Development of a CBM based service indicator for UFD replacements

- An introductory study

Victor Lundqvist & Johan Åkesson

Master's Thesis in
Biomedical Engineering
2017

LUND
UNIVERSITY

Faculty of Engineering LTH
Department of Biomedical Engineering
Supervisor Baxter: Erik Wallenborg
Supervisor LTH: Frida Sandberg

Baxter International Inc.

Abstract

Dialysis is a life-sustaining treatment for many people around the world. In order to meet the high demands on the dialysis machines, replaceable parts must be exchanged in proper time. Condition based maintenance (CBM) bases its service decisions on the actual health status of the component. The goal of this master's thesis was to develop an algorithm based on machine learning, constituting the first step of a CBM based service indicator monitoring the dialysis ultra filter (UFD) of Baxter's AK 98 dialysis machine.

Real treatment data retrieved from ten dialysis machines have been analyzed. Signals believed to be relevant to the UFD status were preprocessed and analyzed. From them different features were extracted whereof some were found in CBM related literature. Two different feature selection methods were used to select 10 out of the 178 available features. Different labeling methods were tested and evaluated together with other relevant parameters in the algorithm.

The final algorithm used a k-nearest neighbors (kNN) classifier with $k = 12$. The classification accuracy was approximately twice as good as a random guess. The main reason for not achieving a better result was that only six features appeared to contain relevant information regarding the UFD status. Furthermore, these features were derived from the same signal and closely related. Despite this, the developed algorithm did show promising result in detecting the UFD degradation level but further development will be needed. The main focus should be to improve the signal quality and/or find more relevant signals and/or features.

Preface

This is a thesis for the degree Master of Science in Biomedical Engineering at the Faculty of Engineering (LTH) of Lund University. The project was carried out at Baxter International Inc., Lund in the spring and summer of 2017. First of all we would like to direct a huge thanks to our supervisor at Baxter, Erik Wallenborg, for your great support and endless stream of ideas and encouragement. We also would like to give a big thanks to our manager at Baxter, Ivan Fulöp, for enabling this project and making sure we were given all the help and support we needed. Furthermore we would like to thank our supervisor at LTH, Frida Sandberg, for your theoretical expertise and help with this report. Finally we would like to thank all other people at Baxter who were involved or showed interest in our project.

Contents

List of Abbreviations

$\boldsymbol{\mu}$	Mean
$\boldsymbol{\sigma}$	Standard deviation
$\boldsymbol{c_{1-10}}$	Sum of the 10 largest Fourier coefficients
$\boldsymbol{c_k}$	Fourier coefficient
$\boldsymbol{c_k IQR}$	Interquartile range of Fourier coefficients
$\boldsymbol{c_k Q_1}$	Lower quartile of Fourier coefficients
$\boldsymbol{c_k Q_2}$	Median of Fourier coefficients
$\boldsymbol{c_k Q_3}$	Upper quartile of Fourier coefficients
$\boldsymbol{L_1}$	Labeling based on time
$\boldsymbol{L_2}$	Labeling based on hypochlorite disinfections
$\boldsymbol{Q_1}$	Lower quartile
$\boldsymbol{Q_2}$	Median
$\boldsymbol{Q_3}$	Upper quartile
BYVA	Bypass valve
CBM	Condition based maintenance
CCM	Maximum value from cross-correlation
CCTD	Cross-correlation time delay
CI	Crest indicator
CKD	Chronic kidney disease
CLI	Clearance indicator

Cond. cell	Conductivity cell
DF	Dominant frequency
DIVA	Direct valve
DRVA	Degass restrictor valve
EPO	Erythropoietin
FFT	Fast Fourier transform
FIVA	Filter valve
GFR	Glomerular filtration rate
HPG	High pressure guard
II	Impulse indicator
IQR	Interquartile range
kNN	k-nearest neighbors
KU	Kurtosis
Max	Maximum value
Min	Minimum value
PD	Pressure dialysis transducer
RMS	Root mean square
RO	Reverse osmosis
RUL	Remaining useful life
SFS	Sequential forward selection
SI	Shape indicator
SK	Skewness
TAVA	Taration valve
TFSWT	Two-stage feature selection and weighting technique
UFD	Dialysis ultra filter
ZEVA	Zeroing valve

List of Figures

List of Tables

1

Introduction

End-stage kidney disease is a serious condition requiring renal replacement therapy either as dialysis treatment or kidney transplant if not to be fatal. As of 2010, 2 million people worldwide required dialysis treatment to stay alive [1]. However, this number likely represented less than 10 % of those who were actually in need of it [2].

In hemodialysis a dialysis machine is used to replace the kidneys' function to remove waste products and excess fluid from the blood. Due to the critical health status of the patients in combination with the fact that the treatment is a time consuming activity, the demands on the dialysis treatment with respect to quality and efficiency are high.

Since a dialysis machine at a clinic is used by multiple patients it is of high importance that the system is carefully disinfected between treatments to avoid contagion. These disinfection programs in combination with the frequent use cause a natural wear on the components. In order to meet the high demands of quality and efficiency on the treatments these components must be replaced in appropriate time.

Condition based maintenance (CBM) is an approach that has received an increased amount of interest in the scientific research area of maintenance during the latest years [3]. It is a preventative approach which bases its service decisions on the actual health status of the component. This can be compared to traditional preventative maintenance which bases its service decisions solely on time or production hours since last service [4]. One desirable measure to retrieve from a CBM algorithm is the remaining useful life (RUL) of the component of interest.

1.1 Aim

According to the current guidelines [5], the replacement of the *dialysis ultrafilter* (UFD) in Baxter's dialysis machine AK 98 should be conducted when one of the following three limits have been exceeded:

- 90 days since last UFD replacement
- 150 heat disinfections
- 12 hypochlorite disinfections

Thereby the actual status of the filter is not taken into consideration. This means that the remaining capacity of the UFD is practically unknown when the filter is replaced. As a consequence of this the UFD may be replaced too early, resulting in increased costs and environmental impact. A method indicating a suitable time to exchange the UFD based on its current status, or RUL, is thereby desirable.

In order to develop a CBM based service indicator for UFD replacements three different steps are necessary to perform. The first step is to develop an algorithm which is able to detect different degradation levels of the UFD and decide the current level. The second step is to perform a long term study of how the UFD is degraded when it is not replaced. The third step is to analyze how well the UFD fulfills its intended purpose during different degradation levels. This analyze should find the level where the intended purpose is no longer fulfilled and from this result map each degradation level of the UFD to a specific RUL.

The data being logged in an AK 98 dialysis machine regarding pump velocities, pressures, temperatures, power usage etcetera is currently almost exclusively used for troubleshooting. The aim of this master's thesis is to perform the first step in the development of the CBM based UFD service indicator mentioned above, based on data analysis of these log files. To achieve this, a machine learning approach will be tested and evaluated.

1.2 Disposition

This report starts off with a basic description of relevant topics in Chapter 2. This includes a section about the physiology and pathology of the kidneys, a general explanation of the hemodialysis treatment and the dialysis machine followed by a more specific description of Baxter's AK 98 dialysis machine which is analyzed in this report. It also includes a section about CBM and previous research in this area as well as a short description of machine learning.

The sensor signals from the AK 98 log files analyzed in this report are described in Chapter 3. Chapter 4 explains how these signals were preprocessed, how features were extracted and selected from these log files and finally how the classification was done. The results from these different steps are presented in Chapter 5, followed by a discussion of the results and possible future work in Chapter 6. Finally, the conclusions from the report are summarized in Chapter 7.

2

Background

2.1 Physiology and pathology of the kidneys

The kidneys are approximately 11 centimeters long, bean-shaped organs located on each side of the vertebral column in the cavity behind the peritoneum. Despite the small size they receive about $20 - 25$ % of the cardiac output which is needed if they should be able to perform their physiological function [6]. These functions can be subdivided into three distinct categories; excretory, regulatory and endocrine function. Briefly, the excretory function is to filter the blood by removing metabolic waste products and toxins and excrete it through the urine. The regulatory function includes the regulation of body fluid, electrolyte balance and acid-base balance. Finally, the endocrine function is made up of the production and activation of different hormones involved in the formation of red blood cells, calcium metabolism and regulation of blood pressure and blood flow.

The dialysis machine tries to mimic some of the functions of the kidney and a basic understanding of the physiology of the kidney is necessary to understand *what* the dialysis machine is intended to do, *why* it is important and finally *how* the dialysis machine works.

2.1.1 Formation of ultrafiltrate

The functional unit of the kidney is the nephron, consisting of a glomerulus and a tubule, see Figure 2.1. The glomerulus is a cluster of small blood vessels surrounded by a capsule called Bowman's capsule and the space in between these two structures is called the Bowman's space. This structure is responsible for the formation of the ultrafiltrate, the first step of urine production. The filtration process is basically the same as in other capillaries in the body, the differences are the greater Starling forces, and most dominantly, higher capillary permeability present in the glomeruli [7]. The Starling forces affecting the filtration rate are the hydrostatic and oncotic pressures present inside as well as outside the glomerular capillaries. The oncotic (or colloid osmotic) pressure is an osmotic pressure exerted by proteins that are not able to pass through a blood vessel wall.

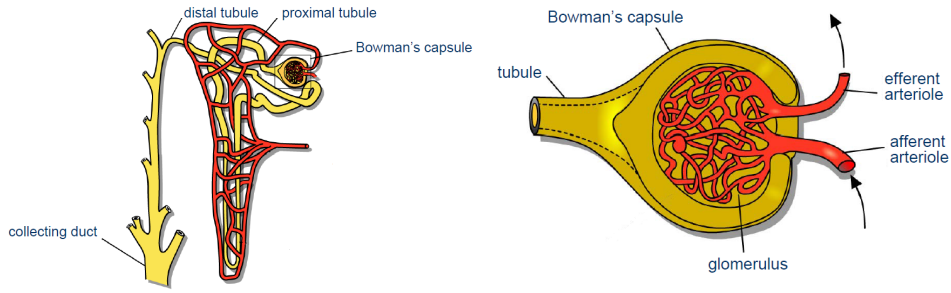


Figure 2.1: A nephron, the functional unit of the kidney to the left and a magnification of Bowman's capsule with the glomerulus to the right. Courtesy of [8].

Out of the four present Starling forces two of them are favoring ultrafiltration and the other two are opposing it. The two forces favoring ultrafiltration is the hydrostatic pressure inside the glomerular capillary (P_{GC}) and the oncotic pressure in the Bowman's space (π_{BS}). The latter is however very small and is sometimes omitted when calculating the glomerular filtration rate (GFR). The two corresponding forces, namely the hydrostatic pressure of the Bowman's space (P_{BS}) and the oncotic pressure of the glomerular capillary (π_{GC}) are the two forces opposing ultrafiltration.

Apart from the net Starling forces (P_{UF}) the ultrafiltration coefficient (K_f) is the other component affecting GFR. The coefficient describes the hydraulic conductance of the glomerular capillary wall and depends on the water permeability per surface area and the total surface area of the glomerular capillary walls. The equation for calculating GFR is given in Equation (2.1a).

$$GFR = K_f \times P_{UF} \tag{2.1a}$$

Where

$$P_{UF} = (P_{GC} - P_{BS}) - (\pi_{GC} - \pi_{BS}) \tag{2.1b}$$

Normal values for P_{GC} and P_{BS} are around 45 and 10 mmHg respectively and π_{GC} is typically around $25 - 40$ mmHg. The value of π_{GC} is increasing along the length of the capillary due to the water loss from ultrafiltration. As previously mentioned π_{BS} is close to zero. The net Starling forces are in the same magnitude as in other capillary beds whereas K_f is more than 100 times larger than in the rest of the systemic capillaries [9]. This explains the huge amount of ultrafiltrate produced in the kidneys each day, 180 liters. Since the human body only excretes around two liters of urine a day, a majority of the fluid is reabsorbed and the composition of the remaining ultrafiltrate is modified in the rest of the nephron [6, 7].

2.1.2 Reabsorption and secretion in the nephron

The ultrafiltrate is collected in the Bowman's space which is contiguous with the tubule, a long tube with different segments, each with its own specific function. The function of the tubule is to modify the composition of the ultrafiltrate through reabsorption and secretion of water and solutes. This occurs through both active and passive transport. The tubules from different nephrons merge to a collecting duct where the last modifications are done, producing final urine. This is then transported through the renal pelvis and ureters to the bladder where it is stored. Thereby no modification of either volume or composition is done in these parts.

The vast majority of the 180 liters ultrafiltrated water is reabsorbed in the tubule. Since sodium is the most abundant electrolyte in the extracellular fluid, the same is true for the blood and thereby the ultrafiltrate. In fact, the amount of sodium filtered by the kidneys each day is equivalent to approximately 1.5 kg of table salt [10]. This demands a very effective reabsorption in the tubule, about 99.6 % of the sodium is reabsorbed. Chloride represents another electrolyte with high concentration in the ultrafiltrate that is reabsorbed effectively. Other solvents that are reabsorbed include phosphate, calcium, glucose, amino acids and proteins (although only a small fraction is filtered in the first place) [11].

Other substances must instead be secreted. The most important example is potassium, which only appears in small concentrations in the blood and thereby in the ultrafiltrate since it is mainly stored intracellular. To compensate for the daily intake, potassium must be actively secreted. Other examples of substances being secreted are small anions e.g. p-aminohippurate [12].

2.1.3 Endocrine function

The kidneys are active in the production and regulation of different hormones with diverse physiological effects. Renin is released into the bloodstream as a response to decreased blood pressure in the afferent arterioles (preglomerular arterioles) which through a complex signaling pathway have the final effects of increased blood volume and blood pressure. Erythropoietin, or EPO, is released by cells in the cortex and outer medulla of the kidney in response to a fall in local tissue P_{O2}, stimulating the production of red blood cells in bone marrow. Cells of the proximal tubule convert inactive vitamin D to its active form which in turn regulates calcium and phosphorus metabolism. In addition to these hormones the kidneys also secrete prostaglandins and various kinins that regulate local circulation within the kidneys [6].

2.1.4 Kidney failure and treatment options

A person can usually sustain a normal life with just one functional kidney. The total kidney function measured as GFR may be as low as 25 % before clear symptoms of renal failure appear [13].

Kidney failure can be of both acute and chronic nature. The most common disease leading to acute kidney failure is acute tubular necrosis which depends on the death of epithelial cells of the tubule. This in turn can be caused by trauma, surgery, exposure to toxins or be secondary to other medical disorders. The treatment of this condition is usually symptomatic and supportive while awaiting spontaneous recovery of kidney function [14].

In chronic kidney disease (CKD) the kidney function is usually lost progressively until it may cease completely, a situation called end-stage kidney disease. The early symptoms are usually vague and include fatigue, nausea and loss of appetite [13]. Major risk factors for developing CKD include diabetes, hypertension and family history of CKD [15].

The nature of the kidney problem, i.e. acute or chronic, the stage of CKD progression as well as individual aspects of the patient affects the choice of treatment method. The three main options in renal replacement therapy are kidney transplant, peritoneal dialysis and hemodialysis, although none of them should be considered a permanent solution.

2.2 Hemodialysis

A hemodialysis machine can be considered as an artificial kidney, at least to some extent. Blood from the patient enters an extracorporeal circuit where a semipermeable membrane is situated. The blood passes by the membrane on one side and the dialysis fluid on the other side and an exchange of solutes occur between the two sides. Metabolic waste products and toxins which should be removed from the body transfer from the blood to the dialysis fluid whereas other solutes which the body is deprived of transfers the other way. In addition to this there is normally a net movement of water from the blood to the dialysis fluid, removing excess fluid from the patient. When the blood has passed by the membrane it reenters the patient's bloodline.

2.2.1 Vascular access

In order to perform a hemodialysis treatment it is necessary get access to the patient's blood circuit, a vascular access. The most crucial criteria to be met by the vascular access is a high blood flow, a minimum of 250 mL/min is desired for a standard dialysis [16] but even higher flows are preferable and as high as 600 mL/min can be used [17]. A dialysis patient who performs dialysis treatments three times a week, each 4 hours long, at high flow may filter a blood volume of up to 400 L/week. As a comparison, the kidneys of a normal person receive a total weekly blood flow of 12 000 L/week. Thus the dialysis machine only filters $3 - 4$ % of the total blood volume that passes the kidneys of a normal person during a week, even when a maximum blood flow through the vascular access is achieved.

2.2.2 Dialysis fluid

The quality standard on normal drinking water is based on a weekly intake of approximately 14 liters, all of it passing through the selective gut barrier. A hemodialysis patient may be exposed to $300 - 500$ liters of dialysis fluid per week through the synthetic dialyzer membrane, a barrier without the protective abilities of the mucous membranes in the gastrointestinal system. Due to this close contact between the dialysis fluid and the patient's blood the purity of normal drinking water is not enough [18].

Water contaminants can be divided into three major groups: particulates, dissolved substances and microorganisms. The particulates are the largest and include minerals and colloids which are responsible for the turbidity of water. Dissolved substances are of both organic and inorganic (ions and salts) nature. Microorganisms are mainly represented by bacteria and their degradation products (endotoxins), but also include fungi and viruses [18].

The dialysis fluid is prepared through two different processes, the first process is responsible for filtering and purifying the water to as high degree as possible. This is done in a central water treatment plant located in the hospital. The second process adds substances to the purified water to give it the correct composition of the final dialysis fluid. This is done separately in each dialysis machine continuously through the treatment and is adjusted to fit each patient's individual need.

The water purifying process conducted at the water treatment plant involves several steps, see Figure 2.2. First the water passes through mechanical filters to remove larger particles. If the water includes chloramines, a substance being used increasingly as a substitute for chlorine as disinfectant, it needs to be removed due to its high toxicity for dialysis patients. When present, chloramines are effectively removed by carbon filters. The next step is removal of calcium and magnesium through a softener before the final purification step which occurs in the reverse osmosis (RO) unit.

During RO, an applied pressure forces water through a semipermeable membrane which is basically only permeable to water. Thereby it removes most of the contaminants present in water, both organic and inorganic. The main purpose of the steps preceding the RO is simply to remove substances that may damage the RO-membrane. After the RO unit the water is stored and ready for distribution to the dialysis machines. The process of turning the purified water into final dialysis fluid is discussed in Section 2.3.1.3, but simplified it includes the addition of two different concentrates (A and B) containing ions and bicarbonate [8, 19].

With the removal of chloramines by the charcoal filters no disinfectant is present in the water and the bacterial growth will increase in the circuit up to the RO unit where it should be removed. However, if the RO unit and the distribution system between it and the dialysis machines are not maintained

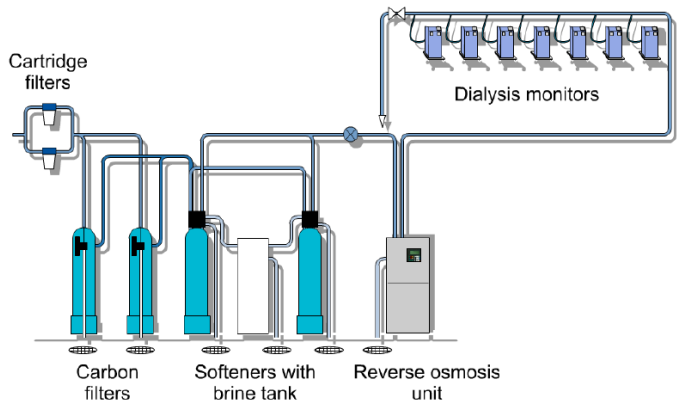


Figure 2.2: The different steps of the water purifying process. Courtesy of [8].

in a proper way a regrowth is possible. Because of this a regular disinfection scheme is of high importance to reduce the microbiological contamination to as high degree as possible [8].

2.2.3 Solute exchange in the dialyzer

The blood that leaves the patient through the vascular access enters the arterial bloodline of the extracorporeal circuit. To prevent the blood from clotting an anticoagulant, often heparin, is administered to the patient. This is done either by injecting it directly into the patient's bloodstream prior to treatment or through a pump, located in the arterial bloodline, continuously during the treatment. The dose is carefully adjusted since both too much and too little anticoagulant possesses a risk to the patient. In addition to the eventual pump with anticoagulant the arterial bloodline also includes a pump and pressure sensor to adjust the blood pressure as well as a clamp that immediately may stop the blood flow to the dialysis machine if problem occurs.

At the end of the arterial bloodline the dialyzer is situated. This is the part of the dialysis machine where the blood and the dialysis fluid interact and an exchange of water and solutes occur, see Figure 2.3.

As previously described the dialyzer is a semipermeable membrane that separates the blood and the dialysis fluid. The membrane is permeable to water and small solutes, e.g. ions, but potentially also microorganisms and toxins which is why the dialysis fluid needs to go through such a thoroughly purification process.

There are mainly two different transport mechanisms across the membrane, ultrafiltration and diffusion. The semipermeable membrane of the dialyzer corresponds to the vessel walls of the glomerulus when the process of ultrafiltration is considered. The same forces are active in both cases, i.e. hydrostatic and oncotic pressures on either side of the membrane and it is the net

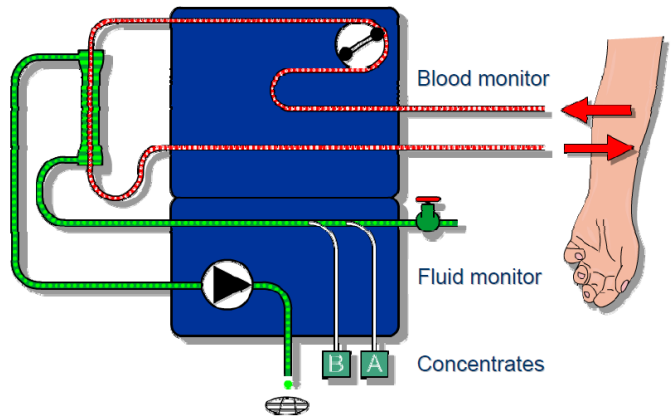


Figure 2.3: The dialyzer, the connection between the arterial bloodline and the dialysis fluid. Courtesy of [8].

Starling force that is guiding the ultrafiltration as described in Section 2.1.1. During dialysis the hydrostatic pressures dominate and the gradient they give rise to is usually deliberately set to favor water movement from blood to dialysis fluid [20]. Since the hydrostatic pressures on either side of the membrane are controlled by the dialysis machine the hydrostatic gradient is easy to regulate as compared to the oncotic gradient which can only be adjusted on the dialysis fluid side.

The other main transport mechanism across the membrane is diffusion which is the net movement of a solute from a higher to a lower concentration. A solute for which the membrane is permeable may move between the two compartments as if the membrane does not exist. The composition of the dialysis fluid decides which way permeable solutes transfer. Thereby it is the composition of the dialysis fluid that ultimately decides the composition of the blood when the treatment section is ended. To maintain a concentration gradient between the blood and dialysis fluid, and thereby the intended movement of solvents, it is important to keep a continuous flow of both fluids. To maximize the concentration gradient and diffusional flow across the entire length of the membrane the dialyzer is constructed so the fluids flow in opposite directions. Diffusion is also responsible for the passive transport that occurs in the tubuli of the nephrons described in Section 2.1.2.

When the blood has passed the dialyzer it is returned to the patient through the venous bloodline. This part of the circuit contains a drip chamber and air detector which protects the patient from potentially dangerous air bubbles. The venous bloodline also contains a pressure sensor and a clamp which may stop the blood from returning to the patient in certain alarm situations, e.g. when air bubbles are detected.

The dialyzer and the hemodialysis machine in general is an acceptable

substitute for the excretory and regulatory functions of the kidney. However, it does not try to replace the kidneys' endocrine function. Hence a patient with kidney failure usually needs injections and medications in addition to the dialysis treatment to compensate for this.

2.3 AK 98 dialysis machine

The AK 98 dialysis machine, Figure 2.4, is intended to perform hemodialysis treatments on patients suffering from renal failure or fluid overload. It may be used both in clinics and in a home care environment [5]. The AK 98 is manufactured by Baxter International Inc., previously Gambro Lundia AB. The first version was taken into commercial use in 2015 and it is a further development of the AK 96.

Figure 2.4: The AK 98 dialysis machine.

The principles of hemodialysis presented above hold true for the AK 98 dialysis machine as well. Below is a technical description of the AK 98 in general with a specific focus on the parts analyzed in this thesis.

2.3.1 Fluid unit

Figure 2.5 shows the flow path of the fluid unit during treatment which is subdivided into the following five main subsystems: [21]

- Water intake and heating system
- Chemical disinfectants intake
- Mixing and conductivity control system
- Degassing/flow pump system
- Fluid output - UF control system

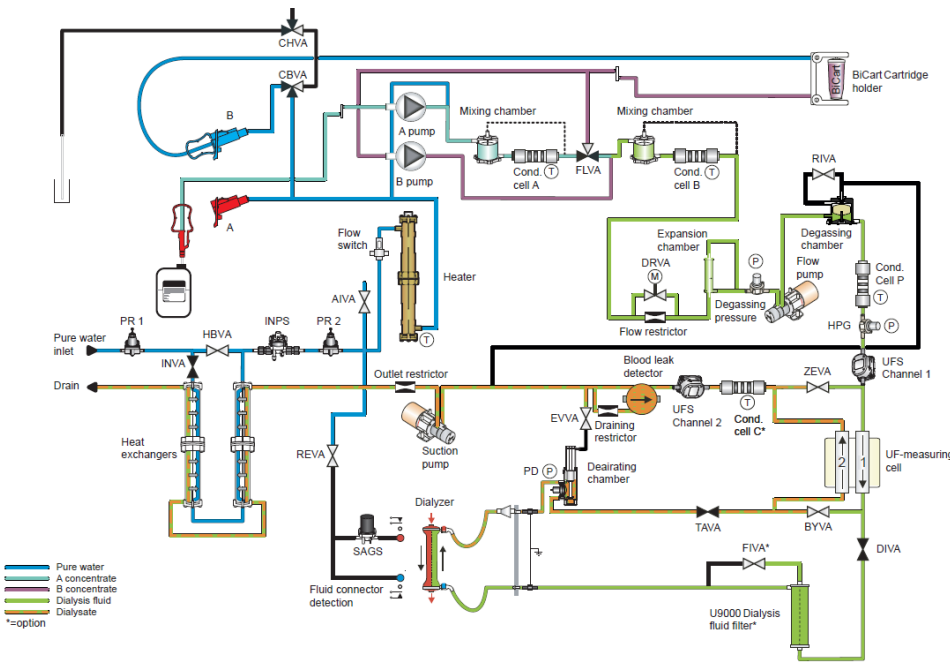


Figure 2.5: Flow path of the fluid unit. Courtesy of [5].

2.3.1.1 Water intake and heating system

The water intake and heating system in the lower left part starts with the *pure water inlet* and ends with the *heater*. It includes two pressure regulators (*PR 1* and *PR 2*) which lower the water supply pressure in two steps to about + 130 mmHg, relative the air pressure, which is monitored by *INPS*. Between the pressure regulators two *heat exchangers* are situated which uses the outgoing water to rise the temperature of the incoming water. The incoming water is raised to its final temperature when it passes through the *heater*. The temperature is regulated both by a temperature transducer at the *heater's* outlet as well as one in *conductivity cell B* (*cond. cell B*), further down the fluid path. The temperature should be around body temperature, the exact value is set manually by the operator.

2.3.1.2 Chemical disinfectants intake

The chemical disinfectants intake is placed on the rear side of the machine, in the upper left of Figure 2.5. For protective reasons it is equipped with two valves (*CHVA* and *CBVA*), controlled separately. Apart from these intakes there is another possible intake for chemicals in a different part of the system, the *BiCart cartridge holder* described in Section 2.3.1.3.

2.3.1.3 Mixing and conductivity control system

The mixing and conductivity system is the part after the *heater* to *cond. cell B*. The main flow goes from the *heater* and follows the circuit to the left and

above *pump A* and *B* where it then fuses with the outlet of *pump A*. This pump controls the addition of concentrate A to the purified water.

The composition of concentrate A may vary and is chosen based on each patient's individual needs. It is stored as a liquid in a canister and contains a number of different electrolytes and sometimes dextrose. The purified water and concentrate A is mixed in the *mixing chamber* forming prepared A dialysis fluid. The conductivity of this fluid is proportional to the amount of A concentrate added. Since the content of concentrate A is known the conductivity measure gives the actual concentration of concentrate A in the prepared A dialysis fluid. This is measured in *conductivity cell A* (*cond. cell A*) which through feedback loops regulates *pump A*.

Concentrate B is made up of bicarbonate and cannot be stored in the same canister as concentrate A since it would generate salt precipitates. It may be delivered to hospitals either dissolved in water in ready-to-use canisters or in solid form as a *BiCart cartridge*. If solid it has to be dissolved before administration to the dialysis fluid. Through slow and continuous addition of purified water the bicarbonate of the *BiCart cartridge* is dissolved and a saturated solution is achieved - concentrate B. The addition of concentrate B to the prepared A dialysis fluid is regulated through *pump B*. The fluid is mixed and the conductivity of the final dialysis fluid is measured by *cond. cell B* which in turn regulates *pump B*.

2.3.1.4 Degassing/flow pump system

The degassing/flow pump system is the part after *cond. cell B* to *conductivity cell P* (*cond. cell P*). The mixing of the A and B concentrates generates free gas, carbon dioxide, which will disturb conductivity and flow measurements if not removed. In addition to the carbon dioxide created, air may enter the dialysis fluid in various parts of the system e.g. through the A concentrate or the *BiCart*. A bypass tube in the top of the second *mixing chamber* diverts the gas away from the fluid temporarily when passing *cond. cell B*, but does not remove it. This is instead done by the degassing circuit. In the *flow restrictor* a negative pressure is created, expanding the gas to larger bubbles which is released through the top of the *degassing chamber*. The *flow pump* situated in between the *flow restrictor* and *degassing chamber* is responsible for generating the negative degassing pressure. *Cond. cell P* measures the conductivity and temperature and acts as a guard which let the dialysis fluid bypass the *dialyzer* if these values differ from the set values.

2.3.1.5 Fluid output - UF control system

The fluid output - UF control system is the remaining part of the fluid unit, after *cond. cell P*. The pressure is monitored in the *high pressure guard* (*HPG*) and then enters channel 1 of the *UF-measuring cell* which measures the fluid flow. If no alarms have been raised previously in the system the fluid is passed through the *Direct Valve* (*DIVA*) allowing it to pass the UFD and the

dialyzer. It then passes the *pressure dialysis transducer* (*PD*), and a *deaerating chamber* before returning through channel 2 of the *UF-measuring cell*. The flow measured in channel 1 is subtracted from the flow of channel 2 giving the actual UF rate. As described in Section 2.2.3 this is mainly dependent on the hydrostatic pressure gradient over the dialyzer which is calculated by subtracting the *PD* value from the pressure of the venous bloodline. The UF rate is regulated by the *suction pump* through its effect on *PD*. *UFS channel 1* and *2* are used by the protective system which is used as supervision of the UF control system among other things. They measure the same flow as the *UF measuring cell* and thereby provide independent flow data for the UF protective system.

In the case of an alarm situation, e.g. the conductivity or temperature is outside its limits, the *dialyzer* needs to be bypassed to not endanger the patient. This is accomplished through the opening of the *Bypass Valve* (BYVA) and the closing of DIVA and the Taration Valve (TAVA).

Conductivity cell C is optional, when present it may calculate the effectiveness of the treatment. The last part of the fluid unit includes a *blood leakage detector*, the aforementioned *suction pump* as well as the *heat exchangers* before the dialysis fluid is lead to the *drain*.

2.3.2 UFD

The physical localization of the optional UFD in the AK 98 is in the base of the machine cabinet, as seen in Figure 2.6a. In the fluid path, when present, it is located between the *UF-measuring cell* and the *dialyzer*. The UFD serves the purpose of being the last outpost in the purifying system of the dialysis fluid before the patient is exposed to it through the dialyzer membrane. This is done by removing possible contamination by bacteria and endotoxins that may still be left in the fluid [5].

On the macroscopic level the UFD is a 35 cm long cylinder with 4 openings, one in each end and two on the side as can be seen in Figure 2.6b. Between the two ends of the cylinder the filter mass is situated and surrounded by a plastic cover. On the microscopic level the filter mass is made up of a large amount of hollow fibers spanning the length of the cylinder. Along these fibers numerous pores are situated, providing an opening to the empty space between the fibers and the plastic cover.

The function of the UFD resembles that of a coffee filter, it separates larger particles from water. The dialysis fluid enters the UFD through the bottom opening where it is spread across the fibers. With the opening in the top being closed the pressure forces water molecules and other small particles through the pores of the fibers into the space between the fibers and the cover. From this space the dialysis fluid is drained through the upper opening on the side, the other being closed. This process leaves the contaminants of the dialysis fluid trapped inside the fibers. During disinfections the opening in the top

(a) Rear view of the AK 98.

(b) Macroscopic view of the UFD.

Figure 2.6: Localization and appearance of the UFD.

is open to rinse the UFD from the larger particles trapped inside the fibers. The use of the four openings of the UFD is illustrated in Figure 2.5 where the opening of the Filter Valve (FIVA) during disinfection enables rinsing. The lower opening on the side, being constantly closed, is left out in this Figure.

During normal usage of the AK 98 dialysis machine the UFD is degraded over time. The precipitates formed by mixing the A and B concentrate gradually clog the UFD, reducing the amount of available fibers and pores. Apart from this, the disinfection programs using either heat or hypochlorite speed up the degradation, especially hypochlorite has a severe effect on the lifetime of the UFD. Hence there are three different parameters at present involved in indicating when to exchange the UFD, namely days since last UFD replacement, number of disinfections using heat and number of disinfections using hypochlorite. As mentioned in Section 1.1 these limits are set to 90 days, 150 runs and 12 runs respectively. During normal usage this leads to a UFD exchange every $1 - 3$ months.

2.3.3 Disinfection programs

A number of different disinfection programs are available in the AK 98 dialysis machine. They differ in their duration, recommended usage intervals and efficiency against different contaminants. A summary of the different disinfection programs available can be seen in Table 2.1.

Table 2.1: Summary of cleaning and decalcification programs. All these programs have high disinfectant efficiency as well. The recommendation is to perform a heat disinfection with or without citric acid after every treatment.

	Decalcification efficiency	Cleaning efficiency	Duration (minutes)	Schedule
CLEANCART A and heat	None	High	47	-
CLEANCART C and heat	High	Medium	47	Every 3rd treatment
Citric acid and heat	High	Medium	50	-
Hypochlorite	None	High	50	Every 7th treatment day

There are two main types of disinfection programs, heat disinfections and chemical disinfections. In heat disinfections the inlet water is heated to 93 degrees Celsius and flushed through the fluid unit a number of cycles. Heat disinfections may be combined with either cleaning or decalcification programs. The purpose of cleaning programs is to remove fat, protein and other organic material whereas decalcification programs remove calcium-carbonate deposits. When performing cleaning in combination with heat, a CLEANCART A cartridge is used. Decalcification in combination with heat is done either with the use of citric acid or a CLEANCART C cartridge.

During chemical disinfections the entire fluid path is filled with a concentrated disinfectant which remains there for a certain dwell time before the fluid path is rinsed and drained. The chemical disinfectants used are based on either hypochlorite or peracetic acid, the latter being used mainly to fill the fluid path during storage when the machine is not intended to be used for 7 days or more. Disinfection with hypochlorite removes organic material.

The recommendation from Baxter is to perform a heat disinfection program with or without citric acid after every treatment. At least after every 3rd treatment the recommendation is to perform a heat disinfection using CLEANCART C and at least every 7th treatment day it is recommended to perform a hypochlorite disinfection after first running a heat disinfection using CLEANCART C.

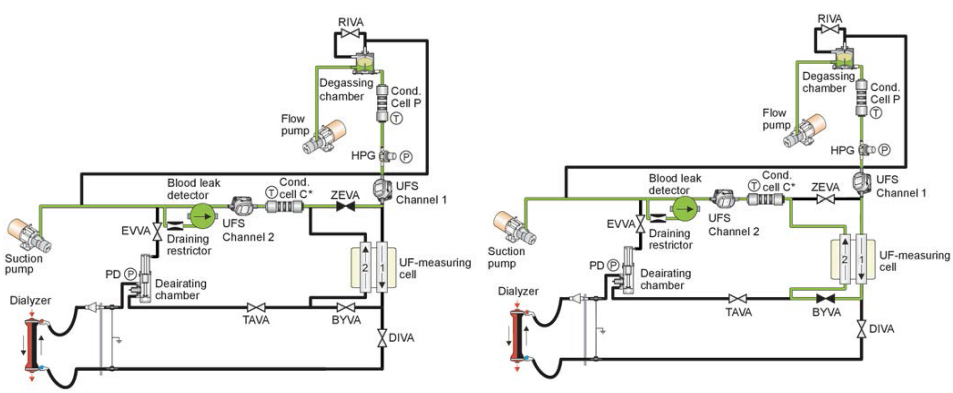
2.3.4 Regulation of main flow

The main dialysis fluid flow can be adjusted between $300 - 700$ mL/min in steps of 20 mL/min. The flow is chosen individually for each patient and is

rarely changed during a treatment. The control of the main flow is done by the adjustable *Degass Restrictor Valve* (*DRVA*) which is connected in parallel with the 200 mL/min *flow restrictor*. As previously mentioned the *flow pump* is responsible for creating the negative degassing pressure in the *flow restrictor*. The control of the *flow pump* is done by comparing the measured degassing pressure to the set point. Thereby the *flow pump* is not directly regulated by the desired main flow. However, a higher main flow will indirectly increase the workload on the *flow pump* since it has to work harder to maintain the same degassing pressure.

2.3.5 Taration

The dialysis fluid passing through channel 2 of the *UF-measuring cell* contains an addition of ultrafiltrate and diffusional products from the patient's blood. Biological substances derived this way will form a deposit called biofilm on the channel walls. Biofilm is made up of bacteria surrounded by protective polysaccharide slime. This biofilm will reduce the area of channel 2, affecting its accuracy which will lead to miscalculations of the UF rate. To compensate for this a calibration takes place every 30 minutes, a process called taration. This process starts with the opening of the *Zeroing Valve* (*ZEVA*) followed by the closing of *DIVA*, *TAVA* and *BYVA*, see Figure 2.7a. This leads to zero flow through both channels of the *UF-measuring cell* and an offset value for each channel is stored. Then *ZEVA* closes and *BYVA* opens, placing both channels in series, see Figure 2.7b. The calibration coefficient for channel 2 is changed until both channels displays the same flow. The two offsets and the calibration coefficient make up the calibrations values. If they differ too much between two tarations the latest is not approved and a new taration is scheduled in 5 minutes instead of 30.



(a) Zero flow phase.

(b) Differential flow phase.

Figure 2.7: The different phases during a taration. Courtesy of [21].

2.4 Condition based maintenance

The traditional ways in performing maintenance of components in machines and larger systems are based on either corrective or preventative maintenance. In corrective maintenance a component is not replaced until it fails. This leads to a maximal utilization of each component's lifetime assuming a degraded component does not reduce the lifetime of other components. However, in complex systems this assumption rarely applies. In a worst case scenario the failure of a single component may lead to catastrophic failure making this approach inappropriate for many systems, among them systems involving human safety e.g. airplanes and medical devices. In addition to this, the approach may also lead to longer downtimes when waiting for spare parts and service technicians to carry out the maintenance [4, 22].

Preventative maintenance aims at replacing components before they fail. Usually this is conducted in a time based manner with scheduled maintenance based on the component's historical failure information. Deciding a correct time interval of the maintenance is crucial if this approach should be successful. This is not as simple as it may seem. If one would schedule the maintenance time as the historical mean time until failure and assuming a normal distributed failure rate, half of the machines would already be broken before maintenance takes place. If the maintenance time is chosen too short it would result in a large amount of unnecessary service episodes of healthy machines. The optimal time period is one that minimizes the cost of both maintenance and failure [4, 22].

Both corrective and preventative maintenance thus have their drawbacks. A third maintenance approach which tries to overcome these has received an increased amount of interest over the years, namely CBM. The main idea is to perform maintenance in a preventative manner, but instead of basing it on time since last service it takes the actual health status of the machine in consideration. In [3] CBM is defined as

> "[...] a decisionmaking strategy to enable real-time diagnosis of impending failures and prognosis of future equipment health, where the decision to perform maintenance is reached by observing the 'condition' of the system and its components."

This definition incorporates two important concepts within CBM, diagnostics and prognostics. Diagnostics is the process of detecting a fault in the monitored system, localizing the fault to a specific component and determine the nature of the fault. Prognostics is the process of predicting failures, i.e. determine whether a failure is impending and estimate when it will occur. This estimate of time left before failure occurs is called remaining useful life [23].

The methods used in diagnostics and prognostics can be divided into three different subgroups with increasing complexity, experience-based, data-driven

and model-based [4]. Experience-based approaches are the most simple and rely on statistical information regarding historical failure rates. Based on this, distributions of failure rates over time are developed which can be used to create a maintenance schedule. However, this is a form of preventative maintenance without any predictive qualities. Hence it could not be considered a true prognostic method.

Model-based approaches exhibit the highest potential when it comes to developing prognostic algorithms. They are based on physics-of-failure models derived from first principles and thereby require an in depth knowledge of the system and what happens when it fails. If such a model is feasible to develop it has the benefit of being able to be used regardless of load or operating conditions. However, in many situations the system under observation is too complex to be described through first principles. In these situations data-driven approaches are a better choice. They make use of historical data of measured signals from the system in all stages from fully functional to failure. Through analysis of the progress of specific features derived from these signals it is possible to develop prognostic algorithms of the systems RUL. This master's thesis focuses on data-driven approaches [4].

2.4.1 Related work

The use of prognostic algorithms in CBM is a rather new phenomenon and the research in this area grows rapidly [3, 23]. The recent developments in CBM are driven by advancements in sensor technologies and improvements in the collection, storage and processing capabilities of sensor data [4]. The majority of the developed models are application specific and not generalized [3]. However, the workflow of general CBM methods all follow the same framework [24]. An illustration of this can be seen in Figure 2.8.

The research literature in CBM is focused into two main categories; mechanical- and electrical engineering [24]. The area of mechanical engineering includes analysis of complete systems as well as smaller subsystems, but they usually narrows down to an analysis of simple, single mechanical components such as ball bearings, bearings and gears. Implementation of CBM in the field of electrical engineering is focused around RUL estimation of batteries.

No literature concerning CBM of fluid filters has been found. However, development of data-driven CBM methods does not only share a common workflow, they are also based on a limited amount of statistical and machine learning algorithms. Hence a review of the current status of this research field is of interest in this master's thesis, although the selection of analyzed signals and the feature extraction from these have to be made without any knowledge based on prior research.

In a review article from 2011, an extensive analysis of the modeling developments for estimating the RUL of that time is presented [25]. The authors

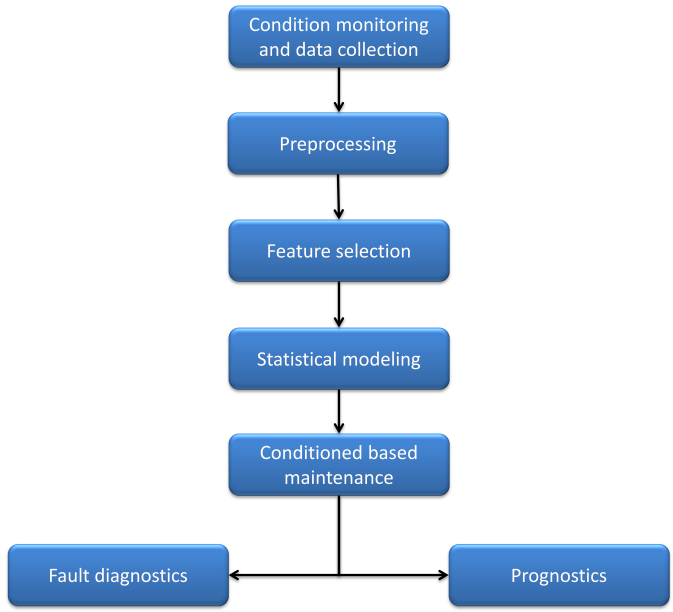


Figure 2.8: Workflow of general CBM methods.

also go through previous review articles within the field of maintenance related issues in general and CBM in particular. From their review of published papers they concluded there was no comprehensive review regarding statistical based data-driven approaches for RUL estimation, hence the focus of their article. For the interested reader, [26–30] provides a summary of the research within the broader area of maintenance related issues and [3, 23, 31–35] gives an overview of older progress regarding CBM.

In a more recent review article from 2015 with focus on data-driven approaches a list with commonly used time-domain features is presented. In addition to this a number of different standard methods within data-driven prognostic are described, among them Markovian process-based models, regression based models and proportional hazard model [24].

With the aim set for the industry and companies in their initial phase of developing prognostic algorithms [36] made an attempt to summarize the prognostic modeling options available. The models proposed are classified as belonging either to knowledge based, life expectancy, artificial neural networks or physical models. The paper does not solely describe the methods available but it also gives a detailed explanation about situations when the methods are suitable and when they should be avoided. An example of this is knowledge based expert systems which can be considered when the problem area is well understood and the operating conditions are stable but it should be avoided when the opposite holds true or a very accurate estimation of the RUL is required.

One example of a model developed for a specific case is presented in [37] where the gear crack level is identified through the use of a k-nearest neighbors (kNN) classifier. In the article a total of 25 different features derived from vibration data are tested for three different gear crack levels. Out of these features, 10 are derived from the time-domain, 4 from the frequency-domain and the remaining 11 features were specially developed for gear damage detection. To select and weight the features best suited for detecting gear crack level in the current experimental setup a two-stage feature selection and weighting technique (TFSWT) via Euclidean distance evaluation technique is used. Seven out of the 25 features are selected and given weights between 0.64 and 1.0. Based on a test/training set of 36/36 or 48/24 samples an accuracy of 86 - 100 % is achieved based on the number of neighbors chosen.

In [38] an algorithm to determine the RUL of an elevator door motion system is demonstrated. Two features derived from an encoder monitoring the door displacement are extracted. Logistic regression is used to map the status of the system from normal (0) to failure (1). This is used together with an ARMA model to estimate the RUL.

In machinery CBM signals such as vibration, temperature and pressure from different components are often used. In [39] a prognostic algorithm deciding the RUL of bearings of a high pressure pump is presented. A total of 10 different statistical parameters from the time domain and 4 parameters from the frequency domain are derived from vibration signals from three accelerometers on the pump housing and tested as features. Out of these, 4 are selected for the final analysis where the bearing failure status is divided into 6 stages from maximal to minimal RUL. The machine learning algorithm support vector machines is used as the classifier. The result of the RUL estimation closely follows the true remaining life of the bearings.

2.5 Machine learning

Machine learning covers the field of developing algorithms that may be trained, or learned, with a data set and based on this are able to predict the identity, or class, of new data. This field can be subdivided into three different areas, supervised learning, unsupervised learning and reinforcement learning [40]. In supervised learning the class of the data is known and may therefore guide in the learning process. This is the area of machine learning that will be used in this master's thesis and it will be described in Section 2.5.1.

In unsupervised learning the class of the data is not known. Thereby the aim is to discover how the data may be organized in different clusters [41]. An example of this could be to find specific areas of the brain associated with a certain task. Another example could be finding expression patterns of genes

associated with cancer tumors.

In reinforcement learning the task is to find suitable actions in a given situation. These algorithms are not given examples of output in the form of classes, but must instead find the solution itself through a trial and error process. Examples of this are algorithms developed to play games like backgammon and chess [40].

2.5.1 Supervised learning

The key steps of supervised learning can be summarized in the steps of feature extraction, labeling, training and classification.

The purpose of feature extraction is to simplify the problem by reducing the total amount of data in each data sample to a small set of descriptive measures, or features, that could be analyzed instead. An example of this could be to reduce the data, in the form of pixels, for an $m \times n$ picture (a data sample) down to just two measures, e.g. mean and maximum intensity. Thereby the analysis may be based solely on these two measures instead of using the intensity value of each pixel. Hence the problem is reduced from $m \times n$ dimensions to only two. In general, the I extracted features of the M samples are stored as an $M \times I$ matrix where each sample may be seen as a point in an I-dimensional feature space.

The process of labeling is to assign a label, or a class, to each data point. This has to be done manually or through a decision rule to get a ground truth for the data. An example could be pictures of a single digit which are labeled with their corresponding digit or to let a computer label documents based on the month they were created. In this report the words *class* and *label* are used interchangeably.

The training of a model is done through the observation of multiple combinations of feature sets and their corresponding label [41]. The aim of this step is to develop a function that may translate a given feature set into a label [40]. Since it is supervised learning the correct label is known during the training phase, thereby it is easy to get an assessment of the function.

When a data set should be analyzed through supervised learning it is often subdivided into a training and test set. The training set is used to develop the model while the test set is used for the actual analysis. Since the training is done for the classifier used in the next step, the implementation of the training step differs depending on which classifier that is chosen.

The final step of the process is the classification. The goal for the trained classifier is to decide the class of unknown feature sets. A number of different classifiers used for supervised learning are available including kNN, support vector machines, decision tree learning and artificial neural networks to mention a few [41]. In this master's thesis kNN is used.

2.5.1.1 k-fold cross validation

One way of splitting the data into a training and test set in an organized way is through k-fold cross validation. This method splits the data into k different groups, using one group for the test set and the remaining $k-1$ groups for the training set. The classification is repeated a total of k times allowing every group to be used as test set once, with the result being the mean of all k classifications. [41]

2.5.1.2 k-nearest neighbors

The classification method kNN is despite its simplicity successful in a wide range of classification problems [41]. The training step is very simple and only includes storing the feature set of the different training data points. To classify a test data point, the Euclidean distances between it and all training data points are calculated. The k closest training data points are chosen and the most frequent label among them determines the label of the test data point.

An example of a two-dimensional feature space including a test data point and training data set with 3 different classes can be seen in Figure 2.9. As can be seen in the figure these two features are enough to separate the three classes into distinct clusters, using only one of these features would not be enough. If a kNN classifier would be used in this case the test point would be classified as class 0, independent of the value of k.

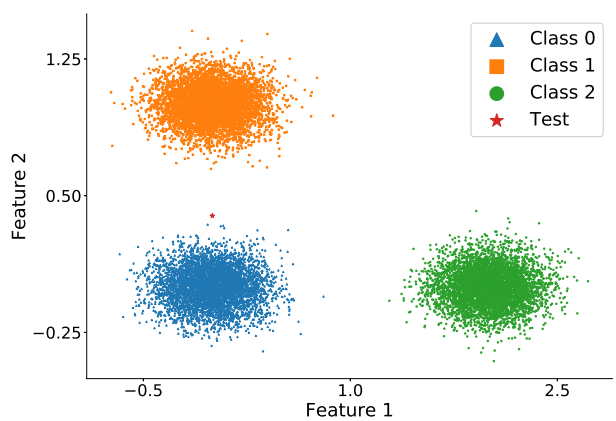


Figure 2.9: Example of a two-dimensional feature space.

Different weighting techniques are available when using kNN. One option is to base the weighting on the distance between the test point and its neighbors, giving closer neighbors a larger weight. This is done by assigning a neighbor at distance d to the test point the weight $1/d$.

2.5.1.3 Evaluation

To evaluate the result the quantity of classifier accuracy can be used. This is achieved through the division of the number of correct classified test points with the total number of test points.

A more sophisticated evaluation method is to use a confusion matrix, see Figure 2.10. It gives a better visualization of the performance, allowing the interpreter to analyze the classification of each label in more depth. The rows of the confusion matrix represent the classification of the actual label, i.e. row 1 shows how test data points with the actual label 1 are classified. In contrast, the columns represent the prediction of labels, i.e. column 1 shows the true label of test points predicted to be label 1. This means that each row sums up to one which is not necessarily true for each column.

Out of the confusion matrix in the figure it can be seen that 9.7 % respectively 10.4 % of the test points with label 1 are wrongly classified as label 0 and label 2, whereas 80.0 % are correctly classified as label 1. Thereby all the correct classified test points are positioned on the diagonal going from the top left to the bottom right corner. A perfect classification would result in a confusion matrix where all squares on the diagonal have the value 1.0 and the remaining squares the value 0.0. Hence, the darker the diagonal is in a confusion matrix, the better the classification is.

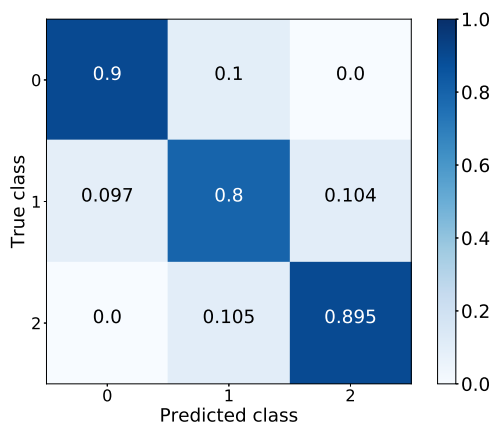


Figure 2.10: Confusion matrix with three different classes.

3

Data

The data analyzed in this master's thesis is derived from the log files of 10 dialysis machines of model AK 98 located at a hospital. Thus the data is derived from treatments of real patients. The log files are compressed to .zip format where each file contains approximately one month of data acquisition. A total of $10 - 13$ .zip files from each machine were analyzed, corresponding to a total time frame of $8 - 12$ months.

3.1 Log files

The internal memory of the AK 98 may contain data for about one month of machine runs during normal usage, i.e. $2-3$ treatments a day of approximately $3-4$ hours. The logging is initiated each time the machine starts. Thereby it is not only actual treatments that are logged but all machine runs, including e.g. disinfections and services. When the memory runs out the oldest entries are overwritten. Because of this a monthly extraction of the .zip files is suitable if no data should be lost.

The process of extracting the .zip files results in the saving of each machine run in a separate file. In addition to this, two other files are generated, containing information about which parameters that have been logged and how this logging was done.

There are three different ways a parameter may be logged. The first option is to log every update which is the standard for e.g. alarms. The second option is to log with a predefined time interval although this option is rarely used. The third option, the delta-method, is to store a new value only when a predefined change from the last logged value has been exceeded. This change, or delta, is chosen based on each signal's individual characteristic. This reduces the amount of data stored but also results in an uneven sampling rate for each signal.

The data from all sensors and valves in Figure 2.5 are available in the log files. In addition to this, a number of different machine processes and states as well as manual input to the machine are logged. This adds up to a total of over 400 different parameters available.

3.2 Analyzed signals

Out of the extensive amount of sensor data available it is only the signals believed to have some relationship to the UFD status that have been analyzed. The sensors responsible for these signals are visualized in Figure 3.1 which is a magnification of the lower right part of Figure 2.5. To be able to perform a suitable preprocessing of these signals it is not only sensor data that has been retrieved from the log files, but also logged information regarding machine states. A brief description of the signals most relevant to this master's thesis will follow below in Sections 3.2.1−3.2.6. An example of each of these signals from a standard run can be seen in Figure 3.2.

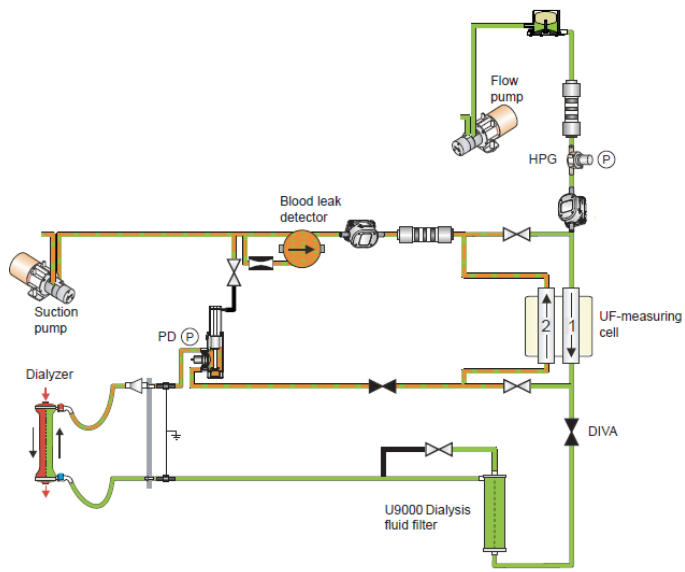


Figure 3.1: A magnification of the most relevant part of the fluid path situated in proximity to the UFD.

3.2.1 Flow pump derived signals

As described in Section 2.3.4, the *flow pump* is responsible for keeping a constant degassing pressure and thereby it is indirectly affected by an increased main flow. The parameter *flow pump cycle* measures the workload of the *flow pump* given in percentage of its maximum capacity. It is logged with the delta-method with a delta of 1.5 percentage points.

The parameter *flow pump current* is related to *flow pump cycle*. It describes the electrical current used by the flow pump and is measured in mA and logged with a delta of 10 mA.

3.2.2 HPG

HPG stands for *high pressure guard*, it is a pressure transducer located before the *UF measuring cell*. As the name implies it serves the purpose of protecting against high pressures if a tube would be blocked. If a blockage appears before the *dialyzer* the tubes could disconnect. If it appears after the *dialyzer* the membrane may be damaged.

The *HPG signal* is measured in mmHg and it is very noisy to its nature, hence it has a high delta of 40 mmHg.

3.2.3 UF channel 1

The main flow of the dialysis fluid is measured in channel 1 of the *UF measuring cell*. Since some of the analyzed signals are affected by the main flow it is necessary to perform a compensation for this in order be able to compare signals with different main flows. The signal *UF channel 1* is used for this purpose. It is measured in mL/min with a delta of 10 mL/min. The process of compensation is further explained in Section 4.7.

3.2.4 PD

The *PD* transducer measures the pressure immediately after the *dialyzer*. It is used by the machine together with the venous pressure to calculate the transmembrane pressure. The PD signal is measured in mmHg with a delta of 20 mmHg.

3.2.5 Blood leak detector

The *blood leak detector* is located after channel 2 of the *UF measuring cell* in order to detect the possible presence of blood in the dialysis fluid, indicating a rupture in the *dialyzer membrane*. The detection is done with an infrared light detector using a LED transducer and a photo transistor as the receiver. The signal from the blood leak detector is measured with a delta of 0.2 percentage points. The alarm is raised if the limit of 3 % is reached, corresponding to a blood leakage of 0.35 mL/min with a hematocrit of 32 %.

3.2.6 Suction pump derived signals

As described in Section 2.3.1.5 the suction pump is responsible for maintaining a correct UF rate through its effect on *PD*. The parameters *suction pump cycle* and *suction pump current* are equivalent to the corresponding *flow pump* parameters, using the same deltas of 1.5 percentage points and 10 mA, respectively.

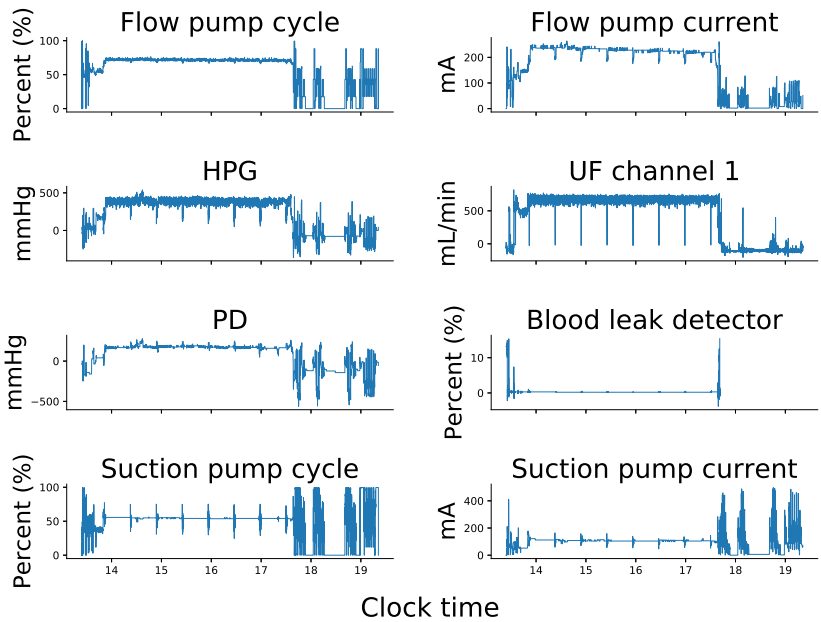


Figure 3.2: An example of the eight most important sensor signals used for the analysis in this master's thesis. The depicted signals are all derived from the same log file. The tarations occurring every 30 minutes during the treatment are easily detected in some of the signals.

4

Method

This project was subdivided into two distinct parts. The first part was the algorithm development which constituted the vast majority. The second, much smaller part was the evaluation of the developed algorithm. In order to avoid biasing, the 10 machines available for analysis were split into two sets with five machines each. One of these sets was used for the algorithm development whereas the other was used for the evaluation of the final algorithm.

The five machines from the algorithm development set were used in the time consuming process of finding the best features and parameter settings. Thus, for the five machines in the evaluation set only these selected features and optimal parameter settings were used during the classification. The final evaluation of the algorithm was therefore based on data that were never used in the development phase.

In the multiple classifications during the algorithm development the training and test data were derived from the five machines of the algorithm development set. In the evaluation of the algorithm the training and test data were derived from the 5 machines of the evaluation set. Thus, the two sets were never mixed. In both cases k-fold cross validation was used.

The method used during the algorithm development was subdivided into shorter steps. The first step was to localize when the UFD replacements and hypochlorite disinfections had taken place in the analyzed machines. Relevant signals were then chosen and preprocessed in order to remove noise and unwanted segments. Based on the found occasions for UFD replacements and hypochlorite disinfections the data was labeled in two different ways. A number of different descriptive features were calculated. To be able to analyze all data together, features that were affected by the main flow were compensated. All features were then normalized and tested to find which ones that were the best to separate different labels. The best features were selected and fed to a classifier which used training data to classify test data. The predicted label was compared to the true label. The classification step was repeated a number of times to test different parameter settings during the algorithm development phase. These steps are further explained in the sections below.

In order to study whether the algorithm was machine specific or not all tests were run in parallel for two different cases, the single respectively multiple machine analyses. In the single machine analysis the machines were analyzed one by one whereas they were analyzed all together in the multiple machine analysis.

4.1 Localizing UFD replacements

Since no information regarding the occurrences of the UFD replacements was available, this had to be retrieved from the log files. When one of the three counters indicating time for UFD replacement has reached its limit an attention is raised. After the UFD replacement the operator needs to press a button to confirm the replacement and remove the attention, thereby resetting the three different counters. This procedure is logged in Process 153 in the state machine, *TimeBetweenUfdChangeProc*, as a state change from any state to state five, *ResetUfdReminder*. The log files containing UFD replacements were saved in a separate list and removed from the rest of the data set.

4.2 Localizing hypochlorite disinfections

To retrieve information regarding when hypochlorite disinfections were conducted the signals $FI\_DisinfTypeID$ and $FI\_DisinfType-StartStop$ were analyzed. All log files were gone through to find occasions when $FI\_DisinfTypeID$ was set to eight, indicating hypochlorite disinfection. The corresponding start and stop times for each disinfection program were collected through $FI\_DisinfTypeStartStop$. In instances when no stop time was found this way the end time was chosen as the end point of the log file. From these times the duration of the disinfection could be calculated, only those exceeding 45 minutes were saved as true disinfections. When a log file contained both a hypochlorite disinfection and a treatment episode the time notifications of each were used to determine how to label the log file.

4.3 Preprocessing of signals

During the extraction of the .zip files each machine run is saved in a separate log file. The first step in the preprocessing was to remove log files corresponding to disinfections, services and shorter treatments. To achieve this the signal *MACHINE_MODE* was used to detect which files that contained a treatment episode and the duration of it. Only files containing a single treatment episode with a duration of more than 90 minutes were further analyzed. Files with multiple treatment episodes or an episode shorter than 90 minutes were discarded due to unstable signal behavior and to remove factory runs.

When the files of interest were selected, the signals retrieved from each file were preprocessed further. Since the logging is initiated when the machine is turned on, each file does not only include data collected during treatment, but also data from start-up, functional check, blood line preparation, pre-treatment, post treatment, disinfections and service. To remove unwanted data only the time frame corresponding to actual treatment was used. This was selected as the time frame when *MACHINE_MODE* was set to 4, which equals treatment, resulting in a continuous signal of at least 90 minutes. This part of the signal included tarations which largely affect some of the signals as can be seen in Figure 3.2. It could also contain episodes when the UFD was bypassed due to e.g. alarms. To remove these unwanted episodes the signal *O_DIVAOPEN* was used. Since some signals exhibited unstable signal behavior around the opening and closing of DIVA an extra cutoff of 4 seconds was used in both cases.

The data remaining after these steps varied greatly between different signals, it could differ as much as a factor 30 in number of samples between different signals of the same treatment episode. To be able to compare the different signals from the same file they had to be uniformly resampled. This was done by using linear interpolation with a sample frequency of 100 Hz, which was enough to avoid aliasing.

Some of the signals behaved unstable in the start-up phase of the treatment. Due to this, the first ten minutes were discarded from all signals. From this point forward, log file refers to a log file that has gone through the preprocessing steps up to this point.

The final, optional, step in the preprocessing was to segment each signal into smaller time windows of equal duration. The length of the time window was set manually by the adjustable parameter *window length*. It was also possible to reject this step and keep the entire treatment as a single segment. In this report window lengths of 5, 15 and 30 minutes have been analyzed in addition to the no windowing option.

The steps of the preprocessing are summarized in chronological order below:

- Discard runs with treatment episodes < 90 min.
- Discard runs with multiple treatment episodes.
- Only keep data corresponding to *MACHINE_MODE* $= 4$ (treatment).
- Only keep data when main flow is passing UFD (*DIVA* = open) with an extra cutoff margin of 4 seconds.
- Resample signals through linear interpolation.
- Discard first 10 minutes of each signal.
- Subdivide signals into shorter segments (optional).

4.4 Labeling of data

Two different methods for labeling the data have been tested. The first method was based on the number of log files between two adjacent UFD replacements. Based on the parameter *number of labels*, which was variable, the log files were divided into equally large groups based on their chronological order. If the number of log files and *number of labels* were not divisible the extra log files were split as evenly as possible. E.g. if there were 62 log files between two adjacent UFD replacements and *number of labels* was set to 3, each label should contain 20.67 log files. However, since the log files were not split the number of log files in each label was rounded to the closest integer. Thus label 0 is assigned to log file 1 to $20.67 \approx 21$, label 1 is assigned to log file 22 to $2 \times 20.67 = 41.33 \approx 41$ and label 2 is thereby assigned to log file 42 to 62, resulting in 21, 20 and 21 log files with label 0, 1 respectively 2. Independent of *number of labels*, label 0 always corresponds to the log files immediately succeeding a UFD replacement, hence the longest RUL of the UFD. Consequently, the last label always corresponds to log files immediately preceding the next UFD replacement, hence the shortest RUL of the UFD.

The second labeling method was based on the number of hypochlorite disinfections the UFD had experienced. Label 0 was assigned to log files with a UFD exposed to zero hypochlorite disinfections. This corresponded to log files situated in between a UFD replacement and the first hypochlorite disinfection. Label 1 was assigned to log files with a UFD exposed to 1 hypochlorite disinfection, corresponding to log files between the first and the second hypochlorite disinfection and so on.

In Figure 4.1 a comparison between the two different labeling methods on a synthetic data set of 63 log files and three labels can be seen. The blue vertical lines symbolize filter replacements and the red lines hypochlorite disinfections. Due to the irregular occurrences of these disinfections this method give rise to some discrepancies in the number of log files in each label. These discrepancies may be large for a small data set but decreases with the size of it due to a reduced impact of randomness. The 63rd log file is situated after the second filter replacement, hence the labeling is reset from 0 for both methods.

Both labeling methods share some common characteristics. When the log files are subdivided into shorter segments, each segment is assigned the same label as the log file it is derived from. Log files preceding the first registered filter replacement and treatments succeeding the last registered filter replacement of each machine are discarded due to uncertainty of how they should be labeled.

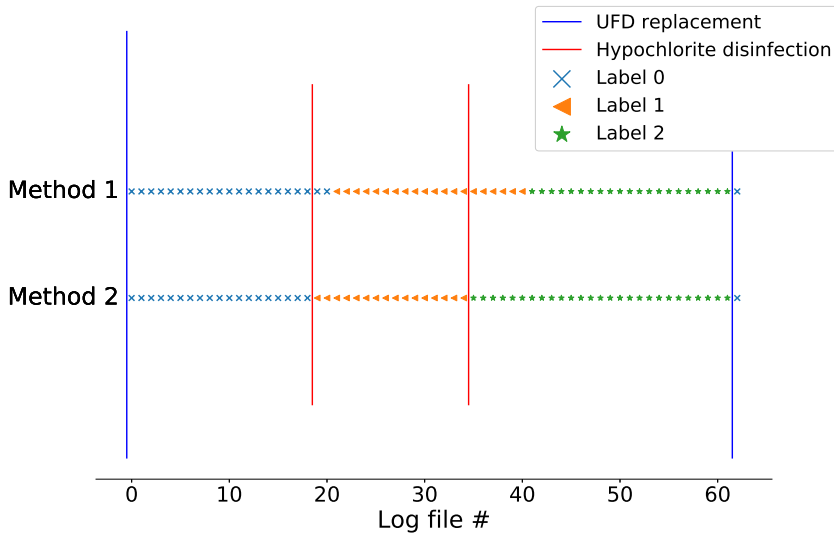


Figure 4.1: Illustration of the two different labeling methods. With method 1 the log files are close to being evenly distributed, the first 21 log files are assigned label 0, the following 20 log files are assigned label 1 and the last 21 are assigned label 2. Method 2 are based on the hypochlorite disinfections (red vertical lines), occurring after the 19th and 35th log files. This results in assigning label 0 to the first 19 log files, label 1 to the succeeding 16 log files (log file number 20−35) and label 2 to the last 27 log files.

4.5 Feature extraction

A total of eight different signals have been analyzed. Seven of these were obtained directly from the log files. These are presented in Figure 3.2, except for the signal *UF channel 1* which is used for the flow compensation. The 8th analyzed signal was created as the difference between two of the signals, the *HPG* and *PD* pressures, and is presented in Figure 4.2. On all of these signals, 21 different measurements have been calculated, where 15 are in the time-domain and 6 in the frequency-domain. In addition to this two different cross-correlation based measurements have been calculated for five different signal pairs.

The combination of a measure and a signal is defined as a feature. With 21 measurements calculated from 8 signals in addition to 2 measurements calculated from 5 signal pairs a total of 178 different features were extracted from each log file. This was done either for the entire signal or individually for each segment when windowing was used.

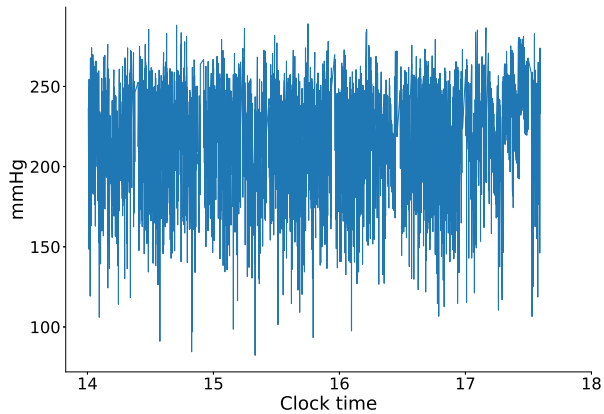


Figure 4.2: An example of the ($HPG-PD$) signal after preprocessing. It is derived from the same treatment episode as the signals in Figure 3.2.

4.5.1 Features derived from time-domain

The first 4 of the measurements derived from the time-domain corresponds to the standardized moments in probability theory, i.e. mean (μ), standard deviation (σ), skewness (SK) and kurtosis (KU) presented in Equations (4.1)–(4.4).

$$\mu = \frac{1}{N}\sum_{n=1}^{N} x_n \tag{4.1}$$

$$\sigma = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \mu)^2} \tag{4.2}$$

$$SK = \frac{\sum_{n=1}^{N}(x_n - \mu)^3}{(N-1)\sigma^3} \tag{4.3}$$

$$KU = \frac{\sum_{n=1}^{N}(x_n - \mu)^4}{(N-1)\sigma^4} \tag{4.4}$$

The next 6 measurements are mentioned in [24] as commonly used time-domain features in CBM. Those are maximum value (max), root mean square (RMS), crest indicator (CI), clearance indicator (CLI), shape indicator (SI) and impulse indicator (II) presented in Equations (4.5) – (4.10).

$$max = \max_{n=1,\ldots,N} x_n \tag{4.5}$$

$$RMS = \sqrt{\frac{1}{N}\sum_{n=1}^{N}(x_n)^2} \tag{4.6}$$

$$CI = \frac{\max|\mathbf{x}|}{RMS} \tag{4.7}$$

$$CLI = \frac{\max|\mathbf{x}|}{(\frac{1}{N}\sum_{n=1}^{N}\sqrt{|x_n|})^2} \tag{4.8}$$

$$SI = \frac{RMS}{\frac{1}{N}\sum_{n=1}^{N}|x_n|} \tag{4.9}$$

$$II = \frac{\max|\mathbf{x}|}{\frac{1}{N}\sum_{n=1}^{N}|x_n|} \tag{4.10}$$

In addition to these measurements another 5 were used, namely minimum value (min), lower quartile (Q_1), median (Q_2), upper quartile (Q_3) and interquartile range (IQR). The equation for calculating the minimum value is presented in Equation (4.11). The median of a data set is the value that separates the higher half of the values from the lower half. The lower quartile is the value that separates the lowest 25 % of the data from the highest 75 %. The upper quartile is correspondingly the value that separates the lowest 75 % of the data from the highest 25 %. The interquartile range is the difference between the upper and the lower quartile.

$$min = \min_{n=1,\ldots,N} x_n \tag{4.11}$$

4.5.2 Features derived from frequency-domain

The fast Fourier transform (FFT) was used to transform the data from time-domain to frequency-domain, see Equation (4.12). To take advantage of the FFT's special feature making it much faster for data of length 2^p, where p is a natural number, zero padding was used.

The measurements used in the frequency domain are the dominant frequency (DF), the sum of the 10 largest Fourier coefficients (c_{1-10}), lower quartile (c_kQ_1), median (c_kQ_2), upper quartile (c_kQ_3) and interquartile range (c_kIQR) of the Fourier coefficient. The equations for calculating the DF and Fourier coefficients (c_k) are given in Equations (4.13)−(4.14).

$$X_k = \sum_{n=0}^{N-1} x_n e^{\frac{-i2\pi nk}{N}}, \quad k = 0, ..., N-1 \tag{4.12}$$

$$DF = \arg\max_k(c_k) \tag{4.13}$$

$$c_k = \frac{|X_k|}{N} \tag{4.14}$$

4.5.3 Features derived through cross-correlation

Cross-correlation is used to detect correspondences between two time series. All signals derived from the fluid unit of the AK 98 are part of the same system subdivided into different smaller control systems. Even if two signals are not in direct connection through a control loop where one of them regulates the other, they may still exhibit some dynamic behavior. The idea behind using features derived through cross-correlation was to analyze signal pairs believed to have some connection and detect if a clogging of the UFD could have an effect on their potential dynamic behavior.

Five different signal pairs were analyzed through cross-correlation. These signal pairs are presented in Table 4.1 and they were chosen based on an expected connectivity. The equation for cross-correlation is presented in Equation (4.15), here assuming signals of equal length. From the result of the cross-correlation two different measures were derived, the maximum value (CCM) and the time delay (CCTD), see Equations (4.16)−(4.17).

Table 4.1: Signal pairs used for cross-correlation.

Signal 1	Signal 2	Abbreviation
Flow pump cycle	*Suction pump cycle*	Pair 1
Flow pump current	*Suction pump current*	Pair 2
Flow pump cycle	*(HPG−PD)*	Pair 3
Flow pump cycle	*HPG*	Pair 4
Suction pump cycle	*PD*	Pair 5

$$\hat{r}_{yx}(k) = \frac{1}{N} \sum_{n=0}^{N-1-k} x_n y_{n+k}, \quad k = \pm 0, ... \pm N-1 \tag{4.15}$$

$$CCM = \max_{k=\pm 0,...,\pm N-1}(|\hat{r}_{yx}(k)|) \tag{4.16}$$

$$TD = \arg\max_k(|\hat{r}_{yx}(k)|)\frac{1}{FS}, \quad FS = \text{sampling frequency} \tag{4.17}$$

4.6 Split of data into training and test sets

When the feature extraction was completed the log files were split into a training and test set using k-fold cross validation [41]. When windowing was used all windows derived from the same log file were kept as a unit belonging to either the training or the test set. When no further explanation is given regarding the split of the data into training and test sets the following holds true.

Since 5 machines were used for the algorithm development it was natural to use $k = 5$ in the k-fold cross validation during the multiple machine analysis. The log files were grouped according to which dialysis machine they were derived from. Thereby the log files from a dialysis machine were never split between the training and test sets and every machine could be used as test set once.

For comparison, $k = 5$ was used in the single machine analysis as well. Thereby 80 % of the log files were assigned to the training set and 20 % to the test set.

4.7 Flow compensation

When the median flow of *UF channel 1* was analyzed for all the log files of a single machine it was found that 3 different main flows were probably used; 500, 600 and 700 mL/min as seen in Figure 4.3. When the different features were plotted against main flow it became apparent that some, but not all, were affected by the flow. To make the algorithm independent of main flow this had to be compensated for.

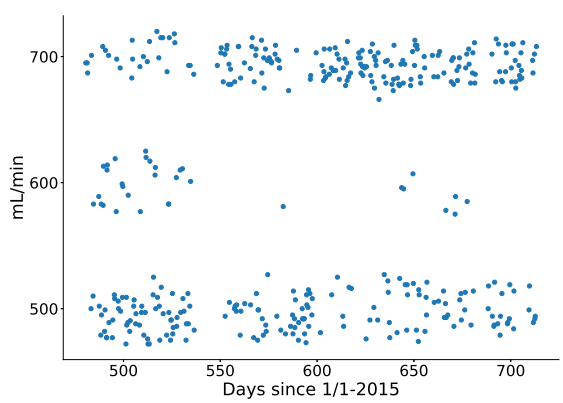


Figure 4.3: Illustration of the median flow of *UF channel 1* for all log files of a single machine when no windowing was used. Each sample corresponds to the median flow of a single log file.

The compensation was done individually for each of the 178 features. Ordinary least squares regression was used to estimate the relationship between the samples of a feature and main flow, see Equation (4.18), where $\mathbf{x}$ is the main flow vector, $\mathbf{y}$ is the samples of the i:th feature ($\mathbf{f_i}$), and $\mathbf{1}$ is a column vector of ones with the same length as $\mathbf{x}$. The sample points were then transferred to a 600 mL/min main flow using $\hat{\beta}_0$, see Equation (4.19).

$$\hat{\boldsymbol{\beta}} = [\hat{\beta}_0 \quad \hat{\beta}_1]^T = (\mathbf{X^T X})^{-1}\mathbf{X^T y} \tag{4.18}$$

Where

$$\mathbf{X} = [\mathbf{x} \quad \mathbf{1}]$$

$$\mathbf{x} = [x_{(0)}, \ldots, x_{(N-1)}]^T$$

And

$$\mathbf{y} = \mathbf{f_i}$$

$$\mathbf{y_{comp}} = \mathbf{y} + (\hat{\beta}_0(600 \times \mathbf{1} - \mathbf{x})) \tag{4.19}$$

An example of this compensation is presented in Figure 5.2 were both original and compensated samples of a feature are plotted. Which features that finally used the compensation are presented in Table 4.2. This decision was based on visual assessment of the plot of original and compensated samples for each feature and the corresponding linear fit.

Table 4.2: The 7 measurements from the 6 signals below constitutes the 42 features that were flow compensated. The corresponding equations are given in brackets.

	Measurements
Flow pump cycle	Mean (4.1), Max (4.5), RMS (4.6), Lower and Higher Quartile, Median and Min (4.11)
Flow pump current	
HPG	
(HPG−PD)	
Suction pump cycle	
Suction pump current	

4.8 Feature normalization

In order to give each feature the same possible impact on the classification result, independent of the scale of the feature, they had to be normalized. This was done through the use of the *Min-Max* normalization method given

in Equation (4.20) where $f'_{m,i}$ is the normalized value of sample m for the feature i. Thereby all features are scaled to the range [0, 1].

$$f'_{m,i} = \frac{f_{m,i} - \min \mathbf{f_i}}{\max \mathbf{f_i} - \min \mathbf{f_i}} \tag{4.20}$$

4.9 Feature selection

Two different methods to perform feature selection were tested and evaluated. The first method was the TFSWT based on Euclidean distance evaluation technique described in [37]. As the name implies it also includes weighting of the selected features but this step was not implemented. The second method was sequential forward selection (SFS).

When classification was done during the TFSWT and SFS the classifier used was kNN with $k = 10$, using the weighting technique based on distance mentioned in Section 2.5.1.2. The labeling was done with the first labeling method, time based labeling, using 12 labels. No windowing was used.

4.9.1 Two-stage feature selection and weighting technique

This method is a comparison of the average distances of the feature samples between and within the different labels. This ratio is higher the better a feature is at separating the different labels, indicating a large distance between samples of different labels and a short distance between samples of the same label.

The first step of this method is to calculate the average distance D between the m samples of the feature i of label c, see Equation (4.21). M_c is the number of samples for label c.

$$D_{c,i} = \sqrt{\frac{1}{M_c \times (M_c - 1)} \sum_{l,m=1}^{M_c} (f_{m,c,i} - f_{l,c,i})^2}, \quad l \neq m \tag{4.21}$$

Then the average distance is calculated within (w) each feature as

$$D_i^{(w)} = \frac{1}{C} \sum_{c=1}^{C} D_{c,i} \tag{4.22}$$

The variance factor for each feature is calculated as

$$V_i^{(w)} = \frac{\max(D_{c,i})}{\min(D_{c,i})} \tag{4.23}$$

The average feature value of all samples for the same label is calculated as

$$a_{c,i} = \frac{1}{M_c} \sum_{m=1}^{M_c} f_{m,c,i} \tag{4.24}$$

Out from this the average distance between (b) samples of different labels is calculated as

$$D_i^{(b)} = \sqrt{\frac{1}{C \times (C-1)} \sum_{c,l=1}^{C} (a_{l,i} - a_{c,i})^2}, \quad c \neq l \tag{4.25}$$

The variance factor between different labels is defined as

$$V_i^{(b)} = \frac{\max(|a_{l,i} - a_{c,i}|)}{\min(|a_{l,i} - a_{c,i}|)}, \quad c, l = 1, 2, \ldots, C, \quad c \neq l \tag{4.26}$$

The variance factor λ for each feature is calculated as

$$\lambda_i = \left(\frac{V_i^{(w)}}{\max(V_i^{(w)})} + \frac{V_i^{(b)}}{\max(V_i^{(b)})} \right)^{-1} \tag{4.27}$$

The ratio of $D_i^{(b)}$ and $D_i^{(w)}$ is calculated and combined with the variance factor as

$$E_i = \lambda_i \frac{D_i^{(b)}}{D_i^{(w)}} \tag{4.28}$$

These values are normalized with the maximum value to obtain the final evaluation criteria

$$\bar{E}_i = \frac{E_i}{\max(E_i)} \tag{4.29}$$

The closer the value of $\bar{E}_i$ is to 1, the better the corresponding feature is to separate the C different labels. With the introduction of a predefined threshold σ, only features fulfilling the criteria $\bar{E}_i \geq \sigma$ may be selected.

The TFSWT was done individually for each of the 5 machines in the algorithm development set during the single machine analysis. To be able to make a comparison with the other feature selection method, SFS, the 10 best features for both methods were chosen without the use of any threshold. Thus 5 feature sets with 10 features each were generated.

Each machine was split into a training and test set corresponding to 80 % respectively 20 % of the log files. All feature sets were used for classification of the test set from the five machines. The mean accuracy for each feature set was calculated and the best feature set was chosen.

In the multiple machines analysis the TFSWT was done once using all the data from the 5 machines in the algorithm development set. Thus the feature set with the 10 best features was obtained at once.

4.9.2 Sequential forward selection

As compared to the previous method which aims at finding the best individual features and then combining them, this method tries to find a combination of features that perform well. SFS is an iterative method starting with an empty feature set. In the first iteration each feature is tested individually and the one generating the highest classification accuracy is added to the empty feature set. In the second iteration all remaining features are tested one by one in combination with the feature from the selected feature set. The feature generating the highest classification accuracy in combination with the already selected feature is added to the feature set and so on. The iterations continue until either a predefined number of features are selected or the improvement between two iterations is below a certain threshold.

The SFS was done individually for each of the 5 machines in the algorithm development set during the single machine analysis. Each machine was split into the same training and test set used in the TFSWT. The test sets were saved for later use whereas the training sets were split once more using k-fold cross validation with $k = 5$. Each of these splits was used to select one feature giving the best accuracy. These 5 accuracies were compared and the feature corresponding to the highest was chosen. The SFS was repeated until 10 features were selected. Thus one feature set was created for each machine. These feature sets were then evaluated using the original test set which was not used to select the features. The mean accuracy of every feature set used on all machines was calculated and the one generating the highest accuracy was chosen.

In the multiple machines analysis the procedure was the same with the exception of the initial split which were neglected. Each machine made up its own group in the 5-fold cross validation. This resulted in a single feature set, therefore no evaluation of different feature sets was necessary.

4.10 Classification using kNN

The classification method used in this report was kNN. The test and training data were vectors of length 10, corresponding to the 10 different features selected through the methods discussed in Section 4.9. The value of k was

varied between 1 and 50. The weighting technique mentioned in Section 2.5.1.2 giving closer neighbors a larger weight was used.

When windowing was used a voting procedure was implemented to decide the classification of each log file as a unit. All windows derived from the same log file, and their classification result, were given one vote and the label receiving the most votes was chosen. If two labels received the same amount of votes the one with the smallest label number was chosen.

4.11 Evaluation

In order to test the developed algorithm on previously unused data the 5 machines from the evaluation set were used. During this final test the specific features and parameters chosen during the algorithm development step for the single respectively multiple machine analyses were used. Both analyses were conducted using k-fold cross validation with $k = 5$, thereby the training data was derived from the machines in the evaluation set.

During the evaluation it was only the training data that was used to calculate the flow compensation parameter which were then, if necessary, used to compensate both the training and test data. In a similar manner, the features from the training set were normalized first and the maximum and minimum value encountered here were saved and used during the normalization of the features in the test set as well.

5

Results

5.1 Preprocessing of signals

The final result of the preprocessing of the signal *HPG-pressure* from a 5 hour long log file is presented in Figure 5.1. The blue part is the raw data whereas the orange parts are the segments remaining after the preprocessing. The upper part shows the signal from the entire machine run, thereby it includes the steps of e.g. start-up, functional check and disinfection. The lower part is a magnification of the part of the signal when $MACHINE\_MODE$ is set to treatment, occurring approximately between 7.15 am and 11.30 am. As can be seen the first part of the treatment episode exhibits somewhat unstable signal behavior, which explains why the first 10 minutes are discarded. The tarations are easily detected as the temporary pressure drops occurring every 30 minutes. As can be seen, these are discarded as well.

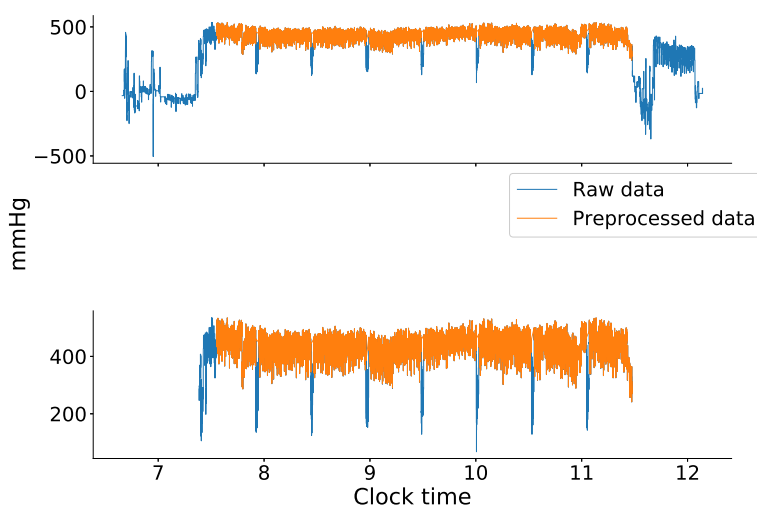


Figure 5.1: Result from the preprocessing of the signal *HPG-pressure*.

5.2 Flow compensation

The compensation of one of the flow sensitive features, *median of (HPG−PD)*, is illustrated in Figure 5.2. The blue points represent the original, uncompensated, samples whereas the orange points represent the compensated ones. The figure shows all the samples from a single machine using a time window of 30 minutes from the aforementioned feature. It is worth noting that the main flows of 500 and 700 mL/min have been used more frequently than the main flow of 600 mL/min in the current machine. This holds true for the other machines as well.

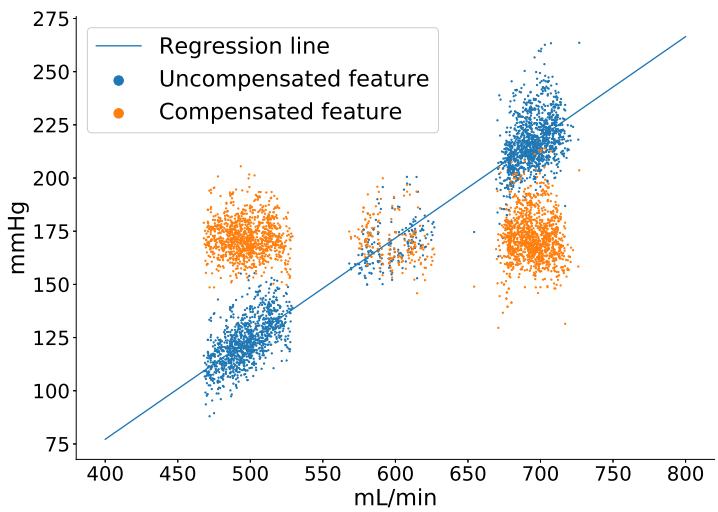


Figure 5.2: Result of the flow compensation for the feature *median of (HPG−PD)*.

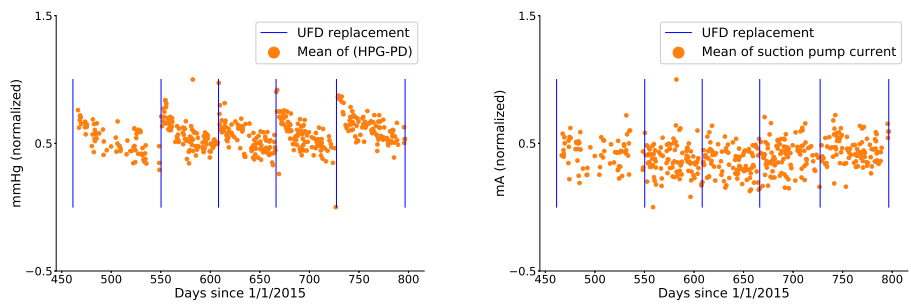
5.3 Feature selection

In order to analyze the extracted features, four different approaches were used. Two different types of plots, time series and box plots, were used to get a subjective visual assessment of the features. In addition to this the objective feature selection techniques of TFSWT and SFS were used.

5.3.1 Time series

To examine how the values of the different features varied during the analyzed time frame a time series plot was used. Since the analyzed machines were produced from 2015 and onward, the date 01/01/2015 was used as day 0. The

date and start hour for each machine run were used as x-value in these plots. The occasions for UFD replacements were marked with a blue vertical line. If a feature exhibiting a periodicity coinciding with the UFD replacements were found this would be a strong indicator that the primary aim of this report, to develop an algorithm which can be used in a service indicator for UFD replacements, would be feasible. A few of the analyzed features did exhibit such a behavior. Those were the *mean*, *max*, *RMS*, *median*, *lower quartile* and *upper quartile* of the signal *(HPG−PD)*. One of these features, *mean of (HPG−PD)*, is shown in Figure 5.3a. The majority did not exhibit any clear relationship with the UFD replacements. For comparison, one of these features, *mean of suction pump current* is shown in Figure 5.3b.



(a) *Mean of (HPG−PD).* (b) *Mean of suction pump current.*

Figure 5.3: Time series of the features *mean of (HPG−PD)* and *mean of suction pump current*. In (a) a clear relationship between the feature and the UFD replacements can be seen. The same pattern with gradually declining values which are "reset" in connection with the UFD replacements are repeated between all UFD replacements. In (b) no such relationship can be seen.

These time series plots were also used to detect other possible time dependencies that had nothing to do with UFD replacements. If such features existed they could possibly damage the developed algorithm, introducing pseudo relationships. As a matter of fact, a couple of features derived from the signal *blood leak detector* did show promising classification results in the initial phase of the algorithm development. However, when they were visualized with time series plots as seen in Figure 5.4, it became apparent that their time dependency had nothing to do with the UFD status but was instead a gradual drift over time. This gradual drift lead to lower feature values for the first labels (just after a UFD replacement) compared to the last (just before the next UFD replacement), no matter which of the UFD replacements they were situated in between.

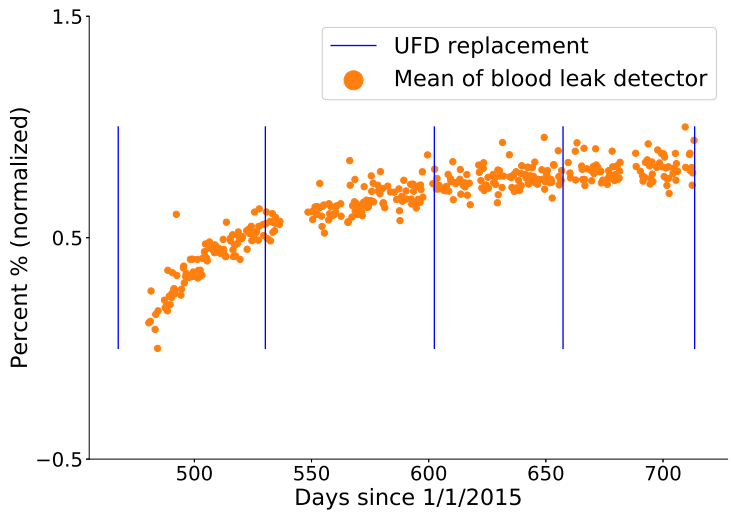


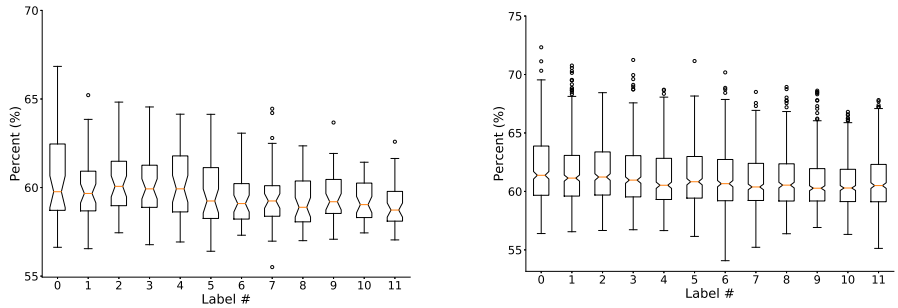
Figure 5.4: Time series of the feature *mean of blood leak detector*. No relationship with UFD replacements can be seen, instead there is a gradual drift over time. The short segment with missing data around day 540 is either due to the fact that the machine has not been used, or more probable, that these log files are overwritten since the monthly extraction is delayed.

This could be compared to a hypothetical feature which simply counted the days since last UFD replacement. Such a feature would always increase its value with higher label number. As such, it would achieve high classification results but the feature would have absolutely no connection with actual status of the UFD, hence be useless from a CBM point of view.

Due to this, the features derived from the signal *blood leak detector* could not be used and they were therefore discarded for the rest of the algorithm development.

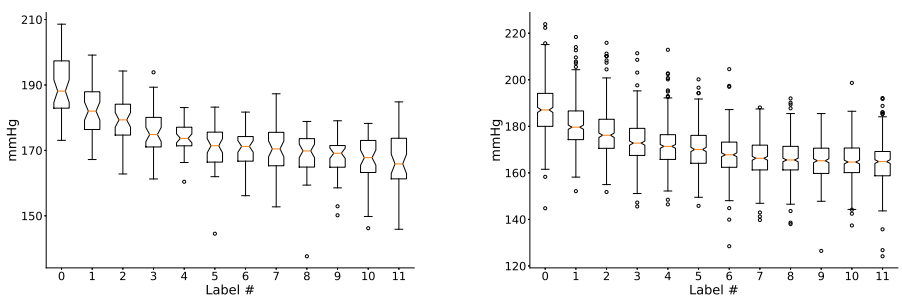
5.3.2 Box plots

In addition to the time series plots, another manual evaluation of the features was done. This was achieved by analyzing the box plots for each feature individually, visualizing a possible difference between the labels. For the majority of the features there were no obvious differences between the labels, one example of such a feature is given in Figure 5.5. The only features that did show a clear both visual and statistical significant difference between labels were the same as those mentioned in Section 5.3.1 found to exhibit a periodicity coinciding with the UFD replacements. This result was achieved both for single machines as well as when the machines were analyzed all together. An example of this is given in Figure 5.6.



(a) One machine. (b) All machines.

Figure 5.5: Box plots of the feature *mean of flow pump cycle*. No obvious relationship between feature value and label number can be seen.



(a) One machine. (b) All machines.

Figure 5.6: Box plots of the feature *mean of (HPG−PD)*. A clear relationship can be seen as a gradual decrease of feature value with label number.

5.3.3 TFSWT

The 10 features receiving the highest normalization scores, $\bar{E}_i$, for each machine in the single and multiple machine analyses are presented in Table 5.1 respectively Table 5.2. The feature set with the highest mean accuracy in the single machine analysis, Table 5.1a, was chosen.

Table 5.1: Result from the single machine TFSWT. The flow pump and suction pump are abbreviated as FP respectively SP. The other abbreviations are given in Section 4.5.

(a) Machine 1

Feature	$\bar{E}_i$
Q_3 (HPG−PD)	1.00
Q_2 (HPG−PD)	0.96
RMS (HPG−PD)	0.95
μ (HPG−PD)	0.94
Q_1 (HPG−PD)	0.88
Q_3 FP cycle	0.30
RMS FP cycle	0.30
μ FP cycle	0.30
Q_2 FP cycle	0.30
Q_1 FP cycle	0.29
Mean acc. (%)	20.3

(b) Machine 2

Feature	$\bar{E}_i$
Q_1 (HPG−PD)	1.00
Q_3 (HPG−PD)	0.73
RMS (HPG−PD)	0.28
Q_3 FP current	0.27
Q_3 FP cycle	0.26
Max (HPG−PD)	0.25
Q_3 HPG	0.25
Max HPG	0.23
RMS HPG	0.23
Min FP current	0.22
Mean acc. (%)	18.4

(c) Machine 3

Feature	$\bar{E}_i$
Q_3 (HPG−PD)	1.00
Q_2 (HPG−PD)	0.99
μ (HPG−PD)	0.75
Q_1 (HPG−PD)	0.70
RMS (HPG−PD)	0.69
CI (HPG−PD)	0.31
II (HPG−PD)	0.30
CLI (HPG−PD)	0.29
Min (HPG−PD)	0.27
Max (HPG−PD)	0.25
Mean acc. (%)	16.9

(d) Machine 4

Feature	$\bar{E}_i$
Q_3 (HPG−PD)	1.00
Q_2 (HPG−PD)	0.91
RMS (HPG−PD)	0.91
μ (HPG−PD)	0.88
Q_1 (HPG−PD)	0.86
Max (HPG−PD)	0.52
Min (HPG−PD)	0.28
CLI (HPG−PD)	0.26
II (HPG−PD)	0.26
CI (HPG−PD)	0.25
Mean acc. (%)	16.9

(e) Machine 5

Feature	$\bar{E}_i$
Q_1 (HPG−PD)	1.00
Q_3 (HPG−PD)	0.70
RMS (HPG−PD)	0.47
Max (HPG−PD)	0.46
μ (HPG−PD)	0.45
SK PD	0.36
RMS FP cycle	0.30
Q_2 SP cycle	0.30
$c_k Q_2$(HPG−PD)	0.30
μ FP cycle	0.29
Mean acc. (%)	19.5

Table 5.2: Result from the multiple machine TFSWT.

Feature	$\bar{E}_i$
RMS (HPG−PD)	1.00
μ (HPG−PD)	0.99
Q_1 (HPG−PD)	0.89
Q_2 (HPH−PD)	0.69
Max (HPG−PD)	0.37
Q_3 (HPG−PD)	0.35
RMS HPG	0.21
Max HPG	0.20
μ HPG	0.19
Q_2 HPG	0.18
Mean acc. (%)	15.4

5.3.4 SFS

The 10 selected features and the accuracy in each step during the SFS for the single and multiple machine analyses are presented in Table 5.3 respectively Table 5.4. These results are visualized in Figure 5.7. Even though the accuracies vary between the different machines an overall behaviour may be noticed which becomes especially clear when the mean (Figure 5.7a) and the result from the multiple machine analysis (Figure 5.7b) are studied. It appears as they all share the common characteristic of a gradual decreased improvement for each added feature. In most of the cases the improvement after feature $6-7$ is practically non-existing. The feature set with the highest mean accuracy in the single machine analysis, Table 5.3e, was chosen.

Table 5.3: Result from the single machine SFS. The features were added one by one resulting in a successively larger feature set. The accuracy at each step corresponds to the accuracy using the entire feature set selected up to that point. It is given in percent (%). The feature in the top of each table was the one selected first.

(a) Machine 1

Feature	Acc. (%)
Q_1 (HPG−PD)	16.5
$CCTD$ Pair 1	22.1
KU SP current	21.6
DF FP cycle	22.3
IQR HPG	22.8
CCM Pair 4	22.6
IQR (HPG−PD)	21.6
c_{1-10} FP cycle	21.4
Max (HPG−PD)	21.4
DF PD	21.9
Mean acc. (%)	16.3

(b) Machine 2

Feature	Acc. (%)
RMS (HPG−PD)	18.9
c_kIQR (HPG−PD)	21.2
SI FP cycle	22.8
DF FP cycle	23.7
IQR FP current	24.3
CCM Pair 1	24.9
Q_1 (HPG−PD)	25.1
SI SP current	24.5
KU FP cycle	24.1
KU (HPG−PD)	24.1
Mean acc. (%)	15.8

(c) Machine 3

Feature	Acc. (%)
Q_2 (HPG−PD)	19.3
c_{1-10} SP current	22.3
DF FP cycle	23.4
KU FP cycle	23.7
DF PD	23.4
KU SP current	22.9
SI SP cycle	22.6
CCM Pair 3	22.6
RMS (HPG−PD)	22.1
μ (HPG−PD)	22.3
Mean acc. (%)	16.8

(d) Machine 4

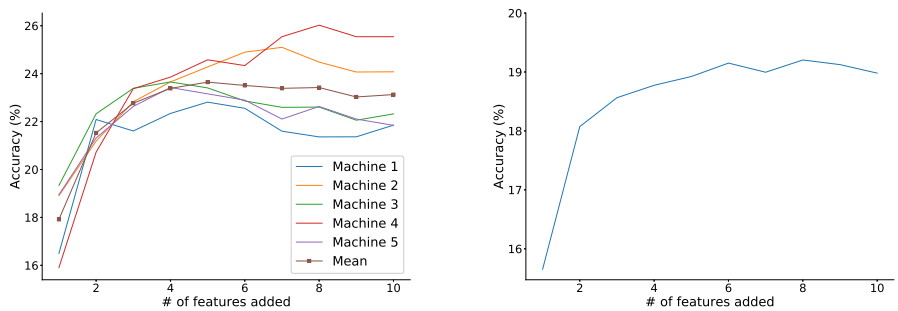
Feature	Acc. (%)
μ (HPG−PD)	15.9
σ FP cycle	20.7
DF SP cycle	23.4
DF PD	23.9
SI FP cycle	24.6
KU (HPG−PD)	24.3
KU FP cycle	25.5
DF FP cycle	26.0
CCM Pair 2	25.5
c_kIQR PD	25.5
Mean acc. (%)	18.1

(e) Machine 5

Feature	Acc. (%)
Q_3 (HPG−PD)	18.9
II FP cycle	21.3
IQR HPG	22.6
CCM Pair 3	23.4
IQR FP cycle	23.2
SI SP current	22.9
SI (HPG−PD)	22.1
CCM Pair 5	22.6
CCM Paie 1	22.1
c_kQ_1 SP cycle	21.8
Mean acc. (%)	19.2

Table 5.4: Result from the multiple machine SFS.

Feature	Acc. (%)
Q_3 (HPG−PD)	15.7
$c_k Q_3$ FP current	18.1
$c_k Q_1$ SP current	18.6
DF FP current	18.8
Q_2 (HPG−PD)	18.9
$c_k Q_2$ SP current	19.2
$c_k Q_2$ FP current	19.0
DF FP cycle	19.2
CCM Pair 2	19.1
KU FP cycle	19.0
Mean acc. (%)	19.0



(a) Single machine analysis.

(b) Multiple machine analysis.

Figure 5.7: Results from the feature selection using SFS in the single and multiple machine analyses. As can be seen for the mean in 5.7a and the single plot in 5.7b the accuracy dropped when 5 respectively 6 features had been selected.

5.3.5 Evaluation of feature selection methods

In Figure 5.8 the result from both the TFSWT and SFS are summarized. As can be seen, the results from the two different methods are hard to separate. Even though the mean accuracy slightly favors the feature set derived through TFSWT in the single machine analysis, the difference is not large enough to say that this method is preferable. However, since it did perform best the feature set derived from the TFSWT (Table 5.1a) was chosen for the single machine analysis.

In the multiple machine analysis the evaluation of the two different methods gave another result. In this case the difference between the two methods was larger and the feature set selected through SFS performed best. Thus this feature set (Table 5.4) was chosen for the multiple machine analysis.

Worth noting, an accuracy of about 20 % should be compared to the random result when using 12 labels which is $1/12 \approx 8.3$ %.

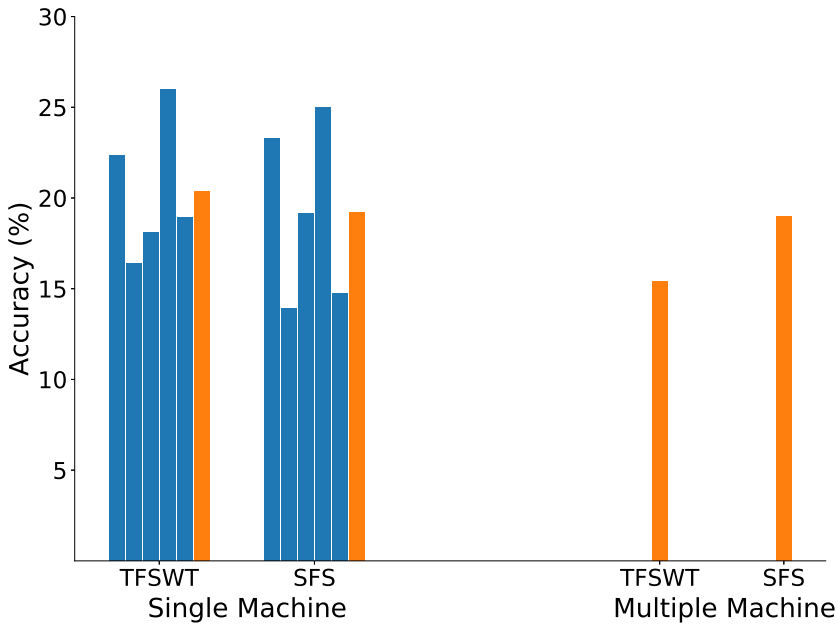


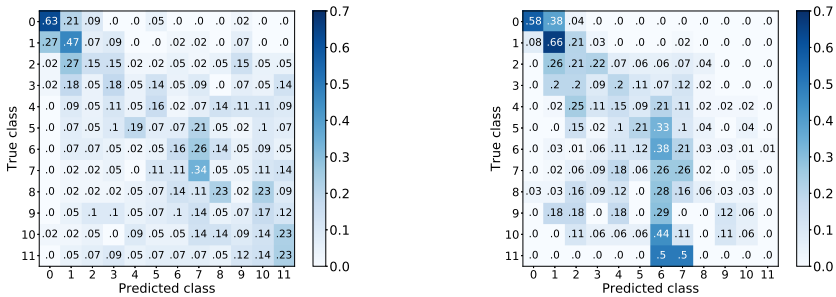
Figure 5.8: Comparison between TFSWT and SFS in the single and multiple machine analyses. The single machine part shows the result of using the selected feature set on each of the five machines (blue) and the mean of these (orange).

5.4 Labeling of data

The evaluation of the different labeling methods was performed with *number of labels* set to 12 respectively 4 in both the single and multiple machine analyses. For the first labeling method (L_1), based on time, the parameter *number of labels* can easily be adjusted. For the second method (L_2), based on hypochlorite disinfections, the value of this parameter is decided by the maximum number of hypochlorite disinfections present between any two UFD replacements for each machine. Hence this parameter is below 12 for most of the machines. When only 4 labels were used with this method the original label 0 and 1 were merged to the new label 0, original label 2 and 3 were merged to the new label 1, original label 4 and 5 were merged to the new label 2 and the remaining original labels (6+) were merged to the new label 3.

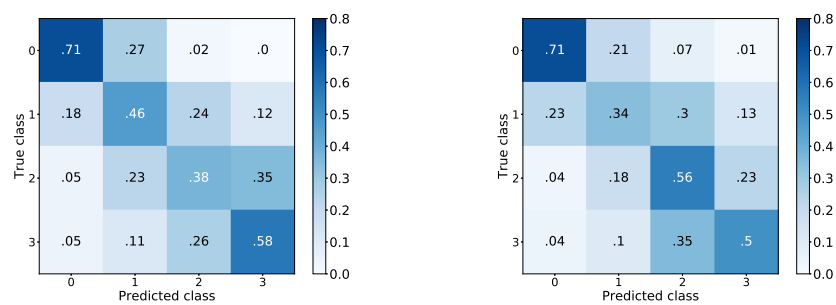
The results from the single machine analysis can be seen in Figures 5.9−5.10. For simplification only one confusion matrix corresponding to one machine is presented in each case. The desired darker diagonal mentioned in Section 2.5.1.3, indicating a good classification result, may be seen in Figure 5.10 for both labeling methods. It is not as apparent in Figure 5.9

even though a slight tendency could be seen. Especially label 0 and 1 achieves a good classification result for both labeling methods.



(a) Time based labeling. (b) Hypochlorite based labeling.

Figure 5.9: Classification results for the different labeling methods using 12 labels in the single machine analysis. The results are derived from machine 4 in Table 5.5.



(a) Time based labeling. (b) Hypochlorite based labeling.

Figure 5.10: Classification results for the different labeling methods using 4 labels in the single machine analysis. The results are derived from machine 2 in Table 5.6.

In Tables 5.5−5.6 the number of samples in each label for the two different cases and the classification accuracies are presented. As can be seen here L_1 is much more evenly sampled as compared to L_2. This is due to the somewhat irregular occurrences of the hypochlorite disinfections.

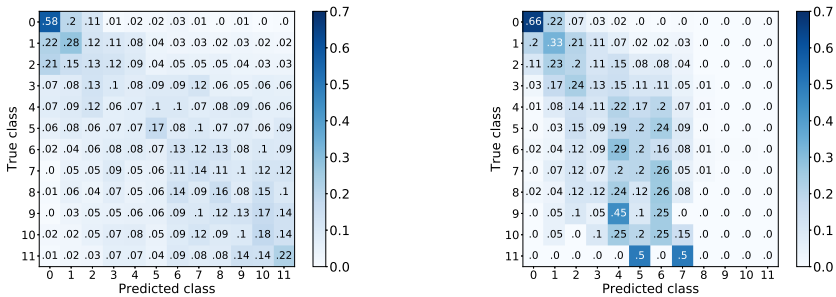
Table 5.5: Classification accuracy given in percent (%) and number of samples present in labels 0−11 for the two different labeling methods for each of the machines in the single machine analysis. For L_1 *number of labels* is set to 12 whereas it varies between 8−12 for L_2. Thereby the classification accuracies are not directly comparable between the two methods except for machine 4.

		Number of samples in each label												
		0	1	2	3	4	5	6	7	8	9	10	11	Acc. (%)
Machine 1	$\boldsymbol{L_1}$	43	43	43	43	43	43	42	43	43	43	43	43	19.4
	$\boldsymbol{L_2}$	70	82	93	67	81	59	33	15	10	3	2	-	35.7
Machine 2	$\boldsymbol{L_1}$	50	51	50	53	49	50	52	49	53	50	51	50	18.6
	$\boldsymbol{L_2}$	58	76	61	70	103	96	88	50	6	-	-	-	31.9
Machine 3	$\boldsymbol{L_1}$	38	41	38	38	40	38	38	40	38	38	41	38	20.8
	$\boldsymbol{L_2}$	49	50	66	61	84	54	53	49	-	-	-	-	31.1
Machine 4	$\boldsymbol{L_1}$	43	45	41	44	44	42	43	44	44	42	44	43	22.7
	$\boldsymbol{L_2}$	26	62	68	56	53	48	72	65	32	17	18	2	27.4
Machine 5	$\boldsymbol{L_1}$	40	39	40	39	40	41	38	40	39	39	40	40	15.2
	$\boldsymbol{L_2}$	59	52	86	45	59	49	83	40	2	-	-	-	25.7

Table 5.6: Classification accuracy given in percent (%) and number of samples present in labels 0−3 for the two different labeling methods for each of the machines in the single machine analysis. For both methods on all machines *number of labels* is set to 4.

		Number of samples in each label				
		0	1	2	3	Acc. (%)
Machine 1	$\boldsymbol{L_1}$	129	129	128	129	51.8
	$\boldsymbol{L_2}$	152	160	140	63	57.3
Machine 2	$\boldsymbol{L_1}$	151	152	154	151	53.0
	$\boldsymbol{L_2}$	134	131	199	144	53.1
Machine 3	$\boldsymbol{L_1}$	117	116	116	117	48.5
	$\boldsymbol{L_2}$	99	127	138	102	49.2
Machine 4	$\boldsymbol{L_1}$	129	130	131	129	51.6
	$\boldsymbol{L_2}$	88	124	101	206	52.6
Machine 5	$\boldsymbol{L_1}$	119	120	117	119	42.1
	$\boldsymbol{L_2}$	111	131	108	125	44.0

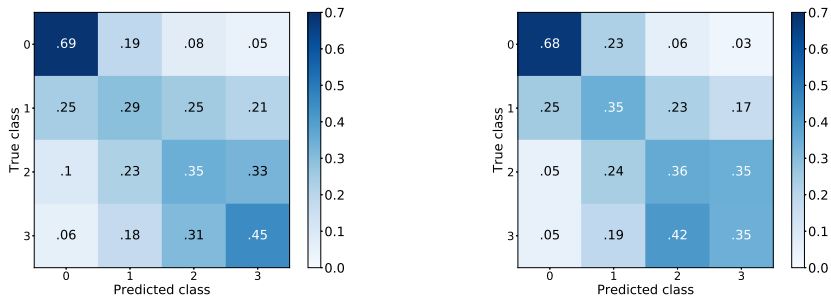
In Figures 5.11−5.12 the results from the multiple machine analysis with 12 respectively 4 labels can be seen. In Table 5.7 the corresponding number of samples in each label and the classification accuracy can be seen. Once again the L_2 method suffers from imbalance in the data with very few samples for the last 4 labels. This can readily be seen in Figure 5.11 where the last 4 columns are almost completely blank. As with the single machine analysis the diagonals are more prominent in the 4 labels case, particularly for L_1.



(a) Time based labeling.

(b) Hypochlorite based labeling.

Figure 5.11: Classification results for the different labeling methods using 12 labels in the multiple machine analysis.



(a) Time based labeling.

(b) Hypochlorite based labeling.

Figure 5.12: Classification results for the different labeling methods using 4 labels in the multiple machine analysis.

The major drawback with L_2 is the imbalanced labeling it generates. Even though most labels contain approximately the same number of samples there is often one or more labels with a significant different number of samples, either higher or lower. Such an imbalance poses a big problem since kNN is known to be sensitive to this phenomenon. This can easily be understood as it is more likely for a test point to be close to one of the 200 training points belonging to class X than any of the 5 training points belonging to class Y (if the classes are not well separated). In a multiple class problem this will in theory lead to that the classifier will have a tendency to more often assign the most frequently occurring labels in the training data to the test data.

An example of this can be seen in Figure 5.9b where a lot of test points are wrongly classified as label 6 which is seen as the colored column above label 6. It comes as no surprise that label 6 has the most samples when L_2 for machine 4 is looked upon in Table 5.5.

Table 5.7: Classification accuracy and number of samples present in labels 0−11 for the two different labeling methods using 12 respectively 4 labels in the multiple machine analysis.

		Number of samples in each label												
		0	1	2	3	4	5	6	7	8	9	10	11	Acc. (%)
12 labels	$\boldsymbol{L_1}$	214	219	212	217	216	214	213	216	217	212	219	214	19.0
	$\boldsymbol{L_2}$	262	322	374	299	380	306	329	219	50	20	20	2	23.3
4 labels	$\boldsymbol{L_1}$	645	647	646	645	-	-	-	-	-	-	-	-	44.4
	$\boldsymbol{L_2}$	584	673	686	640	-	-	-	-	-	-	-	-	42.4

Even though L_2 seems to perform best when classification accuracy is concerned this can be argued. Due to the problem mentioned above the classification accuracies for L_2 could not be completely trusted since they probably are good at classifying the few most frequently occurring labels but bad at the others. For the single machine, 12 label case they also appear better than they are since they often have fewer labels than L_1.

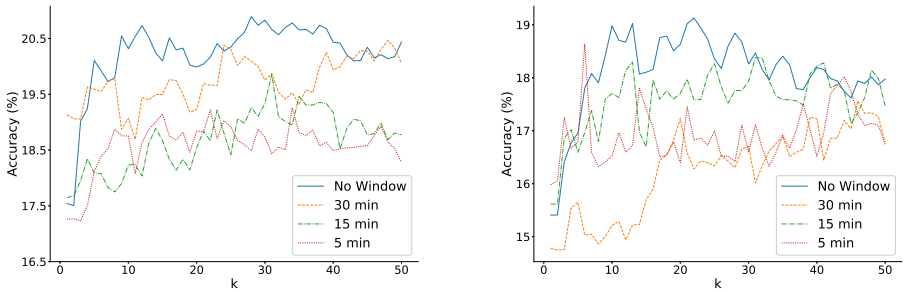
When the 4 label case is analyzed the two methods perform about the same for the single machine analysis, and for the multiple machine analysis L_1 is actually slightly better.

An advantage with L_1 is that it is much more flexible when the parameter *number of labels* is concerned and it can easily be tuned to contain more labels than what is possible with L_2. All this adds up to the decision that L_1, the time based labeling method, is the preferred method in both the single and multiple machine analyses.

5.5 Windowing and choice of k

The value of k, number of neighbors in kNN, was varied between 1−50 for each of the 4 different window lengths tested on the signals. These were 5, 15 and 30 minutes as well as when the signals were kept as a single segment, the no window option. The mean accuracy for the 5 machines in the single machine analysis is presented in Figure 5.13a for each of the tested segment lengths. The corresponding result for the multiple machine analysis is presented in Figure 5.13b.

As can be seen in Figure 5.13 the best performance in both cases is achieved with the no window option. The accuracy follows no clear overall pattern with respect to k, hence it is difficult to choose an optimal value. However, in both cases there are an overall initial increase in accuracy with increasing k before the accuracy flattens out around $k = 10$. These points, $k = 12$ and $k = 13$ were chosen as the final values for k in the single and multiple machine analyses respectively.



(a) Single machine analysis.

(b) Multiple machine analysis.

Figure 5.13: Classification accuracy for different segment lengths and different values of k in kNN.

5.6 Evaluation

The final result of using the developed algorithm on the 5 machines in the evaluation set is visualized in Figure 5.14. Here the classification accuracy for each of the 5 machines in the single machine analysis is displayed as well as their mean together with the result from the multiple machine analysis. The result is presented for three different values of the parameter *number of labels*; 3, 6 and 12. For each case the last bar displays the comparison accuracy which would be achieved by assigning each test point a label at random, assuming uniform distribution. Hence it is calculated as 1 divided by *number of labels*.

As can be seen in the figure all classification accuracies are higher than the comparison accuracy, in many cases almost twice as high. Independent of number of labels, the mean accuracy from the single machine analysis is always higher than the accuracy from the multiple machine analysis. Thereby the single machine approach is selected as the preferred one.

In Figure 5.15 three different confusion matrices derived from machine 10 where the value of *number of labels* have been altered are displayed. In all of the three matrices the classification of label 0 stands out, being correctly classified in 66 %, 57 % and 49 % of the cases when 3, 6 respectively 12 labels were used. When 12 labels are used the comparison accuracy is as low as 8.3 %, making the classification almost 6 times as good as a random guess. The other classification result using 12 labels is not as remarkable. For the 3 and 6 label cases the situation is a bit better, the diagonal is more visible, especially when 3 labels are used. Another thing worth noticing is that in all matrices the upper right and lower left corners are inhabited with low numbers. This is a good thing, meaning that the lower and higher labels are rarely mixed up, which they should not be since they represent completely different levels of the UFD degradation.

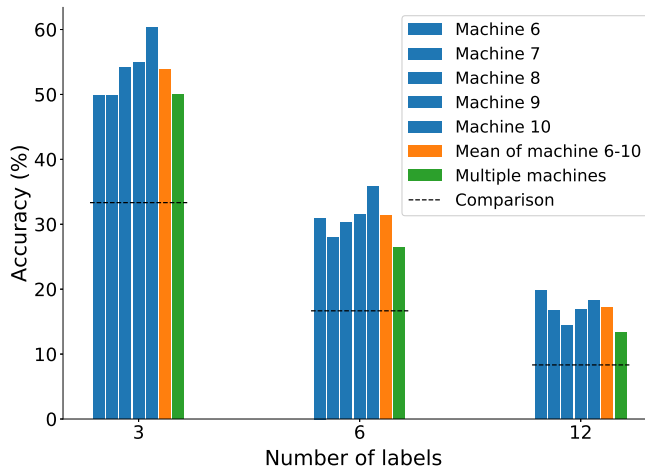
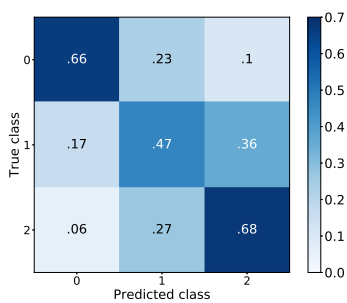
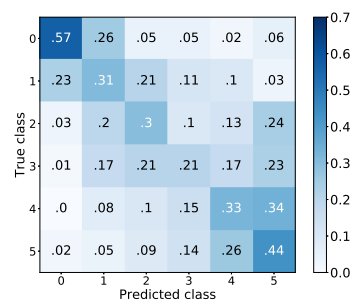


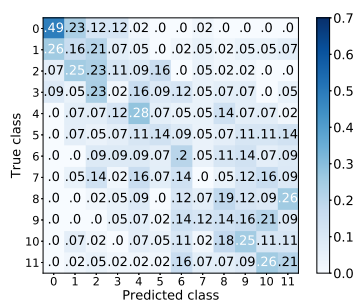
Figure 5.14: Classification accuracy for the 5 machines from the evaluation set. The results from both the single and multiple machine analyses as well as the comparison accuracies are presented.



(a) 3 labels



(b) 6 labels



(c) 12 labels

Figure 5.15: Confusion matrices derived from machine 10 in the single machine analysis with different values of *number of labels*.

To summarize, the developed algorithm uses the 10 features presented in Table 5.1a derived through the TFSWT feature selection method. The labeling method used is L_1, time based labeling. No windowing is used and number of neighbors in the kNN classifier is set to 12. The algorithm should be used on single machines, not mixing training and test data from different machines. With these settings the classifier achieves an accuracy about twice as good as a random guess.

6

Discussion

6.1 Localizing UFD replacements

The information regarding the occurrences of UFD replacements were retrieved manually from the log files. However, it is not actual replacements that are logged but rather when the operator presses the button to confirm a UFD replacement. These are two separate things. When one of the three counters indicating time for UFD replacement has reached its limit an attention is raised. It is still possible to perform a treatment but the attention will remain raised. To remove this attention the only thing the operator has to do is to press the button confirming the UFD is replaced, even if this is not done. There is currently no other way to tell whether the UFD has actually been replaced or not. This poses two different problems. The first problem is that the entire analysis in this report relies on the assumption that these confirmations correspond to actual UFD replacements. If this premise does not hold true the results in this report probably does not hold true either.

However, two things suggest that the confirmation is done in connection with a UFD replacement. The most prominent is the result from the time series which show a clear relationship between some of the features and the UFD replacements. The other is the fact that the analyzed machines are situated at a hospital believed to have high standard on the maintenance of their dialysis machines. This in turn is based on two facts. The first fact is that the clinic has accepted to send all their log files to Baxter. If they would mismanage their machines this could easily be detected. The second fact is their compliance with the recommended disinfection schedule, which was confirmed during the analysis.

To return to the two different problems with the confirmation of UFD replacement, the second problem is that clinics may easily skip these replacements. Whether this is done intentionally by the clinic to save money, or unintentionally by an operator misclicking the confirmation button, this may pose a danger to the patient. If the UFD is not replaced in time it may not fulfill its purpose of purifying the dialysis fluid through removal of a possible contamination by bacteria and endotoxins.

With respect to patient safety, a method to detect true occasions for UFD replacements may be desirable. Even though this was not the aim of the report, the results suggest that this may be a simpler task to solve. Out of the time series plot of e.g. the feature *mean of (HPG−PD)* in Figure 5.3a it is easy to detect probable occasions for UFD replacements even when they are not marked. Based on this a more sophisticated model could be developed, even though the simple visual assessment of this plot probably would do fine in many situations.

6.2 Analyzed signals

When the master's thesis was initiated our hypothesis was that there would be a gradual clogging of the UFD with time. From the vast amount of available signals from the log files only those believed to have some connection with the UFD clogging were analyzed. The signals derived from the two different pumps were believed to show an increased workload and power consumption in response to this. The *HPG* and *PD* pressures was thought to experience a gradual increase respectively decrease with respect to the clogging of the UFD situated in between them. Thereby the signal created as the difference between these two pressures were expected to increase with time. However, the analysis gave the opposite result. The pressure difference did in fact decrease with time as shown in Figures 5.3a and 5.6. The reason for this is explained in Section 6.9.2. As a consequence of this result the signal derived from the blood leak detector was added to the analysis in a later phase. This is explained in Section 6.9.2 as well.

6.3 Preprocessing of signals

Due to the uneven sampling rate of the signals they were resampled using linear interpolation. Thereby it was possible to e.g. retrieve the (HPG−PD) signal and perform FFT. However, interpolation has the downside of introducing errors in the data. With linear interpolation this error is proportional to the squared distance between adjacent points [42]. Thereby it is rather safe to use it on the signals that are frequently sampled, e.g. HPG- and PD pressure. The problem is that especially three of the analyzed signals, *suction pump cycle*, *suction pump current* and *blood leak detector*, have very few samples. In fact, sometimes their sample rate was as low as one sample per hour.

In the case of the features derived from the signal *blood leak detector* this was not a problem since they were discarded due to other reasons explained in Section 5.3.1. However, the features derived from the two different signals from the suction pump were available during the feature selection step. The

TFSWT almost never selected one of these features but they were selected more frequently by the SFS, even though they were never selected as one of the best features. Thereby they did not have much impact on the final result.

The error introduced through interpolation in the time-domain may propagate to all domains demanding a correct time structure. Thereby the features derived from the frequency-domain and through cross-correlation are affected as well.

6.4 Flow compensation

The implementation of the flow compensation was important for the developed algorithm. Although not all extracted features were in need of flow compensation, each and every one of the ten features selected for the final solution were flow compensated. If no compensation would be done these features would hold information regarding both the UFD status and the main flow. Since it is only the UFD status that is of interest, using uncompensated features would blur the result. As a comparison, the mean accuracy using 12 labels in Figure 5.14 is 17.2 % when flow compensation is used, as compared to 14.1 % without flow compensation.

The flow compensation used in this report compensates the features. Another possibility would be to instead compensate the signals before the features are extracted. The main reason for choosing to compensate the features and not the signals was simply that the need of flow compensation was discovered once the features were already extracted. Therefore the implementation could be done faster if these were compensated and not the signals. If the signals would be compensated directly they would first have to be resampled together with the signal *UF channel 1*. If a high enough sampling frequency to avoid aliasing would be used (100 Hz) this would result in several hundred millions of data samples for each signal on every machine and it would not be feasible to perform a flow compensation on such a large amount of data. For comparison, when the features are compensated the data is reduced by a factor 10^6.

Independent of whether it is the signals or features that are compensated there is always a risk with manipulating data in this way, information may be lost or added. However, when the features are plotted against main flow as in figure 5.2 it appears as they have a clear linear relationship. Therefore the procedure of flow compensation could be considered safe to use in these cases.

A simplification that has been done during the flow compensation is to equalize the flow in *UF channel 1* and *2*. This is not totally true since the flow in *UF channel 2* contains the addition of the ultrafiltrate from the patient. The signals derived from the fluid path succeeding the dialyzer should

therefore preferably be compensated with *UF channel 2* instead. The simplification of only using *UF channel 1* is done for two reasons. Firstly, two of the affected signals, *PD-pressure* and *blood leak detector*, were not in need of any flow compensation at all. Secondly, the contribution of the ultrafiltrate is small compared to the volume of the dialysis fluid, mostly in the order of 1−2 %.

6.5 Feature selection

The two different feature selection methods used in this report, TFSWT and SFS, have one fundamental difference worth emphasizing. The TFSWT analyzes all the features one by one, giving each feature a normalization score between 0 and 1. The SFS on the other hand selects a set of features which performs well together. However, it is a greedy algorithm by its nature since it selects the best feature, then the best feature in combination with the first and so on. The combination of features it selects is thereby not necessarily the very best. But still, choosing 10 features which perform well together should logically give a better classification result than choosing the 10 features performing best individually. The reason for this is that several features may be related to each other, describing more or less the same characteristic of the data that is analyzed. If one such feature is selected, adding another of the related features barely improves the result since practically no new information is given to the classifier.

With this taken into consideration a couple of interesting things may be noted from the feature selection step. The most important result is that both methods always select some of features derived from the $(HPG-PD)$ signal as the first feature. Hence this signal should be the one best describing the degradation of the UFD. This result is in consistency with the analysis of the boxplots as well as the time series of each individual feature.

Another interesting result is the features selected in position 2−10, which show large differences between the two methods. The TFSWT has a strong tendency of selecting features derived from the $(HPG-PD)$ signal, with some additions of features derived mainly from the *FP cycle* and *HPG* signals (Tables 5.1 and 5.2). The features selected are almost exclusively from the time-domain. The SFS on the other hand exhibits much more variation in the feature selection (Tables 5.3 and 5.4). All signals are present and from these, features derived from both the time-domain, frequency-domain and through cross-correlation are selected. No clear pattern can be found. This result suggests two things. Firstly, this apparently random selection pattern indicates that the features selected on position 2−10 by the SFS contains little information regarding the degradation of the UFD. Secondly the result indicates that there is a strong relationship among some of the features derived

from the ($HPG-PD$) signal. The reason for this is that even though many of these features received high normalization scores through the TFSWT, usually only one or maybe two of them is present among the 10 features selected by the SFS. This result is reasonable since features such as the mean, RMS and the different quartiles should be closely related.

All this in combination with the ambiguous result in Figure 5.8 regarding which feature selection method that gives the best feature set, points in the same direction. It is only some of the time-domain based features from the signal ($HPG-PD$) that contain relevant information regarding the status of the UFD. Besides, these features are closely related. Thereby, as long as one of these features is present in the feature set it does not make any big difference which the other ones are.

The problem with interpolation described in Section 6.3 makes the features derived from the frequency-domain and through cross-correlation unreliable. Since there was no success in introducing these features they were dismissed on an early stage to instead focus on more promising aspects of the algorithm. Their non-success could either be due to a total lack of information regarding the UFD status or be a consequence of errors introduced through interpolation. Therefore a further analysis needs to be done before features derived from the frequency-domain and through cross-correlation could be said to be of interest or not when analyzing the UFD status.

6.6 Labeling of data

6.6.1 Effects of missing data

Both of the labeling methods analyzed in this report rely on a correct notation of the UFD replacements. As described in Section 6.1 all the notations are believed to correspond to true UFD replacements. However, it is still possible that some replacements are missed. If the monthly extractions of the .zip files from the dialysis machines are delayed and data is overwritten there is a risk that information regarding a UFD replacement is lost. Such episodes of missing data have actually been found on several of the analyzed machines, one example can be seen in Figure 5.4. These episodes vary in length, ranging from a couple of days up to over a month. During these episodes it is impossible to determine by certainty if and when a UFD replacement have taken place. But through analysis of the time series of e.g. the feature *mean of (HPG−PD)* together with the duration between the different UFD replacements it is possible to detect pattern deviations where one can suspect that a UFD replacement is missed.

The analysis in this report is based solely on the confirmed UFD replacements. Thereby it is probable that some replacements are missed. This would result in an incorrect labeling since the labeling is not reset properly. The

possible error introduced this way propagates through the succeeding analysis all the way to the final classification. Since overwritten data is relatively commonly occurring it poses a great problem to the analysis in this report, affecting the final evaluation accuracy achieved by the algorithm.

The labeling method based on hypochlorite disinfections is even more sensitive to missing data. Since these disinfections occur with such short intervals it is almost certain that they are missed during longer episodes with missing data. Even a short episode of just a day or two has a rather high risk of leading to errors in the labeling.

It is not only information regarding UFD replacements and hypochlorite disinfections that are missed. A large amount of treatments are also lost, causing an offset in the labeling. This affects the result even further.

6.6.2 Compensation of unbalanced data

As mentioned in Section 5.4 the labeling method based on hypochlorite disinfections generates unbalanced data, which is a big problem when using kNN. For simplification, the analysis in this report has handled the data as being balanced. However, different ways of dealing with this have been discussed but due to the limited time they have not been implemented. A review on different techniques available are presented in [43]. These include simple approaches as over- and undersampling as well as more sophisticated methods on algorithm level. In oversampling new samples of the minority class are created either synthetically or through replication of existing samples. Undersampling is the opposite where samples of the majority class are randomly eliminated. Examples from the article on the algorithm level are to assign different weights for the classes in a kNN classifier or to implement a cost function with different costs for different misclassifications.

Since the analyzed data for some machines have as few as two samples for some labels it might be difficult to compensate this through over- or undersampling. A compensation on the algorithm level might have greater success. It could even be necessary to merge the minority labels before performing additional compensation.

6.6.3 Other alternatives

None of the tested labeling methods result in a totally consistent labeling. Between filter replacement x and y, label z could correspond to a UFD that has been used between e.g. 50−60 days, whereas it could correspond to a usage of e.g. 58−69 days between filter replacements (x+1) and (y+1). The number of days (and number of hypochlorite disinfections) between two filter replacements affect the labeling. However, since the analyzed dialysis machines have been used in a routinely fashion there are mostly only small variations in e.g. the number of days between UFD replacements and the interval between

hypochlorite disinfections. Thereby the result of using these labeling methods is fairly consistent after all.

There are however methods that would give an even more consistent labeling. One example would be to use the actual usage time of each machine and label the data after strict limits. An even better approach would be to base the labeling on the total volume of dialysis fluid that has passed through the UFD, thereby different main flows would be taken into consideration as well. This could be further developed and include all the different machine modes, not only treatment, and use different weights depending on how demanding the different usages are on the UFD. The complexity of the labeling could be immensely increased.

To be able to analyze other aspects of the algorithm as well, a rather simplified labeling method was finally chosen. Apart from the fairly consistent labeling it generates, its biggest strength is the almost completely balanced data it produces. With this said, a future improvement of the algorithm that probably would have a reasonable impact on its performance would be to base the labeling on the volume of dialysis fluid passing the UFD.

6.7 Windowing and choice of k

With the introduction of windowing the sample size increased substantially. For the five minute time window the sample size increased 40-fold with respect to when each treatment episode was kept as a single sample. The optimal size of the feature set is largely unaffected by the sample size when using kNN [44]. Thereby the same feature set could be used for all window lengths.

The parameter k on the other hand is affected by the sample size. It should also be chosen with consideration taken to the noise level of the data where noisier data is better handled with a larger k [40].

The parameter *number of labels* naturally affects the number of samples in each label. Thus it is possible that the value of k should be chosen differently when *number of labels* is adjusted in the future.

6.8 Algorithm development

The main focus during this master's thesis has been to develop the algorithm, not to optimize it performance-wise. Thereby a couple of simplifications have been made along the way. In addition to this, the algorithm has not been implemented in a strict logical and chronological order. The need of some steps have been thought of in a later stage and they have therefore been implemented there and then, even though the algorithm would have been faster to execute if the code was rewritten. One example of this is the ten minute cutoff used in the preprocessing of each signal. As for now this step

is one of the latest, but it would be more natural to do it before the linear interpolation.

Some other simplifications worth mentioning is the discarding of the log files containing UFD replacements and the ones succeeding the last filter exchange. Since the UFD replacements occurred once every 1−3 month and there were usually three separate treatment episodes each day it is only a small fraction of the log files that are discarded this way. In fact, it was only about half of these log files that did contain a treatment episode. Thereby the fraction of lost data is even smaller. Since the amount of available log files was never a problem this simplification could be accepted.

The reason to discard the log files succeeding the last filter exchange of each machine was due to the problem it arose with the time based labeling. However, the labeling method based on hypochlorite disinfections could easily use these log files. Still they were not used, the reasons are twofold. Firstly the use of these log files would make the data even more imbalanced since it would mainly add samples of the first labels. Secondly it would make the comparison between the labeling methods harder since they would not be based on the same data.

Apart from these simplifications a couple of more could be mentioned but they are only interesting from a code optimization point of view. Since they have no relation to the performance of the algorithm they are not discussed here. However, if someone else in the future would continue the work from this point it could be relevant to know that improvements can be done in this area.

6.9 Evaluation and final thoughts

6.9.1 Single versus multiple machine analysis

During the evaluation the final result of the single versus multiple machine analyses was received, slightly in favor of the single machine analysis. This indicates that the developed algorithm contains machine specific parameters to some extent. There are a couple of possible explanations to this but it all narrows down to the signals derived from the log files. The analyzed machines are obviously not identical but rather exhibit small differences in e.g. their sensor offsets or wear of different components. These differences propagate and could manifest themselves in different parts of the algorithm when the two cases are compared.

One explanation could be that a gradual decrease or increase in feature value with label number is drowned by different sensor offsets in the machines from the multiple machine analysis. This would result in a feature which is interpreted as bad, whereas the single machine analysis would detect the true potential of the feature. However, this does not seem to be the case when the

top selected features are studied.

Another possibility is that the machines react differently to a varied main flow. This would favor the individual flow compensation performed in the single machine analysis. Whether this holds true for the current algorithm is hard to tell based on the results presented but is a possible explanation.

Most likely the fact that the training and test data points are derived from the same machine is the best explanation for the better result with the single machine analysis.

If the developed algorithm based on single machine analysis would be implemented on a brand new machine this would require a couple of months of usage before the algorithm could be trained and a result could be achieved. However, if the multiple machine analysis was to be used instead a classification result could be achieved already for the very first log file. After all, the result when using the multiple machine analysis is almost as good as for the single machine analysis, see Figure 5.14.

6.9.2 Hypothesis regarding the UFD degradation

An interesting result from the evaluation is the high classification accuracy of the very first label, label 0. This result stands out and is achieved independent of the value of *number of labels*. An explanation for this may be found in the boxplots of the most relevant features, see Figure 5.6. Here the largest pressure drop is seen between the first and second label. This pattern was seen among all the six features mentioned in Section 5.3.1 on all machines. The difference could sometimes be even more prominent than in the figure. This indicates that the speed of the degradation of the UFD right after the replacement is high and then declines.

Another very interesting result was the development of the pressure drop across the UFD measured as $(HPG-PD)$. As stated previously our hypothesis was that this would increase with the clogging of the UFD. In reality the pressure drop decreased with time.

These two findings led to the new hypothesis that the UFD was not getting clogged but rather damaged, and the reason for this were the hypochlorite disinfections. This would explain both findings. Since these disinfections are known to severely reduce the lifetime of the UFD it is reasonable to think that the first disinfection would have a large impact. This in combination with the weekly scheduling of these disinfections would result in the first disinfection occurring somewhere around label 0 and 1 which could thereby explain the large difference between these labels. The hypochlorite has the effect of damaging the filter mass by increasing the diameter of the pores. Thereof the decreased pressure drop.

It was this hypothesis that led to both the analysis of the blood leak detector and the implementation of the labeling method based on hypochlorite disinfections in the later stages of the algorithm development. The idea be-

hind the blood leak detector was that the increased diameter of the pores in the UFD would lead to lower filtering capacity, letting more contaminants through, making the dialysis fluid less transparent which could then be detected by the sensitive sensor. Despite its name it was therefore not intended to detect blood. However, the result was not satisfying. Instead a slow drift of the sensor or maybe the wear of some component caused a slowly increased steady-state of the signal, making it inappropriate to use in the algorithm as previously mentioned.

6.10 Ethics

Even though the log files analyzed in this master's thesis are derived from actual treatments of real patients no personal information is available in the log files. Most of the data available in the log files is strictly machine related even though some data is derived directly from, or has a close connection to, the patient e.g. venous blood pressure or ultrafiltered volume. However, this data cannot be used to identify the patient. Besides, only machine related data is used in this report. In addition to this, the analyzed data has been approved for this usage.

The purpose of CBM is to maximize the use of the analyzed component. As compared to preventative maintenance this reduces the number of unnecessary replacements, hence it leads to reduced environmental impact. Another maintenance approach is corrective maintenance where a component is not replaced until it fails. When a UFD is concerned this would not be acceptable since it could endanger the health of the patient. If both ethics regarding sustainability and patient's safety is taken into consideration a CBM approach of UFD replacements would be ideal.

6.11 Future work

This report has shown that the possibility to develop a CBM based service indicator for UFD replacements looks promising. However, a few steps remain. The algorithm presented in this report may detect different degradation levels of the UFD and decide the current level to some extent. Basically it is only based on a few closely related features derived from the same signal. Therefore the most important task in the future development of this algorithm will be to find more signals and/or features containing relevant information regarding the UFD status. Alternatively the information contained in the already found relevant signal could be enhanced. This could possibly be done either through an effective noise reduction or through the implementation of a high, constant sampling frequency of the signals in the AK 98 dialysis machine.

An implementation of a constant sampling frequency may have many positive consequences. It could make the analysis of signals currently sampled at a very low frequency more rewarding and eliminate the need of interpolation. Maybe more important, it would give the analysis of the features derived from the frequency-domain and through cross-correlation much more meaning. A high, constant sampling frequency could simply open a wide range of possibilities. The big problem with this is the limited internal memory of the AK 98. The data collection would thereby have to be done in a different way. A better solution to this would also benefit future studies of this kind.

To get a more fair estimation of the performance of the algorithm it could be tested on data that is not incomplete, or where the missing data is compensated in some way. With such data available it would be interesting to rerun the steps where the different parameter settings were selected as well.

Possible future improvements of the algorithm also include things mentioned previously in the discussion, e.g. compensation of the imbalanced data in the labeling method based on hypochlorite disinfections and the optimization of the algorithm.

An alternative option for the future work could be to develop the algorithm that could detect true occasions for UFD replacements presented in Section 6.1 if there is a need.

As mentioned in the aim, Section 1.1, the development of the current algorithm is just the first out of three steps in the implementation of the CBM service indicator. Before those steps are done nothing can really be said about the RUL of the UFD no matter how sophisticated the algorithm is. With this said a final clarification can be made. This report does not aim at finding an optimal setting for the parameter *number of labels*. What this parameter really does is an adjustment of the time resolution of the algorithm, i.e. if each label should correspond to one day, week or month in the final CBM service indicator. Hence the value of this parameter should be decided later on when the potential of the final service indicator is fully known.

7

Conclusion

This master's thesis resulted in an algorithm based on the machine learning technique kNN with the aim of classifying the degradation level of the UFD filter in Baxter's AK 98 dialysis machine. It was concluded that the analyzed signals derived from the log files of real dialysis treatments were in need of preprocessing in order to segment out the relevant parts. Some of the analyzed signals were affected by the main flow of the dialysis fluid, therefore a flow compensation was needed. The two feature selection methods TFSWT and SFS were used and from their result the conclusion was drawn that out of all the 178 available features only 6 did exhibit a clear relationship with the UFD degradation level. It could further be concluded that they were all closely related and derived from the same signal which gave the pressure drop across the UFD filter. These features were mean, max, RMS, median, lower quartile and upper quartile.

Out of the two tested labeling methods, time based labeling was chosen as the best even though the other method gave promising result but needed further development. It was tested to split the signals into smaller segments of various lengths, but the original length was in the end concluded the best. The number of neighbors in kNN was evaluated and set to 12 and 13 in the single and multiple machine analysis respectively. The single machine analysis performed the best and it was concluded that some machine specific characteristics were present but not so prominent. These setting resulted in a final classification accuracy of 53.8 %, 31.3 %, and 17.2 % when 3, 6 respectively 12 labels were used. Thereby the classification is roughly twice as good as a random guess.

During the development of the algorithm it was discovered that any of the six best features could be used with great certainty to determine when UFD replacements had taken place. This finding could be further developed if there is an interest in implementing such a feature into the AK 98 to be used e.g. if the UFD replacement notification is mistakenly removed.

The final algorithm did show promising result in detecting the UFD degradation level but further development will be needed. As far as the algorithm is concerned, the main focus should be to improve the signal quality and/or

find more relevant signals and/or features. However, the algorithm is just the first step. A study regarding the long term degradation of the UFD as well as an analysis concerning the UFD function in different degradation levels are necessary to be performed before a CBM based service indicator of the UFD could be implemented in the AK 98.

Bibliography

[1] T. Liyanage, T. Ninomiya, V. Jha, B. Neal, H. M. Patrice, I. Okpechi, M.-h. Zhao, J. Lv, A. X. Garg, J. Knight, A. Rodgers, M. Gallagher, S. Kotwal, A. Cass, and V. Perkovic, "Worldwide access to treatment for end-stage kidney disease: a systematic review," *The Lancet*, vol. 385, no. 9981, pp. 1975–1982, 2015.

[2] W. G. Couser, G. Remuzzi, S. Mendis, and M. Tonelli, "The contribution of chronic kidney disease to the global burden of major noncommunicable diseases," *Kidney international*, vol. 80, no. 12, pp. 1258–1270, 2011.

[3] Y. Peng, M. Dong, and M. J. Zuo, "Current status of machine prognostics in condition-based maintenance: a review," *The International Journal of Advanced Manufacturing Technology*, vol. 50, no. 1, pp. 297–313, 2010.

[4] S. Butler, *Prognostic algorithms for condition monitoring and remaining useful life estimation.* PhD thesis, National University of Ireland, Maynooth, 2012.

[5] Baxter International Inc., *AK 98 Dialysis Machine - Operator's Manual*, May 2016.

[6] G. Giebisch and E. Windhager, *Medical Physiology: A Cellular and Molecular approach*, ch. Organization of the Urinary System, pp. 749–766. Saunders Elsevier, 2009.

[7] G. Giebisch and E. Windhager, *Medical Physiology: A Cellular and Molecular approach*, ch. Glomerular Filtration and Renal Blood Flow, pp. 767–781. Saunders Elsevier, 2009.

[8] Gambro Lundia AB (Baxter International Inc.), *Internal education material*, 2009.

[9] L. S. Costanzo, *Physiology.* Saunders Elsevier, Philadelphia, PA, 2006.

[10] G. Giebisch and E. Windhager, *Medical Physiology: A Cellular and Molecular approach*, ch. Transport of Sodium and Chloride, pp. 782–796. Saunders Elsevier, 2009.

[11] G. Giebisch and E. Windhager, *Medical Physiology: A Cellular and Molecular approach*, ch. Transport of Urea, Glucose, Phosphate, Calcium, Magnesium, and Organic Solutes, pp. 797–820. Saunders Elsevier, 2009.

[12] G. Giebisch and E. Windhager, *Medical Physiology: A Cellular and Molecular approach*, ch. Transport of Potassium, pp. 821–834. Saunders Elsevier, 2009.

[13] G. Nyberg and A. Jönsson, *Njursjukvård*. Studentlitteratur, 2004.

[14] C. M. Kjellstrand and B. P. Teehan, *Replacement of Renal Function by Dialysis*, ch. Acute Renal Failure, pp. 822–862. Kluwer Academic Publishers, 1996.

[15] A. S. Levey, J. Coresh, E. Balk, A. T. Kausz, A. Levin, M. W. Steffes, R. J. Hogg, R. D. Perrone, J. Lau, and G. Eknoyan, "National kidney foundation practice guidelines for chronic kidney disease: Evaluation, classification, and stratification," *Annals of Internal Medicine*, vol. 139, no. 2, pp. 137–147, 2003.

[16] A. Santoro, C. Canova, A. Freyrie, and E. Mancini, "Vascular access for hemodialysis.," *Journal of Nephrology*, vol. 19, no. 3, pp. 259–264, 2006.

[17] R. P. Uldall, *Dialysis Therapy*, ch. Vascular Access for Hemodialysis, pp. 5–22. Hanley & Belfus, 1992.

[18] B. J. Canaud and C. M. Mion, *Replacement of Renal Function by Dialysis*, ch. Water Treatment for Contemporary Hemodialysis, pp. 231–255. Kluwer Academic Publishers, 1996.

[19] R. A. Ward, *Dialysis Therapy*, ch. Mechanlical Aspects of Dialysis, pp. 31–64. Hanley & Belfus, 1992.

[20] L. W. Henderson, *Replacement of Renal Function by Dialysis*, ch. Biophysics of Ultrafiltration and Hemofiltration, pp. 114–145. Kluwer Academic Publishers, 1996.

[21] Baxter International Inc., *AK 98 Dialysis Machine - Service Manual*, 10 ed., 2016.

[22] F. Camci, *Process monitoring, diagnostics and prognostics using support vector machines and hidden Markov models*. PhD thesis, Wayne State University, Detroit, 2005.

[23] A. K. Jardine, D. Lin, and D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical systems and signal processing*, vol. 20, no. 7, pp. 1483–1510, 2006.

[24] K. L. Tsui, N. Chen, Q. Zhou, Y. Hai, and W. Wang, "Prognostics and health management: A review on data driven approaches," *Mathematical Problems in Engineering*, vol. 2015, 2015.

[25] X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation–a review on the statistical data driven approaches," *European journal of operational research*, vol. 213, no. 1, pp. 1–14, 2011.

[26] D. I. Cho and M. Parlar, "A survey of maintenance models for multi-unit systems," *European Journal of Operational Research*, vol. 51, no. 1, pp. 1–23, 1991.

[27] R. Reinertsen, "Residual life of technical systems; diagnosis, prediction and life extension," *Reliability Engineering & System Safety*, vol. 54, no. 1, pp. 23–34, 1996.

[28] P. A. Scarf, "On the application of mathematical models in maintenance," *European Journal of Operational Research*, vol. 99, no. 3, pp. 493–506, 1997.

[29] H. Wang, "A survey of maintenance policies of deteriorating systems," *European Journal of Operational Research*, vol. 139, no. 3, pp. 469–489, 2002.

[30] E. Zio, "Reliability engineering: Old problems and new challenges," *Reliability Engineering & System Safety*, vol. 94, no. 2, pp. 125–141, 2009.

[31] R. Kothamasu, S. H. Huang, and W. H. VerDuin, "System health monitoring and prognostics–a review of current paradigms and practices," in *Handbook of maintenance management and engineering*, pp. 337–362, Springer, 2009.

[32] A. Heng, S. Zhang, A. C. Tan, and J. Mathew, "Rotating machinery prognostics: State of the art, challenges and opportunities," *Mechanical Systems and Signal Processing*, vol. 23, no. 3, pp. 724–739, 2009.

[33] N. Gorjian, L. Ma, M. Mittinty, P. Yarlagadda, and Y. Sun, "A review on degradation models in reliability analysis," *Engineering Asset Lifecycle Management*, pp. 369–384, Sept. 2009.

[34] N. Gorjian, L. Ma, M. Mittinty, P. Yarlagadda, and Y. Sun, "A review on reliability models with covariates," *Engineering Asset Lifecycle Management*, pp. 385–397, Sept. 2009.

[35] J. Van Noortwijk, "A survey of the application of gamma processes in maintenance," *Reliability Engineering & System Safety*, vol. 94, no. 1, pp. 2–21, 2009.

[36] J. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mechanical Systems and Signal Processing*, vol. 25, no. 5, pp. 1803–1836, 2011.

[37] Y. Lei and M. J. Zuo, "Gear crack level identification based on weighted k nearest neighbor classification algorithm," *Mechanical Systems and Signal Processing*, vol. 23, no. 5, pp. 1535–1547, 2009.

[38] J. Yan, M. Koc, and J. Lee, "A prognostic algorithm for machine performance assessment and its application," *Production Planning & Control*, vol. 15, no. 8, pp. 796–801, 2004.

[39] H.-E. Kim, A. C. Tan, J. Mathew, E. Y. Kim, and B.-K. Choi, "Machine prognostics based on health state estimation using svm," *Proceedings Third World Congress on Engineering Asset Management and Intelligent Maintenance Systems Conference*, vol. 199, pp. 834–845, 2008.

[40] C. M. Bishop, *Pattern Recognition and Machine Learning.* Springer, 2006.

[41] T. Hastie, R. Tibshirani, and J. Friedman, *Elements of Statistical Learning.* Springer, second edition ed., 2009.

[42] S. L. Eshkabilov, *MATLAB®/Simulink® Essentials: MATLAB®/Simulink® for Engineering Problem Solving and Numerical Analysis*, ch. Experimental Data Analysis, pp. 560–648. Lulu Publishing Services, 2016.

[43] S. Kotsiantis, D. Kanellopoulos, and P. Pintelas, "Handling imbalanced datasets: A review," *GESTS International Transactions on Computer Science and Engineering*, vol. 30, no. 1, pp. 25–36, 2006.

[44] J. Hua, Z. Xiong, J. Lowey, E. Suh, and E. R. Dougherty, "Optimal number of features as a function of sample size for various classification rules," *Bioinformatics*, vol. 21, no. 8, pp. 1509–1515, 2004.