# Hidden Markov Models for indirect classification of occupant behaviour

Jon Liisberg, Jan Kloppenborg, Henrik Madsen DTU Compute, jarl@dtu.dk
Hans Bloem European Commission, JRC
Jordi Cipriano, Gerard Mor CIMNE-UPC

CITIES Centre for IT Intelligent Energy Systems
DTU Compute
Institut for Matematik og Computer Science

## Introduction

- The motivation for this study is based on:
  - Huge variability in energy consumption in similar residential buildings.
  - Limited direct observation of occupancy behaviour.
  - Increasing use of Smart Meters.
- By applying Hidden Markov Models (HMMs) on frequent observations of electricity consumption we can:
  1. Classify the states of the HMM in accordance to occupancy behaviour.
  2. Identify possible covariates/explanatory variables explaining occupancy behaviour.
  3. Forecasting and simulation of the future energy consumption.

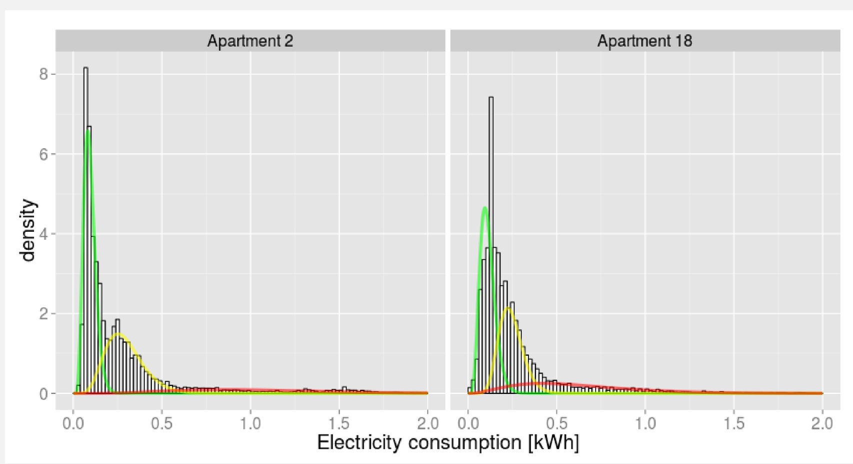
## Data

- Metering and weather data from and nearby an apartment building in Catalonia, with 44 apartments. Also an occupant survey was available.
- Hourly observations from July 2012 to December 2013 consists of:

| Variable | description |
|---|---|
| $x_e$ | Electricity consumption in $kWh$ |
| $x_{sh}$ | Space heating in $kWh$ |
| $x_{hw}$ | Hot water consumption in $kWh$ |
| $x_w$ | Water consumption in litters |
| $T_a$ | Ambient temperature in $°C$ |
| $G$ | Solar radiation in $W/m^2$ |
| $W_s$ | Average wind speed in $m/s$ |
| $W_d$ | Average wind direction in $°$ |
| $P$ | Precipitation in $mm$ |

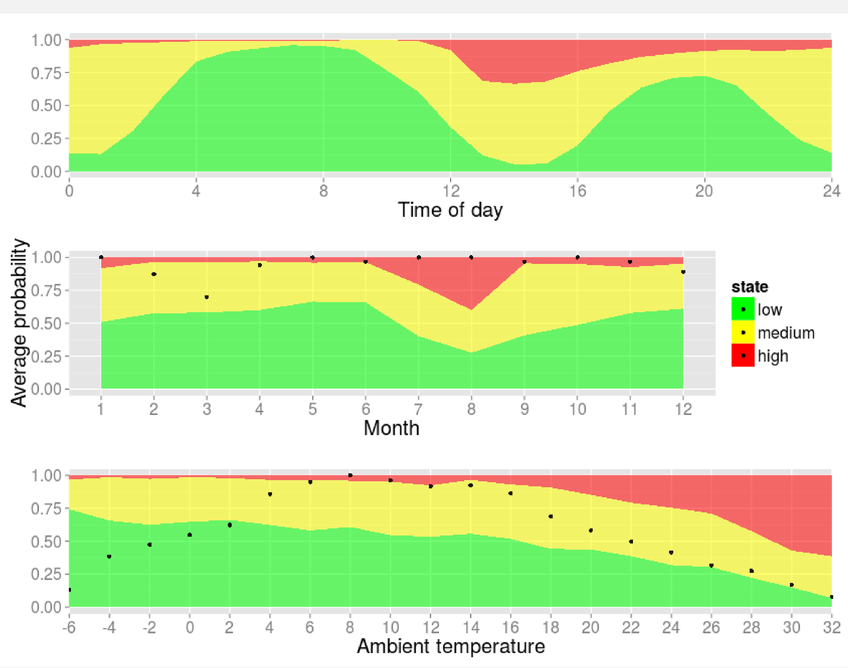
## Homogeneous Hidden Markov Models

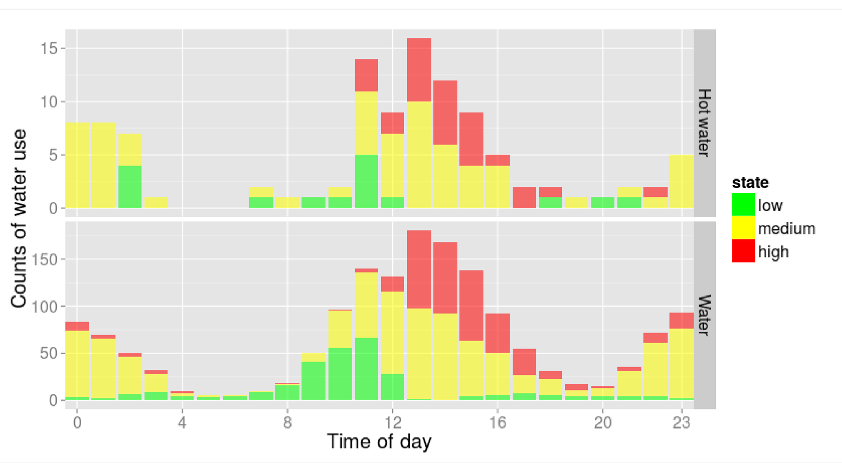
- HMMs were applied on one year of hourly electricity consumption observations, from 14 of the 44 apartments.
- The HMMs returns a predetermined number of state dependent distributions (gamma distributions in this case) and a matrix describing the probability of changing between these states
- For these Apartments, 3-4 states were found adequate.




- Based on the mean values of the state dependent distributions the states are classified as:
  1. Low consumption (green)
  2. Medium consumption (yellow)
  3. High consumption (red)

## Homogeneous HMMs, dependencies

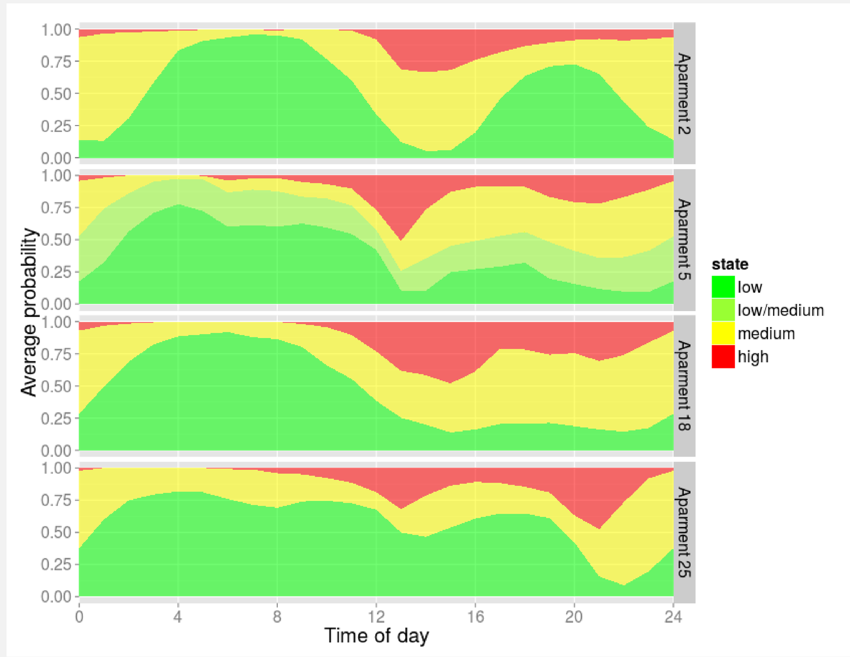


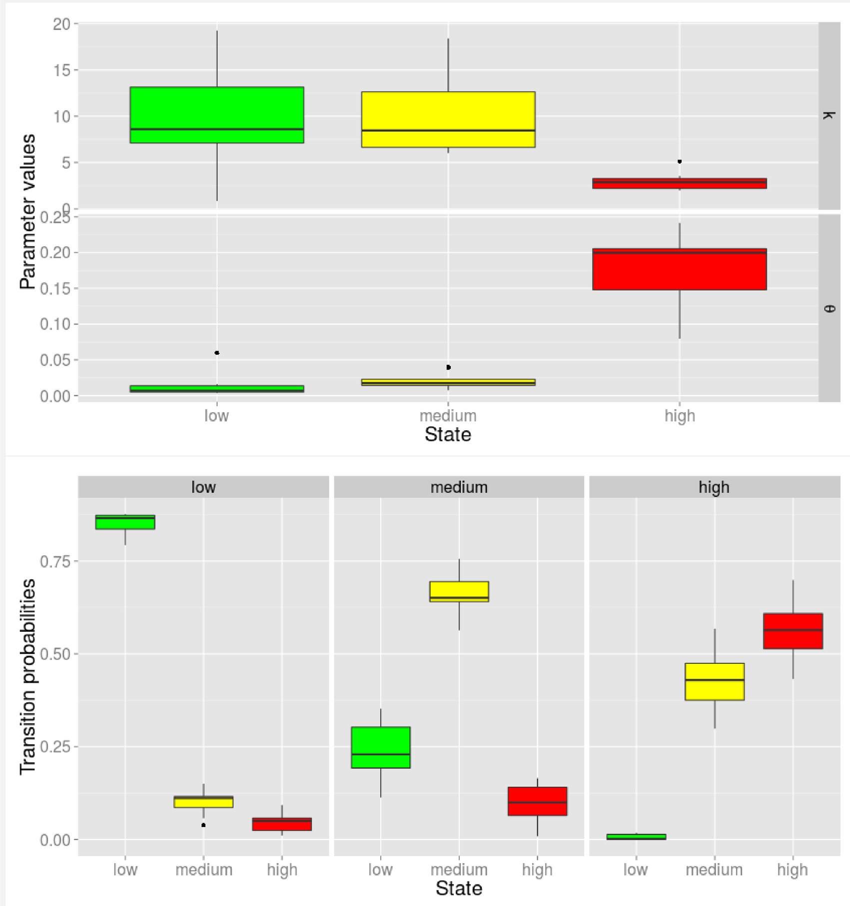

- Based on the daily average probability profile derived from the most likely sequence of states (global decoding), time dependence is observed and the states are classified further as:
  1. Absent or asleep (green)
  2. Home, medium consumption (yellow)
  3. Home, high consumption (red)
- For Apartment 2, clear temperature dependence on high consumption, originating from summer period. In this case due to an air-conditioner.
- These interpretations are compared to water use for validation. Except for the morning hours there is low probability of the low consumption state when using water.

## Homogeneous HMM, distinct patterns

- Four different profiles are observed and classified as:
  - Afternoon/evening absence
  - Equal probability for being home or absent.
  - Mostly at home
  - Mostly absent
- Comparing different patterns with occupant survey, indication of common factors were observed e. g.
  - no. residents
  - Income (work, pension or subsidies)



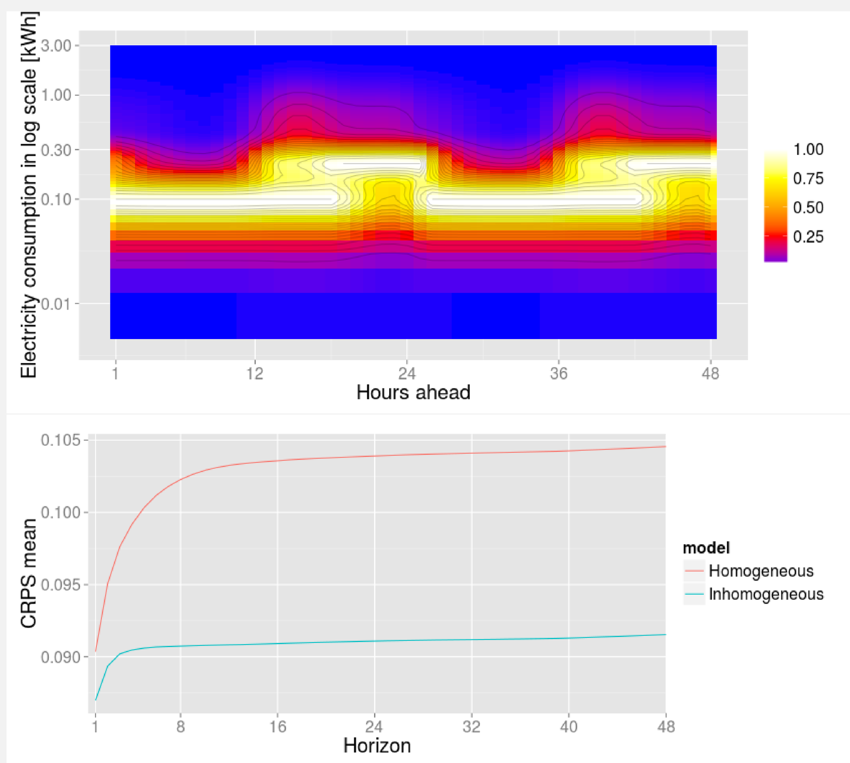

## Homogeneous HMM, common parameters




- Box plots of shape and scale parameters of the state dependent distributions (gamma distributions) and the transition probabilities, indicates common distributions for these parameters.
- This indicates the possibility to collect Apartments in population models.

## Inhomogeneous HMMs

- Time dependence added in the transition probabilities (plots for Apartment 18).
  - Forecasting distributions dependent on time of day obtained.
  - Continuous Rank Probability Score (CRPS) shows improvement of forecast.
  - Indication of common parameters observed for inhomogeneous HMMs.
  - High consumption not captured for some apartments.
  - Changing Behaviour not captured.




## Temperature dependent inhomogeneous HMMs

- Adding temperature as a covariate in the conditional mean of the state where the dependence was observed, did not yield good results due to varying mean values. Smoothing of mean values is suggested.

## Discussion/Conclusion

- Using HMMs we have classified states in accordance to occupancy behaviour
  - "low consumption" and "absent or asleep"
  - "medium consumption" and "home"
  - "high consumption" and "home, high consumption"
- Identified dependency on time and temperature.
- Observed distinct occupancy patterns.
- Improved forecasting by inhomogeneous HMMs
- Found it plausible to model population models, both for homogeneous and inhomogeneous HMMs
- Found model deficiencies
  - High consumption (varying coefficients)
  - Changing behaviour (adaptive methods)
  - Temperature dependent mean values (Smoothing)