# Predictive modelling for sports performance improvement and injury prevention

## Wearable data-driven solutions for performance assessment and injury risk identification in baseball

van der Graaff, L.

**DOI**
10.4233/uuid:755b7cc3-6103-40fa-bab8-63a50f1be359

**Publication date**
2024

**Document Version**
Final published version

**Citation (APA)**
van der Graaff, L. (2024). *Predictive modelling for sports performance improvement and injury prevention: Wearable data-driven solutions for performance assessment and injury risk identification in baseball*. [Dissertation (TU Delft), Delft University of Technology]. https://doi.org/10.4233/uuid:755b7cc3-6103-40fa-bab8-63a50f1be359

# Predictive modelling for sports performance improvement and injury prevention

Wearable data–driven solutions for performance assessment and injury risk identification in baseball pitching

# Predictive modelling for sports performance improvement and injury prevention

## Wearable data–driven solutions for performance assessment and injury risk identification in baseball pitching

### Dissertation

for the purpose of obtaining the degree of doctor
at Delft University of Technology
by the authority of the Rector Magnificus, prof. dr. ir. T.H.J.J. van der Hagen,
chair of the Board for Doctorates
to be defended publicly on
Wednesday 18 December 2024 at 15:00 o'clock

by

**Larisa VAN DER GRAAFF**

Master of Science in Physics and Computer Science Education,
University of Zagreb, Croatia
born in Zagreb, Croatia

This dissertation has been approved by the promotors.

Composition of the doctoral committee:

| | |
|---|---|
| Rector Magnificus, | chairperson |
| Prof. dr. H.E.J. Veeger | Delft University of Technology, *promotor* |
| Prof. dr. ir. F.H. van der Meulen | Vrije Universiteit Amsterdam, *promotor* |

*Independent members:*

| | |
|---|---|
| Prof. dr. F.C.T. van der Helm | Delft University of Technology |
| Prof. dr. ir. G. Jongbloed | Delft University of Technology |
| Prof. dr. E.A.L.M. Verhagen | Vrije Universiteit Amsterdam |
| Dr. J. Söhl | Delft University of Technology |
| Dr. M.J.M. Hoozemans | Vrije Universiteit Amsterdam |

The research in this dissertation is part of the research programme *Citius, Altius, Sanius* with project number P16-28, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

*Keywords:* baseball; prediction; Bayesian multilevel models; latent Markov models; sports injury; wearables
*Printed by:* Ipskamp Printing
*Cover by:* L. van der Graaff



ISBN 978-94-6366-975-7

An electronic copy of this dissertation is available at
`https://repository.tudelft.nl/`.

Dobro se dobrim vraća.

# CONTENTS

# SUMMARY


## ATHLETE MONITORING FOR INJURY PREVENTION

Continuous and prospective monitoring plays an essential role in managing the effects of a prescribed training schedule on an athlete's performance and health. Individual athletes may respond differently to a given training stimulus, and the training load required for adaptation may significantly differ between athletes despite similar training backgrounds. Athlete monitoring aims to maximize positive effects (fitness, performance) and minimize negative effects (injury, illness) of athletic training. The core of injury prevention consists of managing injury risk and appropriate training modifications that can reduce the likelihood of injury occurrence, ensuring consistent and safe sports participation. The challenge lies in providing an adequate training stimulus to enhance performance while keeping the injury risk low. Integrating the data with the domain knowledge asks for the development of new methods and models to feed these ingredients into real–time personalised advice for the athlete.

## INJURIES IN BASEBALL

The injuries commonly seen in baseball pitchers have been attributed to the effects of high levels of energy on the weakest links of the kinetic chain. Poor pitching mechanics and overuse of the pitching arm can negatively affect pitching performance and at the same time put the elbow joint at great risk of injuries. Monitoring pitching mechanics and understanding its effect on elbow load is therefore an important step towards the development of an "early warning system" for safe and efficient pitching. However, monitoring pitching mechanics on the field during practice and competitions is challenging due to the rapid full-body nature of the pitching movement. The PitchPerfect is a multi–sensor system suitable for on-field measurements of pitching mechanics. It captures the (inter)segmental rotation and timing within the kinetic chain essential for performance enhancement and injury prevention.

## INTEGRATED PREDICTIVE MODELS

To establish the relevance of kinematic data for baseball pitchers, it is necessary to develop the translation process between the collected data and personalised information provided to the pitcher. Kinematic data collected through repeated measurements have inherited hierarchical structure where observations are nested within the individual athletes. To account for the similarities within individuals and simultaneously allow the gradation of differences between them, the thesis demonstrates the application of multilevel modelling on data from repeated measurements to a provide tailor-made prediction of performance and health outcomes of interest. The thesis presents a novel approach to individualised performance (chapter 2 and chapter 3) and injury risk prediction (chapter 4 and chapter 5). The methods proposed in each chapter offer solutions for dealing with the data often collected in sports.

## OUTCOMES OF THE THESIS

**Chapter 2** presents a Bayesian multilevel model for individualised ball velocity prediction based on pitching kinematics recorded with wearable sensors. The chapter investigates the added value of individuality to the predictive performance of the developed model. The aim of this study was to predict a ball velocity in baseball pitching such that prediction is tailored to the individual pitcher. The proposed method included the pitcher's body segment rotation, which determines his technique, and the pitcher's height which displays individuality in imparted velocity to a ball. We used multilevel modelling to develop three models with different predictors and examined their predictive performance. By comparing developed models, we investigated the added value of individuality to ball velocity prediction.

**Chapter 3** demonstrates the novel application of machine learning for binary and multiclass classification of pitch types based on wearable sensor data input. The study aimed to establish a methodology for pitch type classification based on biomechanical input from wearables. We used pelvis and trunk peak angular velocity and separation time between them as input and evaluated the performance of five machine learning classifiers in the binary and multiclass classification task. The Naive Bayes algorithm showed the best performance in classifying Fastballs with an accuracy of 71%. Furthermore, in the classification of pitch types as Fastball, Curveball or Change-up, the Random Forest algorithm performed the best with an average accuracy of 61.3% over those three pitch types.

**Chapter 4** presents a Bayesian multilevel model for the external valgus torque prediction, used as a proxy of elbow load, based on (inter)segmental rotation in baseball pitching. The model provides an individualised estimation of the elbow loading based on the pitcher's kinematics during every pitch. This study shows promising results of Bayesian hierarchical models in predicting the external valgus torque based on (inter)segmental rotation in fastball pitching. The results show that it is possible to predict the elbow external valgus torque based on the pelvis and trunk kinematics and separation time. Such an approach allows individualised prediction of the external valgus torque for each pitcher, which has a great practical advantage compared to group-based predictions regarding injury assessment and injury prevention.

**Chapter 5** proposes the multistate injury framework where a latent Markov model is used to predict injury risk. The model is used to analyse longitudinal panel data derived from repeated administration of the OSTRC questionnaire and athletes' time-varying weekly training exposures. The application of latent Markov models allows us to estimate the optimal number of injury states and the influence of included personal characteristics and performance measures on the transition between those states over time. Furthermore, we show that it is possible to predict the injury risk in the form of a probability for occupying each injury state based on different training scenarios.

## GENERAL CONCLUSIONS AND RECOMMENDATIONS

Sport-related injuries occur due to a complex interaction of many internal and external risk factors gathered in a pattern of either positive adaptation (increased fitness), or negative adaptation (injury). The repetitive nature of the high-speed full-body pitching movement exposes the pitcher's elbow to high loads. This thesis illustrates a novel approach to individualised injury risk prediction that accounts for the dynamics of the injury development process. The integration of advanced monitoring techniques plays an important role in the pursuit of high-level sports performance. The utilization of wearable sensors serves that purpose. It allows continuous athlete assessment and provides feedback on the relevant health and performance metrics in real–time. The methods established in the thesis offer solutions for dealing with different quality, time scale and hierarchical data structures collected with high-end wearable sensors, self-reported questionnaires and motion capture systems. Integration of the available data from different sources and implementation of the statistical models that can translate them to the relevant outcome provides actionable insights for performance improvement and injury prevention. Adding these statistical methods is a chance for training and injury-prevention programs to continue to improve.

# SAMENVATTING


## MONITORING VOOR BLESSUREPREVENTIE

Continue en prospectieve monitoring speelt een essentiële rol bij het managen van de effecten van een voorgeschreven trainingsschema op de prestaties en gezondheid van een sporter. Individuele sporters kunnen verschillend reageren op een bepaalde trainingsprikkel. De trainingsbelasting die nodig is voor vooruitgang kan aanzienlijk verschillen tussen sporters, zelfs tussen sporters met vergelijkbare trainingsachtergronden. Het doel van het monitoren van de training van sporters is om de positieve effecten (fitheid, prestaties) te maximaliseren en de negatieve effecten (blessures, ziekten) te minimaliseren. De kern van blessurepreventie bestaat uit het managen van blessurerisico en het aanbrengen van de juiste adaptaties in de training, die de kans op het ontstaan van blessures kunnen verkleinen en een consistente en veilige sportdeelname waarborgen. De uitdaging is om met een adequate trainingsprikkel de prestaties te verbeteren terwijl het blessurerisico laag blijft. Het integreren van gegevens met specifieke kennis uit het domein vraagt om de ontwikkeling van nieuwe methoden en modellen die deze elementen kunnen omzetten in real-time gepersonaliseerd advies voor de sporter.

## BLESSURES IN HONKBAL

Blessures die vaak voorkomen bij honkbalpitchers worden toegeschreven aan de effecten van een hoge belasting op de zwakste schakels in de kinetische keten. Slechte techniek en een overmatig gebruik van de werparm kunnen niet alleen de pitchprestaties negatief beïnvloeden, maar heeft ookeen groot risico op blessures in het elleboogggewricht. Het monitoren van de werptechniek en inzicht in het effect ervan op de belasting van de elleboog is daarom een belangrijke stap in de ontwikkeling van een zogenaamd "waarschuwingssysteem"voor veilig en efficiënt pitchen. Het monitoren van de werptechniek tijdens trainingen en wedstrijden is echter een hele uitdaging vanwege de snelle bewegingen van het lichaam die bij het pitchen komt kijken. PitchPerfect is een multisensor systeem dat geschikt is voor metingen van de werptechniek op het veld. Het systeem legt de (inter)segmentale rotatie en timing van de kinetische keten vast, essentieel voor prestatieverbetering en blessurepreventie.

## GEÏNTEGREERDE VOORSPELLENDE MODELLEN

Om de verzamelde data relevant te maken moet deze vertaald worden naar gepersonaliseerde informatie die aan een sporter verstrekt kan worden. Kinematische data die door middel van herhaalde metingen worden verzameld, hebben een hiërarchische structuur waarin observaties zijn genest binnen de individuele sporters. Om rekening te houden met de overeenkomsten binnen individuen en tegelijkertijd de gradaties van verschillen tussen hen mogelijk te maken, laat dit proefschrift de toepassing zien van multilevel modellering op data van herhaalde metingen, om op maat gemaakte voorspellingen te bieden voor prestaties en gezondheidsuitkomsten. Het proefschrift presenteert een nieuwe aanpak voor geïndividualiseerde prestatievoorspelling (hoofdstuk 2 en hoofdstuk 3) en blessurerisico voorspelling (hoofdstuk 4 en hoofdstuk 5). De in elk hoofdstuk voorgestelde methoden bieden oplossingen voor het omgaan met data die vaak in de sport worden verzameld.

## RESULTATEN VAN HET PROEFSCHRIFT

**Hoofdstuk 2** presenteert een Bayesiaans multilevel model voor het individueel voorspellen van balsnelheid op basis van de werpbeweging die is geregistreerd met draagbare sensoren. In dit hoofdstuk wordt de toegevoegde waarde van het personificeren voor de voorspellende prestaties van het ontwikkelde model onderzocht. Het doel van deze studie was om de balsnelheid bij honkbal pitchen te voorspellen op een manier die is afgestemd op de individuele pitcher. De voorgestelde methode omvat de rotatie van lichaamssegmenten van de pitcher, die zijn techniek bepaalt, en de lengte van de pitcher, die de individuele bijdrage aan de snelheid van de bal weergeeft. We gebruikten multilevel modellering om drie modellen met verschillende voorspellers te ontwikkelen en onderzochten de voorspellende prestaties. Door de ontwikkelde modellen te vergelijken onderzochten we de toegevoegde waarde van personificatie voor de voorspelling van de balsnelheid.

**Hoofdstuk 3** toont de nieuwe toepassing van machine learning voor binaire en multinomiale classificatie van verschillende worpen op basis van gegevens van draagbare sensoren. Het doel van de studie was het ontwikkelen van een methodologie voor de classificatie van verschillende worpen op basis van biomechanische input. We gebruikten de maximale hoeksnelheid van bekken en romp en de tijdsduur tussen deze twee als invoer en evalueerden de prestaties van vijf machine-learning classificatiemodellen voor de binaire en multinomiale classificatietaak. Het Naive Bayes-algoritme leverde de beste prestaties in het classificeren van Fastballs, met een nauwkeurigheid van 71%. Voor het classificeren van de worpen Fastball, Curveball of Change-up, presteerde het Random Forest-algoritme het best, met een gemiddelde nauwkeurigheid van 61,3%.

**Hoofdstuk 4** presenteert een Bayesiaans multilevel model voor de voorspelling van het externe valgus moment op basis van (inter)segmentale rotatie bij pitchen. Het model biedt een geïndividualiseerde schatting van de elleboogbelasting op basis van de kinematica van de pitcher tijdens elke pitch. Deze studie toont veelbelovende resultaten van Bayesiaanse hiërarchische modellen bij het voorspellen van het externe valgus moment op basis van (inter)segmentale rotatie. De resultaten laten zien dat het mogelijk is om het externe valgus moment in de elleboog te voorspellen op basis van de maximale rotatiesnelheid van de bekken en de romp en de intersegementale timing. Deze aanpak maakt een geïndividualiseerde voorspelling voor elke pitcher individueel mogelijk, wat een groot praktisch voordeel biedt ten opzichte voorspellingen gebaseerd op een groepsgemiddelde.

**Hoofdstuk 5** beschrijft een kader waarin een latent Markovmodel wordt gebruikt om blessurerisico te voorspellen aan de hand van een meervoudige status. Dit model wordt toegepast voor de analyse van longitudinale data verkregen uit OSTRC-vragenlijsten. Door gebruik te maken van latente Markovmodellen kunnen we het optimale aantal blessurestatus schatten en de invloed van persoonlijke kenmerken en prestatiemaatstaven op de overgang tussen deze status over de tijd, in kaart brengen. Bovendien tonen we aan dat het mogelijk is om het blessurerisico te voorspellen afhankelijk van verschillende trainingsscenario's in de vorm van een kans op het bereiken of behouden van een bepaalde status.

## ALGEMENE CONCLUSIES EN AANBEVELINGEN

Sportblessures ontstaan door een complexe interactie van vele interne en externe risicofactoren. De herhalende, snelle beweging van een volledige worp stelt het ellebooggewricht van de pitcher bloot aan hoge belastingen. Dit proefschrift illustreert een nieuwe aanpak voor de individuele voorspelling van blessurerisico's waarin rekening wordt gehouden met de dynamiek van het blessureontwikkelingsproces. Het gebruik van draagbare sensoren ondersteunt dit doel door continu atleten te monitoren en real-time feedback te geven op relevante gezondheids- en prestatiefactoren. De methoden ontwikkeld in dit proefschrift bieden oplossingen voor het omgaan met data van verschillende kwaliteit, tijdsschalen en hiërarchische structuren, verzameld met hoogwaardige draagbare sensoren, zelfrapportagevragenlijsten en motion-capture-systemen. Integratie van beschikbare data uit diverse bronnen, en de implementatie van statistische modellen die deze data omzetten naar relevante uitkomsten biedt bruikbare inzichten voor prestatie verbetering en blessurepreventie. Het toepassen van deze statische methodes is een kans om de trainings- en preventieprogramma's verder te verbeteren.

# SAŽETAK


## PRAĆENJE SPORTAŠA U SVRHU PREVENCIJE OZLJEDA

Kontinuirano i prospektivno praćenje sportaša ima važnu ulogu u kontroli posljedica propisanog programa treninga na izvedbu i zdravlje sportaša. Svaki sportaš reagira na jedinstven način na propisani trening, a opterećenje koje je potrebno za njegovu adaptaciju može se značajno razlikovati od sportaša do sportaša bez obzira na sličnu fizičku pripremljenost. Praćenje sportaša nastoji povećati pozitivne utjecaje treninga (fizička priprema, izvedba) te smanjiti njegove negativne posljedice (ozljeda, bolest). Srž prevencije ozljeda u sportu sastoji se od upravljanja rizicima od ozljeda te prilagodbe treninga pojedincu što omogućava smanjenje vjerojatnosti od zadobivanja ozljeda i osigurava kontinuirano i sigurno bavljenje sportom. Najveći izazov u cijelom procesu leži u pronalasku odgovarajućeg opterećenja koje će potaknuti poboljšanje sportske izvedbe uz smanjen rizik od ozljeda. Integracija podataka o sportskoj izvedbi i zdravlju i znanja iz područja sportske medicine i treninga zahtijeva uporabu novih metoda i matematičkih modela koji će nam omogućiti da na temelju prikupljenih podataka pružimo sportašima personalizirani savjet u stvarnom vremenu.

## OZLJEDE U BEJZBOLU

Najčešće ozljede kod bejzbol bacača nastaju uslijed utjecaja visoke razine energije na najslabije karike u kinetičkom lancu. Nepravilna bacačka tehnika i preopterećenje ruke bacača može imati negativan utjecaj na izvedbu bejzbol igrača te istovremeno dovesti zglob lakta u rizik od ozljede. Praćenje i mjerenje mehanike bacača i razumijevanje njenog efekta na opterećenje lakta je važan korak ka razvoju sustava za ranu detekciju rizika koji bi osigurao igračima sigurnu i efikasnu sportsku izvedbu. Međutim, bejzbol bacanje karakteriziraju pokreti cijelog tijela koji dostižu iznimne brzine, što čini mjerenje mehanike bacanja na samom bejzbol terenu tijekom treninga i mečeva velikim izazovom. Jedno od rješenja je PitchPerfect višesenzorski sustav prilagođen mjerenju mehanike bejzbol bacanja na samom terenu. Sustav omogućava mjerenje rotacije između segmenata kinetičkog lanca te njenog trajanja, što je od velike važnosti za poboljšanje izvedbe bacanja i prevenciju ozljeda.

## MATEMATIČKI MODELI

Kako bi se utvrdila relevantnost mjerenja kinematike bejzbol bacača, potrebno je razviti proces prevođenja između prikupljenih podataka i personaliziranih savjeta koji se daju bacaču. Kao posljedica ponovljenih mjerenja kinematičkih veličina tijekom sportske izvedbe, prikupljeni podaci imaju hijerarhijsku strukturu u kojoj su pojedinačna mjerenja ugniježđena unutar individualnih sportaša. Kako bi se uzele u obzir sličnosti unutar pojedinaca i istovremeno omogućilo gradiranje razlika među njima, u ovoj doktorskoj dizertaciji se demonstrira primjena višerazinskog modeliranja na podacima iz ponovljenih mjerenja kako bi se pružilo individualizirano predviđanje mjera koje opisuju sportsku izvedbu i zdravstvene ishode od interesa. Doktorska dizertacija predstavlja novi pristup individualiziranoj sportskoj izvedbi (poglavlje 2 i poglavlje 3) i predviđanju rizika od ozljeda (poglavlje 4 i poglavlje 5). Metode predložene u svakom poglavlju nude rješenja za postupanje s podacima koji se često prikupljaju u sportu.

## ISHODI DOKTORSKE DIZERTACIJE

**Drugo poglavlje** predstavlja Bayesov višerazinski model za individualizirano predviđanje brzine lopte na temelju kinematike bacanja mjerene nosivim senzorima. Ovo poglavlje istražuje koja je dodana vrijednost individualnog pristupa u prediktivnoj izvedbi primijenjenog matematičkog modela. Cilj ovog rada bio je predvidjeti brzinu lopte u bejzbol bacanju tako da je ishod prilagođen pojedinačnom bacaču. Predložena metoda uključivala je rotaciju segmenata tijela bacača, koja određuje njegovu tehniku, i visinu bacača koja predstavlja individualnu karakteristiku bacača u brzini koja se daje lopti. Koristili smo višerazinsko modeliranje za razvoj tri modela s različitim prediktorima i ispitali njihovu prediktivnu izvedbu. Uspoređujući razvijene modele, istražili smo dodanu vrijednost individualnog pristupa u predviđanju brzine lopte.

**Treće poglavlje** demonstrira novu primjenu strojnog učenja za binarnu i višeklasnu klasifikaciju tipova bejzbol bacanja na temelju podataka prikupljenih nosivim senzorima. Cilj rada bio je uspostaviti metodologiju za klasifikaciju vrste bejzbol bacanja na temelju biomehaničkih mjerenja nosivih senzora. Koristili smo vrijednost kutne brzine zdjelice i trupa i vrijeme razdvajanja između njih kao ulazne podatke i procijenili izvedbu pet klasifikatora strojnog učenja u zadatku binarne i višeklasne klasifikacije. Naivni Bayesov algoritam pokazao je najbolju izvedbu u klasifikaciji "Fastball" bacanja s točnošću od 71%. Nadalje, u klasifikaciji vrsta bacanja "Fastball", "Curveball" ili "Change-up", algoritam Random Forest pokazao se najboljim s prosječnom točnošću od 61,3% u odnosu na ta tri tipa bacanja.

**Četvrto poglavlje** predstavlja Bayesov višerazinski model za predviđanje vanjskog valgus momenta, koji se koristi kao zamjena za opterećenje lakta, na temelju (inter)segmentalne rotacije u bejzbol bacanju. Model pruža individualiziranu procjenu opterećenja lakta na temelju kinematike bacača tijekom svakog bacanja. Ovaj rad pokazuje obećavajuće rezultate Bayesovih višerazinskih modela u predviđanju vanjskog valgus momenta na temelju (inter)segmentalne rotacije u "Fastball" tipu bejzbol bacanja. Rezultati pokazuju da je moguće predvidjeti vanjski valgus moment lakta na temelju kinematike zdjelice i trupa te njihova vremena odvajanja. Takav pristup omogućuje individualizirano predviđanje vanjskog valgus momenta za svakog bacača, što ima veliku prednost u praktičnom djelovanju u usporedbi s predviđanjima na temelju podataka o populaciji.

**Peto poglavlje** predlaže novu metodologiju za predviđanje rizika od sportskih ozljeda za svakog pojedinačnog sportaša koja se temelji na primjeni latentnih Markovljevih modela. Model se koristi za analizu longitudinalnih podataka prikupljenih "Oslo Sports Trauma Research Center" (OSTRC) upitnikom i tjednog izlaganja različitim tipovima treninga. Primjena latentnih Markovljevih modela omogućuje nam procjenu optimalnog broja različitih stanja ozljede te utjecaj osobnih karakteristika sportaša i mjera izvedbe na prijelaze između tih stanja tijekom vremena. Nadalje, pokazujemo da je moguće predvidjeti rizik od ozljede u obliku vjerojatnosti svakog pojedinog stanja ozljede na temelju različitih programa treninga.

## OPĆI ZAKLJUČCI I PREPORUKE

Sportske ozljede nastaju zbog složene interakcije mnogih unutarnjih i vanjskih čimbenika rizika okupljenih u obrazac pozitivne prilagodbe (povećana kondicija) ili negativne prilagodbe (ozljeda). Repetitivna priroda brzih pokreta bacanja koje uključuju aktivaciju cijelog tijela izlaže lakat bejzbol bacača velikim opterećenjima. Ova doktorska dizertacija ilustrira novi pristup individualiziranom predviđanju rizika od ozljeda koji uzima u obzir dinamiku procesa razvoja ozljede. Integracija naprednih tehnika praćenja sportaša igra važnu ulogu u postizanju visokih sportskih rezultata. Nosivi senzori omogućavaju kontinuirano praćenje sportaša i daju povratnu informaciju o relevantnim pokazateljima zdravlja i sportske izvedbe u stvarnom vremenu. Metode utvrđene u doktorskoj dizertaciji nude rješenja za analizu podataka prikupljenih korištenjem nosivih senzora, upitnika za samoprocjenu i sustava za snimanje pokreta. Integracija dostupnih podataka iz različitih izvora i implementacija statističkih modela pruža uvid u moguće načine poboljšanja sportske izvedbe i prevencije ozljeda u raznim sportskim disciplinama te doprinijeti unaprijeđenju treninga i programa prevencije ozljeda.

# 1

# INTRODUCTION

## 1.1. BACKGROUND

Knowledge about personal performance is a stimulating factor that contributes to long-term sports engagement and reaching higher performance levels. The integration of advanced monitoring techniques in an athlete's routine hereby plays an essential role.

The aim of athlete monitoring is to maximize positive effects (fitness, performance) and minimize negative effects (injury, illness) of athletic training [1]. Individual athletes may respond differently to a given training stimulus, and the training load required for adaptation may significantly differ between athletes despite similar training backgrounds [2]. Therefore, continuous and prospective monitoring is crucial for managing the effects of a prescribed training schedule on an athlete's performance and health. This includes longitudinal monitoring of physiological and biomechanical load, athlete performance, and well-being [3].

Advancements in sensor technology have increased the list of wearable devices available on the market for athlete monitoring. These devices enable athletes of all performance levels to access their positional, biometric, and biomechanical data and share them with their peers and teams. Sensor technology integrated into watches, sleeves, straps, and fabrics allows data collection on and off the pitch. This creates a data stream fed into built-in algorithms increasingly providing health and performance metrics to athletes in real–time. Such data are often used subjectively to modify training schedules resulting in athletes training "too much too soon". If training loads greatly exceed an athlete's load capacity or progress too rapidly, the risk of injury and illness will increase and directly affect performance through a reduced ability to perform [4–6].

The occurrence of sport-related injuries is a result of a complex interaction between multiple factors (Figure 1.1) gathered in a pattern of either positive (i.e. increased fitness) or negative adaptation (i.e. injury) [7, 8]. Each athlete has its own set of internal risk factors (e.g. age, gender, fitness, and history of injury) that may be minimized if an athlete adapts when exposed to an external risk factor (e.g. training load, contact with another player, training intensity) or a potentially injury-prone situation without sustaining an injury [8]. The core of injury prevention consists of managing injury risk and appropriate training modifications that can reduce the likelihood of injury occurrence, ensuring consistent and safe sports participation. The challenge lies in providing an adequate training stimulus to enhance performance while keeping the injury risk low [9].

The combination of health and performance data leads to a better understanding of that complex interaction but can be hindered due to the different data types. In a time of rapid data production, it is essential to create a link from data collection to direct action on the field. There is

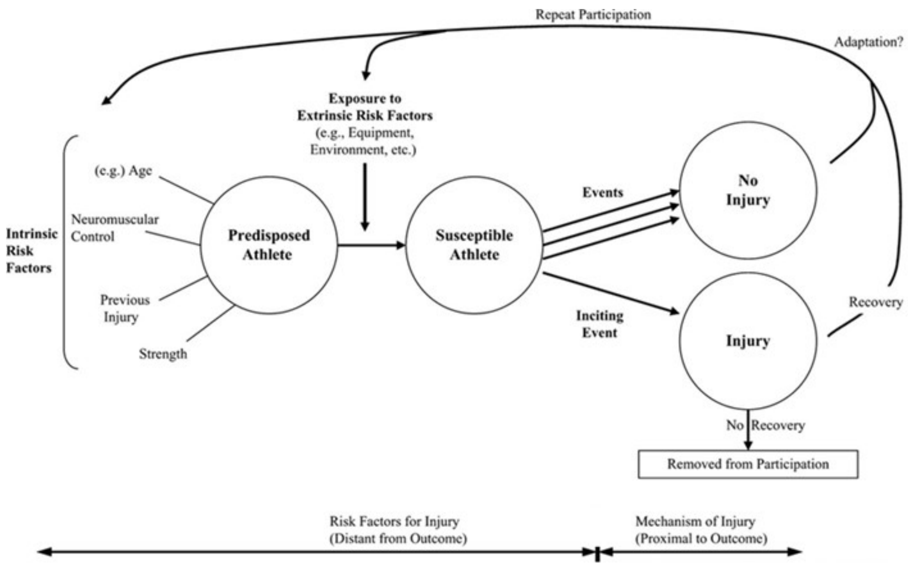


Figure 1.1.: Dynamical model adopted from Meeuwisse (2007).

a need for more objective use of the available health and performance data to support decision-making in training modifications and injury risk management. Integrating the data with domain knowledge asks for the development of new methods and models to feed these ingredients into real–time personalised advice for the athlete.

## 1.2. PROJECT

This thesis has been written within the NWO Perspectief program Citius, Altius, Sanius. The program aimed to strengthen the information chain from sensor technology via data science to informative feedback applications that will allow athletes of all levels of participation to improve their sports performance without sustaining injuries. The focus of the thesis is on data science, linking the data collected with wearable sensors with tailored feedback on the health and performance of an individual athlete. The thesis investigates wearable data-driven solutions for performance assessment and injury risk identification in overhead sports, namely baseball pitching.

Baseball has a long history of game data utilization focusing on predicting outcomes and developing winning teams. The expanse of sabermetrics in the early 2000s resulted in baseball being nowadays one of the biggest data-driven sports worldwide. The performance–driven mindset creates a fertile ground for the implementation of novel technologies for direct feedback on player performances. However, the demands of high performance are making baseball pitchers highly susceptible to pitching injuries. The repetitive and dynamic nature of baseball pitching exposes pitchers to a high load. Most of the injuries arise from the repetitive application of high forces and/or torques to vulnerable tissue [10]. Therefore, there is a need for the development of an early warning system that can provide feedback on injury risk and allow pitchers to modify their training schedule before overloading occurs.

## 1.3. BASEBALL PITCHING

### 1.3.1. PITCHING PERFORMANCE

Baseball pitching is the act of throwing a baseball toward home plate to start the game. The overhead pitching motion is a sequence of body movements starting with the legs, continuing through the pelvis and trunk, and culminating with a whip-like action of the arm to propel the ball toward home plate [11]. The sequential activation of body segments to efficiently impart velocity to the ball follows the summation of speed principle, also known as the kinetic chain. Each segment starts as the adjacent proximal segment reaches top speed, culminating with the top speed of the most distal segment [12].

Optimization of the kinetic chain can result in an improvement of the ball velocity [11] which increases the pitcher's chances of success. By throwing fast pitchers reduce the decision time of the batter whether to strike the ball or not [13] and restricts the ability of the runners to advance bases and score runs [14]. A successful pitcher alters the velocity and trajectory of the ball to keep batters off balance and discourage their anticipation of a particular pitch type. Commonly used pitch types are the fastball, changeup, curveball, and slider. To obtain variation, the pitcher manipulates the grip on the ball at the release point. This causes the seam to catch the air differently and changes the ball's trajectory. Such kinematic, kinetic, and temporal variations in the throwing motion are related to improved velocity and force generation. Optimization of these parameters allows for efficient and consistent transfer of energy from proximal to distal components. Pelvis and trunk kinematics [11, 15, 16] hereby play an essential role. Optimal proximal-to-distal timing between the pelvis and trunk results in maximal ball velocity at release [11, 12, 17]. If this kinematic sequencing is not optimal, energy is dissipated into the upper extremity which results not only in a lower ball velocity [11, 18] but also in a potentially increased risk of injuries.

### 1.3.2. INJURIES

The injuries commonly seen in baseball pitchers have been attributed to the effects of high levels of energy on the weakest links of the kinetic chain, usually more distally at the level of the pitching arm or elbow [19]. Repetitive high–speed movement and excessive ranges of motion expose the elbow joint to intrinsic and extrinsic loads [11, 20, 21] related to the occurrence of overuse injuries [22].

Poor pitching mechanics [23] and overuse of the pitching arm can negatively influence pitching performance and at the same time put the elbow joint at great risk of injuries [10, 24]. Knowledge of the kinetic chain and key kinematic parameters underlying the throwing motion can improve pitching technique that can assist in performance enhancement, return to sport, and injury prevention [11]. Monitoring pitching mechanics and understanding its effect on elbow load is therefore the essential step towards the development of an “early warning system” for safe and efficient pitching.

Despite the knowledge about throwing biomechanics and injury-prone structures, the useful guidelines for the prevention of overload injuries or the optimal throwing technique are still lacking. One of the reasons is absence of a measurement system suitable for capturing fast–pitching movement and intersegmental timing without obstructing the pitcher’s performance.

### 1.3.3. WEARABLE TECHNOLOGY

Monitoring pitching mechanics on the field during practice and competitions is challenging. The extremely rapid pitching motion is not easy to capture without high–speed cameras. Furthermore, the requirement for out–of–lab motion capture is an easy-to-use device that does not obstruct the pitcher’s movement, but at the same time can record the high velocities of (inter)segmental rotations.

PitchPerfect (PitchPerfect, The Netherlands) is a multi-sensor system consisting of two synchronized 3-DOF inertial measurement units (IMU). Each IMU includes a gyroscope that can measure angular velocities up to 2000 deg/s. With optimal pelvis and trunk peak angular velocities recorded in previous studies ranging 600 – 900 deg/s and 900 – 1300 deg/s [25, 26] respectively, the PitchPerfect sensor system is suitable for on-field measurements of pitching mechanics. The outcome of the PitchPerfect system consists of pelvis and trunk peak angular velocities and separation time (i.e. the time interval between the peak angular velocities of pelvis and trunk), capturing the (inter)segmental rotation and timing within the kinetic chain essential for performance enhancement [18, 20, 27] and injury prevention [28].

## 1.4. DATASETS AND METHODOLOGY

The project aimed to use an integrated approach in which longitudinal data collection – algorithm development – practical application would form the central line. To do so, a cohort study was set up at the start of the project in 2019 to facilitate the daily data collection at the KNBSB (The Royal Dutch Baseball and Softball Federation) using wearable sensors to measure baseball players from youth selections (60 players) on every training. Additionally, "traditional" in-lab biomechanical studies using the Vicon motion capture system aimed to strengthen the link with clinical applications and to contribute to the data production necessary for the development of predictive models. To investigate the effects of pitching biomechanics on the development of elbow injuries, the cohort study also would have included weekly data collection of sport-related health problems using the OSTRC questionnaire on health problems integrated in the OptiForm health monitoring system developed by VUmc. However, due to the COVID-19 pandemic, the cohort study had to be cancelled, and longitudinal data were not collected. For this thesis, this resulted in a "plan B" in which the development of algorithms focusing on the prediction of direct performance measures and biomechanical variables from wearable sensor data became the central research theme.

### 1.4.1. LINK BETWEEN REPEATED MEASUREMENT DATA AND INDIVIDUALISED PREDICTIVE MODELLING

Assuming mostly linear dependencies between the biomechanical input variables and performance and health outcomes, classical linear regression seems like a natural choice of statistical analysis. However, if the data vary by group (in this case athlete), a natural thing to do is to predict the desired outcome for an in- or out-of-sample individual athlete. This is not such a straightforward task for a classical regression and if it includes group effects, it still does not have an automatic way of predicting for a new group. To account for the similarities within individuals and at the same time allow for the gradation of differences between them, the thesis demonstrates the application of multilevel modelling on data from repeated measurements to provide tailor-made predictions of performance and health outcomes of interest.

#### FROM CLASSICAL REGRESSION TO MULTILEVEL MODELLING

When dealing with repeated measurements data and their inherent hierarchical structure, it is reasonable to start by fitting simple classical regressions and work our way up to a multilevel model. This section offers two possible starting points, complete pooling, and no pooling, and compares them to multilevel modelling. An example dataset has been generated to highlight the strengths and limitations of the three approaches (complete pooling, no pooling, multilevel modelling) within the Bayesian modelling toolbox in the analysis of repeated measurements data.

We used the dataset from the annual Cherry Blossom Ten Mile race in Washington, D.C. [29]. The data contain running times and ages for a subset of 36 runners that participated in the race in multiple years. With running times and age being measured repeatedly for the same individuals over multiple years in which the race took place, the data have a hierarchical structure where measured running times and age on an individual–level are nested within individual runners on a group–level. Motivated by the examples in [30] and [31], this toy example aims to compare the performance of a complete pooling, no pooling and multilevel Bayesian models in the analysis of repeated measurements data through estimation of running times for an individual runner based on their age.

### COMPLETE POOLING MODEL

The subset of 36 runners, with a mean age of 55.0 ± 2.3 years, participated multiple times over the years in the annual Cherry Blossom race. With each participation, their running time and age have been recorded. Let's assume we want to estimate the running time based on age for each runner by using the complete pooling model, also known as a classical regression model.

The complete pooling model pools data across all participants, assuming that all observations belong to the same individual. It involves fitting a single linear regression to the full dataset under the assumption that the relationship between age and running time is the same for all observed runners. The outcome of the complete pooling model visualized with blue solid lines in Figure 1.2, suggests that all runners as they get older will cross the finish line of the race within almost the same running time as in previous years. The increase in running time is 0.30 min for each year of age and it is the same for all runners. If we focus on observed running times (the black dots in Figure 1.2) for the four selected runners, it is visible that the running times of each individual runner increase with age and at a more rapid rate than suggested by the complete pooling model. Furthermore, it is clear that the increase in running times with age is different for every runner (e.g. the running time of Runner 20 increases more rapidly with age compared to Runner 5).

The reason why complete pooling fails to capture differences between runners is because the model ignores any hierarchical structure or group-level variance by assuming that the relationship between predictors and the outcome is the same across all groups. This approach is appropriate when the data come from a single population or when group-level differences are negligible. However, if there are substantial differences between groups, complete pooling fails to account for within-group correlations or varying group-level effects. Therefore we need to move away from the complete pooling approach and explore the models that account for variation between individuals.

#### NO POOLING MODEL

The alternative to a complete pooling modelling approach is no pooling. Unlike complete pooling, no pooling treats each group separately by fitting an independent model for each group without any shared parameters. No pooling includes group indicators and estimates the model classically.

The no pooling model estimates running time based on age for each individual runner separately and each individual has its own intercept and slope (the red dashed lines in Figure 1.2). This approach is appropriate when there are big differences between groups (i.e. runners) that clearly describe the relationship between predictors and outcomes for each specific group. However, when the number of observations within individuals is small, by using no pooling we are facing a risk of overfitting the data within each individual. In this case, the estimates from the no pooling model overstate the variation among individuals and tend to make the individuals runners seem more different than they actually are.

In the task of estimating running time based on age, the no pooling model performance for selected runners shown in Figure 1.2 seems to be a good choice, but can we use the same model also to predict the running time of a new runner from which we don't have any previous data? The answer is no. Due to the lack of pooling, the no pooling model does not benefit from the information available about other individuals, which can be important for identifying common patterns across individuals. The separately fitted models do not extend beyond individuals included in the initial dataset which means that it is not possible to use the no pooling approach to predict the running time for a new, out-of-sample runner of any age.

Understanding the strengths and limitations of complete pooling and no pooling is crucial for choosing the right modelling strategy based on the characteristics of the available data and formulated research questions. We can conclude that neither of these two approaches alone is fully adequate for the prediction task based on hierarchical data illustrated in this example. What we need is a compromise between complete pooling and no pooling approach in the form of a partial pooling model.

### MULTILEVEL MODEL

Multilevel models are also known as a partial pooling models. A multilevel model represents a compromise between two extremes: complete pooling, in which the group indicators are not included in the model, and no pooling, in which separate models are fit within each group. This approach borrows strength across groups with fewer observations while accounting for individual differences.

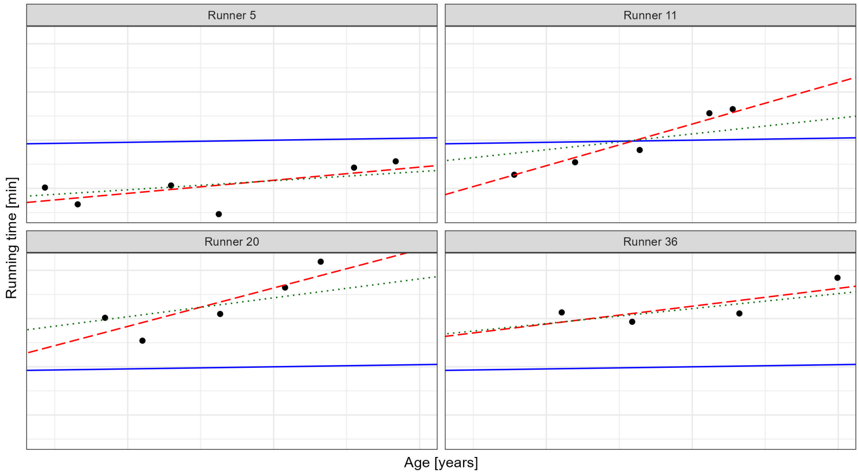


Figure 1.2.: The blue solid line, red dashed line and green dotted line show the complete pooling, no pooling and multilevel estimates respectively.

In the toy example, we fitted the varying–intercept, varying–slope multilevel model which means that every runner has its own intercept and slope. The outcome of the multilevel model is illustrated with the green dotted lines in Figure 1.2. A multilevel model is most effective when it is close to complete pooling, especially for the groups with a small sample size. In this setting, we can allow estimates to vary between groups while still estimating them precisely. For groups with a small sample size the multilevel estimate can be close to complete pooling and close to no pooling for groups with large sample sizes, resulting in good performance for both groups.

## 1.4.2. MODEL PERFORMANCE AND MODEL SELECTION

Evaluating the performance of a Bayesian multilevel model includes assessing the fit of the model to data and its predictive performance. The Posterior Predictive Check (PPC) is a method used to evaluate whether the model fits the data adequately. The PPC process includes

1

data simulation according to the fitted model by sampling from the posterior predictive distribution. This is the distribution of the outcome variable implied by the posterior distribution of the model parameters. The simulated data is then compared to the observed data to highlight the differences. We interpret the generated data as the data sample that we might collect tomorrow if the data collection process remains the same as it initially was. If the model fits the data well, the important features of the observed data should be replicated in the simulations as well [32]. Posterior predictive checks were used to test the performance of the model and visually inspect how much generated data samples match the observed ones [32, 33].

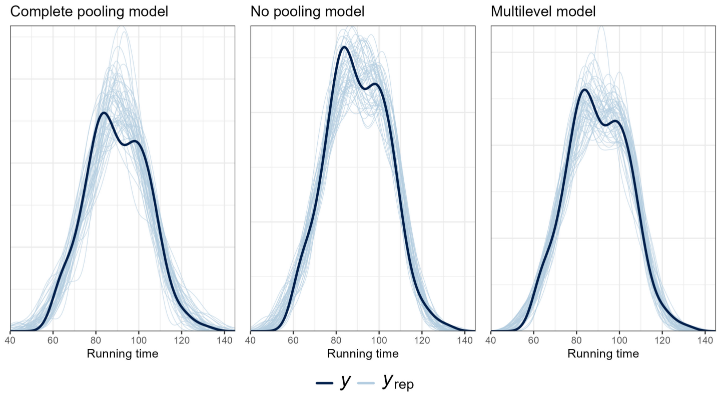


Figure 1.3.: Posterior predictive checks for complete pooling, no pooling and multilevel model. Posterior predictive checks compare the observed outcome variable $y$ to the average of simulated datasets $y_{\text{rep}}$ from the posterior predictive distribution for complete pooling, no pooling and multilevel model respectively.

Figure 1.3 shows the kernel density estimate of the observed data set $y$ (dark blue curve), with density estimates for 100 simulated data sets $y_{\text{rep}}$ drawn from the posterior predictive distribution (light blue curves) [33]. Even though visually no pooling and multilevel model both seem to fit the data well, it is clear that they can simulate new data that are more similar to the observed running time values than the complete pooling model can.

To select the model with the best predictive performance we used leave-one-out (LOO) cross-validation. Following the approach in [34], the predictive performance of a model is defined as the expected log-predictive density (*elpd*). Predictive performance is a useful quantity in assessing a single model. It can be estimated by training the model on all observations except one and then predicting the hold-out observation. This is then repeated for all *n* observations.

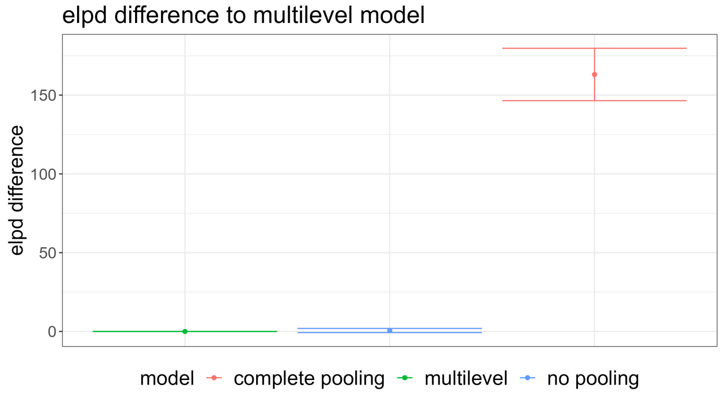


Figure 1.4.: Estimates of absolute expected log-predictive density (*elpd*) difference (dot) using leave-one-out cross-validation. Vertical error bars for each model indicates the standard error of the *elpd* difference estimates. The order on the x-axis follows the ranking starting with the model with best predictive performance on the left.

The expected log-predictive density (*elpd*) was a chosen measure of model fit and was subsequently used to compare models for model selection. The difference in *elpd* of the fitted complete pooling, no pooling and multilevel Bayesian models is shown in Figure 1.4. The ordering of the models reveals that the multilevel model shows the best predictive performance, and it is therefore the selected model.

### 1.4.3. PREDICTION

After assessing the performance of models and selecting the multilevel model as a preferred one, let's use the multilevel model to predict what will be the running time of five in-sample runners when they will be 60 years old.

The posterior predictive model (Figure 1.5) reflects the two sources of uncertainty in the prediction of the running times - the within–group sampling variability (we cannot perfectly predict runner's running time from their mean model) and posterior variability (the parameters defining the runner's relationship between running time and age are unknown and random). Due to the lack of data for the new runner, the third source of uncertainty that is reflected in its running time prediction is a between-group sampling variability (baseline speeds vary between runners) [30].

Based on previously observed trends, it was reasonable to expect

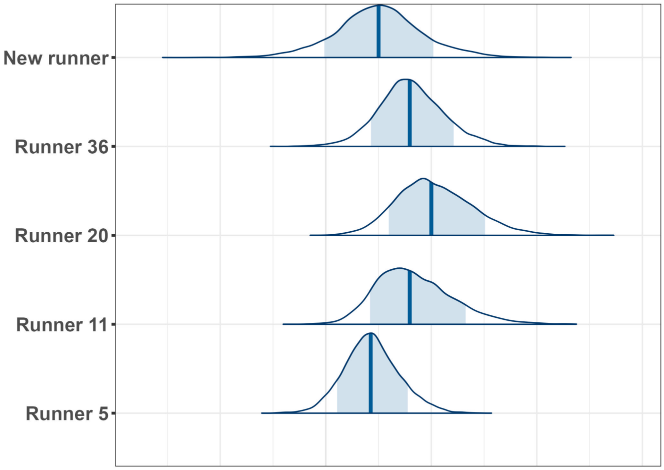


Figure 1.5.: Posterior prediction for the running times of five in-sample runners and one out-of-sample runner at the age of 60.

that at the age of 60 runner 20 will run slower than runner 5. This is also confirmed by the posterior expectation shown in Figure 1.5. The prediction for the new runner is somewhere in between those two extremes. Given that we have no prior information about the new runner, the default assumption is that its running time will be around the average running time of the population included in the analysis. The uncertainty in this assumption is reflected by the relatively wide posterior predictive model (Figure 1.5). Compared to the new runner, the posterior predictive model of other runners is narrower. In conclusion, we are more certain about how fast they will run the Cherry Blossom race at the age of 60 due to the availability of their running times from earlier years.

### 1.4.4. DEVELOPMENT OF INJURY RISK PROFILE

The thesis presents a novel approach to individualized performance (chapter 2, chapter 3) and injury risk prediction (chapter 4, chapter 5). The methods proposed in each chapter offer solutions for dealing with the data with hierarchical structure collected with high-end wearable sensors, self-reported questionnaires and motion capture systems. Integration of the available data from different sources and implementation of the statistical models that can translate them to relevant outcomes, serve as a base for the development of athlete's injury risk profile.

## 1.5. THESIS OUTLINE

Chapter 2 presents a Bayesian multilevel model for individualised prediction of the ball velocity based on pitching kinematics recorded with wearable sensors. The chapter investigates the added value of individuality to the predictive performance of the developed model [35].

Chapter 3 demonstrates the novel application of machine learning for binary and multiclass classification of pitch types based on wearable sensor data input [36].

Chapter 4 presents a Bayesian multilevel model for prediction of the external valgus torque, used as a proxy of elbow load, based on (inter)segmental rotation in baseball pitching. The model provides individualised estimation of the elbow loading based on the pitcher's kinematics during every pitch [37].

Chapter 5 proposes the multistate injury framework where a latent Markov model is used to predict injury risk. The model is used to analyse longitudinal panel data derived from repeated administration of the OSTRC questionnaire and athletes' time-varying weekly training exposures. The proposed approach enables the prediction of injury risk trajectories based on future training scenarios.

# REFERENCES


[1] T. J. Gabbett, G. P. Nassis, E. Oetter, J. Pretorius, N. Johnston, D. Medina, G. Rodas, T. Myslinski, D. Howells, A. Beard, *et al.* *The athlete monitoring cycle: a practical guide to interpreting and applying training monitoring data*. 2017.

[2] T. Soligard, M. P. Schwellnus, J.-M. Alonso, R. Bahr, B. Clarsen, H. P. Dijkstra, T. Gabbett, M. Gleeson, M. Hägglund, M. R. Hutchinson, *et al.* "How much is too much?(Part 1) International Olympic Committee consensus statement on load in sport and risk of injury". In: (2016).

[3] M. Schwellnus, T. Soligard, J.-M. Alonso, R. Bahr, B. Clarsen, H. P. Dijkstra, T. J. Gabbett, M. Gleeson, M. Hägglund, M. R. Hutchinson, *et al.* "How much is too much?(Part 2) International Olympic Committee consensus statement on load in sport and risk of illness". In: *British journal of sports medicine* 50.17 (2016), p. 1043.

[4] T. J. Gabbett. "Debunking the myths about training load, injury and performance: empirical evidence, hot topics and recommendations for practitioners". In: *British journal of sports medicine* (2018), bjsports–2018.

[5] E. Verhagen and T. Gabbett. *Load, capacity and health: critical pieces of the holistic performance puzzle*. 2019.

[6] T. J. Gabbett. "How much? How fast? How soon? Three simple concepts for progressing training loads to minimize injury risk and enhance performance". In: *Journal of orthopaedic & sports physical therapy* 50.10 (2020), pp. 570–573.

[7] N. F. Bittencourt, W. Meeuwisse, L. Mendonça, A. Nettel-Aguirre, J. Ocarino, and S. Fonseca. "Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept". In: *British journal of sports medicine* (2016).

[8] W. H. Meeuwisse, H. Tyreman, B. Hagel, and C. Emery. "A dynamic model of etiology in sport injury: the recursive nature of risk and causation". In: *Clinical journal of sport medicine* 17.3 (2007), pp. 215–219.

[9] M. Roe, S. Malone, C. Blake, K. Collins, C. Gissane, F. Büttner, J. C. Murphy, and E. Delahunt. "A six stage operational framework for individualising injury risk management in sport". In: *Injury epidemiology* 4 (2017), pp. 1–6.

[10] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews. "Baseball Pitching Biomechanics in Relation to Injury Risk and Performance". en. In: *Sports Health: A Multidisciplinary Approach* 1.4 (July 2009), pp. 314–320. issn: 1941-7381, 1941-0921. doi: 10.1177/1941738109338546. url: http://journals.sagepub.com/doi/10.1177/1941738109338546.

[11] S. T. Seroyer, S. J. Nho, B. R. Bach, C. A. Bush-Joseph, G. P. Nicholson, and A. A. Romeo. "The kinetic chain in overhand pitching: its potential role for performance enhancement and injury prevention". In: *Sports health* 2.2 (2010), pp. 135–146.

[12] C. A. Putnam. "Sequential motions of body segments in striking and throwing skills: Descriptions and explanations". en. In: *Journal of Biomechanics* 26 (Jan. 1993), pp. 125–135. issn: 00219290. doi: 10.1016/0021-9290(93)90084-R. url: https://linkinghub.elsevier.com/retrieve/pii/002192909390084R.

[13] J. Hay. *The biomechanics of sports techniques*. Prentice-Hall, 1978.

[14] G. Lehman, E. J. Drinkwater, and D. G. Behm. "Correlation of throwing velocity to the results of lower-body field tests in male college baseball players". In: *The journal of strength & conditioning research* 27.4 (2013), pp. 902–908.

[15] D. F. Stodden, G. S. Fleisig, S. P. McLean, S. L. Lyman, and J. R. Andrews. "Relationship of pelvis and upper torso kinematics to pitched baseball velocity". In: *Journal of applied biomechanics* 17.2 (2001), pp. 164–172.

[16] D. F. Stodden, G. S. Fleisig, S. P. McLean, and J. R. Andrews. "Relationship of biomechanical factors to baseball pitching velocity: within pitcher variation". In: *Journal of applied biomechanics* 21.1 (2005), pp. 44–56.

[17] G. S. Bullock, G. Menon, K. Nicholson, R. J. Butler, N. K. Arden, and S. R. Filbay. "Baseball pitching biomechanics in relation to pain, injury, and surgery: A systematic review". In: *Journal of Science and Medicine in Sport* 24.1 (2021), pp. 13–20.

[18] E. Van Der Graaff, M. ( Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. ( Veeger. "Timing of peak pelvis and thorax rotation velocity in baseball pitching". en. In: *The Journal of Physical Fitness and Sports Medicine* 7.5 (Sept. 2018), pp. 269–277. issn: 2186-8123, 2186-8131. doi: 10.7600/jpfsm.7.269. url: https://www.jstage.jst.go.jp/article/jpfsm/7/5/7_269/_article.

1

[19] A. W. Anz, B. D. Bushnell, L. P. Griffin, T. J. Noonan, M. R. Torry, and R. J. Hawkins. “Correlation of torque and elbow injury in professional baseball pitchers”. In: *The American journal of sports medicine* 38.7 (2010), pp. 1368–1374.

[20] S. L. Werner, M. Suri, J. A. Guido Jr, K. Meister, and D. G. Jones. “Relationships between ball velocity and throwing mechanics in collegiate baseball pitchers”. In: *Journal of shoulder and elbow surgery* 17.6 (2008), pp. 905–908.

[21] G. S. Fleisig, J. R. Andrews, C. J. Dillman, and R. F. Escamilla. “Kinetics of baseball pitching with implications about injury mechanisms”. In: *The American journal of sports medicine* 23.2 (1995), pp. 233–239.

[22] F. S. Chen, A. S. Rokito, and F. W. Jobe. “Medial elbow problems in the overhead-throwing athlete”. In: *JAAOS-Journal of the American Academy of Orthopaedic Surgeons* 9.2 (2001), pp. 99–113.

[23] J. Davis, O. Limpisvasti, D. Fluhme, K. J. Mohr, L. A. Yocum, N. S. ElAttrache, and F. W. Jobe. “The Effect of Pitching Biomechanics on the Upper Extremity in Youth and Adolescent Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 37.8 (Aug. 2009), pp. 1484–1491. issn: 0363-5465, 1552-3365. doi: `10.1177/0363546509340226`. url: `http://journals.sagepub.com/doi/10.1177/0363546509340226`.

[24] G. S. Fleisig, D. S. Kingsley, J. W. Loftice, K. P. Dinnen, R. Ranganathan, S. Dun, R. F. Escamilla, and J. R. Andrews. “Kinetic Comparison among the Fastball, Curveball, Change-up, and Slider in Collegiate Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 34.3 (Mar. 2006), pp. 423–430. issn: 0363-5465, 1552-3365. doi: `10.1177/0363546505280431`. url: `http://journals.sagepub.com/doi/10.1177/0363546505280431`.

[25] A. L. Aguinaldo, J. Buttermore, and H. Chambers. “Effects of Upper Trunk Rotation on Shoulder Joint Torque among Baseball Pitchers of Various Levels”. en. In: *Journal of Applied Biomechanics* 23.1 (Feb. 2007), pp. 42–51. issn: 1065-8483, 1543-2688. doi: `10.1123/jab.23.1.42`. url: `https://journals.humankinetics.com/view/journals/jab/23/1/article-p42.xml`.

[26] D. F. Stodden, S. J. Langendorfer, G. S. Fleisig, and J. R. Andrews. “Kinematic constraints associated with the acquisition of overarm throwing part I: step and trunk actions”. In: *Research quarterly for exercise and sport* 77.4 (2006), pp. 417–427.

[27] B. Dowling, C. Pearl, W. Laughlin, and G. Fleisig. “Relationship of pelvis and trunk kinematics to ball velocity in professional baseball pitchers”. In: *40th American Society of Biomechanics (ASB) annual meeting*. 2016.

[28] A. D. Cohen, E. J. Garibay, and M. J. Solomito. “The association among trunk rotation, ball velocity, and the elbow varus moment in collegiate-level baseball pitchers”. In: *The American journal of sports medicine* 47.12 (2019), pp. 2816–2820.

[29] M. Dogucu, A. Johnson, and M. Ott. *bayesrules: Datasets and Supplemental Functions from Bayes Rules! Book*. R package version 0.0.2.9000. 2021. url: https://github.com/bayes-rules/bayesrules.

[30] A. A. Johnson, M. Q. Ott, and M. Dogucu. *Bayes rules!: An introduction to applied Bayesian modeling*. Chapman and Hall/CRC, 2022.

[31] A. Gelman and J. Hill. *Data analysis using regression and multilevel/hierarchical models*. Cambridge university press, 2007.

[32] C. Muth, Z. Oravecz, and J. Gabry. “User-friendly Bayesian regression modeling: A tutorial with rstanarm and shinystan”. In: *The Quantitative Methods for Psychology* 14.2 (2018), pp. 99–119.

[33] J. Gabry, D. Simpson, A. Vehtari, M. Betancourt, and A. Gelman. “Visualization in Bayesian workflow”. In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 182.2 (2019), pp. 389–402.

[34] A. Vehtari, A. Gelman, and J. Gabry. “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”. In: *Statistics and computing* 27 (2017), pp. 1413–1432.

[35] L. Gomaz, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. “Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data”. en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: 10.3390/s21227442. url: https://www.mdpi.com/1424-8220/21/22/7442.

[36] L. Gomaz, C. Bouwmeester, E. van der Graaff, B. van Trigt, and D. Veeger. “Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors”. In: *Sensors* 23.23 (2023), p. 9373.

[37] L. Gomaz, B. van Trigt, F. van der Meulen, and D. Veeger. “Predicting elbow load based on individual pelvis and trunk (inter)segmental rotations in fastball pitching”. In: *Sports Biomechanics* (2024). doi: 10.1080/14763141.2024.2315230.

# 2

# INDIVIDUALISED BALL SPEED PREDICTION IN BASEBALL PITCHING BASED ON IMU DATA

**Larisa van der Graaff, DirkJan Veeger, Erik van der Graaff, Bart van Trigt, Frank van der Meulen**

Published in *Sensors* **2021** [1].

## 2.1. INTRODUCTION

Proper pitching mechanics play an important role in both success and health of baseball pitchers. In overhead pitching, the lower extremity and trunk generate and transfer energy to the upper extremity. The optimal sequential activation of body parts while pitching, known as the kinetic chain, can result in reduced elbow and shoulder stress and maximise pitching performance [2, 3]. On the other hand, poor mechanics can lead to increased loading of the elbow or shoulder, and increase the injury risk. Injuries of the throwing arm, such as the ones to the shoulder and elbow, are common in the overhead pitching motion of baseball. Major League Baseball pitchers are especially prone to injury because of the throwing velocities commonly seen approaching and even exceeding 100 mph. To create such high ball velocities, high energy levels pass through the components of the kinetic chain that affect the weakest links among them, especially the elbow [4]. Therefore, there is a need for assessment of the throwing technique that enables players to throw fast pitches in the strike zone without an overload.

Throwing velocity plays an important role in a success of a baseball game. Pitchers increase their chances for success by throwing faster and diminishing the hitter's decision time of whether or not to strike the ball [5]. Furthermore, high ball velocities restrict the offense's ability to advance bases and score runs [6]. Among other parameters, ball velocity is considered an important performance measure sought after by coaches and scouts. It enables baseball players to improve their ability to play and to be noticed by coaches and scouts for higher levels of competition. Therefore, every pitcher aims to increase the ball velocity [7, 8].

The pitching biomechanics in baseball is studied to improve players' performance and prevent sport-related injuries. With development of the measurement and analytical tools, pitching coaches and biomechanists can accurately analyse the rapid and complex movements during the pitching motion [9]. Although professional baseball teams have used biomechanical analysis for years already, recent advances in technology give amateur players and clubs opportunities to measure their mechanics and improve performance. Body worn sensors, such as inertial measurement units (IMUs), are a low-cost alternative to motion capture systems with passive markers, with no space limitation or cumbersome setup procedure. Portable, affordable, and easy-to-use, they monitor athlete's performance without obstructing it [10].

As the quality of the throw is mainly determined by the pitcher's throwing mechanics, we can use IMUs to measure kinematic parameters shown to be linked to ball velocity [8, 9, 11]. Enhancing pitching technique through the optimal position and timing of proper pitching mechanics, can result in a fast and accurate throw. Estimating ball velocity based on IMU recordings can be the first step towards assessment of the pitching technique that results in fast throws with reduced injury risk.

Ball velocity is mostly measured in high level games and in training situations. Although a radar gun gives an accurate reading of a ball velocity, a required strict protocol and high price represent a big issue for baseball clubs, especially the smaller ones. On the other hand, IMUs do not need a fixed location on the field for measuring ball velocity, thus they can be used on many different occasions. The previous studies demonstrated the potential use of IMUs for estimation of the ball velocity in different overhead-throwing sports, including baseball [12–14].

The use of IMUs represent a potential for the estimation of the ball speed in different on-field situations based on kinematic parameters measured by the same sensors. However, each pitcher is a unique individual and his individual characteristic may display individuality contributing to imparted velocity to the ball [8, 15]. With IMUs, every throw of an individual pitcher can be recorded: during warm-up, training, before and during the game, which contributes to the element of individualisation. Therefore, in this paper, we present a method for predicting ball velocity in baseball pitching based on pitcher's kinematics measured by IMUs and individual characteristics. We investigate the added value of the individuality to predictive performance of developed models.

## 2.2. MATERIALS AND METHODS

### 2.2.1. PARTICIPANTS

Data were collected from 25 baseball pitchers with a mean age of 14.7 ± 1.5, mean body height 176.91 ± 11.03 cm and mean body weight 65.6 ± 14.4 kg. Participants were recruited from the national U18 baseball team, as well as all six baseball academies in The Netherlands, at which the most talented baseball players of that region train. This research was conducted in accordance with the Declaration of Helsinki and the Department of Human Movement Sciences' local ethical committee approved the measurement protocol [ECB 2013-53]. Both participants and their parents were informed of the procedure and study aims before the start of the measurements. Informed consent was obtained from the parents of the participants before involvement in the study.

### 2.2.2. METHODOLOGY

The measurements were performed at the indoor facilities of the academies. After performing several anthropometric measurements, pitchers were given unlimited amount of time for their standard warm-up. They were instructed to prepare just as if they were going to pitch in a game. Warm-up included a general warm-up, j-band exercises, and longtoss, which is a standard warm-up for baseball pitchers before the game. The pitchers wore sneakers, athletic shorts, and no shirt. They also wore their catching glove to mimic the game situation as much as possible. After warm-up, the pitcher was instructed to perform 10 fastball pitches with maximal effort towards a catcher sitting behind home plate.

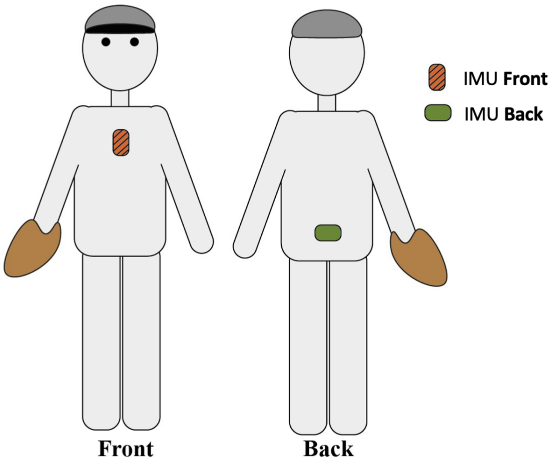


Figure 2.1.: Placement of the sensors.

The pitching motion was recorded using two 9-DOF IMUs (MPU-9150, Invensense, San Jose, CA, USA, Accelerometer ± 16 g, Gyroscope ± 2000 deg/s). Sensors were rigidly attached to pelvis and sternum (Figure 2.1) using double-sided adhesive tape and used to record body segment rotation. Every sensor was embedded in a protective casing together with a battery and SD-card, onto which the data were logged at a sample frequency of 500 Hz. IMU's gyroscope recorded angular velocities continuously throughout the participant's session. Previous studies used peak values of kinematic measures to address their effect on the ball velocity in baseball pitching [3, 7, 8, 11]. Therefore, for the gyroscope signal, we calculated the peak angular velocity as its Euclidean norm. Each recording was manually segmented into parts containing only a single pitch. We performed the segmentation by plotting the entire gyroscope signal and locating the 10 peaks each corresponding to a pitch (see Figure 2.2). This was done in a similar way in [13] for ball velocity data obtained in handball.

The ball velocity (mph) reached during the pitches was measured from behind the home plate using a Stalker pro 2 radar gun (Stalker Radar,

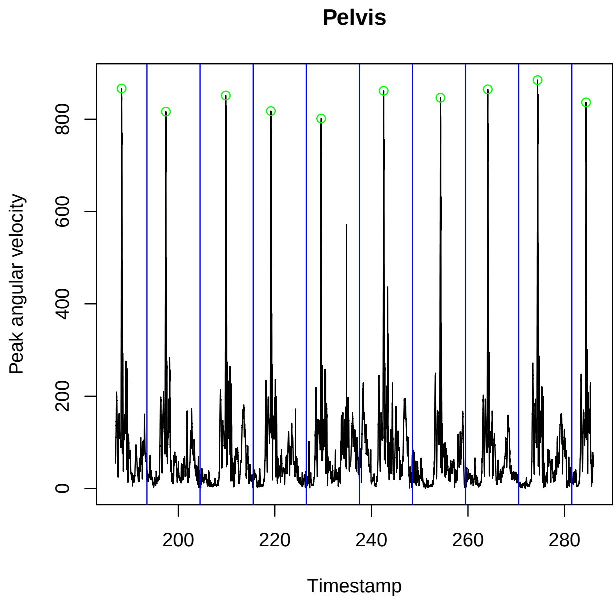


Figure 2.2.: Segmenting the baseball pitches using gyroscope peaks. For the gyroscope signal, we calculated the peak angular velocity as its Euclidean norm. Each recording was manually segmented into parts containing only a single pitch. We performed the segmentation by plotting the entire gyroscope signal and locating the 10 peaks each corresponding to a pitch.

Plano, TX, USA). We coupled recorded ball speed with corresponding peak angular velocities during single pitch.

### 2.2.3. STATISTICAL ANALYSIS

The repeated measurements of individual pitchers can be grouped into a hierarchical structure. The differences between participants arise from differences in personal characteristics, such as age, body weight, and height, that, next to the kinematic parameters of pitching mechanics, may contribute to increased ball velocity [2, 3, 9]. Observations in this study are ball throws nested within different participants and the link between individual- and group-level is participant's indicator (ID).

Statistical models that can deal with units grouped on different levels are known as multilevel models. Multilevel models extend standard regression models to data which are structured in groups and where coefficients are allowed to vary by groups. The feature that distinguishes multilevel models from classical regression is the modeling of variation between groups. This enables us to study the effects that vary by group. Therefore, in this paper we introduce multilevel modeling as the main method for ball velocity prediction in baseball pitching.

At the same time as including repeated measurements of segment rotation per participant, the multilevel approach enables us to examine the added value of the individuality in ball velocity prediction. Group-level predictors were selected among personal characteristics that were collected prior to the measurements. We addressed the high correlation between pitcher's height, weight and age. It is reasonable to expect that older pitchers will be taller and therefore weigh more. To select group-level predictors and avoid poor prediction performance due to correlation of predictors, we applied a random forest (see for instance chapter 8 in [16]), as implemented in the `caret` package [17]. Based on variable importance (Figure 2.3) calculated with `varImp` from a `caret` package [17], we selected pitcher's height as a group-predictor. We developed three multilevel Bayesian regression models for ball velocity prediction using `R 4.0.3` and `rstanarm` [18, 19].

In the following, $y_i$ denotes the ball speed for the observation indexed $i$.

1. Complete-pooling model (*Observations*)

   The complete-pooling model is a single classical regression model completely ignoring group information. In other words, the model treats all ball throws as different observations of the same participant. The model is given by

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \epsilon_i \tag{2.1}$$

   where $\{x_{1i}, x_{2i}\}$ are individual-level predictors, namely peak angular velocity of pelvis and trunk, respectively. The complete-pooling model does not make a distinction between different pitchers and in that way neglects their personal characteristics.

2. Two-level varying-intercept model (*Personal*)

   The two-level varying-intercept model is a regression that opposed to complete-pooling includes indicators for groups. In this model an intercept is calculated for every group and one joint slope is assumed for the entire sample. The model is given by

$$y_i = \alpha_j + \epsilon_i \tag{2.2}$$

$$\alpha_j = \gamma_0 + \gamma_1 \overline{u_j} + \eta_j \tag{2.3}$$

   where $\overline{u_j}$ is a centred group-level predictor, namely pitcher's height. The group membership $j[i]$ denotes pitcher $j$ throwing a ball $i$.

In this model pitching technique is neglected and the outcome depends only on height of an individual pitcher.

3. Two-level varying-intercept, varying-slope model (*Full*)

The varying-intercept, varying-slope model represents the model in which both the intercept and the slope vary by group. The model is given by

$$y_{\mathrm{i}} = \alpha_{\mathrm{j[i]}} + \beta_{\mathrm{1j[i]}} x_{\mathrm{1i}} + \beta_{\mathrm{2j[i]}} x_{\mathrm{2i}} + \epsilon_{\mathrm{i}} \quad (2.4)$$

$$\alpha_{\mathrm{j}} = \gamma_0^{\alpha} + \gamma_1^{\alpha} \overline{u_{\mathrm{j}}} + \eta_{\mathrm{j}}^{\alpha} \quad (2.5)$$

$$\beta_{\mathrm{1j}} = \gamma_0^{\beta_1} + \gamma_1^{\beta_1} \overline{u_{\mathrm{j}}} + \eta_{\mathrm{j}}^{\beta_1} \quad (2.6)$$

$$\beta_{\mathrm{2j}} = \gamma_0^{\beta_2} + \gamma_1^{\beta_2} \overline{u_{\mathrm{j}}} + \eta_{\mathrm{j}}^{\beta_2} \quad (2.7)$$

and includes both individual- and group-level predictors. In (2.3), (2.5), (2.6) and (2.7), the coefficient $\gamma_0$ can be interpreted as the ball speed of a ball thrown without any pelvis and trunk rotation by the pitcher of an average height. The $\epsilon_{\mathrm{i}}$ in (2.1), (2.2) and (2.4) and $\eta_{\mathrm{j}}$ in (2.3), (2.5), (2.6) and (2.7) represent independent error terms at each of the two levels.

All individual- and group-level predictors were rescaled to have sample variance 1. The scaling is done by dividing the centred predictor $\overline{u}$ by its standard deviation. We used scaling to transform the data to comparable values.

We used leave-one-out (LOO) cross-validation to select out of the three proposed models the model with best predictive performance. LOO resulted in a total of 224 folds as 224 pitches from 25 pitchers were included in the analysis. Following the approach in [20], the predictive performance of a model is defined as the expected log-predictive density (elpd). Predictive performance is a useful quantity in assessing a single model. It can be estimated by training the model on all observations except one and then predicting the hold-out observation. This is then repeated for all $n$ observations

$$\mathrm{elpd} = \sum_{i=1}^{n} \log p(y_{\mathrm{i}} \mid y_{\text{-i}}) \quad (2.8)$$

where

$$p(y_{\mathrm{i}} \mid y_{\text{-i}}) = \int p(y_{\mathrm{i}}|\theta) p(\theta|y_{\text{-i}}) d\theta \quad (2.9)$$

is the LOO predictive density upon leaving out the $i$th data point. If the posterior $p(\theta \mid y_{\text{-i}})$ is summarised by $B$ simulation from $\theta^{i,b}$, then we can

2

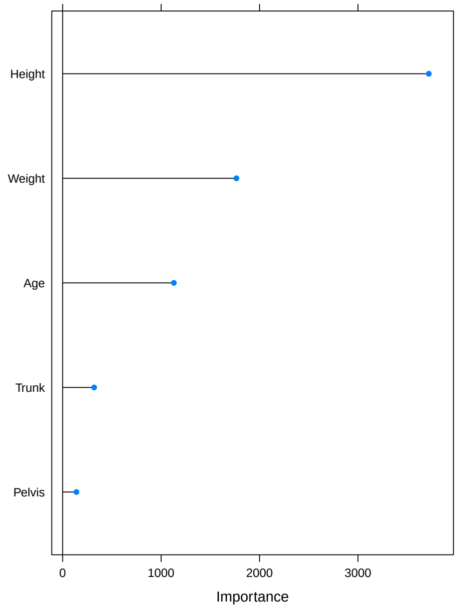


Figure 2.3.: Visual representation of the variable importance calculated by applying random forest. The horizontal axis should be interpreted as a measure for relative importance of predictive variables. The figure reveals *Height* to be the most important predictor for ball speed which is, therefore, selected as group-level predictor.

approximate $\log p(y_{\mathrm{i}} \mid y_{\mathrm{i}=1})$ by

$$\widehat{\mathrm{elpd}}_{\mathrm{i}} = \frac{1}{B} \sum_{b=1}^{B} p(y_{\mathrm{i}} \mid \theta^{i,b})$$

leading to $\widehat{\mathrm{elpd}} = \sum_{i=1}^{n} \widehat{\mathrm{elpd}}_{\mathrm{i}}$ as an estimate for $\widehat{\mathrm{elpd}}$.

Different models can be compared against each other according to their elpd-value. Suppose we wish to compare models $\mathcal{M}_1$ and $\mathcal{M}_2$, with estimated elpd values $\widehat{\mathrm{elpd}}^1$ and $\widehat{\mathrm{elpd}}^2$, respectively.

Since the same set of $n$ data points is being used to fit all models, we can use a paired estimate to compute a standard error of their difference:

$$\mathrm{se}\left(\widehat{\mathrm{elpd}}^1 - \widehat{\mathrm{elpd}}^2\right) = \sqrt{n V_{i=1}^{n}\left(\widehat{\mathrm{elpd}}_{\mathrm{i}}^1 - \widehat{\mathrm{elpd}}_{\mathrm{i}}^2\right)}. \tag{2.10}$$

Here, for numbers $\{a_{\mathrm{i}}\}_{i=1}^{n}$ we define $V_{i=1}^{n} a_{\mathrm{i}} = \frac{1}{n-1} \sum_{i=1}^{n} (a_{\mathrm{i}} - \bar{a}_{\mathrm{n}})^2$.

## 2.3. RESULTS

We included in the analysis 224 pitches from 25 pitchers for which the ball velocity was recorded and sensor clipping did not occur. Characteristics of the measured ball and peak angular velocities of pelvis and trunk are summarised in Table 2.1.



| | **Mean ± Standard Deviation** |
|---|---|
| Peak pelvis angular velocity (°/s) | 690.2 ± 90.9 |
| Peak trunk angular velocity (°/s) | 1172.4 ± 239.5 |
| Ball velocity (mph) | 68.3 ± 6.5 |

Table 2.1.: Summary of measured ball and peak angular velocities.

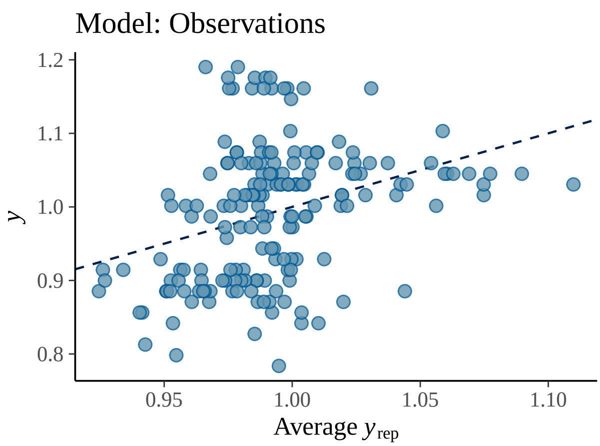


Figure 2.4.: Ball velocity observations (dots) vs. average simulated value of the ball speed (line) from the posterior predictive distribution of the *Observations* model. This graphical representation suggests that *Observations* model leaves a large amount of variation in the data unexplained.

We consider the model called *Observations* model as base model. The other two proposed models, *Personal* and *Full*, are extensions since they have two instead of one level and they introduce the group participation that makes a distinction between pitchers of a different height. Therefore, comparing the developed models determines the contribution of the kinematic parameters related to pitching mechanics and body height of a pitcher to accuracy of ball speed prediction. The graphical representations (Figures 2.4–2.6) show that both the models *Personal* and *Full* provide a good fit to the observed data, while the fit of

*Observations* is unsatisfactory.

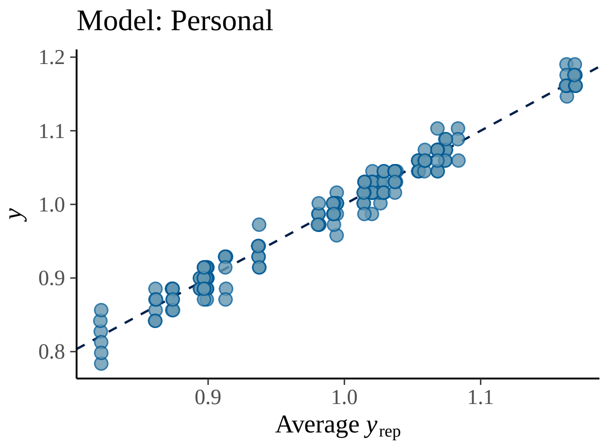


Figure 2.5.: Ball velocity observations (dots) vs. average simulated value of the ball velocity (line) from the posterior predictive distribution of the *Personal* model. This graphical representation suggests that *Personal* model is a good fit to collected data.

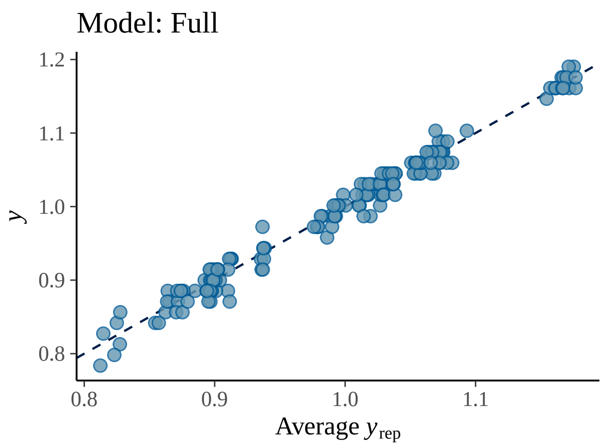


Figure 2.6.: Ball velocity observations (dots) vs. average simulated value of the ball velocity (line) from the posterior predictive distribution of the *Full* model. This graphical representation suggests that *Full* model is a good fit to collected data.

The *Full* model is a preferable model, followed by the *Personal* and *Observations* model (see Tables 2.2 and 2.3).

| | |
|---|---|
| $\widehat{\text{elpd}}^{\text{Full}} - \widehat{\text{elpd}}^{\text{Personal}}$ | −5.5 (3.3) |
| $\widehat{\text{elpd}}^{\text{Full}} - \widehat{\text{elpd}}^{\text{Observations}}$ | −308.3 (13.5) |

Table 2.2.: Comparison of fitted models. The rows show the difference in $\widehat{\text{elpd}}$, with estimated standard error in brackets, between the *Full* model and remaining models (*Personal* and *Observations*).



| | $R^2$ | *RMSE* |
|---|---|---|
| Full | 0.975 | 0.014 |
| Personal | 0.973 | 0.014 |
| Observations | 0.137 | 0.089 |

Table 2.3.: Comparison of fitted models.

## 2.4. DISCUSSION

The aim of this study was to predict a ball velocity in baseball pitching such that prediction is tailored to the individual pitcher. The proposed method included pitcher's body segment rotation, which determines his technique, and pitcher's height that displays individuality in imparted velocity to a ball. We used multilevel modeling to develop three models with different predictors and examined their predictive performance. By comparing developed models, we investigated the added value of individuality to ball velocity prediction.

Ball velocities presented in this study are similar to the ones reported in the previous studies. Pitchers with a mean age of 14.7 ± 1.5 years threw balls with average velocity 30.6 ± 2.9 m/s, while Dun [21] reported average ball velocity of 26.3 ± 3.8 m/s measured in a population of youth pitchers throwing fastballs.

In the overhead pitching, the lower extremity and trunk generate and transfer energy to the upper extremity. To examine the relationship between ball velocity and variations in pitching biomechanics on individual level, previous studies identified maximum pelvis and trunk angular velocity as kinematic parameters linked to ball velocity [8, 9, 11]. Recent technological developments brought IMUs to a spotlight as an alternative to marker-based systems used in a laboratory setting. Since IMU's gyroscope enables measuring body segment rotation, we assessed pitching technique by positioning IMU sensors on pelvis and trunk. Measured peak angular velocity of pelvis of 690.2 ± 90.9 deg/s and trunk of 1172.4 ± 239.5 deg/s supports the findings in previous studies [3, 21]. As the gyroscopes recorded angular velocities continuously throughout the participant's session, manual segmentation was required. In future work, we wish to develop a method for automatic detection of single throws and segmentation of the continuous-time gyroscope signal when the boundaries between different throws are unclear. This will automatise the use of predictive models for predicting ball velocity. Filtering methods from signal processing may prove to be useful for this purpose.

Among the compared models, model *Full* shows the best predictive performance (Table 2.3). Model *Observations* is worse than model *Full* by 308.3 of log predictive probability values. The difference in estimated elpd-values is big compared to estimated standard error of 13.5. Hence, adding pitcher's height to the *Observations* model improves predictive accuracy. Model *Personal* includes only the pitcher's height as a predictor and ignores the pitching technique. The model shows that taller pitchers throw faster and it is possible to already estimate ball velocity only by knowing the pitcher's height. This information can be useful for scouts in search for baseball talents. A pitcher's height compared to other personal characteristics, such as age and weight, is the most important predictor for ball velocity in baseball pitching. On the other hand, of course neither pitchers nor coaches can influence height. The outcome of this paper demonstrates the added value of a pitcher's height to predictive accuracy.

The proposed method can potentially be adopted in baseball practice. IMUs are easy-to-wear low cost sensors that do not influence a pitcher's performance and can be a valuable source of data. It can provide information on pitching performance in every situation and with a method proposed in this paper, gain ball velocity without use of a radar gun. Ball velocity prediction can give a better insight into pitcher's performance and represents a potential for predictions of future throwing speed when pitchers grow taller.

For future studies, we suggest also to include separation time and pitch types in the presented model. Following the concept of a kinetic chain, the relative timing of the moments of pelvis and trunk peak angular velocity, when throwing fastballs, is associated with ball velocity in skilled pitchers [22]. Furthermore, to the best of our knowledge, no study has classified pitch types based on IMU data solely. Classification of pitch types outside the laboratory or game environment provides benefits in designing and outlining training routines and represents a potential research direction in the future. Following the segmentation of continuous gyroscope signals, we suggest extracting additional features next to the peak angular velocities, such as skewness, mean, and difference between minimum and maximum. This would result in more parameters that may be included in the model and improve the classification of different pitch types and the prediction of ball velocity.

# REFERENCES


[1] L. Gomaz, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. "Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data". en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: `10.3390/s21227442`. url: `https://www.mdpi.com/1424-8220/21/22/7442`.

[2] A. M. Pappas, R. M. Zawacki, and T. J. Sullivan. "Biomechanics of baseball pitching: a preliminary report". In: *The American journal of sports medicine* 13.4 (1985), pp. 216–222.

[3] S. T. Seroyer, S. J. Nho, B. R. Bach, C. A. Bush-Joseph, G. P. Nicholson, and A. A. Romeo. "The kinetic chain in overhand pitching: its potential role for performance enhancement and injury prevention". In: *Sports health* 2.2 (2010), pp. 135–146.

[4] A. W. Anz, B. D. Bushnell, L. P. Griffin, T. J. Noonan, M. R. Torry, and R. J. Hawkins. "Correlation of torque and elbow injury in professional baseball pitchers". In: *The American journal of sports medicine* 38.7 (2010), pp. 1368–1374.

[5] J. Hay. *The biomechanics of sports techniques*. Prentice-Hall, 1978.

[6] G. Lehman, E. J. Drinkwater, and D. G. Behm. "Correlation of throwing velocity to the results of lower-body field tests in male college baseball players". In: *The journal of strength & conditioning research* 27.4 (2013), pp. 902–908.

[7] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews. "Baseball Pitching Biomechanics in Relation to Injury Risk and Performance". en. In: *Sports Health: A Multidisciplinary Approach* 1.4 (July 2009), pp. 314–320. issn: 1941-7381, 1941-0921. doi: `10.1177/1941738109338546`. url: `http://journals.sagepub.com/doi/10.1177/1941738109338546`.

[8] S. L. Werner, M. Suri, J. A. Guido Jr, K. Meister, and D. G. Jones. "Relationships between ball velocity and throwing mechanics in collegiate baseball pitchers". In: *Journal of shoulder and elbow surgery* 17.6 (2008), pp. 905–908.

[9] B. Dowling, C. Pearl, W. Laughlin, and G. Fleisig. "Relationship of pelvis and trunk kinematics to ball velocity in professional baseball pitchers". In: *40th American Society of Biomechanics (ASB) annual meeting*. 2016.

2


[10] V. Camomilla, E. Bergamini, S. Fantozzi, and G. Vannozzi. “Trends supporting the in-field use of wearable inertial sensors for sport performance evaluation: A systematic review”. In: *Sensors* 18.3 (2018), p. 873.

[11] E. Van Der Graaff, M. ( Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. ( Veeger. “Timing of peak pelvis and thorax rotation velocity in baseball pitching”. en. In: *The Journal of Physical Fitness and Sports Medicine* 7.5 (Sept. 2018), pp. 269–277. issn: 2186-8123, 2186-8131. doi: 10.7600/jpfsm.7.269. url: https://www.jstage.jst.go.jp/article/jpfsm/7/5/7_269/_article.

[12] J. McGrath, J. Neville, T. Stewart, H. Clinning, and J. Cronin. “Can an inertial measurement unit (IMU) in combination with machine learning measure fast bowling speed and perceived intensity in cricket?” In: *Journal of Sports Sciences* 39.12 (2021), pp. 1402–1409.

[13] S. D. Skejø, J. Bencke, M. Møller, and H. Sørensen. “Estimating throwing speed in handball using a wearable device”. In: *Sensors* 20.17 (2020), p. 4925.

[14] C. Gençoğlu and H. Gümüş. “Standing handball throwing velocity estimation with a single wrist-mounted inertial sensor”. In: *Annals of Applied Sport Science* 8 (2020).

[15] T. Matsuo, R. F. Escamilla, G. S. Fleisig, S. W. Barrentine, and J. R. Andrews. “Comparison of kinematic and temporal parameters between different pitch velocity groups”. In: *Journal of applied biomechanics* 17.1 (2001), pp. 1–13.

[16] M. Kuhn, K. Johnson, *et al. Applied predictive modeling*. Vol. 26. Springer, 2013.

[17] M. Kuhn. “Caret: classification and regression training”. In: *Astrophysics Source Code Library* (2015), ascl–1505.

[18] B. Goodrich, J. Gabry, I. Ali, and S. Brilleman. *rstanarm: Bayesian applied regression modeling via Stan.* R package version 2.21.1. 2020. url: https://mc-stan.org/rstanarm/.

[19] C. Muth, Z. Oravecz, and J. Gabry. “User-friendly Bayesian regression modeling: A tutorial with rstanarm and shinystan”. In: *The Quantitative Methods for Psychology* 14.2 (2018), pp. 99–119.

[20] A. Vehtari, A. Gelman, and J. Gabry. “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”. In: *Statistics and computing* 27 (2017), pp. 1413–1432.

[21] S. Dun, J. Loftice, G. S. Fleisig, D. Kingsley, and J. R. Andrews. "A Biomechanical Comparison of Youth Baseball Pitches: Is the Curveball Potentially Harmful?" en. In: *The American Journal of Sports Medicine* 36.4 (Apr. 2008), pp. 686–692. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546507310074. url: http://journals.sagepub.com/doi/10.1177/0363546507310074.

[22] E. van der Graaff, M. Hoozemans, M. Nijhof, M. Davidson, M. Hoezen, and D. Veeger. "The role of pelvis and thorax rotation velocity in baseball pitching". In: *ISBS-Conference Proceedings Archive*. 2016.

# 3

# MACHINE LEARNING APPROACH FOR PITCH TYPE CLASSIFICATION BASED ON PELVIS AND TRUNK KINEMATICS CAPTURED WITH WEARABLE SENSORS

**Larisa van der Graaff, Celine Bouwmeester, Erik van der Graaff, Bart van Trigt, DirkJan Veeger**

Published in *Sensors* **2023** [1].

## 3.1. INTRODUCTION

Data-driven decision-making is establishing itself in training and high-level sports performance. Data made available through game statistics and technology integrated with training routines serve as the input for big data analytics in sports. Data analysis started in many sports disciplines with some form of video analysis. Currently, a variety of different metrics can be extracted and analysed not only from videos, but also sensors integrated into sleeves, straps, watches, rings, and smart fabrics. For instance, in baseball, for over 100 years, the difference between a slider and a curveball was defined based on previous experience. Following the technological advancements in pitch tracking, the concept of pitch types is quantified and explained by the speed, spin rate, and spin axis of the ball. Information on the ball (Rapsodo), the bat (Blast), and body movement (PITCHPERFECT) has become widely accessible, creating a new flow of data, which are valuable for performance assessment and pitchers' overall success.

The advancements in wearable technology are changing the traditional approach to athlete training and performance monitoring. Wearables enable measurements in a wide range of settings during training and matches. This removes any practical limitation compared to a lab and offers unlimited athlete availability, which results in high numbers of recorded repetitions. While biomechanical measurements in the lab as well as coaching sessions during training are often limited to one athlete at a time, the utilization of wearables ensures that every pitch thrown by the pitchers is recorded, even the ones during warm-up sessions. The use and collection of data from wearables can be performed by any motivated team that might lack the resources available to professional sports teams, and this enables coaches to retrospectively provide feedback to every pitcher. Such performance tracking in terms of pitch counts enables players to pitch without fatigue, directly adhering to the pitch count limit regulated by the federations in order to limit the workload and prevent shoulder and elbow injuries [2].

Next to the pitch count, the pitching mechanics and pitch type are considered the main factors in pitching training, which are relevant not only for pitchers' performance, but also for the prevention of injuries [3–6]. As the pitcher's response to a given training stimulus is highly individualized [7], continuous and prospective individual monitoring is crucial in managing the effect of the intense training and competition schedule on the pitcher's performance and health. The use of wearable sensors may provide the opportunity to achieve this.

Information extracted from wearables creates the opportunity to understand the body mechanics of each pitcher on an individual level. Detailed pitch-to-pitch information can help the pitcher learn safe and efficient pitch mechanics. In general, pitching mechanics follow the kinetic chain principle in which the pelvis and trunk serve as a link in the transfer of the momentum generated by the lower extremities to the upper extremities. Efficient proximal-to-distal timing between the pelvis and trunk allows momentum transfer to the ball, resulting in increased throwing velocity [8–10]. On the contrary, poor pitching mechanics in combination with the repetitive mechanical strain of throwing through a high pitch count can negatively affect pitching performance and, at the same time, put the pitcher at risk of shoulder and elbow injuries [2, 4–6].



To translate training success into game success, pitchers need to translate their movement skills into a variation of pitch trajectories. A successful pitcher alters the velocity and trajectory of the ball to keep the batters off balance and discourage their anticipation of a particular pitch type. To obtain a variation of ball trajectory, in theory, the pitcher manipulates the grip on the ball at the release point, which results in different rotations of the ball out of the hand of the pitcher. The particular seams of a baseball lead to air pressure variations around the ball, which creates the bending, curving, or sliding motions of the pitch. It should be noted though that multiple studies have reported differences in the pelvis and trunk kinematics between pitch types [4, 11–14]. From a strategic point of view, a pitcher may want to achieve similar kinematics among all pitch types to make pitch identification difficult for the batters [12]. If that were the case, it would be unlikely that the pitch type could be distinguished from the body mechanics alone. However, the aforementioned studies acquired their data in a lab setting with highly trained individuals. It can be expected that, at lower levels of play, the movement variation within the individual is even higher.

Except the skill difference, there are obvious differences in financial resources and staff availability as well. Although it is common in youth baseball that a volunteer manually counts the amount of pitches, the tracking of the pitch types is very limited, and in particular, off-speed pitches lead to wildly inaccurate manual classifications given the skill level of the person performing the tagging. Therefore, the automatic detection of pitch types might be extremely beneficial, especially for baseball players who cannot afford expensive camera systems and rely on the manual tracking of pitch types. In this context, it should also be noted that off-speed pitches are associated with an increased risk of shoulder and elbow injuries in youth baseball pitchers. In combination with the increased number of pitches per game and the full baseball calendars, pitchers are at risk of not only acute problems, but also overuse injuries in the later stages of their careers [2].

Translating collected wearables data into actionable insights may bridge the gap between scientific knowledge from biomechanical studies and daily practice. We provide a machine learning approach to the utilization of wearables data through pitch type classification based on the pelvis and trunk peak angular velocity and their separation time recorded using body-worn motion sensors. Machine learning methods showed promising results in pitch type classification investigated in similar contexts [15–21]. Opposed to predicting the next pitch thrown based on the information available prior to that pitch [15–17], our approach relies on inclusion of post-delivery features to detect which pitch was thrown purely based on pitching mechanics. Having pitch type readily available on every pitch, in combination with kinematic data, might help us provide insight into pitching technique to baseball pitchers of various levels. On top of that, overview of such performance metrics can be presented to the athletes in real time, enabling players to track their progress throughout the whole season and empowering them to shape the training accordingly.

To the best of our knowledge, this is the first study investigating baseball pitch type detection based on pelvis and trunk kinematics during pitching and, moreover, based on such data obtained from wearables. This approach allows for workload monitoring, which is important for maintaining safe and efficient pitching performance during the full course of the season. Therefore, this study aims to establish the methodology for pitch type classification based on biomechanical input from wearables by comparing performance of the various classification algorithms.

## 3.2. MATERIALS AND METHODS

### 3.2.1. PARTICIPANTS

Out of 24 pitchers initially participating in the measurements, 19 pitchers were included in this study (age 18.5 ± 3.7 years, height 178.3 ± 11.1 m, weight 71.9 ± 18.3 kg, experience 7.3 ± 3.7 years). The participants were members of the elite youth academies of the Royal Dutch Baseball and Softball Federation (KNBSB). The included pitchers were pain- and injury-free during the course of the measurements. This research was conducted in accordance with the Declaration of Helsinki, and the Ethics Committee of the Delft University of Technology approved the measurement protocol (approval no. ETC_TUDelft_1394). Informed consent was signed by the participants or the general manager of the respective baseball academy.

### 3.2.2. DATA COLLECTION AND DATA PRE-PROCESSING

The data were collected during the pitchers' regular training at the training facilities of the affiliated baseball academy. To maintain pitching-specific routines, warm-up and pitch count were not standardised. After performing their standard warm-up, the pitchers were instructed to throw a selection of pitch types they usually throw during the game, containing a minimum of three different pitch types. The pitchers followed their own training routine in accordance with the training program set by their pitching coach. The bullpen session consisted of a minimum of 20 pitches from mound toward a catcher at the official distance of 18.45 or 16.45 meters, depending on the pitcher's age.

The pitching motion was recorded using the PITCHPERFECT system (PITCHPERFECT, Breda, The Netherlands) consisting of two synchronised 3-DOF IMUs (Gyroscope $\pm 2000$ (°/s)) showed on Figure 3.1. Sensors were taped with Leukoplast FixoMull® stretch (BSN Medical GmbH, Hamburg, Germany) on processus Xiphoideus on the chest and in the middle of the left and right posterior superior iliac spine on the lower back of the pitcher before starting the bullpen (Figure 3.2). Pitch types were manually coded by experienced off-field staff members based on the visual inspection, hand signal and pitcher–catcher agreement prior to each throw. The ball velocity (mph) was measured from behind the pitcher with a Pocket radar Ball coach, Model PR1000-BC (Pocket Radar Inc., Santa Rosa, CA, USA). The accuracy of the pitch was noted, distinguishing only between a wild pitch or not, wherein a wild pitch was noted if the catcher was unable to catch the ball with reasonable effort.

Figure 3.1.: Pitch Perfect sensor system for measuring pelvis and trunk kinematics and separation time between them.

The outcome of the PITCHPERFECT system consists of the pelvis and trunk peak angular velocity and the separation time between them. Pre-processing of the raw sensor signal and computing Euclidean norms from the raw data were conducted by the algorithm developed by the manufacturer (PITCHPERFECT, The Netherlands). Details of the algorithm are property of the manufacturer.

In this study, we used a database created by PITCHPERFECT that characterizes each pitch with three features used directly from the system (Table 3.1). Data were pre-processed and analyzed using the `R` programming language (`version 4.3.1`). Data of five players were excluded from the analyses because their peak angular velocity was below the threshold of 400 (°/s) of the PITCHPERFECT system. Individual pitches were included based on three inclusion criteria: (1) the pitch type is a Fastball (FB), Curveball (CU) or Change-up (CH), as they were the most occurring pitch types among the included pitchers; (2) the thrown ball was not a wild pitch; and (3) all three kinematic parameters (*Pelvis*, *Trunk*, *Separation*) were recorded (i.e., sensor clipping did not occur). All continuous features were scaled and centered.

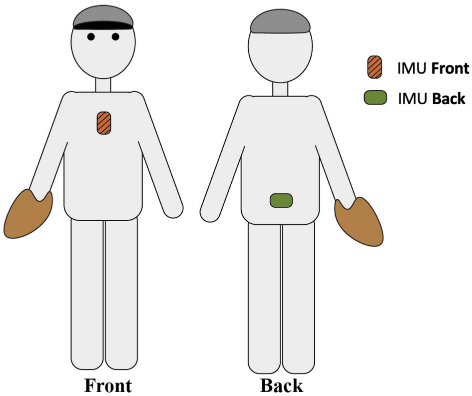


Figure 3.2.: Placement of the sensors. Figure adopted from the study of Gomaz et al. [22].

| Features | Definitions |
|---|---|
| Pelvis (°/s) | Pelvis peak angular velocity available directly from PITCHPERFECT. |
| Trunk (°/s) | Trunk peak angular velocity available directly from PITCHPERFECT. |
| Separation (ms) | The timing between pelvis and trunk peak angular velocity, available directly from PITCHPERFECT. |

Table 3.1.: Included features for pitch type classification.

### 3.2.3. DATA ANALYSIS

The automatic detection of pitch types from sensor data is a classification problem. The goal is to learn a mapping from inputs $x$ to outputs $y$, where $y \in \{1, \ldots, C\}$, with $C$ being the number of classes. Inputs $x$ are the features (Table 3.1) and outputs $y$ are pitch types, where $C$ denotes number of different pitch types.

This study utilized classifiers integrated in the caret package [23] including K–Nearest Neighbors (KNN), Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SVM). We investigated the performance of the classifiers in both binary and multiclass classification, including additional Logistic Regression (LOGREG) for binary and Multinomial Logistic Regression (MNOM) for the multiclass classification task.

Binary classification is a classification task that has two class labels. In this study, it is used to detect whether the pitch was Fastball or not by classifying recorded pitches in one of the two classes—FB and Other (Figure 3.3 (left)). Among the recorded pitches, 48.7% were originally labelled as FB and 51.3% as Other.

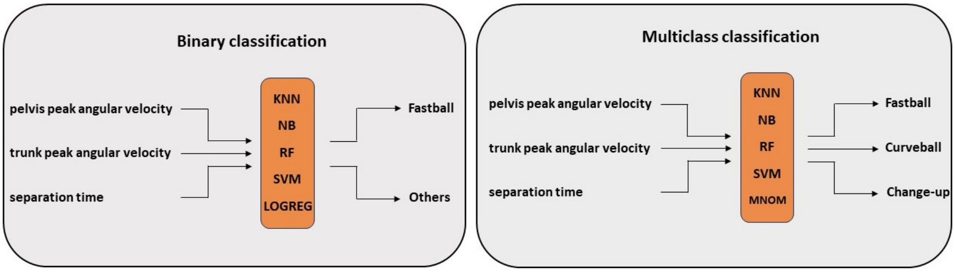


Figure 3.3.: The baseball pitch type classification approaches. (**Left**) The binary classification approach classifies pitch types into two categories—Fastball and Others—based on input from wearables (pelvis and trunk peak angular velocities and separation time). (**Right**) The multiclass classification approach classifies pitch types into three categories—Fastball, Curveball and Change-up—based on input from wearables (pelvis and trunk peak angular velocities and separation time). Both approaches used four classifiers—K-Nearest Neighbors (KNN), Naive Bayes (NB), Random Forest (RF) and Support Vector Machine (SVM)—to assess their classification performance, including additional logistic regression (LOGREG) for binary and multinomial logistic regression (MNOM) for the multiclass classification task.

Multiclass classification refers to classification tasks that have more than two class labels. Unlike binary classification, it classifies non-fastball pitches in different classes and therefore detects whether the pitch was Fastball (FB), Curveball (CU) or Change-up (CH) (Figure 3.3 (right)). Among the recorded pitches, 48.7% were originally labelled as FB, 26.4% as CH and 24.9% as CU. Due to variations in the number and type of off-speed pitches (CU and CH) among pitchers, the collected data show unequal distribution between classes. Such disparity in the frequencies of the observed classes can have a negative impact on model fitting. A possible solution is to subsample the training data in such a way that mitigates the issue (e.g., under- and oversampling). Hence, to address this issue, the minority classes (CU and CH) were up-sampled so that each class was of equal size.

We set up our training and testing cases following the 80% (training) and 20% (testing) split. To achieve a fair understanding of the generalizability of the classifiers, in the designated training set, Leave-One-Group-Out Cross-Validation (LOGO-CV) was carried out. LOGO-CV is a specific type of $k$-fold cross-validation that utilizes data from each individual pitcher as a test set. The number of folds therefore equals the number of pitchers. For every fold, the model is trained on data from $J-1$ pitchers and tested on the data from the one left-out pitcher.

The performance of the classifiers is evaluated by four evaluation criteria—Accuracy (3.1), Sensitivity (3.2), Precision (3.3) and F1-score (3.4)—which can be calculated from the confusion matrix. The confusion matrix provides a summary of the prediction results of a classification algorithm. In the matrix, the numbers of correct and incorrect predictions are summarised with count values and broken down by each class. The output True Positive (TP) represents the number of positives classified correctly, whereas True Negative (TN) represents the number of correctly classified negatives. False Positive (FP) shows the number of negatives that are classified as positives, whereas False Negative (FN) indicates the number of positives classified as negatives.

$$Accuracy = \frac{TP+TN}{Total\ sample} \,, \tag{3.1}$$

$$Sensitivity = \frac{TP}{TP+FN} \,, \tag{3.2}$$

$$Precision = \frac{TP}{TP+FP} \,, \tag{3.3}$$

$$F1 = \frac{2TP}{2TP+FP+FN} \,. \tag{3.4}$$

The hyper-parameters were tuned using grid search, a default method for optimizing tuning parameters in the `caret` package [23]. Feature selection was performed using correlation analysis. Since the correlation between the features was low, the models were trained and tested using all variables derived from the PITCHPERFECT system (Table 3.1).

## 3.3. RESULTS

A total of 353 pitches thrown by 19 pitchers met the inclusion criteria and were included in the study. Descriptive statistics for binary and multiclass classification is presented in Table 3.2 and Table 3.3, respectively. A total of 284 pitches were used for training the models and 69 pitches were used for their testing.



| **Features** | **FB (*n* = 172)** | | **Other (*n* = 181)** | |
|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** |
| Pelvis (°/s) | 737 | 138 | 695 | 120 |
| Trunk (°/s) | 799 | 228 | 827 | 262 |
| Separation (s) | 0.03 | 0.13 | 0.06 | 0.13 |
| Speed (m/s) | 33.1 | 3.82 | 28.6 | 3.81 |

Table 3.2.: Descriptive statistics for binary classification.

| **Features** | **CH (*n* = 93)** | | **CU (*n* = 88)** | | **FB (*n* = 172)** | |
|---|---|---|---|---|---|---|
| | **Mean** | **SD** | **Mean** | **SD** | **Mean** | **SD** |
| Pelvis (°/s) | 708 | 129 | 681 | 109 | 737 | 138 |
| Trunk (°/s) | 831 | 277 | 823 | 247 | 799 | 228 |
| Separation (s) | 0.06 | 0.15 | 0.06 | 0.11 | 0.03 | 0.13 |
| Speed (m/s) | 29.9 | 3.72 | 27.2 | 3.40 | 33.1 | 3.82 |

Table 3.3.: Descriptive statistics for multiclass classification.

### 3.3.1. BINARY CLASSIFICATION

The performance of the K-Nearest Neighbors, Naive Bayes, Random Forest, Support Vector Machine and Logistic Regression algorithms in the binary classification problem was evaluated using four performance metrics (3.1)–(3.4). Among the trained classifiers, the Naive Bayes algorithm performed the best in classifying fastballs among the recorded pitches. The confusion matrix seen in Figure 3.4 shows the summary of the prediction performance for Naive Bayes (Accuracy = 71.0%, Precision = 71.9%, Sensitivity = 67.6%, F1-score = 69.7%). The accuracy of the NB algorithm was 7.2% higher than for KNN, 1.4% higher than for RF, 5.8% higher than for SVM and 20.3% higher than for LOGREG. The sensitivity of the RF algorithm is 11.8% higher than for KNN, 3% higher than for NB, 5.9% higher than for SVM and 17.7% higher than for LOGREG. The precision of the NB algorithm was 7.4% higher than for KNN, 3.3% higher than for RF, 7.2% higher than for SVM and 21.9% higher than for LOGREG. The F1-score of the NB algorithm was 8.2% higher than for KNN, 0.1% higher than for RF, 5.0% higher than for SVM and 18.3% higher than for LOGREG. The confusion matrices with corresponding performance metrics of the remaining algorithms are shown in Appendix A.

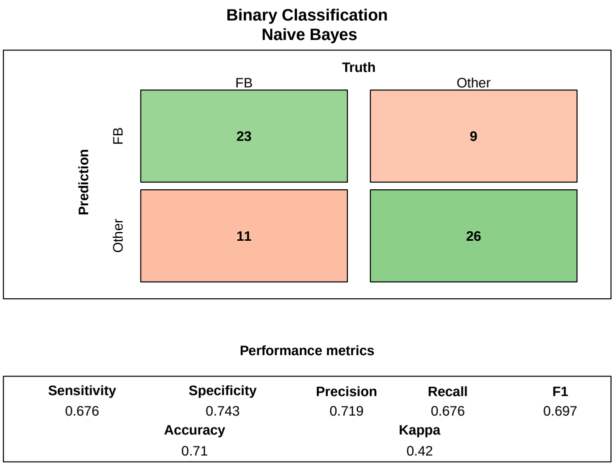


Figure 3.4.: Two-class confusion matrix summarizing the performance of Naive Bayes in classification of fastballs.

### 3.3.2. MULTICLASS CLASSIFICATION

The four metrics are used to evaluate the performance of the K-Nearest Neighbors, Naive Bayes, Random Forest, Support Vector Machine and Multinomial logistic regression algorithms in the multiclass classification problem. Among the trained classifiers, the Random Forest algorithm performed the best in classifying pitches in three different classes of pitch types (FB, CH and CU). The confusion matrix seen in Figure 3.5 shows the summary of prediction performance for Random Forest. The accuracy of the RF algorithm was at 52.2%, which is 7.2% higher than for KNN, 7.2% higher than for NB, 11.6% higher than for SVM and 8.7% higher than for MNOM. The confusion matrices with corresponding performance metrics of the remaining algorithms are shown in Appendix A. Performance metrics of the Random Forest algorithm are reported in Table 3.4.



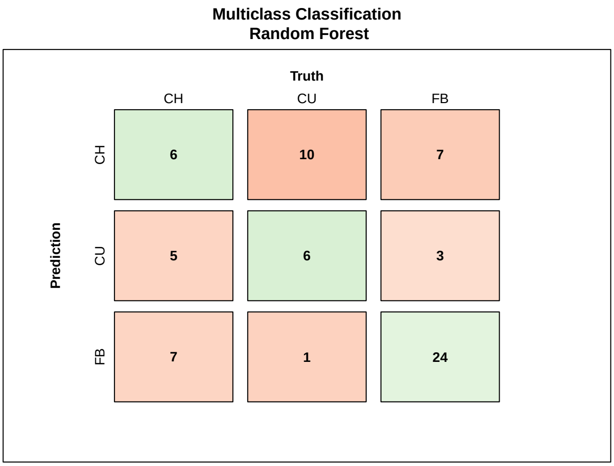


Figure 3.5.: Three-class confusion matrix summarizing the performance of Random Forest by class in classification of baseball pitch types.

| Class | Accuracy | Sensitivity | Precision | F1 |
|---|---|---|---|---|
| CH | 0.500 | 0.333 | 0.261 | 0.293 |
| CU | 0.600 | 0.353 | 0.429 | 0.387 |
| FB | 0.739 | 0.706 | 0.750 | 0.727 |

Table 3.4.: Performance metrics of multiclass Random Forest in classification of three different pitch types.

## 3.4. DISCUSSION

The aim of this study was to establish a methodology for pitch type classification based on biomechanical input from wearables. We used pelvis and trunk peak angular velocity and separation time between them as an input and evaluated the performance of five machine learning classifiers in the binary and multiclass classification task. The Naive Bayes algorithm showed the best performance in classifying Fastballs with an accuracy of 71%. Furthermore, in the classification of pitch types as Fastball, Curveball or Change-up, the Random Forest algorithm performed the best with an average accuracy of 61.3% over those three pitch types.



Binary classification was used to detect whether the pitch was Fastball or not. Fastball can be considered a "normal" throw. Fastball is the most common pitch type thrown, specifically among youth pitchers. This has to do with the physical development of youth pitchers where the Fastball pitch is used to learn proper body mechanics and throwing accuracy before learning more demanding off-speed pitches. Therefore, to explore the possibility of pitch type classification based on pitching mechanics, it makes sense to first investigate whether we can detect fastballs. Previous studies that used a binary approach for pitch type prediction focused on predicting whether the next pitch will be Fastball rather than detecting whether Fastball was thrown [16, 20]. They used pre-pitch ball data as an input, which resulted in accuracies of 70% [16] and 77.45% [20]. Even though such approach offers benefits for choosing the right strategy, it does not contribute to the pitch tracking as part of the workload monitoring for an individual pitcher.

The multiclass classification task classified recorded pitch types into three categories—Fastball (FB), Change-up (CH) and Curveball (CU). It serves as a base for pitch tracking and detects different pitch types thrown. The Random Forest algorithm performed the best with a 50.0% accuracy in classifying CH, a 60.0% accuracy in classifying CU and a 73.9% accuracy in classifying FB. The performance metrics reported in Table 3.4 show the performance of the RF classifier for each pitch type versus the rest. Multiclass classification has been a subject of several studies before, focusing on pitch type classification based on pre-pitch ball data. Compared to the accuracy of the Random Forest algorithm revealed in this paper, those studies reported higher predictive accuracies, from 74.5% [21] for the SVM algorithm to 93.63% for the KNN algorithm with Manhattan distance [18, 19]. This may be due to the sensitive nature of wearable data and inconsistent pitching mechanics of different pitchers among various pitch types. The feature importance for the Random Forest multiclass classifier revealed that the pitcher’s pelvis peak angular velocity is considered most important for the pitch type classification task, whereas the trunk peak angular velocity is considered the least important (Figure A.9).

Although we are confident that the proposed methodology could be key to predict pitch type based on biomechanical data from wearables, the reported accuracies leave much to desire. One limitation of this study was that the amount of collected data was low ($n$ = 353). The proof of methodology provided in this paper could serve for a study on a larger scale. Additionally, due to the small sample size of individual pitchers, we were not able to perform the classification of pitch types per individuals. The data from the pitchers have a hierarchical structure, suggesting that pitching mechanics [22] as well as pitch kinematics [24] among different throws are more similar for an individual pitcher compared to others. Therefore, it may be sensible to classify pitch types for individual pitchers. Pitch type prediction by pitch count and by pitcher showed improved performance in the prediction of the next pitch the pitcher will throw based on features available from the previous throws [16, 17, 19]. Our study would have benefited from longitudinal data collection including kinematic data during the full season. This would allow us to perform classification tasks for different pitch types for individual pitchers. Moreover, matching pitching kinematics with ball speed data may also increase the accuracy of the model.

To the best of our knowledge, this is the first study that uses biomechanical data from wearables to predict pitch types, and thus enriches the available data from an easy-to-use motion sensor system. It is important to clarify that this method is proposed for the classification of the pitch thrown and not the prediction of the next pitch. Pitch prediction uses information available prior the pitch to judge which pitch can be expected. However, pitch type classification uses information available post pitch to determine which pitch type was thrown. Previous studies used post-pitch data from PITCHf/x describing the characteristics of the ball from when it leaves the hand of the pitcher until it crosses the home plate [18, 19]. Defining the pitch type from ball flight data is related to the inherent need of redefining pitch types. Traditional pitch type description is not sufficient any longer, with the newly available data in professional pitching. Our methodology aims to expand this knowledge to situations such as youth baseball, where expensive PITCHf/x systems are not prevalent.

The proposed classification method, based on a limited amount of data from youth baseball pitches, shows promising performance in predicting Fastball vs. off-speed pitches. Application of this binary classification method in youth baseball training can create a major advantage for the development of individual players. Since nowadays pitch count is the only variable that is noted, and mostly manually recorded, the automatic tracking of pitch counts, biomechanical data and pitch types can be of great value to coaches and players. Given that youth players are still learning how to throw different pitch types and their susceptibility to injuries is higher when throwing off-speed pitches [2], implementing the proposed methods in baseball practice may provide a wealth of information relevant for both pitchers and coaches in those situations.

Implementing similar technologies for elite athletes' training could benefit from the aforementioned suggestions to improve the accuracy of the multiclass classification model. However, further studies should determine the necessity of such a system since high-level players often have access to other resources that can measure or calculate pitch trajectory. Indirect pitch type prediction may thus not be needed for players at a high level with many resources at their disposal.



## 3.5. CONCLUSIONS

The accessibility of wearable sensors for performance tracking during both training and games represents a new source of large amounts of data that need powerful algorithms for their analysis, resulting in actionable insights relevant for pitchers' performance and injury risk management. This study established machine learning methods for the detection of the pitch type that was thrown based on pitching mechanics recorded with wearables. The Naive Bayes algorithm showed the best performance in the detection of fastballs, whereas the Random Forest algorithm performed best in the multiclass (FB vs. CH vs. CU) classification task. While these findings demonstrate the potential for the utilisation of wearables in baseball pitching, further development of the classification algorithm, as well as longitudinal data collection, is required. Providing insight into pitch count, pitching mechanics and pitch type enables pitchers to throw safely and efficiently. Through automatic tracking of pitch types, every pitch is counted. Thus, monitoring pitching mechanics and providing an informative feedback to the pitchers may lead to safe and efficient pitching and increase a pitcher's chances of success.

## ABBREVIATIONS

The following abbreviations are used in this chapter:

| | |
|---|---|
| FB | Fastball |
| CH | Change-up |
| CU | Curveball |
| KNN | K-Nearest Neighbors |
| NB | Naive Bayes |
| RF | Random Forest |
| SVM | Support Vector Machine |
| LOGREG | Logistic Regression |
| MNOM | Multinomial Logistic Regression |

# REFERENCES


[1] L. Gomaz, C. Bouwmeester, E. van der Graaff, B. van Trigt, and D. Veeger. "Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors". In: *Sensors* 23.23 (2023), p. 9373.

[2] B. Dowling, M. P. McNally, A. M. Chaudhari, and J. A. Oñate. "A review of workload-monitoring considerations for baseball pitchers". In: *Journal of Athletic Training* 55.9 (2020), pp. 911–917.

[3] S. Lyman, G. S. Fleisig, J. R. Andrews, and E. D. Osinski. "Effect of Pitch Type, Pitch Count, and Pitching Mechanics on Risk of Elbow and Shoulder Pain in Youth Baseball Pitchers". en. In: *The American Journal of Sports Medicine* 30.4 (July 2002), pp. 463–468. issn: 0363-5465, 1552-3365. doi: 10.1177/03635465020300040201. url: https://journals.sagepub.com/doi/10.1177/03635465020300040201.

[4] G. S. Fleisig, D. S. Kingsley, J. W. Loftice, K. P. Dinnen, R. Ranganathan, S. Dun, R. F. Escamilla, and J. R. Andrews. "Kinetic Comparison among the Fastball, Curveball, Change-up, and Slider in Collegiate Baseball Pitchers". en. In: *The American Journal of Sports Medicine* 34.3 (Mar. 2006), pp. 423–430. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546505280431. url: http://journals.sagepub.com/doi/10.1177/0363546505280431.

[5] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews. "Baseball Pitching Biomechanics in Relation to Injury Risk and Performance". en. In: *Sports Health: A Multidisciplinary Approach* 1.4 (July 2009), pp. 314–320. issn: 1941-7381, 1941-0921. doi: 10.1177/1941738109338546. url: http://journals.sagepub.com/doi/10.1177/1941738109338546.

[6] J. Davis, O. Limpisvasti, D. Fluhme, K. J. Mohr, L. A. Yocum, N. S. ElAttrache, and F. W. Jobe. "The Effect of Pitching Biomechanics on the Upper Extremity in Youth and Adolescent Baseball Pitchers". en. In: *The American Journal of Sports Medicine* 37.8 (Aug. 2009), pp. 1484–1491. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546509340226. url: http://journals.sagepub.com/doi/10.1177/0363546509340226.

[7] T. Soligard, M. Schwellnus, J.-M. Alonso, R. Bahr, B. Clarsen, H. P. Dijkstra, T. Gabbett, M. Gleeson, M. Hägglund, M. R. Hutchinson, C. J. van Rensburg, K. M. Khan, R. Meeusen, J. W. Orchard, B. M. Pluim, M. Raftery, R. Budgett, and L. Engebretsen. "How much is too much? (Part 1) International Olympic Committee consensus statement on load in sport and risk of injury". In: *British Journal of Sports Medicine* 50.17 (2016), pp. 1030–1041. issn: 0306-3674. doi: 10.1136/bjsports-2016-096581. eprint: https://bjsm.bmj.com/content/50/17/1030.full.pdf. url: https://bjsm.bmj.com/content/50/17/1030.

[8] A. L. Aguinaldo, J. Buttermore, and H. Chambers. "Effects of Upper Trunk Rotation on Shoulder Joint Torque among Baseball Pitchers of Various Levels". en. In: *Journal of Applied Biomechanics* 23.1 (Feb. 2007), pp. 42–51. issn: 1065-8483, 1543-2688. doi: 10.1123/jab.23.1.42. url: https://journals.humankinetics.com/view/journals/jab/23/1/article-p42.xml.

[9] C. A. Putnam. "Sequential motions of body segments in striking and throwing skills: Descriptions and explanations". en. In: *Journal of Biomechanics* 26 (Jan. 1993), pp. 125–135. issn: 00219290. doi: 10.1016/0021-9290(93)90084-R. url: https://linkinghub.elsevier.com/retrieve/pii/002192909390084R.

[10] E. Van Der Graaff, M. ( Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. ( Veeger. "Timing of peak pelvis and thorax rotation velocity in baseball pitching". en. In: *The Journal of Physical Fitness and Sports Medicine* 7.5 (Sept. 2018), pp. 269–277. issn: 2186-8123, 2186-8131. doi: 10.7600/jpfsm.7.269. url: https://www.jstage.jst.go.jp/article/jpfsm/7/5/7_269/_article.

[11] S. Dun, J. Loftice, G. S. Fleisig, D. Kingsley, and J. R. Andrews. "A Biomechanical Comparison of Youth Baseball Pitches: Is the Curveball Potentially Harmful?" en. In: *The American Journal of Sports Medicine* 36.4 (Apr. 2008), pp. 686–692. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546507310074. url: http://journals.sagepub.com/doi/10.1177/0363546507310074.

[12] R. F. Escamilla, G. S. Fleisig, S. W. Barrentine, N. Zheng, and J. R. Andrews. "Kinematic Comparisons of Throwing Different Types of Baseball Pitches". en. In: *Journal of Applied Biomechanics* 14.1 (Feb. 1998), pp. 1–23. issn: 1065-8483, 1543-2688. doi: 10.1123/jab.14.1.1. url: https://journals.humankinetics.com/view/journals/jab/14/1/article-p1.xml.

[13] R. F. Escamilla, G. S. Fleisig, D. Groeschner, and K. Akizuki. "Biomechanical Comparisons Among Fastball, Slider, Curveball, and Changeup Pitch Types and Between Balls and Strikes in Professional Baseball Pitchers". en. In: *The American Journal of*

3

3

*Sports Medicine* 45.14 (Dec. 2017), pp. 3358–3367. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546517730052. url: http://journals.sagepub.com/doi/10.1177/0363546517730052.

[14] G. S. Fleisig, W. A. Laughlin, K. T. Aune, E. L. Cain, J. R. Dugas, and J. R. Andrews. "Differences among fastball, curveball, and change-up pitching biomechanics across various levels of baseball". en. In: *Sports Biomechanics* 15.2 (Apr. 2016), pp. 128–138. issn: 1476-3141, 1752-6116. doi: 10.1080/14763141.2016.1159319. url: http://www.tandfonline.com/doi/full/10.1080/14763141.2016.1159319.

[15] P. Hoang. "Supervised Learning in Baseball Pitch Prediction and Hepatitis C Diagnosis". PhD thesis. North Carolina State University, 2015. doi: 10.13140/RG.2.2.36217.72800.

[16] G. Ganeshapillai and J. V. Guttag. "Predicting the Next Pitch". In: *MIT Sloan Sports Analytics Conference*. 2012.

[17] G. Sidle and H. Tran. "Using multi-class classification methods to predict baseball pitch types". en. In: *Journal of Sports Analytics* 4.1 (Feb. 2018), pp. 85–93. issn: 2215020X, 22150218. doi: 10.3233/JSA-170171. url: https://www.medra.org/servlet/aliasResolver?alias=iospress&doi=10.3233/JSA-170171.

[18] A. Attarian, G. Danis, J. Gronsbell, G. Iervolino, and H. Tran. "A Comparison of Classification Methods with an Application to Classifying Baseball Pitches". In: *International MultiConference of Engineers and Computer Scientists*. 2013.

[19] M. A. Pane, S. L. Ventura, R. C. Steorts, and A. C. Thomas. *Trouble With The Curve: Improving MLB Pitch Classification*. en. arXiv:1304.1756 [stat]. Apr. 2013. url: http://arxiv.org/abs/1304.1756.

[20] M. Hamilton., P. Hoang., L. Layne., J. Murray., D. Padget., C. Stafford., and H. Tran. "Applying Machine Learning Techniques to Baseball Pitch Prediction:" en. In: *3rd International Conference on Pattern Recognition Applications and Methods*. ESEO, Angers, Loire Valley, France: SCITEPRESS - Science, 2014, pp. 520–527. isbn: 978-989-758-018-5. doi: 10.5220/0004763905200527. url: http://www.scitepress.org/DigitalLibrary/Link.aspx?doi=10.5220/0004763905200527.

[21] J. Bock. "Pitch Sequence Complexity and Long-Term Pitcher Performance". en. In: *Sports* 3.1 (Mar. 2015), pp. 40–55. issn: 2075-4663. doi: 10.3390/sports3010040. url: http://www.mdpi.com/2075-4663/3/1/40.

[22] L. Gomaz, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. “Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data”. en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: 10.3390/s21227442. url: https://www.mdpi.com/1424-8220/21/22/7442.

[23] M. Kuhn. “Building Predictive Models in *R* Using the **caret** Package”. en. In: *Journal of Statistical Software* 28.5 (2008). issn: 1548-7660. doi: 10.18637/jss.v028.i05. url: http://www.jstatsoft.org/v28/i05/.

[24] K. Umemura, T. Yanai, and Y. Nagata. “Application of VBGMM for pitch type classification: analysis of TrackMan’s pitch tracking data”. en. In: *Japanese Journal of Statistics and Data Science* 4.1 (July 2021), pp. 41–71. issn: 2520-8756, 2520-8764. doi: 10.1007/s42081-020-00079-8. url: https://link.springer.com/10.1007/s42081-020-00079-8.

# A

# APPENDIX

# A.1. BINARY CLASSIFICATION



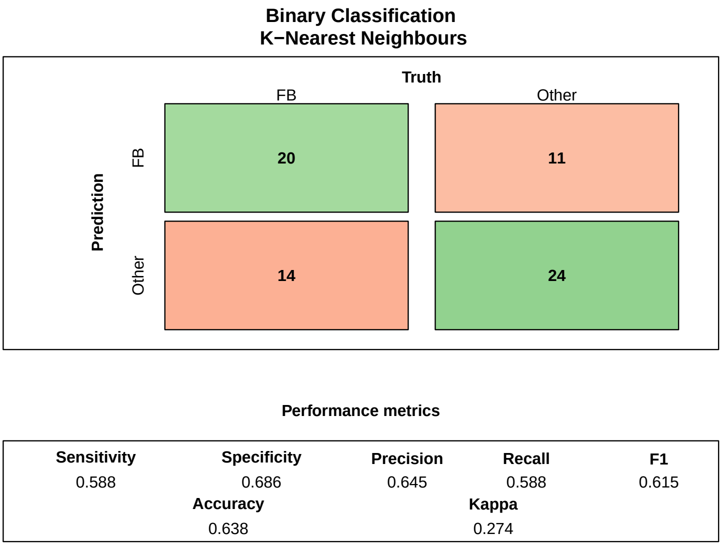


Figure A.1.: Confusion matrix for binary K-Nearest Neighbors algorithm.

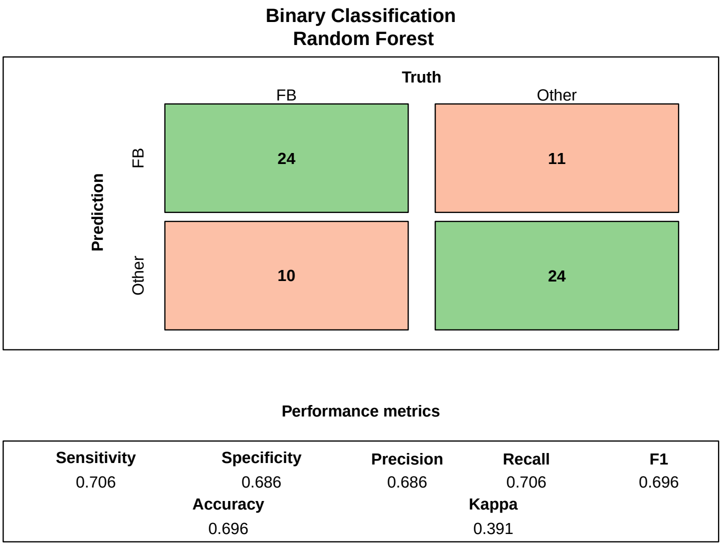


Figure A.2.: Confusion matrix for binary Random Forest algorithm.

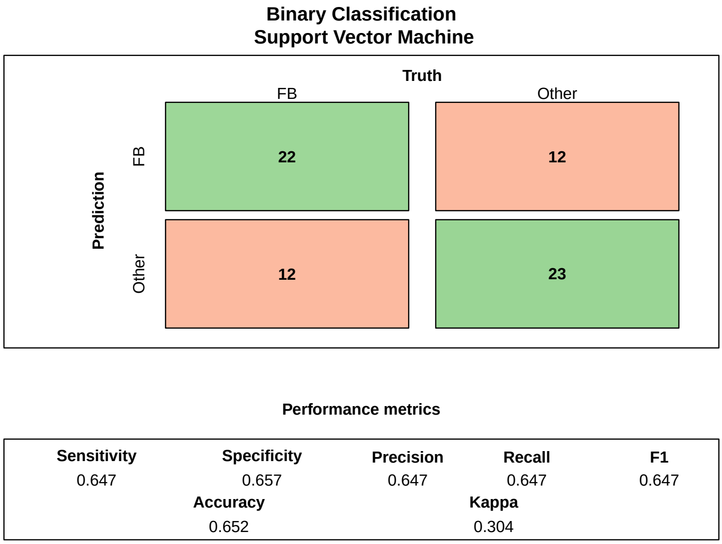


Figure A.3.: Confusion matrix for binary Support Vector Machine algorithm with radial basis kernel function.

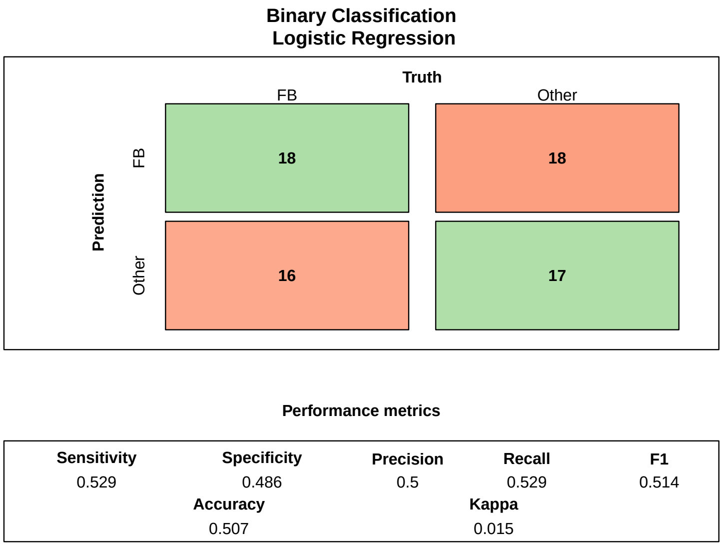


Figure A.4.: Confusion matrix for binary Logistic Regression algorithm with radial basis kernel function.

## A.2. MULTICLASS CLASSIFICATION

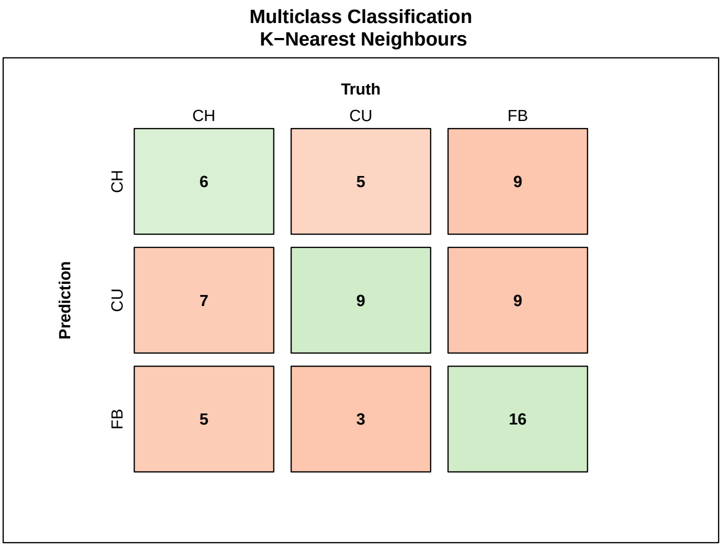


Figure A.5.: Confusion matrix for multiclass K-Nearest Neighbors algorithm.

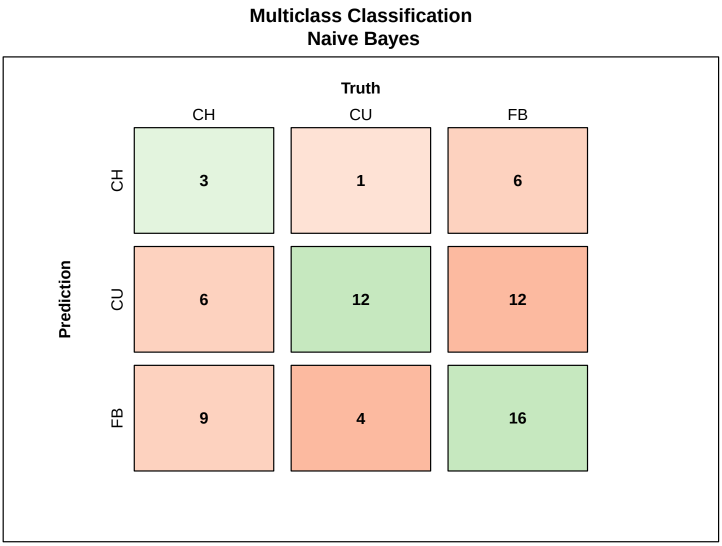


Figure A.6.: Confusion matrix for multiclass Naive Bayes algorithm.

A

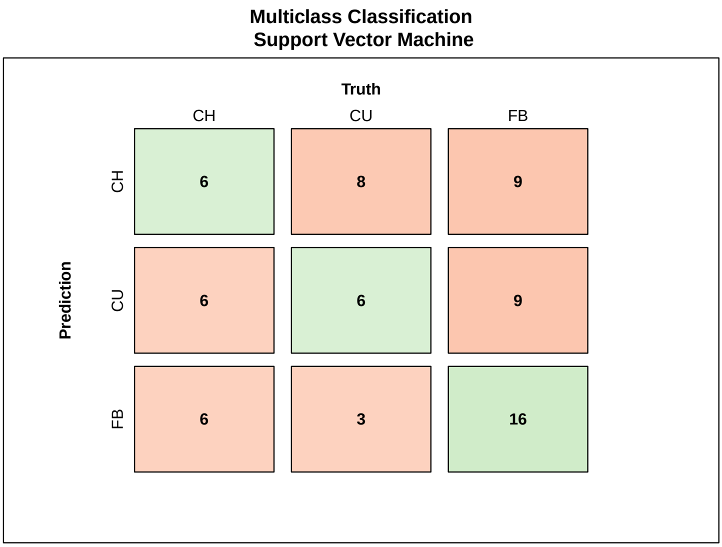


Figure A.7.: Confusion matrix for multiclass Support Vector Machine algorithm with radial basis kernel function.

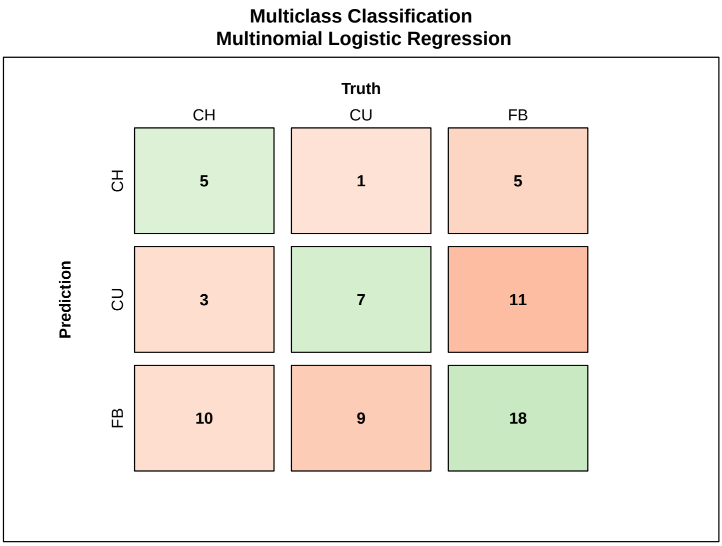


Figure A.8.: Confusion matrix for multiclass Multinomial Logistic Regression algorithm with radial basis kernel function.

## A.3. VARIABLE IMPORTANCE

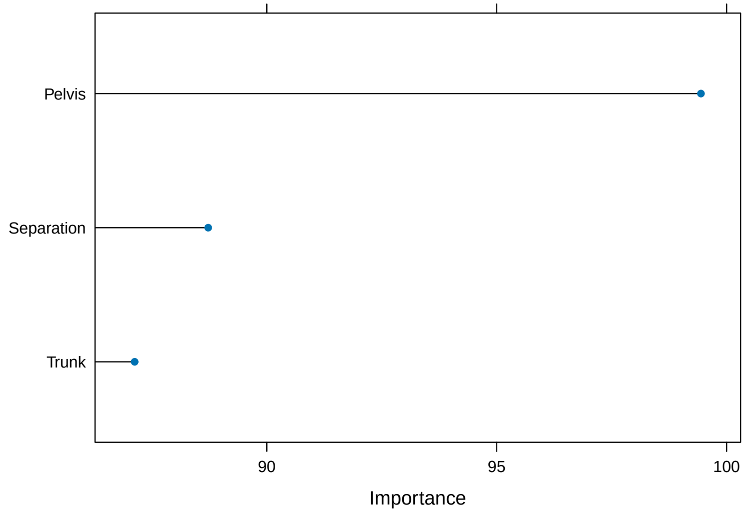


Figure A.9.: Visual representation of the feature importance for Random Forest multiclass classifier calculated with `varImp` from `caret` package. The horizontal axis should be interpreted as a measure for relative importance of predictive variables. The figure reveals *Pelvis* to be considered as the most important for the multiclass classification task, whereas *Trunk* is considered as the least important.

# 4

# PREDICTING ELBOW LOAD BASED ON INDIVIDUAL PELVIS AND TRUNK (INTER)SEGMENTAL ROTATIONS IN FASTBALL PITCHING

**Larisa van der Graaff [+], Bart van Trigt [+], Frank van der Meulen, DirkJan Veeger**

+ these authors contributed equally to this work

Published in *Sports Biomechanics* **2024** [1].

## 4.1. INTRODUCTION

The baseball pitch is a full-body throwing motion that, due to its repetitive nature, exposes the elbow to significant loads [2, 3]. This leads to a high incidence of overused elbow injuries among baseball pitchers at all levels of play [4–6]. The injury aetiology seen in youth and adult pitchers has been linked to high elbow external valgus torques [7, 8]. The external valgus torque imparts a tensile force to the medial elbow structures [9, 10], which in combination with repetitive loading results in injuries to the medially located ulnar collateral ligament (UCL). This indicates that external valgus torque can be used as a proxy of elbow load [11, 12]. Thus, continuous and prospective elbow load monitoring, both in training and in game, plays an essential role in pitchers' performance enhancement whilst minimizing the risk of elbow injuries [13].



To assess the external valgus torque, it is important to understand pitching mechanics. Pitching mechanics can be described by the two well-known biomechanical principles; the summation of speed principle, also known as the kinetic chain, and the principle of optimal coordination of partial momenta [14]. Both principles consider the human body as a linked segment model and explain the biomechanics of pitching in terms of peak angular velocities of body segments and their intersegmental timing. Overhead throwing motion, such as baseball pitching, is more likely to follow the kinetic chain [14]. Regardless of the principle, the high end-point velocities imparted to the ball depend on the contribution of all segments [15].

In the pitching motion, energy is generated in the driving leg and transferred through the stride leg to the pelvis [16]. While part of the energy in the pelvis is transferred back to the stride leg to form a stable base around which the pelvis and trunk can rotate [17], most energy is transferred via the trunk up to the throwing arm [3]. In such complex sequential movement, pelvic and trunk kinematics play an essential role in transferring the momentum generated by the lower extremities to the upper extremity. Optimal proximal-to-distal timing between the pelvis and trunk results in the maximized ball velocity at the most distal end [3, 14]. The timing between the pelvis and trunk peak angular velocities is also referred to as separation time. If this kinematic sequencing or timing is not optimal, energy is dissipated into the upper extremity which results not only in decreased ball velocity [14, 18], but also the potentially increased risk of injuries [19].

Manipulation of biomechanical parameters within the kinetic chain may affect the external valgus torque and help in managing the risk of excessive UCL loading. By increasing trunk peak angular velocity, pitchers may throw faster, but with an increased external valgus torque [20]. There is likely a threshold above which the exceeded external valgus torque represents a significant injury risk. The efficiency of the kinetic chain may contribute to the reduction of external valgus torque levels at this critical point while still maintaining high levels of ball speed and overall pitching performance [11].

We expect that the levels of external valgus torque will differ between pitchers due to variations in anthropometric measures, pitching technique, level of play, and within-individual load variability [21, 22]. Multilevel modelling is well-suited for the analysis of repeated measurements that are considered to be “clustered” within individual pitchers [23]. Such measurements are assumed to be independent as the observations within a cluster are more likely to be similar than observations from different clusters. Since regression- and ANOVA-based techniques do not meet this assumption, they are not fully appropriate for dealing with this type of data structure. Multilevel modelling techniques for repeated measurements allow us to analyse the relationships between data collected at the pitcher- or group-level, and data collected on variables that change with trials at the unit- or individual-level [24].

The aim of the study is to contribute to monitoring the external valgus torque in baseball pitching by developing a prediction model based on the pelvis and trunk peak angular velocities and their separation time. It is hypothesized that external valgus torque for an individual pitcher can be predicted based on the pelvis and trunk peak angular velocity and separation time between them. In addition, we expect that the model including both pelvis and trunk peak angular velocity and their separation time will have the best predictive performance.



## 4.2. MATERIALS AND METHODS

### 4.2.1. PARTICIPANTS

Eleven male Dutch national (AAA) youth elite baseball pitchers participated in the study, with a mean age of 17.4 (± 2.2) years, mean body mass of 80.6 (± 11.7) kg, mean body height of 1.86 (± 6.3) m and mean ball speed was 34.0 ± 1.4 m/s (76.6 ± 3.2 mph). Only participants without present musculoskeletal injuries and who did not have musculoskeletal injuries in the last six months were included in this study. Participants gave written consent to use the data information for analysis and publication after being fully informed. If participants were under 16 years, their parents or guardians were informed about the study and required to sign an informed consent form. This research was conducted as part of a larger study [22] and was performed in accordance with the Declaration of Helsinki and the local ethics committee. The local ethics committee of the Faculty of Behavioral and Movement Sciences (VCWE) approved the study protocol (reference number: VCWE2019-033).

### 4.2.2. PROCEDURE

Data collection was performed in an indoor movement laboratory at the Royal Netherlands Football Association. The participants wore sneakers, athletic stretch shorts, catching gloves, and no shirts. Forty-three reflective markers were attached with double-sided tape on the bony landmarks. Participants performed their regular warming-up, which contained stretching, drills, and several warming-up pitches. Subsequently, they threw several pitches from the mound to become familiar with the research setup. The participants were instructed to throw 25 fastball pitches at full effort toward a squared strike zone (height 0.64m; width 0.38m). The pitching rubber was attached to the top of the mound at 0.55m above the ground and had a distance of 18.44 m to the home plate. The time between each pitch was not controlled but regulated by the pitcher himself, like in a normal game.

### 4.2.3. DATA ACQUISITION

Full body position data of the pitchers were collected with a VICON eight-camera motion capture system. Data were sampled at 400Hz (model V5; Vicon Motion Systems Ltd., Yarnton, UK). The ball speed was measured with a radar gun positioned next to the strike zone (Stalker Radar, Plano, TX, USA).

### 4.2.4. DATA PROCESSING

Three-dimensional position data of the fourteen bony landmarks were used in this study (Table 4.1). The position data were interpolated with a third-order cubic spline polynomial and filtered with a fourth–order Butterworth filter with a cut-off frequency of 12.5 Hz. To calculate the segment angular velocities and the elbow valgus torque an anatomical coordinate system was constructed for the pelvis, trunk, upper arm, forearm, and hand according to the ISB recommendations [25].

The segment angular velocities were computed directly from the rotation matrices following the method described in the study of Zatsiorsky [26]. Subsequently, the Euclidean norm was calculated over all three different axes. The exact moments of peak angular velocities were found analytically by fitting a second-order polynomial function to eleven measured data points. These data points included five samples before and after the samples closest to the maximum angular velocity. The separation time was calculated as the time interval between the pelvis and trunk peak angular velocities [18].

Elbow joint torques were calculated based on the top-down method based on the Newton–Euler equation of motion, starting in the hand of the throwing arm. The segment center of mass position and the moments of inertia were estimated according to Zatsiorsky [26] and De

Leva et al. [27]. The baseball was modelled with a mass of 0.145kg attached to the hand. The mass linearly reduced by 10% over the last ten samples (0.025s) before ball release. Ball release was defined as the moment the wrist exceeded the position of the elbow in the forward direction. The elbow joint coordinate system was expressed in the anatomical coordinate system of the forearm, located in the middle between the medial and lateral humeral epicondyle. The time series of external elbow valgus torque was determined for each individual pitch, covering the duration from foot contact to ball release. Subsequently, the peak external valgus torque was derived from this time series data. The time series of the segment angular velocities and external valgus torque were visually checked for errors and mistakes.



| Bony landmarks |
| --- |
| (1) third proximal interphalangeal |
| (2) ulnar process styloid |
| (3) radial process styloid |
| (4) lateral humeral epicondyle |
| (5) medial humeral epicondyle |
| (6) acromion |
| (7) xiphoid process |
| (8) incisura jugularis |
| (9) 7th cervical vertebrae |
| (10) 8th thoracal vertebrae |
| (11 & 12) anterior superior iliac spine |
| (13 & 14) posterior superior iliac spine |

Table 4.1.: Bony landmarks used in the study.

### 4.2.5. STATISTICAL METHODS AND MODELING

For the $i$-th throw, let $y_i$, $x_{i1}$, $x_{i2}$, $x_{i3}$, $x_{i4}$ and $x_{i5}$ denote the external valgus torque, pelvis angular velocity, trunk angular velocity, separation time, weight and height respectively. Set $\boldsymbol{x}_i = (x_{i1}, x_{i2}, x_{i3})$ and $\boldsymbol{u}_i = (x_{i4}, x_{i5})$. We aim to model the relationship between $y_i$ and $(\boldsymbol{x}_i, \boldsymbol{u}_i)$. The simplest type of model for this is the linear model given by

$$y_i \mid \beta_0, \beta, \sigma^2 \overset{\text{ind}}{\sim} N(\beta_0 + \boldsymbol{\beta}'\boldsymbol{x}_i + \boldsymbol{\gamma}'\boldsymbol{u}_i, \sigma^2) \tag{4.1}$$

where the symbol $\overset{\text{ind}}{\sim}$ denotes "independently distributed as". However, note that the data from repeated measurements such as in this study have the structure in which observations on an individual-level (pelvis and trunk peak angular velocities, separation time, external valgus torque) are nested within baseball pitchers on a group-level. As such, a simple linear model like (4.1) won't be able to take into account that throws by the same pitcher tend to be more similar than throws by different pitchers. This phenomenon is illustrated in Figure 4.1, where we have also included weight and height to see how external valgus torque is affected by these characteristics. This figure strongly suggests



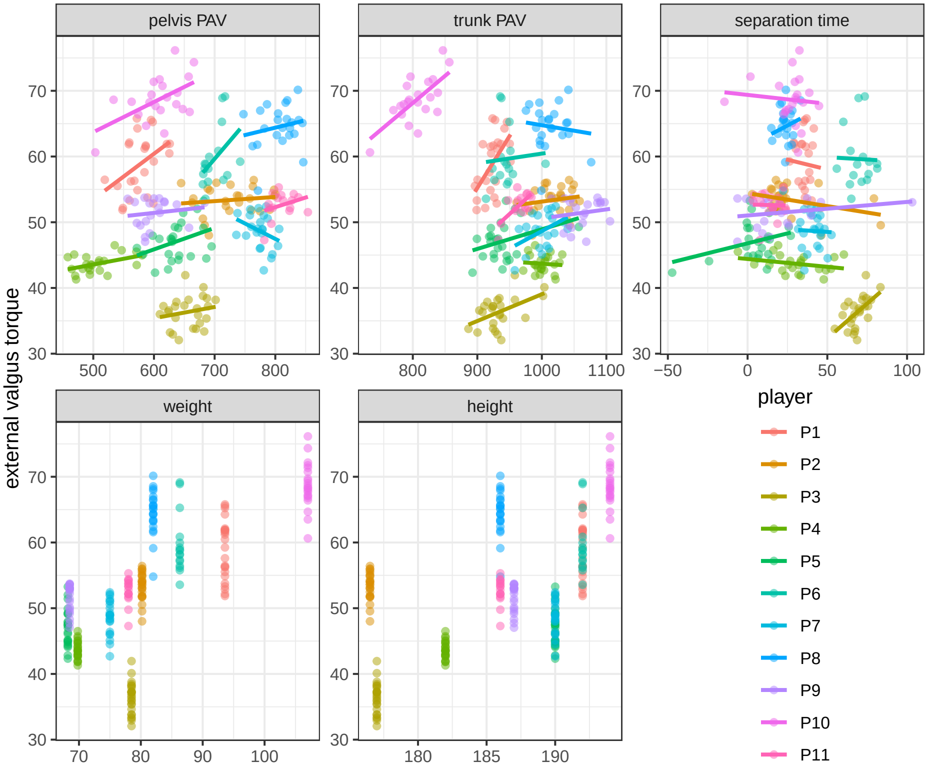


Figure 4.1.: Exploratory data analysis for the relation between external valgus torque and pelvis peak angular velocity (pelvis PAV), trunk peak angular velocity ttrunk PAV) and separation time. In each subpanel, the influence of one predictor on external valgus torque is displayed. In the upper three panels, least-squares fits have been superimposed (separately, for each player).

a *two–level linear model, with both varying intercepts and varying slopes*. The need for such a model is most easily seen from the panel with "Trunk PAV". If we would fit a single line through the data, this would imply a negative relationship between external valgus torque and trunk peak angular velocity (Trunk PAV), whereas for each individual player, this relationship is positive. This can be seen as an instance of Simpson's paradox, well known in statistics. Specifically, we propose the following model:

$$\begin{aligned} y_{\mathrm{i}} \mid \alpha_1, \ldots, \alpha_{\mathrm{J}}, \beta_1, \ldots, \beta_{\mathrm{J}}, \sigma^2 &\overset{\text{ind}}{\sim} N(\mu_{\mathrm{i}} + \gamma' \boldsymbol{u}_{\mathrm{i}}, \sigma^2) \\ \mu_i &= \alpha_{\mathrm{j[i]}} + \boldsymbol{\beta}'_{\mathrm{j[i]}} \boldsymbol{x}_{\mathrm{i}} \end{aligned} \tag{4.2}$$

We have $J = 11$, the total number of pitchers in the study, and $j[i] = k$ if the $i$-th throw corresponds to $k$-th pitcher in the data set. We follow the Bayesian approach to statistics, where unobserved quantities get assigned a prior distribution, reflecting the (lack of) information we have about their values before collecting the data. We impose $\alpha_1, \ldots, \alpha_{\mathrm{J}} \overset{\text{iid}}{\sim} N(0, \sigma^2_\alpha)$, $\beta_1, \ldots, \beta_{\mathrm{J}} \overset{\text{iid}}{\sim} N_3(0, \sigma^2_\beta I_{3\times 3})$ and $\gamma \sim N(0, \sigma_\gamma I_2)$. The symbol $\overset{\text{iid}}{\sim}$ denotes "independent and identically distributed as". We took default values from `rstanarm` [28], which means $\sigma_\alpha = \sigma_\beta = \sigma_\gamma = 2.5$. Taking mean-zero priors is justified as we standardized (i.e. transformed to zero-mean and unit standard deviation) each of the predictors before fitting the model. Also for $\sigma$, $\sigma_\alpha$ and $\sigma_\beta$ we took the default prior mean-one Exponential distribution from `rstanarm` [28].

We used leave–one–group–out cross–validation (LOGO–CV) to select the model with the best predictive performance. LOGO–CV is a specific type of k–fold cross–validation that utilizes data from each individual pitcher as a test set. The number of folds, therefore, equals the number of pitchers. For every fold, the model is trained on data from J-1 pitchers and tested on the data from the one left-out pitcher. Models were compared according to their expected log-predictive density (*elpd*) as described in the work of Vehtari [29, 30].

We used posterior predictive distributions to generate data samples whose average is then compared to the real data. We interpret the generated data as the data sample that we might collect tomorrow if the data collection process remains the same as it initially was. Posterior predictive checks were used to test the performance of the model and visually inspect how much generated data samples match the observed ones.

## 4.3. RESULTS

A total of 240 throws by 11 pitchers were included in the analysis. The number of pitches varied from 19 to 25 throws per pitcher. Descriptive statistics of included variables are shown in Table 4.2.

Expected log–predictive density (*elpd*) was a chosen measure of model fit and it was subsequently used to compare models for model selection. The difference in *elpd* of the fitted two-level varying–intercept, varying–slope Bayesian models is shown in Figure 4.2. Models include

| Variables | Mean ± Standard Deviation |
|---|---|
| Pelvis peak angular velocity [°/s] | 669.87 ± 99.06 |
| Trunk peak angular velocity [°/s] | 964.85 ± 68.61 |
| Separation time [ms] | 32.70 ± 22.98 |
| Weight [kg] | 80.47 ± 11.11 |
| Height [cm] | 186.26 ± 5.85 |
| External valgus torque [Nm] | 52.76 ± 9.59 |

Table 4.2.: Descriptive statistics for the variables included in the analysis.



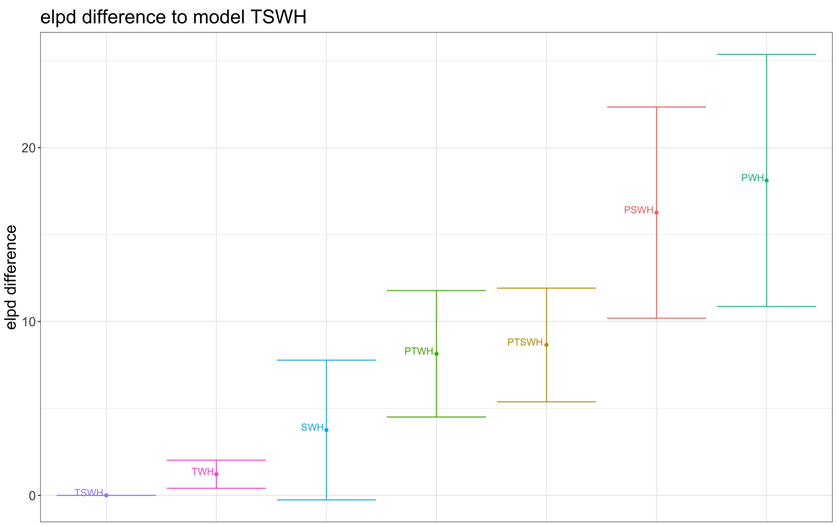


Figure 4.2.: Estimates of absolute expected log–predictive density (*elpd*) difference (dot) using leave–one–group–out cross–validation. Vertical error bar for each model indicates the standard error of the *elpd* difference estimates. The order on the x–axis follows the ranking starting with the model with best predictive performance on the left. Predictors included in the analysis are pelvis peak angular velocity (P), trunk peak angular velocity (T), separation time (S), pitcher's weight (W) and height (H).

various combinations of observed kinematic predictors (P – pelvis peak angular velocity, T – trunk peak angular velocity, S – separation time) with the addition of pitcher's weight (W) and height (H) to all the models. The ordering of the models in Figure 4.2 reveals that the model including a set of predictors TSWH showed the best predictive performance, and it is therefore the selected model. Table 4.3 shows parameter estimates from the selected model TSWH, based on a table generated by `shinystan` [31]. The small *elpd* differences between the selected model TSWH and the second ranked model TWH indicate almost similar performance in predicting external valgus torque.

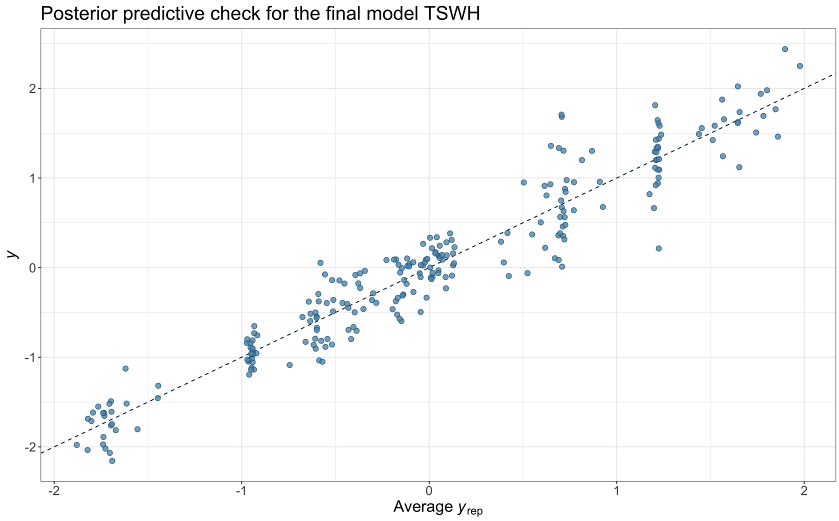


Figure 4.3.: Posterior predictive checks compare the observed outcome variable $y$ to the average of simulated datasets $y_{\text{rep}}$ from the posterior predictive distribution for the selected model TSWH. The model includes a set of predictors of trunk peak angular velocity (T), separation time (S), pitcher's weight (W), and height (H). Bayesian conditional $R^2$ value is 0.916 (95% CI [0.899, 0.931]), and the marginal $R^2$ value 0.927 (95% CI [0.847, 0.969]), where CI is a confidence interval. The marginal $R^2$ considers only the variance of the fixed effects, while the conditional $R^2$ takes both the fixed and random effects into account [32].

The performance of the final model TSWH was tested through a posterior predictive check. In Figure 4.3 the average of the data samples generated from the posterior predictive distributions is compared to the observed data. If the model is a good fit for the data, then observed and simulated data should be aligned. The posterior predictive check shows that the observed data are more dispersed compared to the average of the generated data samples from the posterior predictive distributions. Bayesian conditional $R^2$ value is 0.916 (95% CI [0.899, 0.931]), and marginal $R^2$ value 0.927 (95% CI [0.847, 0.969]), where CI is a confidence interval. The marginal $R^2$ considers only the variance of the fixed effects, while the conditional $R^2$ takes both the fixed and random effects into account [32].

## 4.4. DISCUSSION AND IMPLICATIONS

Poor pitching mechanics [33] and overloading of the pitching arm can negatively affect pitching performance and at the same time put the elbow joint at great risk of injuries [12, 34]. Therefore, estimation of the external valgus torque based on pitching mechanics is an important step toward monitoring the elbow load in the field. This study shows promising results of Bayesian hierarchical models in the prediction of the external valgus torque, used as a proxy of elbow load, based on (inter)segmental rotation in fastball pitching.



The results show that it is possible to predict the elbow external valgus torque based on the pelvis and trunk kinematics and separation time. Although it was hypothesized that the model including all three parameters would have the best performance, according to LOGO-CV the best predictive model is TSWH which includes peak trunk angular velocity, separation time, weight, and height (Bayesian conditional $R^2$ value is 0.916, marginal $R^2$ value is 0.927). The reason why the pelvis angular velocity was not included in the final model might be explained by the fact that the trunk angular velocity contains information from the proximal pelvis segment according to the proximal-to-distal sequence. The contribution of the separation time to the prediction of the external valgus torque indicates the importance of optimal timing between the pelvis and trunk segments in the kinetic chain for safe and efficient pitching. However, it is yet unknown what the “optimal” separation time is. The results in this study showed that certain pitchers exhibited a positive correlation between separation time and external valgus torque, while others demonstrated a negative or no correlation (Figure 4.1). Oyama et al. [35] did not find a relationship between the separation time and external valgus torque on group level. This might indicate that the optimal timing is individually depended, with the proviso that the pelvis and trunk are in sequence. The trunk can produce a lot of power due to its segmental mass, although proper timing is needed for optimal contribution to the ball speed [36, 37]. The increase in trunk rotation does not only increase the ball speed, but it increases the external valgus torque as well [20]. In line with our results, several studies showed a relationship between trunk kinematics and the external valgus torque [7, 20, 36]. In addition, we showed that it is possible to predict the external valgus torque for individual pitchers based on their trunk peak angular velocity and the separation time.

Predictions of the external valgus torque based on the trunk peak angular velocity and the separation time are important in relation to elbow injuries. Manipulation of these biomechanical parameters with training increases the ball speed [38] and may decrease the external valgus torque [20]. However, a pitcher throwing according to an optimal kinetic chain, with a reduced level of external valgus torque is still at risk of sustaining an injury due to repetitive pitching. Therefore, monitoring the external valgus torque is important for managing the risk of excessive elbow loading. Taking into account that the values of external valgus torque vary among pitchers of different ages, levels of play [21], and the variability within-individual pitchers [22], understanding the elbow loading for each pitcher based on his individual characteristics and pitching mechanics may be the base for the development of an "early warning system" for safe and efficient pitching.

This paper introduces the application of Bayesian hierarchical models to repeated measurements of pitching kinematic and temporal parameters. Such models account for the within–pitcher similarity and at the same time allow for the gradation of differences between the pitchers in the prediction of the external valgus torque. The small difference in *elpd* between the selected model TSWH and the model TWH ranked second in terms of LOGO–CV refers to their similar predictive performance (Figure 4.2). In addition, posterior predictive checks reveal similar model fit for the TWH model compared to the TSWH model. From the practical point of view, this means that monitoring external valgus torque is already possible based on the single kinematic variable (trunk peak angular velocity). However, the separation time is related to the efficiency of the kinetic chain and its breakdown may be an indicator of the fatigue [39]. Therefore, considering the practical relevance of both parameters for elbow load monitoring over a longer period, we select the predictive model including trunk peak angular velocity and separation time as the final one. The comparison between the Bayesian and frequentist approach to multilevel analysis and fitting the final TSWH model is discussed in the Appendix B.

One of the limitations of this study is the inclusion of only fastball pitches. Studies have shown that the elbow load is lower in the change-up or breaking balls [40], however, the link between the torso kinematics and elbow load has not been investigated yet. Furthermore, the current study had a very low sample size ($n$ = 11) and included repeated measurements from a single data collection event. The low sample size could affect Bayesian mixed models in terms of overfitting and imprecise inferences. However, the selected model performance criteria based on *elpd* can help mitigate these issues. The lack of longitudinal data collection limits the detection of patterns in elbow loading based on pitching mechanics. A larger data sample including a wider range of age groups and levels of play may improve the predictive performance and lower the uncertainty in predicted external valgus torque. Collecting longitudinal data, including reported injuries, would allow us to link the loading on the elbow joint to injury occurrence in individual pitchers. This information can be used as a base for setting a pitcher's injury threshold. If the elbow loading exceeds the estimated threshold, the pitcher will likely have increased injury risk. Such information may help coaches in training subscription and modification of the pitching technique that leads to reducing the external valgus torque and therefore the risk of elbow injury.

The final model proposed in this paper considered the practical relevance of trunk kinematics and separation time between the pelvis and trunk in managing injury risk and shows its potential utilization for elbow load monitoring on the field. Trunk peak angular velocity and the separation time can be recorded with wearable sensors, like inertial measurement units [23, 41]. Such data recorded with sensors may be used as input for the proposed model and provide actionable insight for injury prevention in baseball pitching.



## 4.5. CONCLUSION

In this study, a model has been proposed to predict elbow load based on the pelvis and trunk peak angular velocities and separation time between them. Application of Bayesian hierarchical models on data including the trunk peak angular velocity and the separation time between the pelvis and trunk peak angular velocities show promising results for the prediction of the external valgus torque in fastball pitching. Such an approach allows individualized prediction of the external valgus torque for each pitcher, which has a great practical advantage compared to group–based predictions in terms of injury assessment and injury prevention.

Table 4.3: Parameter estimates for the final model TSWH. Predictors are trunk peak angular velocity (Trunk_PAV), separation time (Separation), pitcher's weight (Weight) and height (Height). The standard deviation of the errors is called `sigma` and the variance–covariance matrix of the pitcher–specific deviations from the common parameters is called `Sigma`.

| | **mean** | **sd** | **2.5%** | **25%** | **50%** | **75%** | **97.5%** |
|---|---|---|---|---|---|---|---|
| Weight | 0.8 | 0.2 | 0.3 | 0.6 | 0.8 | 0.9 | 1.2 |
| Height | 0.4 | 0.2 | 0 | 0.3 | 0.4 | 0.5 | 0.8 |
| b [(Intercept) Participant: P1] | -0.3 | 0.3 | -0.8 | -0.4 | -0.3 | -0.1 | 0.3 |
| b [(Intercept) Participant: P10] | 0.3 | 0.5 | -0.6 | 0 | 0.3 | 0.6 | 1.2 |
| b [(Intercept) Participant: P11] | 0.2 | 0.1 | -0.1 | 0.1 | 0.2 | 0.2 | 0.4 |
| b [(Intercept) Participant: P2] | 0.6 | 0.4 | -0.1 | 0.4 | 0.6 | 0.9 | 1.4 |
| b [(Intercept) Participant: P3] | -0.8 | 0.4 | -1.5 | -1 | -0.8 | -0.6 | -0.1 |
| b [(Intercept) Participant: P4] | 0.1 | 0.2 | -0.4 | -0.1 | 0.1 | 0.2 | 0.5 |
| b [(Intercept) Participant: P5] | 0.1 | 0.3 | -0.5 | -0.1 | 0.1 | 0.3 | 0.8 |
| b [(Intercept) Participant: P6] | 0 | 0.2 | -0.5 | -0.2 | -0.1 | 0.1 | 0.4 |
| b [(Intercept) Participant: P7] | -0.4 | 0.2 | -0.9 | -0.6 | -0.4 | -0.3 | 0 |
| b [(Intercept) Participant: P8] | 1.1 | 0.1 | 0.9 | 1.1 | 1.1 | 1.2 | 1.4 |
| b [(Intercept) Participant: P9] | 0.5 | 0.3 | -0.1 | 0.3 | 0.5 | 0.7 | 1.1 |
| b [Trunk_PAV Participant: P1] | 0.7 | 0.3 | 0.2 | 0.5 | 0.7 | 0.8 | 1.2 |
| b [Separation Participant: P1] | 0 | 0.1 | -0.3 | -0.1 | 0 | 0 | 0.2 |
| b [Trunk_PAV Participant: P10] | 0.4 | 0.1 | 0.1 | 0.3 | 0.4 | 0.5 | 0.7 |

Continued on next page

4

Table 4.3: Parameter estimates for the final model TSWH. Predictors are trunk peak angular velocity (Trunk_PAV), separation time (Separation), pitcher's weight (Weight) and height (Height). The standard deviation of the errors is called `sigma` and the variance–covariance matrix of the pitcher–specific deviations from the common parameters is called `Sigma`. (Continued)

| | **mean** | **sd** | **2.5%** | **25%** | **50%** | **75%** | **97.5%** |
|---|---|---|---|---|---|---|---|
| b [Separation Participant: P10] | 0 | 0.1 | -0.2 | -0.1 | 0 | 0 | 0.1 |
| b [Trunk_PAV Participant: P11] | 0.3 | 0.2 | -0.1 | 0.1 | 0.3 | 0.5 | 0.8 |
| b [Separation Participant: P11] | 0 | 0.1 | -0.2 | -0.1 | 0 | 0 | 0.2 |
| b [Trunk_PAV Participant: P2] | 0.1 | 0.1 | -0.2 | 0 | 0.1 | 0.2 | 0.4 |
| b [Separation Participant: P2] | 0 | 0.1 | -0.2 | -0.1 | 0 | 0 | 0 |
| b [Trunk_PAV Participant: P3] | 0.3 | 0.1 | 0 | 0.2 | 0.2 | 0.3 | 0.5 |
| b [Separation Participant: P3] | 0 | 0.1 | -0.2 | 0 | 0 | 0 | 0.2 |
| b [Trunk_PAV Participant: P4] | 0 | 0.2 | -0.4 | -0.1 | 0 | 0.1 | 0.4 |
| b [Separation Participant: P4] | 0 | 0.1 | -0.1 | -0.1 | 0 | 0 | 0.1 |
| b [Trunk_PAV Participant: P5] | 0.2 | 0.1 | -0.1 | 0.1 | 0.2 | 0.2 | 0.4 |
| b [Separation Participant: P5] | 0 | 0.1 | -0.1 | 0 | 0 | 0.1 | 0.2 |
| b [Trunk_PAV Participant: P6] | 0.1 | 0.2 | -0.3 | -0.1 | 0.1 | 0.2 | 0.5 |
| b [Separation Participant: P6] | 0 | 0.1 | -0.2 | 0 | 0 | 0 | 0.1 |
| b [Trunk_PAV Participant: P7] | 0.2 | 0.2 | -0.1 | 0.1 | 0.2 | 0.4 | 0.6 |
| b [Separation Participant: P7] | 0 | 0.1 | -0.2 | -0.1 | 0 | 0 | 0.1 |
| b [Trunk_PAV Participant: P8] | 0 | 0.2 | -0.3 | -0.1 | 0 | 0.1 | 0.3 |

Continued on next page

Table 4.3: Parameter estimates for the final model TSWH. Predictors are trunk peak angular velocity (Trunk_PAV), separation time (Separation), pitcher's weight (Weight) and height (Height). The standard deviation of the errors is called `sigma` and the variance–covariance matrix of the pitcher–specific deviations from the common parameters is called `Sigma`. (Continued)

| | **mean** | **sd** | **2.5%** | **25%** | **50%** | **75%** | **97.5%** |
|---|---|---|---|---|---|---|---|
| b [Separation Participant: P8] | 0 | 0.1 | -0.1 | 0 | 0 | 0 | 0.2 |
| b [Trunk_PAV Participant: P9] | 0.1 | 0.1 | -0.2 | 0 | 0.1 | 0.2 | 0.4 |
| b [Separation Participant: P9] | 0 | 0 | -0.1 | 0 | 0 | 0 | 0.1 |
| sigma | 0.3 | 0 | 0.3 | 0.3 | 0.3 | 0.3 | 0.3 |
| Sigma[Participant: (Intercept),(Intercept)] | 0.4 | 0.2 | 0.1 | 0.3 | 0.3 | 0.5 | 0.9 |
| Sigma[Participant: Trunk_PAV,Trunk_PAV] | 0.1 | 0.1 | 0 | 0.1 | 0.1 | 0.2 | 0.3 |
| Sigma[Participant: Separation,Trunk_PAV] | 0 | 0 | -0.1 | 0 | 0 | 0 | 0 |

# REFERENCES


[1] L. Gomaz, B. van Trigt, F. van der Meulen, and D. Veeger. “Predicting elbow load based on individual pelvis and trunk (inter)segmental rotations in fastball pitching”. In: *Sports Biomechanics* (2024). doi: 10.1080/14763141.2024.2315230.

[2] G. S. Fleisig, J. R. Andrews, C. J. Dillman, and R. F. Escamilla. “Kinetics of baseball pitching with implications about injury mechanisms”. In: *The American journal of sports medicine* 23.2 (1995), pp. 233–239.

[3] S. T. Seroyer, S. J. Nho, B. R. Bach, C. A. Bush-Joseph, G. P. Nicholson, and A. A. Romeo. “The kinetic chain in overhand pitching: its potential role for performance enhancement and injury prevention”. In: *Sports health* 2.2 (2010), pp. 135–146.

[4] F. S. Chen, A. S. Rokito, and F. W. Jobe. “Medial elbow problems in the overhead-throwing athlete”. In: *JAAOS-Journal of the American Academy of Orthopaedic Surgeons* 9.2 (2001), pp. 99–113.

[5] B. J. Erickson, B. U. Nwachukwu, S. Rosas, W. W. Schairer, F. M. McCormick, B. R. Bach Jr, C. A. Bush-Joseph, and A. A. Romeo. “Trends in medial ulnar collateral ligament reconstruction in the United States: a retrospective review of a large private-payer database from 2007 to 2011”. In: *The American Journal of Sports Medicine* 43.7 (2015), pp. 1770–1774.

[6] G. S. Fleisig and J. R. Andrews. “Prevention of elbow injuries in youth baseball pitchers”. In: *Sports health* 4.5 (2012), pp. 419–424.

[7] A. L. Aguinaldo and H. Chambers. “Correlation of throwing mechanics with elbow valgus load in adult baseball pitchers”. In: *The American journal of sports medicine* 37.10 (2009), pp. 2043–2048.

[8] F. D. Wilson, J. R. Andrews, T. A. Blackburn, and G. Mccluskey. “Valgus extension overload in the pitching elbow”. In: *The American journal of sports medicine* 11.2 (1983), pp. 83–88.

[9] E. L. Cain Jr, J. R. Dugas, R. S. Wolf, and J. R. Andrews. “Elbow injuries in throwing athletes: a current concepts review”. In: *The American journal of sports medicine* 31.4 (2003), pp. 621–635.

[10] M. Safran, C. S. Ahmad, and N. S. Elattrache. “Ulnar collateral ligament of the elbow”. In: *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 21.11 (2005), pp. 1381–1395.

[11] A. W. Anz, B. D. Bushnell, L. P. Griffin, T. J. Noonan, M. R. Torry, and R. J. Hawkins. “Correlation of torque and elbow injury in professional baseball pitchers”. In: *The American journal of sports medicine* 38.7 (2010), pp. 1368–1374.

[12] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews. “Baseball Pitching Biomechanics in Relation to Injury Risk and Performance”. en. In: *Sports Health: A Multidisciplinary Approach* 1.4 (July 2009), pp. 314–320. issn: 1941-7381, 1941-0921. doi: 10.1177/1941738109338546. url: http://journals.sagepub.com/doi/10.1177/1941738109338546.

[13] J. Vanrenterghem, N. J. Nedergaard, M. A. Robinson, and B. Drust. “Training load monitoring in team sports: a novel framework separating physiological and biomechanical load-adaptation pathways”. In: *Sports medicine* 47 (2017), pp. 2135–2142.

[14] C. A. Putnam. “Sequential motions of body segments in striking and throwing skills: Descriptions and explanations”. en. In: *Journal of Biomechanics* 26 (Jan. 1993), pp. 125–135. issn: 00219290. doi: 10.1016/0021-9290(93)90084-R. url: https://linkinghub.elsevier.com/retrieve/pii/002192909390084R.

[15] A. Leenen, B. v. Trigt, M. Hoozemans, and H. Veeger. “Fastball pitching performance only slightly decreases after mobility impediment of the pelvis and trunk—Do (catch-up) compensation strategies come into play?” In: *Frontiers in Sports and Active Living* 4 (2022), p. 1044616.

[16] A. F. de Swart, B. van Trigt, K. Wasserberger, M. J. Hoozemans, D. H. Veeger, and G. D. Oliver. “Energy flow through the lower extremities in high school baseball pitching”. In: *Sports Biomechanics* (2022), pp. 1–15.

[17] A. Aguinaldo and K. Nicholson. “PELVIS ENERGY FLOW AND GROUND REACTION FORCE PREDICTORS OF ELBOW TORQUE IN HIGH SCHOOL AND COLLEGIATE BASEBALL PITCHERS”. In: *ISBS Proceedings Archive* 40.1 (2022), p. 13.

[18] E. Van Der Graaff, M. ( Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. ( Veeger. “Timing of peak pelvis and thorax rotation velocity in baseball pitching”. en. In: *The Journal of Physical Fitness and Sports Medicine* 7.5 (Sept. 2018), pp. 269–277. issn: 2186-8123, 2186-8131. doi: 10.7600/jpfsm.7.269. url: https://www.jstage.jst.go.jp/article/jpfsm/7/5/7_269/_article.

[19] M. Urbin, G. S. Fleisig, A. Abebe, and J. R. Andrews. “Associations between timing in the baseball pitch and shoulder kinetics, elbow kinetics, and ball speed”. In: *The American journal of sports medicine* 41.2 (2013), pp. 336–342.

[20] A. D. Cohen, E. J. Garibay, and M. J. Solomito. “The association among trunk rotation, ball velocity, and the elbow varus moment in collegiate-level baseball pitchers”. In: *The American journal of sports medicine* 47.12 (2019), pp. 2816–2820.

[21] G. S. Fleisig, S. W. Barrentine, N. Zheng, R. F. Escamilla, and J. R. Andrews. “Kinematic and kinetic comparison of baseball pitching among various levels of development”. en. In: *Journal of Biomechanics* 32.12 (Dec. 1999), pp. 1371–1375. issn: 00219290. doi: 10.1016/S0021-9290(99)00127-X. url: https://linkinghub.elsevier.com/retrieve/pii/S002192909900127X.

[22] B. Van Trigt, F. F. Bouman, A. J. Leenen, M. J. Hoozemans, F. C. Van der Helm, and D. Veeger. “Quantifying Within-Individual Elbow Load Variability in Youth Elite Baseball Pitchers and Its Role in Overuse Injuries”. In: *Applied Sciences* 12.13 (2022), p. 6549.

[23] L. Gomaz, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. “Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data”. en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: 10.3390/s21227442. url: https://www.mdpi.com/1424-8220/21/22/7442.

[24] A. E. Cornelius, B. W. Brewer, and J. L. Van Raalte. “Applications of multilevel modeling in sport injury rehabilitation research”. In: *International Journal of Sport and Exercise Psychology* 5.4 (2007), pp. 387–405.

[25] G. Wu, F. C. Van der Helm, H. D. Veeger, M. Makhsous, P. Van Roy, C. Anglin, J. Nagels, A. R. Karduna, K. McQuade, X. Wang, *et al.* “ISB recommendation on definitions of joint coordinate systems of various joints for the reporting of human joint motion—Part II: shoulder, elbow, wrist and hand”. In: *Journal of biomechanics* 38.5 (2005), pp. 981–992.

[26] V. M. Zatsiorsky. *Kinetics of human motion*. Human kinetics, 2002.

[27] P. De Leva. “Adjustments to Zatsiorsky-Seluyanov’s segment inertia parameters”. In: *Journal of biomechanics* 29.9 (1996), pp. 1223–1230.

[28] B. Goodrich, J. Gabry, I. Ali, and S. Brilleman. *rstanarm: Bayesian applied regression modeling via Stan.* R package version 2.21.1. 2020. url: https://mc-stan.org/rstanarm/.

[29] A. Vehtari, D. Simpson, A. Gelman, Y. Yao, and J. Gabry. “Pareto smoothed importance sampling”. In: *arXiv preprint arXiv:1507.02646* (2015).

[30] A. Vehtari, A. Gelman, and J. Gabry. “Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC”. In: *Statistics and computing* 27 (2017), pp. 1413–1432.

[31] J. Gabry. *Shinystan: Interactive visual and numerical diagnostics and posterior analysis for Bayesian models [Computer software manual]*. 2018.

[32] A. Gelman, B. Goodrich, J. Gabry, and A. Vehtari. “R-squared for Bayesian regression models”. In: *The American Statistician* (2019).

[33] J. Davis, O. Limpisvasti, D. Fluhme, K. J. Mohr, L. A. Yocum, N. S. ElAttrache, and F. W. Jobe. “The Effect of Pitching Biomechanics on the Upper Extremity in Youth and Adolescent Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 37.8 (Aug. 2009), pp. 1484–1491. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546509340226. url: http://journals.sagepub.com/doi/10.1177/0363546509340226.

[34] G. S. Fleisig, D. S. Kingsley, J. W. Loftice, K. P. Dinnen, R. Ranganathan, S. Dun, R. F. Escamilla, and J. R. Andrews. “Kinetic Comparison among the Fastball, Curveball, Change-up, and Slider in Collegiate Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 34.3 (Mar. 2006), pp. 423–430. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546505280431. url: http://journals.sagepub.com/doi/10.1177/0363546505280431.

[35] S. Oyama, B. Yu, J. T. Blackburn, D. A. Padua, L. Li, and J. B. Myers. “Improper trunk rotation sequence is associated with increased maximal shoulder external rotation angle and shoulder joint force in high school baseball pitchers”. In: *The American Journal of Sports Medicine* 42.9 (2014), pp. 2089–2094.

[36] A. Aguinaldo and R. Escamilla. “Segmental power analysis of sequential body motion and elbow valgus loading during baseball pitching: comparison between professional and high school baseball players”. In: *Orthopaedic journal of sports medicine* 7.2 (2019), p. 2325967119827924.

[37] K. Naito, H. Takagi, N. Yamada, S. Hashimoto, and T. Maruyama. “Intersegmental dynamics of 3D upper arm and forearm longitudinal axis rotations during baseball pitching”. In: *Human movement science* 38 (2014), pp. 116–132.

[38] E. Van der Graaff. “Perfecting your Pitch: In search of the perfect baseball pitch and its training”. In: (2019).

[39] B. J. Erickson, T. Sgori, P. N. Chalmers, P. Vignona, M. Lesniak, C. A. Bush-Joseph, N. N. Verma, and A. A. Romeo. “The impact of fatigue on baseball pitching mechanics in adolescent male pitchers”. In: *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 32.5 (2016), pp. 762–771.

[40] G. S. Fleisig, W. A. Laughlin, K. T. Aune, E. L. Cain, J. R. Dugas, and J. R. Andrews. “Differences among fastball, curveball, and change-up pitching biomechanics across various levels of baseball”. en. In: *Sports Biomechanics* 15.2 (Apr. 2016), pp. 128–138. issn: 1476-3141, 1752-6116. doi: 10.1080/14763141.2016.1159319. url: http://www.tandfonline.com/doi/full/10.1080/14763141.2016.1159319.

[41] M. Lapinski, C. Brum Medeiros, D. Moxley Scarborough, E. Berkson, T. J. Gill, T. Kepple, and J. A. Paradiso. “A wide-range, wireless wearable inertial motion sensing system for capturing fast athletic biomechanics in overhead pitching”. In: *Sensors* 19.17 (2019), p. 3637.

# B

# APPENDIX

In the Appendix we compare the outcome of the multilevel linear model fitted within frequentist framework with the outcome of the Bayesian hierarchical model presented in the paper.

We used a frequentist approach to fit the final model (TSWH) including trunk peak angular velocity, separation time, pitcher's weight, and height as predictors. The analysis was performed using the `lme4` `R` package (`version 1.1.26`). When fitting a multilevel model within the frequentist framework using the lme4 package, parameter estimation is done by performing restricted maximum likelihood (REML) estimation. The extended summary including corresponding p-values from the lmerTest R package is listed in Figure B.1.

```
Linear mixed model fit by REML. t-tests use Satterthwaite's method ['lmerModLmerTest']
Formula: Elbow_Adduction_MER ~ 0 + (1 | Participant) + (0 + Trunk_PAV +
    Separation_Pelvis_Trunk | Participant) + Weight + Height
   Data: dT

REML criterion at convergence: 148.5

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.5323 -0.5629  0.0464  0.5520  3.5031

Random effects:
 Groups        Name                    Variance Std.Dev. Corr
 Participant   (Intercept)             0.347006 0.58907
 Participant.1 Trunk_PAV               0.137205 0.37041
               Separation_Pelvis_Trunk 0.001088 0.03299  -1.00
 Residual                              0.081909 0.28620
Number of obs: 240, groups:  Participant, 11

Fixed effects:
       Estimate Std. Error      df t value Pr(>|t|)
Weight   0.8200     0.2114 10.1735   3.879  0.00297 **
Height   0.3935     0.2087  8.6233   1.886  0.09337 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Correlation of Fixed Effects:
       Weight
Height -0.408
```

Figure B.1.: Outcome of the linear mixed-effects TSWH model using `R` package `lme4`. TSWH model includes following set of predictors: trunk peak angular velocity, separation time between pelvis and trunk peak angular velocity, pitcher's weight and height.

The residual within-pitcher standard deviation is estimated as 0.28620. The estimated standard deviations of the pitcher intercepts are 0.58907. The estimated standard deviations of the pitcher slope for Trunk_PAV and

Separation_Pelvis_Trunk are 0.37041 and 0.03299 respectively. The fixed regression slopes for Weight and Height are significant, meaning taller and heavier pitchers have higher external valgus torque. The error term (Variance) for the slope of Trunk_PAV is 0.137205 and for the slope of Separation_Pelvis_Trunk is 0.001088.

Unlike the frequentist approach, the Bayesian approach accounts for all the uncertainty in the parameter estimates when predicting varying intercepts and slopes. We used rstanarm `R` package (`version 2.21.1`) to obtain simulations that summarize uncertainty about coefficients and predictions. Bayesian estimation is performed via Markov Chain Monte Carlo (MCMC) estimation approach whose each step involves random draws from the parameter space.

The summary of the final TSWH model whose parameter estimates are listed in Table 4.3 is shown in Figure B.2.

```
stan_lmer
 family:       gaussian [identity]
 formula:      Elbow_Adduction_MER ~ 0 + (1 | Participant) + (0 + Trunk_PAV +
           Separation_Pelvis_Trunk | Participant) + Weight + Height
 observations: 240
------
       Median MAD_SD
Weight 0.8    0.2
Height 0.4    0.2

Auxiliary parameter(s):
      Median MAD_SD
sigma 0.3    0.0

Error terms:
 Groups      Name                    Std.Dev. Corr
 Participant (Intercept)             0.626
 Participant Trunk_PAV               0.353
             Separation_Pelvis_Trunk 0.092    -0.20
 Residual                            0.290
Num. levels: Participant 11
```

Figure B.2.: Outcome of the linear mixed-effects TSWH model using `R` package `lme4`. TSWH model includes following set of predictors: trunk peak angular velocity, separation time between pelvis and trunk peak angular velocity, pitcher's weight and height.

The estimated standard deviations of the pitcher intercepts are 0.626, which is larger than the REML estimate (0.58907). The estimated standard deviations of the pitcher slopes for the parameters Trunk_PAV and Separation_Pelvis_Trunk are 0.353 and 0.092. Those estimates are both larger than the REML estimates (0.37041 and 0.03299 respectively).

The difference in estimates between frequentist and Bayesian approaches lies in the difference in estimation approaches. While `lme4` uses in this case restricted maximum likelihood (REML) estimation, `rstanarm` performs full Bayesian inference via MCMC. REML tends to underestimate uncertainties due to relying on point estimates of hyperparameters. On the other hand, the Bayesian approach propagates the uncertainty in the hyperparameters throughout all levels of the model.

The advantage of multilevel models fitted within the Bayesian framework is a specification of prior distributions over the regression coefficients and any unknown covariance matrices.

This can help in stabilizing computation as well as in incorporating important information into the analysis that is not included in the data. One of the limitations of multilevel models fitted with `rstanarm` compared to `lme4` is the computation speed. Fitting models with REML tends to be much faster than fitting a similar model using MCMC.

# 5

# FROM INJURY PREDICTION TO RISK ASSESSMENT: LATENT MARKOV MODELS FOR ANALYSING OSTRC QUESTIONNAIRES

**Larisa van der Graaff, Evert Verhagen, Frank van der Meulen**

Submitted in *BMJ Open Sport and Exercise Medicine* **2024** [1].

## 5.1. INTRODUCTION

Sport-related injuries impair individual and team participation and performance. The probability that an athlete will sustain an injury is determined by the interconnection between the internal and external risk factors and (a series of) events that may lead to an injury [2–4]. Constant interaction between the athlete and the environment and changes in training load affect the time needed for the athlete's recovery and training adaptation and change the athlete's susceptibility to injury over time [3, 5, 6].

Injuries are traditionally classified based on the mode of onset. Sudden onset injuries, also known as acute injuries, result from a specific identifiable event. On the other hand, injuries with a gradual onset, known as overuse injuries, are caused by repeated microtrauma [7]. Overuse injuries often go unnoticed in their early development phase due to the lack of associated symptoms (e.g. pain, functional limitation). Consequently, athletes are likely to continue to train and compete despite the increased risk of sustaining a more severe injury in the long run [8, 9]. While overuse injuries are generally considered to be preventable [10, 11], the absence of a specific inciting event and their gradual onset make them difficult to foresee.



In 2020, the International Olympic Committee published a consensus statement offering recommendations for collecting and reporting data on injury and illness in sports [7]. The severity of health problems in sports can be described based on several criteria, including time loss, athlete's self-reported health complaints, clinical extent of the illness or injury and societal cost. The Oslo Sports Trauma Research Center Questionnaire on Health Problems (OSTRC) [9, 12] is a widely adopted monitoring tool that complements time-loss severity measures. Additionally, it records symptoms and functional consequences of injuries. The OSTRC questionnaire is used to evaluate athletes' health status through the domain of reduced sports participation, training modification, performance reduction and symptoms associated with injuries. The responses to each of the four main questions are allocated a numerical value from 0 to 25 that sum up into the severity score ranging from 0 to 100 at specific time points. One of the limitations of the adopted OSTRC outcome is that the severity score has not yet been validated as a proxy for injury severity. Secondly, the severity score represents an ordinal-scale variable with 25 possible outcomes. Consequently, several researchers addressed the analytical benefits of representing various states of athletes' health on an ordinal scale [9, 13, 14].

Given the longitudinal nature of the data collected with the OSTRC questionnaire and the often mixed acute and repetitive mechanisms of injuries, we propose a multistate injury framework as a novel approach to the analysis of the OSTRC data. In this framework, we assume an athlete can be in a set of latent states, each state representing athlete's health status. Athletes can transition between the states many times during the follow-up based on their monthly, weekly, or even daily changes in training exposure. The statistical model we use for an athlete's health status is known as a latent Markov model. Latent Markov models are well established in the literature of longitudinal data analysis [15–17] when the interest lies in describing individual changes with respect to a latent status [18]. The latent status is represented by an unobserved process which is modelled by a first-order Markov chain. In the proposed framework, the latent status of interest is the athlete's injury state. The number of unobserved (latent) injury states is estimated from the response variables of the OSTRC questionnaire.

The application of latent Markov models is demonstrated in data collected by the repeated administration of the OSTRC questionnaire from the Dutch Olympic waterpolo monitoring program [19]. The application investigates how the athlete's injury state changes over time depending on time-varying training exposure. Furthermore, the study aims to demonstrate the application of the proposed model for predicting injury risk in the form of a probability of transition between the injury states under future training scenarios.

Our main contributions are: *(i)* using latent Markov models for analysing the OSTRC-questionnaire and forward prediction under future training schedules, this is done using the `LMest` package in the statistical language `R`; *(ii)* a novel Bayesian analysis for drawing from the posterior distribution using the probabilistic programming language `Turing` in the language `Julia`; *(iii)* application of these methods to data obtained from the Dutch Olympic waterpolo team.

This chapter is organised as follows. In Section 5.2, we describe the dataset from a longitudinal study performed on waterpolo players within the Dutch Olympic monitoring program. Section 5.3 describes the latent Markov model with time-varying covariates. Section 5.4 presents the results of applying the proposed approach using the `LMest` `R` package. In Section 5.5 we fit a submodel of the model fitted with `LMest` using the probabilistic programming language `Turing`. The chapter ends with a discussion, Section 5.6, followed by conclusion in Section 5.7. The Appendix C contains some technical details for the approach used in Section 5.5.

## 5.2. OSTRC OVERUSE INJURIES QUESTIONNAIRE

In order to illustrate the application of latent Markov models to assess changes in athlete's injury status, we consider OSTRC data derived from a longitudinal study on 24 waterpolo players within the Dutch Olympic monitoring program [19]. The study is carried out through the repeated administration of the OSTRC questionnaire that has been filled in by the athletes on a weekly basis over the 72–week–long follow–up. The data contain athletes' weekly exposure (for different types of training and competition [7]) as well as data on athlete's health complaints.

We focus on the dichotomous questionnaire outcomes, which are formulated in a way that responding 1 to any of the questions indicates the presence of a health complaint affecting the corresponding domain. We consider four items (i.e. 4 main questions) — listed in Table 5.1 — and three time–varying covariates:

- time spent on a sport-specific activity in the last 7 days (in hours),
- time spent on a strength training in the last 7 days (in hours),
- time spent on a competition in the last 7 days (in hours).

The interval between consecutive occasions at which the questionnaire was administered is one week, equal for all participants.



Table 5.1.: The questions selected for the injury risk assessment. The last column denotes the percentage of response 1 (which means the answer to the question was "yes") to each question during the follow-up.

| | **Question** | **%** |
|---|---|---|
| Participation: | Have you had any difficulties participating in training due to injury, illness, or other health problems during the past seven days? | 22.36 |
| Modification: | Did you have to modify your training due to injury, illness, or other health problems during the past seven days? | 12.29 |
| Performance: | Have your injury, illness, or other health problems affected your performance during the past seven days? | 14.72 |
| Symptoms: | Have you experienced symptoms/health complaints during the past seven days? | 22.89 |

## 5.3. STATISTICAL METHODS: LATENT MARKOV MODELS

Hidden Markov models, also known as state-space models, are well known statistical models for modelling serial dependence. Such models consist of two components: *(i)* a latent (unobserved) Markov chain which represent a "state" we are interested in; *(ii)* observations at each time that depend on the latent state at that time. We will use this framework as follows: for each subject we assume the hidden process takes values in $\{1,...,k\}$, which represents to what extent the subject is injury prone. Observations at each time are provided by answers to 4 binary questions extracted from the OSTRC questionnaire. Transitions of the latent process over time are possibly affected by characteristics of the subject and training load characteristics. Based on the questionnaire outcomes we then aim to find out how many latent states best fit the data, and how training loads affect state-transitions. Moreover, we aim to reconstruct latent states (this is known as “smoothing” in signal processing literature). A key advantage is that once the model has been fitted, we can simulate future scenarios where we assess the consequences of a particular training schedule to risk of injury. Introducing a model with temporal dependence is therefore of key importance.

In the following, as common in statistics, we denote random quantities by capital letters and their realisations (i.e. observations) by lower case letters. Let $y_{\mathrm{ijt}} \in \{0, 1\}$ denote the response variable to the $j$th question in the OSTRC questionnaire administered by the $i$th subject in week $t$, with $i = 1, \ldots, n$, $j = 1, \ldots, J$ and $t \in 1, \ldots, T_{\mathrm{i}}$. We use the convention that $y_{\mathrm{ijt}} = 1$ means that the question has been answered by “yes”. The response vector for subject $i$ at time $t$ is given by the vector of all answers $\mathbf{y}_{\mathrm{it}} = (y_{i1t}, \ldots, y_{\mathrm{iJt}}) \in \{0, 1\}^J$. For subject $i$, let $\mathbf{x}_{\mathrm{it}}$ be the vector of time-varying covariates at time $t$. In our application, these are the times spent in the last 7 days (in hours) on sport-specific activity, strength training and competition.

Following the latent Markov approach, for each subject $i$, we assume the existence of latent process $(U_{\mathrm{i1}}, \ldots, U_{\mathrm{iT}})$. The latent variable $U_{\mathrm{it}}$ represents the injury status of the $i$th subject in week $t$. The sequence of latent variables $U_{\mathrm{i1}}, \ldots, U_{\mathrm{iT}}$ is assumed to follow a (first–order) Markov chain with state space $\{1, \ldots, k\}$, where $k$ is the number of latent states. For mathematical convenience, we assume that the responses at time $t$ for subject $i$, $Y_{\mathrm{i1t}}, \ldots, Y_{\mathrm{iJt}}$, are independent conditional on the latent state $U_{\mathrm{it}}$.

### 5.3.1. Specification of latent process

The latent process is assumed to be a first order Markov chain. This means that its dynamics are fully governed by specifying the distribution of its initial state and the distribution of transitioning from time $t-1$ to time $t$.

#### Initial probabilities of the latent process

Let $\pi_{\mathrm{i}}$ be the initial probability vector that is the column vector with elements $\pi_i(u)$ for $u \in \{1, \ldots, k\}$.

$$\Pi_{\mathrm{i1}}(u) = P(U_{\mathrm{i1}} = u \mid \mathbf{x}_{\mathrm{i1}}). \tag{5.1}$$

We allow these probabilities to depend on the individual covariates as an adjustment for the differences between the individuals in terms of their injury status at the beginning of the follow-up. As we assume that the latent states are ordered, we require that the initial probabilities depend on the individual covariates $\mathbf{x}_{i1}$ by adopting a multinomial logit parametrization as follows:

$$\log \frac{\Pi_{i1}(d)}{\Pi_{i1}(1)} = \langle \tilde{\mathbf{x}}_{i1}, \boldsymbol{\beta}_u \rangle, \qquad d = 2, \ldots, k, \tag{5.2}$$

where $\tilde{\mathbf{x}}_{it} = (1, \mathbf{x}_{it})$ and $\langle \boldsymbol{a}, \boldsymbol{b} \rangle$ denotes the inner product of two vectors $\boldsymbol{a}$ and $\boldsymbol{b}$. An equivalent description to (5.2) is given by[1]

$$(\Pi_{i1}(1), \Pi_{i1}(2), \ldots, \Pi_{i1}(k)) = \mathrm{softmax}(0, \langle \tilde{\mathbf{x}}_{i1}, \boldsymbol{\beta}_2 \rangle, \ldots, \langle \tilde{\mathbf{x}}_{i1}, \boldsymbol{\beta}_k \rangle).$$

This helps in interpreting $\beta_u$: $\Pi_{i1}(u)$ is an increasing function of $\tilde{\mathbf{x}}_{i1}' \beta_u$, for $u \neq 1$.

5

#### TRANSITION PROBABILITIES OF THE LATENT PROCESS

For each subject $i$, we assume the latent Markov process $\{U_{it}\}_{t=1}^{T}$ is a first order Markov chain where the initial state satisfies (5.1). Let $\Pi_{it}$ be the transition probability matrix with elements $\Pi_{it}(d \mid u)$, $u, d = 1, \ldots, k$ that gives the probabilities of transitioning from state $u$ at time $t-1$ to state $d$ at time $t$. We assume transition probabilities of the latent Markov chain to depend on individual covariates

$$\Pi_{it}(d \mid u) = P(U_{it} = d \mid U_{i,t-1} = u, \mathbf{x}_{it}), \qquad t = 2, \ldots, T_i.$$

Following [20], we adopt the following multinomial logit parametrization

$$\log \frac{\Pi_{it}(d \mid u)}{\Pi_{it}(u \mid u)} = \tilde{\mathbf{x}}_{it}' \boldsymbol{\gamma}_{ud}, \qquad d \in \{1, \ldots, k\} \setminus d. \tag{5.3}$$

### 5.3.2. SPECIFICATION OF CONDITIONAL RESPONSE PROBABILITIES

In case subject $i$ at time $t$ answers “yes” to question $j$, then $y_{ijt} = 1$, else $y_{ijt} = 0$. We assume that the distribution of the response variables depends only on the latent status by imposing

$$p(Y_{ijt} = 1 \mid U_{it} = u, \mathbf{x}_{it}) = \lambda_j(u) \tag{5.4}$$

for each $i$, $j$, $t$ and $u \in \{1, \ldots, k\}$. Moreover, we require these conditional probabilities to satisfy the constraint

$$0 < \lambda_j(1) \leq \lambda_j(2) \leq \ldots \leq \lambda_j(k) < 1 \tag{5.5}$$

for $j = 1, \ldots, J$. This assumption has been used before in [20] and ensures identifiability. The constraint (5.5) implies that the latent states are

[1] The softmax function is given by $\mathrm{softmax}(x_1, \ldots, x_k) = (e^{x_1}, \ldots, e^{x_k}) / \sum_{i=1}^{k} e^{x_i}$.

ordered such that the individuals in the first state are those with the best status (least injury prone) and individuals in the last state are those with the worst injury status. Note that while (5.4) is the same for all subjects, each subject's injury status is modelled by a separate latent process.

### 5.3.3. PARAMETER ESTIMATION AND MODEL SELECTION

Parameter estimation of $\beta_u$ and $\gamma_{ud}$ is done by maximum likelihood in `LMest`. Up to this point, we have considered the number of latent states to be fixed. For choosing it in a data adaptive way, we rely on the Bayesian Information Criterion (BIC) given by

$$BIC = -2\hat{\ell} + p\log(n) \tag{5.6}$$

where $\hat{\ell}$ denotes the maximum log-likelihood of the model, $n$ denotes the number of subjects and $p$ denotes the number of free parameters. The optimal number of latent states is the one corresponding to the minimal value of BIC [20, 21].

5

### 5.3.4. PATH PREDICTION

Given the estimated model, interest lies in predicting the most likely injury state based on the latest state occupied and the scheduled weekly training exposure. The predicted states over time follow a path that illustrates the progression of injury status over time based on the time-varying training exposure. This indicates the predicted injury risk for an individual athlete at the next time point and serves as a support for sports practitioners and coaches in decision-making during the training process.

Assume that at time $T$ parameters have been estimated based on all subjects in the study. Additionally assume for subject $i$ a training schedule has been determined for the upcoming $S$ weeks, i.e. weeks $T+1,...,T+S$. This implies $\mathbf{x}_{\mathrm{i,T+1}},...,\mathbf{x}_{\mathrm{i,T+S}}$ have been specified. Then we can forward simulate scenarios from the latent process $\{U_{\mathrm{it}}\}_{t=T}^{T+S}$ to assess injury risk for subject $i$ under the proposed training schedule. This is simply an instance of Monte Carlo simulation where we use a large number of forward simulations which are initialised according to inferred probabilities for $U_{\mathrm{iT}}$.

## 5.4. APPLICATION USING LMEST PACKAGE

In this section, we show the results obtained from the application of latent Markov models on the OSTRC longitudinal data from Dutch Olympic waterpolo players described in Section 2. The original questionnaire outcome was dichotomized, where 1 indicates the presence of a health complaint. We used the following covariates: the sport-specific training exposure (in hours), the strength training exposure (in hours), and the competition exposure (in hours) in the last 7 days. We consider the latent Markov model, where these time-varying covariates affect both initial and transition probabilities.

We illustrate the application of latent Markov models fitted with `R` package `LMest` [22]. Whereas the `lmest` function within the `LMest` package knows how to deal with missing questionnaire responses, it does not allow missing data in covariates. As the data set we study is very unbalanced, with a different follow-up duration for all participants, it was preprocessed to make it balanced by adding missing values in the response and setting missing covariates equal to zero. This last choice implies that when the covariates are missing we assume transitions of the latent process to all states are equally likely, something which is questionable and can in fact be avoided (see Section 5.5.1).

The first step of the analysis aimed to choose the number of latent states $k$. We fitted the model for increasing values of $k$ until the BIC index decreased with respect to the previous value.

| $k$ | $\hat{\ell}$ | $np$ | **BIC** |
|---|---|---|---|
| 1 | -3526.16 | 4 | 7065.04 |
| 2 | -1707.37 | 20 | 3478.30 |
| 3 | -1625.84 | 44 | 3391.52 |
| 4 | -1596.51 | 76 | 3434.55 |

Table 5.2.: Latent Markov models with covariates: Maximal log-likelihood ($\hat{\ell}$), number of parameters ($np$) and BIC index for a number of latent states $k$ between 1 and 4. The lowest value of the BIC index indicates that the most suitable number of latent states is $k$=3.

For a number of latent states k ranging from 1 to 4, Table 5.2 displays the maximum value of the likelihood ($\hat{\ell}$), the number of parameters ($np$) and the value of the BIC index. We conclude that the lowest value of the BIC index is obtained for the model with $k = 3$ latent states. Therefore, the resulting multivariate latent Markov model with covariates will include three latent states.

The estimates of the conditional response probabilities $\lambda_j(u)$ are reported in Table 5.3 for each latent state $u$ and OSTRC question $j$. The ordered latent states represent three different levels of the injury status for an individual athlete.

Based on the results the first latent state corresponds to athletes with a very low tendency to experience any physical complaints. The

| $j$ | $u = 1$ | $u = 2$ | $u = 3$ |
|---|---|---|---|
| 1 | 0.00 | 0.95 | 1.00 |
| 2 | 0.00 | 0.19 | 0.93 |
| 3 | 0.00 | 0.33 | 0.99 |
| 4 | 0.01 | 0.96 | 1.00 |

Table 5.3.: Estimates of the conditional response probabilities $\lambda_j(u)$ as given in (5.4), where $j$ represents the item (i.e. question from the OSTRC questionnaire) and $u$ represents the latent state (i.e. injury status). As an example $\lambda_2(3) = 0.93$ is to be interpreted as the probability that a subject in state 3 answer "yes" to the second question. It is clear that subjects in state 3 answer "yes" to all of the questions with high probability. In addition, questions 1 and 4 mostly distinguish between either state 1 or states 2 and 3. Questions 2 and 3 better distinguish between states 2 and 3.



second latent state corresponds to athletes with a high tendency to experience difficulties participating in training due to health complaints and related symptoms but a lower tendency to modify their training due to health complaints or have their performance affected by it. The third latent state corresponds to athletes with a high tendency to experience physical complaints.

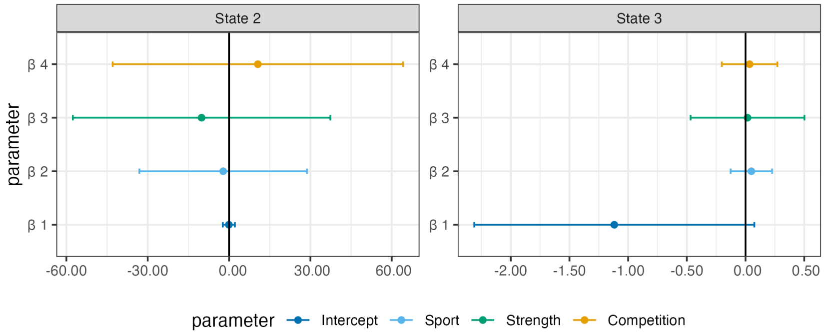


Figure 5.1.: Estimates for $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$, affecting the distribution of the initial probabilities of the latent Markov process (Cf. (5.2)), obtained by `LMest`. The dots represent the estimates; the horizontal line segment corresponding to a dot shows the estimate ± its standard error. Note the different scaling of the horizontal axes in both panels.

Estimates of $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$ that affect the initial probabilities of the latent Markov process are visualised in Figure 5.1. The estimated negative intercepts indicate that, in general, athletes report full participation without health complaints at the beginning of the follow-up. The positive *Competition* estimates indicate that more hours spent in competition leads to a higher probability of a worse initial injury status.

5

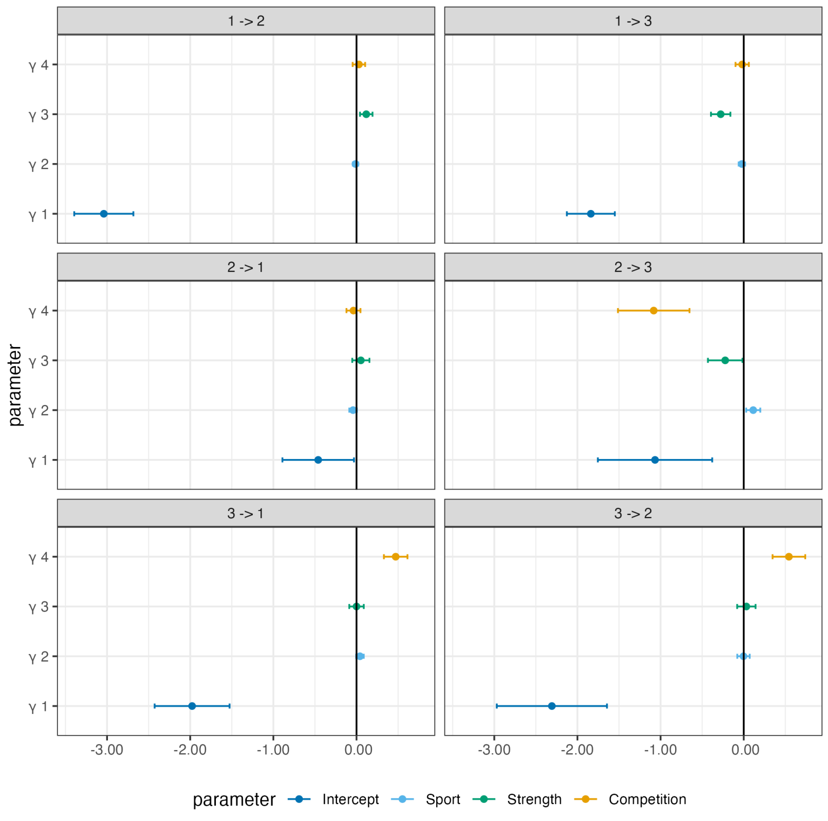


Figure 5.2.: Estimates for $\boldsymbol{\gamma}_{ud}$ ($u \neq d$), the transition probabilities of the latent Markov process (Cf. (5.3)) obtained by `LMest`. The dots represent the estimates; the horizontal line segment corresponding to a dot shows the estimate ± its standard error. As an example, the topright panel shows the estimates for $\boldsymbol{\gamma}_{13}$. It can be seen that the coefficient for strength is negative. This means that an increase of strength will lower the probability of transitioning from state 1 to 3.

Estimates of the regression coefficients for $\boldsymbol{\gamma}_{ud}$ ($u \neq d$), the transition probabilities of the latent Markov process are visualised in Figure 5.2.

In this figure, the scale of each subpanel is fixed. To better read off the coefficients for *Sport*, *Strength* and *Competition*, this figure is also shown in a different way in Figure 5.3. Here, the estimate for intercept is not shown and the scaling of each subpanel varies freely. The given estimates measure the influence of sport-specific training exposure, strength training exposure and competition exposure on the transition between different injury states.

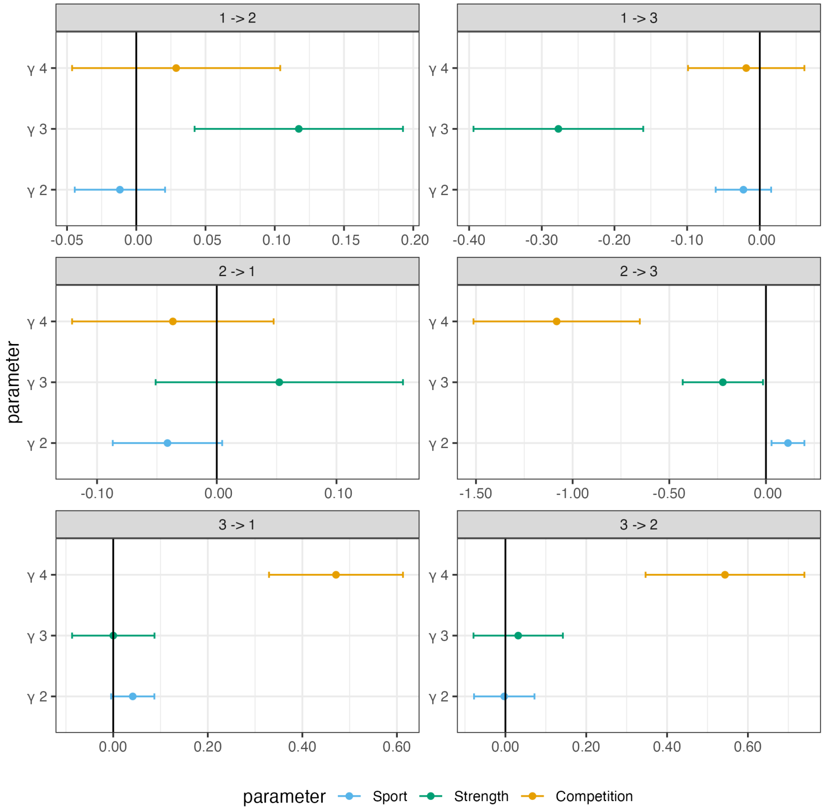


Figure 5.3.: Adjustment of Figure 5.2, where the estimates for the intercepts have been removed and the scales for each subpanel are not fixed.

The influence of the covariate *Sport* is negative. This means that the probability that an athlete moves from the first to the third injury state is lower for athletes with more sport-specific training exposure than for the ones with less (first row in Figure 5.2). On the other hand, *Strength* has a positive effect on the transition from the first to the second state. In other words, more strength training exposure increases the probability that an athlete experiences difficulties in training participation and health complaint - related symptoms. The positive effect of the covariate *Competition* indicates the increased probability of health complaints affecting training participation and causing related symptoms when the competition exposure is higher.

### 5.4.1. PREDICTION OF INJURY RISK UNDER DIFFERENT TRAINING SCENARIOS

By plugging in the parameter estimates, the latent process can be forward simulated under different training schedules. Here, we consider the training schedule shown in Figure 5.4.

5

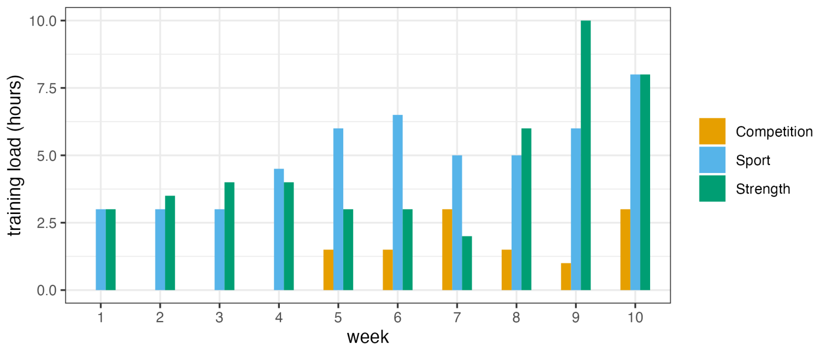


Figure 5.4.: Training schedule for 10 weeks.

In Figure 5.5 we show the marginal distribution over the 3 states obtained from the method outlined in Section 5.3.4. Here, week "0" represents the actual state of the athlete. The top panel shows changes in injury risk for an athlete who was for sure in state 1 at the end of the follow-up. The predicted path indicates that given the weekly training exposure, the athlete's probabilities are changing every week, with the highest probability of staying in state 1 for all the upcoming ten weeks. The middle panel shows how injury risk changes for the athlete who was in state 2 at the end of the follow-up. With the changes in weekly training exposure, the athlete's probabilities of occupying each state also change over time. Modifications in the training exposure result in a reduced probability of staying in the second state and an increased probability of transitioning to the first state. Finally, the bottom panel shows the changes in predicted probabilities in injury risk for an athlete that was in the third state. With the modifications in the weekly training exposure the probability of staying in the same state reduces, with probabilities of occupying the second and the first state increasing accordingly.

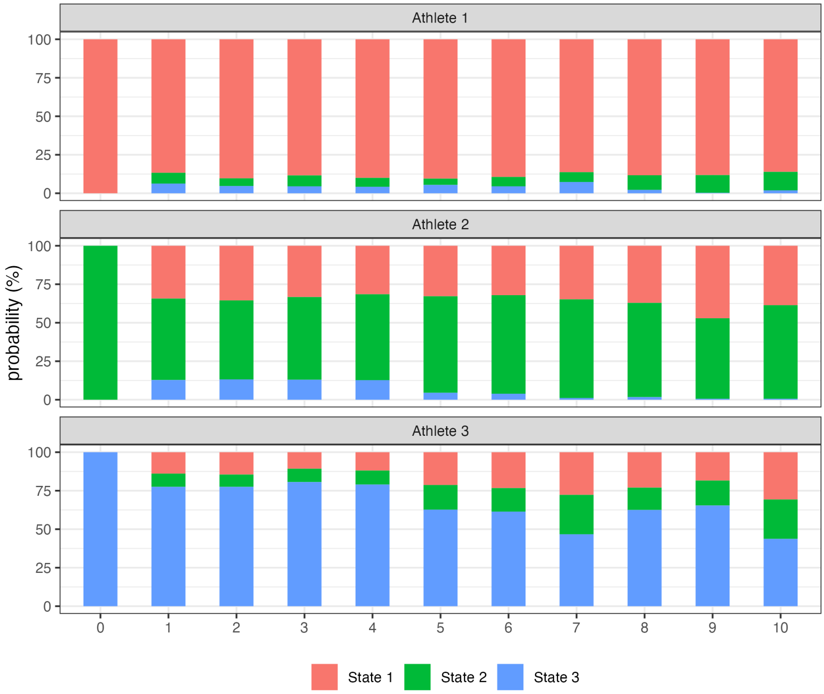


Figure 5.5.: Marginal distribution over the 3 states obtained from the method outlined in Section 5.3.4. Here, week "0" represents the actual state of the athlete.



## 5.5. BAYESIAN APPROACH

In this section we adopt a fully Bayesian approach. Contrary to our earlier approach this incorporates uncertainty in the parameter estimates in path prediction. Moreover, we show that our implementation is more friendly to missing data. In particular, observing different athletes over different time spans poses no restriction and therefore there is no need to artificially add missing data, as was needed to fit the models with `LMest`.

We assume a prior specification on the initial latent state of each subject. In addition we will detail prior distributions on $\boldsymbol{\gamma}_{ud}$ $(u \neq d)$ and the conditional response probabilities in (5.4). Once specified, the Bayesian paradigm postulates that all inferential conclusions are based on the posterior distribution. As this distribution is not available in closed form, we recursively compute the likelihood and use the probabilistic programming language `Turing` [23] within the `Julia`-language ([24]) to sample from the posterior. Details can be found in Section C.

### 5.5.1. TRANSITION PROBABILITIES OF THE LATENT PROCESS

To reduce notation, we assume below that the number of latent states is 3. As before, the number of questions in the questionnaire is denoted by $J$ and hence $J = 4$ in our data example.

The numerical results presented in this section assume a submodel of the model specified in Section 5.3.1, where only transitions to adjacent states are possible. Hence, we parametrise the transition matrix by vectors $\boldsymbol{\gamma}_{12}$, $\boldsymbol{\gamma}_{21}$, $\boldsymbol{\gamma}_{23}$ and $\boldsymbol{\gamma}_{32}$ such that

$$\Pi_{it} = \mathrm{softmax}_{\cdot} \begin{bmatrix} 0 & \langle \tilde{\mathbf{x}}_{it}, \boldsymbol{\gamma}_{12} \rangle & -\infty \\ \langle \tilde{\mathbf{x}}_{it}, \boldsymbol{\gamma}_{21} \rangle & 0 & \langle \tilde{\mathbf{x}}_{it}, \boldsymbol{\gamma}_{23} \rangle \\ -\infty & \langle \tilde{\mathbf{x}}_{it}, \boldsymbol{\gamma}_{32} \rangle & 0 \end{bmatrix}, \tag{5.7}$$

where $\mathrm{softmax}_{\cdot}$ denotes that the softmax function is to be applied to each row of the matrix. If any component of $\boldsymbol{x}_{it}$ is missing, we set $\Pi_{it}$ equal to the identity matrix. This corresponds to assuming the latent state does not change between times $t-1$ and $t$.

5

### 5.5.2. PRIOR SPECIFICATION

For $\boldsymbol{\gamma}_{12}$, $\boldsymbol{\gamma}_{21}$, $\boldsymbol{\gamma}_{23}$ and $\boldsymbol{\gamma}_{32}$ we impose conditionally independent standard multivariate normal priors with covariance matrix $\sigma$ times the identity matrix. We assign $\sigma$ the Exponential distribution with mean 3. The underlying ideas is to provide tractable "uninformative" priors.

For each question $j \in \{1, \ldots, J\}$ we need to specify a prior on $\boldsymbol{\lambda}_j := (\lambda_j(1), \ldots, \lambda_j(3))$ satisfying the ordering constraint in (5.5). We give a construction for that. Let $Z_j(1), Z_j(2), Z_j(3)$ be independent random variables with the standard Exponential distribution. Set $\psi(x) = 2\,\mathrm{logistic}(3x/4) - 1$, where $\mathrm{logistic} = 1/(1 + e^{-x})$ and note that $\psi$ maps $[0, \infty)$ to $[0, 1)$. Then set

$$\lambda_j(\ell) = \psi\left(\sum_{i=1}^{\ell} Z_j(i)\right), \quad j = 1, \ldots, J.$$

In Figure 5.6 we show histograms based on 10_000 samples from the prior. As all $Z_j(i)$ are supported on the positive halfline and $\psi$ is increasing, (5.5) is satisfied.

We assume a uniform prior on the initial state of the latent process, thereby avoiding estimation of $\boldsymbol{\beta}_2$ and $\boldsymbol{\beta}_3$.

### 5.5.3. APPLICATION

The recursive computation of the loglikelihood was implemented in the `Julia`-language in the package `LatentMarkovQuest` (`https://github.com/fmeulen/LatentMarkovQuest`). Subsequently the package `Turing`

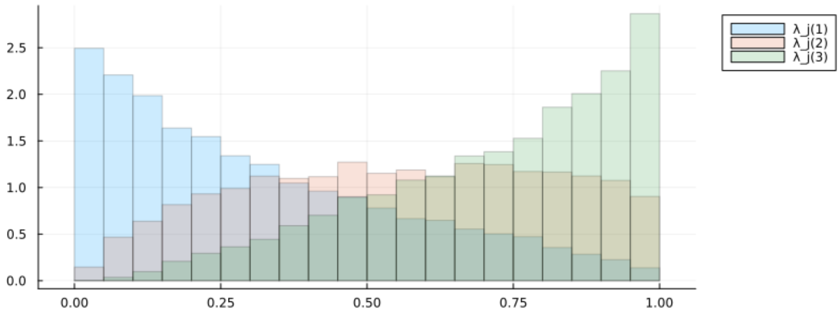


Figure 5.6.: Visualisation of distribution of $\boldsymbol{\lambda}_j$ by Monte-Carlo simulation.

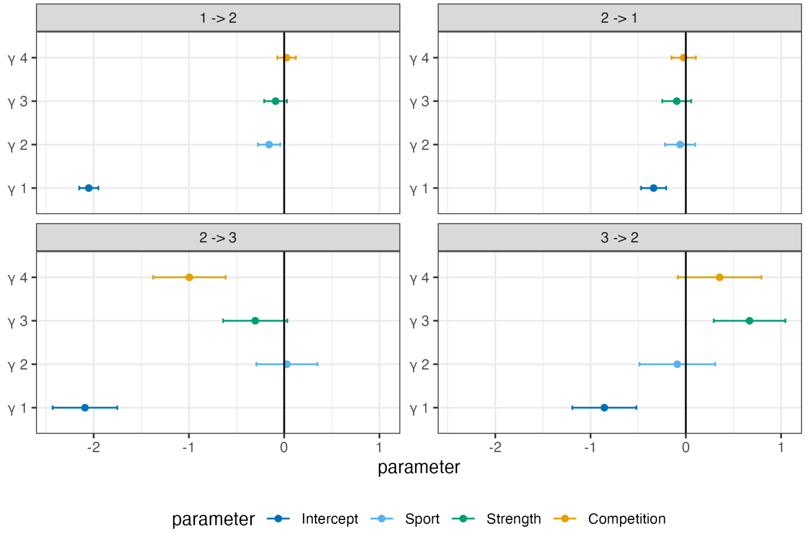


Figure 5.7.: Estimates for $\boldsymbol{\gamma}_{ud}$ $(u \neq d)$ obtained by the Bayesian approach, the transition probabilities of the latent Markov process (Cf. (5.3)). The dots represent the posterior means; the horizontal line segment corresponding to a dot shows the estimate ± its posterior standard deviation.

was used to draw from the posterior using the No-U-Turn-Sampler (NUTS). Cf. [25].

We ran 4 chains for 2000 iterations of which the first half part is considered burning. All Rhat values (Gelman-Rubin diagnostics, see e.g. Chapter 13 in [26]) were very close to one, indicating convergence of the MCMC chain. All density plots of parameter estimates (not included) look bell-shaped. Posterior summary plots are given in Figures 5.7 and 5.8.

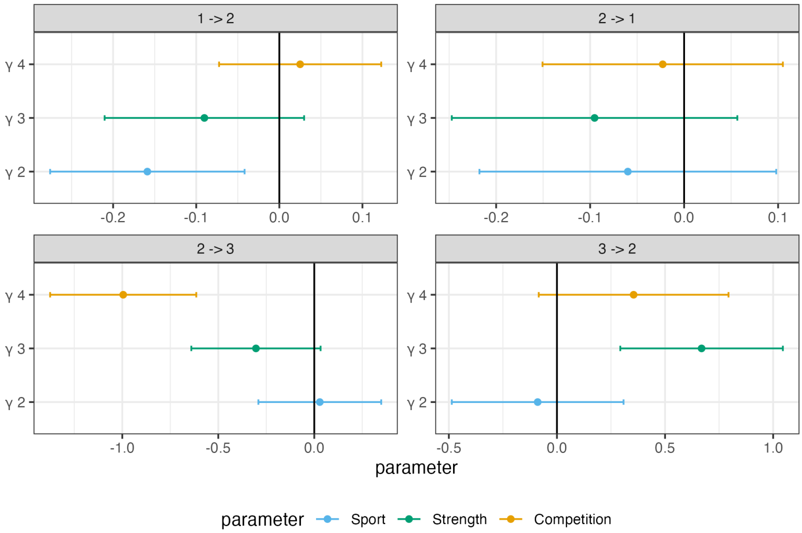


Figure 5.8.: Adjustment of Figure 5.7, where the estimates for the intercepts have been removed and the scales for each subpanel are not fixed.

In Table 5.4 we have converted posterior means estimates for $Z_j(1)$, $Z_j(2)$ and $Z_j(3)$ to conditional response probabilities. These estimates are close to that found using the `LMest` package (see Table 5.3).

| ***j*** | ***u* = 1** | ***u* = 2** | ***u* = 3** |
|---|---|---|---|
| 1 | 0.00 | 0.92 | 0.98 |
| 2 | 0.00 | 0.30 | 0.94 |
| 3 | 0.00 | 0.42 | 0.94 |
| 4 | 0.01 | 0.94 | 0.98 |

Table 5.4.: Estimates of the conditional response probabilities $\lambda_{\mathrm{j}}(u)$ as given in (5.4), where $j$ represents the item (i.e. question from the OSTRC questionnaire) and $u$ represents the latent state (i.e. injury status). Results obtained by the Bayesian method where posterior means for $Z$s were converted to $\lambda$s.

In Figure 5.9 we added predictions for future states, depending on whether the initial state was 1, 2 or 3. This figure is compared to be with the outcomes obtained by `LMest` shown in Figure 5.5 in Section 5.4.

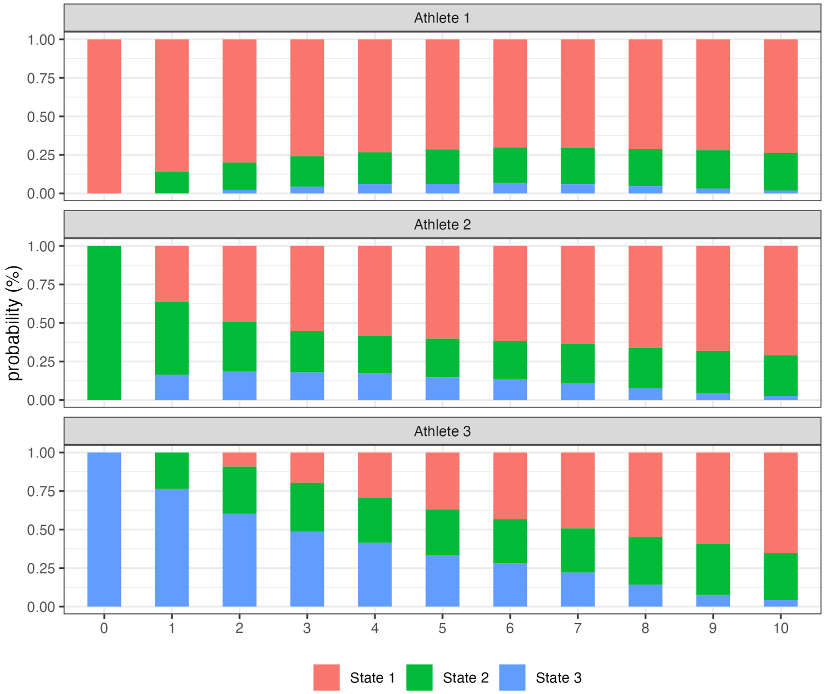


Figure 5.9.: Forward simulated paths using Bayesian model for 3 atheletes.

## 5.6. DISCUSSION

In this paper we propose a multistate injury framework where the latent Markov model is used to predict the risk of injury in the form of a probability. This approach comes with several limitations. In this application, we focused on the binary response variable where reporting 1 denotes the presence of a health complaint. The limitation of such an approach is that it does not capture the full scale on which the health complaint affects the four main domains. The next step would be to test the proposed model on an original OSTRC outcome. Secondly, if fitted by `LMest`, the adopted model does not allow missing data in the covariates. This may affect the parameter estimates describing the effect of each covariate on the probabilities of each injury state. Thirdly, the percentage of reported complaints is low.

The approach proposed in this paper can be adopted in various sports disciplines for any type of panel data as well as time-constant and/or time-varying positional, biometric and biomechanical data coming from various sources (e.g. wearables). Injury risk prediction based on the latent Markov model has the advantage of considering the dynamic nature of the injury state. The outcome of the model can be used to define an injury risk profile for an individual athlete. Such a framework can support clinicians in making real-time risk management decisions in training and (clinical) practice.

The predictions provided by `LMest` and the Bayesian method qualitatively agree, but there are clear differences. These can probably be attributed to either of the following:

- `LMest` does not allow for restricting the model to only transition to neighbouring states, whereas we believe for this application this is a reasonable assumption. The Bayesian approach uses this restriction and therefore assumes a different model.
- Both methods use different handling of missing data. `LMest` requires adding missing data to get the data in rectangular format, i.e. have data for all participants over the same period. In case covariates in a particular week are missing, it is assumed that transitions to all states are equally probable. The Bayesian approach assumes no state transition in this case.
- Estimates on `LMest` are based on maximum likelihood, whereas the Bayesian reported estimates are posterior means. As in any Bayesian analysis with small sample size, the prior may have some influence on the estimates obtained.

There is clearly room for follow-up work. Most notably, we believe model checking should be done. Due to the binary nature of the response this is however much harder than for example in classical regression analysis. As the primary aim of this paper is to show the potential of latent Markov models, we leave this to future research.

## 5.7. CONCLUSION

This paper demonstrates the application of a latent Markov model for injury risk prediction in a multistate injury framework. The application of latent Markov models allows us to estimate the optimal number of injury states and the influence of included personal characteristics and performance measures on the transition between those states over time. Furthermore, we show that it is possible to predict the injury risk in the form of a probability for occupying each of the injury states based on provided individual covariates.

# REFERENCES

[1] L. van der Graaff, E. Verhagen, and F. van der Meulen. “From injury prediction to risk assessment: latent Markov models for analysing OSTRC questionnaires”. In: *BMJ Open Sport and Exercise Medicine* (Submitted).

[2] N. F. Bittencourt, W. Meeuwisse, L. Mendonça, A. Nettel-Aguirre, J. Ocarino, and S. Fonseca. “Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept”. In: *British journal of sports medicine* (2016).

[3] W. H. Meeuwisse, H. Tyreman, B. Hagel, and C. Emery. “A dynamic model of etiology in sport injury: the recursive nature of risk and causation”. In: *Clinical journal of sport medicine* 17.3 (2007), pp. 215–219.

[4] M. Roe, S. Malone, C. Blake, K. Collins, C. Gissane, F. Büttner, J. C. Murphy, and E. Delahunt. “A six stage operational framework for individualising injury risk management in sport”. In: *Injury epidemiology* 4 (2017), pp. 1–6.

[5] W. H. Meeuwisse. *Assessing causation in sport injury: a multifactorial model*. 1994.

[6] E. Verhagen and T. Gabbett. *Load, capacity and health: critical pieces of the holistic performance puzzle*. 2019.

[7] I. O. C. Injury, I. E. C. Group, R. Bahr, B. Clarsen, W. Derman, J. Dvorak, C. A. Emery, C. F. Finch, M. Hägglund, A. Junge, S. Kemp, *et al.* “International Olympic Committee consensus statement: methods for recording and reporting of epidemiological data on injury and illness in sports 2020 (including the STROBE extension for sports injury and illness surveillance (STROBE-SIIS))”. In: *Orthopaedic journal of sports medicine* 8.2 (2020), p. 2325967120902908.

[8] R. Bahr. “No injuries, but plenty of pain? On the methodology for recording overuse symptoms in sports”. In: *British journal of sports medicine* 43.13 (2009), pp. 966–972.

[9] B. Clarsen, R. Bahr, G. Myklebust, S. H. Andersson, S. I. Docking, M. Drew, C. F. Finch, L. V. Fortington, J. Harøy, K. M. Khan, *et al.* "Improved reporting of overuse injuries and health problems in sport: an update of the Oslo sport trauma research center questionnaires". In: *British journal of sports medicine* 54.7 (2020), pp. 390–396.

[10] T. J. Gabbett. "The training—injury prevention paradox: should athletes be training smarter and harder?" In: *British journal of sports medicine* (2016).

[11] C. A. Emery and K. Pasanen. "Current trends in sport injury prevention". In: *Best Practice & Research Clinical Rheumatology* 33.1 (2019), pp. 3–15.

[12] B. Clarsen, O. Rønsen, G. Myklebust, T. W. Flørenes, and R. Bahr. "The Oslo Sports Trauma Research Center questionnaire on health problems: a new approach to prospective monitoring of illness and injury in elite athletes". In: *British journal of sports medicine* (2013).

[13] R. O. Nielsen, M. L. Bertelsen, D. Ramskov, M. Møller, A. Hulme, D. Theisen, C. F. Finch, L. V. Fortington, M. A. Mansournia, and E. T. Parner. "Time-to-event analysis for sports injury research part 1: time-varying exposures". In: *British journal of sports medicine* 53.1 (2019), pp. 61–68.

[14] R. O. Nielsen, M. L. Bertelsen, D. Ramskov, M. Møller, A. Hulme, D. Theisen, C. F. Finch, L. V. Fortington, M. A. Mansournia, and E. T. Parner. "Time-to-event analysis for sports injury research part 2: time-varying outcomes". In: *British journal of sports medicine* 53.1 (2019), pp. 70–78.

[15] F. Rijmen, K. Vansteelandt, and P. De Boeck. "Latent class models for diary method data: Parameter estimation by local computations". In: *Psychometrika* 73.2 (2008), pp. 167–182.

[16] F. Bartolucci, F. Pennoni, and B. Francis. "A latent Markov model for detecting patterns of criminal activity". In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 170.1 (2007), pp. 115–132.

[17] L. J. Paas, J. K. Vermunt, and T. H. Bijmolt. "Discrete time, discrete state latent Markov modelling for assessing and predicting household acquisitions of financial products". In: *Journal of the Royal Statistical Society Series A: Statistics in Society* 170.4 (2007), pp. 955–974.

[18] R. Langeheine and F. Van de Pol. "Latent markov chains". In: *Applied latent class analysis* (2002), pp. 304–341.

[19] E. Verhagen, M. Lang, R. Watson, and M. Moen. “Injuries and Illness in Olympic Level Water Polo Athletes–a Three-Season Prospective Study”. In: *Dtsch Z Sportmed* 72 (2021), pp. 195–202.

[20] F. Bartolucci, M. Lupparelli, and G. E. Montanari. “Latent Markov model for longitudinal binary data: An application to the performance evaluation of nursing homes”. In: *The Annals of Applied Statistics* (2009), pp. 611–636.

[21] F. Bartolucci, A. Farcomeni, and F. Pennoni. “Latent Markov models: a review of a general framework for the analysis of longitudinal data with covariates”. In: *Test* 23 (2014), pp. 433–465.

[22] F. Bartolucci, S. Pandolfi, and F. Pennoni. “LMest: An R package for latent Markov models for longitudinal categorical data”. In: *Journal of Statistical Software* 81 (2017), pp. 1–38.

[23] H. Ge, K. Xu, and Z. Ghahramani. “Turing: a language for flexible probabilistic inference”. In: *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9-11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*. 2018, pp. 1682–1690. url: http://proceedings.mlr.press/v84/ge18b.html.

[24] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah. “Julia: A fresh approach to numerical computing”. In: *SIAM review* 59.1 (2017). Publisher: SIAM, pp. 65–98. url: https://doi.org/10.1137/141000671.

[25] M. D. Hoffman, A. Gelman, *et al.* “The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo.” In: *J. Mach. Learn. Res.* 15.1 (2014), pp. 1593–1623.

[26] B. Lambert. *A student’s guide to Bayesian statistics*. SAGE Publications Ltd, 2018.


5

# C

# APPENDIX

## C.1. DRAWING FROM THE POSTERIOR

In this section we detail how we obtain samples from the posterior distribution. As can be seen, it consists of a simple recursive procedure for computing the loglikelihood. This is subsequently passed on to the probabilistic programming language.

### C.1.1. RECURSIVE LIKELIHOOD COMPUTATION

It is well known that the likelihood can be computed efficiently in a recursive way. Here we propose to use the backward information filter, which can be viewed as a message passing algorithm. This is well known in the literature, see e.g. [1] and [2]. Below, we denote the entrywise product of two vectors by $\odot$: for $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^k$, $\boldsymbol{a} \odot \boldsymbol{b} = (a_1 b_1, \dots, a_k b_k)$. Let $\mathbf{1} \in \mathbb{R}^3$ denote the vector with all elements equal to 1.

To reduce notational overhead, first assume just one subject, with responses $y_1, \dots, y_T$, where $y_t = (y_{t1}, \dots, y_{t4})$, and latent process $u_1, \dots, u_T$.

Define $u \mapsto h_t(u) = P(\boldsymbol{Y}_t = \boldsymbol{y}_t, \dots, \boldsymbol{Y}_T = \boldsymbol{y}_T \mid U_t = u)$. As $u \in \{1, \dots, k\}$ this map can be identified with the vector $\boldsymbol{h}_t = (h_t(1), \dots, h_t(k))$. The *backward information filter* consists of the following steps:

- for $t = 1, \dots, T$, let

$$\boldsymbol{g}_{tj} = \begin{cases} \boldsymbol{\lambda}_j & \text{if} \quad y_{tj} = 1 \\ \mathbf{1} - \boldsymbol{\lambda}_j & \text{if} \quad y_{tj} = 0 \end{cases}$$

  and set $\boldsymbol{g}_t = \odot_{j=1}^{J} \boldsymbol{g}_{tj}$;

- set $\boldsymbol{h}_T = \boldsymbol{g}_T$ and

$$\boldsymbol{h}_t = \boldsymbol{g}_t \odot (\Pi_{i,t+1} \boldsymbol{h}_{t+1}), \qquad t = T-1, \dots, 1; \tag{C.1}$$

- set $h_0 = \Pi_1 \boldsymbol{h}_1$, where $\Pi_1$ is the prior on the initial latent state.

The output of this scheme, $h_0$ is the likelihood. Notationally, we have suppressed any dependence on the parameter vector $\boldsymbol{\theta}$ which under model specification (5.7) is given by the $\boldsymbol{\theta}$ obtained by concatenating $\boldsymbol{\gamma}_{12}$, $\boldsymbol{\gamma}_{21}$, $\boldsymbol{\gamma}_{23}$, $\boldsymbol{\gamma}_{32}$ and $\boldsymbol{\lambda}_1, \dots, \boldsymbol{\lambda}_J$. If we would do so, then indeed $\boldsymbol{\theta} \mapsto h_0(\boldsymbol{\theta})$ is the likelihood function. The extension to multiple subjects/participants is straightforward, we run the backward information filter for each subject and multiply the resulting likelihoods.

If an element in $y_{tj}$ is missing, we can simply set $g_{tj}$ equal to a vector of length $k$ containing only ones. If a covariate vector $x_t$ is missing, we need to specify $\Pi_t$ separately. There are two natural choices here: the identity matrix (meaning no latent state-transition at the time the covariate is missing), or assigning equal probability to all state transitions.

The latter can be obtained by setting the vector of covariates (including the intercept) equal to the zero-vector. In our numerical results we have chosen the former option.

**Remark 1.** Direct implementation of the scheme in C.1 is numerically unstable. Instead, each time $h_t$ is computed we normalise it, i.e. we divide each element in the vector by the sum of all elements. If we denote this sum by $S_{\boldsymbol{\theta}}(h_t)$, then it follows that

C

$$\log L(\boldsymbol{\theta}) = \log h_0(\boldsymbol{\theta}) + \sum_{t=1}^{T} \log S_{\boldsymbol{\theta}}(\boldsymbol{h}_t). \tag{C.2}$$

This avoids numerical underflow problems.

### C.1.2. USING TURING

The basic implementation for inference in `Turing` reads as follows:

```
@model function logtarget(𝒪s, p)
    σ ~ Exponential(3.0)

    γ12 ~ filldist(Normal(0.0, σ), p.DIM_COVARIATES)
    γ23 ~ filldist(Normal(0.0, σ), p.DIM_COVARIATES)
    γ21 ~ filldist(Normal(0.0, σ), p.DIM_COVARIATES)
    γ32 ~ filldist(Normal(0.0, σ), p.DIM_COVARIATES)

    Z1 ~ filldist(Exponential(), p.NUM_HIDDENSTATES)
    Z2 ~ filldist(Exponential(), p.NUM_HIDDENSTATES)
    Z3 ~ filldist(Exponential(), p.NUM_HIDDENSTATES)
    Z4 ~ filldist(Exponential(), p.NUM_HIDDENSTATES)

    θ = ComponentArray(γ12 = γ12, γ23 = γ23, γ21 = γ21, γ32 = γ32,
                        Z1=Z1, Z2=Z2, Z3=Z3, Z4=Z4)

    Turing.@addlogprob! loglik(θ, 𝒪s, p)
end

model = logtarget(𝒪s, p)

map_estimate = maximum_a_posteriori(model)
mle_estimate = maximum_likelihood(model)
chain = sample(model, Turing.NUTS(), MCMCThreads(), 2000, 4)
```

In the first part of the model definition the prior specification is given. The observations are in the data-structure $\mathcal{O}s$ and all that is needed is a function that computes the loglikelihood as outlined in Section C.1.1. This function, called `loglik`, needs to implemented such that automatic-differentiation libraries can operate on it to compute the gradient of the log posterior density. `p` contains the number of covariates, the number of hidden states and number of questions.

Once the model has been specified, the MAP (Maximum A Posteriori)- and MLE (Maximum Likelihood Estimator) estimates can be computed and MCMC-sampling can be carried out to draw from the posterior. Here, we have chosen the No-U-Turn-Sampler.

**Remark 2.** Chapter 29.4.4 in [3] considers Bayesian Hidden Markov Models and remarks that a Gibbs sampler that alternates sampling from the smoothing distribution and updating the parameter $\theta$ may suffer from bad mixing due to high correlation between the latent path and $\theta$. Here, we follow what he calls "collapsed" inference, where the latent states of each person have been integrated out.

# 6

# GENERAL DISCUSSION

The thesis established a methodology for the individualised prediction of direct performance measures and biomechanical variables from wearable sensor data collected through repeated measurements in baseball pitching (chapter 2, chapter 3, chapter 4). By integrating an athlete's training exposure and health outcomes within a multistate injury framework (chapter 5), the thesis strengthens the information chain from training data collected with various data sources (wearable devices, motion capture system, self-reported questionnaires) via statistical models to actionable insights for injury risk management.

## 6.1. REPEATED MEASUREMENTS

6

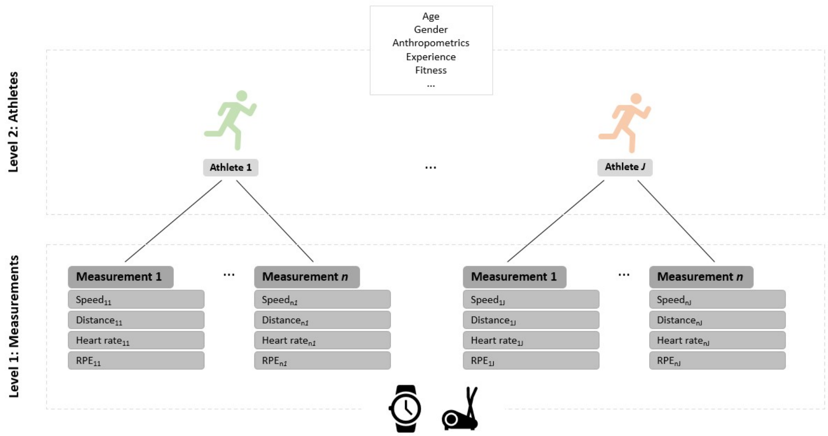


Figure 6.1.: Hierarchical structure of the repeated measures data.

Performance and health monitoring in sports often involves repeated measurements of individual athletes. This entails a single performance-related (e.g. speed, distance) or health-related measurements (e.g. heart rate, RPE) that are repeatedly collected over time on each athlete that we want to monitor. Such repeated measurements on the same individuals establish a two–level data hierarchy with measurements or observations as level 1 units and athletes as level 2 units (Figure 6.1). Measurements are clustered within the groups (athletes), where predictors can be available both on an individual–level (level 1) or a group–level (level 2). The groupings on level 2 often arise due to the differences between the individuals and their individual characteristics [1, 2]. In other words, different athletes may not perform equally in the measurements due to variations in their anthropometric measures, age, gender, experience and level of fitness. Therefore, when analysing a single athlete's data from repeated measurements, it is beneficial to treat each athlete as one group. Implicit to this example is the assumption that measurements from the same athlete are more similar than measurements from other athletes that are monitored at the same time. This supports a strong hierarchy in the structure of repeated measurement data due to generally more variation between individual athletes than between measurements within individuals.

The core of the thesis is an individualised analytical approach placing individual athlete in the centre. Each athlete has its own set of internal risk factors (e.g. age, gender, fitness level, history of injury) that when interacting with extrinsic risk factors, such as training load, determines the athlete's susceptibility to injuries. This means that individual athletes may respond differently to a given training stimulus and the training load required for positive adaptation may significantly differ between athletes despite a similar training background [3].

In the baseball setting, differences between pitchers of different ages and levels of play have already been well established in the literature. College level pitchers generally show higher trunk and pelvis peak angular velocities than youth and high school pitchers [4]. Furthermore, the development of pitchers' throwing abilities follows their physical development. Consequently, the stress applied on the elbow during pitching varies among pitchers of different ages, and levels of play [4] but also within an individual pitcher [5], and exposes pitchers to different levels of injury risk.

Chapter 2 and chapter 4 highlight the inherent between-pitcher differences and establish the methodology for the analysis of repeated measurement data when the interest lies in individualised prediction of future performance or injury outcomes. The chapter outcomes reveal that enriching kinematic data with the pitcher's personal characteristics results in better predictions of ball velocity [6] and elbow loading [7] in a single pitch. Chapter 5 proposes a multistate framework that binds an individual's training exposure and health status as a base for estimated risk of injury. Such an approach may support the coaching team in decision-making for an individual athlete and provide a tailor-made advice for individual performance improvement in a healthy manner.

## 6.2. WEARABLE SENSORS FOR WORKLOAD MONITORING IN BASEBALL PITCHING

Wearable sensors have become a widely used tool for the collection of positional, biometric and biomechanical data. Sensor technology integrated into watches, sleeves, straps and fabrics allows continuous data collection during physical activity and it is suitable for athletes of all levels and all sports disciplines. Such massive data production can be leveraged to create knowledge about the individual performance of athletes. The understanding of an athlete's body response to physical activity enables us to use wearable devices not only for data collection but also to provide personalised training advice to ensure consistent and safe participation in sport.

Since the 1920s, people have been trying to utilize baseball data to their advantage to predict outcomes and create winning teams [8]. Over the years, baseball grew into one of the biggest data-driven sports. However, even though data are driving the decision-making process, the use of data from wearables for workload monitoring and assessment of baseball pitching performance is still limited.

Pitching mechanics, pitch count and pitch type are considered the main factors in pitching training. A successful pitcher translates movement skills into a variation of ball velocities and trajectories to keep the batter off balance and complicate their anticipation of a particular pitch type. Slight changes in pelvis and trunk kinematics may result in a higher or lower ball velocity [9] which also determines the success of the intended pitch type [10–14]. However, the duration of the full pitching sequence is only 0.145s [15] which makes these small modifications in kinematics and timing difficult to assess and capture.



High-end wearable sensors, such as PitchPerfect (PitchPerfect, The Netherlands), are a suitable option for on-field measurements of pitching performance. Equipped with a gyroscope that can record angular velocities up to 2000 °/s, the PitchPerfect sensor system represents a solution for capturing the rapid pitching motion and subtle differences in pelvis and trunk kinematics and (inter)segmental timing. It enables individual pitcher monitoring, both during training and games, and captures every pitch thrown during warm–up, in the bullpen, as well as on the pitch.

To establish the relevance of kinematic data for baseball pitchers, it is necessary to develop the translation process between the collected data and personalised information provided to the pitcher. Kinematic data collected through repeated measurements often consider irregular patterns in the number of observations as well as inherited hierarchical structure. By accounting for hierarchy in the available data, we can apply a set of multilevel modelling techniques that will allow any pattern of measurements while providing statistically efficient parameter estimation [1].

This thesis investigates the application of Bayesian multilevel models [6, 7] and machine learning classification algorithms [16] for performance and workload prediction based on pitching mechanics. The type of data used as a model input is the same as data recorded with the PitchPerfect system, making adopted predictive models suitable for implementation in the system. The implementation offers a possibility for direct on-field feedback through existing mobile applications. The feedback is directed to the individual pitcher and contains pitch type recognition, pitching mechanics assessment and estimated ball velocity and elbow loading based on the PitchPerfect recordings. This creates a chain of information from longitudinal performance data via data science to actionable insight relevant to the improvement of the pitching performance and management of the elbow loading. However, the integration of new technologies and performance feedback is a timely process that requires interdisciplinary collaboration based on knowledge exchange between baseball practitioners, pitchers and scientists. This step is necessary to bridge the gap between science and practice and to bring new insights from pitch–to–pitch data to the field.

## 6.3. PREDICTING PITCHING PERFORMANCE

Detailed pitch-to-pitch information can be used for training adaptations aiming to improve the performance of a pitcher. To translate training success into game success, pitchers need to translate their movement skills into a variation of pitch trajectories.

Ball velocity plays an important role in the success of a baseball game. Throwing faster diminishes the batter's decision time of whether or not to strike the ball and restricts the offence's ability to advance bases and score runs [17, 18]. Therefore, when we talk about performance improvement in baseball pitching, this often implies maximizing ball velocity.

An individual's maximal throwing velocity is the result of optimal pitching mechanics [19]. The baseball pitch is a full-body throwing motion that starts when the pitcher lifts the lead foot, progresses to a linked motion in the pelvis and trunk, and ends with a whip-like action of the throwing arm propelling the ball towards home plate. The maximal ball velocity is a product of pelvis and trunk kinematics and their (inter)segmental timing that leads to the effective transfer of momentum to the baseball [19, 20].

Even if pitchers throw the baseball with a similar speed, the throws from the same pitcher tend to be more similar than the throws by other pitchers. Each pitcher has his or her own set of individual characteristics that enables them to throw faster or move better. This implies that optimal pitching mechanics are not of a unique size that fits all and that for each pitcher the optimal value of pelvis and trunk peak angular velocity may be slightly different. Consequently, inherent between-pitcher differences need to be considered in performance assessment.

Multilevel modelling techniques, such as Bayesian multilevel models adopted in chapter 2 and chapter 4 of the thesis, allow us to analyse the relationships between the data collected on the pitcher-level (height, weight) and the data that change with each measurement trial (pelvis and trunk peak angular velocity). Next to the pelvis and trunk peak angular velocities that ensure the optimal transfer of energy to the ball, including pitcher's length has an added value in assessment of the pitching performance. Pitcher's length is an individual characteristic that in addition to pitching kinematics improves the ball velocity prediction of a single pitch for every pitcher [6]. By using available observations from individual pitchers, Bayesian multilevel models can predict how fast the next single pitch of each of those pitchers will be with a high level of certainty. Additionally, using the group average, the same models can be used to predict how fast a new, out-of-sample pitcher can throw. The level of uncertainty in this case will be greater than if previous data on the new pitcher would already be available. Bayesian multilevel models account for hierarchical structure in the data and are suitable for the prediction of ball velocity not only for the athletes included in the analysis but also for the new, out-of-sample pitchers. This may provide valuable insight to the coaches and scouts in the process of player recruitment.



The implementation of the ball velocity prediction in practice creates more opportunities for monitoring multiple pitchers at the same time. Having an insight into the speed of every pitch contributes to personalised workload monitoring and can help pitchers understand the connection between their movement patterns and ball velocity. The adopted movement skills can then be used to alter ball velocity and ball trajectory for various pitch types.

Manual pitch type annotation is a common way of keeping track of pitches during training. Opposed to the big baseball games where the pitch type is detected based on ball data, in baseball practice the high-tech equipment for an automatic pitch type recognition task is often not available. Furthermore, existing methods for pitch type recognition are based purely on ball data, without considering variations in the pitching mechanics among the pitchers. From a strategic point of view, one could say that pitching kinematics for every pitch type is intended to stay the same to make it difficult for the batter to recognise the pitch [11]. However, in practice, the kinematic differences are visible in the data [10–14] and are expected to be more apparent among youth pitchers who still don't have enough strength and physical ability to properly throw off-speed pitches [21]. In chapter 3 of the thesis, we introduced a method for automatic pitch type recognition from kinematic data recorded with PitchPerfect sensors [16]. Even though the data sample used for training the machine learning model was limited, the outcome showed the potential of pitch type detection in a novel way.

Automatic detection of the fastball pitches has the potential for implementation in youth baseball training. Real-time information on pitch count, pitch type and pitching mechanics can be of great value for coaches and players. It can support youth pitchers in the process of adopting the right movement patterns for throwing fastballs with maximized speed. The detection of the three common pitch types (fastball, curveball, change–up) from kinematic data would have a great benefit for the players during the training process. However, for that further studies need to be conducted with larger data sets to increase the accuracy of the classification.

With the use of wearable devices during training every pitch counts. Adopted predictive models create an opportunity for direct insight into pitching mechanics, ball speed and pitch type and open the door for wearable data utilization on the field in the assessment of pitching performance and workload monitoring.

## 6.4. DEVELOPMENT OF AN INJURY RISK PROFILE

Integrating a pitcher's training schedule with performance measures and health outcomes represents the potential for the development of an injury risk profile. An Injury Risk Profile (IRP) strengthens the information chain from training and health data coming from various data sources via data science to the actionable insights aiming to support athletes and sports practitioners in decision–making and injury risk management process.



Sport-related injuries occur due to a complex interaction of many internal and external risk factors gathered in a pattern of either positive adaptation (increased fitness), or negative adaptation (injury). The repetitive nature of the high-speed full-body pitching movement exposes the pitcher's elbow to high loads. However, when high is too high, it is not easy to determine how much load will result in a positive adaptation and improved performance. Furthermore, measuring the exact value of load applied to the elbow during a single pitch is extremely complicated.

The injury aetiology seen in youth and adult pitchers has been linked to high elbow external valgus torques [22, 23]. The external valgus torque imparts a tensile force to the medial elbow structures [24, 25], which in combination with repetitive loading results in injuries to the medially located ulnar collateral ligament (UCL). This indicates that external valgus torque can be used as a proxy of elbow load [26, 27] and chapter 4 provides a suitable predictive model for estimating the value of the external valgus torque, hence elbow loading. The proposed Bayesian hierarchical model in chapter 4 uses an individualised approach to elbow load prediction based on the pitcher's weight, height, trunk peak angular velocity and the time between peak angular velocities of the pelvis and trunk [7]. Combining elbow load information with performance measures such as ball speed and pitch type would allow athletes to modify their pitching mechanics with an immediate insight into the effects those modifications have on performance metrics and elbow loading. In other words, it would allow a pitcher to fine–tune his movement pattern towards increased ball speed for a desired pitch type and reduced stress that is applied to an elbow.

An IRP can be provided in the form of a probability, carrying the information on the likelihood of sustaining an injury. However, an injury outcome is not always dichotomous. Injuries often arise from the mixture of acute and repetitive mechanisms, where they get overseen in the early stage of development due to the lack of apparent symptoms. Therefore, chapter 5 presents a multistate injury framework that can be the base for the development of an athlete's injury risk profile. An IRP can then be used as an early warning system providing feedback on injury risk which allows pitchers to modify their training schedule while their sport participation is still not affected. The multistate injury framework adopts a latent Markov model to infer the number of injury states based on the athlete's responses to the health questionnaire such as the OSTRC questionnaire. Inferred injury states are not observed and they represent different levels of injury status for an individual athlete. Given the prescribed training load, each athlete has a certain risk level that determines how likely it is to be in a specific injury state. Ideally, the latent Markov model would also have a multilevel structure that provides an individualised effect of prescribed training load in the prediction of injury risk trajectory. However, with available data, the implementation of such a model was not possible within the scope of this thesis.



The thesis illustrates a novel approach to individualised IRP prediction that accounts for the dynamics of the injury development process. The integration of advanced monitoring techniques plays an important role in the pursuit of high-level sports performance. The utilization of wearable sensors serves that purpose. It allows continuous athlete assessment and provides feedback on the relevant health and performance metrics in real–time. The methods established in the thesis offer solutions for dealing with different quality, time scale and hierarchical structures of the data collected with high-end wearable sensors, self-reported questionnaires and motion capture systems. Integration of the available data from different sources and implementation of the statistical models that are able to translate them to the relevant outcome, may provide an actionable insight for performance improvement and injury prevention in a variety of sports disciplines and improve training and injury prevention programs.

# REFERENCES


[1] H. Goldstein. *Multilevel statistical models*. John Wiley & Sons, 2011.

[2] B. Lambert. *A student's guide to Bayesian statistics*. SAGE Publications Ltd, 2018.

[3] T. Soligard, M. P. Schwellnus, J.-M. Alonso, R. Bahr, B. Clarsen, H. P. Dijkstra, T. Gabbett, M. Gleeson, M. Hägglund, M. R. Hutchinson, *et al.* "How much is too much?(Part 1) International Olympic Committee consensus statement on load in sport and risk of injury". In: (2016).

[4] G. S. Fleisig, S. W. Barrentine, N. Zheng, R. F. Escamilla, and J. R. Andrews. "Kinematic and kinetic comparison of baseball pitching among various levels of development". en. In: *Journal of Biomechanics* 32.12 (Dec. 1999), pp. 1371–1375. issn: 00219290. doi: 10.1016/S0021-9290(99)00127-X. url: https://linkinghub.elsevier.com/retrieve/pii/S002192909900127X.

[5] B. Van Trigt, F. F. Bouman, A. J. Leenen, M. J. Hoozemans, F. C. Van der Helm, and D. Veeger. "Quantifying Within-Individual Elbow Load Variability in Youth Elite Baseball Pitchers and Its Role in Overuse Injuries". In: *Applied Sciences* 12.13 (2022), p. 6549.

[6] L. Gomaz, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. "Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data". en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: 10.3390/s21227442. url: https://www.mdpi.com/1424-8220/21/22/7442.

[7] L. Gomaz, B. van Trigt, F. van der Meulen, and D. Veeger. "Predicting elbow load based on individual pelvis and trunk (inter)segmental rotations in fastball pitching". In: *Sports Biomechanics* (2024). doi: 10.1080/14763141.2024.2315230.

[8] J. Mizels, B. Erickson, and P. Chalmers. "Current state of data and analytics research in baseball". In: *Current reviews in musculoskeletal medicine* 15.4 (2022), pp. 283–290.

[9] E. Van der Graaff. "Perfecting your Pitch: In search of the perfect baseball pitch and its training". In: (2019).

[10] S. Dun, J. Loftice, G. S. Fleisig, D. Kingsley, and J. R. Andrews. “A Biomechanical Comparison of Youth Baseball Pitches: Is the Curveball Potentially Harmful?” en. In: *The American Journal of Sports Medicine* 36.4 (Apr. 2008), pp. 686–692. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546507310074. url: http://journals.sagepub.com/doi/10.1177/0363546507310074.

[11] R. F. Escamilla, G. S. Fleisig, S. W. Barrentine, N. Zheng, and J. R. Andrews. “Kinematic Comparisons of Throwing Different Types of Baseball Pitches”. en. In: *Journal of Applied Biomechanics* 14.1 (Feb. 1998), pp. 1–23. issn: 1065-8483, 1543-2688. doi: 10.1123/jab.14.1.1. url: https://journals.humankinetics.com/view/journals/jab/14/1/article-p1.xml.

[12] R. F. Escamilla, G. S. Fleisig, D. Groeschner, and K. Akizuki. “Biomechanical Comparisons Among Fastball, Slider, Curveball, and Changeup Pitch Types and Between Balls and Strikes in Professional Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 45.14 (Dec. 2017), pp. 3358–3367. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546517730052. url: http://journals.sagepub.com/doi/10.1177/0363546517730052.

[13] G. S. Fleisig, D. S. Kingsley, J. W. Loftice, K. P. Dinnen, R. Ranganathan, S. Dun, R. F. Escamilla, and J. R. Andrews. “Kinetic Comparison among the Fastball, Curveball, Change-up, and Slider in Collegiate Baseball Pitchers”. en. In: *The American Journal of Sports Medicine* 34.3 (Mar. 2006), pp. 423–430. issn: 0363-5465, 1552-3365. doi: 10.1177/0363546505280431. url: http://journals.sagepub.com/doi/10.1177/0363546505280431.

[14] G. S. Fleisig, W. A. Laughlin, K. T. Aune, E. L. Cain, J. R. Dugas, and J. R. Andrews. “Differences among fastball, curveball, and change-up pitching biomechanics across various levels of baseball”. en. In: *Sports Biomechanics* 15.2 (Apr. 2016), pp. 128–138. issn: 1476-3141, 1752-6116. doi: 10.1080/14763141.2016.1159319. url: http://www.tandfonline.com/doi/full/10.1080/14763141.2016.1159319.

[15] D. F. Stodden, B. M. Campbell, and T. M. Moyer. “Comparison of trunk kinematics in trunk training exercises and throwing”. In: *The Journal of Strength & Conditioning Research* 22.1 (2008), pp. 112–118.

[16] L. Gomaz, C. Bouwmeester, E. van der Graaff, B. van Trigt, and D. Veeger. “Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors”. In: *Sensors* 23.23 (2023), p. 9373.

[17] J. Hay. *The biomechanics of sports techniques*. Prentice-Hall, 1978.

[18] G. Lehman, E. J. Drinkwater, and D. G. Behm. “Correlation of throwing velocity to the results of lower-body field tests in male college baseball players”. In: *The journal of strength & conditioning research* 27.4 (2013), pp. 902–908.

[19] D. F. Stodden, G. S. Fleisig, S. P. McLean, and J. R. Andrews. “Relationship of biomechanical factors to baseball pitching velocity: within pitcher variation”. In: *Journal of applied biomechanics* 21.1 (2005), pp. 44–56.

[20] E. Van Der Graaff, M. ( Hoozemans, M. Nijhoff, M. Davidson, M. Hoezen, and D. ( Veeger. “Timing of peak pelvis and thorax rotation velocity in baseball pitching”. en. In: *The Journal of Physical Fitness and Sports Medicine* 7.5 (Sept. 2018), pp. 269–277. issn: 2186-8123, 2186-8131. doi: 10.7600/jpfsm.7.269. url: https://www.jstage.jst.go.jp/article/jpfsm/7/5/7_269/_article.

[21] B. Dowling, M. P. McNally, A. M. Chaudhari, and J. A. Oñate. “A review of workload-monitoring considerations for baseball pitchers”. In: *Journal of Athletic Training* 55.9 (2020), pp. 911–917.

[22] A. L. Aguinaldo and H. Chambers. “Correlation of throwing mechanics with elbow valgus load in adult baseball pitchers”. In: *The American journal of sports medicine* 37.10 (2009), pp. 2043–2048.

[23] F. D. Wilson, J. R. Andrews, T. A. Blackburn, and G. Mccluskey. “Valgus extension overload in the pitching elbow”. In: *The American journal of sports medicine* 11.2 (1983), pp. 83–88.

[24] E. L. Cain Jr, J. R. Dugas, R. S. Wolf, and J. R. Andrews. “Elbow injuries in throwing athletes: a current concepts review”. In: *The American journal of sports medicine* 31.4 (2003), pp. 621–635.

[25] M. Safran, C. S. Ahmad, and N. S. Elattrache. “Ulnar collateral ligament of the elbow”. In: *Arthroscopy: The Journal of Arthroscopic & Related Surgery* 21.11 (2005), pp. 1381–1395.

[26] A. W. Anz, B. D. Bushnell, L. P. Griffin, T. J. Noonan, M. R. Torry, and R. J. Hawkins. “Correlation of torque and elbow injury in professional baseball pitchers”. In: *The American journal of sports medicine* 38.7 (2010), pp. 1368–1374.

[27] D. Fortenbaugh, G. S. Fleisig, and J. R. Andrews. “Baseball Pitching Biomechanics in Relation to Injury Risk and Performance”. en. In: *Sports Health: A Multidisciplinary Approach* 1.4 (July 2009), pp. 314–320. issn: 1941-7381, 1941-0921. doi: 10.1177/1941738109338546. url: http://journals.sagepub.com/doi/10.1177/1941738109338546.

6

# DATA AND CODE AVAILABILITY

All research data and code supporting the findings described in this thesis are available in 4TU.ResearchData repository.

# ACKNOWLEDGEMENTS

My PhD journey has been full of challenges, uncertainties and obstacles. Throughout those, it taught me so much about myself, who I am and where I belong. I got to meet new cultures, adopt new traditions, learn a new language and meet incredible people along the way. Sometimes at the start of the journey we don't know where the road we chose is going to take us, and for quite some time I didn't have a clue where I was going. At the end, this road brought me exactly where I needed to be – it brought me home.

I want to thank my colleagues from TU Delft, CAS project and AISS for all the wonderful moments we shared during my PhD. Bart vG thank you for being a wonderful friend and peer mentor, and making me feel welcomed and accepted from the very beginning. Jasper, Andrea, Sebastiano, Simone, Mario, Rik, Andrea, thank you for all the nice office talks and coffee breaks. Martina, thank you for all the support. Lisa and Marian, I really enjoyed our Papendal reunions. Bart vT thank you for all the fruitful discussions. Pablo, Sietske and Justine, you brought a new light into my PhD life. Thank you for being wonderful colleagues and friends and for all the fun we had in and out of office hours.

Najveće hvala ide mojoj obitelji. Da nije bilo vas, ja ne bih bila osoba koja jesam. Mama, tata i Doli, hvala vam što ste bili moj vjetar u leđa kroz sve uspone i padove, što ste uvijek vjerovali u mene i podržali me u mojoj ludoj želji da svoje snove ostvarim u Nizozemskoj. Baka i dida, hvala vam na vašoj beskrajnoj podršci. Hvala vam što ste uvijek vjerovali u mene i čuvali mi leđa. Istinski sam blagoslovljena što imam takvu obitelj, volim vas najviše.

Een grote dank aan mijn schoonouders, schoonzussen en zwagers. Bedankt dat jullie mij in je familie hebt verwelkomd en voor de liefde en steun.

A special thank you goes to my Ljubavi. I am grateful for the life we are building together. Thank you for being you, for all your love and support. I am truly the luckiest person in the world. Volim te. And looking to the future, thank you to our little Coconut. You already changed our world and filled it with wonderful moments full of belly kicks and mermaid dances. We cannot wait to meet you!

# LIST OF PUBLICATIONS

## PUBLICATIONS INCLUDED IN THE THESIS


- **L. Gomaz**, D. Veeger, E. Van Der Graaff, B. Van Trigt, and F. Van Der Meulen. "Individualised Ball Speed Prediction in Baseball Pitching Based on IMU Data". en. In: *Sensors* 21.22 (Nov. 2021), p. 7442. issn: 1424-8220. doi: https://doi.org/10.3390/s21227442. url: https://www.mdpi.com/1424-8220/21/22/7442
- **L. Gomaz**, C. Bouwmeester, E. van der Graaff, B. van Trigt, and D. Veeger. "Machine Learning Approach for Pitch Type Classification Based on Pelvis and Trunk Kinematics Captured with Wearable Sensors". In: *Sensors* 23.23 (2023), p. 9373
- **L. Gomaz**, B. van Trigt, F. van der Meulen, and D. Veeger. "Predicting elbow load based on individual pelvis and trunk (inter)segmental rotations in fastball pitching". In: *Sports Biomechanics* (2024). doi: https://doi.org/10.1080/14763141.2024.2315230
- **L. van der Graaff**, E. Verhagen, and F. van der Meulen. "From injury prediction to risk assessment: latent Markov models for analysing OSTRC questionnaires". In: *BMJ Open Sport and Exercise Medicine* (Submitted)


## OTHER PUBLICATIONS


- E. Verhagen, B. Clarsen, **L. van der Graaff**, and R. Bahr. "Do not neglect injury severity and burden when assessing the effect of sports injury prevention interventions: time to paint the whole picture". In: *British Journal of Sports Medicine* (2024). issn: 0306-3674. doi: https://doi.org/10.1136/bjsports-2024-108215. eprint: https://bjsm.bmj.com/content/bjsports/early/2024/07/05/bjsports-2024-108215.full.pdf. url: https://bjsm.bmj.com/content/58/20/1166

# PHD PORTFOLIO

## PRESENTATIONS AT (INTER)NATIONAL CONFERENCES

- Individualised Ball Speed Prediction in Baseball Pitching based on IMU Data. *ECSS2022*. Sevilla, Spain, 2022

- From injury prevention to risk assessment – are we able to model real-time actionable insights from injury data? *FIMS2022*. Guadalajara, Mexico, 2022

- Scientific Research in Baseball: The past, present and future. *EBCA2022*. Amsterdam, Netherlands, 2022 (invited speaker)

- Data Science for Injury Prevention. *ISB Mid-Year Symposium 2023*, online, 2023 (invited speaker)

- From injury prediction to risk assessment. *GOTS2023*. Luxembourg, 2023

- Development of feedback systems for performance and injury risk. *Nationaal Congres SBG 2023*. Amsterdam, Netherlands, 2023 (invited speaker)

## GUEST LECTURES AND SCIENTIFIC MEETINGS

- Data science for sports performance improvement and injury prevention. — Scientific meeting at *English Institute of Sport*. London, UK, 2020

- Data Science for Injury Prevention — Guest lecture as part of the course Athlete Health Protection. *Vrije Universiteit Amsterdam*, 2022

## AWARDS

- Young Investigator Award 1st place - The 37th World Congress of Sports Medicine; Guadalajara, Mexico (2022)

## MSC STUDENT SUPERVISION

- L. Wu. Predicting Injury Risk with Machine Learning Methods using a Longitudinal Dataset (graduated in July 2020)

- C. Bouwmeester. Pitch type classification based on pelvis and trunk IMU data (graduated in September 2021)

## EDITORIAL BOARD PEER REVIEWED JOURNAL

- Associate Editor at BMJ Open Sport Exercise & Medicine

## WORKSHOPS AND SCIENTIfiC TRAINING

- Valencia International Bayesian Analysis Summer School; Valencia, Spain (2019)
- Study group Mathematics with Industry: Building a mathematical model to estimate joint loading and predict joint tissue damage; *Fontys University of Applied Sciences*, Tilburg (2020)
- Multilevel and longitudinal data analysis (Advanced course); *Leiden University* (2020)
- OpenSim Webinar - Inverse Kinematics: A Bayesian Versus Least – Squares Approach (2021)
- Software Carpentry Workshop (2021)
- Study group Mathematics with Industry: Towards personalized health monitoring; *University of Twente* (2022)

## COURSES

- How to select/make a questionnaire and conduct an interview
- Analysis of Interviews and other unstructured data
- Problem Solving and Decision–making
- Advanced problem solving and decision-making for researchers
- Time management - Fundamentals
- Using creativity to maximize productivity and innovation in your PhD
- Creative tools for scientific writing
- Popular scientific writing
- Speed reading and Mindmapping
- Dutch for non-native speakers
- Mental fitness intervention program
- Managing myself, leading others
- Professional and career development
- Professional and career development