968-

Bepalen van mentale werkbelasting door cardiovasculaire data-analyse

Arjan Stuiver
0970271

29 september 2003

Begeleiders:
dr.ir. L.J.M. Mulder, Psychologie, RuG
dr. F. Cnossen, KI, RuG

Technische Cognitie Wetenschap
Rijksuniversiteit Groningen

Samenvatting

Het vaststellen van de fysiologische effecten van mentale inspanning is één van de eerste stappen die gezet moeten worden bij het ontwikkelen van een adaptieve interface. Het doel hiervan is de taakzwaarte af te stemmen op de capaciteiten van de taakuitvoerder. Bij zware mentale inspanning kan een adaptieve interface reageren door de taak anders vorm te geven of gedeelten van het werk voor een ander moment te bewaren of semi-automatisch te laten uitvoeren.

Een schatting van de mate mentale inspanning kan als input dienen voor een model dat de toestand van de taakuitvoerder bijhoudt, zoals het *operator status model* (OSM) uit het COMPANION project. COMPANION is een samenwerkingsproject tussen de NLR en de RuG met als doel het ontwikkelen van een dynamisch en interactief interfacesysteem.

In dit onderzoek is geprobeerd cardiovasculaire data te classificeren. Tijdens een experiment, dat is uitgevoerd in de Digitale Werkplaats (DWAT) van de afdeling psychologie van de Rijksuniversiteit te Groningen, zijn hartslag en bloeddruksignalen opgenomen. Tijdens het experiment waren er perioden met een lage taaklast en perioden met een hoge taaklast. Het verschil in taaklast moest zorgen voor een verschil in mentale inspanning. De fysiologische respons op de taak, die zou moeten veranderen als gevolg van een verhoogde mentale inspanning, is geanalyseerd om te bepalen of er een hoge of een lage taaklast was. Met multiple regressieanalyse is een model ontwikkeld dat voor elke minuut data aangeeft uit welke periode van het experiment deze data afkomstig is.

Het onderzoek laat zien dat het tot op zekere hoogte mogelijk is de cardiovasculaire data tot de juiste periode te classificeren. Niet alle data kan juist geclassificeerd worden, een gedeelte van de variantie moet verklaard worden door andere oorzaken dan een verschil in taaklast. Hierbij kan gedacht worden aan een zekere mate van stochasticiteit die inherent is aan het fysiologisch systeem of aan een beperkte invloed van taaklast op mentale inspanning.

Het onderzoek laat ook zien dat de classificatie beter is wanneer de data van de proefpersonen individueel geanalyseerd worden. De patronen in de data blijken erg individuspecifiek en zodoende presteert het model op de individuele data significant beter dan op de data van alle proefpersonen samen. Enerzijds wordt dit veroorzaakt door verschillende individuele responspatronen en anderzijds door een verschil belastbaarheid, wat wil zeggen dat de ene persoon bij een lage taaklast al sterk reageert terwijl de andere persoon nog nauwelijks een respons laat zien.

Grafische vormgeving omslag: J.T. van der Beek, 2003

Voorwoord

Tijdens mijn studie is mijn interesse altijd uitgegaan naar het intelligente gedrag van machines en interfaces. Voor mijn afstudeeronderzoek heb ik daarom gekozen voor een onderwerp dat verband houdt met het ontwerpen van intelligente interfaces. Bij het lezen van deze scriptie zal opvallen dat de afstand tussen de beoogde interface en mijn onderzoek erg groot lijkt. Dit mag dan misschien zo zijn, maar er is natuurlijk geen betere plek om te beginnen dan aan de basis.

Ik wil hier van de gelegenheid gebruik maken om een aantal mensen te bedanken die me tijdens mijn studie en afstuderen op wat voor manier dan ook hebben geholpen. Allereerst met betrekking tot het afstuderen zelf natuurlijk Ben Mulder, de regelmatige gesprekken hebben me niet alleen een hoop inzicht verschaft, ze waren ook altijd weer een motivatie om verder te gaan. Daarnaast Fokie Cnossen, enerzijds omdat ze door haar eigen ervaringen blijkbaar precies aan kon geven dat het omgaan met de frustraties die sommige taken geven juist het belangrijkste is dat je moet leren. Anderzijds natuurlijk door het verlagen van de bloeddruk en verhogen van de variabiliteit in de hartslag door het over andere zaken te hebben dan mijn afstuderen. Als laatste, voor het richting geven aan mijn onderzoek waar dit soms een duistere weg in dreigde te slaan, wil ik Lambert Schomaker bedanken.

De studie Technische Cognitie Wetenschap voorzag mij gedurende mijn afstuderen van een werkplek, waar ik erg prettig heb kunnen werken. Ondanks dat de interface zich niet autonoom-adaptief opstelde, kon ik er prima werken, waarschijnlijk mede dankzij Ronald Zwaagstra, die er toch een adaptief karakter aan wist te geven.

Zonder de goede sfeer in de afstudeerkamer was het natuurlijk allemaal nooit gelukt. Dat was de schuld van een grote groep studiegenoten: Albert (budt), Hendria, Jeewee, Karin, Marcia, Peter, Rick, Reindert-Jan en Willem. En Diederik natuurlijk, die altijd speciaal voor de pauzes langs moest komen.

De reden dat ik heb kunnen afstuderen is dat ik ooit begonnen ben met studeren. Mijn ouders hebben ervoor gezorgd dat ik dat kon doen, daar ben ik natuurlijk erg dankbaar voor. Zij zullen ook blij als geen ander zijn dat ik ben afgestudeerd.

Als laatste wil ik Jiska bedanken, zowel voor het tonen van interesse en bieden van steun tijdens mijn afstuderen alsook het tijdig stoppen met vragen: *hoe ging het afstuderen vandaag?*. Daarnaast voor het nalezen van de scriptie en het ontwerpen van de omslag.

Inhoudsopgave

1 Introductie 7
- 1.1 Onderzoeksvraag . . . 8
- 1.2 Opbouw . . . 8

2 Cardiovasculaire signalen 10
- 2.1 Fysiologie . . . 10
- 2.2 Cardiovasculaire tijdreeksen . . . 12
  - 2.2.1 Hartfrequentie en interbeat interval . . . 12
  - 2.2.2 Bloeddruk . . . 13
  - 2.2.3 Samplen en Signaaleigenschappen . . . 13
  - 2.2.4 Artefacten . . . 14
  - 2.2.5 Afgeleide variabelen . . . 14
- 2.3 Spectraalanalyse . . . 15
  - 2.3.1 Frequentiebanden . . . 15
  - 2.3.2 Profielen . . . 16
  - 2.3.3 Afleiding van variabelen . . . 16

3 Multivariate data analyse 21
- 3.1 Overzicht . . . 21
  - 3.1.1 Input-output model . . . 22
- 3.2 Multiple Regressie . . . 23
  - 3.2.1 Residuals en R-kwadraat . . . 25
  - 3.2.2 Aannames van lineariteit en normaliteit . . . 26
  - 3.2.3 Causaliteit . . . 27
  - 3.2.4 Multicollineariteit . . . 28
  - 3.2.5 Voorspellen met behulp van de regressievergelijking . . . 28

4 De Analyse 32
- 4.1 Basis . . . 32
- 4.2 Overgangswaarden . . . 33
- 4.3 Dubbele overgangswaarden . . . 33
- 4.4 Highpass filteren . . . 34
- 4.5 Lopende variantie . . . 37
- 4.6 Combinatie van methoden . . . 38

5 Resultaten 40
5.1 Uitleg resultaten 40
5.2 Resultaten basismethode 42
5.3 Resultaten overgangsmethoden 42
5.4 Resultaten highpass en lopende variantie 43
5.5 Vergelijking verschillende methoden 43
5.6 Voting Neighbors 44
5.7 Individueel versus Samen 45

6 Discussie 47
6.1 Algemeen resultaat 47
6.2 Resultaten aanpassingen 48
6.3 Voting neighbors 49
6.4 Individueel versus samen 49
6.5 Individuele verschillen 51

7 Conclusie en Aanbevelingen 53
7.1 Conclusie 53
7.2 Individu-afhankelijke respons 53
7.3 Neurale netwerken 53
7.4 Online herkennen 54

Hoofdstuk 1

Introductie

Door technologische ontwikkelingen is er niet alleen de mogelijkheid ontstaan om werkomgevingen te creëren die zich aanpassen aan de gebruiker, maar ook is de behoefte aan dergelijke werkomgevingen groter geworden. Langdurige hoge werkdruk blijkt namelijk een ongunstig effect te hebben op de gezondheid [Picard and Klein, 2001]. Vanuit deze achtergrond lijkt het wenselijk de werkdruk te kunnen bepalen. De analyse van fysiologische data kan bijdragen aan het onderzoek naar psychologische fenomenen, zoals bijvoorbeeld mentale inspanning, stellen Cacioppo en Tassinary in *Inferring Psychological Significance from Physiological Signals* [Cacioppo and Tassinary, 1990]. Zo wordt al meerdere decennia onderzocht hoe geschikt cardiovasculaire- en ademhalingsmaten zijn als index voor mentale inspanning.

Voor het ontwikkelen van een adaptieve werkomgeving zijn een aantal ingredienten nodig. Eén ingredient is een model waarin de toestand van de gebruiker bijgehouden wordt, zoals het operator status model(OSM) uit het COMPANION-project. COMPANION is een samenwerkingsproject tussen de NLR en de RuG met als doel het ontwikkelen van een dynamisch en interactief interfacesysteem. Een model zoals het OSM zou gebruik kunnen maken van een schatting van mentale inspanning op moment van uitvoeren van een taak. De eerste stap naar het ontwikkelen van een systeem dat online het niveau van mentale inspanning kan bepalen, is een systeem dat het niveau offline kan bepalen. Met offline wordt hier bedoeld dat de data al geregistreerd is en na afloop van de taak geanalyseerd wordt.

In dit onderzoek is gekeken of op basis van cardiovasculaire tijdreeksen een model kan worden ontwikkeld om twee perioden in een taak van elkaar te onderscheiden. De twee perioden onderscheidden zich door een verschil in taakzwaarte. Er wordt aangenomen dat deze verschillen tot uiting komen in de mate van mentale inspanning. Bij het bestuderen van de relatie tussen de eisen die de taak stelt en de mentale inspanning, moet rekening gehouden worden met een aantal gecompliceerde factoren [Waard, 1996]. Bij een verhoging van de taakzwaarte kan een taakuitvoerder op verschillende manieren reageren. De mentale inspanning kan omhoog gaan wanneer een taakuitvoerder gemotiveerd is om harder te werken. Wanneer de taakuitvoerder niet gemotiveerd is of de taak wordt te zwaar of te moeilijk voor een taakuivoerder, kan de mentale inspanning gelijk blijven of zelfs lager worden bij een verhoogde taakzwaarte. Bij het onderzoek naar de invloed van taaklast op mentale inspanning

moet daarom altijd rekening gehouden worden met de mogelijke invloed van andere factoren.

Het experiment waarin de genoemde taak is onderzocht, is uitgevoerd door I. Venema [Venema, 2003] en A. Kruizenga [Kruizenga, 2003] en heeft plaatsgevonden in de digitale werkplaats DWAT van de afdeling psychologie van de Rijksunversiteit Groningen. De DWAT is een ruimte die is ingericht om zo goed mogelijk een ambulancemeldkamer te simuleren. De proefpersonen vervulden in vijf afzonderlijke sessies gedurende ongeveer twee uur de rol van ambulance meldkamercentralist. Een centralist moet ambulances wegsturen voor geplande ambulanceritten (het vervoeren van patienten zonder spoed) en ambulanceritten naar noodgevallen. Speciaal voor dit experiment was een scenario voor de centralisten ontwikkeld waarin alle ongevallen gepland waren. Daarnaast hadden de centralisten de beschikking over een schema waarin de tijdstippen van de geplande ritten vermeld stonden. Het scenario was zo ontworpen dat het aantal ambulances dat weggestuurd moest worden per periode verschilde. Dit leverde vijf gescheiden perioden op. In twee van deze perioden moesten meer ambulances weggestuurd worden dan in de drie andere perioden. De perioden waarin veel ambulances weggestuurd moesten worden hadden als zodanig een grotere taaklast dan de andere perioden.

Tijdens het uitvoeren van de taak van een centralist werden de hartslag en bloeddruksignalen van de proefpersonen gemeten. Deze hartslag en bloeddruksignalen zijn gebruikt om afgeleide maten te berekenen. Vervolgens is met behulp van multiple regressieanalyse een model ontwikkeld dat kan bepalen uit welke periode de cardiovasculaire data afkomstig is.

1.1 Onderzoeksvraag

De onderzoeksvraag waarop dit verslag een antwoord tracht te geven:

> **Hoe goed kan met behulp van multiple regressieanalyse van cardiovasculaire data bepaald worden of deze uit een periode met hoge of lage taakzwaarte afkomstig is?**

De aannames die hierbij gemaakt worden zijn dat het verschil in taakzwaarte zorgt voor een verschil in mentale inspanning met een bijbehorende fysiologische respons. Het niveau van mentale inspanning wordt verondersteld slecht twee discrete waarden aan te kunnen nemen, namelijk hoog en laag, afgeleid van de twee taaklastniveaus.

1.2 Opbouw

In hoofdstuk 2 wordt de achtergrond van cardiovasculaire signalen en de spectrale analyse van de afgeleide tijdreeksen beschreven. Hoofdstuk 3 beschrijft de multivariate data analyse technieken die gebruikt zijn in de analyse. Er is gekozen voor multiple regressieanalyse als multivariate techniek om een model te ontwikkelen dat

de perioden in de taak kan classificeren. In hoofdstuk 4 wordt de ontwikkelde analyse beschreven. Naast de toegepaste regressieanalyse worden de bewerkingen op de data beschreven die voor een verbetering in de classificatie moeten zorgen. De resultaten van de analyse staan in hoofdstuk 5. In de discussie, die te vinden is in hoofdstuk 6, wordt dieper in gegaan op de verschillen die gevonden worden in de resultaten tussen een benadering waar de data van een proefpersoon afzonderlijk wordt geanalyseerd en een benadering waarin de data van alle proefpersonen tegelijkertijd wordt geanalyseerd. Vervolgens wordt in het laatste hoofdstuk een evaluatie van het onderzoek gegeven en worden er suggesties gedaan voor verder onderzoek op dit gebied.

Hoofdstuk 2

Cardiovasculaire signalen

Onder cardiovasculaire signalen verstaan we hartslag en bloeddruksignalen. In dit gedeelte wordt de oorsprong van de signalen uitgelegd aan de hand van het fysiologische systeem dat eraan ten grondslag ligt. Allereerst wordt de fysiologie besproken om het geheel te kunnen plaatsen. Vervolgens wordt beschreven hoe we hartslag en bloeddruksignalen kunnen onderscheiden en meten. Tenslotte wordt uitgelegd hoe de cardiovasculaire activiteit bestudeerd kan worden door spectraalanalyse van de afgeleide cardiovasculaire tijdreeksen.

2.1 Fysiologie

Een beknopte uitleg over de werking van het onderliggende fysiologische systeem maakt het makkelijker cardiovasculaire signalen te interpreteren. Daarom volgt nu een korte beschrijving van de werking van het cardiovasculair systeem.

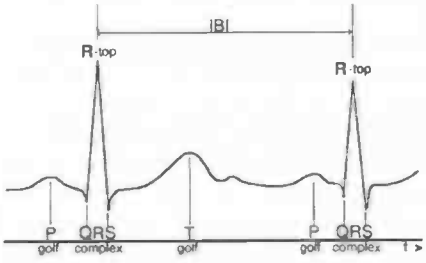
Bij de samentrekking van het hart worden twee toestanden onderscheiden, de systole en de diastole. De toestand waarin het hart het bloed wegpompt wordt systole genoemd; na de systole ontspant het hart zich, deze toestand wordt diastole genoemd.

De sinusknoop probeert samen met de cellen die om het hart voortdurend een samentrekking van het hart te veroorzaken. Samen zorgen ze ervoor dat er regelmatig sterke electrische pulsen aan de hartspier worden gegeven zodat deze samentrekt. De frequentie waarmee de sinusknoop het hart probeert te laten samentrekken wordt beïnvloed door het autonome zenuwstelsel. Het autonome zenuwstelsel kan opgedeeld worden in twee delen die hun eigen (tegengestelde) invloed hebben, namelijk het sympathische en het parasympathische deel van het zenuwstelsel. Het sympathische deel verhoogt de hartfrequentie en de kracht waarmee het hart samentrekt, terwijl het parasympathische deel de frequentie en de kracht juist verlaagt. Het parasympathische deel bestaat voornamelijk uit de *nervus vagus*; parasympathische invloed wordt daarom vaak aangeduid met vagale activatie.

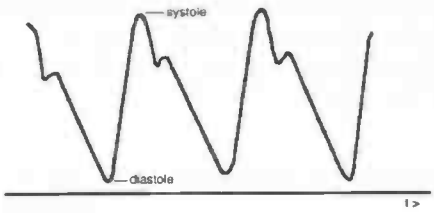
De electrische activiteit die ontstaat rondom het hart kan gemeten worden en geeft het electrocardiogram (ECG). In figuur 2.1 wordt een typisch ECG-signaal getoond. Hierin zijn de P-golf, het QRS-complex en de T-golf te onderscheiden. De P-golf valt samen met de spreiding van electrische lading over de atria en het begin

van de samentrekking daarvan. Het QRS-complex valt samen met de spreiding van electrische lading over de ventrikels en de samentrekking van de ventrikels. De T-golf valt samen met de herstel-periode van de ventrikels.

De bloeddruk wordt uiteindelijk bepaald door de frequentie waarmee het hart slaat, het *slagvolume* oftewel hoeveel bloed het hart in één slag naar buiten pompt en de weerstand van de bloedvaten. Uiteraard zijn er nog andere invloeden, zoals bijvoorbeeld de eigenschappen van vaten die vrij traag op veranderingen in het lichaam reageren, maar deze factoren zijn voor dit type psychologisch onderzoek niet direct interessant.



Figuur 2.1: Een voorbeeld van een ECG-signaal.



Figuur 2.2: Een voorbeeld van een bloeddruksignaal.

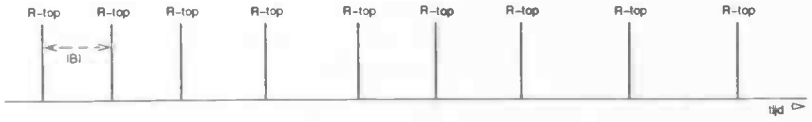
De frequentie waarmee het hart slaat, de bloeddruk die ontstaat door het rondpompen en de toestand in de bloedvaten vertonen veranderingen die interessant zijn voor psychologisch onderzoek. Onderzoek in de laatste decennia heeft laten zien dat de hartfrequentie vaak omhoog gaat door mentale inspanning [Mulder, 1988]. Van Roon et al. laten zien dat mentale inspanning in verschillende onderzoeken dezelfde effecten op de bloeddruk heeft [Roon et al., 1995]. Daarom worden hartfrequentie en bloeddruk veel gebruikt voor de zoektocht naar een fysiologische basis voor psychologische veranderingen. In de volgende sectie worden afgeleide cardiovasculaire tijdreeksen beschreven en uitgelegd.

2.2 Cardiovasculaire tijdreeksen

Tijdens de ambulancemeldkamertaak die de proefpersonen uitvoerden werden het ECG en de bloeddruk opgenomen. Uit deze twee signalen, die beide een functie van de tijd zijn, kunnen verschillende tijdreeksen bepaald worden, die we kunnen gebruiken voor het samenstellen van een complexe index voor mentale inspanning. De directe tijdreeksen die we uit het ECG kunnen halen zijn de *hartfrequentie* en het *interbeat interval*. De *gemiddelde, systolische* en *diastolische bloeddruk* zijn de directe tijdreeksen uit het bloeddruksignaal.

2.2.1 Hartfrequentie en interbeat interval

De frequentie van de hartslag, *hartfrequentie* of *heartrate* (HR) genoemd, kan worden berekend uit het *electrocardiogram* (ECG), zie 2.1. Het interessante onderdeel van het ECG voor het bepalen van de HR is het voorkomen van een R-top in een *QRS-complex*. Elk QRS-complex betekent een hartslag, waarbij Q R en S de verschillende onderdelen van het QRS-complex aanduiden. Eén hartslag kan gerepresenteerd worden door de top van het QRS-complex (R-top). Wanneer we de rest van het signaal weg zouden laten, ontstaat er een puls-trein van R-toppen die aangeven op welk moment er een hartslag was, zie figuur 2.3. De tijdstippen waarop de R-toppen voorkomen, kunnen door een analoge *R-top event detection* routine bepaald worden en vervolgens door een digitale methode die deze events leest naar een file geschreven worden. Dit levert een lijst op met alle R-top tijden die vervolgens kan worden gebruikt voor spectraalanalyse.



Figuur 2.3: R-top pulstrein die ontstaat na R-top event detection op het ECG.

Het aantal pulsen per seconde in deze pulstrein van R-toppen is de hartfrequentie (HR) uitgedrukt in Hz. In figuur 2.3 is ook de term *IBI* aangegeven, dit slaat op de intervaltijd tussen twee opeenvolgende slagen, het zogenaamde *interbeat interval* (IBI). Het IBI is de inverse van de HR en wordt uitgedrukt in *msec*. In principe zijn alle karakteristieke eigenschappen in het spectrum van de HR ook te vinden in het spectrum van het IBI. Beide representaties worden gebruikt voor spectraalanalyse van hartsignalen.

Opgemerkt dient te worden dat er desalniettemin geen lineair verband bestaat tussen HR en IBI. In het kort kan gezegd worden dat de variabiliteit van de hartfrequentie beter bepaald kan worden uit HR dan uit IBI [Rompelman, 1980]. De variabiliteitsmaten die verderop genoemd worden, zijn allemaal bepaald uit hartfrequentie. Een uitzondering is de afleiding van de modulatie index die bepaald wordt uit IBI.

2.2.2 Bloeddruk

Het bloeddruksignaal is net als het ECG een continu signaal waaruit we een tijdreeks willen halen die we kunnen gebruiken naast het HR en/of IBI, zie 2.2. De bloeddruk wordt bepaald voor ieder interval tussen twee R-toppen. Het gemiddelde in een bepaald interval geeft de *gemiddelde bloeddruk*. De *systolische bloeddruk* is de maximumwaarde van het interval. Het minimum, dat voor het maximum in een interval ligt, is de *diastolische bloeddruk*. Deze aanpak levert drie tijdreeksen op die voor elk RR-interval een waarde geven voor gemiddelde, systolische en diastolische bloeddruk in millimeters kwikdruk (mmHg).

2.2.3 Samplen en Signaaleigenschappen

Hartslag en bloeddruksignalen zijn continue signalen, dat betekent dat deze signalen eerst omgezet moeten worden naar hun digitale equivalenten voor er digitale technieken op kunnen worden toegepast.

Zoals beschreven in 2.2.1 worden de R-toppen door een hardware-matige event detectie uit het ECG gehaald. Dit levert eigenlijk direct een tijdreeks van de hartfrequentie op die digitaal geïnterpreteerd kan worden. Alle voorkomens van R-toppen staan vermeld in het signaal, alle andere punten (hoewel niet vermeld) zijn nul. Voor een spectraalanalyse, zoals de *discrete fourier transformatie* (zie sectie 2.3.3) is dit een goede structuur omdat deze methode niet beïnvloed wordt door waarden die nul zijn.

In 2.2.2 werd beschreven hoe de gemiddelde, systolische en diastolische bloeddruk bepaald worden voor de tijdstippen dat er een R-top was. Om precies te zijn, worden de drie bloeddruk-waarden die horen bij een bepaalde R-top, bepaald uit het interval dat voor die R-top ligt.

Niet-equidistante signalen

Hartfrequentie is een niet-equidistante tijdreeks, dat betekent dat de opeenvolgende sample-punten niet op een vaste afstand van elkaar liggen. Dit maakt het bijvoorbeeld onmogelijk om de fourier getransformeerde van de hartfrequentie met een *fast fourier transformatie* (FFT) te berekenen. Er zijn verschillende methoden om toch een fourier getransformeerde te berekenen. Eén manier is om het signaal zo aan te passen dat het wel equidistant is, bijvoorbeeld door het aan te vullen met nullen. Een andere methode is een discrete fourier transformatie (DFT) toe te passen. De discrete fourier transformatie heeft geen problemen met de niet-equidistantheid omdat alle waarden die nul zijn in een discrete fourier transformatie niet bijdragen aan de sommatie. We kunnen het toevoegen van nullen om een equidistante reeks te maken dus achterwege laten en de discrete fourier transformatie direct gebruiken op de lijst met R-top tijden.

2.2.4 Artefacten

Artefacten zijn extremiteiten in het signaal, dit kunnen enkele of meerdere opeenvolgende waarden in een signaal zijn. We kunnen twee soorten artefacten in het signaal onderscheiden: artefacten die ontstaan door gebeurtenissen tijdens het meten en artefacten die een fysiologische oorzaak hebben. De invloed die artefacten hebben op het spectrum is zo groot dat wanneer ze niet gecorrigeerd worden, de variabiliteit voornamelijk afhangt van deze artefacten en niet van het signaal zelf. Het eerste type artefacten moet daarom altijd gevonden en gecorrigeerd worden, extra R-toppen moeten worden verwijderd en missende R-toppen gereconstrueerd. Of artefacten van het tweede type ook gecorrigeerd moeten worden is lastiger te bepalen en hangt af van de oorzaak van het artefact en voor welk onderzoek het signaal bedoeld is. In ons geval zijn de artefacten allemaal zoveel mogelijk gecorrigeerd. Zoals later zal blijken, zijn zowel de gebruikte spectraalanalytische technieken als de multivariate analyse methoden gevoelig voor waarden die sterk afwijken, en dus ook voor artefacten. Het verwijderen van artefacten is dus noodzaak. Wanneer de artefacten gedetecteerd zijn, kunnen ze vervolgens door lineaire interpolatie gecorrigeerd worden. Uit ervaring blijkt dat dit goede resultaten oplevert.

2.2.5 Afgeleide variabelen

Om met behulp van multivariate data-analyse een index voor mentale inspanning te creëren is het noodzakelijk dat we een aantal variabelen kiezen waarvan de waarde gerelateerd is aan het niveau van mentale inspanning. Naast hartfrequentie en hartfrequentievariabiliteit, die al veel onderzocht zijn op hun geschiktheid, is de laatste decennia van verschillende signalen onderzocht of ze geschikt zijn. Een afgeleide variabele is bijvoorbeeld de *baroreflexgevoeligheid* (BRS, baroreflex sensitivity). BRS beschrijft de veranderingen in IBI ten opzichte van de veranderingen in systolische bloeddruk (SBP). Net als hartfrequentievariabiliteit, kan BRS ook d.m.v. spectraalanalyse bepaald worden.

Zoals gezegd, is er al veel onderzoek gedaan naar de geschiktheid van hartfrequentie en hartfrequentievariabiliteit (heart rate variability, HRV) als index. In de meeste gevallen werd bij een verhoogde mentale inspanning een verhoogde HR en een verlaagde HRV gevonden, zie bijvoorbeeld [Mulder, 1988]. Van Roon et al. laten zien dat hartfrequentie en bloeddruk belangrijke informatie kunnen geven over de effecten van mentale inspanning op het cardiovasculaire systeem. De cardiovasculaire effecten die bij mentale inspanning vaak gevonden worden, zijn een toename van de bloeddruk en hartfrequentie en een afname van de bloeddrukvariabiliteit, hartfrequentievariabiliteit en baroreflexgevoeligheid [Roon et al., 1995].

In tabel 2.2.5 wordt een overzicht gegeven van de variabelen die gebruikt zijn. In sectie 2.3.3 wordt iedere variabele beschreven en de wijze waarop deze wordt verkregen.

11 Cardiovasculaire Maten	Afkorting
Directe maten	
1. Hartfrequentie (of Interbeat interval)	HR (of IBI)
2. Systolische bloeddruk	SBP
3. Diastolische bloeddruk	DSP
Hartfrequentie Variabiliteit	
4. midden frequentieband	HRV_M
5. hoge frequentieband	HRV_H
Systolische Bloeddruk Variabiliteit	
6. midden frequentieband	$SBPV_M$
7. hoge frequentieband	$SBPV_H$
Coherentie	
8. midden frequentieband	Coh_M
9. hoge frequentieband	Coh_H
Modulus	
10. midden frequentieband	Mod_M
11. hoge frequentieband	Mod_H

Tabel 2.1: Cardiovasculaire variabelen

2.3 Spectraalanalyse

2.3.1 Frequentiebanden

Door gebruik te maken van spectraalanalyse is het mogelijk variaties in het signaal in verschillende frequentiegebieden afzonderlijk te bestuderen. De variaties in de signalen worden meestal niet veroorzaakt door één bron maar door meerdere, denk alleen al aan het parasympathische deel en het sympathische deel van het zenuwstelsel als twee bronnen met tegengestelde invloed. Om beter grip te krijgen op hoe verschillende bronnen invloeden uitoefenen, kunnen de tijdreeksen bestudeerd worden in verschillende frequentiegebieden [Mulder, 1988]. De verschillende banden die doorgaans gebruikt worden in spectraalanalyse van cardiovasculaire tijdreeksen zijn:

Hele spectrum	0.02-0.50 Hz
Lage frequentieband	0.02-0.06 Hz
Midden frequentieband	0.07-0.14 Hz
Hoge frequentieband	0.15-0.40 Hz

In dit onderzoek is gekozen voor de midden en hoge frequentieband. De lage band weerspiegelt voornamelijk veranderingen ten gevolge van temperatuurveranderingen en langzame aanpassingen aan de taak. Deze aanpassingen aan de taak vinden zo langzaam plaats dat veranderingen in de lage band slecht te gebruiken zijn in een onderzoek dat veranderingen per minuut wil bestuderen. Bovendien zijn relaties tussen hartfrequentie en bloeddruk niet even sterk in alle frequentiebanden. Juist

in de de middenband rond 0.10 Hz en tussen 0.2 en 0.4 Hz, blijken de fluctuatuaties een hoge correlatie te hebben, terwijl de correlatie in de lage frequentieband laag is.

2.3.2 Profielen

Profielen geven het verloop in de tijd weer van een specifieke spectrale maat, zoals bijvoorbeeld de spectrale dichtheid. Een waarde van de betreffende maat wordt bepaald voor een kort tijdsegment (venster). Een serie van dergelijke waarden geeft een nieuwe tijdreeks door elke keer het venster enigszins op te schuiven en dan een nieuwe waarde te bepalen. Onze initiële gedachte was een venster van één minuut zonder overlap te nemen om onafhankelijkheid tussen de waarden te kunnen garanderen, een voorwaarde waaraan voldaan moet worden voor sommige multivariate data analyse technieken. Later bleek dat de grilligheid in het signaal die naar voren komt bij een venster van 1 minuut het noodzakelijk maakte een groter venster van vijf minuten te nemen met een overlap van vier minuten. Bij de analyse moet er daarom op gelet worden dat de afhankelijkheid binnen de variabelen kan zorgen voor een onterechte verhoging van sommige prestaties.

2.3.3 Afleiding van variabelen

Hartfrequentie

Hoe de *hartfrequentie* en *interbeat interval* worden verkregen staat al enigszins beschreven in sectie 2.2.1. Uit het ECG wordt bepaald wanneer er een hartslag plaatsvond door de R-toppen in het signaal te bepalen. Deze punten leveren samen een reeks impulsen op die het voorkomen van hartslagen representeren. Deze reeks impulsen wordt mathematisch beschreven door de volgende functie:

$$x(t) = \sum_{i=1}^{N} \delta(t - t_i) \tag{2.1}$$

met:

i	R-top nummer
N	aantal R-toppen
t_i	tijdstip van R-top i
$\delta(t - t_i)$	unit impulse op punt t_i

Variabiliteit in hartfrequentie

De variabiliteit in het HR signaal (of het IBI-signaal) kan bepaald worden door spectraalanalyse; deze variabiliteit geeft als het ware de fluctuaties in het signaal weer. De aanname is dat voor het bestuderen van het spectrum gekeken moet worden naar het signaal dat bestaat uit de R-toppen en niet het originele signaal.

Vanwege het niet-equidistante karakter van $x(t)$ uit 2.1 is het niet mogelijk een *Fast Fourier Transformatie* (FFT) toe te passen om het spectrum te berekenen. De

Fourier getransformeerde van $x(t)$ uit 2.1 kan bepaald worden met behulp van een *Discrete Fourier Transformatie* (DFT):

$$X_x(f) = \int_0^T x(t)e^{-2\pi jft}dt \tag{2.2}$$

met:
$j^2 = -1$
f = frequentie
T = analysetijd.

Als we $x(t)$ uit 2.1 invullen in 2.2 en de grenzen aanpassen krijgen we de Fourier getransformeerde van $x(t)$:

$$X_x(f) = \sum_{i=1}^{N} e^{-2\pi jft_i} \tag{2.3}$$

Als dit spectrum wordt berekend over het segment $[0, T]$ ontstaat een lijnspectrum op de discrete frequentiepunten f_k met een frequentieresolutie Δf (verschil tussen f_k en f_{k-1}) van $1/T$:

$$X_x(f_k) = \sum_{i=1}^{N} e^{-2\pi jf_kt_i} \tag{2.4}$$

Uit 2.4 kan de spectrale dichtheids functie $S_{xx}(f_k)$ berekend worden met:

$$S_{xx}(f_k) = \frac{2}{T}X_x(f_k)X_x^*(f_k) = \frac{2}{T}\{[\sum_{i=1}^{N} cos(2\pi f_kt_i)]^2 + [\sum_{i=1}^{N} sin(2\pi f_kt_i)]^2]\} \tag{2.5}$$

Het blijkt dat spectrale dichtheids maten afhankelijk zijn van de gemiddelde waarde van IBI [Boer, Karemaker, and Strackee, 1985], [Dellen et al., 1985]. Door de spectrale waarden te normaliseren verdwijnt deze afhankelijkheid. Er ontstaat dan een signaal dat de variaties ten opzichte van het gemiddelde beschrijft en de modulatie index of modulatie diepte wordt genoemd. De genormaliseerde spectra kunnen berekend worden door de spectrale dichtheidsfunctie $S_{xx}(f_k)$ te delen door het gemiddelde van x in het kwadraat ($\bar{x}^2$):

$$S'_{xx}(f_k) = \frac{S_{xx}(f_k)}{\bar{x}^2} \tag{2.6}$$

Systolische bloeddrukvariabiliteit

De beschreven methode om de spectrale dichtheid van HR te berekenen kan algemeen gemaakt worden, zodat deze ook toegepast kan worden op andere signalen dan hartfrequentie, zoals de systolische bloeddruk. De beschrijving van $x(t)$ uit 2.1 moet dan uitgebreid worden met de waarde die het de systolische bloeddruk had in het interval voor de puls i uitgedrukt door de oppervlakte van de SBP in dat interval te nemen:

$$\tilde{x}(t) = \sum_{i=1}^{N} \Delta t_i x_i \delta(t - t_i) \tag{2.7}$$

met:
$\Delta t_i = t_i - t_{i-1}$
$x_i =$ waarde van x voor punt i (R-top i)

De spectrale dichtheidsfunctie voor $x(t)$ uit 2.7 kan analoog aan de spectrale dichtheidsfunctie voor hartfrequentie berekend worden met:

$$\tilde{S}_{xx}(f_k) = \frac{2}{T} X_x(f_k) X_x^*(f_k) = \frac{2}{T}\{[\sum_{i=1}^{N} \Delta t_i x_i cos(2\pi f_k t_i)]^2 + [\sum_{i=1}^{N} \Delta t_i x_i sin(2\pi f_k t_i)]^2]\} \tag{2.8}$$

De normalisatie voor het algemene geval wordt gegeven door:

$$\tilde{S}'_{xx}(f_k) = \frac{\tilde{S}_{xx}(f_k)}{\bar{x}^2} \tag{2.9}$$

Baroreflexgevoeligheid

Baroreflexgevoeligheid (BRS) geeft de veranderingen in IBI ten opzichte van de systolische bloeddruk weer. De baroreflex is een reflex die ontstaat door reksensoren in de lichaams- en halsslagader. De baroreflex probeert de bloeddruk zo te beïnvloeden dat deze stabiel blijft. Bij een verhoogde bloeddruk wordt dus geprobeerd deze omlaag te krijgen en vice versa. Omdat BRS wordt beschouwd als een goede maat voor mentale inspanning is dit één van de variabelen die we hebben gebruikt. Net als HRV en SBPV kan ook de BRS berekend worden door spectraalanalyse. Daarvoor moeten we kijken naar de transferfunctie van de systolische bloeddruk naar het interbeat interval. Vanuit historisch perspectief wordt het interbeat interval en niet de hartfrequentie gebruikt voor het berekenen van de BRS.

Een transferfunctie beschrijft de lineaire relatie tussen twee signalen. Een transferfunctie beschrijft een systeem dat als input signaal $x(t)$ heeft en als output signaal $y(t)$. De transferfunctie is een complexe functie, die beschreven kan worden door z'n modulus en fase.

Het lineaire verband tussen input en output van het systeem dat de transferfunctie beschrijft wordt weergegeven door de coherentie. De coherentie tussen twee signalen $x(t)$ en $y(t)$ is een maat voor in hoeverre de twee signalen in het frequentiedomein correleren. Een coherentie van nul geeft aan dat er geen lineair verband te vinden is in het frequentiedomein, een coherentie van één geeft aan dat de signalen dezelfde frequentie-karakteristiek hebben. Bij een lage coherentie geeft de transferfunctie wel een beschrijving van het systeem maar aangezien input en output weinig verband houden, is de beschrijving van het systeem minder interessant. De transferfunctie is daarom alleen zinvol te interpreteren als de coherentie voldoende hoog is.

Stel dat de lijnspectra van IBI en SBP gegeven worden door $X_x(f_k)$ en $Y_y(f_k)$, dan kan de transferfunctie berekend worden door het kruispectrum:

$$S_{xy}(f_k) = \frac{2}{T} X_x^*(f_k) X_y(f_k) \tag{2.10}$$

De transferfunctie wordt dan:

$$H_{xy}(f_k) = \frac{S_{xy}(f_k)}{S_{xx}(f_k)} \tag{2.11}$$

De baroreflexgevoeligheid wordt dan berekent door de modulus van 2.11 te nemen. De modulus wordt gegeven door:

$$|H_{xy}(f_k)| = \sqrt{[Re\{H_{xy}(f_k)\}]^2 + [Im\{H_{xy}(f_k)\}]^2} \tag{2.12}$$

De coherentie wordt uit de kruisspectra en spectrale dichtheidsfuncties voor IBI en SBP berekend. Stel dat IBI en SBP (weer) respectievelijk beschreven worden door $x(t)$ en $y(t)$, dan wordt de coherentie functie γ_{xy}^2 gegeven door:

$$\gamma_{xy}^2(f_k) = \frac{|S_{xy}(f_k)|^2}{S_{xx}(f_k) S_{yy}(f_k)} \tag{2.13}$$

De coherentie tussen twee sinussen is altijd 1.0, onafhankelijk van het veschil in amplitude en fase. Daarom is de coherentie berekend over één frequentiepunt altijd 1.0. Door middel van een venster over een aantal frequentiepunten te leggen kan een coherentie berekend worden die daadwerkelijk iets zegt over de samenhang. In dit onderzoek werd de coherentie berekend over het gemiddelde van drie frequentiepunten.

Profielen

Vervolgens kan het spectrale profiel berekend worden van een spectrale dichtheidsfunctie door gebruik te maken van een venster. De spectrale dichtheid wordt berekend voor een tijdsegment ter breedte van het *venster*, vervolgens wordt het begin-

en eindpunt van het venster opgeschoven. Door dit voor het hele signaal te doen ontstaat er een functie die de spectrale dichtheid in de tijd weergeeft.

Het profiel kan berekend worden in verschillende spectrale banden. Voor iedere frequentieband $f_l..f_h$ wordt de dichtheid voor segment i, uitgedrukt door $B_{x,i}$, berekend door integratie van de spectrale dichtheid over de specifieke band. Omdat het spectrum alleen uit frequentie-componenten f_k bestaat is een integratie een sommatie van de producten van de componenten f_k in de geselecteerde band met de frequentieresolutie:

$$B_{x,i}(f_l, f_h) = \sum_{f_k=f_l}^{f_h} S_{xx}(f_k)\Delta f \tag{2.14}$$

waarbij $S_{xx}(f_k)$ de spectrale dichtheidsfunctie is waarvan we een profiel willen berekenen. Als we $B_{x,i}(f_l, f_h)$ vervolgens invullen in expressie 2.7, met $\Delta t_{step} = STEP$:

$$P_x(t) = \sum_{i=1}^{N} STEP B_{x,i}(f_l, f_h)\delta(t - t_{step}) \tag{2.15}$$

Hierbij is $P_x(t)$ het signaal dat gebruikt gaat worden in de multivariate analyse.

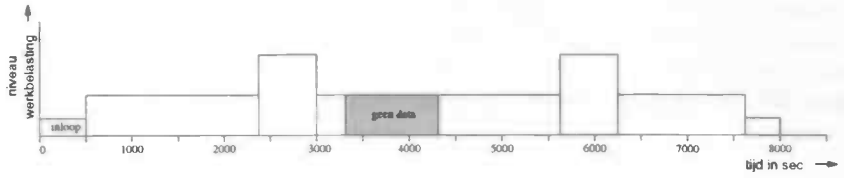
Hoofdstuk 3

Multivariate data analyse

3.1 Overzicht

In dit hoofdstuk worden de multivariate data-analysemethoden beschreven, die toegepast zijn op de tijdreeksen die verkregen zijn na de cardiovasculaire analyses. Na een inleiding over het onderzoeksgebied en de beschrijving van de voorbewerking op de data volgt de beschrijving van de analysemethoden.

In algemene zin kan elke statistische methode waarbij meerdere variabelen geanalyseerd worden een multivariate data analyse genoemd worden. In dit onderzoek hebben we specifiek gekozen voor een type benadering waarbij we door de simultane analyse van meerdere variabelen een uitspraak willen doen over de waarde van één afhankelijke variabele. De onafhankelijke variabelen zijn in dit geval de afgeleide cardiovasculaire tijdreeksen. De afhankelijke variabele is gedefinieerd als een classificatie van een periode in de taak die uitgevoerd werd door de proefpersonen. Het experiment had naast een inloop-periode, perioden met een relatief laag taaklast-niveau en perioden met een hoger taaklast-niveau die elkaar afwisselden, zie figuur 3.1. De afhankelijke variabele is in dit geval dichotomous, en kan de waarden hoog (1) en laag (0) aannemen.



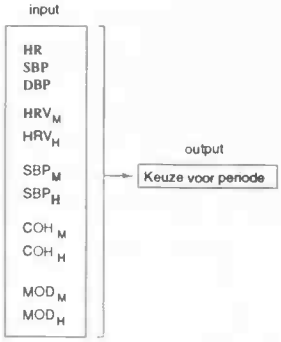
Figuur 3.1: Het taakverloop

Zoals uit figuur 3.1 blijkt, zijn er perioden met een lage en perioden met een hoge taaklast. Een onderscheid maken tussen perioden met een verschil in taaklast wordt in dit onderzoek gezien als een onderscheid in mentale inspanning. De aanname waar deze abstractie op steunt is dat een verschil in taaklast een verschil in

mentale inspanning zal veroorzaken. Zoals eerder al is duidelijk gemaakt, is mentale inspanning niet alleen afhankelijk van taaklast maar van meerdere factoren, zie hoofdstuk 1. Wanneer een persoon niet bereid is extra mentale inspanning te leveren bij een verhoging van de taaklast, zal er geen verhoging van de mentale inspanning optreden. De taakuitvoerder kan bijvoorbeeld geen motivatie hebben om extra inspanning te leveren of niet in staat zijn extra inspanning te leveren doordat de taak te moeilijk is. Wanneer er door het verschil in taaklast geen verschil in mentale inspanning ontstaat, kunnen de perioden niet van elkaar onderscheiden worden op basis van taaklast. Wanneer er wel een onderscheid gemaakt kan worden tussen de perioden, moet gecontroleerd worden of dit onderscheid veroorzaakt wordt door het verschil in taaklast van de perioden. Deze controle kan gedaan worden door alle data random aan een periode toe te wijzen en vervolgens de analyse uit te voeren. Als deze randomdata nog steeds goed geclassificeerd kan worden, kan het onderscheid niet toegeschreven worden aan een verschil in taaklast.

3.1.1 Input-output model

Voor elke methode die gebruikt wordt om dit onderscheid te kunnen maken geldt dat de afgeleide cardiovasculaire maten de input zijn en de categorie van de werkbelasting (hoog of laag) de output:



Figuur 3.2: De cardiovasculaire tijdreeksen zijn de input, de keuze voor de periode de output.

Er zijn verschillende benaderingen om de relatie tussen een aantal onafhankelijke en één afhankelijke variabele te onderzoeken. Drie van zulke methoden zijn:

Multiple regressieanalyse Het doel van multiple regressieanalyse is het voorspellen van de veranderingen in de afhankelijke variabele die samenhangen met de veranderingen in de onafhankelijke variabelen. Meestal wordt gebruik gemaakt van de kleinste-kwadratenmethode om deze relatie te vinden. In de

meestgebruikte vorm van multiple regressieanalyse moeten zowel alle onafhankelijke variabelen als de afhankelijke variabele metrisch van aard zijn [1] De afhankelijke variabele in dit onderzoek is in principe een niet-metrische variabele omdat deze alleen de waarden hoog en laag aan kan nemen. Om toch met multiple regressieanalyse te kunnen werken moet er met een aangepaste variant gewerkt worden.

Multiple discriminantanalyse Multiple discriminantanalyse (MDA) is wat betreft het niet-metrische karakter van de afhankelijke variabele de aangewezen multivariate techniek. Net als bij multiple regressieanalyse moeten de onafhankelijke variabelen metrisch van aard zijn. Discriminant analyse is toepasbaar wanneer de data verdeeld kan worden in groepen op basis van een niet-metrische variabele (in dit geval: rust-hoog-laag). Multiple discriminantanalyse worden veel gebruikt in situaties waar het bepalen tot welke groep een object (persoon, bedrijf of product) behoort, het voornaamste doel is.

Neurale netwerken Neurale netwerken bieden van de hier genoemde drie methoden de meest flexibele vorm van analyse. Met neurale netwerken kunnen relaties tussen variabelen bestudeerd worden. Bovendien kan er data reductie en structuur analyse gedaan worden. Neurale netwerken verschillen van traditionele multivariate data technieken in zowel de wijze waarop ze gedefinieerd worden als de complexe relaties die gemodelleerd kunnen worden. Vanwege het flexibele karakter van neurale netwerken zijn er uiteenlopende toepassingen.

Neurale netwerken verschillen sterk van andere multivariate analysetechnieken, zoals multiple regressieanalyse en multiple discriminantanalyse. Waar andere multivariate analysetechnieken problemen vaak omschrijven in mathematische termen, wordt voor neurale netwerken de basis beschreven in een abstracte vorm van een menselijk neuraal netwerk. Een ander belangrijk verschil is het leeraspect. Neurale netwerken leren op een sequentiele basis; voorbeeld na voorbeeld wordt als input gegeven. Andere multivariate technieken nemen de voorbeelden vaak in één keer als input.

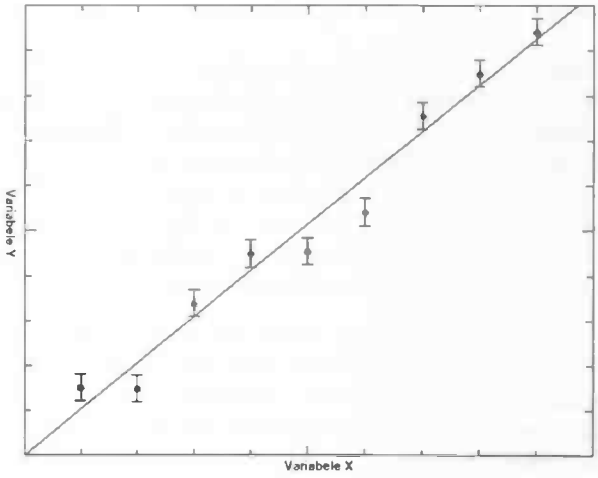
Dit verslag beschrijft verder de analysemethode en de resulaten van de toepassing van multiple regressieanalyse. Deze methode is gekozen omdat het verstandig is met eenvoudige technieken te beginnen. Wanneer er met een eenvoudige techniek namelijk geen enkel resultaat wordt behaald, presteren gecompliceerde technieken vaak ook slecht.

3.2 Multiple Regressie

Zoals gezegd is multiple regressieanalyse een manier om de relatie tussen meerdere onafhankelijke en een afhankelijke variabele te bestuderen. In het eenvoudigere geval

[1] Met metrisch wordt hier bedoeld dat een variabele niet alleen aangeeft of een eigenschap aan- of afwezig is maar in hoeverre een bepaalde eigenschap aan- of afwezig is. Afstand en leeftijd zijn bijvoorbeeld metrisch, i.t.t. bijv. geslacht.

van lineaire regressieanalyse wordt de relatie tussen één onafhankelijke variabele en een afhankelijke variabele bestudeerd. In een regressieanalyse wordt geprobeerd een regressievergelijking te vinden die het verband het best weergeeft. Gegeven deze vergelijking kunnen we het verband in een grafiek laten zien, zie 3.3.



Figuur 3.3: Regressie-lijn tussen variabele X en Y.

Wanneer we meerdere onafhankelijke variabelen hebben is het niet mogelijk dit grafisch weer te geven en moeten we volstaan met de regressievergelijking. Met n observaties en p onafhankelijke variabelen geeft dat een vergelijking voor een observatie van de volgende vorm:

$$y_i = \mathbf{x}_i\boldsymbol{\beta} + \beta_0 + \varepsilon \tag{3.1}$$

met:

$i = 1, 2, ..., n$

$\mathbf{x}_i = [x_1, x_2, ..., x_p]$

$\boldsymbol{\beta} = [\beta_1, \beta_2, ..., \beta_p]$

De regressiecoëfficienten, β_1, β_2,...,β_p, geven aan hoeveel variantie in de afhankelijke variabele y verklaard worden door iedere onafhankelijke variabele. De β-coëfficienten worden partiële coëfficienten genoemd omdat ze gecorrigeerd worden om te voorkomen dat twee variabelen dezelfde variantie in de afhankelijke variabele verklaren. De partiële regressie beschrijft hoe het gedrag van Y en X_1 samen veranderen als X_2 tot en met X_p constant worden gehouden. Met andere woorden, de partiële regressiecoëfficienten geven de unieke bijdrage van elke voorspellende variabele weer. Het equivalent van de simpele regressie coëfficient r tussen twee variabelen X en Y kan ook gegeven worden voor de partiële coëfficienten. De partiële

correlatie tussen Y en X_1 waar $X_2..X_p$ constant worden gehouden wordt geschreven als $r_{X_1Y|X_2..X_p}$. De partiële coëfficienten kunnen worden uitgedrukt in simpele coëfficienten:

$$r_{XY|Z} = \frac{r_{XY} - r_{XZ}r_{YZ}}{\sqrt{(1 - r_{XZ}^2)(1 - r_{YZ}^2)}} \tag{3.2}$$

Als X en Y geen correlatie met Z hebben geldt dus $r_{XY|Z} = r_{XY}$.

3.2.1 Residuals en R-kwadraat

De regressielijn drukt de beste voorspelling van de onafhankelijke variabele y uit gegeven de variabelen **x**. Omdat de voorspelling nooit volledig correct is, moet bepaald worden in hoeverre deze correct is. Dit kan gedaan worden door te kijken naar *residuals*. Een *residual value* is de afwijking van een specifiek punt ten opzichte van de regressie-lijn. Hoe kleiner de variantie in de *residual values*, des te beter beschrijft de lijn de data. Deze variantie van residual values wordt, voor de hand liggend, *residual variance* genoemd. Met residual variance kunnen we bepalen hoe goed de voorspelling van de regressievergelijking is. De standaardafwijking van de residuals wordt gegeven door:

$$s_{res} = \sqrt{\frac{\sum_{i=0}^{n}(Y_i - \bar{Y}_i)}{n-2}} \tag{3.3}$$

met:

Y_i waargenomen uitkomst voor observatie i
$\bar{Y}_i$ voorspelde uitkomst voor observatie i
n het aantal observaties

Een andere manier om aan te geven hoe goed het regressie-model is, is met R-kwadraat. R-kwadraat wordt bepaald door het kwadraat te nemen van de proportie verklaarde variantie R. Als R-kwadraat 1 is dan wordt alle variantie van de oorspronkelijke data verklaard door de regressievergelijking, als R-kwadraat 0 is er juist geen enkele variantie. De proportie verklaarde variantie wordt gegeven door R:

$$R = \frac{verklaarde\ variantie}{totaal\ variantie} = \frac{\sum(Y_{est} - \bar{Y})^2}{\sum(Y - \bar{Y})^2} \tag{3.4}$$

met:

Y waargenomen uitkomst
Y_{est} voorspelde uitkomst
$\bar{Y}$ gemiddelde waargenomen uitkomsten

Het kwadraat van R geeft de R-kwadraat zoals hierboven is beschreven. Omdat het aantal onafhankelijke variabelen dat meegenomen wordt van invloed is op de

proportie verklaarde variantie, is het verstandig hiervoor te compenseren. Dat is precies wat er voor *R-adjusted* is gedaan:

$$R_{adj} = 1 - \frac{onverklaarde\ variantie\ /(n-k-1)}{totaal\ variantie\ /(n-1)} = 1 - \frac{\sum (Y - Y_{est})^2}{\sum (Y - \bar{Y})^2} \frac{(n-1)}{(n-k-1)} \tag{3.5}$$

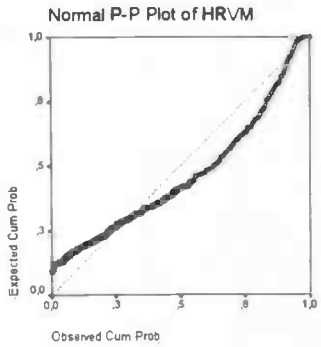
3.2.2 Aannames van lineariteit en normaliteit

Net als bij niet-multivariate lineaire regressieanalyse wordt er bij multiple regressieanalyse vanuit gegaan dat de relatie die bestudeerd wordt lineair van aard is en normaal verdeeld, hoewel multiple regressieanalyse niet erg gevoelig voor niet-lineariteit en scheve verdelingen is. Toch is het niet onverstandig te kijken naar de lineariteit en normaliteit van de variabelen om eventueel een transformatie op de data uit te voeren zodat deze beter verdeeld is.

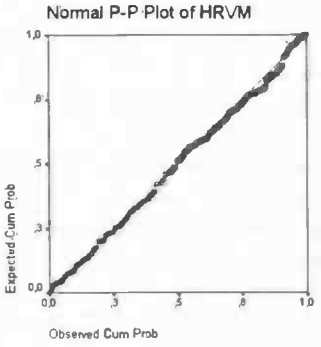
Uit experimentele gegevens blijkt dat spectrale maten van hartfrequentie en bloeddruk vaak niet normaal verdeeld zijn. In onderzoek van Van Roon komt naar voren dat de verdeling van verschillende maten minder afwijkt van een normale verdeling na transformatie [Roon, 1998]. Van Roon laat zien dat hartfrequentievariabiliteit, systolische bloeddrukvariabiliteit en modulatiediepte niet normaal verdeeld zijn. Onder de aanname dat de data χ^2_n verdeeld is, is het mogelijk de data met een logaritmische functie te transformeren zodat de verdeling normaal wordt. De transformatie zal ervoor zorgen dat de tijdreeksen na transformatie in alle spectrale banden normaal verdeeld zijn. Vanuit deze achtergrond hebben we de variabelen die verbeterd kunnen worden getransformeerd. Tabel 3.2.2 geeft een overzicht van welke variabelen getransformeerd zijn. Figuur 3.4 en figuur 3.5 laten als voorbeeld zien hoe de lineariteit van de hartfrequentie voor en na transformatie met een logaritmische functie verschilt. Beide figuren tonen hoe de cumulatieve kans van de data zich verhoudt tot een normale verdeling. In figuur 3.4 is te zien dat de cumulatieve kans van de hartfrequentie duidelijk afwijkt van die van een normale verdeling. In figuur 3.5 is te zien dat na transformatie de hartfrequentie normaler verdeeld is.

Cardiovasculaire Maat	Transformatie
1. Hartfrequentie	geen
2. Systolische bloeddruk	geen
3. Diastolische bloeddruk	geen
Hartfrequentie Variabiliteit	
4. midden frequentieband	logaritmisch
5. hoge frequentieband	logaritmisch
Systolische Bloeddruk Variabiliteit	
6. midden frequentieband	logaritmisch
7. hoge frequentieband	logaritmisch
Coherentie	
8. midden frequentieband	geen
9. hoge frequentieband	geen
Modulus	
10. midden frequentieband	logaritmisch
11. hoge frequentieband	logaritmisch

Tabel 3.1: Cardiovasculaire maten en hun transformatie.



Figuur 3.4: Verschil van de cumulatieve kansverdeling van de hartfrequentie t.o.v. de normale verdeling in de midden band *zonder* logaritmische transformatie.



Figuur 3.5: Verschil van de cumulatieve kansverdeling van de hartfrequentie t.o.v. de normale verdeling in de midden band *na* logaritmische transformatie.

3.2.3 Causaliteit

Wat bij het gebruik van correlatietechnieken zoals multiple regressieanalyse nooit vergeten moet worden is dat de correlatie tussen variabelen niets zegt over de onderliggende causale verbanden. Correlatietechnieken tonen slechts aan dat bepaalde gebeurtenissen samen voorkomen. De cardiovasculaire maten en de index die we daarmee proberen te ontwikkelen zijn bijvoorbeeld slechts *markers* van een verhoogde mentale inspanning. De mentale inspanning zal eerder de oorzaak zijn van

een verhoogde hartfrequentie en bloeddruk dan andersom, maar ook over die causaliteit kan vanuit het multiple regressiemodel niet veel gezegd worden. Spreken over afhankelijke en onafhankelijke variabelen kan aan deze verwarring bijdragen. Daarom wordt in sommige onderzoeken gesproken over voorspellende variabelen en de voorspelde variabelen.

3.2.4 Multicollineariteit

Een lastig probleem in veel correlationele analyses is multicollineariteit, ofwel sterke samenhang tussen verschillende onafhankelijke variabelen. Denk bijvoorbeeld aan de lengte van het linkerbeen en de lengte van het rechterbeen als verschillende variabelen om te voorspellen hoe lang iemand is. Uiteraard is één van de twee redundant omdat deze weinig zal verschillen van de andere variabele.

Het optreden van multicollineariteit valt te verwachten bij verschillende variabelen die in dit onderzoek gebruikt worden. Uit ervaring blijkt bijvoorbeeld dat de informatie die systolische bloeddruk geeft vaak sterk overeen komt met de informatie die diastolische bloeddruk geeft. Dit is niet verwonderlijk omdat een stijging in de bloeddruk zowel in de systolische bloeddruk als in de diastolische bloeddruk naar voren komt. Hetzelfde geldt voor verschillende andere variabelen, hoewel geprobeerd is de variabelen zo te kiezen dat ze allemaal een unieke bijdrage kunnen leveren. De keuze van spectrale banden is bijvoorbeeld gemaakt op basis van veel onderzoek, zodanig dat deze zo veel mogelijk specifieke informatie geven [Mulder, 1988].

Met weinig en heldere variabelen is het nog wel te zien wanneer er multicollineariteit op zal treden, maar met veel en/of moeilijk interpreteerbare variabelen is het een stuk lastiger. Wanneer er veel variabelen gebruikt worden treedt multicollineariteit vaak vanzelf op; het is dan belangrijk te controleren op multicollineariteit.

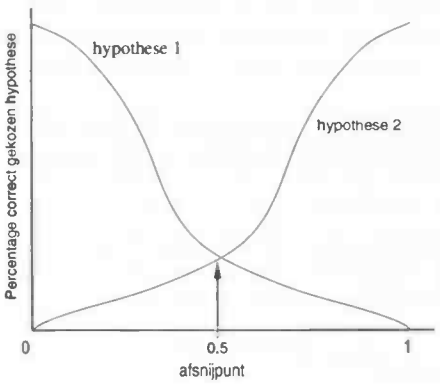
Bij de cardiovasculaire analyse is er voor het creëren van een tijdreeks gebruik gemaakt van profielen. Daarbij werd de keuze gemaakt om overlappende windows te gebruiken. Daarbij werd gezegd dat er een bepaalde afhankelijkheid tussen de waarden van één tijdreeks ontstaat. De multicollineariteit die hier bedoeld wordt, gaat juist om afhankelijkheden tussen variabelen. Desalniettemin kan de afhankelijkheid van waarden binnen een variabele ook een onterecht verhogend effect hebben op de R-kwadraat. Daarom is het verstandig het beoordelen van de regressievergelijking niet te beperken tot een onderzoek naar de R-kwadraat maar ook andere criteria te gebruiken. Er kan bijvoorbeeld getest worden hoe goed het regressiemodel is door te bepalen hoe goed het model *nieuwe* data kan voorspellen.

3.2.5 Voorspellen met behulp van de regressievergelijking

De regressievergelijking kan gezien worden als een beschrijving van de data, maar de vergelijking kan ook gezien worden als een model die voorspelt wat de waarde van de afhankelijke voorspelde variabele is gegeven de onafhankelijke variabelen.

In dit onderzoek is de output van de regressievergelijking een waarde tussen 0 en 1. Aan de hand van deze uitkomst moeten we een keuze maken tussen twee hypothesen: *de uitkomst is* 0 (taaklast was laag) tegenover *de uitkomst is* 1 (taaklast

was hoog). We kiezen een waarde tussen 0 en 1 als afsnijpunt, als de uitkomst van de vergelijking onder het afsnijpunt valt wordt de keuze een 0 en als de uitkomst erboven valt een 1. De afweging die we moeten maken welk snijpunt we kiezen wordt bepaald door de fouten die we maken bij het verwerpen van de verkeerde hypothese. Als het afsnijpunt dichter bij 1 ligt zullen we minder vaak de hypothese dat de uitkomst 1 is ten onrechte verwerpen. Dit betekent wel dat we vaker ten onrechte de andere hypothese dat de uitkomst 0 is verwerpen. In ons geval kost het verwerpen van de ene hypothese evenveel als het verwerpen van de ander, daarom is er gekozen voor een snijpunt dat precies op de helft (0.5) ligt, zie ook figuur 3.6.



Figuur 3.6: Hypothese 1: De uitkomst is 0, hypothese 2: de uitkomst is 1.

Als we vervolgens de uitkomst van de regressievergelijking en de oorspronkelijke classificatie van de data met elkaar vergelijken, kunnen we een uitspraak doen over hoe goed de regressievergelijking de data voorspelt. Als de regressievergelijking bij veel observaties dezelfde classificaties (1 of 0) als de originele data geeft, kunnen we zeggen dat de vergelijking een goed model is. Als de regressievergelijking daarentegen vaak andere uitkomsten geeft dan oorspronkelijk bij de observaties hoorden is het juist een slechte voorspeller. Hoe goed de regressievergelijking als model is kunnen we uitdrukken in het percentage dat goed voorspeld wordt. Deze prestatiemaat kunnen we naast de R-kwadraat gebruiken om te zien hoe goed de regressievergelijking de data beschrijft.

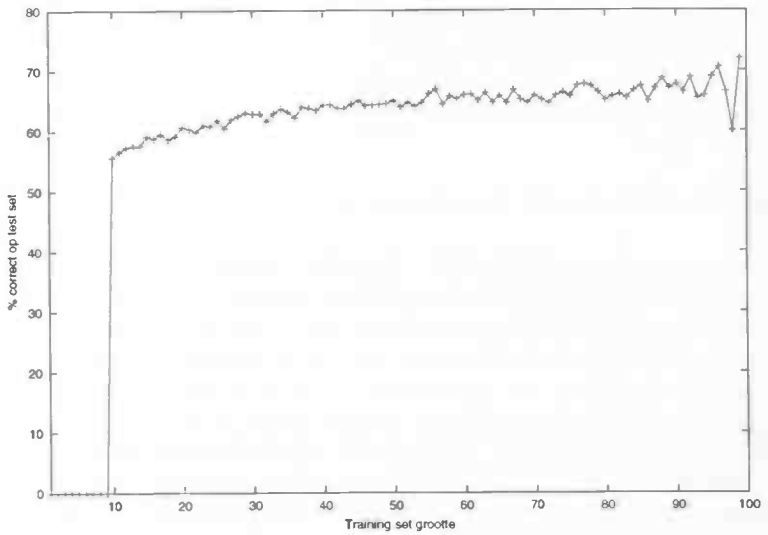
Test- en trainingset

Het is gebruikelijk om prestaties van voorspellende algoritmen te testen met behulp van een gescheiden training- en testset. Wanneer een algoritme op dezelfde sample getest wordt als waarop het getraind is, bestaat het risico dat de prestatie hoger is dan wanneer het algoritme op de populatie zelf getest zou worden. Het algoritme is dan te specifiek voor de trainingset. Dit kan komen door te veel trainen op een set of doordat een algoritme de ene set beter kan voorspellen dan de ander. Om

de resultaten zo algemeen mogelijk te maken en de reproduceerbaarheid van het onderzoek te vergroten wordt de data verdeeld in een test- en een trainingset.

De regressievergelijking wordt bepaald aan de hand van de training set. Vervolgens wordt de prestatie op het voorspellen getest met behulp van de testset. Door een aantal malen willekeurig uit de data een test- en trainingset samen te stellen, wordt de voorspelling betrouwbaarder. De test- en trainingset bestaan uit positieve en negatieve voorbeelden (hoge en lage classificatie) [2]. Het aantal positieve voorbeelden was in dit onderzoek beduidend lager dan het aantal negatieve voorbeelden, daarom moest er een aantal negatieve voorbeelden worden weggelaten om te voorkomen dat de prestatie omhoog zou gaan doordat de groepen niet even groot zijn. Wanneer bijvoorbeeld 80% van de voorbeelden geclassificeerd moet worden met 0, levert een algoritme dat altijd 0 als uitkomst geeft al een score van 80% goed op. Door de groepen precies even groot te maken is de maximum score van zo'n algoritme 50%.

De grootte van de trainingset wordt als een bepaald percentage van de voorbeelden Figuur 3.7 laat zien hoe de score op de testset afhangt van de grootte van de trainingset.



Figuur 3.7: Een leer curve voor de verhouding de grootte van de training- en de testset.

Zoals te zien is de figuur kan er geen score bepaald worden als de trainingset kleiner is dan 10% van de data (dataset bestond uit 120 cases). Na 10% loopt de curve langzaam een beetje op. Aangezien er slechts betrouwbaar een regressievergelijking bepaald kan worden met een minimum hoeveelheid voorbeelden, kan gezegd

[2]Zoals in het volgende hoofdstuk duidelijk zal worden, zijn er ook gevallen waarin er waarden tussen 1 en 0 zijn. In dat geval worden deze cases evenredig over de twee groepen verdeeld.

worden dat er op z'n minst 30% to 40% van de data in de trainingset terecht moet komen (ongeveer 40 cases). Hetzelfde geldt voor de testset: als deze erg klein is, is de score op een dergelijke set niet echt betrouwbaar. In de figuur is dat terug te zien in de toename van de variantie naarmate er minder data overblijft voor de testset. Helemaal rechts in de figuur als bijna alle data in de trainingset is terecht gekomen, en er daarom weinig overblijft voor de testset, vertoont de score op de testset grotere variantie.

Hier is daarom uiteindelijk gekozen voor een verdeling waarbij de helft van de data in de testset en de andere helft in de trainingset terecht kwam. Dit zorgt voor een redelijk betrouwbare vergelijking en een redelijk betrouwbare voorspelling van de prestatie.

Voting neighbors principe

Door gebruik te maken van een methode waarbij de uitkomst voorspeld wordt door meerdere observatiepunten kan een stabielere voorspelling ontstaan. In tegenstelling tot het voorspellen vanuit één observatie $\mathbf{x}_i$, wordt bij een *voting neighbors* voorspelling een voorspelling gedaan vanuit meerdere observaties. Hoeveel een observatie bijdraagt aan de voorspelling wordt bepaald door een coëfficient α_k. Hoe groter α_k des te groter de bijdrage van de observatie. De coëfficienten die de bijdrage van de verschillende observaties bepalen moeten uiteindelijk sommeren tot één ($\sum \alpha_k = 1$).

De uitkomst voor 'observatie' i wordt gegeven door:

$$y_i = \frac{1}{2M+1} \sum_{k=-M}^{M} \alpha_k \mathbf{x}_{(i+k)} + \varepsilon \tag{3.6}$$

met:

$\boldsymbol{x}_i$	De inputvector van de onafhankelijke variabelen.
α_k	De coëfficienten die bepalen hoe zwaar de observaties meetellen.
$2M+1$	Het aantal observaties dat een bijdrage levert.

Hoofdstuk 4

De Analyse

In de vorige hoofdstukken staat beschreven hoe uit gemeten hartslag en bloeddruksignalen verschillende tijdreeksen afgeleid kunnen worden. Daarnaast is uitgelegd hoe met deze tijdreeksen door middel van multiple regressieanalyse een index gevormd kan worden om de data te beschrijven en te voorspellen. Door verschillende manipulaties op de tijdreeksen uit te voeren voordat er (dezelfde) multivariate analyse techniek op wordt toegepast, is het mogelijk verschillende indices te creëren. In dit hoofdstuk staat beschreven hoe verschillende indices met behulp van de beschreven cardiovasculaire maten gevormd kan worden.

Voor de cardiovasculaire data wordt een regressievergelijking bepaald die er als volgt uit ziet:

$$\begin{aligned} Class &= \alpha + \beta_1 \cdot HR + \beta_2 \cdot SBP + \beta_3 \cdot DBP \\ &+ \beta_4 \cdot HRV_m + \beta_5 \cdot HRV_h + \beta_6 \cdot SBPV_m + \beta_7 \cdot SBPV_h \\ &+ \beta_8 \cdot COH_m + \beta_9 \cdot COH_h + \beta_{10} \cdot MOD_m + \beta_{11} \cdot MOD_h \end{aligned}$$

Met deze vergelijking als basis zijn er door het manipuleren van de data verschillende datasets gevormd. Door voor elk van deze datasets een multiple regressie analyse te doen werden er verschillende indices gevormd. Hieronder volgt een beschrijving van de verschillende indices.

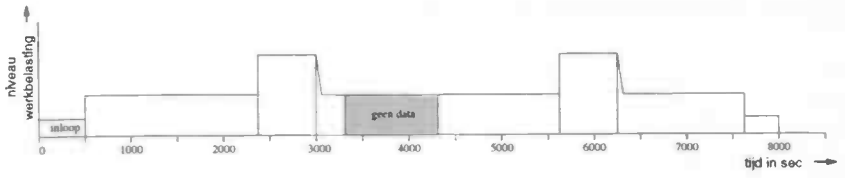
4.1 Basis

Voor deze index worden de cardiovasculaire maten gebruikt zonder daar enige bewerking op toe te passen. Bewerkingen om de normaliteit van de data te verbeteren, zoals eerder beschreven zijn wel toegepast. Alle andere indices ontstaan door verdere bewerkingen op de data in deze vorm.

4.2 Overgangswaarden

Vanuit een psychologische invalshoek zou er kritiek geleverd kunnen worden op de scherpe afscheiding tussen het hoge en lage werkbelastings-niveau zoals dat aangegeven is in figuur 3.1. In de eerste paar minuten na een periode met hogere werkbelasting kan er nog steeds een zwaardere belasting zijn doordat de extra werkdruk, die voor de hogere belasting zorgt, zich niet ineens oplost. Het onderzoek van Venema naar het gedrag en taakprestatie van de proefpersonen in dit experiment laat zien dat de perioden met hogere belasting, waarin veel ambulances weggestuurd moeten worden, gevolgd worden door een periode waarin het aantal langzaam afneemt [Venema, 2003]. Met andere woorden het gedrag van de proefpersonen laat een overgangsfase van een periode met hoge belasting naar een periode met lage belasting zien. Dit levert als hypothese dat een overgangsperiode waarin de werkbelasting niet geclassificeerd wordt als 1 of 0 maar een waarde in het interval $[0, 1]$ de werkbelasting realistischer weergeeft dan een scherpe overgang van 1 naar 0. Dit levert een taakverloop op zoals in figuur 4.1.

De manipulatie op de data betreft dus een aanpassing van de classificatie waarden. De overgang in het taakverloop wordt gerealiseerd door in de overgangsperiode de classificatie langzaam van 1 naar 0 te laten lopen. De overgangsperiode duurt vijf minuten (stappen van 1 minuut) en de waarde van de classificatie neemt zodanig af dat er een soepel verloop ontstaat: [... 1 1 1 0.9 0.7 0.5 0.3 0.1 0 0 0 ...].

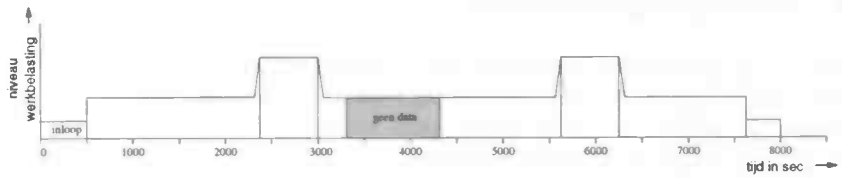


Figuur 4.1: Het taakverloop met een overgangsfase van hoge naar lage werkbelasting.

4.3 Dubbele overgangswaarden

Naast een uitloop van de perioden met een hoge werkbelasting, zou het ook kunnen zijn dat proefpersonen anticiperen op de hogere werkbelasting. Het onderzoek van Venema laat in zeker opzicht ook een dergelijke inloop-periode zien. Dat zou betekenen dat we naast een overgangsperiode na de periode met hogere belasting ook een overgangsperiode voorafgaand aan de perioden met hogere belasting kunnen verwachten. Beide overgangsperioden duren vijf minuten en hebben het zelfde verloop als in het geval met één overgangsperiode; dit levert een taakverloop op zoals in figuur 4.2. Net als bij de eenzijdige overgangsperiode wordt de classificatie van de data aangepast om nieuwe dataset te creëren.

[... 0 0 0.1 0.3 0.5 0.7 0.9 1 1 ... 1 1 0.9 0.7 0.5 0.3 0.1 0 0 ...]

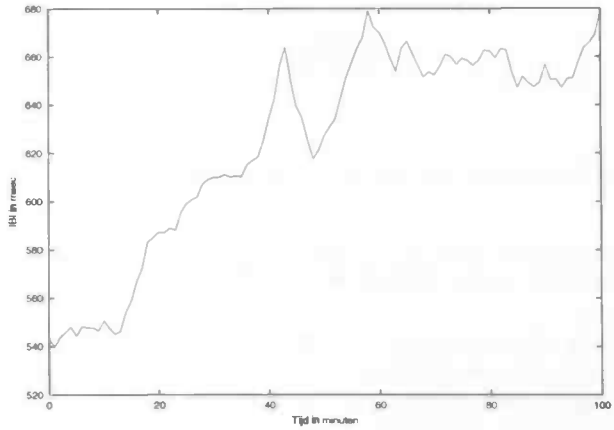


Figuur 4.2: Het taakverloop met een overgangsfase van lage naar hoge en van hoge naar lage werkbelasting.

4.4 Highpass filteren

Langlopende psychologische taken laten regelmatig zien dat bepaalde fysiologische variabelen gedurende de taak langzaam oplopen of afnemen. Een dergelijke trend in de data zorgt ervoor dat de variaties aan het begin niet in verhouding staan tot de variaties verderop in de data.

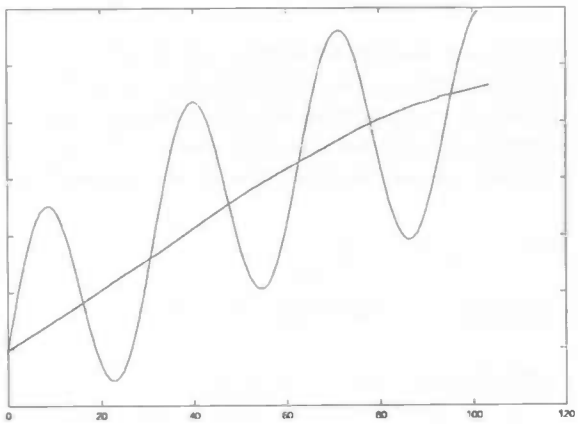
In figuur 4.3 is een interbeat interval signaal te zien met een trend. Aan het begin van de meting zijn de waarden gemiddeld 550 msec; aan het eind zijn de waarden gemiddeld 650 msec. Een afwijking van 7.5 msec (de standaardafwijking) betekent aan het begin van de data dus een percentuele afwijking van 1.36% en aan het eind een afwijking van 1.15%. Een standaardafwijking aan het begin wijkt dus 15% meer af dan een standaardafwijking aan het eind. Dit lijkt een erg onwenselijk effect.



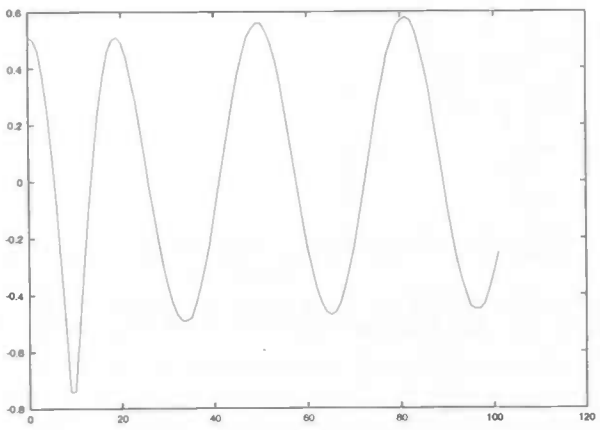
Figuur 4.3: IBI-signaal met trend.

Met een highpass-filter, een filter dat lage frequentie nauwelijks tot slecht doorlaat en hoge frequenties vanaf een bepaald ingesteld frequentie punt goed doorlaat, is het mogelijk een trend uit data te verwijderen. Figuur 4.4 laat een signaal zien dat uit twee sinussen bestaat, een sinus met een hoge frequentie die interessant voor

dit onderzoek is en een traag-frequente sinus die voor dit onderzoek de data alleen maar verstoord. De periode van de langzame sinus is tweemaal zo lang als de analyse periode. Door het signaal in 4.4 te filteren met een highpass-filter ontstaat een signaal dat slechts bestaat uit de sinus waarin we geïnteresseerd zijn, zie figuur 4.5.



Figuur 4.4: Twee sinussen, waarvan 1 de ongewenste trend is.



Figuur 4.5: Het signaal waaruit de ongewenste trend is verwijderd.

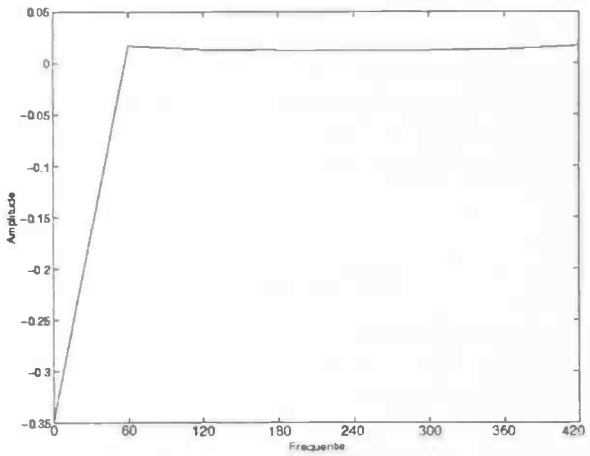
Wanneer in de volgende discrete beschrijving van een bandpassfilter als middenfrequentie (Ω_0) π genomen wordt en vervolgens een ruime bandbreedte van $1.89rad$ levert dit een highpassfilter op dat geschikt is om een langzame trend te verwijderen:

$$h[n] = \frac{1}{n\pi} sin(n\Omega_1) cos(n\Omega_0) \tag{4.1}$$

met:

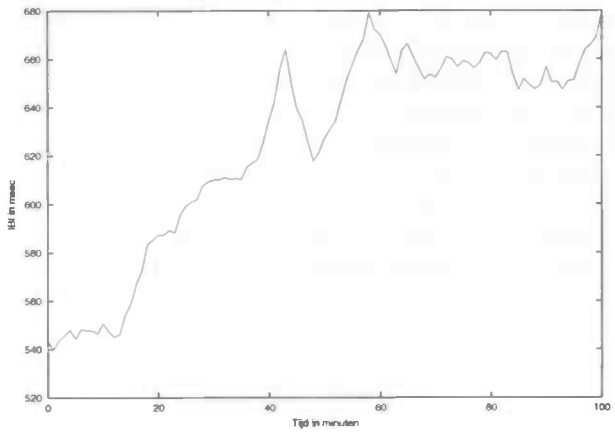
Ω_0 middenfrequentie
Ω_1 bandbreedte

Het filter dat door 4.1 beschreven wordt, met de genoemde parameters, is een niet-recursief hoogdoorlaatfilter. De impulsrespons van het filter wordt beschreven in 4.6 en is een causaal en eindig signaal, een *Finite impulse resonse* filter (FIR). Voor de data in dit onderzoek lijkt een discreet-filter met $n = 21$, $\Omega_0 = \pi$ en $\Omega_1 = 1.89\ rad$ goed te werken. Figuur 4.6 laat het filter zien in het frequentiedomein:

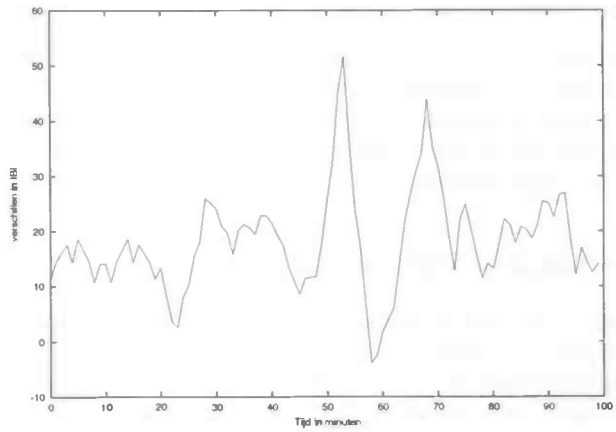


Figuur 4.6: Het highpass FIR-filter in het frequentiedomein.

Waneer we een highpass-filter, zoals uit figuur 4.6 toegepassen op het signaal uit figuur 4.7 ontstaat het signaal zoals dat weergegeven is in figuur 4.8. Als deze twee signalen vergeleken worden is te zien dat de trend verdwenen is. Bovendien is te zien dat de verhouding tussen de variaties aan het begin en aan het einde van het signaal is veranderd.



Figuur 4.7: Op herhaling: Het IBI-signaal laat een duidelijke trend zien.



Figuur 4.8: De ongewenste trend is uit het IBI-signaal verwijderd.

Door de inlooptijd van het filter raakt een deel van de data verloren. Om daarvoor te compenseren wordt er een gedeelte van de data gespiegeld voor de data geplakt. Dit segment is precies even lang als de inlooptijd van het filter, en heeft dus effect op de gehele inlooptijd. Zoals te zien is in figuur 4.5 blijven de effecten van de inlooptijd van het filter wel zichtbaar in het signaal.

4.5 Lopende variantie

De lopende variantie van een signaal geeft aan wat de variantie van het signaal is als functie van de tijd. Hartfrequentievariabiliteit en systolische bloeddrukvariabiliteit zijn voorbeelden van maten die de variabiliteit als functie van de tijd geven. De index

op basis van lopende variantie bestaat uit maten die de variabiliteit als functie van de tijd geven van de cardiovasculaire maten die genoemd zijn in tabel 2.2.5. In het geval van HRV en SBPV wordt dus gekeken naar de hoeveelheid variantie in de variabiliteit van de hartfrequentie resp. de systolische bloeddruk.

De lopende variantie wordt bepaald door een venster met een speciefieke breedte over de data te schuiven met een bepaalde stapgrootte. Vervolgens wordt voor iedere stap de variantie bepaald. De variantie in het venster wordt vervolgens toegeschreven aan het tijdspunt dat in het midden van het venster ligt. De algemene vorm van een rechthoekig venster om de lopende variantie s_i' te berekenen ziet er als volgt uit:

$$s_i' = \frac{1}{2M+1} \sum_{k=-M}^{M} s_{(i+k)} \tag{4.2}$$

met:

i = [1,2,...,n]	n=lengte signaal
s_i	variantie van het signaal op tijdstip i
$2M+1$	breedte van het venster

Omdat een groter venster zorgt voor een lagere tijdsresolutie en kleiner venster voor een onnauwkeuriger schatting van de variantie in het signaal moet een afweging gemaakt worden over de grootte van het venster. In dit onderzoek is gebruik gemaakt van een rechthoekig venster van 7 punten breed, met een stapgrootte van één, waarbij elk punt 1 minuut representeert.

4.6 Combinatie van methoden

Een logische aanpassing van de beschreven methoden om de prestatie te verbeteren lijkt de combinatie van verschillende methoden. Het probleem dat hier meteen opduikt is multicollineariteit. De data die verkregen wordt door twee verschillende methoden, de basis methode en de lopende variantie methode bijvoorbeeld, kan gecombineerd worden tot één dataset. Zo ontstaat er een dataset met tweemaal zoveel onafhankelijke variabelen, waarvan de eerste helft van de variablen ook nog eens sterk lijkt op de tweede helft. Behalve dat de prestatie door het toevoegen van variabelen lastiger te vergelijken is met de andere methoden [1], wordt de kans op afhankelijkheid tussen verschillende onafhankelijke variabelen ook groter.

De enige manier om te voorkomen dat er niet meer onafhankelijke variabelen komen is door niet alle variabelen uit de gecombineerde methode op te nemen. Het probleem dat hierbij ontstaat is de keuze welke variabelen wel en welke niet op te nemen. Een dergelijke keuze zou eigenlijk gemaakt moet worden op basis van de onderliggende theorie over de samenhang tussen het fysiologische systeem en de mentale werkbelasting. Wanneer er een keuze gemaakt wordt op basis van de prestatie van de verschillende variabelen, is er een groot risico dat er een te

[1] De hoogte van R-kwadraat wordt positief beïnvloed door het aantal onafhankelijke variabelen. Hierdoor kan onterecht de waarde van R-kwadraat hoger uitvallen.

dataspecifieke set variabelen wordt samengesteld. De prestatie op de betreffende dataset mag dan wel hoger zijn; of deze op een willekeurige andere set ook hoger is, is dan zeer de vraag. De conclusie die hieruit getrokken kan worden is, dat de mogelijkheid om te generaliseren over de resultaten in het gedrang komt, wanneer er op een dergelijke ad hoc wijze variabelen samengenomen worden om een nieuwe set te vormen.

Hoofdstuk 5

Resultaten

De verschillende methoden die in het vorige hoofdstuk beschreven zijn, zijn getest op de cardiovasculaire data van de tien proefpersonen. Eerst volgt een korte inleiding op de resultaten. Daarna volgen de resultaten zoals die verkregen zijn door de data van iedere proefpersoon individueel te analyseren. Een van de laatste secties is gewijd aan het verschil tussen de indviduele benadering en de benadering waar de data van de proefpersonen samengevoegd wordt.

5.1 Uitleg resultaten

Zoals in het vorige hoofdstuk beschreven staat worden er telkens meerdere subsets uit de data genomen om te testen. Van elk datapunt wordt willekeurig bepaald of deze in de test of in de trainingset terecht kwam. Bij het verdelen van de data is het belangrijk dat de zware-taaklastdata en de lichte-taaklastdata evenredig over de test en trainingset verdeeld wordt, omdat anders de prestatie ten onrechte beter kan worden, zie sectie 3.2.5. Omdat er minder data uit de perioden met zware taaklast beschikbaar was dan data uit de perioden met lichte taaklast werd de zware-taaklastdata volledig opgenomen en willekeurig verdeeld. Van de lichte-taaklastdata werden evenveel datapunten opgenomen als er van de zware-taaklast waren. Bij het bepalen van de regressie-vergelijking is dat gedaan omdat er een verschillende hoeveelheid data beschikbaar is uit de zware- en de lichtetaaklastperioden. Bij het voorspellen van de data zijn de subsets tevens gemaakt om verschillende test en trainingsets te creëren. Na individuele analyse is er vervolgens over de prestaties bij de verschillende proefpersonen gemiddeld.

R-kwadraat De R-kwadraat die telkens genoemd wordt is het gemiddelde van 100 R-kwadraatwaarden die bepaald worden door aselecte trekkingen uit de data waar telkens een regressieanalyse over wordt gedaan. Om aan te geven hoeveel de R-kwadraat om het gemiddelde schommelt is telkens de *standaardfout* vermeld. De standaardfout wordt in de tabellen met Std.Err. aangegeven. Door middel van een onder- en bovengrens (lower en upper bound) is een 99% betrouwbaarheidsinterval gegeven, een interval waarin de waarde van R-kwadraat bij een volgend voorbeeld

met 99% zekerheid zal liggen.

Voorspellen Bij de resultaten van het voorspellen wordt telkens het gemiddelde gegeven van 100 aselecte trekkingen. In dit geval staat het gemiddelde voor het percentage goed geclassificeerde voorbeelden. Ook hier worden standaardfout en een betrouwbaarheidsinterval van 99% aangegeven.

Vergelijken van resultaten Naast de ruwe score van een methode worden ook de verschillen met andere methoden getoond. In deze gevallen gaat het om een verschil tussen de R-kwadraat of de score van de ene methode ten opzichte van een andere methode. Wanneer er op hetzelfde moment meerdere methoden vergeleken worden, moet er rekening gehouden worden met een kans-kapitalisatie. Wanneer er één toets wordt gedaan, is er een kans dat er een fout van de eerste (en/of tweede) soort wordt begaan. Een fout van de eerste soort wordt begaan wanneer een hypothese onterecht verworpen wordt, geen verschil vinden als er wel een verschil is. Een fout van de tweede soort is juist het omgekeerde, een hypothese onterecht aannemen oftewel, een verschil vinden waar er geen verschil is. Wanneer meerdere toetsen worden gedaan, dan tellen de kansen op deze fouten op. De waarschijnlijkheid dat over alle toetsen een fout wordt gemaakt is dan groter dan de waarschijnlijkheid dat er bij één toets een fout wordt gemaakt. Er zijn verschillende oplossingen voor dit probleem, in dit onderzoek is gebruik gemaakt van de Bonferroni inequality. Hierbij wordt het significantieniveau van de individuele toetsen zo aangepast dat er over het geheel van de toetsen een significantieniveau van 0.01 wordt bereikt. Wanneer het aantal toetsen groter is dan 10 wordt de volgende beschrijving van het significantieniveau gebruikt:

$$\alpha_T = 1 - (1 - \alpha)^k \tag{5.1}$$

met:

α	significantieniveau voor individuele toets
α_T	significantieniveau voor het geheel van de toetsen
k	het aantal toetsen

In dit onderzoek zijn er 12 verschillende methoden voor het voorspellen van de data. Naast de basis en random methode met en zonder voting zijn er de 4 andere bewerkingen met en zonder gebruik van voting. In totaal moeten er daarom 12 toetsen uitgevoerd worden, k=12. Als er voor de toetsen een niveau van $\alpha = 0.01$ gebruikt wordt, is het overall niveau $\alpha_T = 1 - (1 - 0.01)^{12} = 0.114$. Dan is te zien dat α_T $0.114/0.01 = 11.36$ maal te groot is. Daarom moet er voor de individuele toetsen $\alpha = 0.01/11.36 = 0.0009$ gebruikt worden. Met $\alpha = 0.0009$ geldt $\alpha_T = 1 - (1 - 0.0009)^{12} = 0.01$.

Random Een model dat data voorspelt zou random data ook daadwerkelijk random moeten voorspellen. Als dit niet zo is, is er blijkbaar een bepaalde bias in de

data of het model aanwezig. Om dit te controleren, is er een dataset voorzien van een random-classificatie in plaats van de originele classificatie. Vervolgens is het model dat wordt bepaald door de regressieanalyse getest op deze data. Aangezien er precies evenveel positieve als negatieve voorbeelden in de data zitten [1] (enen en nullen), is de te verwachten score 50%. De score van het basismodel op de data met random-classificatie is 49.9%, met een significantieniveau van 0.65. Deze test laat dus geen bias in de data of het model zien.

Naast het voorspellen is er ook een R-kwadraat bepaald van het model op deze random-classificatie data, deze is 0.094. De methoden die de werkelijke data moeten beschrijven, moeten dus een R-kwadraat hebben die op z'n minst hoger is dan 0.094.

5.2 Resultaten basismethode

Bij de basismethode werd de regressieanalyse direct toegepast op de data zoals de cardiovasculaire analyse die opleverde, met classificaties zonder overgangswaarden. Tabellen 5.1 en 5.2 laten de resultaten van deze methode zien:

R^2	Std. Error	lower bound	upper bound
.32	.02	.25	.39

Tabel 5.1: De R^2 van de basis methode.

Score	Std. Error	lower bound	upper bound
66.9%	1.41	62.3%	71.4%

Tabel 5.2: De score van de basis methode op het voorspellen van de data.

De lower bound en upper bound geven het 99% betrouwbaarheidsinterval aan. Een eenzijdige vergelijking met de theoretische ondergrens van 50% laat zien dat deze methode 16.9% beter scoort dan 50%, een significante verbetering. De R-kwadraat laat in vergelijking met die van de random-classificatie data een verschil zien van 0.32-0.09=0.22, ook een significant verschil.

5.3 Resultaten overgangsmethoden

De R-kwadraat van de overgangsmethoden zijn beiden significant beter dan de R-kwadraat op de random-classificatie data die $R^2 = 0.09$ was. De R-kwadraat voor de eenzijdige overgang was $R^2 = 0.29$ die voor de dubbelzijdige overgang was $R^2 = 0.27$. De score van deze beide methoden is echter lager dan de basismethode, die immers $R^2 = 0.32$ was. Het verschil tussen de R-kwadraat van de overgangsmethoden en de basismethode is echter niet significant.

[1] De verhouding is precies 50% omdat de groepen even groot gemaakt zijn.

Methode	R^2	Std. Error	lower bound	upper bound
Eenzijdige overgang	.29	.021	.225	.362
Dubbelzijdige overgang	.27	.022	.200	.341

Tabel 5.3: De R^2 van de overgansgwaarden methoden.

Methode	Score	Std. Error	lower bound	upper bound
Eenzijdige overgang	65.4%	1.5%	60.7%	70.2%
Dubbelzijdige overgang	64.7%	1.5%	59.8%	69.6%

Tabel 5.4: De score van de overgangswaarden methoden op het voorspellen van de data.

De lower bound en upper bound geven het 99% betrouwbaarheidsinterval aan.

5.4 Resultaten highpass en lopende variantie

De R-kwadraat afkomstig van de data die highpass-gefilterd werd en de R-kwadraat afkomstig van de lopende variantie vertonen hetzelde patroon als de R-kwadraat van de beide overgangswaarden. Ze zijn beide significant beter dan de R-kwadraat afkomstig van de random-classificatie data, maar het verschil met de R-kwadraat van de basismethode is te klein om significant te noemen. Tabellen 5.5 en 5.6 geven een overzicht van de R-kwadraat en de scores van beide methoden, waarbij het betrouwbaarheidsinterval opnieuw 99% is:

Methode	R^2	Std. Error	lower bound	upper bound
Highpass	.21	.02	.15	.28
Lopende variantie	.25	.02	.20	.31

Tabel 5.5: De R^2 van de highpass en lopende variantie methoden.

Methode	Score	Std. Error	lower bound	upper bound
Highpass	59.1%	1.3	54.9%	63.4%
Lopende variantie	64.3%	1.4	59.9%	68.7%

Tabel 5.6: De score van de highpass en lopende variantie methoden op het voorspellen van de data.

5.5 Vergelijking verschillende methoden

De basismethode lijkt dus uiteindelijk resultaten op te leveren die beter of op z'n minst niet slechter zijn dan de resultaten van de andere methoden. Daarom worden de resulaten van alle methoden in tabel 5.7 met de basismethode vergeleken.

Basismethode R-kwadraat $R^2_{basis} = 0.32$, wordt vergeleken met R^2:						
Methode	R^2	R^2_{basis}-R^2	Std. Error	Sig.	low	upper
Random	0.09	0.22	0.02	0.00	0.11	0.33
Overgang	0.29	0.02	0.01	1.00	-0.04	0.09
Dub. Overgang	0.27	0.05	0.02	0.19	-0.03	0.12
Highpass	0.21	0.11	0.03	0.13	-0.05	0.26
Lop. Variantie	0.25	0.07	0.02	0.34	-0.06	0.18

Tabel 5.7: Het verschil van de R^2 van de basis methode met de R^2 van de andere methoden.

met:

R^2_{basis}	De R-kwadraat van de basis methode
R^2_{new}	De R-kwadraat van de methode waarmee wordt vergeleken
Std. Error	De standaardfout van de vergelijking van de methoden
Sig.	Het significantieniveau van de vergelijking van de methoden
low & upper	De grenzen van het 99% betrouwbaarheidsinterval

Basismethode Score $S_{basis} = 66.9\%$ vergeleken met andere methoden S:						
Methode	S	S_{basis}-S	Std. Error	Sig.	low	upper
Random	49.9	16.9	1.38	.00	8.34	25.56
Overgang	65.4	1.4	.42	.55	-1.22	4.05
Dub. Overgang	64.7	2.2	.71	.91	-2.27	6.62
Highpass	59.1	7.7	1.89	.180	-4.070	19.51
Lop. Variantie	64.3	2.5	1.48	1.00	-6.73	11.79

Tabel 5.8: Het verschil van de score van de basis methode met de scores van de andere methoden.

5.6 Voting Neighbors

Het toepassen van het *voting neighbors* principe zorgt voor een verbetering in de resultaten, zie tabel 5.9. Aangezien het in principe een nieuw model is om mee te voorspellen, is er opnieuw getest of het model op random data niet beter dan random presteert. Uit de test kwam een percentage goed geclassificeerde data naar voren van 50.0%, oftewel het model presteerde niet beter dan de theoretische ondergrens van 50%.

Methode	S	S_{voting}	$S - S_{voting}$	Std. Error	Sig.	low	upper
Random	49.9	50.0	.13	.29	1.00	-1.7	2.0
Basis	66.9	75.4	8.59	.57	.00	5.0	12.2
Overgang	65.4	73.4	7.97	1.08	.00	1.3	14.7
Dub. Over.	64.7	70.9	6.22	1.06	.00	-0.4	12.8
Highpass	59.1	65.7	6.52	.86	.00	1.1	11.9
Lop. Var.	64.3	72.5	8.14	.69	.00	3.8	12.5

Tabel 5.9: Het verschil in score met en zonder gebruik van het voting neighbors principe

met:

S	Het percentage goed geclassificeerd *zonder* voting neighbors
S_{voting}	Het percentage goed geclassificeerd *met* voting neighbors
Std. Error	De standaardfout van de vergelijking
Sig.	Het significantieniveau van de vergelijking
low & upper	De grenzen van een 99% betrouwbaarheidsinterval

5.7 Individueel versus Samen

De resultaten die hierboven staan beschreven zijn ontstaan door de data van de proefpersonen individueel te behandelen en vervolgens over de prestatie van de verschillende proefpersonen te middelen. Het is ook mogelijk een regressie model van de data van alle proefpersonen samen te maken. Vervolgens kan de prestatie op deze data vergeleken worden met de individuele benadering. Dit is gedaan in tabel 5.10:

Score Individueel zonder voting	66.4%
Score Overall zonder voting	58%
Score Individueel met voting	75.4%
Score Overall met voting	58.5%

Tabel 5.10: De score op individuele basis vergeleken met de score over alle proefpersonen.

In de tabel is te zien dat de score voor de individuele aanpak 66.4% zonder voting en 75.4% met voting is, terwijl de score die behaald wordt nadat alle data is samengenomen in beide gevallen ongeveer 58.0% is. Oftewel de individuele aanpak scoort 8.4% beter zonder voting en 17.1% met voting, significante verschillen. Tabellen 5.11 en 5.12 laten zien dat de score op de data voor bijna alle proefpersonen afzonderlijk inderdaad hoger dan 58% is en bovendien tussen nauwe grenzen ligt.

Proefpersoon	Score	lower	upper
2	59.7%	58.2%	61.2%
3	66.4%	64.7%	68.0%
4	60.4%	58.9%	62.1%
5	65.2%	63.9%	66.5%
6	69.3%	67.7%	71.0%
12	74.0%	72.6%	75.4%
13	72.4%	70.7%	74.0%
14	70.3%	68.6%	72.0%
15	62.7%	60.6%	64.7%
16	63.3%	61.8%	64.7%

Tabel 5.11: Individuele scores zonder voting neighbours.

Proefpersoon	Score	lower	upper
2	73.9%	71.3%	76.4%
3	73.8%	71.4%	76.3%
4	67.1%	64.1%	70.0%
5	75.5%	73.2%	77.7%
6	78.8%	76.5%	81.1%
12	83.0%	80.8%	85.2%
13	81.6%	79.0%	84.2%
14	81.8%	79.5%	84.1%
15	69.2%	66.2%	72.3%
16	69.4%	66.3%	72.5%

Tabel 5.12: Individuele scores met voting neighbours.

Hoofdstuk 6

Discussie

In hoofdstuk 5 stonden de resultaten van de analyse beschreven. In dit hoofdstuk zullen de resulaten verder besproken worden. Een aantal resultaten verdient meer aandacht: het algemene resultaat van het onderzoek, welke aanpassingen daadwerkelijk voor verbetering zorgen en het verschil tussen de resulaten binnen en tussen proefpersonen.

6.1 Algemeen resultaat

Dit onderzoek laat zien dat het mogelijk is om de perioden in de taak tot op zekere hoogte van elkaar te onderscheiden. Uiteindelijk is het mogelijk gebleken om ongeveer 75% van de data goed te classificeren. Dat betekent dat nog steeds een kwart niet goed geclassificeerd wordt. Ongeveer 25% kan op basis van een verschil in taaklast gevonden worden. Voor de overige 25% moet gedacht worden aan andere oorzaken. Zoals in de inleiding is beschreven kan de taakuitvoerder bijvoorbeeld geen motivatie hebben om extra inspanning te leveren. Of de taak kan zo moeilijk zijn voor de taakuitvoerder dat het niet mogelijk is extra inspanning te leveren bij een verhoogde taaklast. Bovendien moet er rekening mee worden gehouden dat personen voortdurend willekeurige momentane responspatronen laten zien. Met andere woorden het fysiologische systeem heeft een erg ruizig karakter en ruis is erg moeilijk te modelleren. Hierdoor zal een gedeelte van de fysiologische respons nauwelijks iets met de taaklast en/of de inspanning te maken hebben en moeilijk of niet te classificeren zijn.

In de context van dit onderzoek kunnen voor de marge verschillende verklaringen gegeven worden. Allereerst blijkt uit het onderzoek van Kruizenga [Kruizenga, 2003] dat de verschillen tussen de perioden moeilijk vast te stellen zijn. Over de proefpersonen werden wel duidelijke overeenkomsten aangetroffen, zoals het oplopen van de bloeddruk gedurende het experiment. Dit is zeer opmerkelijk omdat het een langlopend experiment betreft, en de bloeddruk dus gedurende twee uur op bleef lopen. Hetzelfde geldt voor het dalen van de hartfrequentie en de toename in baroreflexgevoeligheid. Dit kunnen effecten van vermoeidheid zijn. Voor het onderscheiden van de perioden dragen deze patronen niet bij. Doordat de belangrijkste patronen niet bijdragen aan de onderscheidbaarheid, valt te verwachten dat de classificatie

moeilijk is.

Een oorzaak voor de moeilijk te interpreteren data kan zijn dat de taak tijdens drukke en rustige perioden niet veel van elkaar verschilt. Het verschil in belasting moet voortkomen uit een verschillend aantal ambulances dat weggestuurd moet worden. Of dat verschil in aantal daadwerkelijk een goede onderscheiding biedt blijft de vraag.

Daarnaast wordt in soortgelijk onderzoek juist vaak een onderscheid gezocht in taak tegenover rust, in plaats van taak tegenover taak. De fysiologische toestanden verschillen gedurende een rustperiode vaak erg sterk van de taakperiode, terwijl de fysiologische toestanden van twee taakperioden veel meer op elkaar lijken. De taakmoeilijkheid tussen de rustige en drukke stukken in het experiment verschilt minder dan gedacht werd. De proefpersonen lijken een strategie te volgen waarbij ze het werk zo plannen dat het verschil in taaklast enigszins genivelleerd wordt.

De trends in de data laten zien dat de fysiologische respons gedurende het experiment veranderd. De bloeddruk wordt bijvoorbeeld, naast de invloed van het verschil in taaklast, ook beïnvloed door het langdurige karakter van het experiment. Het effect van de fysiologsche respons, die te zien is aan het begin van het experiment, werkt dus nog door aan eind van het experiment. Dit effect draagt niet bij aan het onderscheid tussen de perioden.

Ondanks de gecompliceerde data uit het experiment, waar met traditionelere methoden moeilijk uitspraken over gedaan konden worden, blijkt het toch mogelijk 75% goed te voorspellen op individueel niveau.

6.2 Resultaten aanpassingen

Overgangswaarden De resultaten laten zien dat de verschillende bewerkingen op de data geen verbeteringen van de classificatie opleveren. Ondanks dat het onderzoek van Venema laat zien dat er op taakniveau een zekere overgng tussen zwaardere en lichtere taaklast is, blijkt het niet bij te dragen aan een betere classificatie wanneer deze overgang expliciet wordt gemaakt in de data. Dat het verschil in taaklast niet zo groot is, wil blijkbaar niet meteen zeggen dat er een duidelijke overgangsperiode is.

Highpass filteren Ondanks dat het highpass filteren van de data op het eerste gezicht een keurig nieuw signaal oplevert waar de trend duidelijk uit verwijderd is, wordt de classificatie door het filteren niet verbeterd. Blijkbaar is het ontwikkelde regressiemodel prima in staat om de data (tot op zekere hoogte) juist te classificeren, onafhankelijk van of er een trend in het signaal zit of niet. In dat geval verstoord het highpass filter de data slechts.

Lopende variantie Het bepalen van de lopende variantie draagt niet bij aan het verbeteren van de classificatie. Het aantal events (ambulances die weggestuurd moeten worden) tijdens de drukke perioden neemt toe, waardoor er meer events zijn die hun invloed op het signaal uit kunnen oefenen. Als elke event voor een

verstoring in het hart- of bloeddruk signaal zorgt wordt de lopende variantie in de drukke perioden hoger vanwege een groter aantal events in die perioden. Dit verschil kan een verklaring zijn, waarom er met de lopende variantiemethode een gelijke score gehaald kan worden als met de basismethode. Aan de andere kant wordt de afstand tussen de data en bijbehorende classificaties door de lopende variantie te nemen opnieuw met een stap vergroot.

6.3 Voting neighbors

Gebruik maken van het voting neighbors principe is feitelijk de enige aanpassing die voor een significante verbetering van het resultaat zorgt. Zoals te zien is in tabel 5.9 verbetert de classificatie van elke methode door deze manier van voorspellen. De grilligheid die in de data gevonden kan worden blijkt door dit middelende effect goed te worden opgevangen. Het enige nadeel van deze techniek is de extra vertraging die opgelopen wordt wanneer een dergelijk principe wordt toegepast op symmetrische wijze. Dan worden namelijk zowel een aantal buren voor tijdstip t als na tijdstip t meegenomen om de classificatie te bepalen. In dit onderzoek is dat 3 minuten voor en 3 minuten na t. Maar omdat de data offline geanalyseerd wordt is dat geen bezwaar.

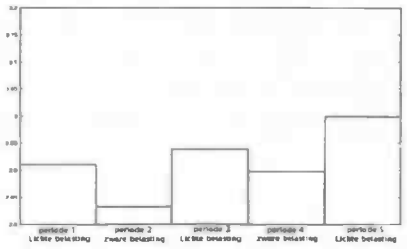
6.4 Individueel versus samen

Een opmerkelijk resultaat is het verschil in prestatie wanneer de data van proefpersonen individueel wordt behandeld en wanneer de data van alle proefpersonen samengenomen wordt. Het verschil in percentage juist geclassificeerde data tussen deze twee benaderingen is 8.4%. Dit resultaat is opmerkelijk omdat het niet gebruikelijk is binnen psychologisch onderzoek om de data van individuele proefpersonen te analyseren. Over het algemeen wordt de data van een groot aantal proefpersonen samen geanalyseerd. Het verschil in prestatie dat hier gevonden is, suggereert duidelijk een individuelere aanpak.

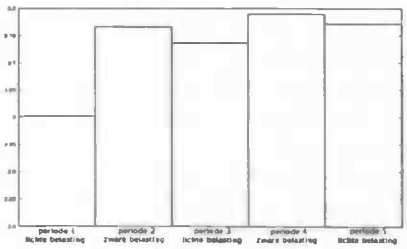
Het verschil tussen een individuele en een gezamenlijke aanpak wordt veroorzaakt doordat niet iedere proefpersoon dezelfde fysiologische respons geeft op de taak. Waar bijvoorbeeld de hartslagvariabiliteit in de middenband bij de ene proefpersoon omhoog gaat, gaat deze bij een ander juist omlaag. Kruizenga laat zien dat op een paar variabelen wel significante verschillen gevonden kunnen worden. Dit verklaard waarom er over alle proefpersonen ook classificatie mogelijk is. Het gaat echter niet om groot aantal patronen, de meeste patronen worden binnen de data van een proefpersoon gevonden. Verschillende oorzaken kunnen verantwoordelijk zijn voor deze individuspecifieke respons. Wanneer een persoon eerder vermoeid raakt zal bij hem of haar eerder een verhoging in variabiliteit van hartslag en bloeddruk verwacht mogen worden. Daarnaast is het persoonlijk hoe moeilijk iemand de taak vindt. Wanneer er een intensievere periode aanbreekt, en de centralist is al zwaar belast dan zal dit een grotere impact hebben dan wanneer de centralist de taak nog betrekkelijk goed aankan. Het laatste en misschien wel de belangrijkste oorzaak

voor individuele verschillen kan een verschil in extra mentale inspanning van de taakuitvoerder zijn. Wanneer een persoon niet bereid is, bij een verhoging van de taaklast, extra inspanning te leveren, zal dat een andere respons op leveren dan wanneer een persoon daartoe wel bereid is. De taakprestatie zal naar beneden gaan wanneer een taakuitvoerder niet bereid is extra inspanning te leveren, maar deze is moeilijk vast te stellen op individueel niveau.

Wanneer de respons van een enkele proefpersoon in verschillende sessies met elkaar vergeleken wordt, blijkt dat elke proefpersoon in alle sessies telkens wel hetzelfde patroon laat zien. Wanneer de bloeddruk oploopt in de ene sessie vanwege een verhoogde taaklast, dan zal deze ook oplopen in de andere sessies. Hieruit volgt dat het patroon individuspecifiek is. Wanneer er echter over individuspecifieke data gemiddeld wordt, blijft er geen weinig over, terwijl de proefpersonen afzonderlijk wel degelijk een effect laten zien.



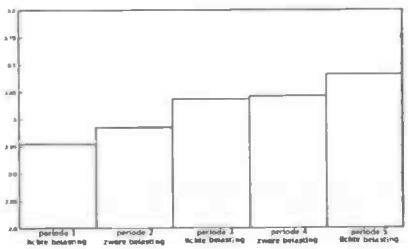
Figuur 6.1: Hartslagvariabiliteit HRV_m proefpersoon 2



Figuur 6.2: Hartslagvariabiliteit HRV_m proefpersoon 13

Figuur 6.1 en 6.2 laten zien dat de fysiologische reactie, in dit geval voor HRV_m, van twee proefpersonen voor dezelfde taak behoorlijk kan verschillen. In de figuren zijn de gemiddelde waarden van de hartslagvariabiliteit in de middenband in een specifieke periode aangegeven. Zowel voor proefpersoon 2 (6.1) als voor proefpersoon 5 (6.2) is er een significant verschil tussen de verschillende perioden. Maar uit de figuren blijkt dat het een effect in verschillende richtingen is. Figuur 6.3 laat zien dat het effect nagenoeg volledig verdwijnt wanneer het gemiddelde van de data van deze twee proefpersonen wordt genomen. Niet alle variabelen laten dit effect zo

sterk zien, vandaar dat de classificatie over de samengenomen data slechter, maar niet onmogelijk is.



Figuur 6.3: Gemiddelde hartslagvariabiliteit HRV_m van proefpersonen 2 en 13 samen.

Als laatste punt dient hier nog opgemerkt te worden dat de classificatie van de samengenomen data niet beter wordt door het voting neighbors principe. Dit wordt grootendeels veroorzaakt door de wijze waarop het principe werkt. Doordat verschillende naast elkaar liggende datapunten meewerken bij de classificatie van 1 punt, is het belangrijk dat deze punten in de tijd ook dicht bij elkaar liggen. Voor de individuele benadering klopt dit bij benadering behoorlijk goed. Omdat er willekeurige trekkingen worden gedaan, liggen de waarden die voor elkaar voorspellen in de tijd niet altijd naast elkaar, maar wel in de buurt. Voor de samengenomen data geldt dit veel minder. Ten eerste omdat er een vrij kleine subset uit de data van één proefpersoon wordt gebruikt, om het totaal aantal datapunten niet te veel te laten afwijken van het aantal datapunten dat wordt gebruikt voor de individuele benadering. Daardoor liggen de datapunten die gebruikt worden in de tijd vaak ver uit elkaar. Ten tweede zijn er situaties waar er van de data van de ene proefpersoon overgeschakeld wordt naar de data van een andere proefpersoon. Op deze punten wordt de voorspelling eerder slechter door voting neighbours dan beter.

6.5 Individuele verschillen

Naast de verschillende respons van een proefpersonen op één variabele, blijkt dat voor elke proefpersoon een andere set variabelen belangrijk is voor het onderscheid. Voor een aantal proefpersonen is een combinatie van de gemiddelde hartfrequentie en de systolische bloeddruk met name indicatief. Voor veel andere de hartslagvariabiliteit in de hoge band, gecombineerd met een andere variabele. De proefpersonen vertonen duidelijk een individuele respons. De patronen die hierboven genoemd zijn, maken het echter niet mogelijk de proefpersonen in te delen in groepen die duidelijk een karakteristieke respons hebben.

De unieke respons van proefpersonen zou verklaard kunnen worden door een verschil in geleverde mentale inspanning. Doordat de ene proefpersoon bereid is meer inspanning te leveren dan een andere, kan er een verschil ontstaan in de fysiologische

respons op een taaklastverandering. Daarnaast zal niet iedere persoon dezelfde fysiologische respons geven op dezelfde mentale inspanning.

Hoofdstuk 7

Conclusie en Aanbevelingen

7.1 Conclusie

Als antwoord op de onderzoeksvraag kan gezegd worden dat het mogelijk is om 75% van de data op individueel niveau juist te classificeren.

Hierbij komt de vraag naar voren in hoeverre deze perioden daadwerkelijk een verschil in mentale inspanning veroorzaken. Om hier meer duidelijkheid over te krijgen zouden er andere methoden gebruikt kunnen worden om het verschil in mentale inspanning te bepalen. Hierbij kan gedacht worden aan subjectieve methoden als aanvulling, waarbij proefpersonen bijvoorbeeld gedurende de taak op vaste tijdstippen op een schaal moeten aangeven hoeveel inspanning ze eraan besteden.

7.2 Individu-afhankelijke respons

Uit dit onderzoek bleek dat de fysiologische respons van proefpersonen onderling sterk verschilt. Om de stabiliteit van deze bevinding en de bruikbaarheid ervan voor classificatie verder te onderzoeken zou gedacht kunnen worden aan een onderzoek waarbij proefpersonen uitgenodigd worden mee te werken aan twee verschillende experimenten. Vervolgens kan gekeken worden of het regressiemodel dat ontwikkeld wordt in het eerste experiment goed presteert op de data van het tweede experiment. Een goede prestatie zou betekenen dat het inderdaad een individu-afhankelijke fysiologische respons betreft. De sterke overeenkomst tussen de data verkregen in verschillende sessies met één proefpersoon zijn suggereren reeds deze individu-afhankelijke fysiologische respons.

7.3 Neurale netwerken

Eerder in dit verslag zijn neurale netwerken genoemd als een zeer flexibele methode om data te classificeren. De kracht van neurale netwerken schuilt in de relatieve ongevoeligheid voor ruis en fouten in de data. De cardiovasculaire data heeft in bepaalde mate een grillig karakter. Dat wordt bevestigd door het positieve effect op de prestatie dat het voting neighbors principe als middelende methode heeft. Het is

dus goed mogelijk dat neurale netwerken erg geschikt zijn om de multiple regressie in dit onderzoek te vervangen.

7.4 Online herkennen

Het doel van dit onderzoek was het ontwikkelen van een model dat een onderscheid kan maken in mentale inspanning op basis van cardiovasculaire data en zo te fungeren als opstap naar een adaptieve werkomgeving. De analyse van de data is in dit onderzoek verricht na afloop van het experiment. Wanneer een model gebruik wil maken van de voorspelling van de hier ontwikkelde methode, zal de data gedurende de taak geanalyseerd moeten worden. Bij het online-analyseren van de data spelen er meer problemen een rol dan bij offline-data-analyse en classificatie. Als eerste kan hierbij gedacht worden aan rekenkracht. Offline is het geen probleem wanneer een analyse twee keer zo lang duurt als de tijdsduur van het experiment. Online is dat uiteraard onmogelijk. Daarnaast zijn er bij de analyse verschillende niet-causale technieken toegepast. Technieken die voor het berekenen van de waarde voor tijdstip t gebruik maken van data van tijdstippen na t. Wanneer deze methoden niet aangepast worden, betekenen deze niet-causale technieken een vertraging in de analyse. Het voting neighbors principe is hier een voorbeeld van. Hierbij wordt gebruik gemaakt van 5 datapunten om een voorspelling voor 1 datapunt t te maken; de voorspelling voor t zelf en de voorspelling van twee punten voor t en twee na t. Dit betekent dat de voorspelling voor tijdstip t pas aan het eind van tijdstip t+2 beschikbaar is. Om geen vertragingen op te lopen bij de online-classificatie kunnen de niet-causale bewerkingen aangepast worden zodat ze wel causaal worden. Dit betekent in het geval van het voting neighbors principe dat er niet langer over tijdstippen na tijdstip t gemiddeld wordt maar slechts over tijdstippen voor t. De prestatie zal hier hoogst waarschijnlijk niet zwaar onder lijden, maar dat dient uiteraard wel eerst uitgezocht te worden.

Referenties

Backs, R. W. (1995). Going beyond heart rate: Modes of autonomic control in the cardiovascular assessment of mental workload. *The International Journal of Aviation Psychology*, 5:25–48.

Backs, R. W., J. K. Lenneman, and J. L. Sicard. (1999). The use of autonomic components to improve cardiovascular assessment of mental workload in flight simulation. *The International Journal of Aviation Psychology*, 9:33–47.

Backs, R. W., A. M. Ryan, and G. F. Wilson. (1994). Psychophysiological measures of workload during continuous and manual performance. *Human Factors*.

Boer, R. W. De, J. M. Karemaker, and J. Strackee. (1985). Spectrum of a series of point events, generated by a integral pulse frequency modulation model. *Med. and Biol. Eng. and Comp.*, 23:138–142.

Cacioppo, J. T. and L. G. Tassinary. (1982). *Perspectives in cardiovascular psychophysiology*. The Guilford Press.

Cacioppo, J. T. and L. G. Tassinary. (1990). Inferring psychological significance from physiological signals. *American Psychology*, 45(1).

Dawson, Michael E., Anne M. Schell, and Diane L. Filion, (1990). "Principles of Psychophysiology: Physical, social and inferential elements". chapter 10, pages 295–324. Cambridge University Press.

Dellen, H. J. Van, J. Aasman, L. J. M. Mulder, and G. Mulder. (1985). Statistical versus spectral measures of heart rate variability. *The psychophysiology of cardiovascular control*, pages 353–374.

Duda, R. O., P. E. Hart, and D. G. Stork. (2001). *Pattern Classification 2nd ed.* Wiley-Interscience Publication: John Wiley and Sons, inc.

Hair, J. F., R. E. Anderson, R. L. Tatham, and W. C. Black. (1998). *Multivariate Data Analysis*. 5th edition. Prentice Hall.

Kruizenga, A. (2003). voorlopige titel:verslag leeronderzoek naar cardiovasculaire effecten van taakbelasting in een gesimuleerde ambulance-meldkamer. Afstudeerscriptie Psychologie, vakgroep Experimentele en Arbeidspsychologie.

Langelaan, S. (2001). Groningen acute stress test - een pilot studie. Master's thesis, University of Groningen.

Mulder, L. J. M., A. van Roon, M. Althaus, H. Veldman, K. Laumann, J. Bos, L. Quispel, and M. Dicke. (2002). Determining dynamic cardiovascular state change using a baro-reflex simulation model. ?

Mulder, L.J.M. (1988). *Assesment of cardiovascular reactivity by means of spectral analysis.* Ph.D. thesis.

Picard, R. W. and J. Klein. (2001). This computer responds to user frustration. *To appear in: Interacting with computers.*

Picard, R. W., E. Vyzas, and J. Healey. (2001). Toward machine emotional intelligence: Analysis of affective physiological state.

Picard, Rosalind W. and Jonathan Klein. (1900). Computers that recognise and respond to user emotion: Theoretical and practical implications. Technical report, MIT Media Lab. TR 538.

Pope, Alan T., Edward H. Bogart, and Debbie S. Bartolome. (1995). Biocybernetic system evaluates indices of operator engagement in automated task. *Biological Psychology,* 40:187–195.

Rompelman, O. (1980). The assesment of fluctuations in heart rate. *The study of heart rate variability.*

Roon, A. M. Van. (1998). *Short-term cardiovascular effect of mental tasks.* Ph.D. thesis.

Roon, A. M. Van, L. J. M. Mulder, J. B. P. Veldman, and G. Mulder. (1995). Beat-to-beat blood-pressure measurements applied in studies on mental workload. *Homeostatis.*

Trujillo, Anna C. (1998). Pilot mental workload with predictive system status information. To be published in Fourth Symposium on Human Interaction with Complex Systems, March.

Venema, I. (2003). Effects of instructions on situation awareness in a simulated ambulance dispatch task. Afstudeerscriptie Psychologie, vakgroep Experimentele en Arbeidspsychologie.

Vyzas, Elias. (1999). *Recognition of emotional and cognitive states using physiological data.* Ph.D. thesis, MIT, June.

Waard, D. De. (1996). *The measurement of drivers' mental workload.* Ph.D. thesis.

Wickens. (1998). *An introduction to human factors engineering.* New York [etc.] : Longman.