Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Thermodynamic Identification of Buildings using Wireless Sensor Networks

E. Toffoli G. Baldan G. Albertin L. Schenato * A. Chiuso ** A. Beghi *

* *Department of Information Engineering, University of Padova emails: {toffolie|gbaldani|galberti|schenato|beghi}@dei.unipd.it*
** *Department of Management and Engineering, University of Padova, Vicenza, Italy, email: chiuso@dei.unipd.it*

Abstract: In this paper we study different strategies for identifying thermodynamic models of buildings using experimental data collected from large scale wireless sensor networks. Wireless sensor networks can easily provide temperature, humidity and solar radiation measurements from tens to hundreds of sensors, thus potentially providing a fine-grained spatial-temporal resolution. In order to cope with such a large number of inputs and outputs, we tested subspace identification algorithms which are suitable for identifying large scale MIMO systems. The identified model can be used to evaluate the thermodynamic efficiency of the building. We also explore different sensor selection strategies in order to choose among all sensors the most informative ones. Using a small set of sensors not only greatly reduces computational burden in the identification algorithms, but can also be used to predict with high accuracy the measurements of the other sensors using Kalman filters. The identification algorithms, the sensor selection strategies, and the Kalman filter adopted have been tested and compared using experimental data collected from 65 sensors deployed in a $80m^2 - 200m^3$ building over an 11 day period.

Keywords: Subspace methods; distributed system identification; sensors networks

1. INTRODUCTION

Steady increasing prices of energy resources and growing environmental concerns about climate changes is putting much attention on energy conservation policies and technologies that can improve energy efficiency with low negative environmental effects. In particular, energy expenditure for temperature control in buildings account for up to 30% of total budget and it is doomed to increase with the proliferation of air conditioning systems. Therefore, there is a need to use energy more efficiently and in a cleaner manner in both new and old buildings, as testified also by a recent European Community Directive 2002/91/EC which imposes several actions in these directions. Therefore, there is a strong need to develop both technologies and tools that can provide:

- a-posteriori thermal efficiency of a building, i.e. a certification based on experimental data
- thermal monitoring and comfort control systems especially in large building
- energy saving quantification after remodeling and energy-specific retrofitting of existing buildings
- automatic fault-detection and monitoring of HVAC systems

Wireless sensor networks (WSNs) seem a particularly useful technology in this prospect. In fact, a WSN is a network of small devices, called motes, provided with sensors (temperature, humidity and solar radiation sensors), a microcontroller, some memory and I/O ports, and a wireless antenna which allow them to communicate with their neighbors. WSNs are easy to deploy since they are battery powered, they do not need to be placed in specific locations since the network is self-configurable and adaptive, they are non-intrusive since each device is smaller than a cigarette packet, and finally they are quite inexpensive. As a consequence a WSN, by avoiding the need of cabling, can be rapidly installed also in existing buildings with minor costs and intrusion, and collect measurements from hundreds of locations for long periods of time. Such measurements thus provide an unprecedented quantity of information that can be used to identify a fine-grained model of the building and to certificate its thermal efficiency. Moreover, it is possible to envision the use of WSNs not only for thermal efficiency certification, but also for collecting data for realtime thermal monitoring and regulation systems especially in large buildings. However, the effective use of WSNs for thermodynamic identification requires the development of novel mathematical tools that can cope with such a large number of sensors. So far thermodynamic identification of buildings have been developed based on data collected from a small number of sensors, mainly due to the fact that measurement collection is expensive and time-consuming. The most popular tools adopted for thermodynamics identifications of buildings are based on ARX, ARMAX and Neural Networks models (see Dodier and Henze (2004)); these two latter model

* This work was partially supported by EU fund MIRC-6-CT-2005-014815 "SENSNET", by the Italian CaRiPaRo Foundation, and by the national project *New techniques and applications of identification and adaptive control* funded by MIUR.

classes turn to be particularly difficult to handle when the number of inputs and outputs grow very large, leading also to ill-conditioned estimation problems when the inputs and outputs are highly correlated, as it is the case for measurements collected from sensors which are closely located. Moreover, the choice of the location of sensors from which data for identification is collected, is generally based on experience and rule of thumbs, and little have been done to experimentally evaluate which are the most informative locations where to place the sensors.

In this work, we propose to use WSNs for collecting data for thermodynamic identification of building since they can be used to rapidly collect measurements from a large number of sensors. In order to cope with large number of sensors we adopted recently developed subspace identification tools that when compared to the traditional methods mentioned above, have the advantage to be numerically efficient also for large scale MIMO systems. To our knowledge this is the first attempt to apply subspace methods for identification of thermodynamical models of buildings. Then we illustrate the problems associated with optimal sensor selection in terms of extracting the most informative sensors from an identification perspective. Finally, we also show how a small number of sensors, if appropriately chosen, can predict with very high accuracy the readings of all the other sensors by using Kalman filters, thus providing a useful tool that can be used to close the loop around a thermoregulation systems. The proposed methodologies were tested using experimental data collected from 65 sensors deployed in a $80m^2-200m^3$ building over an 11 day period. The main limitation of these experiments is that data where collected in open-loop thermodynamical conditions of the building, i.e. the temperature inside the building was not regulated by any heating/cooling system therefore the state of the building evolved due to the external temperature, humidity and solar insulation. Although the very goal of identification of building thermodynamic models is the energy efficiency of the building under closed-loop conditions, i.e. when the heating/cooling regulation is in place, we believe that the methodologies proposed can be readily extended by simply including as inputs also the external loads of the heating/cooling system.

2. EXPERIMENTAL TESTBED AND DATA COLLECTION

As mentioned in the previous section, we tested the identification techniques based on experimental data collected from a real building. The edifice we took into exam is a small two-floor residential building of about 80 m^2 and 200 m^3 whose ichnography and picture are presented in Figure 1. The building is situated in Padova (Italy) at a latitude of 45.41 °N and the climatic zone is characterized by 2383 degrees day referred to nominal temperature of 20 °C. The experimental data was collected through a WSN made of 65 *Tmote-Sky* nodes produced by Moteiv Inc. Each Tmote-Sky is provided with a temperature sensor, a humidity sensor, and a total solar radiation photoreceptor (visible + infrared). In our identification experiments we did not use the humidity sensors.

The data was collected under ideal conditions. In particular, during the data collection period, the building was not inhabited, and all external windows and doors were closed. This prevented the natural thermal dynamics to be disturbed by nonlinear and unpredictable phenomena due to air exchange with the external environment. Also we could not use the building thermoregulation system because it was out of order. This meant that the inputs to the system were given only by the external environment under the form of external temperature and sun radiance. Since the thermal dynamics of the building evolved only due to uncontrollable external conditions, it was not possible to strongly excite the system as it would have been possible with the thermoregulation system in place, thus obtaining data which are ill-conditioned for identification purposes.

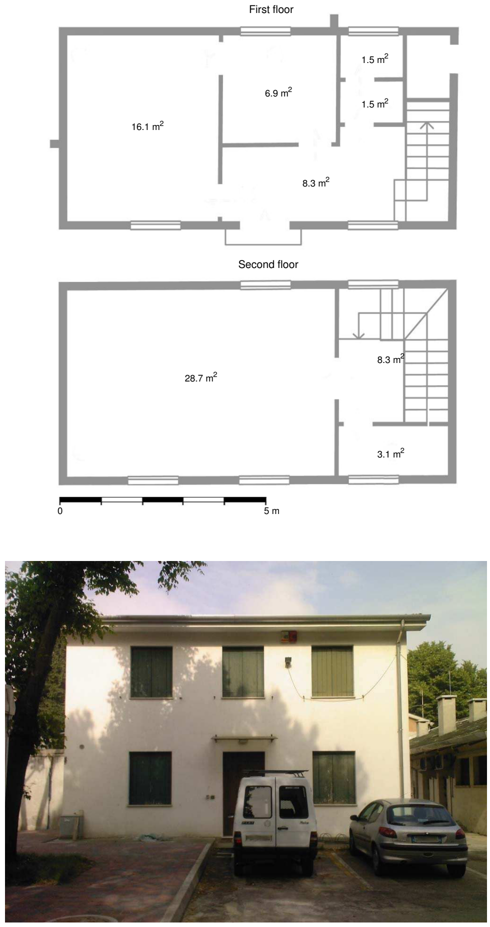


Fig. 1. Ichnography(top) and picture (bottom) of the edifice used as test case.

Out of the 65 sensor nodes, one sensor measuring temperature and one sensor measuring total solar radiation were placed on each wall on the outer surface of the building at an height of about 4.5 m. The remaining 57 sensors, used as temperature sensor only, were positioned inside the edifice. In particular, they were uniformly distributed in the space so that the resulting model could describe precisely the temperature in each part of the edifice, and some of them were placed in the proximity of windows and doors where the heat exchange is larger.

Due to logistical problem, we could perform only a single experimental measurement during a period of 11 days starting at 11.00pm, June 15th, 2007 and lasted until 10.00am, June 26th, 2007. We used a sampling time of 10 minutes. As mentioned above, data collected offered poor model input excitation because during measurement the weather was quite uniform with a nice sun shining every day. The external high temperature was always around 31 °C while the low was about 24 °C. The mean internal temperature of the building increased every day passing from about 25 °C the first day (low temperature due to the perturbation induced by sensor placement activity) to about 28 °C the last one.

3. SUBSPACE IDENTIFICATION WITH INPUT SELECTION

We shall model the thermodynamic behavior of the building as a discrete time, time invariant dynamical linear model in state space form with exogenous inputs. The inputs $u(t) \in \mathbb{R}^m$ of the model are a subset chosen from one or more of the following classes:

- 4 external temperature sensors
- 4 external total solar radiation (visible and infrared spectrum)
- 4 internal temperature sensors placed on the ground floor.

The outputs $y(t) \in \mathbb{R}^l$ are the (remaining) internal temperatures inside the building. All inputs have been properly scaled in order to avoid numerical ill-conditioning. The rational behind the use of temperature sensors as inputs is that they are indirectly related to the heat exchange between the building and the environment.

We assume data $\{y(t), u(t)\}$, $t = 1, ..., N$ are available and we consider the dynamical model (in innovation form)

$$\begin{cases} x(t+1) = A\,x(t) + B\,u(t) + K\,e(t) \\ y(t) \quad\;\; = C\,x(t) + e(t) \end{cases} \tag{1}$$

where $e(t)$ is the one-step-ahead prediction error, which is a zero-mean white noise. We also assumed that there there is one time delay in the transfer function from $u(t)$ to $y(t)$. The dimension of the state space is denoted by n and all matrices are sized accordingly.

The number of inputs and outputs is large and hence we decided to use subspace identification techniques. These methods are based on robust, non iterative and numerically efficient linear algebra tools which, contrary to other methods based on the optimization of some cost function (e.g. Prediction Error Methods, see Ljung (1997)) do not require performing costly iterative minimization thus also avoiding the risk of getting stuck in local minima, see e.g. Van Overschee and De Moor (1996),Chiuso (2007). In particular, we compared the MATLAB® System Identification Toolbox `n4sid.m` routine [1] and a recursive version of the $PBSID_{opt}$ algorithm in Chiuso (2007).

As mentioned in the introduction, using a large number of inputs and outputs data can potentially provide a great wealth of information to obtain a detailed model for the building thermodynamics. However, if the inputs and/or the outputs of the model are highly correlated we may incur in severe numerical problems due to collinearity. This may happen, for instance, when the sensors are closely positioned. This collinearity usually results in estimated models which are very sensitive to the available data (i.e. the estimators' variance will be large). Indeed, some preliminary identification tests confirmed that the model identified with the full set of inputs provided lower prediction performance in fitting the validation data, as compared to models identified with only a subset of total inputs. One solution to address this problem is to force the model to be "simple" by adding some regularization terms or by forcing the model to have only a small number of parameters, e.g. by selecting only a fraction of all possible inputs (or outputs). Here we follow this second approach being more suitable for subspace identification algorithms.

Note that, if one knew the linear model, one could obtain a measure of relative importance of the single input among the others (see the survey by van de Wal and de Jager (2001) for a presentation of various methodologies). However, the system identification step is precisely where the input selection process plays a key role.

Methods for avoiding the collinearity problems include extensions of Principal Component Regression (see e.g. Greenberg (1975)), of PLS (see Wold (1966)) and its dynamic extensions (see e.g. Qin (1998) and references therein).

In this preliminary work we took a simple route which we describe next, leaving to future work analysis and development of more sophisticated techniques. The input selection has been achieved by constructing, for each candidate subset, a linear state space model and electing the best-fitting (on validation data) with respect to some sort of metric. The metric chosen for the model M is the scalar variance of the simulation error, that is

$$\text{fit}(M) = \text{trace}\left(E[e_u(t)e_u^\top(t)]\right). \tag{2}$$

Of course we approximate $E[e_u(t)e_u^\top(t)]$ with the sample variance of the fitting error, i.e. $e_u(t) = y(t) - y_u(t) = y(t) - C\,x_u(t)$, where $x_u(t)$ is the state obtained by setting $e(t) = 0$ and by using the identified initial condition [2] $\hat{x}(0)$ in model dynamics given by equations (1).

Exhaustive search over all possible inputs combinations is a combinatorial problem and is not a viable approach. Instead we adopted an iterative greedy approach to the selection problem. The candidate inputs have been divided in three classes: external temperatures, external solar radiation and internal temperatures. For each class we select the best input in term of the metric proposed above. We remove this input from the candidate input set and we place it into the selected input set for identification. Then we repeat the process until the desired number of inputs have been selected (or until the validation metric increases). Experimental evidences showed us that the chosen subset is mostly independent from the order of class selection. The results based on the previous identification methodologies on validation data set are shown in Figure 2 and Figure 3. In particular, in Figure 2, we compared the simulation performance of the two models identified with `n4sid.m` and $PBSID_{opt}$ by using the best input for

[1] The current implementation in Matlab is actually a mixture of the most well known methods, i.e. N4SID Van Overschee and De Moor (1994), MOESP Verhaegen (1994) and CVA Larimore (1990); different choice are the user parameters are performed based on the options. We used the "default" Matlab choice besides that we had to force stability on some cases.

[2] The initial condition could also have been re-estimated for the validation data.

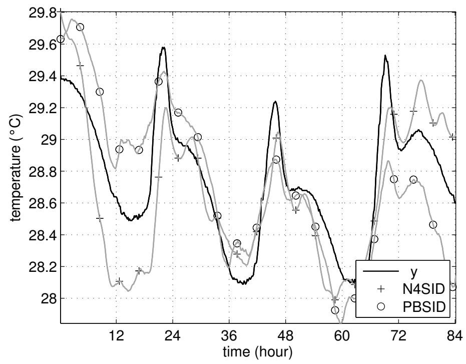


Fig. 2. Temperature measured by one internal sensor, i.e. one entry of output vector y, and simulated temperature y_u using the models identified by `n4sid.m` and $PBSID_{opt}$.

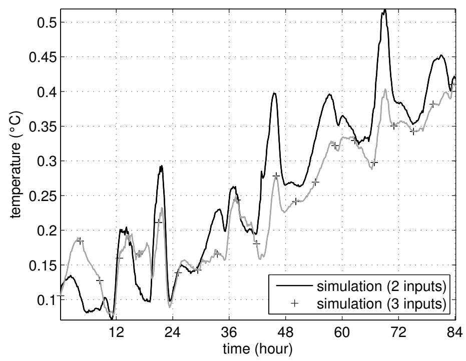


Fig. 3. Simulation error fit(M) based on the model identified by $PBSID_{opt}$ using 2 and 3 inputs.

each of the three classes. In order to have acceptable performance from `n4sid.m` we had to force stability and use a high order model (namely at least as big as the number of outputs) Differently, the model identified by $PBSID_{opt}$ was always asymptotically stable and showed limited output error even with a low order model (namely about one fifth of the number of outputs). It has been observed that $PBSID_{opt}$ performed consistently better than `n4sid.m` under different testing conditions, therefore we shall focus only on $PBSID_{opt}$ in the remaining part of the paper. In Figure 3 we show the simulation performance of the model identified (using $PBSID_{opt}$) as the number of inputs is increased. In the 2-input model we used the most informative external temperature sensor and internal ground temperature sensor, while in the 3-input model we added the most informative radiation sensor. In is interesting to mention that the input selection algorithm found that most informative external temperature sensor was the one placed on the south side and not on the north side as commonly suggested (see e.g. AA. VV. (1994)). This might be the result of the particular data-set we used, however most common sensor placement strategies are based more on experience rather than mathematical analysis and it deserves more investigation.

4. PREDICTION VIA KALMAN FILTERING

It is a well known fact that simulation alone may perform poorly when initial condition are unknown or when "external" excitations, e.g. unmeasured environmental changes in our setup, perturb the system. It has been in fact observed that the simulated output of the model becomes unreliable within few days.

However in many applications it could reasonable to assume that few "internal" sensors are available as measurements and one would like to predict the temperature in other locations of the building. E.g. this could be used to the purpose of designing a controller which maintains the temperature as spatially homogeneous as possible while using only one or few internal sensors.

In this Section we shall design a Kalman estimator, based on the identified model, which use only a subset of the output vector $y(t)$ to predict the output of all the other sensors.

Let us denote with $\mathcal{L} \subseteq \{1, 2, \ldots, l\}$ the subset of outputs which will be used as measurements to estimate the state in the Kalman filter. Let $M_{[\mathcal{R},\mathcal{C}]}$ be the sub-matrix of the matrix $M \in \mathbb{R}^{p\times q}$ obtained by choosing the rows in the set $\mathcal{R} \subseteq \{1, 2, \ldots, p\}$ and columns in the set $\mathcal{C} \subseteq \{1, 2, \ldots, q\}$. Similarly, $V_\mathcal{R}$ will be the vector obtained by choosing entries in set $\mathcal{R}$ of the column vector V. The dot inside a square bracket, as in $M_{[\mathcal{R},\cdot]}$, indicates that all columns of the original matrix have been retained; a similar notation holds for the rows. Based on equation (1), we build the reduced linear system

$$\begin{cases} x(t+1) \;=\; A\,x(t) + B\,u(t) + K_{[\cdot,\mathcal{L}]}\,e_{[\mathcal{L}]}(t) \\ y_{[\mathcal{L}]}(t) \quad\;=\; C_{[\mathcal{L},\cdot]}\,x(t) + e_{[\mathcal{L}]}(t) \end{cases} \tag{3}$$

with model error covariance matrix $\Lambda_{[\mathcal{L},\mathcal{L}]}$, $\Lambda = E[e(t)e^\top(t)]$.

The initial condition μ_0 has been chosen such that it captures the initial "average temperature" (i.e. the mean of the sensor measurements in the set $\mathcal{L}$ at time zero). The initial state covariance matrix P_0 was instead chosen by a trial and error procedure.

Again we face the problem of selecting a small subset of outputs from a large set, yet providing good predictive performance. That is, we have to select $\mathcal{L}$ in such a way to reduce the global error on internal temperature estimates. A natural way to evaluate the quality of the set $\mathcal{L}$ is to measure its ability in predicting all the output measurements. This can be achieved by computing, using the filtering algebraic Riccati equation (ARE) Anderson and Moore (1979), the (steady-state) state error covariance $P(\mathcal{L})$, i.e. the state error covariance when only the measurements in the set $\mathcal{L}$ are used. The corresponding output prediction [3] error variance is given by $\Phi(\mathcal{L}) = CP(\mathcal{L})C^\top + \Lambda$. We then minimize the scalar measure:

$$J_d(\mathcal{L}) = \mathrm{tr}(\Phi(\mathcal{L})). \tag{4}$$

Alternatively one could consider the sample version of the above, i.e.

$$J_f(\mathcal{L}) = \sum_t ||\hat{y}(t) - y(t)||_2. \tag{5}$$

Finding the optimal set of output sensors which minimizes one of the previous metrics is again combinatorial problem, therefore some efficient suboptimal strategy is required. In this work we explored three different strategies. All these

[3] One could, alternatively, consider the filtering error variance.

strategies divide the sensors into two sets: the selected sensor set $\mathcal{S}$ and the remaining ones $\mathcal{R}$.

Greedy Search (GrS) It starts with an empty set $\mathcal{S}$ and then sequentially finds among all the sensors in $\mathcal{R}$ the one that provides best performance when added to the set $\mathcal{S}$. Once this sensor is found, it is removed from $\mathcal{R}$ and placed in $\mathcal{S}$ and the procedure repeated until the desired number of sensors have been selected.

Local Search (LS) The algorithm starts with $\mathcal{S}$ chosen at random. Then it sequentially swaps one sensor between the two sets $\mathcal{R}$ and $\mathcal{S}$ and finds, among all possible swapping combinations, the one that leads to the best performance improvement. Then the best swapping is actually performed between the two sets and the process repeated, avoiding to search previous swapping combination. This procedure is guaranteed to improve performance and every step and it stops when a local minimum is reached or a certain number of iterations have been performed.

Genetic Search (GeS) The previous strategy based on LS is likely to end up in local minima. To reduce this risk we adopted a genetic algorithm, see e.g. Goldberg (1989), to find good sensors swapping between the sets $\mathcal{R}$ and $\mathcal{S}$. In particular, it starts with several candidate sets $\mathcal{S}$, called populations, and then it swaps sensors among them (breeding and mutation) and only the best performing new populations are likely to survive. This process is continued till no major improvements are observed or a certain number of iterations have been performed. Although these algorithms are based on heuristics and are not guaranteed to find the global minimum, they often lead to good performance.

We applied the three previous algorithms in sequence, i.e. we used the solution of GrS to initialize LS, and the solution of LS to initialize GeS. The optimal solution has been used to compare the performance between simulation based only on inputs $u(t)$ and prediction using the Kalman filter based on the outputs $y_{[\mathcal{L}]}$. The dynamical model used for both the simulation and the Kalman predictor was obtained using the $PBSID_{opt}$ algorithm with three inputs chosen based on the greedy algorithm described in the previous section. Figure 4 shows the real temperature of a sensor not included in the 5 outputs $y_{[\mathcal{L}]}$ used by the Kalman filter and the corresponding y_u and $\hat{y}$ given by the open-loop simulation based only on the input u and the Kalman predictor, respectively. The improvements given by the Kalman filter are evident, in fact the use of only 5 sensors is sufficient to reconstruct the temperature of all 57 sensors with high precision.

Figures 5 (single sensor output) and 6 (mean square error) show that the prediction error improves as the number of sensors increase, in particular during the transient period. However, even with only two sensors the mean square error is smaller than half a degree during transient period and smaller than a tenth of degree at steady-state. Of course, such a small error is also a consequence of the specific experimental conditions, i.e. high sensor density and unpopulated building. However, it suggests that the linear model identified by the $PBSID_{opt}$ algorithm is rather effective to describe the behavior of building thermodynamics, in particular when paired with Kalman filtering.

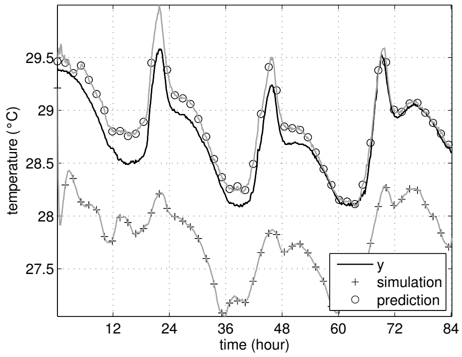


Fig. 4. Temperature y measured by one internal sensor not included in the output set $\mathcal{L}$ vs. y_u simulated using the PBSID$_{opt}$ model and $\hat{y}$ predicted by the Kalman filter using 5 outputs in the set $\mathcal{L}$.

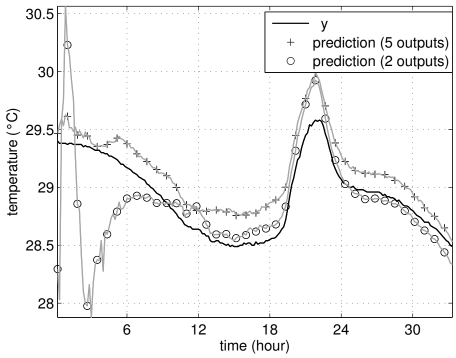


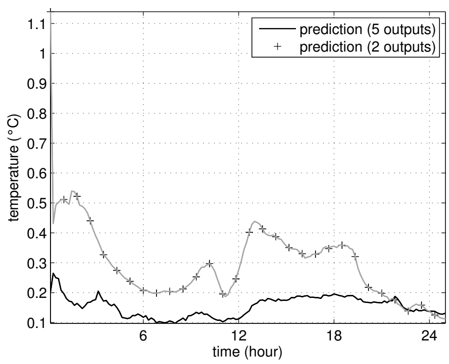
Fig. 5. Output y of a single sensor (not in the set $\mathcal{L}$) vs. the predicted output $\hat{y}$ obtained from Kalman filter based on 2 and 5 outputs in the set $\mathcal{L}$.

Finally, Figure 7 show the mean square error using the 3 best output sensors selected by the three strategies described above. The LS algorithm always offered great enhancements while the GeS presented some problems connected to the parameters calibration. We also noticed that sometimes the application of a random starting solution to the local search produced better results if compared with the ones given by the solution of greedy algorithm.

The previous results are based on the metric $J_d(\mathcal{L})$, however we observed that there was substantial agreement with the empirical cost $J_f(\mathcal{L})$.

5. CONCLUSION AND FUTURE WORK

In this paper, we proposed the use of WSNs in combination with subspace methods for identification of building thermodynamics. In fact, WSNs provide a mean to rapidly and inexpensively collect measurements from hundreds of temperature, humidity and light radiation sensors for long period of times. Although such a great wealth of data potentially provide fine-grained information about building thermodynamics, it also poses novel challenging problems, in particular in terms of model identification. In fact, the number of inputs and outputs provided by WSNs are at least an order of magnitude larger than the number than traditional identification tools for building thermodynamics can handle (AA. VV. (1994)). In this paper we



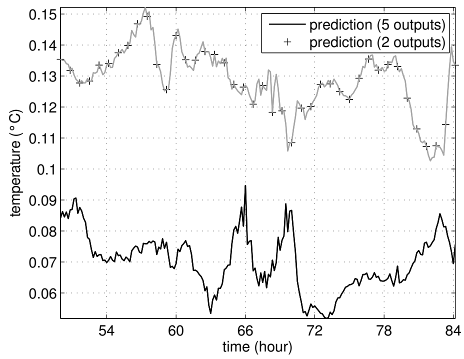


Fig. 6. Average prediction error for all sensors not included in the set $\mathcal{L}$ during transient period (*top*) and at steady-state (*bottom*). Predicted output $\hat{y}(t)$ using the Kalman filter with 2 and 5 outputs.

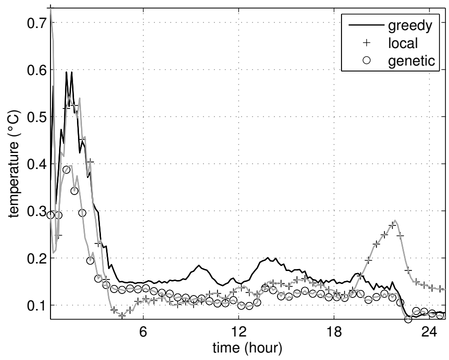


Fig. 7. Predictive performance (on a single output) of Kalman filter with subsets selected by different heuristic algorithms.

proposed to address this problem by adopting subspace identification tools which have been recently developed for identifying large scale MIMO systems. Indeed, we believe that identification of building thermodynamics can be a very useful test bed to evaluate and possible improve subspace identification algorithms. In this work we tested a standard subspace method (the Matlab `n4sid.m`) and a recently developed subspace method (PBSID$_{opt}$, Chiuso (2007)) and we observed that the latter systematically outperformed the former, however a deeper investigation is still required. Also we found that these tools did not perform well when the number of inputs and outputs was very large, mainly due to the fact that inputs and outputs were highly correlated. Therefore we proposed some heuristic for input selection for identification purposes, however more systematic and mathematically sound tools are necessary; we envision that systematic extension of principal component regression (PCR) and Partial Least Squares (PLS) might give significant improvements and hence will be subject of future research.

The simulation and prediction performance obtained by using only a properly chosen subset of inputs and outputs are remarkable. Although this is also a result of the special experimental conditions of the building, we believe that the subspace identification is a viable and effective solution, and we are currently performing more realistic experiments.

Summarizing, this paper poses more questions than answers, mainly due to the fact that WSNs open up new problems and challenges to the identification building thermodynamics community. However, we also believe that major improvements and advancements are to come.

REFERENCES

AA. VV. *System Identification applied to building performance data.* European Community, Luxembourg, 1994.

B.D.O. Anderson and J.B. Moore. *Optimal Filtering.* Prentice Hall, 1979.

A. Chiuso. The role of Vector AutoRegressive modeling in subspace identification. *Automatica*, 43(6):1034–1048, June 2007.

R.H. Dodier and G.P. Henze. Statistical analysis of neural networks as applied to building energy prediction. *J. of Solar Energy Eng.*, 126(1):77–83, 2004.

David E. Goldberg. *Genetic Algorithms in Search, Optimization, and Machine Learning.* Addison-Wesley Professional, 1989.

Edward Greenberg. Minimum variance properties of principal component regression. *Journal of the American Statistical Association*, 70(349):194–197, 1975.

W.E. Larimore. Canonical variate analysis in identification, filtering, and adaptive control. In *Proc. 29th IEEE Conf. Decision & Control*, pages 596–604, Honolulu, 1990.

L. Ljung. *System Identification, Theory for the User.* Prentice Hall, 1997.

J.S. Qin. Partial least squares regression for recursive system identification. In *Proc. of IEEE CDC*, S. Antonio, Texas (USA), 1998.

M. van de Wal and B. de Jager. A review of methods for input/output selection. *Automatica*, 37(4):487–510, 2001.

P. Van Overschee and B. De Moor. N4SID: Subspace algorithms for the identification of combined deterministic–stochastic systems. *Automatica*, 30:75–93, 1994.

P. Van Overschee and B. De Moor. *Subspace Identification for Linear Systems.* Kluwer Academic Publications, 1996.

M. Verhaegen. Identification of the deterministic part of MIMO state space models given in innovations form from input-output data. *Automatica*, 30:61–74, 1994.

H. Wold. *Research Papers in Statistics*, chapter Nonlinear estimation by iterative least squares procedures. Wiley, New York, 1966.