TECHNISCHE UNIVERSITÄT WIEN
VIENNA UNIVERSITY OF TECHNOLOGY

DISSERTATION

Decision support for therapy optimization of heart failure patients based on retrospective home telemonitoring data and expert knowledge

ausgeführt zum Zwecke der Erlangung des akademischen Grades einer Doktorin der Naturwissenschaften unter der Leitung von

Ao.Univ.Prof. Dipl.-Ing.
Dr.rer.nat. Dr.techn. Dr.scient.med. Frank Rattay
Institutsnummer: E101
Institut für Analysis und Scientific Computing

eingereicht an der Technischen Universität Wien
Fakultät für Mathematik und Geoinformation

von

Dipl.-Math. Marija Vuković
geb. am 14.06.1978
Matr. Nr.: 1028880
Viktor Kaplan Strasse 6-8/209
1220 Wien

Wien, im Februar 2014

Abstract

Telemonitoring, as a medical practice, includes measurements of physiological parameters and data transfer of remotely, usually home-located patients. Providing reliable home-telemonitoring based decision support system for therapy of chronic heart failure patients is important to ensure patients' wellbeing, stable health status, enhanced self-care, medical literacy, as well as care givers' satisfaction, productivity and effectiveness in decision making. Furthermore, home-telemonitoring can reduce health care costs and positively affect the economy. Potentially important influences on patient health status include weather conditions, and yet these influences are insufficiently investigated in the current telemonitoring studies.

Currently used techniques to establish relationship between home-telemonitoring and chronic heart failure patients' health status cannot efficiently deal with real time health condition changes of multiple simultaneously monitored physiological parameters, nor provide reliable predictions of possible adverse events. The current systems typically result in large number of false alarms (up to 99.4%), causing care-givers' dissatisfaction to the extent of decreasing quality of medical care, for which reason alarm hazards have been identified as the top health device hazards in recent years.

The research goal was development of a decision support methodology for therapy optimization of heart failure patients based on retrospective home telemonitoring data and expert knowledge. For this purpose, AIT - Austrian Institute of Technology, Department of Safety and Security, Business Unit Information Management, provided 9128 retrospective home telemonitoring records of 65 chronic heart failure patients in Austria - MOBITEL (Mobile telemonitoring in heart failure patients) study.

The dissertation develops innovative patient home telemonitoring data analyses tools (5-level alarm flags, reference state estimation, measurement colour coding, automated alarm generation based on dynamic threshold adjustments) for enhanced interpretation of patients' physiological measurements. Thus, with the application of alarm flag levels, the percentage of true alarms is increased from 8.5% (flag level 1) to 19.2% (flag level 5) as compared to the average of 12.6% true alarms when such a procedure is not applied in the reference MOBITEL study. Most of the alarms occur due to the exceeded systolic or diastolic blood pressure thresholds.

The conducted statistical data analyses establish relationship between measurements of physiological parameters and care givers' therapy decisions. The results show increased accuracy (0.788 vs. 0.782) and specificity (0.789 vs. 0.777), but decreased sensitivity (0.770 vs. 0.942) of the developed automated alarm management system based on optimal dynamically adjustable thresholds, in comparison to the original MOBITEL home-telemonitoring records including manually adjusted alarm thresholds. The system is intended to support the physicians in setting up patient specific alarm thresholds, subject to individual health conditions (e.g. tolerate large number of false alarms preferring not to miss many true alarms if the patient status could be critical; or maximally reduce the number of false alarms even if some true alarms are omitted, if the patients are unlikely to experience adverse events). Such support is achieved through a mathematical model developed to predict the sensitivity (true alarm occurrence) for an arbitrary adjustment

of patient telemonitoring alarm generation thresholds. Another model is then used to calculate specificity (true non-alarm occurrence). The physicians can choose particular alarm threshold values and immediately see the expected impact of such a selection on sensitivity and specificity of the decision support system. The models are obtained using principle component and linear regression and validated (F-values: 315 and 1067 (> 100), respectively) on 52 heart failure patients.

Furthermore, influences of severe weather conditions on heart failure patients' physiological parameters are predicted thanks to the discovered statistically significant ($p < 0.05$) relations with cold and heat stress temperatures. Especially falling temperatures, cold stress days, with mean temperature difference thresholds between 6.4°C and 6.8°C showed statistically significant influences on rising blood pressures (95%CI: (-16, -1) and (-8, 0) mmHg, for systolic and diastolic blood pressure differences, respectively). Although such cases have one end of the 95% confidence intervals comparable to effects of certain blood pressure medications ($\sim$ 10 mmHg), the other end includes trivial values (close to zero) causing the results to remain inconclusive.

The developed methodology and tools could be used in the future home telemonitoring systems to optimize care-givers' decision support management and mitigate possible adverse events of chronic heart failure patients through timely distinction of relevant indications of worsening patient conditions. Applying the developed methodology the physicians would have an enhanced possibility to effectively adjust their therapeutic assistance, reducing health care costs and potentially avoiding ambulatory interventions or hospitalizations.

Last but not least, the results encourage researchers to initiate applications of automated alarm management algorithms, as innovative data interpretation tools for home telemonitoring systems.

Kurzfaßung

Telemonitoring, umfasst in der medizinischen Praxis das Messen physiologischer Parameter und die Übermittlung von Patientendaten von zu Hause aus. Die Bereitstellung zuverlässlicher Telemonitoring-Systeme für die Therapie von Patienten mit chronischer Herzinsuffizienz ist wichtig, um das Wohl der Patienten, den stabilen Gesundheitszustand, die verbesserte Selbstversorgung, die medizinische Kompetenz, sowie die Zufriedenheit der Betreuer und die Produktivität und Effektivität der Entscheidungsfindung zu gewährleisten. Weiters kann Telemonitoring die Gesundheitskosten senken und sich positiv auf die Wirtschaft auswirken. Wichtige eventuelle Einflüsse auf den Gesundheitszustand der Patienten beinhalten auch die Wetterlage, die bisher noch unzureichenend in den aktuellen Telemonitoring-Studien erforscht wurde.

Die derzeit verwendeten Techniken zur Erfassung des Gesundheitszustands von Patienten mit chronischer Herzinsuffizienz per Telemonitoring können weder in Echtzeit mit Zustandsänderungen von mehrfachen, simultan aufgezeichneten Parametern umgehen, noch zuverlässige Vorhersagen über mögliche ungünstige Ereignisse sicherstellen. Typischerweise resultiert das bei aktuellen Systemen in einer Vielzahl von Fehlalarmen (bis zu 99,4%) und verursacht deshalb bei den Pflegekräften starke Frustration, die sich auch in sinkender Qualität der ärztlichen Betreuung äußert. Gefahren bedingt durch Fehlalarme haben sich deshalb in letzter Zeit zu den wichtigsten gesundheitlichen Gefahrenquellen dieser Patientengruppen entwickelt.

Das Ziel der Forschung war es, eine Methodik zur Unterstützung von Entscheidungen zur Optimierung der Therapie von Herzinsuffizienzpatienten zu entwickeln, die auf der retrospektiven Analyse von Daten aus dem Heim-Telemonitoring und auf Expertenwissen basiert. Für diesen Zweck hat das AIT (Austrian Institute of Technology, Department Safety and Security, Geschäftsfeld Information Management) 9128 retrospektive Heim-Telemonitoring-Aufzeichnungen von 65 Patienten mit chronischer Herzinsuffizienz in Österreich bereitgestellt - MOBITEL (Mobile telemonitoring in heart failure patients) Studie.

Die Dissertation entwickelt ein innovatives Datenanalysetool für Heim-Telemonitoring-Patienten (5-Level Alarmfahnen, Referenzzustandsschätzung, Messfarbcodierung, automatisierte Alarmerstellung basierend auf dynamischen Grenzwertanpassungen) das eine verbesserte Interpretation der physiologischen Messungen durch Patienten gewährleisten soll. In der Folge ist mit dem Einsatz von Alarmfahnen der Prozentanteil von gültigen Alarmen von 8,5% (Fahnenlevel 1) auf 19,2% (Fahnenlevel 5) gestiegen, verglichen mit dem Mittel von 12,6% gültiger Alarme bei herkömnmlicher Auswertung der MOBITEL-Studie. Die meisten Alarme entstehen durch die Überschreitungen der Grenzwerte von systolischem und/oder diastolischem Druck.

Die durchgeführte statistischen Datenanalyse stellt die Verbindung zwischen den Messungen physiologischer Parameter und den Therapieentscheidungen der ärztlichen Betreuer her. Das Resultat zeigt im Vergleich zu den ursprünglichen MOBITEL Telemonitoring-Aufzeichnungen (welche manuell angepasse Alarmgrenzwerte beinhalten) erhöhte Genauigkeit (0,788 vs. 0,782) und Spezifität (0,789 vs. 0,777), aber eine reduzierte Sensitivität (0,770 vs. 0,942) des entwickelten automatisierten Alarm-Management-

Systems basierend auf optimalen dynamisch angepassten Schwellwerten. Das System beabsichtigt die Ärzte bei der Anpassung spezifischer Alarmgrenzwerte auf individuelle Gesundheitszustände zu untertützen (z.B. die Vielzahl falscher Alarme zu tolerieren, um möglichst keine echten Alarme mit kritischen Zuständen des Patienten zu verpassen; oder die maximale Reduktion der falschen Alarme bei Patienten mit geringen Risiken, auch wenn einige echte Alarme übersehen werden). Solche Unterstützung wird durch ein mathematisches Modell erreicht, das die Sensitivität (das Eintreten von echten Alarmen) aus den eingestellten Grenzwerten abschätzt. Ein anderes Modell wird benützt, um die Spezifität (das Eintreten von Fehlalarmen) zu berechnen. Die Ärzte können bestimmte Alarmgrenzwerte wählen und sehen sofort die erwarteten Auswirkungen auf die Sensitivität und Spezifität des Systems. Die entsprechenden Modelle wurden durch den Einsatz von Hauptkomponentenregression erstellt und validiert (F-Werte: 315 bzw. 1067 (> 100)) an 52 Patienten mit Herzinsuffizienz.

Weiters werden die Einflüsse extremer Wetterbedingungen auf die physiologischen Parameter von Herzinsuffizienzpatienten aus Hitze- und Kältebelastungen vorhergesagt. Besonders stark fallende Temperaturen (mittlere Temperaturdifferenzen zwischen 6.4 Grad Celsius und 6.8 Grad Celsius) zeigten statistisch signifikante ($p < 0.05$) Einflüsse auf den Blutdruck (95%CI: (-16, -1) und (-8, 0) mmHg, für systolische bzw. diastolische Blutdruckunterschiede). Obwohl das 95%-Konfidenzintervall dieses Effektes den von bestimmten Blutdruckmedikamenten ($\sim$ 10 mmHg) erreicht, umfasst es auch den Nullwert, wodurch die Ergebnisse nicht schlüssig sind.

Die entwickelte Methodik und die Tools können in den zukünftigen Heim-Telemonitoring-Systemen benützt werden, um das Management der Unterstützung der ärztlichen Entscheidung zu optimieren und um die möglichen Nebenwirkungen von Patienten mit chronischer Herzinsuffizienz durch zeitgerechte Erkennung relevanter Anzeichen einer Verschlechterung des Patientenzustandes zu mildern. Durch die Anwendung der entwickelten Methodik können die Ärzte die therapeutische Hilfe effektiver anpassen, und damit die Gesundheitskosten durch die potenzielle Vermeidung ambulanter Eingriffe oder Krankenhausaufenthalte reduzieren.

Nicht zuletzt sollten die Resultate Forscher ermutigen, automatisierte Alarm-Management-Algorithmen als innovatives Dateninterpretationstool für Heim-Telemonitoringsysteme einzusetzen.

Acknowledgements

I would like to express my profound gratitude to my academic advisor Prof Dr. Frank Rattay, for his exceptional supervision, support and guidance during the dissertation work, especially for his broad understanding of Computational Neuroscience and Biomedical Engineering. It was a great pleasure and privilege working with him. My gratitude additionally goes to Prof Dr. Hans J Lohninger, my second adviser, for his readiness to help with inspiring ideas and valuable advices in the most challenging times, as well as for his great course lectures in Biostatistics.

Especially, I would like to thank my husband, Vladimir, for giving me the ultimate support and strength and with whom any mission impossible became possible. At last, but not least, I would like to thank my family and friends in Serbia, for encouraging my efforts, and offering tremendous support.

Finally, I am grateful to the Austrian Institute of Technology, Department Safety and Security, business unit Information Management for their advices, technical and financial support.

Contents

Chapter 1

Introduction

1.1 Motivation

According to the world health organization data approximately 35 million people worldwide died due to chronic diseases in 2005, which presented 60% of the overall mortality (WHO, 2005). Furthermore, while mortality from the other causes was projected to decrease, mortality of people affected by the chronic diseases was projected to increase by 17% by the year of 2015. Consequently, chronic disease share within the overall mortality would rise from 60% to 64% within the 10 year long period. The impacts of chronic diseases such as heart disease, stroke, cancer, chronic respiratory diseases and diabetes are very significant and resolute efforts are required to alleviate adverse outcomes.

Economic impacts of chronic diseases are also very large. For example, cost estimates for a 10 year long period between 2005 and 2015 amount to 558 billion dollars for China (WHO, 2005), 80 billion dollars per year for Canada (Tran et al., 2008), while the US annual costs only due to heart failure are estimated at 33.2 billion dollars (Finkelstein et al., 2010).

Although both, developed and developing countries are affected, about 80% of chronic disease mortality happens in low and middle income countries, which typically have unavailable, unaffordable and low quality health care. Main causes of chronic disease mortality are increased blood pressure (20%), results of tobacco use (14%), raised cholesterol levels (12%), and obesity or overweight (7%) due to unhealthy diet and physical inactivity (WHO, 2005).

1.2 Chronic cardiovascular diseases

Particularly significant chronic diseases are cardiovascular, being the cause of approximately half of the overall chronic disease mortality in 2005 (WHO, 2005). The most costly cardiovascular illness in the US is heart failure (HF), affecting over 5 million Americans (Finkelstein et al., 2010). In Europe, the number is even larger, estimated between 6.5 and 10 million people (Maric et al., 2009). Such numbers are projected to increase in the future and chronic heart failure (CHF) may soon reach epidemic proportions (Tendera, 2005).

Common cardiovascular conditions include (WHF, 2013):

- Rheumatic heart disease – heart valve damage due to rheumatic fever, caused by streptococcal bacterial infection.
- Valvular heart disease – a disease of the heart valves, e.g. narrowing (stenosis), leakage (regurgitation or insufficiency), improper closure (prolapse). Occurs by birth, or due to valve damages by infections, certain medications, radiation.
- Aneurysm – weakness of blood vessel walls. Can occur as a consequence of high blood pressure or weak spots in blood vessels throughout the body.
- Atherosclerosis – thickness and stiffness of arteries due to fatty deposits (plaques). Causes reduction of blood flow throughout the body: coronary (in the heart arteries), peripheral (in the legs). Develops over a period of time, can lead to heart attack or stroke.
- High blood pressure (hypertension) – excessive blood force on the blood vessels. Can cause other cardiovascular diseases, e.g. stroke, heart failure.
- Peripheral arterial disease – manifests as pain in the legs while walking, and eases by rest. Occurs as a consequence of atherosclerosis, shrinkage or blockage of the heart blood supply vessels.
- Angina – chest pain resulting from reduced blood supply to the heart (ischemia). Occurs as a consequence of atherosclerosis, shrinkage or blockage of the heart blood supply vessels. Symptoms include shortness of breath and sweating, relate to physical work and ease by rest.
- Coronary artery disease (ischemic heart disease) – one of the most common heart diseases and the main cause of heart attacks and angina. Occurs as a consequence of atherosclerosis, shrinkage or blockage of the heart blood supply vessels.
- Coronary heart disease – a disease of heart arteries. Can result in angina, heart attack and heart failure.
- Heart attack (myocardial infraction) – blockage of blood flow to the heart, causing heart muscle damage due to oxygen deprivation, or fatal outcome if blood flow is not promptly reestablished. Symptoms include strong central chest pain, extreme shortness of breath, sweating, faintness.
- Sudden death – a consequence of sudden loss of heart functions. Can occur due to a heart attack or severe abnormality in the heart rhythm.
- Cerebral vascular disease – shrinkage or blockage of blood vessels leading to the brain, caused by atherosclerosis. Can result in strokes and transient ischemic attacks.

- Stroke – interruption of blood supply to the brain. Can result in permanent brain damage due to oxygen deprivation, manifesting as weakness, paralysis, speech or vision problems.

- Transient ischemic attacks – temporary blockage of blood flow resulting in loss of blood in the brain. As a consequence, brain function suddenly changes which can manifest as temporary weakness, numbness, blindness, double vision, confusion, loss of balance, severe headache. Although the symptoms do not last long and permanent damage is unlikely, these conditions can indicate increased risk of stroke.

- Cardiomyopathy – a disease of the heart muscle, e.g. enlargement (dilated cardiomyopathy), thickness (hypertrophic cardiomyopathy), muscle loss (ischemic cardiomyopathy). Occurs due to genetic predispositions, infections, or other causes.

- Pericardial disease – a condition affecting the sac in which the heart is located, called pericardium. Can include inflammation (pericarditis), fluid accumulation (pericardial effusion) and stiffness (constructive pericarditis).

- Congenital heart disease – a broad term used to describe structural abnormalities in the heart, e.g. with valves, chambers etc. Occurs by birth due to genetic predispositions or exposure to adverse influences, e.g. medications, alcohol, etc.

- Heart failure – a chronic condition resulting in reduced effectiveness of heart operation and insufficient amounts of blood flow and oxygen levels. Occurs as a result of heart damage which weakens the heart muscle, particularly affecting senior population. Symptoms include shortness of breadth, reduced ability to exercise, swelling of the ankles.

Depending upon severity of experienced symptoms, cardiac disease patients are commonly classified in accordance with the New York Heart Association (NYHA) Functional Classification. Table 1.1 presents overview of patient functional capacities in different NYHA classes (AHA, 2013a; NYHA, 1994).

NYHA Class	Patient's reaction during physical activity
I	No limitation in ordinary physical activity despite of disease.
II	Comfortable at rest, but experience slight limitation in ordinary physical activity, due to fatigue, palpitation, dyspnea or anginal pain.
III	Comfortable at rest, but experience limitation in even less than ordinary physical activity, due to fatigue, palpitation, dyspnea or anginal pain.
IV	Disease symptoms may be present even at rest, with increasing intensity if conducting any physical activity.

Table 1.1: New York Heart Association (NYHA) functional classification of patients with cardiovascular diseases

Commonly used heart medication types including brand names, designated treatment conditions and health effects are presented in Table 1.2.

CARDIAC MEDICATIONS		
Medication type	**Reasons for medication intake**	**Medication effects**
ANTICOAGULANTS: Dalteparin (Fragmin), Danaparoid (Orgaran), Enoxaparin (Lovenox), Heparin (various), Tinzaparin (Innohep), Warfarin (Coumadin)	- Help in prevention of clotting in the blood vessels. - May prevent enlargment of the clots and development of more serious problems. - Prevent the first or recurrent strokes.	Although called blood thinners, they don't really thin the blood. Decrease the ability of clotting (coagulating) the blood, but do not dissolve existing clots. Treat certain blood vessel, heart and lung conditions.
ANTIPLATELET AGENTS: Aspirin, Ticlopidine, Clopidogrel (Plavix®), Dipyridamole	- Help in prevention of clotting in patients who experienced: heart attack, unstable angina, ischemic strokes, TIA (transient ischemic attacks, or "little strokes") and other cardiovascular diseases. - Preventive medication when plaque build-up is obvious but not yet clogging the artery.	Prevent blood platelets from sticking together and forming the blood clots.
ANGIOTENSIN-CONVERTING ENZYME (ACE) INHIBITORS: Benazepril (Lotensin), Captopril (Capoten), Enalapril (Vasotec), Fosinopril (Monopril), Lisinopril (Prinivil, Zestril), Moexipril (Univasc), Perindopril (Aceon), Quinapril (Accupril), Ramipril (Altace), Trandolapril (Mavik)	- Treat cardiovascular conditions together with heart failure and high blood pressure.	Decrease blood flow resistance by lowering levels of angiotensin II. Expand blood vessels enabling easier blood flow and heart work.
ANGIOTENSIN II RECEPTOR BLOCKERS (OR INHIBITORS): Candesartan (Atacand), Eprosartan (Teveten), Irbesartan (Avapro), Losartan (Cozaar), Telmisartan (Micardis), Valsartan (Diovan)	- Treat cardiovascular conditions together with heart failure and high blood pressure.	Prevent angiotensin II from affecting the blood vessels and heart in order to prevent blood pressure from rising.
VASODILATORS: Isosorbide dinitrate (Isordil), Nesiritide (Natrecor), Hydralazine (Apresoline), Nitrates, Minoxidil	- Relieve chest pain (angina).	Increase the blood and oxygen supply to the heart and ease the heart's workload by relaxing the blood vessels.

CARDIAC MEDICATIONS		
Medication type	**Reasons for medication intake**	**Medication effects**
BETA BLOCKERS: Acebutolol (Sectral), Atenolol (Tenormin), Betaxolol (Kerlone), Bisoprolol/hydrochlorothiazide (Ziac), Bisoprolol (Zebeta), Carteolol (Cartrol), Metoprolol (Lopressor, Toprol XL), Nadolol (Corgard), Propranolol (Inderal), Sotalol (Betapace), Timolol (Blocadren)	- Help in lowering blood pressure. - Supplement therapy for cardiac arrhythmias (abnormal heart rhythms) and treat chest pain (angina). - Prevent potential heart attacks in patients who experienced them before.	Decrease the heart rate, lowering heart efforts, and resulting in lower blood pressure.
CALCIUM CHANNEL BLOCKERS: Amlodipine (Norvasc, Lotrel), Bepridil (Vascor), Diltiazem (Cardizem, Tiazac), Felodipine (Plendil), Nifedipine (Adalat, Procardia), Nimodipine (Nimotop), Nisoldipine (Sular), Verapamil (Calan, Isoptin, Verelan)	- Treat high blood pressure and chest pain (angina) due to reduced blood supply to the heart muscle and some arrhythmias (abnormal heart rhythms).	Reduce the movement of calcium into the cells of the heart and blood vessels. Help to relax blood vessels and lower the heart efforts.
DIURETICS: Amiloride (Midamor), Bumetanide (Bumex), Chlorothiazide (Diuril), Chlorthalidone (Hygroton), Furosemide (Lasix), Hydro -chlorothiazide (Esidrix, Hydrodiuril), Indapamide (Lozol), Spironolactone (Aldactone)	- Help lowering blood pressure. - Help reduce swelling (edema) developed from excess build-up of fluids in the body.	Remove excess fluid and soduim from the body through urination. Ease the heart's workload. Decrease accumulation of fluids in the lungs, ankles, legs and other body parts.
DIGITALIS PREPARATIONS: Lanoxin	- Relieve heart failure symptoms, particularly when the patient does not respond to diuretics and ACE inhibitors. - Slow certain types of iregular heartbeat (arrhythmias), mainly atrial fibrillation.	Strengthen heart contractions, benefiting in case of heart failure and irregular heart beats.
STATINS: Statins, resins, nicotinic acid (niacin), gemfibrozil, clofibrate	- Help lower low-density lipoprotein (LDL, or "bad" cholesterol), raise high-density lipoprotein (HDL, or "good" cholesterol) and lower triglyceride levels.	Lower blood cholesterol levels in various ways: affecting the liver, other intestines and disrupting the development of cholesterol.

Table 1.2: Cardiovascular medication types, common names, treatment and effects

In addition to treating cardiovascular patients with medications, the following electrical heart stimulation methods are applied in some cases: spinal cord stimulation (SCS), transcutaneous electrical nerve stimulation (TENS), epidural spinal electrical stimulation (ESES), implantable electrical nerve stimulation (IENS) – pacemaker, implantable cardioverter defibrillator (ICD), electromyostimulation (EMS).

Spinal cord stimulation (SCS) is a therapy applied to refractory angina pectoris (a part of stable angina) patients in which the other treatments have not been successful. The results of SCS reported in literature include significant reduction of ischemic burden, duration and number (frequency) of ischemic episodes (De Jongste et al., 1994; Eriksson et al., 1979).

Transcutaneous electrical nerve stimulation (TENS) has shown as a potentially beneficial treatment of ischemic pain (Tallis et al., 1983) and angina (Mannheimer et al., 1985).

Epidural spinal electrical stimulation (ESES) can improve peripheral blood circulation, alleviating vascular disease of extremities which can cause ischemic pain, skin ulceration or gangrene (Augustinsson et al., 1985). ESES shows promising results in the cases when reconstructive surgery of severe limb ischemia fails or is not possible.

Implantable electrical nerve stimulator (IENS) – pacemaker – is a small battery-powered device placed in patients' chest and used to interrupt or prevent life threatening arrhythmias by providing low-energy electrical pulses (Obel and Bourgeois, 1993). During the arrhythmias heart rate could be too fast (tachycardia), too slow (bradycardia) or irregular, resulting in insufficient blood flow leading to fatigue, shortness of breadth, or fainting.

Implantable cardioverter defibrillator (ICD) is used in prevention of sudden death and cardiac arrest of patients with ventricular tachycardia or fibrillation (AHA, 2013b). This is achieved by using high-energy electrical pulses/shocks which restore normal heart rhythm. Newer ICD devices may also serve as pacemakers (NIH, 2013). In such a case, the device would have a dual role of (1) stimulating heart beats if the heart rate is too slow, using low-energy electrical pulses, and (2) treat life threatening arrhythmias, using high-energy electrical stimulation.

Low-frequency electromyostimulation (EMS) is applied on skeletal muscles to ease physical efforts in chronic heart failure patients (Casillas et al., 2008). As a result of EMS therapy muscles' oxidative capacity and physical performance is improved, similar to the outcomes of physical training. EMS can be an alternative to physical training when such activities would present risk for the patients heart condition. EMS does not have any reported adverse side effects, but can not be applied if IENS or ICD are implanted. EMS is a cheap procedure with beneficial results in rehabilitation of a growing number of CHF patients.

Furthermore, health literacy could highly benefit the heart failure patients. Knowledge of common heart failure diseases, different treatments, health diets, exercise recommendation and active monitoring are important factors of self-care and self-management.

Due to the high and rising prevalence of cardiovascular diseases, the current study focuses on enabling technologies to address some of the present health care issues in providing effective and affordable treatment. Particularly, telemonitoring and tele-homecare offer possibilities to increase the effectiveness of usual health services while at the same time reducing the costs.

1.3 Dissertation overview

Evidence of telemonitoring benefits for the treatment of chronic diseases is presented within the Literature review chapter, particularly focusing on heart failure patients. The chapter also describes some of the current shortcomings and limitations of the existing studies which are further addressed within the dissertation. Methodology for decision support for heart failure patients therapy optimization is developed based on retrospective home telemonitoring data and expert knowledge. The methodology contains definition of the research hypotheses as well as the description of the procedures and tools used within the dissertation. The subsequent chapter deals with Descriptive Statistics and Data Preprocessing, describing the experimental research used to collect and clean the data for further analyses. Alarm Management and Optimization results are presented next, introducing new automated home telemonitoring alarm generation system. This chapter also includes analyses of weather influences on heart failure patient health conditions. The final chapters discuss the results and present conclusions, summary of scientific contributions and outlook for the future work.

Chapter 2

Literature Review

2.1 Telemedicine

Telemedicine is a practice of remotely monitoring health status of patients located at different locations from their health care providers. Typically the patients are located at their homes and use mobile phone technology to transmit the monitoring data. Home telecare medicine is one of the fastest growing areas of healthcare (Koch, 2006). Effectiveness of telemedicine has been thoroughly investigated. As early as 2005, researchers started exploring feasibility of telemedicine and possibilities to use telemonitoring for decision support in order to identify economic and health benefits, satisfaction of patients and care-givers with the telemonitoring systems. Following the feasibility studies, numerous trials were conducted and increasing number of reviews appeared to summarize the findings of conducted studies with respect to the benefits for the patients as well as the society.

A review of 31 existing reviews showed telehealth to be feasible and as effective as in-person care for patients with neurological symptoms, while reductions of hospital admissions and mortality were reported for heart failure patients (Deshpande, 2008). Also identified are improvements of communication with health care providers, quality of disease monitoring and patient quality of life. Another review of 80 published reviews on telemedicine found that application of telemedical tools is mostly effective in cardiovascular and other chronic diseases (Ekeland et al., 2010). Apart from the aforementioned reviews identifying extensive use of telemonitoring for chronic illnesses, also documented are successful applications in the long term psoriasis therapy (Hayn et al., 2009), chronic obstructive pulmonary disease (Basilakis et al., 2010), monitoring asthma (Finkelstein and Hripcsak, 2001), and assisting patients with type one diabetes mellitus (Kollmann et al., 2007). In fact, telemonitoring of heart failure and diabetes was the most common topic among the 98 home telecare studies each including 80 or more patients, published before 2006 (Barlow et al., 2007). However, a review found that application of telemedicine to patients with cardiac diseases is more effective then in the cases of diabetes (Paré et al., 2007).

Out of the cardiovascular diseases, heart failure is the most costly for the health care system and affecting the major part of population (Finkelstein et al., 2010). Therefore, confirmation of telemedicine effectiveness in cases of heart failure patients is particularly significant and documented by the systematic reviews pointing out decreased hospitalization rates, reduced usage of health services and improved patient health outcomes (Barlow et al., 2007; Maric et al., 2009). For example, a study found statistically significant financial and health related benefits of the home-based telemonitoring system resulting in 36% reduction of readmission rates, 31% less episodes of hemodynamic instability, as well as 35% decrease in mean hospital costs (Giordano et al., 2009). Significantly reduced median duration of hospital stay from 10 to 6.5 days ($p = 0.04$) was confirmed by the Mobile Telemonitoring in Heart Failure Patients (MOBITEL) study of 65 patients, conducted in Austria from 2003 to 2008 (Scherr et al., 2009). The study also reported median New York Heart Association (NYHA) class improvement from III to II in the patient group using the telemonitoring. Other reports comparing telemonitoring assisted care with usual care emphasize statistically significant reduction in patient mortality (Cleland et al., 2005; Ferrante et al., 2010).

The economic benefits of home telehealth systems have been confirmed in reviews of studies published prior to 2008, with 91% of the identified publications reporting cost effectiveness (Polisena et al., 2009; Rojas and Gagnon, 2008). However, the methodologies used to calculate costs, as well as the reported economic impacts largely vary among the published studies. The reported savings range between 1.6% and 68.3% (mean 35%, IQR 17.7-55.3%) (Seto, 2008). Additionally, a study pointed out that telemonitoring increased nurses productivity by allowing them to take care of more patients in comparison to the usual care (Weintraub et al., 2010).

Such effects together with the increased compliance of patients with the treatment programs also contribute to the care-givers satisfaction with telemonitoring (Domingo et al., 2012). Apart from the care-givers, positive patient satisfaction is also typically reported in the heart failure telemonitoring programs (Finkelstein et al., 2010), including compliance with telemonitoring (Paré et al., 2007), and/or improvement in the patient quality of life (Domingo et al., 2011; Maric et al., 2009). Typical patient responses indicated satisfaction with the telemonitoring systems due to increased knowledge via automatically provided life style recommendations as well as increased sense of security the patients experienced while staying in contact with their health-care providers using the telemonitoring solutions. Consequently, heart failure patients valued telemedicine and a study found that more than half of the 126 enrolled participants were even willing to pay 20 dollars to have the possibility of using telemedicine devices at home instead of traveling to the physicians (Bradford et al., 2004).

The extent to which the patients appreciated telemedicine and technological advances in this area enabled a new form of patient health-care: the self-care process, instigating adherence to medications, proper diet, exercise, and regular monitoring of physiological conditions (Bui and Fonarow, 2012). According to the American Heart Association, self-care implies application of a decision-making process the patients use in order to maintain physiological stability (Riegel et al., 2009). Similarly, patient self-management implies self-adjustment of the medication therapy, requiring the patients to recognize a change in their health condition (e.g. increasing fluid accumulation and swelling), assess their symptoms, choose to take action, apply a treatment plan (e.g. taking an extra diuretic dose), and evaluate the response to such therapy. As such, self-care and self-management are primarily the responsibility of the patients, rather than the health-care professionals.

Numerous challenges were present in self-care applications. Although daily weight measurements were found to be an important part of HF self-management (Chaudhry et al., 2007), less than one-half of the patients recovering from CHF regularly monitored their weight, even among those recently discharged from the hospital (Moser et al., 2005). Furthermore, an increase in weight of 2 kg over one to 3 days prior to the measurement day had a sensitivity of only 9% in detection of worsening HF symptoms (Lewin et al., 2005). The patients might delay seeking medical care for HF symptoms or fail to mention newly occurring symptoms during periodic examinations by their physicians (Evangelista et al., 2000). Presence of additional diseases often exacerbated the challenges in self-care, limiting the patients understanding of their treatment plans. Such additional disorders or diseases could require extra medications and adjustments of the treatment plans already requiring the heart failure patients to take 9-12 pills each day (Lien et al., 2002). Certain diseases might have similar symptoms to heart failure, such as shortness of breath which also occurs in chronic obstructive pulmonary disease, whereas the other diseases, e.g. diabetes, might impair the occurrence of certain HF symptoms and/or make their interpretation more difficult (Riegel et al., 2009). Additional obstacles that frequently prevent heart failure patients from fully implementing self-management include poor health literacy, social isolation or depression (Bui and Fonarow, 2012). Although self-management has been strongly recommended, this approach has not been well investigated in randomized clinical trials and there is insufficient evidence that heart failure patients benefit from self-management (Bui and Fonarow, 2012). Nevertheless, possible benefits could emerge from the combination of self-management and other interventions, such as telemonitoring.

Table 2.1 presents an overview of existing studies particularly focused on the applications of telemonitoring to chronic heart failure, but also on general telemedicine. The studies mostly discuss economic and health benefits of telemonitoring, typically documenting the evidence of patient satisfaction and in lesser number care-givers satisfaction with the telemonitoring systems. Decision support applications were topic of a lesser number of studies. Decision support was provided through automated adjustment of patient medication intake, as well as automated processing of telemonitoring data to indicate patient alarm status to the health-care providers. Multivariate logistic regression model was used to predict the need for medical intervention with an overall accuracy of 74% (Biddiss et al., 2009). Apart from daily measurements of blood pressure, heart rate and weight, this study considered self-rated health and wellbeing survey responses the patients provided twice a week.

Most of the publications presented in Table 2.1 investigate feasibility and effectiveness of telemonitoring solutions. Following the feasibility and effectiveness studies, focus on methods, tools and technologies used for telemedicine was the topic of fewer research publications. The published reviews focused on methodologies used to assess economic and health benefits or tools used to achieve improved patient symptom management as a decision support measure.

How does the telemonitoring achieve the aforementioned benefits? To answer such a question several key elements of telemonitoring systems can be identified.

	Topic	Benefits		System satisfaction		Decision support
		Economic	Health	Patient	Care-giver	
TELEMEDICINE	Feasibility	**Koch et al. 2005,** Hayn et al. 2009, Finkelstein et al. 2010	Kollmann et al. 2007, **Deshpande et al. 2008,** Hayn et al. 2009, Finkelstein et al. 2010	Dohr et al. 2005, Kollmann et al. 2007, Finkelstein et al. 2010, Domingo et al. 2011	Domingo et al. 2011	Dohr et al. 2005
	Effectiveness	Giordano et al. 2006, **Pare et al. 2007**, **Seto 2007, Rojas & Gagnon 2008,** Weintraub et al. 2009, **Polisena et al. 2009**, **Maric et al. 2009, Ekeland et al. 2010**	Giordano et al. 2006, **Rojas & Gagnon 2008, Deshpande et al. 2008**, Scherr et al. 2009, Ferrante et al. 2009, **Barlow et al. 2007, Maric et al. 2009**, Domingo et al. 2010	Cleland et al. 2005, **Pare et al. 2007**, **Deshpande et al. 2008, Rojas & Gagnon 2008, Tran et al. 2008 Maric et al. 2009, Barlow et al. 2007,** Domingo et al. 2010	**Deshpande et al. 2008,** Weintraub et al. 2009	**Rojas & Gagnon 2008,** Scherr et al. 2009, Weintraub et al. 2009, **Maric et al. 2009,** Domingo et al. 2010, De Vries et al. 2011
	Methods / tools / technologies	**Rojas & Gagnon 2008, Tran et al. 2008**	**Tran et al. 2008,** Basilakis et al. 2010	Dohr et al. 2005, Domingo et al. 2010, Finkelstein et al. 2010, Domingo et al. 2011	Weintraub et al. 2009, Domingo et al. 2011	Biddiss et al. 2009, **Maric et al. 2009,** Basilakis et al. 2010, Vukovic et al. 2012

Table 2.1: Overview of existing studies focusing on application of telemedicine to heart failure patients (plain text - research studies, **bold** - reviews)

1. Patient status management including telemonitoring data analyses provides notification to care-givers about worsening patient conditions via alarm messages (Domingo et al., 2011), as well as trend-threshold based classification of patients into risk groups (Basilakis et al., 2010).

2. Utilization of heart failure telemonitoring data can reassess patients' existing medication therapy and provide medication adjustment suggestions to the physicians (Basilakis et al., 2010; Vries et al., 2011).

3. The most common utilization of telemonitoring in heart failure patients is to provide life style recommendations to the patients through educational materials/videos/messages (Basilakis et al., 2010; Domingo et al., 2011, 2012; Ferrante et al., 2010; Finkelstein et al., 2010; Paré et al., 2007; Vries et al., 2011).

4. Enhanced patient self-care and self-management, as integrated part of telemonitoring, may help in heart failure patients to maintain physiological stability (Riegel et al., 2009).

5. Knowing weather conditions in conjunction with physiological telemonitoring data can play a role in enhancing the patient care. (Vukovic et al., 2012b).

2.2 Limitations of patient health status monitoring

The existing studies also identified significant limitations to the application of telemonitoring in heart failure patients. Thus, measurement signal quality was not often the research topic, but could be included in the data processing algorithm through bio signal analyses (Basilakis et al., 2010). Negative aspects of the telemonitoring systems were identified by both, care-givers as well as patients. The critical limitation of the telemonitoring software is generation of overwhelming number of alarms which the health care providers often found to be very annoying (Domingo et al., 2012). A study reported that only 6.4% of alarms generated by the telemonitoring system could be classified as key medical events (Biddiss et al., 2009). In another study, patients reported dissatisfaction with the frequent measurement related questionnaires, causing them to feel the presence of illness in their home environment (Vries et al., 2011). The same study also pointed out occasional reluctance of care-givers to use telemonitoring assisted health care systems, contributing to the previous findings regarding adverse attitudes of health care practitioners towards automated diagnostic systems (Ridderikhoff and van Herk, 1999).

In clinical practice, device alarms sometimes may actually foster the occurrence of adverse events, rather than preventing them. This is the reason why Emergency Care Research Institute (ECRI) ranked alarm hazards at the top of the list of health device hazards for the years 2012, 2013 and 2014 (ECRI, 2011, 2012, 2013), a status confirmed by the Association for Advancement of Medical Instruments (AAMI) listing the medical device alarm hazard as No. 1 industry challenge of the year 2012 (Ferenc, 2012). The alarm hazards also remained at the second highest position in the ECRI rankings for the years of 2011 (ECRI, 2010) and 2010 (ECRI, 2009).

Problems with clinical alarms have existed since the advent of monitoring and therapy device use in health care and were first reported in the 1974 issue of Health Devices (ECRI, 1974). Patient monitoring alarm shortcomings have been the topic of numerous studies and analysis in the literature. Publications have shown the existence of limitations of current alarm systems (ACCE, 2006). The most reported negative side-effect is the large number of nuisance alarms. A paper on adverse events in low-risk patients with chest pain in emergency department, reported that 99.4% of the alarms were false, not resulting in a change of patient treatment management (Atzema et al., 2006). Another study on intensive care unit monitoring showed that over 90% of the alarms were false or clinically insignificant (Imhoff and Kuhls, 2006). The annoying alarms result from the lack of the systems reliability and accuracy and rarely from an adverse patient condition. Some of the consequences of false or nuisance alarms include interference with patient care resulting in reduced effectiveness of the nursing staff. The large number of false alarms demands substantial caregivers time, patience, full attention, fast reactions and commitment which are not always easy to achieve. All such limitations and poor decision support could lead to dangerous undermining of the true alarms by the clinicians (Zhang, 2003).

Review of the existing studies revealed a differentiation between statistical and artificial intelligence approaches to alarm management (Imhoff and Kuhls, 2006), as presented in Table 2.2. The identified approaches were mostly used in intensive care monitoring (30 studies), less in general monitoring time series (16 studies), whereas telemonitoring alarms were investigated in only 5 studies primarily using statistical approaches of trend detection and curve fitting. Usual alarm generation upon exceeding fixed measurement thresholds results in numerous false alarms (Imhoff and Kuhls, 2006). No standard exists for setting the default alarm thresholds for a particular monitored parameter (Vukovic et al., 2010). Furthermore, there is no gold standard for alarm classification (Scherr et al., 2009; Zhang, 2003). The existing statistical techniques are limited by interpretability of the high-dimensional data (Chambrin, 2001), while the artificial intelligence approaches lack predictability needed for regulatory approvals (Imhoff and Kuhls, 2006).

No studies were found to review which of the available technologies exist to provide the highest degree of patient and care-givers satisfaction. Furthermore, the most reliable/suitable telemonitoring methods and tools to achieve adequate decision support still remain unclear. Although recent publications focused on statistical methods used to provide health care decision support, they typically provided insufficient details to allow repeatability of the applied analyses. As additional work is needed in this area, the current study tries to address some of the issues related to the identified shortcomings, particularly with respect to the methods for reliable decision support and higher degree of care-givers satisfaction.

Additionally, no study was found to combine the telemonitoring of patient physiological parameters with monitoring of weather conditions. Therefore the current research attempts to investigate relations between the telemonitored physiological parameters of heart failure patients and weather conditions.

		Intensive care monitoring	General monitoring and time series	Telemonitoring
STATISTICAL APPROACHES	Improved signal extraction	Charbonnier et al. 2004	Davies et al. 2003, Fried 2004	
	Artifact Filters	Makivitra et al. 1991		
	Statistical Process Control	Kennedy 1995, Hill and Endresen 1978		
	Statistical Pattern Detection with Time-Series Analysis	Imhoff et al. 1997, 1998		
	Dynamic Linear Models and Kalman Filters	Smith et al. 1983	Gordon and Smith 1990, Daumer and Falk 1998	
	Autoregressive (AR) models and self-adjusting thresholds	Imhoff et al. 1997, 1998, 2002		
	Phase-space embedding	Gather et al. 2000	Gather et al. 2002	
	Trend detection and curve fitting	Schoenberg et al. 1999, Koski et al. 1990, Haimowitz et al. 1995	Fried and Imhoff 2004, Brillinger 1989	Basilakis et al. 2010, De Vries et al. 2011, Kollmann et al. 2007
	Multivariate statistical methods		Dahlhaus 2000, Gather et al. 2000, 2001	Biddiss et al. 2009
ARTIFICIAL INTELIGENCE APPROACHES	Knowledge Based Approaches	Westenskow et al. 1992, Koski et al. 1994		
	Knowledge Discovery Based and Machine Learning	Imhoff et al. 1999, Morik et al. 2000, Miksch et al. 1996	Morris 1999	
	Neural Networks	Farrell et al. 1992, Ulbricht et al. 1998, Orr and Westenskow 1994, Guru et al. 2007, Patil and Kumaraswamy 2009	Tsien 2000	
	Fuzzy Logic	Bates and Young 2003, Lowe et al. 2001, Wolf et al. 1996, Zong et al. 2004, Adlassnig et al. 2008, Blacky et al. 2011	Becker et al. 1997, Oberli et al. 1999, Tsipouras et al. 2008	
	Bayesian Networks		Laursen 1994	

Table 2.2: Alarm algorithms in patient health monitoring: intensive care monitoring, general monitoring and time series, telemonitoring

2.3 Weather influence on health

The studies explored the effects of climate on health since Hippocrates who wrote his treatise On Airs, Waters and Places (430 B.C.) as part of the Hippocratic Corpus (Jones, 1923). According to the World Meteorological Organization awareness of associations between human health and weather/climate conditions should be included within the national public meteorological services (Kusch, 2004). The purpose is to enhance public understanding of environmental influences and lessen possible adverse outcomes. The most significant natural hazards are heat/cold waves – prolonged periods of extreme temperature and humidity. As an example, due to the hottest summer in Europe for the past 500 years an estimated 15,000 elderly people died only in France between August 4 and 18, 2003 (Poumadere et al., 2005). Apart from the elderly, susceptible to the thermal stress are also people with extreme body mass and malnutrition - under- or over-weight, including infants, those previously experiencing heat related illnesses, patients with heart disease, high blood pressure or other chronic diseases, diabetes, lung, liver and kidney diseases, socially isolated and poor (Keim et al., 2002; Poumadere et al., 2005). Particularly, many cardiovascular diseases including chronic heart failure were associated with weather and climate (Boulay et al., 1999). During the cold weather blood flow towards periphery of the body would be reduced leading to increasing blood density and higher probability of blood clots and thrombosis (Keatinge et al., 1984). On the other hand, the hot weather would lead to increased body temperature, heart rate and sweating (Keim et al., 2002). Risk of heat stress would increase with decreasing air velocity and increasing environmental humidity and temperature. As the humidity approaches 100% the body would lose the ability to reduce heat by sweating.

Statistically significant increases in hospitalizations and mortality were identified during the cold winter months. Paradoxically, excess winter mortality is more pronounced in countries with relatively mild winters (e.g. Southern European and island countries: Portugal, Spain, Greece, Ireland, UK), than in those with more severe climate (e.g. Northern European and Scandinavian countries: Finland, Sweden, Norway) (Healy, 2003). Such effects are related to the typical thermal properties of dwellings depending upon the housing standards of particular countries. More pronounced impacts of outdoor weather variations on health are noticed in the dwellings with poor thermal insulation, even when the outdoor weather variations are relatively mild.

Mortality associated with environmental heat or cold can be directly caused by hyper- or hypothermia or indirectly by the effects on respiratory and cardiovascular diseases as a physiological result of the human body attempts to acclimatize to the environment (Kysely et al., 2009; Näyhä, 2005). Heat acclimatization should be gradual, followed by moderate physical activities, and happen over the prolonged time period of 8-10 days in adults and 10-14 days in children (Gaffin and Moran, 2001).

In the case of heart failure patients, existing studies considered both hot and cold weather conditions, respectively related to high and low humidity. Consequences of decreasing temperature can include increased coronary risk, higher blood pressure, myocardial infraction and respiratory infections leading to deterioration of health status (Stewart et al., 2002) and higher mortality rate (Gonçalves et al., 2007). A study on relationships between weather and myocardial infraction, conducted in Italy 1998-2002, showed that hot weather conditions increased the hospitalization rates particularly in young people, whereas cold weather boosted the average rate of hospitalizations of the elderly (Morabito et al., 2005).

While many studies explored the influence of environmental temperatures on cardiovascular diseases, influence of atmospheric pressure was less studied (Danet et al., 1999). The researchers recommended using daily atmospheric pressure measurements rather than monthly averages due to greater variability and influence on the patients.

The current study analyzes the influence of daily weather fluctuations on heart failure patients physiological status.

2.4 Published results

The current work builds upon the publications listed in Table 2.3.

Reference	Summary
(Vukovic et al., 2010)	New 5-level alarm flag was proposed to supplement the original algorithm towards more efficient and precise telemonitoring.
(Vukovic et al., 2012a)	Utilization of dynamic threshold adjustments around patients' reference state, obtained using Kalman filtering, achieved optimal balance between sensitivity and specificity of an automated alarm generation.
(Vukovic et al., 2012b)	Identification of weather influence on patient health status was introduced as a new feature in home telemonitoring.
(Vukovic et al., 2012c)	Automated alarm management approach was developed and dynamic automatically adjusted alarm thresholds were identified in comparison to the fixed manually adjusted thresholds.

Table 2.3: Published papers including the dissertation results

Chapter 3

Methodology

The analyses used the data collected in the MOBITEL study during the years 2003 through 2008. The dataset contained 9128 measurement records from 65 patients (in the control group) (Scherr et al., 2009). The patients were required to measure their vital signs and manually enter the measured data into the mobile phone devices used to transfer the obtained information to the telemonitoring centre. The collected data were then transferred to the physicians who were able to review the patients health status on their monitors, manually set and adjust measurement thresholds for each individual patient, and receive the alarms when measurements exceeded such preset thresholds. Four patient health status parameters were telemonitored: systolic and diastolic blood pressure, heart rate and weight.

The available data included measured patients physiological parameters, maximum and minimum thresholds used to automatically alarm physicians when exceeded by the measured parameters, as well as records of physician responses to such alarms. The system allowed physicians to record five types of responses, coded as the following: (1) No action, (2) Threshold adjustment, (3) Patient contact, (4) Medication adjustment, and (5) Other actions. Out of such responses, No action and Threshold adjustment were not related to therapeutic actions, i.e. they were considered to be caused by false alarms. Additionally, false alarms included cases with an alarm occurrence and no responses from physicians. On the other hand, "Patient contact", "Medication adjustment" and "Other actions" were assumed to be the result of true alarms. All alarm occurrences were classified as positive events, while no alarm occurrences were classified as negative events. True negative events meant that neither alarms nor physician responses occurred, while false negative classification included cases when no alarms occurred but physicians recorded responses.

3.1 Description of procedures

The collected measurement data records, alarms and the related physician responses to alarm conditions were used in the present study to develop and test the automated alarm management system for telemonitoring of chronic heart failure patients. The research involved the following general steps (Vukovic et al., 2010, 2012a,b,c):

- Preprocessing, data cleaning,
- Analysis of true positive, false positive, true negative, false negative alarms based on documented physician actions; analysis of documented vs. calculated alarms,
- Analysis of exceeded thresholds (single and multiple) on alarm occurrence,
- New 5-level alarm risk indicator was proposed to supplement the original MOBITEL alarm generation algorithm,
- Smoothing the data using various filters (moving average, Savitsky-Golay, Fast Fourier Transformation (FFT), Kalman) to obtain patient reference state,
- Choosing the optimal data window (interval) size for measurement data smoothing and threshold adjustment: fixed window size, variable window size,
- Replacing fixed thresholds with dynamic thresholds and choosing the optimal threshold values based upon statistical parameters (standard deviation within the window size, 99% CI, mean/median absolute and relative thresholds, ROC curve) to reduce the number of false alarms,
- Graphical visualization tool was developed to indicate trends in the patients' physiological parameters,
- Testing and selection of the optimal procedure for the decision support,
- Analysis of weather influence on patient health status parameters.

Appendix A presents an overview of applied statistical methods and indices used in the data analyses. Precondition for the analyses was data cleaning. The cleaning removed the cases when patients recorded measurements isolated in time and implausible values outside of reasonable measurement ranges. The remaining data included inconsistent number of measurements per day, as patients mostly recorded measurements only once but sometimes several times a day. In such cases new single measurement records were created by averaging and replacing multiple daily measurements. The number of exceeded thresholds was summed for each day, while only a single physician response was considered with a priority given to those responses indicating medical attention. After the cleaning, the number of measured records was reduced to 8114 from 54 patients. However, only 52 patients could be considered in the alarm management optimization due to the insufficient and irregular measurement records of the 2 additionally removed patients.

The subsequent steps focused on providing basic data statistics, which indicated large number of false alarms in the dataset. Therefore, further focus was on the statistical data analysis methods to reduce the number of false alarms. Analyses included alternative procedures for setting up thresholds in comparison to the alarm generation based upon manually adjustable fixed thresholds. Data filtering and smoothing was applied to identify optimal indicator of patient reference states. Investigation of fixed and variable threshold limits around the reference states followed, enhancing the decision support system for patient therapy optimization. The automatically adjusted dynamic thresholds were introduced for each new measurement with respect to the estimated reference state of patient conditions. Additionally, investigation of weather influences on the measurement records was conducted to identify optimal patient health status predictors. Officially recorded weather conditions over the telemonitoring time period, including environmental temperatures, humidity, and atmospheric pressure, were available from the Austrian Central Institute for Meteorology and Geodynamics (ZAMG, 2012). Considering the geographic spread of patients throughout Austria, weather data for each patient were used from one of the five meteorological stations closest to the patient location: Vienna, Graz, Innsbruck, Klagenfurt and Linz.

A "reference state" of four telemonitored variables of chronic heart failure patients was estimated for each measurement day: weight, heart rate, systolic and diastolic blood pressure. The estimates were based on smoothing of the measured data within an appropriate window size of consecutive measurement days preceding the current measurement day. Smoothing of the data was done using statistical measures of location, i.e. mean value (moving average), and Kalman filtering. Investigation of the possible dynamic threshold bounds was based on the calculated means and medians of the existing MOBITEL fixed thresholds. Various window sizes and threshold values (absolute and relative) were used to test the performance of the algorithm.

Finally, the optimized dynamic threshold adjustment methodology (reference state $\pm$ threshold bounds) was applied and tested on the existing data from the MOBITEL study. The difference in occurrence of true and false alarms was observed when dynamically adjusting the alarm thresholds. The automatically generated alarms were classified as true or false based on the comparison to the available physician actions documented in the MOBITEL study. The documented physician actions were used as a reference to estimate sensitivity and specificity of the developed algorithm for automated alarm generation. The results were obtained varying the upper and lower threshold bounds around the reference state in combination with the fixed constant upper and lower threshold limits that will always trigger alarm occurrence regardless of the reference state value. Based on the calculated results, receiver operating characteristic (ROC) curves were used to select the optimal combination of algorithmic upper and lower threshold values for the monitored parameters.

Trend analysis of measurements and of reference state were conducted to develop a tool for graphical visualisation. The tool was designed to indicate potentially critical measurement points approaching the identified thresholds. Such indication included colouring of the measurement background fields within the thresholds.

A comparison of the original manual threshold adjustment method used in the MOBITEL study and the newly developed approach is presented in Figure 3.1.

3.2 Description of used tools

DataLab statistical package (Lohninger, 2013) and R programming language and environment for statistical computing (Cowpertwait and Metcalfe, 2009) were used to perform statistical analyses and create graphics. Additionally, some preprocessing steps, and basic statistics were obtained using Microsoft Excel.

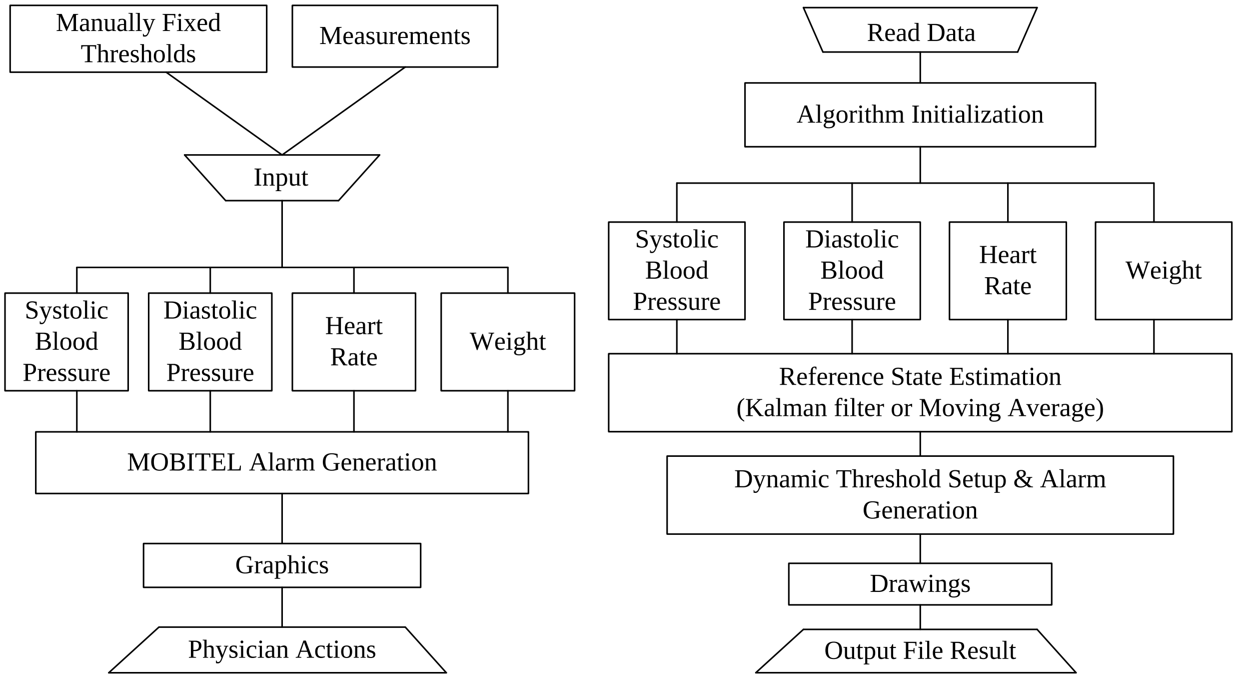


Figure 3.1: Manual (left) and dynamic threshold adjustment (right) procedures

DataLab showed as a user friendly software tool for data analyses and interpretation, offering a variety of data editing and visualization capabilities, including mathematical and statistical procedures. Particularly, DataLab built-in principle component regression capabilities and variance inflation factor calculations were used in the conducted analyses.

The open source R environment, developed and maintained by the R foundation, presented an integrated environment for data analysis, calculation and graphical representation of results. R offered free packages of basic statistical functions as well as dedicated packages focused on particular statistical applications. Apart from the general R functionalities the following dedicated libraries were used:

- RSEIS: Seismic Time Series Analysis Tools (Lees, 2012) - library for general time-series plotting, filtering, and interactive display; functions and algorithms for spectrum analysis, wavelet transforms, particle motion, hodograms. This library was used to remove trend from the analysed data.

- GeneCycle: Identification of periodically expressed genes from time series data (Ahdesmaki et al., 2012). The library provided functions used in conjunction with Fourier transform and spectrum analyses to investigate possibilities for prediction of patient reference state condition.

- sspir: State Space Models in R (Dethlefsen et al., 2009) provided built-in functions for Kalman filtering and smoothing.

- CADStat: GUI to statistical methods for causal assessment (applied in JGR) (Yuan et al., 2012) was used to calculate correlations.

- polycor: Polychoric and Polyserial Correlations (Fox, 2012) library was used to calculate correlations between numeric and ordinal variables.
- Hmisc (Harrell, 2012) library for data analyses and visualization was used to calculate Pearson and Spearman correlations.
- gregmisc (Warnes, 2012) library of miscellaneous functions was used for plotting results of statistical tests.

Furthermore libraries containing statistical functions integrated with user friendly graphical interface were examined:

- R commander (Rcmdr library) (Fox, 2005) is a graphical interface consisting of various built-in functions for statistical data analyses which can be activated graphically but also through a text script command window. R commander has particularly developed capabilities for convenient visualization of typical statistical plots such as: histogram, boxplot, QQ plot, scatterplot, bar graph, pie chart, etc. for which purpose it was primarily used in the current study.
- JGR (read Jaguar) (Helbig et al., 2005) is another graphical interface whose name is an abbreviation of Java Gui for R. Such package also has some built in statistical functions, although not as many as R commander, and can be used for graphical visualization of results.

In general, using any of the available graphical interfaces for R limited the user in terms of availability of implemented functions and scope of provided options for such functions. Thus, usage of R commander and JGR restricted the flexibility of the original R environment, but provided friendlier graphical appearance. Various graphical interfaces offered possibilities for convenient visualization of results. Such tools were very significant as graphical visualization in R is otherwise cumbersome and requires substantial knowledge of various plotting features in order to achieve good control over the graphical appearance.

Appendix B lists various built-in and developed R functions used in the described data analyses.

Chapter 4

Descriptive Statistics and Data Preprocessing

The patients subjected to telemonitoring were spread throughout Austria, as illustrated in Figure 4.1. Majority were located close to Graz, followed by Vienna, Linz, Klagenfurt and Innsbruck.

Figure 4.1: Geographical spread of patients throughout Austria

Basic statistic on the raw and cleaned data are further presented together with the description of procedures used for data preprocessing and cleaning.

4.1 Raw data

Analyses of the raw data was conducted to get an overview of the available patient measurements and physician records.

Figure 4.2 shows the number of days patients conducted measurements (violet) during their participation in the study (blue), as well as the total number of conducted measurements (orange). If the patients conducted one measurement per day each day then all the three bars, violet, blue and orange, would have the same values. However, if the patients missed some days and did not conduct any measurements the blue bar would have higher values then the violet bar. Finally, if the patients conducted multiple measurements per day the orange bar will be higher than the violet bar. For the majority of patients the figure shows that the blue bar is the highest which means that almost every patient had days without any measurements. Also, for the majority of patients orange bar is higher than the violet bar which means that during certain days they had multiple measurements. In extreme cases where the blue bar (the number of total days in the study) is substantially higher than the violet bar, such patient would be removed from the current analysis due to the lack of sufficient measurement information, i.e. large measurement gaps. To overcome the issue with the multiple measurements per day, daily averages of the conducted measurements were calculated. From the figure it can also be noticed that the majority of patients participated in the study and conducted measurements approximately during the period of six months or 180 days.

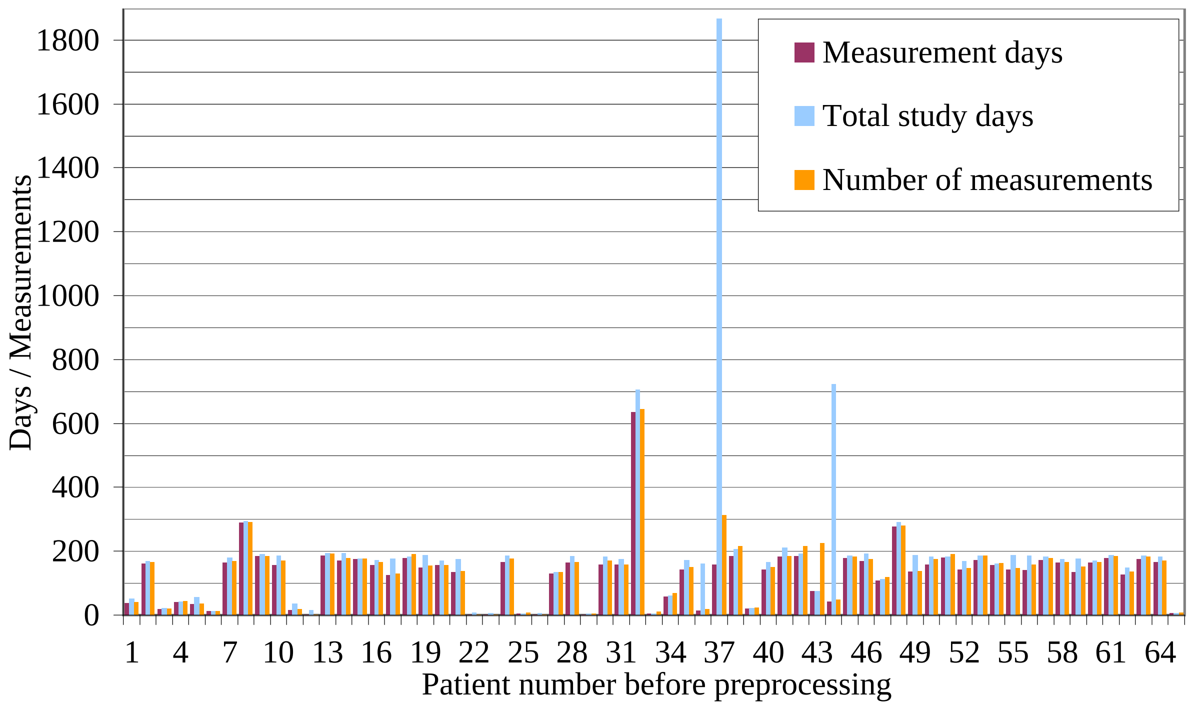


Figure 4.2: Patient measurement statistics on raw data including the number of measurements, measurement days and the number of days between the first and last measurement

4.2 Data preprocessing

Initially, all data were cleaned by removing the outliers defined as follows:

1. diastolic blood pressure greater or equal to systolic, (36 events; 0,4% of measurement records),
2. systolic blood pressure outside of range 50-300 mmHg, (254 events; 2,8%),
3. diastolic blood pressure outside of range 40-150 mmHg, (318 events; 3,5%),
4. heart rate outside of range 20-200 bpm, (264 events; 2,9%),
5. weight outside of range 30-500 kg, (216 events; 2,4%).

Additionally, patients with large measurement gaps and/or insufficient records were removed from the analyses. After the cleaning, the number of patients considered in the analyses was reduced from 65 to 54. To ease the data management patient identification key was introduced between the original recorded Patient IDs and correspondingly assigned Patient numbers, as presented in Table 4.1.

Patient No.	1	2	3	4	5	6	7	8	9
Patient ID	10,131	10,250	10,314	10,402	10,441	10,590	10,601	10,617	10,738
Patient No.	10	11	12	13	14	15	16	17	18
Patient ID	11,061	11,235	11,306	11,322	11,381	11,423	11,490	11,526	11,542
Patient No.	19	20	21	22	23	24	25	26	27
Patient ID	11,777	11,864	12,022	12,041	12,029	12,119	12,145	12,231	12,289
Patient No.	28	29	30	31	32	33	34	35	36
Patient ID	8,532	8,587	8,603	8,779	8,803	8,832	8,844	8,890	8,929
Patient No.	37	38	39	40	41	42	43	44	45
Patient ID	8,994	9,029	9,041	9,051	9,086	9,178	9,190	9,211	9,234
Patient No.	46	47	48	49	50	51	52	53	54
Patient ID	9,247	9,283	9,299	9,448	9,577	9,622	9,715	9,735	9,789

Table 4.1: Identification key between Patient numbers and corresponding Patient IDs

Figure 4.3 presents the distribution of mean blood pressures, heart rates and weights of 54 patients, together with normal distribution curves. According to the Central Limit Theorem, mean of the patient means approaches the mean of the heart failure patient population. As the distribution of patient means resembles normal for all the presented parameters, it can be assumed that the selected sample is representative of the population.

As an illustration of data available after the cleaning, Figure 4.4 presents the home telemonitored vital signs of four selected patients over the course of the study. The figure illustrates variability in patient conditions that can be found in the data set (Patient IDs), e.g. relatively stable (8,832 and 8,587), changing (9,735) or initially changing and then stabilized weight (10,601); measurement gap (9,735); typically high (8,832 and 8,587),

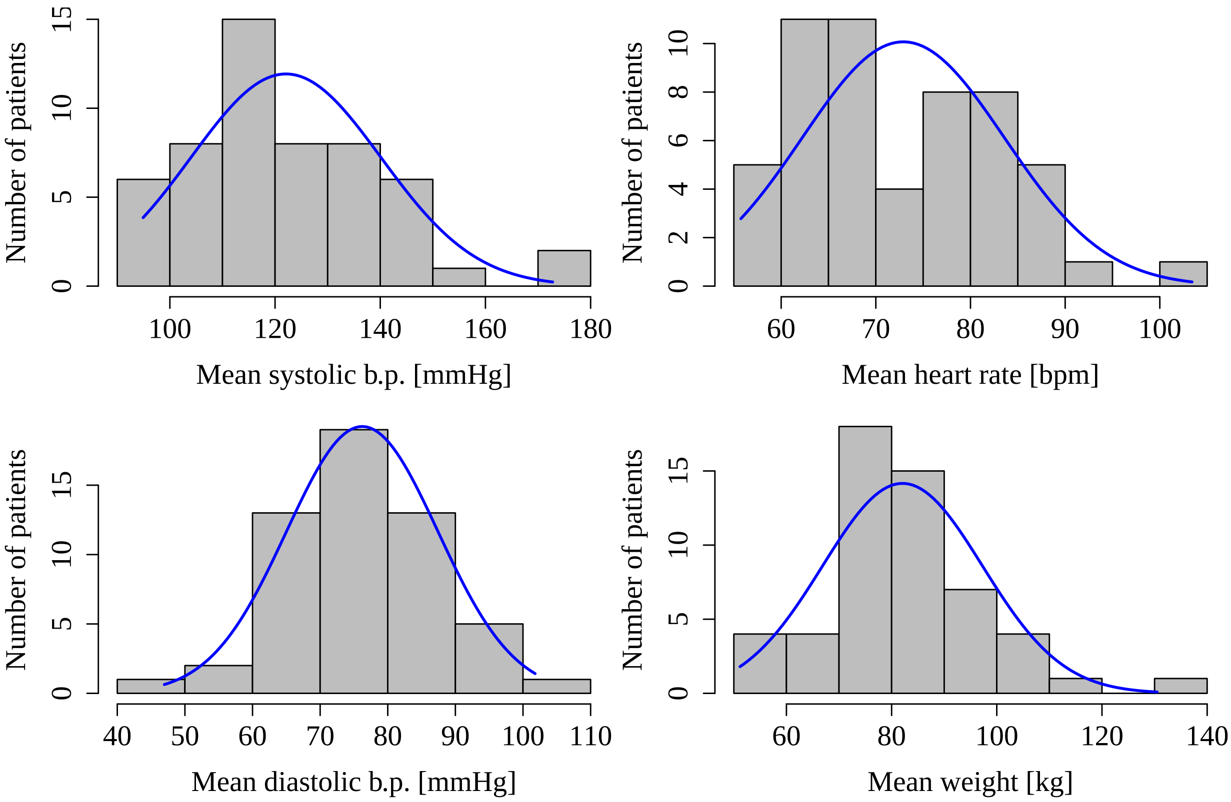


Figure 4.3: Distribution of patient mean blood pressures, heart rates and weights together with normal distribution curves

low (9,735), and stabilized to normal after initially variable (10,601) blood pressures; fluctuations and occasional occurrence of extreme heart rates (all patients).

Figure 4.5 illustrates the occurrence of alarms and associated physician responses for the four selected patients. In the figure, the line "Total alarms" gives a summary of all the weight, heart rate and blood pressure alarms that have occurred. Although in the majority of cases alarms corresponded to physician actions, certain alarms had no associated physician responses (e.g. initial alarms for Patient ID 8,587, or last alarms for Patient ID 9,735). Similarly, some physician responses were recorded even without the indication of alarms (e.g. the last physician action – no action – for Patient ID 10,601).

Table 4.2 presents basic statistics of all the measurements after the cleaning. The table includes minimum, first quartile, median, mean, third quartile and maximum values of the variables of interest.

Analyses of the clean data revealed that 20% of measurement days indicated alarm condition and coincided with physician responses, while no alarm and no responses occurred in 74% of cases, as presented in Figure 4.6. In 1% of the cases physicians recorded responses during the days when no alarms were exhibited, while in 5% of the cases physicians had no responses although the alarms were shown.

Analysis of physician actions to the alarm situations showed that in the majority of cases, 1269 or 78% of all recorded responses to alarms, physicians took no action.

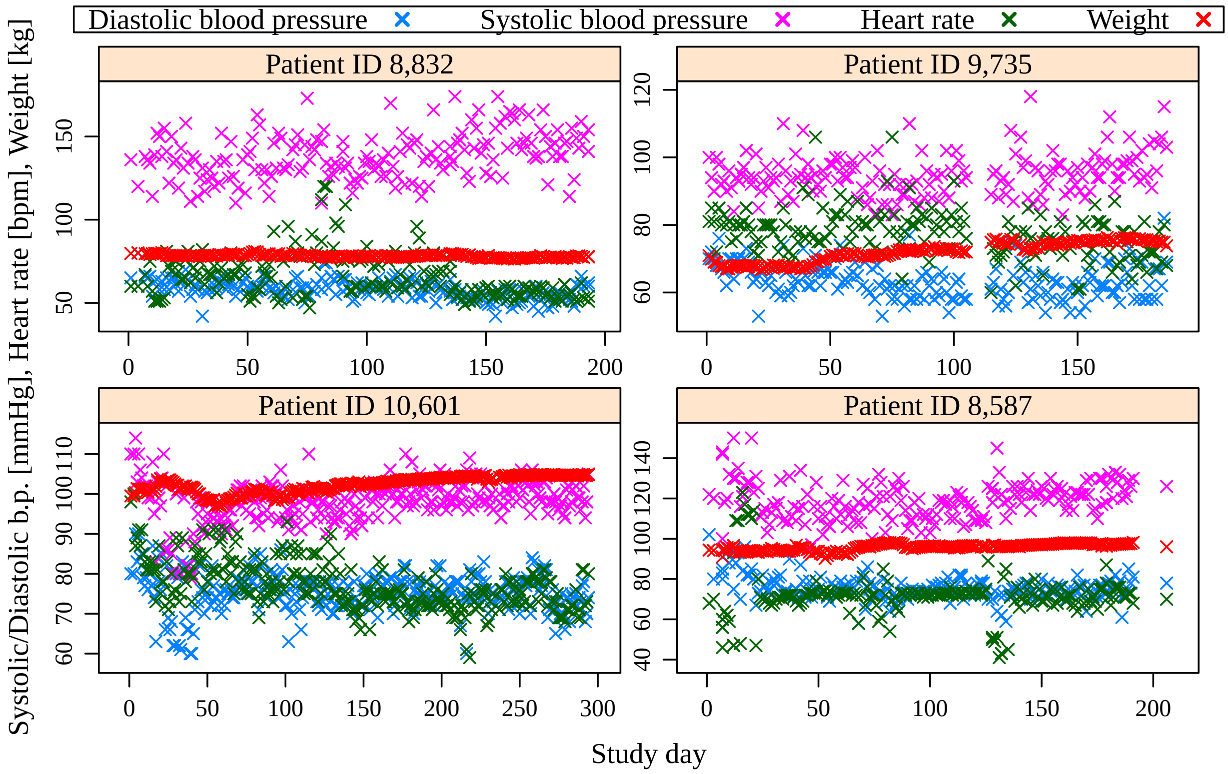


Figure 4.4: Home telemonitored vital signs (systolic and diastolic blood pressure, heart rate, weight) of four selected patients

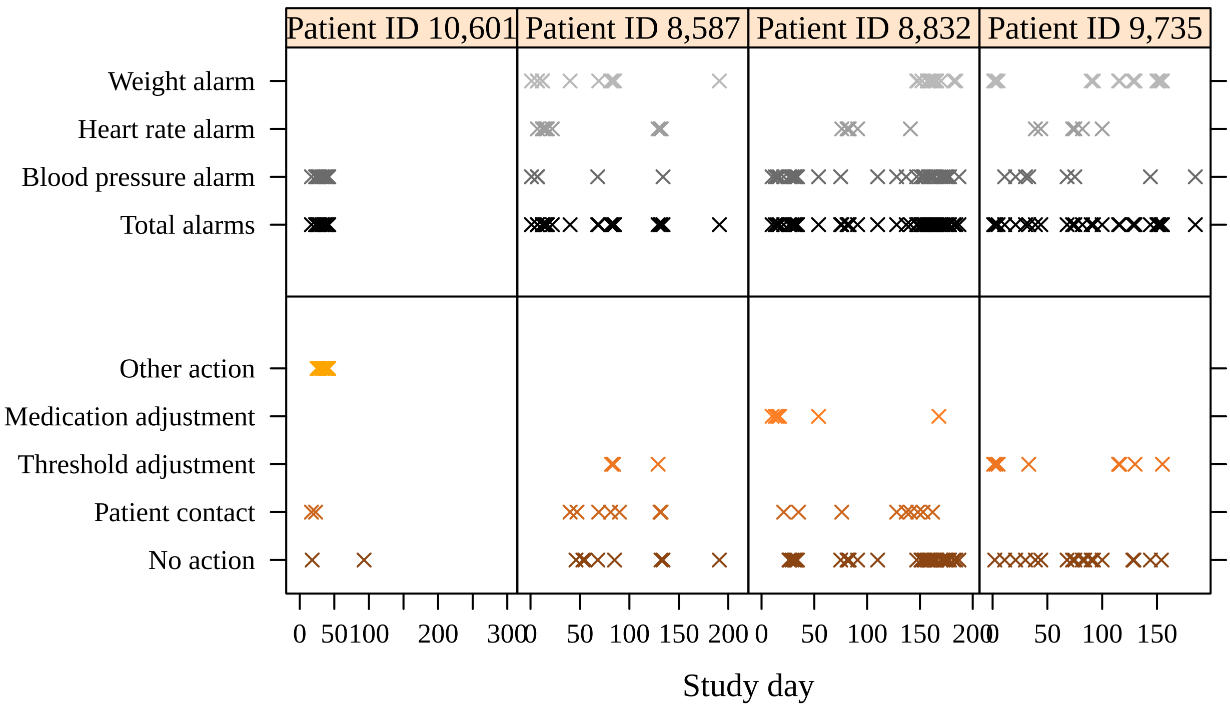


Figure 4.5: Alarms and corresponding physician responses for four selected patients

	Measurement date
Start:	October 9, 2003
1st Qu.:	June 17, 2004
Median:	December 7, 2004
Mean:	June 17, 2005
3rd Qu.:	February 23, 2006
End:	February 19, 2009

	Systolic b.p. [mmHg]	Diastolic b.p. [mmHg]	Heart rate [bpm]	Weight [kg]
Min.:	70	41	29	45.5
1st Qu.:	105	66	63	73
Median:	118	74	71	81.3
Mean:	119.8	73.7	72.1	82.4
3rd Qu.:	131	82	80	92.5
Max.:	220	134	160	200.5

	Minimum systolic b.p. threshold [mmHg]	Maximum systolic b.p. threshold [mmHg]	Minimum diastolic b.p. threshold [mmHg]	Maximum diastolic b.p. threshold [mmHg]	Minimum heart rate threshold [bpm]	Maximum heart rate threshold [bpm]	Minimum weight threshold [kg]	Maximum weight threshold [kg]
Min.:	80	120	35	80	40	60	45	54
1st Qu.:	90	140	55	90	50	90	65	78
Median:	100	150	60	95	50	100	75	83
Mean:	97.2	153.4	57.2	93.7	50.8	96.7	74.7	85.9
3rd Qu.:	100	160	60	100	50	100	84	96
Max.:	120	190	80	110	70	120	117	150

Table 4.2: Minimum, first quartile, median, mean, third quartile and maximum values of the variables of interest (measurement date, systolic/diastolic blood pressure, weight and corresponding minimum and maximum thresholds) in the clean dataset

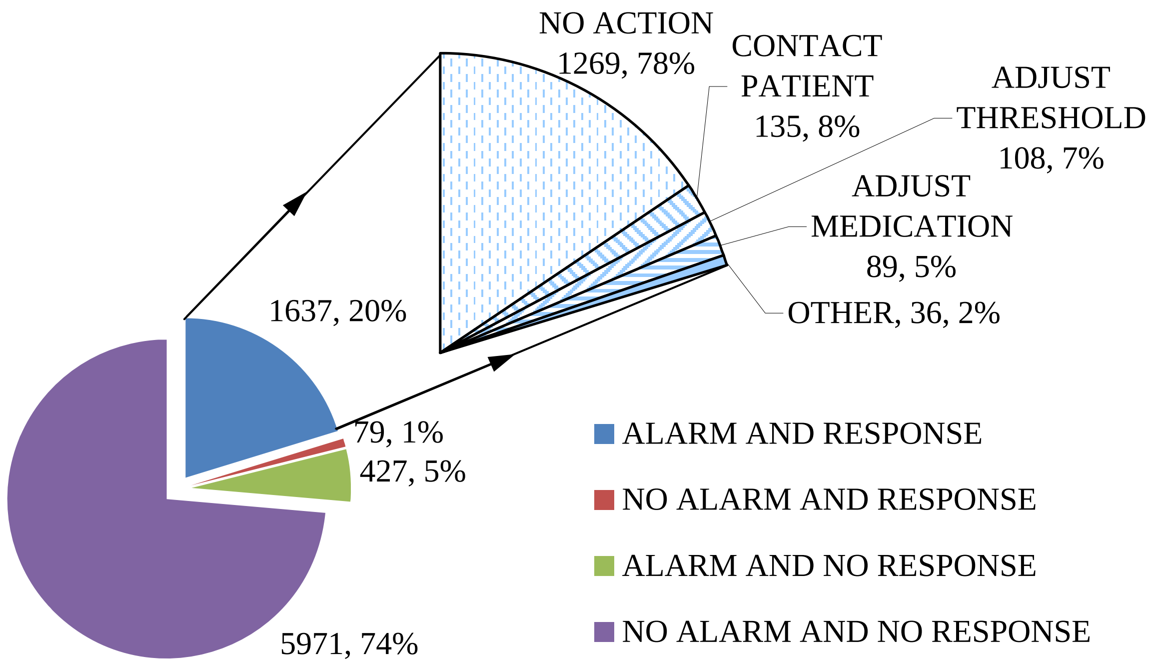


Figure 4.6: Number of alarms and physician actions in the telemonitoring data

Additionally in 427 or 5% of the total measurements no information was recorded regarding the physician actions although the alarms were indicated. If we assumed that physicians simply ignored those alarms then the total number of no actions increased to 1696 or 82% of all detected alarms. In the remaining 18% of detected alarm cases physicians mostly contacted the patients (135 times), adjusted measurement thresholds causing the alarms (108 times), adjusted medications (89 times) or did some other actions (36 times).

Out of the 79 cases when no alarms were indicated but physicians recorded their responses in the database, 59 were “No actions”. Furthermore, 3 cases included threshold adjustment, 6 times patients were contacted, 1 medication adjustment was made, and 10 times other actions were taken.

Looking into the number of exceeded multiple thresholds causing the occurrence of alarms, on average 74% of the alarms were caused by exceeding a single threshold, as presented with a red line in Figure 4.7. Multiple simultaneously exceeded thresholds (double and 3 or more) had the highest impact on true alarms: up to 44% portion, as in the case of “Patient contacts”, opposed to e.g. 23% of “Threshold adjustments” and “No actions” representing false alarms.

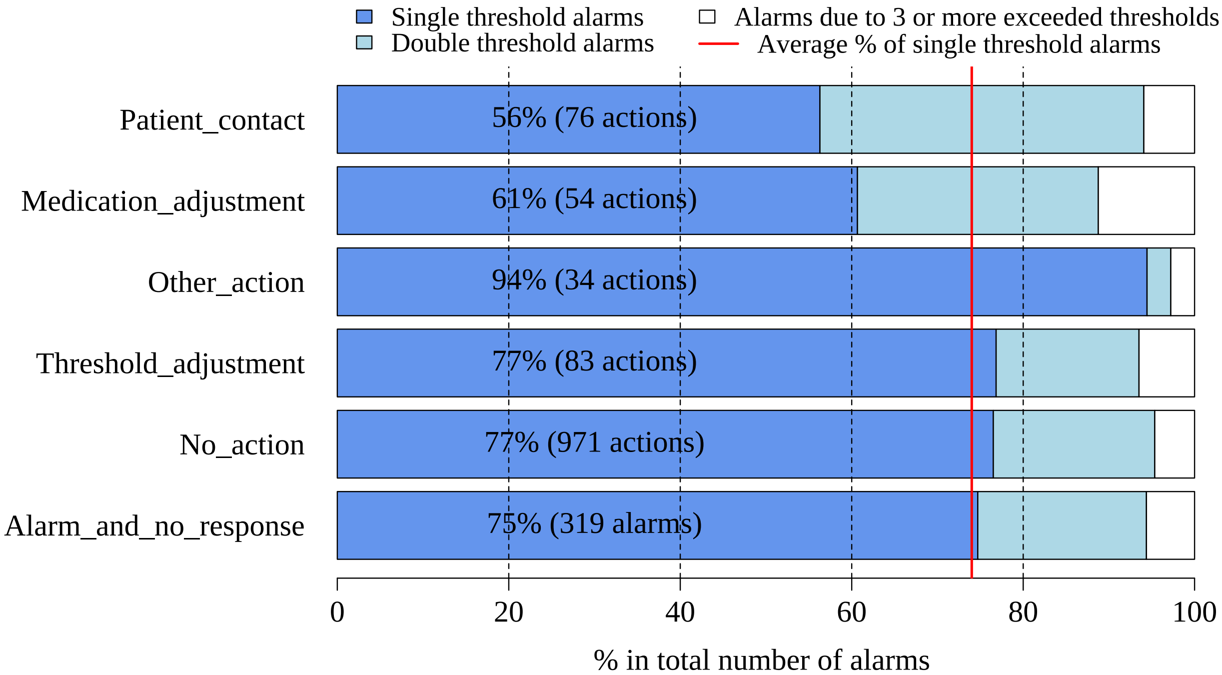


Figure 4.7: Portion of alarms caused by exceeding a single, double and 3 or more thresholds in the total number of alarms

Based on such statistics one can conclude that single exceeding of the thresholds has the highest impact on the alarm occurrence. As the most occurring are false alarms, single threshold exceeding has also the highest influence on the false alarms. Overview of the causes for exceeding single thresholds and associated physician actions to such alarms are presented in Figure 4.8. The figure shows percentages of physician actions in response to the exceeded upper (red bar fields) and lower (yellow bar fields) thresholds. Thus,

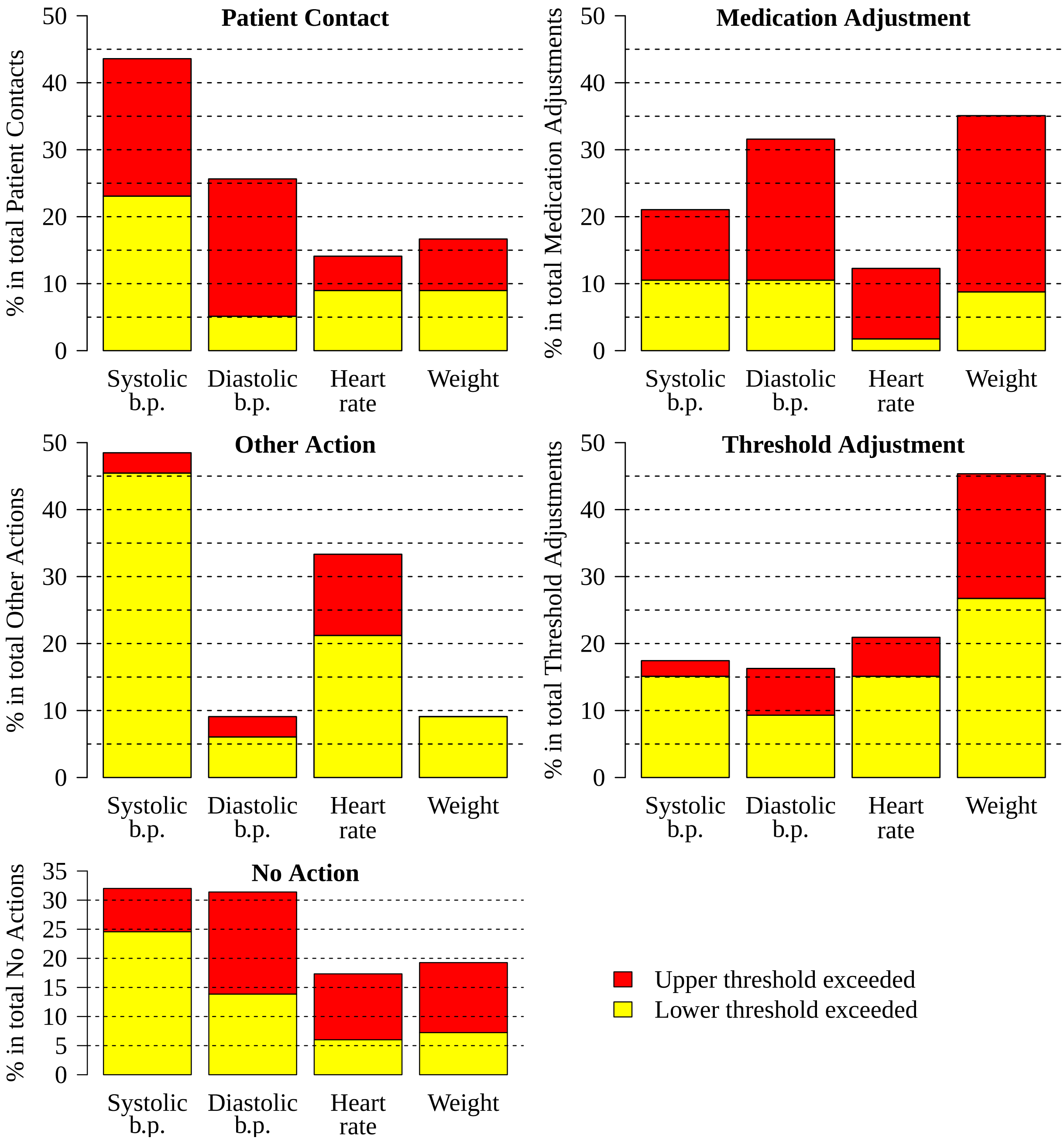


Figure 4.8: Portion of physician actions (Patient contact, Medication adjustment, Other action, Threshold adjustment, No action) caused by exceeding a single threshold (upper or lower)

exceeding systolic blood pressure thresholds causes over 40% of physician actions "Patient contact". Together with the exceeded diastolic thresholds (particularly upper), both systolic and diastolic blood pressures are responsible for nearly 70% of the patient contacts. Physician action "Medication adjustment" mostly (in about 65% of the cases) occurred when upper thresholds were exceeded in the case of all parameters: systolic and diastolic blood pressures, heart rate and weight. Exceeded systolic and diastolic blood pressure thresholds together were responsible for over half of the medication adjustments, whereas the single most influential cause of 35% of medication adjustments was exceeded weight threshold (particularly upper). In the majority of cases (around 80%), physician action "Other action" occurred due to exceeded lower thresholds. Exceeded lower systolic blood pressure threshold was the most prevalent, causing over 45% of other actions, whereas exceeded heart rate thresholds (both upper and lower) added almost 35% of other actions. Physician action "Threshold adjustment" was mostly (45%) caused by exceeded weight thresholds, whereas "No action" mostly occurred when systolic or diastolic blood pressure thresholds were exceeded. Over 30% of no actions occurred due to each of the blood pressures, including exceeded upper or lower thresholds together.

Records of each patient medication intake were also available in the telemonitoring database. To analyze the influence of medications on the patient conditions, Figure 4.9 presents an overview of the number of different medications the patients were given over the course of telemonitoring. The system allowed recording the names and dosage of up to four medications at every given measurement record. In the case when patients took less than four medications, medication names were repeated. As indicated in Figure 4.9, majority of the patients (33 out of 54, or 61%) took four different medications over the course of telemonitoring.

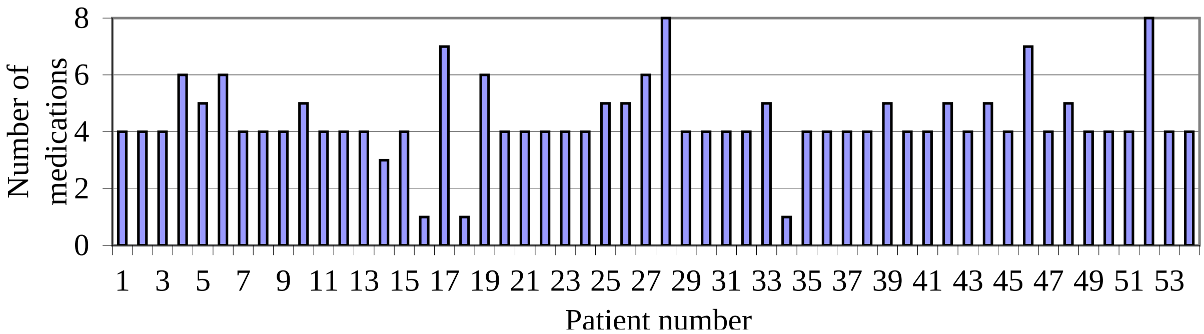


Figure 4.9: Number of different medications patients took over the course of telemonitoring

A change in medication type or dosage between the consecutive measurement records indicates a change in therapy. The number of different therapies for each patient is presented in Figure 4.10.

Despite of the large number of medications the patients took over the course of telemonitoring, as presented in Figure 4.9, change in patient therapies rarely occurred and majority of the patients (34 out of 54, or 63%) had only one medication therapy, as presented in Figure 4.10. Due to the large number of changing therapies and insufficient number of measurements per therapy patient 28 was additionally removed from the dataset and was not considered in further analyses. Also, patient 10 was removed having 2 changes in therapies over a total of 13 monitoring days within a 15 day total monitoring period during which the patient participated in the study.

Figure 4.11 presents the number of patients using a certain type of medications. The most frequently used medications (with more than 50% of the patients) were Lasix and

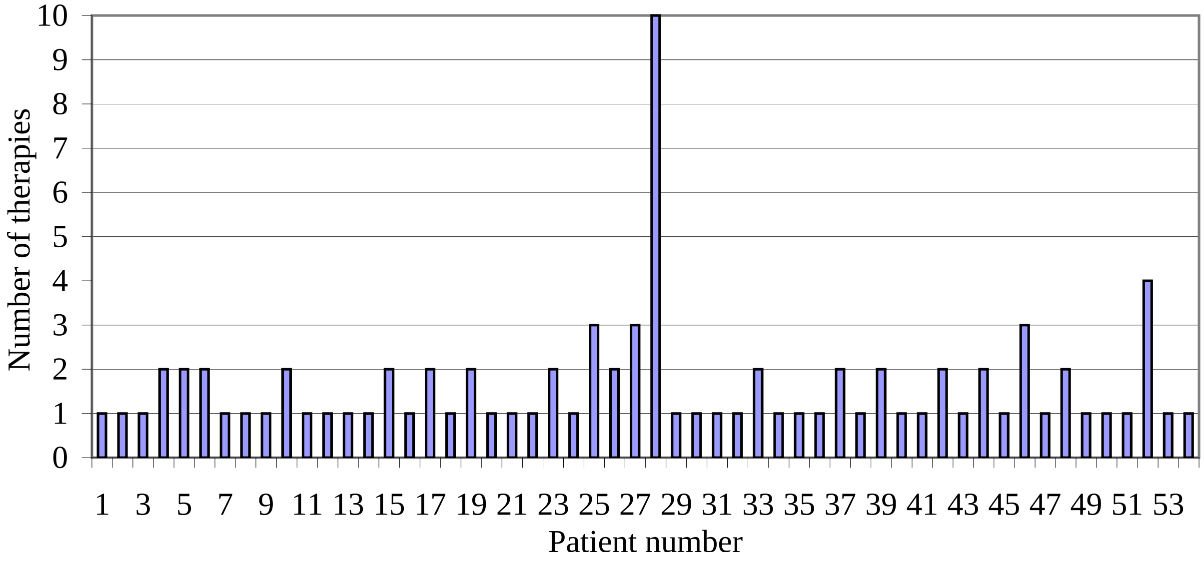


Figure 4.10: Number of different therapies per individual patients over the course of telemonitoring

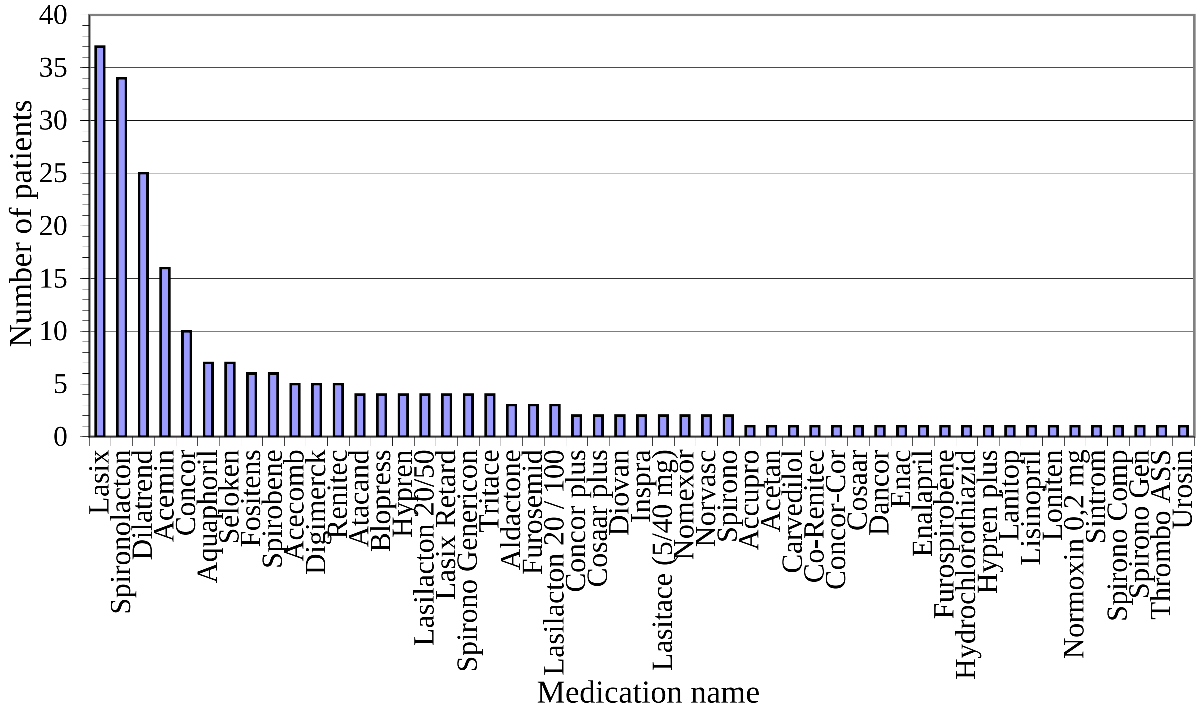


Figure 4.11: Number of different patients taking particular medications over the course of telemonitoring

Spironolacton. These medications are diuretics, used to treat excessive fluid accumulation and swelling (edema) of the body caused by heart failure, cirrhosis, chronic kidney failure, and nephrotic syndrome. As majority of the patients took four different medications over the course of telemonitoring, the two most frequently used medications, Lasix and

Spironolacton, were often prescribed in conjunction with other blood pressure drugs to treat high blood pressure.

To visualize the frequency of prescribing particular medication pairs Table 4.3 shows the number of different medication combinations in which each medication pair occurs. All the possible pairs in which a particular medication is used can be reviewed following the horizontal row next to a particular medication name till the diagonal and then continuing along the upward column till the end of the table. A total of 77 medication combinations were present over the course of telemonitoring. The highlighted diagonal in Table 4.3 presents the number of different medication combinations including each particular medication type. The table considers some of the medication names as different trade names of the same medication type and sums up such occurrences as indicated by the star symbols next to particular medication names.

Similarly to the most frequently used medications by individual patients, Lasix (also traded as Furosemid and Lasix Retard) and Spironolacton (also traded as Aldactone, Spirobene, Spirono Gen and Spirono Genericon), appear as the most commonly used medications in different medication combinations. These medications are present in 54 (70%) and 51 (66%) medication combinations, respectively.

Table 4.4 presents an overview of indications for the most frequently prescribed medications in the study, including their active ingredients and classification type. The presented information is available from the Austrian Agency for Health and Food Safety (AGES, 2013) and UK online prescription drug database (MIMS, 2013). The equality between certain medication names in the table is used to indicate usage of the same active substances in those medications, although the dosage might be different.

As a guiding reference, classifications of hypertension stages is presented in Table 4.5. The classification is based on systolic and diastolic blood pressure bounds and can be applied to ages 18 and older (AHA, 2012; MayoClinic, 2011a,b). If at least one of the blood pressures, systolic or diastolic, falls outside of the desirable category, the person is considered to be outside of the desirable blood pressure range.

As each person's individual life style, weight, possible other diseases etc. may influence blood pressure, Table 4.5 should be only considered as a general guide in decision making. For example, what may be low blood pressure for some people may be normal for others (MayoClinic, 2011b). Apart from the blood pressure levels, sudden changes in blood pressure (e.g. 20 mmHg or more) can also be dangerous.

	Accupro	Acecomb	Acemin*	Aldactone**	Aquaphoril	Atacand***	Carvedilol	Concor	Concor plus	Concor-Cor	Co-Renitec	Cosaar	Cosaar plus	Dancor	Digimerck	Dilatrend	Diovan	Enac	Enalapril****	Fositens	Furosemid*****	Furospirobene 20/50	Hydrochlorothiazid	Hypren	Hypren plus	Inspra	Lanitop	Lasilacton******	Lasitace (5/40 mg)	Loniten	Nomexor	Normoxin 0,2 mg	Norvasc	Seloken	Sintrom	Spirono	Thrombo ASS	Tritace	Urosin
Urosin															1			1			1																		1
Tritace				4				1								2					5													2				5	
Thrombo ASS		1	1																															1			1		
Spirono			3					2													3										1					3			
Sintrom					1				1															1											1				
Seloken		3	5	5		1											1				6							1	1					9					
Norvasc					2						1					1			3		1											1	3						
Normoxin 0,2 mg					1														1													1							
Nomexor			1	1																	3			2		1					3								
Loniten				1			2	2								1														2									
Lasitace (5/40 mg)			1	2												1						1							2										
Lasilacton******		1	6	3		2		3								4			1	1	7						1	10											
Lanitop			1													1											1												
Inspra	1							1													2			1		2													
Hypren plus				1												1					1				1														
Hypren				2	1			1	1						1	1					4			5															
Hydrochlorothiazid			1	1												1							1																
Furospirobene 20/50				1												1						1																	
Furosemid*****	2	4	14	37	4	9		14		1		1	2	1	7	20	1	1	4	5	54																		
Fositens				5	1			1							2	3				6																			
Enalapril****				4	3			1			1				1	2			7																				
Enac															1			1																					
Diovan				2	1				1								2																						
Dilatrend		1	8	21	2	4	1	3				1	1	1	1	29																							
Digimerck		2		4	2	1		1							8																								
Dancor			1											1																									
Cosaar plus				2									2																										
Cosaar					1							1																											
Co-Renitec					1						1																												
Concor-Cor				1	1					1																													
Concor plus				3	2		1		4																														
Concor	2	1	5	10		1	3	18																															
Carvedilol				3			4																																
Atacand***				8	1	9																																	
Aquaphoril		1		5	10																																		
Aldactone**	1	4	10	51																																			
Acemin*		3	21																																				
Acecomb		7																																					
Accupro	2																																						

Acemin* = Acetan = Lisinopril
Aldactone** = Spirobene = Spirono Gen = Spirono Genericon = Spironolacton
Atacand*** = Blopress
Enalapril**** = Renitec
Furosemid***** = Lasix = Lasix Retard
Lasilacton****** = Spirono Comp forte

Table 4.3: Number of different medication combinations in which particular medication pairs occur over the course of telemonitoring

Medication name (Active ingredients)	Indications	Medication type
Lasix = Lasix retard = Furosemid (Furosemide)	Edema due to heart disease (e.g. heart failure)	Loop diuretic
	Fluid retention (edema) due to chronic renal dysfunction (e.g. after burns) and due to certain liver diseases	
	Arterial hypertension – mild to moderate severity	
	Supportive therapy of nephrotic syndrome	
Spironolacton = Spirobene = Aldactone = Spirono Genericon (Spironolactone)	Primary hyperaldosteronism (Conn's syndrome)	Potassium-sparing diuretic (water pill)
	Liver cirrhosis with ascites and edema	
	Water retention caused by heart failure	
	Respiratory failure with chronic pulmonary heart disease	
	Edema associated with chronic renal disease	
	Hypertension	
Dilatrend (Carvedilol)	Stable symptomatic chronic heart failure of any severity, ischemic or non-ischemic origin, in addition to conventional basic therapy with diuretics and ACE inhibitors	Beta-blocker
	Long-term treatment of left ventricular dysfunction after acute myocardial infarction in combination with ACE inhibitors and other standard therapy	
	Essential hypertension	
	Chronic stable angina pectoris	
Acemin = Acetan = Lisinopril (Lisinopril)	Hypertension	Angiotensin Converting Enzyme (ACE) inhibitor
	Congestive Heart failure (in patients who do not respond adequately to diuretics and digitalis)	
	Short-term treatment (6 weeks) of hemodynamically stable patients within 24 hours of an acute myocardial infraction	
	Renal complications of diabetes mellitus	
	Renal disease in hypertensive patients with type 2 diabetes mellitus and incipient nephropathy	
Concor (Bisoprolol hemifumarate)	Essential hypertonia	Beta 1 blocking agent
	Coronary artery disease (angina pectoris)	
Aquaphoril (Xipamide)	Arterial hypertension	Sulfonamide diuretic agent
	Cardiac and renal edema	
Seloken (Metoprolol succinate)	Stable symptomatic heart failure with reduced left ventricular function, in combination with other therapies	Beta-blocker
	Hypertension in adults, Chronic stable angina pectoris	
	Secondary prevention after myocardial infarction	
Fositens (Fosinopril sodium)	Essential hypertension (all grades)	ACE inhibitor
	Heart failure	
Acecomb (Hydrochlorothiazide, Lisinopril)	Essential hypertension when monotherapy with lisinopril or hydrochlorothiazide alone has shown insufficient treatment success	Thiazide diuretic

Medication name (Active ingredients)	Indications	Medication type
Digimerck (Digitoxin)	Chronic heart failure (due to systolic dysfunction)	Heart glycosides of red foxglove plant (Digitalis purpurea)
	Tachycardia arrhythmias especially atrial fibrillation	
Renitec = Enalapril = (Enalapril maleate)	Hypertension	ACE inhibitor
	Symptomatic heart failure	
	Prevention of symptomatic heart failure in patients with asymptomatic left ventricular dysfunction	
Atacand = Blopress (Candesartan cilexetil)	Essential hypertension in adults	Angiotensin II antagonist
	Adult patients with heart failure and impaired left ventricular systolic function (left ventricular ejection fraction ≤ 40%) in addition to treatment with ACE inhibitors or when ACE inhibitors are not tolerated	

Table 4.4: The most frequently prescribed medications in the study: brands, active substances, indications and types

Category	Systolic b.p. [mmHg]	Diastolic b.p. [mmHg]	Health advice
Hypotension	<90	<60	Follow a healthy lifestyle.
Desirable	90-119	60-79	Follow a healthy lifestyle.
Prehypertension	120-139	80-89	Follow a healthy lifestyle.
Stage 1 Hypertension	140-159	90-99	Follow a healthy lifestyle, if desirable b.p. levels aren't reached within 6 months talk to the doctor about taking one or more medications.
Stage 2 Hypertension	160-179	100-109	Follow a healthy lifestyle. Talk to the doctor about taking more than one medication.
Hypertensive Crisis	$\geqslant 180$	$\geqslant 110$	Emergency care needed.

Table 4.5: Classification of hypertension stages for ages 18 and older (based on (MayoClinic, 2011a))

Chapter 5

Alarm Management and Optimization

The large number of false alarms was identified as one of the key factors limiting the applicability of telemonitoring systems. Consequently, alternative alarm management and decision support system optimization strategies were explored to effectively reduce the number of false alarms. The following approaches were developed and applied:

- Multi-threshold alarm flags (described in Section 5.1),
- Patient reference state determination (data smoothing) (described in Section 5.2),
- Trend analyses and visualisation (described in Section 5.5),
- Dynamic threshold type selection (described in Section 5.3),
- Optimal threshold setup (described in Section 5.4), and
- Analysis of weather influence on alarm management (described in Section 5.6.

5.1 Multi-threshold alarm flags

The cleaned data were used to develop several multi-threshold alarm flags supplementing single threshold exceeded alarms. Such flags were based upon counts of exceeded thresholds during the consecutive measurement days. Specifically, 2 or more exceeded thresholds and 3 or more exceeded thresholds were considered during 2 and 3 consecutive days. At the beginning of measurements, it was assumed that no alarms would occur during the preceding days for which no measurement records existed. Considering the increasing sum of true positive and true negative events the following alarm flags were introduced:

Alarm flag Level 1: 1 Day / $\geqslant 1$ exceeded thresholds,

Alarm flag Level 2: 3 Days / $\geqslant 2$ exceeded thresholds,

Alarm flag Level 3: 2 Days / $\geqslant 2$ exceeded thresholds,

Alarm flag Level 4: 3 Days / $\geqslant$ 3 exceeded thresholds,

Alarm flag Level 5: 2 Days / $\geqslant$ 3 exceeded thresholds.

Such new 5-level alarm importance indicator/flag was proposed to supplement the original MOBITEL algorithm towards more efficient and precise telemonitoring system.

Number of various physician actions was counted for each introduced alarm flag level and presented in Table 5.1. These results illustrated the outcomes of replacing the original alarms (flag level 1, 1 day / $\geqslant$ 1 exceeded thresholds) with a different alarm generation criteria (alarm flag levels 2 through 5). They showed between 28% and 78% reduction of false positive (FP) (false alarm) cases depending upon the alarm flag level (2 through 5, as compared to the original alarm flag level 1). The calculated false alarms included cases with exceeded alarm thresholds and no recorded physician response, as well as recorded responses: "No action" and "Threshold adjustment". Apart from reducing the false alarms, the increasing alarm flag levels also resulted in a reduction between 18% and 64% in the true positive (TP) (true alarm) cases, respectively. At the same time, the number of false negative events increased from 17 (0.2%) to 183 (2.3%) of the total 8114 events, corresponding to 0.3% and 2.4% of the total no alarm decisions, respectively. The false negative events represented the sum of physician actions "Patient contact", "Medication adjustment" and "Other action", occurring during the days with no exceeded alarm thresholds.

Physician responses / Type of alarms / Statistics	1 Day / ≥1 exceeded threshold	3 Days / ≥2 exceeded thresholds	2 Days / ≥2 exceeded thresholds	3 Days / ≥3 exceeded thresholds	2 Days / ≥3 exceeded thresholds
(1) No actions	1269 (77.5%)	874 (74.2%)	721 (73.2%)	523 (71.7%)	275 (69.3%)
(2) Threshold adjustment	108 (6.6%)	91 (7.7%)	79 (8.0%)	54 (7.4%)	28 (7.1%)
(3) Patient contact	135 (8.2%)	107 (9.1%)	97 (9.8%)	79 (10.8%)	54 (13.6%)
(4) Medication adjustment	89 (5.4%)	78 (6.6%)	68 (6.9%)	60 (8.2%)	39 (9.8%)
(5) Other actions	36 (2.2%)	28 (2.4%)	20 (2.0%)	13 (1.8%)	1 (0.3%)
True positive TP	260	213	185	152	94
False positive FP	1804	1301	1111	820	395
True negative TN	6033	6536	6726	7017	7442
False negative FN	17	64	92	125	183
Specificity *TN/(TN+FP)*	0.770	0.834	0.858	0.895	0.950
Sensitivity *TP/(TP+FN)*	0.939	0.769	0.668	0.549	0.339
Accuracy	0.776	0.832	0.852	0.884	0.929
Alarm flag	Level 1	Level 2	Level 3	Level 4	Level 5

Table 5.1: Results of the alarm occurrence due to multiple exceeded thresholds over 2 and 3 consecutive days, including specificity, sensitivity, accuracy and alarm flag levels

The results presented in Table 5.1 show the greatest specificity and smallest sensitivity obtained for an alarm indication based on three or more exceeded thresholds over two consecutive days. Nevertheless, accuracy of the overall alarm management algorithm increases with the increasing alarm flag levels, indicating that the physicians should pay more attention to the alarms marked with higher flag levels. Thus the alarm flags present additional information to physicians related to the alarm importance. High level of alarm flags would require immediate attention from the physicians.

Combining the criteria presented in Table 5.1 and assuming supremacy of higher alarm flag levels, Figure 5.1 shows an algorithmic procedure for setting up the introduced alarm flags. Applying the developed algorithm, distribution of flag levels in true, false and total alarms is obtained and presented in Figure 5.2. The numbers indicated within each of the alarm flag level bars represent counts of corresponding true, false and total alarm cases.

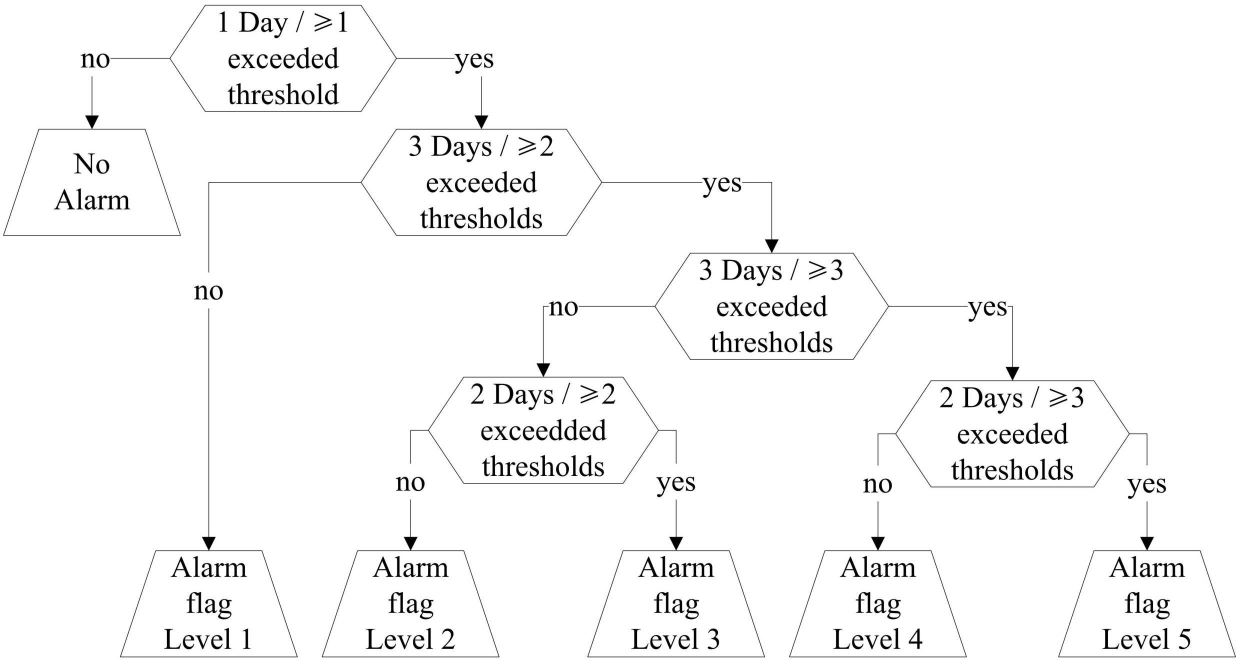


Figure 5.1: Procedure for setting up alarm flag levels for each patient

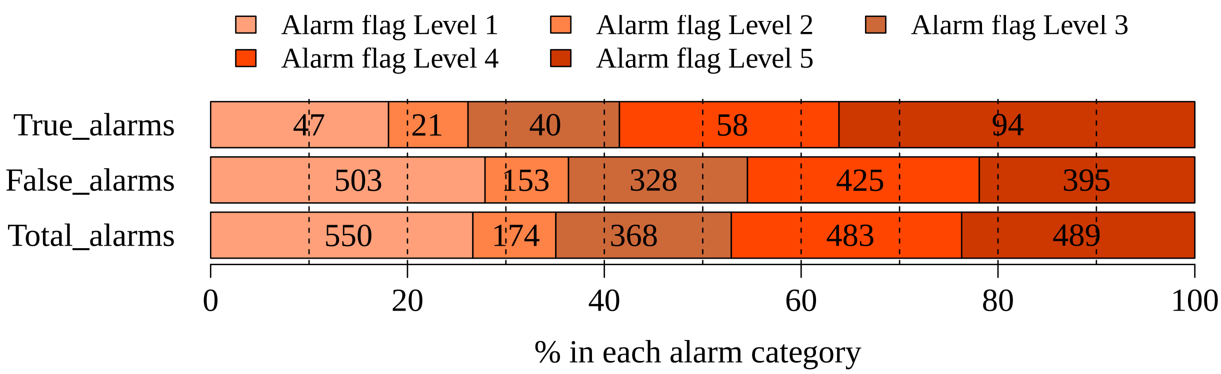


Figure 5.2: Portion of alarm flag levels 1 to 5 in true, false and total number of alarms

5.2 Patient reference state determination (data smoothing)

Apart from the enhanced interpretation of the existing alarms, via multi-threshold considerations, additional efforts focused on the design of new automated patient alarm management procedures. Automated patient alarm management implies that a mathematical decision model is used to improve accuracy of patient alarm generation. The first step in such alarm management was to obtain an indication of the patient's condition which is free from measurement noise, potential seasonal influences and daily, hourly or any other periodic cycles. Such condition will be further considered as patient's reference state. In order to obtain the indication of the patient reference state several procedures were considered:

- Moving average,
- Savitzky-Golay filter,
- Fast Fourier Transformation (FFT), and
- Kalman filter.

Moving average was calculated using average daily measurements of telemonitored parameters over several days prior to the calculation day. The number of considered days varied to find the optimal reference state indication. For example, Figure 5.3 shows the original and smoothed reference state data for patent's systolic blood pressure measurements using 3, 5 and 14 point moving average, calculated according to Equation 5.1 (n_L = 2, 4, and 13, respectively).

$$g_t = \sum_{n=-n_L}^{0} c_n f_{t+n} \tag{5.1}$$

Here g_t represents the reference state point, f_t is the measurement point, whereas n_L denotes the number of points equidistant in time "to the left", i.e. earlier, of the current measurement point. Coefficients c_n correspond to each of the considered measurement points. In the case of moving average, all c_n coefficients are constant and equal to $1/(n_L + 1)$.

Changing the values of coefficients c_n, the same Equation 5.1 can be applied to design Savitzky-Golay filters. For example, 5 point Savitzky-Golay filter will have the following coefficients (Press et al., 2007, p. 769): $c_{-4} = 0.086$, $c_{-3} = -0.143$, $c_{-2} = -0.086$, $c_{-1} = 0.257$, $c_0 = 0.886$. For each measurement point, f_t, Savitzky-Golay reference state estimate given by the aforementioned coefficients is based on the quadratic function approximation using the least squares fit over 5 data points within the moving window. Figure 5.4 shows the original and smoothed reference state data for patent's systolic blood pressure measurements using Savitzky-Golay and FFT filters.

Figures 5.5 and 5.6 show the distribution of systolic blood pressure residuals for a selected patient in the cases of 14 day moving average and FFT filtering, respectively.

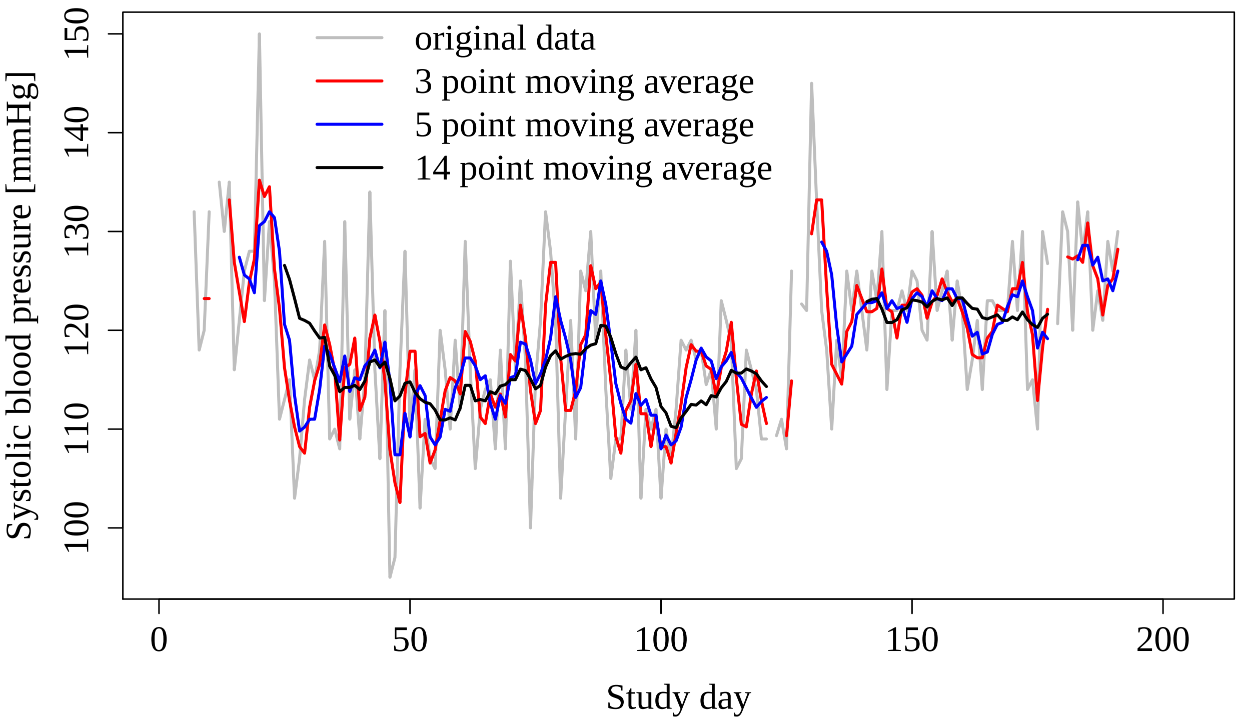


Figure 5.3: Original and smoothed systolic blood pressure data with 3, 5 and 14 point moving average (Patient ID = 8,587)

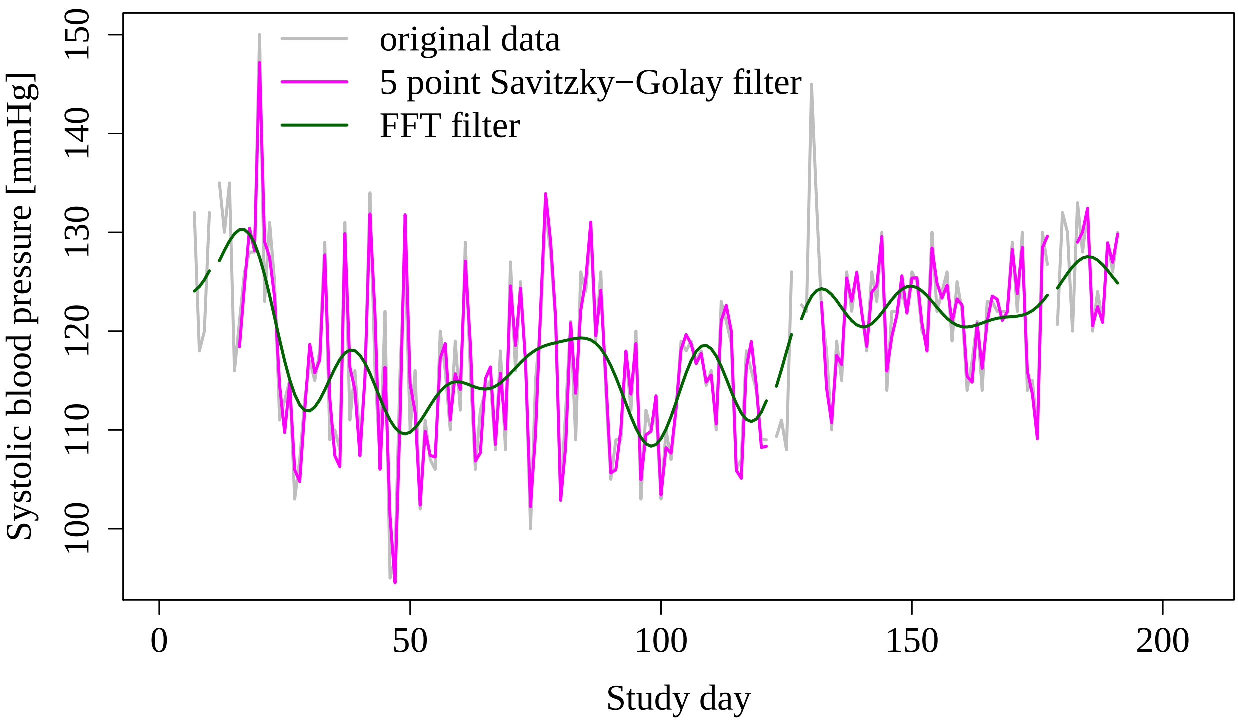


Figure 5.4: Original and smoothed systolic blood pressure data with Savitzky-Golay and FFT filters (Patient ID = 8,587)

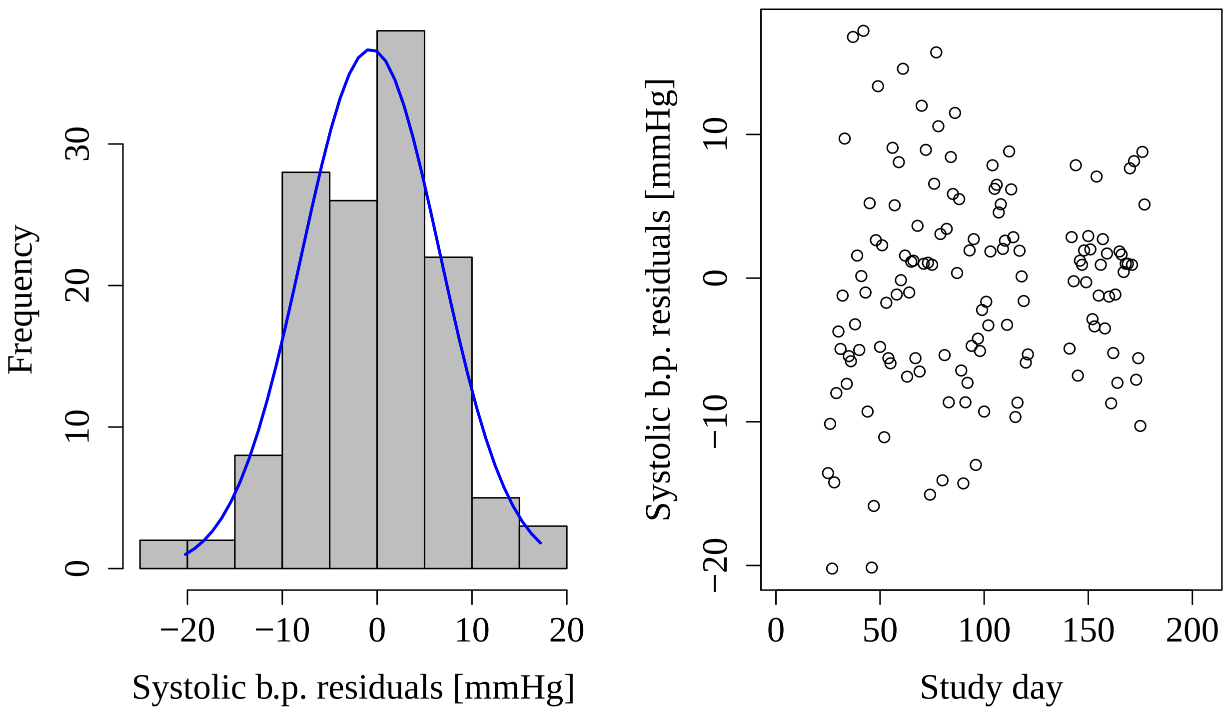


Figure 5.5: Histogram of systolic blood pressure residuals with normal distribution curve (left) and residual plot (right) for 14 day moving average filtering (Patient ID = 8,587)

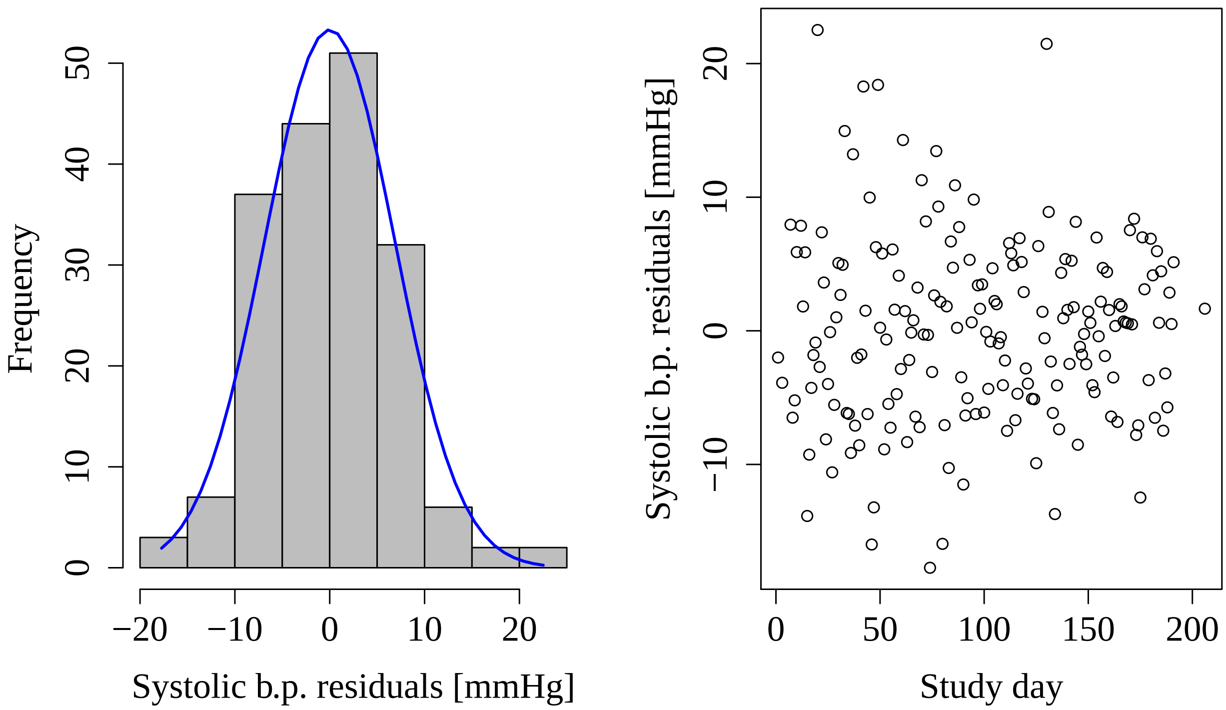


Figure 5.6: Histogram of systolic blood pressure residuals with normal distribution curve (left) and residual plot (right) for FFT filtering (Patient ID = 8,587)

The residuals were calculated as a difference between the average daily measurements and filter estimates.

The `fft` algorithm implementation in R results in real and imaginary vectors of the same length as the input vector. Thus, the second halves of those vectors are mirror images of the first halves minus the first element. Figure 5.7 (left) illustrates such symmetry in the case of real parts of the transformed data.

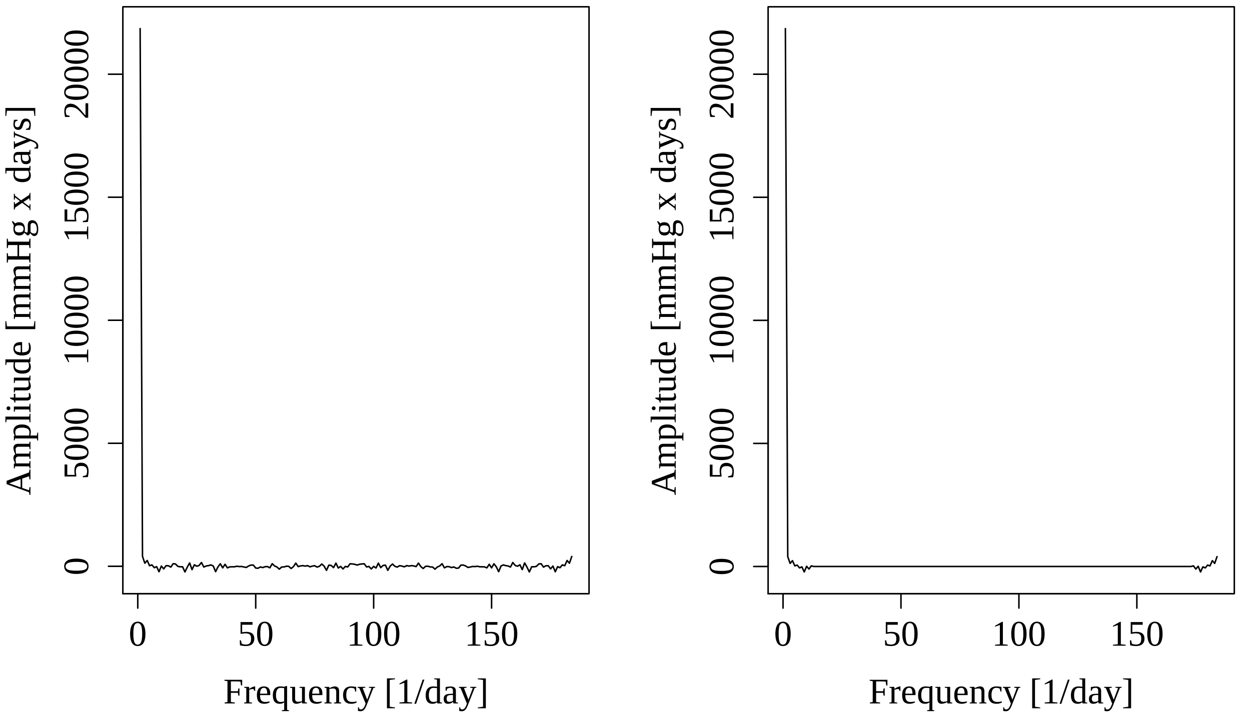


Figure 5.7: Real parts of originally transformed (left) and filtered (right) systolic blood pressure data using FFT (Patient ID = 8,587)

Figure 5.7 (right) presents noise filtering in the data by setting the high frequency (above 12) components equal to zero. The reference state presented in Figure 5.4 is obtained applying the inverse FFT to the filtered frequency domain data and can be adjusted by changing the number of annulated points in the frequency domain. FFT result presented in the figure initially considered application on the whole time interval of measurements, rather than using smaller intervals. The idea was to evaluate the approach on the longer intervals first, as a starting point for developing an algorithm using shorter time spans, e.g. one week, which would be necessary for the real world applications in patient alarm management. Developments of such advanced algorithms will be discussed later in this Chapter.

To quantify effectiveness of the considered filters, variances were compared to the telemonitoring data as measures of uncertainty of the obtained results (for the presented sample patient measurements var = 76.6 mmHg^2). The reference state values obtained with the 5 point Savitzky-Golay filter (var_{SG} = 67.5 mmHg^2) showed substantial variability, as did 3 point moving average (var_3 = 43.6 mmHg^2). The 5 point moving average (var_5 = 31.8 mmHg^2) showed somewhat better results, while the most promising was the application of FFT (var_{FFT} = 29.2 mmHg^2) and 14 point moving average (var_{14} = 15.8 mmHg^2).

The FFT underlying assumption of periodicity is not observed in the recorded patient measurements, as indicated by a sample autocorrelation plot of systolic blood pressures, presented in Figure 5.8 for a selected patient. Similar absence of periodicity was observed for all the other considered parameters and patients. Consequently, FFT equations could not be used to effectively make predictions of the future patient reference state conditions, but only to calculate the reference state based on the existing measurements. Figure 5.8 also shows non-randomness in the data, as statistically significant positive autocorrelations were observed for several lags. The observed non-randomness of the data in the autocorrelation plot justifies application of statistical models to make patient reference state predictions.

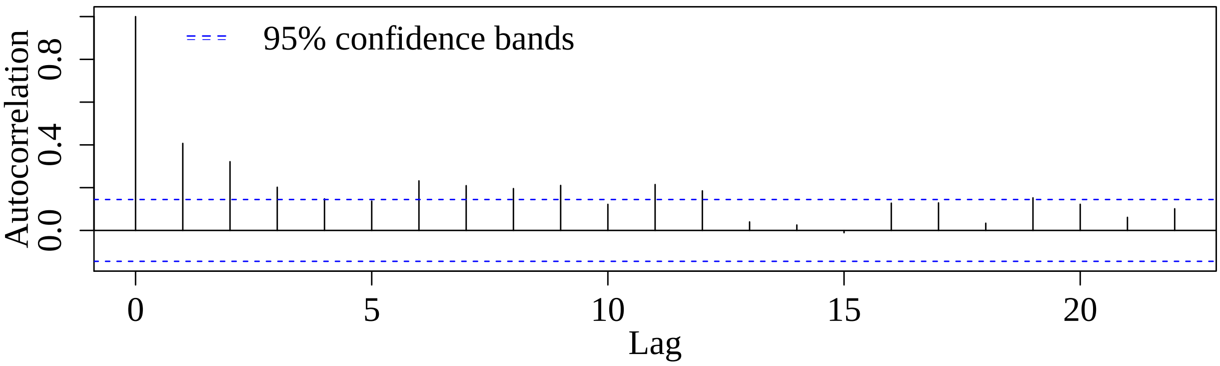


Figure 5.8: Autocorrelations of systolic blood pressure (Patient ID = 8,587) (values outside of the dashed range are statistically significant)

The downside of 14 point moving average filtering was a relatively long moving window size of 2 weeks in order to make valid predictions. If a patient would skip measurements during some of the 14 days prior to the current measurement, the accuracy of the filter would be adversely affected. Consequently, the aforementioned variance is based only on the points having 14 consecutive day predecessors necessary for calculating the moving average, as illustrated in Figure 5.3. Ideally, a desirable procedure would show similar variability of the reference state estimation to the 14 day moving average, but using shorter prior window size. As an alternative, removing the measurement gaps and treating the measurement points around the gaps as consecutive, would provide at least some estimate in the case of irregular data.

The final investigated procedure for obtaining patient reference state included Kalman filtering. Kalman filter was calculated using a built in `kfilter` function of the sspir package in R, which calculates Kalman filters based on the state space formulation. In this formulation several parameters are required in order to obtain valid results of Kalman filter transformation: initial condition or starting value, m0, allowed variance of the Kalman filter results, c0, as well as two transformation matrices, Wmat and Vmat. Additional matrices, Fmat and Gmat, provide the capabilities to consider non-linear models and time varying Kalman filter parameters, which are not used in the current approach. Consequently, Fmat and Gmat were kept constant and equal to 1. To obtain optimal Kalman filter transformation results, orders of magnitude for parameters Vmat and Wmat were determined. The values of Vmat varied between 1 and 100 and Wmat between 0.001 and 1. Figures 5.9, 5.10, 5.11, and 5.12 illustrate the changes in Kalman filter results for Patient ID 8,587 measurements of systolic blood pressure, due to varying m0, c0, Wmat and Vmat parameters, respectively. The considered parameters were varied one at a time while keeping all the rest of algorithmic inputs constant. The black line in all the Figures 5.9 through 5.12 shows initial case considering the following combination of parametric values: m0 = 120, c0 = 10, Wmat = 0.1, and Vmat = 10.

Figures 5.9 and 5.10 show that parameters m0 and c0 only influence the beginning phase of Kalman filtering application. After about a dozen measurements, all the filter curves coincided with each other, regardless of the m0 and c0 parameter values. Also, Figure 5.10 shows almost no difference between the curves in the cases when c0 was 10 and 100. Consequently, the value of c0 = 10 was selected as sufficient initial variance in all the considered Kalman filtering applications.

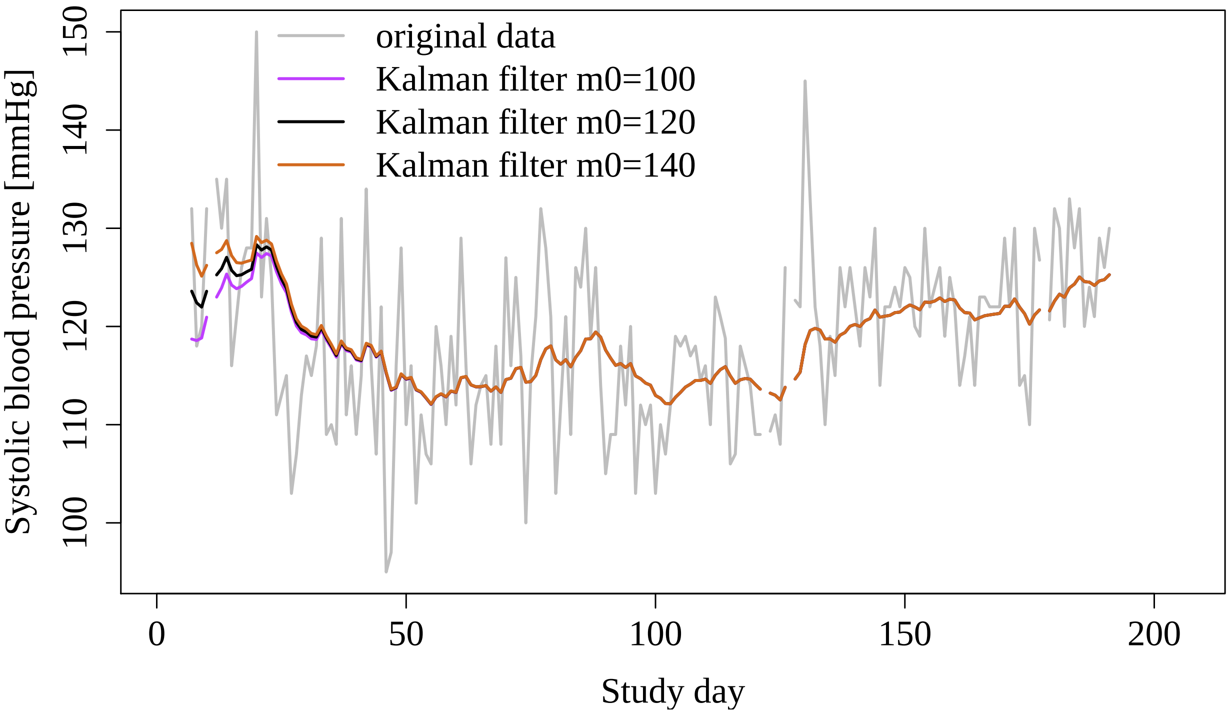


Figure 5.9: Influence of parameter m0 on Kalman filtering of systolic blood pressure (Patient ID = 8,587)

Contrary to m0 and c0, Figures 5.11 and 5.12 show that parameters Wmat and Vmat, mostly influence the later results of Kalman filtering, after the initial application phase. Also, effects of decreasing Wmat seem to be comparable to the effects of increasing Vmat leading to higher smoothness of the resulting curve, and vice versa.

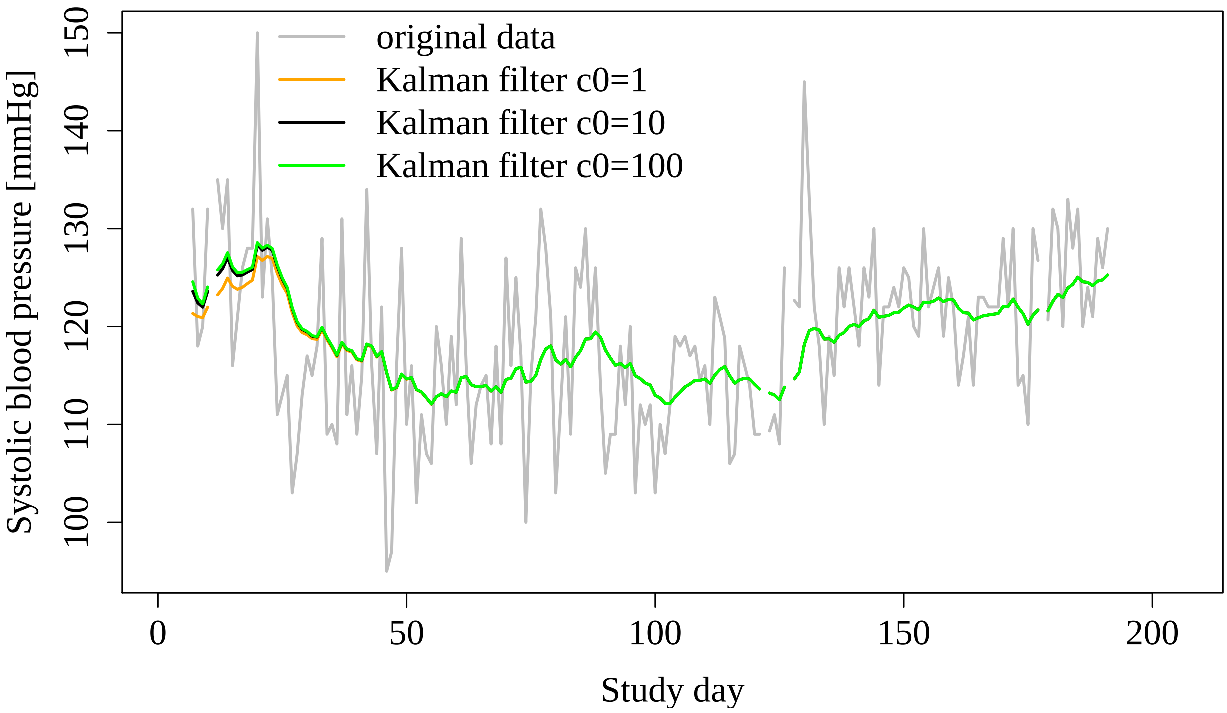


Figure 5.10: Influence of parameter c0 on Kalman filtering of systolic blood pressure (Patient ID = 8,587)

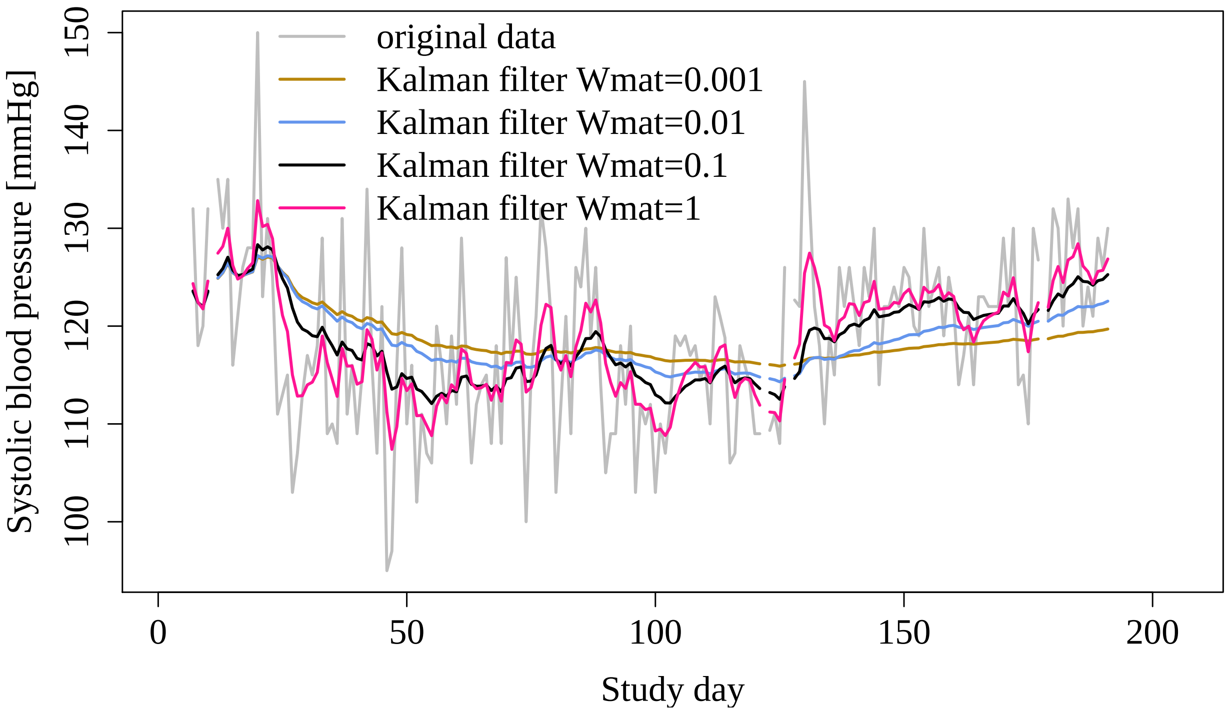


Figure 5.11: Influence of parameter Wmat on Kalman filtering of systolic blood pressure (Patient ID = 8,587)

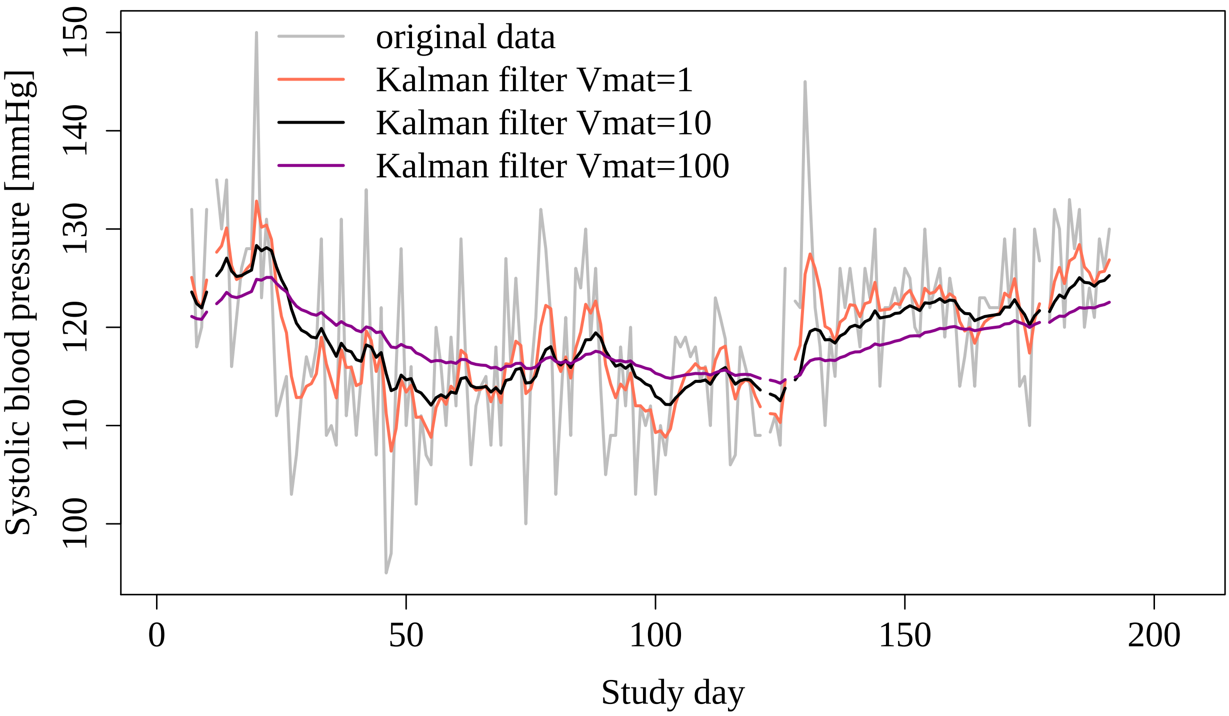


Figure 5.12: Influence of parameter Vmat on Kalman filtering of systolic blood pressure (Patient ID = 8,587)

Based on the analyses of all patient measurement records, Table 5.2 presents the most suitable combination of values identified for obtaining adequate Kalman filtering of the original measurements into the reference state. The initial mean values, m0, for the systolic and diastolic blood pressure (in mmHg), heart rate (in bpm) and weight (in kg) are assumed as typical for healthy individuals.

Parameter	Systolic b.p.	Diastolic b.p.	Heart rate	Weight
m0	120	80	70	75
c0	10	10	10	10
Vmat	100	100	100	1
Wmat	0.01	0.001	0.01	0.01

Table 5.2: Selection of Kalman filter parameters for patient reference state estimation

Using the identified parameters from Table 5.2, Figure 5.13 presents Kalman filtering of systolic blood pressure measurements for a selected patient. With such selection, variability of the calculated patient reference state was the lowest of all the considered data smoothing approaches (var_K = 4.3 mmHg^2).

Figure 5.14 shows the distribution of systolic blood pressure residuals for a selected patient in the case of Kalman filtering. As in the applications of 14 day moving average and FFT filters, presented in Figures 5.5 and 5.6, Kalman filtering residuals also appear nearly normally distributed and symmetric around zero, independent of the study days

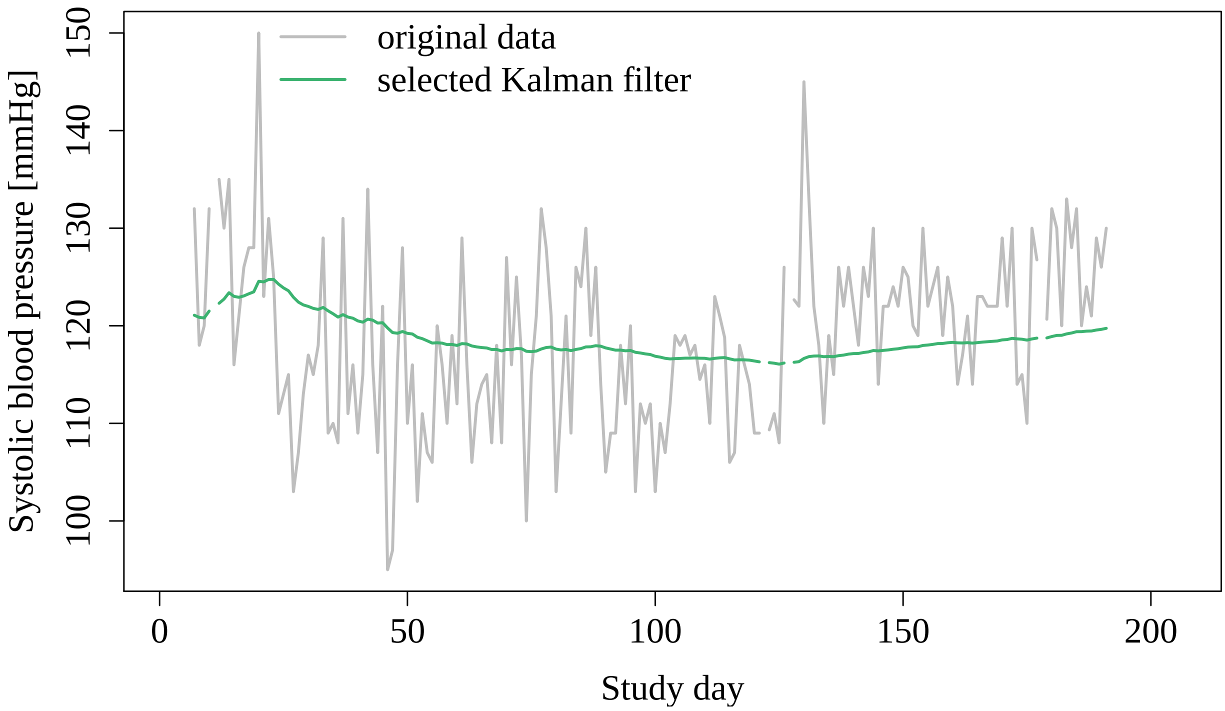


Figure 5.13: Original and smoothed systolic blood pressure data with Kalman filter (Patient ID = 8,587)

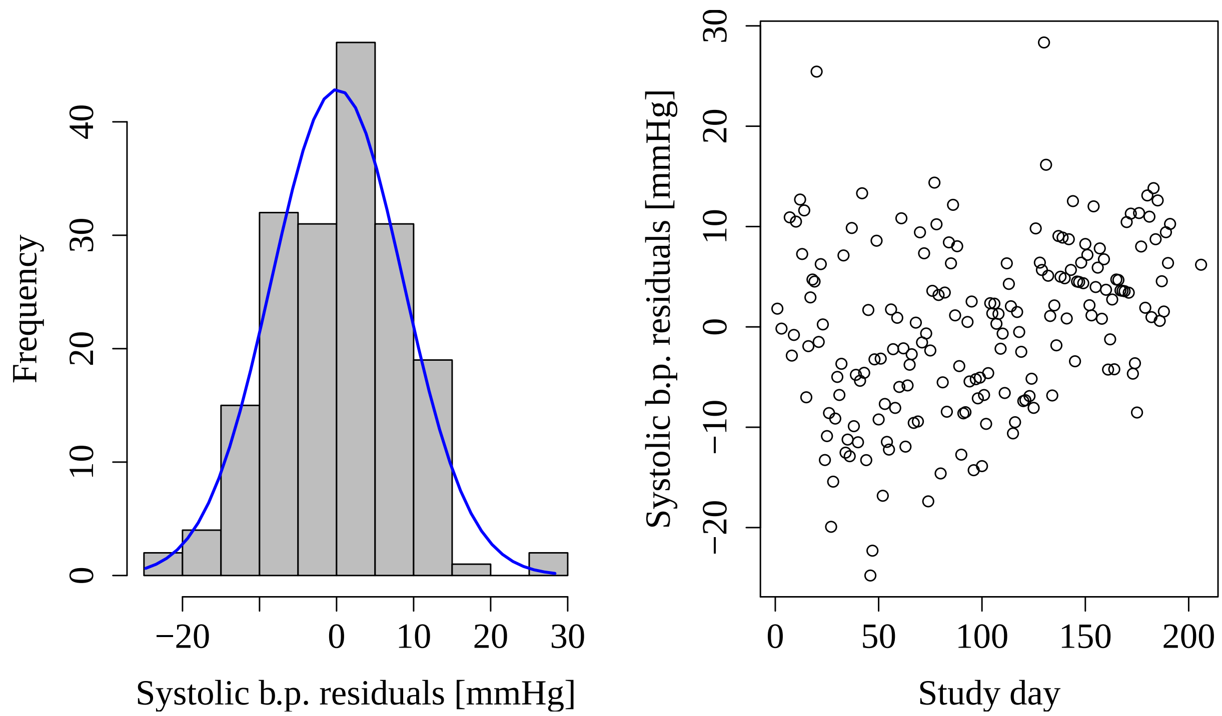


Figure 5.14: Histogram of systolic blood pressure residuals with normal distribution curve (left) and residual plot (right) for Kalman filtering (Patient ID = 8,587)

and uncorrelated, as no patterns or clustering could be observed in Figure 5.14. As the residual plot is relatively homogeneous around zero, homogeneity of residual variances and its independence of the study day can be assumed. Similar behaviour was observed in the case of all considered patients.

5.3 Dynamic alarm thresholds and monitoring window size

Evaluation of the identified methods for determination of patient reference state was performed by comparison of the alarms generated using such reference state to the recorded data. Various mathematical models were considered for the alarm generation specifying thresholds around the patient reference state and calculating occurrence of alarms when measurements exceeded the defined threshold values. As thresholds depended and followed the patient reference state in time, they were considered dynamically adjusted. The thresholds were obtained by increasing and decreasing the reference state values for a certain fixed or variable adjustment. Application of fixed adjustments to the reference state values to obtain upper and lower thresholds was tested versus the variable threshold adjustments by examining the number of newly generated alarms matching the MOBITEL records. In the case of fixed increases/decreases of the reference state to obtain the thresholds, initially the same quantity was added/subtracted for each patient and for each reference state of a certain measured variable. Various increases/decreases were tested starting from the most commonly occurring values derived from the original monitoring data. Table 5.3 represents mean and median ranges for differences between thresholds and reference states obtained for all the monitoring records of all patients and all the measured variables. As the recorded data did not contain information about the patient reference state, the reference state was assumed to be in the center of the range between the lower and upper considered thresholds during the measurements. The values in Table 5.3 were obtained halving all the differences between the upper and lower thresholds recorded by the physicians over the course of patient telemonitoring. Figure 5.15 illustrates the process in the case of systolic blood pressure telemonitoring of a selected patient.

Measured variable	Absolute threshold range		Relative threshold range	
	Mean	Median	Mean	Median
Systolic b.p. [mmHg]	±28	±28	±22%	±23%
Diastolic b.p. [mmHg]	±19	±20	±25%	±25%
Heart rate [bpm]	±24	±25	±31%	±33%
Weight [kg]	±6	±4	±7%	±5%

Table 5.3: Mean and median absolute and relative threshold ranges around reference state from the recorded data

Apart from the described fixed adjustments of the reference state to obtain the alarm thresholds, variable adjustments depending upon the reference state value were also considered. One option for such adjustment was to increase or decrease the reference state

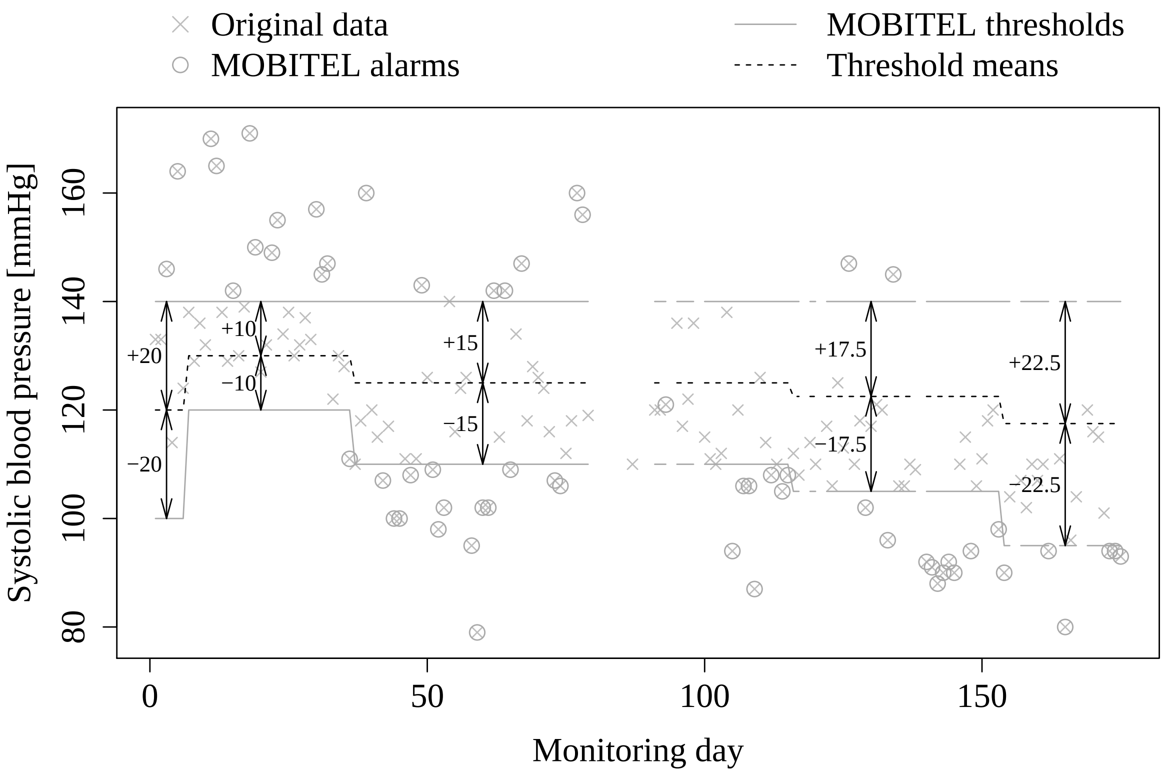


Figure 5.15: Determination of absolute threshold ranges based on the systolic blood pressure telemonitoring records of a selected patient (Patient ID = 12,119)

by a certain percentage of its current value and thus obtain the upper and lower alarm thresholds. As in the case of absolute threshold ranges, considering the relative threshold ranges in the recorded data speeded up the process of finding the most appropriate thresholds. Table 5.3 also presents mean and median relative threshold ranges for a variety of measured variables in the recoded data.

Figure 5.16 illustrates the fixed absolute and relative threshold bounds for systolic blood pressure measurements of a selected patient. The reference state in Figure 5.16 was obtained applying 14-day moving average filtering, after removing the measurement gaps and considering the adjacent measurements around the gaps as consecutive. Upon calculation of the moving average, the gaps were restored as in the original measurements.

Table 5.3 was based on the recorded data including the actual thresholds physicians considered during the telemonitoring. The thresholds were provided as physician inputs to the telemonitoring system rather than calculated by a mathematical model. Another alternative for the dynamic threshold adjustments was to use standard deviation of measurements over the previous days as an indicator of upper and lower threshold bounds in relation to the reference state. The background for consideration of standard deviations is Chebyshev's Theorem.

As an illustration, Figure 5.17 presents the upper and lower threshold bounds determined by increasing and decreasing the reference state for a certain number of measurement standard deviations over 14-day moving window size. The ranges within 1, 2 and 3

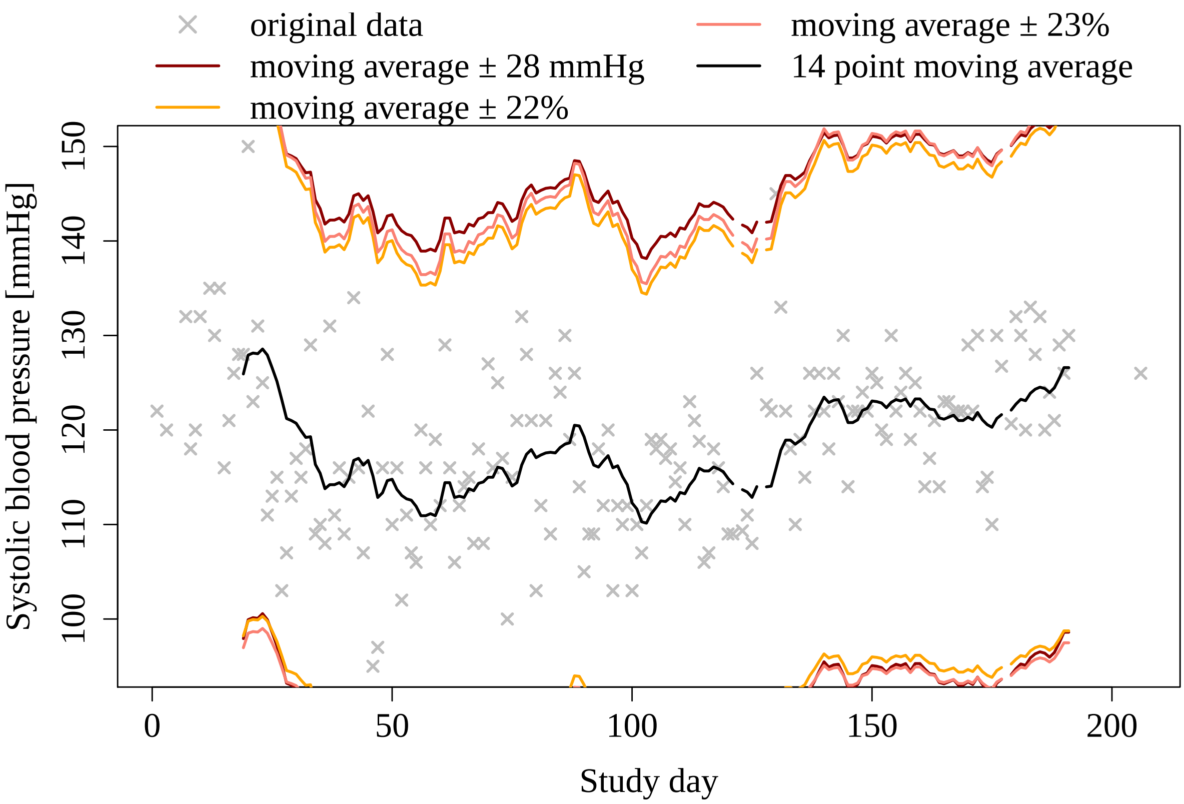


Figure 5.16: Systolic blood pressure measurements and 14-day moving average reference state with dynamic thresholds based on fixed absolute and relative bounds for a selected patient (Patient ID = 8,587)

standard deviations around the reference state are presented.

When considering reference state estimation depending upon the monitoring data selection window size (e.g. moving average), such window size should be additionally varied together with the threshold bounds to yield optimal combination of parameters. Different combinations of upper and lower threshold bounds around the reference state could be used in conjunction with constant threshold limits that will always trigger the alarm occurrence regardless of the reference state value.

In the case of Kalman filtering, which is independent of the data window size after the filter set-up period (corresponding to a couple of initial measurements), only combination of threshold values influences algorithmic accuracy. Receiver operating characteristic (ROC) curves presented in Figure 5.18 show results of varying different threshold parameters around Kalman filtering based reference state estimation for systolic blood pressure. The figure separates influences of lower and upper thresholds on sensitivity and specificity. The presented baseline curves (in pink and dark blue) are obtained varying threshold bounds between ±5 and ±30 mmHg with unit increments around the reference state. The same ranges with additional introduction of constant threshold limits that will always trigger alarm occurrence result in increased sensitivities with minor decrease in specificities. The Figure compares influences of 160 mmHg and 170 mmHg constant

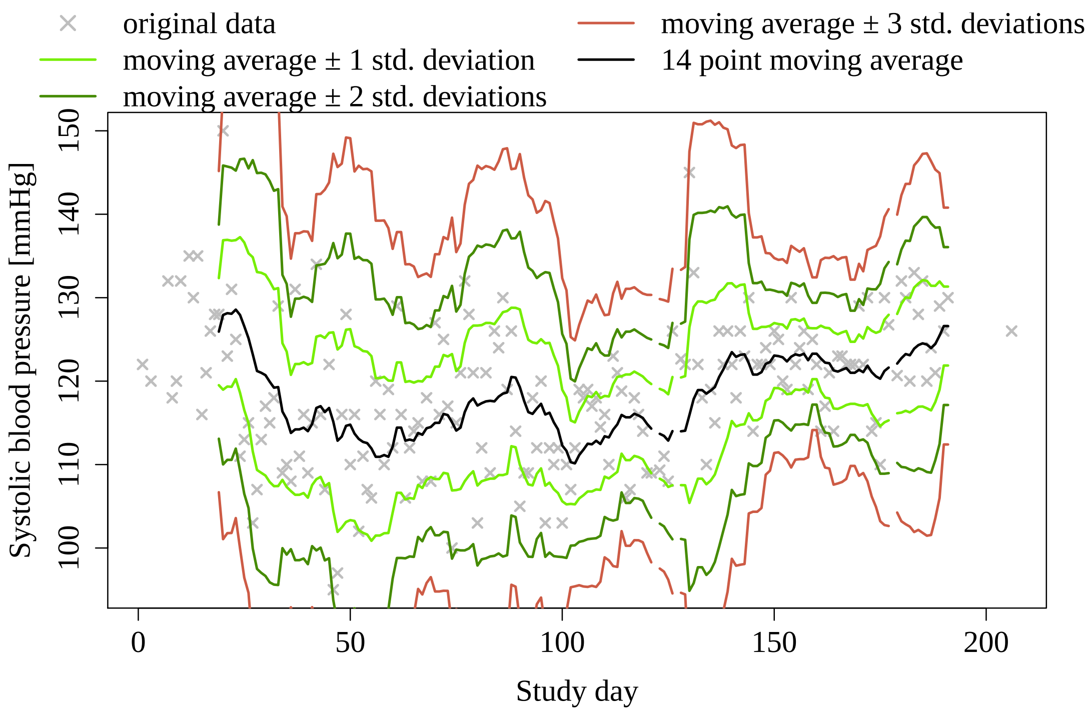
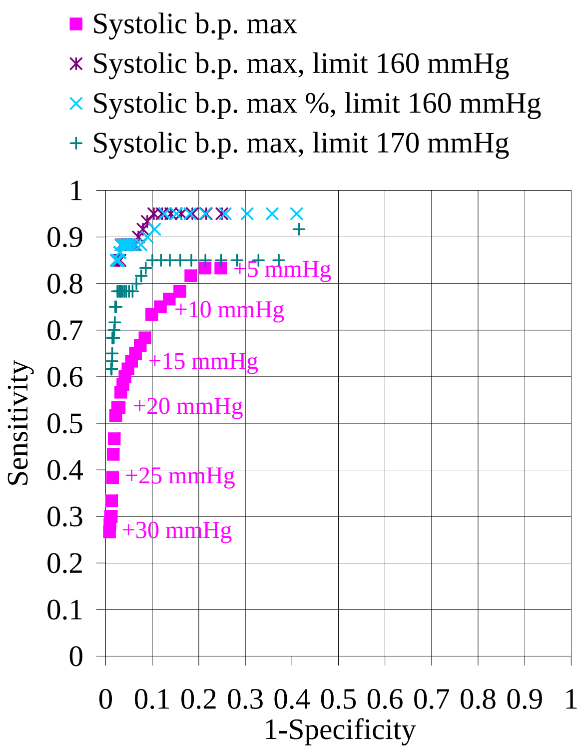


Figure 5.17: Systolic blood pressure measurements and reference state with dynamic thresholds based on 14-day standard deviation for a selected patient (Patient ID = 8,587)



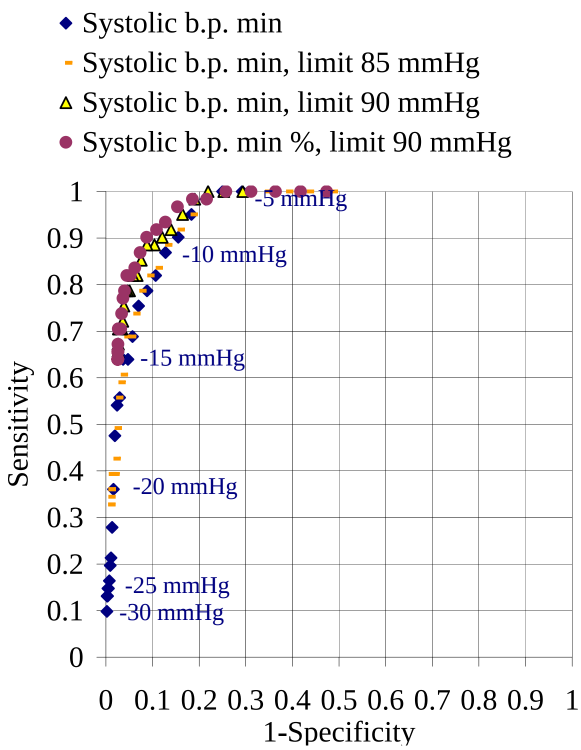


Figure 5.18: ROC curves for varying upper (left) and lower systolic blood pressure thresholds (right)

upper threshold limits, as well as 85 mmHg and 90 mmHg constant lower threshold limits over the considered threshold ranges around the reference state estimate.

ROC curves similar to those presented in Figure 5.18 for systolic blood pressure, can be constructed for all the other monitored parameters: diastolic blood pressure, heart rate and weight. The cumulative effects on alarm occurrence due to the selection of upper and lower threshold bounds and limits for all of these parameters have a negative impact on the overall specificity, with limited benefits towards the increase in sensitivity. Namely, reducing the threshold range beyond a certain value for any of the considered parameters will trigger so many false alarms to outweigh the potential benefits of including omitted true alarm cases. Particularly, such effects can occur as the ROC curves resulting from varying thresholds of individual parameters, such as those presented in Figure 5.18, do not take into account potential true alarms that might have already been encompassed by the specified thresholds of another parameter. For example, Figure 5.18 only presents impacts of systolic blood pressure bounds and limits on sensitivity and specificity, but when considering also diastolic blood pressure, heart rate and weight thresholds, the overall sensitivity and specificity values will be very different. Consequently, careful consideration of impacts on specificity and sensitivity is required when selecting the threshold bounds and limits based on the presented ROC curves. Threshold bound and limit adjustment should be made for the monitored parameter which yields the highest increase in sensitivity with the smallest decrease in specificity. After each such adjustment, a new calculation of the overall sensitivity and specificity values should be performed. Plotting the overall results obtained in this way a ROC curve presented in Figure 5.19 can be constructed.

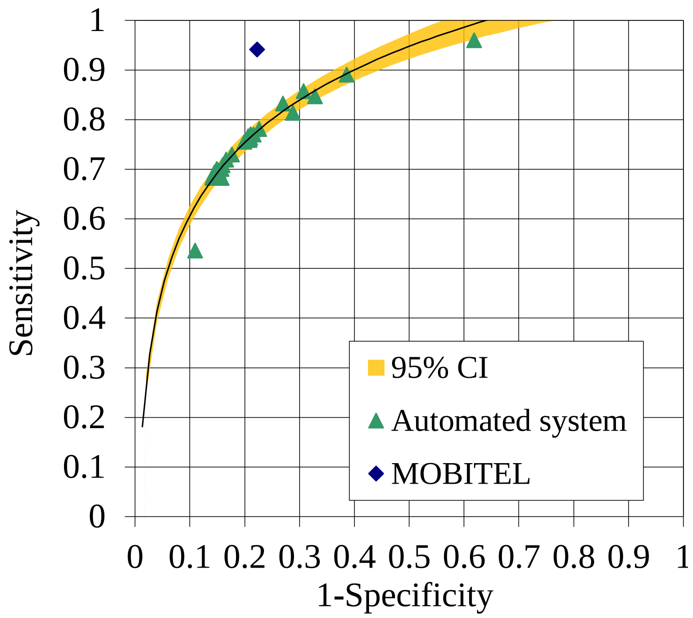


Figure 5.19: ROC curve of the automated alarm generation algorithm in comparison to the recorded MOBITEL data

The presented ROC curve results included variation of absolute and relative threshold bounds and threshold limits around the Kalman filtering based reference state. Equation

5.2 was found to approximate well ($R^2 = 0.98$) the presented ROC curve, where the regression coefficients are given in Table 5.4. Figure 5.19 also shows 95% confidence interval around the estimated regression curve.

Coefficient	Value	Std. error	t-Test	α
k_0	1.094	±0.011	103	<0.0001
k_1	0.212	±0.006	33	<0.0001

Table 5.4: ROC curve regression coefficients

Regression calculations excluded the two outlier points (with minimum and maximum sensitivities) in order to reach no statistical evidence of correlation in the residuals according to the Durbin-Watson statistic of 1.67 (higher than the upper critical value of 1.47). At the same time Lilliefors test indicated no statistically significant deviation in the frequency distribution of the residuals compared to normal distribution ($p = 0.12 > 0.05$), while F-value of 1067 ($p < 0.0001$) was higher than 100 indicating that the model was valid. Standard error presented in Table 5.4 was two orders of magnitude lower than the calculated regression coefficients with the t-Test values significantly different from zero ($\alpha < 0.0001$), indicating significant contribution of parameters to the model.

$$Sensitivity = k_1 \cdot \ln(1 - Specificity) + k_0 \tag{5.2}$$

Integrating Equation 5.2, the area under the ROC curve was approximately calculated to amount 0.85. The point on the curve closest to the targeted ideal value (0, 1) in the left upper corner of the curve (specificity = 1, sensitivity = 1) is identified to correspond to the combination of algorithmic threshold input parameters presented in Table 5.5. Such combination of algorithmic inputs can be considered as optimal in terms of the sensitivity/specificity tradeoff.

Measured variable	Threshold bounds		Threshold limits	
	Upper	Lower	Upper	Lower
Systolic b.p. [mmHg]	+37	−29	160	90
Diastolic b.p. [mmHg]	+13	−14	100	45
Heart rate [bpm]	+16	−20	120	50
Weight [kg]	+1.3	−1.3	135	45

Table 5.5: Optimal thresholds in the automated alarm generation algorithm

Figure 5.20 presents a sample comparison between the MOBITEL and automated alarm generation results for a selected patient. In the presented case the automated threshold adjustment utilizes fixed upper threshold limit for the systolic blood pressure of 160 mmHg, lower threshold limit of 90 mmHg, and illustrates a transition from the variable upper and lower threshold bounds following the reference state to the limiting value when dynamic threshold adjustment around the reference state would exceed the limits. Figure 5.20 also illustrates a large number of false alarms occurring in MOBITEL telemonitoring and reductions with the introduction of dynamic alarms. Note that a

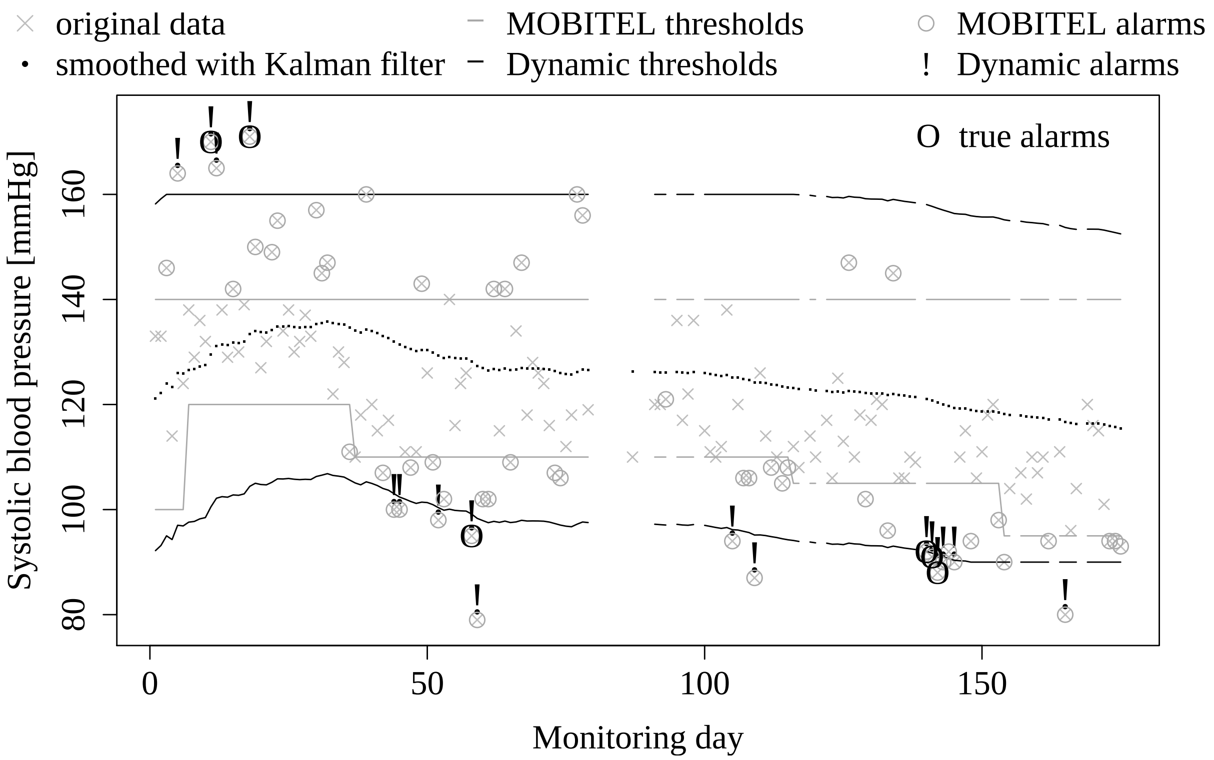


Figure 5.20: Comparison between MOBITEL and automated alarm management applied on systolic blood pressure data of a selected patient (Patient ID = 12,119)

MOBITEL alarm occurring on the 93rd monitoring day is not a mistake, although the measurement appears to fall within the specified thresholds. Namely, this blood pressure value is obtained as a result of an averaging of two measurements conducted during the same day, one of which exceeded the specified MOBITEL thresholds. The averaging is used to provide regular single daily monitoring records, as described in Section 3.1, while the original alarm indications are maintained and used in further analyses.

Table 5.6 shows the performance comparison between the original MOBITEL and automated threshold adjustment algorithm using the identified parameters presented in Table 5.5.

	Physician responses / Type of alarms / Statistics	MOBITEL algorithm	Automated alarm generation algorithm
TRUE ALARMS	Patient contact	133	107
	Medication adjustment	89	70
	Other action	36	34
	Total true alarms (TP)	258	211
FALSE ALARMS	No action	1233	688
	Threshold adjustment	108	56
	No recorded response	387	890
	Total false alarms (FP)	1728	1634
NO ALARMS	False negative (FN)	16	63
	True negative (TN)	6016	6110
	Total no alarms ($TN + FN$)	6032	6173
	$Specificity = \frac{TN}{TN+FP}$	0.777	0.789
	$Sensitivity = \frac{TP}{TP+FN}$	0.942	0.770
	$Accuracy = \frac{TP+TN}{TP+TN+FP+FN}$	0.782	0.788

Table 5.6: Comparison between the MOBITEL and automated alarm generation results

5.4 Algorithmic threshold selection model

In certain cases might be beneficial to specify such combination of algorithmic thresholds to reach higher sensitivity at the expense of specificity and vice versa. To ease the selection of suitable threshold bounds and evaluate the effect of a particular threshold selection, statistical models were constructed between specificity and sensitivity as target variables and all upper/lower threshold bounds as descriptors. During the variation of threshold bounds, the threshold limits presented in Table 5.5 were kept constant as their selection was considered optimal based on the ROC curves for individual physiological measures, such as those presented in Figure 5.18. In order to select suitable statistical model, variance inflation factors were calculated to detect possible multicollinearities between the different threshold bounds, as presented in Table 5.7. The calculated variance inflation factors above 10 indicated high collinearity between the variables. Due to the high collinearity direct application of multivariate linear regression model was not suitable. Principle component regression was used instead. In such approach multivariate linear regression was applied after using principle component analysis to obtain reduced variable domain space of mutually independent principle components.

Principle component regression was applied on both sensitivity and specificity as target variables using the standardized dataset of the constructed ROC curve in Figure 5.19. In both cases two principle components were selected (PC_1 and PC_2), having the eigenvalues higher than 1. Such vectors explained 86% of the total variances, whereas F-values higher than 100 (140 for specificity and 315 for sensitivity models) indicated that the models were valid. However, the resulting coefficients of determination (quality of fit) were different

i	Threshold bounds V_i	Variance inflation factor VIF_i
1	Systolic b.p. upper	7.013
2	Systolic b.p. lower	18.047
3	Diastolic b.p. upper	78.501
4	Diastolic b.p. lower	45.888
5	Heart rate upper	4.141
6	Heart rate lower	1.963
7	Weight upper	50.779
8	Weight lower	57.071

Table 5.7: Variance inflation factors of the considered threshold bounds

for specificity and sensitivity as target variables, amounting to 0.93 and 0.97, respectively. Consequently, the sensitivity model was chosen to describe the relation with algorithmic thresholds, given by Equations 5.3 and A.10.

$$PC_j = \sum_{i=1}^{8} k_{ij} \frac{V_i - \mu_i}{\sigma_i}, j \in \{1, 2\} \tag{5.3}$$

$$Sensitivity = p_0 + p_1 \cdot PC_1 + p_2 \cdot PC_2 \tag{5.4}$$

Values of coefficients k_{i1} and k_{i2} from Equation 5.3 are presented in Table 5.8, together with mean, μ_i, and standard deviation, σ_i, of algorithmic threshold inputs, V_i, used in their standardization. Also presented in Table 5.8 are ranges of algorithmic thresholds considered in the 24 member large data sets. Table 5.9 presents the principle component regression coefficients used in Equation A.10, together with their standard errors, t-Test values and levels of significance, α.

	Threshold variable	PC_1, PC_2 loadings		Mean	Std. deviation	Range
i	V_i	k_{i1}	k_{i2}	μ_i	σ_i	[$\min_i$, $\max_i$]
1	Systolic b.p. upper	0.3599	-0.3619	25.96	8.05	[7, 37]
2	Systolic b.p. lower	-0.4018	0.1380	-25.17	6.89	[-29, -9]
3	Diastolic b.p. upper	0.4105	-0.1589	12.00	2.15	[5, 14]
4	Diastolic b.p. lower	-0.4014	0.1647	-13.29	1.46	[-14, -9]
5	Heart rate upper	0.3494	0.1335	15.83	1.24	[11, 18]
6	Heart rate lower	-0.3114	0.2286	-19.46	2.65	[-29, -13]
7	Weight upper	0.2846	0.6011	1.51	0.41	[0.7, 2.2]
8	Weight lower	-0.2817	-0.6052	-1.57	0.53	[-2.4, -0.7]

Table 5.8: Loadings, means, standard deviations and range of algorithmic threshold variables for principle component transformation

For all the determined principle component regression coefficients standard errors were at least one order of magnitude lower than the calculated regression coefficients, while the t-Test values were significantly different from zero ($\alpha < 0.0001$). Such evidence indicated significant contribution of both principle components, PC_1 and PC_2, to the model.

Coefficient	Value	Std. error	t-Test	α
p_0	0.770	±0.003	290	<0.0001
p_1	-0.028	±0.001	-24	<0.0001
p_2	-0.017	±0.002	-8	<0.0001

Table 5.9: Principle component regression coefficients

Combining Equations 5.3 and A.10 the final relation between sensitivity and algorithmic thresholds could be established by Equation 5.5.

$$Sensitivity = r_0 + \sum_{i=1}^{8} r_i \frac{V_i - \mu_i}{\sigma_i} \tag{5.5}$$

Equation 5.5 used regression coefficients, r_i, defined in Table 5.10.

i	Threshold variable V_i	Regression coefficient r_i
0	–	0.76992
1	Systolic b.p. upper	-0.00397
2	Systolic b.p. lower	0.00890
3	Diastolic b.p. upper	-0.00879
4	Diastolic b.p. lower	0.00844
5	Heart rate upper	-0.01200
6	Heart rate lower	0.00486
7	Weight upper	-0.01804
8	Weight lower	0.01803

Table 5.10: Regression coefficients for sensitivity calculations using algorithmic threshold variables

5.5 Trend analyses and visualization

In the case of home telemonitoring of heart failure patients, trend of measurements or trend of the calculated reference state could be potentially used to enable decisions regarding possible patient interventions and facilitate predictions of future alarm situations. To investigate which of the trends provide better indication of the alarm occurrence tetrachoric correlations were calculated between the binary variables indicating exceeded trend thresholds and alarm occurrence. Such correlations were calculated separately for every patient together with their statistical significance level. The highest correlation was determined based upon the average of statistically significant patient correlations. Furthermore, only the averages from 11 or more statistically significant patient correlations were considered.

A number of cases was examined varying the window size for trend calculations between 2 and 14 days, with daily increments, and trend thresholds between 0.1 and 2.5, with 0.05 increments. Furthermore, trend of the trends was also studied.

The results typically showed the highest correlations using the reference state trend in the calculations. Table 5.11 presents the combination of Kalman filtering based reference state trend calculation parameters resulting in the highest obtained average correlations, for each of the considered physiological measurements.

Physiological parameter	Correlation coefficient	Number of patients	Window size [days]	Trend threshold
Systolic b.p.	0.72	11	2	0.55
Diastolic b.p.	0.79	11	2	0.75
Heart rate	0.79	11	2	0.15
Weight	0.89	26	2	0.15

Table 5.11: Parameters for obtaining the highest correlations between the exceeded Kalman filtering trend of physiological measurements and alarm occurrence

As it could be observed, no perfect match existed between the exceeded trend thresholds and occurrence of alarms. Furthermore, it was noticed that exceeding the trend thresholds more often followed than preceded the alarm occurrence. Also, the window size resulting in the maximum correlation between the exceeded trend thresholds and alarm occurrence considered the minimum possible number of consecutive measurements, including only 2 days.

Consequently, additional indicator of a possible alarm occurrence was considered to always correspond to the alarms. In particular, using the determined threshold ranges and reference state based on Kalman filtering, relative closeness of the measured values to the threshold bounds could be considered as indication of the potential alarm occurrence. Such indication was color coded for each measurement point and added to the background of threshold range plots, as presented in Figure 5.21.

Color generation was obtained applying the R `barplot` function, which implemented separate functions for red, green and blue color components, given by Equations 5.6, 5.7, and 5.8, respectively.

$$f_{red} = \frac{1}{\pi} \cdot arctan(4 \cdot f_{color} + 2) + 0.5 \tag{5.6}$$

$$f_{green} = 2.5 \cdot p(x = 2.125 \cdot f_{color} | x \sim N(0,1)) \tag{5.7}$$

$$f_{blue} = \frac{1}{\pi} \cdot (-arctan(4 \cdot f_{color} - 2)) + 0.5 \tag{5.8}$$

Here, the color function, f_{color}, was obtained as a ratio between the difference in measurement values, f_t, and reference state values, g_t, and the difference between threshold, $V_{t,upper}$ or $V_{t,lower}$, and reference state values, calculated by Equation 5.9.

$$f_{color} = \begin{cases} \frac{f_t - g_t}{V_{t,upper} - g_t} & , \quad f_t \geqslant g_t \\ \\ \frac{f_t - g_t}{g_t - V_{t,lower}} & , \quad f_t < g_t \end{cases} \tag{5.9}$$

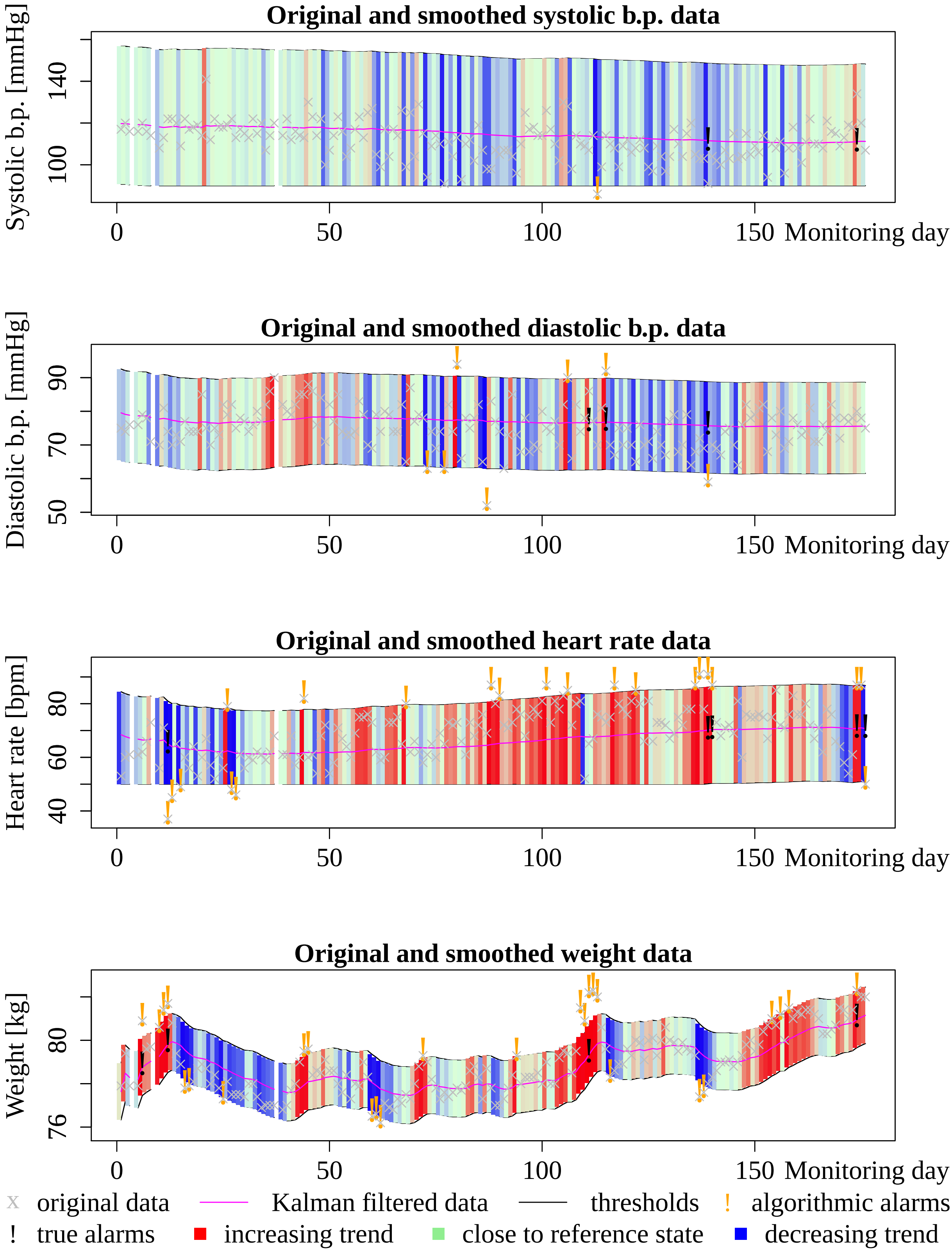


Figure 5.21: Color coded systolic, diastolic blood pressure, heart rate and weight measurements with Kalman filtering based automated alarm indication for a selected patient (Patient ID = 11,322)

As the thresholds were asymmetric around the reference state, the color function algorithm had to account for both upper and lower threshold bounds separately, taking into account their relevance only when the measurements were above or below the reference state, respectively.

The following code generated colors in the case of weight drawings:

```
barplot(
    cex.axis=1.5, #size of the axis lables
    (1-is.na(WEIGHT_NA))* #plot bars only when there are (weight)
                          #measurements, i.e. when no NA elements exist
                          #and is.na function value is 0
                         (pmin(ReWEIGHT+1.3,135)-pmax(ReWEIGHT-1.3,45)),
                          #bar sizes equal the difference between upper
                          #and lower thresholds
    border=NA, #no borders around color bars
    offset=pmax(ReWEIGHT-1.3,45), #offset color bar plotting from the
                                  #lower threshold (defined as a point-
                                  #wise maximum between the fixed 45 kg
                                  #limit and 1.3 kg lower value than the
                                  #reference state)
    col=rgb( #define red, green and blue component values of color bars
            1/pi*atan(4*drawWEIGHT+2)+0.5, #red function
            2.5*dnorm(2.125*drawWEIGHT,mean=0,sd=1), #green function
            1/pi*(-atan(4*drawWEIGHT-2))+0.5 #blue function
            ),
    space=0, #include no spacing between the color bars,
             #i.e. plot them adjacent to each other
    add=T #add color bar plot to the existing figure
    )
```

The following code implemented the color function (drawWEIGHT) for weight drawings:

```
drawWEIGHT<-(WEIGHT_NA-ReWEIGHT) #difference between (weight) measurements
                                 #and reference state
             / #divided by
             pmax( #point-wise maximum of
                  (WEIGHT_NA>=ReWEIGHT)* #when (weight) measurements are
                                         #greater than or equal to the
                                         #reference state then 1,
                                         #otherwise 0 times
                                        (pmin(ReWEIGHT+1.3,135)-ReWEIGHT),
                                        #difference between upper threshold
                                        #and reference state
                  (WEIGHT_NA<ReWEIGHT)* #when (weight) measurements are
                                         #smaller than the reference state
```

```
                                             #then 1, otherwise 0 times
                                            (ReWEIGHT-pmax(ReWEIGHT-1.3,45))
                                            #difference between reference
                                            #state and lower threshold
                    )
```

The maximum color intensities, red or blue, were obtained when the color function exceeded 1 or -1, respectively. Such extreme colors meant that the measurements exceeded the threshold bounds (red color for exceeding the upper threshold bound, blue color for lower). When the color function value equaled 0, green function component was dominant. Table 5.12 presents overview of the considered combinations of red, green and blue function values for obtaining the desired color map.

Color bar	Function values			
	Color	Red	Green	Blue
Red	1	0.947	0.104	0.148
Light green	0	0.852	0.997	0.852
Blue	-1	0.148	0.104	0.947

Table 5.12: Combinations of function values for obtaining the desired background color map in patient measurement plots

Note that the implemented functions for red, green and blue components were continuous, allowing also the color function to take values greater than 1 and lower than -1. However, due to the selected combination of `rgb` function parameters in such cases color change was negligible beyond the considered red and blue extremes, as illustrated in Figure 5.22. Figure 5.22 presents separate variation of red, green and blue components of the `rgb` function and the resulting colors in the case of hypothetical systolic blood pressure measurements varying in such a way to obtain a range of color function (f_{color}) values between -1.2 and 1.2. For simplicity, the selected reference state and thresholds were considered constant (specified as initial values and optimal threshold bounds in the described Kalman filtering algorithm).

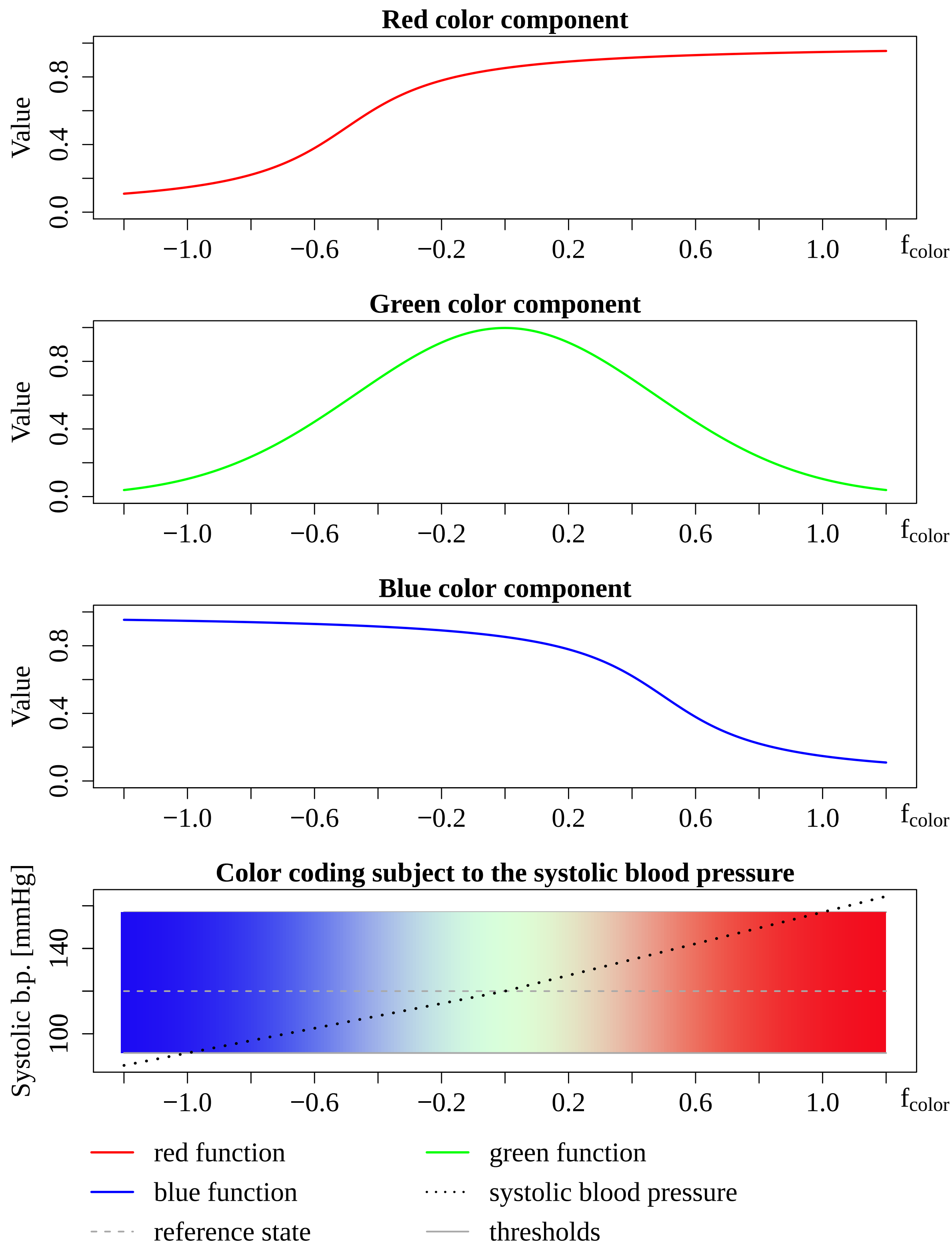


Figure 5.22: Color coding of a hypothetical systolic blood pressure varying within ±1.2 range of color function values, and variation of individual color components

5.6 Weather influence on alarm management

Patient telemonitoring data and associated weather conditions were analysed in two stages. In the first stage, Spearman correlation coefficients were calculated between the average daily patient physiological measurements and recorded weather conditions (temperature T, relative humidity RH, atmospheric pressure).

Spearman correlation coefficient was also used to correlate physiological measurements to effective temperature ($T_{effective}$) and thermal comfort, computed according to Equation 5.10 and Table 5.13 (Ono and Kawamura, 1991), respectively. Additionally, the analyses included calculation of polyserial correlations between the weather conditions and occurrence of patient alarms.

$$T_{effective}[°\mathrm{C}] = T[°\mathrm{C}] - 0.4 \cdot \left(1 - \frac{RH[\%]}{100}\right) \cdot (T[°\mathrm{C}] - 10) \tag{5.10}$$

Thermal sensation	Thermal comfort index	Effective temperature [°C]
Very hot	4	$34 < T_{effective}$
Hot	3	$31 < T_{effective} \leqslant 34$
Moderately hot	2	$28 < T_{effective} \leqslant 31$
Warm	1	$25 < T_{effective} \leqslant 28$
Comfortable	0	$22 < T_{effective} \leqslant 25$
Slightly cool	-1	$19 < T_{effective} \leqslant 22$
Cool	-2	$16 < T_{effective} \leqslant 19$
Cold	-3	$13 < T_{effective} \leqslant 16$
Very cold	-4	$T_{effective} \leqslant 13$

Table 5.13: Relation between thermal sensations, thermal comfort indices and effective temperatures (based on (Fanger, 1970))

Figure 5.23 presents recorded minimum, mean and maximum daily temperatures used in the analyses together with telemonitored physiological parameters and alarms for a selected patient. Also presented are linearly interpolated temperatures to correspond to the exact time of the patient physiological measurement records based on local weather station temperature recordings at 7:00, 14:00 and 19:00 hours on each measurement day.

Figure 5.24 presents calculated effective temperatures based on interpolated temperature values, related thermal comfort indices, relative humidity and atmospheric pressures, for a selected patient location over the course of telemonitoring physiological parameters.

Based on the obtained correlation results, the second stage of the analyses applied tests for determining statistically significant differences between:

1. the averaged daily physiological measurements in two consecutive days when: a) the difference in mean temperatures between those days exceeded preset limits, ranging between 3°C and 8°C; b) one of the two consecutive measurement days had a difference between maximum and minimum temperature higher than preset limits,

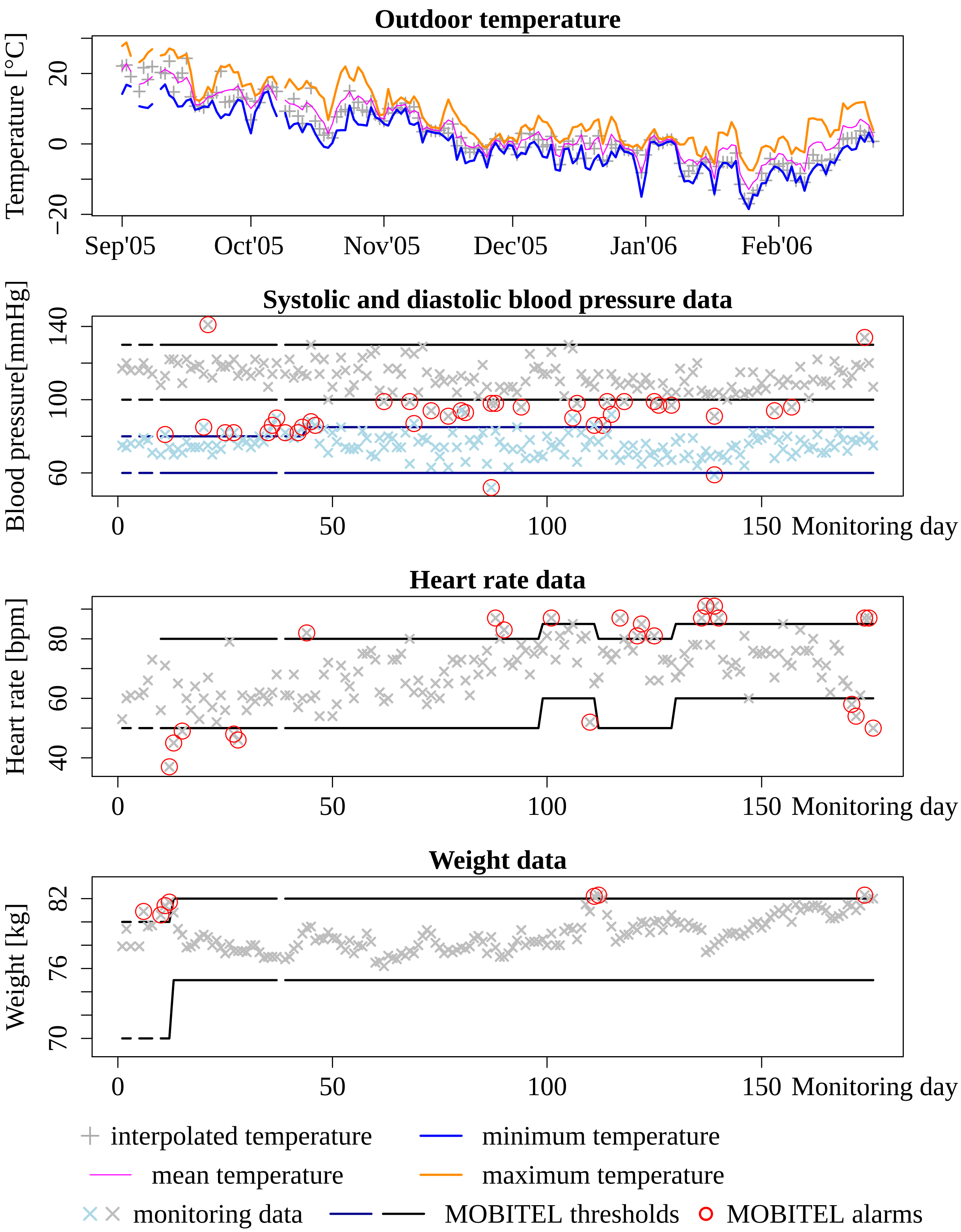


Figure 5.23: Outdoor temperatures and telemonitored physiological parameters (systolic, diastolic blood pressure, heart rate and weight) together with alarm indication for a selected patient (Patient ID = 11,322; Graz)

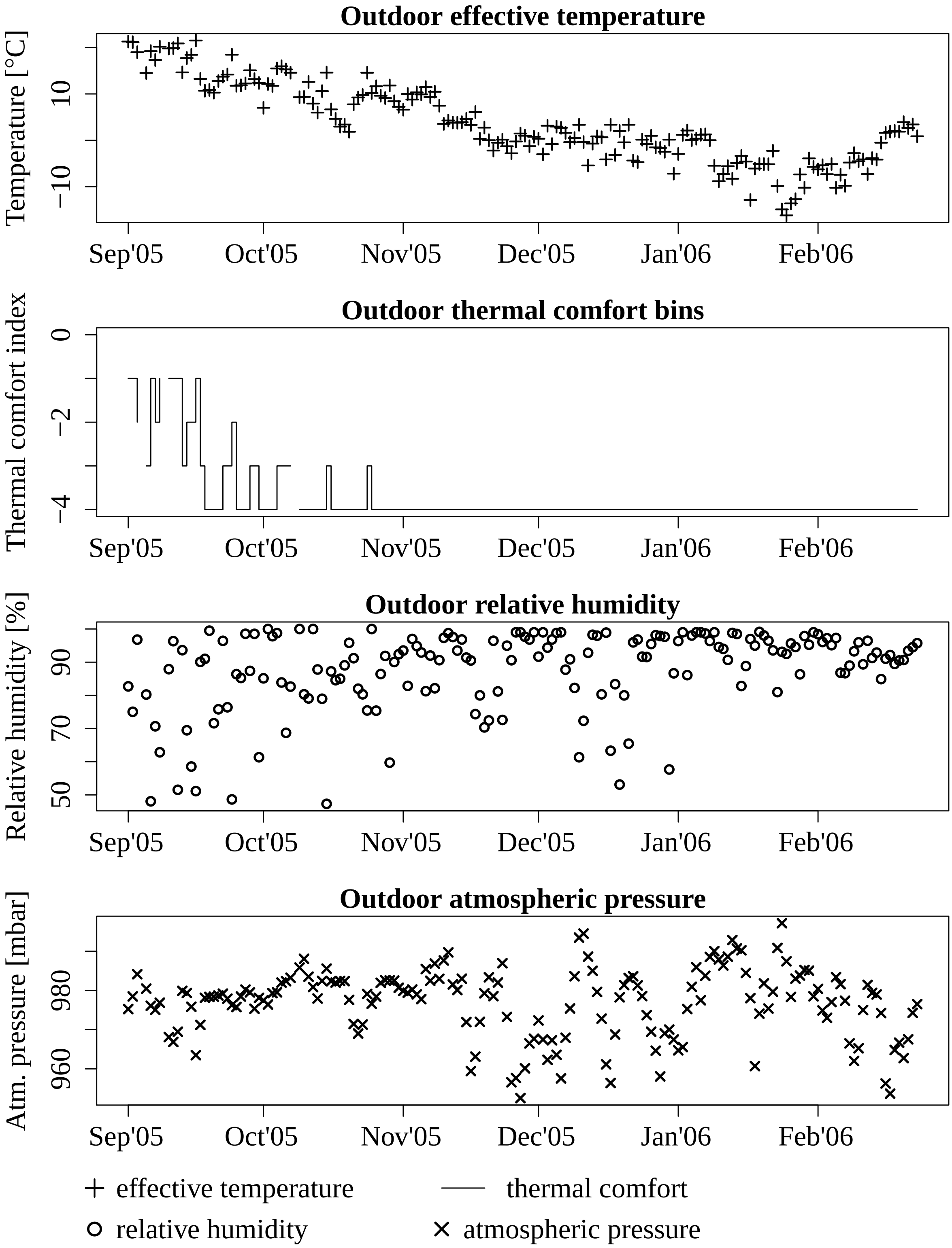


Figure 5.24: Outdoor effective temperatures, thermal comfort relative humidity and atmospheric pressure for a selected patient location (Patient ID = 11,322; Graz)

ranging between 7.2°C and 21°C; c) the difference between maximum and minimum temperatures over the two consecutive measurement days exceeded preset limits, ranging between 10.3°C and 22.3°C;

2. the averaged daily physiological measurements on the first and the third day in a sequence, when the difference in mean daily temperatures between the first two days exceeded preset limits, ranging between 3°C and 8°C;

3. the averaged daily physiological measurements on the days immediately preceding and following a day when the difference between maximum and minimum daily temperatures exceeded preset limits, ranging between 7.2°C and 21°C;

4. the average daily environmental temperatures on the days with and without true alarm occurrences.

All temperature limits were varied at 0.1°C increments, while the considered temperature ranges were based upon the frequency of occurrence of such environmental conditions limiting the extracted sample sizes to a minimum of 10 and maximum of 5,000.

Two-sided u-test and t-test for matched pairs in combination with Shapiro-Wilk normality test were employed to determine the significance of differences between the considered data sets. Apart from p-values the calculations also included 95% confidence intervals (CI).

Correlations were calculated between the telemonitored physiological parameters and weather conditions, effective temperature and comfort. As an illustration, Figure 5.25 summarizes statistically significant correlations between diastolic blood pressure and weather conditions including effective temperature. The highest number of patients in Figure 5.25 had statistically significant correlations with effective temperature ranging between -0.5 and -0.3 (10 patients). Within this correlation range 9 patients had statistically significant diastolic blood pressure correlations with outdoor temperature. Similar influences were observed on the other physiological parameters and alarms.

Following such findings, investigation focused on differences in patient physiological parameters between the days with extreme temperature variation, heat or cold stress. Thermal stress due to the rising temperatures, heat stress, had no statistically significant influences on patient physiological measurements. Thermal stress due to the falling temperatures, cold stress, was statistically significant for systolic and diastolic blood pressure when the mean temperature difference thresholds were between 6.4°C and 6.8°C ($p<0.05$, 95%CI: (-16, -1) and (-8, 0) mmHg, for systolic and diastolic blood pressure differences, respectively, sample sizes: 17 and 16), and additionally statistically significant only for systolic blood pressure when the mean temperature difference thresholds were 6.1°C and 6.2°C ($p<0.05$, 95% CI: (-10, 0) mmHg, sample sizes: 29 and 28). The obtained negative CIs indicated rising blood pressures with the decreasing mean outdoor temperatures. Figure 5.26 presents the obtained results for 6.8°C mean temperature difference threshold.

Considering the day immediately following two consecutive days with large mean temperature variations, statistically significant differences were identified for both, heat and cold stress conditions. Falling temperatures, with the mean daily difference of 5.7°C and

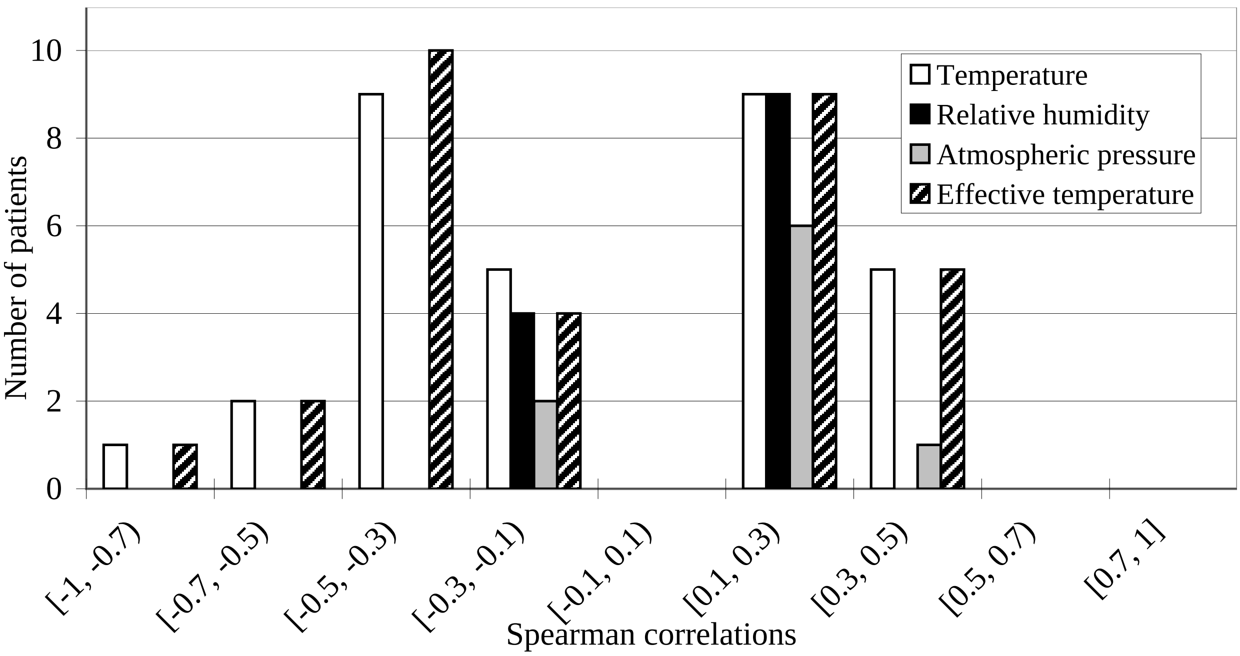


Figure 5.25: Statistically significant correlations between diastolic blood pressure and weather conditions for individual patients (values around zero were not statistically significant)

5.8°C, distinguished significantly different heart rates (p<0.04, 95% CI: (0, 4) bpm, sample sizes 68 and 64). The obtained positive CIs indicated falling heart rates following the decrease in mean outdoor temperatures. Rising temperatures, with the mean daily difference of 6.7°C and 6.8°C, distinguished significantly different diastolic blood pressures (p<0.04, 95% CI: (-4, 0) mmHg, sample size: 15). The obtained negative CIs indicated rising diastolic blood pressures following the increase in mean outdoor temperatures.

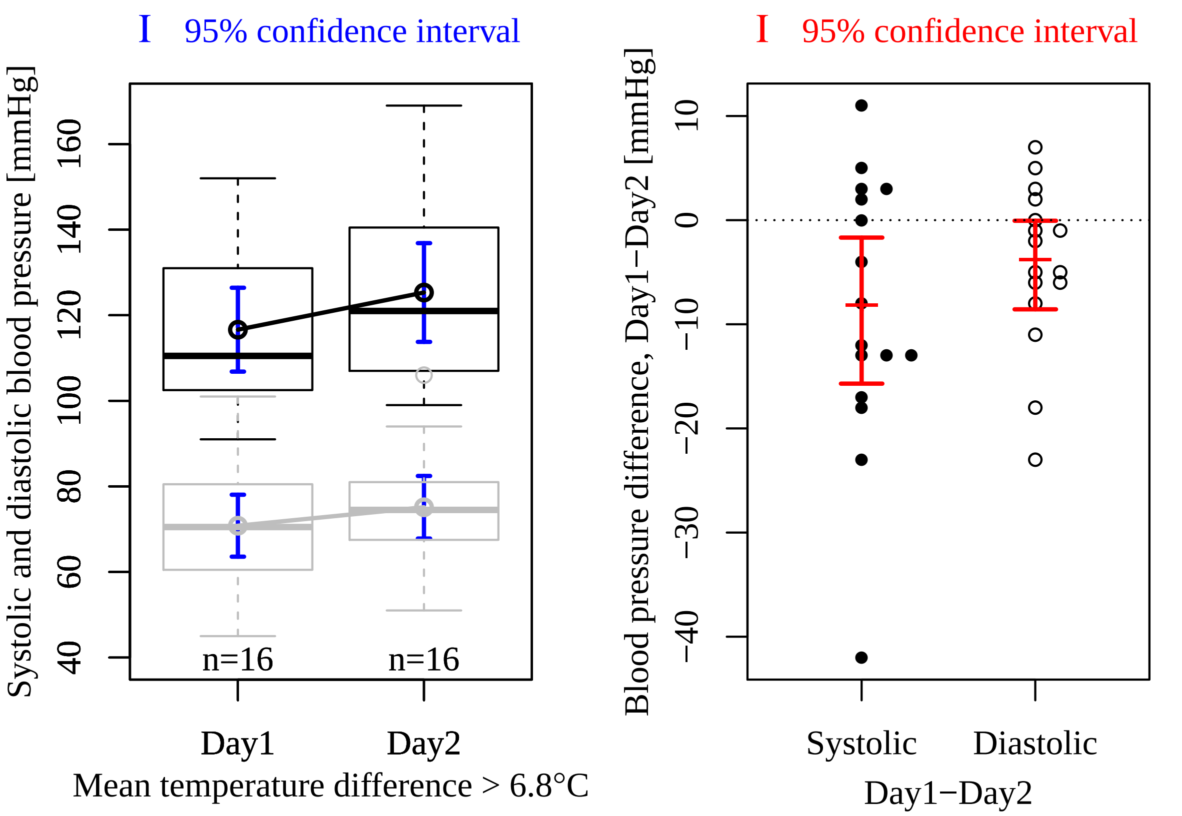


Figure 5.26: Box plots, means and statistically significant differences with confidence intervals for systolic and diastolic blood pressures over two consecutive days with mean temperature difference higher than 6.8°C

Chapter 6

Discussion

The current study used data from MOBITEL telemonitoring database of heart failure patients. Although median improvement from NYHA class III to II over the course of telemonitoring was reported in the literature, there was no information about the type and progress of cardiovascular diseases of each patient in the study. Furthermore, no records about patients' age or gender were available for the data analyses, although it was known that the investigated sample consisted of elderly people. As such individual patient information was not available, it was not possible to determine the effects and dependance of age and gender on the considered physiological parameters.

MOBITEL was the first large scale telemonitoring study conducted by the research team at the Department of Safety and Security, Austrian Institute of Technology. Consequently, not all of the equipment and information transfer was reliable. The telemonitoring database lacked some of the preprocessing capabilities that would ensure higher data quality. For example, indication of the measurement gaps or unrealistic measurement values to the operators could result in timely reaction and mitigation of the possible human or technical causes for insufficient quality of data collection. Eventually, substantial number of patient records was removed from the analyses. Up to 17 out of 65 patients (26%) originally participating in the study had to be removed in different phases of the analyses, primarily due to insufficient data. The patients remaining in the study were those submitting a majority of all the collected telemonitoring records (over 85%).

The following sections comment and describe particular issues and limitations related to the previous dissertation chapters.

6.1 Data statistics

The observed measurement gaps illustrated in Figure 4.2 were removed, rather than conducting any interpolation between the measurements to assume the values of missing data. Such procedure was advised by a data mining expert involved in home telemonitoring (Drobics, 2010). With such preprocessing, the measurements immediately preceding and following the gaps were considered as adjacent. This assumption introduced an error, which would increase with the size of the measurement gap. Such error might have adverse impact on the validity of the algorithm. Consequently, patients with relatively large measurement gaps were eliminated from the database and excluded from the presented analyses. In practical applications, it would be necessary to ensure regular telemonitoring via patient follow-up contacts.

In the case of multiple daily measurements, the presented analyses used daily averages to count with single records each day. Such simplification also introduced an error, as the patient physiological parameters were changing during the day, e.g. weight would typically increase after a meal, heart rate would increase after physical activity etc. However, the introduced averaging was necessary as the considered statistical algorithms required regular measurement records. Ideally, in the future practical applications the patients would be asked to make regular measurements each day, and will be automatically reminded to do so, if they failed to submit the measurements within a certain time frame. The patient reminder levels might also increase with the increasing gap in the measurement records, e.g. initially, the patients could receive an sms, followed by a phone call with an automated voice message, followed by a phone call from the physician, followed by a visit of the medical and/or technical staff. Further improvements of the developed algorithms might be possible to consider multiple regular daily measurements. Such improvements would be particularly important for physiological parameters with larger daily variability such as heart rate or blood pressure, and less important for the parameters with smaller variability, such as weight.

Comparing the Figures 4.4 and 4.5, one can observe high influence of alarm thresholds on the occurrence of alarms. For example, despite of relatively stable weight for Patient IDs 8,832 and 8,587, weight alarms still occurred, whereas no alarms were recorded for the Patient ID 10,601 despite of higher variability in weight measurements. Such situation is the consequence of a wider alarm threshold range setup by the physicians in the case of the latter patient as compared to the former. The alarms presented in Figure 4.5 were those used in the considered analyses. Additionally, MOBITEL study also included weight gradient alarms occurring when the patient weight was changed (increased or decreased) by more than 2 kg over 2 consecutive days (Scherr et al., 2009). Such alarms were relatively rare in the considered telemonitoring records (only 6 unique true alarms representing less than 3% of the total true alarms and less than 0.3% of all the alarms). Consequently, the current analyses did not separately consider such alarms neither did the developed algorithms include the weight gradient rule for automated alarm generation. However, the mentioned weight gradient of 2 kg in 2 days would likely lead to the weight alarm generation according to the proposed Kalman filtering based algorithmic thresholds of ± 1.3 kg presented in Table 5.5. Such occurrences were not additionally investigated.

Possible physician actions in the study included a category “Other action”. However, the information which other actions the physicians undertook in such cases was not available in the database. Nevertheless, this type of physician actions was considered to be related to true alarms, together with patient contacts and medication adjustments.

Statistics presented in Table 4.2 illustrated the frequency of collected measurements, showing that more than half of all the telemonitoring records were collected within the first 14 months of the study (Start: October 9, 2003; Median: December 7, 2004). The remaining 4 years till the end of the study (February 19, 2009) resulted in less frequent measurements. While many reasons could explain such situation, a possibility could be lower enthusiasm and motivation of the patients to regularly monitor their conditions after a certain initial period. Such possibility should be confirmed and additionally investigated together with the measures to stimulate continued patient interest in participation over the long term telemonitoring.

Table 4.2 also showed that the data preprocessing criteria were properly set. The vast majority of the individual patient thresholds fell within the defined ranges of physiological parameters used in data cleaning. Thus the considered physician thresholds were realistic in reference to the recorded measurements.

Figure 4.7 showed that exceeded multiple thresholds (double and 3 or more) had higher influence on physician actions caused by true alarms, than on false alarm responses. Two and more simultaneously exceeded thresholds caused around 40% of "Patient contacts" and "Medication adjustments", as opposed to less than 25% of "Threshold adjustments", "No actions", and "Alarms with no responses". Exceeded single thresholds caused the majority of false alarms (over 75%). The only physician action in response to true alarm occurrences that mostly resulted from exceeded single thresholds was "Other action". However, according to Figure 4.6, the portion of "Other actions" in the total number of considered true alarm cases was relatively small (only 36 (14%) out of the total $135+89+36=260$ cases).

Out of the single threshold alarms, exceeded blood pressure thresholds (both systolic and diastolic together) had the highest influence on resulting physician actions, as presented in Figure 4.8. In both, true ("Patient contact", "Medication adjustment", "Other action") and false positive alarm ("No action") cases, over 50% of physician actions were in response to exceeded systolic or diastolic blood pressure thresholds. The only physician action mostly impacted by the physiological parameters other than blood pressure was "Threshold adjustment", mostly triggered by exceeded weight thresholds. Although not presented in Figure 4.8, the analyses of "Alarm and no response" cases (Figure 4.7) caused by exceeded single thresholds showed the highest influence of alarms due to exceeded upper weight thresholds, responsible for around 60% of all these false alarm occurrences.

During the course of home telemonitoring the patients were given a range of different medications. Heart failure treatment usually assumes that patients take more than one medication, as was the case in 51 (94%) out of 54 patients presented in Figure 4.9. Medication type intake and dosage was mostly kept constant throughout the study, as indicated in Figure 4.10. However, this information may not be accurate in the cases when patients were taking 5 or more different medications at the same time. The telemonitoring software was designed to allow recording of a maximum of 4 different medication types by the physicians, and no records about the possibly larger number of prescribed medications were available. Also, no information was available about the way the physicians might have selected the medications to record in the database in the cases of therapies including more than 4 different medications.

The number of different medication combinations in which each of the prescribed medication types in the study occurs is presented in the diagonal of Table 4.3. This number does not necessarily represent the number of patients taking a particular medication, as each patient might have taken more than one combination of medications in which a particular medication occurs. Also notice that Table 4.3 sums up combinations of certain medication brand names consisting of the same active ingredients, which are all listed separately in Figure 4.11.

6.2 Multi-threshold alarms

The idea behind the alarm flags was to investigate the possibility of replacing the existing alarms using different alarm generation criteria and/or sending additional information to physicians as an indication how important the alarm might be. The results in the case of replacing the existing alarms (flag level 1) using different alarm generation criteria (flag levels 2 through 5) were presented in Table 5.1. The results showed higher accuracy of non-alarm decisions with the increasing alarm flag level, which was reflected in the increasing specificity. Simultaneously, accuracy of alarm decisions was reduced, as indicated by the falling sensitivity. Thus, ever more true alarms remained undetected as the alarm flag levels increased. The specificity was directly proportional to the sum of true positive and true negative events and the alarm flag levels, but inversely proportional to sensitivity. Still, the introduced alarm flags increased the overall accuracy of combined alarm and non-alarm decisions. Consequently, high levels of alarm flags would require immediate attention as it would be more likely that medication adjustment, patient contact or other physician action was necessary.

The increase in alarm flag levels would imply increase in specificity, but also decrease in sensitivity. Specificity vs. sensitivity cost benefit analysis could be based upon the obtained difference and ratio between the reductions (Δ) in false and true positive events. The reductions ΔFP (ΔTP) for each alarm flag level 2 through 5 were calculated by subtracting the number of FP (TP) events from the number or corresponding alarms with a flag level 1, presented in Table 5.1. Thus, $\Delta FP[\%] - \Delta TP[\%]$ reduction difference varied between 10% and 14% and $\Delta FP/\Delta TP$ reduction ratio between 10.7 and 8.5. In relative terms, $\Delta FP[\%]/\Delta TP[\%]$ percentage reduction ratio varied between 1.54% and 1.22%, respectively, for alarm flag levels 2 through 5 compared to the alarm flag level 1, i.e. the original algorithm.

Such reduction difference ($\Delta FP[\%] - \Delta TP[\%]$) could be understood as an absolute percentage benefit reflected in the reduced number of FP events paid by the cost of TP event reductions. The reduction ratio ($\Delta FP/\Delta TP$) described a relative benefit in the form of FP reductions paid by each TP reduction. The percentage reduction ratio ($\Delta FP[\%]/\Delta TP[\%]$) expressed the relative percentage benefit in FP reductions for each percent of TP reductions. The largest absolute benefit, i.e. the greatest difference of 14% between the FP and TP event reduction percentages was achieved for the alarm flag level 5 based on three or more exceeded thresholds over two consecutive days. However, the same maximum alarm flag level 5 corresponded to the smallest relative benefit, i.e. 1.22% decrease in FP events was the minimum achieved for each percent decrease in TP events (or 8.5 false alarms less for each true alarm reduction). The highest relative benefit was in the case of 3 days with 2 or more exceeded thresholds (alarm flag level 2): 1.54% less false alarms for each percent of true alarm reductions (or 10.7 FP events less for each TP event decrease).

The fact that each reduction of false positive events followed the reduction of true positive events meant that reducing physician efforts of analyzing and responding to alarms with flag levels 2 through 5 replacing the alarms with flag level 1, could lead to disregards of some potentially significant patient alarm conditions. Although the introduced alarm flag levels reduced the number of false alarms by up to 78% (from 1804 to 395) with the increasing alarm flag levels from 1 to 5, the percentage of false alarms in the total number of alarms ($TP + FP$) changed only slightly: from 87% (1804 out of $260 + 1804 = 2064$ alarm flags level 1) to 81% (395 out of $94 + 395 = 489$ alarm flags level 5). Thus, only every eighth alarm (13%) with a flag level 1 and every fifth alarm with a flag level 5 (19%) would be true. Such outcome could still lead the physicians to develop distrust in the overall alarm management system.

Consequently, rather than replacing all the generated flag level 1 alarms, they were supplemented by the introduction of the described alarm flag levels 2 through 5. As such, the original alarms (1 day / $\geqslant 1$ exceeded threshold, flag level 1) were maintained, keeping the number of false negative decisions constant (17). Applying the algorithm presented in Figure 5.1 the existing alarms were associated with one of the flag levels 1 though 5. Thus, the distribution of alarm flag levels presented in Figure 5.2 was obtained. The highest alarm flag level 5 was proportionally more present in the true alarms (94 out of 260 cases, or 36%) than in the false alarms (395 out of 1804 cases, or 22%). Consequently, all the other, lower level alarm flags were proportionally more present in the false alarms. Also, the percentage of true alarms increased from 8.5% (47 out of 550 alarm flag level 1 cases) to 19.2% (94 out of 489 alarm flag level 5 cases) as compared to the overall average of 12.6% true alarms (260 out of 2064 total alarms) when such a procedure is not applied in the reference MOBITEL study. According to the algorithm presented in Figure 5.1, the sum of alarms with different flag levels in Figure 5.2 corresponded to the data presented in Table 5.1, as follows:

- The sum of all alarms with all the flag levels in Figure 5.2 were equal to the records for 1 Day / $\geqslant 1$ exceeded threshold in Table 5.1,
- The sum of alarms with flag levels 2 through 5 in Figure 5.2 were equal to the records for 3 Days / $\geqslant 2$ exceeded thresholds in Table 5.1,
- The sum of alarms with flag levels 4 and 5 in Figure 5.2 were equal to the records for 3 Days / $\geqslant 3$ exceeded thresholds in Table 5.1, and
- The data for alarm flag level 5 in Figure 5.2 and Table 5.1 were equivalent.

Such results in Table 5.1 originated from separate investigations of the alarm occurrence for each of the alarm flag levels independently.

6.3 Data smoothing

Further investigation of the alternative procedures that could be used to replace the existing alarm generation algorithm based on manual threshold adjustments, focused on automated specification of optimum alarm thresholds around a certain patient reference state. Such approaches required investigation of various methods to determine the reference state, as already conducted in the preliminary study. Variance of the calculated reference state was compared to the variance of the original measurements as an indication of the influence individual measurements would have on the reference state stability. Fourteen point moving average considering measurements prior to each reference state estimate, FFT filter and Kalman filter were found as the most promising approaches to calculate the reference state. The other investigated (shorter) moving averages and (5 point) Savitzky-Golay filter had much higher variances indicating larger influence of variations in the monitoring data on the reference state points. This was particularly the case for Savitzky-Golay filter which exhibited higher variance than moving average approximation on the same number of measurement points. Thus, to achieve the Savitzky-Golay variance levels comparable to e.g. 14-point moving average, much more than 14 measurements would be required.

The selected reference state estimation models based on each of the three most promising approaches (14-point moving average, FFT and Kalman filters) were statistically appropriate as the distributions of measurement residuals around the reference states were almost normal and symmetric around zero, time independent, and did not exhibit any patterns (uncorrelated), as indicated in Figures 5.5, 5.6 and 5.14. The drawback of the 14-point moving average was a need for relatively large number of monitoring points and correspondingly delayed reaction of the reference state estimation to the monitoring trends. The disadvantage of FFT filtering was assumption of periodicity, which was not observed in the monitoring data.

Thus the procedure of choice for determining patients' reference state was based on Kalman filtering. The analyses were conducted separately for each of the considered physiological measurements: systolic and diastolic blood pressure, heart rate and weight, identifying the most suitable combination of parameters, presented in Table 5.2. These parameters were considered fixed and constant for all the patients during the course of telemonitoring. Fine tuning of the determined Kalman filter parameters could be additionally investigated in the future. The most significant parameters for Kalman filtering were two oscillation parameters (Vmat, Wmat) impacting the variability of the reference state estimation. Further enhancements of Kalman filtering would be possible assuming dynamically changing values of these filter fluctuation parameters during the algorithm runtime over the telemonitoring period. Such fine tuning or dynamic variation in Kalman filter parameters might bring the reference state estimation to closer correspondence with the variability in measurements. Performed setup of all the other Kalman filter parameters only impacted the filter output values for the several initial inputs, whereas after the initial stage the obtained reference state results did not depend on the initial Kalman filter setup. The current approach used mean values typical for healthy individuals as initial assessments for the monitored physiological variables and provided only order of magnitude estimates for the other parameters. Personalized individual assessments could be used in the future for the initial physiological parameter assessments (e.g. weight), subject to each patient's gender, age, height, body frame size etc. For example, it was observed that a substantial number of false alarms occurred relatively early in telemonitoring. In MOBI-TEL, this was possibly subject to corrective actions of the physicians which should have properly adjusted the alarm generation thresholds for individual patients. In the proposed Kalman filtering based algorithm, the initial filter adjustments to the measurements were governed by the considered filter parameters.

6.4 Dynamic thresholds and ROC curve

The proposed reference state estimation using Kalman filtering parameters showed good performance reflected in the comparison of obtained alarms to the original MOBITEL study. The results were obtained using the new alarm generation thresholds dependent upon individual patient reference state provided by Kalman filtering of the measurements.

Using the determined reference state, various dynamic threshold adjustments were investigated to provide increased accuracy of the patient telemonitoring alarm generation. The investigated approaches included absolute and relative increases and decreases of the patient reference state to form threshold bounds, as well as fixed upper and lower thresholds. The most promising approach was identified as a combination of different upper and lower threshold bounds around patients' reference state and fixed upper and lower threshold limits. While the upper and lower threshold bounds were formed by adding or subtracting predetermined values to the patients' reference state, the threshold limits always triggered alarms when exceeded by the measurements, regardless of the patient reference state. Finally, the alarms could be triggered by exceeding either the specified threshold bounds or threshold limits, or both.

The combination of optimal threshold parameters was presented in Table 5.5. When searching for the optimal combination, absolute and relative threshold adjustments were investigated separately. Thus, Table 5.5 presented only absolute threshold adjustments and associated threshold limits which provided the highest accuracy off all the other investigated approaches and proved better than the original MOBITEL algorithm, resulting in higher accuracy presented in Table 5.6. However, potential benefits of combining absolute and relative thresholds adjustments were not thoroughly investigated. The reason was generally equal or worse performance identified using relative threshold bounds, although in certain ranges of absolute adjustments, the corresponding relative thresholds showed minor benefit. Such case was e.g. illustrated in Figure 5.18 (right) for lower systolic blood pressure adjustments ranging between -7 and -15 with fixed limits of 90 mmHg, where the dotted line representing performance in the case of relative percentage adjustments was the closest to the ideal (0, 1) point of the graph. Thus, potential fine tuning of the threshold bounds to include combination of relative (percentage) and absolute threshold adjustments may bring additional accuracy benefits for certain combinations of threshold adjustment parameters in comparison to the existing model. Such enhancements would require the physicians to input either absolute or relative threshold bounds, depending on the values of other inputted adjustments. Despite of the potential to improve the sensitivity-specificity balance, operation of the threshold adjustment algorithm combining absolute and relative threshold bounds might be less feasible and understandable for the physicians in practice.

Apart from the effects of relative and absolute threshold bound adjustments, Figure 5.18 also illustrated the effects of changing threshold limits. In the right hand side graph of

Figure 5.18 the lower systolic blood pressure threshold limit was changed from 85 mmHg, having almost no impact on the baseline ROC curve, to 90 mmHg, bringing the ROC curve closer to the ideal point (0, 1). However, in the case of the upper systolic blood pressure ROC curve presented in the left graph of Figure 5.18, increasing the threshold limits from 160 mmHg to 170 mmHg had quite the opposite effect, distancing the ROC curve further from the ideal point (0, 1). Similar behavior was observed in the case of all analysed thresholds of patient physiological parameters. Thus, threshold limits had optimal values presented in Table 5.5, resulting in the most favorable sensitivity-specificity combinations.

The presented ROC curve results in Figure 5.19 showed variety of sensitivities and specificities that could be achieved changing the threshold parameters of the proposed automated alarm generation algorithm. The identified ROC curve Equation 5.2 parameters given in Table 5.4 excluded the two outlier points with minimum and maximum sensitivities/specificities. This was done to ensure that the residuals showed no statistical evidence of correlation and no statistically significant deviation from normal distribution. If the two outliers would not be omitted the residuals would fail to meet the aforementioned statistical prerequisites for regression. However, even without excluding the two outlier points, the obtained ROC curve equation would still approximate well all the considered points ($R^2 = 0.94$) and would fall within the 95% confidence interval presented in Figure 5.19.

Ideally, the developed alarm generation algorithm should have high sensitivity and specificity, approaching the point (0, 1) on the ROC curve. The point on the curve closest to (0, 1) maximized the sum of specificity and sensitivity values. Such a point was identified and its threshold parameters presented in Table 5.5, combining the adjustable bounds around the patient reference state and fixed extreme upper and lower thresholds as ultimate cut-off values.

Applying such identified optimal combination of thresholds, the results presented in Table 5.6 were obtained. The results showed increased specificity and overall accuracy with decreased sensitivity, compared to the original manually adjusted thresholds by the physicians (MOBITEL algorithm). The number of alarms and corresponding physician actions were different compared to the earlier analyses of alarm flag levels (Table 5.1 alarm flag level 1) as two patients (Patient No. 10 and 28) were removed from the dataset in the meantime. Figure 5.20 illustrated the application of both, automated and manual, threshold adjustment algorithms on a selected patient. Correct identification of all true alarms was followed by a noticeable reduction of false alarms in the case of automated as compared to manually adjusted (MOBITEL) thresholds applied to the selected patient telemonitoring.

Interestingly, the identified constant upper blood pressure threshold limits in combination with reference state dependent thresholds in Table 5.5 coincided with the Stage 2 Hypertension boundaries typical in medical practice, described in Table 4.5. At the same time constant lower blood pressure threshold limits coincided with hypotension boundary for systolic blood pressure (90 mmHg), but were somewhat lower (45 mmHg) than the hypotension boundary in the case of diastolic blood pressure (60 mmHg).

Significance of the presented ROC curve could be viewed in the capabilities to select the appropriate threshold cut-off values to match the algorithm performance to the expected patient conditions at acceptable levels of accuracy. For example, in the case of patients which the physicians might consider unlikely to experience adverse events, selection of thresholds associated with the ROC curve points towards the origin (0, 0) would be appropriate, as otherwise almost all alarm situations would be false. On the contrary, in cases when the physicians would consider patients' status as critical, appropriate selection of thresholds would be associated to the points towards the upper right part of the ROC curve (1, 1); otherwise a large number of cases with no alarm occurrence might be false negative. The ROC curve and proposed automated alarm generation algorithm could thus be viewed as a complementary tool which could help the physicians guiding them to properly adjust the thresholds in the setup of telemonitoring systems.

Quality of the designed model could be estimated calculating the area under the ROC curve (AUC), representing the ability to correctly classify cases with and without the alarms. As a rule of thumb, areas above 0.8 would indicate good accuracy of diagnostic models. As the calculated AUC was higher than 0.8, the presented approach well discriminated between the possible stable and alarm patient conditions. However, the calculated AUC should be considered as an estimate rather than an exact value as not all the points were available on the ROC curve to cover the entire span necessary for exact calculations. Including the ROC curve data points from all the possible combinations of considered threshold values for all the monitored parameters would require prohibitively long computational time. Such computations would include approximately $6.5 \cdot 10^{12}$ possible combinations assuming unit increments in blood pressure and heart rate thresholds, and 0.1 kg increments in weight thresholds. Rather, the fitted ROC curve was used to provide estimate of the model quality.

Based on the used combinations of thresholds to produce the ROC curve, a model to calculate sensitivity was developed applying principle component regression on the 8 threshold variables input domain space. The resulting model coefficients were presented in Table 5.10, whereas the means and standard deviations needed for the standardization of the input values were given in Table 5.8. Negative coefficients next to the upper threshold variables and positive coefficients next to the lower thresholds indicated that increasing values of the thresholds above the means (in absolute terms) would result in reduction of the sensitivity below $r_0 = 0.770$. And vice versa, reducing values of the thresholds below the means (in absolute terms) would result in increase of the sensitivity above $r_0 = 0.770$. Table 5.8 also presented the ranges of input thresholds for which the model was valid. Using these input threshold ranges, sensitivity range [0.276, 0.957] could be obtained. Applying Equation 5.2, this sensitivity range would translate into a specificity range [0.979, 0.476], respectively. One extreme case $(1 - Specificity, Sensitivity) = (0.021, 0.276)$, resulting from the widest threshold ranges, would be appropriate for the patients which the physicians might consider unlikely to experience adverse events. The other extreme case $(1 - Specificity, Sensitivity) = (0.524, 0.957)$, resulting from the narrowest threshold ranges, would be appropriate when the physicians would consider patients' status as critical.

As the analyses were entirely based on the recorded physician responses, uncertainty of diagnoses remained a possible topic for further investigation. Namely, it was questionable whether the classification of true and false alarms would remain the same if the physicians would have a chance to examine all the automatically generated alarms. Particularly, as nearly half (890 of 1845) of all the alarms in the automated system were newly generated, i.e. had no recorded response by the physicians. In the current analysis all such alarms were classified as false and the overall results showed only minor reduction (94, i.e. 5%) of the false alarms in the presented algorithm compared to the original MOBITEL study. However, if the cases with no recorded physician responses would be excluded, the false alarm reduction could reach nearly 45%. At the same time, compared to the MOBITEL study, the presented results of automated alarm management included (47 alarms) 18% less true alarms. Reclassification of the cases with no recorded physician responses might increase the number of true alarms indicated by the developed algorithm, when tested in pilot trials. Potential reclassification of such cases would have a profound impact on the proposed algorithm as well as the MOBITEL study, affecting its false negative occurrences. To properly evaluate usefulness of the proposed algorithm as a potential method of detecting adverse patient health status conditions pilot testing would be required.

6.5 Trend

Trend of the measurements was calculated as a slope of a linear model of measured data over the fixed window size. The linear model was obtained minimizing the sum of squared errors between the line representing the model and the measurements. Similarly, trend of the reference state was obtained by considering the reference state values instead of the measurements. Such trend depended upon the procedure used to calculate the reference state. All the considered procedures described in Section 5.2 were investigated for calculation of the reference state trends. Finally, trend of the trend values was also calculated constructing another linear model on the results of the existing trend calculations. Such double derivation approach showed higher stability as compared to the first order trends.

The calculated trends varied between certain values. An event generated when trend exceeded a preset threshold value was considered as an indicator of possible alarms. The same situation was also applicable to the cases when the trend fell below a certain threshold value. In the current trend analyses absolute upper and lower trend threshold values were considered to be the same. Whether the trend exceeded the thresholds or not, depended on the size of the thresholds but also on the trend itself. The trend value was impacted by the window size used to sample the subset of measurements (reference state, or first order trend values) for obtaining the slope of the linear model. Therefore, an examination of various window sizes to calculate the trend thresholds was conducted to select the most appropriate one. The most appropriate here meant that there was the largest correlation between the occurrence of alarms and the events when trend exceeded the threshold values.

When considering alarms for calculation of the mentioned correlations, two possibilities existed. First, the alarms could be based on measurements (MOBITEL), and second, the alarms could be calculated by the selected automated alarm generation algorithm. The investigations focused on correlations between the exceeded trend threshold values and occurrence of MOBITEL measurement alarms.

Investigation of possibly different trend thresholds subject to the measurement values was also conducted. For example: in patients with prevalently high blood pressures high trend thresholds might be more appropriate as compared to the patients with generally low blood pressures, for which low trend threshold values could be considered. The investigation was conducted varying window sizes between 5 and 21 days with daily increments. The high trend threshold values varied between 0.8 and 1.7 and the low trend thresholds varied between 0.1 and the high trend threshold value, with 0.1 increments. The average value of blood pressure measurements over the first measurement window size was used to assign a patient to a group with low or high trend threshold values. Such average blood pressure, serving as a decision border between the low and high trend thresholds, was also varied between 130 and 160 mmHg in increments of 10 mmHg.

As a final result, no significant advantage was found to use any other described trend calculations (measurement trend, trend of the trends, dual trend thresholds) as compared to the Kalman filtering reference state trend single threshold value for all the patients. On the contrary, single thresholds of the Kalman filtering reference state trends gave at least equal or higher correlations than any of the other trends including the dual trend thresholds, subject to patient classification based on the initial physiological conditions.

Consequently, the variation of window size and single Kalman filtering trend threshold values was examined for all the recorded measurements: systolic and diastolic blood pressures, heart rate and weight. Table 5.11 showed the best combinations of parameters which resulted in the maximum correlations between the exceeded Kalman filtering trend thresholds and the occurrence of alarms. The presented results were averages of at least 11 patients with statistically significant correlations. This number of patients was selected due to the rule of thumb that correlations above 0.8 would be required with 10 pairs of observations in order to be significant. Indeed, the calculated average correlation coefficients for 11 patients were around 0.8 for all the physiological parameters, indicating significant results.

As the identified maximum averages of statistically significant correlations included window sizes of only two consecutive Kalman filtering reference state data points in the cases of all physiological parameters, an alternative approach was used to indicate the proximity of each measurement to the critical values causing the alarm occurrence. With such an approach, background fields behind each measurement were colored based on the relative distance between the measurements and the Kalman filtering reference state, with respect to the distance between the alarm threshold values and the reference state. As the measurements would approach a certain alarm generation threshold, the background color bar would approach color extremes: red for upper, and blue for lower thresholds. Color bars behind the measurements close to the reference state were chosen to be pale green. Continuous smooth functions governed the color variation in order to speed up and simplify the computational algorithm. The selected parameters of the color generation functions ensured negligible color variation close to the color extremes, even in the cases when measurements exceeded the alarm generation thresholds.

Such background colors effectively indicated the trend of the Kalman filtering reference state. Namely, if a new measurement would differ from the previous Kalman filtering reference state value, the new Kalman filter result would follow the value of such a measurement, as would the intensity of the background color bar. For example, a new measurement higher than the previous Kalman filtering reference state would cause that much higher increase in the new Kalman filter reference state estimate, the higher was the difference between the measurement and the previous reference state estimate. Such difference would also mean that the new measurement would be closer to the upper alarm generation threshold value causing the change of the background color towards the red color extreme.

Therefore, when a certain (positive or negative) Kalman filtering trend (difference) would be observed in each two consecutive reference state estimates, this would also translate to an appropriate background color adjustment. Such color coding did not consider particular reference state trend thresholds but rather the actual measurement proximity to the potential patient specific alarm generation thresholds. As such, the color coding would always adequately indicate alarm occurrence in the automated alarm management algorithm. As previously discussed on the other hand, the automated alarm management algorithm parameters (threshold bounds) could be adjusted to correspond to the particular patient conditions the physicians would expect. The supplemental background color coding would provide visualization of the patient reference state trends and could potentially indicate occurrence of alarms in the automated alarm management algorithm, as additional information to the physicians.

6.6 Weather influences

Historic weather records were used from one of the weather stations in Vienna, Graz, Innsbruck, Klagenfurt or Linz, closest to the patient locations. However, weather data from the exact patients' addresses were not available. Maximum distances between the weather station and patient locations amounted up to 80 km. Although occurring in only a few cases, such distances might have adversely affected reliability of weather assignments to particular patient conditions, as local weather conditions might have been different than those at such distant weather stations.

As majority of the patients participated in the study over the period of 6 months (180 days as illustrated in Figure 4.2), it was not possible to take into account the effect of varying weather conditions on each individual patient during a calender year.

The identified statistically significant influences of thermal stress on patient physiological conditions were identified using Spearman correlation coefficients, suitable for determining nonlinear relationships. The considered thermal stress was based on mean temperature differences. Such differences were calculated based on the mean daily temperatures available in the obtained meteorological data. Following the meteorological practice, these values represented average between maximum and minimum recorded daily temperatures. As such, mean daily temperatures could be considered as proxies for temperature variation. Consequently, the obtained statistically significant results could offer practical opportunity to predict heart failure patient health conditions based on the meteorological forecasts, typically including maximum and minimum daily temperatures.

Attempt to relate thermal comfort with patient physiological conditions using effective outdoor temperature did not result in statistically significant correlations. However, such approach did not take into account the actual exposure of patients to particular environmental conditions, dependant also on patient clothing, wind exposure, solar radiation.

Rather, the applied Equation 5.10 only took into account outdoor temperature and relative humidity, as these were available in the weather data records. Numerous other equations existed in the literature to quantify personal thermal comfort, but their application would have required additional data. Such additional approaches could be investigated in the future, providing a more comprehensive telemonitoring data collection. Apart from the additional information about the outdoor weather conditions, records of patient exposure time to such conditions, as well as indoor environmental parameters, might be used in further investigations.

In addition to thermal comfort, the correlation analyses also used the available outdoor weather monitoring records. Although statistically significant correlations were identified for a number of individual patients, such correlation values were in absolute terms rarely above 0.7, which was considered typically desirable in research related to human subjects taking into account person-to-person variability. Similarly, statistically significant but low correlation values were observed between the weather conditions and patient alarm occurrence. In such analyses, linear interpolation was used to derive the weather conditions for the exact time of the stored patient physiological measurements, as only three weather monitoring records were available per each day: at 7:00, 14:00 and 19:00 hours. Such interpolation sometimes resulted in unreasonable estimates, particularly in terms of relative humidity exceeding 100%. All such conditions were replaced with the maximum physically possible relative humidity conditions of 100%. Although more frequent weather monitoring would have provided better insight into the outdoor conditions and alleviated the need for interpolation of the weather data, the influence of weather conditions on the patient physiological parameters would remain questionable. As the patients would be located indoors at the time they submitted the telemonitoring records, they would not be effectively exposed to the outdoor weather conditions.

Therefore, a more tangible approach was to use an overall indicator of particular weather conditions during each telemonitoring day (e.g. mean, or difference between maximum and minimum outdoor temperatures) together with the averaged daily physiological parameters. The environmental temperatures also had the highest observed correlations with the physiological measurements. Therefore, the analysis included differences in patient physiological measurements across the days with pronounced temperature variation, thermal stress. The considered variation in temperature differences for investigation of significant differences in recorded patient physiological measurements were conditioned upon R implementation of statistical testing functions. The investigated cases were designed to take into account a possible delay in manifestation of weather condition influences on the telemonitored patient physiological parameters. Examining 2 and 3 consecutive telemonitoring days, delay of up to 1 day was investigated.

The identified statistically significant correlations between individual patient physiological parameters and weather conditions, such as those for diastolic blood pressure presented in Figure 5.25, revealed high dependency on individual patient conditions and could not be generalized. These results included both, rising and falling outdoor temperatures, as well as relative humidity, atmospheric pressure and calculated effective temperatures. However, statistically significant increase in both, systolic and diastolic blood pressures, were identified in the case of cold stress: two consecutive days with declining mean temperatures. A one day delayed effect of cold stress was found on falling heart rates. Also, heat stress had a one day delayed effect on rising diastolic blood pressures.

The detected statistically significant differences in patient physiological measurements around the thermally stressful days were mostly found between the systolic and diastolic blood pressures. All such cases had 95% confidence intervals ranging from trivial (close to zero) to possibly important differences comparable to effects of certain blood pressure medications (~ 10 mmHg (Wu et al., 2005)). Since the confidence interval ranges included values close to zero, no strong conclusion could be reached about the importance of the observed statistically significant differences in physiological measurements. However, repeated thermal stress over several consecutive days would have multiplicative effect on the patient conditions. The current study did not distinguish such cases but rather focused on any days satisfying the predefined thermal stress criteria.

Although no statistically significant correlations were found between the occurrence of alarms and weather conditions, indirect weather influence on alarm occurrence could be inferred. Namely, statistically significant correlations between the outdoor mean temperature variation and blood pressures were detected, whereas blood pressure alarms were the most prevalent of all the alarm occurrences. Furthermore, they accounted for 60% of all true alarms.

6.7 Limitations

Unavailable patient gender and age records limited the possibility to investigate effects of these parameters in the performed analyses. Furthermore, it could not be verified whether the selection of patients in terms of gender and age distributions formed representative sample of the actual population, although it was assumed that the conducted monitoring study was properly designed in this respect.

Patient distribution map, presented in Figure 4.1, showed that the majority of sampled patients were located in the vicinity of Graz. Such selection of patients failed to ensure representative number of patients for each of the Austria's climatic or geographic regions.

One measurement per day recorded by the majority of patients in the study might not have been a representative measurement for each particular day. Sampling interval of measurements would have to be short enough for the interpolated time series to provide a close approximation to the original continuous signal being monitored, such as the physiological patient conditions. Here, sequential measurements in time over a fixed interval represent a sampling interval while the resulting data form a time series.

Limitation of the presented approach was related to the available alarm records. Namely, the original data collection algorithm allowed physicians to manually adjust threshold values during the course of the measurements. Consequently, the same values of measured parameters might have caused an alarm before the threshold was adjusted while indicating normal condition after such adjustment, and vice versa. Thus, in some cases alarm generations might have been impacted by physician interaction with the monitoring algorithm resulting in inconsistent occurrence of alarms among the patients. Such alarms might not actually reflect any changes in the patient conditions, but rather result from the tuning of the alarm generation algorithm.

Another limiting factor was related to a certain number of physician responses which might have been a consequence of alarm events occurring during the days of previous measurements, rather than the day of the response records. This could be one of the reasons Figure 4.6 showed that 1% of physician responses occurred during the day when no alarms were recorded, while 5% of the total number of measurements indicated alarms without accompanying physician actions. However, such records of a possibly delayed physician actions existed in only a small number of events. Consequently, such cases were considered only when calculating the number of false events for the original data.

Further limitation of the obtained results originated from the comparison to the MOBITEL study, rather than clinical trial tests.

Limitations of the current results were related to the poor documentation of the physician responses to the alarms since no information existed to distinguish physician actions to specific monitoring parameters, irregular or missing measurements. Also, no information was available about potential other contacts the physicians might have had with the patients, not influenced by the telemonitoring system, e.g. periodic patient visits to refill their supply of medications etc.

Although certain descriptive statistics was presented for the prescribed medications, effects of medication intake were not considered in analyzing the patient physiological measurements. Neither medical nor pharmacological experts were available during the course of the current analyses, which may also present a potential limitation.

Finally, the available weather data records included only a limited number of days with severe heat or cold stress conditions. Due to sometimes very small number of records with certain severe mean daily temperature variations, statistically significant health influences of more severe variations in daily weather conditions than presented in Section 5.6 might have remained unproven. In fact, contrary to the expectations, ever more significant findings did not occur with the increasing thermal stress, possibly due to the reduced sizes of the tested data sets and rarer occurrences of such extreme environmental conditions. This could also be one of the reasons that heat stress had no statistically significant influences on patient physiological measurements. Additional explanation could be related to patient avoidance of extreme outdoor temperature variations, during which they might have rather stayed in more temperate indoor environments. Further verification of such assumptions would be needed.

Chapter 7

Conclusions and Future Work

The current study analyzes occurrence of alarms in the existing home telemonitoring records of heart failure patients after an episode of acute decompensation. The study offers ways to improve the alarm generation towards increasing accuracy and reduction of false positive alarms. For the analyses of therapeutic decisions and physiological measurements, R statistical software is used.

New 5-level alarm flag is proposed to supplement the original algorithm towards more efficient and precise home telemonitoring system. The presented approach describes automated alarm generation algorithm for telemonitoring of heart failure patient conditions. The approach considers patient blood pressure, heart rate and weight conditions concurrently within the data analyses algorithm. The proposed algorithm is based on patient reference state estimation using Kalman filtering of monitoring data and automated threshold adjustment around the reference state. The results show that the automated threshold adjustment method could complement the existing telemonitoring systems and help the clinicians to identify potentially important changes in the patient health status.

Although further home telemonitoring trials and testing are suggested, the presented method shows promising capabilities in estimating patient conditions and could make valuable contribution to facilitate the chronic heart failure patient healthcare.

7.1 Scientific contributions

The most significant scientific contributions of the dissertation are the following:

- Introduced new 5-level home telemonitoring alarm risk indicator. The indicator classifies the alarms based on the probability that such alarms would require medical interventions.

- Introduced automated home telemonitoring alarm generation algorithm featuring dynamic threshold adjustments. The algorithm personalizes alarm generation thresholds subject to patient conditions throughout the telemonitoring, reducing the need for manual adjustments by the physicians.

- Developed methodology for selecting algorithmic setup for automated home telemonitoring alarm generation such that the desired levels of sensitivity or specificity could be achieved. The methodology includes mathematical relations for evaluating impacts of particular combinations of alarm generation thresholds.

- Developed tool for graphical visualization of data series trends supplementing the automated home telemonitoring alarm generation system. The provided color visualisation corresponds to and follows the patient physiological conditions indicating approach to potentially critical values causing alarm occurrence.

- Developed home telemonitoring based methodology for analyses of influence of weather conditions (temperature, humidity and atmospheric pressure) on chronic heart failure patients. Statistically significant correlations between variations in systolic blood pressure and cold stress days are identified.

7.2 Open research work

The current scientific contributions open possibilities for future work in home telemonitoring applications. Some of the open research possibilities are: risk parameter modeling, subjective health perception analyses, enhancement of patient health literacy, self-care and self-management, medication therapy modeling, model validation and testing, training for effective usage of the home telemonitoring system, weather influence modeling, pilot and demonstration activities.

In terms of risk parameter modeling, examination of possible risk factors could include combination of two or more physiological parameters (e.g. gain in body weight followed by systolic blood pressure decrease, systolic and diastolic blood pressure difference etc.). Based on the examined risk factors new alarm indicators could be introduced.

As additional parameters to support the model in predicting patient health status, subjective measures could be used in conjunction with the physiological ones (blood pressure, heart rate and weight). Subjective descriptors can explain self-rated wellbeing factors, e.g. self-rated mobility, health, anxiety/depression, the need for extra pillows at night, overeating, swollen ankles, pain/discomfort. Such subjective descriptors could be used to trigger alarms and could be meaningful in patient health monitoring, as additional medical reports. In order to effectively include subjective measures in home telemonitoring based decision support, further research should identify the most appropriate predictors.

Future work should further investigate possibilities for automated physicians' decision support including recommendations for medication therapy. Potential influence of certain medication intake on patient physiological measurements should be also taken into consideration. For research purpose, direct contacts and feedback from the physicians involved in the telemonitoring should be considered. Furthermore, the physician contacts might be useful to provide necessary clarifications of retrospective data records and additional insight into particular decisions, which were not available in the current study.

The presented automated home telemonitoring algorithm should be validated in the future pilot and demonstration studies. Extensive tests are to confirm impacts of the introduced enhancements on system acceptance and effectiveness, number of hospitalizations and health care costs including patient expenses. Such tests should especially target the uncertainty of alarm classification between true and false, as one of the main goals in present day telemedicine. Namely, the physicians should have a chance to examine and verify all the automatically generated alarms. As a result, the system should be sufficiently reliable excluding nuisance false alarms.

Prospective training of medical staff should enable effective usage of the home telemonitoring system to provide timely and appropriate alarm responses, as well as effective therapy management. At the same time, training of the patients should ensure adequate usage of equipment and regular monitoring. Patients should be provided tailored advices about healthy lifestyle (e.g. nutrition, physical activity, stress management), motivational messages, as well as key system performance indicators (e.g. cost effectiveness, clinical benefits) to encourage their continuous participation in home telemonitoring and adherence to the prescribed therapy. Thus, the system should ensure both patients' and care-givers' satisfaction.

Potentially, the biggest benefit of the home telemonitoring technology of chronic heart failure patients can be expected in the countries with low healthcare expenditures (e.g. Portugal, Ireland, Greece which dedicate less than 5% of their GDP on healthcare). As such countries typically also have low hospital bed ratio and short average duration of hospital stay (e.g. Ireland and Spain), the telemonitoring technology offers the possibility to provide enhanced healthcare to more remote patients, reducing their hospitalization rates. Thus, the future studies should consider applicability of the same home telemonitoring system across a range of countries with different healthcare expenditures to demonstrate the results under various social and economic conditions.

Integrating the information on forecasted weather conditions into the home telemonitoring system could additionally support the medical staff in timely decision making concerning patient health status and alarm situations. Heat/cold stress forecasts could be used to generate beneficial warnings to the patients and medical staff with the purpose of prevention and minimization of potentially adverse events. For example, 12, 24 and 48 hours before a forecasted thermal stress period patients could be advised of precautionary actions via their telemonitoring communication devices. The heat/cold stress prevention plan should be developed for the physicians to apply different measures of coping with such climate effects. Future work will focus on defining the levels of weather alerts, with respect to the patient health status, such as: watch, warning, advisory. The identified correlations between patient vital signs and weather conditions should be further investigated and confirmed in a prospective demonstration setting. In order to properly take into account and analyze the variation of annual weather conditions on patient health parameters, each individual patient telemonitoring should last for at least one calender year. Such results could support the creation of short, medium and long term bio-meteorological prognosis to consider the impact of local climatic conditions on chronic heart failure patients. Predicting influence of changing weather conditions on patients' physiological status should also take into account thermal efficiency of the patients' dwellings, e.g. insulation and window properties.

Appendix A

Statistical methods

This section briefly describes statistical methods and indices used in data analyses to achieve the investigation objectives.

Mean and Median are used measures of location (central tendency) to represent the dataset. Mean is calculated as an average of all values in the dataset, whereas the central value of the sorted sequence represents median. Mean and median are equal when data follow normal distribution. As a rule of thumb, median should be used to represent the data which is not normally distributed. As half of the data is smaller than median, it is also called second quartile. Accordingly, the first and third quartiles are values in the sorted sequence of which one and three quarters of the data are smaller, respectively. Interquartile range (IQR) is the difference between the third and first quartiles, containing half of the dataset. IQR together with standard deviation and variance represent measures of data variation.

Sensitivity and specificity indices are often employed to express decision performance of a statistical model in relative terms. Sensitivity represents a ratio between true positive (TP) decisions and actually positive cases, whereas specificity is equal to the ratio between true negative (TN) decisions and actually negative cases, as presented by Equations A.1 and A.2.

$$Sensitivity = \frac{\text{Number of True Positive } (TP) \text{ decisions}}{TP + \text{Number of False Negative } (FN) \text{ cases}} \tag{A.1}$$

$$Specificity = \frac{\text{Number of True Negative } (TN) \text{ decisions}}{TN + \text{Number of False Positive } (FP) \text{ cases}} \tag{A.2}$$

As such, sensitivity and specificity range between 0 and 1 indicating the accuracies for determining positive and negative cases, respectively.

Accuracy of the model is equal to the ratio between the correct decisions and total number of cases, as calculated by Equation A.3. In an ideal case, with no false negative (FN) and no false positive (FP) occurrences, sensitivity, specificity as well as the overall accuracy are equal to 1.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{A.3}$$

Values of sensitivity and specificity may vary in a particular decision support algorithm subject to the changing algorithmic parameters. In such cases it is often convenient to represent dependency between the sensitivities and associated specificities via the Receiver Operating Characteristic (ROC) curves.

ROC curves show the values of 1-Specificity on the abscissa and Sensitivity on the ordinate axis. An ideal point (0, 1) $\equiv$ (Specificity = 1, Sensitivity = 1) is located in the upper left corner of a ROC curve.

Moving average as a smoothing technique for measured data replaces each measurement, f_t, with an average, g_t, over a certain number of neighbouring points. The neighbourhood can include points prior, after, or both, prior and after, each measurement. Equation A.4 defines $(2n+1)$-point moving window size for averaging data around each measurement point (including both points prior and after the measurement).

$$g_t = \sum_{i=-n}^{+n} w_i f_{t+i}, \text{where } w_i = \frac{1}{2n+1} \tag{A.4}$$

Savitzky-Golay filter can be obtained using the same Equation A.4 and changing the values of coefficients w_i according to the higher order approximation (typically quadratic). As opposed to moving average, which approximates the data using a constant (mean) value within a specified moving window, the main advantage of Savitzky-Golay approach is preservation of distribution features such as relative maxima, minima and width, which are usually not invariant to the other filters, e.g. moving average.

FFT is an algorithm to calculate discrete Fourier transformation, converting equally spaced function samples, f_t, into the frequency domain according to Equation A.5. FFT can be used as a smoothing technique to remove the noise from the recorded signals and can be a good tool for identification of the most important frequencies in the data set.

$$F_k = \sum_{t=0}^{N-1} f_t \left(cos\left(\frac{2\pi kt}{N}\right) - isin\left(\frac{2\pi kt}{N}\right)\right) = \sum_{t=0}^{N-1} f_t e^{-i\frac{2\pi kt}{N}} \tag{A.5}$$

In Equation A.5, N is the number of measurement points and k takes the values between 0 and $N/2$ in steps of one (Smith, 1999). Result of the transformation is a vector of complex numbers, F, consisting of real and imaginary parts. Real parts of the vector provide amplitudes and frequencies in terms of cosine functions, whereas imaginary parts provide the same information in terms of sine functions. The number of real and imaginary parts is equal to half the number of original points plus 1.

In the frequency domain, the main components will have the highest amplitudes and low frequencies, whereas the noise will have high frequencies and low amplitudes. Filtering out the noise is performed by setting the high frequency components equal to zero. The downside of FFT procedure is the assumption of periodicity. Namely, regardless of which window size is employed for FFT calculations, the underlying equations assume that the next data point after the considered window size will be equal to the first data point within the considered window size.

Variance as measure of uncertainty can be used to quantify spread of data points within a particular dataset. It is equal to the square of standard deviation, s, and can be calculated by Equation A.6.

$$var = s^2 = \frac{\sum_{i=1}^{N}(f_i - \bar{f})^2}{N-1} \tag{A.6}$$

Here, N is the number of data points f_i, having the mean of $\bar{f}$.

Autocorrelations can be calculated using Equation A.7.

$$r_k = \frac{\sum_{t=max(1,-k)}^{min(N-k,N)}(f_t - \bar{f})(f_{t+k} - \bar{f})}{\sum_{t=1}^{N}(f_t - \bar{f})^2} \tag{A.7}$$

Here k is lag, f_t are some of N data points, and $\bar{f}$ is the mean of all data points. Statistically significant positive autocorrelations for one or more lags indicate non-randomness in the data.

Kalman filter is an iterative algorithm widely used in engineering and natural sciences to capture dynamics and calculate forecasts of time series data. At each measurement time new Kalman filter forecast is obtained from the previous forecast and previous measurement. Characteristic of Kalman filter is minimization of forecast variances (Harvey, 1990). Kalman filter model in the state space formulation assuming linear combination of states can be represented by Equation A.8

$$g_t = \theta_t + \nu_t \tag{A.8}$$

where g_t represents reference state, θ_t is mean state and ν_t is oscillation parameter. At every time instance t, θ_t can be estimated from the values at $t-1$, θ_{t-1}, according to Equation A.9, assuming that the model parameters are constant in time.

$$\theta_t = \theta_{t-1} + w_t \tag{A.9}$$

Here, w_t represents noise oscillation. Initially (at $t = 0$), the mean value θ_t is assumed to be normally distributed $N(m_0, c_0)$, where m_0 is initial mean and c_0 is initial variance of the distribution. Also, the model assumes that the oscillation parameters ν_t and w_t have normal distributions $N(0, \nu)$ and $N(0, w)$, respectively.

Chebyshev's Theorem states that for an arbitrary distribution of N observations, at least $N \cdot (1 - 1/k^2)$ observations will be within the range of k standard deviations around the mean (where $k > 1$) (Lohninger, 2012).

Regression is a statistical process to find the relationship between independent (input, descriptor, x_i) and a dependent (target, response, y) variables. Linear regression assumes linear relationship according to Equation A.10.

$$y = a_0 + \sum_{i=1}^{k} a_i x_i + \epsilon \tag{A.10}$$

Parameters a_0, a_1, ..., a_k are regression coefficients, while ϵ is the error or the residual, having the mean of zero. Regression coefficients are determined minimizing the sum of squared errors. The analysis of variance (ANOVA) Table A.1 for multiple linear regression

Source of variation	Degrees of freedom	Sum of squares	Mean of squares	F-value
Regression	k	$SS_{reg} = \sum(\hat{y}_i - \bar{y})^2$	$MS_{reg} = SS_{reg}/k$	$\frac{MS_{reg}}{MS_{res}}$
Residual	$n-k-1$	$SS_{res} = \sum(y_i - \hat{y}_i)^2$	$MS_{res} = SS_{res}/(n-k-1)$	
Total	$n-1$	$SS_{tot} = \sum(y_i - \bar{y})^2$		

Table A.1: Analysis of variance for multiple linear regression with k independent variables and n observations (y_i are observed responses having a mean $\bar{y}$, while $\hat{y}_i$ are predicted (model) responses)

(MLR) shows calculations with k independent variables and n observations. Here, y_i are observed responses having a mean $\bar{y}$, while $\hat{y}_i$ are predicted (model) responses. Residual sum of squares, SS_{res}, can be used to estimate standard errors of regression coefficients.

F-value represents the result of testing the hypothesis that all regression coefficients are equal to zero. The value follows F-distribution with k and $n-k-1$ degrees of freedom. As a rule of thumb F-values above 100 indicate valid MLR models.

F-distribution is a continuous probability distribution characterized by two degrees of freedom parameters, $F(n_1-1, n_2-1)$. It results from calculating the ratio between variances (s_1 and s_2) of two normal distribution samples with n_1 and n_2 observations, according to Equation A.11.

$$F = \frac{s_1^2}{s_2^2} \tag{A.11}$$

As F-values can be only positive, F-distribution is skewed to the right. Normal, t, and chi-square distribution can be considered as special cases of F-distribution: $F(1,\infty)$, $F(1,n_2)$, and $F(n_1,\infty)$, respectively.

Coefficient of determination, Quality or Goodness of fit, r^2, can be calculated as a square of the Pearson's correlation coefficient. In regression, goodness of fit can be obtained from Equation A.12, using the calculated residual (SS_{res}) and total (SS_{tot}) sum of squares.

$$r^2 = 1 - \frac{SS_{res}}{SS_{tot}} \tag{A.12}$$

Goodness of fit indicates portion of the total variance explained by the regression model.

Standard error of a sample size n can be calculated from the standard deviation, s, according to Equation A.13.

$$SE = \frac{s}{\sqrt{n}} \tag{A.13}$$

In the case of regression, standard error of regression coefficients a_i, SE_{a_i}, can be calculated from Equation A.14.

$$SE_{a_i} = \frac{s_y}{s_{x_i}}\sqrt{\frac{1-r^2}{(1-r_{x_i}^2)(n-k-1)}} \tag{A.14}$$

Here, s_y and s_{x_i} are standard deviations of y and x_i, r^2 is goodness of fit when regressing y onto $X = \{x_1, x_2, ..., x_k\}$, while $r_{x_i}^2$ is goodness of fit when regressing x_i onto $X \backslash x_i$. The estimated standard error follows t-distribution with $n-k-1$ degrees of freedom.

Confidence intervals for each coefficient can be calculated according to Equation A.15.

$$(1-\alpha)\%CI = (a_i - t(\alpha/2, n-k-1)\cdot SE_{a_i}, a_i + t(\alpha/2, n-k-1)\cdot SE_{a_i}) \tag{A.15}$$

Here, $t(\alpha/2, n-k-1)$ is the critical value for probability (level of significance) $\alpha/2$ from t-distribution with $n-k-1$ degrees of freedom.

Student's or t-distribution is formed by calculating the differences between the sample and population means, $\bar{x}$ and μ, according to Equation A.16.

$$t = \frac{\bar{x}-\mu}{s/\sqrt{n}} \tag{A.16}$$

Here, s is the sample standard deviation and n is the sample size.

t-Distribution is lower and wider than normal distribution for a small sample (number of degrees of freedom) and approaches standard normal distribution for a large sample size. As a rule of thumb, large sample is considered to have more than 30 sampling points.

Normal distribution is one of the most important statistical distributions, often occurring in nature. It is symmetric with respect to the mean, μ, and additionally characterized by variance, σ, typically denoted as $N(\mu, \sigma)$. Probability density function of the normal distribution is given by Equation A.17.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{A.17}$$

The term standard normal distribution is used for a normal distribution having a mean of zero and standard deviation of 1.

Variance inflation factor (VIF) for each variable x_i can be calculated from Equation A.18.

$$VIF_{x_i} = \frac{1}{1-r_{x_i}^2} \tag{A.18}$$

VIF shows the influence of collinearity on variance of each variable. For example, VIF equal to 7 would indicate 7 times larger variance than in the case when there is no correlation between the variables. As a rule of thumb, VIF larger than 10 indicates strong collinearity and such variables should not be considered as descriptors in regression. They could be removed, aggregated or preprocessed by principle component regression.

Outliers could present a potential problem for regression. Outliers are extreme, numerically distant values in the sample, that stand out from the rest of the data. As such, they may have substantial influence on statistical analyses and should be removed from the dataset.

Assumptions for linear regression are the following:

- Expected relationship between the dependent and independent variables is linear;
- Measurements are mutually independent, do not exhibit trends or correlation to another variable;

- For each independent variable, dependent variable values are normally distributed and have the same variance.

Checking the data and the residuals should ensure that these assumptions are fulfilled. Particularly, the residuals should:

- be normally distributed and symmetric around zero;
- be uncorrelated (Durbin-Watson test), not having any patterns;
- not exhibit dependence on the independent variable.

Durbin-Watson statistic is used to test correlation between the subsequent residuals. The test is based upon the assumption that successive residuals ϵ_t and ϵ_{t+1} are correlated with a correlation coefficient ρ ($\mid \rho \mid < 1$) according to Equation A.19.

$$\epsilon_{t+1} = \rho\epsilon_t + \omega_{t+1} \tag{A.19}$$

Here, the parameter ω is normally distributed with a mean of zero and constant variance. Durbin-Watson statistic, d, is calculated from Equation A.20, where n is the number of observations (residuals, degrees of freedom).

$$d = \frac{\sum_{t=1}^{n-1}(\epsilon_{t+1} - \epsilon_t)^2}{\sum_{t=1}^{n}\epsilon_t^2} \tag{A.20}$$

If the calculated value of the statistic d is less than the critical value, $d_{L,n}$, the null hypothesis that the residuals are not correlated ($\rho = 0$) is rejected.

Lilliefors test verifies that the tested dataset follows a predefined probability distribution. It is a modification of the Kolmogorov-Smirnov test used for evaluating whether a sample is part of the normal distribution. Lilliefors test is applicable when the parameters of the reference distribution are not exactly known. Typically tested is the null hypothesis that the sample is part of a normal distribution. The hypothesis is rejected if the calculated probability for obtaining the test statistic is lower than the critical level of significance. The test statistic, D, is maximum absolute difference between the predefined (hypothetical) and dataset cumulative distribution functions, specified by Equations A.21 through A.23.

$$D = max\{D^+, D^-\} \tag{A.21}$$

$$D^+ = max_{i=1,\dots,n}\{i/n - p_{(i)}\}, D^- = max_{i=1,\dots,n}\{p_{(i)} - (i-1)/n\} \tag{A.22}$$

$$p_{(i)} = \Phi([x_{(i)} - \overline{x}]/s) \tag{A.23}$$

Here Φ is the predefined (typically standard normal) cumulative distribution function, $\overline{x}$ is the dataset mean, and s is the dataset standard deviation.

Principle component analysis (PCA) is a method for identifying orthogonal directions (axes) along which there is maximum variability of data and then projecting (transforming) the data into such defined multidimensional space. The identified directions (axes) are called principle components. As principle components are orthogonal to each other, correlation between any two of them is equal to zero. Principle components are eigenvectors calculated as a result of eigenanalysis of the square matrix $\mathbf{Z}$=$\mathbf{A}^T\mathbf{A}$, formed out of the data matrix $\mathbf{A}$ in one of the following ways:

- scatter matrix - directly using the data without any scaling,
- variance-covariance matrix - variable means are subtracted from the corresponding data points,
- correlation matrix - each variable is standardized (mean subtracted and the difference divided by standard deviation).

The data matrix, **A**, is formed in such a way that each column represents one of p different variables (descriptors), whereas rows contain different observations or samples.

The identified eigenvectors **E** of the p by p square matrix **Z** satisfy Equation A.24.

$$\mathbf{ZE} = \mathbf{E}diag(\lambda_1, ..., \lambda_p) \tag{A.24}$$

Scalars $\lambda_1, ..., \lambda_p$ are the eigenvalues. As eigenvectors, **E**, are orthogonal to each other product $\mathbf{E}^T\mathbf{E}$ is the identity matrix, **I**.

Principle component regression (PCR) applies multivariate linear regression on principle components. Usually only the first few principle components are sufficient for the statistical model to explain the largest portion of the variance in the target dataset. As such, the reduced dimensionality of the principle component domain has advantages over the MLR:

- noise of the data remains in the residuals, as the eigenvectors with low eigenvalues are excluded from regression analyses,
- regression preconditions are automatically satisfied as eigenvectors are orthogonal and independent from each other.

Trend is a non-periodic systematic change in time series data. The simplest trend representation is linear increase or decrease which could be derived from a linear regression model over a certain window size. As forecasting usually relies on the assumption that the current trends will continue, trend can be used to make reasonable approximations of future events (Cowpertwait and Metcalfe, 2009).

Correlation analysis aims to determine relation between the variables. Measures of correlation are correlation coefficients, taking the values between -1 and +1. Correlation coefficient is a random variable which follows a distribution depending on the sample size and actual correlation in the population. In the case of small samples probability is small that the calculated correlation coefficient will represent the true correlation in the population. For example, correlation of 3 sample pairs from the uncorrelated populations will more likely approach extreme, +1 or -1, than any other values. Correlation of 4 sample pairs from the uncorrelated populations is equally likely to take any value between -1 and +1. As a rule of thumb, only correlation coefficients above 0.8 will present significant results in the case of 10 sample pairs, and above 0.5 in the case of 20 sample pairs (Lohninger, 2012). Furthermore, correlation above 0.7 is considered as typically desirable in research related to human subjects taking into account person-to-person variability (Larose, 2006).

Pearson's correlation coefficient is measure of linear correlation between the two variables (datasets), x and y, with the same number of observations, n. It can be calculated from Equation A.25.

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{A.25}$$

Apart from the linear relation between the variables, calculation of Pearson's correlation coefficient assumes continuous random variables, normally distributed and independent of each other.

Spearman's rank correlation can be applied in the case of non-linear relations between the variables. In such cases, calculation of Pearson's correlation coefficient would provide wrong results. The difference in calculation of Spearman's correlation coefficient is usage of ranks rather than the variable values. Therefore, the same Equation A.25 can be applied, but using variable ranks, $rank(x)$ and $rank(y)$, instead of the observed values, x and y. Taking into account differences in ranks, Δ_i, between the corresponding ranked variables, Spearman correlation can be also calculated from Equation A.26.

$$\rho = 1 - \frac{6 \cdot \sum_{i=1}^{n} \Delta_i^2}{n(n^2 - 1)} \tag{A.26}$$

A minimum of five observations is required ($n > 4$).

Tetrachoric correlations are calculated between the binary (dichotomous) variables. The underlying assumption for the calculation is the normal distribution of the inferred (hypothetical or hidden) variables forming the basis for dichotomy. For example, an alarm indication (binary variable) may occur due to exceeded thresholds by normally distributed measurements (inferred variable).

Binary (dichotomous) variable can only take one of the two possible values (states or categories), e.g. 0 or 1, true or false, -1 or +1 etc. Such variables may occur "naturally" (e.g. gender, pregnancy) or originate from another continuous quantitative variable, classified into two categories (e.g. critical pressure level, warning light for low fuel level in the car).

Polychoric correlations are generalization of tetrachoric correlations in the case when variables can take one of three or more possible values (states, categories). When such categories are ordered we talk about ordinal, and otherwise about categorical variables.

Polyserial correlations are calculated between a quantitative and ordinal variable. The underlying assumption for calculation is bivariate normal distribution of the quantitative and inferred variables.

Central Limit Theorem states that the sum of independent variables with finite variances will have normal distribution. The consequence of the theorem is one of the most important rules in statistics (Lohninger, 2012). Distribution of means of sufficiently large random samples from any population will approach normal distribution, with a mean approaching the mean of the population and a standard deviation equal to the ratio between the standard deviation of the population and a square root of the sample size. Normal distribution is thus the foundation of many statistical methods. As a rule of thumb, a sample size larger than 30 will be normally distributed for most population distributions. Means of smaller samples exhibit t-distribution with a number of degrees of freedom equal to the sample size minus 1. Repeated measurements could result in improved precision.

Statistical testing provides guidance to determine whether it is reasonable to assume that a certain population parameter is equal, larger or smaller than a certain value. The parameters can be population mean or variance. Similar to the court trial where a person is considered innocent until proven guilty beyond the reasonable doubt, in statistical tests we assume that a certain null hypothesis is considered true until proven otherwise with a certain level of significance. Level of significance is the probability that the null hypothesis may be wrongfully rejected. Statistical test is a five-step procedure:

1. Formulate null and alternative hypotheses,
2. Specify level of significance - typically 5%,
3. Calculate test statistic - a value upon which to base decisions,
4. Define rejection region - test statistic critical value for rejecting the null hypothesis,
5. Select appropriate hypothesis.

t-Test is used to determine whether a sample mean is significantly different (two-tailed test), smaller or larger (one-tailed tests) than a certain value. In such cases we talk about one-sample tests. The test can also be used to compare means of two different samples: two-sample test. One-sample t-test statistic is calculated by Equation A.16, with $n-1$ degrees of freedom, where μ is not the population mean, but rather the value to which the sample mean is compared. When comparing two samples, x_1 and x_2, t-test statistic is calculated by Equation A.27.

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{s_p^2(\frac{1}{n_1} + \frac{1}{n_2})}} \tag{A.27}$$

Here, n_1 and n_2 are sample sizes, and s_p^2 is pooled variance, which can be calculated from sample variances, s_1 and s_2, according to Equation A.28.

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2} \tag{A.28}$$

The number of degrees of freedom is equal to $n_1 + n_2 - 2$.

One sample t-test null hypothesis can be one of the following:

- Mean of the underlying population of the sample is greater or equal to a certain value. This hypothesis is to be rejected if the calculated t statistic is smaller than the negative critical value, at the tested level of significance.
- Mean of the underlying population of the sample is smaller or equal to a certain value. This hypothesis is to be rejected if the calculated t statistic is greater than the critical value, at the tested level of significance.

- Mean of the underlying population of the sample is equal to a certain value. This hypothesis is to be rejected if the absolute value of the calculated t statistic is greater than the critical value, at one half of the tested level of significance.

Two sample t-Test null hypothesis can be one of the following:

- Mean of the underlying population of one sample is greater or equal to the mean of the underlying population of the other sample. This hypothesis is to be rejected if the calculated t statistic is smaller than the negative critical value, at the tested level of significance.
- Mean of the underlying population of one sample is smaller or equal to the mean of the underlying population of the other sample. This hypothesis is to be rejected if the calculated t statistic is greater than the critical value, at the tested level of significance.
- Means of the underlying populations of both samples are equal. This hypothesis is to be rejected if the absolute value of the calculated t statistic is greater than the critical value, at one half of the tested level of significance.

Before applying the t-test to compare the two samples, the following precondition steps should be fulfilled:

1. The samples should be independent of each other: if both samples consider the same population members apply t-test for matched pairs, otherwise proceed to step 2.
2. Both samples should follow normal distribution: check by applying Shapiro-Wilk test; if fulfilled, proceed to step 3, otherwise apply two-sample U-test.
3. Both sample variances should be equal: check by applying Chi-Square test; if fulfilled apply t-test, otherwise apply Welch test.

When trying to test the differences between two procedures, methods, treatments etc. applied to the same population, two-sample t-test precondition of independent samples will not be satisfied. In such cases, **t-test for matched pairs** should be applied, using the pairwise calculated differences, d_i, as a variable that follows the t-distribution. The test statistic can be calculated from Equation A.29, where $\bar{d}$ and s_d are respectively the mean and standard deviation of the difference between the matched pairs, n_d is the number of matched pairs and d_0 is the reference value to which the mean difference between the matched samples is compared.

$$t = \frac{\bar{d} - d_0}{s_d / \sqrt{n_d}} \tag{A.29}$$

The number of degrees of freedom is $n_d - 1$.

Null hypothesis of the t-test for matched pairs can be one of the following:

- The mean difference between the two samples is greater or equal to a certain value. This hypothesis is to be rejected if the calculated t statistic is smaller than the negative critical value, at the tested level of significance.

- The mean difference between the two samples is smaller or equal to a certain value. This hypothesis is to be rejected if the calculated t statistic is greater than the critical value, at the tested level of significance.

- The mean difference between the samples is equal to a certain value. This hypothesis is to be rejected if the absolute value of the calculated t statistic is greater than the critical value, at one half of the tested level of significance.

When the t-test precondition for sample normal distributions is not satisfied, **two-sample U-test** should be applied instead of the t-test. U-test evaluates whether the underlying populations of the two samples have the same distributions, leading to the equality of their means. To calculate the U-test statistic both samples are arranged into a single ranked series. Sum of such ranks within the first sample is denoted as R_1, and within the second sample as R_2. Knowing the sample sizes, n_1 and n_2, U-statistic can be calculated from Equations A.30 through A.32.

$$U_1 = n_1 \cdot n_2 + \frac{n_1(n_1+1)}{2} - R_1 \tag{A.30}$$

$$U_2 = n_1 \cdot n_2 + \frac{n_2(n_2+1)}{2} - R_2 \tag{A.31}$$

$$U = min\{U_1, U_2\} \tag{A.32}$$

The tested null hypothesis is that the two sample distributions are the same. This hypothesis is rejected if the calculated U statistic is smaller or equal to the critical value, at the tested level of significance for the given n_1 and n_2 degrees of freedom.

Shapiro-Wilk test is used to evaluate the assumption that a sample belongs to a normal distribution. The test produces good results even in the case of small sample sizes. To calculate the test statistic, the sample should be ordered such that $x_1 < x_2 < ... < x_n$. The test statistic is calculated from Equation A.33, where the sample variance S^2 is calculated from Equation A.34, and b^2 is the slope of the regression line of the QQ plot calculated from Equation A.35.

$$W = \frac{b^2}{S^2} \tag{A.33}$$

$$S^2 = \sum_{i=1}^{n} (x_i - \bar{x})^2 \tag{A.34}$$

$$b = \sum_{i=1}^{k} a_{n-i+1}(x_{n-i+1} - x_i) \tag{A.35}$$

Here, n is the sample size, while $k = n/2$ in the case of the even number of observations in the sample, or $k = (n-1)/2$ otherwise. The coefficient a is determined from Shapiro-Wilk tables depending upon the sample size.

The tested null hypothesis that the sample is part of a normal distribution is rejected if the calculated W statistic is smaller than the critical value, at the tested level of significance.

Chi-Square test is used to compare empirical to a known parametric distribution. The test statistic, χ^2, can be calculated from Equation A.36.

$$\chi^2 = \sum_{i=1}^{k} \frac{(F_i - E_i)^2}{E_i} \tag{A.36}$$

Here, F_i are empirical and E_i are parametric (theoretical) frequencies for k frequency distribution bins.

The tested null hypothesis that the two distributions are equal is rejected if the calculated result is greater than the critical value, at the tested level of significance, with $k-3$ degrees of freedom.

Chi-Square (χ^2) distribution with $n-1$ degrees of freedom is obtained by dividing the variances of n-size random samples to the variance of a normally distributed population. The distribution mean is equal to the number of its degrees of freedom and represents half of its variance. The sum of two variables following χ^2 distribution also follows χ^2 distribution.

Welch test is a modification of t-test providing approximate solution in the case of unequal variances, s_1 and s_2, between the two samples, x_1 and x_2. The modified t-test statistic (t) and degrees of freedom (df) are calculated from Equations A.37 and A.38.

$$t = \frac{\overline{x_1} - \overline{x_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{A.37}$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}} \tag{A.38}$$

The same null hypotheses and rejection criteria hold as in the case of t-testing.

Appendix B

R functions

Table B.1 lists and describes the functions used from various R packages. Apart from R functions, several frequently applied procedures were coded as internal functions: `residual.plot`, `polychor.test`, and `polyserial.test`.

```
#: Function residual.plot displays histogram of input data together with a
#: normal distribution curve

residual.plot <- function(x) {
    x <- x[!is.na(x)]
    h <- hist(x, breaks=10, plot=F)
    xfit <- seq(min(x), max(x), length=100)
    yfit <- dnorm(xfit, mean=mean(x), sd=sd(x))
    yfit <- yfit*diff(h$mids[1:2])*length(x)
    h <- hist(x, breaks=10, col="grey", ylim=c(0, max(max(h$counts),
        max(yfit))), xlab="Input variable [units]", main="Histogram with
        Normal Distribution Curve")
    lines(xfit, yfit, col="blue", lwd=2)
}


#: Functions polychor.test and polyserial.test ease the calculation of
#: p-values for polychoric and polyserial correlations

polychor.test <- function(x, y) {
p <- polychor(x, y)
if (is.na(p)|abs(p)>=.99){
     p$statistic <- NA
     p$p.value <- NA
        p$rho <- p[[1]]
        p$var <- NA
        p$estimate <- NA}
    else {
```

```
        p <- polychor(x, y, std.err = TRUE)
        p$statistic <- p$rho / sqrt(c(p$var))
        p$estimate <- p$rho
        p$p.value = 2 * (1 - pnorm(abs(p$statistic)))
    }
    p
}

polyserial.test <- function(x, y) {
p <- polyserial(x, y)
if (is.na(p)|abs(p)>=.99){
        p$statistic <- NA
        p$p.value <- NA
        p$rho <- p[[1]]
        p$var <- NA
        p$estimate <- NA}
else {
        p <- polyserial(x, y, std.err = TRUE)
        p$statistic <- p$rho / sqrt(c(p$var))
        p$estimate <- p$rho
        p$p.value = 2 * (1 - pnorm(abs(p$statistic)))
    }
    p
}
```

R function	R package	Description
abline	graphics	Add one or more straight lines to a plot
abs	base	Calculate the absolute value
acf	stats	Calculate and plot estimates of autocovariance and autocorrelation function
arrows	graphics	Add arrows to a plot
as.Date	base	Convert to date from characters
as.integer	base	Create integer objects
as.numeric	base	Create numeric objects, i.e. numbers
as.POSIXlt	base	Create calendar date objects – date-time conversion
axis	graphics	Add an axis to a plot
barplot	graphics	Create a bar plot with vertical or horizontal bars
boxplot	graphics	Create box plot of the data
cbind	base	Combine objects by columns
cor	stats	Calculate (Pearson's, Spearman's or Kendall's) correlations between two vectors
cor.test	stats	Test for (Pearson's, Spearman's or Kendall's) correlation between paired samples
data	utils	Load data sets
data.frame	base	Create data frame
detrend	RSEIS	Remove trend from time series vector
dev.off	grDevices	Control multiple graphics devices
diff	base	Calculate differences between internal elements of an object at specified lags
dimnames	base	Retrieve or set the data locations of an object
dnorm	stats	Calculate normal distribution probability (density function)
expression	base	Create expression objects
fft	stats	Perform fast Fourier transformation of an array
filter	stats	Linear filtering of time series
function	base	Use to define functions
hist	graphics	Compute and plot histogram of the data
is.na	base	Indicate missing data in a vector using a logical variable
kfilter	sspir	Kalman filter for Gaussian state space model. Produce the conditional means and variances of the state vectors given the current time point
layout	graphics	Specify arrangements of multiple diagrams within a plot
legend	graphics	Add legend to a plot
length	base	Get or set the length of an object (including lists)
library	base	Load and list R packages
lines	graphics	Add connected line sections to a plot (A generic function taking coordinates given in various ways and joining the corresponding points with line segments)
list	base	Build R list of generic vectors enclosing other objects
lm	stats	Fit linear model to data
matrix	base	Create a matrix from the given set of values

R function	R package	Description
max	base	Return maxima of the input values
min	base	Return minima of the input values
months	base	Get a sequence of months out of date variables
mtext	graphics	Write text to plot margins
par	graphics	Retrieve or set graphical parameters
paste	base	Merge strings
pdf	grDevices	Produce PDF graphics
periodogram	GeneCycle	Calculate power spectral density related to Fourier transform
plot	graphics	Plot objects
plotmeans	gregmisc	Plot group means and confidence intervals
plot.new	graphics	Create a new plot frame
pnorm	stats	Calculate cumulative normal distribution probability
points	graphics	Add points to a plot
polychor	polycor	Calculate polychoric correlation (and its standard error) between two ordinal variables
polyserial	polycor	Calculate polyserial correlation (and its standard error) between a quantitative variable and an ordinal variables, based on the assumption that the joint distribution of the quantitative variable and a latent continuous variable underlying the ordinal variable is bivariate normal
rbind	base	Combine objects by rows
rcorr	Hmisc	Calculate a matrix of correlation (Pearson's or Spearman's) coefficients for all pairs of matrix columns
read.delim	utils	Used to read in delimited text files, where data is organized in a data matrix with rows representing cases and columns representing variables
read.table	utils	Read a file in table format and create a data frame from it, with cases corresponding to lines and variables to columns in the file
rep	base	Replicate elements of vectors and lists
return	base	Return from a function call to the main code. Used in defining new functions
rgb	grDevices	Create colors with given intensity of red, green and blue components
savePlot	grDevices	Save the current plot to a file
setwd	base	Set the working directory
shapiro.test	stats	Perform the Shapiro-Wilk test for normality
spectrum	stats	Spectral density estimation
sqrt	base	Calculate the square root
SS	sspir	Create a Gaussian state space model (for Kalman filtering)
stack	utils	Stack together multiple vectors into a single vector
stl	stats	Decompose time series into seasonal, trend and irregular components
stripchart	graphics	Create dot or scatter plots of the data (alternative to box plot)
sum	base	Return the sum of vector elements

R function	R package	Description
tapply	base	Apply a function over array elements, that is to each (non-empty) group of values
text	graphics	Add text to a plot
title	graphics	Add labels to a plot (plot annotation)
ts	stats	Create time-series object
ts.plot	stat	Plot several time series on a general plot
t.test	stats	Perform one and two sample Student's t-tests on data vectors
unique	base	Extract unique elements (without duplication)
unlist	base	Produce a vector of all the individual components in a variable
windowsFont	grDevices	Translate device independent font into windows font description
write.table	utils	Save data to a file

Table B.1: List of used R functions and their packages

Glossary

AAMI – **A**ssociation for **A**dvancement of **M**edical **I**nstruments.

ACCE – **A**merican **C**ollege of **C**linical **E**ngineering.

ACE – **A**ngiontensin **C**onverting **E**nzyme.

AGES – **A**gentur für **G**esundheit und **E**rnährungs**S**icherheit – Austrian Agency for Health and Food Safety.

AHA – **A**merican **H**eart **A**ssociation.

ANOVA – **AN**alysis **O**f **VA**riance.

Atm. – **Atm**ospheric.

AUC – **A**rea **U**nder the ROC **C**urve – used as a quality indicator of diagnostic models – values above 0.8 indicate good accuracy.

b.p. – **b**lood **p**ressure.

CHF – **C**hronic **H**eart **F**ailure.

CI – **C**onfidence **I**nterval.

ECRI – **E**mergency **C**are **R**esearch **I**nstitute.

EMBS – **E**ngineering in **M**edicine and **B**iology **S**ociety.

EMS – **E**lectromyostimulation.

ESES – **E**pidural **S**pinal **E**lectrical **S**timulation.

FFT – **F**ast **F**ourier **T**ransformation.

FN – **F**alse **N**egative events – cases with no alarms accompanied by one of the following physician actions: patient contact, medication adjustment or other action (excluding threshold adjustment).

FP – **F**alse **P**ositive events – alarms accompanied by one of the following physician actions: threshold adjustment or no action.

GDP – **G**ross **D**omestic **P**roduct.

GUI – **G**raphical **U**ser **I**nterface.

HDL – **H**igh **D**ensity **L**ipoprotein – or "bad" cholesterol.

HF – **H**eart **F**ailure.

IARS – **I**nternational **A**nesthesia **R**esearch **S**ociety.

ICD – **I**mplantable **C**ardioverter **D**efibrillator.

ICT – **I**nformation and **C**ommunication **T**echnology.

IENS – **I**mplantable **E**lectrical **N**erve **S**timulation – pacemaker.

IQR – **I**nter **Q**uartile **R**ange – difference between the values of 3^{rd} and 1^{st} quartiles.

Java – general purpose object oriented programming language.

JGR – **J**ava **G**UI for **R** (read Jaguar).

LDL – **L**ow **D**ensity **L**ipoprotein – or "good" cholesterol.

MIE – **M**edical **I**nformatics **E**urope – international conference.

MIMS – **M**onthly **I**ndex of **M**edical **S**pecialities – free monthly information on prescription medicines sent to registered general practitioners in the UK since 1959.

MLR – **M**ultiple **L**inear **R**egression.

MOBITEL – **MOBI**le **TEL**emonitoring in heart failure patients.

NYHA – **N**ew **Y**ork **H**eart **A**ssociation – Defined heart failure classification scheme ranging from Class I to Class IV, with respect to symptom severity (NYHA, 1994).

Origin – Defines the coordinate frame centre.

PCA – **P**rincipal **C**omponent **A**nalysis.

QQ plot – **Q**uantile-**Q**uantile plot – graphical method for comparing probability distributions.

R – programming language and environment for statistical data analyses.

ROC – **R**eceiver **O**perating **C**haracteristic – curve used to illustrate performance of binary classifiers.

SCS – **S**pinal **C**ord **S**timulation.

sms – **s**hort **m**essage **s**ervice.

TIA – **T**ransient **I**schemic **A**ttacks – or ”little strokes”.

TENS – **T**ranscutaneous **E**lectrical **N**erve **S**timulation.

TN – **T**rue **N**egative events – cases with no alarms accompanied by one of the following physician actions: threshold adjustment or no action.

TP – **T**rue **P**ositive events – alarms accompanied by one of the following physician actions: patient contact, medication adjustment or other action (excluding threshold adjustment).

VIF – **V**ariance **I**nflation **F**actor.

WHF – **W**orld **H**eart **F**ederation.

WHO – **W**orld **H**ealth **O**rganization.

WMO – **W**orld **M**eteorological **O**rganization.

ZAMG – **Z**entral**A**nstalt für **M**eteorologie und **G**eodynamik – Austrian Central Institute for Meteorology and Geodynamics.

Bibliography

ACCE (2006). *Impact of clinical alarms on patient safety.* ACCE Healthcare Technology Foundation.

AGES (2013). PharmaWeb – Austrian medicinal product index. `https://aspregister.basg.gv.at/aspregister/`. Agentur für Gesundheit und Ernährungssicherheit (Austrian Agency for Health and Food Safety). Updated Oct 05, 2013.

AHA (2012). Understanding blood pressure readings. `http://www.heart.org`. Amearican Heart Association. Retrieved Dec 21, 2012.

AHA (2013a). Classes of heart failure. `http://www.heart.org`. Amearican Heart Association. Retrieved Sep 29, 2013.

AHA (2013b). Implantable cardioverter defibrilator. `http://www.heart.org`. Amearican Heart Association. Retrieved Oct 14, 2013.

Ahdesmaki, M., Fokianos, K., and Strimmer, K. (2012). Identification of periodically expressed genes, GeneCycle package. Version: 1.1.2 [2012-Apr-15].

Atzema, C., Schull, M. J., Borgundvaag, B., Slaughter, G. R., and Lee, C. K. (2006). ALARMED: adverse events in low-risk patients with chest pain receiving continuous electrocardiographic monitoring in the emergency department. a pilot study. *The American Journal of Emergency Medicine*, 24(1):62–67.

Augustinsson, L. E., Carlsson, C. A., Holm, J., and Jivegård, L. (1985). Epidural electrical stimulation in severe limb ischemia. pain relief, increased blood flow, and a possible limb-saving effect. *Annals of surgery*, 202(1):104–110.

Barlow, J., Singh, D., Bayer, S., and Curry, R. (2007). A systematic review of the benefits of home telecare for frail elderly people and those with long-term conditions. *Journal of Telemedicine and Telecare*, 13(4):172–179.

Basilakis, J., Lovell, N. H., Redmond, S. J., and Celler, B. G. (2010). Design of a decision-support architecture for management of remotely monitored patients. *IEEE Transactions on Information Technology in Biomedicine*, 14(5):1216–1226.

Biddiss, E., Brownsell, S., and Hawley, M. S. (2009). Predicting need for intervention in individuals with congestive heart failure using a home-based telecare system. *Journal of Telemedicine and Telecare*, 15(5):226–231.

Boulay, F., Berthier, F., Sisteron, O., Gendreike, Y., and Gibelin, P. (1999). Seasonal variation in chronic heart failure hospitalizations and mortality in France. *Circulation*, 100(3):280–286.

Bradford, W. D., Kleit, A. N., Krousel-Wood, M., and Re, R. M. (2004). Willingness to pay for telemedicine assessed by the double-bounded dichotomous choice method. *Journal of telemedicine and telecare*, 10(6):325–330.

Bui, A. L. and Fonarow, G. C. (2012). Home monitoring for heart failure management. *Journal of the American College of Cardiology*, 59(2):97–104.

Casillas, J.-M., Gremeaux, V., Labrunee, M., Troigros, O., Laurent, Y., Deley, G., and Eicher, J.-C. (2008). Low-frequency electromyostimulation and chronic heart failure. In *Annales de réadaptation et de médecine physique*, volume 51, pages 461–472. Elsevier.

Chambrin, M.-C. (2001). Alarms in the intensive care unit: how can the number of false alarms be reduced? *Critical Care*, 5(4):184–188.

Chaudhry, S. I., Wang, Y., Concato, J., Gill, T. M., and Krumholz, H. M. (2007). Patterns of weight change preceding hospitalization for heart failure. *Circulation*, 116(14):1549–1554.

Cleland, J. G. F., Louis, A. A., Rigby, A. S., Janssens, U., and Balk, A. H. M. M. (2005). Noninvasive home telemonitoring for patients with heart failure at high risk of recurrent admission and death: The Trans-European Network-Home-care Management System (TEN-HMS) study. *Journal of the American College of Cardiology*, 45(10):1654–1664.

Cowpertwait, P. S. P. and Metcalfe, A. V. (2009). *Introductory time series with R*. Springer.

Danet, S., Richard, F., Montaye, M., Beauchant, S., Lemaire, B., Graux, C., Cottel, D., Marécaux, N., and Amouyel, P. (1999). Unhealthy effects of atmospheric temperature and pressure on the occurrence of myocardial infarction and coronary deaths – a 10-year survey: The Lille-World Health Organization MONICA project (monitoring trends and determinants in cardiovascular disease). *Circulation*, 100(1):e1–e7.

De Jongste, M. J., Haaksma, J., Hautvast, R. W., Hillege, H. L., Meyler, P. W., Staal, M. J., Sanderson, J. E., and Lie, K. I. (1994). Effects of spinal cord stimulation on myocardial ischaemia during daily life in patients with severe coronary artery disease. a prospective ambulatory electrocardiographic study. *British heart journal*, 71(5):413–418.

Deshpande, A. (2008). *Real-time (synchronous) telehealth in primary care: systematic review of systematic reviews*. Canadian Agency for Drugs and Technologies in Health.

Dethlefsen, C., Lundbye-Christensen, S., and Christensen, A. (2009). State space models in R, sspir package.

Domingo, M., Lupón, J., González, B., Crespo, E., López, R., Ramos, A., Urrutia, A., Pera, G., Verdú, J. M., and Bayes-Genis, A. (2011). Noninvasive remote telemonitoring for ambulatory patients with heart failure: effect on number of hospitalizations, days in hospital, and quality of life. carme (catalan remote management evaluation) study. *Revista Española de Cardiología (English Edition)*, 64(4):277–285.

Domingo, M., Lupón, J., González, B., Crespo, E., López, R., Ramos, A., Urrutia, A., Pera, G., Verdú, J. M., and Bayes-Genis, A. (2012). Evaluation of a telemedicine system for heart failure patients: feasibility, acceptance rate, satisfaction and changes in patient behavior: results from the carme (catalan remote management evaluation) study. *European Journal of Cardiovascular Nursing*, 11(4):410–418.

Drobics, M. (2010). Personal communication.

ECRI (1974). Hazard: Gaymar hypothermia machine. *Health Devices*, 3(9):229–230.

ECRI (2009). 2010 Top 10 technology hazards. *Health Devices*, 38(11):1–10.

ECRI (2010). Top 10 health technology hazards for 2011. *Health Devices*, 39:404–416.

ECRI (2011). Top 10 technology hazards for 2012. The risks that should be at the top of your prevention list. *Health Devices*, 40(11):358–373.

ECRI (2012). Top 10 health technology hazards for 2013. *Health Devices*, 41(11):1–23.

ECRI (2013). Top 10 health technology hazards for 2014. *Health Devices*, 42(11):1–12.

Ekeland, A. G., Bowes, A., and Flottorp, S. (2010). Effectiveness of telemedicine: a systematic review of reviews. *International Journal of Medical Informatics*, 79(11):736–771.

Eriksson, M. B. E., Sjölund, B. H., and Nielzén, S. (1979). Long term results of peripheral conditioning stimulation as an analgesic measure in chronic pain. *Pain*, 6(3):335–347.

Evangelista, L. S., Dracup, K., and Doering, L. V. (2000). Treatment-seeking delays in heart failure patients. *The Journal of heart and lung transplantation*, 19(10):932–938.

Fanger, P. O. (1970). *Thermal comfort. Analysis and applications in environmental engineering*. Copenhagen: Danish Technical Press.

Ferenc, J. (2012). Medical alarm hazards signal call for action. *Health Facilities Management*, 25(4):4.

Ferrante, D., Varini, S., Macchia, A., Soifer, S., Badra, R., Nul, D., Grancelli, H., and Doval, H. (2010). Long-term results after a telephone intervention in chronic heart failure DIAL (randomized trial of phone intervention in chronic heart failure) follow-up. *Journal of the American College of Cardiology*, 56(5):372–378.

Finkelstein, J. and Hripcsak, G. (2001). System and method for remotely monitoring asthma severity. US Patent 6,283,923.

Finkelstein, J., Wood, J., Cha, E., Orlov, A., and Dennison, C. (2010). Feasibility of congestive heart failure telemanagement using a Wii-based telecare platform. In *Engineering in Medicine and Biology Society (EMBS), 2010 Annual International Conference of the IEEE*, pages 2211–2214. IEEE.

Fox, J. (2005). Getting started with the R commander: A basic-statistics graphical user interface to R. *Journal of Statistical Software*, 14(9):1–42.

Fox, J. (2012). Polychoric and polyserial correlations, polycor package. Version: 0.7-8 [2010-Mar-25].

Gaffin, S. L. and Moran, D. S. (2001). Heat-related illnesses. *Wilderness medicine management of wilderness and environmental emergencies*, pages 240–316.

Giordano, A., Scalvini, S., Zanelli, E., Corr, U., Ricci, V., Baiardi, P., and Glisenti, F. (2009). Multicenter randomised trial on home-based telemanagement to prevent hospital readmission of patients with chronic heart failure. *International Journal of Cardiology*, 131(2):192–199.

Gonçalves, F. L. T., Braun, S., Silva Dias, P. L., and Sharovsky, R. (2007). Influences of the weather and air pollutants on cardiovascular disease in the metropolitan area of Sao Paulo. *Environmental Research*, 104(2):275–281.

Harrell, F. E. (2012). Harrell miscellaneous, Hmisc package. Version: 3.9-3 [2012-Mar-27].

Harvey, A. C. (1990). *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press.

Hayn, D., Koller, S., Hofmann-Wellenhof, R., Salmhofer, W., Kastner, P., and Schreier, G. (2009). Mobile phone-based teledermatologic compliance management – preliminary results of the TELECOMP study. *Medical Informatics in a United and Healthy Europe: Proceedings of MIE 2009*, 150:468–472.

Healy, J. D. (2003). Excess winter mortality in Europe: a cross country analysis identifying key risk factors. *Journal of Epidemiology and Community Health*, 57(10):784–789.

Helbig, M., Theus, M., and Urbanek, S. (2005). JGR: JAVA GUI for R. *Journal of Statistical Computing and Statistical Graphics*, 16(2):9–11.

Imhoff, M. and Kuhls, S. (2006). Alarm algorithms in critical care monitoring. *Anesthesia & Analgesia*, 102(5):1525–1537.

Jones, W. H. S. (1923). *Translation of Hippocrates' Air, Waters and Places.* London: Heineman.

Keatinge, W. R., Coleshaw, S. R., Cotter, F., Mattock, M., Murphy, M., and Chelliah, R. (1984). Increases in platelet and red cell counts, blood viscosity, and arterial pressure during mild surface cooling: factors in mortality from coronary and cerebral thrombosis in winter. *British medical journal (Clinical research ed.)*, 289(6456):1405–1408.

Keim, S. M., Guisto, J. A., and Sullivan Jr, J. B. (2002). Environmental thermal stress. *Annals of Agricultural and Environmental Medicine*, 9(1):1–15.

Koch, S. (2006). Home telehealthcurrent state and future trends. *International Journal of Medical Informatics*, 75(8):565–576.

Kollmann, A., Riedl, M., Kastner, P., Schreier, G., and Ludvik, B. (2007). Feasibility of a mobile phone-based data service for functional insulin treatment of type 1 diabetes mellitus patients. *Journal of Medical Internet Research*, 9(5):e36.

Kusch, W. (2004). *Guidelines on biometeorology and air quality forecasts*. WMO.

Kysely, J., Pokorna, L., Kyncl, J., and Kriz, B. (2009). Excess cardiovascular mortality associated with cold spells in the Czech Republic. *BMC public health*, 9(1):19.

Larose, D. T. (2006). *Data mining methods & models*. John Wiley and Sons, Inc.

Lees, J. M. (2012). Seismic time series analysis tools, RSEIS package. Version: 3.0-9 [2012-Jul-24].

Lewin, J., Ledwidge, M., O'Loughlin, C., McNally, C., and McDonald, K. (2005). Clinical deterioration in established heart failure: what is the value of bnp and weight gain in aiding diagnosis? *European journal of heart failure*, 7(6):953–957.

Lien, C. T. C., Gillespie, N. D., Struthers, A. D., and McMurdo, M. E. T. (2002). Heart failure in frail elderly patients: diagnostic difficulties, co-morbidities, polypharmacy and treatment dilemmas. *European journal of heart failure*, 4(1):91–98.

Lohninger, H. (2012). *Fundamentals of statistics*. Epina, Pressbaum.

Lohninger, H. J. (2013). Datalab. `http://www.lohninger.com/datalab/en_home.html`. Version: 3.511 [2013-Mar-16].

Mannheimer, C., Carlsson, C., Emanuelsson, H., Vedin, A., Waagstein, F., and Wilhelmsson, C. (1985). The effects of transcutaneous electrical nerve stimulation in patients with severe angina pectoris. *Circulation*, 71(2):308–316.

Maric, B., Kaan, A., Ignaszewski, A., and Lear, S. A. (2009). A systematic review of telemonitoring technologies in heart failure. *European Journal of Heart Failure*, 11(5):506–517.

MayoClinic (2011a). Blood pressure chart: What your reading means. `http://www.MayoClinic.com`. Retrieved Dec 21, 2012.

MayoClinic (2011b). Low blood pressure (hypotension): Causes. `http://www.MayoClinic.com`. Retrieved Sep 14, 2013.

MIMS (2013). MIMS online – Prescription drug database and drug prescribing guide. `http://www.mims.co.uk/`. Retrieved Oct 03, 2013.

Morabito, M., Modesti, P. A., Cecchi, L., Crisci, A., Orlandini, S., Maracchi, G., and Gensini, G. F. (2005). Relationships between weather and myocardial infarction: a biometeorological approach. *International Journal of Cardiology*, 105(3):288–293.

Moser, D. K., Doering, L. V., and Chung, M. L. (2005). Vulnerabilities of patients recovering from an exacerbation of chronic heart failure. *American heart journal*, 150(5):984–e7.

Näyhä, S. (2005). Environmental temperature and mortality. *International Journal of Circumpolar Health*, 64(5):451–458.

NIH (2013). What is an implantable cardioverter defibrillator? `http://www.nhlbi.nih.gov`. National Institutes of Health. National Heart, Lung and Blood Institue. Retrieved Oct 14, 2013.

NYHA (1994). *Nomenclature and criteria for diagnosis of diseases of the heart and great vessels*. Boston: Little, Brown & Co., 9. edition. 253–256.

Obel, I. W. P. and Bourgeois, I. (1993). Implantable electrical nerve stimulator/pacemaker with ischemia detector for decreasing cardiac workload. US Patent 5,199,428.

Ono, H.-S. P. and Kawamura, T. (1991). Sensible climates in monsoon asia. *International journal of biometeorology*, 35(1):39–47.

Paré, G., Jaana, M., and Sicotte, C. (2007). Systematic review of home telemonitoring for chronic diseases: the evidence base. *Journal of the American Medical Informatics Association*, 14(3):269–277.

Polisena, J., Coyle, D., Coyle, K., and McGill, S. (2009). Home telehealth for chronic disease management: a systematic review and an analysis of economic evaluations. *International Journal of Technology Assessment in Health Care*, 25(03):339–349.

Poumadere, M., Mays, C., Le Mer, S., and Blong, R. (2005). The 2003 heat wave in France: dangerous climate change here and now. *Risk Analysis*, 25(6):1483–1494.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes-The Art of Scientific Computing*. Cambridge University Press, 3. edition.

Ridderikhoff, J. and van Herk, B. (1999). Who is afraid of the system? Doctors' attitude towards diagnostic systems. *International Journal of Medical Informatics*, 53(1):91–100.

Riegel, B., Moser, D. K., Anker, S. D., Appel, L. J., Dunbar, S. B., Grady, K. L., Gurvitz, M. Z., Havranek, E. P., Lee, C. S., Lindenfeld, J., et al. (2009). State of the science promoting self-care in persons with heart failure: a scientific statement from the american heart association. *Circulation*, 120(12):1141–1163.

Rojas, S. V. and Gagnon, M.-P. (2008). A systematic review of the key indicators for assessing telehomecare cost-effectiveness. *Telemedicine and e-Health*, 14(9):896–904.

Scherr, D., Kastner, P., Kollmann, A., Hallas, A., Auer, J., Krappinger, H., Schuchlenz, H., Stark, G., Grander, W., Jakl, G., Schreier, G., Friedrich, F. M., et al. (2009). Effect of home-based telemonitoring using mobile phone technology on the outcome of heart failure patients after an episode of acute decompensation: randomized controlled trial. *Journal of Medical Internet Research*, 11(3):e34.

Seto, E. (2008). Cost comparison between telemonitoring and usual care of heart failure: a systematic review. *Telemedicine and e-Health*, 14(7):679–686.

Smith, S. W. (1999). *The Scientist and Engineer's Guide to Digital Signal Processing*. California Technical Publishing, 2. edition.

Stewart, S., McIntyre, K., Capewell, S., and McMurray, J. J. (2002). Heart failure in a cold climateseasonal variation in heart failure-related morbidity and mortality. *Journal of the American College of Cardiology*, 39(5):760–766.

Tallis, R. C., Illis, L. S., Sedgwick, E. M., Hardwidge, C., and Garfield, J. S. (1983). Spinal cord stimulation in peripheral vascular disease. *Journal of Neurology, Neurosurgery & Psychiatry*, 46(6):478–484.

Tendera, M. (2005). Epidemiology, treatment, and guidelines for the treatment of heart failure in Europe. *European Heart Journal Supplements*, 7(suppl J):J5–J9.

Tran, K., Polisena, J., Coyle, D., Coyle, K., Kluge, E., Cimon, K., McGill, S., Noorani, H., Palmer, K., and Scott, R. (2008). *Home telehealth for chronic disease management*. Canadian Agency for Drugs and Technologies in Health Ottawa.

Vries, A. E. d., Wal, M. H. L., Bedijn, W., Jong, R. M. d., Kraai, I. H., Dijk, R. B. v., Jaarsma, T., and Hillege, H. L. (2011). Follow-up and treatment of an instable patient with heart failure using telemonitoring and an ICT-guided disease management system; a case study. *International Journal of Integrated Care*, 11(6).

Vukovic, M., Drobics, M., Hayn, D., Kastner, P., and Schreier, G. (2010). Optimization of the alarm-management of a heart failure home-monitoring system. In *Computing in Cardiology 2010*, pages 53–56.

Vukovic, M., Drobics, M., Hayn, D., Kreiner, K., and Schreier, G. (2012a). Automated alarm management system for home telemonitoring of chronic heart failure patients. In *10th Int Conf on Inf Communication Tech in Health*.

Vukovic, M., Drobics, M., Hayn, D., Schreier, G., Lohninger, H., and Rattay, F. (2012b). Weather influence on alarm occurrence in home telemonitoring of heart failure patients. In *Computing in Cardiology 2012*, pages 525–528.

Vukovic, M., Drobics, M., Kreiner, K., Hayn, D., and Schreier, G. (2012c). Alarm management in patient health status monitoring. In *Proceedings of the eHealth 2012*, pages 39–44.

Warnes, G. R. (2012). Gregs miscellaneous functions, gregmisc package. Version: 2.1.2 [2011-Sep-02].

Weintraub, A., Gregory, D., Patel, A. R., Levine, D., Venesy, D., Perry, K., Delano, C., and Konstam, M. A. (2010). A multicenter randomized controlled evaluation of automated home monitoring and telephonic disease management in patients recently hospitalized for congestive heart failure: the SPAN-CHF II trial. *Journal of Cardiac Failure*, 16(4):285–292.

WHF (2013). Different heart diseases. `http://http://www.world-heart-federation.org`. World Heart Federation. Retrieved Sep 29, 2013.

WHO (2005). *Preventing Chronic Diseases – A Vital Investment: WHO Global Report*. World Health Organization.

Wu, J., Kraja, A. T., Oberman, A., Lewis, C. E., Ellison, R. C., Arnett, D. K., Heiss, G., Lalouel, J.-M., Turner, S. T., Hunt, S. C., Province, M. A., and Rao, D. C. (2005). A summary of the effects of antihypertensive medications on measured blood pressure. *American Journal of Hypertension*, 18(7):935–942.

Yuan, L., Stockton, T., Bronson, D., Minallah, P., and Fitzgerald, M. (2012). GUI to several statistical methods useful for causal assessment, CADStat package. Version: 2.2-6 [2010-Aug-27].

ZAMG (2012). Zentralanstalt für Meteorologie und Geodynamik Jahrbuch (Austrian Central Institute for Meteorology and Geodynamics yearbook). `http://www.zamg.ac.at/cms/de/klima/klimauebersichten/jahrbuch`. Retrieved Aug 09, 2012.

Zhang, Y. (2003). *Real-time analysis of physiological data and development of alarm algorithms for patient monitoring in the intensive care unit*. PhD thesis, Massachusetts Institute of Technology.

Curriculum Vitae

Personal Information

First name / Surname:	Marija / Vukovic (maiden Filipovic)
Academic degree:	Dipl.-Math.
Nationality:	Serbian
Date Of Birth:	June 14, 1978.
Gender:	Female
Address:	Viktor Kaplan Strasse 6–8/209, 1220 Vienna, Austria
Languages:	Serbian (mother tongue), English, German

Education

2011-2014:	Doctorate (PhD) studies of Mathematics at Vienna University of Technology, Austria Specialisation: Analysis and Scientific Computing
1997-2003:	Graduate Mathematician Degree, (equivalent to Masters degree) at Department of Mathematics, University of Belgrade, Serbia Major: Mathematics and Computer Science Education
1993-1997:	Graduate High School Diploma, at Gymnasium ”Bora Stankovic”, Bor, Serbia
1985-1993:	Primary School Diploma, at ”3. oktobar”, Bor, Serbia

Work Experience

2010-2013:	Junior Scientist at the Austrian Institute of Technology, Vienna, Austria
2009:	Academic Coach at the Tutorworks, San Francisco, CA, USA
2008, 2006:	Visiting Researcher at Psychophysioligy of Movement Laboratory, Department of Kinesiology, The Pennsylvania State University, University Park, PA, USA
2008:	Instructor of Mathematics at Kumon Math and Reading School, State College, PA, USA
2004-2005:	High School Teacher of Mathematics and Computer Science at "9. Beogradska Gimnazija", Belgrade, Serbia
2003-2005:	Teaching Assistant in Mathematics and Computer Science, Cathedra for Mathematics and Physics, Department of Agriculture, University of Belgrade, Serbia
2003:	High School Teacher of Computer Science at "Klasicna Gimnazija", Belgrade, Serbia

Project Work

2012-2013:	New Generation Decision Support System
2010-2011:	SIOPEN-R-NET International Society of Paediatric Oncology European Neuroblastoma Research Network

Publications

Vukovic M., Drobics M., Hayn D., Schreier G., Lohninger H. and Rattay F. (2012). Weather influence on alarm occurrence in home telemonitoring of heart failure patients. In *Proceedings of the Computing in Cardiology Conference*, pages 525–528. Krakow. Poland.

Vukovic M., Drobics M., Hayn D., Kreiner K. and Schreier G. (2012). Automated alarm management system for home telemonitoring of heart failure patients. In: *Prachalias V., editor. Proceedings of the 10th International Conference on Information Communication Technologies in Health.* Samos. Greece.

Vukovic M., Drobics M., Kreiner K., Hayn D. and Schreier, G. (2012). Alarm management in patient health status monitoring. In: *Proceedings of e-Health Conference*, ID: 45. Vienna. Austria.

Vukovic M., Drobics M., Hayn D., Kastner P. and Schreier G. (2010). Optimization of the alarm management of a heart failure home-monitoring system. In *Proceedings*

of the Computing in Cardiology Conference, pages 53–56. Belfast. UK.

Filipovic M., Vukovic V. and Jovanovic A. (2004). Correctness of inference proofs. In *Proceedings of Electronics, Telecommunications, Computers, Automatic Control and Nuclear Engineering (ETRAN) International Conference*. Cacak. Serbia.

Perovic A., Jovanovic A. and **Filipovic M.** (2004). One implementation of automata with certain cognitive functions. In *Proceedings of Electronics, Telecommunications, Computers, Automatic Control and Nuclear Engineering (ETRAN) International Conference*. Cacak. Serbia.