



Dynamisk matematisk modell över glukosmetabolism under träning

A dynamic mathematical model for glucose metabolism during exercise

Kandidatarbete inom civilingenjörsutbildningen vid Chalmers

Hanna Åvall

Isabella Sanchez

Linnea Max

Mauritz Kööhler

Dynamisk matematisk modell över glukosmetabolism under träning

Kandidatarbete i matematik inom civilingenjörsprogrammet Bioteknik vid Chalmers

Hanna Åvall
Isabella Sanchez
Linnea Max
Mauritz Kööhler

Handledare: Sebastian Person
Niek Welkenhuysen

Institutionen för Matematiska vetenskaper
CHALMERS TEKNISKA HÖGSKOLA
GÖTEBORGS UNIVERSITET
Göteborg, Sverige 2021

Förord

Det här projektet har utförts av fyra personer som alla delat arbetsområden under processen. Innebörden av fördelningen var att alla fick en inblick inom både programmering, modellering, litteraturstudie och att skriva vetenskaplig text. Till en början utfördes kodning till stor del av Isabella Sanchez och Hanna Åvall medan Linnea Max och Mauritz Kööhler extraherade data och skrev i rapporten. En bit in i projektet byttes ansvarsområden så att Mauritz och Linnea kodade och modellerade mer medan Isabella och Hanna fördjupade sig i teorin och skrev rapport.

Under arbetet fördes loggbok kontinuerligt i form av veckovis dagbok samt individuell tidslogg av respektive deltagare. Varje vecka hade en person i gruppen ansvar för dagboken och skrev en sammanfattning över vad alla arbetat med under veckans gång.

Vi vill tacka Sebastian Persson, Niek Welkenhuysen och Marija Cvijovic som handlett oss under arbetet. Vi vill också tacka Tobias Rastemo för alla timmars hjälp med kodningen i Python.

Tabell 1 visar vilka avsnitt som tilldelats respektive författare.

Vem	Avsnitt
Linnea	1, 2, 3.1, 3.3, 4.2, 5, 6, Appendix A, Populärvetenskaplig text
Hanna	Sammanfattning, 1, 2, 3.3, 6, Figurer, Förord
Isabella	Sammanfattning, 1, 2, 3.2, 3.3, 4.3, 4.4, 6, Förord.
Mauritz	2, 3.3, 4.1, 4.2, 4.3, 4.4, 5, 6, Appendix A

Tabell 1: Bidragsrapport för kandidatgruppen MVEX01-21-03.

Populärvetenskaplig presentation

Kan man använda matematik för att förutse vilken träningsform en människa behöver för att vara hälsosam? Att göra det var i alla fall det långsiktiga syftet med det här arbetet. För att nå en bit på vägen ville vi först försöka fånga dynamiken i glukosmetabolismen under träning genom att utveckla en matematisk modell.

I ett samhälle där fysiska och psykiska sjukdomar är ett ökande problem bör fokus läggas på att förebygga dessa sjukdomar. Inte minst för att trycket på sjukvården ska minska utan även för att den åldrande befolkningen ska hålla så god hälsa som möjligt. Det är allmänt känt att regelbunden fysisk aktivitet skulle kunna vara en nyckel till detta problem [1]. Men för alla olika människor kanske inte svaret är mer träning för alla? Det svåra i frågan är att för varje individ bedöma vilken mängd och sorts träning som bidrar till bättre hälsa utan att medföra skador. För att göra detta behöver vi förstå vad som händer i våra kroppar när vi tränar. Ett sätt att undersöka det är genom matematisk modellering, där man kan beskriva ett biologiskt system genom en uppsättning matematiska ekvationer. Viktigt att komma ihåg är att en modell inte visar verkligheten men den kan användas för att få en bättre förståelse av verkligheten. En modell kan alltså ge oss en uppfattning om vad som händer i kroppen utan att många komplicerade experiment utförs.

Den process som är central i kroppen under träning är glukosmetabolismen då den förser muskler i rörelse med energi. Glukosmetabolismen är en komplex process med många reaktioner men kan kortfattat beskrivas som glukosets väg från mun till muskel där det omvandlas till energi. För att bygga en modell över systemet behöver förenklingar göras då en komplicerad modell med många delar är väldigt svårt att optimera. En viktig del att ha kvar är dock glukoskoncentration i blodet då det bland annat är här vi ser hur mycket glukos som finns tillgängligt för musklerna att ta upp. Detta är även en lämplig faktor att samla in experimentell data från för validera att modellen följer verkligheten. Det är ett viktigt steg då man annars inte kan dra några pålitliga slutsatser utifrån modellen.

När man har bestämt modellens delar och ekvationer börjar den iterativa processen för att göra modellen bättre och bättre. Det är inte speciellt troligt att man prickar rätt på första försöket utan man får testa sig fram för att se vilka delar som krävs för att kunna beskriva verkligheten med en matematisk modell. I detta projektet innebar det att börja med en enkel modell och sedan bygga på med fler delar för att bättre följa dynamiken i glukosmetabolismen. Resultatet av arbetet blev dock att modellen fortfarande inte följer dynamiken helt utan ytterligare arbete krävs. När man väl har en fungerande modell som validerats mot flera dataset, kanske man faktiskt kan säga något om hur träning påverkar olika personer.

Sammanfattning

För att förbättra den allmänna hälsan, både den mentala och fysiska, är träning ett bra tillvägagångssätt. Det som ger kroppen energi är glukosmetabolismen, ett system av reaktioner som omvandlar maten man äter till energi. När kroppen är under rörelse ökar kroppens energibehov, därför är glukosmetabolismen intressant att observera. I det här arbetet utnyttjades systembiologi för att undersöka glukosmetabolismens mekanismer under träning. Syftet var att fånga glukosmetabolismens dynamik med hjälp av en matematisk modell. Modellen som togs fram bestod av ett antal ODE:er som baserades dels på tidigare studier [2] men även på teori om reaktioner som ingår i glukosmetabolismen. För att lösa modellen användes programmeringsspråket Python och den verifierades med hjälp av experimentell data. Datan visade koncentrationer av glukos och insulin under medelintensiv träning och var hämtad från en studie på hästar[3] som i det här fallet användes som modellorganism. Modellen optimerades genom att uppskatta okända parametrar genom metoden Powell som minimerar avståndet mellan modell och data. En annan metod, Radau, användes för att lösa själva modellen. Resultat visade att en del av dynamiken, främst glukoskoncentrationerna, fångas relativt bra. Det fanns dock delar som inte fångades lika bra av modellen, däribland insulinkoncentrationen.

För att lyckas bättre hade det krävts att optimeraren använde en annan metod för att lösa modellen. Det hade också behövts bättre data som var lättare att matcha mot eftersom datan som användes hade en injektion av insulin att ta hänsyn till. Nästa steg om modellen hade fångat glukosmetabolismens dynamik hade varit att utveckla modellen så den fungerar för olika träningsintensiteter.

Abstract

A way to improve the overall health, both physical and mental, is regular training. Glucose metabolism is a system of reactions in the body that creates energy from food. When the body is in motion, the energy demand increases which makes the glucose metabolism interesting to investigate. In this projekt, we used system biology to investigate the mechanisms in the energy metabolism during exercise. The purpose was to capture the dynamic of the energy metabolism with a dynamic mathematical model. The model that was created, contained a set of ODE:s partially based on earlier stuides [2], but also on theory regarding reactions in the energy metabolism. To solve the model, the programming language Python was used and the model was verified against experimental data. The data in question was taken from a study of horses [3] which in this case were used as a model organism.

The model was optimized by approximating unknown parameters with Powell's method which can minimize the distance between model and data. Another method, the Radau method, was used to solve the ODE system. The result showed that some parts of the dynamic, especially the glucose concentration, was captured relatively well. However, there were parts that were not captured as good, among them was the insulin concentration.

To obtain a better fit, a different method for the optimization would be required to solve the model. It would also be better to use data that is easier to match since the data we used had an injection to take into consideration. Next step in the project, if the model had captured the dynamic of the glucose metabolism, would be to elaborate the model so that it could work with different training intensities.

Innehåll

1	Inledning	1
2	Syfte	2
3	Teori	3
3.1	Glukosmetabolism	3
3.1.1	Hormoner i metabolismen	3
3.1.2	Metabolism under träning	3
3.1.3	Skillnader mellan glukosmetabolism hos hästar och människor	3
3.2	Systembiologi	3
3.3	Matematiska teorier	4
3.3.1	Dynamiska ODE-modeller	4
3.3.2	Parameteruppskattning	5
3.3.3	Val av optimeringsalgoritm	5
3.3.4	Lösning av modellen	5
3.3.5	Latin hypercube sampling	6
3.3.6	Känslighetsanalys	6
3.3.7	Identifierbarhet	7
4	Metod	8
4.1	Val av programmeringsspråk	8
4.2	Matematisk modell	8
4.3	Data	8
4.4	Optimering	9
5	Resultat	11
5.1	GMOL modell	11
5.2	GMOL 2.0 modell	13
6	Diskussion	16
A	Kompletterande resultat till GMOL- och GMOL-modellerna 2.0	i
A.1	Modell GMOL	i
A.2	Modell GMOL2.0	ii
B	GitHub	iv

1 Inledning

Fysisk och mental hälsa är två viktiga områden med stort fokus i dagens samhälle. Fetma och psykiska sjukdomar ökar kontinuerligt och är stora samhällsproblem som påverkar vår levnadsstandard [4]. Dessutom ökar medellivslängden kontinuerligt och det är viktigt att de äldre håller så god hälsa som möjligt både för sin egen skull men också för att minska belastningen på sjukvården [5]. Ett sätt att förbättra den allmänna hälsan är regelbunden fysisk aktivitet då träning kan minska risken för bland annat fetma, kardiovaskulära sjukdomar och depression [1].

Det är dock inte alltid lätt att hitta en balanserad träningsnivå som ger en bättre hälsa utan att föra med sig skador. Både mängd och typ av träning har olika effekter på olika personer och skillnaden kan bland annat kopplas till kroppens metabolism [6]. Under fysisk träning ökas konsumtionen av glukos vilket gör det intressant att följa hur koncentrationer av glukos ändras i kroppen [7].

Ett sätt som kroppen utvinner energi på är via glukosmetabolismen som består av olika reglerade reaktioner som tillsammans bildar ett biologiskt system [8]. För att förstå hur glukosmetabolismen påverkas av träning är det nödvändigt att förstå viktiga mekanismer som ingår i den. En svårighet är att mäta koncentrationer av ämnen i vissa delar av kroppen. De studier som finns mäter ofta ämnen som finns i blodet, men i glukosmetabolismen ingår också bland annat lever, mage och muskler. Ett sätt att komma runt problemet är att göra en matematisk modell över glukosmetabolismens olika reaktioner, som kan approximera okända parametrar som motsvarar de koncentrationer som är svåra att mäta. För att skapa en modell som kan valideras krävs det att den bygger på experimentell data [9]. I tidigare forskning har det skapats en modell för att simulera just glukosmetabolism [2]. En bra modell kan återskapa och förutse verkligheten samt simulera olika förhållanden hos det biologiska systemet [10].

I det här arbetet ska en modell byggas för att simulera hur glukosmetabolismen ser ut under medelintensiv träning. Modellen matchas mot experimentell data från en studie [3] gjord på hästar. Hästar används här som modellorganism på grund av att det inte finns tillgänglig data från studier på människor.

Det här arbetet tar endast hänsyn till den datan som är insamlad på sex hästar som springer och äter en viss kost. En utveckling av projektet hade kunnat innebära att titta på olika sorters träning och intensiteter. Modellen som används är en förenklad bild av verkligheten och innehåller därför inte alla komponenter som ingår i glukosmetabolismen. Då modellen är en förenkling kan man ej dra några slutsatser om alla ingående reaktioner men övergripande dynamik kan beskrivas.

2 Syfte

Syftet med arbetet är att fånga dynamiken i glukosmetabolismen under träning genom att utveckla en matematisk modell. Modellen kommer vara en förenkling av verkligheten och ska efter vidare utveckling kunna användas för att studera olika träningsintensiteter och dess effekt på hälsan.

3 Teori

Teoridelen av rapporten redogör för den teoretiska bakgrund som krävs för att förstå grunderna som projektet baseras på. Inledningsvis förklaras processer i kroppen följt av grunderna inom systembiologi för att avslutas med de matematiska teorier som används.

3.1 Glukosmetabolism

Glukosmetabolismen är den kedja av reaktioner som levande organismer använder för att utvinna energi från mat. Maten bryts ned i magen för att sedan tas upp i blodet i form av mindre molekyler, däribland sockerarten glukos. Molekylerna transporteras till kroppens celler för att genom stegvis oxidation skapa kemisk energi i form av ATP [8]. ATP utnyttjas för att driva energikrävande processer i kroppen, till exempel andning och kroppsliga rörelser. Då ATP är en reaktiv molekyl som inte går att förvara, används istället glykogen för att lagra energi i kroppen, glykogen kan sedan användas för att skapa ATP vid behov [2], [11].

3.1.1 Hormoner i metabolismen

Samverkan mellan olika reaktionsserier i glukosmetabolismen är reglerade genom ett nätverk av hormoner som stimulerar eller inhiberar olika processer. Intag av mat höjer glukosnivåerna i blodet vilket stimulerar utsöndring av insulin, vars främsta funktion är att öka upptaget av glukos från blodplasman till muskel- och fettvävnad. Det innebär att insulin påverkar utnyttjandet av glukos och är en viktig faktor i att kontrollera glukosnivåerna i blodet [12]. I olika delar av kroppen, främst i musklerna och levern, lagras glukos i form av glykogen, vilket är en molekyl som består av flera glukosmolekyler som sitter ihop [13].

Om glukosnivåerna i blodplasman sjunker, kommer även insulinnivåerna sjunka och detta stimulerar hormonet glukagon. Glukagon i sin tur stimulerar nedbrytning av glykogen till glukos i levern, vilket oftast sker i fastande tillstånd eller under träning. Det glukos som bildas i levern utsöndras i blodet och kan med hjälp av höjda insulinnivåer tas upp i musklerna [2].

3.1.2 Metabolism under träning

När kroppen övergår från vila till träning ökar behovet av energi, ATP, i skelettmusklerna kraftigt. Av den anledningen är det intressant att involvera glukosnivåer i aktiverade muskler i modellen då glukos krävs för att skapa ATP. För att bibehålla nivåerna av ATP måste därför ATP-syntesen aktiveras hastigt [7].

3.1.3 Skillnader mellan glukosmetabolism hos hästar och människor

Glukosmetabolismen för människor och hästar har många likheter då de innehåller ungefär samma reaktioner. En av de större skillnaderna är kroppssammansättningen, där hästar generellt har en större andel muskelmassa i förhållande till kroppsvikt. Hästar har även högre syreupptagningsförmåga relativt till sin kroppsmassa. De här faktorerna kan variera från individ till individ för både människor och hästar [3].

Enligt Gabrielsson och Weiner [14] finns det en generell korrelation mellan kroppsvikt och metabolisk omsättning:

$$MR = a \cdot BW^{0,75} \quad (1)$$

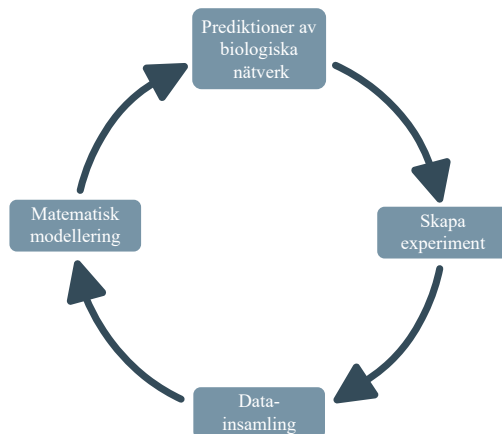
där MR är metabolisk omsättning, a är konstant med lämplig storhet och BW är kroppsvikt. Genom ekvation 1 kan en modell över metabolism översättas från häst till människa.

3.2 Systembiologi

Inom systembiologi studeras interaktioner och funktioner hos biologiska system med hjälp av matematik. Biologiska system kan vara mycket komplexa och därför tillämpas matematiska modeller

i kombination med experimentella studier för att få en bild av systemet. En bra modell kan förutse beteende i det biologiska systemet, till exempel koncentration av ett visst ämne under ett specifikt förhållande [10].

Den systembiologiska arbetsprocessen är en cyklisk process (Figur 1), där modellen uppdateras kontinuerligt utifrån resultat och experimentell data [15]. Systembiologins arbetsprocess börjar utifrån att skapa en frågeställning kring ett biologiskt system man vill undersöka. Data samlas in för att kunna skapa en matematisk en modell. Är modellen bra kan den prediktera systemet. För att undersöka om modellen faktiskt förutsäger verkligheten testas den mot olika dataset. Om det finns brister kan modellen förfinas och arbetscykeln börjar om igen [15].



Figur 1: Den cykliska arbetsprocessen inom systembiologi [15].

3.3 Matematiska teorier

Det här avsnittet fokuserar på att förklara den teoretiska bakgrunden till de matematiska metoder som används i rapporten och varför.

3.3.1 Dynamiska ODE-modeller

För att skapa modeller inom systembiologi kan ordinära differentialekvationer, ODE:er, användas. En ODE kan användas för att beskriva en hastighetsekvation och genom att bygga ett system av ODE:er kan koncentrationsändringar beräknas av olika ämnen i ett biologiskt system. För att beskriva ett exempel på en ODE utgår vi från följande reversibla reaktion,



där A och B är reaktanter som tillsammans bildar produkten C . k_1 och k_2 är hastighetskonstanter där k_1 representerar hastigheten som C bildas med och k_2 hastigheten som C förbrukas med. Genom massverkans lag kan en funktion härledas som visar hur de olika koncentrationerna ändras över tid. Massverkans lag säger att i alla reaktioner där reaktanternas stökiometriska koefficienter är lika med ett, är reaktionshastigheten proportionell mot produkten av reaktanternas koncentrationer [16]. Koncentrationsändringar per tidsenhet kan skrivas som ett system av ODE:er

$$\frac{dA}{dt} = \frac{dB}{dt} = -k_1 \cdot A_t \cdot B_t + k_2 \cdot C_t \quad (3)$$

$$\frac{dC}{dt} = k_1 \cdot A_t \cdot B_t - k_2 \cdot C_t = -\frac{dA}{dt} \quad (4)$$

där A_t , B_t och C_t är de aktuella koncentrationerna vid tiden t . De olika ODE:erna beror inte bara på andra koncentrationer i systemet utan även på tiden och det är tidsberoendet som gör systemet dynamiskt. Genom att integrera ODE:erna fås koncentrationerna fram för varje ämne [16]. Integrationen är svår att lösa analytiskt och därför används ofta numeriska metoder.

3.3.2 Parameteruppskattning

Parameteruppskattning handlar om att matcha en modell efter data och därmed uppskatta de parametrar i modellen som gör att den följer datan så noggrant som möjligt. I det här arbetet har vi en starkt icke-linjär ODE-modell som matchas mot experimentell data med ett mätfel som antas vara normalfördelat. Parametrarna i modellen motsvarar hastighetskonstanterna i reaktionerna. För att göra en parameteruppskattning skapas en kostnadsfunktion som beskriver avståndet mellan modellen och datan i olika tidpunkter. Kostnadsfunktionen härleds utifrån maximum likelihoodteorin. I en modell med givna parametrar, beskriver likelihood sannolikheten att observera tillgänglig mätdata [17]. Likelihooden när oberoende normalfördelade observationer antas beskrivs enligt ekvation 5, där σ^2 är variansen för normalfördelningen, y_i är mätvärdet på datan från experiment i , $\hat{y}$ är modellvärdet med uppsättningen parametrar $\mathbf{k}$, t är vid tidpunkterna $j = 1 \dots m$.

$$\mathcal{L}(\mathbf{k}) = \prod_{i=1}^n \prod_{j=1}^m \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(y_i(t_j) - \hat{y}_i(\mathbf{k}, t_j))^2}{2\sigma_i^2}\right) \quad (5)$$

Utifrån likelihooden går det att härleda minsta kvadratmetoden (ekvation 6) [17].

$$\mathbf{MK}(\tilde{\mathbf{k}}) = \sum_i^n \sum_j^m (y_i(t_j) - \hat{y}_i(t_j, \mathbf{k}))^2 \quad (6)$$

Att minimera minsta kostnadsfunktion (ekvation 6) är ekvivalent med att maximera likelihoodfunktionen. Kostnadsfunktionen minimeras med en iterativ algoritm. Då vi använder lokala algoritmer (en optimerare som inte garanterat hittar det globala minimumet) behöver denna process upprepas flertalet gånger, från olika startpunkter, för att finna det globala minimumet [17].

3.3.3 Val av optimeringsalgoritm

Det finns en rad olika algoritmer för att minimera funktioner och alla har olika för- och nackdelar. Det finns en balans mellan precision och hur tidskrävande processen är. Vissa metoder är till exempel bra på att hitta globala minimum, men är dyra beräkningsmässigt [18]. Powells metod är en adaptiv lösare som är gradientfri vilket innebär att den inte behöver gradienten till modellen för att minimera kostnadsfunktionen. Att ta fram gradienten till kostnadsfunktionen är krävande och svårt att implementera i Python. För att hitta ett minimum för en funktion kräver Powells metod en startgissning samt ett set av startriktningar. Därifrån testas lösaren att stega i de olika riktningarna och letar efter ett minimumvärde för funktionen via en iterativ process [19].

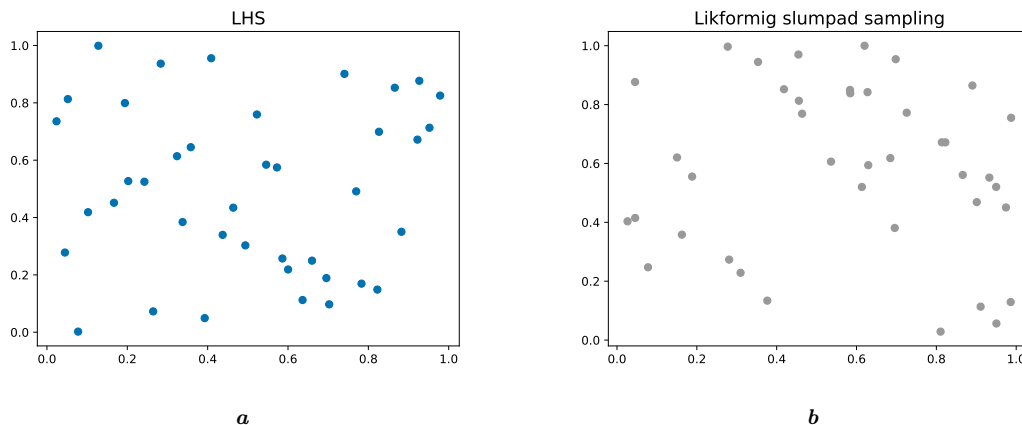
3.3.4 Lösning av modellen

För att minimera kostnadsfunktionen så behöver modellen, ODE:erna, först lösas. I det här projektet används ODE-lösaren Radau som finns i Scipy-biblioteket implementerat i Python. Radau tillhör klassen av implicita Runge-kutta metoder och kan lösa system som är styva (ett system som har snabba förändringar som går snabbt till steady state). Styva system är vanliga inom biologiska system. Radau är en adaptiv lösare, det innebär att den anpassar steglängden kontinuerligt istället för att ta lika stora steg hela tiden. Det här är fördelaktigt för att det ska gå snabbare att lösa modellen då lösaren vet när den kan ta korta och när den kan ta långa steg och ändå nå numerisk precision.

3.3.5 Latin hypercube sampling

När en algoritm ska hitta ett minimum för en funktion finns en risk att den fastnar vid ett lokalt minimum istället för det globala minimum som söks. Orsaken är att algoritmen hittar det minimum som är närmast given startgissning på parametrarna. En startgissning består av ett värde inom valt intervall för varje parameter som ska optimeras. Lösningarna för varje parameter bidrar med en dimension till Lösningensrummet för parametrarna. För att generera flera slumpmässiga startgissningar fördelade över parameterrummet används metoden latin hypercube sampling och på så sätt optimeras chanserna att hitta det globala minimumet [20].

En hyperkub är ett parameterrum där N antal startgissningar genereras i en Latin hypercube sampling, LHS (Figur 2a). Istället för stokastiskt utplacerade startgissningar gör hyperkuben så att startgissningarna är jämt fördelade över parameterrummet. Slumpmässigt valda punkter från en uniform likformig fördelning har en tendens att lägga sig i kluster vilket inte ger en bra spridning (Figur 2b). Genom en god spridning ökar chansen att hitta så många minimum som möjligt och förhoppningsvis även det globala minimumet. Utrymmet delas upp som startgissningar och fördelas i N^2 rutor. Det första startvärdet tilldelas en ruta i första kolumnen där gissningen placeras slumpmässigt inom rutan. På liknande sätt placeras andra gissningen i en ruta i andra raden, dock inte i en kolumn som har blivit vald sedan tidigare. På det här sättet förebygger latin hypercube sampling att två slumpmässigt utplacerade startgissningar är för nära varandra och ger därför en bättre täckning av parameterrummet [20].



Figur 2: a) är 40 samplingar genererade från en Latin hypercube. b) 40 likformigt slumpade utfall. I b) ligger vissa samplingar på varanandra eller nära varandra vilket leder till att det på vissa ytor inte finns några samplingar alls. Medan i a) är samplingarna mer utspridda men fortsatt slumpade.

3.3.6 Känslighetsanalys

I en känslighetsanalys studeras hur förändringar av en enskild parameter i en modell påverkar modellens lösning. Detta uppnås här genom att använda finita differenser [17].

Känsligheten definieras som gradienten av lösningen för en av variablerna i en given tidpunkt [21]. Gradienten kommer innehålla partiella derivator $\frac{\partial \hat{y}_i(t_j)}{\partial k_k}$ för alla $k = 1, \dots, p$ parametrar. För att approximeras dessa partiella derivator används finita differenser:

$$\frac{\partial \hat{y}}{\partial k_i} \approx \frac{\hat{y}(x, k_1, \dots, k_i + \delta, \dots, k_p, t) - \hat{y}(x, k_1, \dots, k_i, \dots, k_p, t)}{\delta} \quad (7)$$

Problemet med ekvation 7 är att steglängden δ måste väljas. Då parametrarna i en biologisk modell varierar kraftigt så betyder det att ett värde på δ kan vara bra för en parameter och dålig för en annan. Trots detta går det att visa att $\delta = \sqrt{\epsilon}$, där ϵ är maskintoleransen, ger en tillräckligt bra approximation [22].

Genom att beräkna sensitiviteten för alla variabler i ett antal tidpunkter kan en matris skapas som kallas för känslighetsmatrisen.

$$S = \begin{pmatrix} \frac{\partial \hat{y}_1(t_1)}{\partial k_1} & \dots & \frac{\partial \hat{y}_1(t_m)}{\partial k_1} & \frac{\partial \hat{y}_2(t_1)}{\partial k_1} & \dots & \frac{\partial \hat{y}_2(t_m)}{\partial k_1} & \frac{\partial \hat{y}_3(t_1)}{\partial k_1} & \dots & \frac{\partial \hat{y}_3(t_m)}{\partial k_1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \hat{y}_1(t_1)}{\partial k_p} & \dots & \frac{\partial \hat{y}_1(t_m)}{\partial k_p} & \frac{\partial \hat{y}_2(t_1)}{\partial k_p} & \dots & \frac{\partial \hat{y}_2(t_m)}{\partial k_p} & \frac{\partial \hat{y}_3(t_1)}{\partial k_p} & \dots & \frac{\partial \hat{y}_3(t_m)}{\partial k_p} \end{pmatrix} \quad (8)$$

Sensitiviteten för varje variabel normeras, summeras och divideras sedan med antalet tidpunkter för att skapa ett RMS-värde vilket är en förkortning av Root Mean Square och på svenska samma som kvadratisk medelvärde.

$$RMS = \sqrt{\frac{1}{N} \sum_{j=1}^m \left(\frac{\partial y_i(t_j)}{\partial p_k} \right)^2 \left(\frac{p_k}{y_i(t_j)} \right)^2} \quad (9)$$

Ett lågt värde relativt de andra kan innebära att den parametern har liten eller ingen betydelse för lösningen och borttagningen av det beskrivna värdet kan således förenkla modellen [21].

3.3.7 Identifierbarhet

Identifierbarhet är en metod för att undersöka om ett dataset är kompatibelt med en modell, det vill säga att den insamlade datan faktiskt kan användas för att unikt bestämma de okända parametrarna. Det som undersöks är om det är möjligt att identifiera en unik vektor av parametervärden för en given modellstruktur eller om många parametervärden passar lika bra till datan. Om datan är rätt och passar till modellen, är den identifierbar vilket innebär att det finns en liten eller ingen osäkerhet i parametrarna. Om modellen inte är identifierbar och osäkerheten är stor betyder det att datan inte innehåller informationen som behövs för att kunna ge tillförlitliga parametervärden. En parameter är identifierbar om varianskoefficienten för parametern är låg. Lågt värde på varianskoefficienten beskriver att parameterns varians är låg om ett annat dataset används. En tumregel är att om det finns fler obekanta variabler än ekvationer kan det leda till att det existerar ett oändligt antal lösningar och resulterar i att modellen inte är identifierbar.

Känslighetsanalys och identifierbarhet har vissa likheter bortsett från att känslighetsanalys tittar på modellparametrar i olika punkter medan identifierbarhet tittar på datapunkter. Om funktionen är okänslig för en parameter så innebär det även att den inte är identifierbar. Att en modell är känslig för alla parametrar betyder dock inte att de alla är identifierbara [21]. Då känslighetsanalys och identifierbarhet är så pass lika kommer endast identifierbarheten att undersökas i det här arbetet.

4 Metod

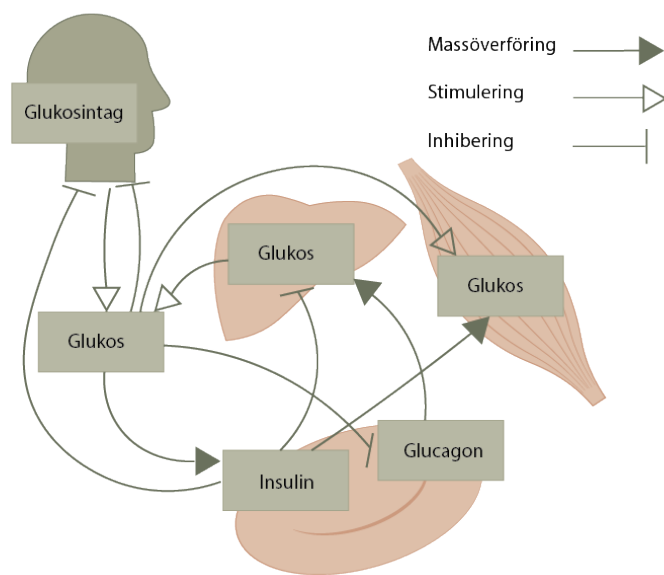
Metoddelen beskriver hur arbetsprocessen har gått tillväga under projektet. Den involverar även vilka modeller som tagits fram och ändringar som gjordes mellan modellerna.

4.1 Val av programmeringsspråk

Initialt i arbetetprocessen beslutades det att använda sig av programmeringsspråket Python på grund av dess generella struktur och användarvänlighet. Versionen av Python som användes var 3.8.5 tillsammans med paketen Scipy (version 1.5.2) samt Numpy (version 1.19.2).

4.2 Matematisk modell

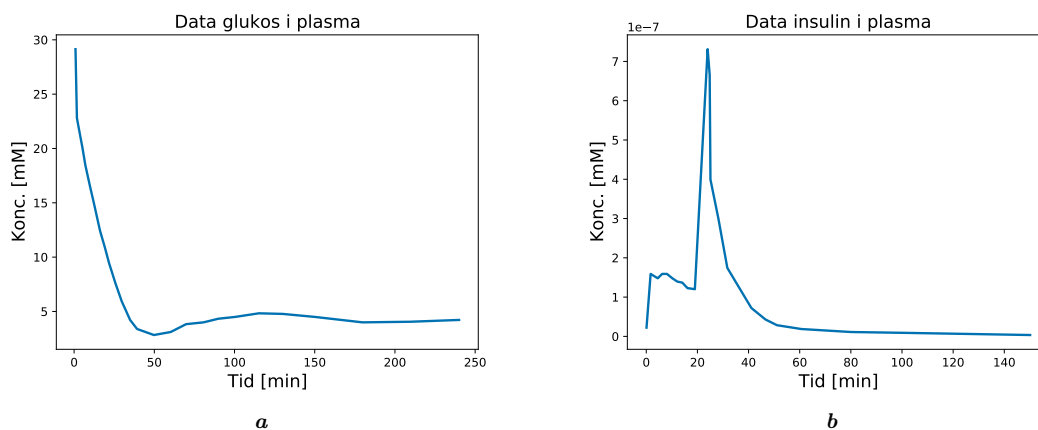
För att beskriva glukosmetabolismen (Figur 3) skapades en modell av typen open loop. Open loop modeller innebär att den har ett inflöde och ett utflöde. Modellen GMOL har ett inflöde av glukos från magen och ett utflöde av glukos som förbrukas i kroppens muskler. Tidigare modeller [2], [11], [23] innehåller många komponenter som skapade komplexa dynamiska system. I den här arbetet gjordes förenklingar av tidigare modell [2], baserat på teori. Essentiella delar ur glukosmetabolismen valdes ut som grund till en första modell. Processen var iterativ och många förändringar gjordes under arbetets gång. Det resulterade slutligen i två modeller, GMOL och GMOL 2.0, av olika komplexitet.



Figur 3: Illustration av biologin bakom modellerna, man kan följa glukosens väg från mun till muskel, där glukos i blodplasman stimulerar insulin som i sin tur stimulerar upptag av glukos i muskeln. Om glukosnivåerna i blodplasman sjunker kommer nivåerna av glukagon stiga och stimulera produktion av glukos i levern för att kunna bibehålla lagom koncentration av glukos i blodet.

4.3 Data

Den uppmätta datan [3], är från en studie av hästar under träning och innehåller koncentrationen glukos i plasma (Figur 4a) och koncentrationen insulin i plasma (Figur 4b).



Figur 4: a) Genomsnittliga glukosvärdet för 6 hästar under träning i 240 minuter där glukos injicerades vid tidpunkten 0 minuter. b) genomsnittliga insulinvärdet för 6 hästar under träning i 150 minuter där insulin injicerades vid tidpunkten 20 minuter. Insulin slutar vid 150 minuter för att den tangerar mot 0 därefter.

I studien ingick 12 hästar som delades in i två lika stora grupper om 6 hästar som fick äta olika kost. Vi undersöker endast datan från en av grupperna där gruppen härnå efter refereras som FF. Innan studien utfördes fick hästarna träna 3 gånger i veckan under en period på 8 månader. När testet utfördes sprang hästarna i en hastighet som räknades som medelintensiv träning. De 6 hästarna i FF hade fri tillgång till foder 12 timmar inför testet och 30 minuter innan testet sprang hästarna långsamt som uppvärmning. Vid start av testet injicerades $3,33 \text{ [mmol glukos / kg bw]}$ och vid 20 minuter injicerades $6,94 \cdot 10^{-8} \text{ [mmol insulin / kg bw]}$.

För att extrahera data från artikeln användes verktyget WebPlotDigitizer och endast data från FF extraherades. Datat som extraherades för glukos var i $[\text{g/l}]$ och insulin var i $[\text{mU/l}]$. Dessa enheter konverterades till millimolar $[\text{mM}]$.

4.4 Optimering

Modellerna i rapporten optimerades genom att minimera kostnadsfunktionen, ekvation 6. Metoden som användes för att optimera modellerna var Powells metod som är inbyggd i Scipy biblioteket. Då Powells metod letar efter parametrar som ska ge det lägsta värdet på kostnadsfunktionen sattes vissa begränsningar för inom vilket spann koncentrationerna i modellen fick vara. Om Powells metod föreslog parametrar som gjorde att koncentrationerna blev större eller mindre än begränsningarna så straffades kostnadsfunktionen med ett tillägg. Om värdet för kostnadsfunktionen ökade, var det inte den optimala uppsättningen parametrar.

Powells metod behöver en startpunkt för att kunna utföra sin optimering. Då Powells metod optimerar med avseende på parametrarna för modellen är det startgissningar på dem som behövs. Startgissningarna genereras utifrån en latin hypercube i lika många dimensioner som parametrar. Sidorna på kuberna byggs upp av längden på intervallen för varje parameter. För att få en uppfattning om storlekarna på intervallen för parametrarna samplades tusen startgissningar från en likformig logaritmisk fördelning med storlek mellan 10^{-3} och 10^3 . Från de startgissningarna beräknades kostnadsfunktionen. Den uppsättningen parametrar som genererade det lägsta värdet på kostnadsfunktionen användes för att generera intervallen för hyperkuben. Intervallet skapades från att välja den övre gränsen till en hel tiopotens över det beräknade parametervärdet. Den undre gränsen sattes till 0 för alla parametrar. Hyperkuben för GMOL (Tabell 2a) och hyperkuben för GMOL 2.0 (Tabell 2b) fick olika långa sidor.

Tabell 2: Dimensionerna för hyperkuben. a) Intervallen för modell GMOL. Enheterna för k_2 och k_4 är $[mM^{-1} \cdot \text{min}^{-1}]$, resten har enheten $[\text{min}^{-1}]$. b) Intervallen för modellen GMOL 2.0. Enheterna för k_2 och k_8 är $[mM^{-1} \cdot \text{min}^{-1}]$, resten har enheten $[\text{min}^{-1}]$.

Parameter	Intervall
k1	0-0.01
k2	0-100
k3	0-10
k4	0-0.001
k5	0-0.01

a

Parameter	Intervall
k1	0-0.1
k2	0-0.01
k3	0-0.01
k4	0-0.01
k5	0-0.01
k6	0-10
k7	0-0.01
k8	0-1

b

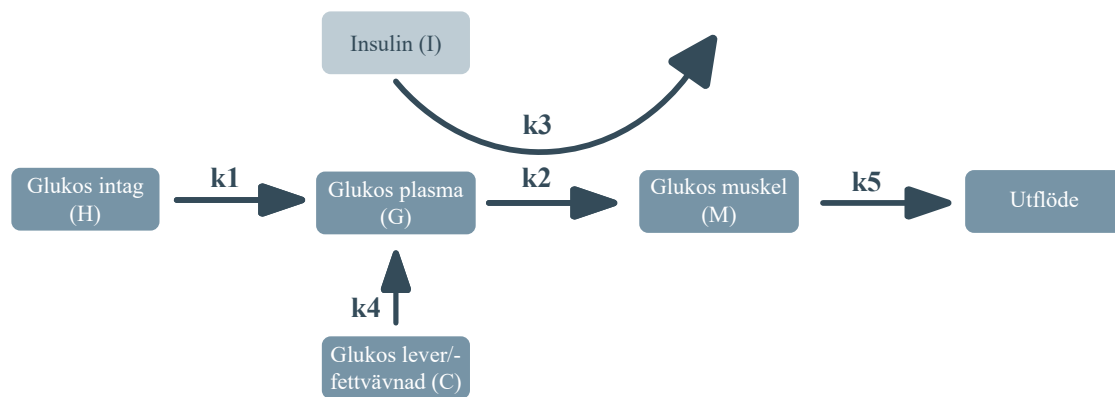
Under optimeringen löses modellerna upprepade gånger för olika parametervärden men vid samma startkoncentrationer för systemet. För glukos och insulin i plasma kunde startkoncentrationer tas direkt från den extraherade datan. För resterande startkoncentrationer användes kvalificerade gissningar.

5 Resultat

Resultatdelen redovisar resultat av två modeller GMOL och GMOL 2.0. För varje modell presenteras ODE:er samt schematisk representation tillsammans med identifierbarhetsanalys av glukos och insulin i plasma. Appendix A, detaljerad analys av andra koncentrationer samt beräknade optimala parametrar.

5.1 GMOL modell

Modellen GMOL (Figur 5) är en enkel modellstruktur som endast tar med de mest essentiella delarna i glukosmetabolismen. Det hormonet som är med i modellen GMOL är insulin, då det ansågs vara viktigast för glukosnivåer i blodet och överförandet av glukos till muskler. De förenklingar av verkligheten som gjordes var att glukos efter intag hamnar direkt i blodplasman, alltså utan mellanstegen mage och tarmar. Ytterligare en förenkling var att fettvävnaden helt togs bort och det antogs att levern kunde symbolisera energireserverna som används under träning och fasta. De delar som är med i modellen är således blodplasma, lever och skelettmuskler. Hela modellen går ut på att visa en så enkel väg som möjligt för glukos från vår mun till skelettmuskeln, där det kommer att förbrukas enligt en parameter som beror på koncentration av glukos i muskeln. För att beskriva dessa processer skapades en dynamisk modell med fem ODE:er och fem parametrar, där varje ODE beskriver en koncentration och varje parameter beskriver hastigheten på en process.



Figur 5: En schematisk bild för de olika delarna i GMOL modellen och hur dessa påverkar varandra. En rak pil betyder masstransport och en böjd pil betyder ett hormon som stimulerar en reaktion. Varje reaktion har en hastighetskonstant k som är parametrarna för modellen.

GMOL-modellens fem ODE:er beskrivs nedan, där G (Ekvation 10a) är glukoskoncentration i plasma, I (Ekvation 10b) är insulinkoncentration i plasma, C (Ekvation 10c) är glukoskoncentration i levern där L är det kontinuerliga glukosbidrag levern utsöndrar, M (Ekvation 10d) är glukoskoncentration i muskeln och H (Ekvation 10e) är intaget av glukos.

$$\frac{dG}{dt} = k_4CI + k_1H - k_2GI \quad (10a)$$

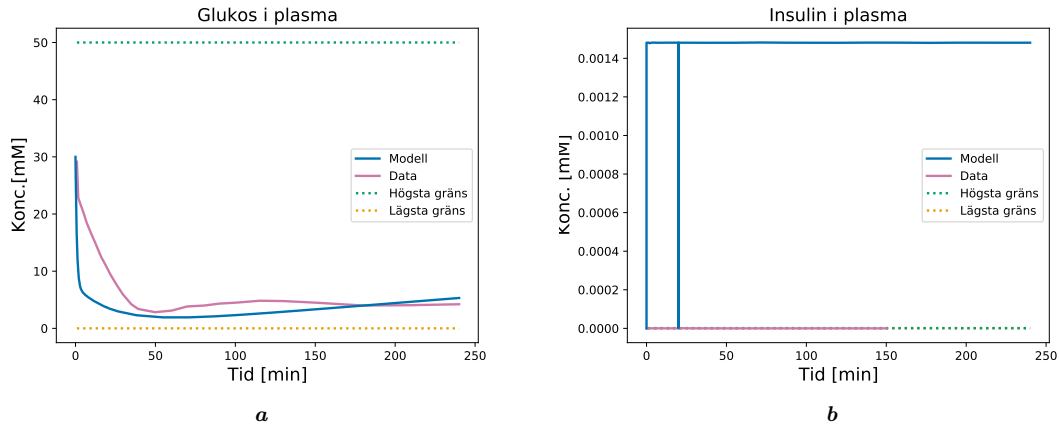
$$\frac{dI}{dt} = k_3G - k_2GI \quad (10b)$$

$$\frac{dC}{dt} = -k_4CI + L \quad (10c)$$

$$\frac{dM}{dt} = k_2GI - k_5M \quad (10d)$$

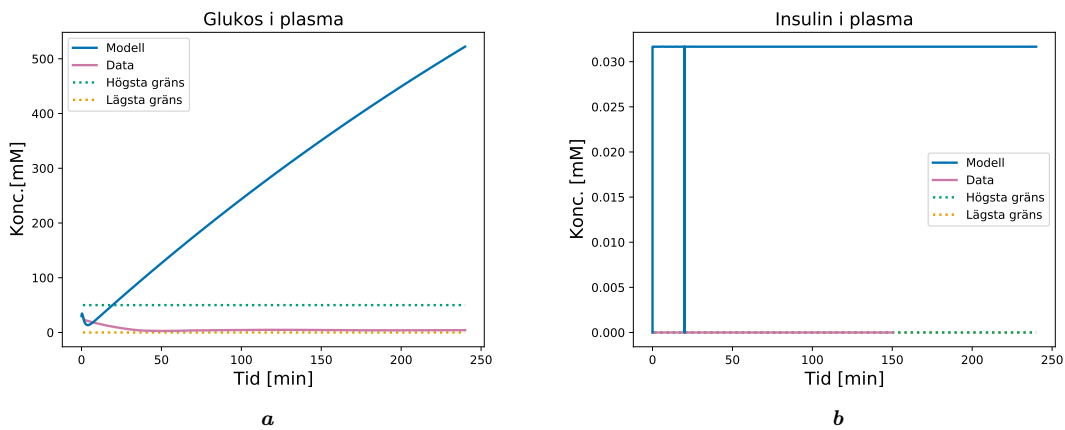
$$\frac{dH}{dt} = -k_1H \quad (10e)$$

För att uppskatta hur stora intervallen för parametrarna borde vara togs 1000 startgissningar från en logaritmisk uniform fördelning med värden mellan 10^{-3} och 10^3 . Intervallet är vanligt för storlekarna för biologiska parametrar. Dessa värden sattes in i kostnadsfunktionen och den uppsättning parametrar som gav lägsta värdet på kostnadsfunktionen användes för att skapa intervallen. Parametrarna som gav det lägsta värdet på kostnadsfunktionen resulterade i koncentrationerna för glukos (Figur 6a) och insulin (Figur 6b).



Figur 6: a) Glukoskoncentrationen i plasman utan optimering i GMOL. b) Insulinkoncentration i plasman utan optimering i GMOL. Den blå linjen visar modellen medan den rosa linjen visar datan för FF [3] där glukos tillsattes vid tidpunkten 0 minuter och insulin tillsattes efter 20 minuter. De streckade linjerna visar de övre och undre gränserna för vad som var tillåtna värden för modellen. Den nedre gränsen i b) ligger på den undre gränsen då koncentrationen av insulin överstiger den övre gränsen så mycket. Det vertikala streck som syns på insulin vid 20 minuter kommer från att vi sätter koncentrationen till värdet på datan för att simulera injektionen.

Utifrån intervallen för parametrarna konstruerades hyperkuben som gav startvärden på parametrarna till optimeringen. Resultatet för glukos (Figur 7a) och insulin (Figur 7b) visar att modellen inte fångar datavärdenas dynamik då förändringarna är för hastiga i jämförelse. Övriga koncentrationer i modellen finns i appendix A.



Figur 7: a) Glukoskoncentrationen i plasman efter optimering i GMOL. b) Insulinkoncentrationen i plasman i GMOL modellen. Den blå linjen visar modellen medan den rosa linjen visar datan för FF [3] där glukos tillsattes vid tidpunkten 0 minuter och insulin tillsattes efter 20 minuter. De streckade linjerna visar de övre och undre gränserna för vad som var tillåtna värden för modellen. Den nedre gränsen i b) ligger på den undre gränsen då koncentrationen av insulin överstiger den övre gränsen så mycket. Det vertikala streck som syns på insulin vid 20 minuter kommer från att vi sätter koncentrationen till värdet på datan för att simulera injektionen.

Varianskoefficienterna för de fem parametrarna i GMOL modellen (Tabell 3) visar att storleken av koefficienterna varierar men är generellt hög för alla parametrar.

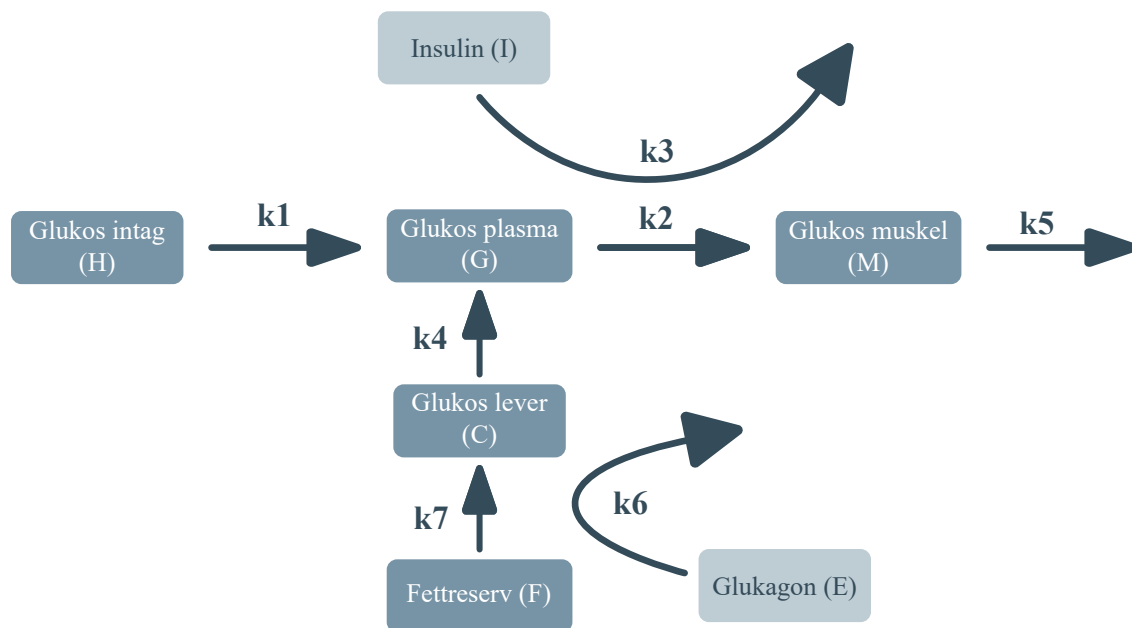
Tabell 3: Varianskoefficienterna för de fem parametrarna i GMOL-modellen.

k1	k2	k3	k4	k5
$9.91 \cdot 10^7$	$3.43 \cdot 10^1$	$2.46 \cdot 10^7$	$2.50 \cdot 10^6$	$1.61 \cdot 10^7$

Sammanfattningsvis, dynamik av glukos (Figur 7a) och insulin (Figur 7b) i plasma tillsammans med varianskoefficienterna (Tabell 3) föreslår att modellen GMOL inte fångar dynamiken väl. För att fånga dynamiken behövs en modifiering av modellen.

5.2 GMOL 2.0 modell

För att se om dynamiken i glukosmetabolismen kunde fångas ännu bättre utvecklades GMOL till GMOL 2.0 (Figur 8). Fettreserverna som tidigare antagits kunna symboliseras av levern lades till som en egen del. Fettreserverna är en del som ökar inflödet av glukos till modellen. Hormonet glukagon, som stiger när insulin sjunker, lades också till. Detta för att öka produktionen av glukos i levern och även på så sätt få in mer glukos i systemet. Modellen GMOL 2.0 innehåller två fler ODE:er, fettreserver och glukagon, och tre fler parametrar vilket gör systemet mer komplext.



Figur 8: Här visas hur de olika delarna i modellen GMOL 2.0 påverkar varandra genom olika parametrar. En rak pil betyder masstransport och en böjd pil betyder ett hormon som ökar på en reaktion

GMOL 2.0 modellens sju ODE:er beskrivs nedan, där G (Ekvation 11a) är glukoskoncentration i plasma, I (Ekvation 11b) är insulinkoncentration i plasma, C (Ekvation 11c) är glukoskoncentration i levern, M (Ekvation 11d) är glukoskoncentration i muskeln, H (Ekvation 11e) är intaget av glukos, E (Ekvation 11f) är koncentration av glukagon i plasma och F (Ekvation 11g) är glukos som transporteras från fettreserver.

$$\frac{dG}{dt} = k_4C + k_1H - k_2GI \quad (11a)$$

$$\frac{dI}{dt} = k_3G - k_2GI \quad (11b)$$

$$\frac{dC}{dt} = -k_4C + k_6E + k_7F \quad (11c)$$

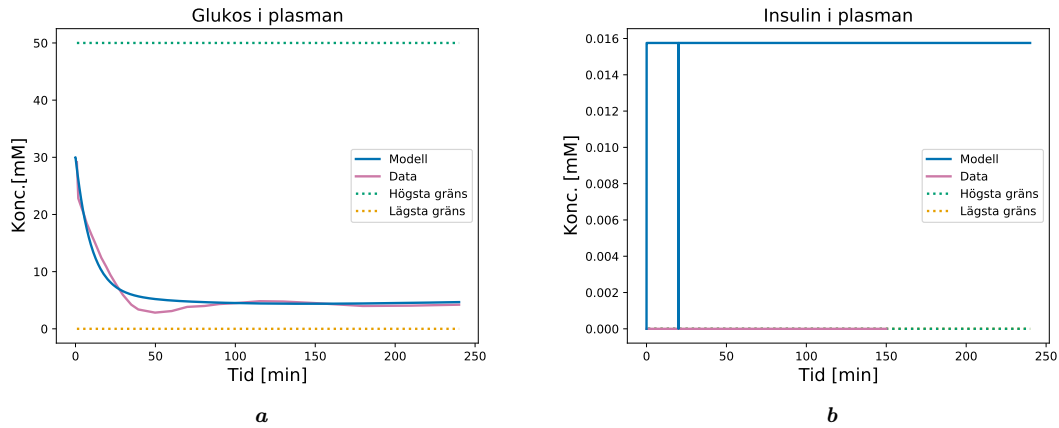
$$\frac{dM}{dt} = k_2GI - k_5M \quad (11d)$$

$$\frac{dH}{dt} = -k_1H \quad (11e)$$

$$\frac{dE}{dt} = k_8 - k_2GE \quad (11f)$$

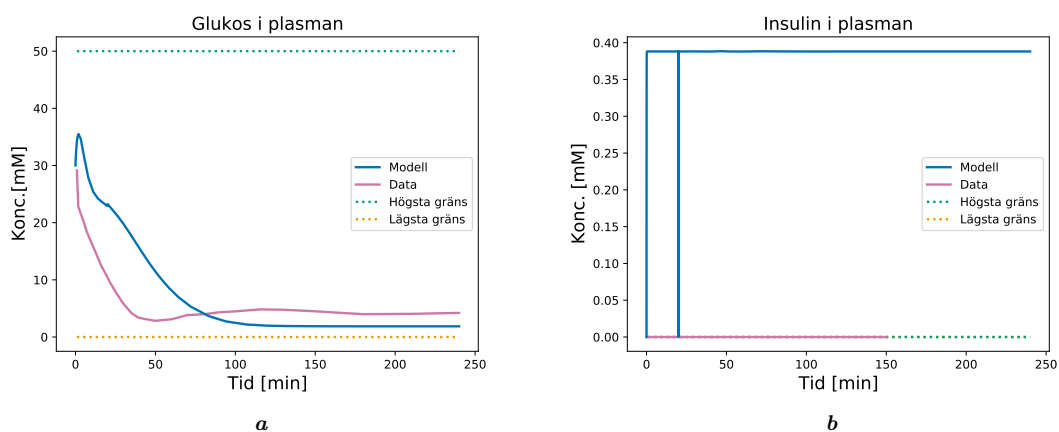
$$\frac{dF}{dt} = -k_7F \quad (11g)$$

För att genomföra optimeringen för modellen behövs en uppskattning av storlekarna på parametrarna. Detta genomfördes som för modell GMOL genom att generera 1000 uppsättningar från en logaritmisk uniform fördelning med storlek mellan 10^{-3} och 10^3 . Den uppsättning parametrar som gav lägst värde på kostnadsfunktionen användes för att plotta upp modellen för glukos (Figur 9a) och insulin (Figur 9b). Koncentrationerna av glukos och insulin i modellen jämförs med tidsseriedata från hästar i FF.



Figur 9: a) Glukoskoncentrationen i plasman utan optimering i GMOL 2.0. b) Insulinkoncentrationen i plasman utan optimering i GMOL 2.0. Den blå linjen visar modellen medan den rosa linjen visar datan för FF [3] där glukos tillsattes vid tidpunkten 0 minuter och insulin tillsattes efter 20 minuter. De streckade linjerna visar de övre och undre gränserna för vad som var tillåtna värden för modellen. Den nedre gränsen i b) ligger på den undre gränsen då koncentrationen av insulin överstiger den övre gränsen så mycket. Det vertikala streck som syns på insulin vid 20 minuter kommer från att vi sätter koncentrationen till värdet på datan för att simulera injektionen.

Parametrarna från fördelningen som gav lägst värdet på kostnadsfunktionen användes för att skapa intervall för varje parameter som utgör hyperkuben i optimeringen. Optimerade värdena för glukos (Figur 10a) och insulin (Figur 10b), tillsammans med datan visar att även här avviker modellen från den experimentiella datan.



Figur 10: a) Glukoskoncentrationen i plasman efter optimering i GMOL 2.0. b) Insulinkoncentrationen i plasman i GMOL 2.0. Den blå linjen visar modellen medan den rosa linjen visar datan för FF [3] där glukos tillsattes vid tidpunkten 0 minuter och insulin tillsattes efter 20 minuter. De streckade linjerna visar de övre och undre gränserna för vad som är tillåtna värden för modellen. Den nedre gränsen i b) ligger på den undre gränsen då koncentrationen av insulin överstiger den övre gränsen så mycket. Det vertikala streck som syns på insulin vid 20 minuter kommer från att vi sätter koncentrationen till värdet på datan för att simulera injektionen.

Variationskoefficienterna för de åtta parametrarna i modellen GMOL 2.0 (Tabell 4) visar att storleken av koefficienterna varierar stort mellan de olika parametrarna.

Tabell 4: Variationskoefficienterna för de åtta parametrarna i GMOL 2.0 modellen

k1	k2	k3	k4	k5	k6	k7	k8
$4.47 \cdot 10^{10}$	$5.02 \cdot 10^5$	$2.57 \cdot 10^{10}$	$1.05 \cdot 10^{17}$	$7.76 \cdot 10^{-1}$	$1.46 \cdot 10^{-1}$	$4.85 \cdot 10^{17}$	$4.80 \cdot 10^{11}$

6 Diskussion

Syftet med arbetet var att bättre förstå glukosmetabolisens dynamik. För detta utvecklades två ODE-modeller med olika komplexitetsgrad. Koncentrationerna av insulin och glukos i modellen matchades mot experimentiella data för att verifiera modellen. Den första modellen, GMOL (Figur 5) hade för hastiga förändringar jämfört med datan. Det har inte nödvändigtvis med modellens struktur att göra utan kan även bero på hur parametrarna togs fram. Jämfört med insulin var modellen något bättre för glukos då den minskar i takt med att datan gör det och sedan lägger sig på en nivå som inte är allt för olik datan. Modellen fångar dock inte helt dynamiken vilket kan tyda på att den är för simpel och saknar någon essentiell del eller att optimeringen inte är optimal för den här typen av modell.

För att fånga dynamiken bättre utvecklades modellen GMOL 2.0 (Figur 8) utifrån den första modellen. Här läggs två faktorer till, fettreserver och glykogen. Detta resulterade i att dynamiken för glukos fångas relativt bra, då modellen har samma nedåtgående kurva som datan under de första 100 minuterna för att sedan plana ut. Att det är så stor skillnad mellan hur de olika modellerna följer datan för glukos tyder på att delarna fettreserver och glykogen är viktiga för att beskriva glukosmetabolismen. Insulin däremot förbättrades inte utan är lik den första modellen. Detta kan bero på att insulins koncentrationer är mycket mindre än glukos och påverkar därför inte kostnadsfunktionen på samma sätt.

Graferna för både GMOL och GMOL 2.0 före optimering (Figur 6, 9) visar en bättre match mellan data och modell jämfört med efter optimeringen (Figur 7, 10). Det påvisar att Powells metod som används i optimeringen inte fungerar. Även om vi skapar ett intervall för varje parameter utifrån den logaritmiska likformiga fördelningen så hittar optimeraren endast sämre lösningar. Powells metod är en adaptiv gradientfri metod vilket gör den effektiv då beräkning av gradienten är resurskrävande [23]. Att den är adaptiv i sin steglängd är bra då vi ser att intervallen för GMOL 2.0 är i olika storleksordningar, något som också är vanligt inom dynamiska biologiska system [20]. Trots detta tog optimeraren endast ett steg från startgissningen genererat av hyperkuben och optimeraren tog uteslutande steg utanför hyperkuben. Det steget som optimeraren tog kan ha gett ett mindre värde från startgissningen men då den inte stegar mer hittar den inte ett bättre minimum. Antagligen är ytan för lösningarna till systemet fullt med lokala minimum och därför har optimeraren svårt att hitta den bästa lösningen. Även om det är resurskrävande att beräkna gradienten kan det vara en bättre lösning att använda sig av en adaptiv optimerare som använder sig av gradienten till systemet.

Ett sätt att hjälpa Powells metod hade varit att konvertera parametrarna till en logskala. Då hade optimeraren lättare hittat passande steglängder och lättare hittat ett minimum. Detta då parametrarna för ett biologiskt system ofta har olika storleksordningar och skulle intervallet för parametrarna logaritmeras minskar man intervallet. En adaptiv lösare har ofta problem att variera sin steglängd om det är för stora skillnader för varje parameter. För att få fram parametrarnas sanna värde är det bara att konvertera tillbaka efter optimeringen.

Resultatet från identifierbarheten visade höga värden på samtliga parametrar. Det pekar på att osäkerheten är hög och att det dataset som används inte innehåller tillräckligt av den informationen som behövs av modellen för att kunna utföra optimeringen. Att datan som används inte är identifierbar kan också vara förklaringen till varför optimeringen inte fungerar så bra.

Ekvation 6 beskriver, i det här fallet, hur mycket datan och modellen för glukos och insulin skiljer sig åt. Då glukos och insulin har en skillnad i respektive datan i storleksordning 10^7 bidrar de olika mycket till värdet på kostnadsfunktionen. Låt säga att en uppsättning parametrar ger att modellen ligger en tiopotens över datan för både insulin och glukos. Det leder till att för varje term skulle glukos bidra 10^{14} gånger mer än insulin till kostnadsfunktionen. Då optimeringen syftar på att minimera kostnadsfunktionen kommer fokuset för optimeringen bli olika för glukos och insulin. Följden av detta är att optimeringen letar efter att minimera avståndet för glukos men inte för insulin (Figur 10). Genom att logaritmera kostnadsfunktionen skulle storleksordningarna på differensen mellan datan och modellen bidra lika mycket till kostnadsfunktionen och därmed skulle bådatermerna bli optimerade.

Ett problem med kodningen av modellen var att Python processerade optimeringen långsamt och det hade därför kunnat vara fördelaktigt att använda ett snabbare programmeringsspråk.

Under projektet gjordes ett försök att reproducera en helkroppsmodell gjord av Uluseker *et al.* [2]. Ett problem med deras helkroppsmodell är att den inte är verifierad mot experimentell data utan istället uppbyggd på intervaller för olika koncentrationer som är baserade på tidigare forskning [2]. En endast intervall-baserad modell har visat sig vara bra på att fånga vissa beteenden hos biologiska system men har låg sannolikhet att simulera verkligheten [24]. En anledning till att modeller inte alltid verifieras mot data är att det kan vara mycket svårt och ibland omöjligt att göra fysiologiska tester på vissa biologiska reaktioner, i det här fallet glukosmetabolismen. Att komma runt det problemet är generellt en stor utmaning inom systembiologin [25].

För att verifiera modellen hämtades data från en studie på hästar under löpning, det fanns dock en del brister i den datan. I experimentet injicerades insulin i hästarna under tiden som de sprang. Injektionen gjorde att datan vek av från hur glukosmetabolismen beter sig naturligt vid löpning. Det blev även svårare att lösa modellen på korrekt sätt då det fanns en pik mitt i datan för insulin att ta hänsyn till.

Sammanfattningsvis var syftet med arbetet att fånga dynamiken hos glukosmetabolism under träning. Modellen som skulle göra det här lyckades fånga en del av dynamiken för glukos men sämre för insulin. Resultatet blev bättre när parametrarna skattades med genererade tal jämfört med användningen av optimeraren. Åtgärder för att få en bättre modell hade framförallt varit att ändra optimeraren till en gradientbaserad lösare för att undvika att fastna vid lokala minimum. Andra åtgärder är att pröva en logaritmisk skala på parametrarna och eventuellt testa ett snabbare programmeringsspråk. Modellen hade även kunnat förbättrats med fler iterationer för att lägga till eller ta bort ekvationer och termer i ekvationerna och undersöka hur de påverkar resultatet.

Om modellen hade uppfyllt syftet att fånga dynamiken hos systemet, hade nästa steg varit att hitta data inom andra träningsformer för att undersöka om modellen kan fånga dynamiken under olika träningsintensiteter. Ideellt är datan tagen från studier på människor men om den fortfarande är baserad på hästar som modellorganism behöver modellen konverteras i efterhand för att fungera på människor. En bra modell kan sedan användas till att forska på hälsoeffekter vid olika nivåer av träning.

Referenser

- [1] J. S. McPhee, D. P. French, D. Jackson, J. Nazroo, N. Pendleton och H. Degens, "Physical activity in older age: perspectives for healthy ageing and frailty", *Biogerontology*, årg. 17, nr 3, s. 567–580, 2016, ISSN: 15736768. DOI: 10.1007/s10522-016-9641-0.
- [2] C. Uluseker, G. Simoni, L. Marchetti, M. Dauriz, A. Matone och C. Priami, "A closed-loop multi-level model of glucose homeostasis", *PLOS ONE*, årg. 13, nr 2, W. Nishimura, utg., e0190627, febr. 2018, ISSN: 1932-6203. DOI: 10.1371/journal.pone.0190627. URL: <https://dx.plos.org/10.1371/journal.pone.0190627>.
- [3] K. H. TREIBER, T. M. HESS, D. S. KRONFELD, R. C. BOSTON, R. J. GEOR, M. FRIERE, A. M. G. B. SILVA och P. A. HARRIS, "Glucose dynamics during exercise: dietary energy sources affect minimal model parameters in trained Arabian geldings during endurance exercise", *Equine Veterinary Journal*, årg. 38, nr S36, s. 631–636, aug. 2006, ISSN: 04251644. DOI: 10.1111/j.2042-3306.2006.tb05617.x. URL: <http://doi.wiley.com/10.1111/j.2042-3306.2006.tb05617.x>.
- [4] Folkhälsomyndigheten, *Övervikt och fetma*, 2021. URL: <https://www.folkhalsomyndigheten.se/livsvillkor-levnadsvanor/fysisk-aktivitet-och-matvanor/overvikt-och-fetma/>.
- [5] —, *Hälsosamt åldrande Seniorguiden WHO : s definition av hälsosamt åldrande*, 2021. URL: <https://www.folkhalsomyndigheten.se/livsvillkor-levnadsvanor/halsa-i-olika-grupper/halsosamt-aldrande/>.
- [6] B. A. Franklin, C. P. D. Thompson, S. S. Al-Zaiti, C. M. Albert, M. F. Hivert, B. D. Levine, F. Lobelo, K. Madan, A. Z. Sharrief och T. M. H. Eijsvogels, *Exercise-related acute cardiovascular events and potential deleterious adaptations following long-term exercise training: Placing the risks into perspective-an update: A scientific statement from the American Heart Association*. 2020, E705–E736, ISBN: 0000000000000. DOI: 10.1161/CIR.0000000000000749.
- [7] Y. Li, R. K. Dash, J. Kim, G. M. Saidel och M. E. Cabrera, "Role of NADH/NAD⁺ transport activity and glycogen store on skeletal muscle energy metabolism during exercise: In silico studies", *American Journal of Physiology - Cell Physiology*, årg. 296, nr 1, 2009, ISSN: 03636143. DOI: 10.1152/ajpcell.00094.2008.
- [8] B. Alberts, A. Johnson, J. Lewis, D. Morgan, M. Raff, K. Roberts och P. Walter, *Molecular Biology of the Cell*, 6. utg. Garland Science, 2014, s. 1–1465, ISBN: 0815344325,9780815344322. URL: <http://gen.lib.rus.ec/book/index.php?md5=c33b3dc8f79ba%5C%5C72066347b43a3cb0114>.
- [9] N. Lai, L. B. Gladden, P. G. Carlier och M. E. Cabrera, "Models of muscle contraction and energetics.", eng, *Drug discovery today. Disease models*, årg. 5, nr 4, s. 273–288, 2008, ISSN: 1740-6757 (Print). DOI: 10.1016/j.ddmod.2009.07.001.
- [10] A. Aderem, "Systems biology: Its practice and challenges", *Cell*, årg. 121, nr 4, s. 511–513, 2005, ISSN: 00928674. DOI: 10.1016/j.cell.2005.04.020. URL: <http://dx.doi.org/10.1016/j.cell.2005.04.020>.
- [11] D. Sundström, "On a bioenergetic four-compartment model for human exercise", *Sports Engineering*, årg. 19, nr 4, s. 251–263, 2016, ISSN: 14602687. DOI: 10.1007/s12283-016-0205-y.
- [12] G. Dimitriadis, P. Mitron, V. Lambadiari, E. Maratou och S. A. Raptis, "Insulin effects in muscle and adipose tissue", *Diabetes Research and Clinical Practice*, årg. 93, nr SUPPL. 1, s. 52–59, 2011, ISSN: 01688227. DOI: 10.1016/S0168-8227(11)70014-6.
- [13] S. L. Berg JM Tymoczko JL, "Glycogen Metabolism", i *Biochemistry*, 5th, New York: W H Freeman, 2002, kap. Chapter 21. URL: <https://www.ncbi.nlm.nih.gov/books/NBK21190/>.
- [14] J. Gabrielsson och D. Weiner, *Pharmacokinetic and pharmacodynamic data analysis: concepts and applications*. CRC press, 2001, ISBN: 9186274929.
- [15] H. Kitano, "Systems biology: A brief overview", *Science*, årg. 295, nr 5560, s. 1662–1664, 2002, ISSN: 00368075. DOI: 10.1126/science.1069492.
- [16] C. R. Clark, "Modeling and Analysis of", nr Eccm V, s. 2014, 2014.
- [17] E. Klipp, L. Wolfgram, W. Christoph och K. Axel, *Systems Biology*, Second. Weinheim, Germany: Wiley-VCH, 2016.
- [18] F. Rothlauf, "Optimization Methods", i *Design of Modern Heuristics: Principles and Application*, Berlin, Heidelberg: Springer Berlin Heidelberg, 2011, s. 45–102, ISBN: 978-3-540-72962-4.

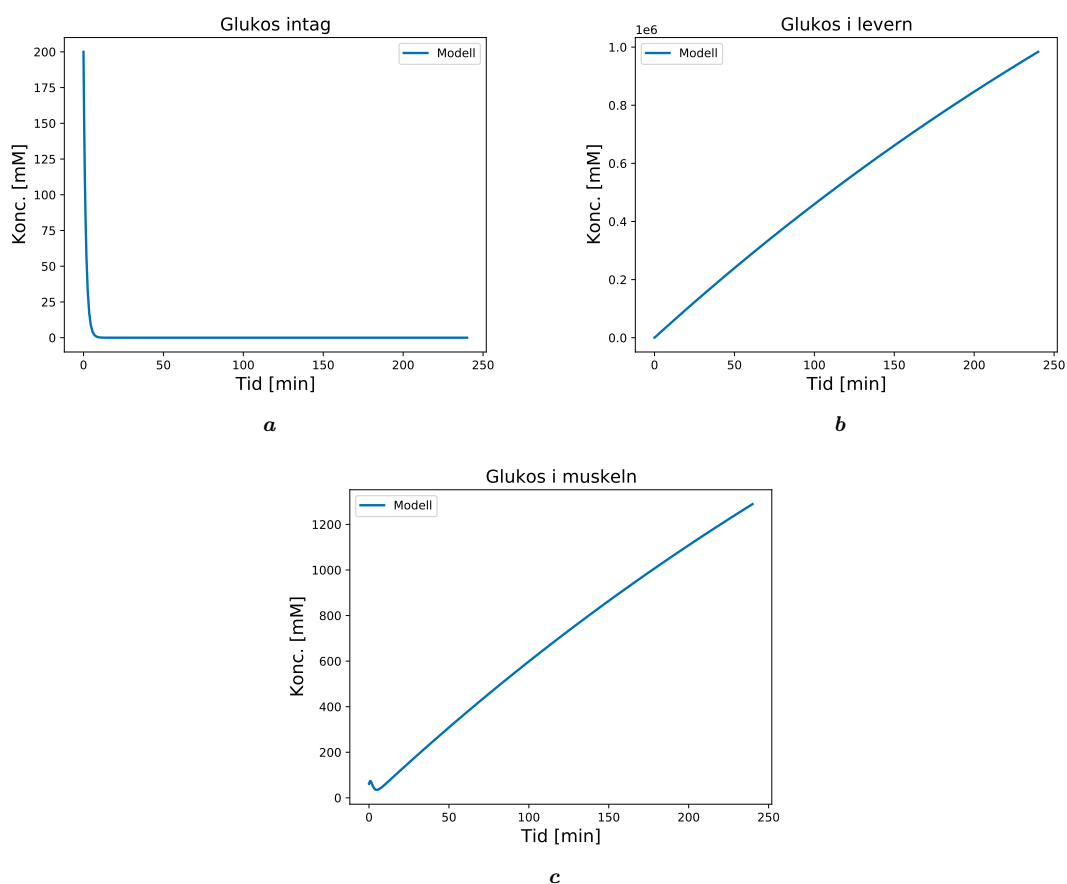
DOI: 10.1007/978-3-540-72962-4_{_}3. URL: https://doi.org/10.1007/978-3-540-72962-4_3.

- [19] V. S. Vassiliadis och R. Conejeros, *Powell method* *Powell Method*, C. A. Floudas och P. M. Pardalos, utg., Boston, MA, 2009. DOI: 10.1007/978-0-387-74759-0_{_}516. URL: https://doi.org/10.1007/978-0-387-74759-0_516.
- [20] A. Raue, M. Schilling, J. Bachmann, A. Matteson, M. Schelke, D. Kaschek, S. Hug, C. Kreutz, B. D. Harms, F. J. Theis, U. Klingmüller och J. Timmer, “Lessons Learned from Quantitative Dynamical Modeling in Systems Biology”, *PLoS ONE*, årg. 8, nr 9, 2013, ISSN: 19326203. DOI: 10.1371/journal.pone.0074335.
- [21] J. Distefano III, *DYNAMIC SYSTEMS BIOLOGY MODELING AND SIMULATION*, First edit. London, Waltham, San Diego: Elsevier, 2014, s. 490–557.
- [22] E. Hairer, S. Norsett och G. Wanner, *Solving Ordinary Differential Equations I*. 1993, vol. 8, ISBN: 978-3-540-56670-0. DOI: 10.1007/978-3-540-78862-1.
- [23] A. Vazeou, “Continuous blood glucose monitoring in diabetes treatment”, *Diabetes Research and Clinical Practice*, årg. 93, nr SUPPL. 1, S125–S130, aug. 2011, ISSN: 01688227. DOI: 10.1016/S0168-8227(11)70028-6.
- [24] D. A. Beard och H. Qian, “Constraint-based modeling of metabolomic systems”, *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*, årg. 3, 2005. DOI: 10.1002/047001153x.g308203.
- [25] A. Finkelstein, J. Hetherington, L. Li, O. Margoninski, P. Saffrey, R. Seymour och A. Warner, “Computational Challenges of Systems Biology”, *Computer*, årg. 37, nr 5, s. 26–33, 2004, ISSN: 00189162. DOI: 10.1109/MC.2004.1297236.

A Kompletterande resultat till GMOL- och GMOL-modellerna 2.0

Här presenteras resterande resultat från modellerna GMOL och GMOL2.0. Resultaten är plottar för resternade koncentrationer och optimala parametervärden för de båda modellerna.

A.1 Modell GMOL

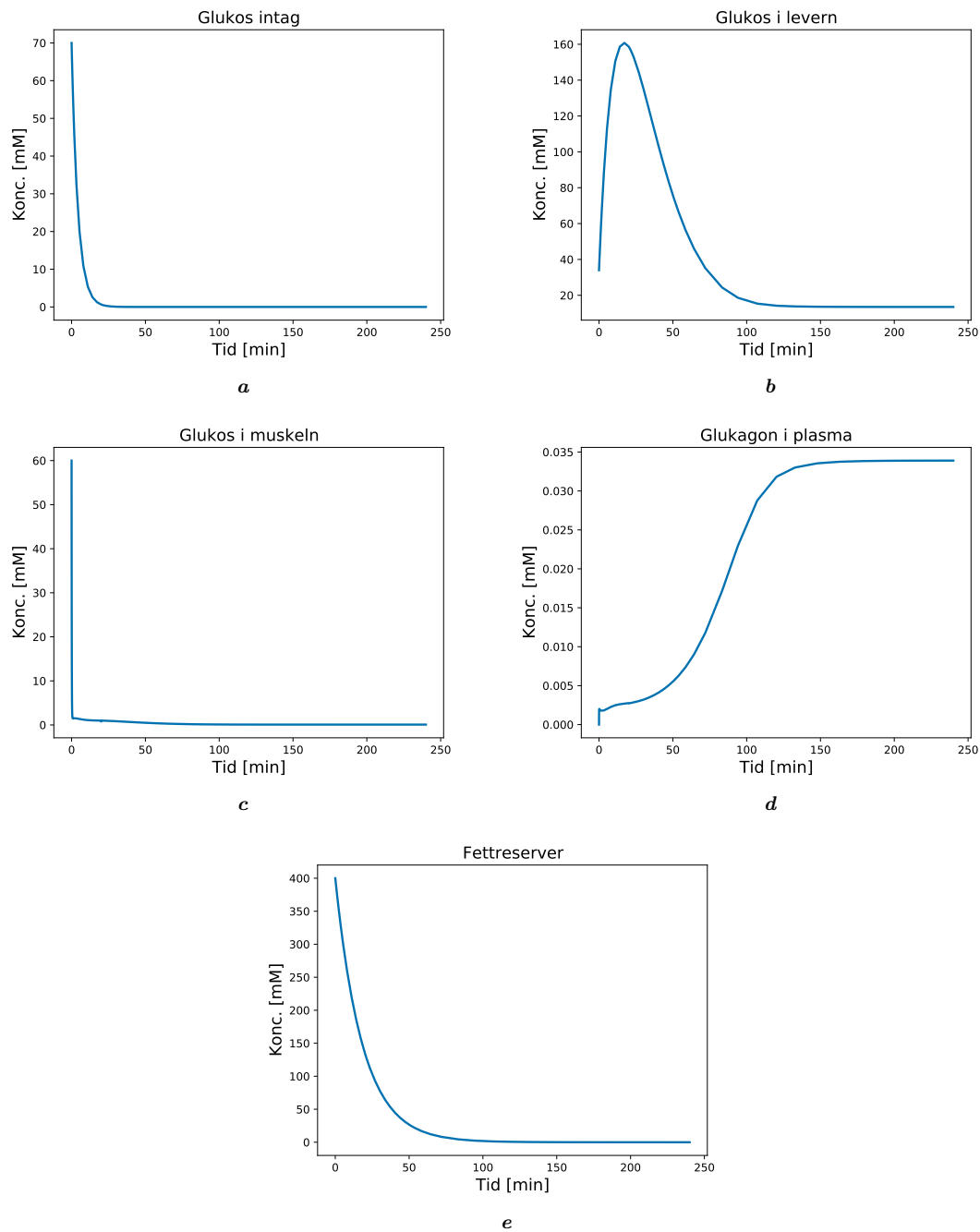


Figur A.1: a) Koncentrationen av glukos in i model GMOL och hur den varierar över tid. b) Koncentrationen av glukos i levern i GMOL och hur den varierar över tid i modellen. c) Koncentrationen av glukos i musklerna i GMOL och hur den varierar över tid.

Tabell A.1: Optimala parametrar för GMOL-modellen. Parametrarna togs fram med Powells metod givet 10 startvärden genererade från Latin Hypercube Samplingen.

k1 [min^{-1}]	k2 [$\text{mM}^{-1} \cdot \text{min}^{-1}$]	k3 [min^{-1}]	k4 [$\text{mM}^{-1} \cdot \text{min}^{-1}$]	k5 [min^{-1}]
$1.11 \cdot 10^0$	$6.28 \cdot 10^2$	$5.26 \cdot 10^2$	$2.21 \cdot 10^1$	$3.61 \cdot 10^1$

A.2 Modell GMOL2.0



Figur A.2: a) Koncentrationerna av glukos in i GMOL 2.0. b) Koncentrationen av glukos i levern i GMOL 2.0 och hur den varierar över tid. c) Koncentrationen av glukos i muskeln i GMOL 2.0 och hur den varierar över tid. d) Koncentrationen av hormonet glukagon i GMOL 2.0 och hur den varierar över tid. e) Koncentrationen av glukos ut ur fettreserverna i GMOL 2.0 och hur den varierar över tid.

Tabell A.2: Optimala parametrar för GMOL-modellen 2.0. Parametrarna togs fram med Powells metod givet 10 startvärden genererade från Latin Hypercube Samplingen.

k1 [min ⁻¹]	k2 [mM ⁻¹ · min ⁻¹]	k3 [min ⁻¹]	k4 [min ⁻¹]	k5 [min ⁻¹]	k6 [min ⁻¹]	k7 [min ⁻¹]	k8 [mM ⁻¹ · min ⁻¹]
2.33·10 ⁻¹	1.01·10 ⁰	3.92·10 ⁻¹	5.42·10 ⁻²	9.17·10 ⁰	2.16·10 ¹	5.42·10 ⁻²	6.39·10 ⁻²

B GitHub

Länk till GitHub finns [**här**](#) där koden för alla modeller finns samt datan som används i arbetet.