# Prediction and Robust Control of Energy Flow in Renewable Energy Systems

**Balázs Csanád Csáji** [*] **András Kovács** [*] **József Váncza** [*,**]

[*] *Fraunhofer Project Center for Production Management & Informatics,*
*Institute for Computer Science and Control,*
*Hungarian Academy of Sciences, Budapest, Hungary*
*email: {balazs.csaji, andras.kovacs, jozsef.vancza}@sztaki.mta.hu*
[**] *Budapest University of Technology and Economics, Hungary*

**Abstract:** The paper is motivated by making use of solar energy in public lighting services via an intermediate battery storage. The aim is to develop algorithms for controlling the energy flow in the system, in such a way that robustness against power outages is guaranteed and the total energy cost is minimized. A novel approach is proposed which predicts energy production and consumption by fitting stochastic models to historic data, and solves the resulting optimization problem on a rolling horizon. Experimental results are also presented, illustrating the behavior of the controlled energy system in typical winter and summer days.

*Keywords:* renewable energy systems, robust control, time-series analysis

## 1. INTRODUCTION

This research is motivated by a work aimed at making use of solar energy in public lighting services. To this end, energy supply from renewable solar sources will be coupled with dynamically changing demand of street lighting by optimized and robust energy flow and storage. By combining street lighting, photovoltaic (PV) solar energy generation, energy storage, and advanced sensor technology with novel data processing, communication and control methods, a so-called *energy-positive community microgrid* (E+grid) will be formulated. E+grid is designed to achieve positive energy balance over a one year horizon, while having power grid connection to allow bi-directional energy flow at variable tariff to handle temporal over- or under-production. So as to be able to provide service also in areas where power outage is a concern, the system should be able to operate for a limited time in island mode, too. Production and use of electricity will be continuously monitored and reported by smart meters that give feedback for remote decisions. Beyond offering lighting services, the system will contribute also to the stability of the electrical grid and become a means for managing peak demand.

The focus of this paper is set on how to collect and process data both from the demand and supply sides, and how to control the energy flow within the system, as well as between E+grid and the overall electricity grid. The major challenges are (1) to predict future supply and demand of electricity heavily depending also on uncertain factors like traffic intensity, weather, energy availability, and (2) to match future supply and demand of electricity on a *finite horizon* in a *profit maximizing* and *robust* way, thus warranting, against all uncertainties, a prescribed level of lighting service. Of course, this would not be possible without intermediate storage of electricity, hence control of the energy flow has to take into account the key technical features of the complete renewable energy system.

Since the system as a whole is embedded in an highly uncertain environment, modelling its future behaviour – particularly, the prediction of its energy supply and demand – will be based *not* on the approximated physical models of its components and their interactions. Instead, by relying on the results of extensive and continuous monitoring activity, predictions will be generated from series of measured historical data. However, so as to keep reality and the model of the controlled system in as a close correspondence as possible, control of its energy flow and prediction of its behaviour will be interleaved: the model will be mapped to (observed) reality time and again, in *model predictive* way of control.

## 2. LITERATURE REVIEW

### 2.1 Optimizing the Energy Flow

Optimization of the energy flow in different household and microgrid energy systems is a widely investigated research area. The key questions addressed in the related papers are rather similar: when and how to charge/discharge batteries or buy/sell electricity in order to maintain the operation of the energy system at minimum cost or with maximal profit. Nevertheless, the applied modeling and solution approaches vary. Many contributions assume that the future energy consumption and production of the system is fully known, and hence, apply a deterministic model. For instance, Vašak et al. (2011a) consider the problem of power flow optimization in an experimental microgrid. The total energy cost is minimized subject to basic technological constraints on power flows and the state of charge in the different storages, using a linear programming (LP) formulation.

Gupta et al. (2011) propose a similar model for a complex hybrid energy system. A rough-cut mathematical model

of each component is given, focusing on the conservation laws (e.g., power flow) and static technological limits in a steady-state system. The model minimizes the total operating cost on a finite, discrete time scale, with several additional objectives, such as minimizing the frequency of diesel generator starts/stops. A time-indexed mixed-integer linear programming model is proposed, but the model is solved using custom dispatching rules. Clastres et al. (2010) present a mathematical model for calculating the optimal operation of a domestic energy system with PV production. A two-step approach is taken: first, an operating plan is computed for the next 24 hours, which determines the schedule of buying and selling electricity from/to the grid, with the goal of maximizing the profit. The resulting active power bid is submitted to the distribution system operator. The second step is the real-time adjustment of the plan to the realization, with the objective of fulfilling the bid.

Among the contributions that assume a probabilistic prediction, Livengood and Larson (2009) investigate the optimal control of electricity usage in a residential or small office environment. They introduce a device called Energy Box that runs as a 24/7 background processor and controls all appropriate appliances in the building, such as refrigerators, water heaters, air conditioners, etc. A probabilistic weather and tariff forecast is assumed on a finite discrete-time scale. A stochastic dynamic programming approach is applied to compute an optimal energy management policy. Zavala et al. (2009) propose an on-line stochastic optimization approach to operate integrated energy systems based on detailed weather forecast data. It is shown that pure reactive strategies (those that disregard weather forecasts) lead to higher operating costs, while employing a weather forecasting model and model predictive control with stochastic optimization can result in 18% cost reduction. The developed system includes a weather forecasting method augmented with a Gaussian process uncertainty model. Constantinescu et al. (2011) take a similar approach to the problem of controlling the production/distribution of a set of thermal power plants in order to compensate for the uncertain production of wind power plants. An integrated model is presented where a probabilistic estimate of wind power production is given using a Numerical Weather Prediction model enhanced with an ensemble-based uncertainty quantification strategy. A stochastic programming formulation is applied.

### 2.2 Predicting PV Production

A key input data for controlling renewable energy systems is the prediction of energy production and consumption. While the prediction of the grid load has been a widely studied problem, production forecast became of interest with the spreading use of renewable energy: fossil and nuclear plants were designed to generate electricity in a stable and controllable way, however, predicting renewable energy production on a short-term horizon of 24-48 hours is considered to be a serious challenge as of now. Marquez and Coimbra (2012) classify short-term PV production prediction models into the following main categories:

- *Clear-sky models* have a single input, the cosine of the solar zenith angle for a given point in time. These

models assume that no meteorological phenomena reduce solar irradiance, and, as the name suggests, work well on clear sky days. Their performance degrades significantly in cloudy weather.
- *Persistence models* assume that current meteorological conditions, e.g., the cloud cover, persist over time during the prediction horizon. Hence, they scale the clear sky estimation for the next point in time with the actual deviation from the estimate.
- *Autoregressive models* use machine learning (ML) methods to estimate future production from a time series of past production. While the standard autoregressive models rely solely on the time series, autoregressive models with exogenous inputs use additional inputs, for example, weather data.

The comparison of the prediction capabilities of several ML modes applied in (Marquez and Coimbra, 2012) showed that ML approaches achieved only marginal improvement compared to baseline clear sky or persistence models. This indicates that production forecast on a clear day is an easy task, while it is very hard under weather conditions changing widely. The problem of predicting the daily solar radiation using a time series approach and artificial neural networks is addressed in (Paoli et al., 2010). Vašak et al. (2011b) propose a complex stochastic method for predicting PV production, which involves analytical approaches as well as neural networks.

## 3. PROBLEM STATEMENT

### 3.1 System Architecture

The full-fledged E+grid architecture contains PV panels for energy production, inverters, multiple batteries for energy storage, charge controllers and a switch box, an ensemble of 100-200 intelligent LED luminaries, as well as a central control system (CCS). Luminaries are equipped with communication and smart local control systems, and they are dimmed according to motion sensor signals. The system is prepared to buy energy from and sell energy to the power grid, at variable tariff rates. Smart meters measure the energy flow between key system components. Recorded measurement data are transmitted to the CCS on a regular basis. CCS makes predictions of energy production and demand, and controls charging/discharging of the batteries. Under actual load conditions these decisions determine the operation of the switch box: if demand of lighting services cannot be covered by overall internal supply, electricity is taken from the grid, and *vice versa*, internal energy can also be directed towards the grid. The system's schematic architecture – highlighting only its energy related components – is presented in Figure 1. This figure shows also the key decision variables of the energy flow optimization problem (see also Table 1).

### 3.2 Robust Control of the Energy Flow

In this section, we formalize the problem of controlling the energy flow in a renewable energy system consisting of uncontrollable generators and loads, a battery, and a bi-directional grid connection. The notations used throughout the paper are summarized in Table 1.
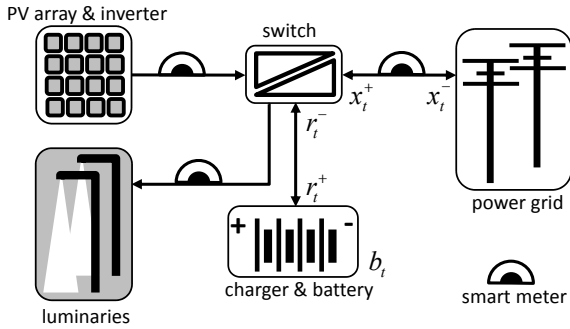
Fig. 1. Schematic architecture of the E+grid system.

An optimal control policy is searched for purchasing and selling electricity, as well as for charging and discharging the battery. The problem is solved on a finite horizon, consisting of a series of discrete time periods, $t = 1, ..., T$. It is assumed that expected future production $C_t^+$ and consumption $C_t^-$, as well as stochastically guaranteed (w.r.t. a given probability) lower bounds on production $\underline{C}_t^+$ and upper bounds on consumption $\overline{C}_t^-$ are available for the complete horizon. The time-varying electricity purchase and feed-in prices $Q_t^+$ and $Q_t^-$ are also known. The battery is characterized by its capacity $\overline{B}$, purchase price $P$, cycle life $L$, maximum charge and discharge rates $R^+$ and $R^-$, initial state of charge $b_0$, and the efficiency of charging $\eta$. The controller must be *robust* in the sense that the battery must be kept at a charge level that ensures an island-mode operation of at least $T_I$ time, even in the worst-case scenario defined by $\underline{C}_t^+$ and $\overline{C}_t^-$.

In the above setting, we would like to minimize the cost of running the energy system by finding the optimal electricity purchase rate $x_t^+$, grid feed-in rate $x_t^-$, as well as the battery charge rate $r_t^+$ and discharge rate $r_t^-$ for each of the time periods. The operating cost also includes the usage-dependent amortization of the battery. According to the common *linear cycle life* assumption, the life time of the battery is specified in the number of full charge cycles, and the deterioration caused by a partial charge cycle is proportional to the charge delivered. Hence, discharging the battery by $r_t^-$ incurs an amortization cost of $\frac{P}{\overline{B}L}r_t^-$.

It is noted that the overall operating cost can be negative, meaning that the system makes a positive profit by selling electricity. The following additional assumptions are made:

- If the energy system uses multiple electric phases, then the problems related to the individual phases can be solved separately. There is no constraint on the phase balance;
- There is no upper bound on the power purchased from, or fed into the grid, since the local load and feedback capabilities of the grid exceed the maximal power output of the microgrid;
- The electricity purchase price is never less than the feed-in price, i.e., $Q_t^+ \geq Q_t^-$, for each period $i$;
- The system is expected to survive at most one power cut with a maximum duration of $T_I$ within the planning horizon. $T_I$ is an integer multiple of the length of the time unit;

| Input Parameters | |
|---|---|
| $T$ | Number of time units |
| $Q_t^+$ | Electricity purchase price in period $t$ |
| $Q_t^-$ | Electricity feed-in price in period $t$ |
| $C_t^+$ | Predicted electricity production in period $t$ |
| $\underline{C}_t^+$ | Lower bound electricity production in period $t$ |
| $C_t^-$ | Predicted electricity consumption in period $t$ |
| $\overline{C}_t^-$ | Upper bound electricity consumption in period $t$ |
| $\overline{B}$ | Battery capacity |
| $P$ | Battery purchase price |
| $L$ | Battery cycle life |
| $R^+$ | Battery maximum charge rate |
| $R^-$ | Battery maximum discharge rate |
| $b_0$ | Battery state of charge in period 0 |
| $\eta$ | Efficiency of battery charging |
| $T_I$ | Required duration of island mode operation |
| Calculated Parameters | |
| $\underline{B}_t$ | Battery minimum state of charge in period $t$ |
| Variables | |
| $x_t^+$ | Electricity purchased from the grid in period $t$ |
| $x_t^-$ | Electricity fed into the grid in period $t$ |
| $r_t^+$ | Battery charge rate in period $t$ |
| $r_t^-$ | Battery discharge rate in period $t$ |
| $b_t$ | Battery state of charge at the end of period $t$ |

Table 1. Notations

- The initial state of charge, the capacity, and the maximum discharge rate of the battery are sufficient for satisfying the requirement on island mode operation (see details in Section 5.1).

The control problem is solved using a *rolling horizon* approach: a plan is computed for a finite horizon, $t = 1, ..., T$. The planned action is executed for the first time unit, $t = 1$, after which an updated plan is computed for a shifted horizon, $t = 2, ..., (T + 1)$, using revised input parameters, which involves updating the predictions.

*3.3 Prediction*

Proactive energy management techniques need to *predict* future energy production and consumption, in order to optimize their energy management policy. However, PV production is affected by weather conditions and consumption is influenced by human behavior, hence, both of them are uncertain and complicated to predict.

*Predicting PV Production* One of the key challenges to be faced for achieving robust control is to compute efficient predictions of the PV production for one day with a resolution of one hour. Not only a deterministic sequence containing an expected behavior, but suitable confidence regions should also be constructed, in order to allow designing a controller which is robust against power outages. More precisely, given a finite trajectory of the past production data $\{C_t^+\}_{t \leq 0}$ together with some a priori known physical characteristics of the system, e.g., GPS coordinates and model-type, we should construct $\{C_t^+\}_{t=1}^T$, a sequence of expected power production, where $T$ is the decision horizon; as well as lower confidence bounds $\{\underline{C}_t^+\}_{t=1}^T$ of the production, given probability $p$.

*Predicting Energy Consumption* Similarly, the energy consumption should also be predicted (expectations and

confidence bounds) with one hour resolution. The luminaries are equipped with movement sensors and ideally, we would also have past consumption data for prediction purposes. However, at the time of this experiment only movement data were available. Therefore, the problem was (i) to first estimate the expected cumulative movements in the area together with its upper confidence bounds; then (ii) to calculate estimated consumption data from this information. More precisely, (a) given a finite trajectory of the past movements $\{m_t\}_{t \leq 0}$, we should construct $\{m_t\}_{t=1}^{T}$, a sequence of expected future movements, where $T$ is the decision horizon; as well as upper confidence bounds $\{\overline{m}_t\}_{t=1}^{T}$, given probability $p$. Then, (b) we should estimate the expected consumption and its confidence bounds based on the estimated movements, more precisely, $C_t^- = g(m_t)$ and $\overline{C_t^-} = g(\overline{m}_t)$, for a suitable function $g$.

## 4. PREDICTING ENERGY PRODUCTION AND CONSUMPTION

In this section, we discuss how to calculate predictions based on past PV production and movement data, which are assumed to be available up to the decision time. In our experiments the original data had one-minute resolution.

### 4.1 Time series analysis

A *time series* is a sequence of data points, typically representing noisy measurements of a dynamical system observed at discrete time steps (Box et al., 1994). Fitting models to time series is one of the fundamental problems of *system identification*, a subfield of control theory and statistics (Ljung, 1999; Söderström and Stoica, 1989).

Discrete-time stochastic systems with exogenous components (inputs) coming from a parametrized family of systems can be typically written in a general form as

$$Y_t \triangleq f(\theta^*; \mathbb{Y}_{t-1}, \mathbb{U}_{t-1}, \mathbb{N}_{t-1}), \qquad (1)$$

where $Y_t$ is the output of the system at time $t$ and $\mathbb{Y}_{t-1}, \mathbb{U}_{t-1}, \mathbb{N}_{t-1}$ are the past outputs, inputs and noises affecting the system *up to and including* time $t-1$, i.e.,

$$\mathbb{Y}_{t-1} \triangleq (Y_{t-1}, Y_{t-2}, \dots), \qquad (2)$$

$$\mathbb{U}_{t-1} \triangleq (U_{t-1}, U_{t-2}, \dots), \qquad (3)$$

$$\mathbb{N}_{t-1} \triangleq (N_{t-1}, N_{t-2}, \dots), \qquad (4)$$

where $Y_k$, $U_k$ and $N_k$ are the output, the input, and the noise at time $k$, respectively. Constant $\theta^*$ is the unknown, true parameter (typically a finite dimensional vector) which needs to be estimated to determine the system.

Standard stochastic models include general LTI (linear time-invariant) systems, which can be formalized as

$$A(\theta^*; z^{-1}) Y_t \triangleq \frac{B(\theta^*; z^{-1})}{F(\theta^*; z^{-1})} U_t + \frac{C(\theta^*; z^{-1})}{D(\theta^*; z^{-1})} N_t, \quad (5)$$

where $A$, $B$, $C$, $D$ and $F$ are (finite) polynomials in $z^{-1}$, the backward *shift* operator (i.e., $z^{-1} Y_t = Y_{t-1}$), and $Y_t$, $U_t$, $N_t$ are as previously. Special cases of LTI systems include FIR (finite impulse response), OE (output error), MA (moving average), ARX (autoregressive exogenous)

and ARMAX (autoregressive moving average exogenous) models (Ljung, 1999; Söderström and Stoica, 1989).

In some cases linear models are not suitable to describe the observations. Standard nonlinear models include Hammerstein, Wiener and NARX (nonlinear autoregressive exogenous) systems (Ljung, 1999). Here, for the sake of brevity, we only describe NARX, which takes the form

$$Y_t \triangleq f(\theta^*; Y_{t-1}, \dots, Y_{t-q}, U_{t-1}, \dots, U_{t-s}) + N_t, \quad (6)$$

where $q$, $s$ are the orders of the system and $f$ is typically a nonlinear function (wavelet, neural network, etc.).

Having selected a model class, the problem of parametric identification is that given a (usually finite) *realization* $D_n$ of the inputs and the outputs up to time $n$, more precisely

$$D_n \triangleq (y_n, y_{n-1}, \dots, u_{n-1}, u_{n-2}, \dots), \qquad (7)$$

a parameter value, $\hat{\theta}_n$ should be found which satisfies

$$\hat{\theta}_n \in \underset{\theta \in \Theta}{\arg\min}\, \varepsilon(\theta, D_n), \qquad (8)$$

where $\varepsilon(\theta, D_n)$ is an *error function*, which describes how well the model fits to the data.

The error function usually has an additive structure, i.e.,

$$\varepsilon(\theta, D_n) \triangleq \sum_t \hat{\varepsilon}_t(\theta, D_t) = \sum_t w_t\, d(y_t, f(\theta, D_t)), \quad (9)$$

where $\{w_t\}$ are weights, $d(\cdot, \cdot)$ is a distance measure and $f(\theta, D_t)$ is defined as, for example,

$$f(\theta, D_t) \triangleq f(\theta; y_{t-1}, y_{t-2}, \dots, u_{t-1}, u_{t-2}, \dots, 0, 0, \dots).$$

A typical choice for $d(\cdot, \cdot)$ is to use $\| y_t - f(\theta, D_t) \|^2$, where $\| \cdot \|$ is the Euclidean norm, which provides the well-known (weighted) least-squares error criterion.

### 4.2 Predicting PV Production

We fitted several dynamical models to the available PV data. The measured quantities were the PV current (A) and PV voltage (V), from which we calculated PV power (voltage × current). We *preprocessed* the data by removing outliers (corrupted measurements), normalized the data and averaged it in windows with one hour length. Averaging helped to decrease the variance and hence to achieve a better signal-to-noise-ratio (SNR).

Since clear-sky type estimations were available, we also applied exogenous models and used the normalized (deterministic) clear-sky predictions as an input.

Several experiments were performed during which stochastic models were fitted to the available PV time series. The experimental results part contains six of those models which allowed multi-step predictions and produced the best results. These are: autoregressive (AR), autoregressive moving average (ARMA), autoregressive exogenous (ARX), state space (STATE), Box-Jenkins (BJ), and nonlinear autoregressive exogenous (NARX) models. The above models were used with several settings, e.g., experiments with different orders were performed. The achieved models were compared according to their prediction errors on different horizons (1-hour, 3-hour, 6-hour, 12-hour and 24-hour) and validated on an independent (not used during the estimation) dataset (see also Section 6.1).

The best performing model for the PV production data was a NARX (6) type of system with combined *affine* and *wavelet* (Goswami and Chan, 2011) type nonlinearities. In this case, function $f$ takes the form

$$f(\theta; x) \triangleq a + b^{\mathrm{T}}x + c(w, x), \qquad (10)$$

where $\theta = (a, b_1, \ldots, b_{q+s}, w_{1,1}, w_{1,2}, \ldots)$ and $c(w, x)$ is a wavelet function with weights $w$, formally

$$c(w, x) \triangleq \sum_{i=1}^{k}\sum_{j=1}^{m} w_{i,j}\psi_{i,j}(x), \qquad (11)$$

where $k$, $m$ are the orders and $\{\psi_{i,j}(\cdot)\}$ are constructed from a suitable mother wavelet (Goswami and Chan, 2011). The orders of the best NARX model were $q = 13$, $s = 2$, cf. (6), while $k$ and $m$ were automatically selected.

Having identified the NARX model, the noise variance is also estimated. For simplicity, the noise is then assumed to be zero-mean Gaussian with the estimated variance. The mean trajectory and the lower confidence bounds are calculated by using Markov Chain Monte Carlo (MCMC) simulations, i.e., the mean prediction is the average of $N$ randomly generated trajectories ($N$ is usually 1000 or 10 000) using the recent sensor measurements as initial conditions (and clear-sky predictions as inputs for exogenous models), while the lower confidence bound for each interval is the smallest number which is larger or equal to at least $1 - p$ portion of the estimations for that interval.

### 4.3 Predicting Energy Consumption

Energy consumption was predicted using data from several motion sensors. The movement data was also preprocessed: the data of all the sensors were aggregated, normalized and averaged in one hour-wide windows.

Based on the past measurements an average behavior was calculated for each hour of the day and it was used as the input for the exogenous models (ARX, BJ, STATE, NARX). The same type of models were applied as in the previous case and compared for different prediction horizons as well as validated on an independent dataset.

According to our experiments, Box-Jenkins (BJ) type models gave the best predictions for movement data from the models we tried. BJ models are LTI, cf. (5), but without polynomial $A$. The orders of the best BJ model were $7, 6, 6, 3$ for polynomials $B, C, D, F$, respectively.

Having identified the model, we again estimated the noise variance and treated the noise as a zero-mean Gaussian. Then, we used MCMC simulations to estimate the mean behavior of the system as well as to get a lower confidence bound for the future movements for each future hour.

Given movement predictions, the expected consumption is calculated as follows. It is assumed that consumption is composed of the constant consumption of the energy system, and the variable consumption of the luminaries proportional to motion intensity, masked by the public lighting calendar. For the latter component, a saturation-type function was used reflecting that a detected movement stimulates several luminaries, but the set of luminaries activated by different movements can overlap. The ratio of the constant and the variable consumption components is about $1 : 5$ during lighting periods.

## 5. ROBUST CONTROL OF THE ENERGY FLOW

### 5.1 Computing the Required Battery Charge

Island mode operation of the system can be ensured by maintaining the appropriate state of charge in the battery. While the prescribed duration of island mode operation is a given constant, the required state of charge varies over time, depending on the future consumption and production. Robust control can be achieved by considering upper bound consumption and lower bound production. In particular, the state of charge of the battery decreases by $C'_t = \max(\overline{C_t^-} - \underline{C_t^+}, -R^+)$ if the system operates in island mode in time unit $t$. A decrease of $-R^+$ (i.e., an increase of $R^+$) arises at the time of production peaks with low consumption, when the battery charge rate limit is hit. Hence, in order to maintain island mode operation in the time interval $[t, t + T_I - 1]$, the battery's state of charge must be at least $\underline{B}_{t-1} = \sum_{u=t}^{t+T_I-1} C'_u$ at the end of the previous time unit, which defines a minimum state of charge constraint in the robust control problem.

Obviously, the island mode operation expectations can be respected only if the battery parameters are adequate, i.e., $\overline{B} \geq \underline{B}_t$, $b_0 \geq \underline{B}_0$, and $R^- \geq C'_t$ $\forall t$.

### 5.2 LP Formulation of the Control Problem

An LP formulation of the control problem is presented below for computing a plan on a finite horizon within one computational step of the rolling horizon scheme.

$$\text{minimize} \quad \sum_{t=1}^{T}\left(Q_t^+ x_t^+ - Q_t^- x_t^- + \frac{P}{BL}r_t^-\right) \qquad (12)$$

subject to

$$C_t^+ - C_t^- + x_t^+ - x_t^- = r_t^+ - r_t^- \quad \forall t \qquad (13)$$
$$\eta r_t^+ - r_t^- = b_t - b_{t-1} \qquad \forall t \qquad (14)$$
$$\underline{B}_t \leq b_t \leq \overline{B} \qquad \forall t \qquad (15)$$
$$0 \leq r_t^+ \leq R^+ \qquad \forall t \qquad (16)$$
$$0 \leq r_t^- \leq R^- \qquad \forall t \qquad (17)$$
$$0 \leq x_t^+, x_t^- \qquad \forall t \qquad (18)$$

The objective (12) is minimizing the total cost, which consists of the difference of the energy purchased and sold, plus the amortization of the battery. Constraint (13) encodes the energy balance in the system. Equality (14) relates the state of charge of the battery to the charge/discharge rate, given the loss on the battery. Constraints (15-17) define the range of the variables, with respect to the pre-computed lower and upper bounds on the battery state of charge and the charge/discharge rates.

## 6. EXPERIMENTAL EVALUATION

In this section, we first show experimental results of fitting dynamical models to PV production and movement data, then present an illustrative example demonstrating the controlled energy system in summer and winter days.

| Estimation Data | | | | | | |
|---|---|---|---|---|---|---|
| Model | Orders | 1-Stp | 3-Stp | 6-Stp | 12-Stp | 24-Stp |
| AR | 5 | 65,05 | 27,99 | 6,78 | -2,45 | -8,35 |
| ARMA | 7 5 | 65,84 | 31,40 | 7,21 | 5,89 | 5,70 |
| ARX | 5 4 | 66,73 | 37,84 | 16,89 | 17,97 | 18,33 |
| BJ | 7 5 5 3 | 68,07 | 41,73 | 27,14 | 31,55 | 44,25 |
| STATE | 5 | 57,69 | 20,34 | 13,53 | 10,99 | 28,15 |
| NARX | 2 13 | 71,27 | 42,77 | 31,19 | 27,10 | 27,21 |
| Validation Data | | | | | | |
| Model | Orders | 1-Stp | 3-Stp | 6-Stp | 12-Stp | 24-Stp |
| AR | 5 | 55,57 | 6,21 | -12,93 | -14,47 | -12,34 |
| ARMA | 7 5 | 56,97 | 10,47 | 2,04 | 0,47 | 1,16 |
| ARX | 5 4 | 58,25 | 20,49 | 16,77 | 14,22 | 15,24 |
| BJ | 7 5 5 3 | 57,87 | 22,56 | 23,86 | 21,93 | 22,83 |
| STATE | 5 | 56,33 | 15,73 | 6,86 | 3,08 | 25,10 |
| NARX | 2 13 | 62,94 | 26,03 | 23,72 | 25,04 | 26,48 |

Table 2. Multi-step prediction of PV energy production data. The input for the exogenous models (ARX, BJ, STATE, NARX) was the prediction of the applied clear-sky type model.

| Estimation Data | | | | | | |
|---|---|---|---|---|---|---|
| Model | Orders | 1-Stp | 3-Stp | 6-Stp | 12-Stp | 24-Stp |
| AR | 7 | 61.75 | 38.12 | 26.10 | 25.51 | 35.55 |
| ARMA | 8 5 | 63.33 | 35.54 | 22.58 | 23.45 | 27.01 |
| ARX | 7 5 | 62.21 | 35.71 | 27.29 | 26.70 | 36.16 |
| BJ | 7 6 6 3 | 73.17 | 47.39 | 36.27 | 31.84 | 52.56 |
| STATE | 5 | 59.62 | 34.12 | 21.66 | 20.99 | 37.23 |
| NARX | 1 4 | 67.58 | 43.63 | 26.86 | 14.61 | 16.68 |
| Validation Data | | | | | | |
| Model | Orders | 1-Stp | 3-Stp | 6-Stp | 12-Stp | 24-Stp |
| AR | 7 | 45.99 | 25.25 | 17.57 | 8.59 | 20.40 |
| ARMA | 8 5 | 45.73 | 26.88 | 16.91 | 7.69 | 20.82 |
| ARX | 7 5 | 46.13 | 27.05 | 18.64 | 12.70 | 30.85 |
| BJ | 7 6 6 3 | 53.06 | 32.33 | 36.86 | 39.18 | 49.67 |
| STATE | 5 | 44.21 | 24.12 | 17.81 | 17.06 | 35.55 |
| NARX | 1 4 | 42.84 | 18.97 | 17.28 | -8.31 | 18.98 |

Table 3. Multi-step prediction of movement data. The input for the exogenous models (ARX, BJ, STATE, NARX) was the averaged amount of past movements in the same hours.

## 6.1 Model Fitting

Several experiments were performed in MATLAB on the preprocessed data. Tables 2 and 3 illustrate the multi-step prediction capabilities of the fitted models, where each "step" corresponds to one hour of averaged data. Each type of model was evaluated with various parameters and model orders, but the tables only contain the fit values of one selected ("best") choice for each model. The vector describing the orders of the models contain the orders in the following structure: AR (autoregressive part), ARMA (autoregressive; moving average), ARX (autoregressive; exogenous), BJ (polynomials $B, C, D, F$), STATE (the same dimension is assumed for the output, input and noise vectors), NARX (autogressive part; exogenous part). The nonlinariy of NARX was as in (10). State space models were identified using a subspace method (N4SID), while prediction error methods were used for the other cases.

The fit values shown in the tables, and denoted by $F(\cdot, \cdot)$ below, are calculated from the (deviation normalized) root-mean-square error. More precisely, they satisfy

$$1 - \frac{F(y, \hat{y})}{100} = \frac{\|y - \hat{y}\|}{\|y - \bar{y}\|} = \frac{\sqrt{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}}{\sqrt{\sum_{t=1}^{n}(y_t - \frac{1}{n}\sum_{k=1}^{n}y_k)^2}}, \quad (19)$$

where $y$ and $\hat{y}$ are $n$-dimensional vectors of the observed and estimated outputs of the system, respectively, $\bar{y}$ is the sample average, and $\|\cdot\|$ is the Euclidean norm.

Hence, $F(y, \hat{y})$ is a real number in $(-\infty, 100]$, where bigger numbers indicate better fits. It is clear from the definition that $F(y, y) = 100$ and $F(y, (\bar{y}, \ldots, \bar{y})^{\mathrm{T}}) = 0$, i.e., the fit value of the sequence itself is 100 and if we used the average as a (time-independent) estimator, we would get zero.

The results of Table 2 indicate that NARX models with wavelet type nonlinearities provide a good fit for the purpose of multi-step PV production prediction, where the exogenous inputs come from a clear-sky model.

Table 3 shows the results for the same models (but with different orders), in case we wanted to predict movement data. The experiments are indicative of the phenomenon that Box-Jenkins type of models can well predict averaged movement data, if enough observations from the immediate past are given as well as the typical movement in the corresponding hours are known and provided as input.

We used the identified NARX and BJ models in our experiments to get production and consumption predictions, respectively. The confidence probability of the (lower and upper) bounds was 95 % in the experiments below.

## 6.2 Illustrative Examples: Controlling the Energy Flow

The proposed LP model for the robust control problem has been implemented in FICO Xpress 7.2. Computational experiments have been performed with data originating from the above presented experimental lighting microgrid. The PV generators in the system have been sized to achieve a slightly positive energy balance over a one-year horizon, and lead-acid batteries ensure 3 hours of island mode operation under almost any consumption and production scenario. A variable energy tariff scheme has been adapted based on data from an Australian public utility company, with $Q_t^+$ varying between 14.08 and 24.86 c/kWh in three steps and $Q_t^- = 7.5$ c/kWh.

Figures 2 and 3 present the optimal energy flow in a summer and in a winter scenario, respectively. Each diagram shows the corresponding battery charge rate (Chrg, positive means charging the battery), grid traffic (Grid, positive means buying electricity), and battery state of charge (BatSoc) curves over the day. In the summer scenario (Fig. 2), early morning consumption is covered from the battery. Increasing PV production during the day is sold to the grid until ca. 14:00, when the system starts to charge the battery, in order to cover the needs of a potential island mode operation during the evening consumption peak. At the end of the planning horizon, surplus energy in the battery is sold to the grid. However, the actual execution of this action is unlikely under the rolling horizon approach, since this section of the plan will be recomputed multiple times. The optimal control in a winter day (Fig. 3) is different due to the fact that consumption dominates
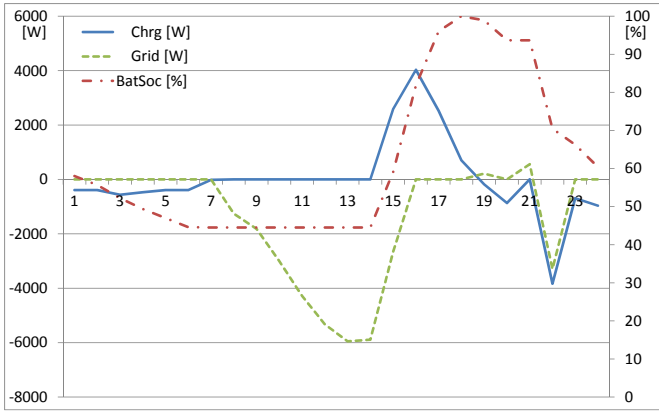
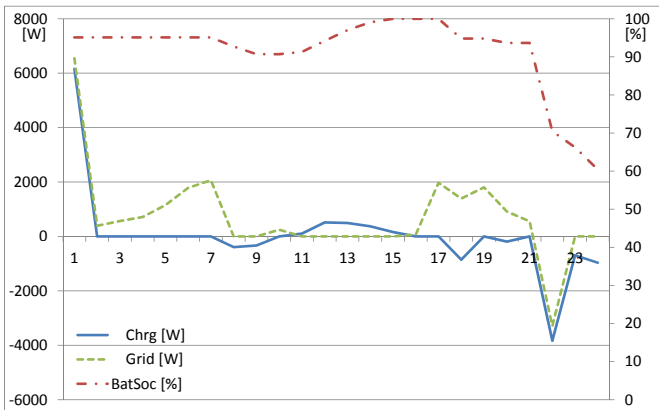Fig. 2. Energy flow in the system on a summer day.



Fig. 3. Energy flow in the system on a winter day.

production, and the battery must be kept at nearly full charge throughout the day. The solution of the proposed small-size LP model takes negligible computation time.

## 7. CONCLUSION AND FUTURE RESEARCH

The paper proposed algorithms for controlling the energy flow in an experimental microgrid for energy-positive street lighting. The algorithms trade with electricity in order to maximize profit subject to a given variable energy tariff, while achieve robustness by maintaining sufficient charge in the batteries for surviving potential power outages. A rolling horizon controller was presented where each time-step involves solving an LP problem. The key inputs of optimization, the prediction of future energy production and consumption, are computed by fitting stochastic models to historic production and movement data. It was demonstrated that sufficiently precise prediction can be achieved on a one-day horizon using NARX and BJ models. As illustrative examples, the behavior of the controlled system was presented on a typical winter and summer day. The results illustrate the possible benefits of smart renewable energy systems in public lighting applications.

The physical E+grid system is currently under construction in cooperation between GE Hungary Ltd, the Institute for Technical Physics and Materials Science, the Budapest University of Technology and Economics, and the Institute for Computer Science and Control. Future work will focus on the improvement of our production and consumption estimators with specialized statistical models as well as improving the controller adopting novel stochastic control and machine learning principles.

## REFERENCES

Box, G.E.P., Jenkins, G.M., and Reinsel, G.C. (1994). *Time Series Analysis*. Prentice-Hall.

Clastres, C., Ha Pham, T.T., Wurtz, F., and Bacha, S. (2010). Ancillary services and optimal household energy management with photovoltaic production. *Energy*, 35(1), 55–64.

Constantinescu, E., Zavala, V., Rocklin, M., Lee, S., and Anitescu, M. (2011). A computational framework for uncertainty quantification and stochastic optimization in unit commitment with wind power generation. *IEEE Transactions on Power Systems*, 26(1), 431–441.

Goswami, J.C. and Chan, A.K. (2011). *Fundamentals of Wavelets: Theory, Algorithms, and Applications*. Wiley.

Gupta, A., Saini, R.P., and Sharma, M.P. (2011). Modelling of hybrid energy system–part I: Problem formulation and model development. *Renewable Energy*, 36(2), 459–465.

Livengood, D. and Larson, R. (2009). The energy box: Locally automated optimal control of residential electricity usage. *Service Science*, 1(1), 1–16.

Ljung, L. (1999). *System Identification: Theory for the User*. Prentice-Hall, 2nd edition.

Marquez, R. and Coimbra, C.F.M. (2012). Comparison of clear-sky models for evaluating solar forecasting skill. In *Proc. of the World Renewable Energy Forum 2012*, 4443–4449.

Paoli, C., Voyant, C., Muselli, M., and Nivet, M.L. (2010). Forecasting of preprocessed daily solar radiation time series using neural networks. *Solar Energy*, 84(12), 2146–2160.

Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall.

Vašak, M., Baotić, M., and Perić, N. (2011a). Deterministic power flow optimization in an experimental microgrid. In *Proceedings of the 5th Global Conference on Power Control & Optimization (PCO 2011)*.

Vašak, M., Gulin, M., Čeović, J., Nikolić, D., Pavlović, T., and Perić, N. (2011b). Meteorological and weather forecast data-based prediction of electrical power delivery of a photovoltaic panel in a stochastic framework. In *MIPRO 2011, Proc. of the 34th Int. Conv.*, 733–738.

Zavala, V.M., Constantinescu, E.M., Krause, T., and Anitescu, M. (2009). Weather forecast-based optimization of integrated energy systems. Technical report, Argonne National Laboratory.