Lukáš Ferkl

Identification and Control of Large-Scale Systems

Czech Technical University in Prague
Faculty of Electrical Engineering
Department of Control Engineering

Identification and Control of Large-Scale Systems

Habilitation thesis

Lukáš Ferkl, 2011

Copyright

Abstract

This habilitation thesis describes the methods used for system identification and control of large-scale systems. These systems require special treatment with respect to their multiple-input, multiple-output nature. In system identification, special attention is paid to subspace identification methods. Subsequently, these models are used in Model-based Predictive Control, which enables efficient control of the modeled systems. This approach is shown in practice for tunnel ventilation control, as well as building heating control.

Anotace

Tato habilitace se zabývá metodami používanými pro identifikaci systémů a řízení rozlehlých systémů. Tyto systémy vyžadují zvláštní přístup s ohledem na jejich více–vstupovou a více–výstupovou povahu. V oblasti identifikace systémů se zaměříme především na metody subspace identifikace. Tyto modely jsou poté použity v prediktivním řízení MPC (Model-based Predictive Control), které umožňuje efektivní řízení modelovaných systémů. Tento přístup je ukázán na příkladech řízení ventilace tunelů a řízení vytápění budov.

Acknowledgements

The topics presented in this book are results of a team work of the group called ProjectEleven, at Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague. I would like to express my special gratitude to all who helped me in writing this book, especially Jiří Cigler, Jakub Kubeček, Ondřej Nývlt, Samuel Prívara and Zdeněk Váňa.

The practical examples throughout this book are published with the kind permission of Energocentrum Plus, a company administrating various HVAC systems, whose curiosity for the novel control approaches gave birth to our first MPC application in buildings; and Satra, a construction company with similar enthusiasm in tunnel ventilation control. Special thanks belong to Jan Široký and Martin Chlupáč, researchers at Energocentrum Plus, and Jan Pořízek and Jiří Zápařka, researchers at Satra, who provided me with all necessary technical support.

I would also like to thank all the other people who have been cooperation with me on my research, especially Gjerrit Meinsma from Twente Universiteit; Lieve Helsen and Clara Verhelst from KU Leuven; Dimitrios Gyalistras and Frauke Oldewurtel from ETH Zürich; Marvin Rausand from NTNU Trondheim; and Josef Komárek and Jan Rakušan from Energoklastr.

Last, but not least, I would like to thank my wife Blanka and my children Jan and Markéta "just for everything".

The research presented in this habilitation thesis has been supported by the following grants:

- *Integration of building systems, research and application of intelligent algorithms with influence on energy consumption of buildings and living houses (MPO, 2A-1TP1/084)*
- *Control systems for energy consumption optimization in low-energy and passive houses (MPO TIP, FR-TI1/517)*
- *Traffic Infrastructure: Mathematical Methods for Environmental Impact Analysis (Financial Mechanisms of EEA/Norway, B/CZ0046/2/0015)*
- *Individual mobility grant (Financial Mechanisms of EEA/Norway, 4/017/09/1)*
- *Methods for Optimization of Energy Consumption in Buildings (Financial Mechanisms of EEA/Norway, B/CZ0046/3/0015)*
- *Energoklastr (OPPI Klastry, 5.1 SPK01/024)*
- *GEOTABS (7^{th} call of EraSME)*

Following companies have also provided their finances to support the research described in this habilitation thesis:

- *Energocentrum Plus, s.r.o.*
- *Feramat Cybernetics, s.r.o.*
- *ICADE, GmBH*
- *Gruner, AG*
- *Satra, s.r.o.*

Contents

List of Appendices

Comments to the Published Works

1 Introduction

Recent trends in carbon dioxide emission reduction establish a socially challenging environment for new ideas for energy consumption optimization, regardless of our individual opinion about the global warming issue. Special attention is paid to various technology advances, such as low-emission engines, green energy generation, etc. Both economical and environmental revenue of such innovations is sometimes subject to fierce discussions. On the other hand, improvements of control system algorithms leading to economic savings represent a solution, which is "green", beyond dispute.

For example, buildings have been subject to energy savings for a long time. Indeed, according to the U. S. Energy Information Administration, in 2005, buildings accounted for 39 % of total energy usage, 12 % of the total water consumption, 68 % of total electricity consumption, and 38 % of the carbon dioxide emissions in the U. S. A. (E.I.A., 2009). However, recent efforts have focused mainly on the construction of the buildings, such as better insulations, double facades, heat-reflecting glass, etc. But large buildings, such as schools, hospitals, or office buildings, are usually equipped with control systems based on general industrial systems and use industrial protocols, PLCs, dense sensor networks, etc., which enable the use of sophisticated modifications of state-of-the-art controllers (PID control, weather-compensated control, …). The use of modern control strategies in buildings has been very limited, with the exception of fuzzy control. More sophisticated methods, such as LQG or MPC control, have been enduring their use in buildings mainly because of the absence of convenient tools for obtaining a model of the building.

This habilitation thesis encloses the major publications of the applicant supported by a short commentary. The work is focused mainly on finding proper models of systems, which is crucial for model-based control. Section 2 presents the main contribution of this thesis, which is system modeling and identification. The role of the models in a model-based controller is shown in Section 3. Practical use of the model-based control is shown in Section 4. Contribution of the work and further work are summarized in Section 5.

Respective publications are supplemented in the appendices.

The first part is about finding optimal control strategy for ventilation in large highway tunnels, which is not necessarily predictive. We begin with the general concept of a static, optimal, model-based controller (Appendix A, **L. Ferkl** and Meinsma (2007)), with the application for tunnel ventilation (Appendix B, Pořízek, Zápařka, and **L. Ferkl** (2008)). As tunnels are systems with high demands for safety, the application of a controller in a real tunnel must be supported by a proper risk analysis, as shown in Appendix C (Nývlt, Prívara, and **L. Ferkl**, 2011).

The second part deals with identification issues with the focus of thermodynamic model identification of buildings. The possibilities of the state-of-the-art identification techniques are discussed in Appendix D (**L. Ferkl** and Široký, 2010). Particular challenges encountered during identification of a large building are described in Appendix E (Váňa, Kubeček, and **L. Ferkl**, 2010). Including some prior information into the subspace identification methods is described in Appendix F (Prívara, Cigler, Váňa, **L. Ferkl**, and Šebek, 2010).

As already stated, the modeling and control part form an indivisible complex in model-based control. Integration of subspace identification models into model-based controllers is described in Appendix K (Prívara, Váňa, Cigler, and **L. Ferkl**, 2011). The implementation of the Model-based Predictive Control is shown in Appendix G (**L. Ferkl**, Široký, and Prívara, 2010), while Appendix H (**L. Ferkl**, Verhelst, Helsen, Ciller, and Komárková, 2011) describes a method of performance estimate of the model-based control. Practical experience with operation of the Model-based Predictive Controller is described in Appendix I (Prívara, Široký, **L. Ferkl**, and Cigler, 2011a), and Appendix J (Prívara, Váňa, Gyalistras, Cigler, Sagerschnig, Morari, and **L. Ferkl**, 2011b) discusses the savings potential of the model-based control for a large office building, showing also a generalized approach to model-based control of large buildings.

2 Modelling and Identification

For a long time, linear system identification was a major field in system identification. In statistical identification, the roots date back to 1801, when Franz Xaver von Zach used the least-squares analysis proposed by Carl Friedrich Gauss (Gauss, 1809) to identify the orbit of the asteroid Ceres. In 1960s, autoregressive models became popular in systems and control theory, thanks to advances by K. J. Åström (e.g. Åström and Eykhoff (1971)) and his successors. The autoregressive models are suitable for single-input, single-output (SISO) systems identification, and many modifications and improvements exist, e.g. enabling incorporation of prior information based on maximum likelihood estimate or frequency domain identification (Ljung, 1999).

Another approach, which is suitable also for multiple-input, multiple-output (MIMO) systems, is represented by identification of linear, time-invariant state-space models by means of subspace identification (4SID). Even though the principles of 4SID has been known for decades – deterministic version was first described by Ho and Kalman (1966) and stochastic by Akaike (1974) – the method gained significant attention in 1990s (Verhaegen and Dewilde (1992), van Overschee and De Moor (1994)) and proved to be useful in many practical applications. Even though some interesting results have been published recently on 4SID, such as incorporation of prior information (Trnka and Havlena, 2009), number of papers being published on linear system identification has been declining since the beginning of the turn of the third millennium. A very good overview on the system identification topics, as of 2010, can be found in an excellent paper by Ljung (2010).

However, a new topic has been opened recently by the introduction of the concept of models that are suitable for control, while classical approach is to aim for models with good open-loop prediction capabilities. This concept is crucial for model-based controllers, such as MPC (Model-based Predictive Control). Even though the idea is not new (one of the first papers on this topic is the famous paper by Åström and Wittenmark (1973)), the concept has been thoroughly described in 1990s (van den Hof and Schrama (1995), Ljung (1999)) and the problem has been specified for the class of models suitable for MPC (Gopaluni et al., 2004; Shook et al., 1992). There has been some work done on closed-loop identification (MacArthur and Zhan, 2007), but the identified models still have very high orders.

In the present work, subspace identification approach is the principal method for finding a suitable model for a model-based controller. The principles can be found in many of the disclosed publications, e.g. Section II.A in Appendix G, Section 2 in Appendix D or Section II in Appendix F. In principle, the classical methods find system matrices, then estimate the system states, which often leads to high order models that have to be reduced afterwards. For MIMO systems, this approach is extremely time consuming. On the other hand, subspace approach uses orthogonal and oblique projections to find Kalman state sequence, and then obtain the system matrices using e.g. least squares method.

The problem of incorporating prior information, which is important for large scale models to be used in model-based controllers (as significant savings can be achieved from the knowledge of DC gains of the controlled system), was partly solved by Trnka and Havlena (2009) for MISO systems. The incorporation of DC gain knowledge for MIMO systems is shown in Appendix F.

3 Model-Based Control

The origins of model-based control, or more specifically the Model-based Predictive Control (MPC), traditionally date back to 1970's, when complex control of chemical processes was needed. Nowadays, it can be found in a wide variety of application areas including chemical industry, petroleum refineries, offshore platforms, food processing, metallurgy, pulp and paper, automotive and aerospace applications (Qin and Badgwell, 2003). The MPC principle, or the receding horizon principle, was proposed by Propoi (1963) back in 1960's and later elaborated by Richalet et al. (1976). Unfortunately, the optimization algorithms based on constrained linear and quadratic programming were a limiting factor in the sense of computational speed.

Recent advances in computation and mathematical optimization techniques have, however, opened new ways of dealing with these problems. One of the simplest is the certainty equivalent model predictive control (CE-MPC) (Bertsekas, 2005) that solves a deterministic optimization problem with stochastic disturbances replaced by their estimates based upon the information available at the time, and proceeds in a receding horizon fashion. Another popular class of control strategies is the affine disturbance feedback policy which turns out to be equivalent to the affine state-sequence feedback policy via a nonlinear transformation similar to the classical Q-design or Youla-Kučera parameterisation (Skaf and Boyd, 2009).

However convenient the paradigm of affine disturbance feedback may be, its use is prohibitive whenever unbounded stochastic disturbances enter the system in the presence of hard control input bounds, since then the linear part necessarily vanishes, which, in effect, renders the policy open-loop. One way to overcome this problem is to use a (saturated) nonlinear disturbance feedback (Skaf and Boyd, 2009), where this approach was developed for the quadratic cost. The upside of this generalization is the fact that the convexity of the cost function associated with the nonlinear disturbance feedback turns out to be independent of the choice of the nonlinear function.

Another branch of approximation techniques bounds the disturbances a priori and solves a robust MPC problem, while guaranteeing an open loop probabilistic bound on the performance (Bertsimas and Brown, 2007). This approach, however, tends to be very conservative, and thus the idea of bounding the disturbances a priori based on their distribution appears more often in the context of chance constraints Oldewurtel et al. (2008).

It is the issue of recursive feasibility of (probabilistic) constraints that has predominantly hampered bridging the gap between stochastic optimal control and constrained model predictive control. The crux of the matter lies in the fact that independent unbounded disturbances additively entering the system cannot give rise to a recursively feasible problem as long as the set of state constraints is compact and control authority bounded. Thus, one has to either develop a backup recovery policy that is triggered when infeasibility occurs, or assume compactly supported disturbances. The former was tackled for instance by Chatterjee et al. (2009), where an optimal solution (in some sense) was developed using dynamic programming techniques, carrying over the inherent computational burden of dynamic programming techniques. The latter was extensively studied in a series of papers (Cannon et al., 2011), where the authors consider various types of constraints and process noise properties, and present multiple techniques to tackle these problems. The common feature is the use of a perturbed linear state feedback (or pre-stabilization), which necessarily limits the number of degrees of freedom, and as a consequence the resulting performance.

While there has been no previous work on model-based control for ventilation in tunnels until the publication of Appendix A (**L. Ferkl** and Meinsma, 2007), the situation in building automation is different. A significant amount of works has been focused on individual aspects of the building automation problems. There are works that describe the MPC control of boilers (Liao and Dexter, 2005), cooling systems (Coffey et al., 2010) etc. Significant effort has been made in simulation and proving that MPC may improve the thermal comfort while maintaining, or even reducing energy consumption of the system; however, these simulations are usually fairly simplified (Huang and Wang, 2008; Moon and Kim, 2010). A more complex view, even though idealized as well, is provided by the OptiControl project (Gyalistras et al., 2010a,b) which claims that use of MPC in buildings can save 10-40 % of energy, while maintaining the thermal comfort. In fact, the first work publishing results of an operational MPC controller of a complex building for an extended period of time (one heating season) is Široký et al. (2010).

4 Case Studies

4.1 Tunnel Ventilation

The first case study of a model-based control is a ventilation control in highway tunnels. The overall concept is described in Appendix A, while the specific case study for the Blanka tunnel is described in Appendix B.

Operational ventilation in tunnels (i.e. not the emergency, fire ventilation) is usually not an issue during tunnel design phase, as usual highway tunnels take advantage of the piston effect of the traffic, which causes air flow velocity for sufficient air exchange. However, in city tunnels, the air has to be exchanged by means of ventilation, as the piston effect is not sufficient enough to comply with all the necessary control constrains. For example, in the Blanka tunnel, there are modes of operation to ensure air quality outside the tunnel that require virtually no air flow out of the tunnel portals – all the air has to be extracted through two major ventilation shafts. As the tunnel has a complex topology, MIMO controllers have to be employed.

As the air flow dynamics are very fast, they have been neglected and a semi-dynamical model was used to model the air flow; the traffic model and exhaust model remain dynamic. Referring to Appendix A and references therein, the air flow model is described by Equations (2) and (3), the traffic model by Equation (1) and the exhaust model by Equations (4)-(6). Based on this model, static optimal control was designed (Sections 4 and 5 of Appendix A).

Even though a brief example is given in Appendix A, a more detailed example is shown in Appendix B.

4.2 Reliability Issues

Appendix C shows a reliability analysis of a tunnel ventilation system. In order to persuade the industrial partners to incorporate the model-based control of tunnel ventilation into their plans, it was necessary to show that the system is safe enough for the risky environment of highway tunnels. As the commonly used methods of reliability analysis in tunnels fail to describe certain probability issues of reliability, a method originally developed by NASA was adopted and modified, and its behavior has been shown for the case of the Strahov tunnel in Prague. Similar analysis for the Blanka tunnel has been under way by the time of submission of this habilitation thesis.

4.3 Building Heating Control

The building heating control originally started by the Czech Ministry of Industry and Commerce grant no. 2A-1TP1/084, "Integration of building systems, research and application of intelligent algorithms with influence on energy consumption of buildings and living houses". Within this grant, a Model-based Predictive Controller (MPC) of a building of Czech Technical University in Prague–Dejvice was developed, followed by other projects: "Control Systems for Optimization of Energy Consumption in Low-Energy and Passive Houses" (Czech Ministry of Industry and Commerce, FR-TI1/517), MPC control of a Hollandsch Huys building in Belgium within the scope of the GEOTABS project (EraSME grant) and savings potential estimation for the Premier building in Munich (OptiPremier project contracted by ICADE, GmBH).

In the first stage, it was necessary to get a reliable model of a building. A study comparing the ARMAX and Subspace identification methods was published (Appendix D). Based on the comparisons, Subspace identification methods were chosen and the process was applied to the CTU building in Prague–Dejvice (Appendix E). However, the state-of-the-art methods were not suitable for the specific building data; prior information had to be incorporated into the methods in order to get them running correctly (Appendix F). At the end, the subspace identification methods were improved and implemented in such a way that a reliable building model could have been obtained (see e.g. Figure 9 in Appendix E, Figure 3 in Appendix F or Section IV in Appendix K).

The outline of the controller is given in Appendix G. The first results with experiments on a real building are given, with estimated savings reaching 20 % compared to the state-of-the-art controllers. It has soon appeared that an assessment tool is needed to estimated savings potential of the model-based control of buildings. A candidate for such a method is described in Appendix H; this method is based on a regular energy audit, as proposed e.g. by the norm EN 12831 (2003). A thorough description of the first season with the Model-based Predictive Control in the CTU building in Prague–Dejvice is described in Appendix I. Last two appendices (Appendix J and Appendix K) describe the concept of implementation of model-based controllers in large modern buildings; this concept is a result of the OptiPremier project with ETH Zürich, an industrial project funded by the French major developer ICADE.

5 Conclusions

5.1 Contribution of the Habilitation Thesis

The main contribution of the research, hereby disclosed as the habilitation thesis, can be summarized into three points:

- The identification methods, particularly the subspace identification methods, have been modified in such a way that identification of LTI systems with prior information of the DC gains is feasible.
- A concept of model-based predictive control for thermodynamical systems, particularly for buildings, has been developed. This concept uses both data from a real system and data from a physical (first principle) model to find an LTI model by means of statistical identification. Such model is used in the model-based controller.
- The model-based controller was implemented on real systems, particularly for ventilation systems in tunnels (where work is still in progress) and for heating control of buildings. The results show significant savings of the model-based control, compared to the traditional controllers.

5.2 Further Work

Even though identification of linear systems is sometimes considered to be a closed issue, recent works by respected scientists (e.g. Ljung (2010)) point out problems that are still open. The application motivated research areas include further incorporation of prior information into subspace identification methods, MPC relevant identification, closed-loop identification or frequency-domain identification.

The general controller design for tunnel ventilation has received some attention, as it is an elegant solution particularly for city tunnels. The opening of the Blanka tunnel (scheduled to spring 2014) and first practical results will show real properties of the model-based controller, and may be an impuls for further implementations.

The savings potential of the HVAC systems in buildings makes the model-based control of such systems a challenging issue. There have been several other ongoing projects in the time of submission of this habilitation thesis, mainly the OptiPremier project (with ETH Zürich) and the GEOTABS project (with KU Leuven, TU Eindhoven, and others). While the concept of the model-based predictive control of buildings has been tuned in last few years, the procedure of controller design is still time consuming and requires non-trivial background of the staff involved in the design procedure. This fact limits the application of the model-based control to a class of buildings with large energy consumption, wherein the savings achieved by the new controller pay off the effort needed for its design and implementation. It is therefore necessary to make the procedure simpler, faster (and hence cheaper), in order to be applicable for a wider class of buildings.

Bibliography

DIN EN 12831:2003 Heating systems in buildings – Method for calculation of the design heat load. International Organization for Standardization, Geneva, Switzerland, 2003.

Buildings and their Impact on the Environment: A Statistical Summary. 2009. Rev. April 22, 2009, available on-line at http://www.epa.gov/greenbuilding/pubs/gbstats.pdf.

K. J. Åström and P. Eykhoff. System Identification – Survey. *Automatica*, 7(2):123–&, 1971. ISSN 0005-1098.

K. J. Åström and B. Wittenmark. On self tuning regulators. *Automatica*, 9(2):185 – 199, 1973. ISSN 0005-1098.

H. Akaike. New Look at Statistical-Model Identification. *IEEE Transactions on Automatic Control*, AC19(6):716–723, 1974. ISSN 0018-9286.

D. P. Bertsekas. Dynamic Programming and Suboptimal Control: A Survey from ADP to MPC, in Fundamental Issues in Control. *European Journal of Control*, 11(4-5):310–334, 2005.

D. Bertsimas and D. Brown. Constrained Stochastic LQC: A Tractable Approach. *IEEE Transactions on Automatic Control*, 52(10):1826–1841, 2007.

M. Cannon, B. Kouvaritakis, S. V. Rakovic, and Q. Cheng. Stochastic tubes in model predictive control with probabilistic constraints. *IEEE Transactions on Automatic Control*, 56(1): 194–200, 2011.

D. Chatterjee, E. Cinquemani, G. Chaloulos, and J. Lygeros. On stochastic control up to a hitting time. In *Proceedings of the IEEE Conference on Decision and Control*, pages 531–536, Shanghai, China, 2009.

B. Coffey, F. Haghighat, E. Morofsky, and E. Kutrowski. A software framework for model predictive control with genopt. *Energy and Buildings*, 42(7):1084–1092, 2010.

C. F. Gauss. *Theoria Motus Corporum Coelestium in sectionibus conicis solem ambientium*. Perthes & Bessen, Hamburg, 1809.

R. B. Gopaluni, R. S. Patwardhan, and S. L. Shah. MPC relevant identification–tuning the noise model. *Journal of Process Control*, 14(6):699–714, 2004. ISSN 0959-1524.

D. Gyalistras, M. Gwerder, F. Oldewurtel, C. N. Jones, M. Morari, B. Lehmann, K. Wirth, and V. Stauch. Analysis of energy savings potentials for integrated room automation. In *10th REHVA World Congress Clima*, Antalya, Turkey, 2010a.

Dimitrios Gyalistras et al. Use of Weather and Occupancy Forecasts for Optimal Building Climate Control. Technical report, ETH Zürich, 2010b.

B. L. Ho and R. E. Kalman. Effective construction of linear, state-variable models from input/output functions. *Regelungstechnik*, 14(12):545–548, 1966.

Gongsheng Huang and Shengwei Wang. Two-loop robust model predictive control for the temperature control of air-handling units. *HVAC&R Research*, 14(4):565–580, 2008.

J. Široký, **L. Ferkl**, and S. Prívara. Model Predictive Control of Building Heating System. In *Proceedings of the 10th REHVA World Congress Clima 2010*, Antalya, Turkey, 2010.

Z. Liao and A.L. Dexter. An experimental study on an inferential control scheme for optimising the control of boilers in multi-zone heating system. *Energy and Buildings*, 37(1):

55–63, 2005.

L. Ljung. Identification for control ďż˝ What is there to learn? In Y. Yamamoto and S. Hara, editors, *Learning, control and hybrid systems*, volume 241, pages 207–225. Springer Berlin / Heidelberg, 1999. ISBN 1-85233-076-7.

L. Ljung. Perspectives on system identification. *Annual Reviews in Control*, 34(1):1–12, 2010. ISSN 1367-5788.

J.W. MacArthur and C. Zhan. A practical global multi-stage method for fully automated closed-loop identification of industrial processes. *Journal of Process Control*, 17(10):770–786, 2007.

Jin Woo Moon and Jong-Jin Kim. Ann-based thermal control models for residential buildings. *Building and Environment*, 45(7):1612–1625, 2010.

O. Nývlt, S. Prívara, and **L. Ferkl**. Probabilistic risk assessment of highway tunnels. *Tunnelling and Underground Space Technology*, 26(1):71–82, 2011. ISSN 0886-7798.

F. Oldewurtel, C. N. Jones, and M. Morari. A tractable approximation of chance constrained stochastic MPC based on affine disturbance feedback. In *Proceedings of the 47th IEEE Conference on Decision and Control*, pages 4731–4736, Cancun, Mexico, 2008.

J. Pořízek, J. Zápařka, and **L. Ferkl**. Ventilation control of the Blanka tunnel: A mathematical programming approach. In Peter J. Sturm and Sabine Minarik, editors, *Proceedings of the 4th Symposium on Tunnel Safety and Ventilation*, pages 192–196, Graz, Austria, 2008. ISBN 978-3-85125-008-4.

A. Propoi. Use of linear programming methods for synthesizing sampled-data automatic systems. *Automation and Remote Control*, 24(7):837–844, 1963.

S. Prívara, J. Cigler, Z. Váňa, **L. Ferkl**, and M. Šebek. Subspace Identification of Poorly Excited Industrial Systems. In *Proceedings of the 49th IEEE Conference on Decision and Control*, pages 4405–4410, Atlanta, USA, 2010.

S. Prívara, J. Široký, **L. Ferkl**, and J. Cigler. Model predictive control of a building heating system: The first experience. *Energy and Buildings*, 43(2-3):564–572, 2011a. ISSN 0378-7788.

S. Prívara, Z. Váňa, D. Gyalistras, J. Cigler, C. Sagerschnig, M. Morari, and **L. Ferkl**. Modeling and Identification of a Large Multi-Zone Office Building. In *Proceedings of the 2011 IEEE International Conference on Control Applications*, Denver, USA, 2011b. Invited lecture.

S. Prívara, Z. Váňa, J. Cigler, and **L. Ferkl**. Subspace Identification: A Path Towards Large-Scale Predictive Controllers of Buildings. In *Submitted to the 50th IEEE Conference on Decision and Control*, Orlando, USA, 2011. Invited lecture.

S. J. Qin and T.A. Badgwell. A survey of industrial model predictive control technology. *Control Engineering Practice*, 11:733–764, 2003.

J. Richalet, A. Rault, J. Testud, and J. Papon. Algorithmic control of industrial processes. In *Proceedings of the 4th IFAC Symposium on Identification and System Parameter Estimation*, page 1119–1167, Tbilisi, 1976.

D. S. Shook, C. Mohtadi, and S. L. Shah. A control–relevant identification strategy for GPC. *Automatic Control, IEEE Transactions on*, 37(7):975–980, 1992. ISSN 0018-9286.

J. Skaf and S. Boyd. Nonlinear Q-Design for Convex Stochastic Control. *IEEE Transactions on Automatic Control*, 54(10):2426–2430, 2009.

L. Ferkl and J. Široký. Celing radiant cooling: Comparison of armax and subspace identification modelling methods. *Building and Environment*, 45(1):205–212, 2010. ISSN 0360-1323.

L. Ferkl and G. Meinsma. Finding optimal control for highway tunnels. *Tunnelling and Underground Space Technology*, 22(2):222–229, 2007. ISSN 0886-7798.

L. Ferkl, C. Verhelst, L. Helsen, A. Ciller, and D. Komárková. Energy Savings Potential of a Model-Based Controller for Heating: A Feasibility Study. In *Proceedings of the 2011 IEEE International Conference on Control Applications*, Denver, USA, 2011. Invited lecture.

L. Ferkl, J. Široký, and S. Prívara. Model Predictive Control of Buildings: The Efficient Way of Heating. In *Proceedings of the 2010 IEEE International Conference on Control Applications*, pages 1922–1926, Yokohama, Japan, 2010. ISBN 978-1-4244-5363-4. Invited lecture.

P. Trnka and V. Havlena. Subspace like identification incorporating prior information. *Automatica*, 45(4):1086–1091, 2009. ISSN 0005-1098.

Z. Váňa, J. Kubeček, and **L. Ferkl**. Notes on Finding Black-Box Model of a Large Building. In *Proceedings of the 2010 IEEE International Conference on Control Applications*, pages 1017–1022, Yokohama, Japan, 2010. ISBN 978-1-4244-5363-4.

P. M. J. van den Hof and R. J. P. Schrama. Identification and control – closed-loop issues. *Automatica*, 31(12):1751 – 1770, 1995. ISSN 0005-1098.

P. van Overschee and B. De Moor. N4SID – Subspace Algorithms for the Identification of Combined Deterministic Stochastic-Systems. *Automatica*, 30(1):75–93, 1994. ISSN 0005-1098.

M. Verhaegen and P. Dewilde. Subspace Model Identification .1. The Output-Error State-Space Model Identification Class Of Algorithms. *International Journal of Control*, 56(5): 1187–1210, 1992. ISSN 0020-7179.

Appendices

Appendix A

Ferkl, Meinsma: Finding Optimal Control for Highway Tunnels

FERKL, L. – MEINSMA, G. Finding Optimal Control for Highway Tunnels. Tunnelling and Underground Space Technology. 2007, 22, 2, pp. 222–229. ISSN 0886-7798.

Number of citations (excluding autocitations):

- WoS: 4
- Scopus: 7
- Google Scholar: 4

Impact factor: 0.917

Available online at www.sciencedirect.com
ScienceDirect
Tunnelling and Underground Space Technology 22 (2007) 222–229
Tunnelling and Underground Space Technology incorporating Trenchless Technology Research
www.elsevier.com/locate/tust

Finding optimal ventilation control for highway tunnels

Lukáš Ferkl [a,c,*], Gjerrit Meinsma [b]

[a] *Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha 6, Czech Republic*
[b] *Department of Applied Mathematics, Faculty of Electrical Engineering, Mathematics and Computer Science, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands*
[c] *Center for Applied Cybernetics, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 4, 182 08 Praha 8, Czech Republic*

Received 4 January 2006; received in revised form 7 April 2006; accepted 20 April 2006
Available online 27 June 2006

Abstract

A control scheme for highway tunnels is designed based on a static model of the highway tunnel. The controller is designed to keep the exhaust levels inside the tunnel below given limits. The control is then simulated on a dynamical model of a highway tunnel.

Keywords: Modelling; Ventilation control; Energy optimization; Pollution control

1. Motivation

There are many large highway tunnels constructed all around the world. The operation of such tunnels has to meet several requirements, e.g. energy consumption optimization or minimal influence on the surrounding environment. As the tunnel system is typically very complex (and may consist of a number of fans, sensors, several ventilation shafts, etc.), a controller designed according to modern control design methods could be very efficient. Surprisingly, the control of the majority of these tunnels is performed by heuristical approaches based on experience of construction engineers.

However, for the tunnels *Mrázovka* (opened August 2004) and *Blanka* (under construction, to be opened in 2011) in Prague, Czech Republic, a different strategy has been chosen. In Ferkl et al. (2005), it was decided to use modern control techniques to meet all the intensive demands associated with highway tunnels in urban areas.

The control task can be stated as follows: the control system should be designed to keep exhaust levels inside the tunnel below limits (that are given), with minimal energy consumption (see Table 1).

2. Tunnel simulation

2.1. System decomposition

A general tunnel system is very complex and no simple model can be designed for it. There is a need for decomposition of this system. To handle the system more easily, both a *functional* and a *spatial* decomposition have been performed.

The *functional* decomposition is very intuitive (Fig. 1). The tunnel model comprises three main functional parts (or subsystems) – ventilation, traffic and exhaust. The inputs and outputs are well defined and the decomposition is quite natural as the system is fully separable.

The separability of the tunnel system into three subsystems can be seen from Fig. 1. It is obvious that ventilation (i.e. jet fans, ventilation shafts) does not effect traffic. Conversely, traffic does have an effect on ventilation because vehicles inside the tunnel generate a piston effect (air flow caused by the moving vehicles) which influences

* Corresponding author.
E-mail addresses: ferkll@control.felk.cvut.cz (L. Ferkl), g.meinsma@math.utwente.nl (G. Meinsma).
URLs: http://dce.felk.cvut.cz (L. Ferkl), http://wwwhome.math.utwente.nl/~meinsmag/ (G. Meinsma).

0886-7798/$ - see front matter
doi:10.1016/j.tust.2006.04.002

Table 1
Variables

Variables	Description	Unit
x	Position of a car (inside tunnel)	m
x_i^t	ith car position at time t	m
$v_{\max}$	Speed limit	m/s
d_{car}	Car length	m
v_{car}	Car speed	m/s
t	Time	s
P	Pressure	Pa
S_{area}	Cross section area	m^2
$c(x,t)$	Exhaust concentration	1
$\mathscr{D}$	Diffusion coefficient	m^2/s
R	Exhaust (percentage) produced by vehicles per second	s^{-1}
ϕ	Traffic–vehicles density	s^{-1}
f	Power input (in %) – jet fans, vent. shafts, etc.	1
p_e	Piston effect caused by vehicles	m/s
σ	Pollution generated by vehicles inside the tunnel	m^{-1}
v_{air}	Air flow velocities	m/s
p	Pollution levels	m^{-3}
a, $\hat{a}$	Linearization coefficient	10^{-2} m/s
v_{air}^*	Air flow velocity caused by other factors than fans	m/s
k	Average exhaust production of one vehicle/s	s^{-1}
N	Number of vehicles passed	1
S_{area}	Tunnel cross section area	m^2
n	Number of vehicles inside the tunnel	1
s	Position of the pollution sensor	m
Q, Q^*	Air flux	m^3/s
M	Number of control sections	1

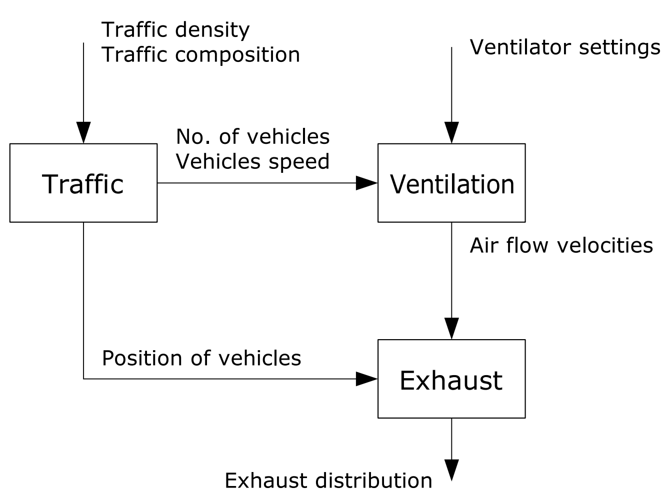


Fig. 1. Functional decomposition of a general tunnel system.

the resulting air flow velocity and hence ventilation. In the same manner, the exhaust generally do not influence neither traffic nor ventilation system.[1] But the traffic influences the exhaust (through exhaust production of the vehicles) and so does ventilation (the exhaust move with the air flow).

Upon the analysis of a general tunnel structure, it has been decided to perform a *spatial* decomposition as well, in order to isolate various ventilation and traffic phenomenon. The spatial decomposition will be described in the following subsections in more detail.

2.2. Traffic

To be able to simulate traffic in a dynamical manner, a microscopic car-following model, which has been described in Rothery (2002), has been used, which was simplified for the purposes of exhaust generation. The effect of acceleration and deceleration of the cars are neglected, because typically cars maintain a constant speed in a tunnel. Every car always tries to keep up with the speed limit $v_{\max}$. Assume that the simulation model discretization is $\delta t = 1$ s and $\delta x = 1$ m. Then the model computes the movement of the cars in the following way:

$$x_i^{t+1} = x_i^t + \min(v_{\max}, 0.5(x_{i-1}^t - x_i^t) - d_{car}) \tag{1}$$

where x_i^t is the ith car position at time t, $v_{\max}$ is speed limit, and d_{car} is the car length.

The meaning of (1) is that a car always watches the preceding car and tries to be at least 2 s behind. If it has enough room, it maintains the maximum speed – the actual speed limit. It can also be seen from (1) that the movement of the cars has to be computed from the end, i.e. from the first car in the tunnel.

A general two-lane module has been made as a building block for the traffic simulation. This way, the tunnel model is modular, which brings all the advantages of a modular approach, well known in the art. In a two-lane situation, a car can overtake if there is not enough room in its lane. A simple algorithm has been chosen for overtaking: if there is not enough room for a car according to (1), a second try is performed for the other lane. The car switches lane, if it allows him to move faster.

The system takes as input the traffic density ϕ, that is, the number of cars per second that enter each of the tunnel entrances, and computes actual distribution of the cars and their speed v_{car}. This model has been solved using MATLAB.[2] It also has some more features, e.g. a car will not overtake if another car is approaching in the other lane, or overtaking can be restricted by introducing a random factor.

2.3. Ventilation

The ventilation subsystem is *spatially* decomposed into several ventilation sections. The desired output of this

[1] The only influence is during some excessively high dust levels in the air, when the drivers slow down because of a decreased visibility range. This emergency situation is not in the scope of this paper.

[2] MATLAB is a commercial language for mathematic and technical computing, commonly used for applied mathematics, control engineering, signal processing, etc.

subsystem is the air flow velocity, which is computed by a set of equations of continuity (2) and Bernoulli equations (3). This approach can be used for fluid flows with low velocities; inside road tunnels, this condition is satisfied. More details on this approach are shown e.g. in Boman et al. (1997) or Douglas et al. (2001).

The equation of continuity (2) is used to connect the ventilation sections together

$$v_{\mathrm{air}} S_{\mathrm{area}} = \text{constant} \tag{2}$$

where v_{air} is the air flow velocity and S_{area} is the tunnel cross section area These may vary per section.

The Bernoulli equation (3) is an equation of energy, however, after suitable manipulation, we get an equation describing pressure changes inside a ventilation section

$$\Delta P_{\mathrm{tot}} = \Delta P_{\mathrm{loc}} + \Delta P_{\mathrm{fric}} \pm \Delta P_{\mathrm{pist}} \pm \Delta P_{\mathrm{fans}} \pm \Delta P_{\mathrm{atm}} \tag{3}$$

where P_{tot} is the total pressure difference. *Pressure drops*: P_{loc} is the local losses, P_{fric} is friction, P_{pist} is vehicles piston effect, P_{fans} is jet fans effect and P_{atm} is the atmospheric conditions.

The set of Bernoulli equations and equations of continuity is a set of non-linear equations. The description of the respective components of (3) is beyond the scope of this paper, more details may be found in Ferkl et al. (2005). Because of the non-linearities, the set of equation has to be solved numerically. We have used an affine scaling trust-region approach described in Bellavia et al. (2003).

However, the Bernoulli equation and equation of continuity describe a steady state of a system. Assuming that the tunnel dynamics is slow compared to that of the pressure dynamics, i.e. the inputs to the tunnel (traffic density, jet fans puissance, etc.) do not change suddenly compared to pressure dynamics, which is a reasonable presumption, we can take a consecutive series of steady states to form long-term dynamics.

2.4. Exhaust

The exhaust levels are the primary output variables from the model. They depend both on vehicles type and distribution and air flow velocity inside the tunnel. The mass of the exhaust is being observed, because it does not depend on the tunnel geometry.

There are three pollutants measured inside the tunnel – nitrogen oxides (NO_x), carbon monoxide (CO) and opacity (OP), which is a formal representation of visibility range and dust particles concentration. Nowadays, this set of three pollutants is considered as standard, see e.g. PIARC (2004).

A mass balance equation for a component with constant density and constant diffusion coefficient is used for exhaust distribution:

$$\frac{\partial c(x,t)}{\partial t} + v_{\mathrm{air}} \frac{\partial c(x,t)}{\partial x} = \mathscr{D} \frac{\partial^2 c(x,t)}{\partial x^2} + R \tag{4}$$

where $c(x,t)$ is the exhaust concentration, v_{air} is the air velocity (as considered earlier), $\mathscr{D}$ is the diffusion coefficient and R is the exhaust produced by vehicles in the tunnel.

This partial differential equation can be discretized in the following way:

$$\frac{c_j^{t+1} - c_j^t}{\delta t} + v_{\mathrm{air}} \frac{c_{j+1}^t - c_j^t}{\delta x} = \mathscr{D} \frac{c_{j+1}^t - 2c_j^t + c_{j-1}^t}{\delta x^2} + R \tag{5}$$

where, as before, superscripts refer to discretized time and subscripts to discretized position. From Eq. (4) to (5), a forward approximation of derivatives has been used. This is rather crude, but in this case, it is stable and provides good results when compared to a real system (Kurka et al., 2005). Assuming that $\delta t = 1$ s and $\delta x = 1$ m, we get an exhaust concentration equation which is suitable for direct implementation into a computer code:

$$c_j^{t+1} = c_j^t + \mathscr{D}(c_{j+1}^t - 2c_j^t + c_{j-1}^t) - v_{\mathrm{air}}(c_{j+1}^t - c_j^t) \tag{6}$$

3. Principles of static control in tunnels

3.1. Problem formulation

The task of the controller is to keep the exhaust limits inside the tunnel below desired limits, with low energy consumption. The measured variables that are available to our controller are the air flow velocities v_{air} measured at various places throughout the tunnel, exhaust levels p also measured at various places and vehicles density ϕ measured at the entrance to the tunnel. The control output is the fans power inputs f which affects the air flow velocity and thereby affects the pollution levels, which is our main concern.

In addition, the control must not effect the lifetime of the ventilation equipment dramatically.

Finding a linear, time invariant model of the ventilation part of the tunnel system is almost impossible. Particularly, during off-peak hours, when traffic is sparse, external atmospheric conditions and various turbulences have major influence on the behavior of the air mass inside the tunnel. Estimating those factors is not an easy task, and, according to Douglas et al. (2001), a real-time dynamical estimation for complex tunnels is not possible at all.

The next step is to decide, whether to choose a static, or a dynamic controller. Here, are some helpful facts about the tunnel:

- To save lifetime of the fans inside the tunnel, the fans should not be switched on or off more often than every 10 min.
- The air flow velocities v_{air} react to any changes almost immediately (within seconds), as the air mass inside a tunnel is considered to be incompressible (for explanation, see Bellasio, 1997).
- The traffic situation ϕ and speed of the cars v_{car} (which have major influence on the exhaust levels) does not notably change over 10 min.

Because of the restriction on the fans switching, a static control of the tunnel is the better option. The above indicates that the dynamics of the air mass can be neglected in this case. The dynamics of the exhaust are more important. However, as the traffic situation changes fairly slowly, the exhaust system is not subject to any fast transitions over a period of 10 min, so the dynamics of the exhaust can be neglected as well.

3.2. Steady-state model of the exhaust

Before turning to the control, a steady-state description of the system is to be found. We want to stress that we use the steady-state description only in order to *derive* the control scheme: once derived we apply it to the more realistic situation. The original system is shown in Fig. 2a. Here, ϕ, v_{car}, f, v_{air} and p are as introduced earlier. In addition, σ denotes pollution generated by vehicles and p_e denotes the piston effect caused by vehicles.

This is not yet in a suitable form. Instead of trying to model the piston effect p_e as a function of traffic density ϕ and speed of the cars v_{car}, we take a different route: since the fans power f operates on a different time scale (faster) than the piston effect and other effects do, we can separate the effect of f on the measured v_{air} from that of the piston effect p_e. Measurements confirm that the effect of f on v_{air} is well approximated as an affine function. That is, we model it as

$$v_{air} = af + v^*_{air} \tag{7}$$

The way this is handled is to estimate a, and then on the basis of this estimate, v^*_{air} is taken equal to $v_{air} - af$. Given the fact that v^*_{air} varies slowly compared to f (and the estimate a, more on this later), it is a reasonable approach to assume that v^*_{air} is constant over, say, the next 10 min after which v^*_{air} is updated for the next 10 min, etc.

Eq. (7) represents block P_1 of Fig. 2b. The linearization coefficient a can be computed, but it can be also set by adaptation, as will be discussed later.

The air flow velocity v_{air} is not computed, it is taken as a direct input. Besides this, the influence of the ventilation equipment f is computed separately.

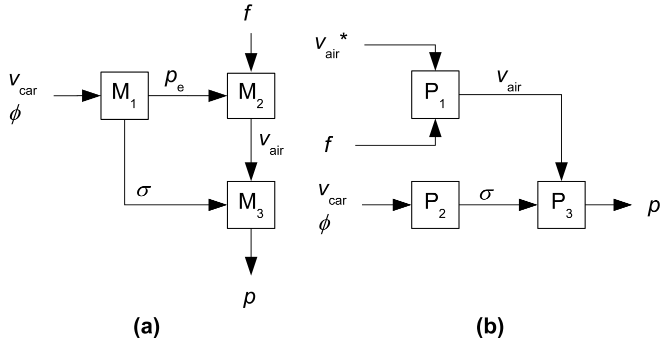


Fig. 2. Original model (a) and a modified model (b) with air flow velocities as direct inputs.

In a one-way highway tunnel, the steady state of the exhaust with zero air flow $v_{air} = 0$ can be described by the following equation (corresponding to block P_2 in Fig. 2b):

$$p = \frac{k \cdot N}{S_{area}} \tag{8}$$

where p is the pollution level introduced earlier, and k is the average exhaust production of one vehicle/s, N is the number of vehicles passed and S_{area} is the tunnel cross section area.

In this situation, the exhaust accumulate inside the tunnel. As the air flow $v_{air} = 0$, to compute the exhaust levels, we have to take *all* the vehicles that have passed through the tunnel, here denoted with capital N. Introducing a non-zero air flow velocity v_{air}, the steady-state exhaust level at some point of the tunnel changes:

$$p = \frac{k \cdot n}{S_{area}} \cdot \frac{s}{v_{air}} \tag{9}$$

where n is the number of vehicles inside the tunnel and s is the position of the pollution sensor.

The first fraction in (9) represents the exhaust production of the vehicles inside the tunnel, as in (8), the second fraction (corresponding to block P_3 in Fig. 2b) adjusts the longitudinal distribution of the exhaust levels. The longitudinal distribution has a general increasing shape, as can be seen in Fig. 4, which is a plot of pollution levels simulation inside the tunnel *Mrázovka* in Prague, Czech Republic. As a result of the non-zero air flow velocity v_{air}, the exhaust levels p depend on the position s of the pollution sensor (from the entrance to the tunnel) and on the number of cars (denoted by small n) inside the tunnel. This simulation has been already verified, as we have previously shown in Kurka et al. (2005), its output values conform with the real data.

3.3. Example of the static model

As an example for the static model, the *Libouchec Tunnel* will be used. This 450 m long tunnel is located in the Czech Republic. It is a very simple tunnel with no branches and constant cross section area. A one-day simulation of the static model is shown in Fig. 3. The air flow velocities in (9) were calculated according to (3) (see e.g. Boman et al., 1997). Here, are some interesting features of the static model that can be seen in the figure:

- The overall distribution of CO is linear (Fig. 3b). It has a triangular shape inside the tunnel, see Fig. 4 as well.
- The exhaust (CO) distribution is inversely proportional to the air flow velocity. If we compare Figs. 3a and 3b, we can see that the difference in CO concentration at the two sensors between 0 and 4 AM (where the traffic is more or less constant).
- The CO levels are proportional to the traffic density (Fig. 3b and 3c). However, the relation is not linear. The vehicles cause a piston effect; the piston effect increases the air flow velocity which decreases the CO levels.

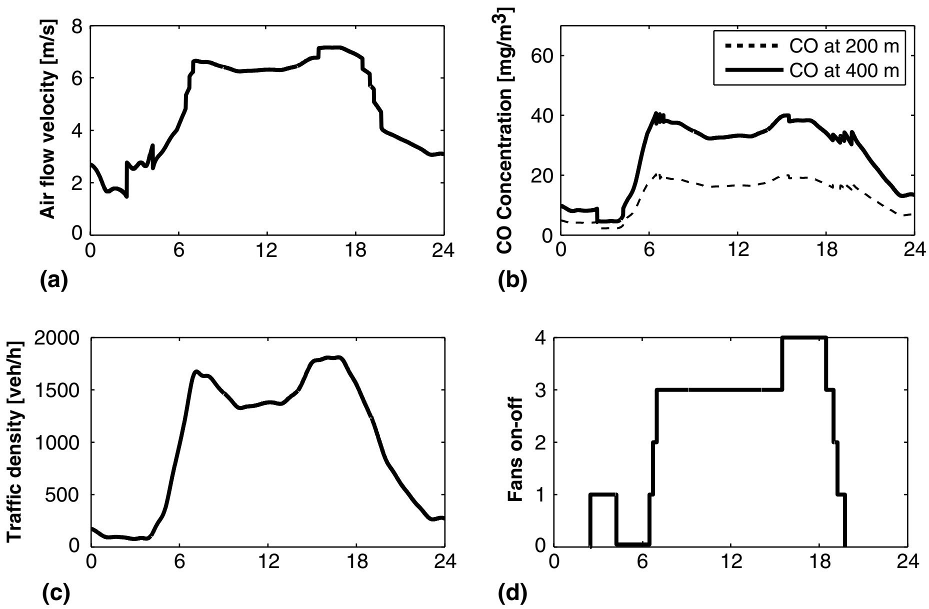


Fig. 3. A static model example (*Libouchec Tunnel*).

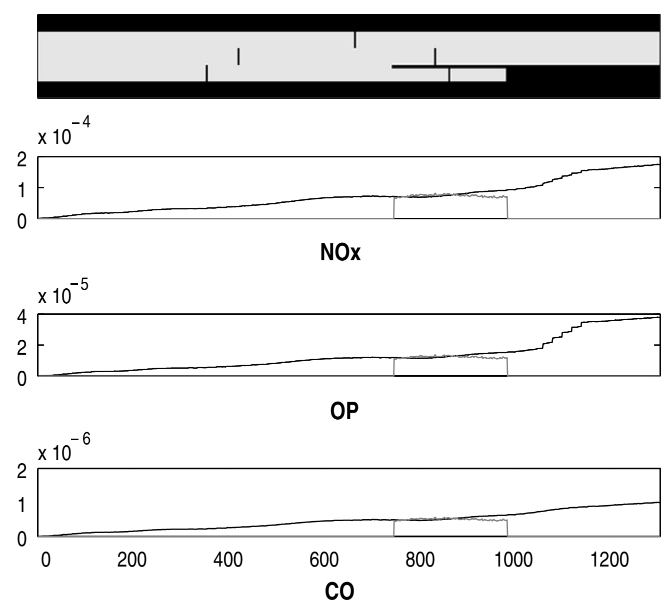


Fig. 4. Exhaust distribution inside the tunnel *Mrázovka*. The uppermost image is a schematic view of the tunnel with cars shown as grey vertical lines. The next three images show distribution of respective exhaust. Darker plots relate to the main tunnel tube, lighter plots relate to the branch. The air flow direction is from right (tunnel entrance) to left (tunnel exit).

- The contribution of the jet fans to the overall air flow velocity is close to linear, as can be seen from Fig. 3a and 3d, especially between 6 and 8 PM.

4. Control of a simple tunnel

4.1. Air flow based control

Now the task to find an optimal control for a highway tunnel becomes simple. Substituting (7) into (9) and isolating f, the optimal fans input power f_o can be computed by substituting p with the desired exhaust, for instance the exhaust limit $\bar{p}$:

$$f_o = \frac{1}{a}\left(\frac{k \cdot n \cdot s}{S_{\text{area}} \cdot \bar{p}} - v^*_{\text{air}}\right) \tag{10}$$

In addition, there are some operational requirements on the fans. To assure a reasonably long lifetime of a fan inside a tunnel, the fan should not be turned on and off very often. Ten minutes is a satisfactory time for this purpose.

The control works in the following way:

- Collect the required average data (traffic density n, fans input power f and air flow velocity v_{air}) over last 10 min.
- If the concentration is higher than the limit $\bar{p}$:
  - Determine the v^*_{air} from (7).
  - Calculate the optimal fans input power f_o from (10).

4.2. Parameter adaptation

In real operation, problems may occur with the parameters k and a. These parameters are uncertain and may vary a little. However, it is difficult to estimate them both at the same time.

Parameter k expresses an average exhaust production of a vehicle. As the technology advances, this parameter is subjected to change, because the vehicles use better engines with lower exhaust production. This parameter changes very slowly (over years).

It is also difficult to estimate it in advance with high accuracy. A way to estimate this parameter is to use the night hours, when the traffic is sparse and there is no need to run the fans inside the tunnel. We can use Eq. (9) to derive a way to estimate the parameter k:

$$k = \frac{v_{\mathrm{air}} S_{\mathrm{area}}}{s \cdot n} p \tag{11}$$

Parameter a expresses the contribution of the jet fans inside the tunnel to the overall air flow velocity. This parameter should not change at all. If this parameter changes, there is a possibility that some of the fans are not working properly. However, there is again a need to adjust this parameter at least during the startup of the control.

The easiest way is to adjust the parameter a according to the pollution levels. If the controller tries to keep some pollution level, but this level is permanently higher (or lower), the parameter a could be lowered (or increased) by a factor of, e.g. 1.02.

5. Control of a complex tunnel

5.1. Air flux based control

The above approach is very useful for simple tunnels. However, for some more complicated situations, it does not allow for extensions.

First of all, it is not very useful to use air flow velocitiesv_{air}, as they vary according to the diameter of the tunnel. It is better to use the air flux

$$Q = v_{\mathrm{air}} S_{\mathrm{area}}$$

because they remain constant throughout the entire tunnel, which is a result of the equation of continuity (2).

First of all, the tunnel can be (but does not need to be) divided into several so-called *control sections*. Considering a situation in some section i of the tunnel, Eq. (9) can be rewritten as follows:

$$p_i = \frac{k_i n_i s_i}{Q} \tag{12}$$

where p_i is the pollution level at the end of the control section i, k_i is exhaust production coefficient, n_i is number of cars present in the tunnel until the end of the control section i and s_i is the position of the end of the control section i with respect to the entire tunnel.

Like for the air flow velocities in (7), the contribution of the fans to the total air flux has to be known. The overall air flux can be partitioned as:

$$Q = Q^f + Q^* = \sum_{i=1}^{M} \hat{a}_i f_i + Q^* \tag{13}$$

where Q^f is the air flux caused by the fans, Q^* is the air flux caused by other factors, $\hat{a}_i$ is linearization coefficient for the fans in a control section i, f_i is fans power input in the control section i and M is the number of control sections.

Now the pollution level in an arbitrary control section i can be rewritten as follows:

$$p_i = \frac{k_i n_i s_i}{\sum_{i=1}^{M} \hat{a}_i f_i + Q^*} \tag{14}$$

5.2. Optimal control through linear programming

Let us assume that there is some exhaust limit $\bar{p}_i$ imposed on the control section i. Given Q^* and $k_i n_i s_i$ there corresponds to each p_i in (14) a unique sum of air flux values, that is, a mapping $\mathscr{P}$ exists such that

$$\sum \hat{a}_i f_i = \mathscr{P}(p_i) \tag{15}$$

So requiring p_i to be in some interval $[0, \bar{p}_i]$ translates into requiring $\sum \hat{a}_i f_i$ to be in some appropriate interval $[\underline{Q^f}, \overline{Q^f}]$. The latter involves our control variables f_i and hence the following linear program, which aims to achieve $p_i \in [0, \bar{p}_i]$ with minimal power is a possible solution to our problem

$$\min_{f_i} \sum_i |f_i| \quad \text{subject to} \quad f_i \in [0, 100], \quad \sum \hat{a}_i f_i \in [\underline{Q^f}, \overline{Q^f}] \tag{16}$$

This controller is a feedback controller – the fans control the air flow velocity that is the input to the controller (through (12), (15) and (16)). In Fig. 6, it can be seen that the control of the pollution levels is achieved by the mapping $\mathscr{P}(p_i)$ from (15), while the linear program (16)

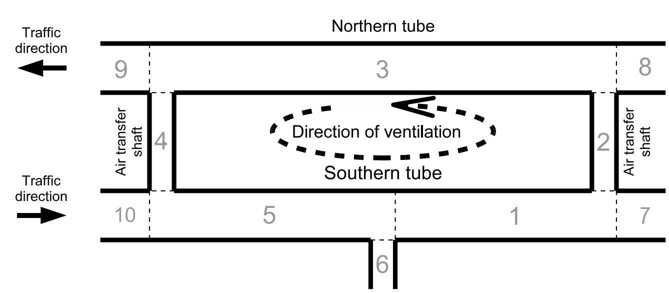


Fig. 5. Complex tunnel example.

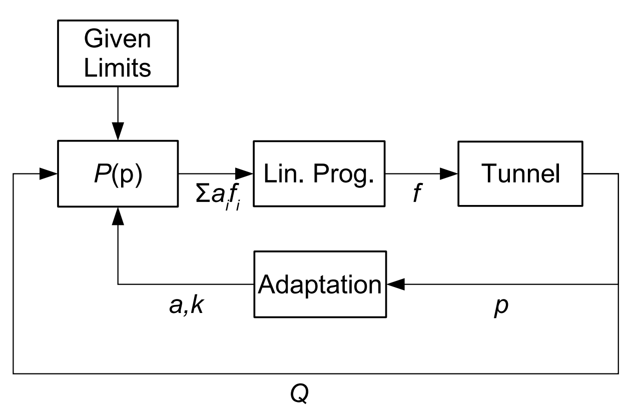


Fig. 6. Feedback controller.

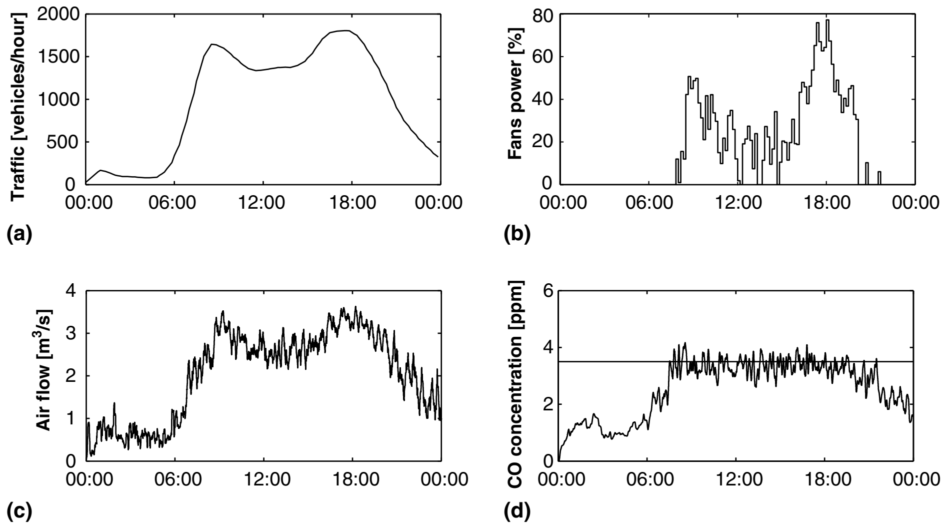


Fig. 7. Simulation results for an example tunnel model: (a) traffic density inside the tunnel (based on real data from tunnel *Mrázovka*, Czech Republic); (b) fans input power, may change every 10 min; (c) air flow velocities inside the tunnel; and (d) CO levels inside the tunnel, which should be maintained below 35 ppm.

optimizes the power of the fans. The greatest advantage of feedback structure is that it works even under perturbed conditions. In the case of the ventilation control in tunnels, it means that the controller reacts "implicitly" to any disturbances to the air flow, e.g. wind, pressure and temperature differences, etc. The control effort is always limited by the power of the fans though.

5.3. Example of a complex tunnel

The linear programming allows us to use any linear constraints that we find useful.[3] As an example, a highway tunnel consisting of two interconnected tubes and ventilation shaft will be given. The tunnel is shown in Fig. 5.

Let us further assume that the tunnel is built in a densely populated urban area. To minimize its impact on the surrounding environment, the polluted air has to leave the tunnel through the ventilation shaft only. This means that the air flux of the sections 7–10 should be always directed into the tunnel.

Some slight modifications have to be made for the static control model. The sections 2 and 4 (air transfer shafts) and 6 (ventilation shaft) are "dummy" tunnels, they contain no traffic. The model virtually starts in the middle of the southern tube, just after the ventilation shaft outlet. As the ventilation shaft does not suck out all the polluted air, the air re-cycles inside the tunnel system. Besides this, not only one air flux Q is used in the tunnel. Eq. (14) for an ith section has to be modified in the following way:

$$
\begin{aligned}
p_i &= \frac{k_i n_i s_i}{\sum_{j\in\mathscr{F}_i}\hat{\alpha}_j f_j + \sum_{k\in\mathscr{I}_i} Q_k^*} + p_t, \\
p_t &= \begin{cases} p_5 \frac{Q_1}{Q_1+Q_6} & \text{for } i \in \{1\ldots 6\} \\ 0 & \text{otherwise} \end{cases}
\end{aligned}
\tag{17}
$$

The sets $\mathscr{F}_i$ and $\mathscr{I}_i$ contain all the indices, that are relevant for the section i. For example, for the section 2 the sets are $\mathscr{F}_2 = \{1, 2, 7\}$ and $\mathscr{I}_2 = \{1, 7\}$. The variable p_t represents the offset of the pollution level inside the tunnel due to the re-circulation of the air.[4] The only place, where the air leaves the tunnel, is the ventilation shaft 6. Not all the polluted air leaves the tunnel, some part re-enters the section 1. Therefore, the pollution p_t has to be taken into account when estimating the exhaust levels in sections 1–6.

The linear program can now be further modified. Let $\mathscr{I} = [1, \ldots, 10]$. Now, if (17) is used to achieve the mapping $\mathscr{P}$ according to (15), the linear program for the present example is

$$
\min_{f_i} \sum_i |f_i| : \begin{cases} f_i \in [0, 100], & i \in \mathscr{I} \\ \sum_{j\in\mathscr{F}_i} \hat{a}_j f_j \in [\underline{Q_i^f}, \overline{Q_i^f}], & i \in \mathscr{I} \\ Q_k \leqslant 0, & k \in \{7, \ldots, 10\} \end{cases}
\tag{18}
$$

6. Simulation results and future work

As an example, a 1000 m long highway tunnel model has been designed, with the fans input power of 2 MW.

[3] Linear programming is a special case of an optimization problem. The designer of the linear program should choose the constrains carefully, because the feasibility set of the optimization criterion may become empty.

[4] The re-circulation of the air is a very complex issue; however, the feedback structure of the controller, as described above, allows us to use this simple approximation.

For simulation purposes, a previously developed dynamical model is used Ferkl et al., 2005. The simulation results can be seen in Fig. 7. For simplicity, only CO levels are being observed. The CO limit has been set to 3.5 ppm for the purpose of the simulation only. In a real tunnel, this limit should be set to the maximum allowable level that applies for the respective tunnel. This level is usually given by law or legal regulation. The results are quite satisfactory, the CO levels stay well around the limit.

Comparing Fig. 7a and c, the air flow velocities strongly depend on the traffic density. This fact could be used to design a predictive controller, that would further minimize the switching of the fans, thus further saving the lifetime of the ventilation equipment.

The controller presented here is suitable for normal operation. However, a completely different controller has to be designed for emergency situations. This does not mean fires or terrorist attacks, but situations, where exhaust is generated rapidly, like sudden traffic jams during car accidents. There is no demand for energy savings, as human health is in danger. In today's tunnels, the ventilation equipment works with full power under emergencies. With the help of the present simulation, it can be found out, whether this really is the best option, and if not, what can be done better.

Acknowledgements

This work was supported by the Ministry of Education of the Czech Republic under Contract No. 1M0567, and partially supported through a European Community Marie Curie Fellowship and in the framework of the CTS, contract number: HPMT-CT-2001-00278.

References

Bellasio, R., 1997. Modelling traffic air pollution in road tunnels. Atmospheric Environment 31, 1539–1551.

Bellavia, S., Macconi, M., Morini, B., 2003. An affine scaling trust-region approach to bound-constrained nonlinear systems. Applied Numerical Mathematics 44, 257–280.

Boman, C.A., Bring, A., Malmström, T.G., 1997. Simulation and measurement of road tunnel ventilation. Tunnelling and Underground Space Technology 12 (3), 417–424.

Douglas, J.F., Gasiorek, J.M., Swaffield, J.A., 2001. Fluid Mechanics, fourth ed. Prentice-Hall, Englewood Cliffs, NJ.

Ferkl, L., Kurka, L., Sládek, O., Pořízek, J., 2005. Simulation of traffic, ventilation and exhaust in a complex road tunnel, In: Proc. of IFAC 2005, Prague.

Kurka, L., Sládek, O., Ferkl, L., Šebek, M., 2005. Model based data verification in a highway tunnel, In: Proc. of WSEAS 2005, Prague.

PIARC, 2004. Road Tunnels: Emissions, Environment, Ventilation. PIARC.

Rothery, R.W., 2002. Car following models, In: Rathi, A.K. (Ed.), Traffic Flow Theory: A State of the Art Report, FWHA.

Appendix B

Pořízek, Zápařka, Ferkl: Ventilation Control of the Blanka Tunnel: A Mathematical Programming Approach

POŘÍZEK, J. – ZÁPAŘKA, J. – FERKL, L. Ventilation Control of the Blanka Tunnel: A Mathematical Programming Approach. In STURM, P. J. – MINARIK, S. (Ed.) proceedings of the 4th Symposium on Tunnel Safety and Ventilation, PP. 192–196, Graz, Austria, 2008. ISBN 978-3-85125-008-4.

Number of citations: 0

VENTILATION CONTROL OF THE BLANKA TUNNEL: A MATHEMATICAL PROGRAMMING APPROACH

Pořízek J.[1], Zápařka J.[1], <u>Ferkl L.</u>[2]
[1]Satra, Czech Republic
[2]Feramat Cybernetics, Czech Republic

ABSTRACT

The Blanka tunnel is a 5.7 km long highway tunnel under construction in Prague, Czech Republic. Because of its complicated topology and very strict environmental restrictions, the synthesis of ventilation control for normal conditions turns out to be fairly challenging.

The air flow is restricted to leave the tunnel through traffic portals – it has to be aspirated by ventilation centers and released by exhaust shafts and chimneys. To achieve this goal, control strategy was designed based on the mathematical programming principles. The designed controller, which is inspired by model-based predictive controllers (MPC) used in heavy industry, is energy optimal by definition, adapts to changes in operational conditions and requires significantly less design time than traditional approaches.

Keywords: ventilation control, city tunnels, model-based predictive control

1. INTRODUCTION

After its opening scheduled for 2011, the Blanka tunnel (see **Figure 1**) will form a part of the inner ring of Prague. Because of its complicated topology, strict operational demands and axial ventilation system, the ventilation control does not have a straightforward solution. After several attempts to use conditional control ("if-then-else" type), we decided to turn to modern control algorithms and to use model-based predictive control (MPC) to achieve the control objectives.

Figure 1: The Blanka tunnel in Prague

2. MPC CONTROL PRINCIPLES

The MPC control is an optimization strategy that minimizes an optimality criterion (cost function) over a finite time horizon. Its first use was in 1970's in oil industry and is widely used for optimal control of "slow" industrial processes today. The main advantages of the MPC control are:

- It handles multivariable control very naturally
- It can take actuator limitations into account
- It allows to define control constraints
- It is very intuitive to tune

For tunnels, the structure of the MPC controller is illustrated on **Figure 2**.

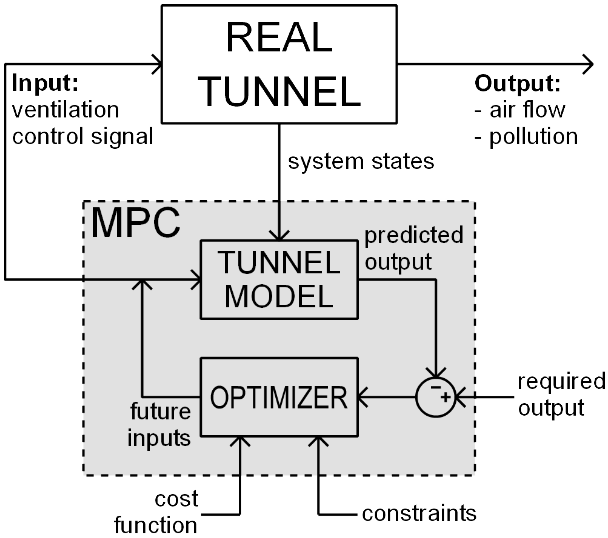


Figure 2: MPC controller block diagram

For the case of tunnel ventilation control, we use a linear model (which will be discussed further), the cost function has a quadratic form

$$F = x^T Qx + u^T Ru \quad (1)$$

wherein x denotes system states and u denotes system control signal. It is obvious that by means of matrices Q and R, we may directly influence the cost of system states and control signal.

The linear constraints have a form

$$Hx + Gu \leq 0 \quad (2)$$

so we can impose physical constraints that exist in the system.

We have to point out that the MPC controller is not a classical, linear controller in the usual sense. It is rather an optimization procedure that optimizes the trajectory of the output signal, while trying to minimize the energy consumption of the system (through the cost function) and maintaining the physical or technological limitations of the system (through constraints).

3. TUNNEL CONTROL MODEL

The basic assumption for the control model is that the air flow has two major contributors – the air flow generated by ventilation system Q^f and the remaining air flow Q^*

$$Q = Q^f + Q^* \tag{3}$$

Moreover, according to measurements from the Mrázovka tunnel (Pořízek, 2007), the air flow can be further expanded to

$$Q = \sum_{i=1}^{M} a_i f_i + Q^* \tag{4}$$

wherein f_i is the power input to the ventilation equipment in the i-th section of the tunnel and a is a suitable linearization coefficient, obtained by simulation or measurement.

The derivation of the simulation model was already presented in (Ferkl, 2007), with the resulting formula for the pollution level in a tunnel section being

$$p_i = \frac{k_i n_i s_i}{\sum_{j \in J_i} a_i f_j + Q^*} \tag{5}$$

wherein k is the exhaust production coefficient for a single vehicle, n is the number of vehicles and s is the length of the i-th tunnel section.

Referring to our previous results (Ferkl, 2007), the optimization process (which turns out to be an MPC controller) that aims to achieve the exhaust inside the tunnel to lie within given limits and the air flow to have the desired direction, is

$$\min_i \left\| \left\| \begin{matrix} \alpha_i f_i(t) \\ \big(f_i(t-1) - f_i(t)\big) \end{matrix} \right\|_{\ell 1} \right\|_{\ell 2} : \begin{cases} f_i \in \langle f_{low}, f_{high} \rangle \\ \sum_{m \in F_i^j} a_m f_m \in \langle Q_{i,low}^f, Q_{i,high}^f \rangle \\ Q_k \leq 0, k \in K \end{cases} \tag{6}$$

The cost function weights the power input to respective ventilation fans (first line) and minimizes the switching of the fans (second line) for enhancing the lifetime of the ventilation equipment. It minimizes the sum of cost functions for all tunnel sections (ℓ_1 norm) according to a quadratic criterion (ℓ_2 norm). Equation (6) is a representation of a mathematical program.

The constraints limit the power input f to the ventilation equipment (first line), imposes the exhaust limits through minimum required air flow Q (second line) and, if needed, requires a negative air flow for tunnel section in a set K (third line).

4. SIMULATIONS

To make the presentation of our results more comprehensive, we will only present the control for the northern tube of the Blanka tunnel only; however, the southern tube is similar to the northern one.

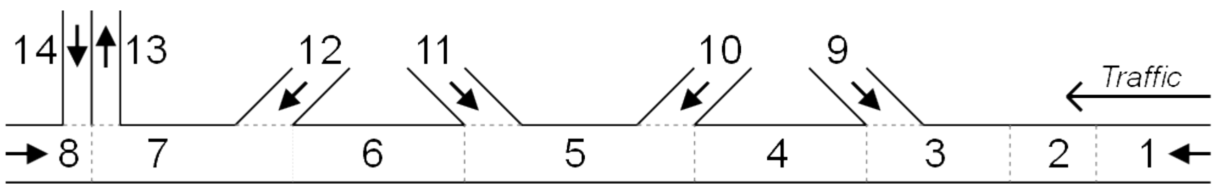


Figure 3: Control sections of the Blanka tunnel, as referred to in the text.

The geometry of the northern tube is shown in **Figure 3**. The tunnel is divided into control sections 1 to 14. Sections no. 13 and 14 represent a ventilation center. The figure also shows the preferred air flow directions for a "closed" mode of operation, wherein the only passage for the air to leave the tunnel is the ventilation center (i.e. section 13).

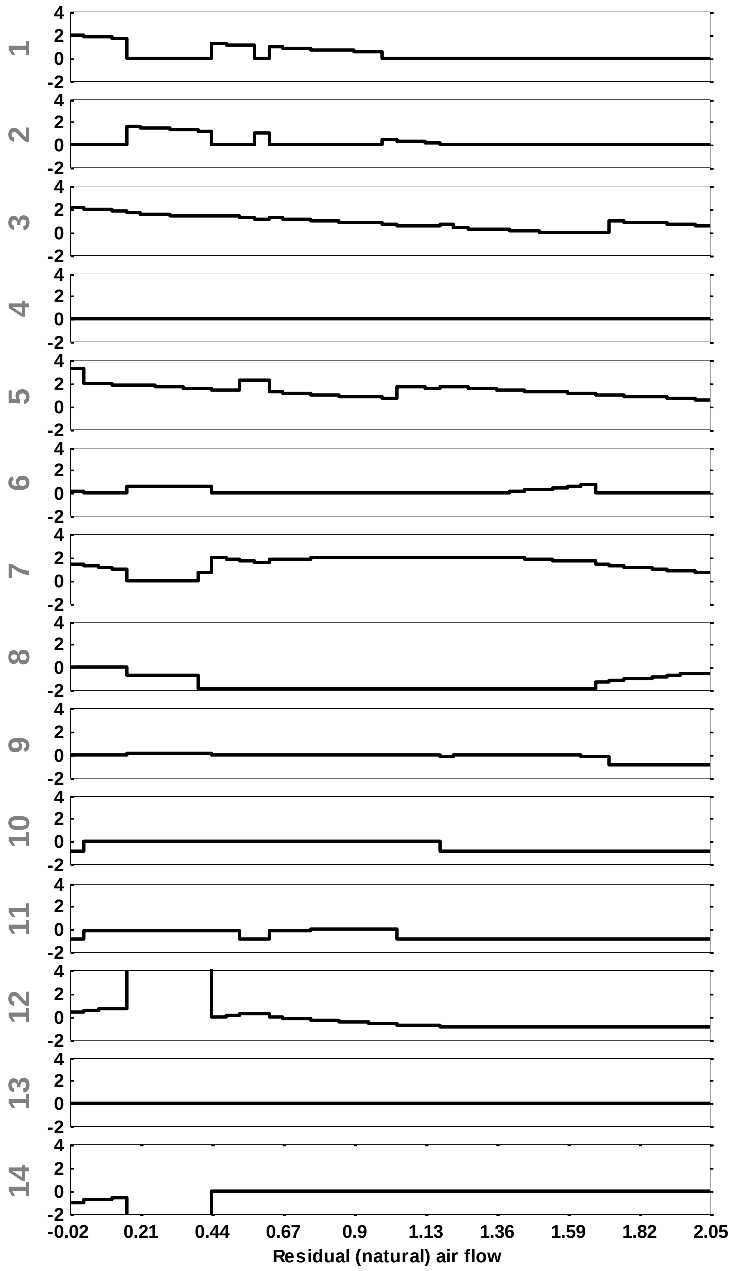


Figure 4: Simulation of MPC control for the Blanka tunnel.

In the following simulations, we use normalized power output equivalent to nominal power of an "average" jet fan installed inside the tunnel. Instead of using time characteristics, we show the results on static characteristics, wherein the power input to the ventilation equipment is the dependent variable and the value of the residual air flow (Q^* in Equation (3), which represents the measured air flow minus the air flow contributed by the ventilation system).

Figure 4 shows the result for the MPC controller without any preferences for the cost function (all power inputs are weighted equally). Unlike for linear controllers (such as PID controllers), the plots are not smooth. This is the result of the constraints – the controller distributes the power according to the capacities of the respective fans, in order to maintain the overall energy consumption minimal. This is something that is very difficult to achieve by purely linear controllers. Numerical difficulties may appear in some cases, as the air flow model is poorly conditioned in principle and the controller sometimes "hesitates", which ventilator to use. It may be seen from the figure that by combining sections 1 and 2 together, we could get a signal that is more "fancy" than the original two separate signals.

Figure 5 shows a comparison of three simulations with different cost functions. The performance of the controller is illustrated by the end-section of the tunnel (sections 7, 8, 12, 14), which is interesting for comparison – because the tunnel operates in a "closed" mode, the air flow in section 8 has to be reversed. In said simulations, the following conditions were set through the cost function:

1. Ventilation in all sections has the same cost.
2. Sections 9-12 (onramps) are penalized, i.e. their use has to be minimized.
3. Sections 9-12 are penalized, while the use of section 14 (the ventilation shaft) is preferred.

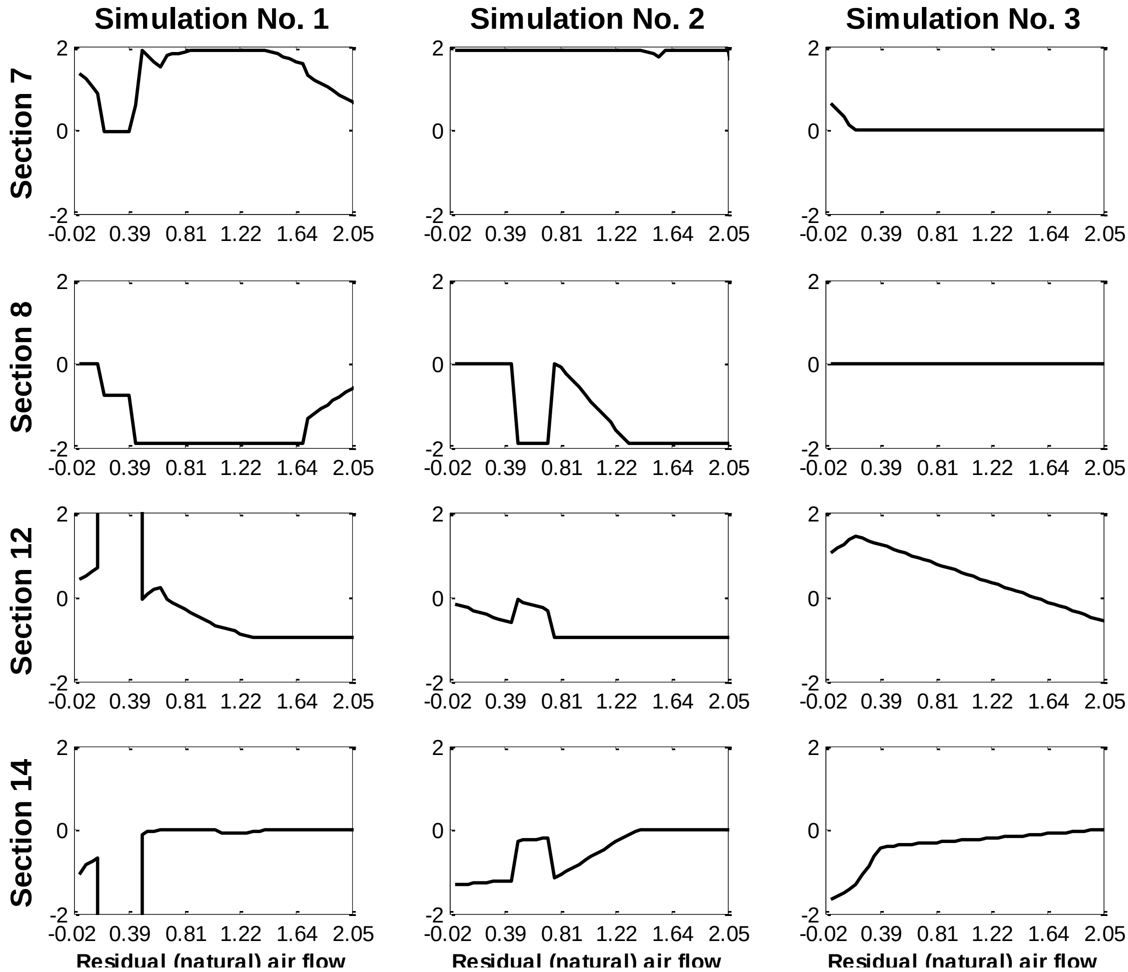


Figure 5: Comparison of controller tuned according to various cost functions.

The results show again a non-smooth behaviour, as the controller tries to balance the energy consumption. The first simulation is not quite desirable, as we can see that the controller counteracts by sections 12 and 14. Indeed, this is the controller with a uniform cost function. The second simulation is much better; we can see the effect of penalizing the onramp (section 12). The third simulation gives the best results, it even has quite smooth signal. The reason may be that preferring the ventilation shaft against other ventilators is natural, so it suits the controller the best.

5. CONCLUSIONS

We have shown an approach to ventilation control, which is based on MPC controller. This type of controller is widely used in industry, especially for large scale systems with multiple inputs and multiple outputs. This makes it an ideal tool for tunnel ventilation control, especially for city tunnels, where special requirements have to be met.

6. REFERENCES

Ferkl L., Meinsma G. (2007) Finding Optimal Control for Highway Tunnels. Tunnelling and Underground Space Technology, vol. 22, issue 2, pp. 222-229. ISSN 0886-7798.

Pořízek J., Ferkl L., Sládek O. (2007) Road Tunnel Ventilation Model: Simulation Analysis. Tunnel Management International, vol. 9, no. 1, pp. 65-72. ISSN 1463-242X.

Appendix C

Nývlt, Prívara, Ferkl: Probabilistic risk assessment of highway tunnels

NÝVLT, O. – PRÍVARA, S. – FERKL, L. Probabilistic risk assessment of highway tunnels. Tunnelling and Underground Space Technology. 2011, 26, 1, pp. 71–82. ISSN 0886-7798.

Number of citations: 0

Impact factor: 0.917

Tunnelling and Underground Space Technology 26 (2011) 71–82

Contents lists available at ScienceDirect

Tunnelling and Underground Space Technology

journal homepage: www.elsevier.com/locate/tust

Probabilistic risk assessment of highway tunnels

Ondřej Nývlt, Samuel Prívara, Lukáš Ferkl *

Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha 6, Czech Republic

ARTICLE INFO

Article history:
Received 11 January 2010
Received in revised form 21 June 2010
Accepted 29 June 2010
Available online 31 July 2010

Keywords:
Risk management
Risk analysis
Road tunnel
Costs effectiveness
Economic optimization

ABSTRACT

Many approaches to risk analysis in tunnels have been proposed by both international and national authorities over the last few years. Many safety problems have been discussed and a large number of important risk factors and hazards in tunnels have been identified. The concept of risk analysis in the scope of tunnel risks is, however, still under development; particularly an overall idea about the risk management concept is still missing. The paper introduces the concept of risk analysis in the scope of risk management and employs methods well-known in aeronautics and aircraft industry, yet, still unused in tunnels. The proposed methodology enables building and refurbishing costs minimization subject to preservation of satisfactory safety level. The outcomes of the proposed method have clear technical and economic interpretation and create a strong support tool for the decision making process. The paper also includes a case study of the Strahov tunnel in Prague, Czech Republic.

1. Introduction

1.1. Risk analyses in tunnels

Risk analysis is a tool developed initially in industries with potentially dangerous applications (chemical plants, nuclear power plants). According to Stamatelatos et al. (2002a) and Rausand and Høyland (2004), the purpose of risk analysis is to establish a proactive safety strategy by investigating potential risks. In last 15 years, risk analysis methods were also adapted in tunnel safety. Risk analysis in tunnel enables comparison of safety measures in terms of risk reduction as well as risk-based cost/effectiveness analysis, which can evaluate the cost of risk reduction.

Even though quite a long time has passed since the first risk analysis methods have been introduced in tunnels and a number of serious tunnel accidents has occurred, there has been no common standard or method used in PIARC[1] member countries (PIARC, 2007; PIARC, 2008). In spite of basic framework of road safety introduced by EU Directive 2004/54/EC (2004) on minimum safety requirements for tunnels in the Trans-Europeans Road Network and several recommendations issued by PIARC, most of the countries use their own methods. Oftentimes, a quantitative approach is chosen to calculate probabilities of respective events/scenarios/…, fire included, which is, according to the majority of available publications (Beard and Cope, 2007; PIARC, 2007; PIARC, 2008) the most serious threat in tunnels. The lack of statistical data for fire occurrence is an ultimate problem most methods are encountering. This is in sharp contrast with the essential statistical requirement for data validity. This paper proposes a method based upon probabilistic risk assessment, yet independent of calculating of fire probability.

The methods introduced in this paper are not fully unknown to tunneling. According to several presentations given at the World Tunnel Congress 2009, the fault trees and event trees have recently been introduced by some companies and institutions, but only for the construction phase of the tunnel, especially for mining (Sander et al., 2009; Yan and Ye, 2009). Sturk et al. (1996) aim to assess risks during the construction phase of a tunnel and to support the decision making process. The case study uses FMEA (Failure Modes and Effects Analysis) and FTA (Fault Tree Analysis) as separate risk analysis tools. On the other hand, Hong et al. (2009) use ETA (Event Tree Analysis) for similar purposes. Eskesen et al. (2004) present general guidelines for performing risk management in tunnels; however, application of specific Risk Analysis methods is not in the scope of their paper. Petelin et al. (2010) provide comparison of the most frequent risk analysis methods used in tunnels – QRAM, TuRisMo and RWQRA; general concept of the risk management is provided, which is similar to our approach, but no detailed information about the background of the specific methods is given. The Austrian TuRisMo approach, described, e.g. by Kohl et al. (2007), uses ETA and consequence analysis for accident scenarios, but again as a stand-alone tool. Holicky (2006) presents an alternative probabilistic risk analysis based on Bayesian networks; however, this approach requires a lot of accurate data, which are usually not available for tunnels.

Even though some of said papers (Eskesen et al., 2004; Hong et al., 2009; Sander et al., 2009; Sturk et al., 1996; Yan and Ye,

* Corresponding author. Tel.: +420 777 855 363; fax: +420 224 916 648.
E-mail address: ferkll@control.felk.cvut.cz (L. Ferkl).

[1] PIARC is an acronym for Permanent International Association of Road Congresses.

0886-7798/$ - see front matter © 2010 Elsevier Ltd. All rights reserved.
doi:10.1016/j.tust.2010.06.010

2009) and documents (EU Directive 2004/54/EC,; PIARC, 2007; PIARC, 2008) use similar risk analysis methods as will be presented in this paper, they do not exploit all possible outcomes they can provide. We will therefore focus on these additional features of risk analysis, in order to present more support for the decision making process.

1.2. Relationship of the risk analysis and risk management

In the state of the art, Risk Analyses (RA) are not considered as stand-alone tools, but are rather incorporated into a more complex Risk Management system (RM), which forms a part of a decision making process (Risk Management ..., 2002; Stamatelatos et al., 2002a). RM provides means for quality management, risk mitigation, production and maintenance planning, safety and reliability analysis, etc.

As illustrated in Fig. 1, the RM process has two major parts, which correspond to the engineering and managing departments of a company. The engineering departments perform the technical analysis which must provide a clear interface for the decision makers in the company management in order to carry out sound decisions.

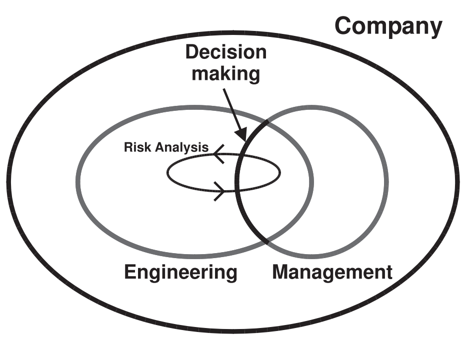


Fig. 1. Risk analysis as a part of a risk management process providing means for sound decision making.

In order to be efficient and to provide meaningful results, the RM process has to be scheduled for the entire lifetime of a system, as illustrated in Fig. 2. It is clear that each phase of the system life stage requires different approaches with respect to corresponding needs of decision making. Another factor is the input data available for the respective RA methods. If properly scheduled, the RM of a system is a continuous process that naturally follows the life cycle of the system (Risk Management ..., 2002). This continuity not only ensures appropriate results of the respective RA methods, but also saves significant amount of effort and resources needed for risk evaluation.

1.3. Risk optimization

One of the primary objectives of any RM process is to balance the cost of safety with the cost of accidents. It is very difficult to achieve as there is only a small evidence about the cost of accidents, while the cost of safety is usually known quite well. The problem is illustrated in Fig. 3.

The principle problem is to evaluate the total system risk. In any RA method, there are two factors that act against each other:

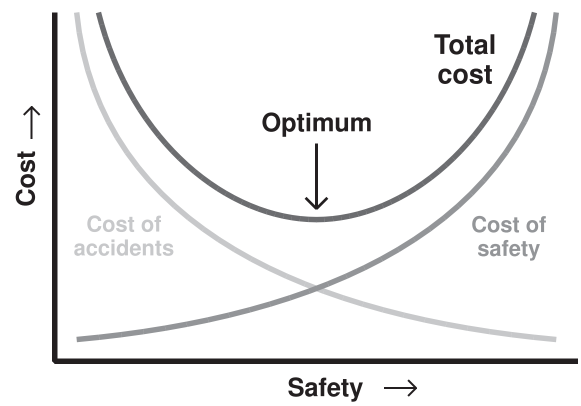


Fig. 3. Risk management – balance of cost of safety and cost of accidents.

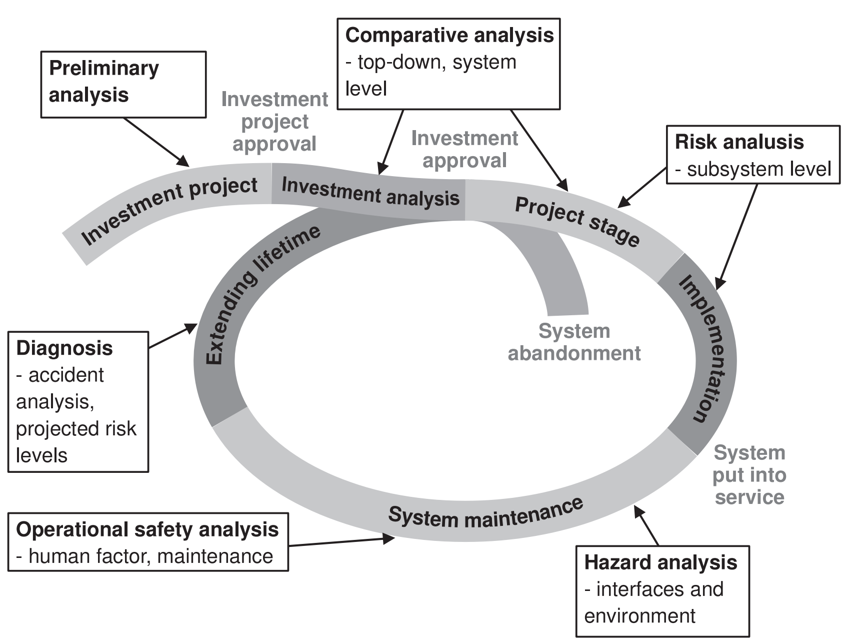


Fig. 2. The role of risk analyses in the system life cycle (Stamatelatos et al., 2002a).

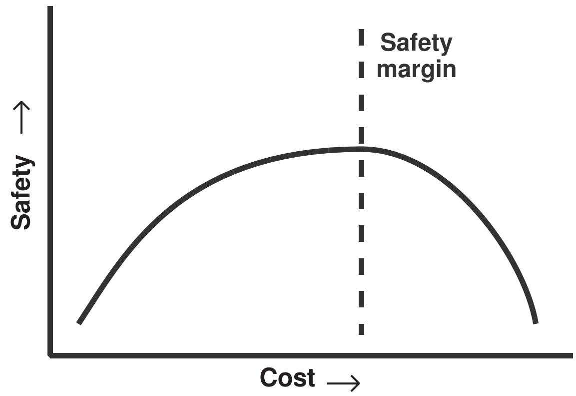


Fig. 4. Safety margin – a "too safe" approach can decrease the overall system safety.

- A risk estimate has to be "on the safe side", i.e. the calculated risk has to be greater or equal to the actual risk.
- The higher the calculated risk is, the higher will be the cost of appropriate mitigation measures.

This implies that complex methods that provide realistic risk estimates result in lower safety costs, because the risks are mitigated in an efficient way. This is a very important point emphasized in all basic safety documents in aerospace industry (e.g. FAA System Safety Handbook, 2000; Stamatelatos et al., 2002a) and is closely related to conditional probabilities.

Another problem is that "being on the safe side" does not necessarily mean "being on the safe side". Incorporating too many safety measures without any reflection of conditional probabilities leads to actual decrease of the real safety, as illustrated in Fig. 4. The method presented in this paper provides clear insight into the trade-off between costs and safety.

The paper is organized as follows: In Section 1, the problem of risk analysis in tunnels is formulated. Section 2 briefly introduces Probabilistic Risk Assessment (PRA) and its adjustments for the needs of this paper. In Section 3, adjusted PRA is performed for Strahov tunnel. Basic idea of the interrelationship between risk analysis and risk management process is introduced in Section 4. Section 5 concludes the paper.

2. Probabilistic risk assessment

As was already stated, the current methods used in RA for road tunnels have several serious problems, such as lack of statistical data on fires, non-quantified results of the analysis and therefore only "experience-based" mitigation of the risk decisions, no unified and standardized approach to the RA, thus an existence of a number of ad hoc methods, etc. The effort of the paper is an introduction of comprehensible, widely applicable and acceptable, yet clear approach to the estimation and assessment of risks in tunnels. The Probabilistic Risk Assessment (PRA) with Fault Tree Analysis (FTA) and Event Tree Analysis (ETA), widely used in aerospace, nuclear, chemical and other industries (FAA System Safety Handbook, 2000; NUREG, 1975; MIL-STD-1629A, 1980; MIL-STD-756B, 1981; Fault Tree Handbook…, 2002b) could be very conveniently applied also for common tunnel analysis. The goal of the usage of these methods is to get clear, comprehensible numerical results for both the RA and costs, i.e. the results should provide an unambiguous decision tool for management. The results should include current risk levels of investigated object, the contribution to the overall risk of its individual components or the sets of components, the numerical decrease/increase of the risk when a safety equipment is added/removed, and above all, also an economic cost of the risk mitigation. All of these afore mentioned criteria are fully satisfiable by using the PRA. An original NASA PRA process adjusted for the paper needs can be characterized in steps as follows (Stamatelatos et al., 2002a):

1. *Definition of the objective.* The objective of the risk assessment must be properly defined and the undesirable consequences, end states (ES_j), are identified. The project success criteria are necessary to define risk assessment end states in Eq. (1).
2. *Familiarization with the system.* All relevant information are gathered to familiarize with the system.
3. *Identification of IEs.* Initial Events (IE) of the event sequences (scenarios) are identified and analyzed by means of Master Logic Diagrams (MLD) or FMEA/FMECA analyses.
4. *Modeling of the scenarios.* Each accident scenario is developed in an inductive manner with probabilistic tool named Event Tree (ET). An ET starts with IE and continues through the scenario (pivotal events) until the end state is reached.
5. *Modeling of the failures.* Each failure of pivotal event in accident scenario is modeled in deductive manner by means of Fault Tree (FT). The top event of FT is given as a negation of the pivotal event defined in an accident scenario. Fig. 6 shows the relationship of the FT and ET.
6. *Collection, analysis and development of the data.* Variety of data types is collected to quantify the accident scenarios and main accident contributors.
7. *Quantification and Integration.* The frequency of occurrence of each end state in the ET is the product of the IE frequency and the (conditional!) probabilities of the pivotal events along the scenario path linking the IE to the end state. Scenarios are grouped according to the end state of the scenario defining the consequences and thereafter end states are grouped and their frequencies are summed up. The mathematically correct way of calculation of the expression for the frequency of a specific scenario, $\Lambda_{j,k}$ is as follows:
$$\Lambda_{j,k} = \Lambda(ES_{j,k}) = \lambda_j p(ES_{j,k}|IE_j), \qquad (1)$$
where λ_j stands for the frequency of the jth IE and $p(ES_{j,k}|IE_j)$ denotes the conditional probability for the end state of the event sequence ES_k (without initiating event IE_j), in the event tree initiated by IE_j given that IE_j has occurred.
8. *Importance ranking.* Ranking of risk scenarios provides insight regarding the contribution of individual events to the total risk. Scenario risk ranking shows the importance of group failures, not the individual events. If the event with significant contribution to the risk is in the structure of many low frequency scenarios, it may be absent in the definition of the dominant risk scenario and scenario risk ranking will not capture the risk importance of this event. To address this issue quantitative importance measures are calculated. When the importance measures are calculated, the events are ranked according to the relative value of the importance measure and treated further with respect to their rank.

2.1. PRA specifications

2.1.1. Initial event

As already mentioned, the fire is considered as the most serious and problematic threat to tunnel safety and therefore it was chosen as an Initial Event (IE) (Fig. 5). Fire occurs in tunnel with some probability (given by variety factors) and has different consequences on tunnel users and tunnel itself depending on incredible broad variety of causes such as tunnel equipment, tunnel users behavior, tunnel operator reactions and the proper reaction and subsequent actions of the rescue services. However, the probability of the occurrence of fire is enormously low, and the statistical data

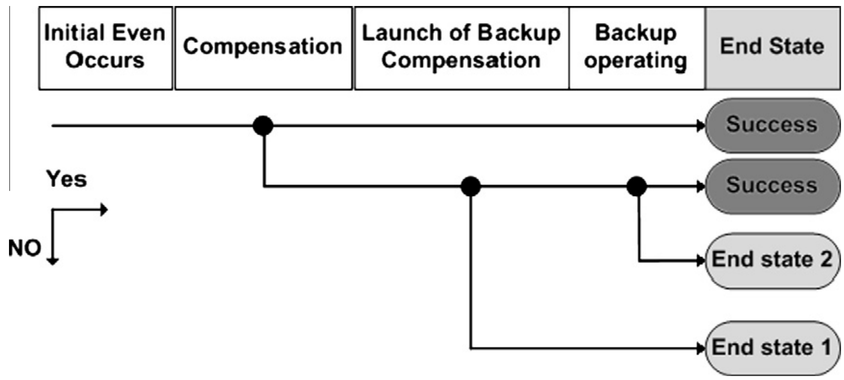


Fig. 5. Event tree.

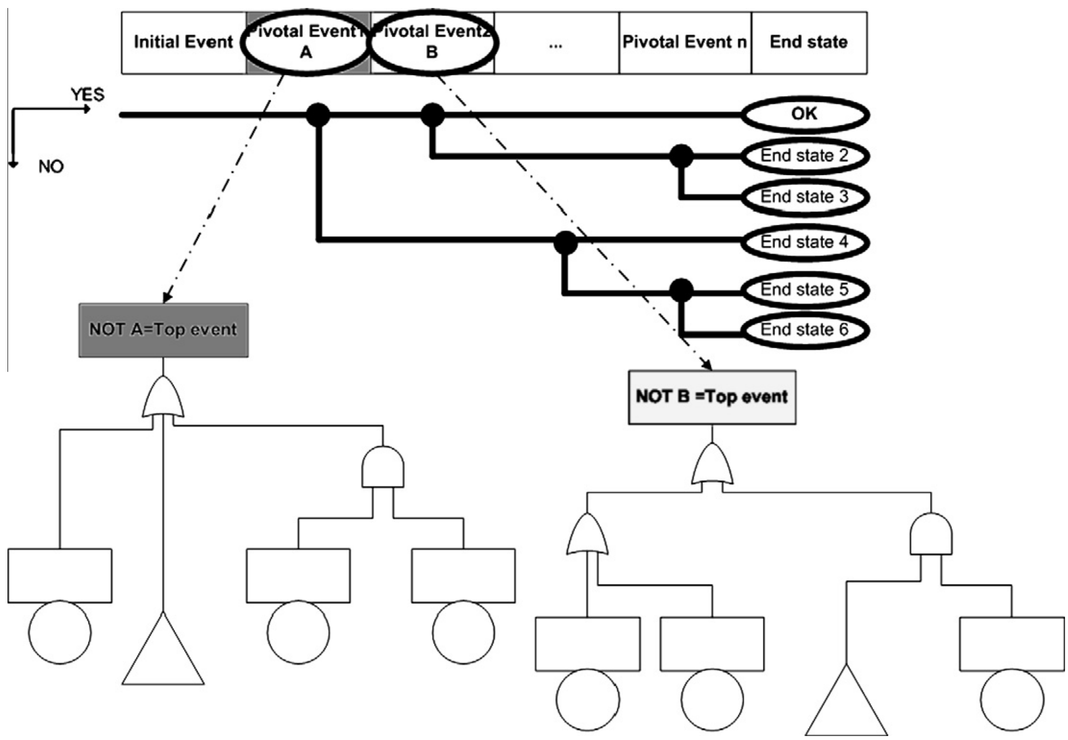


Fig. 6. Event and fault tree relationship.

are insufficient. Many studies have been provided, but neither one had solved the issue of fire in the sufficient manner, e.g. according to the statistical data of the Strahov tunnel, which is the object of the case study of this paper, there was only one fire of insignificant importance since the tunnel had been opened! There were no casualties, nor injuries, and the driver was able to extinguish the fire itself. Therefore, the proposed method uses Eq. (1) in an adjusted form as follows:

$$\Lambda^*_{j,k} = \Lambda^*(ES_{j,k}) = \alpha \cdot p(ES_{j,k}|IE_j), \tag{2}$$

where α stands for unknown probability of fire (IE). Because α is the same for all considered system structure configurations, it is omitted in calculations. In fact, only "tunnel equipment[2]" of the tunnel is taken into account. Therefore, the whole analysis is a comparison of the variety of the safety measures and precautions.

[2] Tunnel users, tunnel operators are also part of considered tunnel equipment.

2.1.2. Minimal cut sets

A Cut Set (CS) is a combination of basic events that can cause the top event. A Minimal Cut Set (MCS) is the smallest such a combination. Because the basic events are the bottom events, the MCSs relate the top event directly to the basic event causes. The structure of MCSs can provide valuable information, e.g. MCS with single event means, that the occurrence of this single event causes the failure of the whole system. The upgrade and prevention actions should be firstly focused on these single event MCSs. The MCSs are very often sorted by probability thus they can provide useful criterion of "severity" of the respective MCSs. Yet another example of MCS usage are importance measures.

2.1.3. Importance measures

As was already said, one of the greatest advantages of the FTA is its ability to express the contribution of an individual event to the overall risk probability. At the time of decision making process, it is useful to have the events sorted according to some criterion/criteria. This is especially useful, e.g. in the case, when (as it is always) the budget is limited and one has to decide which safety measures are crucial to implement and/or which critical elements in the system have to be "neutralized". This ranking is enabled by importance measures. Although Stamatelatos et al. (2002a) introduce four basic types of importance measures, only Fussel–Vesely and Risk Reduction Worth were used in our analysis.

- *Fussel–Vesely importance.* Alternative name for this measure is the Top Contribution Importance and reflects the contribution of individual MCS containing the basic event x_i to the overall risk. The F–V is calculated as follows:

$$I_{x_i}^{FV} = \frac{p\left(\bigcup_j MCS_j^{x_i}\right)}{p\left(\bigcup_j MCS_j\right)} = \frac{p\left(\bigcup_j MCS_j^{x_i}\right)}{p(TE)}, \tag{3}$$

where $p(\bigcup_j MCS_j^{x_i})$ is probability of the union of the MCSs containing event x_i, $p(\bigcup_j MCS_j)$ is probability of the union of all MCSs and $p(TE)$ is the top event probability. The Fussel–Vesely Importance measure shows the conditional probability that at least one MCS containing basic event x_i will occur, given that the system has failed. Alternative calculation of the Fussel–Vesely Importance measures follows as:

$$I_{x_i}^{FV} = \frac{p(TE) - p(TE|x_i = 0)}{p(TE)} \tag{4}$$

- *Risk Reduction Worth.* Alternative name is Top Decrease Sensitivity and implies the decrease of the probability of the top event under assumption of non-occurrence of a given event. For the basic events the Risk Reduction Worth shows the amount by which the risk decreases assuming that respective basic event, i.e. failure, will not occur. The Risk Reduction Worth is calculated by re-quantifying the FT with the probability of the given event set to 0.0 and mathematically as:

$$I_{x_i}^{RRW} = \frac{p(\bigcup_j MCS_j)}{p(TE|x_i = 0)} = \frac{p(TE)}{p(TE|x_i = 0)} \tag{5}$$

Risk Reduction Worth and Fussel–Vesely Importance measures are used to identify hardware elements, that are the biggest contributors to the overall risk. One can see, that there is a relationship between Fussel–Vesely Importance and Risk Reduction Worth, that can be expressed as:

$$I_{x_i}^{FV} = 1 - \frac{1}{I_{x_i}^{RRW}}. \tag{6}$$

2.1.4. Outcomes

Probabilistic risk assessment provides both qualitative and quantitative outcomes. System logic structure and minimal cut sets are produced by the qualitative part of the method. The great advantage of the procedure is its clarity even for non-professionals and great flexibility to new ideas. The quantitative part of the method provides trade-off matrix of ranked costs and safety as well as importance measures results. This enables the most effective trade-off solution between costs and safety (more in Section 3).

3. Case study: the Strahov tunnel

Directives of the European Union (EU Directive 2004/54/EC,) require a new Technical Documentation for tunnels every ten years. The present risk analysis is an integral part of the new technical documentation for the Strahov tunnel. Nowadays, the Strahov tunnel has several safety problems with both aged and missing equipment that is about to be replaced (or newly installed). Among others, e.g. new video surveillance system for transportation of dangerous goods (which is usually strictly prohibited in city tunnels, as it is in the Strahov tunnel), new longitudinal fans with sufficient performance capable to cope with fire up to 30 MW, "soft stop,[3]" and other equipment which is supposed to eliminate the danger of the accident (fire especially), or when it occurs, to suppress it in a sufficient manner, providing enough time for escape. The analysis provides not only the risk levels incurred by the current safety measures, but also evaluates the contribution of new elements in the tunnel. The analysis also enables prioritizing of the safety measures with respect to the risk reduction and the cost, which is a great advantage considering the price of some tunnel equipment (millions of Euros).

3.1. Basic characteristics of the Strahov tunnel

The Strahov tunnel is unidirectional, twin-tube tunnel opened after long (12 years) and difficult process of building and applying the control system. The actual capacity of a tunnel tube is 43,000 vehicles per day (transportation of dangerous goods is prohibited). It is a part of the Prague City Ring which also includes tunnels Mrazovka and Blanka (under construction). Actual length of the Strahov tunnel is 2 km; the length of the Western Tube (WT) is 1997 m, the length of the Middle Tube (MT) is 1990 m with two portals at Malovanka and Plzenska. The daily average number of vehicles in WT is 32,000 and 25,200 in MT, 4% of heavy vehicles included. The Strahov tunnel contains eight emergency exits with fireproof doors (fire resistance 90 min, overpressure 1 kPa), gas proof walls, and one special "open" emergency exit ("the wall hole"), which causes serious ventilation problems. The emergency exits in WT are 73–422 m away from each other and 98–403 m in MT. Currently there is no longitudinal ventilation system in the Strahov tunnel, but it is about to be installed. There are 25 SOS boxes in the Strahov tunnel placed 54–226 m from each other. The tunnel is powered by 22 kV electricity distribution network, and an emergency power supply – two pieces of 220 V, 330 Ah batteries. The Strahov tunnel has its own video surveillance system composed of two separate parts. The first circle is connected to the automatic traffic congestion recognition (63 recording devices). The second video circle is independent on the first one and serves for the operators supervising the tunnel. The tunnel is outfitted with fire detection signalization device, Linear Heat Detection system, which serves for the fire detection.

The Czech and international authorities are requesting the proper quantitative analysis of the tunnel in order to classify the tunnel into a specific risk class. Because the equipment in the tunnel is quite old, there is a great effort to outfit the tunnel with new safety devices. In order to perform this efficiently from both risk reducing and cost effective point of view, a proper analysis had to be performed.

3.2. Safety precautions proposal

Based upon the risk analysis of the current state of the Strahov tunnel, several safety precautions had been proposed. They arose from discussions and brainstorming meetings with tunnel experts, with the support of PIARC documents PIARC (2007, 2008), European Directive 2004/54/EC, 2004 and Czech technical recommendation TP98 (Novák and Přibyl, 2006). The suggested precautions were as follows (see also Tables 4–6):

[3] By means of the soft stop it should be possible to stop drive-in to the tunnel in case of accident.

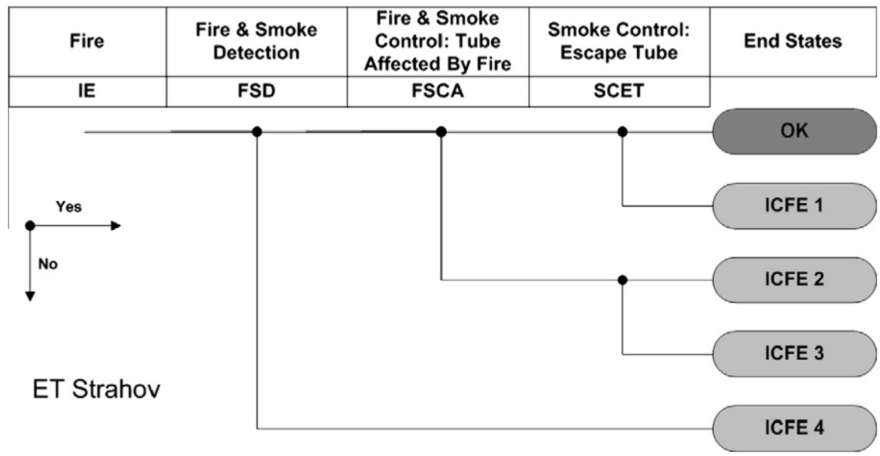


Fig. 7. Event tree of Strahov tunnel.

- Installation of smoke detection devices.
- Installation of longitudinal ventilation (about 8 axial fans per tube).
- Faster fans in the central machinery room of the transversal ventilation system.
- New dampers of the transversal ventilation.
- New dampers and fans in the emergency exits (cross-adits).
- Modification of the walls of the tunnel portals (Malovanka and Plzenska).
- New implementation of the control system software.
- Construction of a division wall inside the tunnel.

Furthermore, two precautions were identified to be necessary *a priori*, and were therefore not incorporated into the analysis:

- Control system hardware upgrade – the current system is not designed to aggressive environment, such as tunnels.
- Operating center staff training.

It is obvious that the important issue was to acquire relevant statistical data for each of the safety measures (both installed and proposed). Here are several examples of how the data were obtained:

- *Smoke detection device.* The device is currently not installed (formally: the probability of failure is 1). The smoke detection is inevitable to successfully identify and localize the initiating fire and therefore the smoke detectors will be installed close to dampers (next to the axial fans). The newly installed detectors have probability of failure lower than 0.01 and the estimated cost is 5 million Kc.[4]
- *Longitudinal ventilation.* The longitudinal fans are currently not installed, but they are key part of successful management of the smoke and fire, thus enable the safe evacuation of people. There will be four longitudinal fans installed at each end of a tube (16 altogether). This precaution costs 10 million Kc and the probability of failure is lower than 0.02.
- *Dampers.* Current dampers are not an integral part of the control system, i.e. they have their own logic of start-thermal fuse cut-outs (75 °C). It has already been proved, that the current system of start up is completely unreliable and unsuitable with probability of failure close to 1. Therefore there will be 48 new dampers installed into the tunnel, with surface of 10 m^2 each. This, however, requires also some construction works and the tunnel roof will not stay intact. The new system of dampers (probability of failure less than 0.02) should be a part of the control system. The new damper system is one of the most expensive solutions for the tunnel with estimated cost of 50 million Kc.
- *Tunnel wall.* There is an open space, a "hole" inside the tunnel that enables free passage from one tube to another. It was built with intention of traffic control for the case of existence of three tubes, however, there is no useful usage of it nowadays. In contrary, it poses serious threat to the tunnel safety, because it enables smoke penetration between the tubes, thus disables effective evacuation in the case of an accident. Because of these reasons, "the hole" should be walled-up, thus lowering the probability of failure from current 0.09 to 0 (according to unpublished CFD fire simulations performed by Satra, s.r.o.) with an approximate cost of 10 million Kc.
- *Staff.* The tunnel operators are not trained on regular basis, they do not have any simulator training or model situations training. It is therefore inevitable to introduce training procedures on regular basis. The estimation of staff reliability was based on a method presented by Stamatelatos et al. (2002a).

3.3. Adjusted probabilistic risk assessment applied for Strahov tunnel

The goal of the analysis is to evaluate current level of safety of the Strahov tunnel. Several options of safety solutions are proposed based upon this analysis providing both quantitative (risk analysis) and cost support. The final table is composed of the respective variants, its risk probabilities and cost of equipment, and serves as a basic decision making and support tool in the process of retrofitting of the Strahov tunnel. Due to the large result dispersion of the variety of analyses devoted to human behavior, it was decided, that only contribution of "technical part" of the tunnel would be taken into an account and analyzed, and therefore the analysis of evacuation and human behavior during the accident was omitted.

The risk analysis for the Strahov tunnel makes use of Event Tree and Fault Tree Analyses (ETA and FTA, respectively). The whole tunnel (analysis) was divided into three separate items:

- Fire and Smoke Detection (*FSD*).
- Fire and Smoke Control: tube Affected by fire (*FSCA*).
- Smoke Control: Escape Tube (*SCET*).

The Event Tree for the tunnel is depicted in Fig. 7.

Due to the problems, reasoned in Section 2, the whole analysis will be performed without the initial event "fire", i.e. the whole

[4] Kc is an abbreviation of Czech Crown (Koruna Ceska), currency unit of the Czech Republic, 1 EUR ~ 25 Kc (as of June 2010).

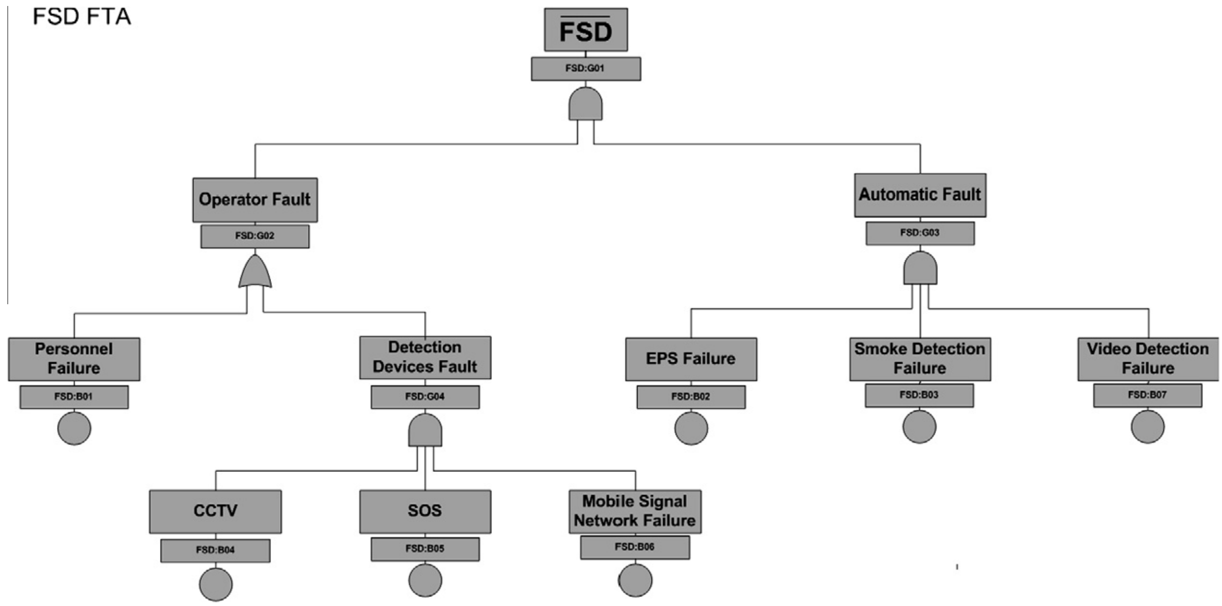


Fig. 8. Fire and Smoke Detection (FSD) fault tree.

analysis will be a comparison of the variety of safety measures and precautions.

The second event in the event sequence (Fig. 7) is Fire and Smoke Detection (FSD). The corresponding FT (Figs. 8–10) was developed for FSD contributing with its probability to the overall probability in the event sequence. The similar logic holds for the consequent events in the event sequence, FSCA and SCET, respectively.

Event sequences can end up to five different states (Fig. 7) depending on the combinations “en route”. The horizontal direction means, that the corresponding subsystem reacted “correctly” (e.g. FSD detected and properly identified fire) whilst the vertical directions means, that the subsystem failed to fulfill its task. The “matrix” of various “yes” and “no” directions represents the options of the scenario development in the tunnel. For the purposes of Strahov tunnel, only state “OK[5]” is of interest. All other states mean, that the lives of people will be somehow endangered. The meanings of the events in FSD, FSCA and SCET are explained in Tables 1–3, respectively.

The probabilities (or probability density functions) are propagated through the fault trees up to the top event. The negations of top event probabilities construct event sequence probabilities as stated above. The numerical results for “OK” state (costs included) are depicted in Tables 4 and 5, respectively. The first column shows an order of scenarios, second column provides the “OK” state probability, then 10 columns with occurrence/non-occurrence of respective equipment follow (each zero means, that the safety measure was not applied, each one means, that the safety item was taken into an account) and the cost of the respective solution is provided in the last column. The results show various possibilities of human and HW fault, because these two, according to the performed analysis, contribute in a greatest manner to the overall probability. In Table 4 one can see, that the overall probability is quite “good” if the new equipment (especially new HW complying with the EU norms) was provided and the staff trained properly. These results are in sharp contrast with Table 5. This figure reflects the current situation, when according to the analysis, every bigger accident that occurs in the tunnel means a major problem.

One of the main advantages of the performed style of risk analysis can be seen in Tables 4 and 5, respectively, when it enables the cost comparison of the results, e.g. by inspection of Table 4, items 11–13 have almost identical or identical probability, but the cost of the measures is completely different.

3.4. Fussel–Vesely importance measures

To help evaluating the significance of the proposed measures for the Strahov tunnel, Fussel–Vesely importance measures were calculated according to Eq. (3) (Table 6).

4. Consequences for the risk management process

It has to be stressed out that the risk analyses are nothing more (and nothing less) than support tools for sound decision making. As shown for the case of the Strahov tunnel, the PRA has been constructed based on some assumptions (see Section 3.3) that impose inevitable simplifications on the tunnel system. It is therefore necessary to combine several views of the problem (FTA, ETA, Fussel–Vesely importance measures, economical aspects, personal experience, legislative recommendations, ...) to make a competent decision.

For the case of the Strahov tunnel, the decision-making process, after the risk analyses were carried out, took several steps as follows (referring to Table 4):

1. A “red line” has been drawn to select combinations of safety measures with “fair” safety impact. The desired value has been chosen as the failure probability of 0.5, which corresponds to about 15% band of uncertainty in the input data.
2. Czech recommendation TP98 (Novák and Přibyl, 2006) was consulted and based on said documents, the combinations including the construction of the division wall and reconstruction of dampers and fans in the emergency exits were selected.
3. The Fussel–Vesely importance measures (Table 6) were taken into account and the selection has been further narrowed to combinations including longitudinal fans.

[5] State OK means, that fire was properly and in time detected, the control system had correctly reacted in less than 10 min in the case of personal vehicle accident or in less than 7 min in the case of heavy goods vehicle.

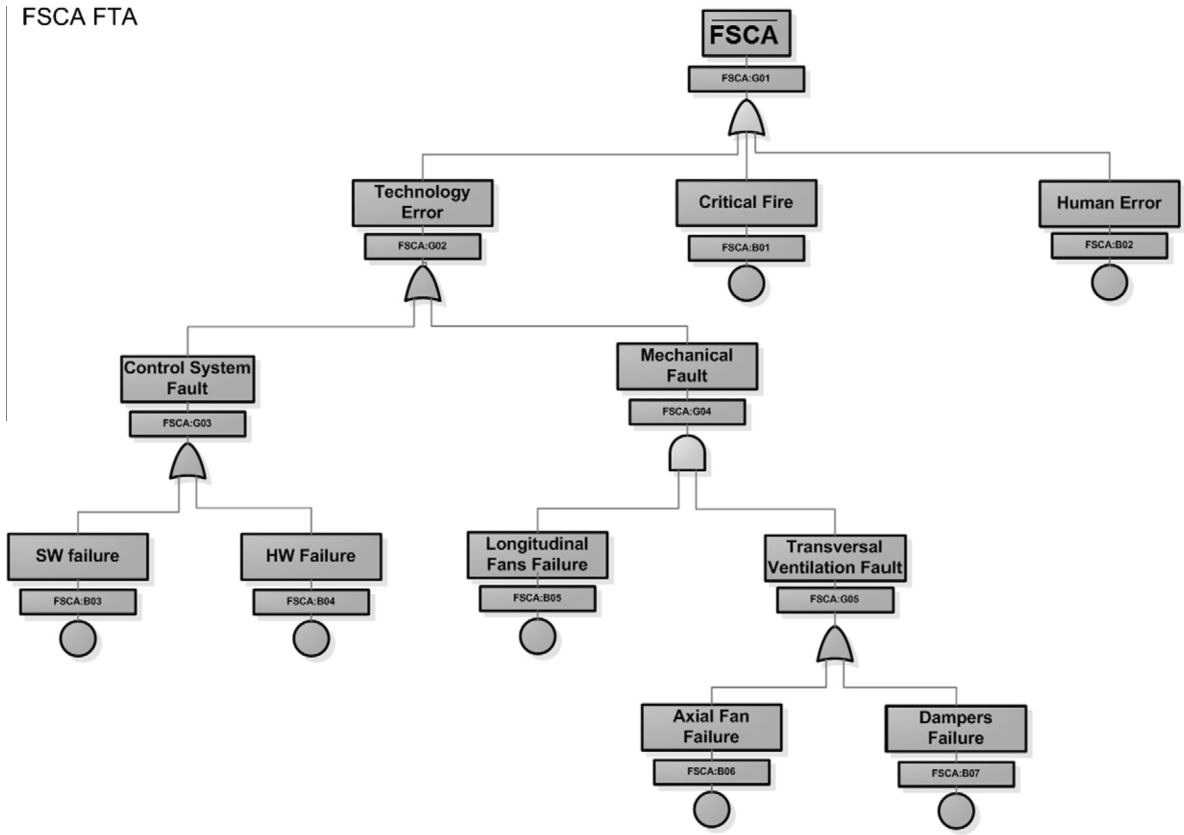


Fig. 9. Fire and Smoke Control: tube Affected by fire (FSCA) fault tree.

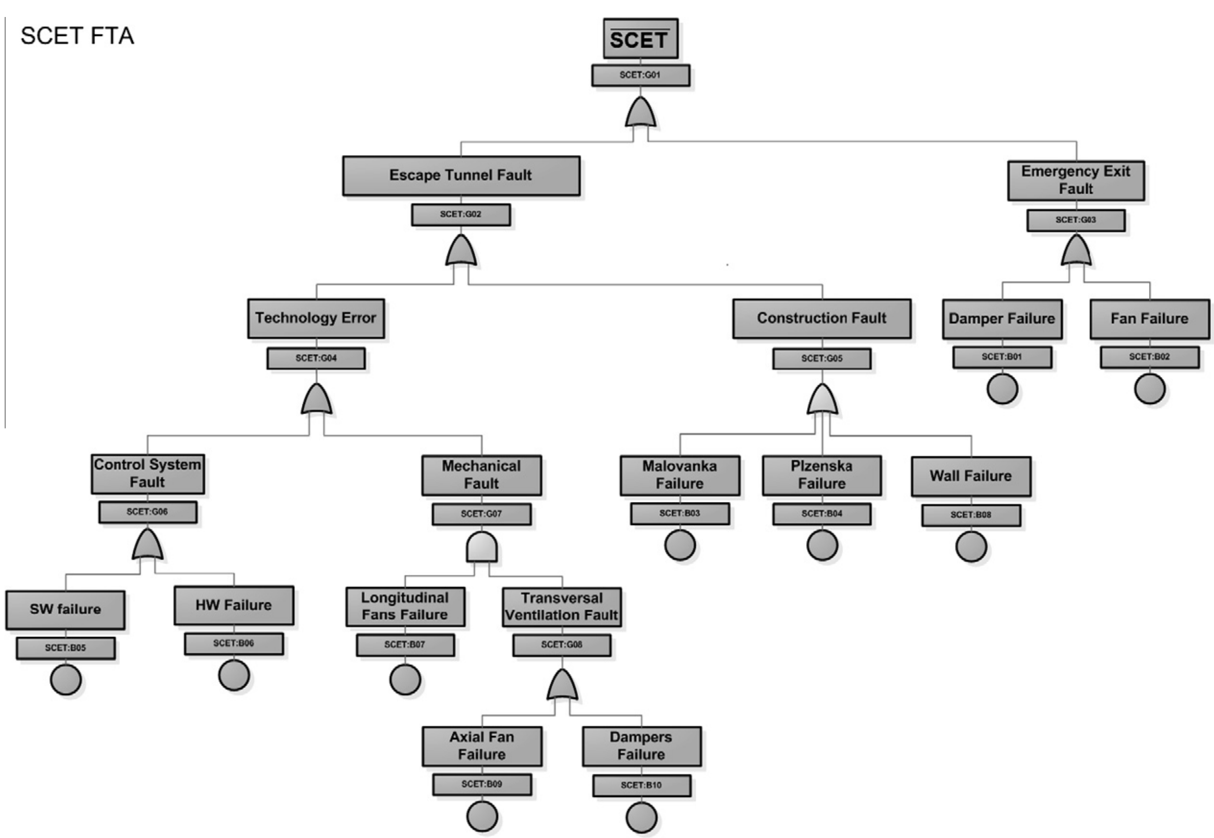


Fig. 10. Smoke Control: Escape Tube (SCET) fault tree.

Table 1
Events of Fire and Smoke Detection (FSD) fault tree.

Code	Name	Description
FSD:G01	Non-FSD	Top event, FSD detected and identified the fire and propagation of smoke correctly and in time
FSD:G02	Operator Fault	Personal fault of the operator – improper reaction or proper reaction not in time
FSD:G03	Automatic Fault	Fault of automatic detection equipment
FSD:G04	Detection Devices Fault	Fault of “other” safety equipment
FSD:B01	Personnel Failure	Improper human reaction
FSD:B02	LHD Failure	Linear heat detection sensor failure
FSD:B03	Smoke Detection Failure	Smoke sensitive sensors failure
FSD:B04	CCTV	Failure of the CCTV
FSD:B05	SOS	SOS box is not working
FSD:B06	Mobile Signal Network Failure	There is no network signal in the tunnel
FSD:B07	Video Detection	Video detection failure

Table 2
Events of Fire and Smoke Control: tube Affected by fire (FSCA) fault tree.

Code	Name	Description
FSCA:G01	Non-FSCA	Top event, the fire and smoke have not developed in affected tube thanks to safety measures
FSCA:G02	Technology Error	Variety of technological equipment failed to control and suppress smoke propagation
FSCA:G03	Control System Fault	Control system failed to launch fans properly
FSCA:G04	Mechanical Fault	Mechanical fault of fans
FSCA:G05	Transversal Ventilation Fault	Mechanical failure of transversal fans or/and dampers
FSCA:B01	Critical Fire	Fire with performance over 30 MW. This performance can not be treated with currently installed ventilation system
FSCA:B02	Human Error	Operator failed to launch the fire sequence properly and in time
FSCA:B03	SW Failure	SW failure of control system
FSCA:B04	HW Failure	HW failure of control system
FSCA:B05	Longitudinal Fans Failure	Mechanical failure of longitudinal fans
FSCA:B06	Axial Fans Failure	Mechanical failure of axial fans
FSCA:B07	Dampers Failure	Mechanical failure of dampers

Table 3
Events of Smoke Control: Escape Tube (SCET) fault tree.

Code	Name	Description
SCET:G01	Non-SCET	Top event, smoke did not reach escape tunnel due to proper operation of safety measures
SCET:G02	Escape tunnel Fault	Smoke in the escape tunnel due to technical or tunnel construction failures
SCET:G03	Emergency Exit Fault	Smoke in the emergency exit due to failure of dampers or fans
SCET:G04	Technology Error	Failure of technology or improper construction caused smoke in the escape tunnel
SCET:G05	Construction Fault	Improper construction at the portals or inside the tunnel (there is a huge passage “hole” between tubes that can cause smoke propagation from affected tube into escape tube)
SCET:G06	Control System Fault	Control system failed to launch fans properly
SCET:G07	Mechanical Fault	Mechanical fault of fans
SCET:G09	Transversal Ventilation Fault	Mechanical failure of transversal fans or/and dampers
SCET:B01	Damper Failure	Damper is unusable in the emergency exit
SCET:B02	Fan Failure	Emergency exit fan failed to operate properly
SCET:B03	Malovanka Failure	See Construction Fault
SCET:B04	Plzenska Failure	See Construction Fault
SCET:B05	SW Failure	SW failure of control system
SCET:B06	HW Failure	HW failure of control system
SCET:B07	Longitudinal Fans Failure	Mechanical failure of longitudinal fans
SCET:B08	Wall Failure	See Construction Fault
SCET:B09	Axial Fan Failure	Mechanical failure of axial fans
SCET:B10	Dampers Failure	Mechanical failure of dampers

4. Based on the prices of the resulting selection of combinations, the combination No. 13 (Table 4) was chosen as the final solution, including the following safety measures:
   - Smoke detection system.
   - Longitudinal ventilation.
   - New dampers and fans in emergency exits.
   - Modification of both portal walls (Malovanka and Plzenska portals).
   - New implementation of the control system software.
   - Construction of the division wall inside the tunnel.

   Later on, the modification of the Plzenska portal was left out, as its construction appeared to be too complicated (compared to its safety effect).

The main problem of the originally proposed improvements was the complexity of the solution – all possible improvements were considered and planned to be used, without their actual impact on the tunnel safety. Our analysis clearly stated the importance of the respective improvements and showed the combinations that were comparable in the sense of overall safety.

Table 4
Numerical results of PRA analysis including cost analysis with probability of human error 0.1 an probability of HW failure 0.1.

		Results of PRA analysis for Strahov tunnel Probabilities: human error = 0.1 and HW failure = 0.1										
		Smoke detection	Longitudinal ventilation	Fast fans	Dampers	Dampers-emergency exits	Fans-emergency exits	Malovanka-wall	Plzenska-wall	SW	Wall in the tunnel	Cost
1	0.6089	1	1	1	1	1	1	1	1	1	1	99,000,000.00 Kc
2	0.6082	0	1	1	1	1	1	1	1	1	1	94,000,000.00 Kc
3	0.6059	1	1	0	1	1	1	1	1	1	1	94,000,000.00 Kc
4	0.6053	0	1	0	1	1	1	1	1	1	1	89,000,000.00 Kc
5	0.5843	1	1	1	1	1	1	1	0	1	1	99,000,000.00 Kc
6	0.5837	0	1	1	1	1	1	1	0	1	1	94,000,000.00 Kc
7	0.5814	1	1	0	1	1	1	1	0	1	1	94,000,000.00 Kc
8	0.5808	0	1	0	1	1	1	1	0	1	1	89,000,000.00 Kc
9	0.5541	1	1	1	1	1	1	1	1	1	0	89,000,000.00 Kc
10	0.5535	1	1	1	1	1	1	0	1	1	1	89,000,000.00 Kc
11	0.5535	0	1	1	1	1	1	1	1	1	0	84,000,000.00 Kc
12	0.5533	1	1	1	0	1	1	1	1	1	1	49,000,000.00 Kc
13	0.5533	1	1	0	0	1	1	1	1	1	1	44,000,000.00 Kc
14	0.553	0	1	1	1	1	1	0	1	1	1	84,000,000.00 Kc
15	0.5528	0	1	1	0	1	1	1	1	1	1	44,000,000.00 Kc
16	0.5528	0	1	0	0	1	1	1	1	1	1	39,000,000.00 Kc
17	0.5514	1	1	0	1	1	1	1	1	1	0	84,000,000.00 Kc
18	0.5508	1	1	0	1	1	1	0	1	1	1	84,000,000.00 Kc
19	0.5508	0	1	0	1	1	1	1	1	1	0	79,000,000.00 Kc
20	0.5502	0	1	0	1	1	1	0	1	1	1	79,000,000.00 Kc
21	0.5317	1	1	1	1	1	1	1	0	1	0	89,000,000.00 Kc
22	0.5314	1	0	1	1	1	1	1	1	1	1	89,000,000.00 Kc
23	0.5312	1	1	1	1	1	1	0	0	1	1	89,000,000.00 Kc
24	0.5311	0	1	1	1	1	1	1	0	1	0	84,000,000.00 Kc
25	0.531	1	1	1	0	1	1	1	0	1	1	49,000,000.00 Kc
26	0.531	1	1	0	0	1	1	1	0	1	1	44,000,000.00 Kc
27	0.5309	0	0	1	1	1	1	1	1	1	1	84,000,000.00 Kc
28	0.5306	0	1	1	1	1	1	0	0	1	1	84,000,000.00 Kc
29	0.5304	0	1	1	0	1	1	1	0	1	1	44,000,000.00 Kc
30	0.5304	0	1	0	0	1	1	1	0	1	1	39,000,000.00 Kc
31	0.5291	1	1	0	1	1	1	1	0	1	0	84,000,000.00 Kc
32	0.5285	1	1	0	1	1	1	0	0	1	1	84,000,000.00 Kc
33	0.5285	0	1	0	1	1	1	1	0	1	0	79,000,000.00 Kc
34	0.528	0	1	0	1	1	1	0	0	1	1	79,000,000.00 Kc
35	0.5099	1	0	1	1	1	1	1	0	1	1	89,000,000.00 Kc
36	0.5094	0	0	1	1	1	1	1	0	1	1	84,000,000.00 Kc
37	0.5037	1	1	1	1	1	1	0	1	1	0	79,000,000.00 Kc
38	0.5035	1	1	1	0	1	1	1	1	1	0	39,000,000.00 Kc
39	0.5035	1	1	0	0	1	1	1	1	1	0	34,000,000.00 Kc
40	0.5032	0	1	1	1	1	1	0	1	1	0	74,000,000.00 Kc
41	0.503	1	1	1	0	1	1	0	1	1	1	39,000,000.00 Kc
42	0.503	1	1	0	0	1	1	0	1	1	1	34,000,000.00 Kc
43	0.503	0	1	1	0	1	1	1	1	1	0	34,000,000.00 Kc
44	0.503	0	1	0	0	1	1	1	1	1	0	29,000,000.00 Kc
45	0.5025	0	1	1	0	1	1	0	1	1	1	34,000,000.00 Kc
46	0.5025	0	1	0	0	1	1	0	1	1	1	29,000,000.00 Kc
47	0.5012	1	1	0	1	1	1	0	1	1	0	74,000,000.00 Kc
48	0.5007	0	1	0	1	1	1	0	1	1	0	69,000,000.00 Kc
49	0.4836	1	0	1	1	1	1	1	1	1	0	79,000,000.00 Kc
50	0.4834	1	1	1	1	1	1	0	0	1	0	79,000,000.00 Kc

Table 5
Numerical results of PRA analysis including cost analysis with probability of human error 0.7 an probability of HW failure 0.5

		Results of PRA analysis for Strahov tunnel Probabilities: human error = 0.7 and HW failure = 0.5										
		Smoke detection	Longitudinal ventilation	Fast fans	Dampers	Dampers-emergency exits	Fans-emergency exits	Malovanka-wall	Plzenska-wall	SW	Wall in the tunnel	Cost
1	0.0626	1	1	1	1	1	1	1	1	1	1	99,000,000.00 Kc
2	0.0623	1	1	0	1	1	1	1	1	1	1	94,000,000.00 Kc
3	0.0622	0	1	1	1	1	1	1	1	1	1	94,000,000.00 Kc
4	0.0619	0	1	0	1	1	1	1	1	1	1	89,000,000.00 Kc
5	0.0601	1	1	1	1	1	1	1	0	1	1	99,000,000.00 Kc
6	0.0598	1	1	0	1	1	1	1	0	1	1	94,000,000.00 Kc
7	0.0597	0	1	1	1	1	1	1	0	1	1	94,000,000.00 Kc
8	0.0594	0	1	0	1	1	1	1	0	1	1	89,000,000.00 Kc
9	0.057	1	1	1	1	1	1	1	1	1	0	89,000,000.00 Kc
10	0.0569	1	1	1	1	1	1	0	1	1	1	89,000,000.00 Kc
11	0.0569	1	1	1	0	1	1	1	1	1	1	49,000,000.00 Kc
12	0.0569	1	1	0	0	1	1	1	1	1	1	44,000,000.00 Kc
13	0.0567	1	1	0	1	1	1	1	1	1	0	84,000,000.00 Kc
14	0.0567	1	1	0	1	1	1	0	1	1	1	84,000,000.00 Kc
15	0.0566	0	1	1	1	1	1	1	1	1	0	84,000,000.00 Kc
16	0.0565	0	1	1	1	1	1	0	1	1	1	84,000,000.00 Kc
17	0.0565	0	1	1	0	1	1	1	1	1	1	44,000,000.00 kc
18	0.0565	0	1	0	0	1	1	1	1	1	1	39,000,000.00 Kc
19	0.0563	0	1	0	1	1	1	1	1	1	0	79,000,000.00 Kc
20	0.0563	0	1	0	1	1	1	0	1	1	1	79,000,000.00 Kc
21	0.0547	1	1	1	1	1	1	1	0	1	0	89,000,000.00 Kc
22	0.0547	1	0	1	1	1	1	1	1	1	1	89,000,000.00 Kc
23	0.0546	1	1	1	1	1	1	0	0	1	1	89,000,000.00 Kc
24	0.0546	1	1	1	0	1	1	1	0	1	1	49,000,000.00 Kc
25	0.0546	1	1	0	0	1	1	1	0	1	1	44,000,000.00 Kc
26	0.0544	1	1	0	1	1	1	1	0	1	0	84,000,000.00 Kc
27	0.0544	1	1	0	1	1	1	0	0	1	1	84,000,000.00 Kc
28	0.0543	0	1	1	1	1	1	1	0	1	0	84,000,000.00 Kc
29	0.0543	0	0	1	1	1	1	1	1	1	1	84,000,000.00 Kc
30	0.0543	0	1	1	1	1	1	0	0	1	1	84,000,000.00 Kc
31	0.0542	0	1	1	0	1	1	1	0	1	1	44,000,000.00 Kc
32	0.0542	0	1	0	0	1	1	1	0	1	1	39,000,000.00 Kc
33	0.0541	0	1	0	1	1	1	1	0	1	0	79,000,000.00 Kc
34	0.054	0	1	0	1	1	1	0	0	1	1	79,000,000.00 Kc
35	0.0525	1	0	1	1	1	1	1	0	1	1	89,000,000.00 Kc
36	0.0521	0	0	1	1	1	1	1	0	1	1	84,000,000.00 Kc
37	0.0518	1	1	1	1	1	1	0	1	1	0	79,000,000.00 Kc
38	0.0518	1	1	1	0	1	1	1	1	1	0	39,000,000.00 Kc
39	0.0518	1	1	0	0	1	1	1	1	1	0	34,000,000.00 Kc
40	0.0517	1	1	1	0	1	1	0	1	1	1	39,000,000.00 Kc
41	0.0517	1	1	0	0	1	1	0	1	1	1	34,000,000.00 Kc
42	0.0516	1	1	0	1	1	1	0	1	1	0	74,000,000.00 Kc
43	0.0515	0	1	1	1	1	1	0	1	1	0	74,000,000.00 Kc
44	0.0514	0	1	1	0	1	1	1	1	1	0	34,000,000.00 Kc
45	0.0514	0	1	0	0	1	1	1	1	1	0	29,000,000.00 Kc
46	0.0514	0	1	1	0	1	1	0	1	1	1	34,000,000.00 Kc
47	0.0514	0	1	0	0	1	1	0	1	1	1	29,000,000.00 Kc
48	0.0512	0	1	0	1	1	1	0	1	1	0	69,000,000.00 Kc
49	0.0497	1	0	1	1	1	1	1	1	1	0	79,000,000.00 Kc
50	0.0497	1	1	1	1	1	1	0	0	1	0	79,000,000.00 Kc

Table 6
Fussel–Vesely importance measures for the Strahov tunnel. The codes correspond to labels in the respective fault trees. Safety measures typed in bold face are the proposed new measures.

Safety measure	F–V	Code
Division wall	0.000	SCET:B08
Longitudinal fans	0.032	FSCA:B05, SCET:B07
CCTV	0.035	FSD:B04
SOS box	0.035	FSD:B05
GSM signal	0.035	FSD:B06
Critical fire	0.042	FSCA:B01
Wall–Malovanka	0.051	SCET:B03
Wall–Plzenska	0.051	SCET:B04
Emergency exits fans	0.101	SCET:B08
Dampers	0.110	FSCA:B07, SCET:B10
Software	0.408	FSCA:B03, SCET:B05
Axial fans	0.442	FSCA:B06, SCET:B09
Hardware	0.711	FSCA:B04, SCET:B06
Personnel 1	0.969	FSD:B01
Personnel 2	0.982	FSCA:B02, FSD:B01
LHD cable	1.000	FSD:B02
Smoke detection	1.000	FSD:B03
Video detection	1.000	FSD:B07

The savings achieved by our analysis reached about 67% of the original price. It was difficult to convince the tunnel operating company about the accuracy of our analysis; however, the PRA method has been used for years in aerospace industry and besides relevant mathematical algorithms, it also comprises means for comprehensive presentation of results for the decision makers (such as importance measures, financial tables, etc.) This turned out to be the crucial point in convincing the tunnel operating company, as they became familiar with the results and our solution has been finally accepted.

5. Conclusions

The paper introduces concept of adjusted Probabilistic Risk Assessment into the framework of advanced risk analysis and enables incorporation into risk management process. It provides support material for decision makers and risk managers to choose sound and cost effective solution. In the case of Strahov tunnel, the proposed analysis was able to cut almost 67% of originally calculated costs.

The principal advantage of the proposed PRA method is its independency on the object of the method, i.e. it is applicable to any type of tunnel, with any equipment, mode of traffic, etc. Compared to the state-of-the-art methods (TuRisMo, QRAM, ...), PRA uses accurate mathematical background suitable for data with high level of uncertainty; the lack of relevant statistical data is taken into account in the results by means of, e.g. specific importance measures. The drawback of the method is a "free" structure – it is more a framework than a guideline and each application has to be treated very carefully.

The method used in this paper originates mainly from aerospace industry, where many relevant data are available. Unfortunately, this is not the case of tunnels. The method proposed in this paper therefore requires further investigations towards qualitative results, which would bring comparative measures into the decision making process.

Acknowledgment

This work was supported by the Czech Ministry of Education, Youth, and Sports under the EuSophos Project No. 2C06010.

References

Beard, A., Cope, D., 2007. Assessment of the safety of tunnels. In: Background Paper for the Workshop "Assessment of the Safety of Tunnels", Brussels.
Eskesen, S.D., Tengborg, P., Kampmann, J., Veicherts, T.H., 2004. Guidelines for tunnelling risk management: international tunnelling association, Working Group No. 2. Tunnelling and Underground Space Technology 19(3), 217–237.
EU Directive 2004/54/EC, 2004. Directive 2004/54/EC of the European Parliament and of the Council on minimum safety requirements for tunnels in the Trans-European Road Network. European Commission, Directorate-General for Energy and Transport, Brussels.
FAA System Safety Handbook, 2000. FAA System Safety Handbook. Federal Aviation Administration, Washington, DC.
Holicky, M., 2006. Risk analysis and optimization of road tunnels. In: Brebbia, C.A., Popov, V. (Eds.), Risk Analysis V: Simulation and Hazard Mitigation, Wit Transactions on Ecology and the Environment, vol. 91, pp. 57–66, 5th International Conference on Simulation in Risk Analysis and Hazard Mitigation, Malta, June 19–21, 2006.
Hong, E.-S., Lee, I.-M., Shin, H.-S., Nam, S.-W., Kong, J.-S., 2009. Quantitative risk evaluation based on event tree analysis technique: application to the design of shield TBM. Tunnelling and Underground Space Technology 24 (3), 269–277.
Kohl, B., Botschek, K., Horhan, R., 2007. Austrian risk analysis for road tunnels development of a new method for the risk assessment of road tunnels. In: Bartak, J., Hrdina, I., Romancov, G., Zlamal, J. (Eds.), Underground Space – The 4th Dimension of Metropolises, vols. 1–3. Proceedings and Monographs in Engineering, Water and Earth Sciences, pp. 1835–1840. 33rd ITA-AITES World Tunnel Congress, Prague, Czech Republic, May 05–10, 2007.
MIL-STD-1629A, 1980. MIL-STD-1629A: Procedures for Performing a Failure Mode, Effects and Criticality Analysis. Department of Defence of the USA, Washington, DC.
MIL-STD-756B, 1981. MIL-STD-756B: Reliability Modeling and Prediction. Department of Defence of the USA, Washington, DC.
Novák, M., Přibyl, P., 2006. TP98 – technological equipments of road tunnels. Ministry of Transportation of the Czech Republic, Prague.
NUREG, 1975. Reactor safety study. Report WASH-1400, Nuclear Regulatory Commission.
Petelin, S., Luin, B., Vidmar, P., 2010. Risk analysis methodology for road tunnels and alternative routes. Strojniski Vestnik – Journal of Mechanical Engineering 56 (1), 41–51.
PIARC, 2007. Integrated Approach to Road Tunnel Safety. PIARC, La Defense.
PIARC, 2008. Risk Analysis for Road Tunnels. PIARC, La Defense.
Rausand, M., Høyland, A., 2004. System Reliability Theory, Models, Statistical Methods, and Applications. John Wiley & Sons, Inc., Hoboken, NJ, USA.
Risk Management, 2002. Risk Management Procedural Requirements. NASA, Office of Safety and Mission Assurance, Washington, DC, NASA Procedural Requirements NPR 8000.4.
Sander, P., Spiegl, M., Schneider, E., 2009. Risk management for large infrastructure projects using probabilistic methods. In: Proceedings of the ITA-AITES World Tunnel Congress 2009, Budapest, Hungary.
Stamatelatos, M., Apostolakis, G., Dezfuli, H., Everline, C., Guarro, S., Moieni, P., Mosleh, A., Paulos, T., Youngblood, R., 2002a. Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners. NASA Office of Safety and Mission Assurance, Washington, DC (version 1.1, August).
Stamatelatos, M., Vesely, W., Dugan, J., Fragola, J., Minarick, J., Railsback, J., 2002b. Fault Tree Handbook with Aerospace Applications. NASA Office of Safety and Mission Assurance, Washington, DC (version 1.1, August).
Sturk, R., Olsson, L., Johansson, J., 1996. Risk and decision analysis for large underground projects, as applied to the stockholm ring road tunnels. Tunnelling and Underground Space Technology 11 (2), 157–164.
Yan, J., Ye, S., 2009. Risk evaluation and control of Yesanguan railway tunnel on Yiwan railway line, China. In: Proceedings of the ITA-AITES World Tunnel Congress 2009, Budapest, Hungary.

Appendix D

Ferkl, Široký: Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods

FERKL, L. – ŠIROKÝ, J. Celing radiant cooling: Comparison of ARMAX and subspace identification modelling methods. Building and Environment. 2010, 45, 1, pp. 205–212. ISSN 0360-1323.

Number of citations (excluding autocitations):

- WoS: 1
- Scopus: 4
- Google Scholar: 3

Impact factor: 2.129

Building and Environment 45 (2010) 205–212

Contents lists available at ScienceDirect

Building and Environment

journal homepage: www.elsevier.com/locate/buildenv

Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods

Lukáš Ferkl [a,*], Jan Široký [b]

[a] Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha 6, Czech Republic
[b] Energocentrum Plus s.r.o., Prague, Czech Republic

ARTICLE INFO

Article history:
Received 2 March 2009
Received in revised form
4 May 2009
Accepted 5 June 2009

Keywords:
Cooling
Building management systems
System identification
Subspace methods
ARMAX
Modelling

ABSTRACT

The building of the Faculty of Mechanical Engineering, Czech Technical University in Prague – Dejvice, has been heated by a ceiling radiant heating system ("Crittall" system) since 1961. One building block has been equipped by means which enable the use of the pipes of the radiant heating system for ceiling radiant cooling. As optimization of this system is desired, the first step to be performed is the identification. Because of a complicated physical description of the entire system, and a set of measured data which is large enough to perform statistical identification, a decision was made to identify the system by ARMAX model identification and subspace identification methods. The resulting identified model shows a standard deviation of 0.2–0.3 °C on the verification data. The model identified by the subspace methods, enhanced by a Kalman filter, shows a standard deviation of 0.063 °C for output tracking. Such a model is ready to be used in modern multidimensional controllers, such as model-based predictive controllers (MPC).

1. Motivation

Rapidly growing costs of energy in recent years have shown us the importance of optimization of the heating and cooling of buildings. Besides technical improvements in building technologies (such as thermal insulation), new methods of technology control and regulation have been employed that use smart algorithms to decrease the energy consumption of buildings, while maintaining a comfortable environment inside the building [1]. Said modern methods make use of computer models of the controlled systems – of the buildings.

There are various approaches to computer modelling of buildings and their internal environment. For projecting and optimizing the building systems, simulation tools based on a mathematical description of the physical system are very popular [2,3]. The simulation parameters are physical parameters of the system, and the measured data can be used for model calibration. A good overview of this class of modelling methods can be found e.g. in [4].

This paper presents a different modelling approach. The simulation model, which will be further used for control prediction and energy optimization, is based on a general description of an unknown ("black-box") system. Identification of such a general system may be achieved by statistical methods that find the system dynamics and parameters from measured input and output data.

The model of the building is intended for two purposes. The first one is a pure prediction – the HVAC system operators may see the effects of their control efforts; the model is a tool to assist their personal experience with the system. The second purpose of the model is that it can be incorporated into a Model Predictive Controller (MPC), a mathematical control method based on quadratic programming that optimizes the control efforts with respect to the prediction of the system behaviour. Even though the second option brings more comfort and optimal performance of the system, it requires essential changes to the control system, which may not be desirable.

In the first part of this paper, the presented methods of statistical identification – ARMAX model identification and subspace identification – are briefly introduced. In the second part, our system of interest – the building of the Faculty of Mechanical Engineering, Czech Technical University in Prague, Czech Republic – is described. In the third part, practical results of the identification experiments are shown. The fourth part of this paper summarizes the results, and outlines further work.

Abbreviations: 4SID, Subspace State Space IDentification; ARMAX, Auto-Regressive Moving Average with eXogenous inputs; MIMO, multiple-input, multiple-output system; SISO, single-input, single-output system; PEM, prediction error method; CTU, Czech Technical University in Prague; HVAC, heating, ventilating and air conditioning; PID, proportional–integral–derivative controller.

* Corresponding author. Tel.: +420 777 855 363; fax: +420 224 916 648.
E-mail address: ferkll@control.felk.cvut.cz (L. Ferkl).

0360-1323/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.buildenv.2009.06.004

2. Overview of modelling methods

2.1. General notation and overview

2.1.1. Input–output model

For the identification of the ARMAX model, we have to define the ARMAX model (Auto-Regressive Moving Average model with eXogenous inputs), which is based on a general discrete-time transfer function model

$$y(k) = G(d)u(k) = \frac{b(d)}{a(d)}u(k) \tag{1}$$

where k is the discrete time, $y(k)$ is the model output, $u(k)$ is the model input, and

$$a(d) = 1 + a_1 d + \cdots + a_{n_a} d^{n_a} \tag{2}$$

$$b(d) = b_0 + b_1 d + \cdots + b_{n_b} d^{n_b} \tag{3}$$

are polynomials of degree n_a and n_b respectively in d-space. The d-operator can be seen as an inverse z-operator ($d = z^{-1}$) in the $\mathcal{Z}$ transform representing a unity delay. Technically, this is a simplification, the difference between d and z^{-1} can be found e.g. in [5].

We further need to introduce the noise signal $e(t)$ to Equation (1), which is achieved by

$$\begin{aligned} y(k) + a_1 y(k-1) + \cdots + a_{n_a} = {} & b_0 u(k) + b_1 u(k-1) + \cdots \\ & + b_{n_b} u(k-n_b) + e(k) \\ & + c_1 e(k-1) + \cdots + c_{n_c} e(k-n_c) \end{aligned} \tag{4}$$

where

$$c(d) = 1 + c_1 d + \cdots + c_{n_c} d^{n_c} \tag{5}$$

is a polynomial of degree n_c in d-space. Equation (4) can be written in a compact way

$$a(d)y(k) = b(d)u(k) + c(d)e(k) \tag{6}$$

or equivalently

$$y(k) = \frac{b(d)}{a(d)}u(k) + \frac{c(d)}{a(d)}e(k) \tag{7}$$

or

$$y(k) = G(d)u(k) + H(d)e(k) \tag{8}$$

which is the input–output model of the ARMAX process. As will be further described, the ARMAX model is usually derived for SISO (single-input, single-output) cases only, as the MIMO (multiple-input, multiple-output) case does not have any suitable canonical form.

2.1.2. State space model

For the needs of subspace identification, we will consider a lumped, discrete time, linear, time-invariant state space model in the form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k) \end{aligned} \tag{9}$$

where $u(k) \in \mathbb{R}^m$ is the m-dimensional input vector at time instant k, $y(k) \in \mathbb{R}^l$ is the l-dimensional output vector, $x(k) \in \mathbb{R}^n$ is the n-dimensional state vector. The vectors $w(k) \in \mathbb{R}^n$ and $v(k) \in \mathbb{R}^l$ are process and measurement noise signals respectively. We assume that these signals are zero mean, stationary, white noise signals. We further assume that the expected value of the noise signals satisfies

$$\mathbf{E}\left[\begin{pmatrix} w_p \\ v_p \end{pmatrix}\begin{pmatrix} w_q^T & v_q^T \end{pmatrix}\right] = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}\delta_{pq} \geq 0 \tag{10}$$

where δ_{pq} is Kronecker delta and $Q \in \mathbb{R}^{n\times n}$, $R \in \mathbb{R}^{l\times l}$, $S \in \mathbb{R}^{n\times l}$ are covariance matrices of the noise signals $w(k)$ and $v(k)$.

Let us further assume that the model (9) is observable and controllable.

2.1.3. Relationship of input–output and state space models

It is clear that the auto-regressive model (7) and state space model (9) should describe the same system. By applying $\mathcal{Z}$ transform onto the deterministic part of model (9), we get the transfer function

$$G(z) = C(zI - A)^{-1}B + D \tag{11}$$

Note that the initial state x_0 cannot be represented by the transfer function.

The output spectral density S_{yy} of the stochastic part of model (9) is

$$s_{yy}(z) = C(zI - A)^{-1}Q\left(z^{-1}I - A\right)^{-T}C^T + R = H(z)\textstyle\sum_e H^T\left(z^{-1}\right) \tag{12}$$

where H(z) is an impulse response matrix and $\sum_e$ is a covariance matrix of the stochastic part of the model (9). Now the general structure of the input–output linear, stochastic system is

$$y(k) = G(d)u(k) + H(d)e(k) \tag{13}$$

where

$$G(d) = G_0 + G_1 d + G_2 d^2 + \cdots \tag{14}$$

$$H(d) = I + H_1 d + H_2 d^2 + \cdots \tag{15}$$

are impulse characteristic matrices of the deterministic transfer function, and the noise shaping filter, respectively. For a SISO case, Equation (13) is similar to Equation (8).

2.2. ARMAX model identification

The ARMAX model (7), first introduced by [6], is one of the most widely used identification models. Its properties make it an ideal tool for the identification of SISO systems. There are many ARMAX model identification algorithms available: extended least squares, recursive maximum likelihood, instrumental variable, prediction error method, etc. (see e.g. [7] or [8]). In this paper, we will briefly introduce the principle of the Prediction Error Method (PEM), which we will further use for our building system identification.

The Prediction Error Method, applied to the ARMAX model (7), estimates the parameters θ of the model, which comprise the parameters a, b and c of the polynomials (2), (3) and (5), by means of the model prediction error $e(k)$, as illustrated in Fig. 1. The prediction error $e(k)$ is assumed to be white noise with zero mean.

It can be shown that the predictor form of the ARMAX model (7) can be rewritten as

$$\hat{y}(k|\theta) = \frac{b(d)}{a(d)}u(k) + \left(1 - \frac{a(d)}{c(d)}\right)y(k) \tag{16}$$

The parameters of this model can be found by minimizing the criterion

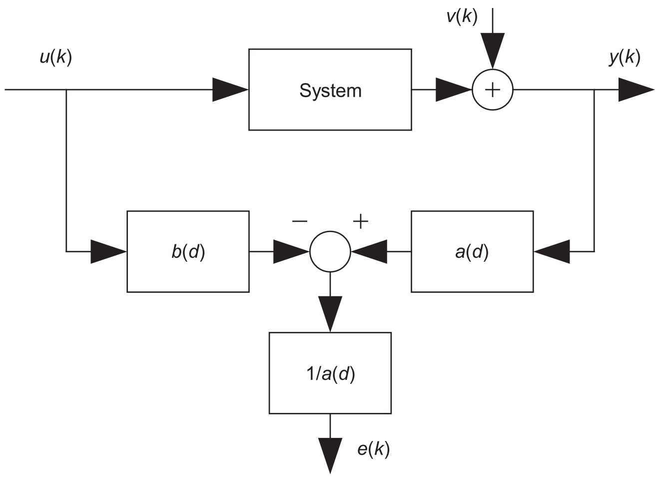


Fig. 1. Error generation of the ARMAX model.

$$V_{\mathrm{ARMAX}} = \frac{1}{N}\sum_{k=1}^{N}\left(\frac{1}{c(d)}[a(d)y(k) - b(d)u(k)]\right) \tag{17}$$

The minimum of this criterion can by found using the quasilinearization technique, as described e.g. by [7] or [8].

The advantages and disadvantages of the ARMAX model identification method are as follows:

- Advantages:
  + It can be applied to a wide spectrum of model parameterizations.
  + It gives models with excellent asymptotic properties, due to its relationship to maximum likelihood.
  + It can handle systems that operate in a closed loop (the input is partly determined as output feedback, when the data is collected) without any special techniques.
  + Recursive PEM algorithms are available for on-line identification.
- Disadvantages:
  – Need a priori information on the model structure (model type and orders of each term).
  – The search for the parameters that gives the best output prediction fit may be tedious, and involves search surfaces that have many local minima.
  – Structural inconsistency may lead to biased parameter estimates.
  – The PEM can be extended to MIMO identification; however, lack of an appropriate canonical form for the MIMO ARMAX model leads to an over-parameterized model structure. The industrial practice for MIMO identification is to separate the model into MISO (multi-input single-output) subsystems, perform MISO identification independently, and combine the results.

2.3. Subspace identification

The main contribution of subspace methods (4SID – Subspace State Space IDentification) is that the sequence of states x can be determined without knowing the system matrices A, B, C, D from the Equation (9). The advantage of this approach is that for identification of system states x, we can use robust methods of linear algebra (such as QR or SVD decomposition), and the identification of the system matrices A, B, C, D becomes a least squares problem – we do not need any iterative algorithms (no convergence problems), which makes the subspace identification methods numerically fast and robust. It is therefore suitable for large data sets and large scale systems (number of outputs m, system order n, number of inputs l are "large").

We will now briefly introduce the principles of deterministic subspace identification algorithms. Given a model (9), the mathematical problem of subspace identification is – according to [9] – as follows.

Given s consecutive input and output observations $u(0),\ldots,$ $u(s-1)$, and $y(0),\ldots,$ $y(s-1)$, find an appropriate order n and the system matrices A, B, C, D.

The subspace algorithms are basically performed in two steps:

1. Determine the model order n and estimate the state sequence $\hat{x}(i), \hat{x}(i+1), \ldots, \hat{x}(i+j)$.
2. From the estimated state sequence $\hat{x}$, input data u and output data y, determine the system matrices A, B, C, D.

First, we construct a block Hankel matrix from input and output data:

$$U(0|2i-1) \overset{\text{def}}{=} \left(\begin{array}{cccc} u(0) & u(1) & \cdots & u(j-1) \\ u(1) & u(2) & \cdots & u(j) \\ \vdots & \vdots & \ddots & \vdots \\ u(i-1) & u(i) & \cdots & u(i+j-2) \\ \hline u(i) & u(i+1) & \cdots & u(i+j-1) \\ u(i+1) & u(i+2) & \cdots & u(i+j) \\ \vdots & \vdots & \ddots & \vdots \\ u(2i-1) & u(2i) & \cdots & u(2i+j-2) \end{array}\right) = \begin{pmatrix} U(0|i-1) \\ U(i|2i-1) \end{pmatrix} = \begin{pmatrix} U_p \\ U_f \end{pmatrix} \tag{18}$$

The number of block rows i is a user-defined index that is "large enough" – it is one of the parameters for tuning the identification. The number of columns is usually set automatically by the implemented algorithm to $s-2i+1$, such that all s available data samples are used.

The output block Hankel matrices $Y(0|2i-1)$, Y_p, Y_f are defined in a similar way.

It can be shown (see [9]) that the system order n can be obtained by the equation

$$n = \operatorname{rank}\begin{pmatrix} U_p \\ Y_p \end{pmatrix} + \operatorname{rank}\begin{pmatrix} U_f \\ Y_f \end{pmatrix} - \operatorname{rank}\begin{pmatrix} U_p \\ Y_p \\ U_f \\ Y_f \end{pmatrix} \tag{19}$$

It can also be shown that the state sequence may be obtained from the intersection of past and future inputs and outputs. This can be illustrated by the following scheme:

$$\begin{pmatrix} U(0|2i-1) \\ \vdots \\ Y(0|2i-1) \end{pmatrix} \rightarrow \text{Here is the intersection} - \text{state } x$$

The explanation of why this happens is very simple – if there exists any linear relationship between the inputs and outputs (i.e. the inputs and outputs were measured on the same linear system), the intersection between them must be linearly dependent on each other. The linear dependency reduces the row rank of the Hankel matrix $\begin{pmatrix} U(0|2i-1) \\ Y(0|2i-1) \end{pmatrix}$ and the intersection (i.e. the linearly dependent rows of the Hankel matrix) forms a basis which represents the state sequence X. The intersection can be found using LQ or SVD decomposition (see [10]).

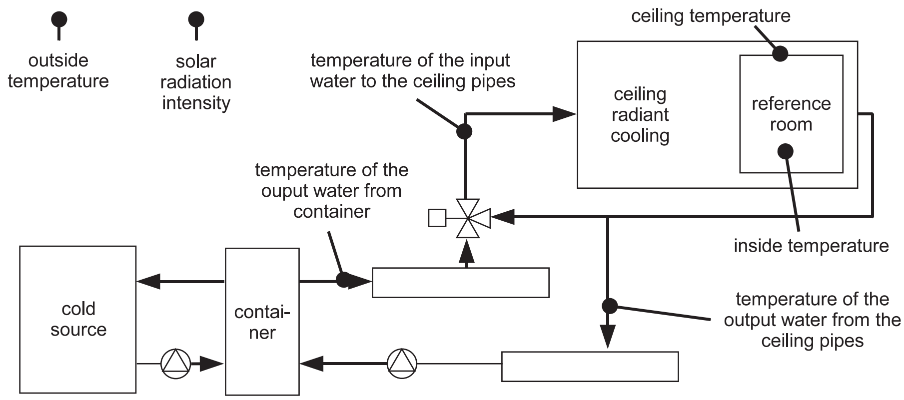


Fig. 2. Simplified diagram of the water circulation network for the ceiling radiant cooling system.

Knowing the inputs U, outputs Y and states X, it is now easy to find the state space representation (9) of the system, e.g. by least squares or total least squares methods, from the following equation:

$$\underbrace{\begin{pmatrix} X(i+1) \\ Y(i|i) \end{pmatrix}}_{\text{known}} = \underbrace{\begin{pmatrix} A & B \\ C & D \end{pmatrix}}_{\text{to be found}} \underbrace{\begin{pmatrix} X(i) \\ U(i|i) \end{pmatrix}}_{\text{known}} \tag{20}$$

where $U(i|i)$, $Y(i|i)$ are block Hankel matrices with only one block row of inputs and outputs respectively (i.e. $U(i|i) = (u(i)\ u(i+1)\ldots u(i+j-1))$, $Y(i|i)$ is constructed in a similar way).

The subspace identification method described in this paper is a *deterministic subspace identification* – we assume that the original system is ideally linear and time invariant and subject to zero noise. Various modifications of said basic method for less ideal systems are described in relevant literature, such as [10] or [9].

According to [7,9,11], there are several advantages and disadvantages of the subspace identification methods:

- Advantages:
  + **MIMO systems identification**. The complexity of the identification for large MIMO (Multiple-Inputs Multiple-Outputs) systems is the same as for SISO (Single-Input Single-Output) systems. There is no need for an extensive MIMO structure parametrization.
  + **Numerical robustness**. 4SID can be implemented with QR and SVD factorizations, which have well known properties, and very good numerical robustness.
  + **Few user parameters**. In fact, there is only one parameter, and that is the number of block rows of Hankel matrices. There is no need for complex parametrization even for MIMO systems, because 4SID methods give a state space model.
  + **Model order reduction**. The algorithms of 4SID incorporate implicit model order reduction. This is useful for real-world data, where noises and disturbances play an important role, and increase the order of the estimated model.
- Disadvantages:
  – **Need a large set of input/output data**. The statistical properties of geometrical methods used in 4SID are the reason for the fact, that they need a large amount of input/output data samples.
  – **Difficult recursification**. The basic algorithms were developed to identify the system parameters from off-line data. Extending this algorithm for on-line identification is not straightforward.
  – **Prior knowledge cannot be easily incorporated into 4SID**. These methods have a black-box approach to the identified system.

The solution to the last two disadvantages is known today, as there have recently been some papers published on the problems of recursification and incorporating prior knowledge, e.g. [12]; however, the solution is not straightforward.

One of the biggest advantages over other identification methods is the small number of user-defined parameters, which makes this method suitable even for users that need to identify some systems only occasionally. In particular, the following parameters must be entered by the user:

- Input and output data U, Y
- System order N (which may be estimated by the 4SID algorithm)
- Number of block rows of the Hankel matrices

The identification process is in fact tuned only by the number of block rows of the Hankel matrices, which depends on the characteristics of the input and output data, but can be found on a trial-and-error basis.

3. System overview

3.1. The building

The "Crittall" system, invented in 1927 by Crittall and Musgrave [13], was a favorite heating system in the Czech Republic during 60s for large buildings. The heating (or cooling) beams of this system are embedded into the concrete ceiling. Even though the principle of ceiling radiant cooling is quite old, it is still very popular and subject to intensive research (e.g. [14–16]).

The building of the Faculty of Mechanical Engineering, CTU, was built in 1961 (see Fig. 3). As the original construction included aluminium window frames and glass panel facing with poor thermal insulation characteristics, the building was recently refurbished. In addition to new insulation, air conditioning and ventilation systems were upgraded [17].

In said building, it is not possible to control room temperatures individually. The control is therefore carried out for one entire building block, i.e. the same control effort is applied to all rooms of the building block.

There are several types of blocks in the building – lecture halls, laboratories and office blocks. The laboratories do not have any cooling systems, as they are built in a shadow of other buildings, and the lecture halls have new air conditioning systems. The cooling of the office blocks, which are of our interest, is achieved by said Crittall system.

Fig. 3. The building of the Czech Technical University in Prague.

It is difficult to control the radiant cooling (and heating) system, because it has a huge accumulation mass, and hence a large thermal capacity. Moreover, the control is limited by several factors – the environment in the cooled rooms has to comply with specific temperature demands, and the surface temperature of the ceiling is limited by the dew point of the inner air, such that no condensation of water on the ceiling occurs.

A simplified scheme of the ceiling radiant cooling system is illustrated in Fig. 2. The source of cold is a compact cooling unit (chiller), which supplies the cooling water to the water container. A mixing occurs here, and the water temperature supplied to the respective cooling circuits is higher than the water temperature supplied by the chiller. An accurate temperature control of the cooling water for respective building segments is achieved by a three-port valve with a servo drive. The cooling water is then supplied to the respective ceiling beams. There is one measurement point in a reference room for every segment, wherein the ceiling temperature is also measured. The setpoint of the control valve is therefore the control variable for the ceiling radiant cooling system in each segment.

3.2. The data

As an identification object, we have chosen the A2 block of the CTU building in Prague, Czech Republic. The data is collected by the RcWare Vision[1] system and stored in an SQL database with a sampling period of 3 min. The following data is measured on the system (see Fig. 2):

- Reference room ambient temperature
- Reference room ceiling temperature
- Temperature of the input water to the ceiling pipes
- Temperature of the output water from the ceiling pipes
- Temperature of the output water from the container
- Valve position of the three-port control valve
- Outside temperature
- Solar radiation intensity

All temperatures are measured by Sensit Ni1000 industrial thermistors. The outside temperature is calculated as a mean value from 11 temperature sensors, distributed throughout the area of the CTU building. The solar radiation intensity is calibrated by precise measurement carried out by the Solar Laboratory[2] at the Faculty of Mechanical Engineering, CTU, which resides in the same building.

For the identification, data measured between June 1 and August 7, 2008, was available; there were enough days with an outside temperature exceeding 30 °C in this data set, when the ceiling radiant cooling was active. Unfortunately, the data contained a large amount of drop-outs. The frequency of sampling periods T_s is shown in Table 1 – we can see that the majority of data was sampled with a period of around 3 min, but we have 23 drop-outs that are longer than 1 h.

Table 1
Frequency of sampling periods T_s in the measured data.

T_s [min]	0–2	2–3	3–4	4–60	Over 60
Samples	116	21482	5370	151	23

From the data, we have selected some consistent data sets with the following features:

- There must not be a drop-out longer than 2 h in the data set
- The data set must contain at least 3000 samples (approximately one week)

Three data sets comply with the above criteria (6/13–6/19, 6/24–7/1 and 7/4–7/14, 2008). There was cold weather in the first data set, so the ceiling radiant cooling system was not active. So we ended up with two data sets, one for identification and the other one for verification, which is quite sufficient for our purposes.

The three-port valve is a standard device which can be controlled by a classic controller (e.g. PID controller, see [18]); it supplies any possible temperature of the cooling water with high accuracy. For ceiling radiant cooling system identification, we will use the following input and output variables:

- Input variables:
  - Temperature of the input water to the ceiling pipes
  - Outside temperatures
  - Solar radiation intensity
- Output variables:
  - Reference room ceiling temperature
  - Reference room ambient temperature

The next step was to decide which data set to use for **identification**, and which one for its **verification**. A general requirement for statistical identification says that the system must be excited in as many modes (a mode is basic dynamical behavior entity of a system, corresponding e.g. to a resonance frequency; for further explanation, see e.g. [5]) as possible [10]. To be more specific, the input vector U and output vector Y must yield a state sequence X which is of full row rank. A rule of thumb translates said

Table 2
Mean values and standard deviations of the sampling period T_d in the data sets used for identification and verification.

Data set	Mean value	Stand. dev.
Identification	3.00	0.02
Verification	3.01	0.19

All numbers are in minutes.

[1] http://www.rcware.eu.

[2] http://solab.fs.cvut.cz.

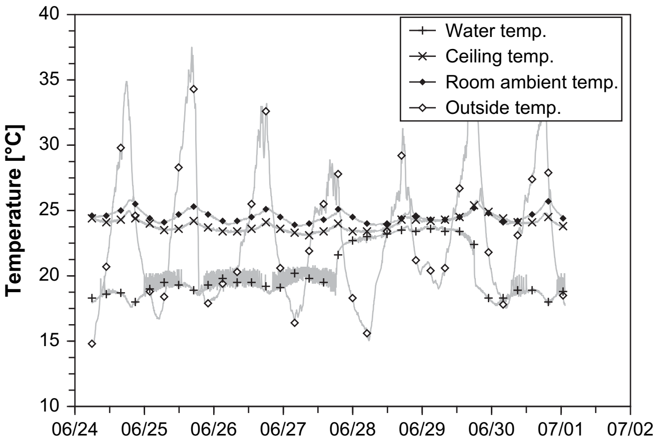


Fig. 4. Data set for the system identification.

requirements that the data must be “varied” enough, which is typically satisfied in practical experiments. From this point of view, the second data set (6/24–7/1, 2008 see Fig. 4) seems to be better, it also has a better sampling period deviation (see Table 2). The third data set will be used for verification.

4. Identification experiments

4.1. ARMAX model identification

ARMAX model identification is a very convenient method for simple systems which are subject to noise, as all the system and noise parameters can be identified by setting up the appropriate polynomial degrees in Equation (7). However, for more complex systems, finding the “appropriate” orders is tedious work, as the identification algorithms are very sensitive to any changes in the ARMAX model parameters. Another disadvantage is that the transfer function of the ARMAX model is a MISO model (multiple-input, single-output); for our system with three inputs and two outputs, we actually have to identify 2 individual ARMAX models, one for each output.

Each ARMAX model comprises 8 parameters in total (refer to Equation (7)):

- $1\times$ the degree of the denominator $a(d)$
- $3\times$ the degree of the numerator $b(d)$ (the model part of the ARMAX model)
- $1\times$ the degree of the numerator $c(d)$ (the noise filter part of the ARMAX model)
- $3\times$ the traffic delay n_k for inputs (for the PEM identification method)

For example, if we try to identify the ARMAX model by nested for cycles, assuming that the model order is less than ten, 18, 080, 425 cycles would be needed (assuming that $\deg(a(d))\geq\deg(b(d))$, $\deg(a(d))\geq\deg(c(d))$ and $\deg(a(d))\geq n_k$). As stated in Table 6, about 1 s is needed for one identification cycle – the identification in a nested for cycle would take about 209 days for one ARMAX model. This illustrates the necessity of an “engineering” approach, and experience of the ARMAX model identification process.

Table 3
ARMAX identification parameters.

Parameter	Ceiling temp.	Room amb. temp.
$\deg(a(d))$	8	8
$\deg(b(d))$	[7,7,2]	[3,4,4]
$\deg(b(d))$	5	5
n_k	[3,3,3]	[3,3,3]
Error std. deviation	0.247	0.214

In our example, the best ARMAX parameters found are listed in Table 3.

4.2. Subspace identification

Table 4
4SID identification parameters.

Parameter	Ceiling temp.	Room amb. temp.	Both
Model order N	10	4	8
Block rows D	70	40	65
Error std. dev.	0.206	0.281	0.212 (ceil.) 0.481 (room)

As already mentioned, one of the advantages of the subspace identification methods is the small number of tuning parameters. We only need to know the number of block rows D in Hankel matrices (18), and, if desired, the system order N. The actual computation is very fast, enabling an automated trial-and-error estimation process, which is implemented in many commercial implementations of the 4SID identification.

For our identification problem, we used a low-level implementation that enabled us to tune both parameters, i.e. D and N. We carried out three identification experiments:

1. All inputs, ceiling temperature as an output
2. All inputs, room ambient temperature as an output
3. All inputs, both temperatures as outputs

For an identification by nested for cycles, 2 cycles would be needed that search the parameters e.g. from 1 to 10 (for N) and from 20 to 100 (for D). The D can be changed with a step of 5, as the 4SID algorithms are not sensitive to minor parameter changes. This would result in 160 for cycles which need, according to Table 6, a period of 40 s (compared to 209 days for ARMAX model identification).

The best parameters found are listed in Table 4. It can be seen here that the cascade model consisting of the separate ceiling and room model yield better results than the full MIMO model. The reason might be that the ceiling and ambient room temperatures are subject to cross-correlated noises, so one of the stochastic 4SID method conditions is not fulfilled (the theoretical background of

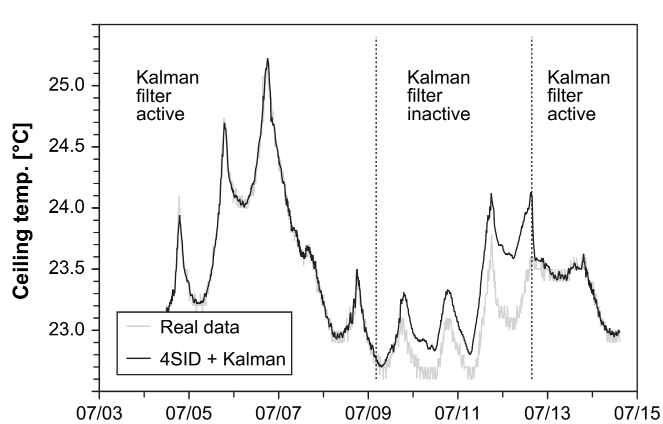


Fig. 5. The effect of simultaneous use of the model and Kalman filter, as identified by the 4SID methods.

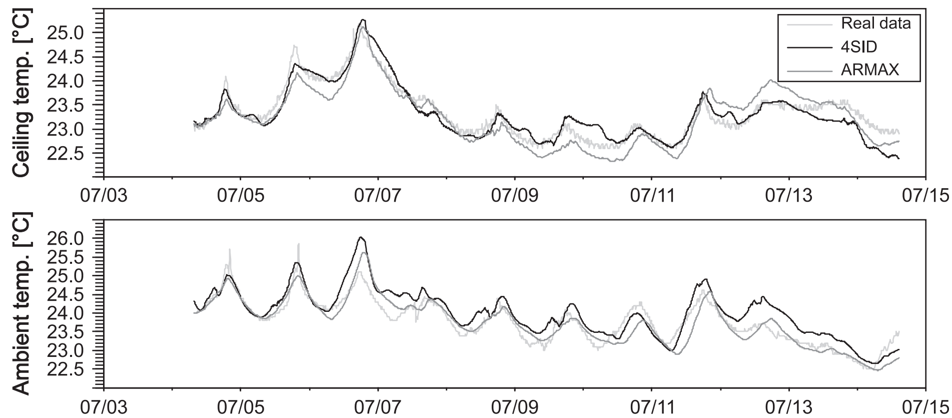


Fig. 6. Comparison of the real data, ARMAX and 4SID identification models. The 4SID & Kalman data are not shown, as their estimation error is negligible in the scale of the figure.

Table 5
Standard deviations of the models.

Output	ARMAX	4SID	4SID & Kalman
Ceiling temp.	0.247	0.206	0.052
Room ambient temp.	0.214	0.281	0.063

Table 6
Time requirements for ARMAX and 4SID methods. Tested on Intel Mobile Core2 Duo T5600, 1.83 GHz, 2048 kB of L2-cache, 1024 MB RAM.

Method	Tuning time (approx.)	Computational time	For cycle ident.
ARMAX	10 h	955 ms	209 days
4SID	0.5 h	234 ms	40 s

stochastic identifications is not in the scope of this article, more information can be found e.g. in [10]). However, if the system is decomposed into two subsystem, we get one output noise only, and the cross-correlation is no longer a problem.

5. Improvement of subspace identification results by Kalman filtering

A significant improvement of the behaviour of a mathematical model can be achieved by employing a Kalman filter (first published by [19]). The Kalman filter is a very important case of an optimal state observer. The theory behind the Kalman filter is not in the scope of this paper, an excellent source on this topic is [20]. The basic idea of the Kalman filter is that the model output $\widehat{u}(k)$ is compared with the real system output $u(k)$ and the model error (also called prediction error) is supplied to a gain matrix K (the Kalman gain matrix) and further to the model state $\widehat{x}(k)$. The state is thus updated by the real measurement, and the model keeps track of the behavior of the system.

The Kalman filter is calculated in two steps – "predict" and "correct". Referring to the state space model (9), the model state is predicted in the "predict" phase:

$$\widehat{x}^{-}(k) = A\widehat{x}(k-1) + Bu(k-1) \tag{21}$$

The system state is then updated from measurement in the "correct" phase:

$$\widehat{x}(k) = \widehat{x}^{-}(k) + K\left(u(k) - \widehat{u}(k)\right) \tag{22}$$

The Kalman gain K can be continuously updated; however, for industrial systems, wherein the noise characteristics are assumed to be time invariant, it is advantageous to use a constant Kalman gain which does not change over time.

The Kalman filter is tuned by the noise covariance matrices Q, R, S (Equation (10)). The measurement of the noise parameters is quite difficult, and for complex systems, such as HVAC systems, they only provide a rough estimate. It is therefore necessary to tune the Q, R and S matrices manually, which may be tedious work.

One of the advantages of the subspace identification methods, as described by [21], is that there are algorithms that can actually estimate the Q, R, S matrices from the measurement history data, and calculate the time-invariant Kalman gain K.

An example of the use of the Kalman filter is illustrated in Fig. 5. The Kalman filter keeps tracking the real system with almost zero *estimation* error. If a *prediction* of the real system behavior is needed, the Kalman filter is disabled from the model, and the output prediction (here the ceiling temperature, based on weather prediction and cooling water temperature schedule) is carried out. Once enabled, the Kalman filter can correct the system state within the discrete time (i.e. number of steps) equal to the model order.

6. Comparison of results

The results of the ARMAX and 4SID identification can be seen in Fig. 6, the standard deviation of the models are listed in Table 5 (along with the combination of 4SID & Kalman filter). First of all, we have to say that the results are very good for both identification methods – the standard deviations are less than 0.3 °C. The ceiling temperature is subject to lower noise – the concrete structure acts as a very effective filter – and the 4SID method yields better results here. On the other hand, the room ambient temperature is subject to very high noise and disturbances (doors and windows open and closed, people moving around,...) and the ARMAX method can make full use of its noise filter terms (the $c(d)$ polynomial in Equation (7)).

Another aspect is the complexity of the identification process with respect to time required to tune and to actually calculate the models, shown in Table 6. Here the 4SID approach is clearly more convenient than the ARMAX identification. As already mentioned, the reason is that for 4SID, we only need to find 2 parameters, whereas the ARMAX identification needs 8 parameters in total, which is time consuming. Moreover, the ARMAX model is much more sensitive to the respective parameter changes than the 4SID methods. The "tuning time" in Table 6 represents the subjective time needed for the model tuning (excluding the work on importing the measurement data) experienced by the authors.

7. Conclusions

The two statistical approaches to model identifications presented in this paper – ARMAX model identification and subspace identification – present two main streams in model identification. The example of the university building shows the advantages and disadvantages of said models – while the subspace identification is faster, easier to implement, and more accurate for systems with white noise input/output signals, an ARMAX model can still achieve better results for systems, wherein the parameters of the respective noise signals are far from ideal.

Acknowledgment

The project of building identification has been supported from the state budget of the Czech Republic, through the Ministry of industry and commerce, in the scope of grant No. 2A-1TP1/084, *"Integration of building systems, research and application of intelligent algorithms with influence on energy consumption of buildings and living houses"*.

References

[1] Gwerder M, Lehmann B, Tödtli J, Dorer V, Renggli F. Control of thermally-activated building systems (tabs). Applied Energy 2008;85(7):565–81.
[2] Olesen BW, Sommer K, Dutchting B. Control of slab heating and cooling systems studied by dynamic computer simulations. ASHRAE Transactions 2002;108(2):698–707.
[3] Tian Z, Love JA. Energy performance optimization of radiant slab cooling using building simulation and field measurements. Energy and Buildings 2009;41(3):320–30.
[4] Underwood CP. HVAC control systems: modelling, analysis and design. London, Great Britain: E & FN Spon; 1999.
[5] Antsaklis PJ, Michel AN. Linear systems. 2nd ed. New York, USA.: Birkhäuser Boston; 2006.
[6] Åström KJ, Bohlin T. Numerical identification of linear dynamic systems from normal operating records. In: Proceedings of the 2nd IFAC symposium on the theory of self-adaptive control systems, Teddington, Great Britain; 1965.
[7] Zhu Y. Multivariable system identification for process control. In: Verhaegen Michel, editor. Subspace model identification of MIMO processes. New York, USA: Elsevier Science Inc.; 2001. p. 199–216.
[8] Ljung L, editor. System identification: theory for the user. 2nd ed. Upper Saddle River, NJ, USA: Prentice Hall; 1999.
[9] De Cock K, De Moor B. Subspace identification methods, Control systems, robotics and automation. In: EOLSS, UNESCO encyclopedia of life support systems, vol. 1. Oxford, Great Britain: Eolss Publishers Co., Ltd; 2003. p. 933–79.
[10] van Overschee P, De Moor BLR. Subspace identification for linear systems: theory, implementation, applications. Dordrecht, The Netherlands: Kluwer Academic Publishers Group; 1996.
[11] Trnka P. Subspace identification methods, Ph.D. thesis, Czech Technical University in Prague, Prague, Czech Republic; 2007.
[12] Trnka P, Havlena V. Subspace identification method incorporating prior information. In: Proceedings of the 46th conference on decision and control, New Orleans, USA; 2007. p. 4968–73.
[13] R.G. Crittall, J.L. Musgrave, Heating and cooling of buildings, GB Patent No. 210880 (April 1927).
[14] Catalina T, Virgone J, Kuznik F. Evaluation of thermal comfort using combined CFD and experimentation study in a test room equipped with a cooling ceiling. Building and Environment 2009;44(8):1740–50.
[15] Paulauskaite S, Valancius K, Parfentieva N. Results of the investigation of microclimate created by radiant heating/cooling system. In: Proc. of 7th international conference on environmental engineering, vols. 1–3, Vilnius, Lithuania; 2008. p. 859–63.
[16] Lihua Z, Li L, Yi C. Simulation and analysis of the cooling ability of ceiling radiant cooling system. In: Proceedings of the 5th international symposium on heating, ventilating and air conditioning, vols. I and II, Beijing, Peoples R. China; 2007. p. 124–9.
[17] Lain M, Drkal F, Hensen J, Zmrhal V. Low energy cooling techniques for retrofitted office buildings in central Europe. In: Proc. of AIVC, Prague; 2004. p. 79–84.
[18] Åström KJ, Hägglund T. Advanced PID control. Research Triangle Park, NC, USA: ISA – The Instrumentation, Systems, and Automation Society; 2005.
[19] Kalman RE. A new approach to linear filtering and prediction problems. Transactions of the ASME – Journal of Basic Engineering 1960;82(Series D):35–45.
[20] Simon D. Optimal state estimation: Kalman, H_∞, and nonlinear approaches. Hoboken, NJ, USA: Wiley & Sons, Inc; 2006.
[21] van Overschee P, De Moor BLR. Subspace algorithms for the stochastic identification problem. Automatica 1993;29(3):649–60.

Appendix E

Váňa, Kubeček, Ferkl: Notes on Finding Black-Box Model of a Large Building

VÁŇA, Z. – KUBEČEK, J. – FERKL, L. Notes on Finding Black-Box Model of a Large Building. In Proceedings of the 2010 IEEE International Conference on Control Applications, pp. 1017–1022, Yokohama, Japan, 2010. ISBN 978-1-4244-5363-4.

Number of citations: 0

Notes on Finding Black-Box Model of a Large Building

Zdeněk Váňa, Jakub Kubeček, Lukáš Ferkl

***Abstract*— Finding a suitable dynamic, linear, time-invariant model is a major obstacle for the use of model-predictive control of buildings. While finding models based on physical properties of the system is time consuming, statistical models need the system to be excited, which is not always possible. This paper presents possibilities for finding a suitable model based on subspace identification methods for unexcited data and data containing a specially designed identification experiment, and presents practical experiences gained while finding suitable models for a large building. Finally, some other possibilities of finding a model by incorporating prior information to the identification process are discussed, and the performance of the models is evaluated with respect to their use as a part of an MPC controller.**

I. Introduction

Recent trends in carbon dioxide emission reduction establish a socially challenging environment for new ideas for energy consumption optimization, regardless of our individual opinion about the global warming issue. Special attention is paid to various technology advances, such as low-emission engines, green energy generation, etc. Both economical and environmental revenue of such innovations is sometimes subject to fierce discussions. On the other hand, improvements of control system algorithms leading to economical savings represent a solution, which is "green", beyond dispute.

Buildings have been subject to energy savings for a long time. Indeed, according to the U. S. Energy Information Administration, in 2005, buildings accounted for 39 % of total energy usage, 12 % of the total water consumption, 68 % of total electricity consumption, and 38 % of the carbon dioxide emissions in the U. S. A. [1]. However, recent efforts have focused mainly on the construction of the buildings, such as better insulations, double facades, heat-reflecting glass, etc. But large buildings, such as schools, hospitals, or office buildings, are usually equipped by control systems based on general industrial systems and use industrial protocols, PLCs, dense sensor networks, etc., which enable the use of sophisticated modifications of state-of-the-art controllers (PID control, weather-compensated control, ...).

The use of modern control strategies has been very limited, with the exception of fuzzy control. More analytic methods, such as LQG or MPC control, have been enduring their use in buildings mainly because of the absence of convenient tools for obtaining a model of the building. Basically, there are three approaches to find a dynamic, linear, time-invariant model of a system:

- white-box
- grey-box
- black-box

For buildings, finding a white-box model, i.e. a model based on physical properties of the building, is feasible, but time consuming and it is definitely not useful for commercial application. Statistical identification procedures, often called black-box models, represented by e.g. ARMAX models or subspace identification, need "sufficient" excitation of all modes of the system. This is seldom possible for real systems, as most identification experiments are costly and bring thermal uncomfort to the persons inside the building. A compromise between those two approaches is to incorporate some prior information, such as static gain or predefined structure, into the model, which is further identified by means of statistical identification. This approach is well known for single-input, single-output (SISO) systems, but only few attempts have been made for the case of multiple-inputs, multiple-outputs (MIMO) systems. A procedure is known that introduces prior information to models identified by subspace identification, but works only on MISO systems.

In this paper, we will present practical experience gained while finding a model for a large building – the building of Faculty of Electrical Engineering and Faculty of Mechanical Engineering, Czech Technical University in Prague (see Figure I).

II. Model identification

One of the crucial contributors to the quality of the control is a well identified model which will be later on used for control in MPC algorithm. There are several completely different approaches to system identification including physical modeling, CFD modeling or statistical identification. As traditional methods are rather time consuming for buildings,

Fig. 1. The building of the Czech Technical University in Prague that was used for model identification experiments and Model Predictive Control.

Z. Váňa, J. Kubeček, L. Ferkl are with the Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic {zdenek.vana}{jakub.kubecek}{lukas.ferkl} @fel.cvut.cz

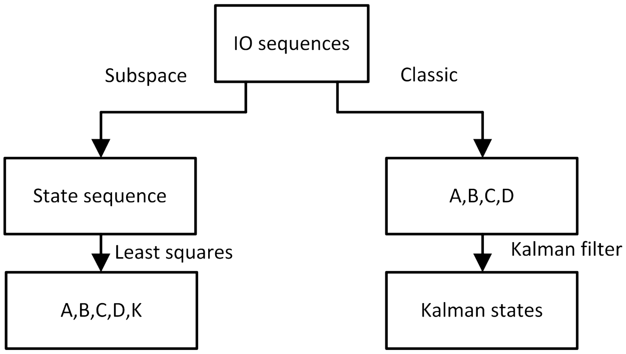


Fig. 2. Comparison of classical and subspace identification methods

we have turned towards statistical identification methods, and, more specifically, towards subspace methods [2], [3], [4].

The objective of the subspace algorithm is to find a linear, time invariant, discrete time model in an innovative form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k), \end{aligned} \quad (1)$$

where A, B, C, and D are system matrices, K is Kalman gain – derived from the Algebraic Riccati Equation (ARE) [5], and e is a white noise sequence. This model is equivalent to the well-known stochastic model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k), \end{aligned} \quad (2)$$

with

$$\begin{aligned} \operatorname{cov}(w, v) &= E\left(\begin{bmatrix} w(p) \\ v(p) \end{bmatrix} \begin{bmatrix} w^T(q) & v^T(q) \end{bmatrix} \right) = \\ &= \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \geq 0, \end{aligned} \quad (3)$$

wherein matrices Q, S and R are covariance matrices of process and measurement noise sequences, respectively. Loosely speaking, the objective of the algorithm is to determine the system order n and to find the matrices A, B, C, D and K.

The main difference between classical and subspace identification is, given the sequence of input data $u(k)$ and output data $y(k)$, as follows:

- **Classical approach.** Find system matrices, then estimate the system states, which often leads to high order models that have to be reduced thereafter.
- **Subspace approach.** Use orthogonal and oblique projections to find Kalman state sequence (see [5]), then obtain the system matrices using least squares method.

The differences in the approaches can be seen in Fig. 2[6]. The entry point to the algorithm are input-output equations as follows:

$$\begin{aligned} Y_p &= \Gamma_i X_p^d + H_i^d U_p + Y_p^s \\ Y_f &= \Gamma_i X_f^d + H_i^d U_f + Y_f^s \\ X_f^d &= A^i X_p^d + \Delta_i^d U_p, \end{aligned} \quad (4)$$

where Y_p and Y_f are the Hankel matrices of past and future outputs, U_p and U_f are the Hankel matrices of past and future inputs, X_p^d and X_f^d are the deterministic Kalman state sequences, Y_p^s and Y_f^s are the stochastic Hankel matrices of past and future outputs, H_i^d is the lower block triangular Toeplitz matrix for the deterministic subsystem (which contains all matrices A, B, C and D), Γ_i is the extended system observability matrix (which contains the system matrices A and C) and Δ_i^d is the deterministic reversed extended controllability matrix (which contains system matrices A and B). More details about constructing said matrices can be found in [2], [6], [7]. Using combined algorithm described in [7], we get

$$\begin{bmatrix} \widetilde{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \widetilde{X}_i \\ U_{i|i} \end{bmatrix} + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}, \quad (5)$$

with $\widetilde{X}_i = \widetilde{X}_{i_{\widehat{X}_0, P_0}}$, where $\widehat{X}_0$ is oblique projection defined by [7] as:

$$\widehat{X}_0 = X_p^d /_{U_f} U_p, \quad (6)$$

where P_0 is state covariance matrix, and we can determine noise sequence covariance matrices Q, S and R from the residuals, which are defined by Eq. (3). Solving Eq. (5) using least squares methods, we get the state space system description of the system, namely the system in the innovation form (Eq. (1)) with matrices A, B, C, D and K.

III. Description of the Building

The building of the Czech Technical University in Prague uses a Crittall [8] type ceiling radiant heating and cooling system. The Crittall system, invented in 1927 by R. G. Crittall and J. L. Musgrave, was a favorite heating system in the Czech Republic during 1960s for large buildings. In this system, the heating (or cooling) beams are embedded into the concrete ceiling. The control of individual rooms is very complicated due to the technical state of the control elements in all rooms. The control is therefore carried out for one entire building block, i.e. the same control effort is applied to all rooms of the building block. There are three building blocks with the same construction and orientation. Therefore, this situation is ideal for comparison of different control strategies, as depicted in Fig. **??**.

A simplified scheme of the ceiling radiant heating system is illustrated in Fig. 4. The source of heat is a vapor-liquid heat exchanger, which supplies the heating water to the water container. A mixing occurs here, and the water is supplied to the respective heating circuits. An accurate temperature control of the heating water for respective circuits is achieved by a three-port valve with a servo drive. The heating water is then supplied to the respective ceiling beams. There is one measurement point in a reference room for every circuit. The setpoint of the control valve is therefore the control variable for the ceiling radiant heating system in each circuit.

IV. Description of the model

The ceiling radiant heating system was simplified into a linear, time invariant mathematical model. Outside temperature prediction and heating water temperature were used

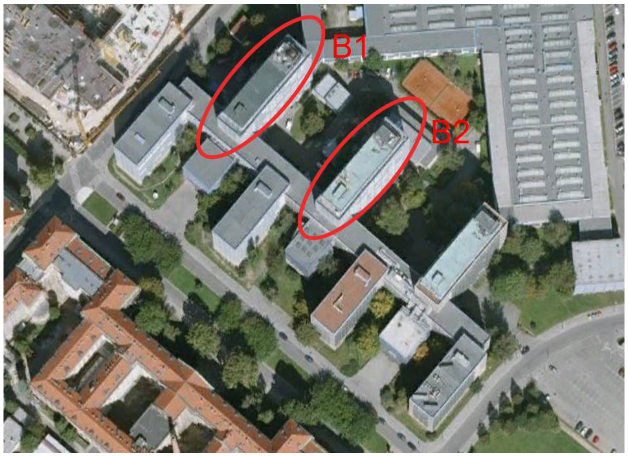


Fig. 3. The Czech Technical University building with the building blocks B1 and B2. Because of their very similar characteristics, they were chosen as ideal objects for comparison of the behavior of the Model-Predictive Controller, wherein the models described in this paper were used.

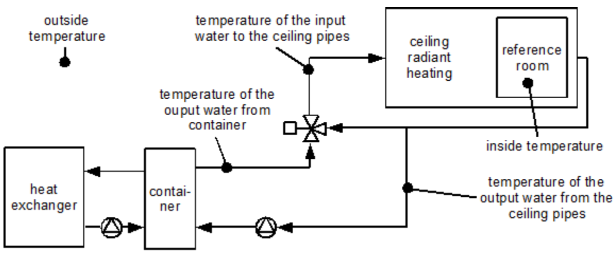


Fig. 4. Simplified scheme of the ceiling radiant heating system.

as the model inputs. Prediction of the outside temperature is composed of two values, T_{max} and T_{min}, defining a confidence interval. The only output of the model was the inside temperature. This can be formalized according to Eq. (2) as

$$x(k+1) = Ax + B \begin{bmatrix} T_{min} \\ T_{max} \\ T_{hw} \end{bmatrix}$$

$$T_{in} = Cx + D \begin{bmatrix} T_{min} \\ T_{max} \\ T_{hw} \end{bmatrix}, \qquad (7)$$

where T_{hw} is the temperature of the heating water and T_{in} denotes the inner temperature. The state x has no physical interpretation, when identified by means of the subspace identification. System order is determined by the identification algorithm.

V. Winter 2008/09: The First Attempt

For the implementation of the identification algorithms, Scilab has been chosen, which is a free software for numerical calculations created by the French scientific institutions INRIA and ENPC. Its license allows for free use, but does not meet the conditions of the Open Source Initiative or Free Software Foundation. Therefore, it can be used in commercial applications free of charge. Another option was to use the more popular Matlab, but it was rejected after a discussion with the industrial partner of the project, because it is expensive, slower than Scilab and Subspace

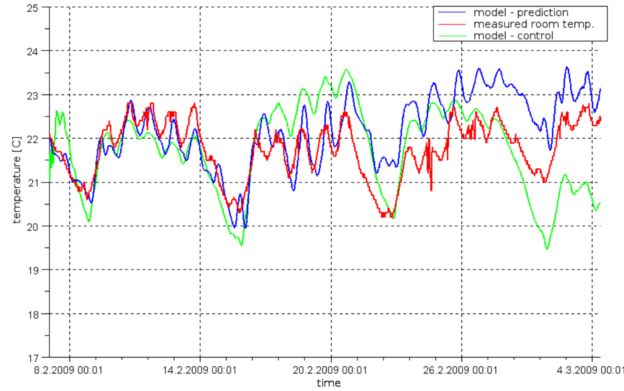


Fig. 5. Subspace model verified by real data, the first attempts. The green line represents the reference room temperature as predicted by the model later used in the MPC controller.

Identification fails to provide good support and variability to computations (in other words, you do not need much knowledge to use Matlab, but if you have such a knowledge, Scilab allows you to play around with the identification procedure). In the beginning, the Scilab's standard System Identification toolbox was used.

After the data from the building became available, sequences suitable for identification had to be found. Unfortunately, the data showed frequent minor failures and incomplete records. It was therefore necessary to select the data sections which did not contain too long drop-outs. The data, originally sampled in the period of 3 minutes, were resampled to 10 minutes, which is enough for the control and no important information is lost, given the estimate of the dynamics of the building.

As a first step, a rough selection of about 20 models obtained by subspace identification with the smallest relative error was made. At this moment, a rather surprising discovery was made; the models identified by the identification data set showed pretty performance on the verification data set, however, it turned out that the inputs and outputs had been switched. In fact, a known phenomenon was proven that the subspace identification represent a good approximator, regardless on the value of the estimated data. It was not surprising that the plots obtained by correct inputs and outputs did not look as pretty as the wrong ones.

Later on, it turned out that model selection according to the conformity of its response to the real data is not applicable, because majority of the models that have been selected in such way were not suitable in conjunction with the MPC. It was therefore necessary to choose the model according its behavior directly in conjunction with the MPC, which represented a process with considerable time consumption.

From many models, which had to be managed individually, a model with subspace parameters $N = 6$, $S = 25$ was chosen. Response of this model can be seen in Figure 5.

MPC controller, which was described e.g. by [9], was tuned and tested on data obtained from February 22nd to March 2nd, 2009. The schedule of the MPC implementation, which required sufficient computer tests before its use on the

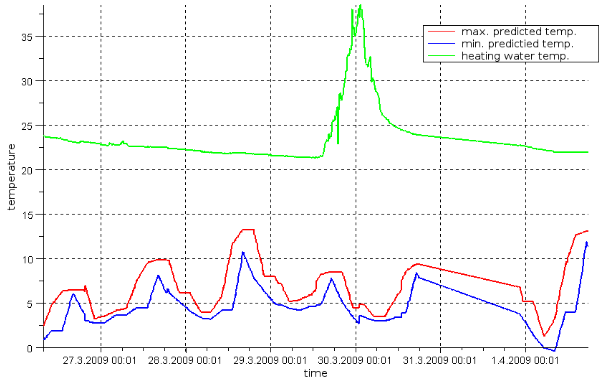


Fig. 6. The inputs to the model used for the MPC.

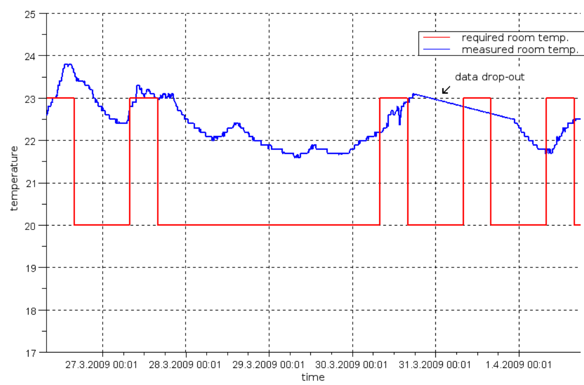


Fig. 7. The outputs from the model used for the MPC.

real building, was as follows:

1) The model was tested by computer simulations only.
2) The model was implemented into the control system, in a stand-by mode, wherein the conformity of the model and the real system was being observed.
3) The MPC controller was implemented and calculated values of the regulator were checked. The control was still done by the old system.
4) After acceptable results were obtained from the MPC controller, it was fully involved into the control.

It turned out that the model identified by the subspace identification, used in the MPC controller, provides acceptable performance, and was set operational. The behavior of this control can be seen in Fig. 6 (the inputs of the model) and Fig. 7 (the outputs).

VI. Winter 2009/10: The First Success

The first experiments with the MPC control were successful and the comparison of the MPC and the original control showed that the savings achieved by the MPC reach 10 %, which finally lead to an extensive collaboration between our team and the company responsible for the heating system of the Czech Technical University. However, it was clear that the results of the first MPC were not as brilliant as originally expected. Detailed analysis of the model showed that it is only suitable for outside temperatures well above 0 °C, which was the case of March 2009, and high savings achieved by the MPC could be a mere coincidence.

The rather misleading behavior of the first model resulting from the detailed analysis was caused by not quite correct use of subspace identification methods and by practical limitations that do not satisfy the theoretical assumptions. These assumptions are:

- I/O data length goes to infinity (which is obviously not true)
- White noises are considered in subspace algorithms (which cannot be verified)
- The system must be “sufficiently” excited by I/O data (which will be discussed further)

There could be some other problems with the I/O data as well, such as their mutual dependency (which can be detected in the covariance matrix). The effect of such data is that outputs are influenced by some inputs only and some inputs have no influence to any output. Obviously, this is undesirable. According to the theory, there is one way to improve the model – to excite the system sufficiently. Moreover, it has to be done in such a way that main inputs excite the system itself; then the dynamics reflecting these inputs appear.

As already mentioned, it is not very practical to perform identification experiments on real buildings. It was very fortunate that the building for this experiment was a university building, which is almost empty during any vacations. The identification experiment had been proposed and was realized during the Christmas vacation period (Fig. 8). New I/O data were obtained and new models were identified. After the first tests, the problems discussed above seemed to disappear, but the deviations in the process were increased. It could mean the system is not time invariant and its identification is more complicated. We can discuss two extreme ways of identification in this case:

- *Short-time data identification:* Short-time data means that we can consider the data were measured on a linear system. The identified model is accurate on said data, but in the real process, the deviations are increased significantly.

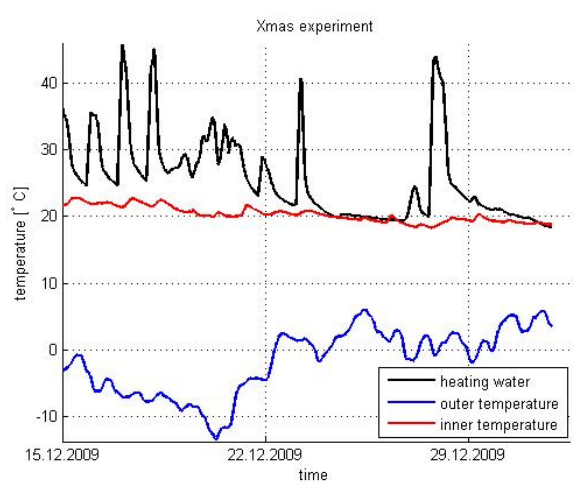


Fig. 8. The data acquired from the Christmas identification experiment.

- *Long-time data identification:* This is the opposite case. The identified model does not match the data accurately, but the deviations are practically uniform – they depend on the length of the data.

Both ways are extreme; in practice, we have to use an engineering sense for a proper choice, which lies somewhere in between. Unfortunately, no choice leads to the best model, because each model has its pros and cons.

Now let us have a look at some methods which can be useful for proper control:

- *Model updating:* Regularly identified model (i.e. each month), always from last data of length N (because of computational rate).
- *Recursive subspace identification:* Up-dating model matrices by force of new incoming I/O data. Unlike the previous method, the entire identification process is not started, only model matrices are up-dated via special algorithm.
- *Incorporated prior information:* With any basic knowledge of system this information could be incorporate into the identification algorithm. A prior information can be expressed in several ways, i.e. the covariance matrices, the static gain, the ratio of gains of outputs, the order of the system, the main dynamics or the structure of the system. Special kind of a prior information is usage of artificial data including a prior information.

The results from the Christmas identification experiment can be seen in Fig. 9. Even though the results are not as pretty as of the model from the previous season, the models are based more solid assumptions and achieve better performance with the MPC controller. Grey-box model based on analogy with RC-circuits was finally used in the MPC controller, as it had slightly better performance and was less demanding on the level of excitation of the system. However, the subspace identification algorithm has been modified and subspace model is in operation in the time of submission of this paper. The results from the MPC control can be seen in Fig. 10 – the savings achieved by the MPC control were about 20 % compared to the original weather-compensated control.

VII. Conclusions

In the example of the application of subspace identification methods, it can be seen that underestimation of theoretical assumptions can result in fairly misleading results. We can also conclude that subspace identification can be suitable for models included in MPC control of (at least some) buildings.

VIII. Acknowledgment

This work has been supported from the state budget of the Czech Republic, through the Ministry of industry and commerce, in the scope of grant No. FR-TI1/517, *"Control Systems for Optimization of Energy Consumption in Low-Energy and Passive Houses"*.

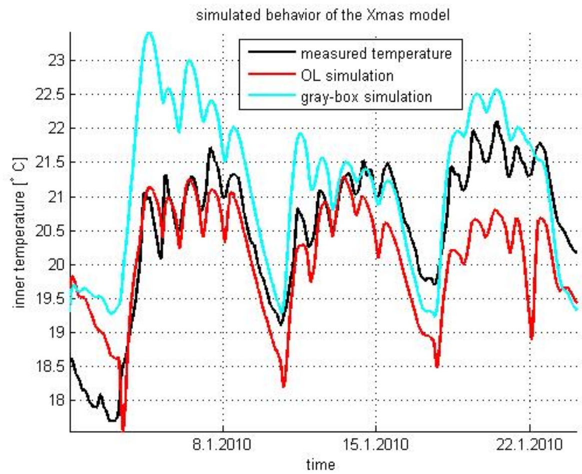


Fig. 9. Models from the Christmas identification experiment – open loop (OL) simulation of the subspace model and the alternative gray-box model, which achieved better performance with the MPC.

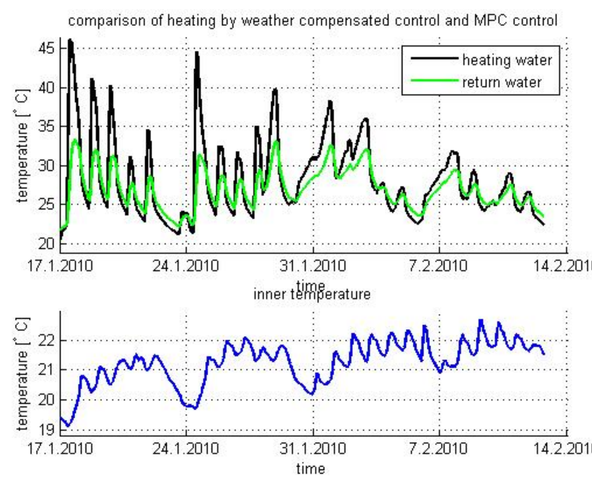


Fig. 10. Results from the MPC control. The original control was based on weather-compensated control, MPC was set operational on January 30th and achieves about 20 % savings compared to the original control.

References

[1] *Buildings and their Impact on the Environment: A Statistical Summary*, 2009, rev. April 22, 2009, available on-line at http://www.epa.gov/greenbuilding/pubs/gbstats.pdf.

[2] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, 1999.

[3] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2007.

[4] L. Ferkl and J. Široký, "Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods," *BUILDING AND ENVIRONMENT*, vol. 45, no. 1, Sp. Iss. SI, pp. 205–212, 2010.

[5] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. John Wiley & Sons, 1972.

[6] K. De Cock and B. De Moor, "Subspace identification methods," in *EOLSS, UNESCO Encyclopedia of Life Support Systems*. Oxford, UK: Eolss Publishers Co., Ltd., 2003, vol. 1, ch. Control systems, robotics and automation, pp. 933–979.

[7] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 1999.

[8] R. G. Crittall and J. L. Musgrave, "Heating and cooling of buildings," GB Patent No. 210880, April 1927.

[9] L. Ferkl and J. Široký, "Model of a hot water full circulation mode in a building," in *Proceedings of the 7th IEEE Conference on Control and Automation*, Christchurch, 2009.

Appendix F

Prívara, Cigler, Váňa, Ferkl, Šebek: Subspace Identification of Poorly Excited Industrial Systems

PRÍVARA, S. – CIGLER, J. – VÁŇA, Z. – FERKL, L. – ŠEBEK, M. Subspace Identification of Poorly Excited Industrial Systems. In Proceedings of the 49th IEEE Conference on Decision and Control, pp. 4405–4410, Atlanta, USA, 2010.

Number of citations: 0

Subspace Identification of Poorly Excited Industrial Systems

Samuel Prívara, Jiří Cigler, Zdeněk Váňa, Lukáš Ferkl and Michael Šebek

***Abstract*— Most of the industrial applications are multiple-input multiple-output (MIMO) systems, that can be be identified using knowledge of the system's physics or from measured data employing statistical methods. Currently, there is the only class of statistical identification methods capable of handling the issue of vast MIMO systems – subspace identification methods. These methods, however, as all statistical methods, need data of certain quality, i.e. excitation of corresponding order, no data corruption, etc. Nevertheless, the combination of statistical methods and physical knowledge of the system could significantly improve system identification. This paper presents a new algorithm which provides remedy to insufficient data quality of certain kind through incorporating of prior information, e.g. known static gain or input-output feedthrough. The presented algorithm naturally extends classical subspace identification algorithms, that is, it adds extra equations into the computation of system matrices. The performance of the algorithm is shown on a case study and compared to current methods, where the model is used for an MPC control of a large building heating system.**

***Index Terms*— Subspace methods; Identification for control**

I. Motivation

With 39 %, buildings contribute significantly to total energy usage in 2005, as stated by the U. S. Energy Information Administration [1]. This poses strong motivation for creation of advanced and energy saving HVAC (Heating, Ventilating, and Air Conditioning) systems [2]. Significant amount of energy can be saved using predictive control strategies (see project OptiControl[1]) compared to the conventional strategies. Widely used control strategy, weather-compensated control, can lead to poor energy management or reduced thermal comfort even if properly set up, because it utilizes current outside temperatures only. In case of sharp change of weather, there is an improper control action due to the energy accumulation in large buildings, resulting in over- or underheating of the building. Even though HVAC control systems have been improved significantly during recent years, predictive controller described in [3] introduces a different approach to the heating system control design. There is, however, a crucial condition for the successful control, that is, properly identified model of the system. Model identification can be performed by variety of methods, physical modeling or statistical approach among others.

This paper presents incorporation of prior information into subspace identification methods. These methods originally emerged as a conjunction of linear algebra, geometry and system theory and compared to the classical identification methods [4], they provide user with several advantages such as numerical robustness, natural extension to MIMO systems, etc. There are, however, also some drawbacks, e.g. lack of satisfactory number of data samples, proper order of excitation or strong noise contamination can lead to poor identification results [4], [5]. Some problems coupled to these methods such as identification of stable, positive real models, etc. using regularization can be found in [6], [7] or formulated as constrained optimization in [8]. Black-box identification, such as subspace identification methods, rely only on experimental data, that is, they may result in biased models [9], or fail in giving a proper model (this problem is addressed in [10], [11]).

Prior information can significantly improve identification results, however, current algorithms are not able to provide satisfactory results for MIMO systems. Previous works such as [12] count with single-input single-output (SISO) system only. Even [9] using Bayesian framework approach did not present method which would treat MIMO system in a satisfactory manner. This paper, in contrary, presents a new algorithm of incorporation prior information, which is built-in directly into system matrices B and D and does not make use of the covariance matrix. Proposed algorithm enables treating MIMO systems in a natural way using state-space approach.

The paper is organized as follows: Section II provides an insight into the building-up of the matrices used in subspace algorithms and formulates the general identification algorithm. Section III describes incorporation of prior information (PI) in subspace identification framework and shows two special cases of PI, knowledge of static gain and input-output feedthrough. Section IV presents identification results of previously described algorithms. The objective of the identification was creation of proper model (in sense of fit and controllability) of a real, eight-floor building. Future development is outlined in Section V and the paper is concluded with Section VI.

II. Subspace Identification

A. Problem Statement

The objective of the subspace algorithm is to find a linear, time invariant, discrete time model in an innovative form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k), \end{aligned} \qquad (1)$$

based on given measurements of the input $u(k) \in \mathbb{R}^m$ and the output $y(k) \in \mathbb{R}^l$ generated by an unknown stochastic

S. Prívara, J. Cigler, Z. Váňa, L. Ferkl and M. Šebek are with Department of Control Engineering, Faculty of Electrical Engineering of Czech Technical University in Prague, Karlovo náměstí 13, Czech Republic `samuel.privara@fel.cvut.cz`

[1]http://www.opticontrol.ethz.ch/

system of order n, which is equivalent to the well-known stochastic model as defined in e.g. [13], [14]. Loosely speaking, the objective of the algorithm is to determine the system order n and to find the matrices A, B, C, D and K.

B. Matrices Used in Subspace Algorithm

Notation and building-up of the matrices as follows further on were adopted as in [15]. Upper index d and s denotes deterministic and stochastic subsystems, respectively.

1) Data Matrices: Input block Hankel matrix is built-up from input data as follows:

$$U_{0|2i-1} = \left(\begin{array}{ccccc} u_0 & u_1 & u_2 & \cdots & u_{j-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{i-1} & u_i & u_{i+1} & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & u_{i+2} & \cdots & u_{i+j-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{2i-1} & u_{2i} & u_{2i+1} & \cdots & u_{2i+j-2} \end{array}\right). \quad (2)$$

Input and output Hankel matrices can be grouped as follows:

$$W_p = \left(\frac{U_p}{Y_p}\right) \quad , \quad W_p^+ = \left(\frac{U_p^+}{Y_p^+}\right), \quad (3)$$

where $U_{0|i-1} = U_p$ and $U_{i|2i-1} = U_f$ with U_p and U_f denoting the past and future inputs, respectively. The same logic holds for outputs $y(k)$ and noise $e(k)$. Change of indices results in $U_{0|i} = U_p^+$ and $U_{i+1|2i-1} = U_f^-$, respectively.

2) System Related Matrices: Extended ($i > n$) observability (Γ_i) and reversed extended controllability (Δ_i) matrices for deterministic and stochastic subsystems, respectively are defined as follows:

$$\Gamma_i = \begin{pmatrix} C^T & (CA)^T & \dots & (CA^{i-1})^T \end{pmatrix}^T \quad (4)$$
$$\Delta_i^d = \begin{pmatrix} A^{i-1}B & A^{i-2}B & \dots & AB & B \end{pmatrix} \quad (5)$$
$$\Delta_i^s = \begin{pmatrix} A^{i-1}K & A^{i-2}K & \dots & AK & K \end{pmatrix} \quad (6)$$

The lower block triangular Toeplitz matrix for deterministic and stochastic subsystem, respectively are defined as

$$H_i^d = \begin{pmatrix} D & 0 & \dots & 0 \\ CB & D & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{i-2}B & CA^{i-3}B & \dots & D \end{pmatrix}, \quad (7)$$

$$H_i^s = \begin{pmatrix} I & 0 & \dots & 0 \\ CK & I & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ CA^{i-2}K & CA^{i-3}K & \dots & I \end{pmatrix}, \quad (8)$$

and Kalman state sequence as a sequence generated by a bank of non-steady state Kalman filters [16], working in parallel on each other of the columns of the matrix W_p.

C. General Algorithm

The entry point to the algorithm are input-output equations as follows:

$$\begin{aligned} Y_p &= \Gamma_i X_p + H_i^d U_p + H_i^s E_p \\ Y_f &= \Gamma_i X_f + H_i^d U_f + H_i^s E_f \\ X_f &= A^i X_p + \Delta_i^d U_p + \Delta_i^s E_p. \end{aligned} \quad (9)$$

Oblique projection as described in [17], [15] is the main tool used in subspace methods. It is defined as follows:

$$\mathcal{O}_i = Y_f \mathop{/}_{U_f} W_p, \quad (10)$$

or, equivalently,

$$\mathcal{O}_i = Y_f \begin{pmatrix} W_p^T & U_f^T \end{pmatrix} \begin{pmatrix} W_p W_p^T & W_p U_f^T \\ U_f W_p^T & U_f U_f^T \end{pmatrix}^\dagger \begin{pmatrix} I_{l\times l} \\ 0 \end{pmatrix} W_p, \quad (11)$$

where l is a number of outputs and $(\bullet)^\dagger$ is Moore-Penrose pseudoinverse. It holds, that $\mathcal{O}_i = \Gamma_i \widetilde{X}_i$ [15], where $\widetilde{X}_i$ is Kalman filter state sequence. The order of the system can be determined from analysis of singular values obtained using singular value decomposition (SVD) of $W_1 \mathcal{O}_i W_2$, where W_i are weighting matrices of appropriate size and determine resulting state space basis as well as importance of particular element of $\mathcal{O}_i$. This decomposition also yields extended observability matrix Γ_i and Kalman filter states $\widetilde{X}_i$.

Algorithm continues from either Γ_i or $\widetilde{X}_i$ in a slightly different manner depending on particular subspace identification algorithm, however, both ways lead to a computation of system matrices A and C using least squares method.

Computation of system matrices B and D is the next step, such that matrices A and C acquired in previous step. Different approaches for matrices determination are addressed in detail in [15]. This step is crucial for the incorporation of the prior information and will be discussed in detail in the following section.

The algorithm concludes with computation of Kalman gain matrix K. The essential condition for optimal filter running is the knowledge of the noise covariance matrices Q and R (state and measurement noise covariance matrices). These two matrices are used for calculation of Kalman gain K. In early 70s' Mehra's publications on covariance matrices estimation were published [18]. Then, for a large period, the estimation of covariance matrices was largely overlooked and only in 2006 Odelson's article [19] was published that offered a new method for Q and R estimation called ALS [19]. A few more modifications of this method can be found in [20], [21]. Kalman gain matrix K is the computed in a standard way using the state and noise covariance matrices computed using ALS as described in [19].

III. Incorporation of Prior Information

Prior information (PI) is a good tool for improvement of identification results. Its incorporation can be considered as a bridge between classical identification approaches based on time response of unknown system on e.g. step or impulse response, and statistical based identification methods [4].

System properties such as steady state gain, settling time, asymptotic stability, dominant time constants, smoothness of step response etc. can be used in classical approach to determine the unknown system. The question is, how to involve at least some of these properties into statistical based identification, and in particular, into 4SID methods.

Several methods dealing with above problem have been proposed. They can be generally classified into four groups.

1) Bayesian framework: This method can be characterized as a natural way for incorporation of PI because it allows inference of prior estimate of unknowns system parameters with information retrieved from measured data. Resulting posterior conditional probability function can be obtained using Bayesian rule $p(\theta|y) \propto l(\theta|y)p(\theta)$ [22], where $p(\theta)$ is prior probability density function of parameters and $l(\theta|y)$ the likelihood function for measured data.

Although many satisfactory results were proposed for incorporation of PI into ARX or ARMAX model identification [22], similar strategies do not work well for the class of 4SID methods. This problem is treated in [9], but favorable results are given only for multiple input single output (MISO) systems, because presented algorithm is based on structured weighted lower rank approximation (SWLRA)[23] which provides optimal solution only for MISO systems. However, at least suboptimal algorithm is presented in [9], but the level of optimality is not guaranteed.

2) Direct incorporation of system properties into 4SID algorithms: The following section tries to sketch out identification algorithm in simplified way. The incorporation of all conceivable kinds of PI is shown.

- Computation of extended observability matrix and state vector sequence $W_1 \mathcal{O}_i W_2 = \Gamma_i X_i$. Different 4SID algorithms make use of different rules for computation of these matrices (see e.g. [15]).
- Computation of system matrices A and C based on Γ_i using least squares method.
- Determination of matrices B and D and possible incorporation of prior information in this step. Solution will be addressed in Section III-A.
- Kalman gain computation.

3) Artificial data: Generation of data with desired properties is yet another approach how to deal with the weak point of 4SID, its black-box character (and associated statistical problems). Such data can contain trends that represent system in a decoupled form (connection of particular input to particular output etc.). As the ratio between artificial and measured data is unknown, the only way how to address this problem is trial and error method.

4) Frequency domain identification methods: Yet another approach for system identification is use of frequency domain methods. It was shown that this approach leads to maximum likelihood formulation of the frequency domain estimation problem [24]. Even though there were some proposals (e.g. [25]) how to incorporate prior information into identification algorithm, it is still an open problem and a topic of ongoing research.

In the following incorporation of prior information will be addressed:

A. *Knowledge of Static Gain*

Subspace identification process consists of several parts. Each of them corresponds to a particular property of resulting system. Matrix A contains dynamics of states, while matrix C transfers dynamics to the outputs. Therefore, the system input/output structure is influenced mainly by determination of matrices B and D, with A and C fixed. Hence, the key idea is to involve prior information about steady state gain into latter matrices.

Let matrices A and C have already been computed by some 4SID algorithm (e.g. [15]). Knowledge of these matrices can be exploited to compute such matrices B and D, that lead to desired steady state behavior. This is possible thanks to the fact, that the sum of elements of impulse response is equal to the steady state:

$$D + CB + CAB + CA^2B + \ldots = G, \tag{12}$$

$$\begin{pmatrix} I_{l\times l} & \sum_{k=0}^{\infty} CA^k \end{pmatrix} \begin{pmatrix} D \\ B \end{pmatrix} = G, \tag{13}$$

where G is a matrix of steady state gains (g_{ij} is a steady state gain from the j-th input to i-th output)

$$G = \begin{pmatrix} g_{11} & g_{12} & \cdots & g_{1m} \\ \vdots & \vdots & \ddots & \vdots \\ g_{l1} & g_{l2} & \cdots & g_{lm} \end{pmatrix}. \tag{14}$$

In case of asymptotically stable matrix A, the following holds (Neumann series convergency theorem [26]):

$$(I_{n\times n} - A)^{-1} = \sum_{k=0}^{\infty} A^k. \tag{15}$$

Finally, we get resulting formula, which represents the additional set of constraints that have to be fulfilled:

$$\underbrace{\begin{pmatrix} I_{l\times l} & C(I_{n\times n} - A)^{-1} \end{pmatrix}}_{\Gamma_s} \begin{pmatrix} D \\ B \end{pmatrix} = G, \tag{16}$$

Consider any 4SID algorithm that computes matrices B and D after A and C being already computed. The computation is performed using least squares method as follows:

$$B, D = \arg\min_{B,D} \left\{ \left|\left| M - L \begin{pmatrix} D \\ B \end{pmatrix} \right|\right|_F \right\}, \tag{17}$$

where $||\bullet||_F$ denotes Frobenius norm, and M and L are appropriate size matrices defined in [15]. It must be said in this place, that these matrices are defined differently for each 4SID algorithm.

Incorporating constraints (16) can be done in two possible ways:

- Solve least squares problem with equality constraints

$$B, D = \arg\min_{B,D} \left\{ \left|\left| M - L \begin{pmatrix} D \\ B \end{pmatrix} \right|\right|_F : \Gamma_s \begin{pmatrix} D \\ B \end{pmatrix} = G \right\}. \tag{18}$$

- Solve weighted least squares problem

$$B, D = \arg\min_{B,D} \left\{ \left|\left| \begin{pmatrix} M \\ G \end{pmatrix} - \begin{pmatrix} L \\ \Gamma_s \end{pmatrix} \begin{pmatrix} D \\ B \end{pmatrix} \right|\right|_{F,W} \right\}, \tag{19}$$

where W is user-defined weighting matrix that guarantees the desired steady state behavior. In case that only a submatrix G_{sub} of the gain matrix G is known, the constraints (16) can be modified by two square diagonal matrices S_r, S_c of appropriate size as

$$S_r \Gamma_s \begin{pmatrix} D \\ B \end{pmatrix} S_c = G_{sub}. \tag{20}$$

Matrices S_r, S_c are "selectors" of relevant rows (S_r) and columns (S_c), and contain only ones and zeros for retaining and disposal of the known gain, respectively.

Computation of matrices B and D in some 4SID algorithms is based on vectorization and Kronecker product, that is:

$$B, D = \arg\min_{B,D} \left\{ \left|\left| \text{vec}\, \overline{M} - \overline{L}\ \text{vec} \begin{pmatrix} D \\ B \end{pmatrix} \right|\right|_F \right\}. \tag{21}$$

Using vectorization and Kronecker product, the set of equality constraints (16) can be expressed in following manner:

$$(I_{m \times m} \otimes \Gamma_s)\text{vec} \begin{pmatrix} D \\ B \end{pmatrix} = \text{vec}\, G\,, \tag{22}$$

which can be readily included in either (18) or (19). Each input-output channel gain is in one row and therefore, in case of partial knowledge of G, it is easy to omit respective rows for unknown gains.

B. Knowledge of input-output feedthrough

Oftentimes in industrial applications, the input-output feedthrough of the system to be identified is known in advance. In fact, it is not a rare phenomenon, that there is no input-output feedthrough present in the system, that is, the system matrix D is equal to zero. This will be treated in the following:

Consider again (17), the computation of matrices B and D, that is, the very last step of subspace identification algorithm as proposed in [15]. Matrix D can be forced to be zero by a computation of (17) or (21) using modified matrix L by the elimination of the columns corresponding to matrix D. The set of omitted columns differs for two distinct algorithms:

1) Without Kronecker product: Solution to this problem is given by omitting first l columns of matrix L in (17), where l corresponds to the number of outputs.

2) With Kronecker product: This situation is more complicated than in the first case due to vectorization and Kronecker product, nevertheless the selection of columns of $\overline{L}$ is determined by indices $k = 1, 2, \dots m$ in set I, given as:

$$I = \{k(l+n) - n + 1, k(l+n) - n + 2, \dots k(l+n)\}. \tag{23}$$

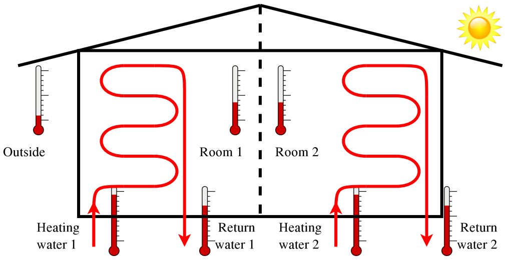


Fig. 1. Simplified scheme of model identification setup.

IV. Identification Results

Proposed algorithms were implemented and then applied to data gathered from HVAC system of the building of the Czech Technical University in Prague. The simplified scheme of one building block consisting of three inputs (outside temperature, heating water 1, heating water 2) and four outputs (room temperature 1, return water 1, room temperature 2, return water 2) is depicted in Fig. 1.

Data from such an industrial environment do not always have sufficient quality, they suffer from strong noise contamination, occurrence of outliers, low excitation, etc. In our case, there is a strong multi-collinearity present in the data, that is, the conventional control strategies, which have been used for maintenance of desired temperature levels, drive both courses (north and south course, as well) of heating water, so that return waters and room temperatures had similar behavior and were strongly correlated. Black-box identification approach was not able to carry out this problem. Prior information about system structure i.e. steady state gain or/and no presence of input-output feedthrough had to be incorporated to get desired results. This can be seen in Fig. 2, where the step responses of models identified by different 4SID approaches are shown. Prior knowledge about steady state gain was in this case selected as follows:

$$G = \begin{pmatrix} 0.5 & 0.75 & 0.15 \\ 0 & 0.9 & 0 \\ 0.5 & 0.15 & 0.75 \\ 0 & 0 & 0.9 \end{pmatrix}.$$

These 4SID methods come, in general, from robust combined deterministic and stochastic algorithm as introduced in [15]. Moreover two methods from [9] are mentioned for comparison: *i*) 4SID – version without changes, *ii*) 4SID+PI – Steady state gain was included using (22). Matrix D is not set to zero, *iii*) 4SID-D – Matrix D is set to zero but steady state gain is not included, *iv*) 4SID-D+PI – Both types of PI information, i.e. zero D and steady state gain are incorporated, *v*) Kung and SWLRA – PI incorporated as in [9], Kung's respectively SWLRA realization algorithms are used to get system matrices. The models retrieved from the proposed algorithm were verified against validation data by open-loop simulation, see Fig. 3. SWLRA and Kung's algorithms produce poor results therefore the open-loop responses are not mentioned.

Both figures prove the superiority of the identification

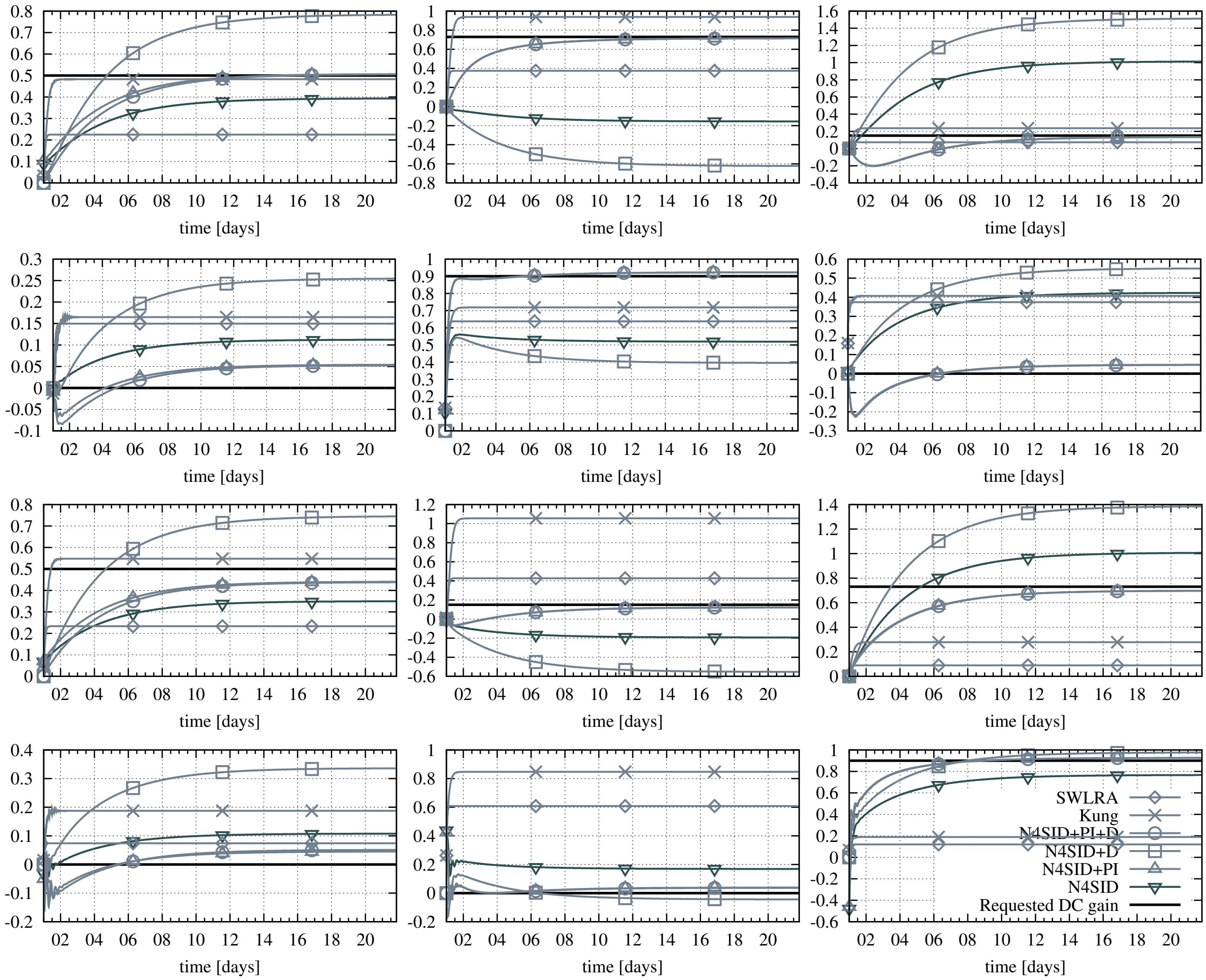


Fig. 2. Comparison of step responses of systems identified using different algorithms. There is significant improvement in identification results using prior information about steady state gain.

algorithm with PI included. Identification results can be summed-up as follows:

- *Zero D matrix.* There is almost no difference in results between robust combined algorithm (full matrix D) and algorithm with zero D matrix. This is useful especially in cases, when nonzero matrix D has no physical meaning in many industrial applications.
- *Prior information in matrices B and D.* The incorporation of known static gain into identification algorithm has different consequences for deterministic and stochastic (in sense of system with noise) algorithms. In case of deterministic algorithm, the prior information is able to substitute the lack of information caused by noise (no presence of Kalman filter) and significantly improves identification results. In many cases, it is even not possible to identify system with noise using deterministic algorithm without knowledge of prior information due to the insufficient information and noise contamination and this can be rectified using prior information. In case of stochastic algorithm, the differences in fit between algorithm with and without prior information is not major, however, the incorporation of prior information enables creation of the model which has properties equivalent to real physical system and is valid for control.
- *Sensitivity of true value of PI.* The price for the better identification performance in case of PI incorporation must be paid by greater sensitivity to the changes in parameters, that is, even the slight change in parameters aggravates identification results (in sense of fit). The importance of prior information in respective parameters can be decided by weighting matrix in (19).

V. Future Development

As mentioned in Section III and shown in Section IV there is no SWLRA algorithm for MIMO systems working properly. This is still a topic of ongoing research. In case of successful solving of this problem, the prior information could be incorporated by means of Bayesian network as proposed by [9] even to MIMO systems. Yet another approach was presented in this article via direct incorporation PI into system matrices B and D. There is, however, PI of certain

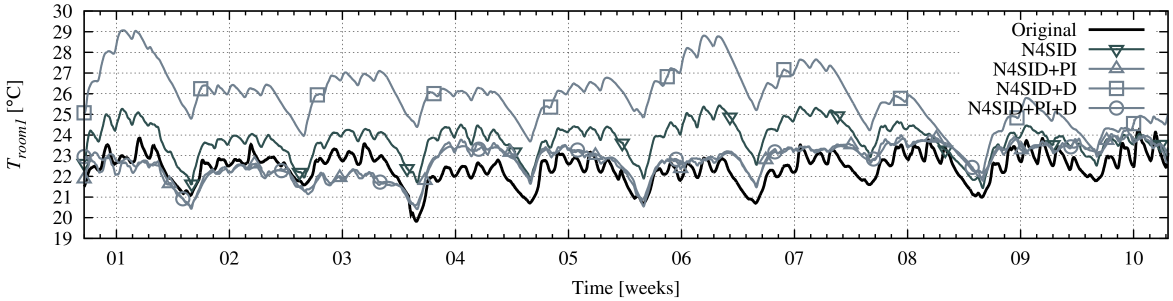


Fig. 3. Comparison of different identification strategies: open-loop simulation.

type (e.g. dynamics) which could be incorporated directly into matrices A or C, however, this approach is still unknown and topic of possible research as well.

VI. Conclusions

The proposed algorithm presents incorporation of PI into the subspace identification methods. The incorporation is performed directly into system matrices B and D, thus enables certain type of prior information, e.g. static gain. The incorporated PI is able to significantly improve identification results and substitute the lack of information in input-output data. Moreover, it notably improves model for control purposes by approaching to physical system structure. However, the quality of identification is sensitive to the accuracy of prior estimate of parameters. The constructed model has been used for temperature control in real operation of the 8-floor building of the Czech Technical University in Prague. The predictive control with model identified using algorithm proposed in this paper proved to save 23% of energy required by the weather-compensated controller.

Acknowledgment

This work has been supported from the state budget of the Czech Republic, through *i)* the Ministry of Industry and Commerce in the scope of grant No. 2A-1TP1/084, *"Integration of building systems, research and application of intelligent algorithms with influence on energy consumption of buildings and living houses" ii)* the Ministry of Education in scope of project No. 1M0567; Grant Agency of Czech Republic in the scope of grant No. 102/08/0186 and Czech Technical University through grant No. SGS10/283/OHK3/3T/13.

References

[1] *Buildings and their Impact on the Environment: A Statistical Summary*, 2009, rev. April 22, 2009, available on-line at http://www.epa.gov/greenbuilding/pubs/gbstats.pdf.

[2] C. P. Underwood, *HVAC Control Systems: Modelling, Analysis and Design*. E & FN Spon, 1999.

[3] J. Cigler and S. Prívara, "Subspace identification of large systems and model predictive control for buildings," in *The 11th International Conference on Control, Automation, Robotics and Vision*. Singapore: Nanyang Technological University, 2010.

[4] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA, 1999.

[5] J. C. Willems, P. Rapisarda, I. Markovsky, and B. L. D. Moor, "A note on persistency of excitation," *Systems & Control Letters*, vol. 54, no. 4, pp. 325 – 329, 2005.

[6] T. Van Gestel, J. Suykens, P. Van Dooren, and B. De Moor, "Identification of Positive Real Models in Subspace Identification with Regularization," 2000.

[7] I. Goethals, T. Van Gestel, J. Suykens, P. Van Dooren, and B. De Moor, "Identifying positive real models in subspace identification by using regularization," in *Proceedings of the 13th System Identification Symposium, Rotterdam, The Netherlands*. Citeseer, 2003, pp. 1411–1416.

[8] S. Lacy and D. Bernstein, "Subspace identification with guaranteed stability using constrained optimization," *IEEE Transactions on automatic control*, vol. 48, no. 7, pp. 1259–1263, 2003.

[9] P. Trnka and V. Havlena, "Subspace like identification incorporating prior information," *AUTOMATICA*, vol. 45, no. 4, pp. 1086–1091, APR 2009.

[10] M. Gevers, B. Euler, and A. Lemaıtre, "Identification for control: Achievements and open problems," *Dynamics and Control of Process Systems 2004*, p. 401, 2005.

[11] C. Rojas, M. Barenthin, J. Welsh, and H. Hjalmarsson, "The cost of complexity in identification of FIR systems," in *16th IFAC World Congress. Seoul, South Korea*, 2008.

[12] E. Bai and S. Sastry, "Parameter identification using prior information," *International Journal of Control*, vol. 44, no. 2, pp. 455–473, 1986.

[13] F. Lewis, *Optimal estimation with an introduction to stochastic control theory*. Wiley New York et al., 1986.

[14] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[15] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Pard, Nowell, Massachusetts: Kluwer Academic Publishers, 1999.

[16] K. De Cock and B. De Moor, "Subspace identification methods," in *EOLSS, UNESCO Encyclopedia of Life Support Systems*. Oxford, UK: Eolss Publishers Co., Ltd., 2003, vol. 1, ch. Control systems, robotics and automation, pp. 933–979.

[17] P. Trnka, "Subspace identification methods," Ph.D. dissertation, Czech Technical University in Prague, 2007.

[18] R. Mehra, "On the identification of variances and adaptive kalman filtering," *IEEE Transactions on Automatic Control*, vol. AC-15, pp. 175–184, 1970.

[19] B. J. Odelson, M. R. Rajamani, and J. B. Rawlings, "A new autocovariance least squares method for estimating noise covariances," *Automatica*, vol. 42, no. 2, pp. 303–308, 2005.

[20] M. R. Rajamani and J. B. Rawlings, "Estimation of the disturbance structure from data using semidefinite programming and optimal weighting," *Automatica*, vol. 45, no. 1, pp. 142–148, 2009.

[21] B. M. Akesson, J. B. Jorgensen, N. K. Poulsen, and S. B. Jorgensen, "A generalized autocovariance least squares method for kalman filter tuning," *Journal of Process control*, vol. 18, pp. 769–779, 2008.

[22] V. Peterka, "Predictor-based self-tunning algorithm," *Automatica*, vol. 20, pp. 39–50, 1984.

[23] M. Schuermans, P. Lemmerling, and H. S.V, "Block-row hankel weighted low rank approximation," *Numerical Linear Algebra with Applications*, vol. 13, pp. 293 – 302, 2006.

[24] T. McKelvey, "Frequency domain identification methods," *Circuits, Systems, and Signal Processing*, vol. 21, 2002.

[25] J. Gillberg and L. Ljung, "Frequency domain identification of continuous-time output error models, Part I: Uniformly sampled data and frequency function approximation," *Automatica*, 2009.

[26] G. W. Stewart, *Matrix Algorithms: Volume 1, Basic Decompositions*. Society for Industrial Mathematics, 1998.

Appendix G

Ferkl, Široký, Prívara: Model Predictive Control of Buildings: The Efficient Way of Heating

FERKL, L. – ŠIROKÝ, J. – PRÍVARA, S. Model Predictive Control of Buildings: The Efficient Way of Heating. In Proceedings of the 2010 IEEE International Conference on Control Applications, pp. 1922–1926, Yokohama, Japan, 2010. ISBN 978-1-4244-5363-4.

Invited lecture

Number of citations: 0

Model Predictive Control of Buildings: The Efficient Way of Heating

Lukáš Ferkl, Jan Široký, Samuel Prívara

Abstract— **The implicit model predictive control based on models identified by subspace identification methods was implemented and tested on a large university building. The control was improved by incorporating the weather prediction into the model. The performance of said controller was estimated in an experiment, wherein two almost identical building blocks were compared – one controlled by the model predictive control, and the other one by the existing weather-compensated heating controller. The model predictive control achieved energy consumption lower by approximately 10 %. Based on the positive results, an implementation was developed, which is suitable for commercial applications.**

I. Introduction

According to the U. S. Energy Information Administration, in 2005, buildings accounted for 39 % of total energy usage, 12 % of the total water consumption, 68 % of total electricity consumption, and 38 % of the carbon dioxide emissions in the U. S. A. [1]. Although the energy efficiency of systems and components for heating, ventilating, and air conditioning (HVAC) has improved considerably over recent years, there is still potential for substantial improvements. This article deals with an advanced control technique, that can provide significant energy savings in comparison with conventional, non-predictive techniques.

Widely used control strategy of water heating systems is the weather-compensated control. This feedforward control can lead to poor energy management or reduced thermal comfort even if properly set up, because it utilizes current outside temperatures only. Weather conditions, however, can change dramatically in few hours; and due to the heat accumulation in large buildings, it can lead to underheating or overheating of the building easily.

During recent years, significant advances have been done for the HVAC control systems. For instance, continuous adaptation of control parameters, optimal start-stop algorithms, or inclusion of free heat gains in the control algorithm are particular improvements of the building heating system. The model predictive controller presented in this article introduces a different approach to the heating system control design. As the outside temperature is one of the most influential quantity for the building heating system, weather forecast is employed in the predictive controller. It enables to predict inside temperature trends according to the selected control strategy. The aims of the control can be expressed in natural form as thermal comfort and economy trade off.

S. Prívara, L. Ferkl are with the Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Czech Republic {samuel.privara}{lukas.ferkl}@fel.cvut.cz

Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia in Pilsen, Czech Republic jan.siroky@rcware.eu

II. Model predictive control

Model (Based) Predictive Control (MPC) is a method of advanced control originated in late seventies and early eighties in the process industries (oil refineries, chemical plants, ...) ([2], [3], [4], [5]). The MPC is not a single strategy, but a vast class of control methods with the model of the process explicitly expressed trying to obtain control signal by minimizing objective function subject to (in general) some constraints.

The minimization is performed in an iterative manner on some finite optimization horizon to acquire N step ahead-prediction of a control signal that leads to the minimum value of the criterion, subject to all constraints. This, however, carries lots of drawbacks such as no feedback, no robustness, no stability guarantee, etc. Many of these drawbacks can be overcome by applying so-called receding horizon, i.e. at each iteration only the first step of the control strategy is implemented and the control signal is calculated again, thus, in fact, the prediction horizon keeps being shifted forward. Stability of the constrained receding horizon has been discussed in [6], [7], or yet another approach using robust control design approach in [8].

There were several attempts made to utilize predictive control concept in HVAC in the last decade [9], [10], [11]. Complex view into area of optimal building control gives the project OptiControl[1]. Besides its own results, it also provides a wide range of references to the related articles. Another project worth to mention is the Predictive Networked Building Control that deals with predictive control of the thermal energy storage on the campus of the UC-Berkeley[2]. Most of the articles devoted to the HVAC predictive control conclude results just by numerical simulations. On the contrary, this article describes MPC being tested on the ***real eight-floor building*** (see Fig. II).

A. Model identification

One of the crucial contributors to the quality of the control is a well identified model which will be later on used for control in MPC algorithm. There are several completely different approaches to system identification including physical modeling, CFD modeling or statistical identification. As traditional methods are rather time consuming for buildings,

[1] http://www.opticontrol.ethz.ch

[2] http://sites.google.com/site/mpclaboratory/research/predictive-networked-building-control-1

Fig. 1. The building of the Czech Technical University in Prague that was used for MPC application

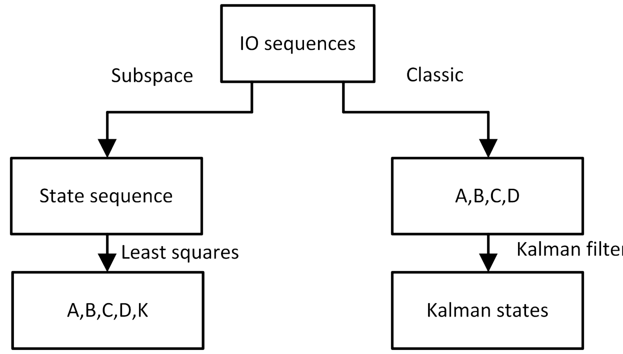


Fig. 2. Comparison of classical and subspace identification methods

we have turned towards statistical identification methods, and, more specifically, towards subspace methods [12], [13], [14].

The objective of the subspace algorithm is to find a linear, time invariant, discrete time model in an innovative form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k), \end{aligned} \tag{1}$$

where A, B, C, and D are system matrices, K is Kalman gain – derived from the Algebraic Riccati Equation (ARE) [15], and e is a white noise sequence. This model is equivalent to the well-known stochastic model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k), \end{aligned} \tag{2}$$

with

$$\begin{aligned} \mathrm{cov}(w,v) &= E\left(\begin{bmatrix} w(p) \\ v(p) \end{bmatrix} \begin{bmatrix} w^T(q) & v^T(q) \end{bmatrix}\right) = \\ &= \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \geq 0, \end{aligned} \tag{3}$$

wherein matrices Q, S and R are covariance matrices of process and measurement noise sequences, respectively. Loosely speaking, the objective of the algorithm is to determine the system order n and to find the matrices A, B, C, D and K.

The main difference between classical and subspace identification is, given the sequence of input data $u(k)$ and output data $y(k)$, as follows:

- **Classical approach.** Find system matrices, then estimate the system states, which often leads to high order models that have to be reduced thereafter.
- **Subspace approach.** Use orthogonal and oblique projections to find Kalman state sequence (see [15]), then obtain the system matrices using least squares method.

The differences in the approaches can be seen in Fig. 2[16]. The entry point to the algorithm are input-output equations as follows:

$$\begin{aligned} Y_p &= \Gamma_i X_p^d + H_i^d U_p + Y_p^s \\ Y_f &= \Gamma_i X_f^d + H_i^d U_f + Y_f^s \\ X_f^d &= A^i X_p^d + \Delta_i^d U_p, \end{aligned} \tag{4}$$

where Y_p and Y_f are the Hankel matrices of past and future outputs, U_p and U_f are the Hankel matrices of past and future inputs, X_p^d and X_f^d are the deterministic Kalman state sequences, Y_p^s and Y_f^s are the stochastic Hankel matrices of past and future outputs, H_i^d is the lower block triangular Toeplitz matrix for the deterministic subsystem (which contains all matrices A, B, C and D), Γ_i is the extended system observability matrix (which contains the system matrices A and C) and Δ_i^d is the deterministic reversed extended controllability matrix (which contains system matrices A and B). More details about constructing said matrices can be found in [12], [16], [17]. Using combined algorithm described in [17], we get

$$\begin{bmatrix} \widetilde{X}_{i+1} \\ Y_{i|i} \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \widetilde{X}_i \\ U_{i|i} \end{bmatrix} + \begin{bmatrix} \rho_w \\ \rho_v \end{bmatrix}, \tag{5}$$

with $\widetilde{X}_i = \widetilde{X}_{i_{\widehat{X}_0, P_0}}$, where $\widehat{X}_0$ is oblique projection defined by [17] as:

$$\widehat{X}_0 = X_p^d /_{U_f} U_p, \tag{6}$$

where P_0 is state covariance matrix, and we can determine noise sequence covariance matrices Q, S and R from the residuals, which are defined by Eq. (3). Solving Eq. (5) using least squares methods, we get the state space system description of the system, namely the system in the innovation form (Eq. (1)) with matrices A, B, C, D and K.

B. Predictive controller

MPC strategy. The MPC strategy comprises two basic steps:

- The future outputs are predicted in an open-loop manner using the model provided information about past inputs, outputs and future signals, which are about to be calculated. The future control signals are calculated by optimizing the objective function, i.e. chosen criterion, which is usually in the form of quadratic function. The criterion constituents can be as follows:
  - errors between the predicted signal and the reference trajectory $r(k)$
  - control effort
  - rate of change in control signals

- The first component of the control sequence $u(k)$ is sent to the system, whilst the rest of the sequence is disposed. At the next time instant, new output $y(k+1)$ is measured and the control sequence is recalculated, first component $u(k+1)$ is applied to the system and the rest is disposed. This principle is repeated continuously (receding horizon).

The reference trajectory $r(k)$, temperature in the room in our case, is known in advance as a schedule. The major advantage of MPC is the ability of computing the outputs $y(k)$ and corresponding input signals $u(k)$ in advance, that is, it is possible to avoid sudden changes in control signal and avoid the undesired effect of delay in system response.

MPC problem formulation. For given linear, time invariant, discrete model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{7}$$

find the optimal control sequence on the horizon of prediction (length T) by minimizing the objective function

$$J = \sum_{k=0}^{T-1} q(t)\left(y(k) - r\right)^2 + r(k)u(k)^2, \tag{8}$$

subject to

$$\begin{array}{rcl} u_{min} \leq & u & \leq u_{max} \\ (y-r)_{min} \leq & (y-r) & \leq (y-r)_{max} \\ \Delta_{min}\xi \leq & \Delta\xi & \leq \Delta_{max}\xi \end{array} \tag{9}$$

where constraints u_{min}, $(y-r)_{min}$, $\Delta\xi_{min}$ and u_{max}, $(y-r)_{max}$, $\Delta\xi_{max}$ are minimum and maximum values of the control signal, error of the output signal from reference and a rate of change of control signal or error of the output signal from reference, respectively. J is the value of the objective function, r is the reference trajectory, y and u are output and control signal (denoted without specification of a time instant k). The criterion (Eq. (8)) can be rewritten into a matrix form

$$J = (y-r)^T Q (y-r) + u^T R u, \tag{10}$$

where Q and R are weighting matrices of output error and control effort, respectively. The trajectory of the output is given as:

$$\begin{aligned} \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(T-1) \end{bmatrix} &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{T-1} \end{bmatrix} x_0 + \\ &+ \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & & \ddots & \\ CA^{T-2}B & \dots & CB & D \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(T-1) \end{bmatrix}, \end{aligned} \tag{11}$$

i.e.

$$y = \Gamma x_0 + Hu, \tag{12}$$

where Γ is observability matrix and H is matrix of impulse responses. Using Eq. (12), we can rewrite Eq. (10) as follows:

$$J = (\Gamma x_0 + Hu - r)^T Q (\Gamma x_0 + Hu - r) + u^T R u. \tag{13}$$

If the constraints (Eq. (9)) are not taken into account, model of the system is linear, and the criterion is quadratic, minimization problem Eq. (8) can be solved analytically – using methods such as completing the square or Lagrange multipliers, etc., leading to the optimal control sequence in the form of

$$u = -(H^T Q H + R)^{-1} H^T Q (\Gamma x_0 + Hu - r), \tag{14}$$

otherwise iterative methods have to be used. In the case of quadratic criterion and constrained optimization problem, a quadratic programming is one of the most popular solutions built in the most of the modern optimization packages. Alongside to the many commercial software, such as APC Library (Siemens), Pavilion8 (Pavilion Technologies), ADMC & DMCX1 (Cutlertech), RMP(Honeywell), etc., there is also a free software, e.g. Scilab[3] and its optimization routines.

III. Case Study

The methods described in the previous section were tested in spring 2009 at the building of the Czech Technical University in Prague. All algorithms were implemented in Scilab.

A. Description of the Building

The building of the Czech Technical University in Prague uses a Crittall [18] type ceiling radiant heating and cooling system. The Crittall system, invented in 1927 by R. G. Crittall and J. L. Musgrave, was a favorite heating system in the Czech Republic during 1960s for large buildings. In this system, the heating (or cooling) beams are embedded into the concrete ceiling. The control of individual rooms is very complicated due to the technical state of the control elements in all rooms. The control is therefore carried out for one entire building block, i.e. the same control effort is applied to all rooms of the building block. There are three building blocks with the same construction and orientation. Therefore, this situation is ideal for comparison of different control strategies, as depicted in Fig. 4.

A simplified scheme of the ceiling radiant heating system is illustrated in Fig. 3. The source of heat is a vapor-liquid heat exchanger, which supplies the heating water to the water container. A mixing occurs here, and the water is supplied to the respective heating circuits. An accurate temperature control of the heating water for respective circuits is achieved by a three-port valve with a servo drive. The heating water is then supplied to the respective ceiling beams. There is one measurement point in a reference room for every circuit. The setpoint of the control valve is therefore the control variable for the ceiling radiant heating system in each circuit.

[3] Open source scientific software package for numerical computations (http://www.scilab.org/)

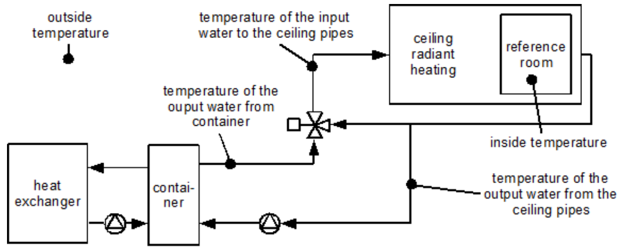


Fig. 3. Simplified scheme of the ceiling radiant heating system.

B. Description of the model

The ceiling radiant heating system was simplified into a linear, time invariant mathematical model. Outside temperature prediction and heating water temperature were used as the model inputs. Prediction of the outside temperature is composed of two values, T_{max} and T_{min}, defining a confidence interval. The only output of the model was the inside temperature. This can be formalized according to Eq. (7) as

$$x(k+1) = Ax + B\begin{bmatrix} T_{min} \\ T_{max} \\ T_{hw} \end{bmatrix}$$
$$T_{in} = Cx + D\begin{bmatrix} T_{min} \\ T_{max} \\ T_{hw} \end{bmatrix}, \qquad (15)$$

where T_{hw} is the temperature of the heating water and T_{in} denotes the inside temperature. The state x has no physical interpretation, when identified by means of the subspace identification. System order is determined by the identification algorithm.

Modeling of the heating system of the CTU building is discussed in detail in [19].

C. Results

Two nearly identical blocks of the CTU building were used for testing. The first block was controlled by the weather-compensated controller, while the second one by predictive controller. The current weather compensated control is well tuned up thanks to long term experience with the building. Testing was performed from March 24 to March 30, 2009, which was the end of the heating season in the Czech republic. That highlights the advantages of the predictive control, because there are days when there is no need of heating at all. The model predictive controller was able to identify those gaps where no control was needed, thus save energy. This can be seen in Fig. 4.

The efficiency of the predictive control was superior to the weather-compensated controller by about 10 %.

IV. Remarks to Future Development

As already stated, subspace identification represents one of the black-box approaches to system modeling. This, alongside with its advantages, carries also some drawbacks:

- The system might not be excited enough [12], i.e. the input of the system does not excite the system on satisfactory number of frequencies, thus identification algorithms lack considerable amount of information.
- User may have knowledge of some key feature or characteristics of the physical essence of the system, which is "lost" in the number of data.
- Natural characteristics of the data might pose considerable statistical problem.

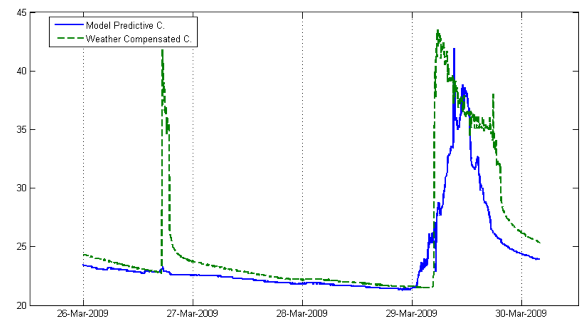


Fig. 4. Measured trends of the heating water temperature controlled by the weather compensated control and the predictive controller

Future development of the identification algorithm will try to remedy the above-mentioned problems. One way which is still not full discovered[4] is incorporating of prior information. This approach make use of Bayesian approach [20] to the system identification and changes nature of the subspace identification method towards grey-box identification.

Another aspect of the identification is the persistency of the excitation or the excitation itself. Data gathered from the measurement lack some important physical characteristics of the building. One of the possible approaches how to deal with this weak point is to generate artificial data that already contain the desired properties. There is also another possibility, more expensive though – a specially proposed experiment on real building, which is about to be done in the beginning of 2010.

V. Acknowledgment

This work has been supported from the state budget of the Czech Republic, through the Ministry of industry and commerce, in the scope of grant No. FR-TI1/517, *"Control Systems for Optimization of Energy Consumption in Low-Energy and Passive Houses"*.

VI. Conclusion

It is obvious, that predictive control has a great potential in the area of building heating control. Especially in case of buildings with great heat accumulation capabilities. The results from one week testing in spring 2009 are very encouraging. Testing confirmed our empirical experiences and the efficiency of the predictive controller was superior to the present well tuned weather compensated controller. However, it is a complicated technique and launching of

[4]Some methods (SWLRA) algorithms are capable of handling incorporation of prior information to SISO system, but algorithm for MIMO system is still missing

the predictive controller requires deep knowledge of identification methods and predictive control. It is questionable, whether this drawback can be overcome in the future and the tuning of the predictive controller would be feasible for wide range of practitioners.

REFERENCES

[1] *Buildings and their Impact on the Environment: A Statistical Summary*, 2009, rev. April 22, 2009, available on-line at http://www.epa.gov/greenbuilding/pubs/gbstats.pdf.

[2] J. Richalet, A. Rault, J. L. Testud, and J. Papon, "Algoritmic control of industrial process," in *Proceedings: Symposium on Identification and System Parameter Estimation*. Tbilisi: IFAC, 1976.

[3] V. Peterka, "Predictor-based self-tunning algorithm," *Automatica*, vol. 20, pp. 39–50, 1984.

[4] J. Richalet, A. Rault, J. Testud, and J. Papon, "Model heuristic predictive control," *Automatica*, vol. 14, pp. 413–428, 1978.

[5] W. H. Kwon, A. M. Bruckstein, and T. Kailath, "Stabilizing state feedback design via the moving horizon method," *International Journal of Control*, vol. 37, p. 631643, 1983.

[6] J. Rawlings and K. Muske, "Stability of constrained receding horizon," *IEEE Transaction on Automatic Control*, vol. 38, 1993.

[7] A. Zheng and M. Morari, "Stability of model predictive control with soft constraints," *Internal Report.Californian Institute of Technology*, 1994.

[8] P. J. Campo and M. Morari, "Robust model predictive control," *American Control Conference*, 1987.

[9] T. Y. Chen, "Application of adaptive predictive control to a floor heating system with a large thermal lag," *Energy and Buildings*, vol. 34, no. 1, pp. 45 – 51, 2002.

[10] R. Z. Freire, G. H. Oliveira, and N. Mendes, "Predictive controllers for thermal comfort optimization and energy savings," *Energy and Buildings*, vol. 40, no. 7, pp. 1353 – 1365, 2008.

[11] M. Gwerder, J. Tödtli, B. Lehmann, F. Renggli, and V. Dorer, "Control of thermally activated building systems," in *Proceedings: 9th REHVA World Congress Clima*, 2007.

[12] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, 1999.

[13] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge University Press, 2007.

[14] L. Ferkl and J. Široký, "Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods," *BUILDING AND ENVIRONMENT*, vol. 45, no. 1, Sp. Iss. SI, pp. 205–212, 2010.

[15] H. Kwakernaak and R. Sivan, *Linear Optimal Control Systems*. John Wiley & Sons, 1972.

[16] K. De Cock and B. De Moor, "Subspace identification methods," in *EOLSS, UNESCO Encyclopedia of Life Support Systems*. Oxford, UK: Eolss Publishers Co., Ltd., 2003, vol. 1, ch. Control systems, robotics and automation, pp. 933–979.

[17] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 1999.

[18] R. G. Crittall and J. L. Musgrave, "Heating and cooling of buildings," GB Patent No. 210880, April 1927.

[19] L. Ferkl and J. Široký, "Model of a hot water full circulation mode in a building," in *7th International Conference on Control and Automation*, Christchurch, 2009.

[20] P. Trnka and V. Havlena, "Subspace like identification incorporating prior information," *AUTOMATICA*, vol. 45, no. 4, pp. 1086–1091, 2009.

Appendix H

Ferkl, Verhelst, Helsen, Ciller, Komárková: Energy Savings Potential of a Model-Based Controller for Heating: A Feasibility Study

FERKL, L. – VERHELST, C. – HELSEN, L. – CILLER, A. – KOMÁRKOVÁ, D. Energy Savings Potential of a Model-Based Controller for Heating: A Feasibility Study. In Proceedings of the 2011 IEEE International Conference on Control Applications, Denver, USA, 2011.

Invited lecture

Number of citations: 0

Energy Savings Potential of a Model-Based Controller for Heating: A Feasibility Study

Lukáš Ferkl, Clara Verhelst, Lieve Helsen, Alexandr Ciller, Dana Komárková

Abstract— **In order to provide decision-making mechanism for application of model-based control of heating systems in buildings, a method has been proposed that assumes the energy balance calculation as the performance bound for the model-based controller. If the performance bound is significantly lower than the actual energy consumption, a more accurate estimate is made by means of dynamic model identification and model-based controller simulation. Said method has been tested as a case study on a health care center located in Prague, Czech Republic.**

I. Motivation

Previous research has already shown a great potential of savings achieved by model-based controllers, both in theory and practice. While theoretical results of the OptiControl project at ETH Zürich predict savings of 16-41 % [1] [2], practical results are very promising as well (27 % for TABS heating in Prague [3] and 24.5 % for cooling in Berkeley [4]. In order to promote more applications on building HVAC systems, it has to be noted that not all buildings are suitable for model-based control. The main criterion is the savings potential – if it is not high enough, the investment into the model-based control implementation does not return back in a reasonably short time (e.g. 5 years, which considered reasonable in construction, with respect to other investments) and the landlord is going to be rather disappointed. We have therefore proposed the following guideline, which should help to evaluate the savings potential for a specific building and provide a reasonable output, such that a sound decision can be made on whether to upgrade the building control system by a model-based controller or not.

The guideline is as follows:

1) Calculate the energy balance of the building using some standard means of calculation (such as energy audit or energy label [5]).
2) Compare the energy balance (i.e. static building energy losses) with the actual energy consumption.
3) If savings potential is reasonable (e.g. 15 % or more), identify a dynamic model of the building.
4) Design a model-based controller.
5) Compare the savings achieved by the model-based controller with the energy balance. If the results comply to each other, the building is suitable for model-based controller implementation.

L. Ferkl and A. Ciller are with Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technicka 2, 166 27 Praha 6, Czech Republic lukas.ferkl@fel.cvut.cz

C. Verhelst and L. Helsen are with Division Applied Mechanics and Energy Conversion, Department of Mechanical Engineering, K.U.Leuven, Celestijnenlaan 300A – post box 2421, B-3001 Leuven (Heverlee), Belgium {clara.verhelst, lieve.helsen}@mech.kuleuven.be

D. Komárková is with Technofiber, Lazaretní 1/7, 615 00 Brno dkomar@fi.muni.cz

II. Necessary Simplifications

In order to make the feasibility analysis simple enough and straightforward, several assumptions had to be made.

- *Continuously working heat pumps.* Even though heat pumps usually work in an on-off regime, we assume that they operate continuously. In practice, this would be achieved by a PWM modulation of the control signal.
- *No weather prediction.* The model-based controller will operate one a one-step ahead control horizon, as we assume that historical weather-prediction data are not available. Even though outside temperatures are usually stored within building energy management systems (BMS, EMS), the weather forecast is not. Moreover, the weather data are very costly in the Czech Republic (about 800 EUR/year for a single location), we will be on the safe side of the analysis even without the weather data.
- *A single simulation & control model.* We will use the identified model for both simulation and control. Even though this breaks one of the basic rules of controller design, we in fact have no other option, as i) two models have to be constructed with different methods, which is costly for the purposes of a feasibility study, and ii) we want to calculate a performance bound, therefore assuming we have a good model, the single-model simulation would approach the performance of a two-model simulation.
- *Coefficient of Performance (COP) of the heat pumps.* Coefficient of Performance is an efficiency measure for heat pumps defined as

$$COP = \frac{\dot{Q}_{\text{hp}}}{P_{\text{hp}}} \tag{1}$$

wherein $\dot{Q}_{\text{hp}}$ is the heat delivered by the heat pump and P_{hp} is the heat pump compressor power.
The international standard for COP calculation [6] is not specific enough and as has been shown by [7], manufacturers of the heat pumps often overestimate their coefficients of performance. Unfortunately, we did not have data available that would allow us to calculate the real COP; therefore, we rely on the heat pump manufacturer's data.

III. The BIOREGENA Building

BIOREGENA is a private health care center, located in Prague, Czech Republic. Due to the climatic environment, cooling is not a concern, as temperatures in the Czech Republic seldom require massive cooling of buildings. The building was built in 1970's and is made of concrete shell. In 2006, a general refurbishment took place – new windows and facade insulation have been installed, as well as new gas boilers and heat pumps (air-to-water) along with a digital control system. The building is heated by radiators, the heating system has three branches that go into three separate parts of the building. This study will only focus on one block of the building, as the other blocks are similar from the construction point of view. The building is shown in Fig. 1.

IV. Performance Bound

In order to calculate the performance bound for the heating system, we will estimate the energy losses based on energy balance of the building using the "envelope method". This method is a basis e.g. for energy label, as described in [5]. The result of the calculation is yearly energy losses in GJ, the equations are based on physical properties of the construction elements, such as:

- geographical location and orientation of the building
- material, thickness of the walls and windows
- thermal bridges
- occupancy
- window ventilation
- etc.

Basically, there are tens of algebraic equations needed to estimate the yearly energy losses; the entire set of equations is described in [5]. Unlike the ISO norm, we also considered uncertainties in the estimates of physical parameters based on the expertise of the construction engineers involved in the project. We considered several values of uncertainties:

- 0 % – wall thickness, window areas, etc.
- 10 % – e.g. wall and roof heat conductivity
- 15 % – e.g. losses through ground
- 20 % – e.g. air exchange ratio

For the BIOREGENA building, the energy losses were calculated as 807 $\pm$ 286 GJ/year. This corresponds to total heat losses of 4513 $\pm$ 750 W/K, or equivalently to a heat transfer coefficient of 0.85 $\pm$ 0.14 W/(m^{2}K), which means that the building complies with the requirements for the C-class building – see Fig. 2, which is very common for pre-1990, re-insulated buildings in the Czech Republic.

Fig. 1. Bioregena, the building of our concern

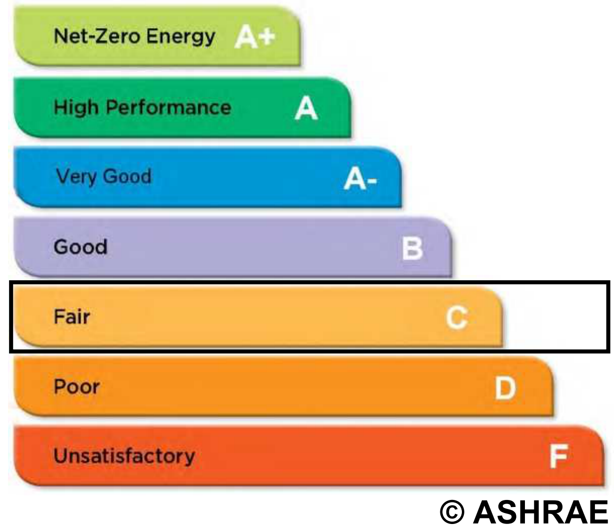


Fig. 2. Energy label of the BIOREGENA building

V. Dynamic Model Estimation

There are several ways to estimate the dynamic model of the building. The first option is to use measured data to obtain an LTI model using some statistical methods, such as ARX model estimation or subspace identification [8][9]. An example comparing both approaches can be found e.g. in [10].

For the case of BIOREGENA building, we have decided to identify a physical (or first-principle) model, in an off-line identification process. The main reason was that we had measured data available and the system is small-scale (it only contains one heating branch), so first-principle model should yield fair results with less effort than statistical identification, which was the experience gained in our previous work, described e.g. in [3].

We chose the description of the system according to [11] in the following way:

$$\begin{aligned}\begin{bmatrix} \dot{T}_{\text{return}}(t) \\ \dot{T}_{\text{zone}}(t) \end{bmatrix} &= \\ &= \begin{bmatrix} -\frac{k_{wz}}{\rho_w c_w V_w} & \frac{k_{wz}}{\rho_w c_w V_w} \\ \frac{k_{wz}}{k_b \tau_b} & -\frac{k_{wz}+k_b}{k_b \tau_b} \end{bmatrix} \begin{bmatrix} T_{\text{return}}(t) \\ T_{\text{zone}}(t) \end{bmatrix} + \\ &+ \begin{bmatrix} 0 & \frac{1}{\rho_w c_w V_w} \\ \frac{1}{\tau_b} & 0 \end{bmatrix} \begin{bmatrix} T_{\text{out}} \\ \dot{Q}_{\text{tot}} \end{bmatrix}\end{aligned} \tag{2}$$

where

T_{return}	return water temperature
T_{zone}	zone temperature
T_{out}	outside temperature
$\dot{Q}_{\text{tot}}$	total heating power input
k_{wz}	heat transfer coefficient between the water and the zone
ρ_w	water density
c_w	water specific heat capacity
V_w	water volume in the heating system
k_b	transfer coefficient between the zone and the ambient air
τ_b	principal time constant of the building

In the sense of model-based control, the total heating power input $\dot{Q}_{\text{tot}}$ is our manipulated variable (MV, system control input), outside temperature T_{out} is a disturbance variable (DV), zone temperature T_{zone} is a controlled variable (CV, system control output) and return water temperature T_{return} is a measured system state.

However, we do not need to identify every single parameter of the building model. If we look at Equation (2) carefully, we can see that there are in fact only four structural parameters and the equation can be simplified in the following way by substitution:

$$\begin{bmatrix} \dot{T}_{\text{return}}(t) \\ \dot{T}_{\text{zone}}(t) \end{bmatrix} = \\ = \begin{bmatrix} -a & a \\ c & -(c+d) \end{bmatrix} \begin{bmatrix} T_{\text{return}}(t) \\ T_{\text{zone}}(t) \end{bmatrix} + \\ + \begin{bmatrix} 0 & b \\ d & 0 \end{bmatrix} \begin{bmatrix} T_{\text{out}} \\ \dot{Q}_{\text{tot}} \end{bmatrix} \tag{3}$$

In the yearly data, 5 representative data samples were chosen, such that the entire temperature range was covered. As an identification tool, we used the ACADO Toolkit [12], which was used for control calculation as well.

The numerical results from the system identification can be seen in Tab. I. We can see that the respective models show differences, which are significant for some of the parameters. Parameters a and b do not vary a lot, which means that the identification of the model corresponding to the return water temperature T_{return} is quite accurate. However, variation of parameters c and d is significant; on the other hand, the overall fit for one-step ahead prediction of the models was satisfactory for control purposes.

The graphical results can be seen in Fig. 3 and Fig. 4. Figure 3 shows the identification procedure on the data set covering the outside temperature T_{out} around 0 °C; we can see that for an identification over one-day data, the match for the next day is reasonable. We have also observed that the best identification result is achieved if identification stops close the first heating impulse in the morning, which might contain high-order dynamics not present in the model (2) and hence mislead the identification algorithm. However, this effect does not occur when identifying the model with a sampling period of 60 minutes (Fig. 4).

TABLE I

COMPARISON OF DIFFERENT DATA SAMPLES AND THE IDENTIFICATION RESULTS.

Data set	Sampling $[min]$	No. of samples $[-]$	a $\times 10^{-1}$	b $\times 10^{-1}$	c $\times 10^{-2}$	d $\times 10^{-3}$
1.	10	101	5.836	3.808	1.300	6.306
2.	10	156	4.702	2.624	3.705	8.300
3.	10	174	4.329	2.904	0.456	0.804
4.	10	107	4.875	3.292	1.310	5.091
5.	10	121	4.825	3.437	0.163	0.815
6.	60	67	8.000	5.383	10.45	0.345

We can see that for the model identified with the 60 minutes sampling period shows a good prediction over about 5 days. But after some discussions with the facility manager of the building, we decided to not use it as the controller model, as the control action has to be changed more often due to operational characteristics of the building. Moreover, the weather prediction is not good enough for the Prague region (see e.g. [10]) for the period of 5 days ahead, and in reality, this would be a significantly limiting factor.

VI. Controller Design

As already stated, the model-based controller will be used for comparison with the static losses computation, as introduced in the previous sections. We have mentioned that the method introduced in this paper is meant to be used as a feasibility study for implementation of advanced controllers. As historical weather prediction data are usually not commonly available, we have decided to use a model-based controller with a one-step ahead control horizon (i.e. MPC with one-step prediction). It is obvious that this controller is on the save side – the savings achieved by longer prediction horizons will be even larger.

First of all, we have to determine the coefficient of performance COP of the heat pumps:

$$COP = f(T_{\text{return}}, T_{\text{out}}) \tag{4}$$

Based on technical sheet data, there were six COPs known for two return water temperatures T_{return} (35 and 50 °C) and three outside temperatures T_{out} (-15, 0 and 7 °C).

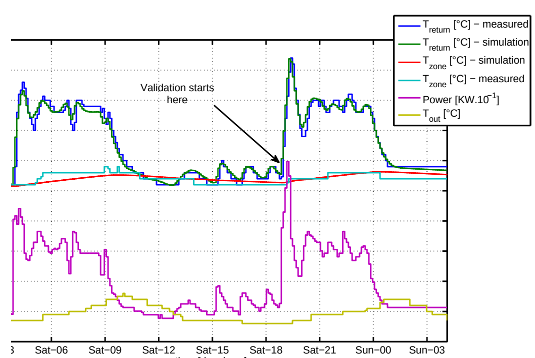


Fig. 3. Identification and validation of the model, sampling period 10 minutes.

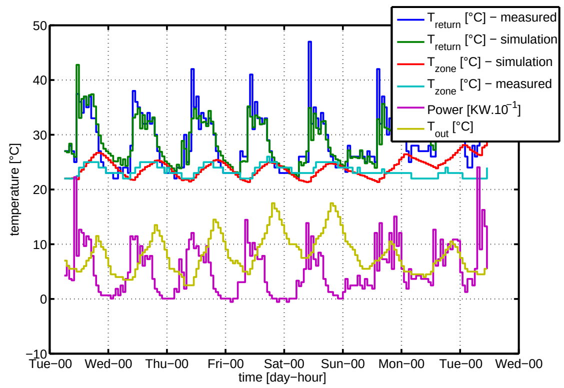


Fig. 4. Validation of the model, sampling period 1 hour

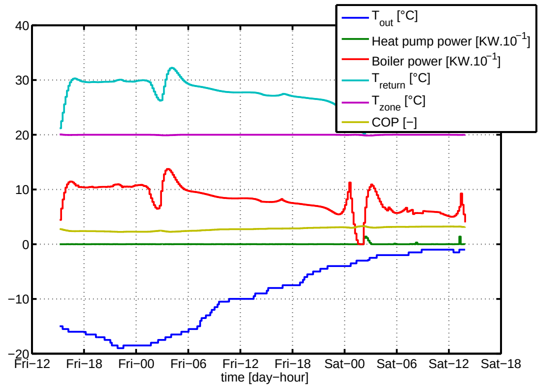


Fig. 5. Control, outside temperature T_{out} well below 0 °C.

These data were interpolated by a quadratic function, which was further limited by two constraints:

1) $COP \geq 1$
2) COP does not rise anymore with outside temperatures $T_{\text{out}} \geq 15$ °C.

In order that the controller gives us smooth control action, we have two possibilities. One is to incorporate this requirement into the optimization solver (most of the solvers enable this option), the other one is to define a new dummy state x_3, which will reflect the change in the control action $\dot{Q}_{\text{tot}}(t)$:

$$\dot{x}_3(t) = \tau_3 \cdot (\dot{Q}_{\text{tot}}(t) - x_3(t)) \tag{5}$$

where τ_3 is an arbitrary time constant which may be used to tune the "smoothness" of the control input $\dot{Q}_{\text{tot}}(t)$; τ_3 was tuned to 1 in our case, as further smoothing of the control input was not necessary.

Now we can introduce the control cost function. We will minimize the criterion

$$\min_{\dot{Q}_{\text{hp}}(t), \dot{Q}_{\text{gas}}(t)} \{J_1 + J_2 + J_3\} \tag{6}$$

wherein J_1, J_2 and J_3 are defined as follows.

The first criterion term minimizes the price of the heat (in terms of money):

$$J_1 = \{(1-K) \cdot \ldots$$

$$\ldots \cdot \left(\frac{\dot{Q}_{\text{hp}}(t) \cdot \text{electricity price}}{COP(t)} + \text{gas price} \cdot \dot{Q}_{\text{gas}}(t) \right) \Bigg\} \tag{7}$$

where K is used for tuning, electricity and gas price are calculated in such a way that electricity price+gas price $= 1$, $\dot{Q}_{\text{hp}}(t)$ is the electrical power in [kW] and $\dot{Q}_{\text{gas}}(t)$ is the gas heat power in [kW], assuming that burning of 1 [m^3] of natural gas yields 10.6 [kWh] of heat. The prices were taken as fixed, according to the current state in the BIOREGENA building (5 CZK/kWh[1] for electricity and 15 CZK/m^3 for gas); however, use of variable prices (e.g. day and night) is straightforward. Now we get that

$$\dot{Q}_{\text{tot}}(t) = \dot{Q}_{\text{hp}}(t) + \dot{Q}_{\text{gas}}(t).$$

[1] 1 EUR = 24.5 CZK (Czech Crown), as of February 2011.

The next criterion term has to minimize the difference between the required and actual zone temperature T_{zone}:

$$J_2 = \left\{ L \cdot K \cdot (T_{\text{zone}}(t) - T_{\text{zone REF}}(t))^2 \right\} \tag{8}$$

where L is yet another tuning parameter used to set proportion between criteria (8) and (9). The last requirement we have is the smoothness of the control action, which is achieved (with reference to Eq. 5) by the following criterion term:

$$J_3 = \left\{ K \cdot (L-1)(\dot{Q}_{\text{tot}}(t) - x_3(t))^2 \right\} \tag{9}$$

According to the above equations, the controller will be defined as

$$\dot{Q}_{\text{tot}}(t) = \arg \left\{ \min_{\dot{Q}_{\text{hp}}(t), \dot{Q}_{\text{gas}}(t)} \{J_1 + J_2 + J_3\} \right\} \tag{10}$$

The results of the controller simulation are shown in Fig. 6 for temperatures around 0 °C and in Fig. 5 for temperatures well below 0 °C. It is around 0-5 °C where the heat pumps take over the heating of the building, while they remain inactive below 0 °C. This is due to currently low price of natural gas and expensive electricity. The usage of heat pumps would be even narrower if proper COP were available – as already noted, heat pumps manufacturers tend to overestimate the COP due to imprecise requirements of the respective standard [6].

I has to be noted again that the same model was used for both simulation and control, as we have explained in Section II.

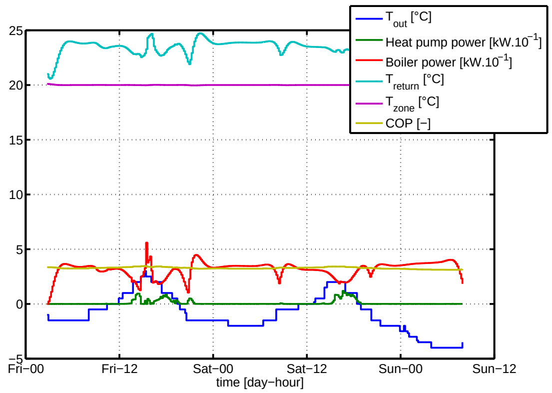


Fig. 6. Control, outside temperature T_{out} around 0 °C.

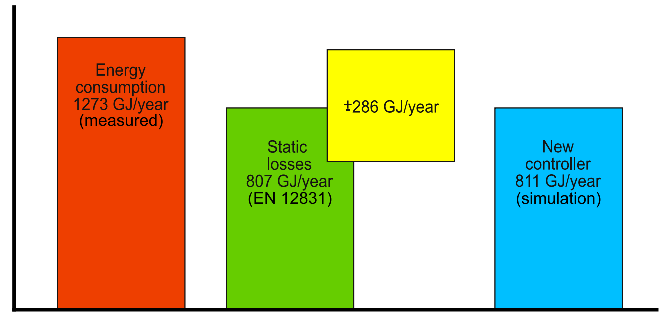


Fig. 7. Achieved savings

VII. Comparison of the Static and Dynamic Performance

Now we have to compare the results from the static and dynamic performance analysis. For evaluation, the yearly data were divided into 5 groups corresponding to the identified models and yearly energy consumption of the model-based controller was calculated. No cooling is installed on the building. The model-based controller predicts average savings of 36.25 % per year; the simulation also shows that maximum savings (close to 50 %) can be achieved for outside temperatures around $T_{\text{out}} = 7$ °C. This corresponds to the experience with model-based predictive controller, which was implemented in 2010 on an office building and has been in full operation for more than one year now (see [3]); also in this real case, the MPC has the largest savings potential around 5-10 °C. The reason of this is subject to ongoing research and the problem is still open, there are two major hypotheses at the moment:

1) State-of-the-art controllers (PID, heating curve) tend to perform frequent switching and get "confused" easily; comfort range tracking is also troublesome for said controllers, which might be additional clue to the problem.
2) The building insulation is never perfect and dew point becomes an issue for said temperatures; this is not reflected by the state-of-the-art controllers, but the information is somehow incorporated in the statistically identified models. This could also explain variations of the c and d parameters in Tab. I.

As a result, the model-based controller showed energy consumption 811 GJ/year, while the static losses were calculated to 807 GJ/year (Fig. 7). Even though the match looks perfect, it has not to be overestimated, due to the uncertainty region of the static losses and simplifications mentioned in Section II.

The figures also show that the model-based controller does not use the heat pumps, when outside temperature is below ca. 5 °C, due to the ratio of the electricity and gas price (gas is much cheaper than electricity). One could find the ratio of the electricity and gas price, which would make the investment into a heat pump economical, but this was not the purpose of this paper.

VIII. Conclusions and Next Steps

Our aim was to validate a method to estimate the energy savings achievable by modern controllers, such as MPC, compared to state-of-the-art controllers. The method comprises of two steps: First, the energy consumption compared to the static losses based on the widely used EN 12831 standard [5]. If the savings potential is significant, the second step is to identify a dynamic model of the building and to perform a simulation with a model-based controller, as proposed in this paper. If the results correspond to each other and the savings are significant, the building would be suitable for application of a modern controller, such as MPC.

The results for the case of the BIOREGENA building show that the estimated static heat losses indeed correspond to the simulation results for the dynamic building model with an ideal, optimal controller. However promising this result is, it has to be evaluated on more buildings for general conclusions. One should also keep in mind the simplifications of Section II, which make the savings potential rather optimistic.

The next step would be implementation of the controller, with all the neglected factors mentioned in Section II taken into account. Unfortunately, there are technical problems with hydraulic separation of the heating circuits for the BIOREGENA building which hindered implementation of MPC during the 2010/11 heating season (all the heating system has to be drained to fix the problem, which is impossible during winter). We hope that application in the next season, along with more feasibility studies, will provide more confidence in the method described in this paper and help to spread implementations of advanced controllers in building automation.

IX. Acknowledgments

This work has been supported from the state budget of the Czech Republic, through the Ministry of industry and commerce, in the scope of grant No. FR-TI1/517, "Control Systems for Optimization of Energy Consumption in Low-Energy and Passive Houses", and by a CTU internal grant SGS 10-802830, "Optimization methods for modeling and control and their applications".

References

[1] D. Gyalistras *et al.*, "Use of Weather and Occupancy Forecasts for Optimal Building Climate Control," ETH Zürich, Tech. Rep., 2010.

[2] D. Gyalistras, M. Gwerder, F. Oldewurtel, C. Jones, M. Morari, B. Lehmann, K. Wirth, and V. Stauch, "Analysis of energy savings potentials for integrated room automation," in *10th REHVA World Congress Clima*, 2010.

[3] S. Prívara, J. Široký, L. Ferkl, and J. Cigler, "Model predictive control of a building heating system: The first experience," *Energy and Buildings*, vol. 43, no. 2-3, pp. 564 – 572, 2011.

[4] Y. Ma, F. Borrelli, B. Hencey, A. Packard, and S. Bortoff, "Model predictive control of thermal energy storage in building cooling systems," in *Decision and Control, 2009 held jointly with the 2009 28th Chinese Control Conference. CDC/CCC 2009. Proceedings of the 48th IEEE Conference on*, 2009, pp. 392 –397.

[5] *DIN EN 12831:2003 Heating systems in buildings – Method for calculation of the design heat load*. Geneva, Switzerland: International Organization for Standardization, 2003.

[6] *EN 14511-2:2004 Air conditioners, liquid chilling packages and heat pumps with electrically driven compressors for space heating and cooling. Test conditions*. Geneva, Switzerland: International Organization for Standardization, 2004.

[7] R. Matějka and R. Dlabaja, "Heat pumps testing," *TZB Info*, June 2010, (in Czech). [Online]. Available: http://vytapeni.tzb-info.cz/tepelna-cerpadla/6543-zkouseni-tepelnych-cerpadel

[8] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA, 1999.

[9] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Pard, Nowell, Massachusetts: Kluwer Academic Publishers, 1999.

[10] L. Ferkl and J. Široký, "Ceiling radiant cooling: Comparison of ARMAX and subspace identification modelling methods," *Building and Environment*, vol. 45, no. 1, pp. 205–212, 2010.

[11] M. A. Bianchi, "Adaptive modellbasierte prädiktive regelung einer kleinwärmepumpenanlage," Ph.D. dissertation, ETH Zürich, 2006, (in German).

[12] B. Houska, H. J. Ferreau, and M. Diehl, "Acado toolkit – an open-source framework for automatic control and dynamic optimization," *Optimal Control Applications and Methods*, 2010, article in press. [Online]. Available: http://dx.doi.org/10.1002/oca.939

Appendix I

Prívara, Široký, Ferkl, Cigler: Model Predictive Control of a Building Heating System: The First Experience

PRÍVARA, S. – ŠIROKÝ, J. – FERKL, L. – CIGLER, J. Model Predictive Control of a Building Heating System: The First Experience. Energy and Buildings. 2011, 43, 2-3, pp. 564–572. ISSN 0378-7788.

Number of citations (excluding autocitations):

- WoS: 1
- Scopus: 2
- Google Scholar: 1

Impact factor: 2.041

Energy and Buildings 43 (2011) 564–572

Contents lists available at ScienceDirect

Energy and Buildings

journal homepage: www.elsevier.com/locate/enbuild

Model predictive control of a building heating system: The first experience

Samuel Prívara[a,*], Jan Široký[b], Lukáš Ferkl[a], Jiří Cigler[a]

[a] Department of Control Engineering, Faculty of Electrical Engineering, Czech Technical University in Prague, Technická 2, 166 27 Praha 6, Czech Republic
[b] Department of Cybernetics, Faculty of Applied Sciences, University of West Bohemia in Pilsen, Czech Republic

ARTICLE INFO

Article history:
Received 16 February 2010
Received in revised form 12 October 2010
Accepted 18 October 2010

Keywords:
Optimal control
Subspace identification
Model predictive control
Costs effectiveness

ABSTRACT

This paper presents model predictive controller (MPC) applied to the temperature control of real building. Conventional control strategies of a building heating system such as weather-compensated control cannot make use of the energy supplied to a building (e.g. solar gain in case of sunny day). Moreover dropout of outside temperature can lead to underheating of a building. Presented predictive controller uses both weather forecast and thermal model of a building to inside temperature control. By this, it can utilize thermal capacity of a building and minimize energy consumption. It can also maintain inside temperature at desired level independent of outside weather conditions. Nevertheless, proper identification of the building model is crucial. The models of multiple input multiple output systems (MIMO) can be identified by means of subspace methods. Oftentimes, the measured data used for identification are not satisfactory and need special treatment. During the 2009/2010 heating season, the controller was tested on a large university building and achieved savings of 17–24% compared to the present controller.

1. Introduction

According to the U.S. Energy Information Administration, in 2005, buildings accounted for 39% of total energy usage, 12% of the total water consumption, 68% of total electricity consumption, and 38% of the carbon dioxide emissions in the U.S.A. [1]. Although the energy efficiency of systems and components for heating, ventilating, and air conditioning (HVAC) has improved considerably over recent years, there is still potential for substantial improvements. This article deals with an advanced control technique, that can provide significant energy savings in comparison with conventional, non-predictive techniques.

Widely used control strategy of water heating systems is the weather-compensated control. This feedforward control can lead to poor energy management or reduced thermal comfort even if properly set up, because it utilizes current outside temperatures only. Weather conditions, however, can change dramatically in few hours; and due to the heat accumulation in large buildings, it can lead to underheating or overheating of the building easily.

During recent years, significant advances have been done for the HVAC control systems [2–6]. For instance, continuous adaptation of control parameters, optimal start–stop algorithms, optimization of energy loads shifting [7], or inclusion of free heat gains in the control algorithm are particular improvements of the building heating system [8]. Some new concepts have been verified by simulations [9,10], nevertheless they are still waiting for real operations. The model predictive control, [11–15] (MPC) presented in this article introduces a different approach to the heating system control design. As the outside temperature is one of the most influential quantity for the building heating system, weather forecast is employed in the predictive controller. It enables to predict inside temperature trends according to the selected control strategy. The aims of the control can be expressed in natural form as thermal comfort and economy trade off. Unfortunately, this concept has some drawbacks, such as extensive computational requirements or necessity of a mathematical model of the physical system (building in this case).

All these issues are discussed in detail in following sections, which are organized as follows. Section 2 compares the current control techniques with MPC. Section 3 introduces model predictive control concept more in detail and explains the mathematical background of this technique. This section also addresses new modified zone model predictive controller. Problem of the model identification is discussed as well. Application results are summarized in Section 4. Remarks to future development are outlined in Section 5. The last section concludes the work.

List of abbreviations used throughout the article is mentioned in Table 1.

2. Current heating control strategies

Let us briefly compare the major state-of-the-art heating control techniques – on–off room temperature control, weather-

* Corresponding author. Tel.: +420 776 697 672.
E-mail addresses: samuel.privara@fel.cvut.cz (S. Prívara), jan.siroky@rcware.eu (J. Široký), lukas.ferkl@fel.cvut.cz (L. Ferkl), jiri.cigler@fel.cvut.cz (J. Cigler).

0378-7788/$ – see front matter © 2010 Elsevier B.V. All rights reserved.
doi:10.1016/j.enbuild.2010.10.022

Table 1
Notation.

Abbrev.	Meaning
ARX	Auto-regressive model with external inputs
ARMAX	Auto-regressive, moving average model with external inputs
CTU	Czech Technical University in Prague
HVAC	Heating, ventilation and air-conditioning systems
MIMO	Multiple-input, multiple-output systems
MPC	Model predictive control
OE	Output error model
PID	Proportional – integrative – derivative controller
SISO	Single-input, single-output systems
WC	Weather-compensated control

compensated control, and PID [8] control – with the proposed application of MPC.

The *on–off room temperature control* is the simplest type of control; the heating devices in a room are switched on and off (device state S) according to some room temperature error ($e_t = t_{\text{required}} - t_{\text{room}}$) threshold, usually implemented as a suitable hysteresis curve $f_{\text{on–off}}$:

$$S = f_{\text{on–off}}(e_t) \tag{1}$$

This is a very simple feedback control, which does not contain any information about the dynamics of the building. The main advantage is its simplicity.

On the contrary, the *weather-compensated control* is a feedforward control, which also does not contain any information about the building dynamics. The temperature of the heating medium, such as water (t_{water}), is set according to the outside temperature t_{outside} by means of predetermined heating curves $f_{\text{w-c}}$, that is

$$t_{\text{water}} = f_{\text{w-c}}(t_{\text{outside}}) \tag{2}$$

In spite of the lack of dynamics in the control, this is a long used and proven control strategy; its advantage is its robustness and simple tuning.

PID control is one of the most favorite strategies of control engineers [16,17]. It is a feedback control with some information about the system dynamics, that is, the heating water temperature t_{water} is determined according to the room temperature error e_t and "some" history:

$$t_{\text{water}} = f_{\text{PID}}(e_t, \text{history}) \tag{3}$$

PID controllers are robust and allow accurate tuning, but they cannot reflect the outside temperature effects. This is the reason why PIDs in HVAC control are not as common as in other control applications.

Even though all the above controllers are easy to tune for single-input, single-output (SISO) systems, their tuning for multiple-input multiple-output (MIMO, sometimes called multidimensional) systems becomes very difficult or even impossible in practice. The PID control can be applied to MIMO systems only in very rare occasions, in case of specially structured (input–output decoupled) systems.

We would therefore appreciate some control strategy, which would have a feedback (i.e. the room temperature error e_t is used), use as much information as possible (the outside temperature t_{outside}, the weather forecast $t_{\text{predicted}}$, and others x) and include some system dynamics ("history") as well. This can be formalized – in the spirit of the above Eqs. (1)–(3) – as

$$t_{\text{water}} = f_{\text{MPC}}(e_t, t_{\text{outside}}, t_{\text{predicted}}, x, \text{history}) \tag{4}$$

These requirements are satisfied by a so-called model (based) predictive controller (MPC), which is specially suitable for systems with multiple inputs and multiple outputs, which is very typical for heating systems. Its main drawbacks are high demands for computational resources and non-trivial mathematical background, especially in the "Model" part of the controller.

Fig. 1. The building of the Czech Technical University in Prague that was used for MPC application.

3. Model predictive control

3.1. State of the art

Model (based) predictive control (MPC) is a method of advanced control originated in late seventies and early eighties in the process industries (oil refineries, chemical plants, etc.) [11]. The MPC is not a single strategy, but a vast class of control methods with the model of the process explicitly expressed trying to obtain control signal by minimizing objective function subject to (in general) some constraints [18]. The minimization is performed in an iterative manner on some finite optimization horizon to acquire N step ahead prediction of control signal that leads to minimum criterion subject to all constraints. This, however, carries lots of drawbacks such as no feedback, no robustness, and no stability guarantee. Many of these drawbacks can be overcome by applying so-called receding horizon, i.e. at each iteration only the first step of the control strategy is implemented and the control signal is calculated again, thus, in fact, the prediction horizon keeps being shifted forward. Stability of the constrained receding horizon has been discussed in Refs. [13,14], or yet another approach using robust control design approach [15].

There were several attempts made to utilize predictive control concept in HVAC in the last decade [19,9,20,21,10]. Complex view into area of optimal building control gives the project OptiControl.[1] Besides its own results, it also provides a wide range of references to the related articles. Another project worth to mention is the predictive networked building control that deals with predictive control of the thermal energy storage on the campus of the UC-Berkeley.[2] Most of the articles devoted to the HVAC predictive control conclude results just by numerical simulations. On the contrary, this article describes MPC being tested on the real eight-floor building (see Fig. 1).

3.2. Principles

We will now briefly describe the basic ideas lying behind the MPC. To be more illustrative, we will take the course of the MPC implementation in our own project; even though the experienced practitioners in heating control are rather conservative in their field, they can accept new method, such as MPC, if performed in small, consecutive steps, which helps them to get acquainted with its principles.

[1] http://www.opticontrol.ethz.ch.
[2] http://sites.google.com/site/mpclaboratory/research/predictive-networked-building-control-1.

Having a well working control, such as weather-compensated control of a building, it is often unwise to change it to something novel, but unproven. However, it can be very advantageous to provide a “tool” that would enhance the possibilities of the existing system. A mathematical model can be such a “tool”, allowing the system operators to predict the behavior of the building. If the model is accurate enough (e.g. as a one-day predictor), another feature can be added—the operator can experiment with the model and try some “what if” scenarios. The next step is obviously implementation of an algorithm that proposes the best scenarios; it is still a “tool”, the model and algorithm are not involved in the control loop. That would be the last step – after the operator begins to trust the algorithm, he begins to ask for the closer of the control loop incorporating what we call model predictive control.

To be more precise, the first step is to find a dynamic model P

$$y = P(u, t) \tag{5}$$

where y is the output, u is the input (both can be vectors) and t is time. Inputs u may be entered by the operator in the beginning, such that he can see the expected behavior of the system, as seen on outputs y. The next step is finding the optimal inputs u automatically. This can be achieved by introducing an optimality criterion $J(y, u, t)$, wherein the control demands are described in the language of mathematics. Substituting from (5), the optimal control inputs can be found by computing

$$u_{\text{optimal}} = \min_u J(P(u, t), u, t) \tag{6}$$

subject to “some” constraints. This very basic idea will now be discussed more in detail.

3.3. Model identification

One of the crucial contributors to the quality of the control is a well identified model which will be later on used for control in MPC algorithm. There are several completely different approaches to system identification (see e.g. [22,23]). Some of them use knowledge of system physics, while others exploit statistical data, such as grey-box [24,25] (some prior information such as system structure is known in advance) or black-box identification. Grey box methods using models such as ARX, ARMAX, OE and others are well established among the practitioners as well as theoreticians. There is, however, a significant problem, when multiple input multiple output (MIMO) systems are considered. The standard input–output identification methods are not capable of dealing with such a model, thus one has to either reformulate the problem to several single-output cases, or to use state-space identification methods, such as subspace methods. The first approach, including computer modeling of the building, as well as comparison of ARMAX model and subspace methods, was briefly described in [26].

The main difference between classical and subspace identification can be seen in Fig. 2 (see Ref. [27]). Given the sequence of input and output data, $u(k)$ and $y(k)$, respectively, do:

- **Classical approach**. Find system matrices, then estimate the system states, which often leads to high order models that have to be reduced thereafter.
- **Subspace approach**. Use orthogonal and oblique projections to find Kalman state sequence, then obtain the system matrices using least squares method. Here we introduce the latter—subspace identification methods.

The objective of the subspace algorithm is to find a linear, time invariant, discrete time model in an innovative form:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k), \end{aligned} \tag{7}$$

based on given measurements of the input $u(k) \in \mathbb{R}^m$ and the output $y(k) \in \mathbb{R}^l$ generated by an unknown stochastic system of order n, which is equivalent to the well-known stochastic model:

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + w(k) \\ y(k) &= Cx(k) + Du(k) + v(k), \end{aligned} \tag{8}$$

with covariance matrices Q, S and R of process and measurement noise sequences as follows:

$$\operatorname{cov}(w, v) = E\left(\begin{bmatrix} w(p) \\ v(p) \end{bmatrix} \begin{bmatrix} w^T(q) & v^T(q) \end{bmatrix}\right) = \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \delta_{pq} \geq 0, \tag{9}$$

and with A, B, C, and D denoting system matrices and K and e in (7) is Kalman gain – derived from the Algebraic Riccati Equation (ARE) [28], and white noise sequence, respectively. Loosely speaking, the objective of the algorithm is to determine the system order n and to find the matrices A, B, C, D and K.

3.3.1. Data matrices for subspace algorithm

The following matrices are necessary to form for subspace algorithm. Notation was adapted as in Ref. [27]. Upper index d denotes deterministic subsystem, while the upper index s denotes stochastic subsystem. Two kinds of matrices are used for subspace algorithm, data and system related matrices.

- **Data matrices**. Input and output block Hankel matrix are formed from input and output data as follows:

$$U_{0|2i-1} = \left(\begin{array}{ccccc} u_0 & u_1 & u_2 & \cdots & u_{j-1} \\ u_1 & u_2 & u_3 & \cdots & u_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{i-1} & u_i & u_{i+1} & \cdots & u_{i+j-2} \\ \hline u_i & u_{i+1} & u_{i+2} & \cdots & u_{i+j-1} \\ u_{i+1} & u_{i+2} & u_{i+3} & \cdots & u_{i+j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ u_{2i-1} & u_{2i} & u_{2i+1} & \cdots & u_{2i+j-2} \end{array}\right) \quad Y_{0|2i-1} = \left(\begin{array}{ccccc} y_0 & y_1 & y_2 & \cdots & y_{j-1} \\ y_1 & y_2 & y_3 & \cdots & y_j \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{i-1} & y_i & y_{i+1} & \cdots & y_{i+j-2} \\ \hline y_i & y_{i+1} & y_{i+2} & \cdots & y_{i+j-1} \\ y_{i+1} & y_{i+2} & y_{i+3} & \cdots & y_{i+j} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{2i-1} & y_{2i} & y_{2i+1} & \cdots & y_{2i+j-2} \end{array}\right), \tag{10}$$

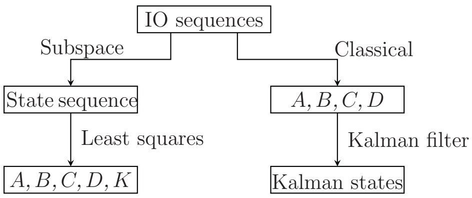


Fig. 2. Comparison between classical and subspace identification methods.

which can be written in shorten form as follows:

$$\begin{pmatrix} \frac{U_{0|i-1}}{U_{i|2i-1}} \end{pmatrix} = \begin{pmatrix} \frac{U_p}{U_f} \end{pmatrix} \\ \begin{pmatrix} \frac{Y_{0|i-1}}{Y_{i|2i-1}} \end{pmatrix} = \begin{pmatrix} \frac{Y_p}{Y_f} \end{pmatrix}, \tag{11}$$

where matrices U_p and U_f represent past and future inputs, respectively. Outputs $y(k)$ and noise $e(k)$ related matrices can be formed in similar manner. Grouped data matrix consisting of past input and past output data is formed as follows:

$$W_p = W_{0|i-1} = \begin{pmatrix} \frac{U_{0|i-1}}{Y_{0|i-1}} \end{pmatrix}.$$

- **System related matrices**. Extended ($i > n$) observability (Γ_i) and reversed extended controllability (Δ_i) matrices for deterministic and stochastic subsystems, respectively are defined as follows:

$$\Gamma_i = \begin{pmatrix} C \\ CA \\ \vdots \\ CA^{i-1} \end{pmatrix} \tag{12}$$

$$\Delta_i^d = \begin{pmatrix} A^{i-1}B & A^{i-2}B & \dots AB & B \end{pmatrix} \tag{13}$$

$$\Delta_i^s = \begin{pmatrix} A^{i-1}K & A^{i-2}K & \dots AK & K \end{pmatrix} \tag{14}$$

Algorithm also uses lower block triangular Toeplitz matrix for deterministic and stochastic subsystem, respectively:

$$H_i^d = \begin{pmatrix} D & 0 & 0 & \dots & 0 \\ CB & D & 0 & \dots & 0 \\ CAB & CB & D & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{i-2}B & CA^{i-3}B & CA^{i-4}B & \dots & D \end{pmatrix}$$

$$H_i^s = \begin{pmatrix} I & 0 & 0 & \dots & 0 \\ CK & I & 0 & \dots & 0 \\ CAK & CK & I & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ CA^{i-2}K & CA^{i-3}K & CA^{i-4}K & \dots & I \end{pmatrix}. \tag{15}$$

3.3.2. Subspace algorithm

The entry point to the algorithm are input–output equations as follows:

$$\begin{aligned} Y_p &= \Gamma_i X_p + H_i^d U_p + H_i^s E_p \\ Y_f &= \Gamma_i X_f + H_i^d U_f + H_i^s E_f \\ X_f &= A^i X_p + \Delta_i^d U_p + \Delta_i^s E_p. \end{aligned} \tag{16}$$

Oblique projection as described in Refs. [29,27] is the main tool used in subspace methods is defined as follows:

$$\mathcal{O}_i = Y_f \underset{U_f}{/} W_p. \tag{17}$$

The order of the system can be determined from analysis of singular values obtained using singular value decomposition (SVD) of $W_1\mathcal{O}_i W_2$, where W_i are weighting matrices of particular size and determine resulting state space basis. It has been shown [27], that $\mathcal{O}_i = \Gamma_i \tilde{X}_i$, where $\tilde{X}_i$ is Kalman filter state sequence. This factorization also yields extended observability matrix Γ_i and Kalman filter states $\tilde{X}_i$.

Algorithm continues from either Γ_i or $\tilde{X}_i$ in a slightly different manner depending on particular subspace identification algorithm, however, both ways lead to a computation of system matrices A and C using least squares method.

Computation of system matrices B and D follows subject to matrices A and C computed in previous step. Different approaches for matrices determination are addressed in detail in Ref. [27].

The algorithm concludes with computation of Kalman gain matrix K in a standard way using state and output noise covariance matrices (9) which are computed from residuals of the previous computations.

The model structure used in MPC is the state-space model (7) identified by subspace identification (described in Section 3.3) from measured data. The application of the model will become apparent later in this section.

3.4. Predictive controller

3.4.1. MPC strategy

The MPC strategy comprises two basic steps:

- The future outputs are predicted in an open-loop manner using the model provided information about past inputs, outputs and future signals, which are to be calculated. The future control signals are calculated by optimizing the objective function, i.e. chosen criterion, which is usually in the form of quadratic function. The criterion constituents can be as follows:
  - errors between the predicted signal and the reference trajectory $y_r(k)$;
  - control effort;
  - rate of change in control signals.
- The first component of the optimal control sequence $u(k)$ is sent to the system, whilst the rest of the sequence is disposed. At the next time instant, new output $y(k+1)$ is measured and the control sequence is recalculated, first component $u(k+1)$ is applied to the system and the rest is disposed. This principle is repeated *ad infinitum* (receding horizon).

The reference trajectory $y_r(k)$, room temperature in our case, is known prior, as a schedule. The major advantage of MPC is the ability of computing the outputs $y(k)$ and corresponding input signals $u(k)$ in advance, that is, it is possible to avoid sudden changes in control signal and undesired effects of delays in system response.

Standard formulation of criterion for MPC is as follows:

$$J = \sum_{k=0}^{N-1} q(k)(y(k) - y_r(k))^2 + r(k)u(k)^2, \tag{18}$$

where $q(k)$ is weight for difference between system output $y(k)$ and reference $y_r(k)$ at time instant k, while $r(k)$ is the weight of the displacement of control signal $u(k)$. If the future desired output value is known in advance, then this criterion leads to such an optimal system input, which minimizes weighted square of $y(k) - y_r(k)$. By

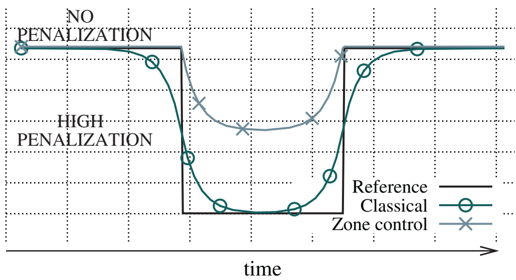


Fig. 3. Comparison between classical and zone predictive strategy. Weighting of entirely negative errors makes predictive controller to follow accurately the upper part of reference trajectory. When step down of desired value occurs, the system output drops to the reference value with a minimum control effort.

this, the area delimited by the system output below desired value is same as the area above it. This is depicted in Fig. 3 by line marked with circles. Such a behavior is satisfactory for most of the common control problems but not for temperature control of a building. The aim of the control is to adhere the upper desired value from its beginning to its end. Resulting behavior of the output is delineated in Fig. 3 by line with crosses.

This unusual problem can be solved by several approaches:

- The intuitive method is to use dynamic weights $q(k)$ and $r(k)$ at time, i.e. to make them time-dependant. These weights then depend on the shape of the reference trajectory – if there is a step-up/down on a prediction horizon, then weight $q(k)$ is set to be greater than $r(k)$ for k when the reference trajectory is on upper level, whilst $q(k) < r(k)$ for the rest of k on prediction horizon. This simple procedure ends if there exists more reference trajectory levels than two (but in this case is the best way how to solve such a problem).
- The second approach is as follows: In the minimization of the criterion (18) the reference trajectory y_r can be substituted with "artificial" reference w, which can be some approximation from the actual output y to real reference y_r. This can be done using following convex combination [30]:

$$\begin{aligned} w(k) &= y(k) \\ w(k+i) &= \alpha w(k+i-1) + (1-\alpha) y_r(k+i), \end{aligned} \tag{19}$$

  where $i = 1 \ldots N$ and $\alpha \in \langle 0;1\rangle$ is a parameter, that determines the smoothness (and speed) of the approaching of the real output to the real reference. (19) can be also restated as follows:

$$\begin{aligned} w(k) &= y(k) \\ w(k+i) &= \alpha r(k+i) - \alpha^i (y(k) - y_r(k)). \end{aligned} \tag{20}$$

  Making use of artificial reference may be very helpful in the case of number of "steps" in reference trajectory with need of its precise tracking by the actual output.
- Completely different way is to reformulate the part of criterion (18), which refer to the desired value error. If $y(k) < y_r(k)$ then weight the square of this difference using $q(k)$, otherwise the error is not weighted. This can be treaded by using the concept of zone control (also called funnel MPC) [18] where the reference error is not weighted in a specified interval while the weighting out is made in a common way. The lower bound of the interval is in our case desired value, whilst the upper bound is not specified due to the fact, that the building naturally tends to underheat providing the weighted output. Such a method can be used for tracking of reference trajectory with arbitrary number of levels.

The last approach will be discussed in detail.

3.4.2. MPC problem formulation

For a given linear, time invariant, discrete-time deterministic model

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) + Du(k) \end{aligned} \tag{21}$$

find the optimal control sequence on the horizon of prediction (length N) by minimizing the objective function

$$J = \sum_{k=0}^{N-1} q(k)(y(k) - z(k))^2 + r(k)u(k)^2, \tag{22}$$

subject to

$$\begin{aligned} u_{\min} \le \quad & u(k) \quad \le u_{\max} \\ y_r(k) \le \quad & z(k) \\ \Delta_{\max} \ge \quad & |u(k) - u(k-1)| \end{aligned} \tag{23}$$

where constraints $u_{\min}$, $u_{\max}$ are the minimum and maximum values of the control signal, $y_r(k)$ is desired value, thus lower bound for $z(k)$ and $\Delta_{\max}$ is a maximum rate of change of the control signal.

The objective function J (in (22)) can be rewritten into a matrix form (denoted without specification of a time instant)

$$J = (y - z)^T Q (y - z) + u^T R u, \tag{24}$$

where Q and R are weighting matrices of output error and control effort, respectively. The trajectory of the output is given as:

$$\begin{aligned} \begin{bmatrix} y(0) \\ y(1) \\ \vdots \\ y(N-1) \end{bmatrix} &= \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{N-1} \end{bmatrix} x(0) \\ &+ \begin{bmatrix} D & & & \\ CB & D & & \\ \vdots & & \ddots & \\ CA^{N-2}B & \ldots & CB & D \end{bmatrix} \begin{bmatrix} u(0) \\ u(1) \\ \vdots \\ u(N-1) \end{bmatrix}, \end{aligned} \tag{25}$$

i.e.

$$y = \Gamma x(0) + Hu, \tag{26}$$

where Γ is extended observability matrix and H is a matrix of impulse responses. Let $\tilde{y} = \Gamma x(0)$, then using (26), we can rewrite (24) as follows:

$$J = (\tilde{y} + Hu - z)^T Q (\tilde{y} + Hu - z) + u^T R u. \tag{27}$$

Minimization of such an objective function is a nonlinear programming problem, which can be readily rewritten into quadratic programming problem

$$\begin{aligned} J &= \begin{bmatrix} u^T & z^T \end{bmatrix} \begin{bmatrix} H^T Q H + R & -H^T Q \\ -QH & Q \end{bmatrix} \begin{bmatrix} u \\ z \end{bmatrix} + \\ &+ 2 \begin{bmatrix} \tilde{y}^T Q H & -\tilde{y}^T Q \end{bmatrix} \begin{bmatrix} u \\ z \end{bmatrix} + \tilde{y}^T Q \tilde{y} \end{aligned} \tag{28}$$

This yields the optimization problem $\min_{u,z} J$, which can be effectively solved using some of the computer algebra systems. The resulting problem has $(m+p) \cdot T$ variables which is a greater dimension compared to the classical one (described by criterion (18)) with $m \cdot T$ variables, where m and p are number of inputs and outputs respectively.

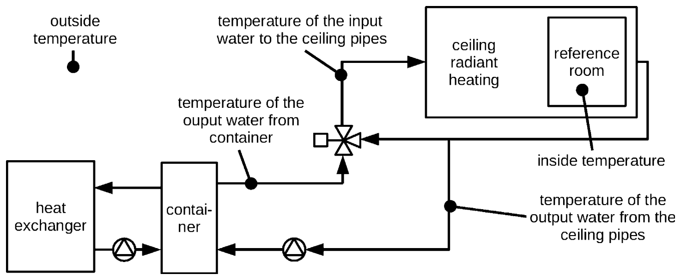


Fig. 4. Simplified scheme of the ceiling radiant heating system.

4. Application

The methods described in the previous sections were tested through December 2009 and January 2010 and the the real run of control application using proposed control strategy started in February 2010 at the building of the Czech Technical University in Prague. As of February 2010 the whole building consisting of 7 control blocks is controlled by presented MPC algorithm. All algorithms were implemented in Scilab.[3]

4.1. Description of the building

The building of the Czech Technical University in Prague uses a "Crittall" type ceiling radiant heating and cooling system. The "Crittall" system, invented in 1927 by R.G. Crittall and J.L. Musgrave [31], was a favorite heating system in the Czech Republic during 1960s for large buildings. In this system, the heating (or cooling) beams are embedded into the concrete ceiling. The control of individual rooms is very complicated due to the technical state of the control elements in all rooms. The control is therefore carried out for one entire building block, i.e. the same control effort is applied to all rooms of the building block. There are two (out of seven control blocks) building blocks with the same construction and orientation. Therefore, this situation is ideal for comparison of different control strategies, as depicted in Fig. 5.

A simplified scheme of the ceiling radiant heating system is illustrated in Fig. 4. The source of heat is a vapor–liquid heat exchanger, which supplies the heating water to the water container. A mixing occurs here, and the water is supplied to the respective heating circuits. An accurate temperature control of the heating water for respective circuits is achieved by a three-port valve with a servo drive. The heating water is then supplied to the respective ceiling beams. There is one measurement point in a reference room for every circuit. The setpoint of the control valve is therefore the control variable for the ceiling radiant heating system in each circuit.

4.2. Description of the model

The ceiling radiant heating system was modeled by a discrete-time linear time invariant stochastic model. We can consider this model as a Kalman filter giving an estimate of $\hat{x}(k)$ and $\hat{y}(k)$. Outside temperature prediction and heating water temperature were used as the model inputs. Prediction of outside temperature is composed of two values $T_{\max}$ and $T_{\min}$ defining a confidence interval. The outputs of the model are estimates of inside temperature $\hat{T}_{in}$ and of return water[4] $\hat{T}_{rw}$. This can be formalized according to (21) as

$$\begin{aligned} \hat{x}(k+1) &= A\hat{x}(k) + B\begin{bmatrix} T_{\min}(k) \\ T_{\max}(k) \\ T_{hw}(k) \end{bmatrix} + K(y(k) - C\hat{x}(k)) \\ \begin{bmatrix} \hat{T}_{in}(k) \\ \hat{T}_{rw}(k) \end{bmatrix} &= C\hat{x}(k) + D\begin{bmatrix} T_{\min}(k) \\ T_{\max}(k) \\ T_{hw}(k) \end{bmatrix}, \end{aligned} \tag{29}$$

where T_{hw} is temperature of the heating water and T_{in} denotes the inside temperature. System matrices A, B, C and D are to be identified using subspace methods as was described in Section 3.3.2. The state $\hat{x}(k)$ has no physical interpretation, when identified by means of the subspace identification. System order is determined by the identification algorithm. Modeling of the heating system of the CTU building is discussed in detail in Ref. [32].

4.3. Results

We have employed two methods of estimating the savings achieved on the building, based on comparison with a finely tuned weather-compensated controller (which also took weather forecast into account).

The first one was a cross-comparison of energy consumption in particular building blocks based on the difference between the heating and return water temperatures (this is directly proportional to the heat consumption provided that the pumps have a constant flow). In the period from mid-February to the end of the heating season (end of March), the overall savings reached 17–24%, depending on the particular building block.

The second method was based on comparison of calorimeter measurements for the entire building for MPC and said weather-compensated control. The measurements were normalized by outside temperatures and ambient temperature set-points to achieve reliable results. For said period of measurement, MPC achieved 29% savings according to this method.

It should be noted that the heating and return water temperature is being measured by standard industrial thermometers, which suffer from measurement errors, such as noise or offset. This introduces some uncertainty into the results. On the other hand, the calorimeters are installed by the heat provider, so we expect them to be well calibrated (or, at least, they do not measure less than the actual heat); heat payments are also based on the calorimeters. So in the terms of finances, the money savings of were also 29% (there is a flat rate on heat for the building).

Measurement of thermal comfort is always difficult and highly individual. As there are some 1500 employees and 8000 students in the building and there are always some people who complain about the ambient temperature, we decided to take the number of complains as the thermal comfort measure. To achieve objective results, the building occupants were not told about the new heating strategy. Under such conditions, the change in the number of complains was insignificant during the test period.

The results are depicted in Fig. 5. The upper part shows outside temperature, whilst the lower compares reference tracking for weather-compensated and predictive controllers. It can be seen, that the predictive controller heats in advance in order to perform optimal reference tracking, that is, inside comfort, and minimum energy consumption. Two last subfigures compare the efficiency

[3] Open source scientific software package for numerical computations (http://www.scilab.org/).

[4] It is crucial to model return water as an output because it gives a significant information about energy accumulated in the building, moreover it represents the interconnection between heating water and room temperature. Omitting the return water would lead to significant lost of information.

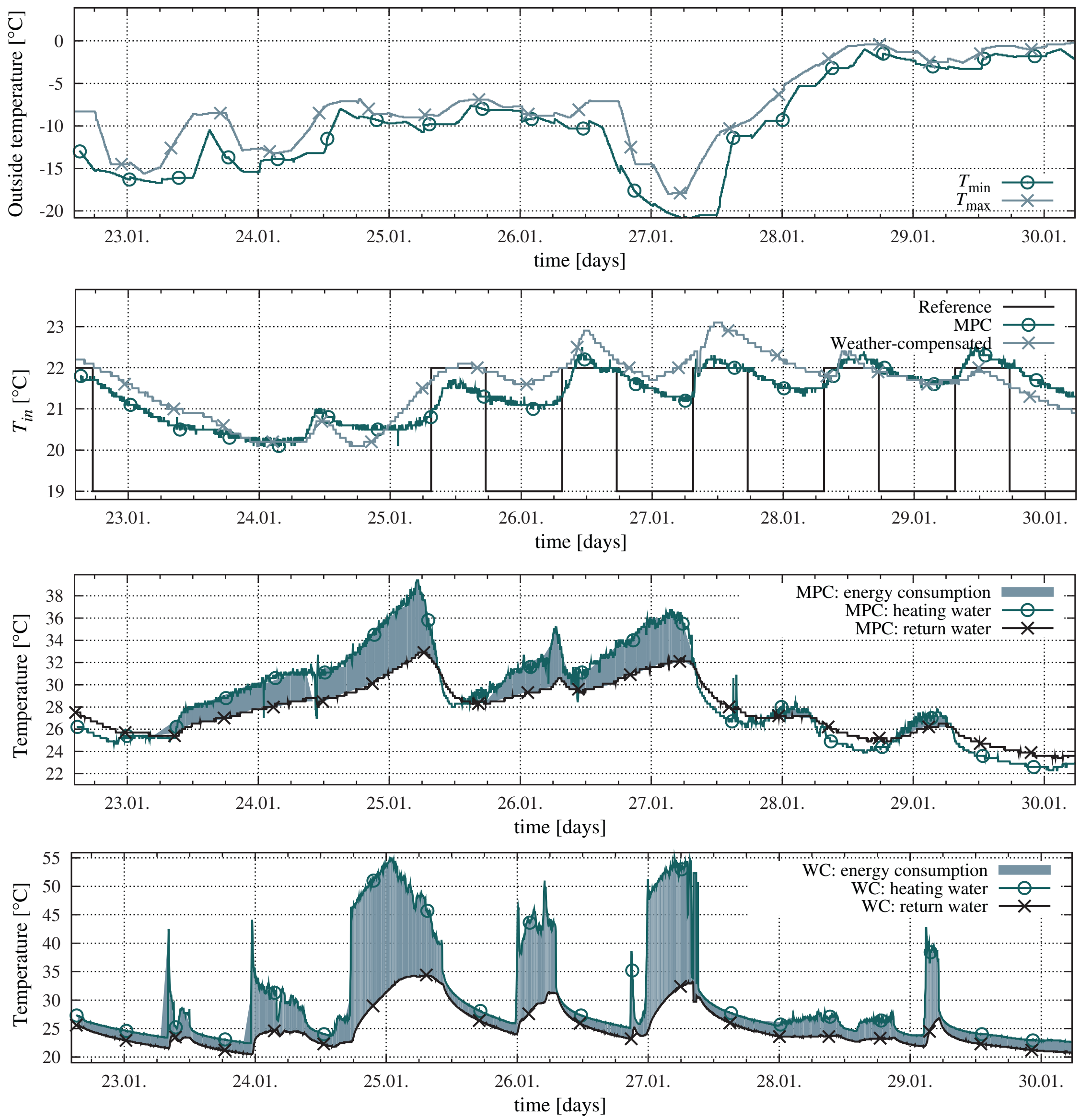


Fig. 5. Different control strategies: comparison of weather-compensated (WC) and predictive control (MPC) of heating water temperature and the room temperature controlled by MPC.

of control measured by energy consumption. The efficiency of the predictive control was superior to the weather-compensated controller, even if the active heating was necessary.

As mentioned before, the building has up to 12 hour heating delay. During weekends, the building cools down and classical heating has to be launched approximately one day before Monday 8 am, depending on the outside temperature.

5. Remarks to future development

Subspace identification methods represent black-box approach to the system modeling. This, alongside with its advantages carries also some drawbacks:

- The system might not be excited enough [22], i.e. the input of the system does not excite the system on satisfactory number of frequencies, thus identification algorithms lack considerable amount of information.
- User may have knowledge of some key feature or characteristics of the physical essence of the system, which is “lost” in the number of data.
- Natural character of the data might pose considerable statistical problem.

One of the most important aspects of the identification is the persistency of the excitation or the excitation itself. Data gathered from the measurement lack some important physical characteristics of the building. One of the possible approaches how to deal with this weak point is generation of artificial data that already contains desired properties. There is also another possibility, more expensive though—specially proposed experiment. It was decided to perform an experiment on real building in through late December 2009 and early January 2010. The comparison of model identification results is depicted in Fig. 6.

It is obvious, that experimental data significantly improved the identification fit. Yet another approach (and much cheaper) how to

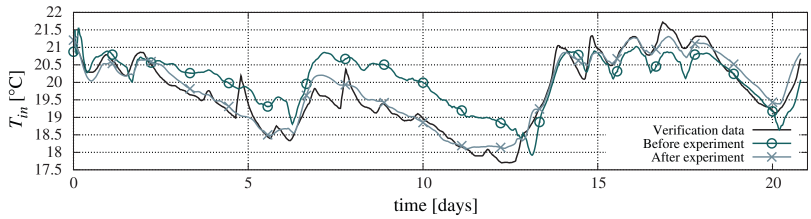


Fig. 6. Validation of model identified from data before and after experiment.

deal with lack of data quality is prior information and its incorporation to the subspace algorithm. Current methods [33] proposed algorithm how to incorporate PI into the algorithm using Bayesian framework. This algorithm makes use of Structured Weighted Lower Rank Approximation (SWLRA) [34] to decompose the projection matrix in order to save special structure, thus keep PI. However, this approach is able to deal with single input single output (SISO) and single input multiple output (MISO) systems only.

Future development of the identification algorithm will try to remedy the above-mentioned problems. Speaking generally, there several approaches to this problem:

- *Bayesian framework*. This approach requires extension to SWLRA algorithm to effectively solve MIMO systems.
- *Incorporation of PI into subspace algorithm*. This approach requires such an computation in subspace identification procedure which enables direct incorporation of PI into system matrices. This approach is the topic of ongoing research.
- *Spectral identification methods*. In robust control, analysis in frequency domain is very popular. The prior information could be incorporated by means of user-defined “filters”. This methodology is also topic of current research.
- *Artificial data*. Generation of data with desired properties is yet another approach. The user incorporates required properties and the knowledge of the physical essence into artificial data which are then used for regular identification. This approach, however, does not explicitly say, how to choose the ratio between artificial and measured data and, therefore, it is only of experimental nature.

In this paper, we treated only predictions of outside temperature because it has dominant influence out of all disturbances affecting the inside temperature. There are, however, other energy sources (like sun intensity, occupancy of the building, etc.). Taking them into account would provide better MPC performance as well as further savings.

6. Conclusion

Predictive control proved to have a great potential in the area of building heating control. The results from real operation on a large university building are very promissing and proved the supremacy of predictive controller over a well tuned weather-compensated control, with the savings of 17–24%. The MPC implementation discussed in the present paper is able to track the desired temperature very accurately, thus maintaining the heating comfort of the building.

However, the MPC strategy requires some extra effort. The crucial part of the controller is the mathematical model of the building. This is not possible by traditional system identification techniques based on statistical identification, as the building data usually do not have the desired statistical properties. On the other hand, finding first principle models is time consuming and not suitable for commercial application. We have shown that a proper identification experiment can provide data suitable for statistical identification, with the help of certain modifications of the standard identification algorithms. Numerical issues of the identification process must be treated very carefully, especially for large-scale systems.

Fortunately, once an appropriate model is found, the MPC tuning is very intuitive and desired properties of the control system can be achieved in a short term. The energy peaks are reduced and the controller does not make fast changes to the control input of the system, which also saves the lifetime of the equipment and reduces the peak energy demands. If desired, it also enables to take different energy prices into account by introducing time-variable tuning parameters into the optimization criterion.

Finally, the decision whether to implement the MPC or not depends largely on the return time of the investments. Even though this largely depends on air temperatures and sunshine during the heating season, the return time for our building is estimated to 2 years. As the identification effort does not really depend on the size of the building, this time will be shorter for large buildings with expensive heating and longer for small buildings with cheap heating.

Acknowledgment

This work has been supported from the state budget of the Czech Republic, through the Ministry of Industry and Commerce, in the scope of grant no. FR-TI1/517, “*Control systems for energy consumption optimization in low-energy and passive houses*.”.

References

[1] Buildings and their Impact on the Environment: A Statistical Summary, 2009. rev. April 22, 2009, available on-line at http://www.epa.gov/greenbuilding/pubs/gbstats.pdf.

[2] H.J. Han, Y.I. Jeon, S.H. Lim, W.W. Kim, K. Chen, New developments in illumination, heating and cooling technologies for energy-efficient buildings, Energy 35 (6) (2010) 2647–2653.

[3] J. Wong, H. Li, S. Wang, Intelligent building research: a review, Automation in Construction 14 (1) (2005) 143–159, doi:10.1016/j.autcon.2004.06.001.

[4] L. Peeters, J. Van der Veken, H. Hens, L. Helsen, W. D'haeseleer, Control of heating systems in residential buildings: current practice, Energy and Buildings 40 (8) (2008) 1446–1455, doi:10.1016/j.enbuild.2008.02.016.

[5] S. Soyguder, M. Karakose, H. Alli, Design and simulation of self-tuning pid-type fuzzy adaptive control for an expert hvac system, Expert Systems with Applications 36 (3 Part 1) (2009) 4566–4573, doi:10.1016/j.eswa.2008.05.031.

[6] J. Bai, S. Wang, X. Zhang, Development of an adaptive smith predictor-based self-tuning pi controller for an hvac system in a test room, Energy and Buildings 40 (12) (2008) 2244–2252, doi:10.1016/j. enbuild.2008.07.002.

[7] C. Ghiaus, I. Hazyuk, Calculation of optimal thermal load of intermittently heated buildings, Energy and Buildings 42 (8) (2010) 1248–1258, doi:10.1016/j.enbuild.2010.02.017.

[8] C.P. Underwood, HVAC Control Systems: Modelling, Analysis and Design, E & FN Spon, Great Britain, London, 1999.

[9] R.Z. Freire, G.H. Oliveira, N. Mendes, Predictive controllers for thermal comfort optimization and energy savings, Energy and Buildings 40 (7) (2008) 1353–1365.

[10] P.-D. Morosan, R. Bourdais, D. Dumur, J. Buisson, Building temperature regulation using a distributed model predictive control, Energy and Buildings 42 (9) (2010) 1445–1452, doi:10.1016/j.enbuild.2010.03.014.

[11] J. Richalet, A. Rault, J.L. Testud, J. Papon, Algoritmic control of industrial process, in: Proceedings: Symposium on Identification and System Parameter Estimation, IFAC, Tbilisi, 1976.

[12] W.H. Kwon, A.M. Bruckstein, T. Kailath, Stabilizing state feedback design via the moving horizon method, International Journal of Control 37 (1983) 631–643.

[13] J.B. Rawlings, J.R. Muske, The Stability of Constrained Receding Horizon Control, IEEE Transaction on Automatic Control 38 (10) (1993) 1512–1516.
[14] A. Zheng, M. Morari, Stability of Model Predictive Control with Soft Constraints, Internal Report, Californian Institute of Technology, 1994.
[15] P.J. Campo, M. Morari, Robust model predictive control, in: American Control Conference, 1987.
[16] K. Ang, G. Chong, Y. Li, PID control system analysis, design, and technology, IEEE Transactions on Control Systems Technology 13 (4) (2005) 559–576, doi:10.1109/TCST.2005.847331.
[17] Y. Li, K. Ang, C. Chong, PID control system analysis and design—problems, remedies, and future directions, IEEE Control Systems Magazine 26 (1) (2006) 32–41.
[18] J.M. Maciejowski, Predictive Control with Constraints, Prentice Hall, 2002.
[19] T.Y. Chen, Application of adaptive predictive control to a floor heating system with a large thermal lag, Energy and Buildings 34 (1) (2002) 45–51.
[20] M. Gwerder, J. Tödtli, B. Lehmann, F. Renggli, V. Dorer, Control of thermally activated building systems, in: Proceedings: 9th REHVA World Congress Clima, 2007.
[21] Z. Liao, A. Dexter, An inferential model-based predictive control scheme for optimizing the operation of boilers in building space-heating systems, IEEE Transactions on Control Systems Technology 18 (5) (2010) 1092–1102, doi:10.1109/TCST.2009.2033667.
[22] L. Ljung, System Identification: Theory for User, Prentice-Hall, 1999.
[23] M. Verhaegen, V. Verdult, Filtering and System Identification: A Least Squares Approach, Cambridge University Press, 2007.
[24] T. Bohlin, Practical Grey Box Process Identification, Springer, 2006.
[25] V. Peterka, Bayesian system identification, Automatica 17 (1) (1981) 41–53.
[26] L. Ferkl, J. Široký, Ceiling radiant cooling: comparison of ARMAX and subspace identification modelling methods, Building and Environment 45 (1, Sp. Iss. SI) (2010) 205–212.
[27] P. Van Overschee, B. De Moor, Subspace Identification for Linear Systems, Kluwer Academic Publishers, 1999.
[28] H. Kwakernaak, R. Sivan, Linear Optimal Control Systems, John Wiley & Sons, 1972.
[29] P. Trnka, Subspace identification methods, Ph.D. Thesis, Czech Technical University in Prague, 2007.
[30] E.F. Camacho, C. Bordons, Model Predictive Control, Springer, London, 1999.
[31] R.G. Crittall, J.L. Musgrave, Heating and cooling of buildings, GB Patent No. 210880 (April 1927).
[32] J. Široký, L. Ferkl, Model of a hot water full circulation mode in a building, in: IEEE International Conference on Control and Automation. ICCA 2009, 2009, pp. 585–590, doi:10.1109/ICCA.2009.5410421.
[33] P. Trnka, V. Havlena, Subspace like identification incorporating prior information, Automatica 45 (4) (2009) 1086–1091.
[34] M. Schuermans, P. Lemmerling, S.V. Huffel, Block-row hankel weighted low rank approximation, Numerical Linear Algebra with Applications 13 (2006) 293–302.

Appendix J

Prívara, Váňa, Gyalistras, Cigler, Sagerschnig, Morari, Ferkl: Modeling and Identification of a Large Multi-Zone Office Building

PRÍVARA, S. – VÁŇA, Z. – GYALISTRAS, D. – CIGLER, J. – SAGERSCHNIG, C. – MORARI, M. – FERKL, L. Modeling and Identification of a Large Multi-Zone Office Building. In Proceedings of the 2011 IEEE International Conference on Control Applications, Denver, USA, 2011.

Invited lecture

Number of citations: 0

Modeling and Identification of a Large Multi-Zone Office Building

Samuel Prívara, Zdeněk Váňa, Dimitrios Gyalistras, Jiří Cigler, Carina Sagerschnig,
Manfred Morari, Lukáš Ferkl

Abstract— **Predictive control in buildings has undergone an intensive research in the past years. Model identification plays a central role in a predictive control approach. This paper presents a comprehensive study of modeling of a large multi-zone office building. Many of the common methods used for modeling of the buildings, such as a detailed modeling of the physical properties, RC modeling, etc., appeared to be unfeasible because of the complexity of the problem. Moreover, most of the research papers dealing with this topic presents identification (and control) of either a single-zone building, or a single building sub-system. On contrary, we proposed a novel approach combining a detailed modeling by a building-design software with a black-box subspace identification. The uniqueness of the presented approach is not only in the size of the problem, but also in the way of getting the model and interconnecting several computational and simulation tools.**

I. Introduction

Climate changes, diminishing world supplies of the "traditional" fuels, ecological as well as economical aspects are only some of the many factors of a huge effort of today to save energy. Besides significant focus on renewable energy sources broaden, the goals can be reached only if the energy consumption is optimized. As the buildings account for about 40% of total final energy consumption (and its amount has been increasing at a rate 0.5–5 % *per annum* in developed countries [1]), the efficient building climate control can significantly contribute to reduction of the power consumption as well as the greenhouse gas emissions. Energy savings with minimal additional cost can be achieved by improvement of building automation system (BAS), which can nowadays control both heating, ventilation and air conditioning (HVAC) systems and the blind positioning and lighting systems [2], [3].

One of the control strategies suitable for building automation is the *Model Predictive Control* (MPC); unfortunately, the modeling and identification is rather difficult and time-consuming, not only in MPC. The special requirement for MPC is that the model should be reasonably simple and have good prediction properties on the control-relevant frequency range (see e.g. [4], [5], [6], [7]). One approach is to use the first-principle models (see [8], [9]), which are often used on systems such as TRNSYS, EnergyPlus (EP), ESP-r, etc., but these models are not explicit and cannot be used for control directly. The alternative is to use statistically-based, i.e. data-driven models [10]; in this approach, problems with sufficient excitation of the system modes arise.

In this work, we combined the benefits of both above mentioned approaches. A physical model in a building simulation software was created, such that it describes the real building as close as possible. Then identification signals were fed into the simulation software to obtain the high-quality identification data, and consequently these were used for obtaining a suitable control-oriented model.

The main contribution of this paper is twofold: Firstly, it presents in a detail the unique two step modeling procedure (real building $\rightarrow$ EnergyPlus model $\rightarrow$ linear-time invariant model for control), secondly it handles set-up of a large variety of tools used in different communities to deal with a problem of extraordinary size.

This paper is structured as follows: In the following section we will describe the problem and introduce the basic setup. Section III deals with identification and modeling procedures, describes the tools and algorithms used for obtaining the model. Section IV provides the results of the presented approach. Section V concludes the paper.

II. Problem description and setup

A. Description of the building

The 20 000 m^2 office building has six floors above ground. For this study, the entire third floor (as depicted in Figure 1), which is representative for all office floors, was modeled. Based on usage, façade orientation and HVAC supply, the floor can be divided into 24 zones which are mutually interconnected. Most zones are used as open-space offices; for modeling reasons, single offices were always lumped to a bigger zone.

The total floor area of the simulation model is approx. 2 800 m^2. The façade of the building has a window-to-wall ratio of approx. 70 %. Façades to the atrium have a glazing ratio of approx. 50 %. Roughly 50 % of the windows have interior blinds, remaining blinds are in-between-glass blinds of double windows.

There are the following actuators installed in the building:

- Convectors: individual convector control is possible.
- Radiant ceiling panels for cooling and heating: for control purposes ceiling panels of the floor are grouped into 24 zones that are controlled independently of each other.

S. Prívara and J. Cigler are with Department of Control Engineering (DCE), Faculty of Electrical Engineering (FEE) of Czech Technical University (CTU) in Prague, Czech Republic and with Automatic Control Laboratory (ACL), Department of Information Technology and Electrical Engineering (DITEE), ETH Zurich, Switzerland samuel.privara@fel.cvut.cz

Zdeněk Váňa and Lukáš Ferkl are with DCE, FEE CTU in Prague, Czech Republic

Carina Sagerschnig is with Gruner AG, Gellertstrasse 55, CH-4020 Basel, Switzerland

Dimitrios Gyalistras and Manfred Morari are with Automatic Control Laboratory, DITEE, ETH Zurich, Switzerland

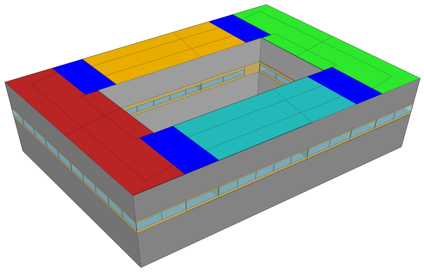


Fig. 1. 3D simulation model of the building: Investigated zones were on the third floor, other floors are greyed out. The zone layout is shown on top of the model for clarity. Zones of the same sub-system are colored alike. The core zones enabling the decoupling are dark blue.

- Ventilation: There are two air handling units (AHU)s – for the north, and the south. The temperature of supply air can be set independently in both AHUs.
- Venetian blinds are available for all windows in all zones. Controllable blinds of individual windows within the same zone are grouped together as one control input.

Energy supply, i.e. hot and chilled water supply for the entire building, is provided by a central heating and cooling plant, which is located partly in the basement and partly on the roof. District heating is used for the building's heat supply. Chilled water is provided locally by mechanical chillers.

B. Choice of a modeling strategy

As already stated, one of the objectives of this project was to find a convenient MPC-oriented modeling strategy suited for buildings, which would balance accuracy with the design-time demand.

The first possible approach is based on detailed physical modeling, represented by e.g. equivalent RC-network [11]. Unfortunately, fitting of parameters of differential equations is infeasible for large-scale problems [12].

The second approach is based purely on measurements collected during the building operation, which are used for input-output statistical identification. Even though this procedure looks simple, the results are oftentimes far from good – some important assumptions, such as persistent excitation [13], are nearly always violated during building's normal operation. The identification procedure can be improved by including some prior information [14] or by carrying out the identification experiment on the building which would excite all important system modes. Depending on the building size, the experiment might be rather expensive, but can bring high improvements to the resulting model [10].

When the building is brand new, real data are not available. From the aforementioned discussion, the only possible approach might be modeling of the building using RC network, which is quite challenging for multiple zone buildings.

Therefore a new approach had to be introduced, which yields a model of a large multi-zone building. A very promising strategy might be a combination of a building simulation software (to have an implicit model of the building) used for identification experiments to get data for a standard statistical identification procedure. We used EP as the building

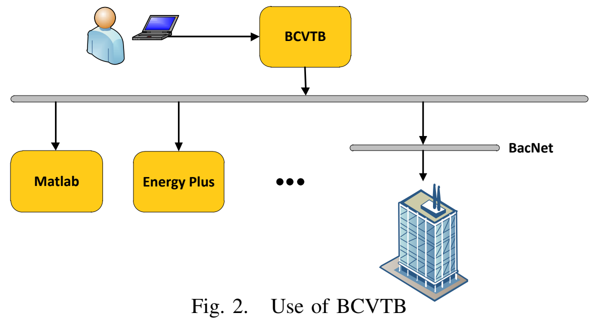


Fig. 2. Use of BCVTB

simulation software and Building Controls Virtual Test Bed (BCVTB) as the middleware between EP and a controller written in Matlab (in terms of excitation signal generator).

C. Software tools

A widely used tool for building energy performance simulation is EnergyPlus by the Lawrence Berkeley National Laboratory, which can be used for thermal load simulation and energy analysis of buildings. Besides the simulation itself, EP has a built-in energy management system that allows for integration of a rule-based control. Traditionally, EP is a stand-alone simulation engine which processes text-based input files. For developing and testing MPC models, co-simulation was necessary to allow more flexible simulation input. Co-simulation describes the integration of different tools by run-time coupling. This allows for example to couple building energy performance simulation tools to Matlab, and thus provide new possibilities to building simulation. Co-simulation fundamentals for building simulation such as coupling strategies and data transfer are described in [15]. Extensive capabilities for coupling simulation tools are provided by the Building Controls Virtual Testbed (BCVTB) [16]. It is a middleware tool that allows to couple different simulation programs for distributed simulation. Programs to be linked via the BCVTB are EP, Matlab, Modelica and Radiance. Data exchange with BACnet building automation systems is also featured. The BCVTB plays a master role in the data exchange, as depicted in Figure 2. For the entire simulation study, hourly weather data for Munich were used. The statistical weather data used were provided by the weather database of the US Department of Energy and prepared by ASHRAE[1] based on International Weather for Energy Calculations (IWEC) data.

III. Identification and Modeling

As was mentioned in previous sections, the identification and modeling is one of the most demanding tasks. We will describe the whole procedure of getting a building model in the following steps. Firstly, we describe the choice of suitable inputs and outputs for the identification, then we present software tools needed for handling and keeping all information about system consistent, and finally mathematical tools necessary for successful system identification (SID).

[1] American Society of Heating, Refrigerating and Air-Conditioning Engineers

A. Choice of model inputs and outputs

The choice of model inputs and outputs plays an important role for the particular identification procedure. According to the physical relationship between chosen inputs and outputs, one should opt for a suitable procedure which is able to handle underlying physics. In other words: if the input-output relation is non-linear, then linear identification methods may fail. The size of the problem is also quite an important factor – especially in the presented problem.

Based on the aforementioned observations, we decided to choose heat fluxes affecting zone temperatures as system inputs and temperatures as outputs. The key benefit is that underlying physics is linear. Complete sets of inputs and outputs are described in Table I. Note that the model inputs are different from the inputs of the detailed EP model – direct manipulation of some heat fluxes is not allowed, and therefore we have to provide signals on a lower level (see Table I). The input set was divided into two categories: the first group represents the actuator heat fluxes, whilst the second represents disturbances affecting the system. The identification procedure does not distinguish between disturbances and manipulated variables, however, is needed for user orientation and consequent control as well.

B. Step-by-Step to get a model

Each of the following steps is actually a stand-alone software package which enables a specific task as follows.

1) GenEI: The main task of GenEI is a generation of sufficiently exciting input signals. Such signals are needed to satisfy key theoretical assumptions on reliable statistical identification – persistent exciting signals [17]. In real operation, this request is almost infeasible due to technical, physical or economical limitations. As the image of the building modeled in EP is at hand, a proper identification experiment can be designed. Obviously, when the objective is to build-up of a model suitable for control, the generated inputs do not need to cover the entire frequency domain, but rather some control-relevant selection. The prior knowledge of the time constants of the system is often known or at least possible to estimate using some preliminary tests, thus the input signal is generated according to this information. We have proposed three different kinds of input signals, pseudo-random binary signal (PRBS), sum of sinusoids (SINE) and multilevel pseudo-random signal (MPRS). All of them have similar settings as follows. Let τ_H, τ_L denote the slowest and the fastest systems time constants, respectively. Then the required frequency spectrum to be covered by the generated signal is (ω_*, ω^*) and the following equation holds:

$$\omega_* = \frac{1}{\beta\tau_H} \leq \omega \leq \frac{\alpha}{\tau_L}\omega^*, \tag{1}$$

where α defines the ratio of closed-and-open loop responses and β specifies the settling time. Typical values are $\alpha = 2$ and $\beta = 3$, which corresponds to 95 % of settling time [18]. Due to the Nyquist-Shannon-Kotelnikov theorem, frequency range of the generated signal cannot be as in (1), but must be larger, and the range (1) should bear majority of the power

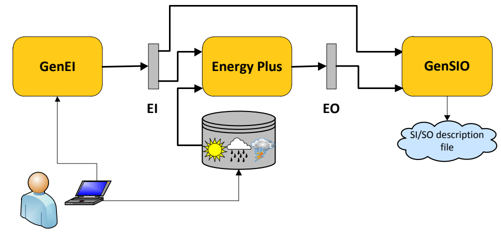


Fig. 3. Preparation of data for identification.

of the signal. Furthermore, the choice of switching time is based on $T_s \leq \frac{2.78}{\omega^*}$.

In case of MPRS, the input sequence is computed by Galoise fields [18] with the number of shift registers n and the length q, which defines the maximum possible multiple of harmonics to be suppressed. In the opposite way, let h be the maximum possible multiple of the harmonics to be suppressed. Then q has to be chosen such that $q \geq 2^h - 1$. Next, the length n can be computed according to $\omega_* \geq \frac{2\pi}{T_s(q^n-1)}$. Further on, the length of a signal cycle is $N_{cyc} = q^n - 1$, which, in time domain, represents a signal of duration $T_{cyc} = N_{cyc} \cdot T_s$. The number m of the signals to be generated has to be considered as well, but in practical applications, it is sufficient to generate only 1 signal, and shift it $(m - 1)$ times afterwards. It is indeed a very suitable solution as the signal generation is time consuming. Moreover, this technique guarantees the sufficient lack of cross-correlation between the respective signals [19].

2) GenSIO.: This block processes outputs produced by GenEI (inputs to EP), outputs of EP and some variables from schedules and databases, and produces the input and output data sets used in SID. The respective inputs, outputs and disturbances, as used in identification, are described in Table I and the procedure of data generation and preparation for identification is schematically depicted in Figure 3 and Figure 4, respectively.

3) Splitter.: Even the powerful servers (64bit machines, 16 cores @ 2.6 GHz and 24 GB RAM) are not able to compute the identification procedure because of the size of the problem. However, due to the floor layout, façade orientation and floor usage, the model of the third floor can be looked at as four decoupled subsystems. The ring-shaped layout of the floor houses also four cores (hosting infrastructural supply such as elevators, staircases, etc.) which separate the office spaces from each other. Thermal coupling of the investigated office zones via the cores is very loose and can be neglected, as far as control issues are concerned. The distributed heating, cooling and ventilation supply of the zones also support the idea of the system division said four subsystems. Consequently, each of the subsystems contains its zone-relevant signals and a copy of the signal, which was originally common for all subsystems. In other words, when a satisfactory computational power is at hand, the proposed procedure does not request special knowledge of the building's physics, etc. On the other hand, it does not exclude the possibility of the system division due to computational or other reasons.

TABLE I

NOTATION OF THE VARIABLES USED FOR SYSTEM IDENTIFICATION

ID	Variable Category	Type	Zone relevant	EP equivalent
Q_{CONV}	Convector heating rate	Input	Yes	Same quantity, power can be arbitrarily set within limits
ZCPCR	Zone ceiling panel cooling rate	Input	Yes	Supply water temperature and mass flow rate through plumbing can be adjusted. Together with return water temperature, they stand for heat flux of radiant ceiling
ZCPHR	Zone ceiling panel heating rate	Input	Yes	Same as ZCPCR
LG	Lighting gains	Input	Yes	Same quantity, power can be arbitrarily set within limits
NRF	Net radiation flux	Disturbance	Yes	Partly by means of blinds control
FP	Fan power	Input	Yes	Air flow rate (which is either 55 or 0 m^3/h) and supply air temperature. Together with return air temperature, they stand for heat flux of fans.
ODBT	Outdoor dry bulb temperature	Disturbance	No	Same quantity
EG	Equipment gains	Disturbance	Yes	Same quantity
OG	Occupancy gains	Disturbance	Yes	Same quantity
ZT	Zone temperature	Output	Yes	Same quantity
ZI	Zone interior illuminance	Output	Yes	Same quantity

4) SID.: The choice of the identification method was determined by the factors described in previous sections, namely the size of the problem with a vast number of inputs and outputs, implying the multiple input multiple output (MIMO) system, and on the other hand, a huge set of generated (and/or measured) data suggesting the use of statistical identification procedures. Two different choices for subspace identification algorithm were implemented

- `N4SID` function from System Identification toolbox for Matlab, (see [20], [13]).
- Combined deterministic-stochastic algorithm [17], [21].

5) Joiner.: The four resulting subsystems are merged together, when all the subsystems retain their zone-specific signals, whilst the common signals are joined.

6) Verification and validation.: Each of the identified system was verified using the verification data sets and residual and correlation analyzes. The joined system was repeatedly verified.

7) LTI2MPC.: For optimization requirements, there are several variables added to the model, e.g. total Energy Power Demand (totEPD[2]) or total Heat Power Demand (totHPD[3]). For this reason, the model provided by Joiner must be transformed according to control requirements. This transformation is actually determined by the MPC variant, e.g. optimization objective. Furthermore, for purposes of predictive optimization (cost function and particular bounds), the B and D matrices must be split for the deterministic (manipulated variables) and stochastic (disturbance variables) counterparts, as the SID identified all the inputs (no matter the deterministic and stochastic parts) together.

IV. IDENTIFICATION RESULTS

Because of the size of the system matrices, we will omit them here and show only the resulting model validations, which were carried out using comparison of k-step ahead predictions, as well as by the analysis of the system structure. The first identification attempts, which seem to be straightforward, were to excite the system by generated SINE signals with $\tau_L = 60$ and $\tau_H = 240$ minutes, which is sufficient in

[2]totEPD is a sum of lighting and equipment gains and ceiling cooling

[3]totHPD is a sum of convectors' heating rate and a ceiling heating

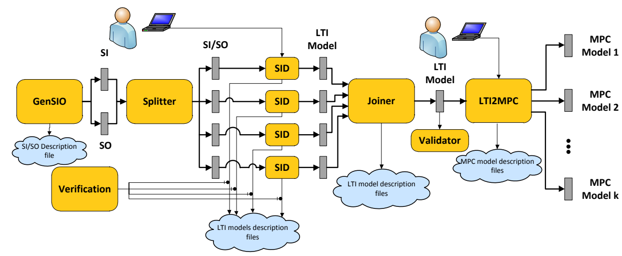


Fig. 4. System identification procedure

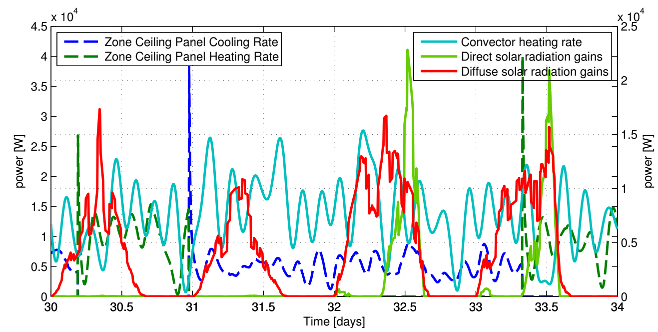


Fig. 5. Part of model inputs.

sense of building dynamics. Since EP needs different input signals than our identified model (Table I), not all model inputs are able to excite the system in an arbitrary way (see Figure 5 and Figure 6). Anyway, these data are still excited enough to describe the system behavior well. This statement can be deduced from the response of the identified model to the verification data set. The part of output data corresponding to the time axes of Figure 5 and Figure 6 is depicted in Figure 7. The model, or to be more specific "1 of the 4 submodels", has an order around 20. This value depends on the type of input excitation, identification data length, time period of the year for which the identification is computed, focus on either simulation or prediction, and on the users choice (since the data is disturbed by noises). Even after joining partial submodels into one big model, the verification response stays great not only for 1-step ahead prediction (Kalman filtering), but for longer predictions as well – see comparison for all zones in Figure 8.

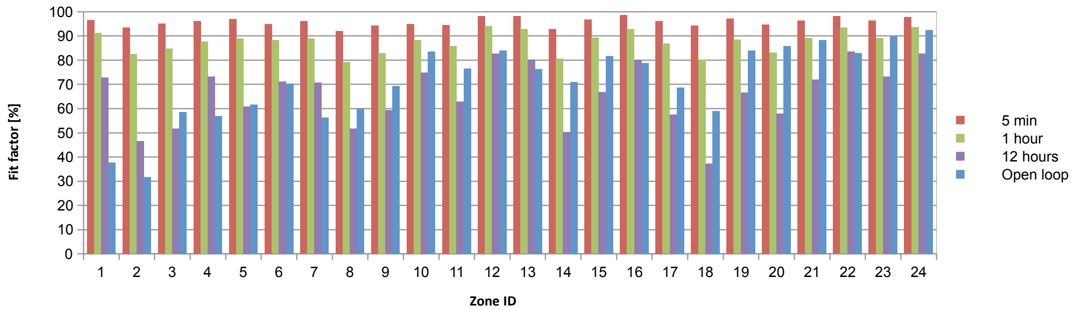
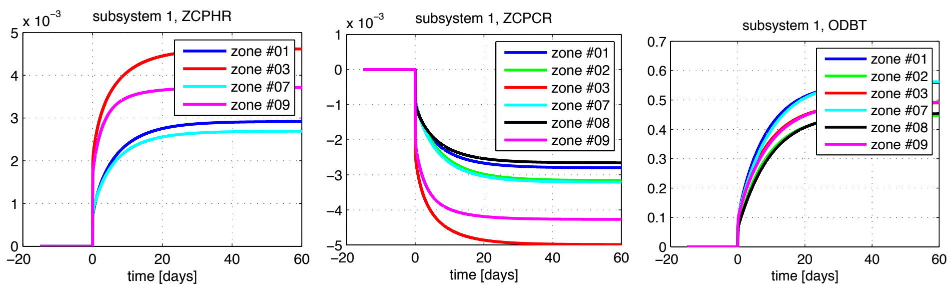


Fig. 8. Fit-factors for all zones for different k-step ahead predictions



(a) Zone ceiling panel heating rate (b) Zone ceiling panel cooling rate (c) Outdoor dry bulb temperature

Fig. 9. Step responses from a subset of inputs to particular zone temperature (inputs and outputs are paired by zone name)

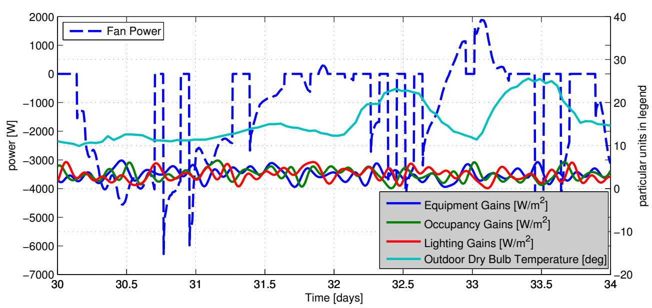


Fig. 6. Part of model disturbances.

However the system response to verification data is nice, a significant (and surprising) drawback of the statistical identification has come out. The step responses did not satisfy our expectations in both DC gains and signs. To ensure not only good verification response, but step response as well, it was necessary to reshape the experiment inputs. So far, all inputs have been excited in parallel, and therefore partial zones in one subsystem could affect each other. Thus for the new identification, the experiment inputs for one subsystem have been torn in time in such a way, that at any time instant, only one input category in only one zone is excited; all other inputs are set to a constant "stand-by" value. "Stand-by" values have been selected to ensure no active heating nor cooling into the building, but only natural behavior. Improvement of the resulting system structure was significant, and the resulting model is now valid also from physical point of view, which can be illustrated by step responses from subset of system inputs to zone temperatures in Figure 9. All the zones have correct step response dynamics as well as sign. The step responses from a specific energy source or outdoor temperature do not have the same impact for all zones (each zone has different size, orientation and equipment), but should be similar:

- Ceiling heating rates (see Figure 9(a)) present correct structure with an appropriate impact of energy sources – the larger the zone is, the smaller the temperature impact of 1 W of input signal.
- Ceiling cooling (see Figure 9(b)) has, in all cases, correct sign of the step response (positive power demand should affect the zone negatively).
- From Figure 9(c), we can see quite a high impact of the outdoor temperature on the zone temperatures. It has of course, slower dynamics than ceiling panels shown in Figure 9(b) and Figure 9(a), respectively.

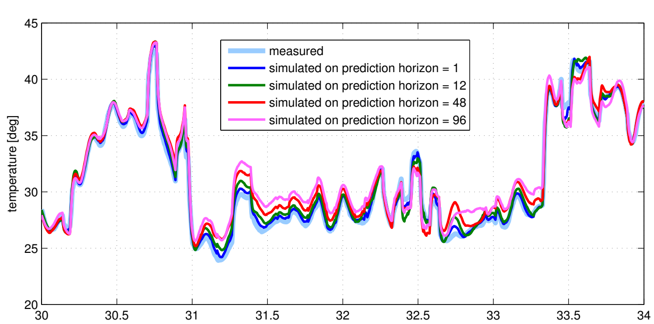


Fig. 7. Part of model outputs.

V. Conclusions and and future development

This paper has introduced a new methodology of interconnecting building simulation software and traditional identification methods in order to avoid the statistical problems with data gathered from the real building. The building was modeled using EnergyPlus, which was excited by specially proposed signals to get data of a good quality. Then the subspace identification approach (with some modifications) was applied to acquire a model suitable for predictive control. To the authors' best knowledge, there was no detailed building modeling intended for predictive control of such a size. The last step of preparation of the model for control are adjustments of inputs and outputs, in order to obtain a model corresponding to the variety of MPC problems (according to the control criteria).

Acknowledgments

The work was supported from the state budget of the Czech Republic, through the Ministry of industry and commerce, in the scope of grant 5.1 SPK01/024 ENERGOKLASTR.

References

[1] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.
[2] J. Wong, H. Li, and S. Wang, "Intelligent building research: a review," *Automation in Construction*, vol. 14, no. 1, pp. 143–159, jan 2005.
[3] H. Han, Y. Jeon, S. Lim, W. Kim, and K. Chen, "New developments in illumination, heating and cooling technologies for energy-efficient buildings," *Energy*, 2009.
[4] H. Hjalmarsson, "System identification of complex and structured systems," *European journal of control*, vol. 15.
[5] R. Gopaluni, R. Patwardhan, and S. Shah, "MPC relevant identification–tuning the noise model," *Journal of Process Control*, vol. 14, no. 6, pp. 699–714, 2004.
[6] D. Lauri, J. Salcedo, S. Garcia-Nieto, and M. Martinez, "Model predictive control relevant identification: multiple input multiple output against multiple input single output," *Control Theory & Applications, IET*, vol. 4, no. 9, pp. 1756–1766, 2010.
[7] D. Shook, C. Mohtadi, and S. Shah, "A control-relevant identification strategy for GPC," *IEEE Transactions on Automatic Control*, vol. 37, no. 7, pp. 975–980, 2002.
[8] H. Madsen and J. Holst, "Estimation of continuous-time models for the heat dynamics of a building," *Energy and Buildings*, vol. 22, no. 1, pp. 67–79, 1995.
[9] K. Andersen, H. Madsen, and L. Hansen, "Modelling the heat dynamics of a building using stochastic differential equations," *Energy and Buildings*, vol. 31, no. 1, pp. 13–24, 2000.
[10] J. Cigler and S. Prívara, "Subspace identification and model predictive control for buildings," in *The 11th International Conference on Control, Automation, Robotics and Vision – ICARCV2010*, pp. 750–755.
[11] M. Gwerder, B. Lehmann, J. Tödtli, V. Dorer, and F. Renggli, "Control of thermally-activated building systems (tabs)," *Applied energy*, vol. 85, no. 7, pp. 565–581, 2008.
[12] T. Bohlin and S. Graebe, "Issues in nonlinear stochastic grey box identification," *International Journal of Adaptive Control and Signal Processing*, vol. 9, no. 6, pp. 465–490, 2007.
[13] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA, 1999.
[14] S. Prívara, J. Cigler, Z. Váňa, L. Ferkl, and M. Šebek, "Subspace identification of poorly excited industrial systems," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 4405–4410.
[15] M. Trčka, J. Hensen, and M. Wetter, "Co-simulation of innovative integrated HVAC systems in buildings," *Journal of Building Performance Simulation*, vol. 2, no. 3, pp. 209–230, Sept. 2009.
[16] W. M. and H. P., "A modular building controls virtual test bed for the integration of heterogenous systems," in *Third National Conference of IBPSA-USA Berkeley/California*, 2008.
[17] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Pard, Nowell, Massachusetts: Kluwer Academic Publishers, 1999.
[18] M. Braun, D. Rivera, A. Stenman, W. Foslien, and C. Hrenya, "Multi-level pseudo-random signal design and model-on-demand estimation applied to nonlinear identification of a RTP wafer reactor," in *American Control Conference, 1999. Proceedings of the 1999*, vol. 3. IEEE, 2002, pp. 1573–1577.
[19] S. V. Gaikwad and D. E. Rivera, "Control-relevant input signal design for multivariable system identification: Application to high-purity distillation," 1996.
[20] P. Van Overschee and B. De Moor, "N4SID: Subspace algorithms for the identification of combined deterministic-stochastic systems," *Automatica*, vol. 30, no. 1, pp. 75–93, 1994.
[21] ——, "A unifying theorem for three subspace system identification algorithms," *Automatica*, vol. 31, no. 12, pp. 1853–1864, 1995.

Appendix K

Prívara, Váňa, Cigler, Ferkl: Subspace Identification: A Path Towards Large-Scale Predictive Controllers of Buildings

PRÍVARA, S. – VÁŇA, Z. – CIGLER, J. – FERKL, L. Subspace Identification: A Path Towards Large-Scale Predictive Controllers of Buildings. Submitted to the 50th IEEE Conference on Decision and Control, Orlando, USA, 2011.

Invited lecture

Number of citations: 0

Subspace Identification: A Path Towards Large-Scale Predictive Controllers of Buildings

Samuel Prívara, Zdeněk Váňa, Jiří Cigler, Lukáš Ferkl

***Abstract*— Even though the advanced control has emerged into numerous parts of our world, building automation is still a field where the position of the classical control is almost exclusive. The main reason is, that for the synthesis of an advanced controller, a decent model, "model for control", is needed. In this paper, a building model identification procedure is presented, wherein the building model is built-up as a first-principle model using a simulation software, and then a state-space model is identified by means of the subspace identification methods. The main focus of the paper is a case study of a large office building, and the entire process of its identification. The paper will show, that the model needs not to be *precise* at the whole frequency range, however, it is crucial to have good properties at frequencies where the control is intended. Moreover, the major difficulties encountered during the model identification process are described and possible solutions are outlined.**

I. Introduction

In recent years, there has been a significant effort to revert or at least diminish the effect of the climate changes or the climate changes themselves. Moreover, there is a permanent effort for energy savings in the most of the developed countries. In addition, the European Union (EU) presented targets concerning energy cuts defining goals until 2020 [1]: *i)* Reduction in EU greenhouse gas emissions at least 20 % below the 1990 levels *ii)* 20 % of EU energy consumption to come from renewable resources *iii)* 20 % reduction in primary energy use compared to projected levels, to be achieved by improving energy efficiency. Even though there are fierce debates about the real impact of the renewable resources, because of e.g. the power grid issues, energy savings achieved by smart control algorithms are free of political controversy.

As the buildings account for about 40% of total final energy consumption (and its amount has been increasing at a rate 0.5–5 % *per annum*)[2], an efficient building climate control can significantly contribute to reduction of the power consumption as well as the greenhouse gas emissions. Energy savings with minimal additional cost can be achieved by improvement of building automation system (BAS), which can nowadays control heating, ventilation and air conditioning (HVAC) systems, as well as the blind positioning and lighting systems [3], [4].

One of the control strategies suitable for building automation is the *Model Predictive Control* (MPC); unfortunately, the modeling and identification of buildings is rather difficult and time-consuming. The special requirement for MPC is that the model is reasonably simple and has good prediction properties on the control-relevant frequency range (see e.g. [5], [6], [7], [8]). One approach is to use the first-principles, however, most of the papers devoted to the modeling from the first principles provide only two-room example. We would like to model a large office building with tens of rooms (i.e. tens of inputs and outputs) which can be modeled in building simulation software such as TRNSYS, EnergyPlus (EP), ESP-r, etc. (see [9], [10], [11], [12]), but these models are not explicit and cannot be used for predictive control. Therefore we have decided to use statistically-based, i.e. data-driven models [13], [14].

In this work, we combined the benefits of both above mentioned approaches. A physical model in a building simulation software (EP) was created, such that it describes the real building as closely as possible. Then identification signals were fed into the simulation software to obtain the high-quality identification data, and consequently these were used for obtaining a suitable control-oriented model. The uniqueness of this approach is in the combination of real building data, the first-principle model of the building and the identification algorithms, which make use of both.

The next section provides the motivation and explains the background of the problem to be solved. Section II describes the setup of the building and formulates the problem. The modeling and identification are dealt with in Section III. Section IV describes the problems encountered during the identification procedure and suggest some solutions. Section V concludes the paper.

II. Problem Description and Setup

A. Description of the building

The analyses to be described all deal with the third floor of a large office building in Munich (20 000 m^2 and six above-ground floors, see Fig. 2 and Fig. 1). Based on the usage, façade orientation and HVAC supply, the floor can be divided into 24 mutually interconnected zones. The total floor area of the simulation model is approx. 2 800 m^2. The façade of the building has a window-to-wall ratio of approx. 70 %. Façades to the atrium have a glazing ratio of approx. 50 %. Roughly 50 % of the windows have interior blinds, remaining blinds are in-between-glass blinds of double windows.

The building automation system contains several actuators, namely individually controlled convectors, 24 independently controlled radiant ceiling panels for cooling and heating, two air handling units (AHU) for control of the ventilation, and venetian blinds for all windows in all zones. Energy supply,

S. Prívara, Z. Váňa, J. Cigler and L. Ferkl are with Department of Control Engineering (DCE), Faculty of Electrical Engineering (FEE) of Czech Technical University (CTU) in Prague, Technická 2, 166 27 Praha 6, Czech Republic

Fig. 1. Office building in Munich

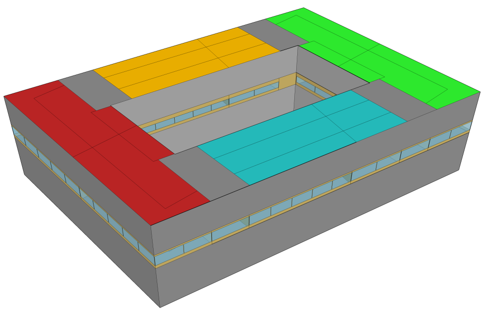


Fig. 2. 3D simulation model: Investigated zones were on the third floor, other floors are grayed out and used for shading purposes only. The zone layout is shown on top of the model for clarity. Zones of the same subsystem are colored alike. Core areas are grey.

i.e. hot and chilled water supply for the entire building, is provided by a central heating and cooling plant, which is located partly in the basement and partly on the roof. District heating is used for the building's heat supply. Chilled water is provided locally by mechanical chillers.

B. Choice of modeling strategy and model inputs and outputs

The choice of model inputs and outputs is strongly dependent on the choice of the identification method which is determined by the application. As was briefly mentioned before, the usage of the first-principle models (represented by e.g. an equivalent RC-network [15]) for a problem of such dimensions as one examined in this paper, is very limited [16] (if possible at all), and the well-known simulation software packages do not provide model in an explicit form suitable for control. Therefore, we have decided for statistical identification, namely *subspace identification methods* (4SID). Even though this procedure looks simple, the results are oftentimes far from good – some important assumptions, such as persistent excitation [17], are nearly always violated during building's normal operation. The identification procedure can be improved by including some prior information [18] or by carrying out the identification experiment on a building which would have all important system modes excited. Depending on the building size, the experiment might be rather expensive, but can bring high improvements to the resulting model [14], [13]. Therefore a new approach had to be introduced, which yields a model of a large multi-zone building. A very promising strategy might be a combination of a building simulation software (to have an implicit model of the building) used for identification experiments to get data for a standard statistical identification procedure. We used EnergyPlus as the building simulation software and Building Controls Virtual Test Bed (BCVTB) as the middleware between EnergyPlus and a controller written in Matlab (in terms of excitation signal generator). The description of the interconnection of the various simulation and computational tools is described more in detail in [19].

Based on the chosen identification method, 4SID, the model inputs and outptus were selected as follows. The heat fluxes affecting zone temperatures were selected as system inputs and temperatures as outputs. The key benefit is that the underlying physics is linear. The complete set of inputs and outputs is described in Table I. Note that the model inputs are different from the inputs on side of the detailed EnergyPlus model – direct manipulation of some heat fluxes is not allowed, and therefore, we have to provide signals on lower level (description of these are given also in Table I). Note also that the input set was divided into two categories: the former group represents the actuator heat fluxes and the latter disturbances affecting the system. Identification procedure does not distinguish between disturbances and manipulated variables, however, this categorizing is good for user orientation as well as consequent control.

III. Identification and Modeling

A. Problem statement

In the last two decades, the subspace algorithms (4SID) have become an important tool for the system identification (SID). The objective of the 4SID, as will be used further on, is to find a linear, time invariant, discrete time state space model in an innovative form

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ke(k) \\ y(k) &= Cx(k) + Du(k) + e(k), \end{aligned} \tag{1}$$

given the measurements of the input $u(k) \in \mathbb{R}^m$ and the output $y(k) \in \mathbb{R}^l$ with e being zero-mean white noise. In other words, we want to determine the system order n and to find the matrices A, B, C, D and K. The set of data

$$Z^N = (u(t), y(t))_{t=1}^N, \tag{2}$$

is generated by an unknown stochastic system of order n, which is equivalent to the well-known stochastic model as defined in e.g. [20], [21].

B. General algorithm

The entry point to the algorithm are input-output equations

$$\begin{aligned} Y_p &= \Gamma_i X_p + H_i^d U_p + H_i^s E_p \\ Y_f &= \Gamma_i X_f + H_i^d U_f + H_i^s E_f \\ X_f &= A^i X_p + \Delta_i^d U_p + \Delta_i^s E_p, \end{aligned} \tag{3}$$

where all the corresponding symbols are explained in Table II.

TABLE I

NOTATION OF THE VARIABLES USED FOR SYSTEM IDENTIFICATION

ID	Variable Category	Type	Zone relevant	EP equivalent
Q_{CONV}	Convector heating rate	Input	Yes	Same quantity, power can be arbitrary set within limits
ZCPCR	Zone ceiling panel cooling rate	Input	Yes	Supply water temperature and mass flow rate through pipes can be adjusted. Together with return water temperature, they stand for heat flux of radiant ceiling
ZCPHR	Zone ceiling panel heating rate	Input	Yes	Same as ZCPCR
LG	Lighting gains	Input	Yes	Same quantity, power can be arbitrary set within limits
DSRV	Direct solar radiation gains	Input	Yes	By means of blind control (position and angle), we can adjust solar gains influencing zone temperature.
DFSRV	Diffuse solar radiation gains	Input	Yes	Same as DSRV
FP	Fan power	Input	Yes	Air flow rate (which is either 55 or 0 m^3/h) and supply air temperature. Together with return air temperature, they stand for heat flux of fans.
ODBT	Outdoor dry bulb temperature	Disturbance	No	Same quantity
EG	Equipment gains	Disturbance	Yes	Same quantity
OG	Occupancy gains	Disturbance	Yes	Same quantity
ZT	Zone temperature	Output	Yes	Same quantity
ZI	Zone interior illuminance	Output	Yes	Same quantity

TABLE II

SYMBOLS AND THEIR MEANING USED FOR SID ALGORITHM.

Symbol	Meaning
Y_p	Hankel matrix of the past outputs
Y_f	Hankel matrix of the future outputs
X_f	Hankel matrix of the future states
U_p	Hankel matrix of the past inputs
U_f	Hankel matrix of the future inputs
Γ	Extended observability matrix
H^d	Markov parameter matrix (deterministic part (DP))
H^s	Markov parameter matrix (stochastic part (SP))
Δ^d	Reversed extended controllability matrix (DP)
Δ^s	Reversed extended controllability matrix (SP)
E_p	Hankel matrix of the past noise
E_f	Hankel matrix of the future noise
i	Prediction horizon
N	Number of data samples

The basic idea of the algorithm is to eliminate the input and the noise matrices in Eq. (3) by finding the appropriate projection and instruments matrices. The main tool of SID, an oblique projection, is defined (see [22], [23]) as follows:

$$\mathcal{O} = Y_f \begin{bmatrix} W_p^T & U_f^T \end{bmatrix} \begin{bmatrix} W_p W_p^T & W_p U_f^T \\ U_f W_p^T & U_f U_f^T \end{bmatrix}^{\dagger} \begin{bmatrix} I_{l\times l} \\ 0 \end{bmatrix} W_p, \quad (4)$$

where l is the number of outputs and $(\bullet)^{\dagger}$ is Moore-Penrose pseudo-inverse. Eq. (4) is in literature often referred to as $\mathcal{O} = Y_f \underset{U_f}{/} W_p$. Then it can be shown (see e.g. [23], [24]), that $\mathcal{O} = \Gamma X$, where X is a Kalman filter state sequence, i.e. the oblique projection is a tool how to get rid of the input and noise matrices in Eq. (3). The order of the system can be determined from analysis of the singular values σ_i obtained using a singular value decomposition (SVD) of $W_1 \mathcal{O} W_2$, where W_i are the weighting matrices of an appropriate size and determine the resulting state space basis as well as importance of the particular element of $\mathcal{O}$. A general formula for order determination does not exist, however, one should look for big changes in amplitude between two neighboring singular values. We have proposed the following heuristic formula for automatic selection of the order n:

$$\begin{aligned} f(\sigma_j) &= \text{grad} \log \begin{bmatrix} \sigma_1, & \sigma_2, & \dots, & \sigma_{i\cdot l} \end{bmatrix} \\ n &= \arg\min_j f(\sigma_j) \end{aligned} \quad (5)$$

Algorithm continues from either Γ or X in a slightly different manner depending on the particular 4SID algorithm, however, both ways lead to a computation of the system matrices A and C using a least squares method. In the following, we will use an approach of [24] and using Matlab like notation for the selection of the sub-matrix of a given matrix:

$$\hat{C} = \Gamma(1:l,:), \quad \hat{A} = \Gamma(1:(i-1)\cdot l,:)^{-1}\Gamma(l+1:i\cdot l,:) \quad (6)$$

Given the matrices $\hat{A}$ and $\hat{C}$, the estimate of the system matrices B and D (and initial state x_0) is performed in a number of different ways, see e.g. [23], [17], [25], [26], [27]. We will adopt the idea from [24]. The system output equation can be written as

$$y(k) = CA^k x(0) + \sum_{j=0}^{k-1} CA^{k-j-1}Bu(j) + Du(k) + e(k), \quad (7)$$

with $e(k)$ being the noise contributions. Then Eq. (7) can be readily rewritten using an operator of vectorization vec as follows:

$$\begin{aligned} y(k) = CA^k x(0) + \left(\sum_{j=0}^{k-1} u(j)^T \otimes CA^{k-j-1} \right) \text{vec}\,(B) + \\ + (u(k)^T \otimes I_l)\text{vec}\,(D) + e(k). \end{aligned} \quad (8)$$

The optimization problem can be then formulated using a matrix form as

$$\theta^* = \arg\min_\theta \|\mathcal{Y} - \mathcal{Z}^T\theta\|_2^2 \quad (9)$$

where $\mathcal{Y}$ represents the vectors $y(k)$ stacked onto each other, $\mathcal{Z} = (\varphi(1), \dots, \varphi(N))$, with $\varphi^T(k) = \begin{pmatrix} \hat{C}\hat{A}^k & \sum_{j=0}^{k-1} u(j)^T \otimes \hat{C}\hat{A}^{k-j-1} & u(k)^T \otimes I_l \end{pmatrix}$ and $\theta^T =$

$\Big(x(0)^T \mathrm{vec}\,(\hat{B})^T \mathrm{vec}\,(\hat{D})^T\Big)$. Finally, given the estimates of the system matrices A, B, C, D the Kalman gain matrix K can be computed. If an estimate of a state sequence X is known, the problem can be solved by computing the Algebraic Riccati Equation (ARE) in which the covariance matrices are determined from the residuals as follows:

$$\begin{bmatrix} W \\ V \end{bmatrix} = \begin{bmatrix} X_{k+1} \\ Y_k \end{bmatrix} - \begin{bmatrix} \hat{A} & \hat{B} \\ \hat{C} & \hat{D} \end{bmatrix} \begin{bmatrix} X_k \\ U_k \end{bmatrix}, \quad (10)$$

where

$$\begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} = \frac{1}{N}\left(\begin{bmatrix} W \\ V \end{bmatrix}\begin{bmatrix} W^T & V^T \end{bmatrix}\right). \quad (11)$$

IV. Identification Algorithm Properties, Issues and Proposed Solutions

A. Input signals

Generation of sufficiently exciting input signals is one of the key theoretical assumptions enabling reliable statistical identification. Under real operation, this request is almost infeasible due to technical, physical or economical constraints and limitations. As the image of the building modeled in EP is at hand, the identification experiment was proposed as follows.

Three different kinds of input signals have been constructed, namely pseudo-random binary signal (PRBS), sum of sinusoids (SINE) and multilevel pseudo-random signal (MPRS). Let τ_H, τ_L denote the slowest and the fastest time constants of the system, respectively. Then the frequency spectrum to be covered is (ω_*, ω^*) with $\omega_* = \frac{1}{\beta\tau_H} \leq \omega \leq \frac{\alpha}{\tau_L}\omega^*$, where α defines how fast will the closed-loop be with respect to the open-loop response, and β specifies low frequency information corresponding to the settling time. The typical values are $\alpha = 2$ and $\beta = 3$, which corresponds to 95 % of the settling time [28]. In case of MPRS, the input sequence is computed by Galoise fields [28] with the number of shift registers n and the length q, which defines the maximum possible multiple of harmonics to be suppressed. In the opposite way, let h be the maximum possible multiple of the harmonics to be suppressed. Then q has to be chosen such that $q \geq 2^h - 1$ holds and the length n is computed as

$$\omega_* \geq \frac{2\pi}{T_s(q^n - 1)}. \quad (12)$$

The length of a signal cycle is $N_{cyc} = q^n - 1$, which (in time domain) represents a signal of duration $T_{cyc} = N_{cyc} \cdot T_s$. The number of the signals to be generated (m) does not need to be considered, as it is sufficient to generate a single signal and shift it $(m-1)$ times, which guarantees good statistical properties of the generated signals [29].

B. Analysis of model linearity

The fact that the underlying physics of the process is linear (Section II-B) does not mean, that the EP model is indeed linear. The linearity should be hence verified in the first place. To cope with this issue, one can design special input consisting of several multilevel steps, and afterwards compare their particular effect on output. Fig. 3 depicts the design of the multilevel steps as input, as well as the output from the EP. Next, Fig. 4 shows appropriately scaled and shifted step responses. We can see, that the step responses corresponding to the different operating points are the same, thus the underlying system is indeed linear.

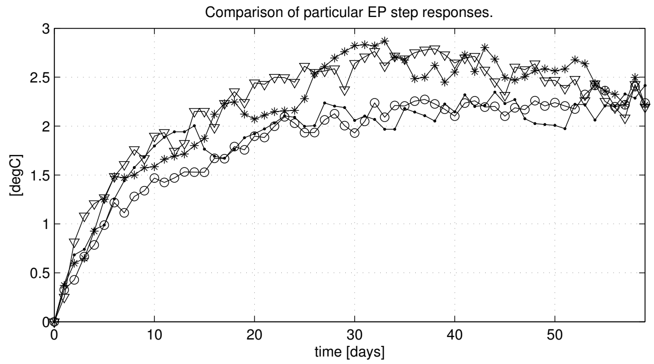


Fig. 4. Linearity verification: comparison of particular steps.

C. Analysis of size of Hankel matrices and step responses

The next step in identification is to decide, how to choose the parameters that enter the algorithm, e.g. identification algorithm, model order or size of the Hankel matrices, which is given by the number i of the block rows in said matrices. Due to the assumed (and verified) linearity of the EP model and the physically-based assumption that the change of temperature is the 1^{st} order process, the whole subsystem consisting of 6 zones is considered to be of the 6^{th} order[1]. [23] showed, that the number i of block rows of the Hankel matrices must be larger than the maximum order of the system to be identified. Essentially, the number i means how much into the past or future of the measured data are we looking, therefore it may appear that the greater i is, the better the result will be. However, a compromise has to be done concerning the computation difficulties, especially for the case of a large MIMO system identified (such as a building) .

Several numbers i had been selected and their effect on the identification results were analyzed. Fig. 5 shows the step responses of several inputs for $i = 6, 9, 15$ and 20. For i up to, let say, 12 (two times the considered order), the step responses possess good properties, such as reliable dynamics and the sign of the effect, as well as its nominal value. On the other hand, greater i causes unexpected "nervous-like" behavior at higher frequencies, wrong sign and different DC-gains (even for different i) of the contribution to output.

D. Analysis of extended observability matrix singular values

Yet another question arises and should be examined: Is the chosen algorithm suitable and powerful enough to identify such a large system? The first step in this analysis was to identify the model from a "simple" input dataset, and, if

[1]This was also verified by subspace algorithm selecting order according to Eq. (5)

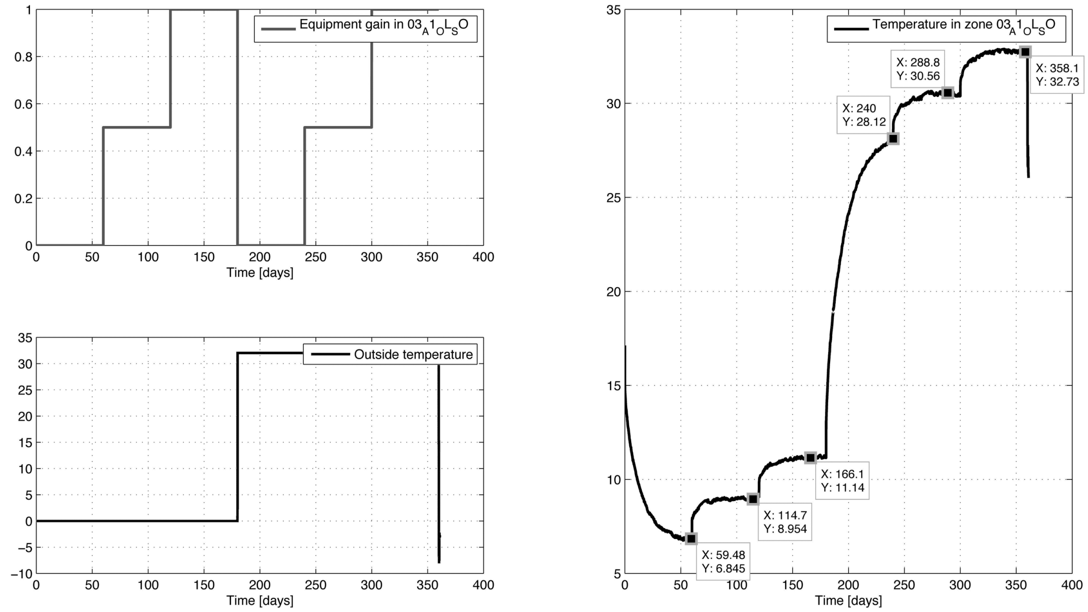


Fig. 3. Linearity verification: experiment design.

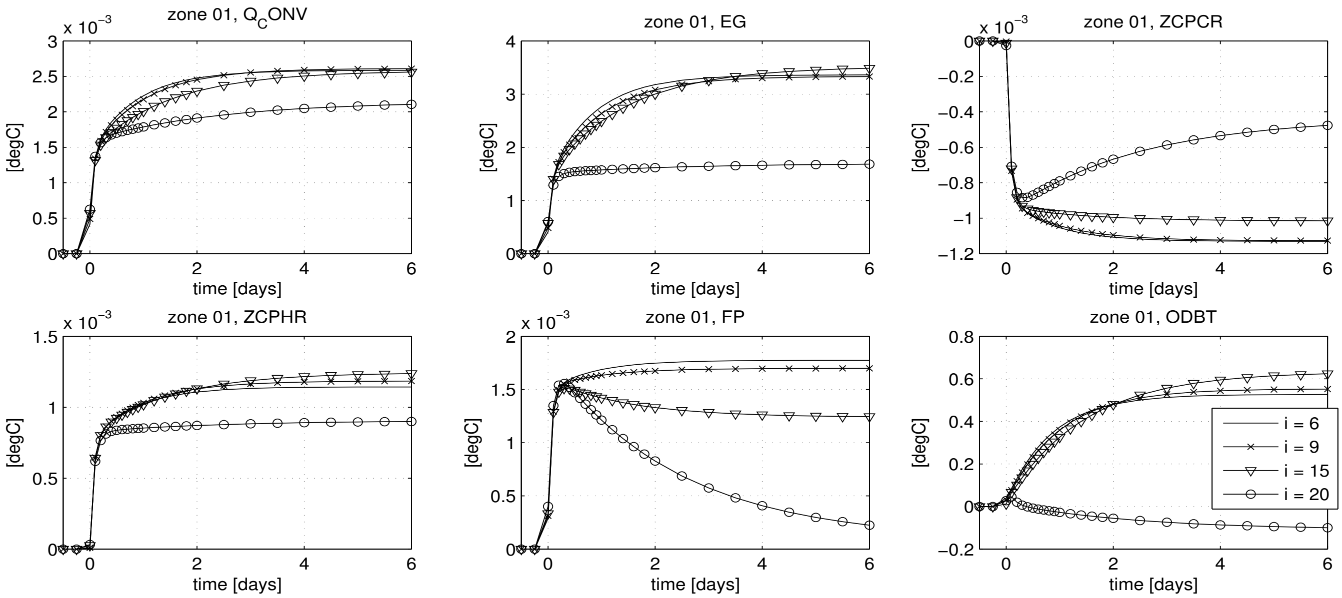


Fig. 5. Step responses of several inputs in zone 1 for different is. Vertical axes are particular contributions to zone temperatures.

successful, a deeper analysis follows. The "simple" input dataset has been chosen such that all inputs are constant inputs, except of one input, where a long step function was introduced. This input dataset is assumed to lead the system into a steady state, and the effect of the single-input-step shows up in all outputs. The model identified from this dataset identifies both the DC-gain, including the sign, and the time constant (Fig. 6).

Matrix $\mathcal{O}$ (see Eq. (4)) is used to select the order of the model – 4SID algorithms often rely on the mere distribution of its singular values (see for instance the presented algorithm which determines the order n according to Eq. (5)). For deterministic systems, the singular values distribution has a big step in magnitude, while for stochastic systems, it holds that the bigger noise covariance is, the smaller is the step in magnitude of singular values; for certain covariances, there is no step and one can hardly identify the correct order.

The analysis of singular values is of special interest when some of the 4SID algorithm assumptions do not hold (refer to Section III-A). Unfortunately, this is a common situation in real systems identification; the system may be e.g. subject to a non-white noise, or, in other words, by an unknown disturbance input.

We performed two different experiments to see the impact

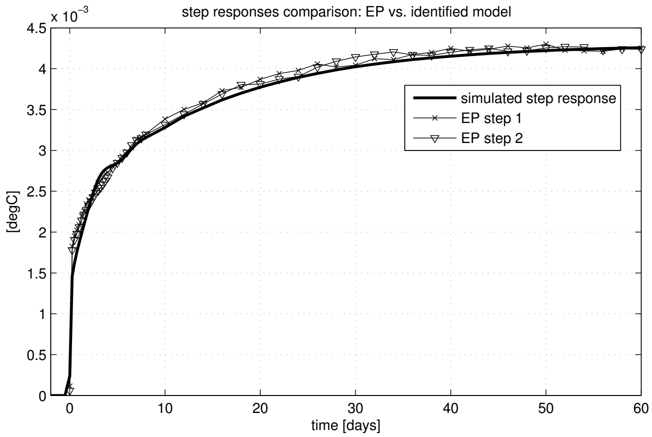


Fig. 6. Comparison of step responses of EP and identified model. Step applied at convectors, output is contribution to zone temperature.

of possible unknown disturbance inputs. In the first example, we created a simple two-zone model of a building (according to [15]). The inputs to the system were convector heating rates in both zones, common heating rate caused by occupants, equipment or solar radiation, and the last system input was the outside temperature. The system had six states, each zone was characterized by three states (one state for air in the zone with fast dynamics and two representing wall layers). Only two states were measured – zone air temperatures. The data were generated according to Section IV-A for three different setups:

- Common heating rate was set to zero, as well as outside temperature.
- Common heating rate was set to a constant value, as well as outside temperature.
- Common heating rate was generated using MPRS, whilst outside temperature was chosen according to winter weather in Prague, Czech Republic i.e. sinus character, $-4\,^{\circ}$C at night and $4\,^{\circ}$C during the day.

The convector heating rates were generated using PRBS in all setups. We had taken into account only a subset of inputs, namely convector heating rates, and then performed identification using generated input-output data. Analyzing the singular values of extended observability matrix in Fig. 7, we can see that the more exciting the unknown signal is, the more states are included in the system matrix A in order to track the trajectory of the output perfectly. Constant value on unknown inputs adds one extra state located in 1.0; MPRS and sine on unknown inputs adds even more states. The fit of the resulting model is perfect on the identification dataset, but in situation where unknown inputs change, the model completely fails.

From the first example, one can see the importance of knowledge of all inputs affecting the system. The second example addresses analysis of singular values (see Fig. 8) for data gathered using EP simulation environment and artificial model with a similar structure (in terms of number of zones, as well as level of detail). Note the different scaling of the

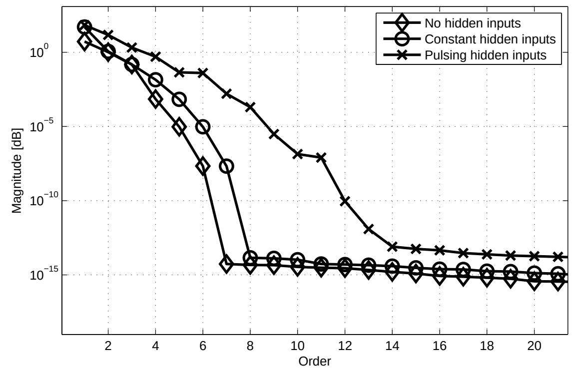


Fig. 7. Comparison of singular values of extended observability matrix for different levels of unknown inputs.

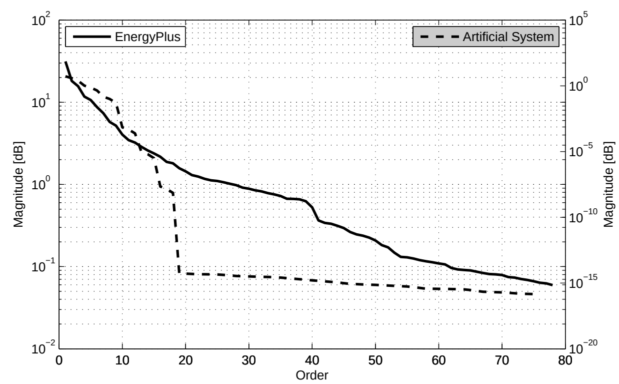


Fig. 8. Comparison of singular values of extended observability matrix composed of EnergyPlus or Artificial system input output data.

both plots. I/O data generated by EP leads to singular values distribution which is strongly affected by:

- Nonlinearity of EP computations (even though the I/O behavior should be linear).
- The relation might be linear, however, there are some internal EP controllers which cannot be switched off, which can also cause the failure of the identification procedure.
- There are some extra inputs affecting the system which are not considered, even though Table I contains a comprehensive selection of I/O for 4SID.
- There is a pseudo-random noise additively superposed to the system outputs.

Detailed analysis of all aforementioned points may lead to the holy grail i.e. getting a linear time invariant model suitable for control using a simplification of a detailed model in some building simulation software.

V. Consequences for Control and Conclusions

A. Concluding remarks

This paper has introduced a new methodology of interconnecting building simulation software and traditional identification methods, in order to avoid the statistical problems with data gathered from the real building. The building was modeled using EnergyPlus, which was excited by specially proposed signals to get data of a good quality. Then subspace identification approach (with some modifications) was applied to acquire a model suitable for predictive control. The last step of preparation of the model for control are adjustments of inputs and outputs for obtaining the model corresponding to the variety of MPC problems (according to the control criteria). We have investigated a number of properties and parameters of the identification algorithm and provided some hints for identification of the large MIMO systems such as buildings. The predictive control of the large building will follow-up in a separate paper.

VI. Acknowledgments

The project has been supported by the Ministry of Industry and Commerce of the Czech Republic, grant No. FR-TI1/517, *Control systems for energy consumption optimization in low-energy and passive houses*.

References

[1] "Communication from the commission to the European Parliament, the Council, the European Economic and Social Comittee and the Comittee of the regions," Brussels, p. 14, 2010.

[2] L. Perez-Lombard, J. Ortiz, and C. Pout, "A review on buildings energy consumption information," *Energy and buildings*, vol. 40, no. 3, pp. 394–398, 2008.

[3] J. Wong, H. Li, and S. Wang, "Intelligent building research: a review," *Automation in Construction*, vol. 14, no. 1, pp. 143–159, jan 2005.

[4] H. Han, Y. Jeon, S. Lim, W. Kim, and K. Chen, "New developments in illumination, heating and cooling technologies for energy-efficient buildings," *Energy*, 2009.

[5] H. Hjalmarsson, "System identification of complex and structured systems," *European journal of control*, vol. 15.

[6] R. Gopaluni, R. Patwardhan, and S. Shah, "MPC relevant identification–tuning the noise model," *Journal of Process Control*, vol. 14, no. 6, pp. 699–714, 2004.

[7] D. Lauri, J. Salcedo, S. Garcia-Nieto, and M. Martinez, "Model predictive control relevant identification: multiple input multiple output against multiple input single output," *Control Theory & Applications, IET*, vol. 4, no. 9, pp. 1756–1766, 2010.

[8] D. Shook, C. Mohtadi, and S. Shah, "A control-relevant identification strategy for GPC," *IEEE Transactions on Automatic Control*, vol. 37, no. 7, pp. 975–980, 2002.

[9] H. Madsen and J. Holst, "Estimation of continuous-time models for the heat dynamics of a building," *Energy and Buildings*, vol. 22, no. 1, pp. 67–79, 1995.

[10] K. Andersen, H. Madsen, and L. Hansen, "Modelling the heat dynamics of a building using stochastic differential equations," *Energy and Buildings*, vol. 31, no. 1, pp. 13–24, 2000.

[11] M. Gwerder and J. Toedtli, "Predictive control for integrated room automation," in *8th REHVA World Congress Clima*, 2005.

[12] M. Trčka, J. Hensen, and M. Wetter, "Co-simulation of innovative integrated HVAC systems in buildings," *Journal of Building Performance Simulation*, vol. 2, no. 3, pp. 209–230, Sep. 2009.

[13] J. Cigler and S. Prívara, "Subspace identification and model predictive control for buildings," in *The 11th International Conference on Control, Automation, Robotics and Vision – ICARCV2010*, pp. 750–755.

[14] Z. Váňa, J. Kubeček, and L. Ferkl, "Notes on finding black-box model of a large building," in *Control Applications (CCA), 2010 IEEE International Conference on*, 9 2010, pp. 1017–1022.

[15] M. Gwerder, B. Lehmann, J. Tödtli, V. Dorer, and F. Renggli, "Control of thermally-activated building systems (TABS)," *Applied energy*, vol. 85, no. 7, pp. 565–581, 2008.

[16] T. Bohlin and S. Graebe, "Issues in nonlinear stochastic grey box identification," *International Journal of Adaptive Control and Signal Processing*, vol. 9, no. 6, pp. 465–490, 2007.

[17] L. Ljung, *System Identification: Theory for user*. Prentice-Hall, Inc., Upper Saddle River, New Jersey, USA, 1999.

[18] S. Prívara, J. Cigler, Z. Váňa, L. Ferkl, and M. Šebek, "Subspace identification of poorly excited industrial systems," in *Proceedings of the 49th IEEE Conference on Decision and Control*, 2010, pp. 4405–4410.

[19] S. Prívara, Z. Váňa, D. Gyalistras, J. Cigler, C. Sagerschnig, M. Morari, and L. Ferkl, "Modeling and identification of a large multizone office building," in *2011 IEEE Multi-Conference on Systems and Control, Denver, CO, USA*, 2011.

[20] F. Lewis, *Optimal estimation with an introduction to stochastic control theory*. Wiley New York et al., 1986.

[21] R. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[22] P. Trnka, "Subspace identification methods," Ph.D. dissertation, Czech Technical University in Prague, 2007.

[23] P. Van Overschee and B. De Moor, *Subspace Identification for Linear Systems*. Kluwer Academic Publishers, 101 Philip Drive, Assinippi Pard, Nowell, Massachusetts: Kluwer Academic Publishers, 1999.

[24] C. Lyzell, "Initialization methods for system identification," Licentiate Thesis no. 1426, Department of Electrical Engineering, Linköping University, SE-581 83 Linkőping, Sweden, Nov. 2009.

[25] G. V. D. Veen, J.-w. V. Wingerden, and M. Verhaegen, "Closed-loop MOESP subspace model identification with parametrisable disturbances," in *49th IEEE Conference on Decision and Control*, Atlanta, 2010, pp. 2813–2818.

[26] M. Pouliquen, O. Gehan, and E. Pigeon, "An indirect closed loop subspace identification method," in *49th IEEE Conference on Decision and Control*, no. 3, Atlanta, 2010, pp. 4417–4422.

[27] D. N. Miller and R. A. D. Callafon, "Subspace Identification Using Dynamic Invariance in Shifted Time-Domain Data," in *49th IEEE Conference on Decision and Control*, no. 0, Atlanta, 2010, pp. 2035–2040.

[28] M. Braun, D. Rivera, A. Stenman, W. Foslien, and C. Hrenya, "Multi-level pseudo-random signal design and model-on-demand estimation applied to nonlinear identification of a RTP wafer reactor," in *American Control Conference, 1999. Proceedings of the 1999*, vol. 3. IEEE, 2002, pp. 1573–1577.

[29] S. V. Gaikwad and D. E. Rivera, "Control-relevant input signal design for multivariable system identification: Application to high-purity distillation," 1996.