather data is already factored into the load and wind-supply forecast, thus it does not add new information to the model. We compute estimates of the weekly and daily available supply from the information on outages and generator bids publicly available[10].

From Eirgrid, we use information on planned outages to estimate the weekly maximum available supply based on the maximum capacity of the available generating plants. From SEMO day-ahead generator bids, we extract features on daily available supply. For example, we set thresholds on the maximum bid price (e.g., €40) in order to obtain estimates of expected cheap supply. The

---

maximum price thresholds of bids are set at €40, €50 and €60, based on the bids and empirical SMP distribution on the validation set. Once the data required for preparing features is processed, we scale all features and use an SVM with an RBF (i.e., Gaussian) kernel for learning. The RBF-SVM works by mapping examples into a higher dimensional space and computing a model in that space, and is known to deliver very good results in many applications[11].

We use the epsilon-SVR implementation of the LIB-SVM package [19] (widely accepted as state-of-the-art for SVM implementations). It has 3 main parameters which can be tuned to improve performance: C that controls the amount of regularization of the loss function, γ that controls the kernel width, and ϵ that sets the tolerance of the termination criterion. We tune these parameters on the validation set, using grid search. All the data collected and the scripts for reproducing our forecasting experiments are online[12].

The second model, *FM2*, aims to predict the difference from the average price, rather than the price directly. It builds on the following observation: the actual historical SMP (i.e., *past prices*) are a good indicator for the *average electricity price* at a given half-hour, but do not capture the particular behavior of a given day in terms of the magnitude of the SMP peaks and valleys. Due to the particular features of the next tradeday (e.g., strong wind, lower load, enough cheap supply), the SMP may diverge from its average value (e.g., below/above the average).

We can estimate the characteristics of the next tradeday using the publicly available forecasts. We then compute the SMP as a sum of a locally computed average-SMP (e.g., over the last 7 days) and a *learned SMP-difference* from the average, estimated from the training set, capturing whether the SMP is going up or down with respect to the average. For example, for forecasting the SMP on 1st of January 2010, 7AM (equal to €31.04), we use the local average-SMP (equal to €29.57) over the most recent seven days, as a first component. The second component is the learned-difference between the actual SMP and the average-SMP. As learning features, we use the difference between the forecasted tradeday characteristics (load, wind-supply, shadow price) from their local averages. Intuitively, lower than average load and higher than average wind, should trigger a decrease in price, thus a negative difference of SMP from the average. A regression model can estimate the SMP-difference (from its average) from the differences of its features.

As shown in Table 3 (and further detailed in [1]), when optimizing these models for regression metrics such as the MSE, they give a 28% improvement over the SEMO forecast. Nevertheless, when plugged into cost-aware schedul-

Table 3: Mean squared error (MSE) for SEMO forecast versus learned forecasts FM1 and FM2.

Method	MSE
SEMO	1086.25
FM1	821.01
FM2	**781.72**

ing applications, the improvement in price estimation is not reflected in the schedule-cost. We use these two models as the base for our study of error metrics that are more appropriate for scheduling-driven price forecasts (as detailed in Section 5).

4. Scheduling Applications

In this work we analyse cost-aware scheduling models for two applications: feedmill manufacturing and home energy management systems. The following subsections present the corresponding models. For simplicity the models below describe cost in each time period in terms of power usage rather than energy usage. However the average power rating and fixed durations enable us to compute energy usage in each time period, and this was implemented for the applications.

4.1. Feedmill Manufacturing: Energy Cost-Aware Scheduling Model

In order to test the quality of the price forecasts on a realistic scheduling problem, we adapted a variant of the feedmill scheduling problem from [20]. The schedule is generated from orders on the current day for delivery in the next morning. Tasks i are scheduled on four disjunctive press lines with their allocated machine m_i, duration d_i, power requirement p_i and due date e_i, satisfying an overall power limit l_t at each time point t.

We express the problem as a Mixed Integer Linear Programming (MILP) minimization problem following [21], where the main decision variables are 0/1 integers x_{it} indicating whether task i starts at time t, and non-negative, continuous variables pr_t denoting the power use at time t. For this evaluation we choose the MILP formulation over a more conventional constraint programming model, as it allows us to find the optimal solutions for the core problem.

The objective function is based on the predicted price v_t, while the evaluation of the quality uses the actual price a_t. This corresponds to a scenario whereby a forecast price is available 24 hours in advance, but the price paid will be the actual market price. As executing the schedule requires significant preparation work, we cannot continuously reschedule based on the current, actual price. Therefore, the energy cost of an optimal schedule is computed

---

[11]http://svms.org/tutorials/

[12]http://4c.ucc.ie/~gifrim/Irish-electricity-market

as:

$$\text{cost} = \sum_t pr_t^* a_t \tag{1}$$

where pr_t^* is the power use at time t of the optimal solution to the following MILP problem:

$$\min \sum_t pr_t v_t \tag{2}$$

subject to:

$$\forall_i : \qquad \sum_t x_{it} = 1 \tag{3}$$

$$\forall_t : \quad \sum_i \sum_{t-d_i+1 \leq t' \leq t} p_i x_{it'} = pr_t \leq l_t \tag{4}$$

$$\forall_m \forall_t : \quad \sum_{i|m_i=m} \sum_{t-d_i+1 \leq t' \leq t} x_{it'} \leq 1 \tag{5}$$

$$\forall_i \forall_{t|t+d_i>e_i} : \qquad x_{it} = 0 \tag{6}$$

Constraints (3-6) can be described as follows. Each task must be assigned a start time. The overall capacity cannot be exceeded in any time period. No two tasks running on the same machine can overlap in their processing time. The final constraint (6) states that no task can start after its due date.

4.2. Home-Energy Management Systems: Energy-Cost-Aware Scheduling Model

Energy consumption in the residential sector accounts for a significant proportion of national usage. For example, in Ireland this sector accounted for 32% of total electricity usage in 2008, and over 44% of thermal energy usage [22]. The electricity demand in this sector is expected to increase significantly in the coming years due to the influx of plug-in electric vehicles (PEVs) [23]. A number of methods have been proposed to improve the energy efficiency of homes such as smart metering and better building insulation. The pricing strategy used by the utility can also affect user behavior. In particular, time-variable tariffs such as Time-of-Use Pricing and Real Time Pricing encourage load shifting on the part of the user from peak to off-peak times. However, it is often impractical for the user to manage their energy usage in reaction to constantly changing price information.

A Home Energy Management System (*HEMS*), illustrated in Fig. 3, is an energy efficiency tool which can be used to address this issue, automating energy usage of certain home appliances with respect to time-varying prices. In this work, we model the HEMS problem as follows: The objectives are to minimize the electricity cost of the home over a fixed horizon (e.g., 24 hours), while maximizing user comfort. The latter involves user preferences with regard to the ambient temperature and completion times of certain activities (e.g., dishwasher, washing machine, etc.).

There are three primary components for scheduling in the HEMS problem that we discuss in detail in the following subsections.
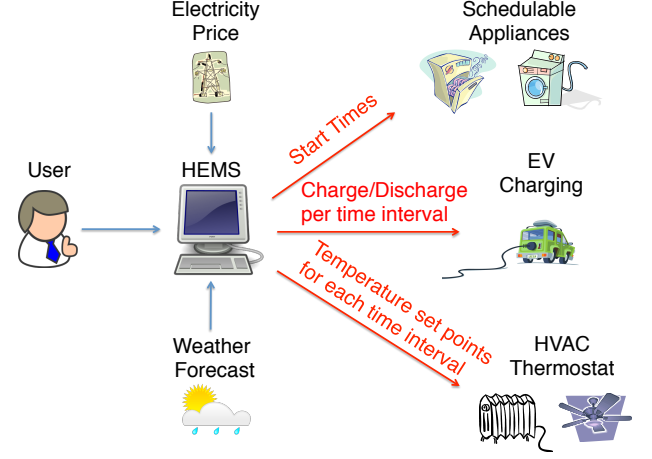

Figure 3: HEMS Overview

4.2.1. Deferrable Activities

A deferrable activity is an activity that can be scheduled to run at some point in the fixed horizon. Let A be the set of activities to be run, and A^r be the set of activities on resource r. For simplicity we will consider all activities i on a resource to have the same fixed duration, d_i, and will use the resource power rating, p_i, as a constant across the duration. Furthermore we assume the activities are non-preemptive, i.e., once the activity has started it must be run to completion.

The user supplies the earliest, latest and preferred starting time (s_i^e, s_i^l, s_i^{best} respectively) for each requested activity i. The start time, s_i, of activity i is constrained by the user preferences (constraint (7)). A resource can only handle one activity at a time (8). The total energy usage of the deferrable appliances, E_t^{app}, in each time period t is given by (9).

$$s_i^e \leq s_i \leq s_i^l \qquad \forall i \in A^r \tag{7}$$

$$(s_i + d_i \leq s_j) \vee \qquad \forall i,j \in A^r, i \neq j$$
$$(s_j + d_j \leq s_i) \tag{8}$$

$$E_t^{app} = \left(\sum_i p_i\right) \qquad \forall i \text{ active in } t \tag{9}$$

We define the user comfort relative to a deferrable appliance to be the difference between the preferred and actual starting time. We further differentiate between the early and late completion of an activity as the user may have a preference which is reflected in the appliance cost f_i:

$$f_i = \begin{cases} \delta_A^- (s_i^{best} - s_i) & if \ s_i \leq s_i^{best} \\ \delta_A^+ (s_i - s_i^{best}) & if \ s_i > s_i^{best} \end{cases} \tag{10}$$

where δ_A^-/δ_A^+ is a (user-inputted) cost coefficient for the early/late completion of an activity, reflecting the strength of the preference the user has for both.

4.2.2. Electric Vehicle

A widespread introduction of EVs would see an increase in demand during off peak hours as users charge their vehicles overnight. One benefit of EVs is their capacity to operate as energy storage devices in a vehicle-to-home (V2H) setting. This can result in a flattening of the demand curve as batteries are charged during off-peak hours and battery power is supplied to the home during peak times [24].

The user inputs for the EV battery charging of each EV i for each charge request are as follows:

- Arrival time, t_i^{arr}, and estimated state-of-charge (SOC) upon arrival, l_i^{arr} (in kWh).

- Departure time, t_i^{dep}, and required SOC for departure, l_i^{dep}.

The charge state of the EV battery of each EV i is subject to the following constraints for all time periods $t \in (t_i^{arr} + 1, \ldots, t_i^{dep})$, where $p_{ev,i,t}^+/p_{ev,i,t}^-$ is the power rating at which the battery of EV i is charged/discharged in time period t. These power ratings are set to 0 for all other time periods. For each EV i, constraint (11) ensures that the battery cannot be charged and discharged in the same time period, constraints (14) and (15) enforce that the battery should not be discharged beyond some minimum level, l_i^{min}; while constraint (12) ensures that the battery is charged to the required amount at the departure time. The battery level, $l_{i,t}$, is calculated using Equation (16), where z_i^c is the charge efficiency of the battery. Finally the net energy usage of the electric vehicle, E_t^{ev}, in time period t is given by Equation (17), where z_i^d is the discharge efficiency of the EV battery i.

$$(p_{ev,i,t}^+ == 0) \quad \vee \quad (p_{ev,i,t}^- == 0) \tag{11}$$

$$0 \leq \quad l_{i,t} \quad \leq l_i^{max} \tag{12}$$

$$l_{i,t_i^{dep}} \quad \geq \quad l_i^{dep} \tag{13}$$

$$(l_{i,t-1} \leq l_i^{min}) \quad => \quad p_{ev,i,t}^- = 0 \tag{14}$$

$$(l_{i,t-1} \geq l_i^{min}) \quad => \quad l_{i,t} \geq l_i^{min} \tag{15}$$

$$l_{i,t} \quad = \quad l_{i,t-1} + \\ (z_i^c(p_{ev,i,t}^+) - (p_{ev,i,t}^-)) \tag{16}$$

$$E_t^{ev} \quad = \quad \sum_i (p_{ev,i,t}^+ - z_i^d p_{ev,i,t}^-) \tag{17}$$

4.2.3. Heating, Ventilation and Air Conditioning (HVAC)

The thermal model we use is a simplified estimate based on the "Irish official method for calculating and rating energy performance of dwellings" [25]. A house-specific learning component will be added and will be used to refine the coefficients model once sufficient data has been gathered. We make the following assumptions: thermal losses between internal zones are negligible; and the thermal mass of internal elements are negligible compared to the thermal mass of the fabric surfaces, e.g., walls, roof, floor. We describe the model for a single zone below, which can easily be extended to handle multiple zones.

The HVAC system cannot operate in both heating and cooling modes in the same time period (18), where $p_{h,j}^+/p_{h,j}^-$ is the power used to heat/cool the room. The indoor temperature (T_t^{in}) at time point t is calculated using Equation (19), where w is a conversion factor to calculate average power rating (in Watts) from the energy used (in Joules) in the time period, and H^{cap} is the heat capacity of the room.

The heat loss (H_t^{loss}) due to the difference between internal and external temperatures at time point t is given in Equation (20). This comprises: the losses across the external surfaces S, where ΔT_t^{ext} is the temperature difference between indoor and outdoor temperatures, A_j is the area of external surface j and u_j is the *u-value* (rate of heat loss through a material) of the external surface j; and losses due to ventilation, where V_{air} is the volume of air in the zone, *vf* is the ventilation factor and *ac* is the rate of air change. The three latter factors are considered to be constant for all time periods.

We consider heat transfer over multiple time periods to account for systems such as the underfloor heating. The net heat gain (H_t^{gen}) from the HVAC system in time period i is the sum of the power used to heat/cool the room ($p_{h,j}^+/p_{h,j}^-$) given in Equation (21), where λ_k is the proportion of heating/cooling from period $(t - k)$ which is still dissipating into the room, and h^{eff}/c^{eff} is the efficiency of the heating/cooling.

$$(p_{h,t}^+ == 0) \quad \vee \quad (p_{h,t}^- == 0) \tag{18}$$

$$T_t^{in} \quad = \quad T_{t-1}^{in} + w(H_{t-1}^{gen} - H_{t-1}^{loss})/H^{cap} \tag{19}$$

$$H_t^{loss} \quad = \quad (\Delta T_t^{ext} \sum_{j \in S}(A_j u_j)) + \\ (\Delta T_t^{ext} * V_{air} * \text{vf} * \text{ac}) \tag{20}$$

$$H_t^{gen} \quad = \quad \sum_{j=t-k}^{t} \lambda_k((h^{eff} * p_{h,j}^+) - \\ (c^{eff} * p_{h,j}^-)) \tag{21}$$

$$E_t^{hvac} \quad = \quad (p_{h,t}^+ + p_{h,t}^-) \tag{22}$$

The user supplies their preferred temperature values for certain time intervals. More formally, let R be the set of temperature requests; tr_j^{st}, tr_j^{end} be the start and end time periods respectively for request j; and T_j^{min}, T_j^{max}, T_j^{best} the minimum, maximum and preferred temperatures for request j. The indoor temperature must be between the

minimum and maximum temperature for each temperature request interval.

We define the *thermal user comfort* to be the difference between the preferred and actual temperature values. Similarly to the deferrable appliances, we further differentiate between lower/higher temperatures than preferred as the user may have a stronger dislike for one over the other which is reflected in the penalty g_t:

$$g_t = \begin{cases} \delta_T^-(T_j^{best} - T_t^{in}) & if \ T_t^{in} \leq T_j^{best} \\ \delta_T^+(T_t^{in} - T_j^{best}) & if \ T_t^{in} > T_j^{best} \end{cases}$$
$$\forall j \in R, \forall t \in (tr_j^{st}, ..., tr_j^{end}) \quad (23)$$

where, similarly to the activity start times, δ_T^-/δ_T^+ is a (user-inputted) cost coefficient for a cooler/warmer temperature than requested.

4.2.4. Objective Function

The total electricity consumed in each time period t is given by:

$$E_t^{Tot} = E_t^{app} + E_t^{ev} + E_t^{hvac} + E_t^{bl} \quad (24)$$

where E_t^{bl} is the predicted base load during time period t. The objective function to be minimized is a weighted sum of the electricity cost and user discomfort components over a fixed horizon N:

$$\alpha \sum_{t=1}^{N} (E_t^{Tot} v_t) + \beta \sum_{a \in S} f_a + \gamma \sum_{t=1}^{N} g_t \quad (25)$$

where v_t is the (predicted) unit cost of electricity in time period t; and α, β, and γ are (user-inputted) coefficients representing the priority of the objectives from the user's perspective. Standard techniques are used to reformulate the problem as a MILP.

5. Price Forecasting for Energy Cost-Aware Scheduling

In [1] we showed that optimizing price forecasting models for regression metrics (e.g., MSE) does not automatically lead to improved schedule-cost. In fact, the best price forecast model (i.e., best MSE), had worst schedule-cost. Here, we further analyze those forecasts with respect to both price estimation metrics and scheduling cost, and clarify that behavior. We then propose better scheduling-driven price forecasts.

5.1. Learning Price Forecasts

In the price forecasting framework presented in Section 3 we build regression vectors for each half-hour, based on historical and forecasted data. In this paper we use a similar price forecasting framework: we build kernel-SVM models and tune model parameters to optimize a set of

error metrics. The novelty of this work is in showing what are the trade-offs of different regression error metrics and how these trade-offs affect the end scheduling-cost in two different scheduling applications.

5.2. Evaluation Methodology

We evaluate price forecasts with respect to metrics defined below, on the set of 88 test days (3 months of 2011 as test data) described in Section 3.

In [1] the parameters of our models are optimized for minimizing the MSE (Equation 26), but they can be calibrated for any quality measure. The reference price forecast used for comparison is the Irish market operator's forecast, which we refer to as SEMO from now on[13]. MSE is a classical measure of both bias and variance of regression models [26]. We also compute the Mean Absolute Error (MAE, Equation 27). We show that MSE is less useful than MAE as a price estimation metric, because MSE is too sensitive to outliers. Nevertheless, reducing the MAE alone, is still not sufficient to guarantee significant impact on schedule-cost. In [1], we observed that underestimating the price had higher effect on scheduling-cost, than over-estimating, but did not quantify this effect or present ways to exploit this for building better forecasts. Here, we decompose the MAE into an over-estimation and an under-estimation component (MAEover, MAEunder, Equations 28-29), and compute, via experiments, their different weight on scheduling-cost.

Furthermore, we employ the Spearman rank correlation [27] (Equation 30), to characterize how well does a forecast estimate the correct *ranking of time periods* with respect to the true price. We show that optimizing price forecasts for these 3 metrics, MAEover, MAEunder and Spearman rank correlation, is important to achieve cheaper scheduling. In Equations 26-30, x_obs stands for the actual, observed, price, while x_pred stands for the predicted price.

$$MSE = \frac{1}{n} \sum_{i=1}^{n} (x_{pred,i} - x_{obs,i})^2 \quad (26)$$

$$MAE = \frac{1}{n} \sum_{i=1}^{n} |x_{pred,i} - x_{obs,i}| \quad (27)$$

$$MAEover = \frac{1}{n} \sum_{\{i|x_{pred,i} > x_{obs,i}\}}^{n} |x_{pred,i} - x_{obs,i}| \quad (28)$$

$$MAEunder = \frac{1}{n} \sum_{\{i|x_{pred,i} < x_{obs,i}\}}^{n} |x_{pred,i} - x_{obs,i}| \quad (29)$$

---

[13]Although SEMO is also short for the market operators' name, from here on, SEMO will refer to the operator's forecast. Clarification will be given when required.

$$Spearcor =$$
$$\frac{\sum_{i=1}^{n}(rx_{obs,i} - \bar{rx}_{obs,i})(rx_{pred,i} - \bar{rx}_{pred,i})}{\sqrt{\sum_{i=1}^{n}(rx_{obs,i} - \bar{rx}_{obs,i})^2 \sum_{i=1}^{n}(rx_{pred,i} - \bar{rx}_{pred,i})^2}} \quad (30)$$

Spearcor stands for the Spearman rank correlation between the forecasted prices and the true prices each day, $rx_{obs,i}/rx_{pred,i}$ is the true/predicted price rank of time period i. A rank correlation of 1 means that the rankings are identical, while a value of -1 means that the two rankings are the exact opposite. Further details on how this metric is computed can be found in [27].

5.3. Baseline Price Forecasts

To analyze the benefits of building energy cost-aware schedules based on the true market price (and its forecasts), we will compare our forecasts with the following baseline schedules:

- Minimize schedule-cost with respect to true price to get the best-case scheduling cost where the true market price is known (*BestCase*).

- Maximize schedule-cost with respect to the true price to get the worst-case scheduling cost, where energy-intensive tasks would be placed at high price time slots (*WorstCase*).

- Minimize schedule-cost using an informed[14] day-night tariff as a price forecast to build and evaluate the schedule. This day/night-tariff price forecast is computed as the average true price over the 88 test days for day and night periods (day 9h-23h, night 24h-8h). This schedule gives an idea about the cost when using a day-night tariff that encourages simply shifting the big energy consumers to night-tariff periods (*DayNight-DayNight*).

- Minimize schedule-cost using the same day-night tariff as above, but evaluate with the true price (*DayNight-True*). The difference between the *DayNight-DayNight* and *DayNight-True* is only in the evaluation of the schedule (they are both built using the same forecast), the first is evaluated using the forecast, the second is evaluated using the true price.

- Minimise schedule-cost based on simple price forecasts, e.g., averages of historical true price over a time period in the past (*PrevDay-True*, *WeeklyAvg-True*, *MonthlyAvg-True*, *YearlyAvg-True*). Evaluation of cost is done using the true price. *PrevDay-True* refers to the price at the same half-hour for the most recent day in the past for which we know the true price. *WeeklyAvg-True* is the averaged price at each half-hour over the most recent 7 days in the past, for which we know the true price. Similarly for the *MonthlyAvg-True* and *YearlyAvg-True*, we compute averages over past 30 and 365 days respectively.

6. Feedmill Scheduling Results

In this section we analyze the impact of various price forecasts for two scenarios of the feedmill scheduling application. The first is a simplified scenario where task durations can take only a single half-hour (*"feedmill-single"*), and due dates and total capacity constraints are ignored. This scenario allows us to quickly analyze the impact of various price error metrics on the schedule cost. The optimal schedule can be obtained by simply sorting the tasks with respect to energy requirements (descending), and the electricity price (ascending), so we assign energy-intensive tasks to low energy price time slots. The second scenario (*"feedmill-across"*), is more realistic and it is modeled after a real-world running plant (COSYTEC, UK), but is much more computationally intensive to find optimal solutions.

We generated problem instances randomly, filling each of 4 production lines to capacity for 24 hours. The power requirements for each task were uniformly chosen between 0 kW (as there are a small number of tasks in feedmill production that require very little power) and 199 kW.

Regarding duration of tasks, we analyse and present experiments for the two above-mentioned scenarios, feedmill-single and feedmill-across. In the feedmill-single case all task durations are set to exactly 30 minutes (a single time period). The feedmill-across scenario uses random durations uniformly chosen between 25 and 100 minutes, the last task generated being truncated to fit into 24 hours. The time resolution was set to 5 minutes so that optimal solutions could be found within a 10 minute time limit.

Due to the computational cost of solving to optimality, ten instances were generated for the feedmill-across scenario which, combined with 88 prediction days, resulted in a total of 880 runs for each forecast. For the feedmill-single scenario, we tested on both a set of 100 instances and a subset of 10 instances to confirm that the number of instances had no significant bearing on the results.

For each instance we computed the actual cost based on an optimal solution for the actual price, and for each forecast. The schedule based on the actual price provides a lower bound, but since the actual price is not known in advance, it is not realizable. The state-of-the-art commercial MIP solver IBM ILOG CPLEX V12.5.1[15] was used

---

[14]Built using information of the prices on the actual test data.

[15]http://www-01.ibm.com/software/commerce/optimization/cplex-optimizer/

10

to find the optimal solution for each day in the feedmill-across instances, while the statistical software R^{16} was used for computing the optimal solution to the feedmill-single instances employing the sorting method described above. R was also used to compute all statistics given in the following sections and in Section 7.

6.1. Feedmill-Single

We present results for the schedule-cost when the schedule is built using baseline price forecasts, versus using learned price forecasts, for the first feedmill-single scheduling scenario. Unless otherwise stated, the following scheduling results are on the set of 100 instances, tested with 88 test days of price forecasts.

6.1.1. Baselines

Table 4 shows the average schedule-cost over all instances and all test days for the above baselines in the feedmill-single scenario, *ordered from worst to best schedule-cost* (column %Worse sorted in decreasing order). The **%Worse** column stands for how much worse is the daily schedule-cost using a price forecast, as compared to using the actual true price, for designing the schedule. The percentage is computed for each instance for each day and the average is given in the following tables. This measure helps compare the impact of price forecasts on schedule-cost for different scheduling applications without regard to the schedule-cost scale, which is very different for the feedmill and HEMS applications (e.g., schedule-cost is in hundreds of € for the feedmill versus tens of € for HEMS).

Table 4: **Feedmill-Single:** Schedule-Cost (euro) of Baseline Price Forecasts.

Baseline Schedules	Avg Cost	%Worse
WorstCase	684.67	36.18%
DayNight-DayNight	549.72	10.55%
DayNight-True	548.07	8.48%
PrevDay-True	535.72	5.91%
YearlyAvg-True	532.66	5.38%
WeeklyAvg-True	530.27	4.83%
MonthlyAvg-True	526.82	4.19%
BestCase	505.90	-

The results given in Table 4 show that even with fairly simple price forecasting baselines, the schedule-cost comes close to optimal cost. For example, the schedule built using a price forecast based on the monthly average price and evaluated with the true price (*MonthlyAvg-True*) is

only 4.19% worse on average than using the actual price for building the schedule. We note that even when using an informed *day-night tariff* (based on the true prices), building a schedule with such a price forecast has worse schedule-cost than using a simple *market-price forecast* based on the average historical prices. This is due to the high variability of the true market price for different time periods.

6.1.2. Scheduling-driven Price Forecasts

We present price forecasts optimized for different error metrics, and analyze the impact of those metrics on schedule-cost. We start by comparing the price forecasts (FM1 and FM2) presented in Section 3, with the market operator's forecast (SEMO) as a strong baseline.

Table 5: **Feedmill-Single:** Schedule-Cost (euro) of Price Forecasts.

Price	Min	Mdn	**Avg**	Max	**%Worse**
FM2	358.05	505.76	**522.66**	858.94	**3.28%**
SEMO	358.23	502.58	**521.95**	886.07	**3.10%**
FM1	352.67	504.09	**520.99**	868.93	**2.97%**
Actual	343.12	492.40	**505.90**	838.48	-

Statistics of the scheduling cost for the different price forecasts are given in Table 5 (as we show later, differences in average behaviour between the different forecasts were statistically significant in all cases). We find that these three forecasts further improve on the baseline results of Table 4, with %Worse scheduling cost down to 2.97%. Nevertheless, the price forecast with the best MSE (FM2, as shown in Figure 4a) has significantly worse schedule-cost than the other two (SEMO, FM1). In Figure 4 we analyze these price forecasts with regards to the price error metrics they optimize and the effect they have on schedule cost. Figure 4 shows the distribution of MSE, MAE, Spearman rank correlation and schedule-cost across the 88 test days (and across the 100 instances for the schedule-cost), for the 3 forecasts, via boxplots[17]. Each circle on the Y axis of the boxplots, represents one test day, and in the case of schedule-cost, an instance of the test day. The colored line in each box shows the total average and the 95% confidence interval for each metric.

Both FM1 and FM2 outperform SEMO, regarding *price estimation metrics*, MSE and MAE (Figure 4a-4b). Figure 4c compares the forecasts with respect to *price ranking* performance, via the average Spearman rank correlation. We note that although FM2 is best with regards to MSE, it is worst with regards to daily Spearman rank correlation. FM1, on the other hand, is best from the perspec-
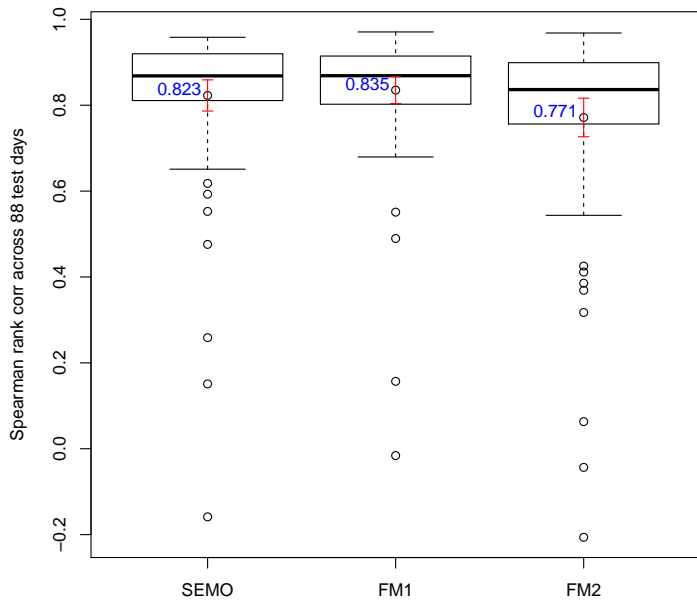
---

[16]http://www.r-project.org

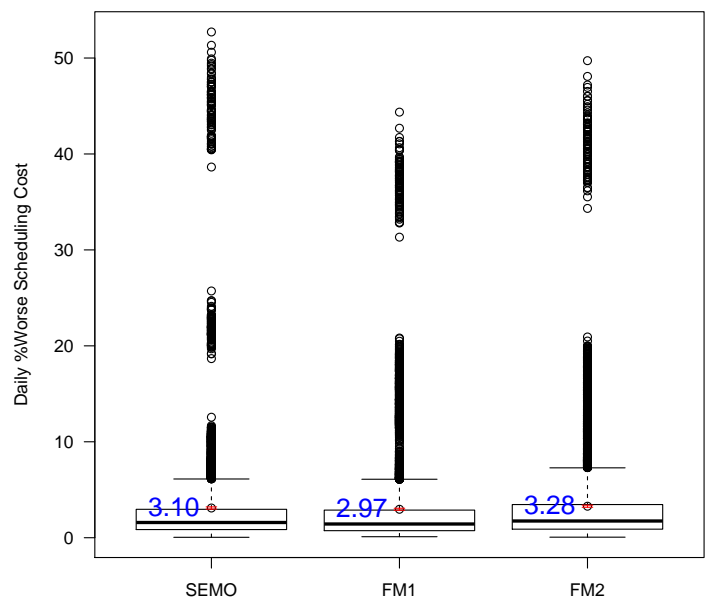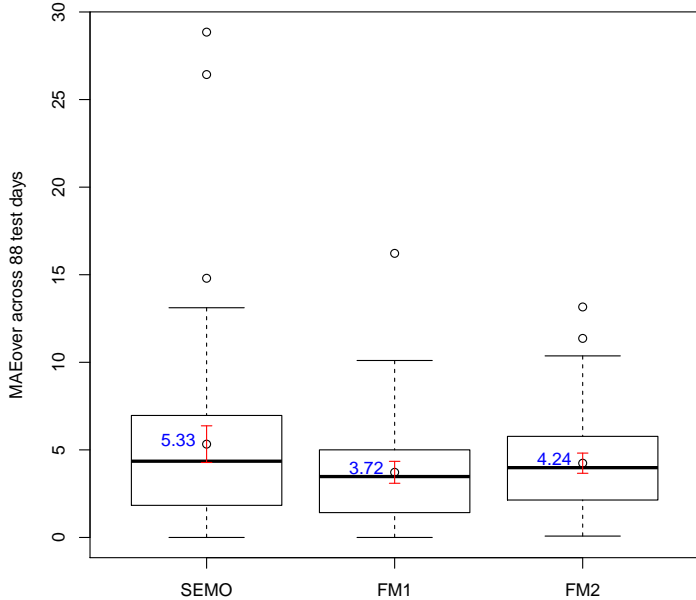[17]The default in R is that the box range is 1.5 · IQR.
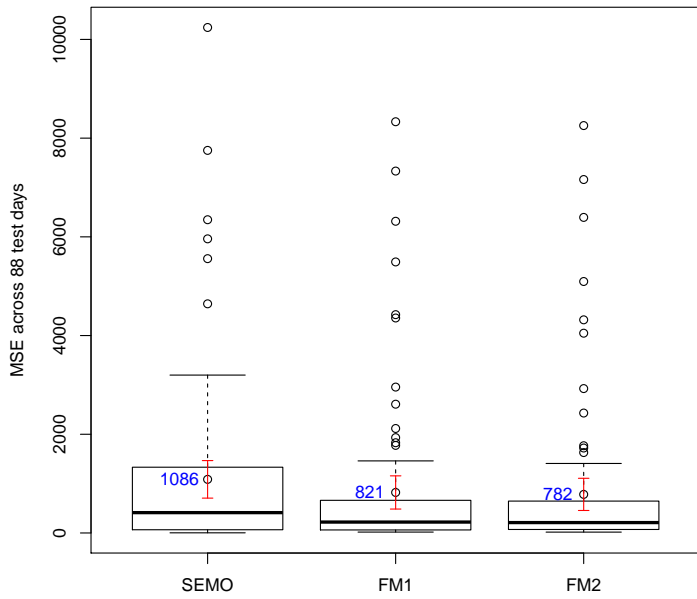
(a) MSE
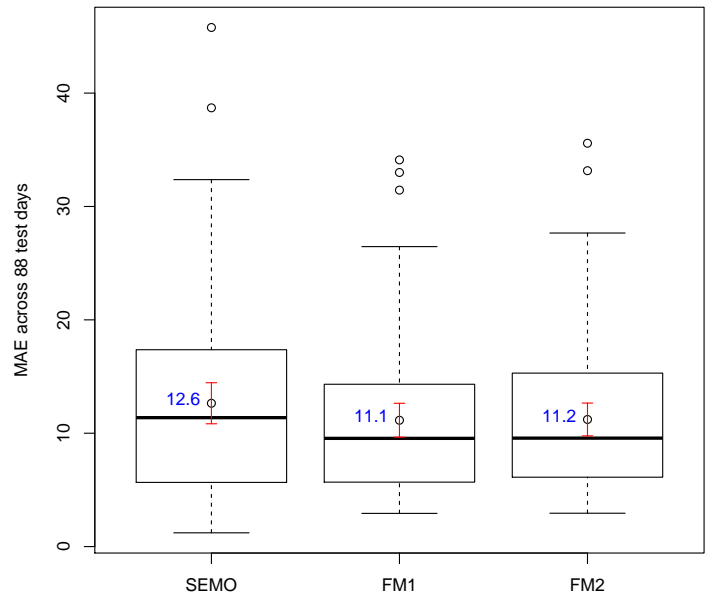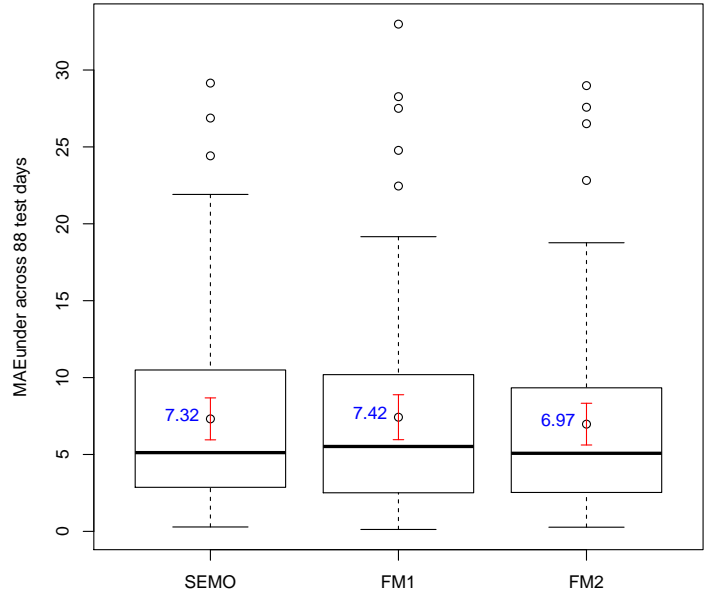
(b) MAE

(c) Average Spearman Rank Correlation

(d) %Worse scheduling cost per instance per day

Figure 4: Price forecasts compared wrt MSE, MAE, Rank Correlation and Scheduling-Cost.

(a) MAEover

(b) MAEunder

Figure 5: Price forecasts compared wrt MAEover and MAEunder.

tive of MAE and Spearman rank correlation, and only marginally worse than FM2 in terms of MSE.

Lastly, Figure 4d focuses on the schedule-cost of the 3 forecasts (rather than price forecasting metrics), and shows the average scheduling-cost reported as %Worse (relative to using the true price to design the schedule). FM1 performed best. The outliers above 30% come from the same day (Feb 4th, 2011) for the three forecasts. The night time prices were nearly always twice the average daytime price, possibly due to unexpected generator outages. This was the only day for SEMO and FM1 where the Spearman correlation was negative, whereas there were a number of days where the SEMO forecast had worse MSE.

Figure 4 shows an interesting point: price estimation metrics (e.g., MSE, MAE) cannot alone explain the effect of different price forecasts on schedule-cost. Price *ranking* metrics (e.g., Spearman rank correlation) bring light into new properties that a price forecast should have, to benefit scheduling: it is not enough to get the price estimate right, the forecast has to get the rank of prices right too. Nevertheless, ranking alone is not enough to deliver a forecast that benefits scheduling, since similar price ranking can be achieved for very different price forecasts, e.g., a *hypothetical forecast* that swaps a €50 time slot with a €51 time slot, versus another one that swaps a €50 time slot with a €200 time slot. With respect to ranking, the two forecasts would be the same, but the penalty implied for price estimation is very high, since we may schedule an energy-intensive task during a very high priced time

slot. Thus, both price magnitude and price rank have to be forecasted correctly by a price forecasting model, to positively impact the schedule-cost.

To better understand the trade-offs of each price forecast, in Figure 5 we decompose the MAE into an over-estimation and an under-estimation component (MAEover, MAEunder). This decomposition allows us to zoom-in on the type of price estimation error made by each price forecast. Figure 5 shows that all forecasts are better at predicting low values than high values (they don't overestimate as much as they under-estimate the price). Additionally, although FM2 has both MAEover and MAEunder error values considerably lower than SEMO, due to having a bad Spearman rank correlation it leads to worst schedule-cost. The above discussion sheds light into the key point for building price forecasts for cost-aware scheduling: *good price forecasts should co-optimize price estimation and price ranking metrics.*

We show next how to build such forecasts. We first evaluate a simple metric based on MAE and Spearman rank correlation: maximizing the geometric mean of (-MAE) and average daily Spearman correlation. This essentially favors price forecasts that minimize MAE and maximize Spearman rank correlation. The geometric mean allows combining quantities expressed on different scales (e.g., MAE of 11.0 versus Spearman correlation of 0.83). We tune the kernel-SVM model parameters (i.e., C, γ and ϵ) to optimize this new metric (rather than the MSE), using the learning framework of FM1 (since this has the
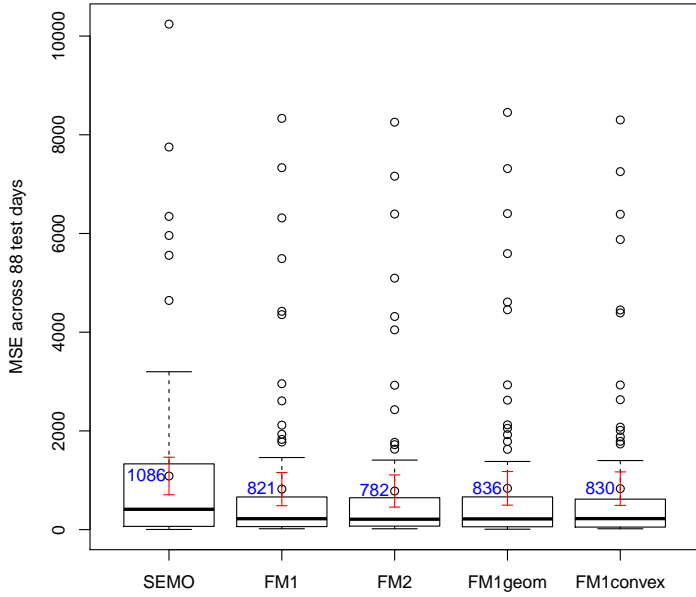
(a) MSE
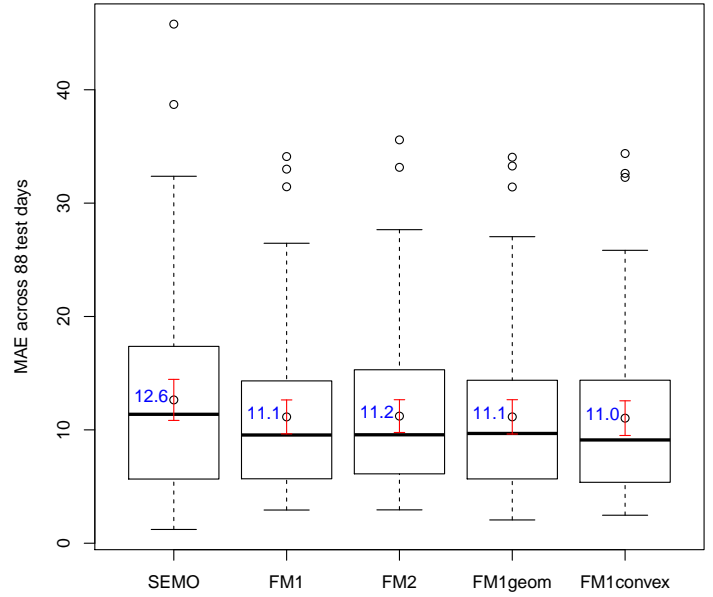
(b) MAE
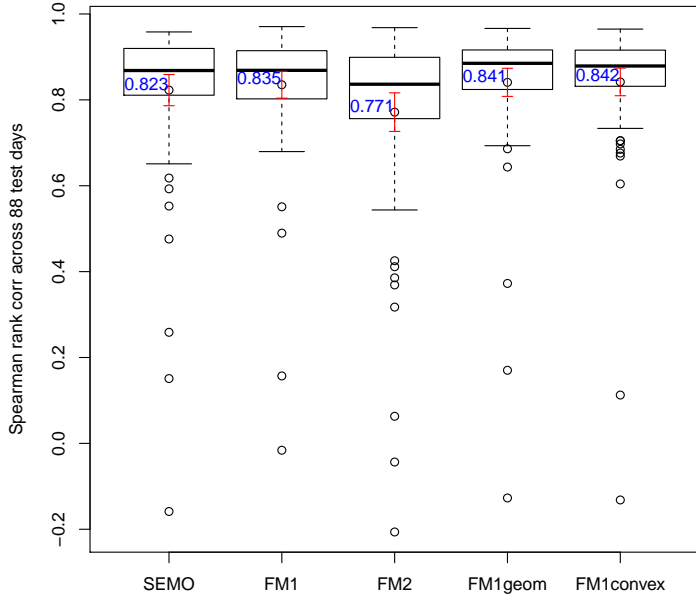
(c) Spearman Rank Correlation
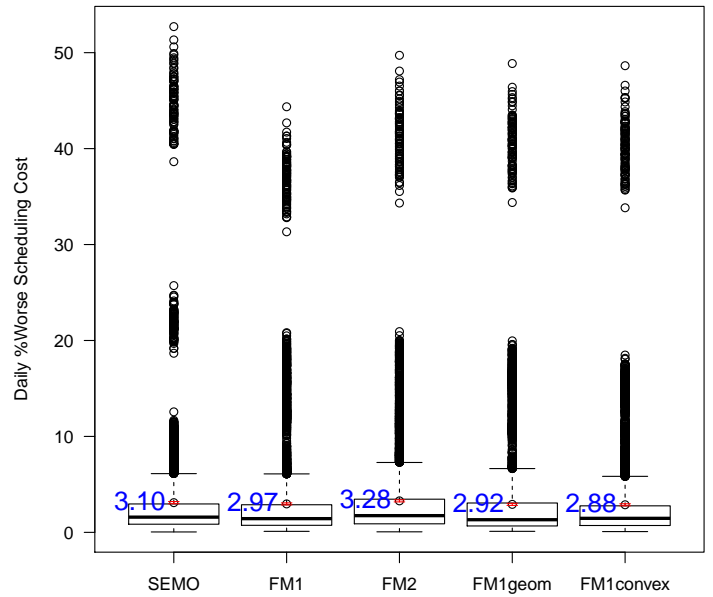
(d) %Worse scheduling cost per instance per day

Figure 6: Price forecasts compared wrt MSE, MAE, Rank Correlation and (Feedmill-Single) Scheduling Cost.

best schedule-cost). We refer to the resulting forecast as *FM1geom*. Next, we analyze approaches to capture finer trade-offs when optimizing the price forecast, e.g., via co-optimizing the decomposed MAE (MAEover, MAEunder) and Spearman rank correlation. In order to allow each metric to have a (metric-importance) weight, we op-timize the convex combination of these scaled metrics[18]. We use the previously gathered set of forecasts and their

---

[18]Each metric is computed as an average over test days and is nor-malized by its maximum value over test days. Convex combination refers to the weighted linear combination, weights sum up to 1.

14

Table 6: Feedmill-Single: Schedule-Cost (euro) of Scheduling-Driven Price Forecasts.

Price	Min	Mdn	**Avg**	Max	%**Worse**
FM2	358.05	505.76	**522.66**	858.94	**3.28%**
SEMO	358.23	502.58	**521.95**	886.07	**3.10%**
FM1	352.67	504.09	**520.99**	868.93	**2.97%**
FM1geom	351.76	503.74	**520.76**	866.37	**2.92%**
FM1convex	350.15	503.20	**520.63**	861.51	**2.88%**
FM1scost	362.78	509.60	**527.63**	858.22	**4.35%**
Actual	343.12	492.40	**505.90**	838.48	-

Table 7: Confidence intervals (95%) and p-values for **feedmill-single** scheduling-cost with new price forecasts.

Price		SEMO	FM1	FM2	**FM1geom**	**FM1convex**
Actual	L	-16.67	-15.63	-17.33	-15.43	-15.29
	U	-15.43	-14.55	-16.18	-14.28	-14.17
	p	$2.2e-16$	$2.2e-16$	$2.2e-16$	$2.2e-16$	$2.2e-16$
SEMO	L	-	0.71	-0.91	0.96	1.10
	U	-	1.21	-0.51	1.42	1.54
	p	-	$3.2e-14$	$9.0e-12$	$2.2e-16$	$2.2e-16$
FM1	L	-	-	-1.87	0.15	0.23
	U	-	-	-1.47	0.31	0.49
	p	-	-	$2.2e-16$	$1.5e-8$	$6.0e-8$

scheduling-cost, to learn the weights of the 3 metrics by fitting a linear regression model (e.g., with the 3 metrics as features, to predict the %Worse scheduling-cost as a regression target). Modeled this way, the learned weights (rescaled to sum to 1) are: 0.573 for the average Spearman correlation, 0.305 for MAEunder and 0.122 for MAEover. This shows that the 3 metrics all play a role (since the weight > 0) in capturing the potential of a price forecast for reducing scheduling-cost, with the most important metric being rank correlation, followed by MAEunder (correctly estimating high prices) and MAEover (correctly estimating low prices). We build a new price forecast by tuning the kernel-SVM parameters to optimize this convex combination (denoted by *FM1convex*). We test this same price forecast for the other two scheduling scenarios (Section 6.2 and Section 7) to see if it generalizes. i.e., still leads to lower schedule-cost.

Table 6 and Figure 6 show FM1geom and FM1convex, compared to the previous 3 forecasts: SEMO, FM1, FM2. The two new forecasts outperform the previous forecasts in terms of scheduling cost. We further note that co-optimizing the convex combination of MAEover, MAEunder and average Spearman correlation (FM1convex) performed marginally better than FM1geom with regards to scheduling cost. Table 7 provides paired t-test results for the 5 forecasts, showing that the improvements of

FM1geom and FM1convex over the previous 3 forecasts were statistically significant.

For the feedmill-single simplified scheduling scenario, we can build price forecasts that optimize directly the scheduling cost. This is normally expensive to compute for more realistic (i.e., more constrained) problem applications. For feedmill-single, we simply sort the tasks by decreasing energy demand, sort the price forecast by increasing price, and use the forecast to order the true price and compute the end schedule-cost. This gives a fast way to compute the schedule-cost for a price forecast, and thus can be used for parameter tuning. In Table 6 we also show the results of optimizing directly the scheduling cost to create a forecast (*FM1scost*) for this simple scenario. Surprisingly, we find that FM1scost results in worse scheduling cost than either FM1geom or FM1convex. Indeed, FM1scost resulted in considerably worse performance than any of the other forecasts in Table 6. We discuss next the reasons behind this behavior.

Consider the following hypothetical example where there are three tasks to schedule, with energy requirements 200MWh, 100MWh, 50MWh, and durations of 30 minutes each. Assume the true price of the three time periods is €50/MWh, €100/MWh, €50/MWh. We compare two sample price forecasts for this setting: €40/MWh, €80/MWh, €40/MWh and €0/MWh,

Table 8: **Feedmill-Single:** Average Schedule-Cost (euro) and %Worse over 100 instances versus over 10 instances.

Forecast	Cost 100	Cost 10	%Worse 100	%Worse 10	%Worse Diff
WorstCase	684.67	688.62	36.18	35.99	-0.19
DayNight-DayNight	549.72	553.91	10.55	10.62	0.07
DayNight-True	548.07	552.21	8.48	8.55	0.07
Prevday-True	535.72	540.39	5.91	6.10	0.19
YearlyAvg-True	532.66	536.64	5.38	5.43	0.05
WeeklyAvg-True	530.27	534.10	4.83	4.85	0.02
MonthlyAvg-True	526.82	530.74	4.19	4.23	0.04
FM2	522.66	526.47	3.28	3.31	0.03
SEMO	521.95	526.06	3.10	3.19	0.09
FM1	520.99	524.82	2.97	3.00	0.03
FM1geom	520.76	524.60	2.92	2.96	0.04
FM1convex	520.63	524.46	2.88	2.91	0.03
FM1scost	527.63	531.43	4.35	4.37	0.02
BestCase	505.90	509.47	-	-	-

€1/MWh, €0/MWh for this setting. The procedure for computing the schedule-cost, is to sort the forecasted price from lowest to highest, and use the forecast ordering to induce an ordering on the true price, which we then use to compute the end schedule-cost. If we sort the true price according to either forecast, we get the same ordering €50/MWh, €50/MWh, €100/MWh, thus the same schedule-cost: $50 * 200 + 50 * 100 + 100 * 50$. Nevertheless, these two price forecasts are very different, and while the first one is reasonably close to the true price, the second one is quite far off. This issue suggests that simply optimizing the schedule cost may not be well-suited for parameter tuning of the price forecasting model, since we need to also constrain the price estimation error, not only the ranking of prices, thus again having to co-optimize schedule-cost (similar to optimizing price ranking) and MAE (for price estimation).

When we build such a price forecast we get the same result as for the previous FM1geom that was co-optimizing MAE and Spearman rank correlation. Therefore, we can see from this experiment that our proxy metric combining MAE and Spearman rank correlation is quite effective for building price forecasts that are good for cost-aware scheduling. Furthermore, running a realistic schedule with each price forecast is normally a time consuming process, thus a proxy metric that only involves price forecast properties (without having to run the schedule for each price forecast and parameter configuration) is quite useful.

Finally, Table 8 provides results for each forecast over 100 instances and over the first 10 instances. We note a decrease in average schedule-cost of approximately 4 euro for all forecasts, when going from 10 to 100 instances. However, there is no difference in the ranking of the dif-

ferent forecasts with respect to average schedule-cost, and very little difference in performance in terms of %Worse cost, as can be seen from the final column.

6.2. Feedmill-Across

In this section, we compare the previous 5 forecasts for a more realistic scheduling scenario, where tasks can take between 25 to 100 minutes, thus can be spread over more than one time slot or only part of a time slot. We test on a set of 10 instances (as described in Section 6), in combination with the same 88 test days of forecasts.

6.2.1. Baselines

Table 9 shows the average schedule-cost over the test days for the baseline price forecasts on the *feedmill-across* instances, where tasks can span a fraction of one or several time periods.

Table 9: **Feedmill-Across**: Schedule-Cost (euro) of Baseline Price Forecasts.

Baseline Schedules	Avg Cost	%Worse
WorstCase	680.83	34.08
DayNight-DayNight	549.06	8.13
DayNight-True	548.08	7.94
PrevDay-True	535.97	5.55
YearlyAvg-True	533.15	5.00
WeeklyAvg-True	530.85	4.54
MonthlyAvg-True	527.74	3.93
BestCase	507.78	-

Table 10: **Feedmill-Across**: Schedule-Cost (euro) of Scheduling-Driven Price Forecasts.

Price	Min	Mdn	**Avg**	Max	%**Worse**
FM2	381.14	507.57	**523.62**	838.21	**3.07%**
SEMO	375.64	504.58	**522.73**	851.57	**2.87%**
FM1	374.68	504.97	**522.24**	838.44	**2.82%**
FM1geom	374.11	505.24	**521.98**	844.45	**2.77%**
FM1convex	371.82	504.36	**522.04**	838.29	**2.76%**
Actual	365.31	494.54	**507.78**	817.16	-

Table 11: Confidence intervals (95%) and p-values for **feedmill-across** price-aware scheduling costs comparing optimal solutions priced using actual price, and each of the analysed price forecasts.

Price		SEMO	FM1	FM2	**FM1geom**	**FM1convex**
Actual	L	-16.71	-16.15	-17.61	-15.97	-16.00
	U	-13.19	-12.78	-14.07	-12.44	-12.52
	p	$2.2e-16$	$2.2e-16$	$2.2e-16$	$2.2e-16$	$2.2e-16$
SEMO	L	-	-0.13	-1.44	0.15	0.12
	U	-	1.11	-0.32	1.34	1.25
	p	-	$1.2e-1$	$1.9e-3$	$1.4e-2$	$1.6e-2$
FM1	L	-	-	-2.00	-0.00	-0.20
	U	-	-	-0.74	0.52	0.60
	p	-	-	$1.9e-5$	$5.4e-2$	$3.3e-1$

Note that the costs are quite similar to the feedmill-single scenario because both types of instances were generated such that each production line must have a task running at all times. The results follow the same pattern as those found with the baselines on the feedmill-single instances (cf. Table 4), with a day-night rule of thumb approach again performing much worse than the simple forecasts based on historical prices.

6.2.2. Scheduling-driven Price Forecasts
Table 10 and Figure 7 compare the schedule-cost of the 5 forecasts for the feedmill-across scheduling scenario. The dots on the Y axis of the boxplot represent %Worse schedule-cost for the scheduling instances (880 data points, 10 instances for each of 88 test days), built using the forecast versus built based on the true price. We observe that the scheduling-cost when the schedule is built using the FM1geom and FM1convex price forecasts remains best when compared to the other forecasts. The FM1convex price forecast is the exact same as for the previous section, to test how that forecast behaves for other scheduling scenarios.

The outlier points are those with worst scheduling-cost, and they correspond to 2-3 days in which, as discussed in the previous section, prices were abnormally high across the night period or there were electricity price peaks at unusual times (e.g. peak occurring on 06/04/11 at 6:30

a.m. was €323.77/MWh). Interestingly, for the former day type the rank correlation was lowest for all 3 forecasts as previously discussed, while for the latter the rank correlation was high for all 3 forecasts (> 0.77) but still resulted in poor performance as the price at the 6:30a..m. peak, mentioned above, was approximately 5 times larger than the average price for the day. This further illustrates the need to co-optimize price ranking and estimation error metrics.

Table 11 gives t-tests results for the schedule-cost of the 5 forecasts. The statistical significance results for 10 instances per day are not as strong as for the 100 instances of the feedmill-single scenario, although they are similar to the t-test results for the first 10 instances of the feedmill-single scenario.

The main observation for this experimental setup (i.e., feedmill-across vs feedmill-single) is that the ranking of price forecasts with respect to reducing scheduling-cost is preserved, with the price forecasts that are built by co-optimizing MAE and Spearman rank correlation providing better schedule-cost.

7. HEMS Scheduling Results

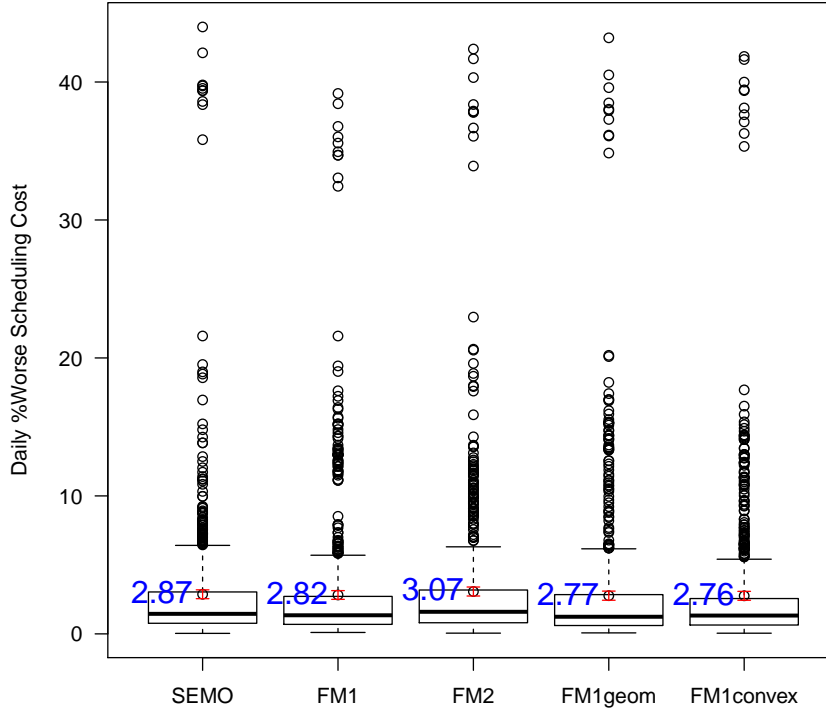The HEMS was run on a subset of 40 of the 88 test days, according to the scheduling scenario described in [28].

Figure 7: Feedmill-Across: Price forecasts compared wrt %Worse Scheduling Cost.

This accounted for ten weeks of consecutive Mon-Fri, midday to midday forecasts. The inputs which changed from day to day were the price forecast and external weather data (taken from Weather Underground[19]).

The daily (actual) schedule-cost was computed in an identical manner to that in the Feedmill scheduling application, and the state-of-the-art commercial MIP solver IBM ILOG CPLEX V12.5.1 was again used to find the optimal solution for each day for each forecast.

For this application the forecast with best MSE (i.e., FM2) delivered good scheduling-cost, but by co-optimizing price estimation error and Spearman rank correlation, we further reduced the cost. Table 12 shows statistics for the cost across all instances over the 40 test-days. Table 13 gives the t-test results for comparing the scheduling-cost by using different price forecasts. Figure 8 shows the 5 forecasts for this scheduling scenario.

We note that the FM1convex forecast described in Section 6.1 does not deliver better cost here (e.g., as compared to FM1geom) and there was no statistical significant difference between FM1convex and FM1. The reason for this could be that the importance weight of MSE or MAE is higher for HEMS, than it was for Feedmill, relative to the importance of the Spearman correlation metric. This is further supported by the fact that FM2 with best MSE, has good schedule-cost for this application (while it had worst schedule-cost for Feedmill).

Nevertheless, in order to re-learn metric importance weights, one needs enough samples of forecasts and their schedule cost. It is much simpler to use the geometric mean of MAE and Spearman correlation to co-optimize the price estimation and ranking. If enough examples of price forecasts and their scheduling-cost are available, learning metric-importance weights may further reduce the schedule-cost (of FM1geom), but testing this hypothesis is beyond the scope of this paper.

Note that this scheduling application is quite different from the feedmill scheduling problem. Firstly the optimality criterion involves not only energy cost, but also user discomfort factors related to user temperature and appliance start time preferences in a weighted sum objective function.

Secondly, as illustrated in Figure 9, there is typically considerably more slack in the HEMS experimental setup than in the feedmill instances. In the former, there are large time periods during the day where there is little if any energy consumption to be scheduled. On the other hand, production lines are typically run non-stop in feedmill production. Note that the feedmill instances generated had no slack in that there were no time periods where

---

Table 12: **HEMS**: Schedule-Cost (euro) of Price Forecasts.

Price	Min	Mdn	Avg	Max	%Worse
SEMO	1.50	1.96	2.10	2.93	6.50%
FM1	1.55	1.94	2.06	2.96	4.69%
FM1convex	1.49	1.96	2.05	2.93	4.24%
FM2	1.53	1.96	2.05	2.94	4.06%
FM1geom	1.49	1.93	2.04	2.94	3.70%
Actual	1.35	1.90	1.97	2.84	-

Table 13: Confidence intervals (95%) and p-values for **HEMS** price-aware scheduling costs.

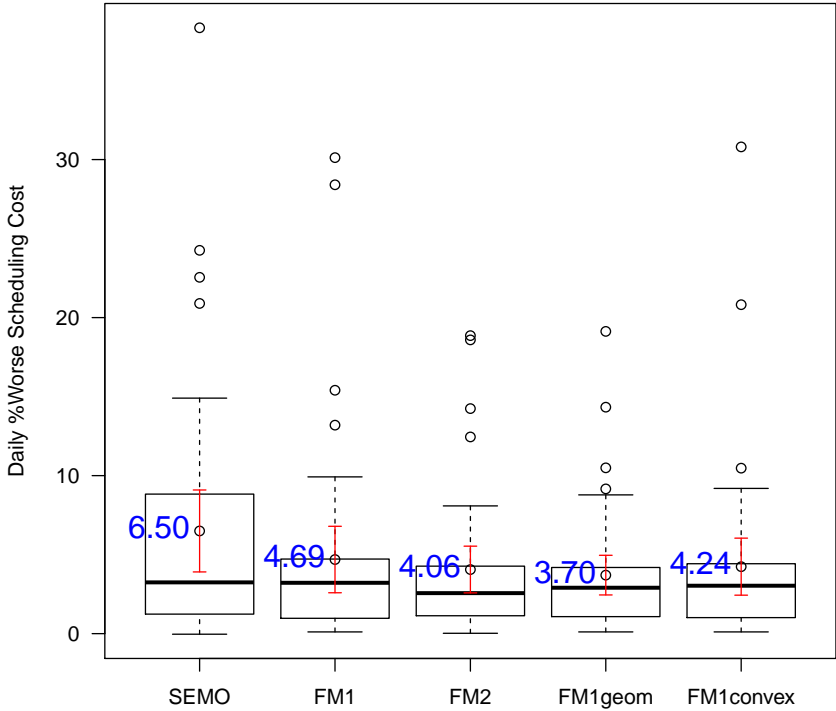| Price		SEMO	FM1	FM2	**FM1geom**	**FM1convex**
Actual	L	-0.18	-0.13	-0.10	-0.09	-0.12
	U	-0.07	-0.05	-0.05	-0.04	-0.04
	p	$1.4e-5$	$7.3e-5$	$9.6e-7$	$8.7e-7$	$5.0e-5$
SEMO	L	-	-0.003	0.01	0.01	0.01
	U	-	0.07	0.08	0.09	0.07
	p	-	$7.2e-2$	$7.7e-3$	$3.2e-3$	$7.5e-3$
FM1	L	-	-	-0.01	-0.003	-0.01
	U	-	-	0.04	0.04	0.03
	p	-	-	$2.9e-1$	$9.9e-2$	0.5

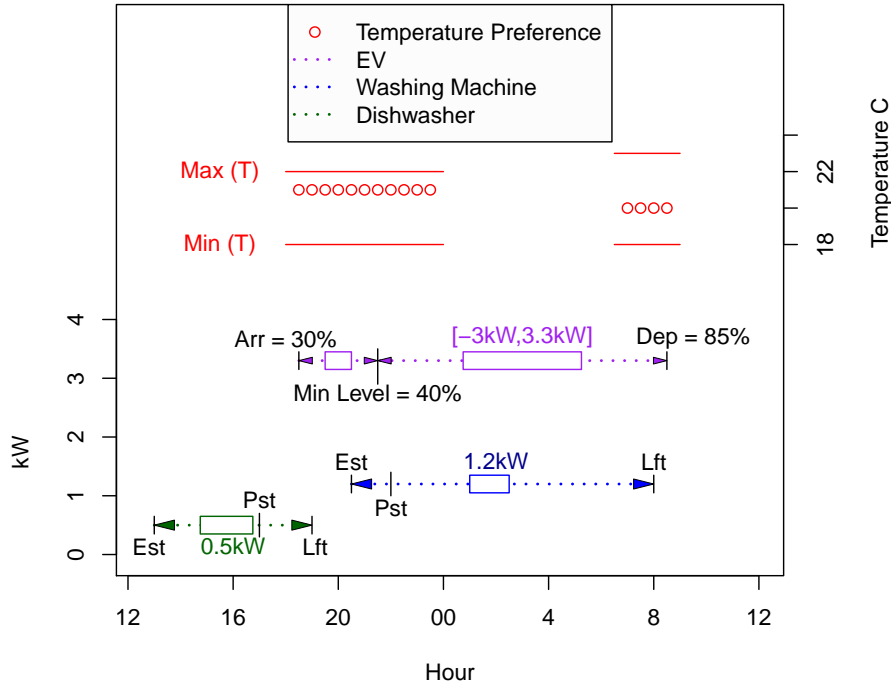Figure 8: HEMS: Price forecasts compared wrt %Worse Scheduling Cost.

Figure 9: HEMS Experimental Setup: Est/Pst/Lft refers to earliest starting/preferred starting/latest finishing time for a deferrable appliance. The bottom half of the figure, along with left y-axis, shows user preferences and energy consumption requirements for an EV and schedulable appliances; the top half, along with right y-axis, shows the user temperature requests.

a task wasn't running, albeit some tasks may have had a power rating of 0kWh. Indeed, as can be observed in the figure, there are time windows (e.g., 8:30 - 12:30) when the only possible energy consumer that could be scheduled is the HVAC. Since the next temperature request after 8:30 is at 18:00, it is highly unlikely that preheating will occur in this time window, given heat losses.

Finally, the impact of the price has different characteristics for the Feedmill scheduling problem as for the HEMS. In the former the tasks are scheduled based on the relative price, adding a constant to all prices will result in an identical schedule. In the HEMS, there are energy losses to overcome when preheating the home and when availing of V2H capabilities. Preheating / discharging will only occur if the price differential between time periods is above some certain threshold value. Therefore the daily energy consumption in the HEMS instances is dependent on the electricity prices, energy efficiency of heating/battery-charging, etc., whereas each feedmill instance has a fixed energy consumption total.

For example let us consider the EV battery charging requirements in the HEMS. There is an efficiency rating for charging (z^c) and discharging (z^d) the battery. In our experiments the battery must be charged from 40% to 85% between 21:00 and 08:00 the next morning (so 22 available time periods). For a battery of capacity 30kW and maximum charge rate of 3.3kW, this means that at least the 9 *cheapest* time periods will be used for this charge. In the remaining 13 time periods, V2H will only occur if the ratio of price in the discharge period to price in the "charge-back" period is greater than $1/(z^c * z^d)$, to

overcome the energy losses due to charging/discharging the battery. This is further constrained by the bounds on the battery capacity, l^{min}/l^{max} which cannot be violated.

8. Conclusion

In this paper we study methods for building effective price forecasts for cost-aware scheduling. We analyze the Irish electricity market, present electricity price forecasting models and test our forecasts for two different scheduling applications: animal feed manufacturing and home energy management systems. We show that building cost-aware schedules using forecasts of the actual electricity market price, is feasible, and leads to significant cost savings, as compared to cost-unware or day-night-tarrif based schedules. This is an encouraging finding in the context of increased electricity prices and increased interest from electricity stakeholders (operators, suppliers, consumers) to start charging/paying the actual market electricity price, instead of constant tarrifs plans.

One of the key technical insights from this work is that optimizing price forecasts for regression metrics alone (e.g., estimating the price magnitude-error via the MSE or MAE), does not suffice to guarantee significant reductions in schedule-cost (as already shown in our prior work [1] for manufacturing scheduling). Adding to the state-of-the-art (e.g., [1]), in this paper we show that price ranking is also an important metric, and present forecasts that co-optimize price estimation and price ranking to deliver improved schedule-cost across all scheduling scenarios studied. Depending on the scheduling setting and application,

the price estimation (e.g., MSE or MAE) seems to have a different importance-weight relative to price ranking (e.g., Spearman rank correlation). In this work we have shown two ways to combine price estimation and ranking metrics, and have shown how to learn the importance weights of each metric, from given examples of price forecasts with known schedule-cost.

In the future we intend to further study the relative importance of various price forecasting metrics in different scheduling scenarios, e.g., how do the importance-weights of price metrics change with various scheduling constraints.

References

[1] G. Ifrim, B. OSullivan, H. Simonis, Properties of energy-price forecasts for scheduling, in: M. Milano (Ed.), Principles and Practice of Constraint Programming, Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2012, pp. 957–972.
URL `http://dx.doi.org/10.1007/978-3-642-33558-7_68`

[2] P. Conlon, Development of Domestic and SME Time of Use Tariff Structures for a Smart Metering Program in Ireland, 2008.

[3] H. Allcott, Real-Time Pricing and Electricity Market Design, New York University, 2013.

[4] S. K. Aggarwal, L. M. Saini, A. Kumar, Electricity price forecasting in deregulated markets: A review and evaluation, International Journal of Electrical Power and Energy Systems 31 (1) (2009) 13 – 22.
URL `http://www.sciencedirect.com/science/article/pii/S0142061508000884`

[5] J. L. Berral, I. n. Goiri, R. Nou, F. Julià, J. Guitart, R. Gavaldà, J. Torres, Towards energy-aware scheduling in data centers using machine learning, in: Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking, e-Energy '10, ACM, New York, NY, USA, 2010, pp. 215–224.
URL `http://doi.acm.org/10.1145/1791314.1791349`

[6] D. Aikema, C. Kiddle, R. Simmonds, Energy-cost-aware scheduling of hpc workloads, in: IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM), 2011, 2011, pp. 1 –7.

[7] B. Guenter, N. Jain, C. Williams, Managing cost, performance, and reliability tradeoffs for energy-aware server provisioning, in: Proceedings of the IEEE International Conference on Computer Communications (INFOCOM), 2011, pp. 1332 –1340.

[8] A. Qureshi, R. Weber, H. Balakrishnan, J. Guttag, B. Maggs, Cutting the electric bill for internet-scale systems, in: Proceedings of the ACM SIGCOMM 2009 conference on Data communication, SIGCOMM '09, ACM, New York, NY, USA, 2009, pp. 123–134.
URL `http://dx.doi.org/10.1145/1592568.1592584`

[9] P. Bodík, R. Griffith, C. Sutton, A. Fox, M. Jordan, D. Patterson, Statistical machine learning makes automatic control practical for internet datacenters, in: Proceedings of the 2009 conference on Hot topics in cloud computing, HotCloud'09, USENIX Association, Berkeley, CA, USA, 2009.
URL `http://dl.acm.org/citation.cfm?id=1855533.1855545`

[10] S. Srikantaiah, A. Kansal, F. Zhao, Energy aware consolidation for cloud computing, in: Proceedings of the 2008 conference on Power aware computing and systems, HotPower'08, USENIX Association, Berkeley, CA, USA, 2008, pp. 10–10.
URL `http://dl.acm.org/citation.cfm?id=1855610.1855620`

[11] S. Albers, Algorithms for Energy Saving, in: S. Albers, H. Alt, S. Nher (Eds.), Efficient Algorithms, Vol. 5760 of Lecture Notes in Computer Science, Springer Berlin Heidelberg, 2009, pp. 173–186.
URL `http://dx.doi.org/10.1007/978-3-642-03456-5_12`

[12] J. Kulkarni, K. Munagala, Algorithms for cost-aware scheduling, in: WAOA, 2012, pp. 201–214.

[13] H. Simonis, T. Hadzic, A family of resource constraints for energy cost aware scheduling, in: Third International Workshop on Constraint Reasoning and Optimization for Computational Sustainability, St. Andrews, Scotland,UK, 2010.

[14] S. Mani, S. Rao, Operating cost aware scheduling model for distributed servers based on global power pricing policies, in: Proceedings of the Fourth Annual ACM Bangalore Conference, COMPUTE

'11, ACM, New York, NY, USA, 2011, pp. 12:1–12:8. URL http://doi.acm.org/10.1145/1980422.1980434

[15] X. Chen, T. Wei, S. Hu, Uncertainty-aware household appliance scheduling considering dynamic electricity pricing in smart home, IEEE Trans. Smart Grid 4 (2) (2013) 932–941.

[16] D. Connolly, H. Lund, B. Mathiesen, M. Leahy, Modelling the existing irish energy-system to identify future energy costs and the maximum wind penetration feasible, Energy 35 (5) (2010) 2164 – 2173.
URL http://www.sciencedirect.com/science/article/pii/S0360544210000484

[17] P. Finn, C. Fitzpatrick, D. Connolly, M. Leahy, L. Relihan, Facilitation of renewable electricity using price based appliance control in Ireland's electricity market, Energy 36 (5) (2011) 2952 – 2960.
URL http://www.sciencedirect.com/science/article/pii/S0360544211001162

[18] M. Jablonska, A. Mayrhofer, J. P. Gleeson, Stochastic simulation of the uplift process for the Irish electricity market, Mathematics-in-Industry Case Studies Journal 2 (pp 86-110).
URL http://www.micsjournal.ca/index.php/mics/article/view/23/23

[19] C.-C. Chang, C.-J. Lin, LIBSVM: a library for support vector machines, ACM Transactions on Intelligent Systems and Technology (TIST) 2 (3) (2011) 27.

[20] H. Simonis, Models for global constraint applications, Constraints 12 (1) (2007) 63–92.

[21] J. Hooker, Integrated Methods for Optimization, Springer, New York, 2007.

[22] M. Howley, B. Gallachóir, E. Dennehy, Energy in Ireland, 2009.

[23] A. Foley, B. Tyther, P. Calnan, B. Ó. Gallachóir, Impacts of Electric Vehicle charging under electricity market operations , Applied Energy 101 (0) (2013) 93 – 102, sustainable Development of Energy, Water and Environment Systems. doi:http://dx.doi.org/10.1016/j.apenergy.2012.06.052.
URL http://www.sciencedirect.com/science/article/pii/S0306261912004977

[24] G. Haines, A. McGordon, P. Jennings, N. Butcher, The Simulation of Vehicle-to-Home Systems–Using Electric Vehicle Battery Storage to Smooth Domestic Electricity Demand, 2009.

[25] *Sustainable Energy Authority Ireland*, Dwelling Energy Assessment Procedure, Version 3.2 (2009).

[26] S. Pelland, G. Galanis, G. Kallos, Solar and photovoltaic forecasting through post-processing of the global environmental multiscale numerical weather prediction model, Progress in Photovoltaics: Research and Applications.
URL http://dx.doi.org/10.1002/pip.1180

[27] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, Numerical recipes in C (2nd ed.): the art of scientific computing, Cambridge University Press, New York, NY, USA, 1992.

[28] D. Grimes, H. Simonis, A. Pratt, C. Sheridan, Automated Energy Usage Optimization for the Residential Sector: Impact of Price Tariffs, in: 3rd International Conference on Computational Sustainability (CompSust'12), 2012.