UNIVERSITÀ
DEGLI STUDI
DI PADOVA

DIPARTIMENTO
DI INGEGNERIA
DELL'INFORMAZIONE

Energy Efficient Control and Fault Detection for HVAC Systems

Ph.D. Candidate

Francesco Simmini

Advisor

Prof. Alessandro Beghi

School Director

Prof. Matteo Bertocco

Coordinator

Prof. Carlo Ferrari

University of Padova

Department of Information Engineering

Ph.D. School in Information Engineering

XXVI Series, 2014

UNIVERSITÀ
DEGLI STUDI
DI PADOVA

Sede Amministrativa: Università degli Studi di Padova

Dipartimento di Ingegneria dell'Informazione

SCUOLA DI DOTTORATO DI RICERCA IN: INGEGNERIA DELL'INFORMAZIONE
INDIRIZZO: SCIENZA E TECNOLOGIA DELL'INFORMAZIONE
CICLO XXVI

Energy Efficient Control and Fault Detection for HVAC Systems

Direttore della Scuola: Ch.mo Prof. Matteo Bertocco

Coordinatore d'indirizzo: Ch.mo Prof. Carlo Ferrari

Supervisore: Ch.mo Prof. Alessandro Beghi

Dottorando: Francesco Simmini

Summary

The interest in HVAC (Heating, Ventilation and Air-Conditioning) technology has rapidly increased in the last years. HVAC systems have become important in the design of medium-large buildings in order to ensure thermal comfort in the environments with respect to the temperature and humidity of the air. Control, optimisation and maintenance procedures are fundamental in HVAC systems in order to guarantee people comfort and energy efficient solutions in their management.

Two different topics are covered in this thesis.

Energy Efficient Control of Ice Thermal Energy Storage Systems

HVAC plants have recently begun to be matched with thermal energy storage systems. If properly designed, installed, and maintained, these systems can be used to store energy when its cost is low and exploiting it when the price increases. In particular, in HVAC cooling systems, a common thermal storage medium is ice. From a control and optimisation point of view, a cooling plant with ice storage proves to be a complex system. Standard control strategies seem not to be able to achieve the right trade-off between energy efficiency and demand satisfaction.

In this thesis, in order to design efficient control strategies for storage systems, a HVAC model with ice storage is developed in a simulation environment. The thermal behaviour of the HVAC system is derived from the mass and energy conservation equations; in particular the ice storage is considered a hybrid system, thus taking into consideration both sensible and latent heat. Three standard control methods are compared with a non-linear predictive con-

trol strategy. The simulations results show that the implemented non-linear predictive control strategy provides the best control for the efficient energy management of ice storage systems.

Fault Detection in HVAC Systems

Operating problems associated with degraded equipment, poor maintenance, and improperly implemented controls, plague many HVAC systems. Fault detection methods can therefore play a key role in monitoring complex HVAC plants, detecting anomalous behaviours in such a way as to keep the systems in their best operational conditions with minimum costs.

In this thesis, fault detection and diagnosis methods on variable air volume (VAV) systems are first designed. To this aim, a VAV system model with two zones is developed; the control of system is obtained with a direct feedback linearisation technique. Supervised classification methods are used to detect and diagnose the simulated faults in the model. The simulations results show the good performances of the classification in the detection and diagnosis of the most common faults in VAV systems.

Detection methods are then developed for the most relevant faults affecting chillers. To this aim, data collected in the research project 1043-RP promoted by ASHRAE (American Society of Heating, Refrigerating and Air Conditioning Engineers) are used. In this project experimental studies were conducted on a centrifugal water-cooled chiller in order to collect data in both normal and faulty situations. The developed technique is based on one-class classification methods with a novelty detection approach, where only normal data are used to characterize the correct system behaviour. The classification results confirm the effectiveness of the proposed method for the detection of the most common faults in chillers.

Sommario

Negli ultimi anni l'interesse per la tecnologia HVAC (*Heating, Ventilation and Air-Conditioning*) è rapidamente cresciuto. I sistemi HVAC sono diventati importanti nella progettazione di edifici medio-grandi al fine di assicurare comfort termico negli ambienti rispetto alla temperatura e umidità dell'aria. Procedure di controllo, ottimizzazione e manutenzione sono fondamentali nei sistemi HVAC al fine di garantire il comfort delle persone e soluzioni energeticamente efficienti nella loro gestione.

In questa tesi vengono trattati due diversi argomenti.

Controllo Energeticamente Efficiente di Sistemi ad Accumulo di Energia Termica a Ghiaccio

Recentemente si è cominciato ad abbinare impianti HVAC con sistemi ad accumulo di energia termica. Se adeguatamente progettati, installati e mantenuti, questi sistemi possono essere usati per accumulare energia quando il suo costo è basso e sfruttandola quando il prezzo aumenta. In particolare, nei sistemi HVAC per raffreddamento, un mezzo diffuso per accumulo termico è il ghiaccio. Da un punto di vista di controllo e ottimizzazione, un impianto per raffreddamento con accumulo a ghiaccio si dimostra essere un sistema complesso. Strategie di controllo standard non sembrano essere capaci di ottenere il giusto compromesso tra efficienza energetica e soddisfacimento della domanda.

In questa tesi, al fine di progettare strategie di controllo efficiente per sistemi con accumulo, si sviluppa in un ambiente di simulazione un modello HVAC con accumulo a ghiaccio. Il comportamento termico del sistema HVAC viene derivato dalle equazioni di conservazione di massa e energia; in parti-

colare l'accumulo a ghiaccio viene considerato un sistema ibrido, tenendo così in considerazione sia il calore sensibile sia quello latente. Tre metodi di controllo standard vengono confrontati con una strategia di controllo predittivo non-lineare. I risultati delle simulazioni mostrano che la strategia di controllo predittivo non-lineare implementata fornisce il migliore controllo per una gestione energeticamente efficiente di sistemi con accumulo a ghiaccio.

Rilevamento Guasti in Sistemi HVAC

Problemi di funzionamento associati ad apparati degradati, scarsa manutenzione, e controlli erroneamente implementati, affliggono molti sistemi HVAC. Metodi di rilevamento guasti possono dunque giocare un ruolo chiave nel monitorare impianti HVAC complessi, rilevando comportamenti anomali in modo da mantenere i sistemi nelle loro migliori condizioni operative a costo minimo.

In questa tesi vengono dapprima progettati metodi di rilevamento e diagnosi guasti su sistemi a tutta aria a portata variabile (VAV- *Variable Air Volume*). A questo scopo, viene sviluppato un modello di un sistema VAV con due zone; il controllo del sistema viene ottenuto con una tecnica di *direct feedback linearisation*. Metodi di classificazione supervisionati vengono usati per rilevare e diagnosticare i guasti simulati nel modello. I risultati delle simulazioni mostrano le buone performance della classificazione nel rilevamento e diagnosi dei guasti più diffusi nei sistemi VAV.

Vengono successivamente sviluppati metodi di rilevamento per i guasti più rilevanti che affliggono i refrigeratori di liquido (*chiller*). A questo scopo, vengono utilizzati i dati raccolti nel progetto di ricerca 1043-RP promosso da ASHRAE (*American Society of Heating, Refrigerating and Air Conditioning Engineers*). In questo progetto sono stati condotti studi sperimentali su un *chiller* centrifugo raffreddato ad acqua al fine di raccogliere dati sia in situazioni normali sia in situazioni in presenza di guasto. La tecnica sviluppata si basa su metodi di classificazione a una classe con un approccio di *novelty detection*, dove solamente dati normali vengono utilizzati per caratterizzare il comportamento corretto del sistema. I risultati di classificazione confermano l'efficacia del metodo proposto per il rilevamento dei guasti più diffusi nei *chiller*.

Contents

HVAC Systems

HVAC (Heating, Ventilation, and Air Conditioning) systems are important in the design of medium and large buildings, such as hospitals, factories, etc. where safe and healthy conditions are regulated with respect to temperature and humidity. HVAC systems operating principles are based on the laws of thermodynamics, and on many discoveries and inventions that have been made since the end of 1800s.

In this introductory chapter, first the main components of the HVAC systems are described. Many of the listed components will be considered for study and modeling in this thesis. A possible distinction between HVAC systems is then done, based on the fluid that is used to exchange heat with the environment [29]. Since in this thesis HVAC cooling mode operations will be considered, the vapour-compression cycle is finally described, since it is used in most HVAC refrigerating units.

HVAC Components

Boilers A boiler is a closed container in which the water, or another fluid, is heated under pressure. The fluid is then circulated out of the boiler in a system of ducts to be used in different applications related to heating.

Chillers A chiller is a machine that removes heat from a liquid through a vapour compression cycle, or through a heat absorption cycle. The liquid that is cooled is usually water, although there can be other substances in a variable percentage, as inhibitors of corrosion or other additives. The cold outlet water from the chiller is then used to cool and dehumidify the air in commercial, industrial and institutional buildings of medium and large size or for other applications in the field of refrigeration. Different plant configurations can be found: some chillers are "assembled" on site, since the chiller components (compressor, heat exchangers, etc.) are purchased separately and then assembled in a specific place; mono-bloc chillers (more recently used) are instead sold as single machines already assembled that are simply connected to the system. Depending on the mode of heat exchange with the external environment, chillers are also divided into air-condensed ones (heat is transferred to the outside air, Figure 1) and water-condensed ones, where the heat is extracted from the environment to be transferred to a second water circuit with cooling towers.

Figure 1: Air-condensed chiller by Rhoss S.P.A.

Air conditioners An air conditioner is a device designed to extract or provide heat in the environments. Air conditioners are mainly used as heating and cooling household systems and for the cooling inside vehicles. A system that, in addition to conditioning, also guarantees the control of heating and ventilation becomes a HVAC system.

Air ducts Air ducts are used to provide and remove air in the environments. Appropriate air quality and speed are very important in order to ensure people thermal comfort. To this end, air ducts usually have filtering systems to remove dust and bacteria in the air that is introduced in the zones.

Diffusers A diffuser is a mechanical device that is designed to control the characteristics of a fluid at the entrance of the rooms. Diffusers are used to slow the fluid's velocity, to enhance its mixing into the surrounding fluid and then to distribute the air flow in the desired directions in the rooms. Diffusers usually have different forms, depending on the specific application.

Air-filters Filters are useful for removing contaminants from air, mainly solid particles such as dust, pollen, mold and bacteria. The materials used for filtering may be of different types. They range from fibers, such as cotton, to other materials that attract particles by means of static electricity.

Grids Grids are part of the terminals of an air distribution system, being mainly used as input units of the exhaust air in the ducts. In some cases, grids can also be used as exit points for the air that is supplied into the zones.

Radiant panels Radiant panel heating systems provide heat through pipes located behind the surfaces of the environments. Three kinds of radiant heating

can be distinguished:

- radiant floor heating: a hydraulic circuit with water temperature of 30/40°C flows under the floor; floor panels are mainly used in residential buildings;
- radiant wall heating: wall panels are mainly used as integration of other heating systems such as those behind the floor, for balancing the high heat losses in certain environments;
- radiant ceiling heating: ceiling panels are used where there is the need to heat very large rooms (industrial buildings, etc.).

Pumps A pump is a device that moves fluid by mechanical action. Pumps usually operate by reciprocating or rotary mechanism and they can operate at a constant speed, or offer a variety of speeds, depending on the specific application.

Heat exchangers A heat exchanger is a device built to exchange heat efficiently from one fluid to another, without any mixing between the two fluids. Heat exchangers often take the name of cooling coils (or heating), for the form often adopted by the pipes. The most commonly employed fluids are water, a solution of glycol and water, steam or refrigerant, depending on the particular application.

Cooling towers Cooling towers (Figure 2) are heat removal devices used to transfer process heat to the atmosphere. Cooling towers use the evaporation of water to remove heat and cool the working fluid to near the wet-bulb air temperature. They are mainly used coupled with water-condensed chiller systems in order to reject to the atmosphere the heat from the building zones.

Pipes Pipes are used for the transport of fluids. Pipe materials may be different: glass, aluminium, steel, plastic. In the pipes there may be different components, such as valves or other equipment, which typically measure and control the pressure, flow and temperature of the fluid.

Humidifiers A humidifier is a device that increases humidity in the environments. The most common humidifier, the “evaporative” one, consists of few basic parts: a reservoir, a wick and a fan. The wick is made of a porous

Figure 2: Cooling tower by Rhoss S.P.A.

material that absorbs water from the reservoir and provides a large surface area. The fan blows air onto the wick to ensure the evaporation of the water.

Dehumidifiers A dehumidifier is a device which reduces the level of humidity in the air, usually for health or comfort reasons, or to eliminate dank smells. Dehumidifiers extract water from the air, by cooling it and reaching the condensation temperature. A disposal system is needed to discard the condensate water.

Air handling units An air handler, or Air Handling Unit (AHU), is a device used to handle the air in a heating, ventilation, and air-conditioning system. An air handler is usually a large metal box containing a blower, heating or cooling elements, filters, sound attenuators, and dampers. The air that is treated in air handling units is usually a mixture of recirculated air from environments and external air. This solution allows both the renewal of the air in the rooms (through the contribution of external air) and the reduction of energy consumption. Usually the air handler is connected to ducts for air distribution, carrying air throughout the building and then returning it to the AHU.

Valves A valve is a device that regulates, directs or controls the flow of a fluid by opening, closing, or partially obstructing its passageways.

Fans A mechanical fan is a device used to create flow within a fluid (gas or air). A fan usually consists of rotating blades which act to move the fluid. Depending on their use, fans may operate at a single speed or offer a variety of speeds.

Dampers A damper is a particular valve or plate that stops or regulates the air flow inside a duct, or other air handling equipments. In particular, dampers are used to control the air flow that enters in the zones of a building.

Tanks Tanks are often employed within HVAC systems to provide additional system fluid volume in order to increase system inertia and prevent short cycling of heating or cooling apparatus.

Thermal energy storages Thermal energy storage technology is used to store energy (with a storage medium) when energy costs are high. The stored energy is used during the on-peak period when the energy price increases. If the thermal energy storage is well managed through smart control and optimisation procedures, it can ensure reduced energy costs and energy consumption, and decreased maintenance costs for the HVAC system.

HVAC Systems Classification

The classification of air conditioning systems can be made according to the fluid used to regulate the temperature and humidity in the environments. Using this criterion we can distinguish [29]:

All-air systems

All-air systems are the most common installations for heating and cooling applications. They control the air temperature of the zones by supplying heating or cooling air from a central source via a network of air ducts. These systems increase or decrease the room temperatures by changing either the volume or temperature of the supplied air. Since air velocity is important for thermal comfort, most buildings that require cooling employ all-air systems, providing the best control of the quality, temperature and humidity of the air. The components of an all-air HVAC system usually include an air handling unit which includes a fan, heating and cooling coils, filters to clean the air, and often elements to humidify it. Dehumidification, when required, is accomplished by cooling the air below the dew-point temperature and then reheating it. The conditioned air from the AHU is supplied to the spaces by a network of air ducts and then air returns from the conditioned spaces through a parallel network of return-air ducts. All-air systems also include a duct which supplies

external air to the AHU and one which pulls out some or all of the exhaust air to the external environment.

All-water systems

Water can be used for carrying energy in both heating and cooling systems. Water actually carries more energy per unit volume than air and therefore requires the least space for piping. All-water distribution systems have a relatively low installation cost when compared to all-air systems. The minimal space required for distribution piping makes them an excellent choice for retrofit installation in existing buildings or in structures with significant space constraints. The disadvantage of these systems is that they provide little or no control over air quality and humidity.

Air and water systems

All-water systems usually provide low humidity levels during winter, while humidity values are closely related to internal loads during summer. Therefore, when an accurate humidity control is required, mixed air-water installations can be used [29].

The Refrigeration Cycle

The refrigeration cycle is the key process for HVAC applications that require cooling and dehumidification. The vapour-compression refrigeration cycle is used in most household air-conditioning systems as well as in many large commercial and industrial chiller systems. Figure 3 shows the vapour-compression refrigeration cycle: it consists of compressor, condenser, expansion valve and evaporator.

Compressor

The mechanical compressor has two main tasks:

- to pump refrigerant through the cooling system;
- to compress gaseous refrigerant in the system so that it can be condensed to liquid releasing the absorbed heat from the building environments.

Condenser

Condensers are typically refrigerant-to-air (in air-cooled systems) or refrigerant-to-water (in water-cooled systems) heat exchangers which are used to get rid of the heat extracted from the interior of the buildings. The condenser receives the high-pressure gas from the compressor and convert this gas to a liquid: air or water passing over the condenser coils carries off the heat and as a consequence the gas condenses.

Expansion valve

Expansion valves are flow-restricting devices that cause a pressure drop of the refrigerant. In fact, the refrigerant enters the valve as a high-pressure liquid. The refrigerant flow is then restricted by an orifice, causing its change from a high-pressure liquid to a low-pressure liquid. After leaving the expansion valve the refrigerant enters the evaporator.

Evaporator

The evaporator is a heat exchanger that allows the refrigerant to evaporate from liquid to gas while absorbing heat coming from the environments: when the liquid refrigerant reaches the evaporator its pressure has been reduced by the expansion valve, dissipating its heat content; this causes the refrigerant to absorb heat from the environments and then to vaporize. The absorbed heat is then carried by the refrigerant from the evaporator to the compressor, from which the whole refrigeration cycle is repeated.

The thermodynamic properties of the cycle can be analysed on the diagram shown in Figure 4, [29]. From point 1 to point 2, the vapour is compressed at constant entropy and exits the compressor superheated. From point 2 to point 3 and on to point 4, the superheated vapour travels through the condenser which first cools the vapour and then condenses it into a liquid by removing additional heat at constant pressure and temperature. Between points 4 and 5, the liquid refrigerant goes through the expansion valve where its pressure abruptly decreases. That results in a mixture of liquid and vapour at a lower temperature and pressure as shown at point 5. The cold liquid-vapour mixture then travels through the evaporator coils and it becomes vaporized by cooling the warm air. The resulting refrigerant vapour returns to the compressor at point 1 to complete the thermodynamic cycle.

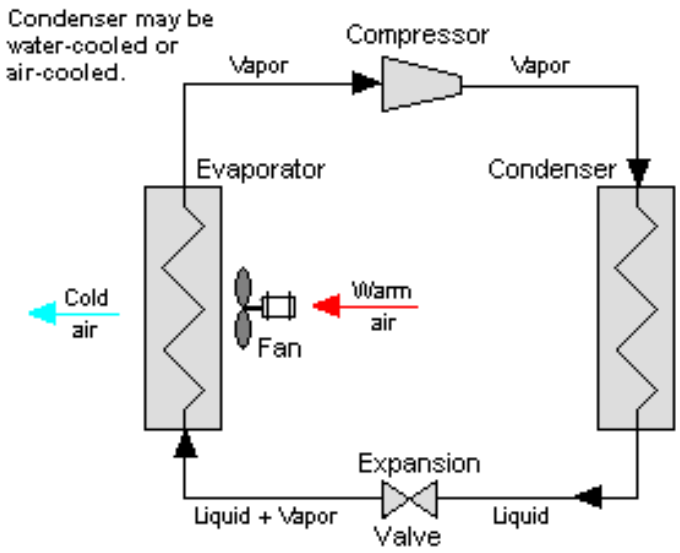


Figure 3: Vapour-compression refrigeration system.

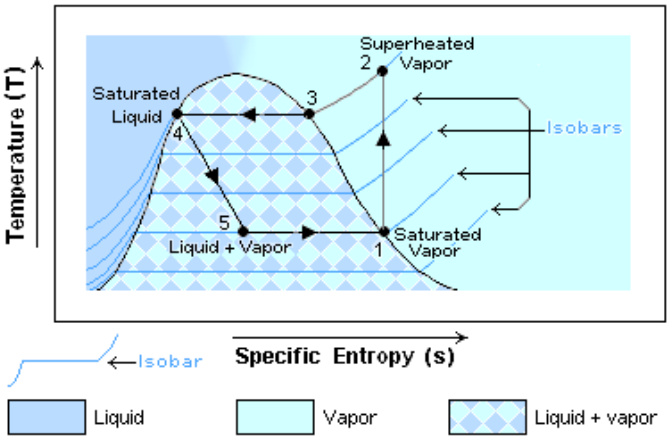


Figure 4: Thermodynamics of the vapour-compression cycle.

Part I

Energy Efficient Control of Ice Thermal Energy Storage Systems

1

Introduction

HVAC (Heating, Ventilation, and Air Conditioning) systems are designed to ensure safe and healthy conditions in the environments of medium and large buildings. The three tasks of Heating, Ventilation, and Air-Conditioning are interrelated, guaranteeing people thermal comfort and providing good air quality.

A fundamental point to consider for HVAC plants is their energy consumption, making extremely important to obtain the best use of the operative units. To minimise energy consumptions, optimisation procedures are required, whose objective is to define the operational procedures that lead to the fulfillment of the desired conditions by minimising operating costs of the system. HVAC systems have to provide healthy and comfortable conditions for the occupants, but this must be done efficiently by using the available resources and reducing the emission of pollutants into the air and water. It comes as no surprise that much of HVAC control and optimisation is a compromise: a trade-off that results in reasonable comfort at minimum energy use (Figure 1.1) and financial costs.

Recently, energy efficiency politics have encouraged the adoption of different time slots of energy prices. In order to exploit low price slots, HVAC plants have begun to be matched with Thermal Energy Storage (TES) systems. Such systems allow accumulating energy and using it in subsequent moments. The use of TES systems often results in significant benefits, such as: reduced energy costs and energy consumption, increased flexibility of operation, decreased initial and maintenance costs, reduced equipment size, more efficient and effective utilization of equipment, conservation of fossil fuels (by facilitating more efficient energy use) and reduced pollutant emissions (e.g. CO_2) [9], [30].

In particular, in HVAC cooling systems, a popular thermal storage medium

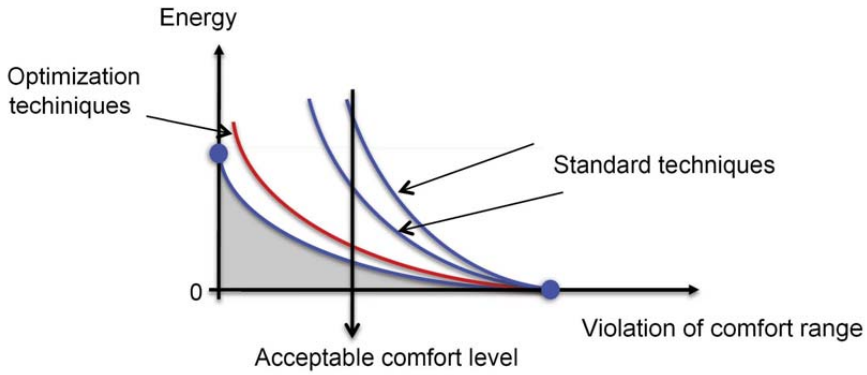


Figure 1.1: Energy-Violation of comfort: optimisation techniques ensure acceptable comfort levels with reduced energy costs.

is ice: the cooling capacity of an ice Cold TES (ice-CTES) system under total freezing is 18 times as high as that of a water-CTES system operating between 12°C and 7°C [10]. An ice-CTES has operating phases, namely, a charging phase where (typically at night) heat is removed from water to produce ice, and a discharging phase, where, when the building requires cooling, heat is removed from the building and added to the ice. The melted ice is reused during the next charging period. The advantage of this cooling scheme is that the main electrically driven device in cooling systems, namely, the compressor, is operated during low-electrical cost periods, i.e. at night [10]. However, experience with operating TES systems demonstrates that poor design and operation of the control systems can lead to bad energy efficiency [1]. It is worthwhile noting that a cooling plant with TES is a complex system. Highly non-linear behaviour and strong cross coupling of inputs and outputs make its modeling and control a non-trivial task. Classical control design methods seem not to be able to achieve the right trade-off between energy efficiency and demand satisfaction: they can not provide reasonable comfort at minimum energy use and financial costs.

Different approaches have been proposed in the literature in order to find suitable controls for TES systems. In [15] a simulation environment for the analysis of ice storage controls is presented. In [16] Henze et al. provide some guidelines to obtain an improvement of TES system energy performances; these guidelines are derived from the analysis of optimal control and its comparison to TES system standard control strategies. In [14] a predictive control design for a three-story office building equipped with two chillers with constant

coefficient of performance and a thermal energy storage system is illustrated.

In this part of the thesis, a model-based approach is developed to design efficient control strategies for HVAC systems equipped with ice-CTES. The thermal behaviour of the HVAC plant is analysed by resorting to a lumped formulation of the conservation equations and a simulation environment is designed accordingly. The contribution of this work is twofold. The first contribution regards modeling of the ice-CTES. Differently from what has been proposed so far in the literature, the ice-CTES is modelled as a hybrid system, where the water phase transitions (solid-melting-liquid, liquid-freezing-solid) are described by combining continuous and discrete dynamics, so that both latent and sensible heats are considered. It is worth noting that the ice latent heat represents a large portion of the total storage energy (i.e. 90%). However it is appropriate to consider also the sensible heat portion in order to adequately evaluate the energy efficiency of the HVAC system when operated with different control strategies. The second contribution is the comparison, in the developed simulation environment, of three standard control strategies (constant-proportion control, chiller-priority control and storage-priority control) and an advanced control strategy based on Model Predictive Control (MPC), which has been successfully used in building cooling systems with water-TES, [23] and [24]. The optimisation step is developed by using a stochastic technique, the Particle Swarm Optimisation (PSO) algorithm, that has already been proved to be a practical solution in energy-related industrial applications [3]. The simulations show that model predictive control provides the best control solution for the efficient management of ice-CTES systems.

The first part of thesis is organized as follows:

- In Chapter 2 the thermal energy storage technology is presented. The differences between sensible and latent TES systems are highlighted. Finally, two different management solutions (full-storage and partial-storage) for TES systems are described.

- In Chapter 3 the considered ice cold thermal energy storage plant is modelled. Standard control strategies for TES systems are compared with a predictive control strategy, proving that MPC provides the best control in terms of energy efficiency and demand satisfaction.

- Some concluding remarks are given in Chapter 4.

2

Thermal Energy Storage

Thermal Energy Storage (TES) is one of the key technologies for energy conservation, and therefore, it is of great practical importance for heating and cooling thermal applications. TES technology is perhaps as old as civilization itself. Since recorded time, people have harvested ice and stored it for subsequent use. For example, during the fourth century BC, Persian engineers mastered the technique of storing ice in the "Yakhchal" in the desert during summer. The ice could be brought in during the winters from the nearby mountains, but more often a wall close to the "Yakhchal" along east-west direction was made. In winter, the water was channeled to the north side of the wall. The cold temperatures and the shadow of the wall made the water freeze quickly. The "Yakhchal" constructions had a domed shape above ground (Figure 2.1) but had a subterranean storage space up to 5000 m^3 that had thick walls made out of a special mortar composed of sand, clay, egg whites, lime, goat hair, and ash in specific proportions, and which were resistant to heat transfer. The ice was then used to chill food during hot summer days and to make the "Faloodeh", the traditional Persian frozen dessert.

TES systems have been employed in more recent history for many applications, ranging from solar hot water storage to building air conditioning systems. The TES technology has only recently been developed to a point where it can have a significant impact on modern technology. It is a key component of many successful thermal systems: a good TES should allow little thermal losses, leading to energy savings, while permitting the highest reasonable extraction efficiency of the stored thermal energy.

Figure 2.1: "Yakhchal" in Yazd, Iran.

2.1 Cold Thermal Energy Storage

Cold Thermal Energy Storage (CTES) technology is an innovative way of storing night-time off-peak energy for daytime peak use for cooling applications, becoming one of the primary means of addressing the electrical power imbalance between high daytime demand and high night-time abundance. In many locations, demand for electrical power peaks during summer. Often, at night, the electricity is much less expensive. As a consequence, it is convenient to store energy during night and release it during daytime, shifting peak cooling loads to off-peak periods.

A CTES consists of a storage cold medium, a container, and input/output devices. Containers must both retain the storage material and prevent losses of thermal energy. Cooling energy can be stored either by chilling or freezing water (or other materials as glycol and eutectic salts). Water is the storage material of choice for a variety of practical and thermodynamic reasons, including its ready availability, relative harmlessness, and its compatibility with different equipments.

Two different CTES system typologies can be distinguished in HVAC plants:

- sensible CTES;
- latent CTES.

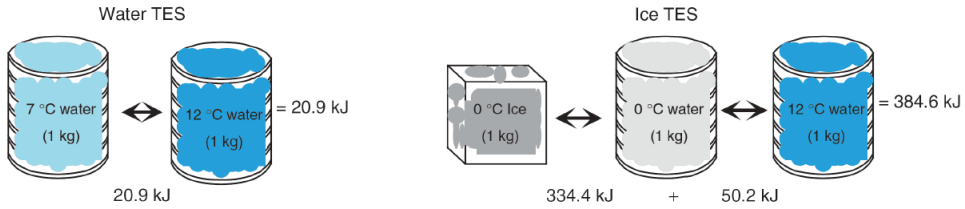


Figure 2.2: Water and ice capacities.

2.1.1 Sensible CTES Systems

In sensible CTES, energy is stored or released by changing the temperature of the storage medium. Each medium has its own advantages and disadvantages. The high heat capacity of water (4.186 kJ/kg K) often makes water tanks a logical choice for CTES systems that operate in a temperature range for building cooling applications. The amount of energy stored in a sensible CTES can be expressed as:

$$Q = mc_p\Delta T = \rho c_p V \Delta T\,, \tag{2.1}$$

where c_p is the specific heat of the storage medium, ΔT is the temperature change, V is the volume of the medium and ρ is its density.

2.1.2 Latent CTES Systems

In latent CTES, the heat transfer occurs when a substance changes from one phase to another and the energy is stored by allowing the substance to change its phase at a constant temperature. The main advantage of latent CTES systems is a higher capacity with respect to sensible CTES systems (bringing to a dimension reduction of the storage), and a smaller temperature operative range since the heat transfer occurs at constant temperature. Figure 2.2 shows that the cooling capacity of an ice-CTES system under total freezing is 18 times as high as that of a water-CTES operating between 12°C and 7°C. It can be seen that a big amount of energy can be exploited when water temperature is 0°C, when it is in two-phase conditions (transition time). The amount of energy that can be exploited during the transition phases is:

$$Q_{tran} = m \cdot \lambda\,, \tag{2.2}$$

where m is the mass of the medium and λ is the specific latent heat (e.g. in the water λ = 334,4 kJ/kg).

2.2 CTES Systems Management

Two different management solutions can be adopted when setting the operative conditions of a CTES system.

2.2.1 Full-Storage CTES

A full-storage strategy shifts the entire peak cooling load to off-peak hours. The system is typically designed to operate during night at full capacity during all non-peak hours in order to charge the storage. Full-storage (load-shifting) designs use storage to fully decouple the operation of cooling generating equipment (e.g. chiller) from the peak cooling load. The peak cooling load is met through the use (i.e. discharging) of the storage while the cooling generating equipment is idle.

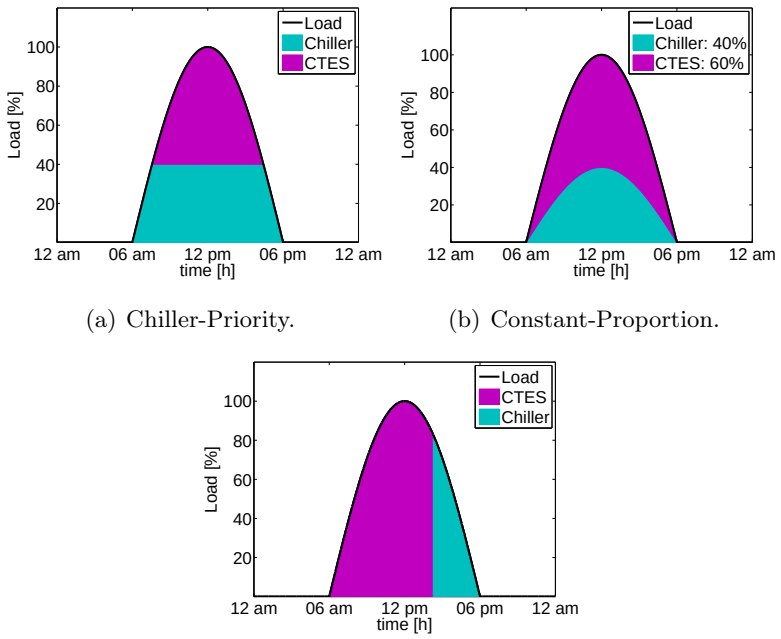
2.2.2 Partial-Storage CTES

In a partial-storage strategy, the cooling generating equipment (e.g. chiller) operates to meet part of the peak-period cooling load, and the rest is met by the storage. The chiller is sized at a smaller capacity than the design load. Although partial-storage does not shift as much load (on a design day) as a full-storage system, partial-storage systems can have lower initial costs, particularly if the design incorporates smaller equipment by using low-temperature water and cold-air distribution systems.

The following standard control strategies are usually adopted to manage partial-storage CTES systems :

- Chiller-Priority Control (Figure 2.3(a)): The simplest of the existing control strategies for thermal energy storage is chiller-priority control. Here the chiller runs continuously under conventional chiller control (direct cooling), subject to a power-limit while the storage provides the remaining cooling power if required. In this strategy a complete discharge of the storage is not guaranteed.

- Constant-Proportion Control (Figure 2.3(b)): This strategy implies that the storage meets a constant fraction of the cooling load under all conditions. Thus, neither chiller nor the storage have priority in providing cooling. Constant-proportion control is rather easy to implement by assigning a fixed fraction of the total temperature difference between the



(a) Chiller-Priority.

(b) Constant-Proportion.

(c) Storage-Priority.

Figure 2.3: Partial-storage standard control strategies.

water supply and return flow to be realized by the storage and the remainder by the chiller. Finding the best load fraction for each application is a matter of trial and error. Caution should be exercised that the chiller can always meet its load fraction.

- Storage-Priority Control (Figure 2.3(c)). Storage-priority control exploits as much stored energy as possible during the on-peak period. It is generally defined as that control strategy that aims at fully discharging the available storage capacity over the on-peak period. Chiller begins to meet load only after the full discharge of the CTES.

3

Modeling and Control of Ice-CTES Systems

Modeling and control of TES systems play a key role in HVAC systems research. By resorting to control and optimisation techniques, an efficient energy management of TES systems can be achieved. Significant energy savings can be obtained if the TES is well managed, whereas the electrical energy consumption markedly increases without proper control and optimisation procedures. An ice-CTES system is taken as a reference in this study. A partial-storage management is adopted: during the night the storage is fully charged, while during the daytime both storage and chiller are used to satisfy cooling demand. A non-linear predictive control strategy is adopted to efficiently manage the CTES discharging, whereas a simple PI-controller is used during the charging phase: in fact in this system control design it is always useful to recharge the ice-CTES completely for both high load and low load conditions coming in the following daytime due to the low energy price during the night. A simple controller can be used during the charging phase.

3.1 Plant Description and Modeling

In Figure 3.1 the block structure of the considered system is reported. We can clearly distinguish:

- The energy production section: it consists of two parallel air-condensed chillers. One chiller is used during night in order to charge the ice cold thermal energy storage; during day the other one and the ice-CTES are used together to satisfy cooling demand.
- The hydraulic section: a common primary-secondary pumping arrangement with by-pass is adopted. The hydraulic section is composed of

pipes, a temperature-modulating control valve and a diverting valve that allows charging/discharging operations.

- The load section: the building thermal load and capacity are represented in the simulation scheme by the load block and a water tank of suitable capacity.

The following scenario is an example of a partial-storage system at nominal conditions. During off-peak nighttime hours the nighttime-chiller (ChN) works as ice maker: a glycol solution (e.g. 25% ethylene glycol) is pumped through the chiller coils and the CTES in the chilled-water loop (VA-CD = 0, Figure 3.1). The -5°C ethylene glycol produced by the chiller freezes the water contained inside the CTES and charges it for use during the next day's cooling. Ice-making has the effect of de-rating the nominal chiller capacity by approximately 30-35%. Compressor efficiency, however, varies only slightly because lower nighttime temperatures result in cooler condensation temperatures and help to keep the unit operating efficiently (reducing the compressor pressure ratio). A full charging cycle of an ice tank requires approximately 6-12 hours, depending on its size. During the discharge cycle in the following day (VA-CD = 1, Figure 3.1), the glycol solution pre-cooled by the daytime-chiller (ChD) is further cooled by the ice-CTES. The temperature-modulating valve, in the bypass loop around the TES, permits a sufficient quantity of glycol solution to bypass the storage, mixed with solution arriving from the CTES, and allows one to achieve the desired temperature of the supplied glycol solution that is distributed to the cooling devices (e.g. air-handler coils). The solution leaving the building re-enters the chiller and is cooled again.

The thermal behaviour of the plant can be usefully analysed by a lumped formulation of the conservation equations. The elements of the plant are simulated through blocks, and the heat transfer processes are considered as concentrated inside the blocks. Furthermore, the following hypotheses are introduced:

- The water thermal properties are considered constant;
- The water is considered incompressible;
- The plant has constant water mass flow rate;
- Water tank and piping are considered adiabatic.

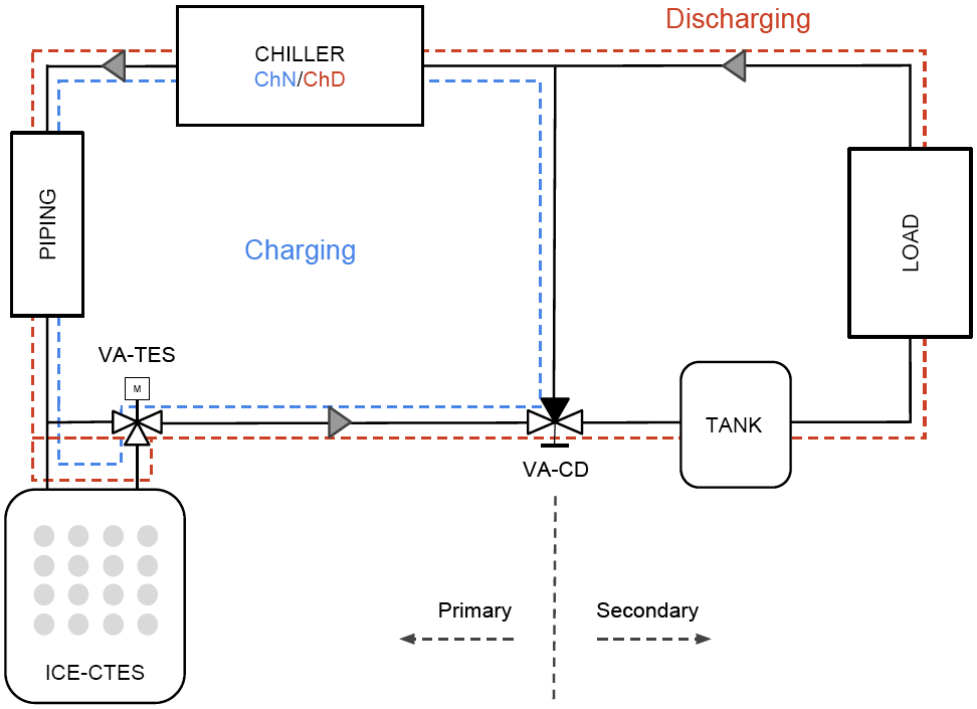


Figure 3.1: Plant structure of the HVAC system.

The system dynamics are governed by the mass and energy conservation laws, which are implemented as block equations for each component of the plant. The dynamic behaviour of the plant is thus obtained solving the fluid flow problem and the energy balance. The fluid flow problem consists only in the determination of the mass flow rate and the equations for the k-th block may be simply written as follows:

$$\dot{m}_{i,k} - \dot{m}_{o,k} = 0\,, \tag{3.1}$$

and therefore, for all k,

$$\dot{m}_k := \dot{m}_{i,k} = \dot{m}_{o,k}\,, \tag{3.2}$$

where dependence on the time variable τ is omitted for notation convenience. The energy equation is deduced from the conservation laws; for each block the energy equation at time τ can be written as follows:

$$\dot{Q}_{g,k} + \dot{m}_{i,k} c_{p_l} T_{i,k} - \dot{m}_{o,k} c_{p_l} T_{o,k} - \dot{Q}_{w,k} - \dot{Q}_{s,k} = 0\,. \tag{3.3}$$

where:

- $\dot{Q}_{g,k}$ is the amount of cooling power generated by the block;
- $\dot{Q}_{s,k}$ is the time derivative of the amount of energy stored in the block;

Nomenclature - Scripts.

$\mathcal{C}$	price of electricity [€ $W^{-1}min^{-1}$]	V	volume [m^3]
ChN	nighttime-chiller	VA-CD	charging/discharging valve
ChD	daytime-chiller	VA-TES	modulating control valve
c_{pl}	water specific heat [$J\,kg^{-1}K^{-1}$]	Z	part load factor [-]
c_{ps}	ice specific heat [$J\,kg^{-1}K^{-1}$]	α	tuning coefficient [€ K^{-2}]
h	tank well-mixed section fraction [-]	β	moving window weight factor [-]
HS	CTES Hybrid System	Δ	difference operator [-]
$\dot{m}$	mass flow rate [$kg\,s^{-1}$]	ϵ	efficiency [-]
N	moving window size [-]	λ	specific latent heat [$J\,kg^{-1}$]
PLR	Part Load Ratio [-]	ν	VA-TES valve opening [-]
Q	energy [J]	ρ	water density [$kg\,m^{-3}$]
T	water temperature [°C]	σ	loss coefficient [$K^{-1}\,s^{-1}$]
T_{air}	air temperature [°C]	τ	time [min]
$\mathfrak{T}_c$	control horizon [min]	τ_c	tank section time constant [min]
$\mathfrak{T}_p$	prediction horizon [min]		

Nomenclature - Subscripts.

ch	chiller	n	nominal
e	electrical	o	outlet
f	water tank or piping	s	accumulation or storage
g	generation	TES	thermal energy storage
i	inlet	$tran$	transition
k	block index	v	VA-TES valve
l	load	w	wastage

- $\dot{Q}_{w,k}$ is the wasted power, if the block is not considered adiabatic;
- $T_{o,k}$ is the outlet water temperature from the block;
- $T_{i,k}$ is the inlet water temperature in the block;
- c_{pl} is the water specific heat.

3.1.1 Chiller

In the hypothesis of using air condensed chillers, the nominal cooling power and the nominal electric power consumption, at full (100%) load conditions, are expressed (by using a multiple linear regression) as a function of the inlet chiller water temperature, external air temperature and water mass flow rate:

$$\dot{Q}_{g,ch_{100\%}} = a_g + b_g T_{i,ch} + c_g T_{air} + d_g \dot{m}_{ch} + e_g T_{i,ch} \dot{m}_{ch}\,, \tag{3.4a}$$

$$\dot{Q}_{e,ch_{100\%}} = a_e + b_e T_{i,ch} + c_e T_{air} + d_e \dot{m}_{ch} + e_e T_{i,ch} \dot{m}_{ch}\,. \tag{3.4b}$$

The coefficients in (3.4a) and (3.4b) can be obtained from manufacturer's data. The energy performances of the chiller at full load conditions are evaluated through the Energy Efficiency Ratio (EER), i.e. the ratio of the nominal cooling capacity and nominal power absorption:

$$EER_{100\%} = \frac{\dot{Q}_{g,ch_{100\%}}}{\dot{Q}_{e,ch_{100\%}}} . \tag{3.5}$$

The part load operation influence is taken into account by multiplying the full cooling capacity by Part Load Ratio (PLR, defined as the chiller cooling load divided by its maximum capacity for the given values of air and glycol inlet temperature) as well as by multiplying the full load power consumption by a part load factor Z, which is calculated as a function of PLR and air temperature, [4]:

$$\dot{Q}_{g,ch} = PLR \cdot \dot{Q}_{g,ch_{100\%}} , \qquad \dot{Q}_{e,ch} = Z \cdot \dot{Q}_{e,ch_{100\%}} . \tag{3.6}$$

The Energy Efficiency Ratio at part load conditions is simply:

$$EER = \frac{\dot{Q}_{g,ch}}{\dot{Q}_{e,ch}} . \tag{3.7}$$

In this HVAC plant, chiller block is considered adiabatic and its water content is neglected; the energy balance equation (3.3), at time τ, can thus be written as follows:

$$\dot{Q}_{g,ch} = \dot{m}_{ch} c_{p_l} \left(T_{o,ch} - T_{i,ch} \right) . \tag{3.8}$$

3.1.2 CTES

The ice-CTES is modelled as a hybrid system, where the water phase transitions are described by combining continuous and discrete dynamics. Although the ice latent heat represents a large portion of the total storage energy in the TES (in order of 90%), we also consider here the sensible heat portion to exhaustively evaluate the energy efficiency of the HVAC system when operated with different control strategies.

In the CTES an event changes the phase from solid to melting, thereby also changing the continuous dynamics of the water: the absorbed thermal energy melts ice and the temperature of the water does not change during melting since water is in a two-phase condition. The change from a single-phase condition (liquid, solid) to a two-phase condition (melting, freezing) happens when a given temperature level is reached (i.e. the saturation temperature

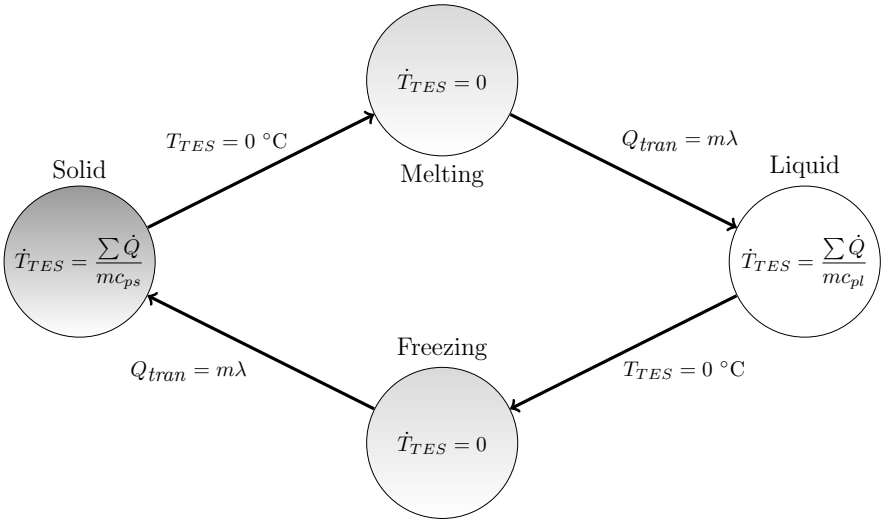


Figure 3.2: Ice-CTES Hybrid System (HS): the phases of the water.

corresponding to the water pressure). Instead, a transition from a two-phase condition to a single-phase condition occurs when a given energy (i.e. latent heat fusion energy) is added or subtracted to the water. Figure 3.2 shows the different phases (and their corresponding continuous dynamics) and the transitions between them.

From (3.3), we can depict the ice-CTES as a heat exchanger, where the variation of the state of charge can be expressed as the sum of the exchanged powers:

$$\dot{Q}_{s,TES} = \dot{m}_{TES} c_{p_l} (T_{i,TES} - T_{o,TES}) - \dot{Q}_{w,TES} \,. \tag{3.9}$$

The first term on the right side of (3.9) represents the power exchange between the water solution and the ice-CTES during process operations, which is:

$$\dot{m}_{TES} c_{p_l} (T_{i,TES} - T_{o,TES}) = \epsilon \dot{m}_{TES} c_{p_l} (T_{i,TES} - T_{TES}) \,. \tag{3.10}$$

According to the lumped formulation, T_{TES} is the average temperature of the water/ice in the CTES. To take into account the efficiency of the heat exchange, the coefficient ϵ, i.e. the ratio of the actual rate of heat transfer to the maximum possible rate, is added in (3.10). Such coefficient varies during the charging and discharging processes (due to the change of heat transfer coefficient and transfer area) and can be expressed as a function of the nominal capacity, the state of charge of the TES and the difference between the glycol and ice temperatures.

Due to the storage dimension and the high difference between water/ice temperature in the storage and the external air temperature, it is important

to consider also the energy dispersion in the TES. In order to take into account the heat transfer to the ice from the surroundings, the second term on the right side of (3.9), involves TES energy wastage:

$$\dot{Q}_{w,TES} = \sigma Q_{n,TES}(T_{TES} - T_{air})\,. \tag{3.11}$$

The loss coefficient σ is related to the degree of isolation of the storage with respect to the external environment.

3.1.3 Modulating Valve

In the bypass loop around the TES, the temperature-modulating valve (VA-TES, Figure 3.1), by manipulating its opening ν, allows a quantity of glycol (from chiller) to bypass the ice storage and to be mixed with the outlet TES glycol, achieving the chilled glycol setpoint value for the building. At each time τ, material and energy balance equations yield:

$$\dot{m}_v c_{p_l} T_{o,v} = (1-\nu)\dot{m}_v c_{p_l} T_{o,ch} + \nu \dot{m}_v c_{p_l} T_{o,TES}\,. \tag{3.12}$$

3.1.4 Water Tank and Piping

The water tank is modelled as two separate parts connected in series. In the first part a well-mixed condition is assumed and the energy balance equation is:

$$\dot{m}_f c_{p_l} T_{i,f}(\tau) = \dot{m}_f c_{p_l} T_f(\tau) + h\rho V_f c_{p_l} \frac{dT_f(\tau)}{d\tau}\,, \tag{3.13}$$

where h is the well-mixed section fraction of the tank total volume. In the second part a perfect stratification condition is considered:

$$\dot{m}_f c_{p_l} T_{o,f}(\tau) = \dot{m}_f c_{p_l} T_f(\tau - \tau_c)\,, \tag{3.14}$$

where τ_c is the tank section time constant defined as:

$$\tau_c = (1-h)\frac{\rho V_f}{\dot{m}_f}\,. \tag{3.15}$$

The same approach is used to model piping blocks, although water mixing is negligible due to the low water velocity. More details about the derivation of the previous equations can be found in [2].

3.1.5 Building Thermal Load

For the building load block, at time τ, the energy balance equation is:

$$\dot{Q}_{g,l} = \dot{m}_l c_{p_l}\left(T_{o,l} - T_{i,l}\right)\,. \tag{3.16}$$

Table 3.1: Charging/discharging operations.

	charging	discharging
chiller	ChN	ChD
VA-CD	0	1
Eqs.	(3.6)-(3.15), HS	(3.6)-(3.16), HS
ν	1	$[0, 1]$
y	$T_{o,ch}$	$T_{i,l}$
$\mathbf{u}$	PLR	$[PLR; \nu]$
$\mathbf{d}$	T_{air}	$[\dot{Q}_{g,l}; T_{air}]$
$\mathbf{x}$	$[T_{o,ch}; Q_{s,TES}]$	$[T_{i,l}; Q_{s,TES}]$

3.1.6 Model Summary

The overall discrete dynamical system (with a one-minute sampling time) during charging/discharging operations (according to Table 3.1) can be summarised as follows:

$$\boldsymbol{\Sigma}: \quad \mathbf{x}(\tau+1) = \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau), \mathbf{d}(\tau)), \tag{3.17a}$$

$$y(\tau) = h(\mathbf{x}(\tau)), \tag{3.17b}$$

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{3.17c}$$

$$\tau \geq 0, \tag{3.17d}$$

$$y \in \mathbb{Y} := [y_{\min}, y_{\max}] \subset \mathbb{R}, \tag{3.17e}$$

$$\mathbf{u} \in \mathbb{U} := [\mathbf{u}_{\min}, \mathbf{u}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{u})}, \tag{3.17f}$$

$$\mathbf{d} \in \mathbb{D} := [\mathbf{d}_{\min}, \mathbf{d}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{d})}, \tag{3.17g}$$

$$\mathbf{x} \in \mathbb{X} := [\mathbf{x}_{\min}, \mathbf{x}_{\max}] \subset \mathbb{R}^2, \tag{3.17h}$$

where $\mathbf{x}, \mathbf{u}, \mathbf{d}$, are the vectors of state, control variables and disturbances, respectively, y is the output variable, $\mathbf{f}$ is a non-linear map, h is a simple map that selects the first element of $\mathbf{x}$. The sets $\mathbb{Y}, \mathbb{U}, \mathbb{D}, \mathbb{X}$ are the feasible domains.

3.2 CTES Charging

During TES charging, only the Part Load Ratio is used in order to control the outlet chiller water temperature and all the glycol solution coming from the chiller enters the storage to charge it (Table 3.1). In detail, a PI-controller

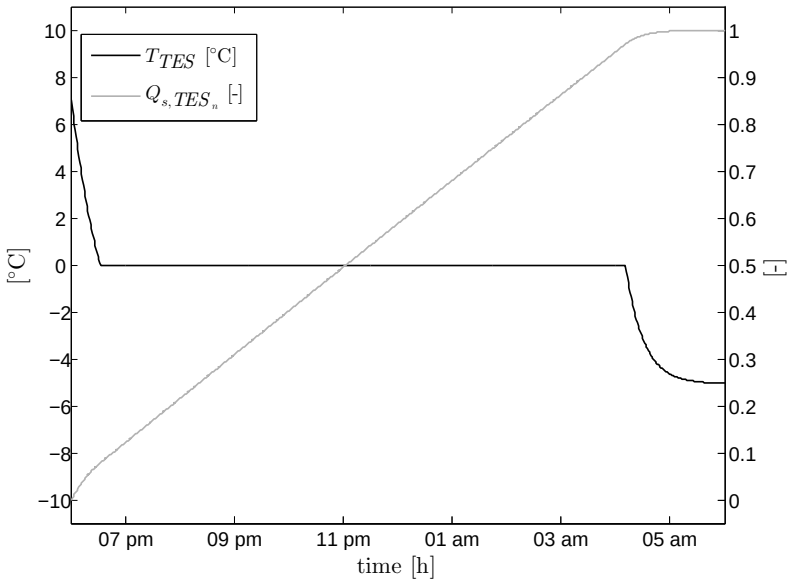


Figure 3.3: TES temperature and Normalised TES energy during charging time.

with feed-forward is used: knowledge about the external air temperature can be fed forward and combined with the PI-controller to improve the system performance; the reference trajectory for the outlet chiller water temperature is set to -5°C. Figure 3.3 shows temperature inside the CTES and the normalised TES energy during a 12-h charging cycle (from 06 pm to 06 am). The normalised TES energy is set to 0 when storage water temperature is 7°C while it is set to 1 when storage ice temperature is -5°C. Since the energy price is low during the night, it is useful to consider also the ice/water sensible heat and to charge the storage until its temperature reaches the outlet chiller water temperature.

3.3 CTES Discharging

In order to satisfy cooling demand during working hours, standard control strategies and a non-linear model predictive control approach are compared. In this section, the non-linear predictive control problem is first described, then simulations results for TES discharging are provided.

3.3.1 NLMPC Problem Formulation

The controller consists of three fundamental elements: a predictor that predicts the outputs based on the model and process information, a cost function, and an algorithm to solve a constrained non-linear optimisation problem. The aim

of the proposed Non-Linear MPC[1] (NLMPC) is to minimise electricity costs during building operating hours, while keeping the cooling demand satisfied. We define the electricity bill as:

$$\mathcal{B}(\tau) = \mathcal{C}(\tau)\dot{Q}_{e,ch}(\tau)\Delta\tau\,. \tag{3.18}$$

The model predictive control action, at time τ, is obtained by solving the optimisation problem:

$$\text{find} \qquad \arg\min_{\bar{\mathbf{u}}} J(\bar{\mathbf{x}}, \bar{\mathbf{u}}, \bar{\mathbf{d}}, \mathcal{B}, \mathfrak{T}_c, \mathfrak{T}_p)\,, \tag{3.19a}$$

$$\text{with} \qquad J = \sum_{\tau=\mathfrak{t}}^{\mathfrak{t}+\mathfrak{T}_p} \mathcal{B}(\tau) + \alpha\,|\bar{y}(\tau) - \bar{r}(\tau)|^2\,, \tag{3.19b}$$

$$\text{s.t.} \quad \bar{\mathbf{x}}(\tau+1) = \mathbf{f}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau), \bar{\mathbf{d}}(\tau))\,, \tag{3.19c}$$

$$\bar{y}(\tau) = h(\bar{\mathbf{x}}(\tau))\,, \tag{3.19d}$$

$$\bar{\mathbf{u}}(\tau) \in \mathbb{U}\,, \quad \forall\tau \in [\mathfrak{t}, \mathfrak{t}+\mathfrak{T}_c]\,, \tag{3.19e}$$

$$\bar{\mathbf{u}}(\tau) = \bar{\mathbf{u}}(\mathfrak{t}+\mathfrak{T}_c)\,, \forall\tau \in [\mathfrak{t}+\mathfrak{T}_c, \mathfrak{t}+\mathfrak{T}_p]\,, \tag{3.19f}$$

$$\bar{\mathbf{d}}(\tau) \in \mathbb{D}\,, \quad \forall\tau \in [\mathfrak{t}, \mathfrak{t}+\mathfrak{T}_p]\,, \tag{3.19g}$$

$$\bar{\mathbf{x}}(\tau) \in \mathbb{X}\,, \quad \forall\tau \in [\mathfrak{t}, \mathfrak{t}+\mathfrak{T}_p]\,, \tag{3.19h}$$

$$\bar{y}(\tau) \in \mathbb{Y}\,, \quad \forall\tau \in [\mathfrak{t}, \mathfrak{t}+\mathfrak{T}_p]\,, \tag{3.19i}$$

where $\mathfrak{T}_p$ and $\mathfrak{T}_c$ are the prediction and the control horizon with $\mathfrak{T}_c \leq \mathfrak{T}_p$; the term r refers to the reference trajectory; α is a weight coefficient that determines the trade-off between energy efficiency (first term in (3.19b)) and demand satisfaction (second term in (3.19b)). The vectors of state, input, disturbances and output $(\mathbf{x}, \mathbf{u}, \mathbf{d}, y)$ are given in Table 3.1. In order to distinguish the real system from the system model used to predict the future within the controller, we use a bar to denote the internal variables in the controller (e.g. $\bar{\mathbf{x}}$). The control action sent to the plant is the first element of the optimal sequence $\bar{\mathbf{u}}$.

3.3.1.1 Disturbances Forecasting

To solve (3.19a), disturbances forecasting (cooling load and external air temperature) are required. Therefore, the strategy proposed in [5] is implemented: an ANN (Artificial Neural Network[2]) algorithm is used to obtain accurate

[1] A brief description of the MPC approach is provided in Appendix A

[2] Artificial neural networks are explained in Appendix B

predictions by exploiting a limited dataset of standard measures available in HVAC installations like the one considered here. In particular, it is assumed that load forecasting for the next day is made at midnight. In the first step, training of the ANN on data from previous days is performed. Then, input predictions are used as the inputs of the ANN model to forecast the coming 24-h cooling load profile. Only cooling devices (e.g. air handling unit) schedule, external dry-bulb and wet-bulb temperatures are used as inputs of the ANN model to obtain a cooling load forecasting. In fact, in practical situations, information on internal loads (lighting, computers, and occupancy) is not available. However since internal gains are highly correlated to the cooling devices schedule, a satisfactory load prediction performance can be achieved: e.g. in Figure 3.4 actual and predicted load data are compared for two weeks in a typical commercial building.

Forecasting of the external air temperature, $\hat{T}_{air}$, is made with a weighted moving average, since we suppose that weather forecast is not available:

$$\begin{aligned}\hat{T}_{air,j+1}(\tau) =& \frac{1}{N}\Big[\beta_j T_{air,j}(\tau) + \beta_{j-1} T_{air,j-1}(\tau) + \dots \\ & \dots + \beta_{j-N+1} T_{air,j-N+1}(\tau)\Big] ,\end{aligned} \tag{3.20}$$

where $T_{air,j}$ refers to the temperature of the j-th day. Figure 3.5 shows an example of air temperature prediction for two weeks.

3.3.1.2 Optimisation

To solve the non-linear optimisation problem (3.19a) a nature-inspired, metaheuristic algorithm (PSO) is used, [18]. Such choice is motivated as follows. First, PSO algorithm makes few or no assumptions about the optimisation problem to be solved; it does not need gradient information of the objective function under consideration, as required by classic optimisation methods such as gradient descent and quasi-Newton methods, and it can search very large spaces of candidate solutions. Moreover, the PSO algorithm is suitable for avoiding local minima. With respect to other stochastic optimisation techniques (e.g. Genetic Algorithms) PSO is more advantageous because there are few parameters to set. Furthermore, it can be easily implemented and it is not computationally expensive, since its memory and CPU speed requirements are low. The implementation of the PSO algorithm is detailed in Appendix C.

There are many PSO variants, and in this thesis, to determine the optimal

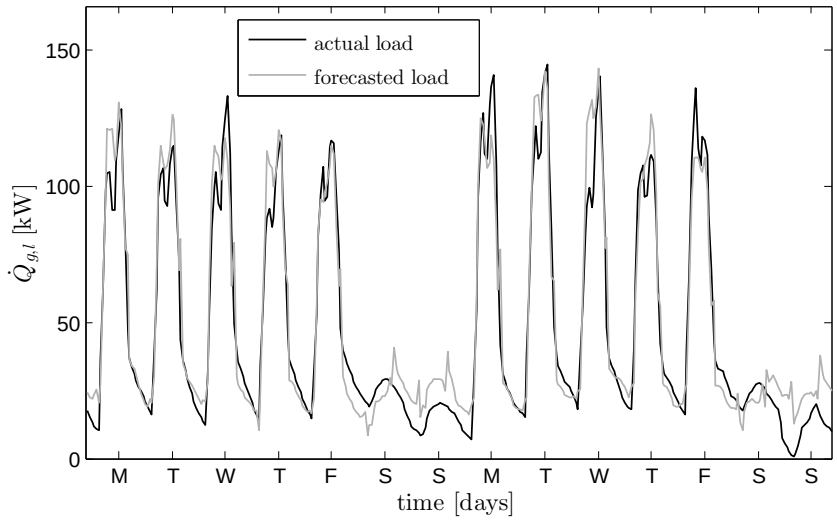


Figure 3.4: Example of cooling load forecasting of a commercial building.

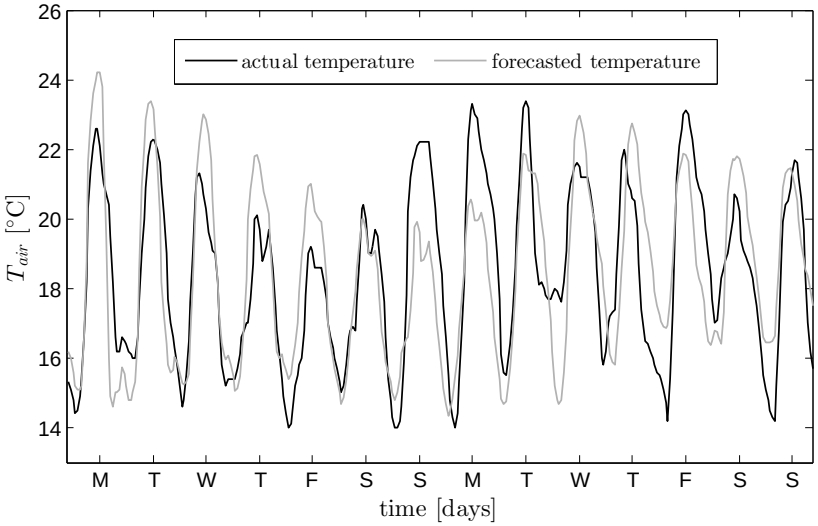


Figure 3.5: Example of temperature prediction.

future control inputs, a version of PSO algorithm which deals with constraints is used. The sub-optimal solution for the optimisation problem (3.19a) obtained by applying the PSO algorithm has been validated with a brute-force search method. The computations required by the optimisation process can be performed within a fixed time length thus granting on-line implementation.

3.3.1.3 NLMPC Implementation

The prediction and control horizons depend on the length of time window where disturbances predictions are available (until the end of the working hours [5]). The control sampling time is set to 15 minutes. To reduce the computational burden, a move blocking strategy is used by imposing inputs (PLR and ν) to be constant over several time-steps in the prediction horizon [7]. It is worth noting that to ensure feasibility for the NLMPC problem, the storage is assumed to always have enough energy to satisfy time-varying, uncertain building load demands.

3.3.2 Simulations Results

To illustrate the performance of the non-linear MPC we compare the predictive control with the three standard control strategies:

P) Constant-Proportion Control (e.g. Chiller 40% and ice-CTES 60%);

C) Chiller-Priority Control;

S) Storage-Priority Control.

In the hypothesis that the ice-CTES is fully recharged by the chiller during nighttime, we report simulation results during daytime, when CTES discharge occurs (from 06 am to 06 pm). We set $\bar{r}(\tau)$ equal to 7°C for each τ. We introduce three time slots of electricity price: F1 (high price, on-peak), F2 (medium price, on-peak) and F3 (low price, off-peak). Two building cooling load case studies are considered: low load ($\boldsymbol{L}$) and high load ($\boldsymbol{H}$) conditions, Figure 3.6(a) and 3.6(b), respectively.

L) Figures 3.7(a) and 3.8(a) show the manipulated variables with the different control strategies. Figure 3.8(a) includes time slot F3 meaning that TES energy has been produced during the night (off-peak period). In Figure 3.9(a) the inlet load-side water temperature is reported. Finally,

Figure 3.10(a) and 3.11(a) show the temperature inside the CTES and the normalised TES energy, respectively. It can be seen that in this case all control strategies meet cooling building load (Figure 3.9(a)). However, NLMPC controller and storage-priority controller satisfy demand by using only storage energy, while the remaining strategies use chiller power and do not discharge the CTES. In this way, part of the energy produced during the night is not used to satisfy cooling demand, causing energy wastage.

H) It can be seen that proportion control and storage-priority control do not satisfy cooling demand (Figure 3.9(b)): the resulting discharging process turns out to be too fast (Figures 3.10(b) and 3.11(b)), and the chiller cannot satisfy cooling demand. Instead, chiller control priority strategy and predictive control always provide sufficient cooling power. However, in chiller-priority case complete discharge of the CTES is not achieved.

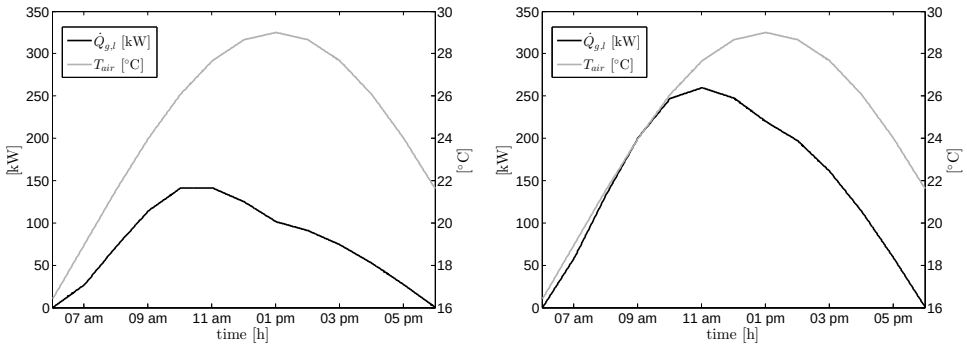
Some comments on the NLMPC simulation results are in order. From information given by disturbances forecasting, predictive control strategy can completely discharge the ice storage within the end of the working hours. During F1 time slot (high price) the NLMPC algorithm prefers using TES energy (produced during the night), while during F2 time slot (medium price), NLMPC tends to exploit chiller power instead of the storage energy (Figures 3.7(b) and 3.8(b)). The increase of the VA-TES valve opening in the late afternoon is explained by the need of exploiting of the last energy in the storage, since its temperature has exceeded 0°C.

In Table 3.2 the normalised electricity bills for all control strategies are given, also including nighttime bill during the next charging. Nighttime-chiller has a lower Energy Efficiency Ratio (EER) with respect to the one used during the day due to the fact that cooler water temperature is needed to recharge the TES. However it is convenient to fully recharge the TES to exploit low electricity price. Bills are normalised on the bill of the chiller-priority strategy (100, high case) and they are not reported when the demand is not satisfied. It can be seen that the NLMPC provides the lowest energy bill, always satisfying cooling demand.

The NLMPC controller gives the best performance due to the fact that we consider the actual disturbances equal to the predicted ones. The case in which the actual load and external air temperature differ from the predicted

Table 3.2: Electricity bills.

	L	***H***
Constant-Proportion Control	59	-
Chiller-Priority Control	67	100
Storage-Priority Control	41	-
Non-Linear Model Predictive Control	41	94



(a) Low cooling load case

(b) High cooling load case

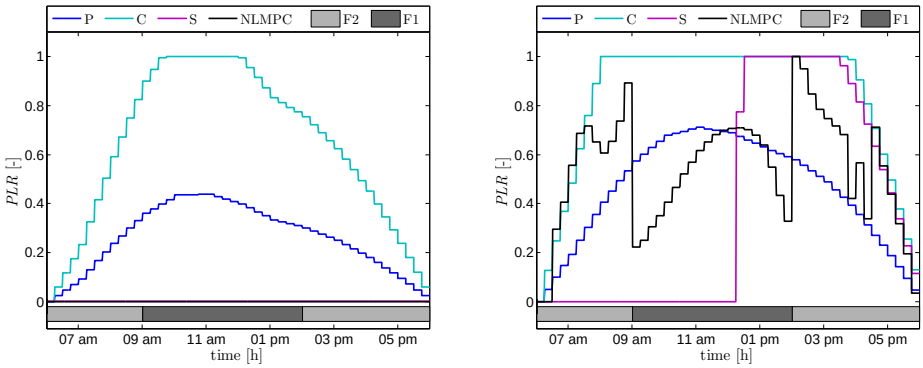
Figure 3.6: Building cooling load and external air temperature.

values is also considered in this study. When actual and forecast values do not match, the NLMPC algorithm uses $\bar{\mathbf{d}}_c(\tau)$ for each τ instead of $\bar{\mathbf{d}}(\tau)$ where:

$$\bar{\mathbf{d}}_c(\tau) = \bar{\mathbf{d}}(\tau) + \bar{\mathbf{q}}(\tau)\,, \tag{3.21}$$

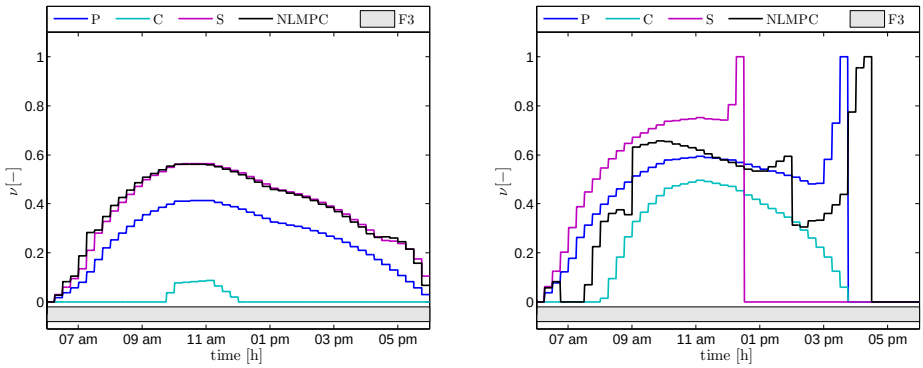
and $\bar{\mathbf{q}}$ is an appropriate function that allows smoothly connecting the actual values of $\bar{\mathbf{d}}(\mathrm{t})$ with its future values as previously predicted.

In the following example we consider an unpredictable load disturbance from 01 pm to 03 pm, Figure 3.12. The inlet load side water temperature trajectory performed by the NLMPC is depicted in the same figure. It can be seen that updating predictions results in good performances since the error from the reference trajectory (7°C) is negligible for typical HVAC building applications.



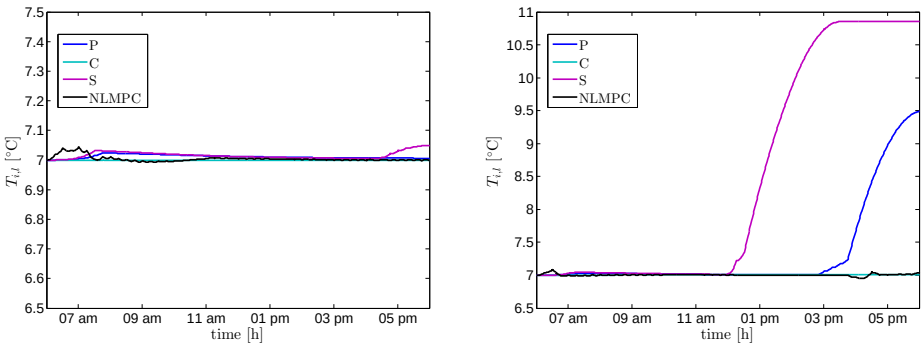
(a) Low cooling load case (b) High cooling load case

Figure 3.7: Part Load Ratio.



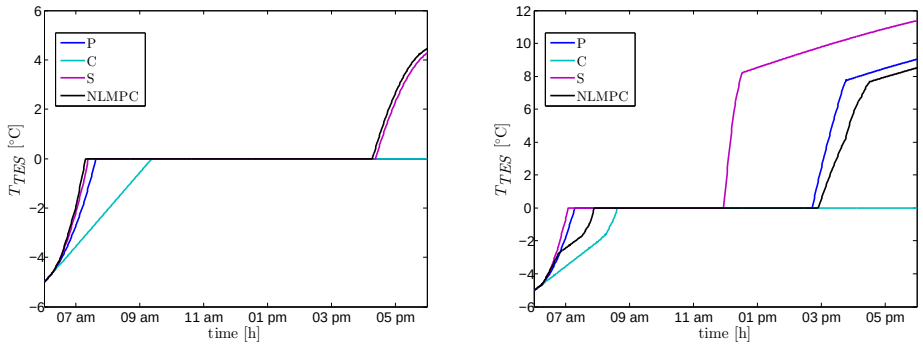
(a) Low cooling load case (b) High cooling load case

Figure 3.8: VA-TES valve opening.



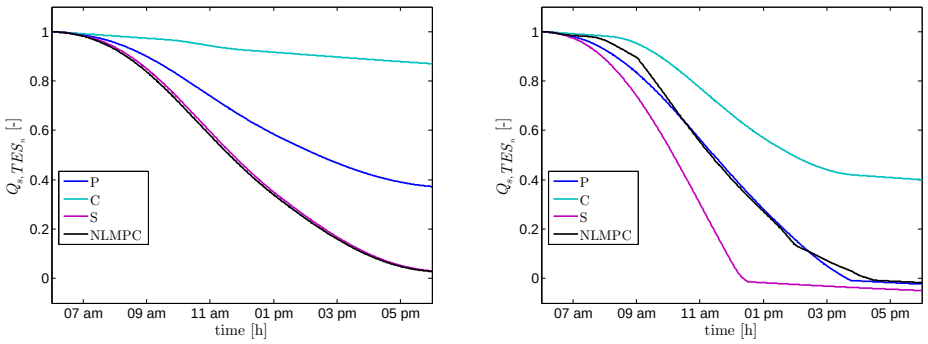
(a) Low cooling load case (b) High cooling load case

Figure 3.9: Inlet load-side water temperature.



(a) Low cooling load case (b) High cooling load case

Figure 3.10: TES temperature.



(a) Low cooling load case (b) High cooling load case

Figure 3.11: Normalised TES energy.

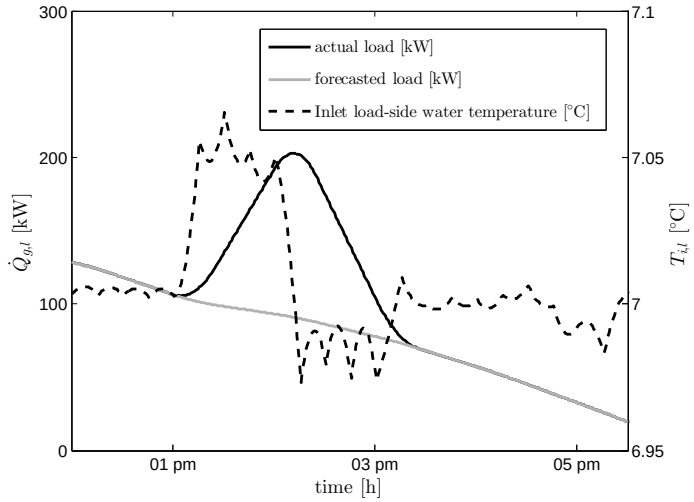


Figure 3.12: Inlet load-side water temperature with load prediction update.

4

Conclusions

Cold Thermal Energy Storage (CTES) systems nowadays play a key role in HVAC plants. CTES technology can be used to shift peak loads from on-peak to off-peak periods, storing energy during the night when the energy price is low, exploiting it in the following daytime when the energy price is higher. Due to CTES systems complexity, control and optimisation procedures are fundamental in order to ensure energy efficiency and load demand satisfaction.

In this part of the thesis an ice-Cold TES system is considered. Ice storage technology ensures a reduction of the storage dimension with respect to water-CTES, due to the amount of energy that can be extracted during the transition phases. The thermal behaviour of the plant is analysed with a lumped formulation of the conservation equations (energy and mass). In particular, the ice-CTES is described in terms of a hybrid system, thus taking into consideration the dynamics of both sensible and latent heat. Standard control strategies are compared with a NLMPC-based approach. The NLMPC is developed by using a predictor based on plant model and process information, a cost function, and a PSO algorithm. The predicted outputs of the plant are also obtained from cooling devices scheduling as a mean for inferring information on the internal loads, which is in general not available in practice. Extensive dynamic simulations show that NLMPC exhibits better performance (in terms of energy efficiency and demand satisfaction) than conventional strategies.

The use of a stochastic algorithm like PSO (which does not use derivative information) for the optimisation task gives the possibility to easily extend the predictive control strategy to other plant configuration (e.g. chiller downstream configuration) or inserting extra terms in the performance index (e.g. comfort index) in future developments of the research.

Part II

Fault Detection in HVAC Systems

1 Introduction

The discipline of process control has made many advances in the last three decades with the advent of computerization. Low-level control actions, which used to be performed by human operators, are now routinely performed in an automated manner with the aid of computers. With progress in advanced control systems, the benefits to different industries (e.g. chemical, petrochemical, cement, steel) have been enormous. However, a very important control task in managing process plants still remains largely a manual activity, performed by human operators: this is the task of responding to abnormal events in a process. This involves the timely detection of an abnormal event, diagnosing its origins and then taking appropriate supervisory control decisions and actions to bring the process back to a normal, safe and operating conditions.

Fault Detection and Diagnosis (FDD) is a sub-field of control engineering which deals with monitoring a system, identifying when a fault occurs, and pinpointing the type of fault and its location. "Fault" is a deviation of at least one characteristic property or parameter of the system from the usual or standard condition. "Detection" refers to the discovery of the anomaly that occurs in the system. "Diagnosis" means "locating" the place where the fault occurs and identifying its origin. The first studies on fault detection and diagnosis go back to the early 70s, but research in applying it to HVAC systems began to develop only in the late 80s - early 90s. It is in fact found that different faults may occur in HVAC systems, which can cause discomfort and waste of energy. However, efficient FDD methods for HVAC systems still remain as a challenge and commercial FDD systems have appeared in the market in recent years only. To this aim, it is particularly relevant to access data that describe both regular and faulty behaviour of the system. Such data can be available from actual plants, or it is possible to generate artificial data

by means of simulations, a common practice nowadays. This can be done by resorting to models which are able to describe the response of the system in both normal and faulty operating conditions.

FDD techniques can be divided into model-based [20] [21], when they are based on a model tied to the physics that characterizes the system, and process history-based (or data-driven), when the FDD methods are based on data without considering structures a priori known. The first case has the advantage of working with a physics-based model but the disadvantage of making approximations of different types, since modern systems are very complex and often highly non-linear. Model-based techniques can be divided into quantitative methods and qualitative-based ones. Examples of quantitative techniques are state space representations or transfer functions, in which for example FDD methods can be performed resorting to Kalman filters such as in [35], [34]. Qualitative model-based approaches are related to logic and rule-based techniques. Process history-based methods include for example black-box representations [38], artificial neural networks [19], but also more recent tools such as machine learning techniques (support vector machines, k-nearest neighbour, etc.) and often need a large amount of data.

A data-based approach is chosen in this thesis to develop fault detection methods for HVAC systems; in particular the second part of the dissertation deals with two different fault detection problems.

The first topic tackles the fault detection and diagnosis of the most common faults usually affecting Variable Air Volume Air Conditioning (VAVAC) plants. These systems are designed to provide good comfort performances with low energy costs. However different faults can happen in VAVAC plants, causing discomfort for the occupants and energy wastage. Since a real VAV plant is not available to perform experimental tests, a simulation environment, based on a dynamic model of a controlled two-zone VAVAC system, is developed in the thesis: it can simulate the behaviour of the system in both normal and faulty operating conditions. A supervised classification method is used to detect and diagnose the simulated faults in the plant: support vector machines coupled with a posterior probability are used to perform classification. The proposed supervised learning technique is shown to be effective in the detection and diagnosis of the simulated faults.

The second topic approaches the fault detection problem with application to vapour-compression chillers. In this case, fault detection is performed on the

datasets from the research project 1043-RP of the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE). In this project experimental studies were conducted on a centrifugal water-cooled chiller in order to collect data in both normal and faulty situations. Only data relating to the correct behaviour of the chiller are used to identify the reference model of functioning by resorting to one-class classification and a novelty detection approach with the help of principal component analysis. The proposed method is tested in the detection of the most relevant faults plaguing chiller systems.

The second part of the thesis is organized as follows:

- Chapter 2 describes the variable air volume air conditioning technology. A lumped-parameter model of a two-zone VAVAC system is developed. A direct feedback linearisation technique is used to control the system. Two-class support vector machines are used to perform detection and diagnosis of the simulated faults in the VAVAC system.
- In Chapter 3 the ASHRAE research project 1043-RP is presented. A steady-state detector is employed in order to select the subset of the ASHRAE database in stationary conditions. One-class classification coupled with principal component analysis is used for the detection of the most common faults in vapour-compression chillers. ROC analysis is used to evaluate the detection results.
- Final observations are given in Chapter 4.

Machine Learning, classification and novelty detection are of critical importance in this dissertation and are introduced in the following sections.

1.1 Machine Learning and Classification

Machine learning concerns the construction and study of systems that can learn from data. Machine learning algorithms are classified into three main categories:

- Unsupervised learning algorithms operate on unlabelled examples, i.e. input data where the corresponding output is unknown. The problem of unsupervised learning is trying to find hidden structures in unlabelled data, leading to one of the most difficult task in machine learning theory.

- Semi-supervised learning is a class of learning techniques that make use of labelled data in addition to unlabelled data. Semi-supervised learning falls between unsupervised learning (without any labelled training data) and supervised learning (where all the training data are considered labelled).

- Supervised learning algorithms are trained on labelled examples, i.e. input where the output is known. Supervised learning algorithms attempt to identify a function or mapping from inputs to outputs which can then be used to speculatively generate an output for previously unseen inputs.

A very important branch of machine learning theory is classification, i.e. the problem of identifying the set to which a new observation belongs, on the basis of a training set of data containing observations whose category membership is known. The core of classification deals with representation and generalisation. Representation is the property of well representing data; generalisation property quantifies the ability of the system to perform well on unseen data. Finding the right trade-off between the previous characteristics is a key object of study in the field of learning theory.

In classification algorithms the individual observations are analysed in terms of a set of quantifiable properties, known as different explanatory variables, features, etc.; these properties may variously be categorical (e.g. A, B, AB or 0 for blood type), ordinal (e.g. large, medium or small), integer-valued (e.g. the number of occurrences of a word in a text) or real-valued (e.g. a measurement of blood pressure). Some algorithms work only in terms of discrete data and require real-valued or integer-valued data discretised into groups (e.g. less than 5, between 5 and 10, or greater than 10). An algorithm that implements classification is known as a classifier. The term classifier sometimes also refers to the mathematical function, implemented by a classification algorithm, that maps input data to a category. Examples of classifiers in the literature are neural networks algorithms, support vector machines, k-nearest neighbours, Gaussian mixture models, and decision trees.

Finally, it is worth noting that classification algorithms are suitable for fault detection and diagnosis problems: labelling measured data with fault-free or faulty conditions can be seen as a binary classification problem.

In this part of the thesis machine learning algorithms are chosen to deal with the FDD problems; support vector machines will be used to perform classification: they have the capability of dealing with high-dimensional dataset

and they can provide flexible solutions and complex classification rules.

1.2 Novelty Detection

Novelty detection is the identification of new or unknown data/situations that a machine learning system is not aware of during the training phase [25]. Novelty detection is becoming fundamental in different applications such as fault detection and diagnosis: it is becoming a key issue as engineering systems are becoming more and more sophisticated (and more "data-rich") and downtime due to unexpected faults should generally be minimised [33]. Novelty detection belongs to the class of unsupervised or semi-supervised problems. In this context faults are seen as novelties for the system, i.e. events that have not been observed in past time. In fact, getting labelled data in FDD applications can be very arduous. Labelled data contain qualitative information related to the functioning condition of the system (normal or faulty for example) and they can sometimes be:

- costly because labelling is done manually by a human expert;
- unfeasible because anomalous data may be rarely seen in the system, where none, or more generally, few of the potential anomalies already happened in the past.

Several statistical approaches are employed in the literature for novelty detection, amongst them: Gaussian mixture models, hidden Markov models, hypothesis testing; however, all the aforementioned methods assume that data distributions are Gaussian in nature: this, as in the modeling of HVAC systems, can be a strict assumption. Other approaches that do not impose restrictions on data distribution are based on: k-nearest neighbour, statistical clustering, one-class support vector machines. One-class support vector machines are used for one-class classification, and, in this part of the thesis, they will be used to tackle the novelty detection problem for FDD applications.

2

Fault Detection and Diagnosis for VAVAC Systems

Heating, Ventilation and Air Conditioning (HVAC) systems are designed to provide thermal comfort for people in the buildings. Comfort can be achieved by maintaining a desired level of humidity, pressure, air motion and quality. One of the most popular HVAC applications, Variable Air Volume Air Conditioning (VAVAC), is designed to provide low energy cost, low maintenance and good comfort performances. In VAV systems, the supply air is kept at a constant temperature and internal room temperatures are controlled by varying the volume of air through damper actuators. In a control system design perspective, VAVAC systems are relatively complex due to their intrinsically multi-variable, coupled, and non-linear characteristics. Zone temperature sensors are used to control the damper, maintaining the room temperatures at the desired value.

A single-duct VAVAC plant is illustrated in Figure 2.1. Starting from the left side of the figure we can see that a mixture of external air (1) and recirculated air (13) enters the Air Handling Unit (AHU) (2). The main components of this unit are the filter (3), the cooling coil (4), the heating coil (5) and the supply fan (6) that supplies the primary air towards individual rooms. The speed of the supply fan is controlled by an AC drive (7). The temperature of each room is measured by a thermostat (9), and is controlled by adjusting the position of the damper of the corresponding VAV box (8). The return air fan (10), which is controlled by the AC drive (11), pulls out exhaust air outside the building (12), and part of the air (13) is recirculated to the air handling unit.

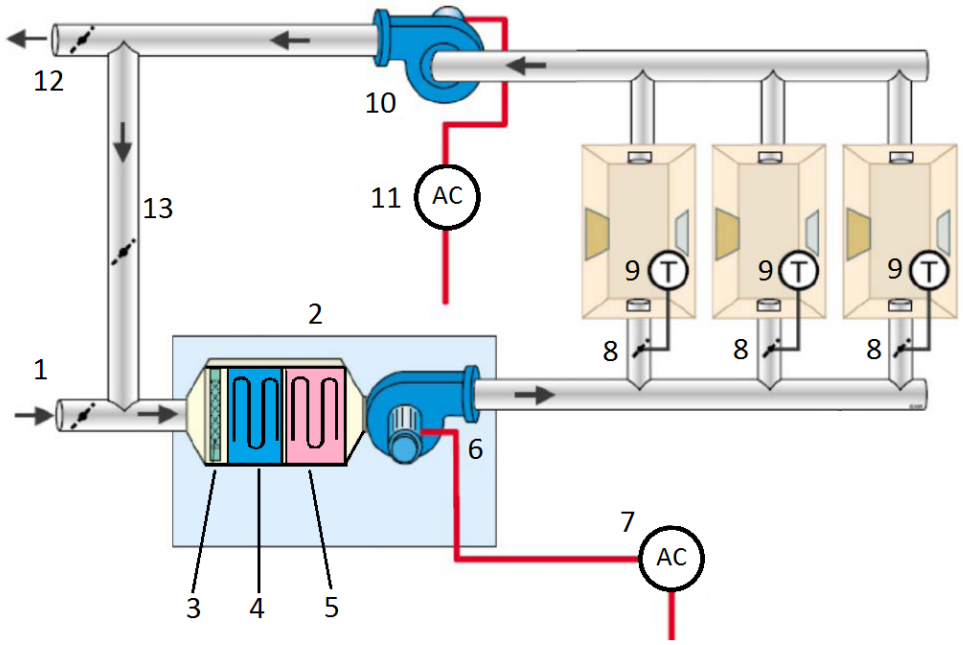


Figure 2.1: Single-duct variable air volume air conditioning system.

2.1 VAVAC Components

In the following a description of the main components of a VAV system is provided.

2.1.1 Air Handling Unit

An air handling unit is typically composed of a large metal shell which contains a supply fan, heating and cooling coils, filters, valves. The AHU is connected to the ducts for air distribution, that carry air in the zones, and bring the exhaust air back to the unit. If heating is required, the unit may contain a burner or a heating coil with steam or hot water that are supplied by a boiler. In case of cooling, the air-handler contains coils with cold water coming from the cooling generating equipment.

The main components of an air handling unit are described in the following:

Filter Air filtration provides clean and dust-free air for the comfort of the building occupants.

Heating and/or cooling coils Air handling units need to provide heating, cooling, or both to regulate the supply air temperature; to this aim, heat

exchanger coils are placed within the unit: hot or cold water circulates in the coils and exchange heat with the air in the unit.

If dehumidification is required, then the cooling coil is employed to over-cool so that the dew point is reached and condensation occurs. A heating coil placed after the cooling coil re-heats the air (therefore known as a re-heat coil) to the desired supply temperature. This has the effect of reducing the relative humidity level of the supply air. A humidifier is often needed in case of continuous heating, in order to ensure a comfortable level of air humidity.

Fans Air handlers typically employ fans driven by an AC induction electric motor to move the air. Fans may operate at a single speed, offer a variety of speeds, or be driven by a variable frequency drive to allow a wide range of air flow rates.

2.1.2 Ducts

The ducts system brings the handled air from the air handling unit to the internal rooms and then brings back the exhaust air from the zones to the unit. The ducts system also allows the connection with the external environment for the renewal of the air and the removal of a part of the exhaust air, while the remaining part is usually recirculated to the air handling unit.

2.1.3 VAV Terminal Units

A VAV terminal unit is a sheet-metal assembly installed upstream of its respective zone. The unit consists of an air-modulation device, control hardware and, depending on the system application, possibly a heating coil, a filter, and a small terminal mixing fan. Modulating the airflow to each individual space is accomplished using a temperature-controlled damper that varies the airflow resistance in the supply duct to that space. Typically, either a pneumatic or electric controller can be used to adjust the damper.

2.1.4 Mixing Boxes

In a mixing box flows with different characteristics (temperature, humidity, air mass flow rate) mix producing a homogeneous flow. A mixing box is placed before the air handling unit to mix the external air flow with the recirculated air one.

2.2 Two-Zone, Single-Duct VAVAC System Modeling

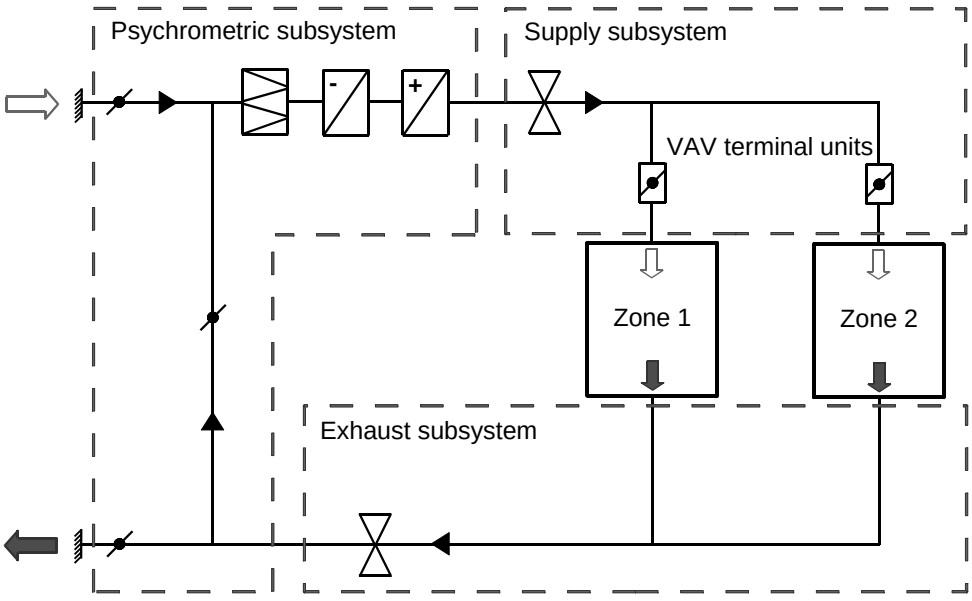


Figure 2.2: Two-zone, single-duct VAV system.

In this thesis a dynamic model of a two-zone, single-duct VAVAC plant is developed. In Figure 2.2 the block structure of the considered VAV system is reported. We can distinguish:

- the Psychrometric subsystem, whose elements are the plant items that are involved in the air conditioning process;
- the Supply subsystem, whose elements are the variable-duty supply fan and a network of air distribution ducts, VAV terminal units, and air terminal devices;
- the Zone subsystem, which comprises the individual zones;
- the Exhaust subsystem, whose individual elements are a variable-duty extract/return air fan, return air diffusers, a network of extract ducts, and the extract air terminal devices in the zone.

The thermal behaviour of the VAVAC system is analysed with the energy and mass conservation equations. The dynamic behaviour of each subsystem is thus obtained solving the fluid flow and the energy balance problems.

Nomenclature.

$A_{z,j}$	wall area of j-th zone [m^2]
C_{pa}	heat capacity of air [$Jkg^{-1}K^{-1}$]
C_{pc}	heat capacity of cooling coil [$Jkg^{-1}K^{-1}$]
C_{pw}	heat capacity of water [$Jkg^{-1}K^{-1}$]
$\dot{m}_a$	air mass flow rate [kgs^{-1}]
$\dot{m}_{a,j}$	air mass flow rate of j-th zone [kgs^{-1}]
$\dot{m}_{ext}$	external air mass flow rate [kgs^{-1}]
$\dot{m}_m = \dot{m}_a$	mixed air mass flow rate [kgs^{-1}]
$\dot{m}_r$	recirculated air mass flow rate [kgs^{-1}]
$\dot{m}_w$	chilled water mass flow rate [kgs^{-1}]
M_c	mass of cooling coil [kg]
$\dot{Q}_j$	internal and external heat source of j-th zone [W]
$\dot{Q}_{int,j}$	internal heat source of j-th zone [W]
$\dot{Q}_{ext,j}$	external heat source of j-th zone [W]
$\dot{Q}_{in,j}$	inlet power contribution in the j-th zone [W]
$\dot{Q}_{out,j}$	outlet power contribution from the j-th zone [W]
r	percentage of recirculation [%]
T_{ai}	inlet air temperature of cooling coil [°C]
$T_{ao} = T_{sa}$	outlet air temperature of cooling coil (supply)[°C]
T_{ext}	external air temperature [°C]
$T_m = T_{ai}$	mixed air temperature [°C]
T_r	recirculated air temperature [°C]
T_{wi}	inlet water temperature of cooling coil [°C]
T_{wo}	outlet water temperature of cooling coil [°C]
$T_{z,j}$	temperature of j-th zone [°C]
$(UA)_c$	heat exchange coefficient [WK^{-1}]
$U_{z,j}$	thermal transmittance of j-th zone [$Wm^{-2}K^{-1}$]
$V_{z,j}$	volume of j-th zone [m^3]
ρ_a	density of air [kgm^{-3}]
τ	time [min]

2.2.1 Zone Subsystem

Without loss of generality, the system is considered in cooling mode. As the zone temperature increases, the controller opens the damper to supply cold air, whereas it closes the damper when the temperature decreases. In this way, the heat generated by internal loads, such as people occupancy, lights, electronic devices, etc., and external loads (i.e. solar thermal gains) is dissipated. The following assumptions are made:

- the air in the zone is fully mixed, so that the zone air temperature distribution is uniform;
- the density of the air is constant and is not influenced by temperature changes.

Under the above assumptions, and by using energy governing equations, the j-th zone thermal dynamics are governed by the following equation:

$$\dot{Q}_{in,j} + \dot{Q}_{int,j} + \dot{Q}_{ext,j} + \dot{Q}_{d,j} + \dot{Q}_{out,j} = V_{z,j}\rho_a C_{pa}\frac{dT_{z,j}}{dt}\,. \tag{2.1}$$

The inlet power contribution $\dot{Q}_{in,j}$ depends on the air mass flow rate of each zone and the difference between the room temperature and the supply air temperature:

$$\dot{Q}_{in,j} = \dot{m}_{a,j}C_{pa}(T_{sa} - T_{z,j})\,. \tag{2.2}$$

The outlet power contribution $\dot{Q}_{out,j}$ is null since the outlet air temperature is equal to the zone one. The power exchange with outside ($\dot{Q}_{d,j}$) is proportional to the difference between the external environment and internal air temperatures:

$$\dot{Q}_{d,j} = U_{z,j}A_{z,j}(T_{ext} - T_{z,j})\,, \tag{2.3}$$

where $U_{z,j}$ is the thermal transmittance of exterior walls (through the $A_{z,j}$ contact area). $\dot{Q}_{int,j}$ is the contribution of internal loads (lighting, computers, occupancy) while $\dot{Q}_{ext,j}$ refers to the solar power contribution (from windows). By substituting (2.3) and (2.2) into (2.1), the dynamic of the temperature of each zone is obtained:

$$\frac{dT_{z,j}}{d\tau} = \frac{1}{V_{z,j}\rho_a C_{pa}}[\dot{Q}_j + \dot{m}_{a,j}C_{pa}(T_{sa} - T_{z,j}) + U_{z,j}A_{z,j}(T_{ext} - T_{z,j})]\,. \tag{2.4}$$

$\dot{Q}_j$ is the sum of internal and external heat sources ($\dot{Q}_{int,j}, \dot{Q}_{ext,j}$).

2.2.2 Exhaust Subsystem

In the exhaust subsystem the two flows coming from the rooms mix; the thermal properties of this part are described by the following mixing box equations:

$$\dot{m}_a C_{pa} T_r = \dot{m}_{a,1}C_{pa}T_{z,1} + \dot{m}_{a,2}C_{pa}T_{z,2}\,, \tag{2.5}$$

$$\dot{m}_a = \dot{m}_{a,1} + \dot{m}_{a,2}\,. \tag{2.6}$$

$\dot{m}_a$ is the air mass flow rate in the ducts of the VAVAC system while T_r is the return air temperature.

2.2.3 Psychrometric Subsystem

The Psychrometric subsystem includes the mixing box and the cooling coil model:

Mixing box In the mixing box the recirculated air mass flow mix with the external air producing an homogeneous flow with temperature T_m. The energy and mass balance equations imply:

$$\dot{m}_m = \dot{m}_r + \dot{m}_{ext}\,, \tag{2.7}$$

$$T_m = rT_r + (1-r)T_{ext}\,, \tag{2.8}$$

where:

$$r := \frac{\dot{m}_r}{\dot{m}_m}\,. \tag{2.9}$$

Cooling Coil Model The cooling coil provides conditioned air with a water-to-air heat exchange. The heat balance equations describe the heat transfer between hot air and cold water in the coil. The following assumptions, during the thermal exchange, are made:

- the temperature of air inside the coil is constant and equal to the inlet air temperature;
- constant air and water flow rate;
- direct relationship between coil mass temperature and outlet water temperature;
- steady-state heat transfer between coil mass and air flow;
- the thermodynamic properties of the fluids are independent of temperature variations;
- in the cooling coils the inlet water temperature is assumed constant, supposing that the water comes from the cooling generating equipment (e.g. chiller).

Following [32] and [36], the following model is adopted for the cooling coil:

$$M_c C_{pc}\frac{dT_{wo}}{dt} = \dot{m}_w C_{pw}(T_{wi} - T_{wo}) + (UA)_c(T_{ai} - T_{wo})\,. \tag{2.10}$$

$$\dot{m}_a C_{pa}(T_{ao} - T_{ai}) = (UA)_c(T_{wo} - T_{ai})\,. \tag{2.11}$$

The manipulation of (2.10) and (2.11) leads to:

$$\frac{dT_{ao}}{dt} = \left[\frac{\dot{m}_w C_{pw}}{\dot{m}_a C_{pa}}(T_{wi} - T_{wo}) + (T_{ai} - T_{ao})\right] \frac{(UA)_c}{M_c C_{pc}}\,, \tag{2.12}$$

$$\frac{dT_{wo}}{dt} = \begin{cases} \dfrac{[\dot{m}_w C_{pw}(T_{wi} - T_{wo}) + \dot{m}_a C_{pa}(T_{ai} - T_{ao})]}{M_c C_{pc}}\,, \text{ if } \dot{m}_w \neq 0\,, & (2.13) \\ \xi \cdot (T_{wi} - T_{wo})\,, \text{ if } \dot{m}_w = 0\,. & (2.14) \end{cases}$$

It is worthwhile to note that $\dot{m}_a$ is greater than zero to ensure that a minimum volume flow rate is drawn into the VAV system. The condition (2.14) is added to ensure that T_{wo} reaches T_{wi} when $\dot{m}_w = 0$ (with $\xi[s^{-1}] > 0$).

2.2.4 Supply Subsystem

In the supply subsystem the air coming from the air handling unit divides into two flows for the air-conditioning of each room. The two flows have the same temperature T_{sa} and the mass balance equation is:

$$\dot{m}_{a,1} + \dot{m}_{a,2} = \dot{m}_a\,. \tag{2.15}$$

2.2.5 Model Summary

Let the states, the inputs and the disturbances of the system be expressed as:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} T_{z,1} \\ T_{z,2} \\ T_{sa} \\ T_{wo} \end{bmatrix}\,, \qquad \mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} \dot{m}_{a,1} \\ \dot{m}_{a,2} \\ \dot{m}_w \end{bmatrix}\,, \qquad \mathbf{d} = \begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix} = \begin{bmatrix} \dot{Q}_1 \\ \dot{Q}_2 \\ T_{ext} \end{bmatrix}\,. \tag{2.16}$$

We define:

$$\boxed{\begin{aligned} a_1 &= \frac{1}{V_{z,1}\rho_a}\,, & b_1 &= \frac{1}{V_{z,2}\rho_a}\,, \\ a_2 &= \frac{1}{V_{z,1}\rho_a C_{pa}}\,, & b_2 &= \frac{1}{V_{z,2}\rho_a C_{pa}}\,, \\ a_3 &= \frac{U_{z,1}A_{z,1}}{V_{z,1}\rho_a C_{pa}}\,, & b_3 &= \frac{U_{z,2}A_{z,2}}{V_{z,2}\rho_a C_{pa}}\,. \end{aligned}} \tag{2.17}$$

The resulting non-linear MIMO system described by (2.4)-(2.15) is characterized by the following equations:

$$\dot{x}_1 = a_1 u_1 (x_3 - x_1) + a_2 d_1 + a_3 (d_3 - x_1)\,, \tag{2.18}$$

$$\dot{x}_2 = b_1 u_2 (x_3 - x_2) + b_2 d_2 + b_3 (d_3 - x_2)\,, \tag{2.19}$$

$$\dot{x}_3 = \left[\frac{C_{pw}}{C_{pa}} \frac{u_3}{(u_1+u_2)} (T_{wi} - x_4) + \left(r \frac{x_1 u_1 + x_2 u_2}{u_1 + u_2} + (1-r) d_3 - x_3 \right)\right] \frac{(UA)_c}{M_c C_{pc}}\,, \tag{2.20}$$

$$\dot{x}_4 = \begin{cases} \left[C_{pw} u_3 (T_{wi} - x_4) + C_{pa} (u_1 + u_2) \left(r \frac{x_1 u_1 + x_2 u_2}{u_1 + u_2} + (1-r) d_3 - x_3 \right) \right] \cdot \frac{1}{(M_c C_{pc})}\,, & \text{if } u_3 \neq 0\,, \quad (2.21) \\ \xi \cdot (T_{wi} - x_4)\,, & \text{if } u_3 = 0\,. \quad (2.22) \end{cases}$$

We remember that the inlet water temperature T_{wi} is considered constant. Finally, the first three states are considered also outputs of the VAVAC system:

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\,. \tag{2.23}$$

Defining the non-linear function $\mathbf{f}$ which comprises all the differential equations, the overall dynamical system can be summarised as:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{u}, \mathbf{d})\,, \tag{2.24a}$$

$$\mathbf{y} = \mathbf{h}(\mathbf{x})\,, \tag{2.24b}$$

where the map $\mathbf{h}$ selects the first three elements of $\mathbf{x}$.

2.2.6 Direct Feedback Linearisation Control

The model described in (2.24) is a non-linear coupled system. In this thesis a direct feedback linearisation technique ([11],[37]) is used to transform the original system model into a simpler equivalent form, exploiting the Dini's implicit function theorem on non-linear differential equations (2.18)-(2.22).

Considering the non-linear system (2.24), the Dini's theorem ensures that, if:

$$\exists\, \mathbf{x_0} \,:\, \left.\frac{\partial \mathbf{f}}{\partial \mathbf{u}}\right|_{\mathbf{x_0}} \neq \mathbf{0}\,, \tag{2.25}$$

there exists a function $\mathbf{g}$ and a new input $\mathbf{v}$ in a neighbourhood of $\mathbf{x}_0$ such that:

$$\mathbf{u} = \mathbf{g}(\mathbf{x}, \mathbf{v}, \mathbf{d}) \,, \tag{2.26}$$

and

$$\dot{\mathbf{y}} = \mathbf{v} \,. \tag{2.27}$$

The VAVAC model has three outputs, so three linear control loops are defined with the new input:

$$\mathbf{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} := \begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{bmatrix} . \tag{2.28}$$

In this way a linear, decoupled system is obtained:

$$\begin{bmatrix} \dot{y}_1 \\ \dot{y}_2 \\ \dot{y}_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} . \tag{2.29}$$

The following inputs are applied in order to linearise the system:

$$u_1 = \frac{v_1 - a_2 d_1 - a_3(d_3 - x_1)}{a_1(x_3 - x_1)} \,, \tag{2.30}$$

$$u_2 = \frac{v_2 - b_2 d_2 - b_3(d_3 - x_2)}{b_1(x_3 - x_2)} \,, \tag{2.31}$$

$$u_3 = \left[\frac{M_c C_{pc}}{(UA)_c} v_3 - (r \frac{x_1 u_1 + x_2 u_2}{u_1 + u_2} + (1 - r) d_3 - x_3) \right] \frac{C_{pa}(u_1 + u_2)}{C_{pw}(T_{wi} - x_4)} \,. \tag{2.32}$$

In the decoupled system the following feasibility conditions are assumed:

$$x_1 \neq x_3 \,, \quad x_2 \neq x_3 \,, \quad x_4 \neq T_{wi} \,. \tag{2.33}$$

The previous conditions are not restrictive assumptions; in fact the supply air temperature is usually set to a lower value with respect to the zones air temperature. Furthermore the outlet water temperature of the cooling coils (x_4) is always higher than the inlet water temperature since we are in cooling mode operation.

A conventional closed-loop proportional control scheme is adopted and with the appropriate choice of $k_{i=\{1,2,3\}}$, we can arbitrarily place the closed-loop poles of the linearised system:

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = \begin{bmatrix} k_1(y_{1,set} - y_1) \\ k_2(y_{2,set} - y_2) \\ k_3(y_{3,set} - y_3) \end{bmatrix} . \tag{2.34}$$

2.2.7 Simulations Results

Tests on the two-zone controlled VAVAC system are reported in this section. In the simulations results the volumes of the zones are $V_{z,1} = 36$ m^3 and $V_{z,2} = 90$ m^3. The external wall areas are $A_{z,1} = 12$ m^2 and $A_{z,2} = 18$ m^2. The portion of exhaust air is equal to 25% ($r = 0.25$). The external air temperature is equal to 27°C, if not otherwise specified. The inlet chilled water temperature T_{wi} is set to 7°C. It is assumed that the supply air temperature (T_{sa}) setpoint is equal to 15°C. The initial indoor air temperatures $T_{z,1}$ and $T_{z,2}$ as well as their setpoints are equal to 22°C. The disturbances are assumed constant if not otherwise specified.

The following cases are considered:

1. In Figure 2.3 a step change in air temperature setpoints occurs in each zone. In the first zone, at time 200 s the set point increases from 22°C to 23°C, while in the second zone at time 600 s the setpoint decreases from 22°C to 21°C. The considerable times of the responses are due to the presence of saturated actuators. It is worthwhile to note that the coupling between the loops of the system is eliminated: a temperature setpoint variation in each zone has no influence on the other zone.

2. In Figure 2.4 the responses to step changes of the supply air temperature T_{sa} are shown. The control exhibits good performance and the times of the responses are small (around 10 s).

3. In Figure 2.5 simulation time is 10 hours. Outdoor external temperature disturbance and internal heat gain load profiles (which are the same for the two zones) are modelled as sine waves (Figure 2.5(a)). The influence of external temperature and internal loads variations on outputs and control signals are shown in Figure 2.5(b).

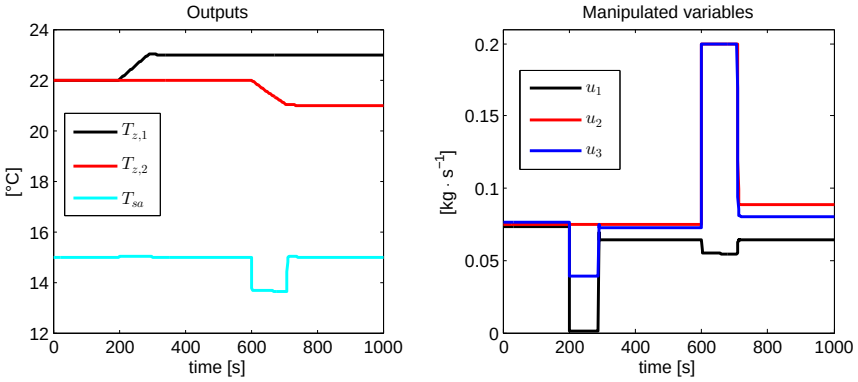


Figure 2.3: Two-zone, single-duct VAV modeling: room temperatures setpoint variations.

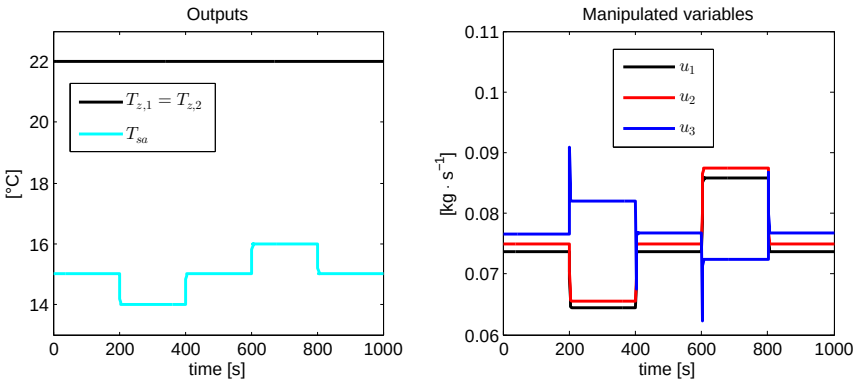
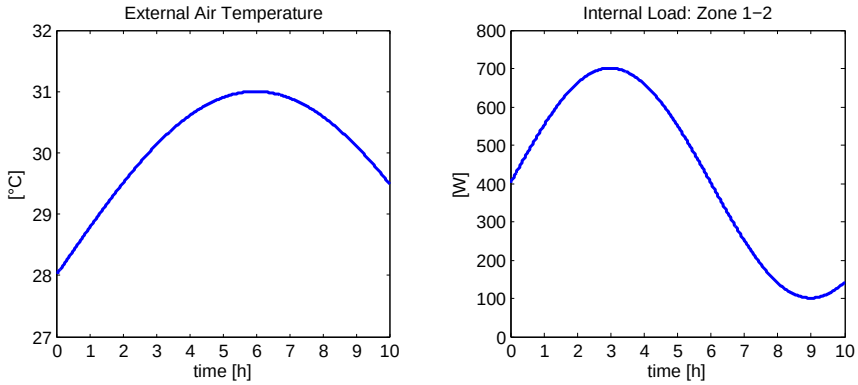
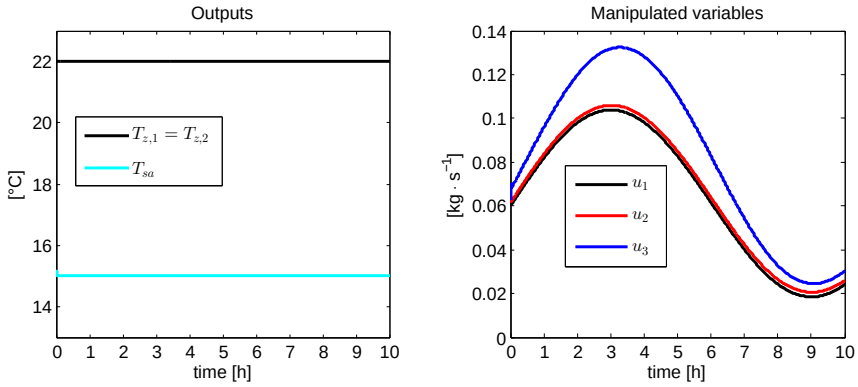


Figure 2.4: Two-zone, single-duct VAV modeling: step changes of the supply air temperature setpoint.



(a) External temperature and loads.



(b) Outputs and control signals.

Figure 2.5: Two-zone, single-duct VAV modeling: varying disturbances.

2.3 Support Vector Machines

Support Vector Machines (SVMs) are supervised learning models that analyse data and recognize patterns, and they are widely used for classification and regression analysis. Given a set of training samples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new observations to one category or the other, making it a non-probabilistic binary linear classifier. A SVMs model is a representation of the samples as points in space, where samples of the separate categories are divided by a gap that is as wide as possible. New observations are then mapped into that same space and predicted to belong to a category based on which side of the gap they fall on. Support vector machines have been introduced within the context of statistical learning theory and structural risk minimisation. Some remarkable features of SVMs are: the use of non-linear kernels which allows separation of non-linearly separable data and the use of structural risk minimisation that improves the ability of generalisation even with a reduced number of data and avoids over-fitting of the samples [13]. SVMs are gaining popularity in a wide range of applications. They are largely used as classifier and a fault detection and diagnosis problem can be formulated as a classification problem. In the following, we briefly describe SVMs and their tuning. The support vector classification theory is detailed in Appendix D.

Given a sample training dataset of n points $S = \{(\mathbf{x}_i, \mu_i)\}$, where $\mathbf{x}_i$ is the i-th input vector and $\mu_i \in \{-1, +1\}$ is its label (in this case -1 means faulty, 1 means fault-free), an optimal hyperplane is computed in a feature space to construct the SVM. The basic idea of the SVM learning algorithm can be summarised in two steps. First, the vector input space is transformed to a higher dimensional feature space by a non-linear transformation function $\boldsymbol{\phi}$. Then, the optimal separating hyperplane can be constructed in this space by solving:

$$\arg\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||^2 + C\sum_i \xi_i \,, \tag{2.35a}$$

subject to:

$$\xi_i \geq 0 \,, \quad \mu_i f(\mathbf{x}_i) \geq 1 - \xi_i \,, \quad \forall i \,, \tag{2.35b}$$

with

$$f(\mathbf{x}_i) = \mathbf{w}^T \boldsymbol{\phi}(\mathbf{x}_i) + b \,. \tag{2.35c}$$

$\mathbf{w}$ and b are the hyperplane parameters. $\{\xi_i\}$ are the slack variables. $f(\mathbf{x}_i)$ provides the signed distance of $\mathbf{x}_i$ from the separating hyperplane. The classification rule induced by f is:

$$G(\mathbf{x}_i) = \text{sign}[f(\mathbf{x}_i)]\,. \tag{2.36}$$

The non-linear transformation $\boldsymbol{\phi}$ can be realised by defining a proper kernel function k:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \boldsymbol{\phi}(\mathbf{x}_i), \boldsymbol{\phi}(\mathbf{x}_j) \rangle\,. \tag{2.37}$$

Common choices for the inner product $k(\mathbf{x}_i, \mathbf{x}_j)$ are:

$$\text{Radial Basis (RBF)} : \exp[-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)]\,, \tag{2.38}$$

$$\text{Polynomial} : (1 + \langle \mathbf{x}_i, \mathbf{x}_j \rangle)^p\,, \tag{2.39}$$

$$\text{Neural Network} : \tanh(a_1 \langle \mathbf{x}_i, \mathbf{x}_j \rangle + a_2)\,. \tag{2.40}$$

For its simplicity, the RBF function is the most popular kernel used in support vector machine classification: with a RBF kernel it is possible to build classification models able to well represent non-uniform data shapes, without over-fitting the samples. Using a RBF kernel function, SVM optimisation results strongly depend on the proper setting of the parameter σ, which has a significant effect on accuracy and generalisation performances. With a large value of σ, the number of support vectors[1] is low, resulting smooth decision boundaries. As σ decreases, the support vectors increases, leading to a greater curvature of the decision boundaries.

Moreover, the parameter C in (2.35a) determines the magnitude of the SVM margin. For a large value of C a large penalty is assigned to slack variables focusing attention more on (correctly classified) points near the decision boundary. A smaller value of C provides a much larger margin, focusing attention on generalisation performances [13].

Posterior probability for support vector machines. Standard support vector machines do not provide a posterior probability after assigning labels to data. However a posterior probability could be useful in fault detection and diagnosis methods. A probabilistic information related to the distance of the points from the separating hyperplane can be introduced: the underlying idea is that observations with large distance from the separating hyperplane will

[1] See Appendix D

have a high probability of faulty or fault-free conditions, while the points close to the separating hyperplane will have a higher uncertainty in the decision.

Platt [28] suggests using the sigmoid function as a posterior probability:

$$P[\mu = 1|\mathbf{x}] \approx P_{A,B}(f) := \frac{1}{1+\exp(Af+B)}\,, \tag{2.41}$$

$$f := f(\mathbf{x})\,. \tag{2.42}$$

In [28] the best setting of parameters A and B is determined by solving the following regularised maximum likelihood problem (with N_+ number of the positive labels, and N_- number of the negative ones):

$$\arg\min_{A,B} \; -\sum_{i=1}^{n} \big(t_i \log(p_i) + (1-t_i)\log(1-p_i)\big)\,, \tag{2.43}$$

$$p_i = p(\mathbf{x}_i) = P_{A,B}(f(\mathbf{x}_i))\,, \quad t_i = \begin{cases} \dfrac{N_+ + 1}{N_+ + 2} & \text{if } \mu_i = +1\,, \\ \\ \dfrac{1}{N_- + 2} & \text{if } \mu_i = -1\,. \end{cases} \tag{2.44}$$

In [22] (2.43) is proved to be a convex (unconstrained) optimisation problem.

2.4 Fault Detection and Diagnosis

The model described in Section 2.2 is able to simulate the response of the two-zone controlled VAVAC system in both normal and faulty operating conditions. The following most common and relevant faults are investigated:

- **Stuck Chilled Water Valve** The chilled water valve that regulates the water mass flow rate is considered blocked; it is worth noting that in normal conditions the control signal u_3 is equal to $\dot{m}_w$ (see Section 2.2) whereas during the fault the corresponding actuator is considered blocked, bringing to a different value of the water mass flow rate with respect to the value provided by the control system. In Figure 2.6 an example of stuck chilled water valve is illustrated: the profile of the fault is assumed with a ramp shape (the fault starts after two hours of normal conditions, then gradually increases for 2 hours to reach its maximum value, then it remains in such state for 2 hours). During the fault the chilled water mass flow rate $\dot{m}_w$ decreases by 40% with respect to normal conditions. As a consequence the supply air temperature increases,

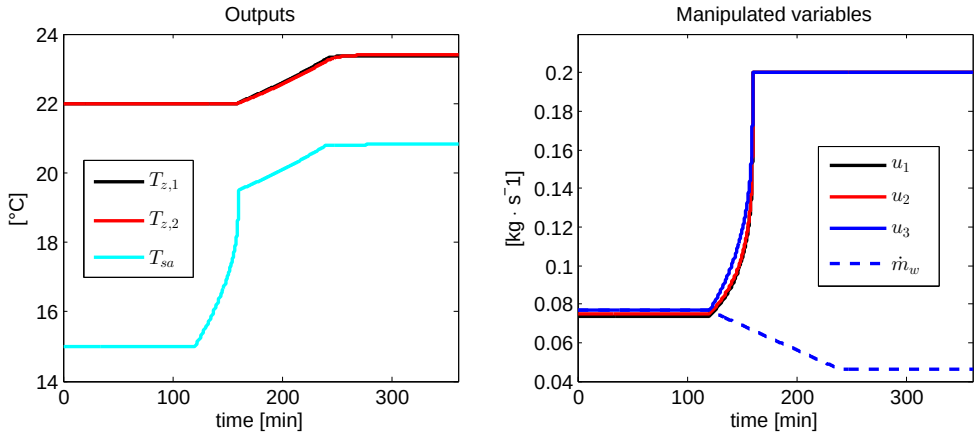


Figure 2.6: Two-zone, single-duct VAVAC system: Stuck Chilled Water Valve (-40%).

bringing hotter air towards the zones. The control system cannot keep the setpoint values in the two rooms: air mass flow rates of the two zones increase in order to maintain the desired temperatures in the rooms, but they reach the saturation values. The control signal of the water mass flow rate u_3 acts in opposite direction with respect to the fault, but it is not an actual input of the system (the real input value is $\dot{m}_w$).

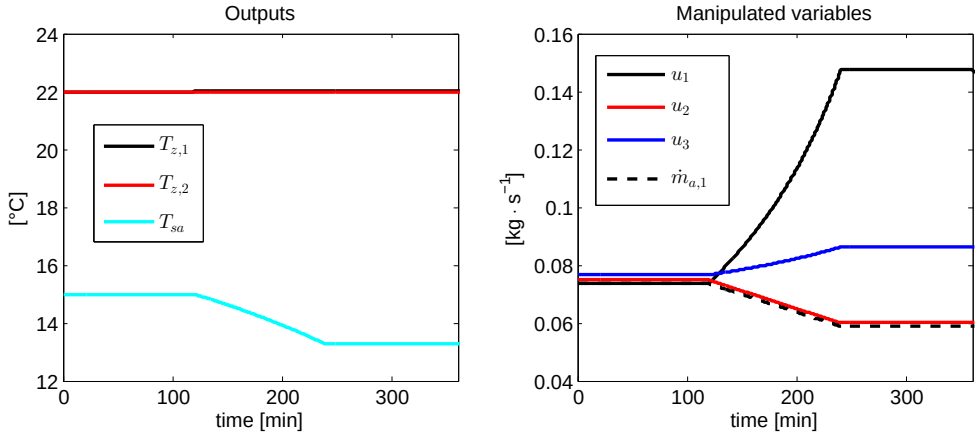


Figure 2.7: Two-zone, single-duct VAVAC system: Stuck VAV Box Damper Zone 1 (-20%).

- **Stuck VAV Box Damper** The position of the damper of the VAV terminal unit is in a wrong position, leading to a different air mass flow rate in the zone; in Figure 2.7 we assume that VAV box damper of zone 1 is blocked: its position is incorrect (-20%) with respect to normal conditions (air mass flow $\dot{m}_{a,1}$ decreases). In order to keep the setpoint

in zone 1, u_1 signal acts in opposite direction with respect to the fault but it is not a real input of the system. The chilled water mass flow rate increases in order to ensure lower supply air temperature and the control signal u_2 decreases in order to guarantee the setpoint value in the second zone.

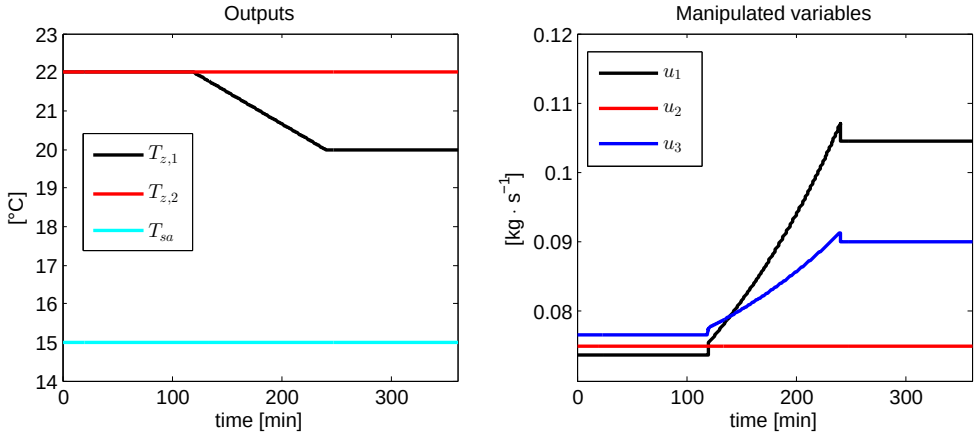


Figure 2.8: Two-zone, single-duct VAVAC system: Temperature Sensor Offset Zone 1 (+2°C).

- **Temperature Sensor Offset** One sensor has an offset in the measured temperature; in Figure 2.8 it is assumed that the sensor of zone 1 has an offset of +2°C with respect to actual conditions. Due to the temperature offset the control system acts in order to maintain the measured temperature at 22°C value (with an actual temperature of 20°C).

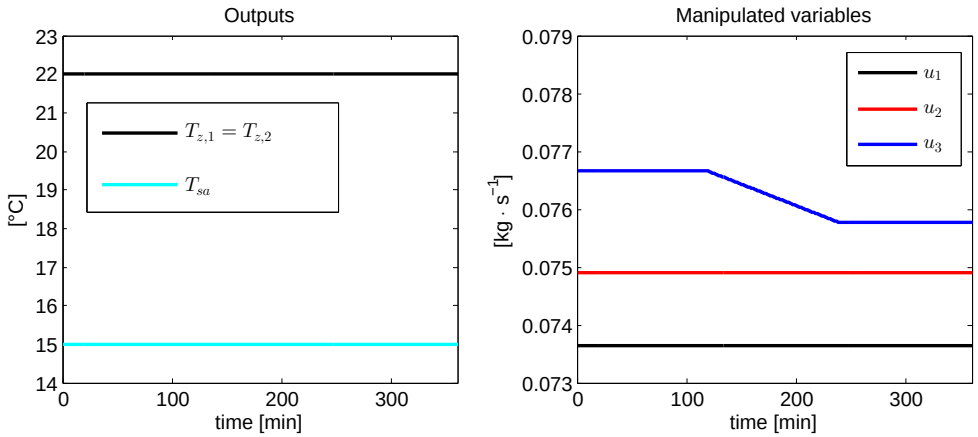


Figure 2.9: Two-zone, single-duct VAVAC system: Stuck Recirculation Damper (+10%).

- **Stuck Recirculation Damper** After the malfunctioning of the recir-

culation damper a different fraction of the external air mass flow is introduced in the zones; for example in Figure 2.9 the recirculation damper position is assumed wrong (+10%) with respect to normal conditions. In order to guarantee load satisfaction in the zones the chilled water mass flow rate decreases during the fault.

2.4.1 Training Data for SVMs Learning

An important step in machine learning methods is the selection of the data to be used in the training phase. With the VAV simulation environment, a database including fault-free data and data under the different faulty conditions is developed. The following variables are collected for the SVMs learning:

- States:
  - Supplied air temperature;
  - Air temperatures of the zones;
  - Outlet water temperature from the cooling coils;
- Manipulated variables:
  - Chilled water mass flow rate;
  - Air mass flow rates of the zones;
- Disturbances:
  - Cooling loads in the zones;
  - External air temperature;
- Other data:
  - Recirculated air temperature;
  - Mixed air temperature.

The database includes simulations with different operative conditions (different setpoints for the controlled variables and different disturbances) and simulating both normal and faulty behaviours at different levels of severity. Finally, only steady-state data are chosen for the SVMs training; in this way the samples in the SVMs space represent stationary conditions of the system: dynamic behaviours and delays would add useless complexity to the classification model.

Table 2.1: Faults and Affected Parameters.

Faults	Affected Parameters	SVM
stuck chilled water valve	chilled water mass flow rate	1
stuck damper zone 1	air mass flow rate zone 1	2
stuck damper zone 2	air mass flow rate zone 2	3
sensor zone 1	temperature sensor offset zone 1	4
sensor zone 2	temperature sensor offset zone 2	5
stuck recirculation damper	percentage of air recirculation	6

2.4.2 FDD Results

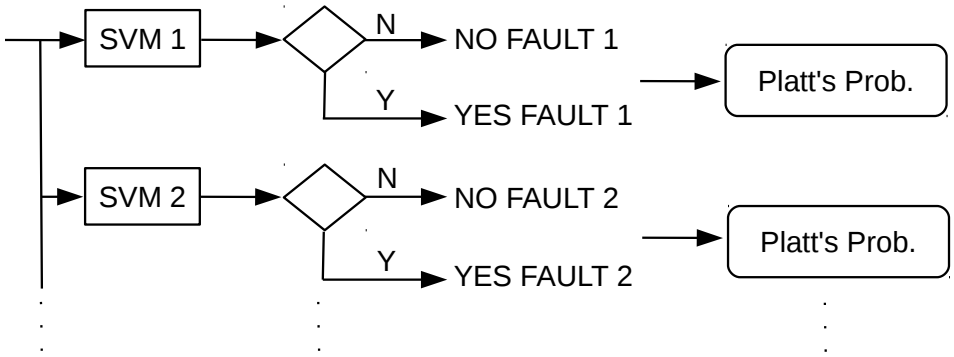


Figure 2.10: Multi-layer SVMs: parallel configuration.

In the following FDD results for the VAV simulated faults are provided, assuming that only one fault at a time may occur in the plant.

Since SVM-based technique is formulated for two-class classification problems, in order to deal with different faulty conditions, a multi-layer SVMs structure is implemented, Figure 2.10. Six SVMs are trained according to the simulated faults, as shown in Table 2.1. For each investigated fault, half of the dataset is used for training, while the rest of data are used for fault diagnosis validation.

In order to achieve an optimal trade-off between generalisation and representation performances, the parameters C and σ of each SVM are tuned by solving an optimisation problem. The goal is to find the parameters values that minimise the diagnosis error:

$$\arg\min_{C,\sigma} \frac{1}{n} \sum_{i}^{n} \mathcal{H}(-\mu_i f(\mathbf{x}_i)) \,, \tag{2.45}$$

where $\mathcal{H}$ is the Heaviside step function. The minimisation problem (2.45) is

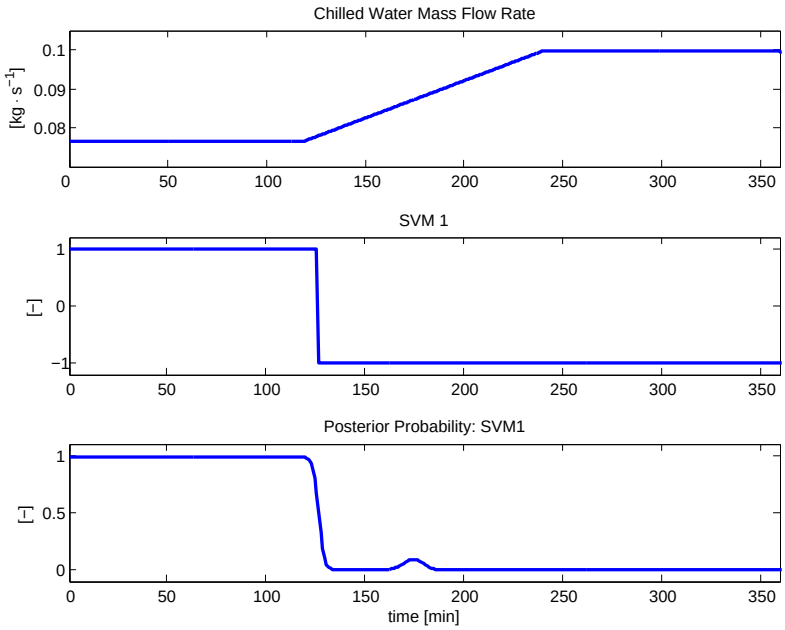


Figure 2.11: Stuck Chilled Water Valve: +30%.

solved by using the particle swarm optimisation algorithm[2].

The Platt's probability is finally provided for each of the simulated fault. In order to solve the optimisation problem (2.43) the solution proposed in [22] is used, where a simple Newton's method is chosen.

2.4.2.1 Stuck Chilled Water Valve

To evaluate the performance of the FDD system we start with the stuck chilled water valve fault. We assume that the chilled water mass flow rate increases by 30% with respect to normal conditions (fault-free), Figure 2.11. SVMs correctly detect and diagnose the fault: SVM 1 is -1 during faulty condition. The posterior probability, which is closed to zero, confirms the correct classification (we remember that Platt's method measures the probability of fault-free conditions). False alarms do not occur during the fault since other SVMs are 1.

In Figure 2.12 the case of analogous fault by -30% degree is depicted: detection and diagnosis are correct.

[2]Details about the optimisation algorithm are given in Appendix C

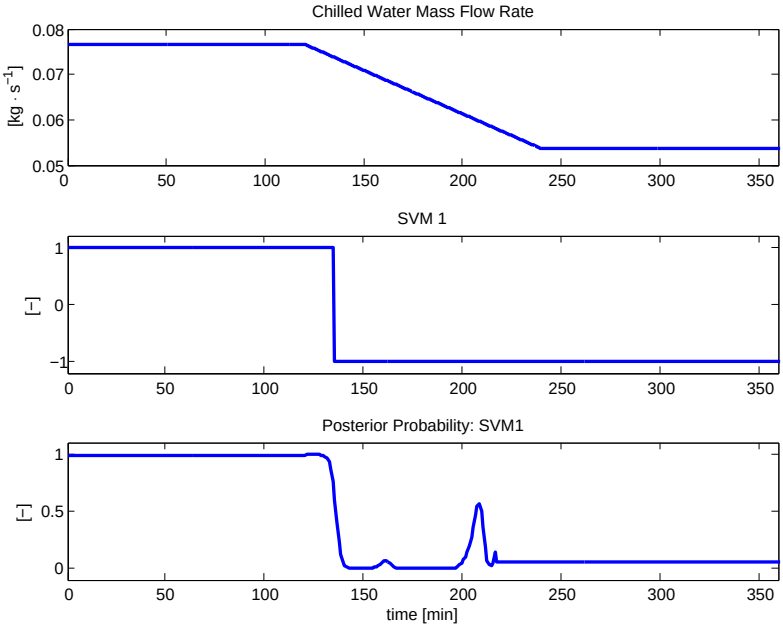


Figure 2.12: Stuck Chilled Water Valve: -30%.

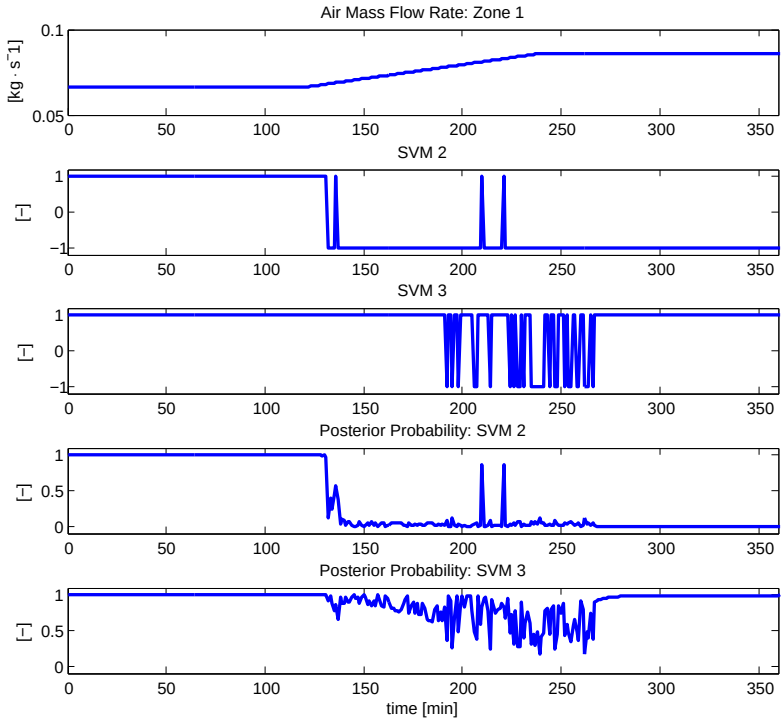


Figure 2.13: Stuck VAV Box Damper Zone 1: +30%.

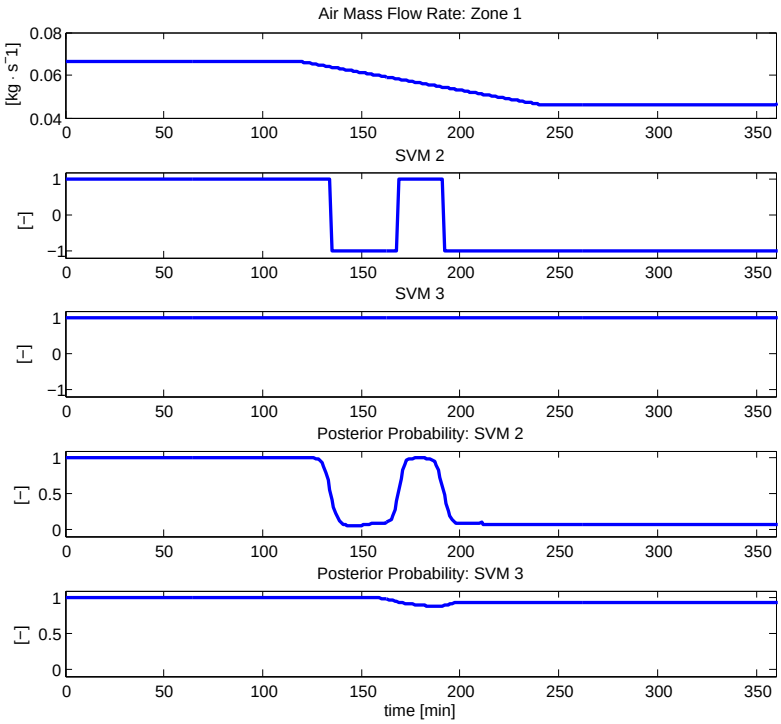


Figure 2.14: Stuck VAV Box Damper Zone 1: -30%.

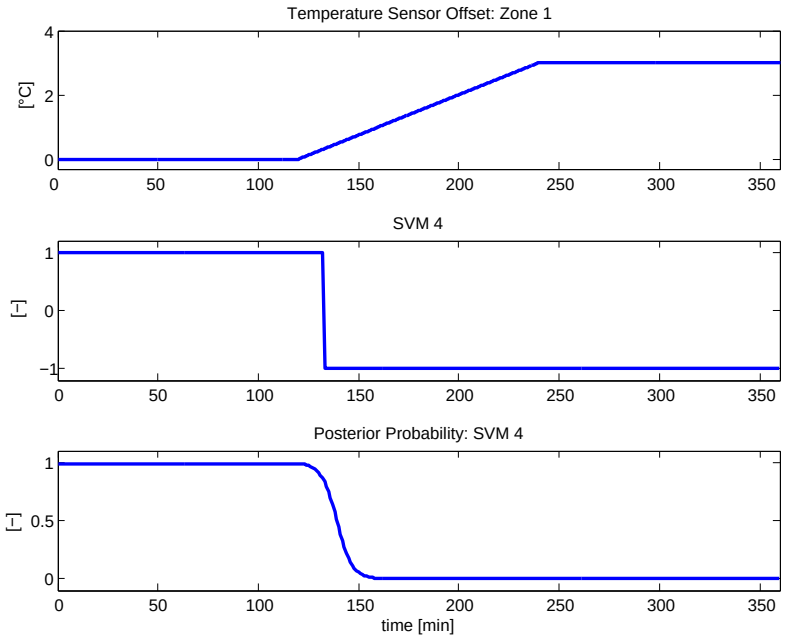


Figure 2.15: Temperature Sensor Offset Zone 1: +3°C.

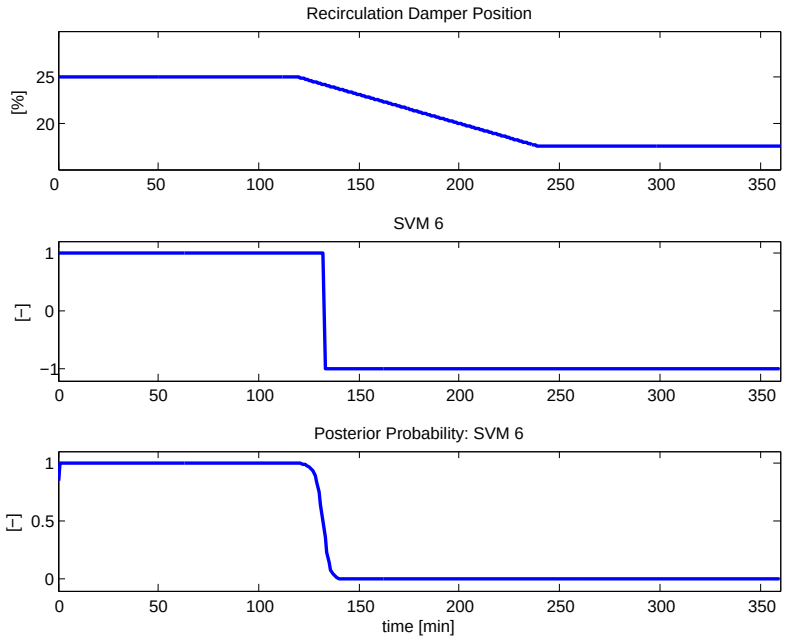


Figure 2.16: Stuck Recirculation Damper: -30%.

2.4.2.2 Stuck VAV Box Damper

It is assumed that VAV box damper of zone 1 is blocked: its position is incorrect (+30%) with respect to normal conditions (air mass flow increases). Since the system is coupled both SVM 2 and SVM 3 are reported, Figure 2.13. In this case, we can appreciate the help of the posterior probability, since the two SVMs signals have not the same values during faulty conditions. SVM 2 Platt's probability is closed to zero and SVM 3 one is almost always greater than 0.5 during faulty behaviour: the posterior probability can help the final decision in case of false alarms.

The case of analogous fault by -30% degree is illustrated in Figure 2.14: posterior probabilities support the classification results.

2.4.2.3 Zone Air Temperature Sensor Offset

It is assumed that the sensor of zone 1 has an offset of +3°C with respect to actual conditions. In this case SVMs correctly detect the fault of zone 1 sensor temperature: SVM 4 is -1 during faulty condition and the posterior probability confirms the diagnosis, Figure 2.15.

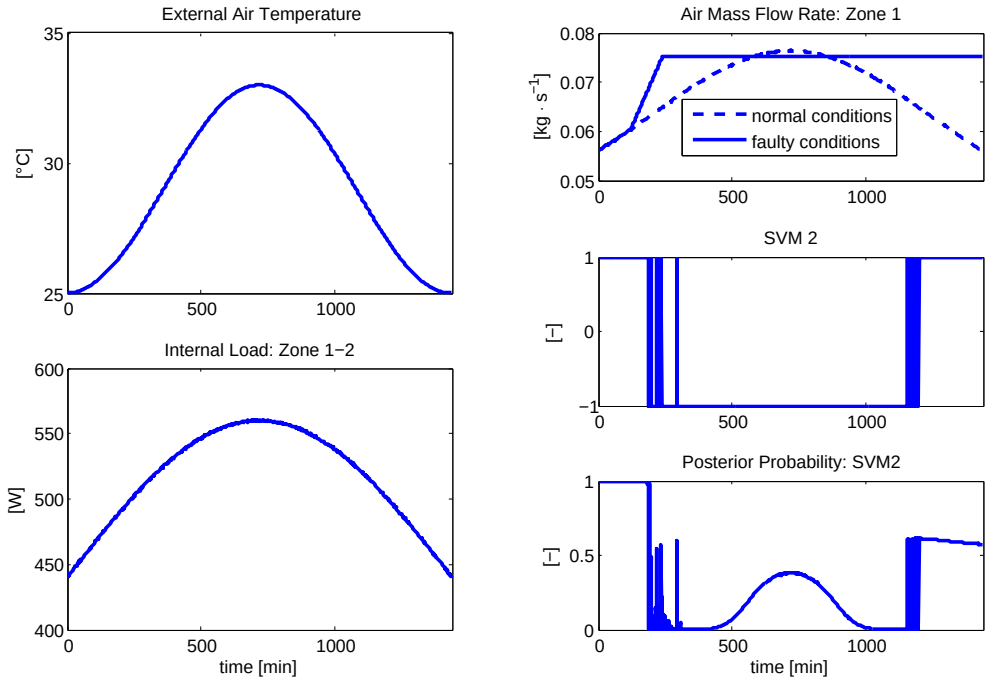


Figure 2.17: Stuck VAV Box Damper Zone 1: varying disturbances.

2.4.2.4 Stuck Recirculation Damper

The recirculation damper position is assumed wrong (-30%) with respect to normal conditions. SVMs and posterior probabilities guarantee that the fault is correctly detected and diagnosed: SVM 6 is -1 and its Platt's probability is null during the presence of the fault, Figure 2.16.

2.4.2.5 Fault Under Varying Disturbances

The case of varying disturbances is also considered. Let us suppose that the outside temperature varies sinusoidally (from 25°C to 33°C) and internal loads vary as depicted in Figure 2.17. The case of a fault on VAV box damper position is investigated: its position is assumed wrong (Figure 2.17, air mass flow, continuous line) with respect to normal conditions (Figure 2.17, air mass flow, dashed line). SVMs system correctly detects and diagnoses the fault. This is confirmed by the value of SVM 2 posterior probability; it is worth noting that probability information adds uncertainty to the decision when the loads and the external air temperature peaks, since the actual air mass flow rate is similar to normal one.

3

Fault Detection for Chiller Systems

Although there is a large body of literature on fault detection and diagnostics for applications in critical processes, relatively little exists for applications in chillers or other vapour compression equipments. With the increased instrumentation of chillers, the interest in FDD systems for refrigeration units is rapidly growing. The benefits of applying FDD to chillers can include less expensive repairs, timely maintenance, and shorter downtime. In the 90s the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) compiled a survey of common faults in chillers and ranked them according to the frequency of occurrence and cost to repair. The major American chiller manufacturers participated in this survey. A large number of possible faults and failures were identified in this study. However, not all of them were considered for further examination in detection and diagnostics methods. For example, most electrical and compressor failures do not need sophisticated detection methods, since their presence is easily detected (e.g. motor burnout). However the survey also showed that degradation faults, on the other hand, generally led to a loss of performance, and were not easily detected (since the chiller was often still operational). After the previous survey, ASHRAE engineers performed experimental tests on a 316-kW centrifugal water-cooled chiller. The goal was to develop a rich database of measurements under different faulty conditions for the development and evaluation of fault detection and diagnostic methods applied to refrigeration units.

3.1 Chiller Data

Data collected in the ASHRAE research project included both transient and steady-state conditions for the following faults: loss of water flow in the evap-

Table 3.1: Considered chiller faults.

Fault Type with Symbol	Normal Operation	SL1	SL2	SL3	SL4
Reduced condenser water flow (fwc)	264-270 gpm	234-250 gpm	209-219 gpm	187-190 gpm	159-166 gpm
Reduced evaporator water flow (fwe)	214-216 gpm	194-196 gpm	175-177 gpm	155-156 gpm	137-141 gpm
Refrigerant leak (rl)	300 lbs	270 lbs	240 lbs	210 lbs	180 lbs
Refrigerant overcharge (ro)	300 lbs	330 lbs	360 lbs	390 lbs	420 lbs
Excess oil (eo)	22 lbs	25 lbs	29 lbs	33 lbs	37 lbs
Condenser fouling (cf)	164 tubes	144 tubes	131 tubes	115 tubes	90 tubes
Non-condensables in refrigerant (nc)	0% N	1.0% N	1.7% N	2.4% N	5.7% N

orator and condenser, refrigerant leak, refrigerant overcharge, presence of excess oil, condenser fouling, presence of non-condensables in the refrigerant, and faulty expansion valve. The faulty conditions in the chiller were tested at four levels of severity (Table 3.1), with the exception of the defective pilot valve, which was a preexisting fault found in the chiller. In detail, the reduced condenser and evaporator water flow were simulated by reducing the base water flow rate, respectively 270 and 216 gpm (US gallons per minute), by about 10% at each fault level. ASHRAE engineers removed and added refrigerant from the system in order to simulate respectively the refrigerant leak and refrigerant overcharge; the base refrigerant charge in the system was 300 pounds (lbs); each fault level changed the refrigerant charge by about 10%. Oil was added to the system to test the excess oil conditions; the base oil charge in the system was 22 pounds and each fault level increased the oil charge by varying amounts as shown in Table 3.1. Tubes were plugged in the condenser to test the condenser fouling; the condenser contained 164 tubes; Table 3.1 shows the number of available tubes for each fault level. Nitrogen was added to the system to simulate non-condensables in the refrigerant; Table 3.1 shows the amount of nitrogen in the system for each fault level. Finally, the defective pilot valve fault was suspected during the commissioning phase of the chiller, when the superheat could not be properly adjusted. Once the pilot valve was replaced, a substantial improvement in the superheat control as well as overall capacity confirmed the suspicion.

Moreover, ASHRAE engineers conducted also normal experimental tests, collecting data describing the correct functioning of the chiller. These tests were used as "benchmarks" for the normal operational values with respect to faulty-operating conditions.

3.1.1 Test Matrix

Research project engineers chose the following three control variables to perform the tests: the outlet evaporator water temperature (T_{eo}), the inlet con-

denser water temperature (T_{ci}), and the chiller cooling capacity; for each fault 27 experimental tests were performed by varying the control variables as shown in Table 3.2. A range was given for the load capacity, since cooling load could not be directly controlled during the tests. Both steady-state and transient data were collected during the tests. The transient data were ensured by the changes in the operating setpoints. On the other hand, stationary operative points could be reached by providing enough time to arrive at steady-state conditions.

Table 3.2: Control variables and corresponding values during the 27 tests in the ASHRAE research project.

T_{eo} [°C]	T_{ci} [°C]	Capacity [%]
50	85	90-100
50	85	50-60
50	85	25-40
50	75	90-100
50	75	50-60
50	75	25-40
50	70	70-80
50	65	45-50
50	62	25-35
45	85	90-100
45	85	50-60
45	85	25-40
45	75	90-100
45	75	50-60
45	75	25-40
45	70	70-90
45	62	45-50
45	62	25-35
40	80	90-100
40	80	50-60
40	80	25-40
40	70	90-100
40	70	50-60
40	70	25-40
40	65	70-80
40	62	45-50
40	62	25-35

3.1.2 Steady-State Detection

Steady-state detection is a critical step for optimisation, control and performance evaluation of processes. Many methods for analysis and processing of data require that the system is in steady-state conditions. The system is consid-

ered in steady-state conditions when all its variables are around an operational value that does not vary unless significant events such as the application of an input, working points changes or a disturbance action, occur.

In this thesis the following variables are used for steady-state detection on ASHRAE data:

- the chilled water supply temperature;
- the pressure of refrigerant in evaporator;
- the entering condenser water temperature;
- the pressure of refrigerant in condenser.

The previous characteristics can provide adequate information about the functioning of the chiller. The two water temperatures are related to the two external water cycles. The two pressures provide information about the refrigerant. For each of the previous variables, the unknown parameters in the following linear regression model are estimated (with a fixed moving time window length):

$$y = mt + b\,. \tag{3.1}$$

If $\|m\|$ is below a chosen threshold the variable is considered in steady-state conditions. The threshold T is chosen as a function of the standard deviation of the data σ:

$$T = a_0 + a_1\sigma\,. \tag{3.2}$$

The optimal coefficients a_0 and a_1 are chosen empirically by trials.

Finally, only the subset where all the four variables are found in steady-state, is considered in steady-state conditions. At the end, about 60% of the original datasets are considered in stationary conditions.

3.1.3 Chiller Features Selection

ASHRAE datasets contain data evolutions of 66 variables of the chiller. Experience gained from past studies [8], [26] indicates that anomaly/fault detection can be more sensitive if certain characteristic quantities or characteristic parameters are used instead of the basic sensor measurements: ASHRAE datasets contain many features that provide redundant information since they are highly correlated; other features provide no useful information for anomaly detection

Table 3.3: Characteristic features employed in the detection analysis.

Characteristic Features
Evaporator Water Temperature Difference
Condenser Water Temperature Difference
Calculated Condenser Heat Rejection Rate
Calculated Evaporator Cooling Rate
Refrigerant Suction Superheat Temperature
Refrigerant Discharge Superheat Temperature
Liquid-line Refrigerant Subcooling from Condenser
Compressor Power
Calculated Compressor Efficiency
Evaporator Approach Temperature
Condenser Approach Temperature
Oil Feed minus Oil Vent Pressure
Oil in Sump minus Oil Feed Temperature
Pressure of Refrigerant in Evaporator
Pressure of Refrigerant in Condenser

analysis. Definitions of the $d = 15$ characteristic features used for fault detection are provided in Table 3.3. These characteristic features can be directly deduced from the sensor measurements using arithmetic operations and thermodynamic refrigerant properties.

3.2 The One-Class Classifier

In this section the reference model that characterizes the correct system behaviour of the chiller is derived: the analysis is developed working on the steady-state data of the 15 variables in Table 3.3 from the ASHRAE research project. A semi-supervised approach is adopted, where only fault-free data are supposed labelled (in fault-free conditions) and are used to identify the reference model of functioning of the system. In this context the fault detection problem is tackled with a novelty detection approach, underlying that we want to classify data, which have not been observed in the past (in faulty behaviour) and are seen as novelties for the machine learning system. To this aim, a one-class classifier is developed, which learns from training sets containing only fault-free data. As a consequence, a key observation is that, unlike the FDD

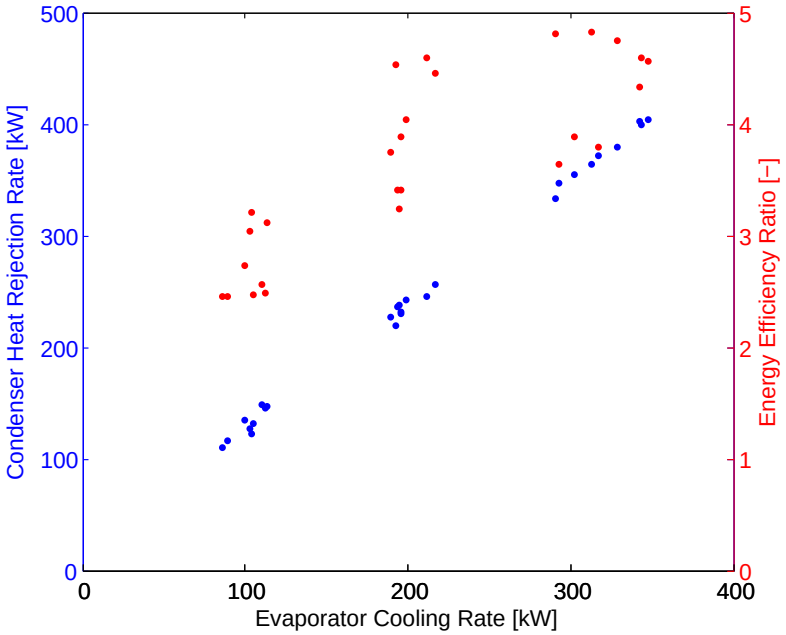


Figure 3.1: Chiller operating points at normal (i.e. fault-free) conditions. The Energy Efficiency Ratio (EER) is defined as the ratio of heat removal from the evaporator over energy input to the compressor. ASHRAE 1043-RP project.

system developed for the VAV plant, fault diagnosis is not performed in this case, since faulty data are not used for training. The goal will be to distinguish a faulty condition from a normal one, without identifying the typology of the fault.

Moreover, it is worth saying that most of the variability in the selected data (Table 3.3) describing normal working behaviour is due to information relating to the operative conditions of the system; for example Figure 3.1 depicts some important chiller operating points in normal conditions. In this section, Principal Components Analysis (PCA) is employed to group in the first principal components the dominant variability of the data that is associated with the behaviour of system through the different operating conditions. Differently from the common use of principal component analysis, in which the components associated with the highest eigenvalues are usually considered in data analysis, in this case these components will be those less informative for the novelty detection problem.

3.2.1 Principal Component Analysis

A classification problem becomes significantly harder as the dimensionality of the data increases. Sometimes data are sparse in the space they occupy leading to difficulties for the learning phase. A high data dimensionality is a problem for many classification algorithms given the consequent high computational cost and memory usage. Moreover, huge dimensionality in the data can lead to poor understanding of the describing model [13].

In this work, Principal Component Analysis (PCA) is employed for dimensionality reduction. PCA is a linear projection-based method that transforms a set of variables into a new set of uncorrelated variables, named Principal Components (PCs). PCA is run for a dataset defined by an $n \times d$ design matrix X where the d columns are variables and the n rows are observations. X is written in terms of the $n \times l$ scores matrix T, where $l \leq d$, and the $d \times l$ loadings matrix P, plus a residual matrix E, as follows:

$$X = TP^T + E = \sum_{i=1}^{l} \mathbf{t}_i \mathbf{p}_i^T + E\,, \tag{3.3}$$

where $\mathbf{t}_i = X\mathbf{p}_i$. The vectors $\{\mathbf{p}_i\}$ are the PCs and if $l = d$, then $E = 0$. PCs are arranged in order of magnitude of variability of X explained: the first PC, $\mathbf{p}_1$, can be geometrically interpreted as the direction where most of the variability lies, then other PCs define orthogonal directions where less and less variability is contained. PCA is called also eigenvalue decomposition, as each of the PCs is related to an eigenvalue of the matrix in exam, ordered in terms of magnitude. The transformation induced by PCA can therefore be employed for reducing the dimensionality of the problem, as just $l < d$ variables can be employed to express a certain amount of variability in the input dataset.

In this thesis, principal component analysis is applied on the data belonging to the 15 characteristic features (Table 3.3) of the normal dataset. In Figure 3.2 the cumulative explained variance of the 15 considered principal components is depicted. It is worth noting that the first three components contain more than 95% of the variability.

3.2.2 One-Class SVMs

One-Class Support Vector Machines (OCSVMs) are a powerful tool to perform one-class classification. OCSVMs may be viewed as standard two-class SVMs[1],

[1]SVMs theory is detailed in Appendix D

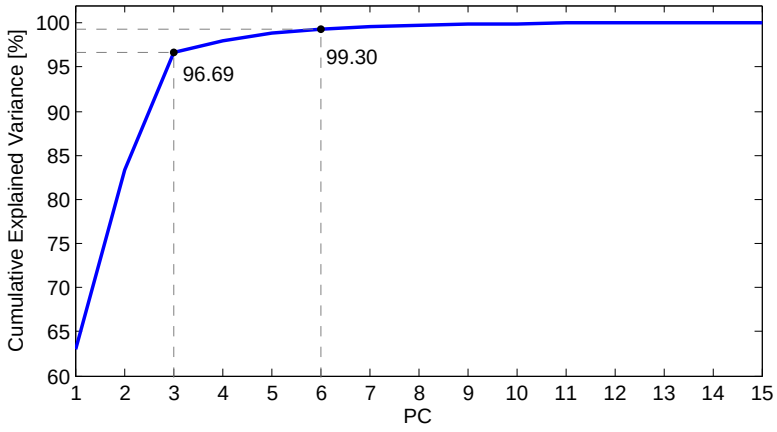


Figure 3.2: Cumulative explained variance at the increase of the PCs considered.

where all the training data lie in the first class, and only the origin is taken as member of the second class. Given a training dataset $X = \{\mathbf{x}_1, .., \mathbf{x}_n\}$, $\mathbf{x}_i \in \mathbb{R}^d$, the OCSVMs algorithm maps the data into a higher dimensional feature space and finds a hyperplane to separate all the data objects from the origin with maximum margin, by solving the following quadratic programming problem:

$$\arg\min_{\mathbf{w},\xi,\rho} \frac{1}{2}\|\mathbf{w}\|^2 + \frac{1}{\nu n}\sum_{i=1}^{n}\xi_i - \rho\,, \tag{3.4}$$

$$\text{subject to} \quad \begin{cases} \langle \mathbf{w}, \boldsymbol{\phi}(\mathbf{x}_i)\rangle \geq \rho - \xi_i \\ \xi_i \geq 0 \end{cases}, \tag{3.5}$$

where n is the number of training samples, $\xi = [\xi_1 \ldots \xi_n]$ and $\boldsymbol{\phi}$ is generally a non-linear map. $\mathbf{w}$ represents the hyperplane vector and ρ is its offset in the feature space. The slack variable ξ_i measures the degree of misclassification of the data. The trade-off parameter $\nu \in (0, 1]$ is an upper bound on the fraction of training samples outside the decision boundaries and a lower bound on the fraction of support vectors [31].

As in the two-class SVMs system, the RBF function is used as kernel to perform one-class classification:

$$k(\mathbf{x}_i, \mathbf{x}_j) = \langle \boldsymbol{\phi}(\mathbf{x}_i), \boldsymbol{\phi}(\mathbf{x}_j)\rangle = \exp[-\|\mathbf{x}_i - \mathbf{x}_j\|^2/(2\sigma^2)]\,. \tag{3.6}$$

OCSVMs Tuning In the considered semi-supervised context, the goal is to identify the correct region of functioning of the chiller by setting the respective boundaries (which correspond to the separating hyperplane in the enlarged space).

An optimal parameters tuning is needed for the OCSVMs: we remember that the width parameter σ of the Gaussian kernel has a significant impact on accuracy and generalisation performance. As σ increases, the number of support vectors decreases and the decision boundaries become looser. In addition, the parameter ν also affects the shape of the decision boundaries of the OCSVMs: as ν increases, the number of support vectors increases and the number of misclassified training samples grows.

Since ν is closely related to the fraction of misclassified training samples (i.e. points representing correct functioning but outside the fault-free region provided by the classifier), it is usually set to a small value to ensure a small misclassification error rate on the training phase: therefore, choosing an appropriate value for σ is the main challenge of building a satisfactory OCSVM model. To this aim, the heuristic approach proposed by [17], which chooses σ via *Tightness Detecting*, is employed. Based on the assumption that training samples are representative, an ideal decision boundary of OCSVMs should be neither tight to ensure the generalisation of classifiers, nor loose to ensure the sensitivity to faulty data. Since the relationship between the tightness of the boundaries and σ is monotonous, an iterative algorithm is used to choose the width of the Gaussian kernel via *Tightness Detecting* so that an appropriate tightness of decision boundaries is found. The steps of the algorithm are:

1. find the initial upper bound σ_u and lower bound σ_l;
2. take $\sigma = (\sigma_u + \sigma_l)/2$. Operate the OCSVM algorithm to obtain the decision boundaries;
3. evaluate the tightness of the boundaries via *Tightness Detecting*. If the *Tightness Detecting* result turns out to be tight, let $\sigma_l = \sigma$; if it is loose, $\sigma_u = \sigma$; if neither tight nor loose, stop iteration;
4. if $\sigma_u - \sigma_l$ is less than the given threshold, stop iteration. Otherwise return to step 2.

Tightness Detecting To evaluate the tightness of the decision boundaries, Want et al. [17] propose the following algorithm. The underlying idea is that if there exists one large hole inside the boundaries, these will be considered loose; if the boundaries nearby two neighbouring samples are concave, they will be considered tight.

The algorithm implementation is as follows; given the training set $\mathbb{T} \subset \mathbb{R}^d$ and the signed distance f from the separating hyperplane in the enlarged space we define the threshold T. Furthermore we calculate all distances between any two training samples; we define the nearest distance D_i for each sample $\mathbf{x}_i$ and then the parameter D:

$$D = \max_i\{D_i\} = \max_i\{\min_j \|\mathbf{x}_i - \mathbf{x}_j\|\}\,. \tag{3.7}$$

1. *Evaluate if the boundaries are loose.* For all pairs $(\mathbf{x}_i, \mathbf{x}_j)$ such that $\|\mathbf{x}_i - \mathbf{x}_j\| > T, \mathbf{x}_i, \mathbf{x}_j \in \mathbb{T}, i \neq j$, if there exists at least one (or more) pair whose midpoint $\bar{\mathbf{x}}$ is inside the decision boundaries:

$$f(\bar{\mathbf{x}}) \geq 0\,, \tag{3.8}$$

and

$$\mathbb{T} \cap \{\mathbf{x} |\, \|\mathbf{x} - \bar{\mathbf{x}}\| \leq \frac{D}{2}, \mathbf{x} \in \mathbb{R}^d\} = \emptyset\,, \tag{3.9}$$

the decision boundaries are considered loose and then go to step 2. If no judgment is made after checking all the above conditioned pairs, the boundaries are considered not loose and go to step 2.

2. *Evaluate if the boundaries are tight.* For all pairs $(\mathbf{x}_i, \mathbf{x}_j)$ such that $\|\mathbf{x}_i - \mathbf{x}_j\| \leq T, \mathbf{x}_i, \mathbf{x}_j \in \mathbb{T}, i \neq j$, if there exists one (or more) pair whose midpoint $\bar{\mathbf{x}}$ is outside the decision boundaries:

$$f(\bar{\mathbf{x}}) < 0\,, \tag{3.10}$$

the boundaries are judged tight, and then go to step 3. If no judgment is made after checking all the above conditioned pairs, the boundaries are judged not tight and go to step 3.

3. *Give a judgment for tightness of decision boundaries.* Referring to the judgment provided in 1. and 2., give a final judgment. There are two special cases:

   - considering the possibility that the distribution of the training samples is non-uniform, the boundaries may be judged loose and tight simultaneously; when this situation happens, the boundaries are preferred to be considered tight;
   - the other case is that the boundaries are judged neither loose nor tight; when this situation happens, the optimal σ is found.

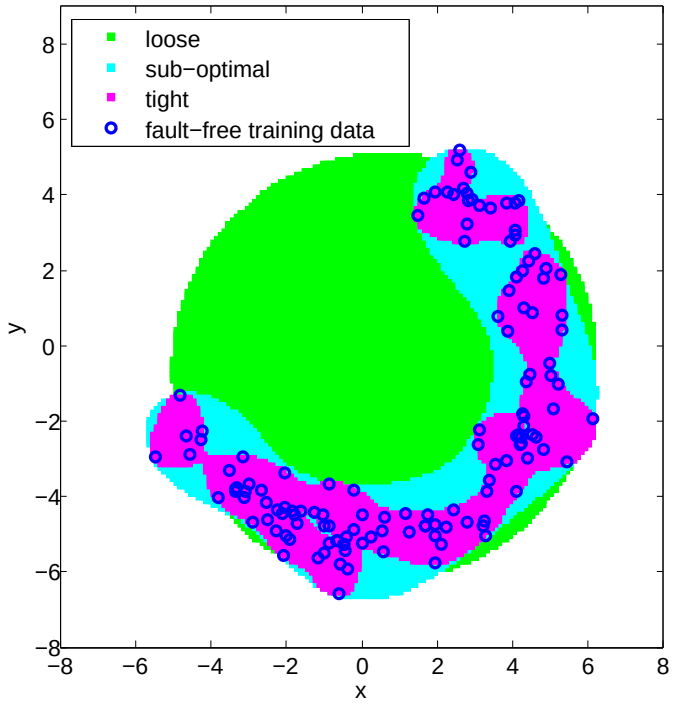


Figure 3.3: Examples of tightness of the decision boundaries; the sub-optimal classifier can be seen as a trade-off between the loose (that barely adapts to the data at hand) and the tight solution (that clearly over-fits the samples available).

To evaluate the performance of the *Tightness Detecting* algorithm, the following example is provided. We consider a dataset of points belonging to a Rosenbrock banana set, whose points are supposed to belong to the normal functioning of a system. Figure 3.3 depicts the training data fed to the *Tightness Detecting* algorithm. Three different boundaries are shown. In detail, the green boundary is considered loose, whereas the magenta boundary is considered tight; the light blue one is considered sub-optimal, i.e. neither loose nor tight, and is the output of the *Tightness Detecting* algorithm.

In order to validate the decision boundaries, we use as validation data different fault-free working points and we suppose that a Gaussian distribution represents a faulty behaviour of the system (Figure 3.4). The circles represents normal working points while the points represent faulty situations. There are four possibilities:

- Faulty data correctly classified;
- Fault-free data correctly classified;
- Faulty data classified as fault-free data;

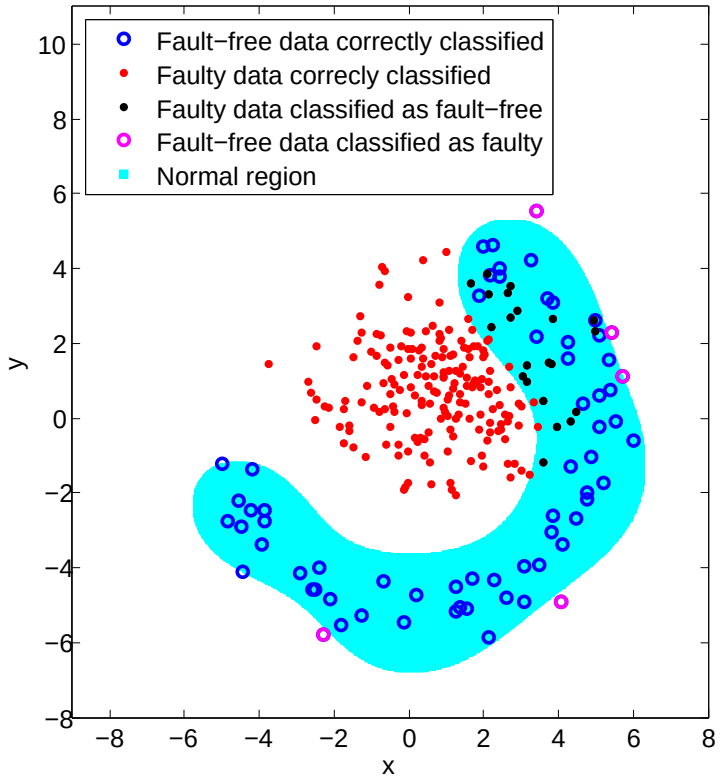


Figure 3.4: Validation phase for the *Tightness Detecting* algorithm.

- Fault-free data classified as faulty data.

It is worth noting that the boundaries play a key role in classification results. With looser boundaries there would be a higher number of faulty situations classified as normal with a lower number of fault-free data classified as faulty. Conversely, the opposite occurs with tighter boundaries.

3.2.3 Reference Model

In the following the reference model that characterizes the baseline system behaviour of the chiller is derived. The dataset of the principal components of fault-free data is divided into training (67% of samples) and validation (the remaining 33%); training principal components are then fed to the one-class classifier. The tuning phase of the one-class SVM is tackled as follows: the parameter ν is set to a small value to ensure a small misclassification error rate in the training phase, as motivated in Section 3.2.2, whereas the parameter σ is chosen via *tightness detecting*.

Different classifiers are computed depending on the input subset selection considered, adopting a smart reduction of the principal components fed to the

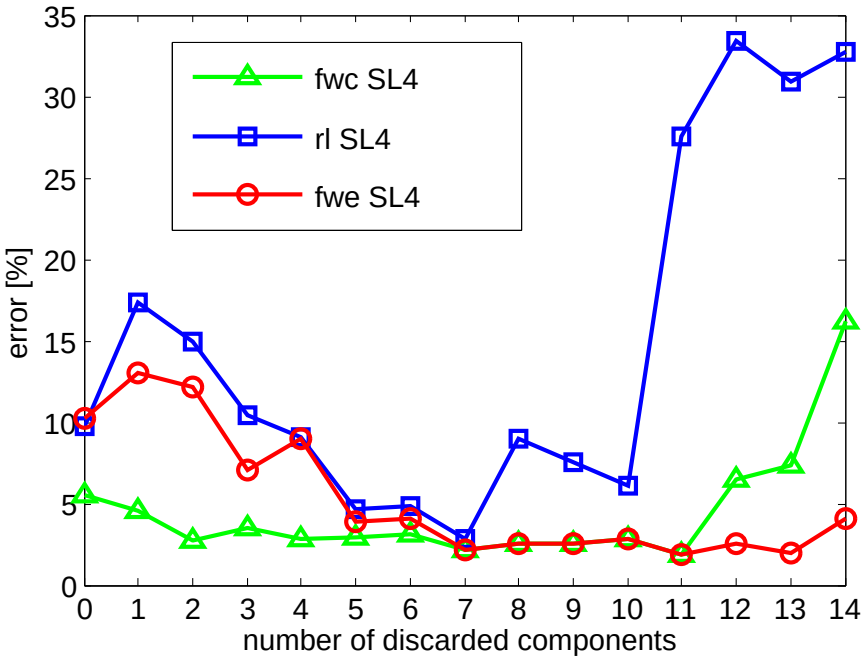


Figure 3.5: Error rate as function of the number of discarded top eigenvectors.

one-class classifier. Classifiers performances are then tested on new data with faulty (see Table 3.1) and fault-free related runs.

Two experiment types are developed based on different choice of the classifier input; in the first case, the classifiers $\{G^{(i)}\}_{i=0}^{d-1}$ are computed by adopting different input spaces $\{U_i\}$ as follows:

$$G^{(i)}(U_i)\,, \qquad U_i = [\mathbf{p}_{i+1} \dots \mathbf{p}_d]\,, \tag{3.11}$$

for $i = 0, \dots, d-1$. The performances, in terms of misclassification error rates, are provided in Figure 3.5 for three different types of faults: the error rate is defined as the ratio of wrongly predicted data on total testing data. The abscissa of Figure 3.5 represents the number of PCs discarded in the input space; more precisely, the leftmost point corresponds to the error rate without dimensionality reduction, while the next point corresponds to the error rate when data are projected onto the space spanned by all eigenvectors except that associated with the largest eigenvalue, and so on. It is worth noticing that the curves generally have a U-shape: the minimum error rate is never achieved when the PCs that contain most of the input variability are included into the input space; better results are obtained once the first PCs are discarded. This can be interpreted as an improvement of the classification performance once the most of the variability (mainly the one related to operating conditions

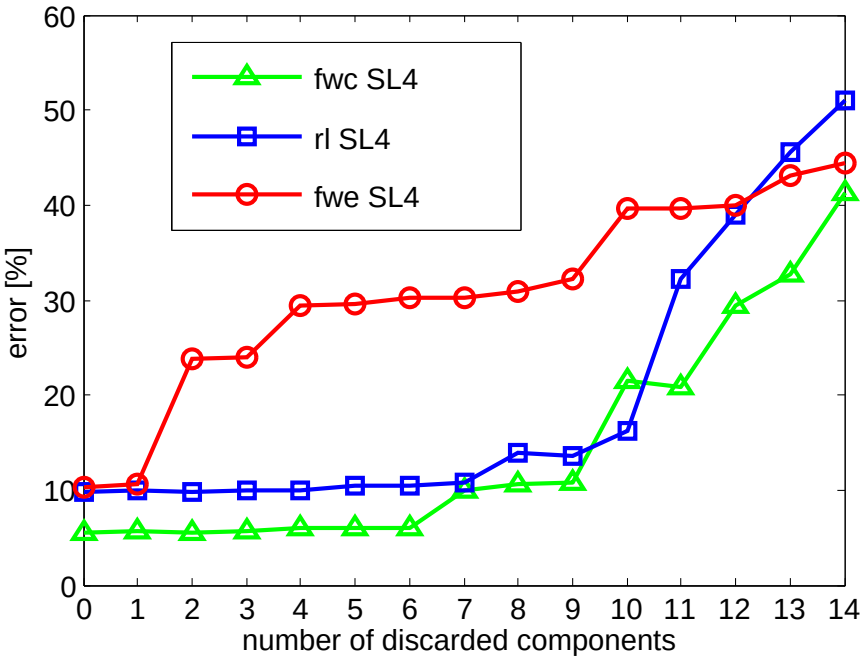


Figure 3.6: Error rate as function of the number of discarded minor eigenvectors.

changes) is left out from the input space. This result motivates the use of PCA analysis as procedure for ordering the variability of the system in exam and allowing the concealed anomalies related variability to be highlighted and more easily identified by the classifier. Projecting the data onto the directions of eigenvectors associated with smaller eigenvalues (i.e. from the 8th to the last one) and then feeding the classifier with the corresponding principal components, the average error rate is at least halved, compared to the error without dimension reduction.

The idea that the first PCs describe most of the usual operating condition variability is also supported by a second type of experiment, where the classifier input space is changed again; the new classifiers $\left\{H^{(i)}\right\}_{i=0}^{d-1}$ differ for the working input space V_i employed:

$$H^{(i)}(V_i)\,, \qquad V_i = [\mathbf{p}_1 \dots \mathbf{p}_{d-i}]\,, \tag{3.12}$$

for $i = 0, \dots, d-1$. The results for this second type of experiments are reported in Figure 3.6: it can be seen how the misclassication error rate increases as the number of discarded PCs associated with the smallest eigenvalues is augmented. This proves how first PCs are not informative features for novelty detection.

This intuition can be also corroborated by the following observation. Fig-

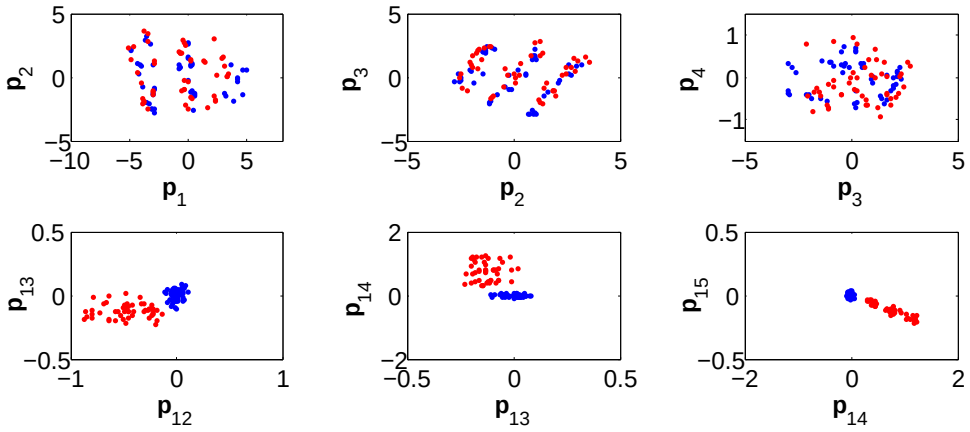


Figure 3.7: Principal components: fault-free (blue points) and fwe SL4 (red points).

ure 3.7 depicts normal validation data (blue points) and faulty data (i.e fwe data, red points) projected onto the first four principal components and the last four principal components (the principal components are coupled in 2-D dimensions). It is worth noting that while the first principal components do not exhibit substantial differences, normal and faulty data are easily separable in the last principal components.

Similar results can be appreciated for the other faults: first principal components are not easily separable; they add useless information and complexity to the one-class classification problem; discarding them helps the different values of the smallest principal components to be highlighted in the classification task.

Following the experiments outcome, the classifier $G^{(7)}$ is chosen to evaluate the classification performances in Section 3.3, as it exhibits the lowest average error rate with the considered faults.

3.3 Classification Results

In order to assess the performances of the one-class classifier, data in faulty conditions (Table 3.1) coupled with normal validation data are used. We evaluate the ability of the classification model to distinguish normal behaviour from anomalous one with Receiving Operating Characteristic (ROC) analysis, exploiting the Area Under the ROC Curve (AUC) as indicator of discriminatory power.

3.3.1 ROC Analysis

In a two-class prediction problem (binary classification), in which the outcomes are either positive (P, in this case in fault-free conditions) or negative (N, faulty conditions), there are four possibilities. If the outcome from a prediction is P and the actual value is also P, then it is called a True Positive (TP); however if the actual value is N then it is said to be a False Positive (FP). Conversely, a True Negative (TN) occurs when both the prediction outcome and the actual value are N, and False Negative (FN) happens when the prediction outcome is N while the actual value is P. The four outcomes can be formulated in a 2×2 confusion matrix, as follows:

Table 3.4: Confusion matrix

	Condition positive	Condition negative
Test outcome positive	True Positive (TP)	False Positive (FP)
Test outcome negative	False Negative (FN)	True Negative (TN)

The following parameters are usually adopted to evaluate the performance of the classification:

- the True Positive Rate (TPR) measures the proportion of positives which are correctly identified:

$$\text{TPR} = \frac{TP}{TP + FN}; \tag{3.13}$$

- the True Negative Rate (TNR) measures the proportion of negatives which are correctly identified:

$$\text{TNR} = \frac{TN}{FP + TN}; \tag{3.14}$$

- the Positive Predictive Value (PPV) is the proportion of positive outcomes which are true positive:

$$\text{PPV} = \frac{TP}{TP + FP}\,; \tag{3.15}$$

- the Negative Predictive Value (NPV) is the proportion of negative outcomes which are true negatives:

$$\text{NPV} = \frac{TN}{TN + FN}\,; \tag{3.16}$$

- the False Positive Rate (FPR) measures the proportion of negatives which are wrongly identified:

$$\text{FPR} = \frac{FP}{TN + FP}\,; \tag{3.17}$$

- the False Negative Rate (FNR) measures the proportion of positives which are wrongly identified:

$$\text{FNR} = \frac{FN}{TP + FN}\,; \tag{3.18}$$

- the Accuracy (ACC) is the proportion of outcomes which are correctly identified:

$$\text{ACC} = (TP + TN)/(P + N)\,. \tag{3.19}$$

A Receiver Operating Characteristic (ROC), or simply ROC curve, is a graphical plot which illustrates the performance of a binary classifier system as its discrimination threshold varies. The corresponding Area Under the ROC curve (AUC) measures the ability to correctly classify data.

For example in Figure 3.8 we suppose that normal and faulty behaviours are described with two Gaussian distributions with different standard deviations and different means. Assuming that the threshold provided by the classifier is the black vertical continuous line, that separates the FN-TN values from the TP-FP ones, the ROC curve plots the true positive rate and the false positive rate at different thresholds settings (dashed line). In Figure 3.9 examples of ROC curves are plotted. The red line, which passes through the point (0,1), represents a perfect test in which 100% of the data are correctly classified (AUC = 1). The dashed line refers to the performance of a random classifier (AUC = 0.5). The blue line represents a ROC curve representing a good classification result since the curve is above the diagonal (the classification results work better that random classifier).

3.3.2 Reduced Condenser and Evaporator Water Flow

The classification results for the reduced evaporator and condenser water flows are now provided. In Figures 3.10 and 3.11, we compute the ROC curves to analyse the classification model on sets composed by normal validation data and data related to the two anomalies at the four severity levels. It is worth noting that the classification task has excellent performances: the area under the ROC curve is always higher than 0.9.

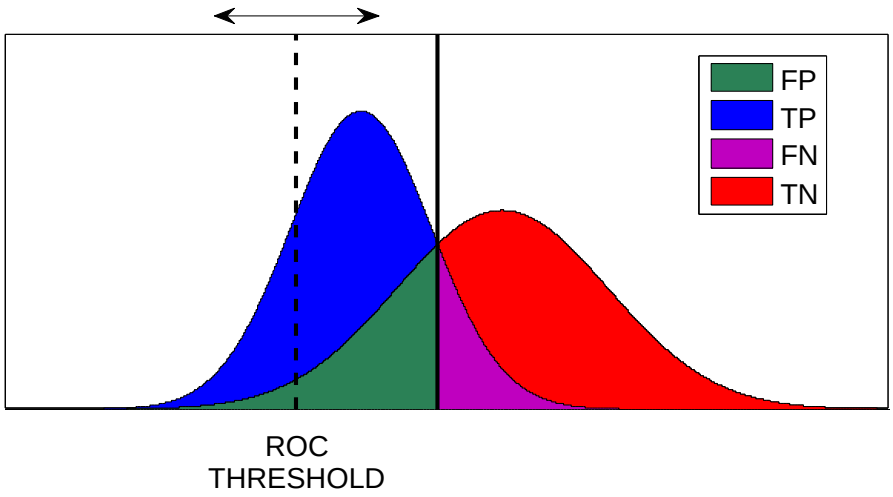


Figure 3.8: Faulty and fault-free data: two different Gaussian distributions.

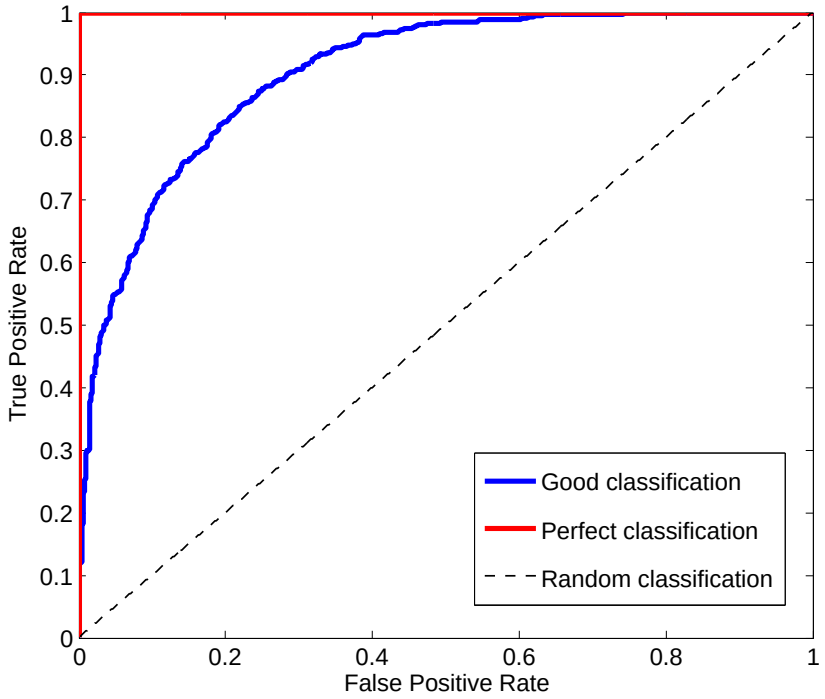


Figure 3.9: Examples of ROC curves.

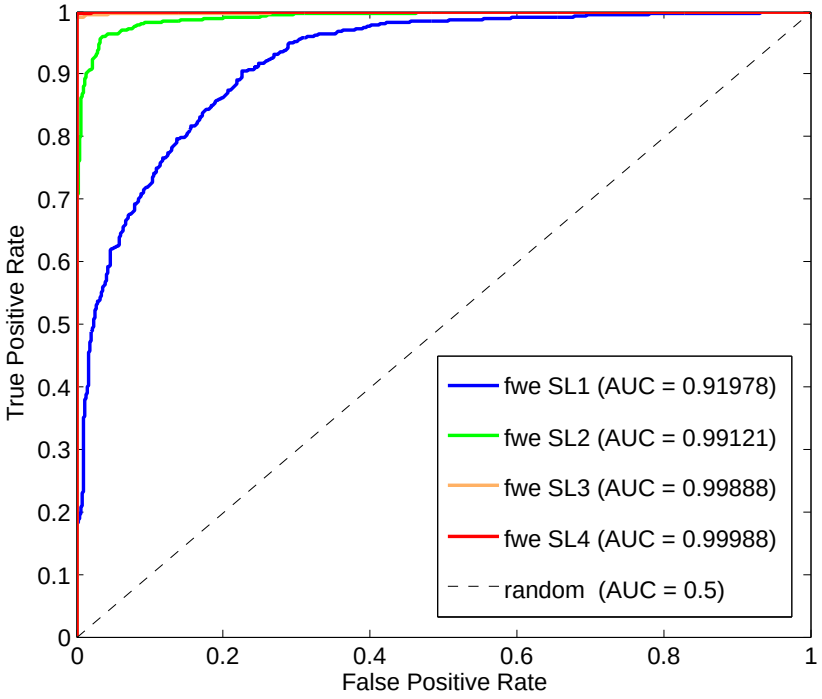


Figure 3.10: ROC analysis: reduced evaporator water flow.

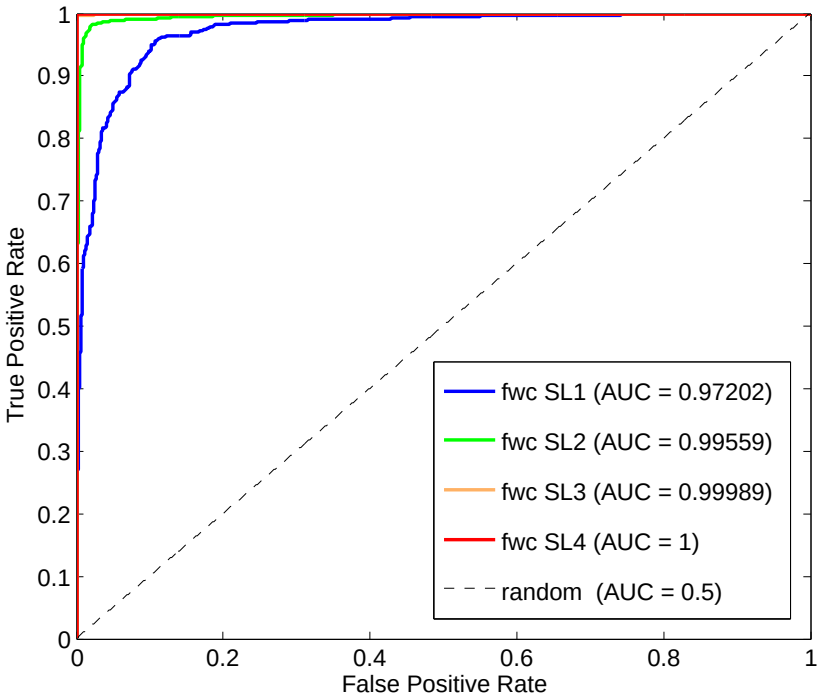


Figure 3.11: ROC analysis: reduced condenser water flow.

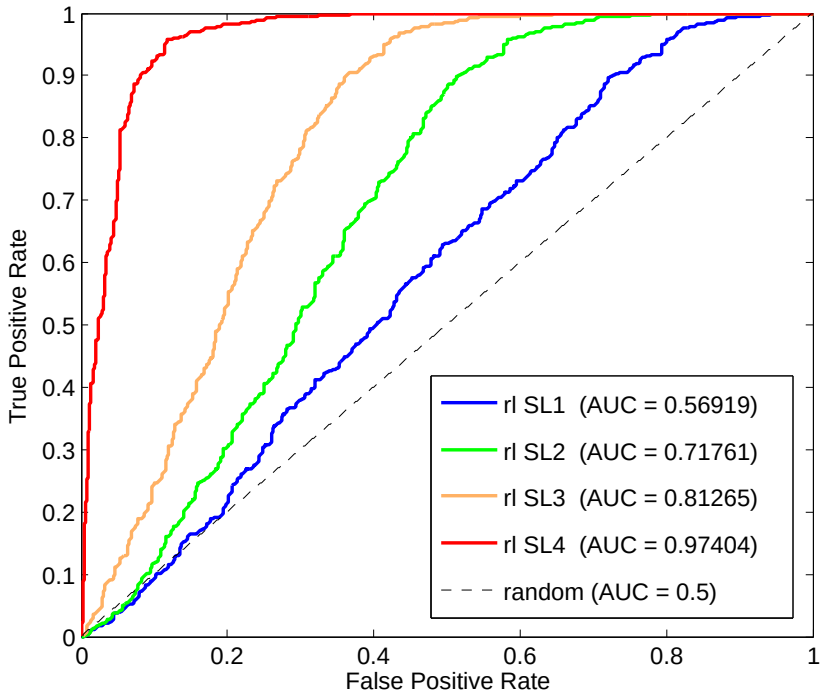


Figure 3.12: ROC analysis: refrigerant leak.

3.3.3 Refrigerant Leak and Overcharge

The performances of the classification of refrigerant leak and refrigerant overcharge faults are given by ROC curves in Figures 3.12 and 3.13. All the ROC curves are above the diagonal representing a good classification result. It is worth noting that the classification score markedly increases as the severity levels raises: high intensity faults are easier to be detected than low intensity faults (as intuitively expected).

3.3.4 Excess Oil

In this case the SL2 and SL3 faults are perfectly detected while the SL1 fault has good performances (Figure 3.14). Data on SL4 severity level reach stationary conditions in few points, and they are not given to the classifier for performance analysis.

3.3.5 Condenser Fouling and Non-Condensables in Refrigerant

Excellent results are achieved for the condenser fouling and non-condensables in refrigerant (Figures 3.15 and 3.16), where classification works optimally even

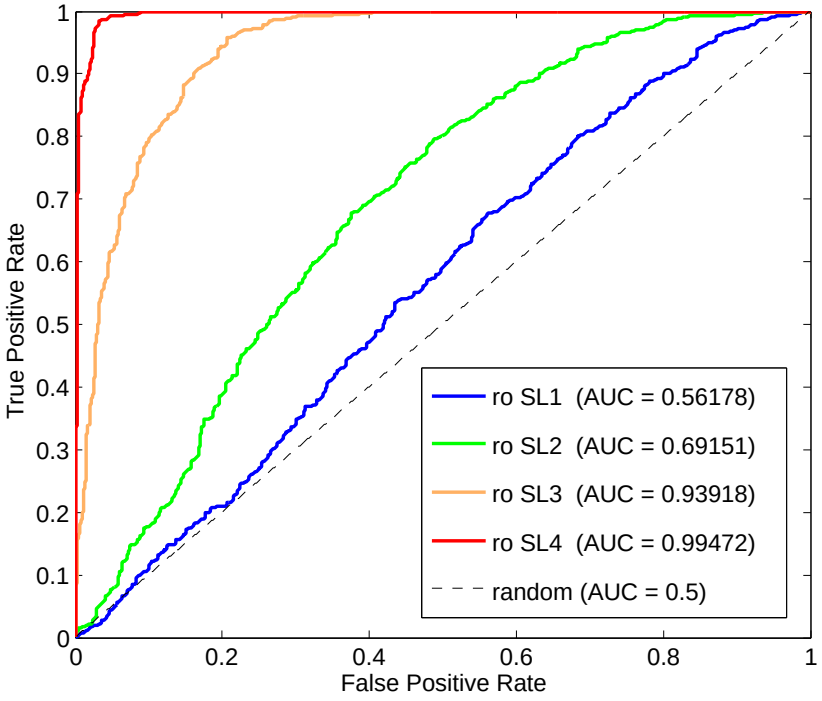


Figure 3.13: ROC analysis: refrigerant overcharge.

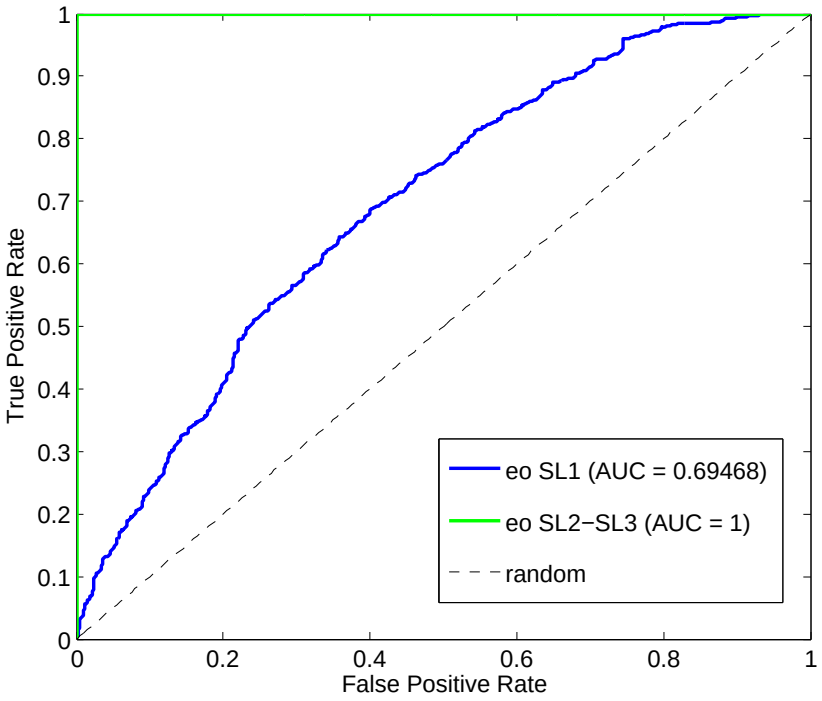


Figure 3.14: ROC analysis: excess oil.

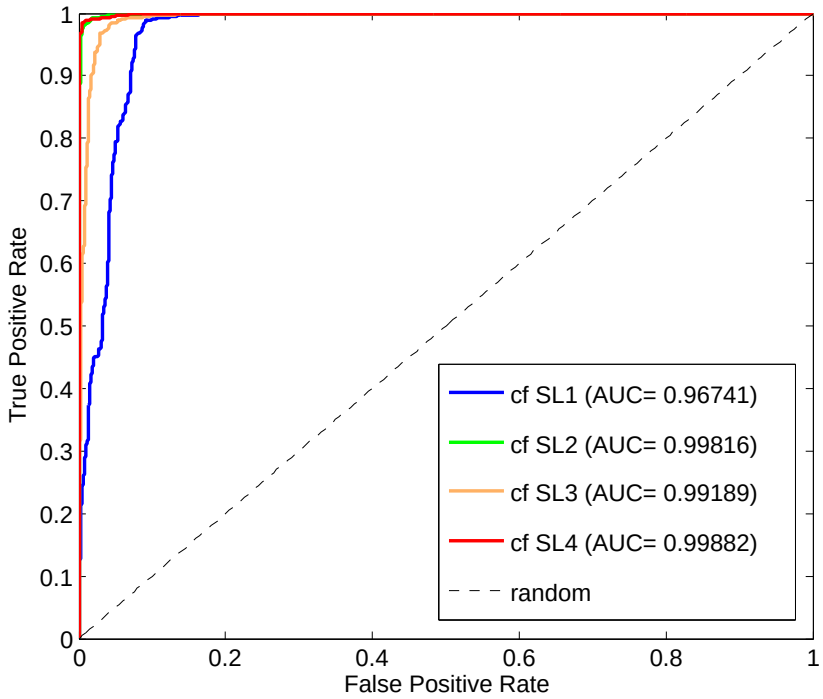


Figure 3.15: ROC analysis: condenser fouling.

in low severity levels. Only the first three severity levels results are reported in Figure 3.16 since the fourth severity level test could not reach all desired operating conditions during the tests in the project.

3.3.6 Comments

The classification has exhibited good results with all the considered faults. It is not easy to make a comparison between the results for the different typologies of faults, since their severity levels do not have a common scale of comparison. However, the ROC analysis confirms that both the reduced condenser and evaporator water flow rate are easily detected, [8].

Furthermore, it is interesting to exploit the distance of the data in exam from the OCSVM decision boundary as a mean for inferring information about anomaly severity: the underlying idea is that observations with small fault nature will be close to the normal data (and therefore to the decision boundary), while severe faults related samples will be distant from normally classified data and the OCSVM decision boundary.

Figure 3.17 shows the signed distances (in the 8D dimensional space) of all the test points belonging to two anomalies (reduced evaporator water flow and reduced condenser water flow) where only two principal components ($\mathbf{p}_{11}$-$\mathbf{p}_{12}$)

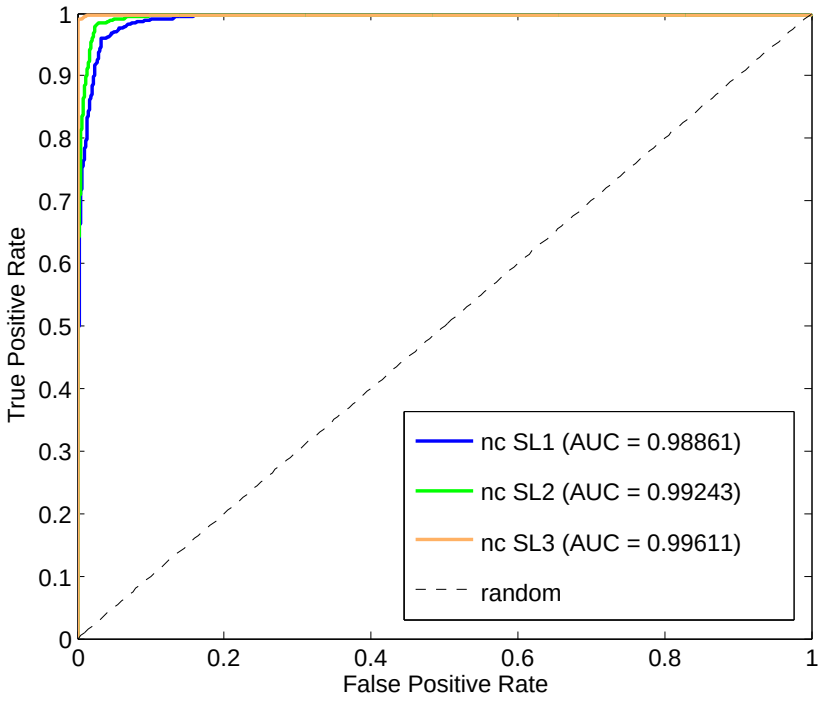


Figure 3.16: ROC analysis: non-condensables in refrigerant.

are used for visualization. A positive distance value corresponds to a test point classified as normal, whereas a negative one refers to an anomalous behaviour; it can be seen that as the severity level of fault rises, the magnitude of negative distance increases. This is confirmed by the boxplot representation depicted in Figure 3.18, where statistical properties of the data, related to the reduced evaporator water flow rate, at different severity levels, are shown. This suggests that the distance from the decision boundary could be exploited for predictive maintenance purposes.

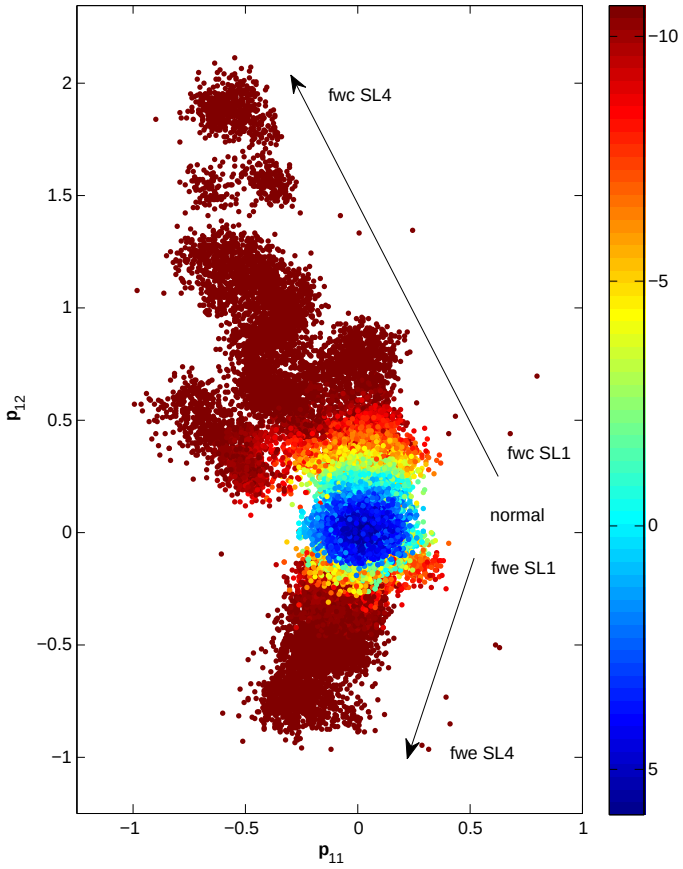


Figure 3.17: Signed distance related to test points corresponding to the reduced condenser water flow and reduced evaporator water flow at different levels of severity.

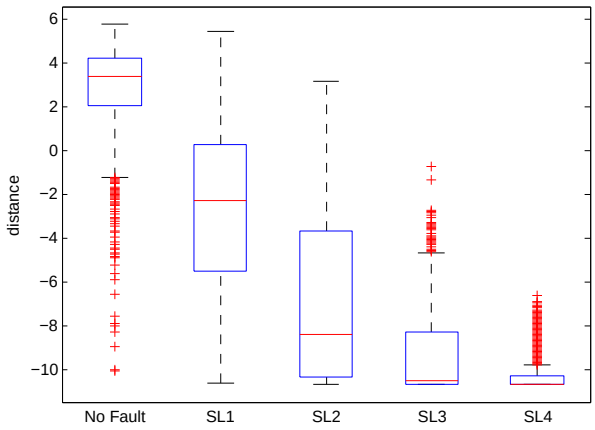


Figure 3.18: Boxplot related to the signed distance of points corresponding to the reduced evaporator water flow at different levels of severity.

4

Conclusions

Faulty operations of Heating, Ventilation and Air Conditioning (HVAC) systems can lead to discomfort for the occupants, energy wastage, unreliability and shorter equipment life. Cost-effective Fault Detection and Diagnosis (FDD) methods can therefore ensure an increase in the system uptime, reliability, and overall efficiency.

In this thesis the FDD task has been approached as a binary classification problem. Support vector machines have been used to perform classification: they can deal with high dimensionality of the data and can provide non-linear decision boundaries in the classification model.

First, the most common relevant faults affecting VAVAC systems have been investigated, with the help of a simulation environment which allows to generate data that describe both regular and faulty system behaviour. A process history-based FDD method, which combines a supervised approach with a two-class SVMs system and a posterior probability has been employed. Extensive simulations have shown how the proposed FDD method correctly detects and diagnoses the different kinds of faults which plague VAVAC systems.

Then, a semi-supervised approach has been used to detect the most common chiller systems faults, under the assumption of unavailability of faulty labelled data. A novelty detection tool has been presented, being able to identify anomalous situations using only fault-free data in the training phase. The proposed combination of one-class classification (with one-class SVMs) and principal component analysis for discarding the high variability related to usual operating conditions changes has been shown to be effective in the detection performance.

Future works of the research can include the design of FDD methods, which could take advantage from an on-line version of the classification, being able to

detect the system faults under real time working conditions and comprehending adaptive solutions for automatic faults labelling. A interesting FDD research topic is the multi-faults detection and diagnosis for HVAC plants, i.e. the development of FDD techniques that could be able to understand multiple deviations of the data from normal conditions.

Final Remarks

Two different arguments have been covered in this thesis:

Energy Efficient Control of Ice Thermal Energy Storage Systems

The first part of this dissertation has corroborated the use of predictive control for the energy efficient management of ice Cold Thermal Energy Storage (ice-CTES) systems.

A model-based approach has been developed to design the ice-CTES plant. In detail, the storage has been modelled as a hybrid system, thus taking into consideration both sensible and latent heat. Three standard control strategies (constant-proportion control, chiller-priority control and storage-priority control) and a Non-Linear Model Predictive Control (NLMPC) approach have been compared in the management of the ice-CTES system. The NLMPC implementation needs the prediction of both outputs and disturbances of the plants. The disturbances of the plant have been obtained with an artificial neural network algorithm that can achieve sufficiently accurate predictions by measuring quantities that are typically available in standard HVAC installations, as explained in:

- A. Beghi, L. Cecchinato, M. Rampazzo, and F. Simmini. Load Forecasting for the Efficient Energy Management of HVAC Systems. *In IEEE International Conference on Sustainable Energy Technologies (ICSET)*, 2010.

Particle Swarm Optimisation (PSO) algorithm has been used to minimise the objective function associated with the predictive control problem. Nonlinear model predictive control has been shown to be the better control strategy for the efficient energy management of ice-CTES systems. The results of this research have been published in:

- A. Beghi, L. Cecchinato, M. Rampazzo, and F. Simmini. Modeling and Control of HVAC Systems with Ice Cold Thermal Energy Storage. In *IEEE 52nd Annual Conference on Decision and Control (CDC)*, 2013.

- A. Beghi, L. Cecchinato, M. Rampazzo, and F. Simmini. Energy efficient control of HVAC systems with ice cold thermal energy storage. *Journal of Process Control*, 24(6):773-781, 2014.

Fault Detection in HVAC Systems

In the second part of the dissertation fault detection methods for HVAC systems have been developed.

First a Fault Detection and Diagnosis (FDD) system has been designed to detect and diagnose the most common faults affecting Variable Air Volume Air Conditioning (VAVAC) plants. To this aim, a lumped-parameter model of a two-zone VAVAC system has been developed; a direct feedback linearisation technique has been adopted to control the resulting non-linear and coupled system. The FDD task has been approached as a supervised classification problem. The proposed supervised learning technique has exhibited good performance in the detection and diagnosis of the VAVAC model simulated faults.

Fault detection algorithms have been later developed to detect the most relevant faults plaguing chiller systems. Data from the research project 1043-RP of the American Society of Heating, Refrigerating and Air Conditioning Engineers (ASHRAE) have been used. The developed fault detection technique is based on one-class classification coupled with a novelty detection approach. The one-class classifier has shown to be effective in the fault detection of the chiller faults. Research results of the second part of the thesis have been published in:

- A. Beghi, L. Cecchinato, L. Corso, M. Rampazzo, and F. Simmini. Process History-Based Fault Detection and Diagnosis for VAVAC Systems. In *IEEE International Conference on Control Applications (CCA), Part of the IEEE Multi-Conference on Systems and Control (MSC)*, 2013.

- A. Beghi, L. Cecchinato, C. Corazzol, M. Rampazzo, F. Simmini, and G.A. Susto. A One-Class SVM Based Tool for Machine Learning Novelty Detection in HVAC Chiller Systems. In *19th IFAC World Congress (accepted)*, 2014.

Appendices

A

Non-Linear Model Predictive Control Strategy

Model Predictive Control (MPC) is an advanced method of process control originated in the late seventies. The term Model Predictive Control refers to a very wide range of control methods which make an explicit use of a model of the process to obtain the control signal by minimising an objective function. There are different MPC algorithms, that differ amongst themselves in the model used to represent the process, the noises and the cost function to be minimised. However, the ideas appearing in all the predictive control families are:

- use of a model to predict the process output at future time instants;
- minimisation of an objective function in order to calculate a sequence of future control inputs;
- the MPC is a receding control strategy, i.e. only the first control signal of the calculated sequence is applied to the system.

MPC control has been widely received by the academic world and by industry. There are many applications of predictive control successfully in use at the present time, not only in the process industry but also in different applications ranging from robot manipulators to biomedical systems. The good performance of these applications shows the capacity of the MPC to achieve highly efficient control systems able to operate during long periods of time.

A.1 Model Predictive Control Principles

The methodology of all the controllers belonging to the MPC family is characterized by the following principles:

- The future outputs for a determined horizon T_p, called the prediction horizon, are predicted at each instant τ, using the process model. These predicted outputs depend on the known values up to instant τ (past inputs and outputs) and on the future control signals, which are those to be calculated.
- The set of future control signals is calculated by optimising an objective function in order to keep the process as close as possible to the reference trajectory. The performance index usually takes the form of a quadratic function of the errors between the predicted output signal and the reference trajectory. The control effort is included in the objective function in most cases.
- Only the first control signal is sent to the process while the other control signals are discarded.

A.2 Mathematical Formulation of NLMPC

Actual systems often contain non-linear dynamics. When dealing with non-linear models, predictive control strategy is known as Non-Linear Model Predictive Control (NLMPC). Generally, a discrete non-linear system can be summarized as follows:

$$\mathbf{x}(\tau+1) = \mathbf{f}(\mathbf{x}(\tau), \mathbf{u}(\tau), \mathbf{d}(\tau)), \tag{A.1a}$$

$$\mathbf{y}(\tau) = \mathbf{h}(\mathbf{x}(\tau), \mathbf{u}(\tau)), \tag{A.1b}$$

$$\mathbf{x}(0) = \mathbf{x}_0, \tag{A.1c}$$

$$\tau \geq 0, \tag{A.1d}$$

$$\mathbf{y} \in \mathbb{Y} := [\mathbf{y}_{\min}, \mathbf{y}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{y})}, \tag{A.1e}$$

$$\mathbf{u} \in \mathbb{U} := [\mathbf{u}_{\min}, \mathbf{u}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{u})}, \tag{A.1f}$$

$$\mathbf{d} \in \mathbb{D} := [\mathbf{d}_{\min}, \mathbf{d}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{d})}, \tag{A.1g}$$

$$\mathbf{x} \in \mathbb{X} := [\mathbf{x}_{\min}, \mathbf{x}_{\max}] \subset \mathbb{R}^{\dim(\mathbf{x})}, \tag{A.1h}$$

where $\mathbf{x}, \mathbf{u}, \mathbf{d}, \mathbf{y}$, are the vectors of state, control variables, disturbances and output variables, respectively, $\mathbf{f}$ and $\mathbf{h}$ are general non-linear maps. The sets $\mathbb{Y}, \mathbb{U}, \mathbb{D}$, $\mathbb{X}$ are the feasible domains (system constraints).

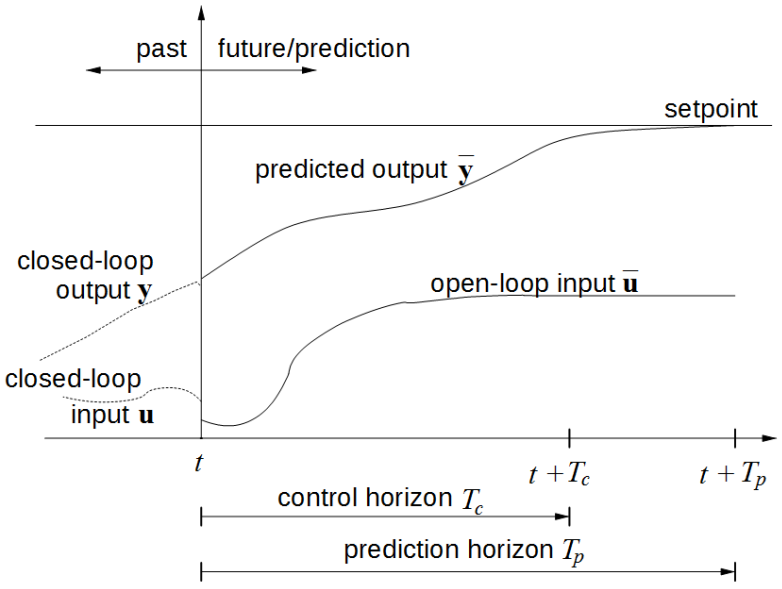


Figure A.1: Principle of predictive control.

Many formulations for the objective function can be found in MPC literature. One of the most used non-linear model predictive control cost function is written as follows:

$$\arg\min_{\bar{\mathbf{u}}} \sum_{\tau=t+1}^{t+T_p} \alpha(\tau) \left\|\bar{\mathbf{y}}(\tau) - \bar{\mathbf{r}}(\tau)\right\|^2 + \sum_{\tau=t+1}^{t+T_c} \beta(\tau) \left\|\bar{\mathbf{u}}(\tau)\right\|^2 + \gamma(\tau) \left\|\Delta\bar{\mathbf{u}}(\tau)\right\|^2 , \tag{A.2a}$$

$$\text{s.t.} \quad \bar{\mathbf{x}}(\tau+1) = \mathbf{f}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau), \bar{\mathbf{d}}(\tau)) , \tag{A.2b}$$

$$\bar{\mathbf{y}}(\tau) = \mathbf{h}(\bar{\mathbf{x}}(\tau), \bar{\mathbf{u}}(\tau)) , \tag{A.2c}$$

$$\bar{\mathbf{u}}(\tau) \in \mathbb{U} , \quad \forall\tau \in [t, t+T_c] , \tag{A.2d}$$

$$\bar{\mathbf{u}}(\tau) = \bar{\mathbf{u}}(t+T_c) , \forall\tau \in [t+T_c, t+T_p] , \tag{A.2e}$$

$$\bar{\mathbf{d}}(\tau) \in \mathbb{D} , \quad \forall\tau \in [t, t+T_p] , \tag{A.2f}$$

$$\bar{\mathbf{x}}(\tau) \in \mathbb{X} , \quad \forall\tau \in [t, t+T_p] , \tag{A.2g}$$

$$\bar{\mathbf{y}}(\tau) \in \mathbb{Y} , \quad \forall\tau \in [t, t+T_p] , \tag{A.2h}$$

where T_p and T_c are the prediction and the control horizons with $T_c \leq T_p$; $\alpha(\cdot)$, $\beta(\cdot)$ and $\gamma(\cdot)$ are weight coefficients; Δ is the difference operator; $\|\cdot\|$ is a norm chosen in such a way as to weigh the trajectory error (from the reference $\bar{\mathbf{r}}$) and the control effort. To distinguish the real system from the system model used to predict the future within the controller, in (A.2) a bar is used to denote

the internal variables in the controller.

The implementation concepts of the predictive control strategy are represented in Figure A.1. The system model is used to predict the future plant outputs, based on past and current information and on the proposed optimal future control actions, which are calculated by minimising the cost function taking into account the constraints. Only the first control input is applied to the plant. The procedure is repeated at each iteration. The cost function (A.2a) is generally non-convex, leading to a non-convex optimisation problem to be solved.

A in-depth analysis of NLMPC theory can be found in [12].

B

Artificial Neural Networks

Artificial Neural Networks (ANNs) are computational models inspired by the nervous systems (e.g. the brain). Artificial neural networks are generally presented as systems of interconnected "neurons"; they can be used to solve engineering problems of artificial intelligence as those that arise in different technological fields (in electronics, computer science, and other disciplines). In most cases, an artificial neural network is an adaptive system that changes its structure on external inputs/outputs and the information that flows through the network during the learning phase. In this way the neural network is able to understand complex relationships between inputs and outputs that other analytic functions fail to represent. An artificial neural network receives external signals on a layer of nodes, each of which is connected with many internal nodes, in a multiple levels organization. Each node processes the received signals and transmits the result to the subsequent nodes.

A neural network is composed by three kinds of layers:

- an input layer;
- hidden layers;
- an output layer.

There are different kinds of artificial neural networks. One of the most important is the feed-forward neural network. It has been shown that a feed-forward neural network with one hidden layer can approximate any function, and this is the most used scheme amongst ANNs in black-box identification, [13]. In Figure B.1 a general scheme for a feed-forward artificial neural network with one hidden layer is shown. Nodes represent variables while arches are

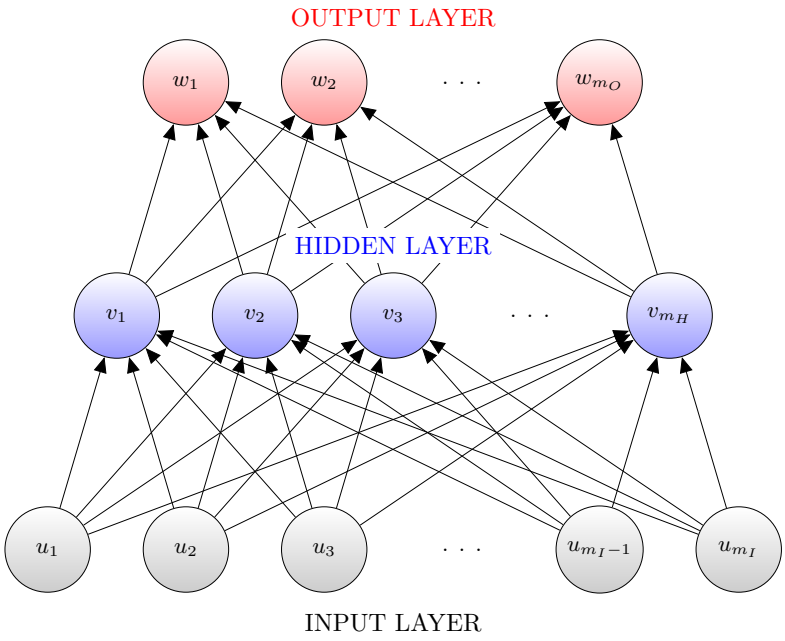


Figure B.1: Single hidden layer neural network.

associated to functions that describe interconnection between variables. In the scheme $\{u_i\}_{i=1}^{m_I}$ are the inputs and $\{w_i\}_{i=1}^{m_O}$ are the outputs. Features $\{v_i\}_{i=1}^{m_H}$ are created from combinations of the inputs:

$$v_i = h_a(\alpha_{0i} + \alpha_i^T U)\,, \quad i = 1, \ldots, m_H\,, \tag{B.1}$$

where U is the matrix of the inputs, while outputs are created from combinations of the created features:

$$w_i = h_b(\beta_{0i} + \beta_i^T V)\,, \quad i = 1, \ldots, m_O\,, \tag{B.2}$$

where V is the matrix of hidden features. The activation function $h_a(\cdot)$ is usually chosen to be non-linear (sigmoid, arctan, radial basis function), while the output function $h_b(\cdot)$ is typically chosen linear. For more details the reader is referred to [13].

Coefficients $\alpha.$ and $\beta.$ are called weights and are chosen in such a way as to minimise the Mean Squared Error (MSE):

$$\text{MSE} = \frac{\sum_{j=1}^{n}(\hat{\mathbf{w}}_j - \mathbf{w}_j)^2}{n}\,, \tag{B.3}$$

where, n is the number of observations, $\mathbf{w}$ is the real output vector and $\hat{\mathbf{w}}$ is the predicted one.

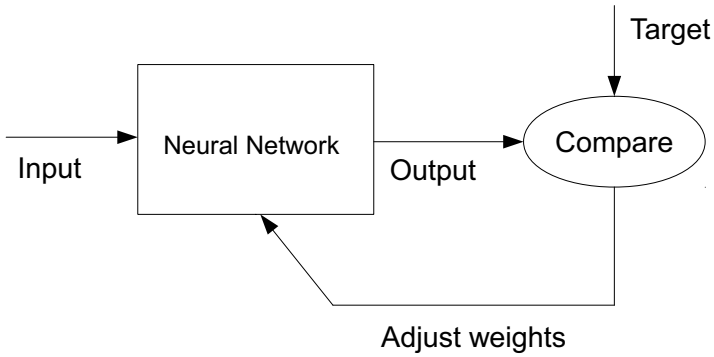


Figure B.2: ANN back-propagation algorithm.

The algorithm that is commonly employed for the training of the ANNs is the back-propagation algorithm, where weights are computed in a two-phase procedure where, after an initial guess, the prediction error is computed and then propagated backwards in the ANN structure to correct weights; then with the new weights the new prediction errors are calculated (Figure B.2). This procedure is iterated several times in order to reach the minimum value of (B.3). Further information about the ANNs tool can be found in [13].

C

Particle Swarm Optimisation Algorithm

Collective behaviours are a key concept in several studies of modern research in many disciplines, from econometric studies to social sciences, from computer science to control and optimisation theory. Particle Swarm Optimisation (PSO) is a population-based optimisation technique inspired by the motion of bird flocks (Figure C.1), [18].

The key idea in the PSO method is in the way to combine two different types of perspectives (individual and social perspectives) to predict the best update for the position in the next iteration [27], maintaining a balance between exploration and exploitation in the space. At each iteration, each particle combines information from the current motion vector, an individual correction, and a social correction. The contribution of these components results in the prediction of the position in the next iteration.

The basic algorithm implementation is as follows. Let $f : S \subseteq \mathbb{R}^n \to \mathbb{R}$ be the cost function to be minimised. The particles are manipulated according to the following equations related to velocity and position of the j-th particle:

$$\begin{aligned}\dot{\mathbf{p}}_j(t+1) = w\dot{\mathbf{p}}_j(t) &+ c_1\mathbf{R}_1(t)(\mathbf{p}_{j,\mathbf{P}}(t) - \mathbf{p}_j(t)) \\ &+ c_2\mathbf{R}_2(t)(\mathbf{p}_G(t) - \mathbf{p}_j(t))\,, \end{aligned} \tag{C.1}$$

$$\mathbf{p}_j(t+1) = \mathbf{p}_j(t) + \dot{\mathbf{p}}_j(t+1)\,. \tag{C.2}$$

The positions $\mathbf{p}_\mathbf{G}(t)$ and $\mathbf{p}_{j,\mathbf{P}}(t)$ represent the global and local best particles positions at iteration t respectively. w is an inertia weight, $c_1 > 0$ and $c_2 > 0$ are the cognitive and social parameters, respectively, and $\mathbf{R}_1$ and $\mathbf{R}_2$ are $n \times n$ diagonal matrices of random numbers uniformly distributed in $[0, 1]$. The

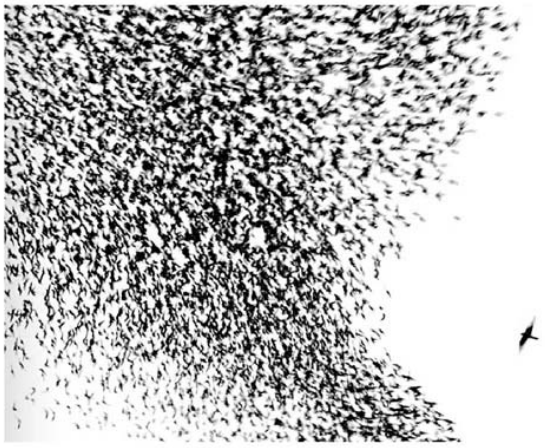


Figure C.1: Bird flock.

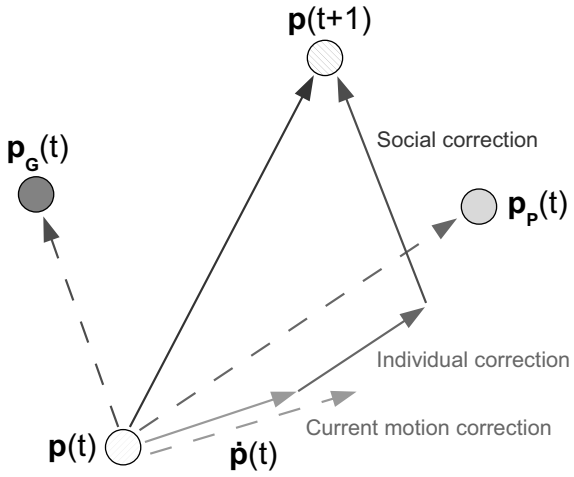


Figure C.2: PSO, graphical illustration of position update.

velocity is updated according to the following components, Figure C.2:

- previous velocity, $\dot{\mathbf{p}}(t)$: it represents the inertia term, that avoids sudden changes in the direction of the particle;
- individual correction, $\mathbf{p_P}(t) - \mathbf{p}(t)$: it quantifies the dynamic of the particle with respect to past performances. It takes care of the memory of the individual particle, and its desire to go back to its best position;
- social correction, $\mathbf{p_G}(t) - \mathbf{p}(t)$: it quantifies the dynamic of the particle with respect to the performance of the swarm; the particle tends to move towards the best position of the swarm.

The steps of the PSO algorithm can be summarised as follows:

1. Set values for swarm size, inertia weight, cognitive and social parameters. Initialize a population of particles with random positions and velocities. Initialize each particle's best position to its initial position. Determine the initial global best position among all particles.
2. Update the velocity and position of each particle according to equations (C.1) and (C.2).
3. For each particle, compute the value of $f(\mathbf{p}_j)$.
4. Compare each particle current value with the corresponding local best value. If the current value is better, then update the local best value.
5. Determine the particle of the current swarm with the best objective function value. If such value is better than $\mathbf{p}_G$, then update $\mathbf{p}_G$.

The algorithm ends when a stopping criterion is met, for example if the objective function no longer improves or after a certain number of iterations. $\mathbf{p}_G$ is the solution provided by the algorithm.

C.1 Constrained Optimisation in PSO

A constrained optimisation problem in particle swarm optimisation algorithm can be written in the following form:

$$\text{find} \qquad \arg\min_{\mathbf{p}\in S} f(\mathbf{p})\,, \tag{C.3a}$$

$$\text{subject to} \qquad g_i(\mathbf{p}) \leq 0\,, \quad i = 1, \dots q\,, \tag{C.3b}$$

$$h_j(\mathbf{p}) = 0\,, \quad j = q+1, \dots m\,, \tag{C.3c}$$

where $\mathbf{p}$ is the position of the particle and S is the whole search space. There are q inequalities and $m - q$ equality constraints.

Different approaches can be adopted to manage constraints in PSO optimisation. In order to decide whether a particle is in the feasible region, a penalty function can be introduced, for example:

$$\phi(\mathbf{p}) = \prod_i e^{\max(0, g_i(\mathbf{p}))} \prod_j e^{\max(0, |h_j(\mathbf{p})|)}\,. \tag{C.4}$$

It is worth noticing that $\phi(\mathbf{p}) \geq 1$, and all the positions such that $\phi(\mathbf{p}) = 1$ are feasible. By minimising $\phi(\mathbf{p})$, the particle tends to move towards the feasible region. When the particle moves away from the feasible region, assuming that the previous particle position is in the feasible region, one possible strategy (as explained in [3]) is to let the particle move towards a position between the previous position and the just calculated position minimising the objective function f:

$$\bar{\mathbf{p}}(t+1) = \mathbf{p}(t) + \beta(\mathbf{p}(t+1) - \mathbf{p}(t))\,. \tag{C.5}$$

The new position can be found by solving the following optimisation problem:

$$\arg\min_{\beta\in[0,1)} \phi(\bar{\mathbf{p}}(t+1))\,. \tag{C.6}$$

Since the values of $\mathbf{p}(t)$ and $\mathbf{p}(t+1)$ are known, a value of β is selected to get the minimum of the function ϕ. Since β is only a scalar coefficient, the problem can be dealt as an one-dimensional search optimisation problem and can be easily solved by means of, e.g., golden section search methods [6].

D

The Support Vector Classifier

The mathematical theory for the support vector classifier is briefly described in the following. The reader is referred to [13] for a in-depth analysis of the support vector classification problem.

D.1 Separating Hyperplanes

The training data for the classifier are n pairs $(\mathbf{x}_1, \mu_1), ..., (\mathbf{x}_n, \mu_n)$, where $\mathbf{x}_i \in \mathbb{R}^d$ and $\mu_i \in \{-1, 1\}$. Training data can thus be divided into two classes with respect to the value of μ.

Defining the following hyperplane:

$$\{\mathbf{x} : f(\mathbf{x}) = \mathbf{x}^T\mathbf{w} + b = 0\}\,, \tag{D.1}$$

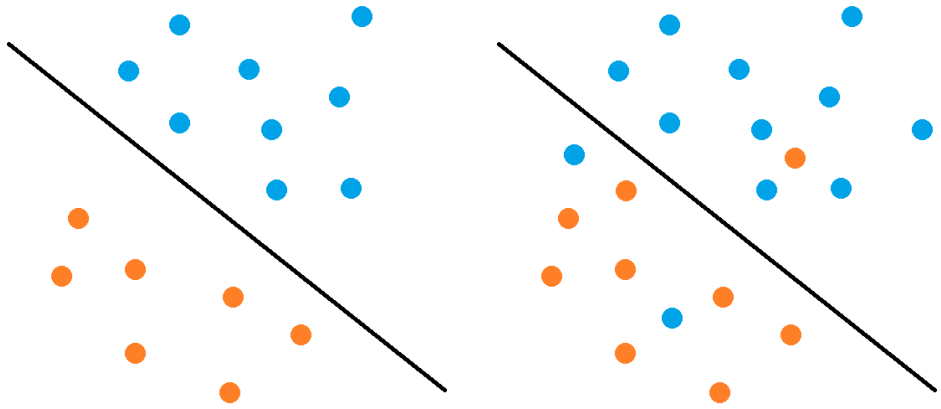
where $\mathbf{w}$ is a unit vector ($||\mathbf{w}|| = 1$), a possible induced classification rule can be:

$$G(\mathbf{x}) = \text{sign}[\mathbf{x}^T\mathbf{w} + b]\,. \tag{D.2}$$

From geometry we know that, given the point $\mathbf{x}$, $f(\mathbf{x})$ provides the signed distance of $\mathbf{x}$ from the hyperplane.

From the previous observation we can deduce that if the two classes are separable (see Figure D.1(a)) it is possible to find a function $f(\mathbf{x}) = \mathbf{x}^T\mathbf{w} + b$ with $\mu_i f(\mathbf{x}_i) > 0\ \forall i$. Therefore the hyperplane that creates the biggest margin M between the training points for class 1 and -1 can be found; the optimisation problem for the linearly separable case can thus be written:

$$\max_{\mathbf{w},b,||\mathbf{w}||=1} M\,, \tag{D.3a}$$



(a) Separable case. (b) Non-Separable case.

Figure D.1: Support vector classifiers.

subject to:

$$\mu_i(\mathbf{x}_i^T\mathbf{w} + b) \geq M\,, \quad i = 1, \ldots, n\,. \tag{D.3b}$$

If we set $M = 1/||\mathbf{w}||$ the previous optimisation problem can be easily written in the equivalent form [13]:

$$\min_{\mathbf{w},b} ||\mathbf{w}||\,, \tag{D.4a}$$

subject to:

$$\mu_i(\mathbf{x}_i^T\mathbf{w} + b) \geq 1\,, \quad i = 1, \ldots, n\,, \tag{D.4b}$$

which is the standard way of writing the support vector problem for separable data.

We suppose now that data are not linearly separable. In this case M is still maximized, but some points are on the other side of the margin (Figure D.1(b)). The slack variables $(\xi_1, \xi_2, \ldots \xi_n)$ are defined and the constraints in (D.3) are modified in the following way [13]:

$$\mu_i(\mathbf{x}_i^T\mathbf{w} + b) \geq M(1 - \xi_i)\,, \tag{D.5}$$

$\forall i,\ \xi_i \geq 0,\ \sum_{i=1}^n \xi_i \leq$ constant. As in the previous case this problem can be more conveniently written as follows:

$$\min_{\mathbf{w},b} ||\mathbf{w}||\,, \tag{D.6a}$$

subject to:

$$\mu_i(\mathbf{x}_i^T\mathbf{w} + b) \geq 1 - \xi_i \quad \forall i\,, \tag{D.6b}$$

$$\xi_i \geq 0\,, \sum \xi_i \leq \text{constant}\,. \tag{D.6c}$$

D.2 Optimisation

The problem (D.6) is a convex optimisation task, since it is quadratic with linear inequality constraints. A quadratic programming solution is shown using Lagrange multipliers [13]. To this aim, equation (D.6) can be re-expressed in the equivalent form adding the cost parameter C:

$$\min_{\mathbf{w},b} \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n} \xi_i\,, \tag{D.7a}$$

subject to:

$$\xi_i \geq 0\,, \quad \mu_i(\mathbf{x}_i^T\mathbf{w} + b) \geq 1 - \xi_i \quad \forall i\,. \tag{D.7b}$$

The Lagrange (primal) function of the previous optimisation problem is:

$$L_P = \frac{1}{2}||\mathbf{w}||^2 + C\sum_{i=1}^{n} \xi_i - \sum_{i=1}^{n} \alpha_i[\mu_i(\mathbf{x}_i^T\mathbf{w} + b) - (1 - \xi_i)] - \sum_{i=1}^{n} \gamma_i\xi_i\,, \tag{D.8}$$

which has to be minimised with respect to $\mathbf{w}, b$ and ξ_i. If the respective derivatives are set to zero, the following equations are obtained:

$$\mathbf{w} = \sum_{i=1}^{n} \alpha_i\mu_i\mathbf{x}_i\,, \tag{D.9}$$

$$0 = \sum_{i=1}^{n} \alpha_i\mu_i\,, \tag{D.10}$$

$$\alpha_i = C - \gamma_i\,, \quad \forall i\,, \tag{D.11}$$

with the positivity constraints α_i, γ_i, $\xi_i \geq 0$, $\forall i$. By substituting equations (D.9)-(D.11) into (D.8), the Lagrangian dual objective function is found:

$$L_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2}\sum_{i=1}^{n}\sum_{i'=1}^{n} \alpha_i\alpha_{i'}\mu_i\mu_{i'}\mathbf{x}_i^T\mathbf{x}_{i'}\,, \tag{D.12}$$

giving a lower bound on the objective function (D.7a) for the feasible set. The dual function L_D has to be maximized subject to $0 \leq \alpha_i \leq C$ and $\sum_{i=1}^{n} \alpha_i\mu_i =$

0. In addition to (D.9)-(D.11) the Karush-Kuhn-Tucker conditions comprise the constraints:

$$\alpha_i[\mu_i(\mathbf{x}_i^T\mathbf{w}+b)-(1-\xi_i)]=0\,, \tag{D.13}$$

$$\gamma_i\xi_i=0\,, \tag{D.14}$$

$$\mu_i(\mathbf{x}_i^T\mathbf{w}+b)-(1-\xi_i)\geq 0\,, \tag{D.15}$$

for $i=1,\ldots,n$. The equations (D.9)-(D.15) uniquely define the solution to the primal and dual problem. From (D.9) we see that the solution for $\mathbf{w}$ has the form:

$$\hat{\mathbf{w}}=\sum_{i=1}^{n}\hat{\alpha}_i\mu_i\mathbf{x}_i\,, \tag{D.16}$$

with nonzero coefficients $\hat{\alpha}_i$ only for the observations i for which the constraints in (D.15) are exactly satisfied (due to (D.13)). Therefore, only these observations, called the support vectors, determine the solution. From (D.13) b can be found with any of the margin points (which are characterized by $0<\hat{\alpha}_i$, $\hat{\xi}_i=0$, [13]), and an average of all the solutions is usually used. The dual (D.12) maximisation is a simpler convex quadratic programming problem with respect to primal problem (D.8), and can be solved with standard optimisation techniques.

Once obtained the solutions $\hat{b}$ and $\hat{\mathbf{w}}$, the classification rule can be finally written as:

$$\hat{G}(\mathbf{x})=\text{sign}[\hat{f}(\mathbf{x})]=\text{sign}[\mathbf{x}^T\hat{\mathbf{w}}+\hat{b}]\,. \tag{D.17}$$

D.3 Kernels

The described support vector classifier can find linear boundaries in the input feature space. However it is possible to make the procedure more flexible: the idea is to transform the vector input space into a higher dimensional feature space using new basis functions $\phi_m(\mathbf{x})$, $m=1,\ldots,M$. The support vector classifier is then computed using $\boldsymbol{\phi}(\mathbf{x}_i)=(\phi_1(\mathbf{x}_i),\phi_2(\mathbf{x}_i),\ldots,\phi_M(\mathbf{x}_i))$, $i=1,\ldots,n$ as input features and produce the (non-linear) function $\hat{f}(\mathbf{x})=\boldsymbol{\phi}(\mathbf{x})^T\hat{\mathbf{w}}+\hat{b}$. Finally the classifier is $\hat{G}(\mathbf{x})=\text{sign}[\hat{f}(\mathbf{x})]$.

In detail, it is possible to represent the optimisation problem (D.8) and its solution involving the input features only via inner products. Using a basis expansion the Lagrange dual function is:

$$L_D = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{i'=1}^{n} \alpha_i \alpha_{i'} \mu_i \mu_{i'} \langle \boldsymbol{\phi}(\mathbf{x}_i), \boldsymbol{\phi}(\mathbf{x}_{i'}) \rangle \ . \tag{D.18}$$

From (D.9), the solution $f(\mathbf{x})$ can be written as:

$$f(\mathbf{x}) = \boldsymbol{\phi}(\mathbf{x})^T \mathbf{w} + b = \sum_{i=1}^{n} \alpha_i \mu_i \langle \boldsymbol{\phi}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{x}_i) \rangle + b \, . \tag{D.19}$$

Given α_i, b can be determined as in the previous case.

From (D.18) and (D.19) we see that the transformation $\boldsymbol{\phi}(\mathbf{x})$ is not needed to be specified, but only knowledge of the kernel function is required:

$$k(\mathbf{x}, \mathbf{x}') = \left\langle \boldsymbol{\phi}(\mathbf{x}), \boldsymbol{\phi}(\mathbf{x}') \right\rangle \ , \tag{D.20}$$

that computes inner products in the enlarged space. k is a symmetric positive (semi-) definite function [13]. Common choices for the inner product k are:

- p-th degree polynomial: $k(\mathbf{x}, \mathbf{x}') = (1 + \langle \mathbf{x}, \mathbf{x}' \rangle)^p$,
- Radial basis: $k(\mathbf{x}, \mathbf{x}') = \exp[- \left\| \mathbf{x} - \mathbf{x}' \right\|^2 / (2\sigma^2)]$,
- Neural network: $k(\mathbf{x}, \mathbf{x}') = \tanh(a_1 \langle \mathbf{x}, \mathbf{x}' \rangle) + a_2)$.

References

[1] H. Akbari and O. Sezgen. Case Studies of Thermal Energy Storage (TES) Systems: Evaluation and Verification of System Performance. Technical report, Lawrence Berkeley Laboratory, University of California (United States), 1992.

[2] M. Albieri, A. Beghi, C. Bodo, and L. Cecchinato. Advanced control systems for single compressor chiller units. *International Journal of Refrigeration*, 32(5):1068 –1076, 2009.

[3] A. Beghi, L. Cecchinato, G. Cosi, and M. Rampazzo. A PSO-based algorithm for optimal multiple chiller systems operation. *Applied Thermal Energy*, 32:31 – 40, 2012.

[4] A. Beghi, L. Cecchinato, and M. Rampazzo. A multi-phase genetic algorithm for the efficient management of multi-chiller systems. *Energy Conversion and Management*, 52(3):1650 – 1661, 2011.

[5] A. Beghi, L. Cecchinato, M. Rampazzo, and F. Simmini. Load Forecasting for the Efficient Energy Management of HVAC Systems. In *IEEE International Conference on Sustainable Energy Technologies (ICSET)*, 2010.

[6] R.P. Brent. *Algorithms for Minimization without Derivatives*. Prentice-Hall, 1973.

[7] R. Cagienard, P. Grieder, E.C. Kerrigan, and M. Morari. Move blocking strategies in receding horizon control. *Journal of Process Control*, 17(6):563 – 570, 2007.

[8] M.C. Comstock and J.E. Braun. Development of analysis tools for the evaluation of fault detection and diagnostics for chillers. Technical report, ASHRAE Research Project 1043-RP, HL 99-20 Report 4036-3, 1999.

[9] I. Dincer, S. Dost, and X. Li. Performance analyses of sensible heat storage systems for thermal applications. *International Journal of Energy Research*, 21(12):1157–1171, 1997.

[10] I. Dincer and M.A. Rosen. *Thermal Energy Storage: Systems and Applications*. John Wiley & Sons, Ltd, West Sussex, United Kingdom, 2011.

[11] L. Gao, L. Chen, Y. Fan, and H. Ma. A Nonlinear Control Design for Power Systems. *Automatica*, 28(5):975–979, 1992.

[12] L. Grüne and J. Pannek. *Nonlinear Model Predictive Control*. Springer, 2011.

[13] T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer New York, 2009.

[14] G.P. Henze, C. Felsmann, and G. Knabe. Evaluation of optimal control for active and passive building thermal storage. *International Journal of Thermal Sciences*, 43(2):173 –183, 2004.

[15] G.P. Henze, M. Krarti, and M.J. Brandemuehl. A Simulation Environment for the Analysis of Ice Storage Controls. *HVAC&R Research*, 3(2):128–148, 1997.

[16] G.P. Henze, M. Krarti, and M.J. Brandemuehl. Guidelines for improved performance of ice storage systems. *Energy and Buildings*, 35(2):111–127, 2003.

[17] W. Huangang, Z. Lin, X. Yingchao, and X. Wenli. An Approach to Choosing Gaussian Kernel Parameter for One-Class SVMs via Tightness Detecting. In *4th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*, volume 2, pages 318–323, 2012.

[18] J. Kennedy and R. Eberhart. Particle Swarm Optimization. In *IEEE International Conference on Neural Networks*, volume 4, pages 1942–1948, 1995.

[19] W.-Y. Lee, J.M. House, and N.-H. Kyong. Subsystem level fault diagnosis of a building's air-handling unit using general regression neural networks. *Applied Energy*, 77(2):153–170, 2004.

[20] S. Li. *A Model-Based Fault Detection and Diagnostic Methodology for Secondary HVAC Systems*. PhD thesis, Drexel University, 2009.

[21] J. Liang and R. Du. Model-based Fault Detection and Diagnosis of HVAC systems using Support Vector Machine method. *International Journal of Refrigeration*, 30(6):1104–1114, 2007.

[22] H.-T. Lin, C.-J. Lin, and R.C. Weng. A Note on Platt's Probabilistic Outputs for Support Vector Machines. *Machine Learning*, 68(3):267–276, 2007.

[23] Y. Ma, F. Borrelli, B. Hencey, B. Coffey, S. Bengea, and P. Haves. Model Predictive Control for the Operation of Building Cooling Systems. *IEEE Transactions on Control Systems Technology*, 20(3):796–803, 2012.

[24] Y. Ma, A. Kelman, A. Daly, and F. Borrelli. Predictive Control for Energy Efficient Buildings with Thermal Storage: Modeling, Simulation, and Experiments. *IEEE Control Systems Magazine*, 32(1):44 – 64, 2012.

[25] M. Markou and S. Singh. Novelty detection: a review - part 1: statistical approaches. *Signal Processing*, 83(12):2481–2497, 2003.

[26] I.B.D. McIntosh, J.W. Mitchell, and W.A. Beckman. Fault detection and diagnosis in chillers–Part 1: Model development and application. *ASHRAE Transactions*, 106(2):268–282, 2000.

[27] S.M. Mikki and A.A. Kishk. *Particle Swarm Optimization: A Physics-Based Approach*. Morgan and Claypool, Arizona State University, 2008.

[28] J.C. Platt. Probabilistic Outputs for Support Vector Machines and Comparisons to Regularized Likelihood Methods. In *Advances in Large Margin Classifiers*, pages 61–74. MIT Press, 1999.

[29] M. Rampazzo. *Efficient Management of HVAC Systems*. PhD thesis, Ph.D. School in Information Engineering, University of Padova, 2010.

[30] M.A. Rosen, I. Dincer, and N. Pedinelli. Thermodynamic Performance of Ice Thermal Energy Storage Systems. *ASME-Journal of Energy Resources Technology*, 122(4):205–211, 2000.

[31] B. Schölkopf, J.C. Platt, J.C. Shawe-Taylor, A.J. Smola, and R.C. Williamson. Estimating the Support of a High-Dimensional Distribution. *Neural Computation*, 13(7):1443–1471, 2001.

[32] B.R. Sorensen. Modelling and simulation of a HVAC system. Technical report, Narvik Intitute of Technology, 1999.

[33] G.A. Susto, A. Beghi, and C. De Luca. A Predictive Maintenance System for Epitaxy Processes Based on Filtering and Prediction Techniques. *IEEE Transactions on Semiconductor Manufacturing*, 25(4):638–649, 2012.

[34] N. Tudoroiu and M. Zaheeruddin. Fault detection and diagnosis of valve actuators in HVAC systems. In *Proceedings of IEEE Conference on Control Applications*, 2005.

[35] N. Tudoroiu, M. Zaheeruddin, E.-R. Tudoroiu, and V. Jeflea. Fault Detection and Diagnosis (FDD) in Heating Ventilation Air Conditioning Systems (HVAC) Using an Interactive Multiple Model Augmented Unscented Kalman Filter (IMMAUKF). In *Conference on Human System Interactions*, 2008.

[36] J. Wang, Y. Wang, and H. Shao. Performance improvement of VAV air conditioning control system through diagonal matrix decoupling and Lonworks technology. *Energy and Buildings*, 37(9):911–919, 2005.

[37] L. Xuquan, S. Zhigang, and H. Songtao. A Novel Control Method of a Variable Volume Air Conditiong System for Indoor Thermal Environment. In *2nd International Conference on Computer Engineering and Technology (ICCET)*, volume 2, 2010.

[38] H. Yoshida, S. Kumar, and Y. Morita. Online fault detection and diagnosis in VAV air handling unit by RARX modeling. *Energy and Buildings*, 33(4):391–401, 2001.