



System Identification of a Non-Uniformly Sampled Multi-Rate System in Aluminium Electrolysis Cells

H. Viumdal¹ S. Mylvaganam² D. Di Ruscio²

¹ *Tel-Tek, N-3918 Porsgrunn, Norway and Telemark University College, Faculty of Technology, N-3918 Porsgrunn, Norway. E-mail: hakon.viumdal@hit.no*

² *Telemark University College, Faculty of Technology, N-3918 Porsgrunn, Norway. E-mail: saba.mylvaganam@hit.no and david.di.ruscio@hit.no*

Abstract

Standard system identification algorithms are usually designed to generate mathematical models with equidistant sampling instants, that are equal for both input variables and output variables. Unfortunately, real industrial data sets are often disrupted by missing samples, variations of sampling rates in the different variables (also known as multi-rate systems), and intermittent measurements. In industries with varying events based maintenance or manual operational measures, intermittent measurements are performed leading to uneven sampling rates. Such is the case with aluminium smelters, where in addition the materials fed into the cell create even more irregularity in sampling. Both measurements and feeding are mostly manually controlled. A simplified simulation of the metal level in an aluminium electrolysis cell is performed based on mass balance considerations. System identification methods based on Prediction Error Methods (PEM) such as Ordinary Least Squares (OLS), and the sub-space method combined Deterministic and Stochastic system identification and Realization (DSR), and its variants are applied to the model of a single electrolysis cell as found in the aluminium smelters. Aliasing phenomena due to large sampling intervals can be crucial in avoiding unsuitable models, but with knowledge about the system dynamics, it is easier to optimize the sampling performance, and hence achieve successful models. The results based on the simulation studies of molten aluminium height in the cells using the various algorithms give results which tally well with the synthetic data sets used. System identification on a smaller data set from a real plant is also implemented in this work. Finally, some concrete suggestions are made for using these models in the smelters.

Keywords: Height measurements, aluminium electrolysis, system identification

1 Introduction

Many industrial processes involve systems where two or more physical processes have strongly differing temporal characteristics, i.e large differences in time constants. Simulations of such systems may be solved without excessive use of small time steps, by using two different models, one for the rapid variations, assuming the slow varying process to be constant, and another model for the slow variations, where the rapid

varying process is neglected. This is a possible way of attacking the system identification problem of such systems, also known as stiff systems. In some cases a model should be determined based upon already existing data from industrial processes. In such cases the researcher may be confronted with lacking data samples or outliers that should be deleted, hence making the system identification problem some more complicated. In addition multi-rate system widely exists in chemical process industries, typically with a slower sampling

Table 1: Nomenclature

Symbol	Quantity
A	Cross section of aluminium
CE	Current Efficiency
F	Faraday constant
I	Line current
K	Kalman gain matrix
M	Sampling difference factor between T_f and T_s
M_{Al}	Molecular mass of aluminium
N	Number of data samples
P	Covariance matrix
Q	Charge
T_f	Fast sampling time
T_s	Slow sampling time
V	Volume of aluminium
V_t	Criterion function
h	Height of molten aluminium
b_1 and b_2	Parameters in mechanistic model
f_s	Sampling frequency
f_m	System frequency
$\Delta\hat{h}$	Predicted change of Aluminium height
\bar{h}	Aluminium height prediction
h	Aluminium height measurement
k	Discrete time
m	Mass of molten aluminium
n_i	Discrete sampling instant of the i -th output measurement
t	Continuous time
w_1	Weight put on the model
w_2	Weight put on the measurements
z	Charge number of an ion
α	Forgetting factor
Δt	Sampling time of input measurements
ϵ	Prediction error
θ	Parameter vector of OLS models
λ	Parameter weight in OLS model
ρ	Density
Φ	Measurement vector for fast sampling
ψ	Measurement vector for slow sampling
ACD	Anode Cathode Distance
CCA	Canonical Correlation Analysis
DSR	Deterministic and Stochastic system identification and Realization
MDSR	Multiple time series DSR
MSE	Mean Square Error
NMSE	Normalized Mean Square Error
NRMSE	Normalized Root Mean Square Error
NUSM	Non-Uniformly Sampled Multi-rate syst.
OLS	Ordinary Least Squares
PEM	Prediction Error Method
ROLS	Recursive Ordinary Least Squares
SNR	Signal to Noise Ratio
SSM	State Space Model

rate for the outputs, compared to the inputs of the system. Some of the measurements may also be manually performed, causing rare and intermittent sampling instants. All these factors complicate the system identification process. As the standard methods of system identifications are based upon equidistant sampling instants, these methods cannot be directly applied for such data sets.

During periods without measurement data, gathered data will consist of several data segments. Simply concatenating the data segments may lead to false transients in their connection points. A more consistent option is to calculate model parameters for each of the segments, using the identical model structure. Finally, the segment estimates should somehow be merged into the resulting model of the system. In [Ljung \(1999\)](#) the parameter estimates of the segments are given weight according to their estimated inverse covariance matrices. In 1957 Kranc introduced a method of replacing multi-rate sampled systems with single-rate models, using z transform methods ([Kranc, 1957](#)). In Kranc's approach the sampling and updating instants are synchronized but operate with different sampling periods. The sampling periods could all be expressed as unit fractions of an overall sampling rate T , later known as the frame period ([Sheng et al., 2002](#)). T is the least common multiple of the periods of the sampling and updating pattern of the system. Within each frame period T , there is one or more "sub"-sampling periods. The most common case is the system in which the inputs of the system are sampled on a higher rate than the output. The extracted single-rate model of the multi-rate system is designed with the sampling rate of the frame period T . This method is therefore known as the lifting technique ([Sheng et al., 2002](#)), as sampling periods are lifted to a higher and mutual level for all the variables. A dual-rate modeling case study on continuous catalytic reforming in the oil industry, using lifting technique, is presented in [Li et al. \(2003\)](#). A least squares method was used by Lu and Fisher to estimate the inter-sample outputs, for applications where the outputs are sampled at a slower rate, compared to the inputs ([Lu and Grant Fisher, 1989](#)). In [Li et al. \(1999\)](#) system identification of multi-rate systems, using subspace methods are discussed. In particular the systems considered are systems where the sampling rate of the input variables are n times faster than the sampling rate of the output variables, where n is an integer.

The lifting method has been used to generate a minimum variance predictor with the fast sampling rate of the input variables. This predictor shows enhanced performance compared to a similar predictor using the slow output sampling rate. The controllability and ob-

servability of lifted systems are discussed in Wang et al. (2004) and Ding et al. (2009), in addition to a recursive auxiliary least square model based upon a dual-rate system. A lifted state space model with an implemented Kalman filter for Non-Uniformly Sampled Multi-rate (NUSM) systems is presented in Li et al. (2008).

A simplified simulated model of the aluminium metal height in a pre-baked aluminium electrolysis process is used as a representative example throughout this paper. Aluminium electrolysis processes applying pre-baked anodes are the main procedure to manufacture primary aluminium to day. This is an old method heavily based upon frequent operator interventions involving a plethora of manual operations involving materials feeds, handling of measurands, removing crusts and excessive alumina of cryolite. Although there has been a considerable improvement by the introductions of the automated control system and some few automatically sampled measurements, the majority of the measurements are still manually performed. Regular sensor applications are normally not suitable for this electrolysis process due to the high temperature, the invasive corrosive environment of the bath and the generally harsh environment of the plants. In this paper the focus will be upon estimating the molten aluminium height, also called the metal height. By improving the monitoring of this vital variable, the process can be stabilized at more optimal conditions, which are beneficial to the environmental and economical performance of the whole plant. In other words undesired variations of the process should be reduced or eliminated.

The main focus in this paper is on using system identification methods on non-uniformly sampled multi-rate systems, and in particular on a simplified model of the aluminium electrolysis cell. The main contributions in this paper are as itemized in the following:

- Deriving a simplified model of the metal height in an aluminium electrolysis cell.
- Deriving and testing a Prediction Error Method (PEM) (Ljung, 1999) model for a non-uniformly sampled multi-rate system.
- Reducing the impact of the computer precision on updating the covariance matrix of the Multiple Input Single Output (MISO) system, in the Recursive Ordinary Least Square (ROLS) method (Di Ruscio, 2001), by rearranging the updating equation.
- Testing the combined Deterministic and Stochastic Realization (DSR) algorithm (Di Ruscio, 1996) on a data set with multi-rate sampling.

- Testing the Multiple time series Deterministic and Stochastic Realization (MDSR) algorithm (Di Ruscio, 1997) on a data set with multi-rate sampling and multiple time series.
- Comparing different system identification methods on the simplified discrete model of the metal height, showing saw-tooth-shaped behavior.

The rest of the paper is organized as follows. In Sec. 2 we derive a simplified discrete model of the metal height in an aluminium electrolysis cell, which represents the synthetic system that is to be identified throughout the rest of the paper. In Sec. 3 the black, white and gray model approaches in system identification are emphasized. Data sampling for system identification methods are discussed in Sec. 4. In Sec. 5 prediction error methods are described for use on the given synthetic system, assuming non-uniformly multi-rate sampling. In Sec. 6 the sub space method DSR and a corresponding version for multiple series are tested, assuming synchronous but multi-rate sampling. Sampled data from a real plant is utilized in Sec. 7. The results from the different methods are compared and analyzed in Sec. 8, whereas this paper is summarized and concluded in Sec. 9.

2 Simplified model of the metal height

There are several model approaches describing the level and shape of the metal/bath interface in aluminium electrolysis cells. The interface is not totally flat, but consists of gravitational standing waves, and unstable propagating waves (Kurenkov et al., 2004). The waves are generated by the high electrical and magnetic fields (Lorentz forces), and the Kelvin-Helmholtz instability of the mean flow. Under stable conditions, the interface waves are estimated to have an amplitude of about 0.5 mm, at fixed positions (for a 500 kA cell), but these waves can increase by variations in stability affecting parameters, like the Anode Cathode Distance (ACD) and the height of the liquid metal (Bojarevics and Pericleous, 2006, 2009). Larger deviations are present along the horizontal axis, due to the Lorentz forces. In this work we will consider a fixed horizontal position of the metal/bath interface, and are focusing on its mean value.

Consider the sketch of an aluminium electrolysis cell of Figure 1. A simplified model of the metal height, i.e. h , will be derived in this section, based upon the aluminium mass balance.

Due to the spatial mass balance of molten aluminium the time gradient of mass equals

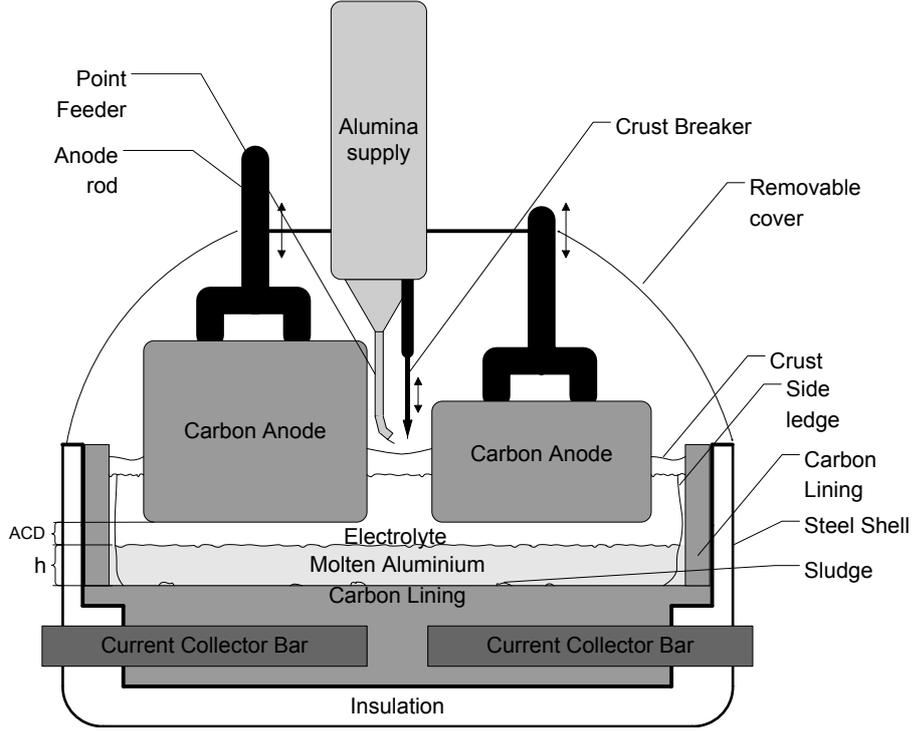


Figure 1: Schematic drawing of an aluminium electrolysis cell with the different structures in it forming the system under consideration. Our model focuses on temporal variations of molten metal height h .

$$\frac{dm}{dt} = m_{gen}(t) + m_{in}(t) - m_{out}(t). \quad (1)$$

where $m_{gen}(t)$ is the instant mass rate of molten aluminium generated inside the cell, whereas $m_{in}(t)$ and $m_{out}(t)$ represents the mass flow of molten aluminium into and out of the electrolysis cell, respectively. For small values of Δt , a good approximation of the mass balance is the discrete model

$$\Delta m(k) = \Delta t(m_{gen}(k) + m_{in}(k) - m_{out}(k)), \quad (2)$$

where $k = \{1, 2, 3, \dots\}$ represents the discrete time, whereas Δt is the sampling time. Normally, there are no inflow of molten aluminium to the system, as the aluminium is generated within the electrolysis cell, which is regarded as the system of interest.

The generated mass of aluminium per time unit is given by Faraday's laws of electrolysis (Grjotheim, 1993);

$$m_{gen}(k) = \frac{\Delta Q(k) \cdot CE \cdot M_{Al}}{F \cdot z \cdot \Delta t} \quad (3)$$

where CE is the current efficiency of the cell which in the following simulation is assumed to be constant for all t . $M_{Al} = 26.98g/mol$ is the molecular mass of aluminium, F the Faraday constant, whereas z is

the number of electrons involved in the electrode reaction generating one single aluminium atom from a aluminium-ion. The charge transferred from time $k-1$ to k , i.e. $\Delta Q(k)$ is given by the time integral of the current

$$\Delta Q(k) = \int_{k-1}^k I(\tau) d\tau \quad (4)$$

Assuming constant value of the current within each Δt , i.e. zero-order hold, the generated mass of aluminium per time unit equals

$$m_{gen}(k) = \frac{CE \cdot M_{Al}}{F \cdot z} I(k), \quad (5)$$

There is usually no inflow of molten aluminium to the process, hence

$$m_{in}(k) = 0 \quad (6)$$

The outflow of molten aluminium per time unit relates to the tapping proceeding. As the tapped mass of aluminium, and not its mass flow is measured in the tapping procedure, it is more convenient to put

$$m_{out}(k) = \Delta t \cdot \dot{m}_{out}(k), \quad (7)$$

in Eq. (2), hence assuming a constant flow rate within each time step.

Assuming a crisp interface between the electrolyte and the molten aluminium, and that all the generated molten aluminium will be located at the bottom of the cell, i.e. assuming no time delay due to transportation of aluminium in the electrolyte, and no precipitations of other substances within this area, and assuming a homogeneous pure aluminium, $\Delta m(k)$ equals

$$\Delta m(k) = \rho \cdot \Delta V(k). \quad (8)$$

Due to side ledge formation of frozen electrolyte at the side walls of the cell, the area of the horizontal section of molten aluminium varies with both height and time. Here the side ledge profile (Figure 1) is assumed to be straight vertical. Hence, the change in volume of the molten aluminium in one time step is, according to Figure 2, given by

$$\Delta V(k) = \Delta h(k) \cdot A(k) + h(k-1) \cdot \Delta A(k) \quad (9)$$

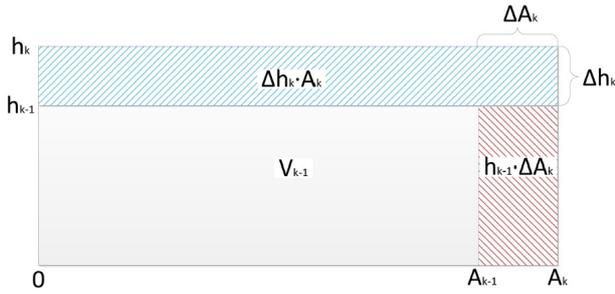


Figure 2: Schematic figure of the volume of the metal pad, showing change in volume of the molten metal described by the variables A and h .

Consequently, in one time step, the net mass aluminium accumulated is

$$\Delta m(k) = \rho (\Delta h(k) \cdot A(k) + h(k-1) \cdot \Delta A(k)). \quad (10)$$

Based upon Eq. 10, the change in metal height in one time step is given by

$$\Delta h(k) = \frac{\Delta m(k)}{\rho \cdot A(k)} - \frac{h(k-1) \cdot \Delta A(k)}{A(k)}. \quad (11)$$

Inserting Eq. (2), (5), (7) into Eq. (11), leads to this simplified model of the height of aluminium in the cell:

$$\Delta h(k) = \frac{1}{\rho \cdot A(k)} \left[\Delta t \frac{CE \cdot M_{Al}}{F \cdot z} I(k) - m_{out}(k) \right] - \frac{h(k-1) \cdot \Delta A(k)}{A(k)}, \quad (12)$$

The model in Eq. (12) is used for generating the synthetic data set simulating the aluminium height throughout this paper.

An even simpler model of the metal height is derived, by assuming that the area A is constant. Hence, the last term in Eq. (12) will disappear. By further collecting the constants into two constants, a very simple finite impulse response (FIR) model is formed

$$\Delta h(k) = b_1 \cdot \Delta t \cdot I(k) - b_2 \cdot m_{out}(k), \quad (13)$$

where

$$b_1 = \frac{CE \cdot M_{Al}}{\rho \cdot A(k) \cdot F \cdot z}, \quad (14)$$

and

$$b_2 = \frac{1}{\rho \cdot A(k)}. \quad (15)$$

This model is the basis of the PEM models used for the system identification problem in Sec. 5.

The assumptions used in this model are:

1. Homogeneous layers of molten cryolite (electrolyte) and molten aluminium.
2. Perfect crisp interface between the layers. No mixing of the materials in the different layers.
3. No transportation delay for any of the materials in the system.
4. The process variables are fixed for each sampling interval Δt .
5. A straight vertical side ledge profile.
6. Constant current efficiency.
7. No precipitates or impurities in the aluminium volume, only pure aluminium.
8. Constant density of the molten aluminium.

The aluminium electrolysis process requires a very stable energy balance, as the operation temperature of the electrolyte within the cell is close to its freezing point, so that the top surface of the bath is covered with frozen electrolyte, as well as the side ledges. The frozen electrolyte has a desired heat insulating effect, as well as it protects the side shell of the cell from corrosion. Simultaneously it is important that the active electrolyte is in molten condition. To maintain these steady conditions of the heat transfer, and hence the side ledge profiles, the variation of the metal height varies with only about some centimeters between maximum and minimum position of the interface. However, the total height of the aluminium layer is normally close to 20 cm.

2.1 Simulation of the molten metal height

A simulation of the aluminium height was performed using the simplified model derived in Eq. (12). Normal distributed noise with mean value $\mu = 0$ and standard deviation $\sigma = 0.75$ cm was added to the height measurement, and rounded to nearest half centimeter. This is a very high, but a realistic measurement uncertainty of the manually performed height measurements in aluminium electrolysis plants. The time of the daily performed tapping procedure is randomly varied from day to day, as seen in in Figure 3. In the simulation typical variations of the variables I and m_{out} are included. To make the simulation more dynamic, a variation of the horizontal cross sectional area A has been implemented as a sine-curve. The simplified model is simulated with $\Delta t = 5min$. This simulation approach will be used throughout this paper, examining different system identification methods.

3 Black, white and gray box models

There are several ways of categorizing different system identification methods. In this paper we use the box system to categorize the utilized algorithms, to have a better overview of the benefits and drawbacks of each method.

3.1 Introduction to model types

There are two main approaches to generate mathematical models of industrial processes. Mechanistic models, also called white box models, are merging several well-know physical relations between the different variables of concern to achieve a reasonable model of the process. With set of mechanistic model variables that are assumed to influence the total system, the outputs may be calculated with these models. Although the basic models are well established and may work excellent for small processes, the many assumptions and uncertainties induced by the often huge number of relations integrated in the model, the final simulation of the variables may not be satisfying. The deviations of the simulated variables from the real values, regarded as random noise is normally not white noise, and some of the noise might be eliminated by improving the models.

The optional approach of creating mechanistic models is a conglomerate of different empirical methods that all depend on sampled data from the particular system of interest. In that sense such systems are more individually calibrated to the actual system, whereas the mechanistic approach induces more general models. System identification, neural networks, multivariate data analysis are all groups of algorithms that

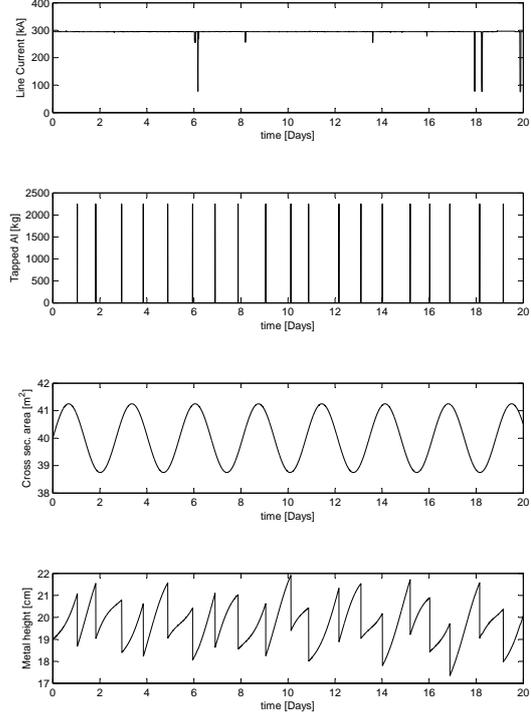


Figure 3: Simulations of the three input variables (Line current, Tapped Aluminium, and The horizontal Cross sectional area of the metal pad) and the output variable (Metal height) considered in the simplified model of an aluminium electrolysis cell.

are belonging to algorithms based on empirical reasoning/analysis. Models where the model structure is not influenced by physical relations is often called black box models (Ljung, 1999), because parameters are adjusted to fit the input and output data-sets, without reflecting physical considerations of the system. The model structure inside is not reflecting the structure of the real system, whereas the mechanistic models attempt to have identical structure to the real systems, and hence called white models.

It is also possible to take advantage of the empirical methods in conjunction with the mechanistic models, by utilizing the well known physical relations and calibrate the physical parameters by empirical methods, called parameter estimation. With this approach, all the variables are still easily available as the model structure is identical or similar to the mechanistic

model, but simultaneously more individually fitted, due to the empirical methods. There is a large range of such gray models, varying from models with almost any physically based model structures to mechanistic models where some parameters are estimated. Hence, gray model is a collective term of such mixed models.

The main focus in this approach will be parameter estimation using least square methods, which can be considered as a gray box, as the structure is designed based on physical insight, and the parameters calibrated according to empirical data.

Many testing techniques of different processes have been developed over the years based on functional, process and system modeling (among a plethora of other possibilities), aiming at testing, verifying and predicting system behavior robustly and as accurately as possible. The different modeling approaches display diverse characteristics as shown in Table 2. Naturally, some solutions may involve a series of white box, black box, and gray box in tandem depending on the system architecture and the needed set of outputs.

3.2 White Box Models

White box models are based on knowledge of the internal behavior of the system. The available form of knowledge can be in terms of equations, coherent data with the related temporal and spatial associations etc. This approach may increase effectiveness, reveal internal structures, but is essentially a general approach, that is not calibrated to the actual process. White box modeling is widely performed in simulation purposes, as no measurement of the process has to be performed, to generate the model. The process does not need to exist at all, which makes this modeling approach ideal for designing and modifying systems. The simplified model of the metal height, derived in Sec. 2 is an example of a white box model.

3.3 Black Box Models

Black box models are based on the insurmountable fact of not having any knowledge about the internal behavior of the system under scrutiny. Black box model does not have any information of the system architecture or any underlying equations describing the internal behavior of the variables involved. A typical black box model is based on the study of a set of inputs provided by the user to the system and outputs from the system oblivious to where, when and how these inputs were operated inside the system to deliver the observed outputs. In other words, how these outputs are generated or what is inside the black box representing the system are not important or unknown to the user of the black box model. The main advantages may be in the

ease of the usage and implementation of the model, and the process specific approach. On the other side, generating the model requires consistent measurement of the system, measurements that can both be cumbersome and expensive. The DSR and MDSR model of the metal height, that will be presented in Sec. 6 are examples of black box models.

3.4 Gray Box Models

Gray box models address systems with limited knowledge of the internal behavior of the system under scrutiny. As the name implies, the model will have the strategies of white box and black box models in analyzing a given system. This method may have the advantage of harvesting from both white box and black box analysis of a given system. For complicated systems like the one we have in the case of aluminium electrolysis cell, this method may be an unavoidable option, as models addressing all the phenomena in the cells are not available or incomplete. The OLS and ROLS model of the metal height, that will be presented in Sec. 5 are examples of gray box models.

4 Data sampling

Table 3 provides a set of the most common variables that are measured in the aluminium electrolysis process.

In the aluminium electrolysis process, there is a large variety in sampling intervals for the different variables. The online measurements; line current and cell voltage are the most regular performed measurements often given as mean values every 5 minutes. The tapping procedure is intermittently performed typically once a day. The height of the molten aluminium in industrial plants is only measured just before the tapping proceeding, and may influence the amount of aluminium to be tapped. In case of generating a model of this process, if just implementing these few measurements into a black box model, without adding any additional information of the system, a complete unfeasible model of the aluminium height will be the result. A black box model will “see” a folded version of the system as the dynamic of the saw-tooth shaped periodic oscillation is not sampled. By further including a realistic normal distributed measurement error with mean $\mu = 0$, and standard deviation $\sigma = 0.75$ cm, the identification process will become even worse, as can be seen in Figure 4. The uncertainty of the metal height measurements will add random dynamics to the already poor folded measurements.

Tests with real data from the industry show that the simple prediction model;

Table 2: Different “Box” Systems in conjunction with modeling and their overall characteristics

Model Type	Physical insight	Model structure	Validity
White Box	Crucial	Rigid	General
Black Box	Less important	Flexible	Specific
Gray Box	Important	Adjustable	Fairly Specific

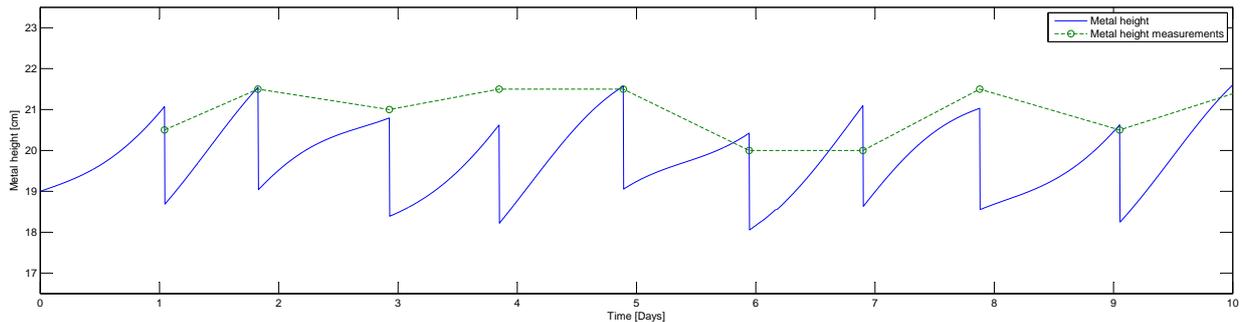


Figure 4: Synthetic data set showing metal height steadily increasing with sudden drops in its values, due to tapping, over a period of time, in this case about 10 days. It is common to measure the metal height just before the metal tapping, once a day.

$$\bar{h}(k+1) = h(k), \quad (16)$$

provides smaller error predictions than many advanced system identification models, because of the lack of information of the dynamic system the few sampling instants gives. To overcome these challenges both more physical insight of the system and more measurements have to be considered.

In cases where the output variables make stepwise jumps, like the saw-tooth jumps of the aluminium height, making the measurements at the right time instants is even more important than just increasing the number of measurements in general. The Nyquist theorem states that the sampling frequency f_s should at least be twice the frequency of the variable to be measured (Shannon, 1949, 1998);

$$f_s > 2f_m \quad (17)$$

In many reel applications it is more common to sample with a sample frequency about 10 times the system frequency. In the aluminium electrolysis example $f_s = f_m$, which according to both the Nyquist theorem and simulations provides a too poor sampling rate for system identification purposes. At the other side, increased number of manually performed metal height measurements might be too time demanding, and hence expensive for the plant, as the sampling range might

span some weeks. Especially in cases where recursive models are considered, as few manual measurements as possible should be included.

A more cost efficient solution to the high sampling rate is to use knowledge of the system to decide crucial measurement instants, i.e. including more information on the physical and chemical connections between the variables to establish a more realistic model with improved predictability. Based on sensible reasoning of the system, or if a mechanistic model of the system is available, it is possible to set up decisive intermittent measurement instants without performing a lot of measurements with a high sampling frequency. One measurement just before, and one just after the tapping procedure would have been a reasonable minimum requirement for detecting the model of the aluminium height. If the mechanistic model structure is assumed to be of high accuracy, a prediction-error-method (PEM) would be the first choice, to estimate its parameters (Ljung, 1999).

When deciding the crucial measurement instants, variables with rapid short term variations are isolated from variables with long term slow variations. After this segregation of variables, it is possible to utilize different system identification methods for each of the two variable groups. This will be considered in the following example from the aluminium electrolysis cell.

5 Prediction Error Methods (PEM) 5.1 Ordinary Least Square (OLS)

In this section different PEMs are used to predict the metal height, based on the synthetic data set described in Sec. 2. To make it realistic, the input variables are assumed sampled with a “fast” sampling rate T_f , while the output is sampled with a “slow” variable sampling rate $T_s(i)$. $T_s(i)$ will vary with time, due to manual intervention in performing the measurements. For convenience $T_s(i)$ is adjusted to fit a multiple of T_f , as shown in Figure 5. n_i is the discrete time of output measurement number i , while the inputs are sampled at each k .

The Ordinary Least Square (OLS) algorithm is designed for both these sampling rates, entailing two OLS-models. The model with sampling rate $T_s(t)$ is used for determining the parameter vector θ , while the model with sampling rate T_f is used for predicting the inter-sample outputs. It is utilized that the identified parameter vector θ is identical for both models.

Based upon the simple discrete model in Eq. (13), a linear regression model is constructed;

$$\Delta\hat{h}(k|\theta) = \underbrace{\left[\begin{array}{cc} \Phi^T(k) & \\ \Delta t \cdot I(k) & m_{out}(k) \end{array} \right]}_{\Phi^T(k)} \underbrace{\left[\begin{array}{c} \theta \\ \theta_1 \\ \theta_2 \end{array} \right]}_{\theta} \quad (18)$$

where $\Delta\hat{h}(k|\theta)$ is the predicted change of the metal height from time step $k - 1$ to k , given the parameter vector θ . The parameters are estimated by the least square method, where the optimal θ , given the defined model structure and a specific data set with N input and output samples, is the parameter vector $\hat{\theta}_N^{LS}$ that minimizes the least square criterion for the linear regression

$$V_N(\theta) = \frac{1}{N} \sum_{k=1}^N \epsilon^T(k) \lambda \epsilon(k), \quad (19)$$

where $\epsilon(k) = \Delta h(k) - \Delta\hat{h}(k|\theta)$ describes the prediction error at time step k .

Table 3: Common measured variables in aluminium electrolysis plants. The statistics are based upon process data from one single aluminium electrolysis cell generated over one year. The measurements can be categorized into two main categories; on-line measurements (\oplus) and measurements performed with manual interactions (\ominus).

Index	Variable	Unit	Category	Average sample time	Average	Standard deviation	Minimum value	Maximum value
1	Metal height	cm	\ominus	24 h	20.1	1.2	17	28
2	Cell Voltage	V	\oplus	5 min	4.49	0.29	0.44	28.3
3	Line current	A	\oplus	5 min	293.4	3.7	0	315.6
4	Bath height	cm	\ominus	24 h	20.7	1.5	14	32
5	Anode position 1	$\frac{rotations}{day}$	\oplus	78 min	0.0	1679	150	2569
6	Anode position 2	$\frac{rotations}{day}$	\oplus	78 min	0.0	1677	112	2634
7	Alumina feeding	$\frac{kg}{day}$	\oplus	15 min	4137	323	3044	5121
8	Tapped aluminium	$\frac{kg}{day}$	\ominus	24 h	2182	301	0	3125
9	Acidity	%	\ominus	48 h	10.7	1.7	6.3	15.2
10	Added bath	$\frac{kg}{day}$	\ominus	10 days	590	264	350	1750
11	Tapped bath	$\frac{kg}{day}$	\ominus	7 days	665	354	150	2100
12	Bath temperature	C	\ominus	7.5 h	962	9.0	939	998
13	Added CaF2	$\frac{kg}{day}$	\ominus	24 days	35	12.7	25	50
14	CaF2 concentration	%	\ominus	56 h	5.1	0.2	4.3	5.7
15	Anode effects	$\frac{1}{day}$	\ominus	24 hours	0.19	0.6	0	5
16	Feeder defects	$\frac{1}{day}$	\ominus	13 days	0.080	0.31	0	2
17	Superheat	C	\ominus	4 days	7.58	5.19	1.9	36

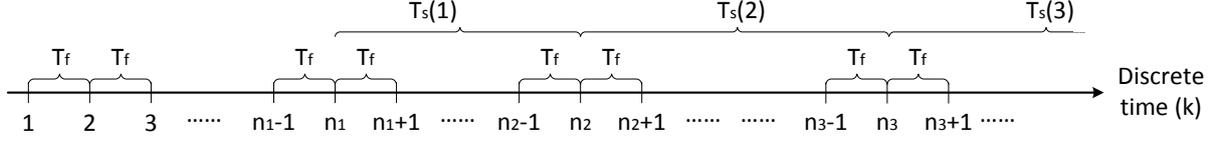


Figure 5: Time line showing the fast sampling rate T_f used for the input variables, and the slow sampling rate $T_s(i)$ used for the output measurement (the aluminium height). As indicated $T_s(i)$ will vary with time, due to manual intervention, but it is adjusted to fit a variable multiple of T_f . i indicates the i -th element of the output measurements sequence, while the inputs are sampled at each k .

The analytical solution of the optimal $\hat{\theta}_N^{LS}$ is given by

$$\hat{\theta}_N^{LS} = \left(\sum_{k=1}^N \Phi(k) \Phi^T(k) \right)^{-1} \sum_{k=1}^N \Phi(k) \Delta h(k), \quad (20)$$

if the inverse of the indicated matrix exists. As this is a MISO (multiple input, single output) system, the weight λ is a scalar, hence $\lambda^{-1}\lambda = 1$, and λ is therefore not included in Eq. (20);

Consider a data set of input and output variables describing a period of 20 days. Assuming a regular consistent measurement regime, the number of metal height measurements will be $N_h = 20$, whereas the number of line current measurements are $N_I = 5760$. Although the tapping is performed 20 times, the number of samples in the ‘‘tapping’’ vector is for convenience set to $N_t = 5760$ to match N_h , where the sample values between the tapping incidents are zero. Due to the manual intervention of the height measurements and tapping procedure, variations in both the number and time instants are common. The LS-model of Eq. (18) has to incorporate these properties, hence become more flexible in utilizing intermittent measurements from real plants.

The elements of the regression vector $\Phi(k)$ are defined as $\phi_1(k) = \Delta t \cdot I(k)$ and $\phi_2(k) = m_{out}(k)$. The predicted change of the metal height between two metal height measurements, i.e. from time step n_i to n_{i+1} are due to the LS-model in Eq. (18), given by

$$\begin{aligned} \sum_{k=n_i+1}^{n_{i+1}} \Delta \hat{h}(k | \hat{\theta}_N^{LS}) = & \\ & \theta_1 [\phi_1(n_i + 1) + \phi_1(n_i + 2) + \dots \\ & \quad + \phi_1(n_{i+1} - 1) + \phi_1(n_{i+1})] \\ & + \theta_2 [\phi_2(n_i + 1) + \phi_2(n_i + 2) + \dots \\ & \quad + \phi_2(n_{i+1} - 1) + \phi_2(n_{i+1})] \end{aligned} \quad (21)$$

To simplify the notation, we write

$$\Delta \hat{h}(n_{i+1} | \hat{\theta}_N^{LS}) = \sum_{k=n_i+1}^{n_{i+1}} \Delta \hat{h}(k | \hat{\theta}_N^{LS}) \quad (22)$$

By summing up the right-hand side of Eq. (21), a prediction model with linear regression structure is achieved;

$$\begin{aligned} \Delta \hat{h}(n_{i+1} | \hat{\theta}_N^{LS}) = & \quad (23) \\ & \underbrace{\left[\sum_{k=n_i+1}^{n_{i+1}} \phi_1(k) \quad \sum_{k=n_i+1}^{n_{i+1}} \phi_2(k) \right]}_{\Psi_{n_{i+1}}^T} \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \end{aligned}$$

Although the linear regression model of Eq. (23) is used for estimating the least square parameter vector θ , this parameter vector is mutual for both Eq. (23) and Eq. (18), hence it can be used to predict the metal height by Eq. (18).

5.1.1 Simulation of the metal height

As the tapping proceeding is a rapid process compared to the slow electrolysis process where molten aluminium is generated, height measurements should ideally be made both just before and immediately after the metal tapping. Hence, the top and at the bottom level of the metal height would be measured in each cycle. In that way, the rapid changes of the metal height, due to metal tapping, is isolated into a small temporal regression vector Ψ_{n_i} . The next regression vector $\Psi_{n_{i+1}}$ will isolate the influence of the slower electrolysis process. However, the second measurement instant should be carefully chosen due to an eventually time lag. With this approach every second Ψ -vector in the considered simulation, the last element of this vector, representing the amount of tapped metal, is zero.

5.1.2 The prediction algorithm

With the predictor model of Eq. (18), the measurement based update of the predicted metal height is very rare. Hence, the predictions of $\Delta h(k)$ will be updated only by the predictor model for all the time instants between two metal height measurements. The estimated height of the metal height is given by;

$$\bar{h}(k+1) = \bar{h}(k) + \Delta \hat{h}(k+1) \quad (24)$$

Although the parameter estimation may be good, noise and uncertainties in the model may cause the predicted output of the aluminium height to drift away from the real value of the aluminium height. On the other side the manually performed metal height measurements are likely to be corrupted by decisive noise. Therefore, a weighting algorithm has been incorporated in the predictor at each time step a new metal height measurement is performed;

$$\bar{h}(k+1) = w_1 \bar{h}(k) + w_2 \tilde{h}(k) + \Delta \hat{h}(k+1), \quad (25)$$

where $\tilde{h}(k)$ is the metal height measurement at time step k . The weights w_1 and w_2 are related so that $w_1 + w_2 = 1$. The weights are adjusted by trial and error. Generally w_1 should be close to 1, if the model is assumed to be of proper quality and simultaneously the measurements having large uncertainties. w_1 should be reduced if the model seems unreliable or the measurements more reliable. In the following example $w_1 = 0.75$.

The data sets given in Figure 3, were used to generate the θ -parameters of the OLS-algorithm. The predictions are compared to the results of simulated model and measurements in Figure 6. The predicted metal height is following the real metal height quite well, in spite of large uncertainties in the metal height measurements.

5.2 Recursive Ordinary Least Square (ROLS)

Variations in the cell performance may cause a need for updating the parameters of the model. In this case where the measurements are both rare and displaying low accuracy, updating the model has to be very slow, in order to avoid rapid variations of the model parameters, due to measurement uncertainties. Hence, short term variations of the plant will have minor influence on the model, but prospective seasonal and aging variations will have impact on the parameter variations of the model. In an OLS or a recursive ordinary least square (ROLS) method, by increasing the number of

samples, each sample will have less influence on the estimated parameters. In systems where it is likely that there will be long term variations of the model parameters, a forgetting factor α could be included (Ljung, 1999). The criterion function at time t will then obtain the following form;

$$V_t(\theta) = \frac{1}{t} \sum_{k=1}^t \alpha^{t-k} \epsilon^T(k) \lambda \epsilon(k), \quad (26)$$

Hence, the newest measurements are weighted more than the older ones. The choice of α is a trade off between reducing the sensitivity of the model regarding measurement noise, and simultaneously be able to adopt to time variations of the system parameters. Values between 0.95 and 0.99 are common choices. For a given forgetting factor, e.g. $\alpha = 0.95$, the weight assigned to a sample-value 30 samples before the current sample, is reduced to approximately 21% ($0.95^{30} \approx 0.21$) of the weight put on the current sample-value, whereas defining $\alpha = 0.99$, the representative weight is reduced to approximately 74% ($0.99^{30} \approx 0.74$).

According to Eq. (20), an optimal parameter vector given a data sample with discrete time $[1, 2, \dots, t-2, t-1]$ is given by

$$\hat{\theta}_t = \left(\underbrace{\sum_{k=1}^t \Phi(k) \lambda \Phi^T(k)}_{P_t} \right)^{-1} \sum_{k=1}^t \Phi(k) \lambda \Delta h(k), \quad (27)$$

if the inverse of P_t exists.

A recursive ordinary least square (ROLS) method with a forgetting factor is possible to include in the parameter estimation in the following way (Di Ruscio, 2001):

1. Initial values of the covariance matrix P_t and the parameter vector θ_t are defined. It is common to let

$$P_0 = \delta \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \delta I$$

$$\theta_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

where δ is a "large" number, e.g. 10 000.

2. The regression vector Ψ is updated at each metal height measurement:

$$\Psi_{n_{i+1}} = \left[\sum_{k=n_i+1}^{n_{i+1}} \phi_1(k) \quad \sum_{k=n_i+1}^{n_{i+1}} \phi_2(k) \right]^T \quad (28)$$

3. Updating the inverse of the covariance matrix

$$P_{n_{i+1}}^{-1} = \alpha P_{n_i}^{-1} + \Psi_{n_{i+1}} \lambda \Psi_{n_{i+1}}^T \quad (29)$$

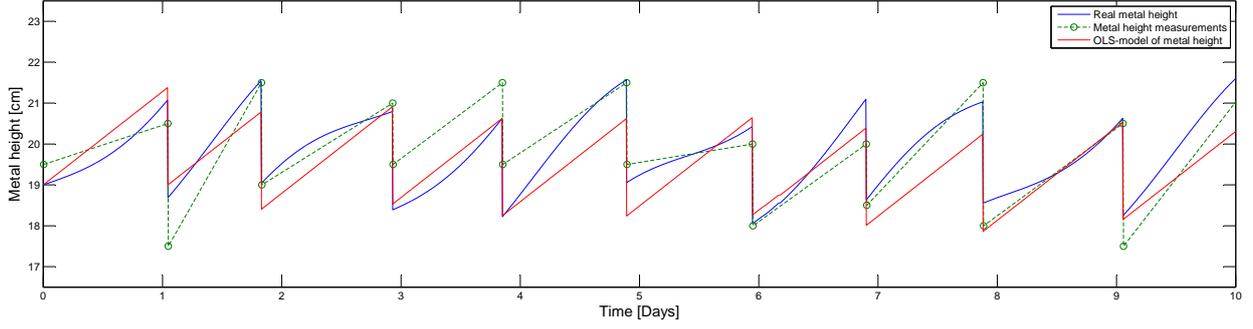


Figure 6: Prediction of metal height for the earlier presented synthetic data set showing fairly well correlation between the simulated and predicted values. The prediction model is based on the OLS algorithm and the noisy metal height measurements, just before and just after metal tapping, shown with the green circles in the figure.

and the “Kalman” gain matrix

$$K_{n_{i+1}} = P_{n_{i+1}} \Psi_{n_{i+1}} \lambda \quad (30)$$

after each metal height measurement.

4. Updating the present parameter vector at each metal height measurement

$$\theta_{n_{i+1}} = \theta_{n_i} + K_{n_{i+1}} (\Delta h_{n_{i+1}} - \Psi_{n_{i+1}} \theta_{n_i}) \quad (31)$$

where $\Delta h_{n_{i+1}} = h(n_{i+1}) - h(n_i)$

5. Computing the covariance matrix after each metal height measurement:

$$P_{n_{i+1}} = \left(\alpha P_{n_i}^{-1} + \Psi_{n_{i+1}} \lambda \Psi_{n_{i+1}}^T \right)^{-1} \quad (32)$$

6. At each time instant k the predicted change of metal height is given by

$$\Delta h(k) = \Phi(k) \theta^T \quad (33)$$

5.2.1 Matrix inversion lemma

An equivalent form of Eq. (29) and Eq. (32) that is more suited for rapid computation is given by the Sherman-Morrison-Woodbury formula (Golub and Loan, 1996), also called the matrix inversion lemma;

$$(A + UV^T)^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1}U)^{-1}V^T A^{-1} \quad (34)$$

Applying Eq. (34) to Eq. (32), where

$$\begin{aligned} A &= \alpha P_{n_i}^{-1} \\ U &= \Psi_{n_{i+1}} \\ V^T &= \Psi_{n_{i+1}}^T \end{aligned}$$

gives

$$P_{n_{i+1}} = \quad (35)$$

$$\frac{1}{\alpha} P_{n_i} - \left(\frac{1}{\alpha} P_{n_i} \Psi_{n_{i+1}} \right) \cdot \left(\left(I + \Psi_{n_{i+1}}^T \frac{1}{\alpha} P_{n_i} \Psi_{n_{i+1}} \right)^{-1} \Psi_{n_{i+1}}^T \frac{1}{\alpha} P_{n_i} \right)$$

The matrix inversion in Eq. (32) is modified using Eq. (34), hence the inversion of a matrix is now replaced by a inversion of a scalar, i.e. the denominator in the following deduced equation:

$$P_{n_{i+1}} = \quad (36)$$

$$\frac{1}{\alpha} P_{n_i} \left(I - \frac{\Psi_{n_{i+1}} \Psi_{n_{i+1}}^T P_{n_i}}{\alpha + \Psi_{n_{i+1}}^T P_{n_i} \Psi_{n_{i+1}}} \right)$$

Note that this way of calculating $P_{n_{i+1}}$ is more sensitive regarding selection of initial values of P_0 when working with regression vectors where one element is zero, and the other is “large”. In the simulated system that is considered, only proper parameter estimation is achieved for values of $\delta \leq 10^{-4}$. For larger values of δ , the first element of the P-matrix will become zero, and hence θ_1 will also remain at zero throughout the whole estimation period. For values of $\delta \leq 10^{-8}$, P and hence θ will converge very slowly towards its proper values.

If the normal routine of letting the initial value of P_{11} be “large” is followed, and in addition ψ_1 is “large”, the impact of $\alpha \in (0, 1]$, will easily be neglected due to the limitations of the computer precision. Assuming that $\alpha \ll \Psi_1^T P_0 \Psi_1$, the first estimate of P will become a lower triangular matrix due to the computer precision;

$$\begin{aligned}
 P_1 &= \frac{1}{\alpha} P_0 \left(I - \frac{\Psi_1 \Psi_1^T P}{\alpha + \Psi_1^T P_0 \Psi_1} \right) \\
 &= \frac{1}{\alpha} P_0 \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \frac{1}{\alpha} P_0 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\
 &= \frac{1}{\alpha} P_0 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \quad (37)
 \end{aligned}$$

However, by rearranging Eq. (36) in the following way, proper values of θ will be estimated also for larger values of δ :

$$\begin{aligned}
 P_1 &= \frac{1}{\alpha} P_0 \left(\frac{\alpha}{\alpha + \Psi_1^T P_0 \Psi_1} I + \frac{\Psi_1^T P_0 \Psi_1 I - \Psi_1 \Psi_1^T P}{\alpha + \Psi_1^T P_0 \Psi_1} \right) \quad (38) \\
 &= \frac{1}{\alpha} P_0 \left(\begin{bmatrix} \frac{\alpha}{\Psi_1^T P_0 \Psi_1} & 0 \\ 0 & \frac{\alpha}{\Psi_1^T P_0 \Psi_1} \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \right) \\
 &= \frac{1}{\alpha} P_0 \begin{bmatrix} \frac{\alpha}{\Psi_1^T P_0 \Psi_1} & 0 \\ 0 & 1 \end{bmatrix} \quad (39)
 \end{aligned}$$

P_1 will in this case remain a diagonal matrix, which is crucial. The reason why, the matrix inversion lemma had to be rearranged is that displaying $\frac{\Psi_1 \Psi_1^T P}{\alpha + \Psi_1^T P_0 \Psi_1}$ requires higher computer precision than displaying $\frac{\alpha}{\alpha + \Psi_1^T P_0 \Psi_1}$, since the former has a very small perturbation from 1, whereas the other has a small perturbation from 0.

Given a realistic example from the system considered in this paper, where $\Psi_1^T = [2.5 \cdot 10^{10} \ 0]$ and $P_0 = 10^4 I$. Computing P_1 will then involve computation with numbers close to the precision limits of the computer. By using the ordinary matrix inversion lemma, only values where $10^{-7} < \delta < 10^{-4}$ caused reasonable results. Model predictions outside this area where unstable as one of the parameters maintained zero throughout the regression. For larger values P_{11} and θ_1 became zero for the ordinary matrix inversion lemma, whereas the modified inversion lemma performed proper estimations for all larger values of P_{11} .

5.2.2 Prediction of the metal height

The data sets given in Figure 3, were used to generate the θ -parameters of the ROLS-algorithm. The prediction is compared to the simulated model and measurements in Figure 7. The predicted metal height is following the real metal height quite well, in spite of large uncertainties in the metal height measurements.

In Figure 8 the temporal variation of the regression parameters is shown. For this simulation the recursive variant of the OLS is superfluous, as these parameters are not following the constant system parameters very well.

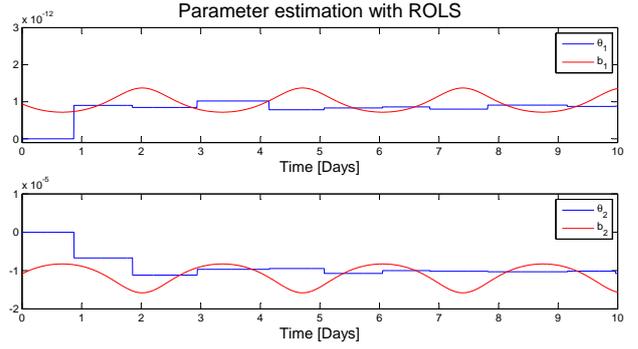


Figure 8: Simulating the values of the estimated parameters, calculated in the ROLS model, using the earlier presented synthetic data set. The "rapid" oscillation of the parameters is not detected by the ROLS-model. The estimated parameters are close to the mean value of the parameters, and will only follow long term variations of the parameters.

6 System Identification using DSR

Assume that there are dynamics within the system that are not detected by the strict mechanistic model used with PEM, as how the temperature is influencing the area A . An alternative is to identify black box models, like the DSR (Deterministic and Stochastic Realization) model. The DSR method is based upon a linear discrete time invariant State Space Model (SSM)

$$x_{k+1} = Ax_k + Bu_k + Ce_k \quad (40)$$

$$y_k = Dx_k + Eu_k + Fe_k, \quad (41)$$

and is explained in [Di Ruscio \(1996\)](#).

6.1 Regular DSR

The sample scenario used with the OLS and ROLS algorithms in the former section is not consistent with the SSM in Eq. (40). Further the number of measurements has to be increased to extract information on the short term system dynamics between the tapping instants. Hence, a new measurement regime has been set up for the identification of a DSR model, as described in the following.

Assume that the system of interest might be described by the linear SSM, and that the inputs are sampled with one fast sampling rate, T_f , whereas the output is known at a slow sampling rate, T_s . Contrary to Sec. 5, T_s is constant in this section. The input sampling rate is M times faster than the output sampling rate, hence $T_s = MT_f$. If it is not possible to increase T_s , the *modus operandi* is to consider only the

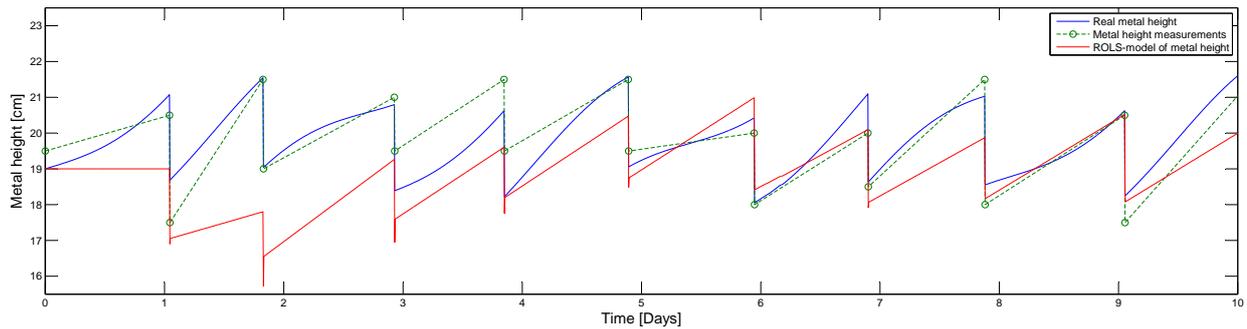


Figure 7: Simulating the predicted metal height using ROLS. Prediction of metal height for the earlier presented synthetic data set showing good correlation between the simulated and predicted values. The prediction model is based on the ROLS algorithm and the noisy metal height measurements shown with the green circles in the figure. Although the initial value is set to high, the predicted values “soon” reaches the desirable level. ROLS needs some time to adjust its model parameters and hence the variations in its initial phase.

inputs at a sampling rate of T_s , in order to achieve a common sampling rate, which is required to use the standard DSR-method (Di Ruscio, 2000). Then, to get a discrete SSM with the fast sampling rate, T_f , the generated discrete SSM is first transformed into a continuous SSM, utilizing e.g. the MATLAB function $[A_c, B_c] = d2c(A, B, T_s)$. Next the continuous model is transformed back to a discrete model, but now with another sampling rate by the MATLAB function $[A_d, B_d] = c2d(A_c, B_c, T_f)$. A periodical variation in the cell temperature introducing a variation in the width of the metal bath has been introduced in the following simulations. Hence, in addition to the sawtooth variation, already explained, the metal height will fluctuate with a period of about 10 hours. Assuming that the three input variables I , m_{out} and T (bath temperature) are sampled every 5 minutes, and the metal height measured every 3rd hour. To be able to utilize the DSR algorithm the mean values every third hour are used as inputs, to generate the state space model. The model is then transformed from a sampling time of 3 hours to 5 minutes as already described. Figure 9 shows the result of a simulation of the metal height in the aluminium electrolysis cell, using this method. The simulation shows a fairly good estimation, in spite of the poor metal height measurements also shown in the figure. By running several additional simulations with random generated measurement noise, some of the concomitant predictions had a tendency of drifting away from the real metal height. This could be compensated by including measurement in the prediction algorithm, e.g. including a Kalman filter, when running the prediction model online. This could be a useful estimation technique if implementing automatic metal height

measurements as discussed in (Viumdal et al., 2010; Viumdal and Mylvaganam, 2010). To be able to utilize the DSR algorithm the mean values every 3rd hour are used as inputs, to generate the state space model. The model is then transformed from a sampling time of 3 hours to 5 minutes as already described. This transformation of the model unfortunately more often involves less robust models.

6.2 DSR in case of multiple time series

Another approach to identify a model of the metal height is to use the MDSR, as addressed in (Di Ruscio, 1997). MDSR is designed to handle multiple time series, e.g. an industrial process where the time series are interrupted by shutdowns in the production line, or where each time series represents a batch process. By just placing measurements from different time series successively in a large data matrix, the initial states of each time series is not computed when using ordinary DSR algorithm. By using MDSR, initial state values are calculated for each time series, in addition to the overall state space model. In the aluminium electrolysis example, the increase in metal height is “interrupted” by the metal taps. Instead of measuring the metal height every 3rd hour, as assumed with the DSR-algorithm, another measurement scenario is applied for the MDSR approach. These measurement series are divided into several sub time series, each starting immediately after the metal tapping. As the tapping has an intermittent sampling interval, there will be a shift in the sampling incidents between each of the time series. Within each sub time series, new metal height measurements are sampled each 3rd hour. The mean value

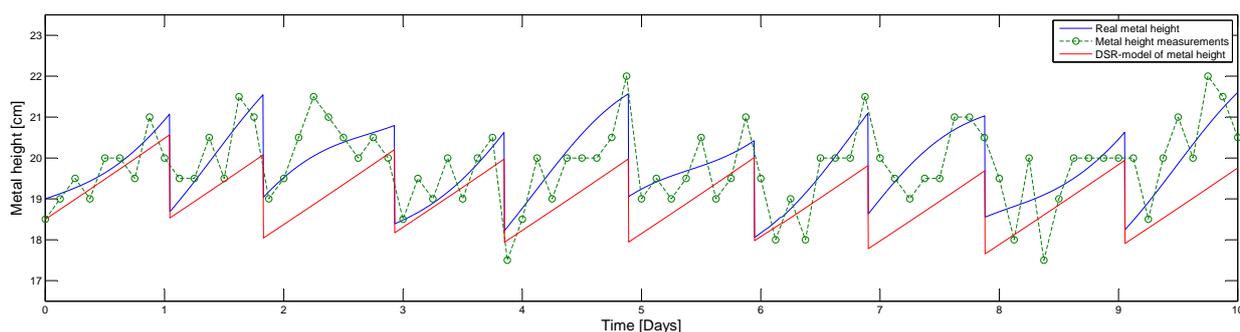


Figure 9: Prediction of the metal height, using DSR on the earlier synthetic data set. Unlike the assumed measurement regime for the OLS and ROLS models, the metal height is here assumed to be measured every 3rd hour. The noisy metal height measurements are shown with the green circles in the figure. In general, the simulated values are closer to the predicted than the measured values.

of the input variables are taken for each time step as done with the DSR model. The result is a new model, with as many initial states as there are sub time series. Figure 10 shows an example where MDSR is used, still with a sampling time between each metal height measurement of 3 hour, but here with sampling occurrence according to the latest metal tap instance. Attempts of transforming the MDSR model with a sampling rate of 3 hours to a sampling rate of 5 minutes failed, therefore the graph in Figure 10 is marked with dots, at the locations where model has calculated the predictions.

However, in a real aluminium electrolysis process, the metal height is still measured manually in an intermittent manner, and normally only just before the tapping proceeding. Hence, the PEM-algorithms seem to be the most realistic approach at the moment. Using MDSR online is not straightforward, one alternative is to use any PEM-algorithm to predict the drop in metal height due to the metal tapping, and using the MDSR as the model describing the metal height between the tapping proceeding.

7 System identification using real measurements

A measurement sequence from an Aluminium reduction cell in Norsk Hydro, Årdal, was performed with a time span of 76 hours. The metal height was measured just before, and just after the tapping instants, and then approximately every 4th hour. The input variables “Line current” and “Tapped bath” were sampled each 5 minute. Two different OLS models of the process was generated for this data set. The first model was only based on the metal height measurements performed in conjunction with the metal tapping, while

the inter sampled measurements are included in the second OLS model. Figure 11 shows the predictions of the two OLS models, the measurements, and a simulation of the mechanistic model. The mechanistic model is based upon geometrical parameters of the aluminium reduction cell, and the initial value is adjusted to fit the measurements. The element values of the parameter vector θ are about 36% to 50% less than the corresponding parameter values in the mechanistic model. The parameter vector of the two OLS models are similar, but as the first model is based on fewer measurements than the second model, the former is expected to be less reliable. However, the main difference between these two models, is that the second OLS model is updated more often by new measurements. It is difficult to verify which model is closest to the real system based of these few measurements, particularly as the measurements are corrupted with so much noise. More reliable models can be achieved by extending the time span of the sampling, and by introducing improved measurement systems, which are beyond the scope of this paper.

8 Comparing performance of the models in metal height predictions

In addition to the methods already described, the problem was also organized as a lifted model approach, using the same data set, as used for the DSR approach, i.e. the sampling time for the three input variables were every 5 minute, whereas the sampling time of the output was every 3rd hour. Hence, the inputs were sampled 36 times as often as the output, resulting in an input matrix with 72 (2x36) variables. Due to the

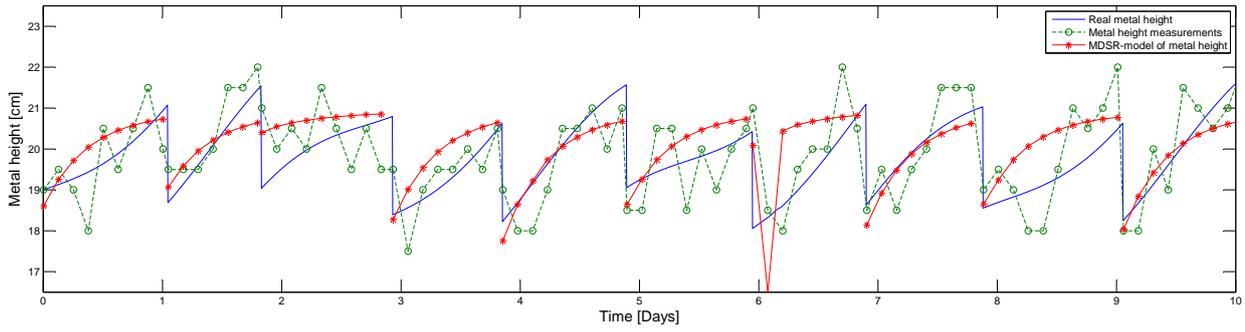


Figure 10: Prediction of the metal height, using MDSR on the earlier synthetic data set. Unlike the assumed measurement regime for the OLS and ROLS models, the metal height is here assumed to be measured every 3rd hour, but as opposed to the measurement regime using the DSR, here the measurements are considered as multiple time series, each starting just after the tapping proceeding. The noisy metal height measurements are shown with the green circles in the figure. In general, the simulated values are closer to the predicted than the measured values.

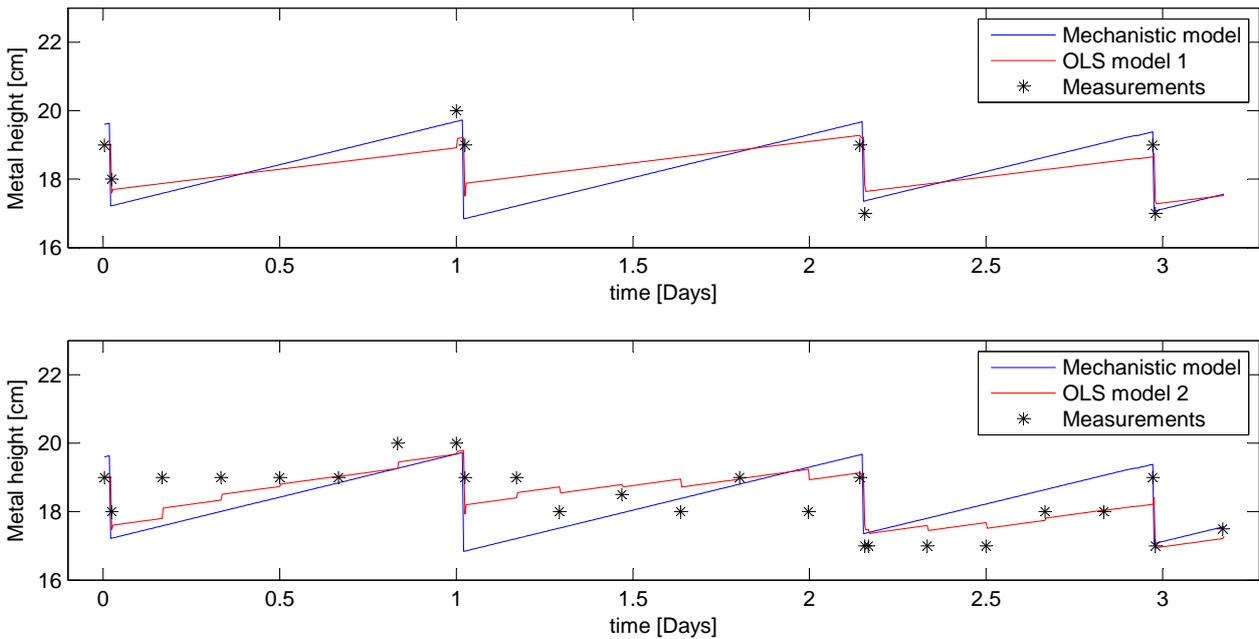


Figure 11: Predictions of the metal height, using two different OLS models based on the data from a real aluminium plant. OLS model 1 is based upon the measurements just before and just after the metal tap instants. OLS model 2 is in addition based on inter sampled measurements every 4 hour as indicated in the figure. The mechanistic model is based upon the physical parameters of the aluminium reduction cell, and the initial value is adjusted to fit the measurements.

rare tapping operations, 21 variables having no variance, were deleted from the input matrix. They were all representing the original metal tap variable. As proposed by Li et al. (2003), the Canonical Correlation

Analysis (CCA) initially developed by Hotelling was applied, to reduce the number of variables with poor correlation to the output. Running this test, all the 25 best correlation results were representing the

line current. Hence, a model built on this approach would not include the impact of the metal tapping. The main problems related to the lifted model in this approach is most likely related to the poor variance of the input matrix, that is even poorer in the lifted input matrix. In addition the saw tooth like metal height graph, is difficult to identify, as the tap proceedings are not synchronized, and the noise level was rather high. Of that reason, the lifted model was not further considered in this work. For the other models considered in this work, a new data set was generated for validation purpose. In addition to the mechanistic model, predictions of the system identification models were simulated. The selected results from the simulations are shown in Figure 12. The validation of the models was analyzed using Mean Square Error (MSE), Normalized Mean Square Error (NMSE), and Normalized Root Mean Square Error (NRMSE). On contrary to the MSE, where the numerical value should be as small as possible, the numerical values of the NMSE and NRMSE are in the range $(-\infty, 1]$, where 1 is representing a perfect model. The numerical performance of these models are shown in Table 4. The validation data set has been extract for the period after the ROLS model has stabilized.

The results show that the OLS model and ROLS model are following the mechanistic model quite well. So are also the results based on DSR, but the DSR model fails in the validation data set. In contrast to the OLS model and ROLS model, where the predictions are updated by new measurements, only the initial state vector is calculated in the DSR model. Thus, the states are not updated by new measurements in the DSR model, inducing a tendency of drifting away. By utilizing new measurements of the metal height in a implemented state estimator, an online version of the DSR, will enhance the performance, by reducing the drift of the model. The MDSR predictions are updated by the new measurements, but the model does not recreate the dynamics in a proper way. In systems where the main dynamics is well known, and the variations in the measurement instant make subspace methods difficult, an OLS approach seems to be the optimal choice of system identification algorithm.

In real applications the DSR model often has the advantage to include more input variables to the model, with the intention of improving the model. The need for looking for such “lurking” input variables can be revealed by analyzing an error prediction plot or residual plot. The prediction errors for the OLS model is plotted in Figure 13, together with a sine curve with identical periodicity of the not measured input variable “Cross section area”. As this is a “lurking”variable, indicating that it has not been involved in the identi-

fication problem (as it is not measured), it will not be detected by the OLS model. Similar error predictions are seen by the other models, but as mentioned a DSR model will easily include new input measurements of the system. Although the variable “Cross section area” is difficult to measure on a regular basis, some of the “lost” dynamics can perhaps be detected by the other variables that are more easily available. It would then be possible to identify a model describing this additional dynamic by eg. a DSR model. Finally, the OLS and the new DSR model could be merged to a gray box model, hence gaining advantage of both model types. This will not be treated in the present paper, as it is beyond the scope of this work, but the reader is referred to [Draper and Smith \(1998\)](#) for more details for utilizing the error prediction or residual plots.

9 Conclusion

This paper addresses challenges using ordinary system identification methods in applications with rare and intermittent sampling frequencies, and if the sampling frequencies are different for the input and output variables. In particular the challenge with identifying a robust predictor of the metal height in aluminium electrolysis cells is under scrutiny. When dealing with so few regularly performed measurements as in the electrolysis process, using historical data to model the metal height is insufficient. With measurements just before, and just after the tapping of metal, it is possible to achieve quite good predictors, by just using the two main inputs; line current I and the amount of tapped aluminium m_{out} . This can be achieved in spite of the intermittent tapping instants, by defining two linear models with common parameter vector θ . The first with a slow sampling rate, that is consistent with the metal height measurement, where the input variables are summed between each metal height measurements. This regression model is used for determining θ . The other linear model should have a sampling rate identical to the input variables, and serves as a predictor of the metal height between the measurements. To determine more complex dynamic between the tapping, more measurements have to be taken, in particular metal height measurements, but also more input variables with regularly performed measurements. If measuring the metal height every 3rd hour throughout 20 days, some acceptable predictions of the metal height was achieved using ordinary DSR-algorithms, although it often created rather unstable models, due to the transformation from models with 3 hours sampling time to 5 minutes sampling time. If instead considering metal height measurements every 3rd hour, but now adjusted in time after each metal tapping procedure,

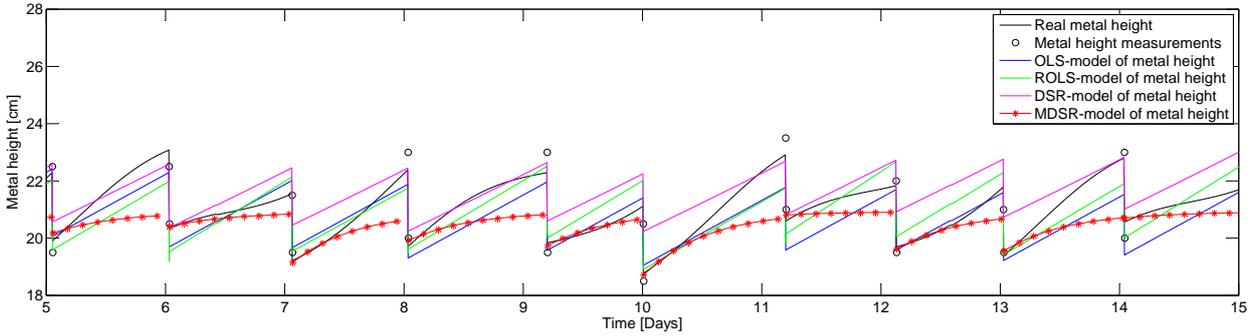


Figure 12: Simulations of the different models, based on the validation data set. The OLS and ROLS predictions are following the “real” metal height quite well, in spite of the large measurement uncertainty. The DSR has a tendency of drifting away as it is not updated by the new metal height measurements, as in the case of the former models. The MDSR has problems in reproducing the system dynamic, but updates the predicted metal height very well at each new measurement.

Table 4: Numerical validation of the models

Algorithm	MSE		NMSE		NRMSE	
	Identification	Validation	Identification	Validation	Identification	Validation
OLS	$6.52 \cdot e^{-5}$	$5.03 \cdot e^{-5}$	0.16	0.49	0.83	0.29
ROLS	$3.59 \cdot e^{-5}$	$3.79 \cdot e^{-5}$	0.54	0.62	0.32	0.38
DSR	$12.8 \cdot e^{-5}$	$9.97 \cdot e^{-5}$	-0.65	-0.0054	-0.28	0.0027
MDSR	$5.33 \cdot e^{-5}$	$7.93 \cdot e^{-5}$	0.37	0.080	0.21	0.041

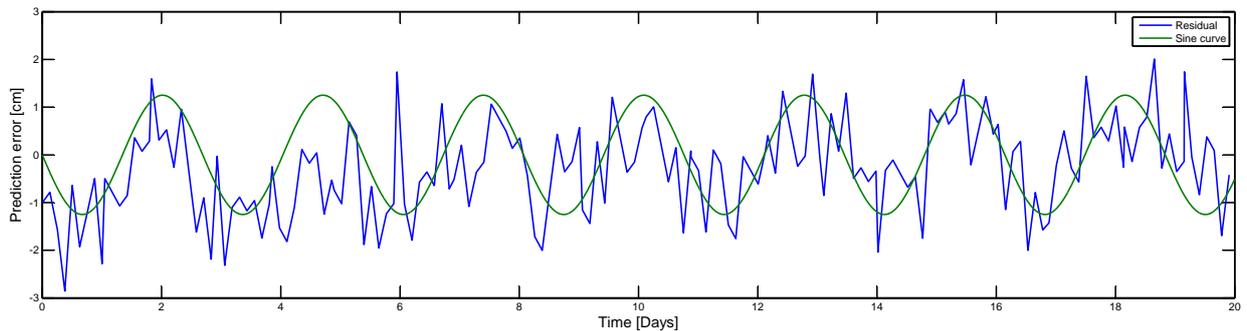


Figure 13: The prediction errors of the OLS prediction shows a periodic tendency as indicated by the sine function plotted with identical periodicity of the not measured input variable “Cross section area”. This indicate that there are dynamics in the data set not identified by the OLS model. This is related to a “lurking” variable, that is not included in the data set used for identification.

the DSR did not work, due to the time-lag. In this case MDSR, which is the DSR for multiple time series provided acceptable predictions. As far as there are so few regularly performed measurements in the aluminium electrolysis process, parameter estimation approaches using PEM-algorithms seem to be the most satisfying and reliable approach to predict the metal height in the aluminium electrolysis process. This conclusion is mainly based on the experience with this simplified mechanistic model given in Eq. (12), as the real measurements are few in their numbers and have considerable uncertainties.

Acknowledgments

This work was supported by the Research Council of Norway in conjunction with a strategic institute program (Multiprocon). We appreciate the collaboration with Morten Liane and Sven Olof Ryman at Norsk Hydro, Årdal, Norway, and are thankful to them for taking the metal height measurements used in this paper.

References

- Bojarevics, V. and Pericleous, K. Comparison of MHD models for aluminium reduction cells. In *TMS annual meeting and exhibition*. pages 347–352, 2006.
- Bojarevics, V. and Pericleous, K. Solutions for the metal-bath interface in aluminium electrolysis cells. In *TMS annual meeting and exhibition*, volume 1. pages 569–574, 2009.
- Di Ruscio, D. Combined Deterministic and Stochastic System Identification and Realization: DSR - A Subspace Approach Based on Observations. *Modeling, Identification and Control*, 1996. 17(3):193–230. doi:[10.4173/mic.1996.3.3](https://doi.org/10.4173/mic.1996.3.3).
- Di Ruscio, D. A Three Dimensional DSR Algorithm. Technical report, Telemark University College, Faculty of Technology, 1997.
- Di Ruscio, D. Subspace system identification of multiple time series and different sampling rates. *Unpublished report*, 2000. pages 1–18.
- Di Ruscio, D. Model Predictive Control and Optimization. Technical report, Telemark University College, Faculty of Technology, 2001. Lecture Notes.
- Ding, F., Qiu, L., and Chen, T. Reconstruction of continuous-time systems from their non-uniformly sampled discrete-time systems. *Automatica*, 2009. 45(2):324 – 332. doi:[10.1016/j.automatica.2008.08.007](https://doi.org/10.1016/j.automatica.2008.08.007).
- Draper, N. and Smith, H. *Applied regression analysis*. Wiley series in probability and mathematical statistics. Applied probability and statistics. Wiley, 3rd edition, 1998.
- Golub, G. H. and Loan, C. F. V. *Matrix Computations*. The John Hopkins University Press, 3rd edition, 1996.
- Grjotheim, K. *Introduction to Aluminium Electrolysis: Understanding the Hall-Héroult Process*. Aluminium-Verlag, 1993. URL <http://www.beuth.de/en/publication/introduction-to-aluminium-electrolysis/143079846>.
- Kranc, G. Input-output analysis of multirate feedback systems. *Automatic Control, IRE Transactions on*, 1957. 3(1):21–28. doi:[10.1109/TAC.1957.1104783](https://doi.org/10.1109/TAC.1957.1104783).
- Kurenkov, A., Thess, A., Zikanov, O., Segatz, M., Droste, C., and Vogelsang, D. Stability of aluminium reduction cells with mean flow. *Magneto-hydrodynamics*, 2004. 40(2):203–212.
- Li, D., Shah, S., Chen, T., and Patwardhan, R. System identification and long-range predictive control of multi-rate systems. In *American Control Conference, 1999. Proceedings of the 1999*, volume 1. pages 336–340 vol.1, 1999. doi:[10.1109/ACC.1999.782796](https://doi.org/10.1109/ACC.1999.782796).
- Li, D., Shah, S., Chen, T., and Qi, K. Application of dual-rate modeling to CCR octane quality inferential control. *Control Systems Technology, IEEE Transactions on*, 2003. 11(1):43–51. doi:[10.1109/TCST.2002.806433](https://doi.org/10.1109/TCST.2002.806433).
- Li, W., Shah, S. L., and Xiao, D. Kalman filters in non-uniformly sampled multirate systems: For FDI and beyond. *Automatica*, 2008. 44(1):199 – 208. doi:[10.1016/j.automatica.2007.05.009](https://doi.org/10.1016/j.automatica.2007.05.009).
- Ljung, L. *System Identification -Theory for the User*. Prentice Hall, 2nd edition, 1999.
- Lu, W. and Grant Fisher, D. Least-squares output estimation with multirate sampling. *Automatic Control, IEEE Transactions on*, 1989. 34(6):669–672. doi:[10.1109/9.24247](https://doi.org/10.1109/9.24247).
- Shannon, C. Communication in the presence of noise. *Proceedings of the IRE*, 1949. 37(1):10–21. doi:[10.1109/JRPROC.1949.232969](https://doi.org/10.1109/JRPROC.1949.232969).
- Shannon, C. Communication in the presence of noise. *Proceedings of the IEEE*, 1998. 86(2):447–457. doi:[10.1109/JPROC.1998.659497](https://doi.org/10.1109/JPROC.1998.659497).

- Sheng, J., Chen, T., and Shah, S. L. Generalized predictive control for non-uniformly sampled systems. *Journal of Process Control*, 2002. 12(8):875 – 885. doi:[10.1016/S0959-1524\(02\)00009-4](https://doi.org/10.1016/S0959-1524(02)00009-4).
- Viumdal, H. and Mylvaganam, S. Beyond the dip stick: Level measurements in aluminum electrolysis. *JOM*, 2010. 62(11):18–25. doi:[10.1007/s11837-010-0161-0](https://doi.org/10.1007/s11837-010-0161-0).
- Viumdal, H., Yan, R., Liane, M., Moxnes, B., and Mylvaganam, S. Multi sensor data fusion for aluminum cell health monitoring and control. In *TMS annual meeting and exhibition supplemental proceedings*, volume 3. pages 149–159, 2010. ISBN Number 978-0-87339-753-7.
- Wang, J., Chen, T., and Huang, B. Multirate sampled-data systems: computing fast-rate models. *Journal of Process Control*, 2004. 14(1):79 – 88. doi:[10.1016/S0959-1524\(03\)00033-7](https://doi.org/10.1016/S0959-1524(03)00033-7).