Randomized Model Predictive Control for HVAC Systems

Alessandra Parisio, Damiano Varagnolo, Daniel Risberg, Giorgio Pattarello, Marco Molinari, Karl H. Johansson
ACCESS Linnaeus Centre, School of Electrical Engineering, KTH Royal Institute of Technology
{ parisio | damiano | dri | pattarello } @kth.se, marco.molinari@byv.kth.se, kallej@kth.se
*

Abstract

Heating, Ventilation and Air Conditioning (HVAC) systems play a fundamental role in maintaining acceptable thermal comfort and Indoor Air Quality (IAQ) levels, essentials for occupants well-being. Since performing this task implies high energy requirements, there is the need for improving the energetic efficiency of existing buildings. A possible solution is to develop effective control strategies for HVAC systems, but this is complicated by the inherent uncertainty of the to-be-controlled system. To cope with this problem, we design a stochastic Model Predictive Control (MPC) strategy that dynamically learns the statistics of the building occupancy and weather conditions and uses them to build probabilistic constraints on the indoor temperature and CO_2 concentration levels. More specifically, we propose a randomization technique that finds suboptimal solutions to the generally non-convex stochastic MPC problem. The main advantage of this method is the absence of apriori assumptions on the distributions of the uncertain variables, and that it can be applied to any type of building. We investigate the proposed approach by means of numerical simulations and real tests on a student laboratory, and show its practical effectiveness and computational tractability.

Keywords

Randomized model predictive control, HVAC control, Copulas, statistics

1 Introduction

It is well known that Heating, Ventilation and Air Conditioning (HVAC) systems, necessary technologies to guarantee acceptable Indoor Air Quality (IAQ) and thermal comfort levels, come with high energy requirements. How to reduce the energy use of HVAC systems, while satisfying occupants comfort requirements, is a relevant research topic.

An effective controller for HVAC systems should incorporate time-dependent energy costs, bounds on the control actions, targets on the IAQ and thermal conditions, as well as account for system uncertainties, i.e., weather conditions and occupancy. By doing so the buildings thermal storage capacities can be effectively utilized.

A natural scheme that achieves the systematic integration of all the aforementioned elements is the so-called Stochastic Model Predictive Control (SMPC) [19]. Since the stochastic laws ruling the occupancy and weather patterns are geographically and time varying, it is desirable that the controller can learn the statistics of the random variables from the experience.

Literature review

The literature on Model Predictive Control (MPC) for indoor climate control is flourishing. Several studies show that predictive controllers may significantly decrease energy consumptions when endowed with real-time measurements, weather conditions, and occupancy forecasts [7, 16, 24, 10, 9]. This is confirmed by experimental results on real buildings, where MPCs yield better energy use and comfort levels performance than current practices [26, 12].

There is nonetheless still room for improvements: these controllers consider deterministic forecasts for the disturbances, and disregard information on the statistics on the unavoidable forecasts errors. A common opinion is that actually this is an issue: as current standards explicitly state, rooms temperatures should be kept within a comfort range *with a predefined probability* [2]. Thus, building climate control leads *naturally* to probabilistic constraints.

A stochastic version of MPC including probabilistic constraints can address this issue and explicitly account for system uncertainties. Several SMPC schemes with probabilistic constraints, generally called *chance constraints*, have already been proposed in literature [15, 17, 21, 18]. E.g., [18] incorporates stochastic occupancy models within the control loop, while [15, 17] propose stochastic predictive building temperature regulators where weather and load disturbances are modeled as Gaussian processes. The resultant nonlinear program is then solved with a tailored sequential quadratic programming which exploits the sparsity of the

*This work is supported by the European Institute of Technology (EIT) Information and Communication Technology (ICT) Labs, the Swedish Energy Agency, the Swedish Governmental Agency for Innovation Systems (VINNOVA) and the Knut and Alice Wallenberg Foundation.

Buildsys'13, November 14-15, 2013, Rome, Italy.

quadratic sub-problems. Also [21] integrates weather predictions into an SMPC. Here the control action is computed by solving a non-convex problem which exploits linearizations around nominal trajectories, and then by applying a disturbance feedback. Remarkably, [21] uses deterministic predictions of the internal gains; the only prediction for which uncertainties (assumed Gaussians) are accounted for is the weather one. Actually this is a common feature of all the SMPC schemes described in this paragraph: disturbances are Gaussians and additive processes. Further, generally the proposed SMPC controllers do not explicitly control the indoor air quality considering the uncertainty in the occupancy.

At the best of our knowledge, only a few proposals depart from these Gaussian assumptions. One is our [22], where the controller exploits a scenario-based tractable approximation of the chance constrained MPC problem, and where the scenarios are i.i.d. samples extracted from general probability distributions. The other one is [28], where the bilinear building model is iteratively linearized around nominal trajectories and where occupancy scenarios are sampled from a set of measurement data collected in eight single offices equipped with motion sensors.

The numerical simulations performed in [28] suggest that scenarios-based techniques outperform other predictive methods and that the number of scenarios required to obtain reliable solutions can be prohibitive for the building case, while using a small number of scenarios fails in obtaining effective actuation levels.

Statement of contributions

Our aim is to develop effective control laws that do not require demanding installation costs. The big vision is to pair advanced control schemes with learning technologies, and obtain easily deployable HVAC control schemes. Here we move along this direction, and propose a stochastic MPC for HVAC systems, which employs a learning module that continuously and dynamically infers the statistics of the uncertainties from real data. The results from the learning module are incorporated in an MPC problem with probabilistic constraints on the indoor temperature and CO_2 concentration levels.

The control target is to minimize the energy use while satisfying both thermal comfort and air quality requirements.

Randomized techniques are applied in order to find suboptimal solutions to the generally non-convex chance-constrained problem; in the rest of the paper we indicate this novel scheme with the acronym Randomized Model Predictive Control (RMPC).

With respect to the existing literature we introduce some major novelties:

- we show that appying a randomized technique to the chance constrained MPC for HVAC systems can improve the control of these systems;
- we extend the statistics learning scheme by adding some parametric families as plausible distributions for the stochastic variables;
- we present results of the implementation of the scheme on a real testbed located in Stockholm, Sweden.

Organization of the manuscript

In Section 2 we presents the predictive controller and the related system model. Section 3 outlines instead the learning module that dynamically infers the statistics of the uncertainties from actual data. Section 4 provides and discusses the experimental results, and Section 5 eventually summarizes our conclusions and proposes some future extensions.

2 Implementation of Randomized MPC for HVAC systems

In this section we first describe the model of the system, then we outline the structure of the MPC problem.

The inputs of the overall MPC scheme are, at every time step, weather conditions, occupancy scenarios, and measurements of the current state of the system. The output is instead a heating, cooling and ventilation plan for the next N hours, where N is the prediction horizon. Conforming with the MPC paradigm, only the first step of this control plan is applied to the HVAC system. After that, the whole procedure is repeated. This introduces feedback into the system, since the optimal control problem is a function of the current state and of any disturbance acting on the building at the current time step. More precisely, the outputs computed at each time k are a mass air flow rate $\dot{m}_{\text{venting}}(k)$, a ventilation system air temperature $T_{\text{sa}}(k)$, and a radiators mean radiant temperature T_{mr}.

The independence of the air quality dynamics from the thermal ones allow us to decouple the control of the temperature and of the air quality in two separated subproblems: (*i*) the IAQ-RMPC, which aims at satisfying the required air quality at a minimum energy use, and computes the optimal sequence of the mass air flow rates over a given prediction horizon; (*ii*) the T-RMPC, which handles the indoor temperature. By doing so, the computational tractability of the overall control problem will be improved.

Since the air quality requirements have priority over the thermal comfort, the solution computed by the IAQ-RMPC lower bounds the air flow rate of the T-RMPC.

2.1 Modeling

Since the overall building energy usage is commonly computed as the sum the energy usages of the single thermal zones [10], here we focus on the control of a single thermal zone (or room). As the structure of this subsection suggests, we employ two different models: one for the thermal evolution of the environment, and one for the dynamics of the concentration of CO_2.

Model for the thermal dynamics

We consider a thermal Resistive-Capacitive (RC) network of first-order systems, where the nodes are the states representing the room, the walls, the floor and the ceiling temperatures. Each state is associated to a heat transfer differential equation. We assume that we can control two different heat flows: Q_{venting}, representing the contribute due to the ventilation system, and Q_{heating}, representing the radiators. We consider the outside temperature, the radiation, the internal gains, the heat flows due to occupancy, equipments and lightings as disturbances. See [22] for additional details.

The control inputs are expressed as

$$Q_{\text{venting}} = \dot{m}_{\text{venting}} c_{\text{pa}} \left(\Delta T_h - \Delta T_c\right) = c_{\text{pa}}(u_h - u_c), \quad (1)$$

$$Q_{\text{heating}} = A_{\text{rad}} h_{\text{rad}} \Delta T_{\text{h,rad}} = A_{\text{rad}} h_{\text{rad}} \left(T_{\text{mr}} - T_{\text{room}}\right), (2)$$

where $\dot{m}_{\text{venting}}$ is the ventilation mass flow, c_{pa} is the specific heat of the dry air, $\Delta T_h = \left(T_{\text{sa}} - T_{\text{room}}\right)$ and $\Delta T_c = \left(T_{\text{room}} - T_{\text{sa}}\right)$ are respectively the temperature difference through the heating and cooling coils, T_{sa} is the temperature of the air supplied by the ventilation system, A_{rad} is the emission area of the radiators, and h_{rad} is the heat transfer coefficient of the radiators and T_{mr} is the mean radiant temperature of the radiators. Notice that $c_{\text{pa}} u_h(k)$ and $c_{\text{pa}} u_c(k)$ model the portion of the ventilation heat flow due to respectively heating and cooling.

We model the room temperature dynamics with the discrete-time Linear Time Invariant (LTI) system

$$\begin{aligned} x_{\text{T}}(k+1) &= A_{\text{T}} x_{\text{T}}(k) + B_{\text{T}} u_{\text{T}}(k) + E_{\text{T}} w_{\text{T}}(k) \\ y_{\text{T}}(k) &= C_{\text{T}} x_{\text{T}}(k), \end{aligned} \qquad (3)$$

where $x_{\text{T}}(k)$ is the state vector containing the room temperature and the inner and outer temperatures of all the walls, $u_{\text{T}}(k) := \left(u_h(k), u_c(k), \Delta T_{h,\text{rad}}(k)\right)$ is the input vector, $w_{\text{T}}(k)$ is the vector of random disturbances containing the outside temperature, the solar radiation and the internal heat gain at time k, and the matrices $A_{\text{T}}, B_{\text{T}}, E_{\text{T}}, C_{\text{T}}$ are of appropriate sizes. The output $y_{\text{T}}(k)$ is the room temperature at time k.

Hence, the mass air flow rate and the supply air temperature at each k are easily computed from the obtained values of either $u_h(k)$ or $u_c(k)$ considering both the requirements on the air quality and the comfort requirements on the supply air temperature.

Model for the CO_2 concentration dynamics

The model is derived from a CO_2 balance equation accounting for the fresh air from the ventilation system and the amount of CO_2 generated per occupant. The state of the model is the nonnegative difference between the CO_2 concentration in the room and inlet air CO_2 concentration (assumed equal to outdoor CO_2 concentration), and is indicated with $x_{\text{CO}_2} = \Delta\text{CO}_2$. We assume that we can control the mass air flow from the ventilation system, while the number of occupants is considered a disturbance.

The resulting model is bilinear in the state and in the control input. To simplify the problem formulation we then derive an equivalent linear model by replacing the bilinear term $\dot{m}_{\text{venting}} \cdot x_{\text{CO}_2}$ with u_{CO_2} and by adding the constraint

$$\dot{m}_{\text{venting}}^{\min} \cdot x_{\text{CO}_2}(k) \leq u_{\text{CO}_2}(k) \leq \dot{m}_{\text{venting}}^{\max} \cdot x_{\text{CO}_2}(k) \qquad (4)$$

on the new input $u_{\text{CO}_2}(k)$. These constraints guarantee that the physical bounds on the control input in the original nonlinear model are always satisfied. The original input, at each k and for $x_{\text{CO}_2}(k) > 0$, can eventually be obtained as

$$\dot{m}_{\text{venting}}(k) = \frac{u_{\text{CO}_2}(k)}{x_{\text{CO}_2}(k)}.$$

Then, the CO_2 concentration dynamics can be described by the discrete time Linear Time Invariant (LTI) system

$$\begin{aligned} x_{\text{CO}_2}(k+1) &= a x_{\text{CO}_2}(k) + b u_{\text{CO}_2}(k) + e w_{\text{CO}_2}(k) \\ y_{\text{CO}_2}(k) &= x_{\text{CO}_2}(k). \end{aligned} \qquad (5)$$

2.2 Randomized MPC

Here we describe the design of the two controllers, Temperature (T)-RMPC and IAQ-RMPC, which use models (3) and (5) respectively.

Since both models are LTI and both controllers need to handle hard constraints on the inputs and probabilistic constraints on the outputs, we can uniform the notation and develop both the controllers following similar steps.

We thus indicate both models simultaneously with

$$\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + Ew(k) \\ y(k) &= Cx(k), \end{aligned} \qquad (6)$$

where $x(k) \in \mathcal{R}^n$, $u(k) \in \mathcal{R}^m$, $w(k) \in \mathcal{R}^r$ and $y(k) \in \mathcal{R}^p$. The model in (6) represents either (3) or (5), depending on the controller under consideration (T-RMPC or IAQ-RMPC). We notice that the bound on the room temperature are generally time-varying, since the comfort levels can be relaxed during no-occupancy periods.

Let thus x_t be the current state of system (6). The output trajectories over the prediction horizon N can then be written as

$$y(t+k|t) = CA^k x_t + \sum_{i=0}^{k-1} CA^{k-i-1} Bu(i) + \sum_{i=0}^{k-1} CA^{k-i-1} Ew(i). \qquad (7)$$

Given (7), we can then express the output $\boldsymbol{Y}_t \in \mathbb{R}^{pN}$ over the whole prediction horizon as a function of the initial state x_t as $\boldsymbol{Y}_t = \boldsymbol{C}(\boldsymbol{A}x_t + \boldsymbol{B}\boldsymbol{U}_t + \boldsymbol{E}\boldsymbol{W}_t)$,where the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{E}$ and $\boldsymbol{C}$ are built applying (7) recursively N times, $\boldsymbol{U}_t \in \mathbb{R}^{mN}$ are the control inputs, and $\boldsymbol{W}_T \in \mathbb{R}^{rN}$ and the disturbances over the prediction horizon.

Letting $\boldsymbol{G}_x := [\boldsymbol{CA}]$, $\boldsymbol{G}_u := [\boldsymbol{CB}]$, $\boldsymbol{G}_w := [\boldsymbol{CE}]$, $\widetilde{\boldsymbol{g}} := \left[-y_{\min}(k)^{\text{T}} \cdots - y_{\min}(k)^{\text{T}} y_{\max}(k)^{\text{T}} \cdots y_{\max}(k)^{\text{T}}\right]^{\text{T}}$, $\boldsymbol{g} := \widetilde{\boldsymbol{g}} - \boldsymbol{G}_x x_t$, $\boldsymbol{F} := \begin{bmatrix} -\boldsymbol{I} \\ \boldsymbol{I} \end{bmatrix}$, $\boldsymbol{f} := \left[-u_{\min}^{\text{T}} \cdots - u_{\min}^{\text{T}}\, u_{\max}^{\text{T}} \cdots u_{\max}^{\text{T}}\right]^{\text{T}}$, with $\boldsymbol{0}$ and $\boldsymbol{I}$ opportunely dimensioned zero and identity matrices, the inputs and outputs constraints over the whole prediction horizon N become

$$\boldsymbol{G}_u \boldsymbol{U}_t + \boldsymbol{G}_w \boldsymbol{W}_t \leq \boldsymbol{g}, \qquad \boldsymbol{F}\boldsymbol{U}_t \leq \boldsymbol{f}.$$

Problem 1 (Chance Constrained MPC for HVAC Control) The MPC problem can be formulated as

$$\begin{aligned} \min_{\boldsymbol{U}_t} \quad & \boldsymbol{c}^{\text{T}} \boldsymbol{U}_t \Delta k \\ \text{s.t.} \quad & \mathbb{P}\left[\boldsymbol{G}_u \boldsymbol{U}_t + \boldsymbol{G}_w \boldsymbol{W}_t - \boldsymbol{g} \leq 0\right] \geq 1 - \alpha, \quad \boldsymbol{F}\boldsymbol{U}_t \leq \boldsymbol{f} \end{aligned}$$

where $1-\alpha$ is the desired probability level for constraint satisfaction, Δk is the sampling period, $\boldsymbol{c}^{\text{T}} \boldsymbol{U}_t$ is the energy use vector over the whole prediction horizon, $\boldsymbol{c} \in \mathbb{R}^{mN}$ is the cost vector, containing either only ones for the IAQ-RMPC case, or the specific heat of the dry air c_{pa} and the product $A_{\text{rad}} h_{\text{rad}}$ between the emission area and the heat transfer coefficient of the radiators for the T-RMPC case.

Chance constrained problems like 1 are generally intractable unless the uncertainties follow specific distributions, e.g., Gaussian or log-concave; in these cases, it is possible to obtain equivalent convex –and thus computationally efficient– reformulations [14].

However, as described later, Gaussian assumptions are rather restrictive. To overcome this limitation, but still obtain a solvable MPC problem, we propose to apply *randomized* approaches [3], that do not require the specification of particular probability distributions for the uncertainties but only the capability of randomly extracting from them.

The approach is as follows: let $\boldsymbol{W}_{t,1},\ldots,\boldsymbol{W}_{t,M}$ be a set of M i.i.d. disturbances samples (called *scenarios*), $\boldsymbol{W}_{t,i} := \left[w_i^{\mathrm{T}}(t),\ldots,w_i^{\mathrm{T}}(t+N-1)\right]^{\mathrm{T}}$, $i=1,\ldots M$. Then, the chance constraints in Problem 1 are replaced with the following set of deterministic constraints

$$\boldsymbol{G}_u\boldsymbol{U}_t + \boldsymbol{G}_w\boldsymbol{W}_{t,i} - \boldsymbol{g} \leq \boldsymbol{0}, \quad i=1,\ldots,M.$$

Since the only constraint that is required to be satisfied is

$$\boldsymbol{G}_u\boldsymbol{U}_t \leq \boldsymbol{g} - \max_{i=1,\ldots M} \boldsymbol{G}_w\boldsymbol{W}_{t,i},$$

where the max applies element-wise to $\boldsymbol{G}_w\boldsymbol{W}_{t,i}$, most of the constraints in (8) are redundant.

Letting $d=mN$ be the number of decision variables, to choose the number of scenarios M to be generated one may exploit the sufficient condition

$$M \geq \frac{2}{\alpha}\left(ln\left(\frac{1}{\beta}\right)+d\right), \tag{8}$$

that guarantees that solving constraints (8) will lead to a feasible solution for Problem 2 with a confidence level $(1-\beta)\in(0,1)$ [3, 4] (with β an user-defined parameter).

Further, to guarantee that the problem with sampled constraints is always feasible, we soften the constraints in (8) by introducing the slack variables $s(k)\in\mathbb{R}^p$ at each time step k. The number of possible constraint violations can then be tuned by introducing a parameter that weights the slack variables in the objective function. If the optimal solution can be obtained without violations of the softened constraints, the slack variables will be set to zero. The designer can thus considerably penalize constraint violations by assigning to the weighting factor a value that is orders of magnitude greater than the other coefficients parameters.

Eventually we thus formulate the random convex problem embedded in the MPC scheme as

Problem 2 (RMPC for HVAC Control)

$$\begin{aligned} \min_{\boldsymbol{U}_t} \quad & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{U}_t\Delta k + \rho\boldsymbol{1}^{\mathrm{T}}\boldsymbol{s} \\ \text{s.t.} \quad & \boldsymbol{G}_u\boldsymbol{U}_t \leq \boldsymbol{g} + \boldsymbol{s} - \max\nolimits_{i=1,\ldots M}\boldsymbol{G}_w\boldsymbol{W}_{t,i}, \;\; \boldsymbol{F}\boldsymbol{U}_t \leq \boldsymbol{f} \end{aligned} \tag{9}$$

where $\boldsymbol{s}$ is the vector containing all the slack variables, ρ is the weight on the slack variables, and $\boldsymbol{1}$ is a matrix of ones with appropriate dimensions.

Our experience indicates that (8) may be overly pessimistic. E.g., we ran numerical simulations with $\alpha=0.05$ and $\beta=0.001$ and computed the empirical probability of constraint violation over 2400 different i.i.d. instances of the random convex problem (2). Applying condition (8), we set $M=3157$ and empirically reported a constraints violations probability of 0.0044. Halving the indication given by (8) ($M=1579$) instead led to an empirical probability of constraint violations of 0.042, much closer to the confidence level required initially.

Further, when compared to an ideal case endowed with error-free forecasts, used as a theoretical benchmark, our RMPC yields an almost neglectable amount of violations of the thermal bound and an increase of only 2.5% in the energy use.

3 Learning how to generate the scenarios

We now describe the approach used to learn the scenarios generation rules used by the above RMPC strategy. We start motivating the technological choice, then briefly introduce the mathematical concepts and the theory used.

3.1 Motivations

To model the distributions of the disturbances a first approach is to apply apriori considerations, e.g., physics based, that do not account for the actual measurements seen in the field. An alternative paradigm is instead to learn from the experience. If correctly implemented, the learning-based approaches give robustness and adaptability to different environments, necessary qualities if the technology wants to reach the market.

But how to do this learning step? As reported in the literature review, a classical approach is to pose Gaussianity assumptions, and then exploit the data to estimate the means and autocovariances. Unfortunately, Gaussianity induces limitations in the kind of dependencies that can be captured. I.e., Gaussianity restricts the plausible dependencies in the tails of the marginal distributions, see Figures 1 and 2 and their captions.

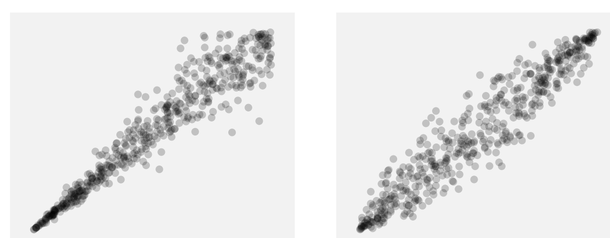

Figure 1. Samples from bidimensional Clayton (left) and Gaussian (right) copulas with uniform marginal densities. The Clayton samples (x,y) show strong left-tail dependency (x small induces y small) but weak right-tail dependency (x big does not induce y big). Gaussian samples instead have the same degree of dependency for both left- and right-tails.

Another classical approach is to represent the forecast quantities using Markov chains formalisms, but this requires some form of discretization processes (e.g., temperatures that may take values only on multiples of 0.5°C). Our opinion is that it is preferable to do not treat random processes like temperature or solar radiations as discrete quantities but rather maintain their natural continuous nature.

We thus consider *copulas*, mathematical objects famous specially in finance, hydrology, and wind forecasting, that

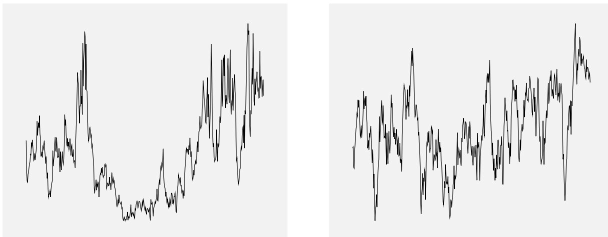

Figure 2. Effect of different left- and right-tail dependencies in time series. The Clayton samples $x(k)$ (left) show small variability when their value is small, and big variability when their value is big. The Gaussian samples (right) instead show an uniform variability over the whole range of values.

naturally capture every kind of dependence, allow far more flexibility than Gaussian processes assumptions, can manage both continuous and discrete random processes, and come with robust, tested and reliable learning algorithms.

The drawbacks are in the major computational requirements needed to handle the generation of scenarios w.r.t. Gaussian cases; nonetheless the feeling is that this is not anymore a concern, given the technological advancements in the capabilities of modern processors. Moreover, although theoretical foundations of copulas might seem complex, practical implementations and estimations are relatively straightforward. For more complete treatments on the subject we send the interested reader to [13, 20, 27]. For some specialized literature on copula methods for forecasting multivariate time series we suggest instead [23].

3.2 Notation and basic definitions

We use $\mathbb{P}[*]$ to indicate the probability of the generic event $*$. Letting $w(k)$ be a generic random variable of interest, we denote its Cumulative Distribution Function (CDF) with $\mathbb{F}_{w(k)}(a_k) := \mathbb{P}[w(k) \leq a_k]$, and its quantile with

$$\mathbb{F}^{-1}_{w(k)}(u_k) := \inf_{a_k} \left\{ a_k \mid \mathbb{F}_{w(k)}(a_k) \geq u_k \right\}. \tag{10}$$

We recall that $\mathbb{F}^{-1}$ is the inverse of $\mathbb{F}$ in the sense that if $\mathbb{F}_{w(k)}(a_k)$ is absolutely continuous and strictly monotone then $a_k = \mathbb{F}^{-1}_{w(k)}\left(\mathbb{F}_{w_k}(a_k)\right)$ for all a_k. We moreover recall the so-called *probability integral transform*, that is that particular property ensuring every continuous random variable $w(k) \sim \mathbb{F}_{w(k)}(a_k)$ to be transformable into $\omega_k = \mathbb{F}_{w(k)}(w(k)) \sim \mathcal{U}[0,1]$, i.e., an uniform r.v. Letting $\boldsymbol{w} := [w(1), \ldots, w(K)]$ be a generic random vector of interest, we denote its joint CDF with

$$\mathbb{F}_{\boldsymbol{w}}(a_1, \ldots, a_K) = \mathbb{P}[w(1) \leq a_1, \ldots, w(K) \leq a_K]. \tag{11}$$

Given (11), we call $\mathbb{F}_{w(k)}(a_k)$ the *marginal* distribution of $w(k)$.

3.3 Copulas

A *copula* is simply a function from the unitary hypercube to the unitary segment, i.e., $\mathbb{C} : [0,1]^K \mapsto [0,1]$, that satisfies three conditions: (*i*) $\mathbb{C}(1, \ldots, 1, u_k, 1, \ldots, 1) = u_k$ for every k and $u_k \in [0,1]$; (*ii*) if at least one u_k is zero then $\mathbb{C}(u_1, \ldots, u_K) = 0$; (*iii*) $\mathbb{C}$ is a K-increasing function. In words, a copula is a K-dimensional joint CDF of a random vector whose scalar components have all uniform marginals. I.e., every copula is an opportune CDF

$$\mathbb{C}(u_1, \ldots, u_K) = \mathbb{P}[\omega(1) \leq u_1, \ldots, \omega(K) \leq u_K] \tag{12}$$

where $\omega(k) \sim \mathcal{U}[0,1]$, for each k. Thus every different $\mathbb{C}$ can be considered a different way to impose dependencies between a set of K random variables $\omega(k)$ that, when considered by themselves, are uniformly distributed in $[0,1]$.

The previous concept can be extended to handle generic r.v.s: due to the probability integral transform, each $\omega(k)$ can be considered the transformation of an other $w(k)$, i.e., one can think that $\omega_k = \mathbb{F}_{w(k)}(w(k))$. This means that (12) can be rewritten as follows: choose K generic continuous marginals $\mathbb{F}_{w(1)}(\cdot), \ldots, \mathbb{F}_{w(K)}(\cdot)$, and let

$$\mathbb{C}(u_1, \ldots, u_K) = \mathbb{P}\left[\mathbb{F}_{w(1)}(w(1)) \leq u_1, \ldots, \mathbb{F}_{w(K)}(w(K)) \leq u_K\right]. \tag{13}$$

Since $\mathbb{F}_{w(k)}(w(k)) \leq u_k$ is equivalent to $w(k) \leq \mathbb{F}^{-1}_{w(k)}(u_k)$, it follows that

$$\mathbb{C}(u_1, \ldots, u_K) = \mathbb{P}\left[w(1) \leq \mathbb{F}^{-1}_{w(1)}(u_1), \ldots, w(K) \leq \mathbb{F}^{-1}_{w(K)}(u_K)\right]. \tag{14}$$

Let then $a_k = \mathbb{F}^{-1}_{w(k)}(u_k)$. This implies $u_k = \mathbb{F}_{w(k)}(a_k)$, and thus

$$\begin{aligned} \mathbb{F}_{\boldsymbol{w}}(a_1, \ldots, a_K) &= \mathbb{P}[w(1) \leq a_1, \ldots, w(K) \leq a_K] \\ &= \mathbb{C}\left(\mathbb{F}_{w(1)}(a_1), \ldots, \mathbb{F}_{w(K)}(a_K)\right). \end{aligned} \tag{15}$$

Thus if the random variables are continuous [1] one can always decompose the joint probability distribution $\mathbb{F}_{\boldsymbol{w}}(\cdot, \ldots, \cdot)$ in two distinct terms: the set of marginals $\mathbb{F}_{w(1)}(\cdot), \ldots, \mathbb{F}_{w(K)}(\cdot)$, that describe the statistical behavior of the random variables $w(k)$ when considered independently, and the copula $\mathbb{C}$, that captures the statistical dependency between the various $w(k)$. To summarize in words, copulas allow the researchers to specify separately the marginal distributions and the dependence structure, without losing any flexibility in the model, as instead Gaussian processes do.

3.4 Learning copulas

Assume to have measured N K-dimensional vectors $\boldsymbol{w}_n = [w_n(1), \ldots, w_n(K)]$ from some past observations (e.g., external temperatures for several days). One may thus use the N samples $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_N$ to learn the joint CDF $\mathbb{F}_{\boldsymbol{w}}(a_1, \ldots, a_K)$, and then use this estimated CDF to generate the scenarios needed by the RMPC. As said before, our approach is to learn $\mathbb{F}_{\boldsymbol{w}}(a_1, \ldots, a_K)$ by exploiting the copula - marginals decomposition.

The learning step can now be performed constructing empirical copulas and marginals directly from the data, as

[1] Incidentally, we recall that Sklar's representation theorem [25] ensures that if the $w(k)$'s are continuous random variables then the $\mathbb{C}$ in (15) exists unique. If the random variables are mixed then the uniqueness is not ensured anymore, while the existence is preserved. This means that removing the continuity assumptions leads to complications when proving theoretical results, but does not affect the effectivity of practical and empirical estimation schemes.

in [22]. The empirical method nonetheless suffers whenever the $w_n(k)$ are not i.i.d. In this case it is preferable to let the various distributions (both marginals and the copula) belong to some parametric family, and explicit this dependence by writing the joint CDF for $\boldsymbol{w} = [w(1), \dots, w(K)]$ as

$$\mathbb{C}\left(\mathbb{F}_{w(1)}(a_1 \,;\, \boldsymbol{\beta}_1), \dots, \mathbb{F}_{w(K)}(a_K \,;\, \boldsymbol{\beta}_K) \,;\, \boldsymbol{\theta}\right). \tag{16}$$

(16) specifies that the marginals $\mathbb{F}_{w(k)}$ and the copula $\mathbb{C}$ depend respectively on the parameters $\boldsymbol{\beta}_k$ and $\boldsymbol{\theta}$. For a through list of possibilities see, e.g., [20].

Specifying probability distributions in parametric forms like (16) induces two questions, addressed in the next subsections:

1. given one specific parametric family for the $\mathbb{F}_{w(k)}$'s and one specific family for $\mathbb{C}$, how should one estimate $\boldsymbol{\beta}_k$ and $\boldsymbol{\theta}$ from the data?
2. given various different parametric families for the $\mathbb{F}_{w(k)}$'s and for $\mathbb{C}$, how should one choose which is the best family from the data?

3.4.1 Learning the parameters from the data

Delegating to the specific literature for more detailed descriptions, we notice that this task is usually solved using Maximum Likelihood (ML) approaches. I.e., denoting the likelihood of the dataset of the measurements $\boldsymbol{w}_1, \dots, \boldsymbol{w}_N$ as a function of some unknown parameters

$$\mathcal{L}(\boldsymbol{w}_1, \dots, \boldsymbol{w}_N \,;\, \boldsymbol{\beta}_1, \dots, \boldsymbol{\beta}_K, \boldsymbol{\theta}) \tag{17}$$

then one aims to find that particular vector of $\boldsymbol{\beta}_1^*, \dots, \boldsymbol{\beta}_K^*, \boldsymbol{\theta}^*$ that maximizes $\mathcal{L}$. We notice that, thanks to the separation between marginals and dependence introduced by the copulas formalism, it is often numerically convenient to adopt *inference functions for margins* approaches [13], i.e., estimate the $\boldsymbol{\beta}_k^*$'s (the marginals) separately by maximizing the marginal likelihood

$$\sum_{n=1}^{N} \left(\left. \frac{\partial \mathbb{F}_{w(k)}(a_k \,;\, \boldsymbol{\beta}_k)}{\partial a_k} \right|_{w_n(k)} \right) \tag{18}$$

with respect to $\boldsymbol{\beta}_k$, then insert these $\boldsymbol{\beta}_k^*$ in (17), and then eventually find the best $\boldsymbol{\theta}$.

We notice that these maximization steps are usually performed numerically by means of Newton or quasi-Newton methods, and that they can be performed online, i.e., incrementally as soon as new data arrive [11].

3.4.2 Selecting the proper copula family

Every particular choice for $\mathbb{C}$ induces a particular statistical dependency among the various $w(k)$: since there is no always-valid solution, each to-be-modeled quantity needs tailored considerations. Sending back the interest reader to [8, 5, 27, 1], we report that given a dataset $\boldsymbol{w}_1, \dots, \boldsymbol{w}_N$ and two parametric copulas $\mathbb{C}_1(\cdot \,;\, \boldsymbol{\theta}_1)$, $\mathbb{C}_2(\cdot \,;\, \boldsymbol{\theta}_2)$ as plausible hypotheses, then an approach for deciding which one to choose is to: *(i)* start computing an empirical copula $\widehat{\mathbb{C}}$ from the data; *(ii)* compute the optimal (given the data) parameters $\boldsymbol{\theta}_1^*$, $\boldsymbol{\theta}_2^*$ for respectively $\mathbb{C}_1$ and $\mathbb{C}_2$; *(iii)* choose between $\mathbb{C}_1$ and $\mathbb{C}_2$ that $\mathbb{C}_j$, $j = 1, 2$, that is closer to $\widehat{\mathbb{C}}$ in terms of an opportune metric, e.g., the quadratic residuals

$$\sum_{n=1}^{N} \left(\widehat{\mathbb{C}}\left(w_n(1), \dots, w_n(K)\right) - \mathbb{C}_j\left(w_n(1), \dots, w_n(K) \,;\, \boldsymbol{\theta}_j^*\right) \right)^2.$$

3.5 Extraction of samples from copulas

To extract a i.i.d. sample from a copula $\mathbb{C}$ corresponds to extract a scenario for the considered process. This can be done exploiting the general scheme: letting $\mathbb{C}_k(u_1, \dots, u_k) := \mathbb{C}(u_1, \dots, u_k, 1, \dots, 1)$ denote the k-dimensional margin for $\mathbb{C}$ and $\mathbb{C}_k(u_k \mid u_1, \dots, u_{k-1})$ the corresponding conditional distribution, then

- extract $\Omega_1 \sim \mathcal{U}[0, 1]$;
- extract $\nu_2 \sim \mathcal{U}[0, 1]$, and then compute that Ω_2 that satisfies $\nu_2 = \mathbb{C}_2(\Omega_2 \,;\, \Omega_1)$;
- ...
- extract $\nu_K \sim \mathcal{U}[0, 1]$, and then compute that Ω_K that satisfies $\nu_K = \mathbb{C}_K(\Omega_K \,;\, \Omega_{T-1}, \dots, \Omega_1)$.

The equations $\nu_k = \mathbb{C}_k(\Omega_k \,;\, \Omega_{k-1}, \dots, \Omega_1)$ are generally solved with numerical root-finding procedures. But if $\mathbb{C}$ belongs to some particular parametric family (e.g., Gaussian, T, Archimedean) then opportune closed forms lead to fast and reliable extraction procedures [6, Chap. 6].

4 Experimental Results

Description of the experimental setup

The testbed is located in the KTH Royal Institute of Technology campus in Stockholm. Its HVAC system is composed by two parts, see also Figure 3: the *ventilation system*, supplying fresh air, and the *heating system*, providing hot water to the radiators. The first pre-conditions fresh air from outside, canalizing it into a ventilation duct at a temperature of about 21°C. Part of this air is pushed directly into the room, part may be heated/cooled by a chiller circuit. The exhaust air is ejected by an additional duct. The actuators are dumpers for both the inflow / outflow ducts and the chiller circuit valve. The *heating system* is composed by radiators; the hot water flowing inside is regulated by means of a valve and is provided by a central system.

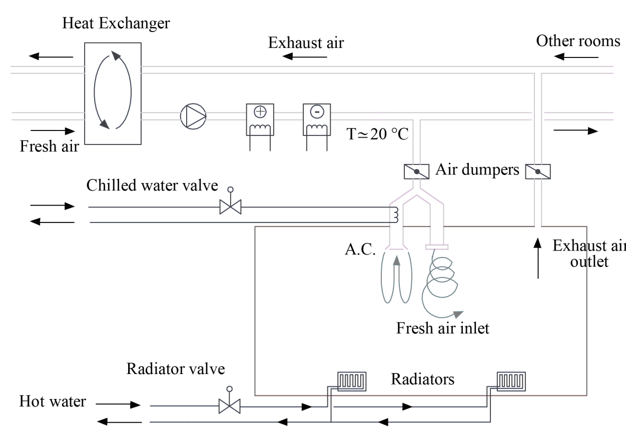

Figure 3. Scheme of the HVAC system of the testbed.

Figures 4 and 5 validate models (3) and (5) against data collected during the end of July 2013. We notice that the models capture the main dynamics, even if with a generalized smoothing effect. We believe that this error is induced by the map "damper opening percentage $\mapsto$ mass air flow $\dot{m}_{\text{venting}}$", provided for the test, which was not sufficiently accurate.

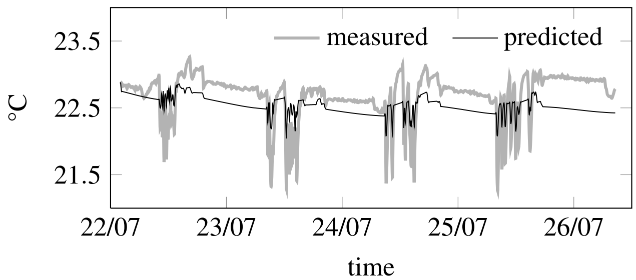

Figure 4. Validation of the thermal using the measured temperatures collected from the testbed.

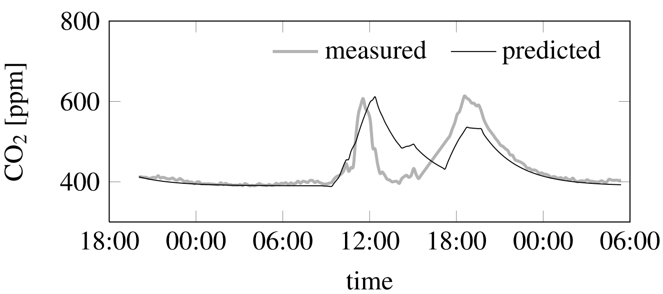

Figure 5. Validation of the CO_2 concentration model using the measured concentrations collected from the testbed.

Definition of the performance indexes

Out indexes are the total energy usage and the level of violations of the comfort bounds, calculated respectively as

$$E_{\text{tot}} = c_{\text{pa}} \sum_{k=0}^{N-1} \dot{m}_{\text{venting}}(k)\big(T_{\text{sa}}(k) - T_{\text{room}}(k)\big)\Delta k \quad [\text{kWh}],$$

$$C_h = \sum_{k \text{ s.t. } T_{\text{room}}(k) > T_{\text{UB}}} \big(T_{\text{room}}(k) - T_{\text{UB}}\big)\Delta k \quad [°\text{C h}].$$

T_{UB} in the equations above is the upper bound temperature of the comfort level, while Δk is the time between two samples.

Summary of the results

We compare two controllers: the current practice, a simple control logic with distinct PI control loops and switching logic, indicated by the acronym "AHC" (from Akademiska Hus, the company managing the building of the testbed), and our RMPC scheme. The controllers are tested respectively on August 5 and 6, 2013, both from 9:00 to 14:00, under similar occupancy patterns and with equivalent external weather conditions (sunny Swedish summer days). The sampling time for the RMPC was 10 minutes, while the predictions horizon for the weather, occupancy and solar radiance processes was 8 hours.

The results shown in Figure 6 clearly indicate that our RMPC controller outperforms the current practice in terms of both energy use and violations of the thermal comfort range ($21 - 23$ °C).

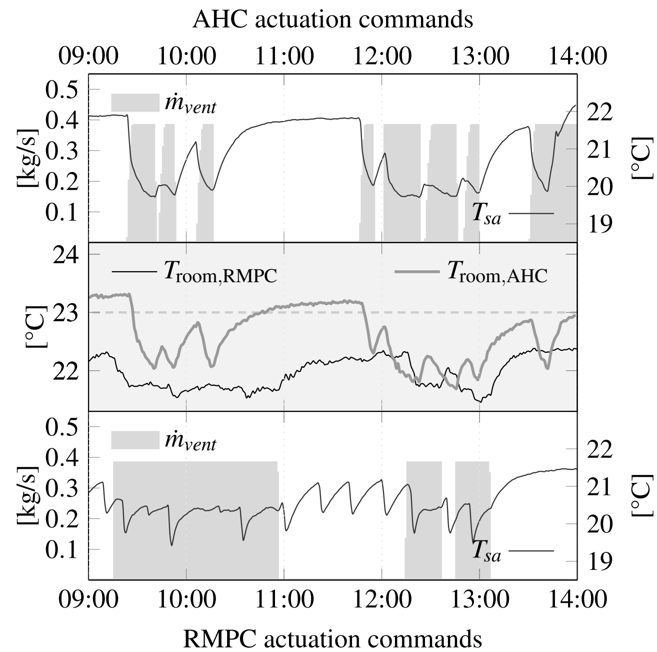

Figure 6. Comparison of the actuation levels computed by the AHC and the RMPC controller. Numerically, for the RMPC $E_{\text{tot}} = 1.275$kWh while for the AHC $E_{\text{tot}} = 1.392$kWh, approximatively 8.4% higher. At the same time, for the RMPC $C_h = 0$, while for the AHC $C_h = 0.2662$ °C h.

Namely, in Figure 6, it can be seen that the RMPC controller does not yield violations of the thermal comfort band, while the Proportional Integrative (PI) controller from Akademiska Hus has violations of the upper bound on the temperature . Moreover, the temperature variations are much smaller with RMPC, which is a more favorable behavior in terms of comfort.

The improvements can be explained by the control input profiles depicted in Figure 6, where it is shown the precooling effect. The ventilation system was scheduled to operate during the period with the lowest temperature (roughly from 9:00 to 11:00) so that the variations of the temperature profile of the inlet air, T_{sa}, are maintained as small as possible and less cooling energy could be used in the next hours.

5 Conclusions

We proposed a Stochastic Model Predictive Control (SMPC) controller for Heating, Ventilation and Air Conditioning (HVAC) systems, aiming to diminish the energy required to maintain indoor thermal comfort and good air quality levels. The mechanism to account for the probabilistic nature of the disturbances affecting the comfort indicators is a scenario-based one: the controller starts by sampling from the probability distributions of the disturbances, and then constructs from those samples some constraints on the evolution of the state of the system.

For robustness purposes, we endowed the algorithm with a learning module that infers the statistics of the disturbances from the data. This choice follows the trend of developing general control schemes, that can be installed without high or time-consuming deployment phases. Again for the sake of generality, we choose not to exploit Gaussian assumptions for the statistics of the disturbances, and opted for using copulas, a more computationally demanding but very flexible formalism that can handle every form of stochastic dependency among the various disturbances.

The strategy has then been implemented and tested on a real office, showing simultaneously that: *(i)* the computational burden of the SMPC plus the learning scheme can be managed by off-the-shelf devices; *(ii)* the actuation laws computed in this way are more effective than the current practice.

The good results achieved in real experimentations motivate efforts to improve the method. Probably the most important direction is towards the generalization of the control scheme to the case of whole buildings, which leads to increased complexity for both the models and the costs. Another very important achievement is to extend the learning capabilities of the scheme to arrive to a fully self-tunable and adaptable controller.

We eventually notice that there is still the need of measuring precisely and extensively the amount of energy savings / comfort maintaining performance of the strategy, to correctly evaluate, also monetarily, the degree of the improvements brought to the current practice.

6 References

[1] T. Ane and C. Kharoubi. Dependence Structure and Risk Measure. *The journal of business*, 76(3):411–438, 2003.

[2] BSI. En 15251:2007: Indoor environmental input parameters for design and assessment of energy performance of buildings addressing indoor air quality, thermal environment, lighting and acoustics. Technical report, British Standards Institute, 2008.

[3] G. Calafiore. Random convex programs. *SIAM Journal on Optimization*, 20(6):3427–3464, 2010.

[4] G. Calafiore and M. Campi. The scenario approach to robust control design. *IEEE Transactions on Automatic Control*, 51(5):742–753, 2008.

[5] X. Chen and Y. Fan. Estimation and model selection of semiparametric copula-based multivariate dynamic models under copula misspecification. *Journal of econometrics*, 135(1 - 2):125 – 154, 2006.

[6] U. Cherubini, E. Luciano, and W. Vecchiato. *Copula methods in finance*. Wiley, 2004.

[7] V. Erickson, Y. Lin, A. Kamthe, R. Brahme, A. Surana, A. Cerpa, M. Sohn, and S.Narayanan. Energy efficient building environment control strategies using real-time occupancy measurements. In *BuildSys2009*, pages 19–24, November 2009.

[8] C. Genest and L. Rivest. Statistical inference procedures for bivariate Archimedean copulas. *Journal of the American statistical association*, 88(423):1034–1043, 1993.

[9] S. Goyal, H. Ingley, and P. Barooah. Zone-level control algorithms based on occupancy information for energy efficient buildings. In *American Control Conference*, pages 3063–3068, June 2012.

[10] M. Gwerder and J. Toedtli. Predictive control for integrated room automation. In *CLIMA 2005*, 2005.

[11] A. Harvey. Dynamic distributions and changing copulas. Technical report, Cambridge Working Papers in Economics, University of Cambridge, 2008.

[12] J.Široký, F. Oldewurtel, J. Ciglerc, and S. Prívarac. Experimental analysis of model predictive control for an energy efficient building heating system. *Applied Energy*, 88(9):3079–3087, 2011.

[13] H. Joe. *Multivariate models and multivariate dependence concepts*. Chapman & Hall / CRC, 1997.

[14] P. Kall and J. Mayer. *Stochastic Linear Programming: Models, Theory, and Computation*. Springer-Verlag, 2005.

[15] Y. Ma and F. Borrelli. Fast stochastic predictive control for building temperature regulation. In *American Control Conference*, pages 3075–3080, June 2012.

[16] Y. Ma, F. Borrelli, B. Hencey, A. Packard, and S. Bortoff. Model predictive control of thermal energy storage in building cooling systems. In *48th IEEE Conference on Decision and Control and 28th Chinese Control Conference*, 2009.

[17] Y. Ma, S. Vichik, and F. Borrelli. Fast stochastic MPC with optimal risk allocation applied to building control systems. In *Conference on Decision and Control (CDC)*, pages 7559–7564, December 2012.

[18] A. E. D. Mady, G. Provan, C. Ryan, and K. Brown. Stochastic model predictive controller for the integration of building use and temperature regulation. In *Proceedings of the Twenty-Fifth AAAI Conference on Artificial Intelligence*, August 2011.

[19] M. Morari, J. Lee, and C. Garcia. *Model Predictive Control*. Prentice Hall, 2001.

[20] R. B. Nelsen. *An Introduction to Copulas*. Springer, second edition, 2006.

[21] F. Oldewurtel, A. Parisio, C. Jones, D. Gyalistras, M. Gwerder, V. Stauch, B. Lehmann, and M. Morari. Use of model predictive control and weather forecasts for energy efficient building climate control. *Energy and Buildings*, (45):15–27, February 2012.

[22] A. Parisio, M. Molinari, D. Varagnolo, and K. Johansson. A scenario-based predictive control approach to Building HVAC managementsSystems. In *IEEE Conference on Automation Science and Engineering*, August 2013.

[23] A. Patton. Copula methods for forecasting multivariate time series. In *Handbook of Economic Forecasting*, volume 2, pages 1 – 77. Elsevier, Oxford, second edition, May 2012.

[24] T. Salsbury, P. Mhaskar, and S. Qin. Predictive control methods to improve energy efficiency and reduce demand in buildings. *Computers & Chemical Engineering*, 51:77–85, August 2012.

[25] A. Sklar. Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, 8:229 – 231, 1959.

[26] D. Sturzenegger, D. Gyalistras, M. Gwerder, C. Sagerschnig, M. Morari, and R. S. Smith. Model Predictive Control of a Swiss office building. In *Clima-RHEVA World Congress*, June 2013.

[27] P. K. Trivedi and D. M. Zimmer. Copula Modeling: An Introduction for Practitioners. *Foundations and Trends in Econometrics*, 1(1):1 – 111, 2006.

[28] X. Zhang, G. Schildbach, D. Sturzenegger, and M. Morari. Scenario-based MPC for energy-efficient building climate control under weather and occupancy uncertainty. In *European Control Conference*, July 2013.