

Missing data estimation for energy resources management in tertiary buildings



Antoine GARNIER
a.garnier@pyres.com

PYRESCOM
PROMES laboratory

Julien EYNARD
Matthieu CAUSSANEL
Stéphane GRIEU

PROMES laboratory



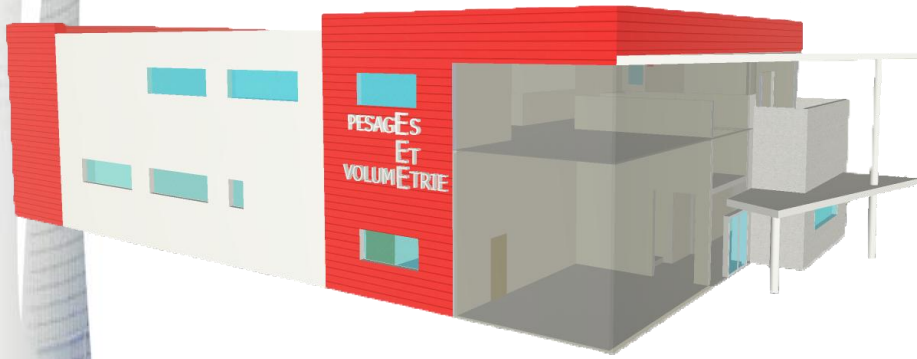
Summary

- BATNRJ Project
- Objectives and constraints
- Solar radiation estimation
- Temperature estimation
- Conclusion
- Prospects

BATNRJ project

- Monitoring system designed to **improve energy efficiency of buildings**
- Meteorological parameters acquisition
- Energy measurement





Instrumented sites

| Place | Building type |
|--------------------------|---|
| Perpignan, Saint-Charles | Offices + 200 m ² manufacturing area |
| Meudon, Paris region | 23 000 m ² individual offices or open spaces, positive energy building |
| Montpellier | Offices + datacenter |



Sensors



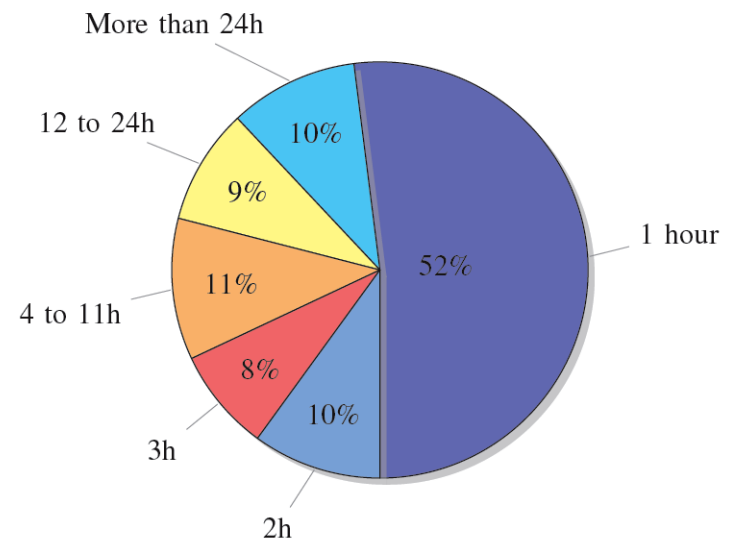
About 400 sensors of various types installed:

- Temperature (ambient and radiative)(120)
- Energy consumption (electricity, gas, water)(50)
- Air speed (80)
- Humidity (60)
- CO₂ (30)
- Solar radiation (10)



Sensor failures

- During the system development phase, failures occur due to:
 - Sensor batteries discharged
 - Network error
 - Sensor malfunction
- About 2% data missing
- Variable duration



Objectives and constraints

- Knowing the mean duration of failures, goal is to **develop a correction system for solar radiation and temperature** capable of:
 - 24-hour estimation
 - Efficient estimation between 1 and 3 hours
- **Minimal calculation time** to integrate algorithm into an existing server without penalizing other installed applications



Solar radiation estimation

- **Low priority:** estimation can be done at the end of the day
- **Interpolation made by fitting the available values to a function with a least square criterion**
- Data measured each 30 seconds, then averaged over a period of one hour: equivalent to a low-pass filtering that decreases influence of rapid clouds

Equations used for fitting

- Gaussian: based on repartition function equation

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} * \exp^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$



$$p(t, A) = A_1 * \exp^{-\left(\frac{t-A_2}{A_3}\right)^2} + A_4$$

- Cosine: based on theoretical solar radiation equation

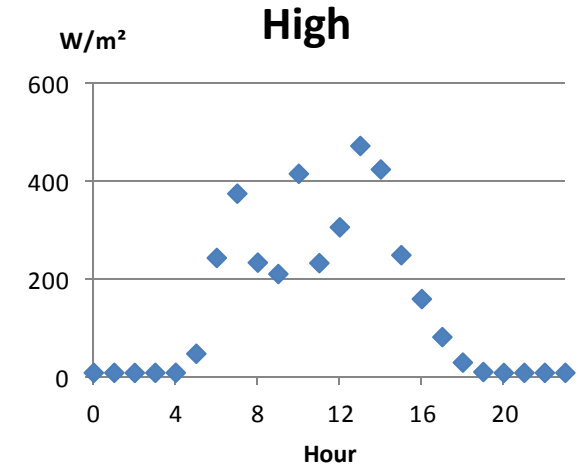
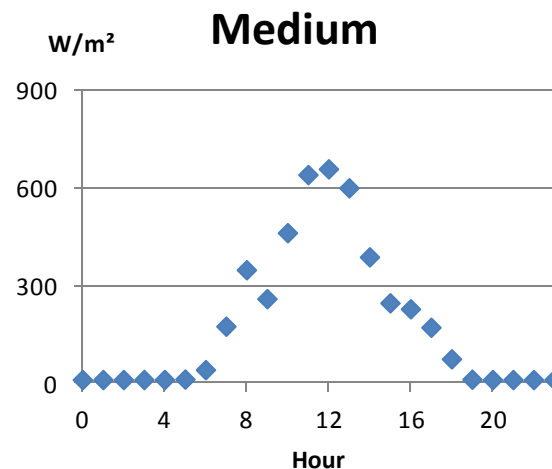
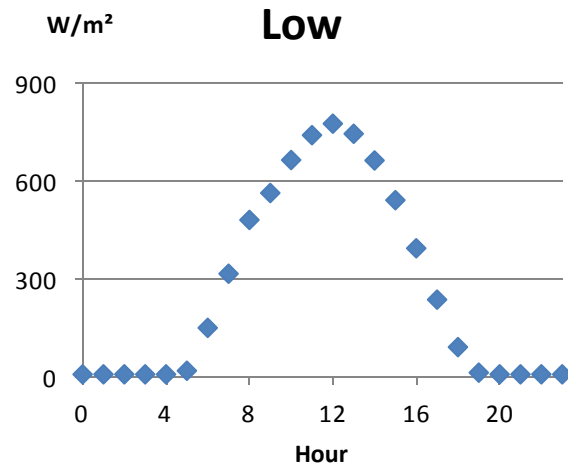
$$G_0 = \varepsilon_0 * (1 + 0.0334 * \cos(w * (j - 2))) * \sin(h)$$



$$p(t, A) = A_1 * \cos(A_3(t - A_2)) + A_4$$

Data set splitting according to a perturbation level

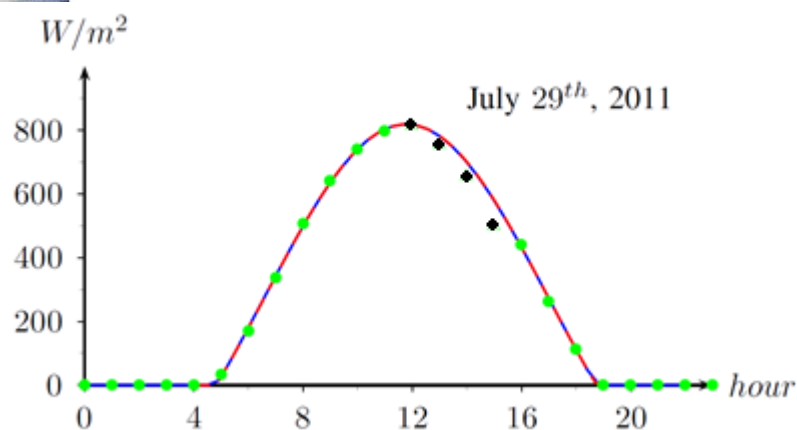
| Perturbation level | Occurrence percentage |
|--------------------|-----------------------|
| Low | 50% |
| Medium | 30% |
| High | 20% |



Data estimation for a 4-hour failure

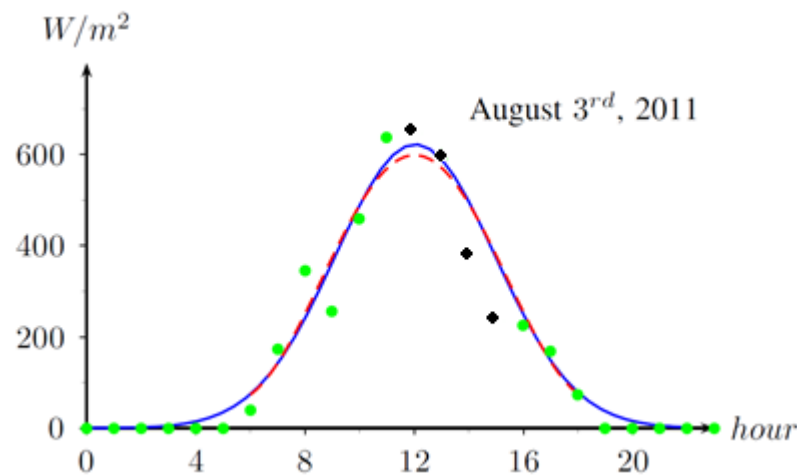
Low perturbation

$$\epsilon_{GAUSSIAN} = 7.9\%; \epsilon_{COS} = 7\%$$



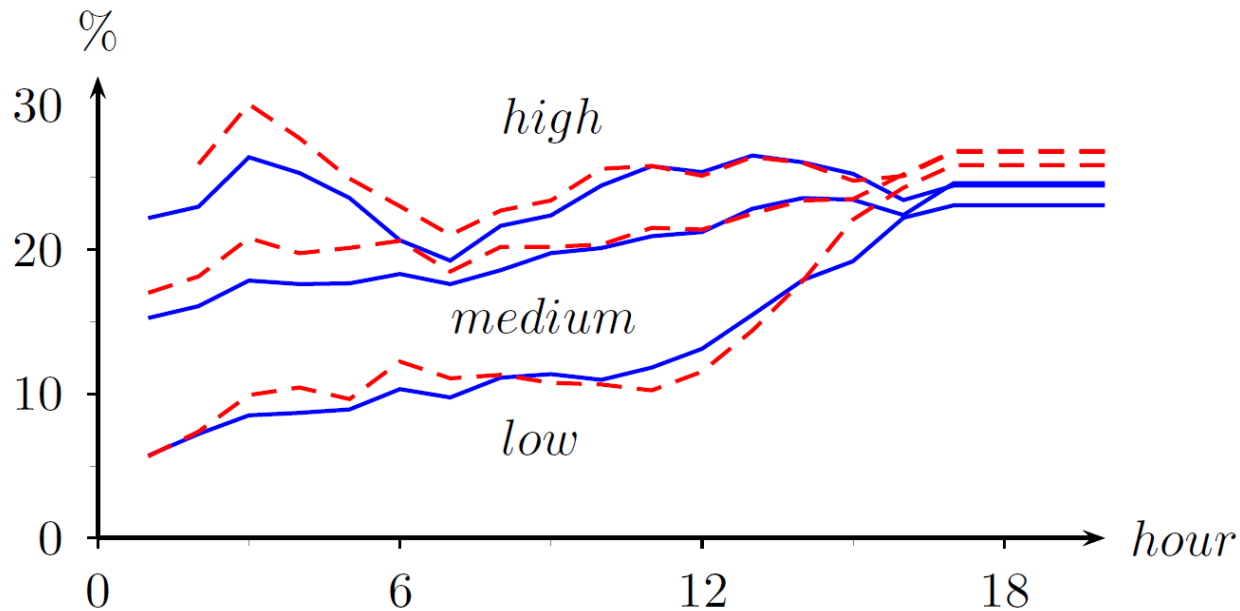
Medium perturbation

$$\epsilon_{GAUSSIAN} = 22\%; \epsilon_{COS} = 24\%$$



Green circles: data used for fitting; Black circles: missing data;
Blue line: Gaussian approach; Red dashed line: cosine approach

Mean relative error



Mean relative error according to failure duration for each level of perturbation

Blue line: Gaussian approach

Red dashed line: cosine approach

Overall mean relative error

- Average error weighted by failure distribution

| Perturbation level | Gaussian fit error | Cosine fit error |
|--------------------|--------------------|------------------|
| Low | 8.0% | 8.33% |
| Medium | 16.8% | 18.6% |
| High | 23.1% | 26.2% |

- Gaussian approach more accurate

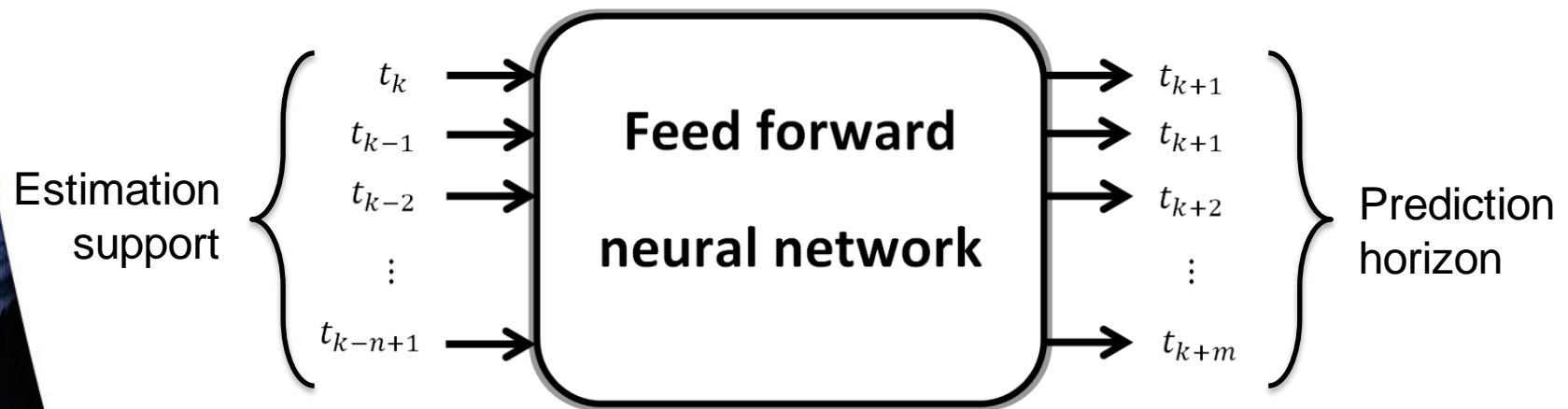


Temperature estimation

- Estimation of indoor or outdoor temperature
- Real-time estimation based on the concept of time series
- Flexibility, no need to wait failure ending
- Developed with artificial neural networks (multilayer Perceptron)

Concept of time series

- Neural network handled by *GNU Octave* and *nnet* package



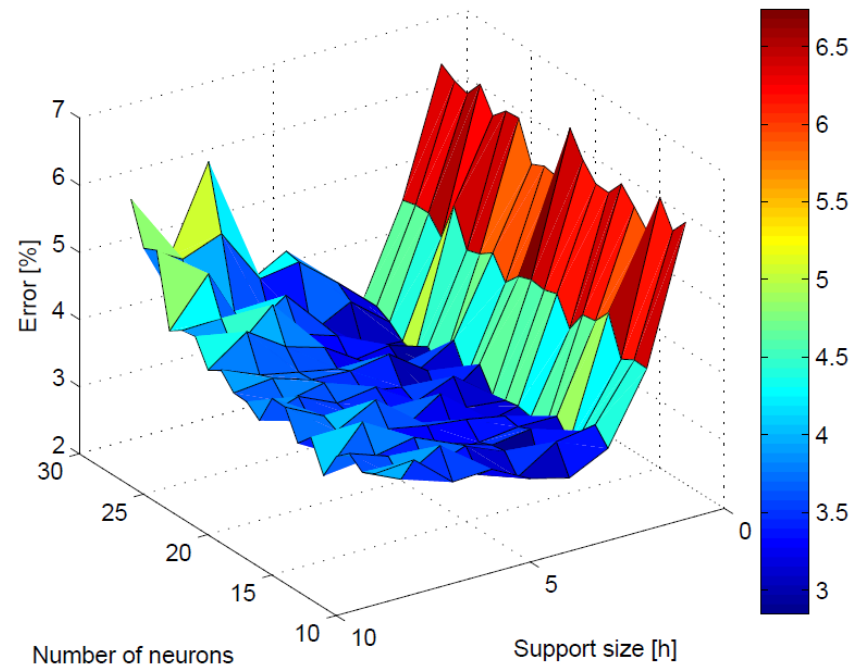


Neural network training

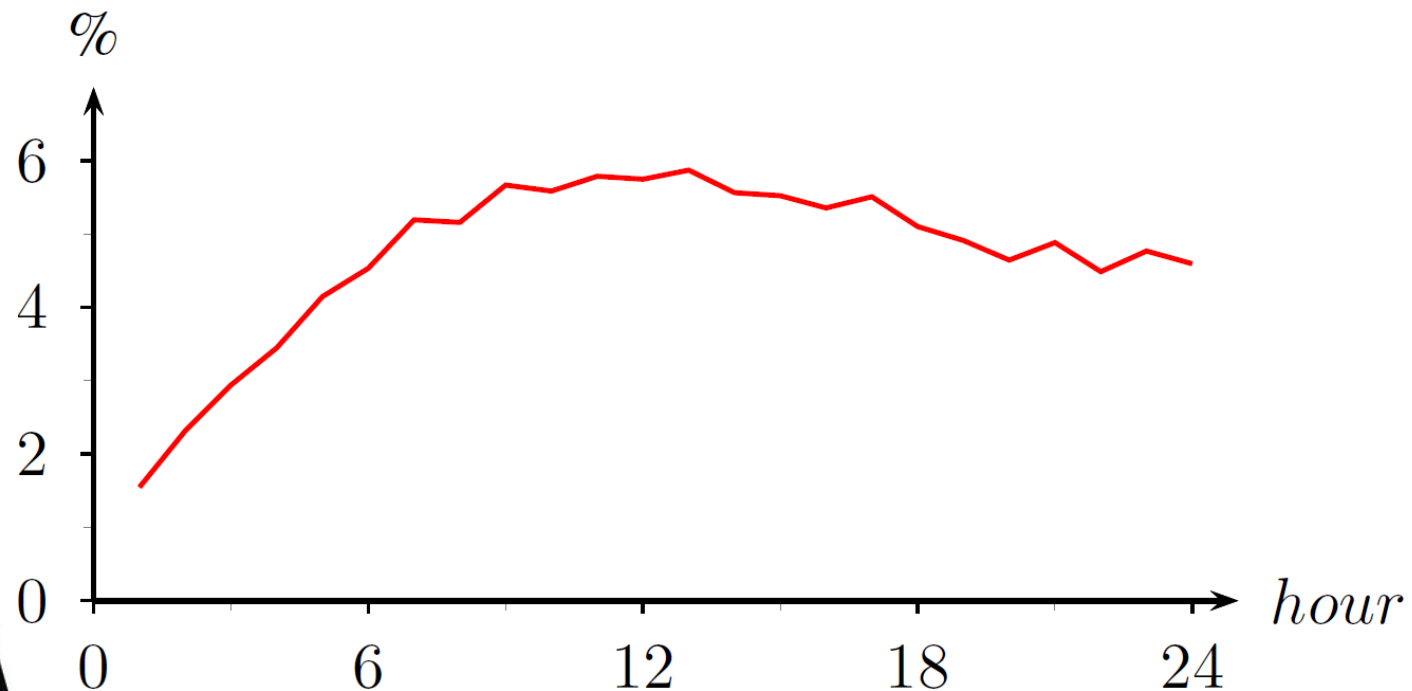
- Training phase using Levenberg-Marquardt algorithm
- Training examples covering a whole year (January to August)
- Parametric study to find optimal network topology by adjusting:
 - Size of the estimation support
 - Number of hidden neurons
 - Number of training examples

Mean relative error

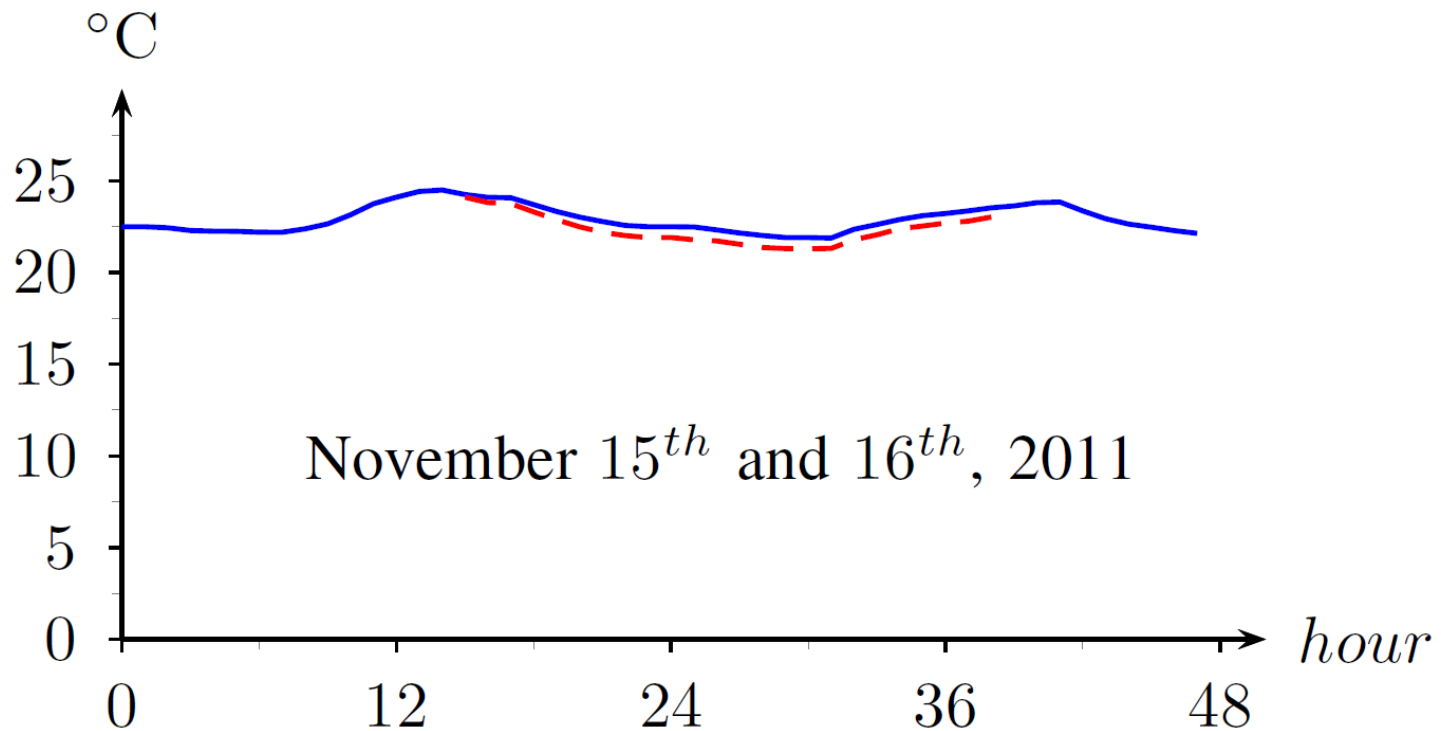
Best results with **200 training examples**, a **3-hour estimation support** and **15 to 20 hidden neurons**



Estimation error according to horizon with optimal topology



Data estimation for a 24-hour failure



Blue line: measured data; Red dashed line: estimated data

Conclusion

- Solar radiation estimation:
 - Efficient tool
 - For short failures, mean relative error ranges from 3 to 8%
 - At least 4 or 5 valid values to preserve accuracy
 - Difficulties to model very cloudy days

Temperature estimation:

- Good accuracy up to a 24-hour failure
- With optimal topology, error lower than 6%



Prospects

- Humidity estimation to compute thermal comfort
- Building modelling using *Energy+* software
- Predictive HVAC system regulation in tertiary buildings



Thanks for your attention

Any questions?