Data-driven models for demand-side management

Peder Bacher

IEA Symposium on Demand Flexibility and RES Integration

May 9, 2016

Data analysis and statistics

- How can we gain as much as possible useful information from data?
- Statistical inference: the process of drawing conclusions from data that is subject to random variation
- Time-series models for describing a dynamical system

Model complexity

- Einstein: *"Everything should be made as simple as possible, but not simpler"*
- Fundamental question: "Which model and how complex should it be for *optimally* for providing the answers?"

Model complexity

- Einstein: *"Everything should be made as simple as possible, but not simpler"*
- Fundamental question: "Which model and how complex should it be for *optimally* for providing the answers?"
- Answer: it depends on the data!
  - 'simple data' $\Rightarrow$ 'simple model'
  - 'complex data' $\Rightarrow$ 'complex model'
- It is a matter of what we need to know or simply economical investment. Which sensors are needed for providing the needed information?
- or the other way around what can be achieved with current resources

Model complexity

From statistical theory a wide range of techniques are available:

- Find the most suitable model to describe the data
- Estimate the uncertainty
- Validate the model fit (likelihood)

Model structure

Different types of *time series* models:

- Static models, *no dynamics* (e.g. for daily values)
- ARMAX, discrete models based on transfer functions, *black-box dynamics, however for control and steady-state parameters (e.g. UA-value, gA-value) fully applicable*
- Grey-box models. *Continuous (or discrete) time models, combination of physics and statistics*

Model structure

Different types of *time series* models:

- Static models, *no dynamics* (e.g. for daily values)
- ARMAX, discrete models based on transfer functions, *black-box dynamics, however for control and steady-state parameters (e.g. UA-value, gA-value) fully applicable*
- Grey-box models. *Continuous (or discrete) time models, combination of physics and statistics*

Static models (linear function):

$$\text{Measurements} = \text{Function}(\text{Inputs}) + \text{Residual}$$

Model structure

Different types of *time series* models:

- Static models, *no dynamics* (e.g. for daily values)
- ARMAX, discrete models based on transfer functions, *black-box dynamics, however for control and steady-state parameters (e.g. UA-value, gA-value) fully applicable*
- Grey-box models. *Continuous (or discrete) time models, combination of physics and statistics*

ARMAX model:

$$\text{Measurements} = \text{Transferfun}_1(\text{Inputs}) + \text{Transferfun}_2(\text{Error})$$

MODEL STRUCTURE

Different types of *time series* models:

- Static models, *no dynamics* (e.g. for daily values)
- ARMAX, discrete models based on transfer functions, *black-box dynamics, however for control and steady-state parameters (e.g. UA-value, gA-value) fully applicable*
- Grey-box models. *Continuous (or discrete) time models, combination of physics and statistics*

Grey-box model:

$$\begin{aligned} \text{States} &= \text{Fun}_1(\text{States}, \text{Inputs}) + \text{Fun}_2(\text{SystemError}) \\ \text{Measurements} &= \text{Fun}_3(\text{States}, \text{Inputs}) + \text{Fun}_4(\text{MeasurementError}) \end{aligned}$$

MODEL STRUCTURE

Different types of *time series* models:

- Static models, *no dynamics* (e.g. for daily values)
- ARMAX, discrete models based on transfer functions, *black-box dynamics, however for control and steady-state parameters (e.g. UA-value, gA-value) fully applicable*
- Grey-box models. *Continuous (or discrete) time models, combination of physics and statistics*

Grey-box model:

$$\begin{aligned} \text{States} &= \text{Fun}_1(\text{States}, \text{Inputs}) + \text{Fun}_2(\text{SystemError}) \\ \text{Measurements} &= \text{Fun}_3(\text{States}, \text{Inputs}) + \text{Fun}_4(\text{MeasurementError}) \end{aligned}$$

Note that part of the model is a description of the error(s)

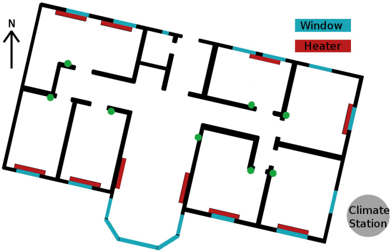
Annex 58: Guidelines

- In Annex 58 we developed guidelines (focus on *energy performance assessment*)
- 'Physical guidelines': setup of measuring campaign and experiments
- 'Statistical guidelines': models for data from buildings (unoccupied, e.g. from a test sequence run 3-7 days):
  - Static, ARX and grey-box models
  - Model selection procedure
  - Examples and implementations i R

TEST CASE: ONE FLOORED 120 M^2 BUILDING

Objective

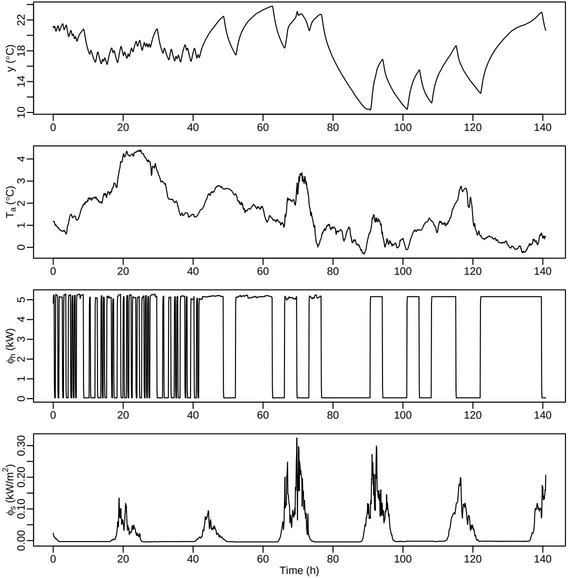
Find the best model describing the heat dynamics of this building



DATA

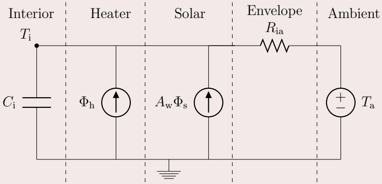
Measurements of:

- y_t Indoor air temperature
- T_a Ambient temperature
- Φ_h Heat input
- Φ_s Global irradiance



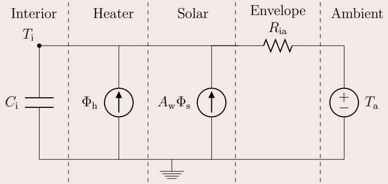
IDENTIFY THE BEST PHYSICAL MODEL FOR THE DATA

Simplest model

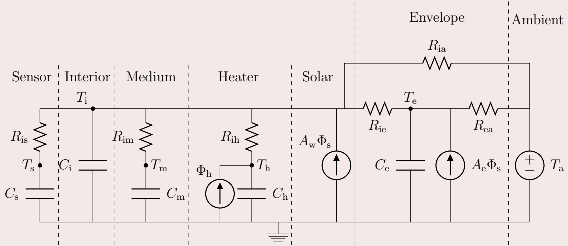


Identify the best physical model for the data

Simplest model

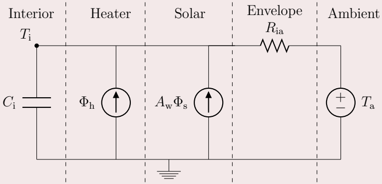


Most complex model applied



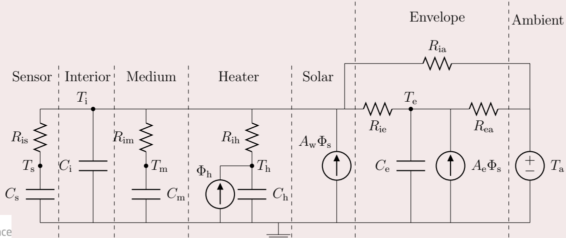
Identify the best physical model for the data

Simplest model



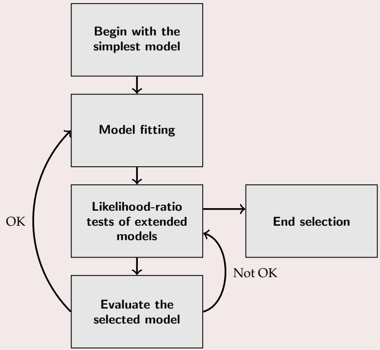
The best model for the given data is probably in between

Most complex model applied

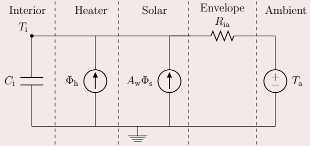


SELECTION PROCEDURE

Iterative procedure using statistical tests

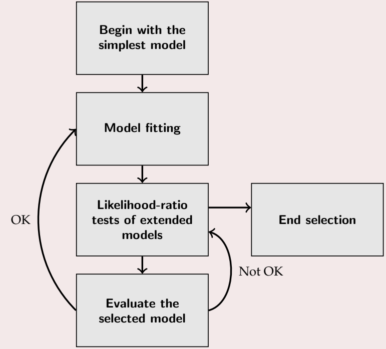


Simplest model

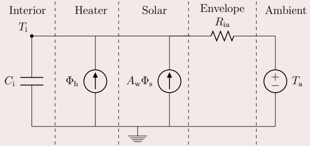


SELECTION PROCEDURE

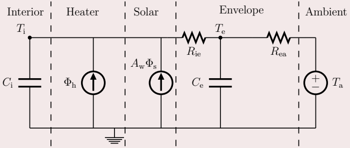
Iterative procedure using statistical tests



Simplest model

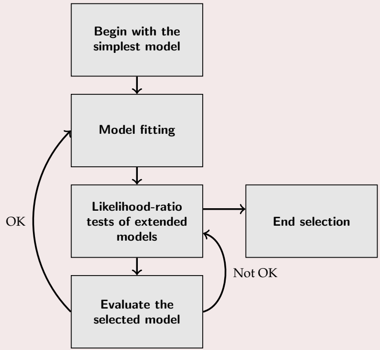


First extension: building envelope part

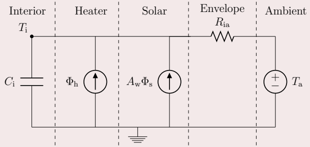


SELECTION PROCEDURE

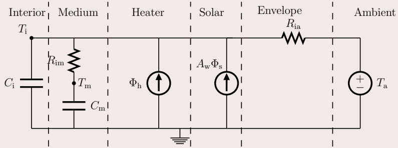
Iterative procedure using statistical tests



Simplest model

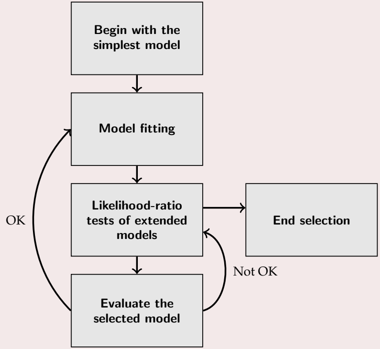


First extension: indoor medium part

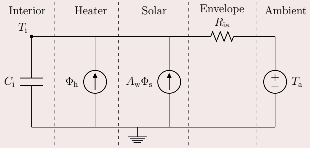


SELECTION PROCEDURE

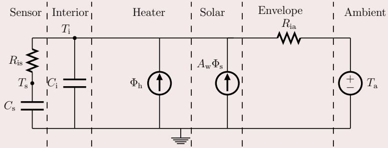
Iterative procedure using statistical tests



Simplest model

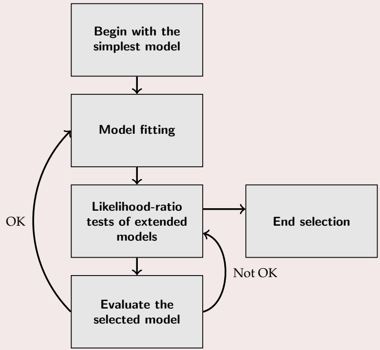


First extension: sensor part

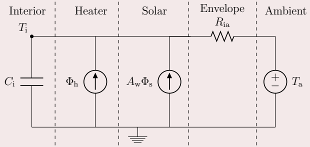


SELECTION PROCEDURE

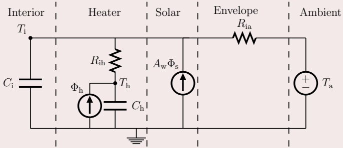
Iterative procedure using statistical tests



Simplest model

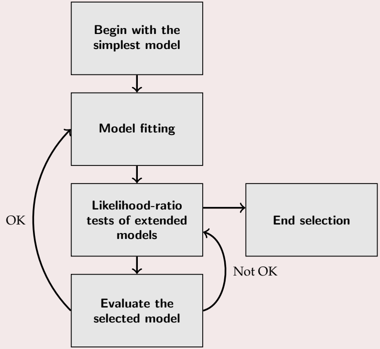


First extension: heater part

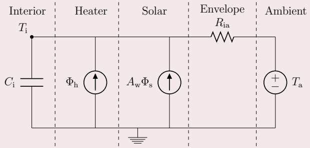


Selection procedure

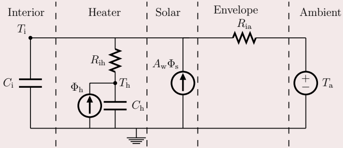
Iterative procedure using statistical tests



Simplest model

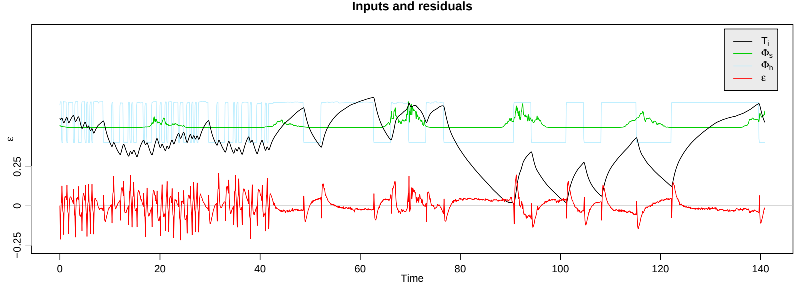


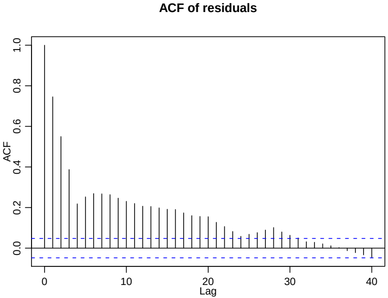
First extension: heater part

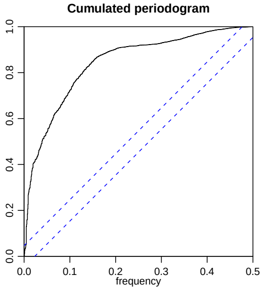


Start	$Model_{\text{Ti}}$			
$l(\theta;\mathcal{Y}_N)$	2482.6			
m	6			
1	$Model_{\text{TiTe}}$	$Model_{\text{TiTm}}$	$Model_{\text{TiTs}}$	$Model_{\text{TiTh}}$
$l(\theta;\mathcal{Y}_N)$	3628.0	3639.4	3884.4	3911.1
m	10	10	10	10
2 ...				

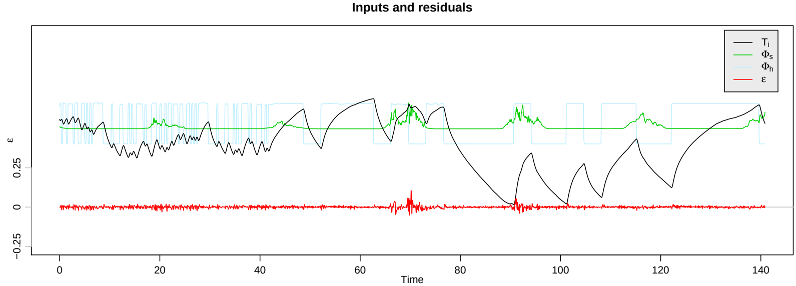
Evaluate the simplest model

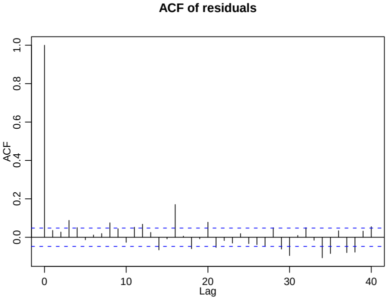


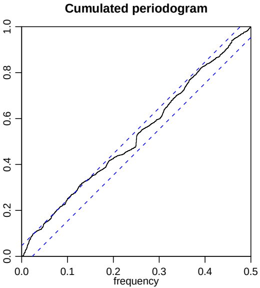




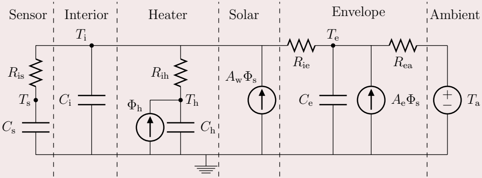
EVALUATE THE SELECTED MODEL



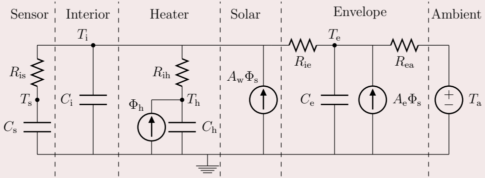




Selected model



Selected model



Estimated parameters

$\hat{C}_i$	0.0928	(kWh/C)
$\hat{C}_e$	3.32	-
$\hat{C}_h$	0.889	-
$\hat{C}_s$	0.0549	-
$\hat{R}_{ie}$	0.897	(°C/kW)
$\hat{R}_{ea}$	4.38	
$\hat{R}_{ih}$	0.146	-
$\hat{R}_{is}$	1.89	-
$\hat{A}_w$	5.75	(m^2)
$\hat{A}_e$	3.87	-

Estimated time constants

$\hat{\tau}_1$	0.0102	hours
$\hat{\tau}_2$	0.105	-
$\hat{\tau}_3$	0.788	-
$\hat{\tau}_4$	19.3	-

Important points

- Need to excite the dynamics of the system!
- Hence you need data with variation in the inputs:
  - Turn on/off the heaters
  - Low ambient temperature preferable
  - You need direct solar radiation
- Data from buildings with thermostatic control wont work (flexibility can be with hot water tank)

MPC

- We have a model to predict the indoor temperature:
  - Input: heating and climate
  - Output: indoor temperature
- Model Predictive Control (MPC):
  - Setup a cost function (e.g. monetary and indoor climate)
  - Constrains (max heating etc.)
  - Use weather forecasts and calculate an optimized heat input

More time series modelling techniques

- Model selection (likelihood-ratio test, AIC, BIC)
- Parametric, semi-parametric and non-parametric models:
  - splines, kernels, regression trees, neural-networks, ...

More time series modelling techniques

- Model selection (likelihood-ratio test, AIC, BIC)
- Parametric, semi-parametric and non-parametric models:
  - splines, kernels, regression trees, neural-networks, ...
- Kalman filtering (grey-box models)
- Hidden Markov models (regime models)
- Robust estimation and outlier detection
- Time adaptive models

We are setting up a new Annex: focus models for occupied buildings(contact Staf Roels, KU Leuven)

Some links

- Annex 58 Statistical Guidelines
- Summer school on these matters (time-series modelling for buildings), 19. to 24. June, Grenada, Spain
- DTU Compute, Dynamical Systems
  - Solar and wind forecasting, load forecasting, data-driven models for: buildings, user behaviour, EVs, district heating, grids
  - MPC and optimization
- CITIES project
- Send me a mail pbac@dtu.dk

Thanks for your time!