Statistics and Crystall Ball Techniques

... on the use of Hidden Markov Models

Henrik Madsen
Professor

www.henrikmadsen.org
Contents

Applications of Hidden Markov Models for:

- Geolocation of fish
- Energy optimization of buildings
- Automatic dosing of insulin
- Energy systems integration
Case Study No. 1

Geolocation of Fish
Geolocation of fish

- Goal: Identify models for the movements of fish.
- GPS systems do not work under water.
- 'Data storage tags' for measuring the pressure (depth under the surface).
- Data gets available at capture of the fish.
**Hidden Markov Model**

- **$x(t)$**: *States* (are not observed) – describes the evolution in time of the system
- **$y(t)$**: *Observations*

**Typical application**: Find the values for $x$!
Hidden Markov Model

The probability $\Phi$ for, that the fish at time $t$ is in position $x$, is:

Systemet (generelt): $\frac{\partial \phi}{\partial t} = -\nabla (u\phi - D\nabla \phi)$

Systemet (her): $\frac{\partial \phi}{\partial t} = D \left( \frac{\partial^2 \phi}{\partial x_1^2} + \frac{\partial^2 \phi}{\partial x_2^2} \right)$

Data is (as mentioned):

Observationer: $Y_k$ : Dybden (til tidspunkt $t_k$)
Further information

- Bathymetry (depths)
- Time and place for release and capture
- Information about the tide system – see the graph
Observations

Measured sequence of depths from release to capture:

Where has the fish been?
Case Study No. 2

Characterising the Energy Performance of Buildings
Example

U=0.86 W/m²K       U=0.21 W/m²K

Consequence of good or bad workmanship (theoretical value is U=0.16W/m²K)
Examples (2)

Whole House Heat Loss - Measured versus Predicted for New Build UK Dwellings (n=18)

Measured versus predicted energy consumption for different dwellings
Energy Labelling of Buildings

- Today building experts make judgements of the energy performance of buildings based on drawings and prior knowledge.

- This leads to 'Energy labelling' of the building

- However, it is noticed that two independent experts can predict very different consumptions for the same house.
Model for the heat dynamics

Measurements:
- Indoor air temp
- Radiator heat sup.
- Ambient air temp
- Solar radiations

Hidden states are:
- Heat accumulated in the building
- $k$: Fraction of solar radiation entering the interior
Measurements of:

\( y_t \) Indoor air temperature

\( T_a \) Ambient temperature

\( \Phi_h \) Heat input

\( \Phi_s \) Global irradiance
Selection procedure

Iterative procedure using statistical tests

1. Begin with the simplest model
2. Model fitting
3. Likelihood-ratio tests of extended models
4. Evaluate the selected model
5. End selection

Simplest model

First extension: heater part

Start

\[ l(\theta; \mathcal{Y}_N) \]

\[
\begin{array}{c|c|c|c|c}
\text{Model}_{Ti} & \text{Model}_{TTe} & \text{Model}_{TTh} & \text{Model}_{TTh}\text{m} \\
\hline
\text{m} & 2 & 3628.0 & 3639.4 & 3884.4 \\
1 & 10 & 10 & 10 & 10 \\
\hline
2 & \ldots & & & \\
\end{array}
\]
EVALUATE THE SIMPLEST MODEL

Inputs and residuals

ACF of residuals

Cumulated periodogram

DTU Compute
Institut for Matematik og Computer Science
**Final model for the heat dynamics**

- **Again: HMM - or state space - model**

  \[
  \text{Systemet: } \quad dx_t = f(x_t, u_k, t, \theta) dt + \sigma(u_t, t, \theta) d\omega_t \\
  \text{Observationer: } \quad y_k = h(x_k, u_k, t_k, \theta) + e_k
  \]

- **Model found using statistical modelling:**
Perspectives

- Identification of most problematic buildings
- Automatic energy labelling
- Recommendations:
  - Should they replace the windows?
  - Or put more insulation on the roof?
  - Or tighten the building?
  - Should the wall against north be further insulated?
  - ......
- Better control of the heat supply
Better utilization of renewable energy (solar and wind power)
Case Study No. 3

Insulin – Glucose Models
Today a diabetic person must measure the blood sugar in order to provide the correct dose of insulin.

A correct dose depends on a lot of factors like: activity, stress level, hormonal state, meals, etc.
Intelligent device (A) for optimal dosing of insulin (B) based on measurements of glucose (C and D)
Traditional Dynamic model

- Ordinary Differential Equation:

\[ dA = -KA \, dt \]
\[ Y = A + \epsilon \]
Stochastic Dynamic Model

Stochastic Differential Equation:

\[ \frac{dA}{dt} = -KA_t + \sigma dw \]

\[ Y = A + e \]
HMM or State Space Model

Model for a description of glucose-insulin relation:

Systemet: \[ d\mathbf{x}_t = f(\mathbf{x}_t, \mathbf{u}_k, t, \theta)dt + \sigma(\mathbf{u}_t, t, \theta)d\mathbf{\omega}_t \]

Observationer: \[ y_k = h(x_k, u_k, t_k, \theta) + e_k \]

Using appropriate statistical methods we can again:

- Find the best model and the hidden states
Grey-box modelling concept

<table>
<thead>
<tr>
<th>White</th>
<th>Grey</th>
<th>Black</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deterministic equations</td>
<td>Prior Knowledge</td>
<td>Databased</td>
</tr>
<tr>
<td>Physical knowledge</td>
<td></td>
<td>Input–output</td>
</tr>
<tr>
<td>Detailed submodels</td>
<td></td>
<td>representation</td>
</tr>
</tbody>
</table>

- Combines prior physical knowledge with information in data
- Equations and parameters are physically interpretable
Best model found ...
Case Study No. 4

Intelligent and Integrated Energy Systems
Quote by B. Obama at the Climate Summit 2014 in New York:

*We are the first generation affected by climate changes,*

*and we are the last generation able to do something about it!*
**Scenario:** We want to cover the world's entire need for power using wind power.

- How large an area should be covered by wind turbines?
**Potentials and Challenges for renewable energy**

**Scenario:** We want to cover the world’s entire need for power using wind power.

How large an area should be covered by wind turbines?

**Conclusion:** Use data intelligence ....


Total of power consumption - Worldwide
Energy Systems Integration

Energy system integration (ESI) = the process of optimizing energy systems across multiple pathways and scales

Single Building

Community, City

Region, Country

Data Pathway: Information and communication technologies allow a better understanding and control of systems by linking sensor data from multiple locations to control centers.
ESI – Hypothesis

The central hypothesis of ESI is that by intelligently integrating currently distinct energy flows (heat, power, gas and biomass) in we can enable very large shares of renewables, and consequently obtain substantial reductions in CO2 emissions.

Intelligent integration will (for instance) enable lossless ‘virtual’ storage on a number of different time scales.
ESI – Research Challenges

To establish methodologies and models for operation of integrated electrical, thermal, fuel pathways at all scales
Modelling

Use of **Hidden Markov Models** for operation of future integrated energy system.
Smart-Energy OS
Energy Flexibility in Wastewater Treatment

1. Elspot price forecast
2. Sewer system
3. WWTP
4. STAR
5. R
   - Sensor filtering and time delay
   - CTSM-R
     - Model parameter estimation
   - Closed loop prediction
minimize overflow $+ p^T_{\text{elspot}} f(Q)$
Sewer System Annual Elspot Savings

Savings [1000 DKK]

Storage volume [m³]

0  10^2  10^3  10^4  10^5
Hidden Markov Models

- Eksemples considered:
  - Geo-location of fish
  - Energy labelling and optimization
  - Control of insulin injection
  - Intelligent and integrated energy systems

- In general: Hidden Markov Models are useful for observing phenomenas that we cannot otherwise observe
Thank you for your attention ...

And thanks to:

Søren Klim, Andreas Bærentzen, Sten Frandsen, Peder Bacher, Martin Wæver Petersen, John Bagterp Jørgensen, ....