



CHALMERS
UNIVERSITY OF TECHNOLOGY

Matematisk modellering av hjärntumörer
Dataanalys och utvidgning av modell av glioblastom
Kandidatarbete inom civilingenjörsutbildningen vid Chalmers

Sebastian Berg
William Danielsson
Erik Nilsson
William Orton Sörensen

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborg 2017

Matematisk modellering av hjärntumörer

Dataanalys och utvidgning av modell av glioblastom

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk matematik vid Chalmers

Sebastian Berg

William Danielsson

Erik Nilsson

William Orton Sörensen

Handledare: Philip Gerlee Matematiska Vetenskaper
Examinator: Maria Roginskaya
 Marina Axelson-Fisk

Institutionen för matematiska vetenskaper
Chalmers tekniska högskola
Göteborg 2017

Glioblastom - hjärncancern matematik vill besegra

Genom en upptäckt angående cancer glioblastoms karaktär togs en ny matematisk modell fram år 2012 för att beskriva dess cancer cellutveckling. Ett första försök att matcha modellen med verklig patientdata har nu gjorts med lovande resultat. Enligt modellen har det visat sig att cancer cellerna oftast får en minskad rörlighet som följd av behandling medan cellernas delningsförmåga verkar vara oförändrad. Detta motiverar till djupare och bredare forskning om denna modell.

Glioblastom dödar omkring 80 000 människor per år och patienter har en dyster prognos på cirka 12 månaders överlevnad efter diagnos. Cancern är väldigt elakartad och sprider sig så pass snabbt i hjärnvävnaden att det bästa resultatet patienten kan förvänta sig av både cellgift- och strålningsbehandling är en för tillfället stagnerad tillväxt av cancer. Otäckt nog är det just denna snabba tillväxt hos cancer som tillåter en effektiv matematisk modell.

Biologer har upptäckt att en given glioblastom cell alltid befinner sig i en av två olika faser. En fas där cellen enbart kan dela sig och en annan fas där den enbart kan röra på sig. Det unika med den nya modellen är att den tar hänsyn till cellernas fasbyten. För att koppla samman modellen med verkligheten krävs data från drabbade patienter. Det behövs också matematiska metoder för att ta fram de egenskaper i modellen som kan mätas med hjälp av MR-röntgen. Dessa metoder har nu tagits fram och tillämpats på 20 patienter med glioblastom.

Det är svårt att dra generella slutsatser om modellens riktighet eftersom datan innehåller väldigt få patienter vars behandling har varierat från fall till fall. Det huvudsakliga bidraget till forskningen kring glioblastom är de metoder som används för att koppla modellen till verkligheten. Det krävs mer omfattande arbete och en mycket större mängd patienter för att dra slutsatser kring hur väl modellen beskriver verkligheten. Däremot tror vi att vi har byggt en bra grund för framtida upptäckter inom ämnet.

Sammandrag

Hjärncancern glioblastom är den mest svårstoppade och maligna av alla typer av hjärncancer. Cancern tillåter sig dock modelleras matematiskt på grund av denna aggressiva karaktär. En vanligt förekommande modell för cancers utveckling baseras på Fisher-ekvationen $\partial_t u = D\nabla^2 u + \rho u(1 - u)$. Matematiker Philip Gerlee och cancerbiolog Sven Nelander postulerade en ny modell för cancerutvecklingen av glioblastom år 2012, vilken vidare analyserades år 2016. Modellen bygger på att ta hänsyn till ett experimentellt verifierat bytesmönster (växelverkan mellan rörande och förökande stadie) hos cancerceller, vilket leder till ett system av kopplade partiella differentialekvationer. I vår rapport lägger vi till en linjär dödsterm i modellen för att ta hänsyn till effekten av en extern cancerbehandling, som endast påverkar cellerna som delar på sig. Det resulterande systemets lösningar simuleras numeriskt, och för respektive lösning beräknas en våghastighet och en lutning med numeriska metoder. Under vissa antaganden på modellen tillämpas dels perturbationsteknik för att erhålla en analytisk lösning och dels fasrumsanalys för att erhålla en analytisk våghastighet. Under nya antaganden anpassas rådata av patienters cancerutveckling till modellparametrar gällande tillväxthastighet och diffusionskonstant av cancerceller. Anpassningen visas väldigt bra och indikerar främst att parametrarna för 60% av alla patienter minskar eller hålls konstant över behandlingsfaserna.

Abstract

Glioblastoma may because of its aggressive behaviour be modelled mathematically. A frequently used model for simulation of cancer growth is based on the Fisher-equation $\partial_t u = D\nabla^2 u + \rho u(1 - u)$. One such model for specific glioblastoma growth based on phenotypic switching was postulated 2012, and later analysed 2016 by mathematician Philip Gerlee and cancer biologist Sven Nelander. Via the addition of a linear term corresponding to apoptosis of the cancer due to treatment on the proliferating cells, said model is adjusted to produce slightly different behaviour characterised by its underlying parameters. The solution of this new system is analysed numerically and methods to calculate the solution's corresponding propagation speed and slope has been found. Furthermore a perturbation approach is applied to find a closed form analytic approximation of the solution, while phase space analysis allows for the discovery of a limiting wave speed. We bring these results to scrutiny under the lens of a set of real world data and find that an optimal set of parameters provides for a good fit between model and reality. The fit foremost indicates that the parameters for over 60% of the total number of patients decreases or are kept constant over the treatment phases.

Förord

Författarna vill tacka sin handledare Philip Gerlee för hans tid, tillgänglighet och hjälpsamma vägledning.

Genomförande

De första veckorna gick åt till att komma underfund med den underliggande teorin samt till att skriva projektplan. Samtliga medlemmar inhämtade och studerade relevanta artiklar för arbetet, däri inkluderat tidigare arbeten av Philip Gerlee m. fl. samt andra artiklar om matematisk modellering inom biologi.

Alla medlemmar har bidragit till skapande av metoder under hela projektets gång. Till en början fokuserade W.O.S., W.D. och S.B. på numeriska metoder medan E.N. lade mer tid på de analytiska uttrycken. W.D. och W.O.S. fokuserade även på analytiska metoder när numeriska metoder började bli klart. Alla medlemmar bidrog med kreativa idéer under projektets gång gällande metoder, problemlösning och analys.

Efter mitten av mars var de numeriska och analytiska resultaten för våghastighet och lutning färdiga. Fokus skiftades då till att analysera den kliniska datan genom att upprätta effektiva metoder för att lösa ut parametervärden. E.N. påbörjade rapportskrivandet medan W.O.S., W.D. och S.B. började arbeta på metoder för dataanalysen. Arbetet att framställa resultat har samtliga medlemmar bidragit till. När det kommer till diskussion och slutsats har alla deltagare varit med och bidragit till värdefulla insikter.

Viktigt att poängtera är att ovanstående beskrivning är en väldigt enkel redovisning av arbetsfördelningen. Vi vill betona att alla projektdeltagare har hjälpt varandra under projektets gång. Därför är det svårt att säga exakt vem som är ansvarig för projektets olika delmoment. Speciellt är det viktigt att nämna att även fast vissa delar av rapporten huvudsakligen skrivs av enskilda individer har alla bidragit till material och innehållet i texterna, genom att t.ex. ha gett åsikter och råd på förbättringar.

För en mer ingående redogörelse av vad respektive medlem har bidragit med hänvisas läsaren till loggboken där varje enskild individs prestation har antecknats.

Huvudförfattare per sektion

William Orton Sörensen - Numeriska metoder, populärvetenskaplig text, Förord, Analytiska resultat, Diskussion och slutsatser, Felkällor.

Erik Nilsson - Analytiska härledning, Analytiska resultat, Bakgrund, Syfte, Sammandrag/Abstract, Redogörelse av Gerlees kopplade system, Bilaga A, Bilaga B, Diskussion och slutsatser, Populärvetenskaplig text.

William Danielsson - Metoder för anpassning av modell till rådata, Resultat, Diskussion och slutsatser, Sammandrag, Förord, Syfte, Bilaga D, Bilaga E.

Sebastian Berg - Bilaga C, Diskussion och slutsatser.

Samtliga deltagare har granskning av övriga deltagares texter och genomgående bidragit med värdefulla kommentarer.

Innehåll

1	Inledning	1
1.1	Bakgrund	1
1.2	Syfte	2
2	Redogörelse av Gerlees kopplade system	2
3	Metoder	3
3.1	Numeriska metoder	3
3.1.1	Våghastighet med medianmetoden	4
3.1.2	Lutning med histogrammetoden	6
3.2	Analytiska härledningar	7
3.2.1	Våghastighet via analys av stationära punkter	7
3.2.2	Lutning genom perturbationslösning	10
3.3	Metoder för anpassning av modell till rådata	13
3.3.1	Minimeringsproblem baserad på numeriska resultat	13
3.3.2	Minimeringsproblem baserad på analytiska resultat	14
4	Resultat	15
4.1	Analytiska uttryck	16
4.2	Jämförelse av \hat{a} och D_v före och efter behandling	16
5	Diskussion och slutsatser	17
5.1	Analytisk och numerisk dataanpassning	17
5.2	Kopplingar mellan diverse patientdata och resultatet	18
5.3	Felkällor och avgränsningar	20
5.4	Avslutande kommentarer	21
Bilagor		22
A	Beskrivning & förenkling av ekvationer	22
A.1	Fisher's ekvation	22
A.2	Gerlee och Nelander's ekvationer	22
B	Motivation för gräns- och randvärden	23
C	Beskrivning av rådata	25
D	Analytiska resultat i jämförelse med numeriska	25
E	Jämförelse av numeriskt kontra analytiskt minimeringsproblem	26
E.1	Jämförelse av optimala parametrar för dataanalysmetoderna	27
E.2	Datorintensivitet och effektivitet av minimiseringsproblemen	28
E.3	Optimala parametrar före behandling samt explicit målfunktion	29
F	Kod: Matlab och Mathematica	30

1 Inledning

I detta kapitel introduceras läsaren till de centrala koncept som är viktiga att förstå för att följa med i resten av rapporten. Vi beskriver syftet med arbetet och lägger en grund för analysen.

1.1 Bakgrund

Glioblastom multiforme är en form av hjärncancer som är mycket elakartad. Bland de olika formerna av hjärncancer som finns är den klassad som den mest maligna [2].

Det har gjorts försök till att modellera glioblastom via den s.k. Fisher-ekvationen, där $u(x, t)$ är en funktion som representerar cancers normerade densitet, alltså andelen cancerceller per volymsenhet, vid tiden t och positionen x från cancers centrum,

$$\frac{\partial}{\partial t}u(x, t) = D\nabla^2u(x, t) + \rho u(x, t)(1 - u(x, t)). \quad (1)$$

Modellen har visat sig vara användbar i prognosen av patientens tumörutveckling [3] [4] [5]. I kopplingen mellan modell och verklighet brukar ekvationen beskrivas radiellt istället för (3+1) dimensionellt. Detta är motiverat eftersom de olika representationerna har samma asymptotiska beteende. Vi gör samma förenkling i denna rapport. Ekvationen har tillhörande *färdande våglösningar*, $u(x, t) = V(x - ct)$, som ger upphov till en färdande vågfront vars konstanta hastighet är av speciellt intresse. Den ges av $c = 2\sqrt{D\rho}$ för rimliga initialvillkor. Parametrarna D och ρ , som återfinns i (1) ovan, kan bestämmas utifrån uttryck för cancers utbredningshastighet och lutning [9] som i sin tur kan erhållas från MR-röntgen (magnetisk resonanstomografi).

Fisher-ekvationen är emellertid gammal sett till de nya upptäckter som gjorts angående glioblastomens egenskaper. Det har bekräftats experimentellt att canceren har en 'go-or-grow' karaktär, som betyder att en cell aldrig kan både dela och röra på sig samtidigt. Denna 'go-or-grow' karaktär kallas även för ett *fenotypiskt bytesmönster*. År 2012 skrevs en artikel av Gerlee och Nelander där dessa egenskaper förenades med Fisher's ekvation i form av en ny matematisk modell av glioblastom [6]. Modellerna beskrivs i mer detalj under kapitel 2 samt bilaga A men Gerlees modell har sitt mest generella utseende som,

$$\begin{cases} \frac{\partial p}{\partial t} = D_\alpha(1 - p - m)\frac{\partial^2 p}{\partial x^2} + \alpha p(1 - p - m) - (q_m + \mu)p + q_p m \\ \frac{\partial m}{\partial t} = D_v((1 - p)\frac{\partial^2 m}{\partial x^2} + m\frac{\partial^2 p}{\partial x^2}) - (q_p + \mu)m + q_m p \end{cases}, \quad (2)$$

där $d = p + m$ motsvarar cancers totala normerade densitet.

Den primära rådatan vi analyserar kommer ifrån en amerikansk studie, som mer ingående beskrivs i bilaga C. Datan som bland annat innehåller parametrarna D och ρ togs fram under behandlingsprocessen av 20 patienter. Patienterna fick genomgå både kirurgiska ingrepp såsom strålning- och cellgiftsbehandling. Tiden som patienterna genomgick behandlingen varierar. Genom att mäta parametrarna D och ρ före och efter behandlingen kan man utvärdera huruvida den var gynnsam eller inte. Låga värden på D och ρ är bra för patienten sett till cancers utbredning, medan höga värden innebär motsatsen.

Förväntningen på utvärderingen var att patienternas D - och ρ -värden skulle minska. Så var dock inte fallet, utan istället noterades att tillväxtintensiteten ρ hade en tendens att minska men att diffusionskonstanten D ökade hos vissa patienter men minskade hos andra. Man funderade över om de oväntade resultaten kunde härstamma från en ändring i cancers fenotyp.

På så sätt frågar vi oss om de kopplade ekvationerna (2) skulle kunna beskriva cancers utveckling hos de 20 patienterna bättre.

1.2 Syfte

En matematisk modell som kan fånga glioblastom-tumörens alla egenskaper är mycket eftertraktad. Vårt syfte är att bidra till sökandet av en sådan modell genom att fortsätta analysen som Gerlee m. fl. påbörjade år 2012. I denna föregående analys inkluderades ej effekten av en extern cancerbehandling. Genom att addera en linjär dödsterm till modellen vill vi utföra en utökad analys av modellen där effekten av behandling på glioblastom beaktas. Vi utvecklar analytiska och numeriska metoder för att relatera modellen till rådata om patienters cancerutveckling. I följd drar vi slutsatser angående effektiviteten av behandlingen.

2 Redogörelse av Gerlees kopplade system

För de mer ingående beskrivningarna av termer och ekvationer, samt hur man kommer fram till dem, hänvisas läsaren till bilaga A.

Det slutgiltiga uttryck som Gerlee m. fl. får sina analytiska resultat ifrån ser ut som följande, där α är tillväxthastigheten, D_v är diffusionskonstanten och q_m samt q_p är övergångskonstanter,

$$\begin{cases} \frac{\partial p}{\partial t} = \alpha p(1 - p - m) - q_m p + q_p m \\ \frac{\partial m}{\partial t} = D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p \end{cases} .$$

Summan $d(x, t) = p(x, t) + m(x, t)$ beskriver, på grund av det fenotypiska bytesmönstret, andelen cancer i en liten del av vävnaden. Vi kommer fortsättningsvis att kalla denna andel (som egentligen är en normerad densitet) för just *densitet*. De prolifererande (delande) cellerna beskrivs av p medan de diffunderande (migrerande) cellerna beskrivs av m . Modellen utgår ifrån en stokastisk modell där cancercellerna befinner sig på ett rutnät och migrerar från ruta till ruta genom en Markov-process. Ekvationerna i systemet ovan är förenklade, tidskontinuerliga varianter av sannolikheten att hitta en cell av respektive typ i en viss ruta.

Ekvationen ovan är en modell på densiteten av glioblastom utan behandling. Vi använder dödskonstanten μ för att beskriva apoptosintensiteten, det vill säga takten som celler dör på. Eftersom cellgifter samt strålningsbehandling angriper cancercellernas arvs massa (och därigenom deras celledning) är det rimligt att antaga att deras påverkan på de migrerande cellerna är försumbar. Vi inleder således vår uppgift med följande system som utgångspunkt, som vi kallar det *analytiska utgångssystemet*,

$$\begin{cases} \frac{\partial p}{\partial t} = \alpha p(1 - p - m) - q_m p + q_p m - \mu p \\ \frac{\partial m}{\partial t} = D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p \end{cases} . \quad (3)$$

Vi gör nu en kortfattad beskrivning och enhetsanalys av parametrarna till (3) i tabell 1. För detaljer hänvisas läsaren till artikeln [6]. Enheterna cellcykel och cellbredd har typiska värden på 24 timmar respektive $20 \mu\text{m}$.

Tabell 1: Systemets parametrar, deras enhet och innebörd.

Parameter	Enhet	Beskrivning
α	$[\frac{1}{(\text{cellcykel})}]$	Tillväxthastigheten för cancern. Infinner sig alltså bara i ekvationen för p , det vill säga i ekvationen för de stilla prolifererande cellerna. Den senare i kapitel 3.2.1 definierade parametern $\hat{\alpha} = \alpha - \mu$ kallar vi för proliferationstakten.
μ	$[\frac{1}{(\text{cellcykel})}]$	Apoptoskonstanten svarar för cancerns död på grund av behandling. Konstanterna som hör till cancerns tillväxt, nämligen α och μ , förhåller sig till varandra så att $0 \leq \mu \leq \alpha$. Detta härleds i bilaga B.
D_v	$[\frac{(\text{cellbredd})^2}{(\text{cellcykel})}]$	Diffusionskonstanten motsvarar cellens motilitet (rörlighet). Högre värde på D_v innebär mer sporadisk diffusion.
$q_{p,m}$	$[\frac{1}{(\text{cellcykel})}]$	Övergångskonstanterna representerar cellövergång mellan de två fenotyperna. Det kan argumenteras utifrån den stokastiska modellen[6] att $q_{p,m}$ har en övre begränsning så att $q_{p,m} < 24 [\frac{1}{(\text{cellcykel})}]$.

Vi kommer fortsättningsvis antaga vissa standardvärden på konstanterna för att ge en enhetlig bild av resultaten. Vi har testat olika varianter och våra resultat gäller oavsett. Standardvärdena är följande: $\alpha = 1$, $\mu = 0.1$, $D_v = 25$, $q_{m,p} = 20$.

3 Metoder

Analys av modellen utgår från ett numeriskt och ett analytiskt perspektiv, varav dessa sedan kopplas till rådata via en avslutande dataanalys. Med numeriska metoder beräknas våghastigheter och lutningar av propagerande lösningar, vilket beskrivs mer detaljerat i kapitel 3.1. Analytiska metoder går att tillämpa under vissa antaganden på modellen. Dels fasrumsanalys för att erhålla våghastigheter av lösningar och dels perturbationsteknik för att erhålla analytiska uttryck på lösningar. Detta utförs i kapitel 3.2. Slutligen härleds två numeriska minimeringsproblem som möjliggör en koppling mellan modell och rådata, vilka kan användas för dataanalys av patienters cancerutveckling före och efter behandling.

3.1 Numeriska metoder

Betrakta det analytiska utgångssystemet i domänen $x \in [0, x_M]$ och $t \in R^+$,

$$\begin{cases} \frac{\partial p}{\partial t} &= \alpha p(1 - p - m) - (q_m + \mu)p + q_p m \\ \frac{\partial m}{\partial t} &= D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p \end{cases} \quad (3)$$

Där x_M är det maximala avståndet som cancercellerna kan sprida sig till. Följande begynnelse- och randvillkor använd för de numeriska metoderna

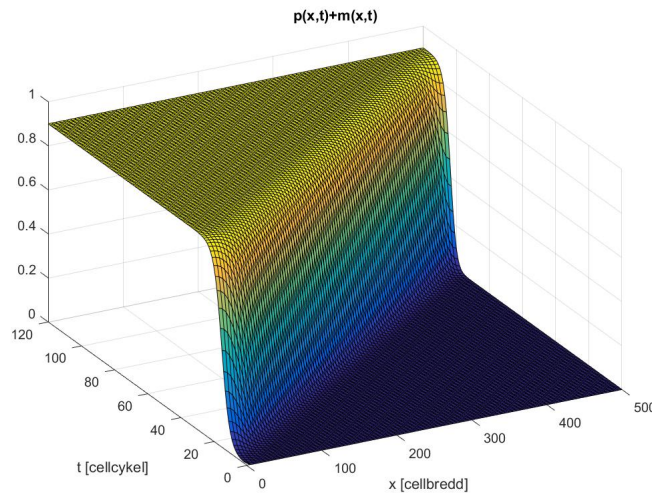
$$\begin{cases} p(x, 0) &= 0.01e^{-10x} \\ m(x, 0) &= 0 \\ \frac{\partial p}{\partial x}(0, t) &= \frac{\partial m}{\partial x}(0, t) = \frac{\partial p}{\partial x}(x_M, t) = \frac{\partial m}{\partial x}(x_M, t) = 0 \end{cases}.$$

Begynnelsevillkoren säger att det till en början enbart existerar en liten mängd cancerceller i proliferationsfasen. Cancerceller kan inte sprida sig hur långt som helst. Rumsderivatan för p och m väljs därför till 0 i ändpunktern. Dessa randvillkor och begynnelsevillkor har ingen väsentlig betydelse för de numeriska metoderna. Vi är inte intresserade av vad som händer i

ändpunkterna av den observerade mängden utan hur $p(x, t)$ och $m(x, t)$ beter sig däremellan. En numerisk metod för att lösa (3) är finita differensmetoden (FDM). I detta fallet går FDM ut på att diskretisera rummet och tiden till ett rutnät av punkter varefter derivator skrivs om till finita differenser. I vårt fall betraktas följande diskretisering,

$$\Omega = [0 = x_1, x_2, \dots, x_M] \times [0 = t_1, t_2, \dots, t_N]. \quad (4)$$

Genom att lösa system (3) med FDM fås en matris $U \in R^{M \times N}$ som representerar cancers densitet $d(x, t) = p(x, t) + m(x, t)$ som funktion av avståndet x från cancers centrum och t från tiden 0. Radnumreringen i U svarar direkt mot de diskreta tidpunkterna $[0 = t_1, t_2, \dots, t_N]$ och motsvarande numrering för kolonner representerar de diskreta x-värdena $[0 = x_1, x_2, \dots, x_M]$. Element $U_{i,j}$ är alltså den numeriska approximationen av $d(x_i, t_j) = p(x_i, t_j) + m(x_i, t_j)$. Figuren nedan visar hur den numeriska lösningen till system (3) ser ut med en angiven uppsättning parametervärden. MATLABs funktion *pdepe* har använts för att beräkna U .

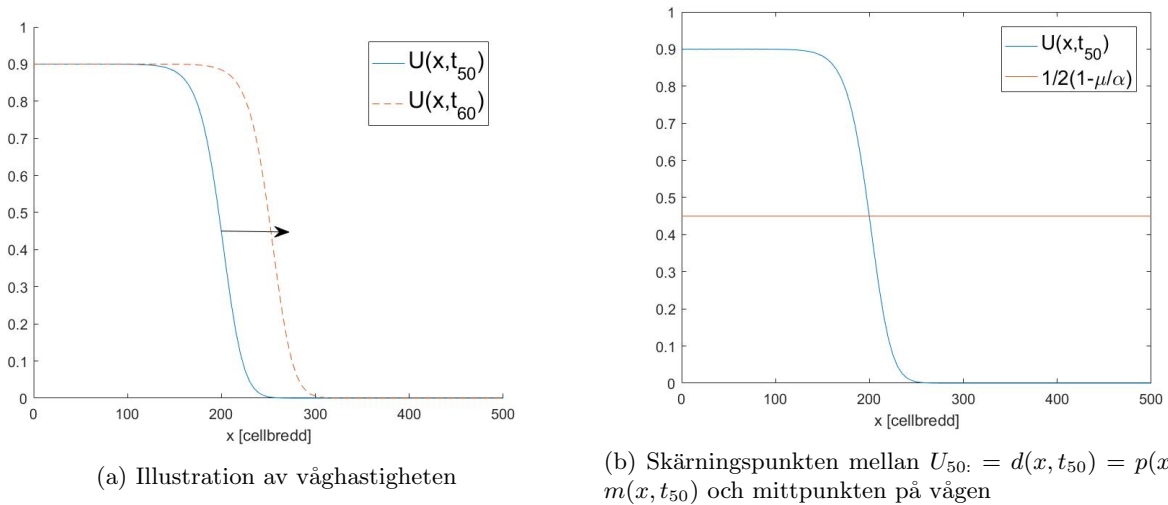


Figur 1: Matrisen U med hjälp av *pdepe* i MATLAB. Parametervärden $\alpha = 1$, $D_v = 25$, $q_m = 20$, $q_p = 20$, $\mu = 0.1$. Steglängder $h_x = 5$, $h_t = 1$

Viktigt att nämna här är att när man jämför de numeriska resultaten med de analytiska krävs en hög precision på de numeriska uträkningarna. Hög precision uppnås genom att ha en liten steglängd mellan de diskreta punkterna $[0 = x_1, x_2, \dots, x_M]$ och $[0 = t_1, t_2, \dots, t_N]$. Vi definierar steglängderna $h_x = x_{i+1} - x_i$, $i = 1, 2, \dots, M - 1$ och $h_t = t_{i+1} - t_i$, $i = 1, 2, \dots, N - 1$.

3.1.1 Våghastighet med medianmetoden

Vi definierar mittpunkten på den färdande vågen som halva värdet av vågens maximala höjd vid en viss tid, dvs $\frac{1}{2} \max U(x, t_i)$. Våghastigheten beräknas då som mittpunktens hastighet i x-led. Vilken punkt man mäter spelar ingen roll eftersom hela vågfronten har samma hastighet. Figur 2a nedan visar en visualisering av våghastigheten.

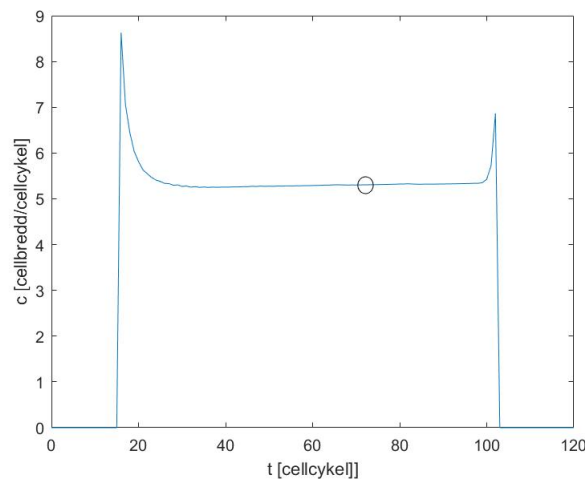


Figur 2: Parametervärdena $\alpha = 1$, $D_v = 25$, $q_m = 20$, $q_p = 20$, $\mu = 0.1$. Steglängder $h_x = 5$, $h_t = 1$

För att numeriskt beräkna våghastigheten utgår vi ifrån den tidigare nämnda lösningsmatrisen $U_{ij} = d(x_i, t_j) = p(x_i, t_j) + m(x_i, t_j)$. Vi låter U_i definiera alla kolonner på rad i . För varje tidpunkt t_1, \dots, t_{N-1} låter vi θ_i definiera skärningspunkten mellan mittpunkten och radvektorn U_i : (se figur 2b). Vi definierar c som en vektor med element

$$c_i = \frac{\theta_{i+1} - \theta_i}{t_{i+1} - t_i}, \quad (5)$$

som representerar våghastigheten vid tiden t_i . Eftersom vi har en färdande våglösning förväntar vi oss ett konstant värde på våghastigheten. Följande figur visar ett typiskt exempel på hur vektorn c ser ut.



Figur 3: Den numeriskt beräknade våghastigheten som funktion av tiden. Medianen markeras med en svart cirkel. Parametervärden $\alpha = 1$, $D_v = 25$, $q_m = 20$, $q_p = 20$, $\mu = 0.1$. Steglängder $h_x = 5$, $h_t = 1$

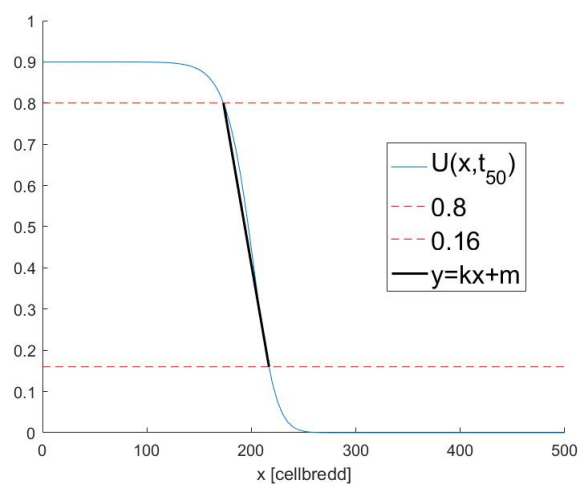
Genom att välja nollskilda element i vektorn c och ta medianen av dessa fås ett representativt värde för den konstanta våghastigheten. I exemplet ovan skulle alltså våghastigheten väljas till $c = 5.3055$.

3.1.2 Lutning med histogrammetoden

För att numeriskt beräkna lutningen utgår vi ifrån att en lösningsmatris $U_{ij} = p(x_i, t_j) + m(x_i, t_j)$ har beräknats enligt FDM. Lutningen s är definierad på följande vis för varje diskret tidpunkt (d.v.s. för varje rad i matrisen U).

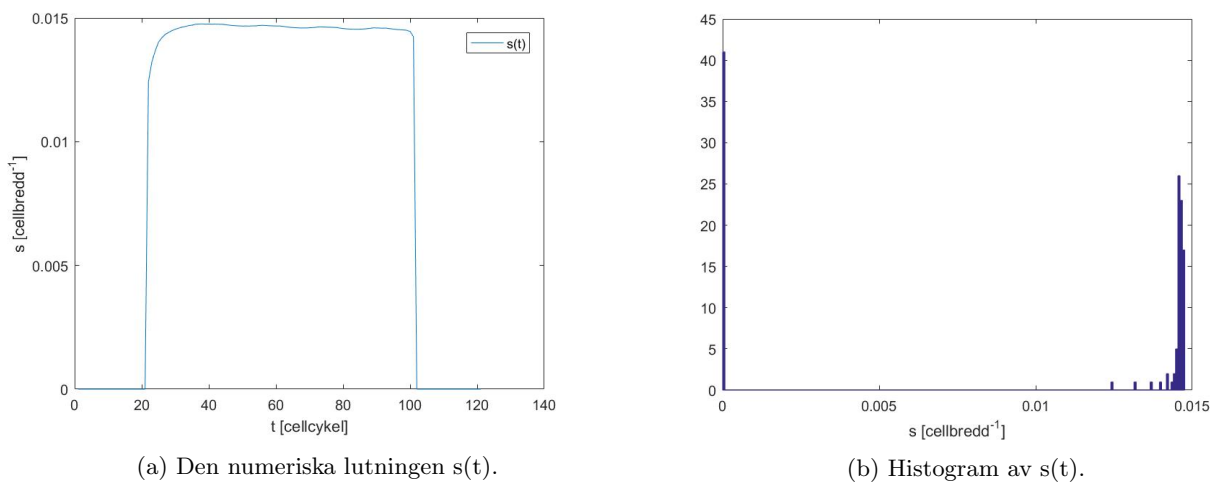
$$s_i = -\frac{0.16 - 0.80}{x_{i,016} - x_{i,080}}, \quad (6)$$

där $x_{i,0.16}$ och $x_{i,0.80}$ är skärningspunkterna mellan vågen U_i och värdena 0.16 respektive 0.80 för tidssteg t_i (se figur 4). Lutningen s är alltså en vektor med ett element för varje tidssteg (se figur 5a). Motiveringen till denna definition ligger i att det är detta värde som kan beräknas utifrån rådatan.



Figur 4: En grafisk illustration för att tydliggöra definitionen av lutningen. Parametervärden $\alpha = 1$, $D_v = 25$, $q_m = 20$, $q_p = 20$, $\mu = 0.1$. Steglängder $h_x = 5$, $h_t = 1$

Precis som för våghastigheten så önskas ett konstant representativt värde för lutningen.



Figur 5: Parametervärdena $\alpha = 1$, $D_v = 25$, $q_m = 20$, $q_p = 20$, $\mu = 0.1$. Steglängder $h_x = 5$, $h_t = 1$

För att erhålla ett värde för lutningen utan att behöva studera grafen för varje tidssteg används vad vi kallar för histogrammetoden. Histogrammetoden gör ett histogram av vektorn

s och returnerar medelvärden av det intervall som har flest värden. Detta värde är alltså det representativa värdet för lutningen s hos den färdande våglösningen. I figur 5b kan vi se att intervallet som inkluderar värdet 0 förekommer flest gånger. Värdet 0 är ingen lutning av intresse och därför försummas det intervallet. I ovanstående fall fås därför en konstant lutning på $s = 0.0145$.

3.2 Analytiska härledningar

För att kunna dra paralleller mellan modell och verklighet krävs, på grund av datorns nutida beräkningskraft, i regel inte längre exakta analytiska resonemang. Det kopplade systemet (3) råkar ändå tillåta (om något approximativa) analytiska härledningar vad gäller precis de egenskaperna i modellen som kan jämföras med verkligheten. Det visar sig i kapitel 3.3 samt 4.2 vara väldigt användbart. Utöver det jämför vi den analytiska metodiken för (3) med motsvarande i [7], för att se exakt vilken inverkan $\mu \neq 0$ har.

3.2.1 Vågshastighet via analys av stationära punkter

Kolmogorov m. fl. visade att Fisher-ekvationen har en begränsande utbredningshastighet som alla andra utbredningshastigheter konvergerar mot [10]. Gerlee m. fl. hittade [7] en motsvarande begränsande vågshastighet för systemet (3) med $\mu = 0$. Via samma metodik, som går ut på att föra över (3) till ett system av ODE som sedan görs autonomt för att slutligen göra en fasplansanalys, hittar även vi en begränsande vågshastighet.

Vi betraktar det analytiska utgångssystemet (3),

$$\begin{cases} \frac{\partial p}{\partial t} &= \alpha p(1 - p - m) - (q_m + \mu)p + q_p m \\ \frac{\partial m}{\partial t} &= D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p \end{cases} \quad (3)$$

Vi vet av de numeriska uträkningarna att våra lösningar erhåller färdande våglösningar (se kapitel 3.1), alltså inskränker vi oss fortsättningsvis ytterligare genom att definiera $z := x - ct$, där c är vågshastigheten, så att vi ifrån $P(z) = p(x, t)$ samt $M(z) = m(x, t)$ kan föra över (3) till ett system av kopplade ODE. Vi låter således $D(z) := P(z) + M(z)$. Genom kedjeregeln får vi de ordinära differentialekvationerna,

$$\begin{cases} -c \frac{dP}{dz} &= \alpha P(1 - P - M) - (q_m + \mu)P + q_p M \\ -c \frac{dM}{dz} &= D_v \frac{d^2 M}{dz^2} - q_p M + q_m P \end{cases} \quad (7)$$

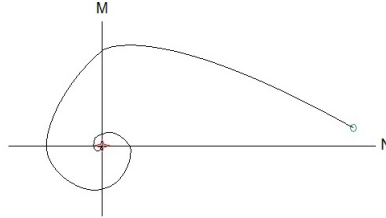
Vi vill ha ett första ordningens autonomt system för att kunna utnyttja den teorin, alltså följer vi Gerlees exempel [7] och introducerar en ny variabel N så att $M' = N$. Vi lämnas med,

$$\begin{cases} M' &= N, \\ N' &= 1/D_v(-cN + q_p M - q_m P) \\ P' &= -1/c(\alpha P(1 - P - M) - (q_m + \mu)P + q_p M) \end{cases} \quad (8)$$

För (8) har vi följande stationära punkter: den triviala $(M, N, P) = (0, 0, 0)$ samt den så kallade invaderade $(M, N, P) = (\frac{q_m}{\alpha(q_m + q_p)}(\alpha - \mu), 0, \frac{q_p}{\alpha(q_m + q_p)}(\alpha - \mu))$.

Det finns många banor till (8) som är intressanta analytiskt. Av speciellt intresse rent biologiskt är banan som går från den invaderade till den triviala stationära punkten, eftersom det är mellan dessa värden som vi har ett densitetsbeteende ($0 \leq P \leq 1$). Den triviala stationära punkten kan dessutom inte ha ett periodiskt beteende då våra lösningar i så fall skulle anta negativa värden, se Fig. 6 för en illustration i 2D. Med det i åtanke fortsätter vi analysen genom

att titta specifikt på Jacobianen $J(s)$ till (8) i den triviala stationära punkten. Eftersom en linjärisering kring punkten beskriver systemet lokalt (Grobman-Hartman) vill vi hitta $J(0)$'s egenvärden då det är dessa som i sin tur beskriver linjäriseringens beteende. Vi frågar oss alltså när $J(0)$'s karakteristiska polynom får *reella* rötter.



Figur 6: 2D system med periodisk bana i dess triviala stationära punkt

Vi deriverar således våra ekvationer i (8) partiellt och får på så sätt för $s = (M, N, P)$,

$$J(s) = \begin{bmatrix} 0 & 1 & 0 \\ q_p/D_v & -c/D_v & -q_m/D_v \\ 1/c(\alpha P - q_p) & 0 & 1/c(q_m - (\alpha - \mu) + \alpha(2P + M)) \end{bmatrix}.$$

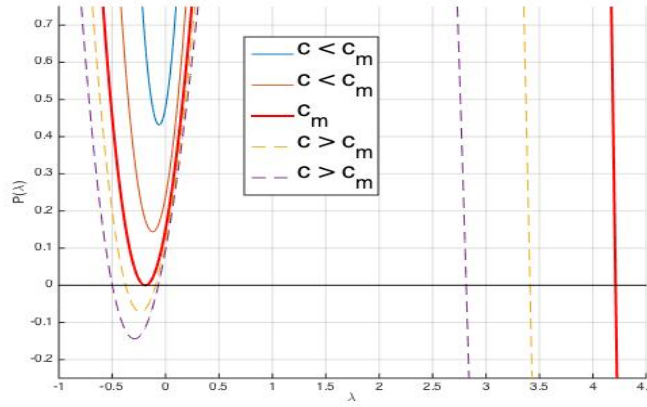
I origo (den triviala stationära punkten) gäller att,

$$J(0) = \begin{bmatrix} 0 & 1 & 0 \\ q_p/D_v & -c/D_v & -q_m/D_v \\ -q_p/c & 0 & 1/c(q_m - (\alpha - \mu)) \end{bmatrix}. \quad (9)$$

Egenvärdena ges som rötterna till det karakteristiska polynomet. Vi kan nu räkna ut det karakteristiska polynomet och därigenom i enighet med Gerlee [7] få fram ett krav på våghastigheten så att vi för våghastigheter som uppfyller kravet på den resulterande tredjegradaren i λ för reella rötter¹. Vi kallar denna begränsande våghastighet för c_m så att kravet för reella rötter (och biologiskt relevanta lösningar) blir $c > c_m$. Kolmogorov m. fl. visade att våghastigheter skilda från c_m för den enklare ekvationen $\frac{\partial v}{\partial t} - \frac{\partial^2 v}{\partial x^2} = F(v)$ konvergerar mot c_m [10]. Vi visar inte detta för (3) utan fortsätter heuristiskt.

Fig. 7 visar hur rötterna till det karakteristiska polynomet blir reella exakt då $c = c_m$, det vill säga då polynomet har tre reella rötter och en dubbelrot.

¹I själva verket undersöks diskriminanten till tredjegradaren eftersom tredjegradaren blir för komplicerad. Fallet med reella rötter och en dubbelrot (vilket sett från Fig. 7 är precis fallet vi är ute efter) svarar precis mot diskriminanten lika med noll.



Figur 7: Karakteristiska polynomet av Jacobianen i origo, ritad för olika värden av våghastigheten c . $D_v = 25$, $\alpha = 1$, $q_m = q_p = 20$ samt $\mu = 0.1$.

Notera nu att Jacobianen i origo (9) är *exakt* samma Jacobian som Gerlee får [7], förutom att α har bytts ut mot $\alpha - \mu$. Om vi alltså skapar en ny parameter $\hat{\alpha} = \alpha - \mu$, proliferationstakten, så ser vi att resten av uträkningarna blir fullständigt analoga till Gerlees fall då $\mu = 0$. Därav får vi den nedre begränsande våghastigheten som,

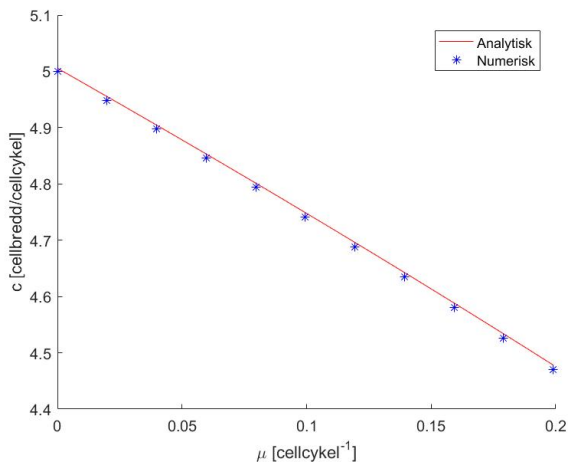
$$c_m = \sqrt{\frac{D_v}{J} \sqrt{K^2 + I} - K}, \quad (10)$$

där

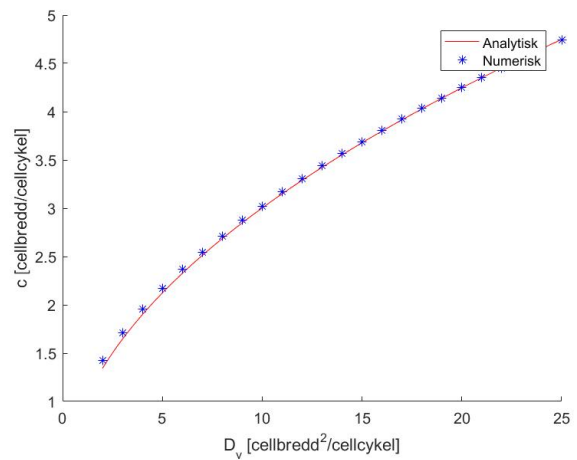
$$\begin{aligned} K &= A^2 E^2 - 18AEq_p \hat{\alpha} - 27q_p^2 \hat{\alpha}^2 + 12A^2 q_p \hat{\alpha}, \\ I &= 16A^3 q_p \hat{\alpha} (4E^3 + 18Eq_p \hat{\alpha} - 12Aq_p \hat{\alpha} - 2AE^2), \\ J &= 2(4E^3 + 18Eq_p \hat{\alpha} - 12Aq_p \hat{\alpha} - 2AE^2), \end{aligned}$$

och i sin tur $A = q_m - \hat{\alpha}$ samt $E = q_m + q_p - \hat{\alpha}$.

En jämförelse mellan ovan analytiska våghastighet och den vi får av den numeriska lösningen ges i Fig. 8a samt Fig. 8b. Vi låter i Fig. 8a μ variera medan de andra parametrarna hålls konstanta, medan vi i Fig. 8b låter D_v variera.



(a) $D_v = 25$, $\alpha = 1$ samt $q_m = q_p = 20$.
 μ varierar.



(b) $\mu = 0.1$, $\alpha = 1$ samt $q_m = q_p = 20$.
 D_v varierar.

Figur 8: Röd hastighet - analytisk, blå hastighet - numerisk.

Eftersom vårt fall är helt analogt kan vi även räkna den förenklade våghastigheten på precis samma sätt som Gerlee m. fl. [7]. Följande uttryck är endast giltigt då $q = q_m = q_p \gg \hat{\alpha}$.

$$c_m^* = \sqrt{\frac{D_v \hat{\alpha}}{1 - \hat{\alpha}/(3q)}}, \quad q = q_m = q_p. \quad (11)$$

3.2.2 Lutning genom perturbationslösning

För bekvämlighetens skull skriver vi åter igen systemet (7) nedan,

$$\begin{cases} cP' + \alpha P(1 - P - M) - (q_m + \mu)P + q_p M & = 0 \\ cM' + D_v M'' - q_p M + q_m P & = 0 \end{cases}. \quad (7)$$

Vi vill perturbera en liten parameter (som är < 1), vilket innebär att vi vill serietveckla² lösningarna i denna parameter så att högre ordningens termer blir av relativt liten betydelse. Vi antar utan stor inskränkning att våra lösningar är av snäll karaktär och att vi kan derivera dem ett oändligt antal gånger.

Vi väljer att perturbera proliferationstakten $\hat{\alpha} := \alpha - \mu$ eftersom vi kan se denna parameter som vår nya tillväxtparameter. Vi skalar om rummet med $\xi = \hat{\alpha}z$ så att $\frac{dP}{dz} = \hat{\alpha} \frac{dP}{d\xi}$ och $\frac{dM}{dz} = \hat{\alpha} \frac{dM}{d\xi}$. Vi låter vidare $f(\xi) := P(z)$ samt $g(\xi) := M(z)$ och får efter att ha skrivit om (7) i $\hat{\alpha}$ följande uttryck,

$$\begin{cases} \hat{\alpha} c f' + \hat{\alpha} f - \alpha f(f + g) - q_m f + q_p g & = 0 \\ \hat{\alpha} c g' + D_v g'' - q_p g + q_m f & = 0 \end{cases}. \quad (12)$$

I nästa steg utför vi serietveckling på $f(\xi)$ och $g(\xi)$ i parametern³ $\hat{\alpha}$, och kan skriva

$$\begin{cases} f(\xi) = f_0(\xi) + \hat{\alpha} f_1(\xi) + \frac{\hat{\alpha}^2}{2} f_2(\xi) + HOT \\ g(\xi) = g_0(\xi) + \hat{\alpha} g_1(\xi) + \frac{\hat{\alpha}^2}{2} g_2(\xi) + HOT \end{cases}. \quad (13)$$

Genom att expandera f och g på samma vis som (13) i (12) får vi efter matchande av storleksordningar följande sex ekvationer, först tre för f och sedan tre för g ,

$$\begin{aligned} O(1)_f &: q_p g_0 - q_m f_0 - \alpha f_0(f_0 + g_0) = 0, \\ O(\hat{\alpha})_f &: q_p g_1 - q_m f_1 + c f_0' + f_0 - \alpha(f_0(f_1 + g_1) + f_1(f_0 + g_0)) = 0, \\ O(\hat{\alpha}^2)_f &: \frac{q_p}{2} g_2 - \frac{q_m}{2} f_2 + c f_1' + f_1 + \\ &\quad - \alpha \left(\frac{f_0}{2} (f_2 + g_2) + \frac{f_2}{2} (f_0 + g_0) + f_1(f_1 + g_1) \right) = 0, \\ O(1)_g &: q_m f_0 - q_p g_0 = 0, \\ O(\hat{\alpha})_g &: q_m f_1 - q_p g_1 + c g_0' = 0, \\ O(\hat{\alpha}^2)_g &: \frac{q_m}{2} f_2 - \frac{q_p}{2} g_2 + c g_1' + D_v g_0'' = 0. \end{aligned}$$

Om vi stoppar in resultatet från $O(1)_g$ i ekvationen $O(1)_f$ ges att $\alpha f_0^2(1 + q_m/q_p) = 0$, vilket betyder (eftersom $q_m, q_p \geq 0$) att $f_0 \equiv 0$. Därför gäller också att $g_0 \equiv 0$, $f_0' = 0$ samt $g_0' = 0$. Vi

²Konvention är att baka in $1/2$ i f_2 och på så sätt göra en potensserietveckling, det finns ingen särskild anledning till varför vi valde att göra en Taylor-utveckling istället - de är ekvivalenta i det här fallet.

³Här ska nämnas att man rent stringent borde utveckla i en dimensionslös variabel för att utvecklingen ska vara logisk. Vi borde alltså dela med q_m och utveckla i $\hat{\alpha} = (\alpha - \mu)/q_m$ men eftersom det ger samma svar i slutändan låter vi bli för tydlighetens skull.

har nu följande ekvationer kvar,

$$\begin{aligned} O(\hat{\alpha}^2)_f &: \frac{q_p}{2}g_2 - \frac{q_m}{2}f_2 + cf'_1 + f_1 - \alpha(f_1(f_1 + g_1)) = 0, \\ O(\hat{\alpha})_g &\&O(\hat{\alpha})_f : q_m f_1 - q_p g_1 = 0, \\ O(\hat{\alpha}^2)_g &: \frac{q_m}{2}f_2 - \frac{q_p}{2}g_2 + cg'_1 = 0. \end{aligned}$$

Vi har nu att $g_1 = \frac{q_m}{q_p}f_1$ och därför gäller även $g'_1 = \frac{q_m}{q_p}f'_1$. Uttrycker vi g'_1 på det sättet i ekvationen $O(\hat{\alpha}^2)_g$ ges följande samband, $g_2 = \frac{q_m}{q_p}f_2 + \frac{2cq_m}{q_p^2}f'_1$. Detta uttryck för g_2 samt motsvarande uttryck för g_1 stoppar vi sedan in i ekvationen $O(\hat{\alpha}^2)_f$ och får

$$\begin{aligned} 0 &= \frac{q_p}{2} \left(\frac{q_m}{q_p}f_2 + \frac{2cq_m}{q_p^2}f'_1 \right) - \frac{q_m}{2}f_2 + cf'_1 + f_1 - \alpha \left(f_1 \left(f_1 + \frac{q_m}{q_p}f_1 \right) \right) \\ \iff 0 &= c(1 + \nu)f'_1 + f_1 - \alpha(1 + \nu)f_1^2, \quad \nu := \frac{q_m}{q_p}. \end{aligned}$$

Denna (logistiska) första ordningens olinjära ODE har den generella lösningen

$$f_1(\xi) = \frac{1}{e^{\frac{1}{c(1+\nu)}\xi+k} + \alpha(1+\nu)}, \text{ där } k \text{ är en godtycklig konstant.}$$

Härifrån har vi alltså genom $\xi = \hat{\alpha}z$ även att $g_1(\xi) = g_1(z) = \frac{q_m}{q_p}f_1(z) = \nu f_1(z)$. Således ges för den totala densiteten $D(z) := f(z) + g(z) = \hat{\alpha}(f_1(z) + g_1(z)) + \hat{\alpha}^2(f_2(z) + g_2(z)) + HOT$, att

$$D(z) = \frac{\hat{\alpha}(1 + \nu)}{e^{\frac{\hat{\alpha}}{c(1+\nu)}z+k} + \alpha(1 + \nu)} + \hat{\alpha}^2(f_2(z) + g_2(z)) + HOT.$$

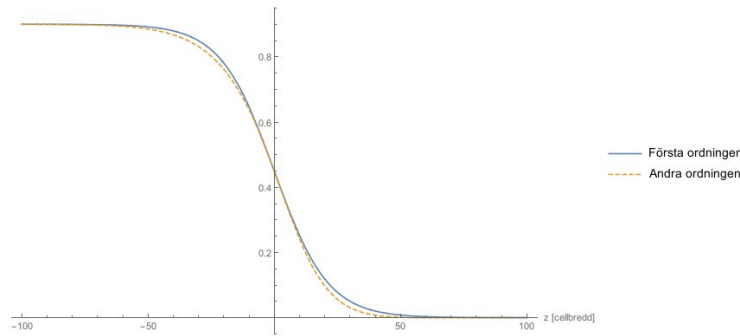
Vi vill nu kontrollera att våra gräns- och randvärden stämmer och därigenom bestämma k . En redogörelse för vilka randvärden och gränsvärden som $d = f + g$ antar kan fås i bilaga B. Med vetskapen om att $f_0 \equiv g_0 \equiv 0$ samt att lösningar till f_1 & g_1 kan hittas bestämmer vi följande rand- och gränsvärden,

$$\left\{ \begin{array}{ll} \lim_{z \rightarrow -\infty} \hat{\alpha}(f_1(z) + g_1(z)) = 1 - \mu/\alpha = \frac{\hat{\alpha}}{\alpha} & \lim_{z \rightarrow -\infty} f_2(z) + g_2(z) = 0 \\ \hat{\alpha}(f_1(0) + g_1(0)) = \frac{1}{2}\frac{\hat{\alpha}}{\alpha} & f_2(0) + g_2(0) = 0 \\ \lim_{z \rightarrow \infty} D(z) = 0 & \lim_{z \rightarrow \infty} f_2(z) + g_2(z) = 0 \end{array} \right\}$$

Det återstår nu att undersöka ifall vår lösning uppfyller dessa. Vi låter

$$\overline{D(z)} := \hat{\alpha}(1 + \nu)f_1(z) = \frac{\hat{\alpha}(1 + \nu)}{e^{\frac{\hat{\alpha}}{c(1+\nu)}z+k} + \alpha(1 + \nu)},$$

och noterar att $\lim_{z \rightarrow \infty} \overline{D(z)} = 0$, $\lim_{z \rightarrow -\infty} \overline{D(z)} = \frac{\hat{\alpha}}{\alpha}$. Med andra ord är gränsvärdena uppfyllda direkt. För att uppfylla randvärdet i $z = 0$ tvingar vi helt enkelt k att bestämmas så att randvärdet gäller. Speciellt får vi $k = \log(\alpha(1 + \nu))$. Alltså uppfyller $\overline{D(z)}$ alla de gräns- och randvärden som den per dess definition måste uppfylla.



Figur 9: Olika ordningars approximation av densitet som funktion av z , vi noterar en väldigt liten skillnad. $D_v = 25$, $\alpha = 1$, $\mu = 0.1$ samt $q_m = q_p = 20$.

Genom att utveckla f och g till $O(\hat{\alpha}^3)$ går det genom att addera de två nya ekvationerna för att få ett uttryck för f_2 samt g_2 . Rent resultatmässigt uppfyller detta ingen funktion då dessa uttryck för f_2 och g_2 ger ett minimalt bidrag till värdet av $D(z)$. Detta illustreras i Fig. 9. För själva uttrycken f_2 och g_2 i sig, samt uträkningarna som leder till dem, hänvisas läsaren till Mathematica-koden i bilaga F. Vi låter alltså $D(z)$ approximeras av $\overline{D(z)}$ och nöjer oss efter viss förenkling, det vill säga

$$D(z) \approx \overline{D(z)} = \frac{\hat{\alpha}}{\alpha(e^{\frac{\hat{\alpha}}{c(1+\nu)}z} + 1)}. \quad (14)$$

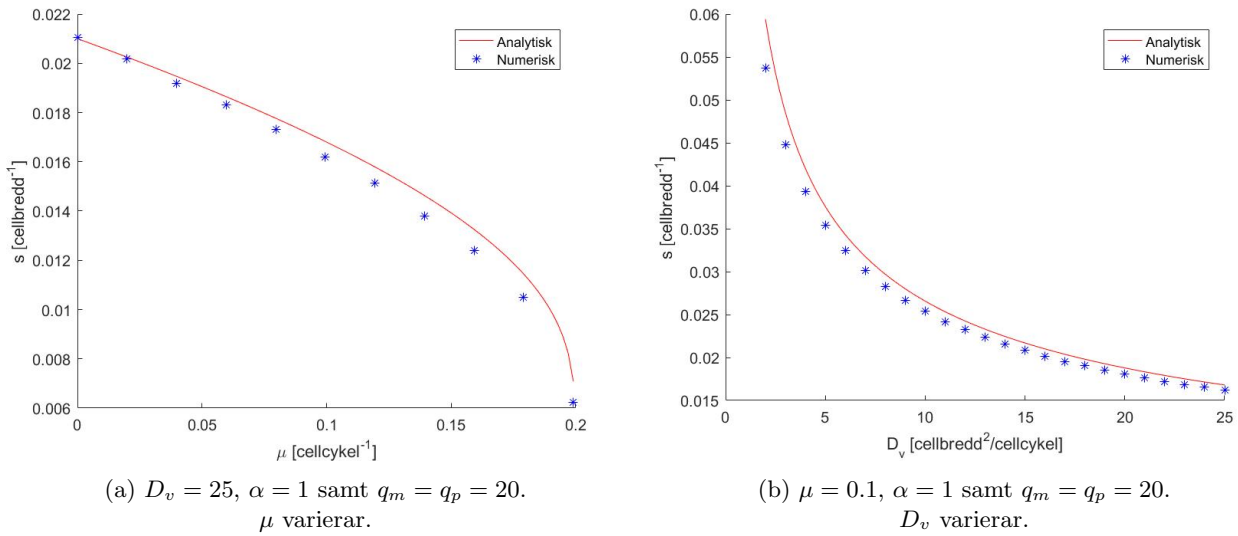
Eftersom de numeriska metoderna räknar ut en approximativ lutning på formen av en sekant $-\frac{0.16-0.8}{x_{0.16}-x_{0.8}}$ är vi intresserade av att finna inversen i x till (14). Vi noterar först att $\overline{D(z)} \approx d(x, t)$ och skriver $d(x, t) = \frac{\hat{\alpha}}{\alpha(e^{\frac{\hat{\alpha}(x-ct)}{c(1+\nu)}} + 1)}$. Vi kan på grund av dess enkla form få x som funktion av d och t ,

$$x(d, t) = ct + \frac{c(1+\nu)}{\hat{\alpha}} \log\left(\frac{\hat{\alpha}}{d\alpha} - 1\right). \quad (15)$$

Med hjälp av uttrycket (15) ovan för x får vi en motsvarande analytisk sekant, det vill säga

$$s = -\frac{0.16 - 0.8}{x(0.16, t) - x(0.8, t)}. \quad (16)$$

Notera att tidsparametern inte spelar någon roll då den oavsett värde kommer slås ut i nämnaren till (16). Vi kan ta reda på hur bra denna perturbationslutning stämmer överens med vad som ges av det numeriska (för det numeriska se kapitel 3.1.2). Vi illustrerar resultatet i Fig. 10a och 10b och noterar en väldigt stark korrelation.



Figur 10: Röd lutning - analytisk (16), blå lutning - numerisk.

3.3 Metoder för anpassning av modell till rådata

Numeriska- och analytiska metoder, härledda i kapitel 3.1 och 3.2, vill vi nu använda för att koppla modell till rådata. För detta kommer två dataanalysmetoder att härledas i kommande sektioner. Den ena baserad på numeriska metoder och den andra baserad på analytiska uttryck. Läsaren intresserad av datorintensiviteten och effektiviteten av dataanalysmetoderna i praktiken hänvisas till bilaga E.2.

3.3.1 Minimeringsproblem baserad på numeriska resultat

Huvudmetoden för anpassning av modell till rådata i detta kapitel bygger på att numeriskt simulera en familj av lösningar med olika ingående parametrar. Ur denna familj bestäms med en normerad minstakvadratmetod sedan den lösning som främst överensstämmer med rådata. Därefter antas att parametrarna hos lösningen utgör den bästa anpassningen, modell och rådata emellan. Minimeringsproblemet kommer att beskrivas konceptuellt. Anledningen är främst att den exakta implementationen i praktiken är väldigt ingående. Detta är en konsekvens av att numeriska lösningar har väldigt skilda beteenden emellan varandra, och kräver skilda kodstycken för att implementeras så korrekt som möjligt.

Betrakta nu utgångssystemet

$$\begin{cases} \frac{\partial p}{\partial t} &= \alpha p(1 - p - m) - (q_m + \mu)p + q_p m \\ \frac{\partial m}{\partial t} &= D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p \end{cases}. \quad (3)$$

Av främst intresse är lösningarnas karaktärer relativt varandra då α , D_v och μ tillåts variera. Därför ansätts fortsättningsvis $q_m = q_p$ till ett fixt värde⁴. Generellt har systemet tre frihetsgrader, α , D_v och μ , vilka tillåts variera över biologiskt motiverbara värden. För att anpassa numeriska lösningar efter rådata låtes

$$\mathcal{F} := \{f_1, f_2, \dots, f_k, \dots, f_N\}, \quad (17)$$

ange mängden av numeriska lösningar simulerade med olika ingående värden på (α, D_v, μ) . Generellt är beteendet av numeriska lösningar beroende av val på initialvillkor och sluttid på

⁴Metoder fungerar generellt analogt utan detta antagandet.

simulerade lösningar. Varierande (α, D_v, μ) medför lösningar som kräver olika lång tid för att erhålla vågliknande karaktärer, vilket är det som eftersöks. Därför är det viktigt att en godtycklig numerisk lösning simuleras med tillräckligt lång sluttid innan den inkluderas i mängden \mathcal{F} . För biologiskt förankrade initialvillkor antar vi att

$$p(x, 0) = \gamma \exp(-\beta x), \beta, \gamma > 0, m(x, 0) = 0$$

används. Det vill säga en initialt exponentiellt avtagande fördelning av prolifererande cancer-celler. Det är även den typ av initialvillkor som används i kapitel 3.1 och i simuleringar som ger resultatet av dataanalysen i kapitel 4.2. Fortsättningsvis antas γ och β ta de värden som angavs i sektion 3.1. Det har dock experimentellt visat sig att dynamiken av numeriska lösningar asymptotiskt, med avseende på tiden, är oberoende av valet på γ och β , vilket är av främst intresse för oss. Sålänge de antar positiva värden och att $\gamma \leq 1$. För en godtycklig numerisk lösning $f_i \in \mathcal{F}$ kan sedan dess associerade våghastighet och lutning beräknas med numeriska metoder redogjorda i avsnitt 3.1.

Anpassning av numeriska lösningar till rådata baseras på införandet av ett minimeringsproblem. Låt nu s_k, c_k beteckna lutning respektive hastighet av tumörradien hos en godtycklig patient med index k , uppmätt från rådata. En anmärkning är att dessa mått, ur ett praktiskt perspektiv, kan betraktas som riktiga värden, det vill säga de verkliga approximativa värdena på tumörens hastighet och form. Speciellt, vill vi relatera s_k, c_k till mängden av numeriska lösningar. Av denna anledning definieras målfunktionen för minimeringsproblemet enligt

$$M_k(f_i) := \frac{(s_k - s_{f_i})^2}{s_k^2} + \frac{(c_k - c_{f_i})^2}{c_k^2}, \quad (18)$$

där s_{f_i}, c_{f_i} är lutning respektive hastighet av den numeriska lösningen f_i . Notera att målfunktionen är ett mått på den relativa avvikelsen då lutning och hastighet av f_i approximerar s_k och c_k . Minimeringsproblemet definieras då enligt

$$\min_{f_i \in \mathcal{F}} M_k(f_i). \quad (19)$$

Genom att diskretisera intervall av α, D_v respektive μ enligt önskad noggrannhet och i följd lösa problem (19) kan optimala parametrar lösas ut. I fasen före behandling ansätts $\mu = 0$ och problemet bestäms då av endast två frihetsgrader.

I den implementation som medför resultaten som tas upp i 4.2 har $\alpha \in [0, 1], D_v \in [25, 2000]$ diskretiserats med en steglängd på 0.01 respektive 100. I fasen efter behandling antar $\mu \in [0, 0.19\alpha]$ en diskretisering med steglängd $\frac{0.19\alpha}{5}$. Detta val av diskretisering täcker en bred räckvidd av värden på (α, D_v, μ) , utan att ta upp för mycket simulationstid.

3.3.2 Minimeringsproblem baserad på analytiska resultat

En anpassning av modell till rådata kan utföras via två analytiska huvudresultat. För dessa påminns läsaren av den analytiska förenklade våghastigheten,

$$c(\alpha, \mu, D_v) = \sqrt{\frac{D_v \hat{\alpha}}{1 - \hat{\alpha}/(3q)}}, \quad q = q_m = q_p, \quad (11)$$

som beskriver beteendet av våghastigheten asymptotiskt, under vissa antaganden redogjorda i kapitel 3.2.1. Även den analytiska lutningen kommer att användas, given av

$$s(\alpha, \mu, D_v) = -\frac{0.16 - 0.8}{x(0.16, t) - x(0.8, t)}, \quad x(d, t, \alpha, \mu, D_v) = ct + \frac{c(1 + \nu)}{\hat{\alpha}} \log\left(\frac{\hat{\alpha}}{d\alpha} - 1\right). \quad (16)$$

Vi vill nu relatera dessa uttryck till rådata. Låt därför c_k , s_k beteckna hastighet respektive lutning av tumörradien för en godtycklig patient med index k . Med dessa ekvationer (11), (16) av icke-trivial karaktär och tre obekanta visar det sig generellt analytiskt omöjligt att lösa ut för c_k , s_k explicit. Av denna anledning är det istället lämpligt att introducera ett flervariabelt minimeringsproblem. Huvudmålet blir att lösa ut de *bästa* parametrarna $(\alpha_k, \mu_k, D_{v,k})$ för varje patient. För ett mått på felet av att approximera c_k , s_k med Ekv. (11) respektive Ekv. (16) introduceras en målfunktion lik den använd i kapitel 3.3.1. Målfunktionen definieras enligt

$$E(\alpha, D_v, \mu) = \frac{(s_k - s(\alpha, D_v, \mu))^2}{s_k^2} + \frac{(c_k - c(\alpha, D_v, \mu))^2}{c_k^2}. \quad (20)$$

Likt målfunktionen för minimeringsproblemet i kapitel 3.3.1 mäter den relativ avvikelse. Bivillkoren för problemet blir att endast betrakta icke-negativa värden på parametrarna, vilket är biologiskt motiverat. Dessutom, för att undvika komplexa lösningar, krävs det att $5\mu < \alpha$. Detta är dels matematiskt motiverat av logaritmens argument i högra uttrycket i (16), vilket måste anta ett positivt värde, och dels biologiskt motiverat eftersom $1 - \frac{\mu}{\alpha} > 0.8 \Rightarrow \mu < 0.2\alpha$. Det vill säga, en skärning av cancerandelsfunktionen $p(x, t) + m(x, t)$ med $y = 0.8$ måste existera för att metoden ska fungera. Minimeringsproblemet definieras då enligt

$$\begin{aligned} & \underset{(\alpha, D_v, \mu) \geq (0, 25, 0)}{\text{minimera}} && E(\alpha, D_v, \mu), \\ & \text{så att} && 5\mu < \alpha. \end{aligned} \quad (21)$$

Av detta kan optimala parametrar för varje patient lösas ut.

Vi vill även notera som ett specialfall att i fasen före behandling (karaktäriserat av $\mu = 0$) kan faktiskt optimala parametrar $(\alpha_k, D_{v,k})$ för en godtycklig patient lösas ut explicit, enligt

$$\alpha_k = \frac{2s_k c_k \log(21)}{0.64}, \quad (22)$$

$$D_{v,k} = \frac{c_k^2}{\alpha_k} \left(1 - \frac{\alpha_k}{3q}\right). \quad (23)$$

För en detaljerad redogörelse av detta samt härledning till ett explicit uttryck för målfunktionen (20) hänvisas läsaren till bilaga E.3. Implementation av denna metod baseras på att helt enkelt använda färdiga optimeringsalgoritmer i Mathematica och Matlab. Eftersom en implementation i kod är relativt enkel för denna dataanalysmetoden, bifogar vi med associerad kod i bilaga F.

Det visar sig även att målfunktionerna antar väldigt små värden, som funktion av optimala parametrar för varje patient. Detta gäller för båda dataanalysmetoderna. En direkt konsekvens är att de bästa parametrarna för respektive patient är väldigt bra anpassade. Mer om detta i bilaga E.2.

4 Resultat

För snabb referens katalogiseras våra huvudsakliga resultat här. För en mer ingående redogörelse av härledningarna ledandes till analytiska resultat hänvisas läsaren till kapitlen 3.2.1 för våghastigheten samt 3.2.2 för lutningen. Vad gäller mer diskussion av graferna i därtill nästkommande kapitel 4.2, hänvisas läsaren först till kapitel 5.1 och sedan för mer ingående analys till bilaga E.

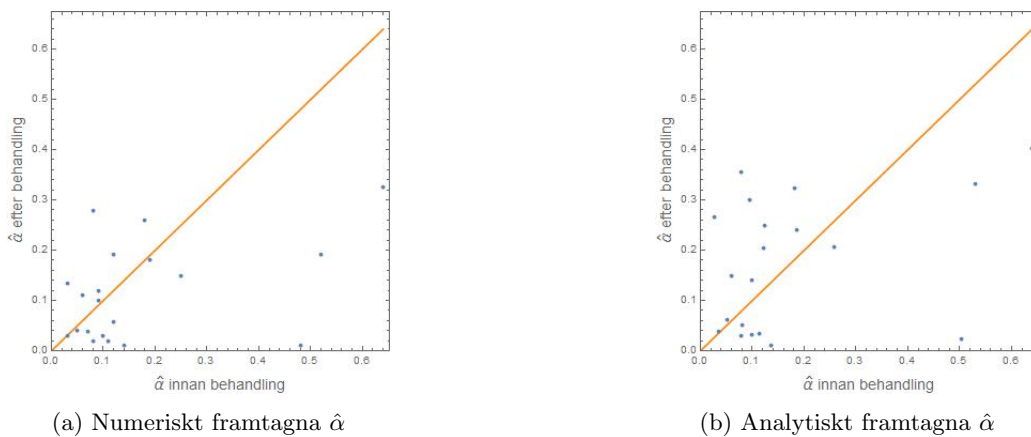
4.1 Analytiska uttryck

Tabell 2: Analytiska uttryck. Notera att $\hat{\alpha} = \alpha - \mu$.

Ekvation	Benämning	Beskrivning
$c = \sqrt{\frac{D_v}{J} \sqrt{K^2 + I} - K}$ $A = q_m - \hat{\alpha}, \quad E = q_m + q_p - \hat{\alpha}$ $K = A^2 E^2 - 18AEq_p \hat{\alpha} - 27q_p^2 \hat{\alpha}^2 + 12A^2 q_p \hat{\alpha}$ $I = 16A^3 q_p \hat{\alpha} (4E^3 + 18Eq_p \hat{\alpha} - 12Aq_p \hat{\alpha} - 2AE^2)$ $J = 2(4E^3 + 18Eq_p \hat{\alpha} - 12Aq_p \hat{\alpha} - 2AE^2)$	Våghastighet	Hastighet av propagerande vågor.
$c = \sqrt{\frac{D_v \hat{\alpha}}{1 - \hat{\alpha}} (3q)}, \quad q = q_m = q_p \gg \hat{\alpha}$	Förenklad våghastighet	Hastighet av propagerande vågor, $q_m = q_p$.
$d(x, t) = \frac{\hat{\alpha}}{\alpha} (e^{\frac{\hat{\alpha}}{c(1+\nu)}(x-ct)} + 1)^{-1}$	Perturbationslösning	Sluten form av propagerande våglösningar.
$s = -\frac{0.16-0.8}{x(0.16,t)-x(0.8,t)}, \quad x(d, t) = ct + \frac{c(1+\nu)}{\hat{\alpha}} \log\left(\frac{\hat{\alpha}}{d\alpha} - 1\right)$	Lutning	Branthet på propagerande lösningar.

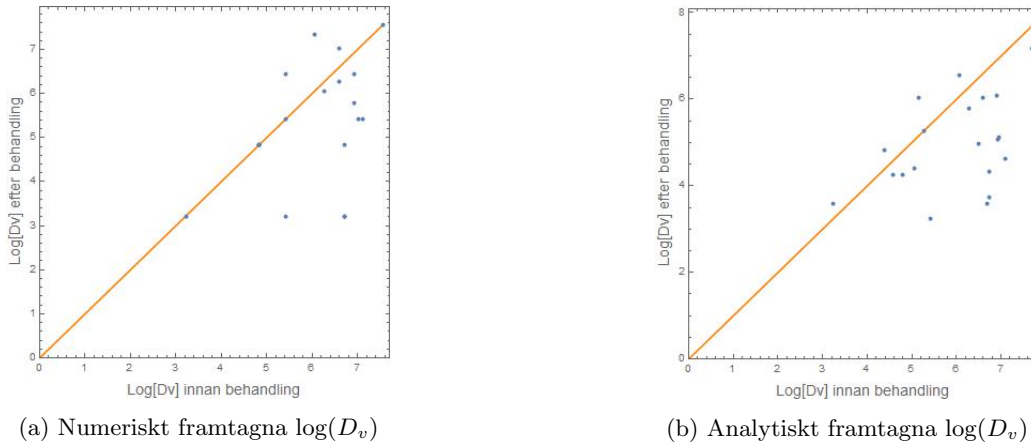
4.2 Jämförelse av $\hat{\alpha}$ och D_v före och efter behandling

I detta kapitel dokumenterar vi resultatet av kopplingen mellan systemet (3) och patientdatan. Kommande nämnda procentsatser refererar till numerisk framtaget resultat. Detta av anledning att numeriskt och analytiskt framtaget resultat är nästintill identiskt, med avseende på procentsatser om andel som påverkas positivt eller negativt över behandlingsfaserna.



Figur 11: Proliferationstakten $\hat{\alpha}$, före och efter behandling. Notera att en trend är svår att se, men att oroligt många patienter har en ökad proliferationstakt trots behandling.

Patienternas respektive behandling tycks ha en mild inverkan på canceren. Av alla patienter får 60% av dem en minskad eller stagnerad tillväxthastighet över faserna, där 25% av patienterna har en över 50% minskad tillväxthastighet. Det är en indikation att cancerbehandlingen verkar mycket väl för ett mindre antal patienter. Ett viktigt resultat är att $\hat{\alpha}$, och då speciellt även α , är väldigt lågt. Detta indikerar en långsam delningstakt hos cancercellerna.



Figur 12: Logaritmen av diffusionskonstanten D_v före och efter behandling. D_v tenderar att minska som effekt av behandling.

Enligt figur 12 kan vi se att parametervärdet D_v tenderar att minska hos patienterna. Av alla patienter genomgår ca 50% en minskad diffusionskonstant, vilket tyder på en god effekt på patienter kollektivt. En anmärkning är att diffusionskonstanten förblir ungefär oförändrad för 35% av alla patienter, vilket biologiskt tyder på effektlös cancerbehandling med avseende på cancercellers motilitet. Resterande andel av patienter har en diffusionskonstant som ökar över behandlingsfaserna, vilket antyder en svårighet för cancerbehandlingen att angripa motiliteten hos cancercellerna.

5 Diskussion och slutsatser

Vi har lyckats både numeriskt och analytiskt hitta en anpassning mellan modellen (3) och patientdatan. Det återstår att analysera resultaten. I det här kapitlet redogör vi för hur vi tolkar dessa resultat. Vi gör sedan en koppling mellan modell och verklighet för att hitta samband mellan systemets förmåga som modell och rådatans kvalitet. Slutligen dokumenterar vi eventuella felkällor samt de avgränsningar vi tvingats göra under projektets gång. För en djupare analys av dataanalysmetoderna och hur de presterar relativt varandra hänvisas läsaren till bilaga E. Vid intresse av en mer detaljerad analys av perturbationslösning mot numerisk lösning hänvisas läsaren till bilaga D.

5.1 Analytisk och numerisk dataanpassning

Anpassningen av modellen till rådata har varit lyckad. Med våra metoder för dataanalys har parametrarna $\hat{\alpha} = \alpha - \mu$ och D_v , som bäst beskriver cancerutvecklingen för respektive patient hittats. Detta har utförts för båda dataanalysmetoderna med generellt lika resultat.

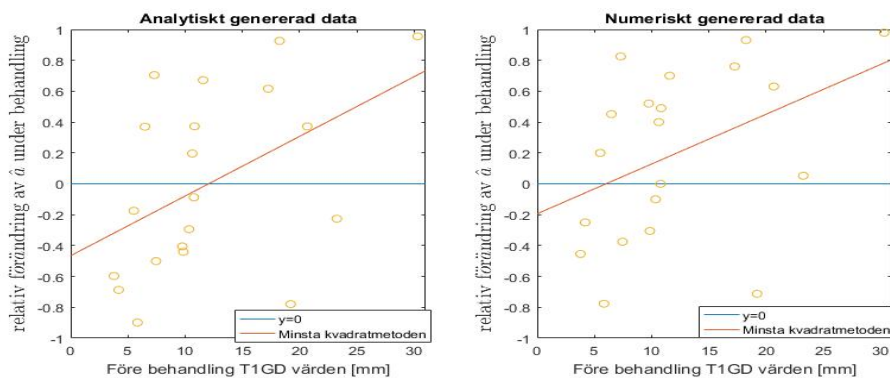
Det är emellertid utifrån resultaten från dataanalysen som den färdande vågansatsen kan anses rimlig. Detta gäller eftersom den numeriska dataanpassningen är helt oberoende vågansatsen, och att resultaten för den numeriska och analytiska dataanpassningen är generellt lika, med någon avvikelse. Vi noterar att vågansatsen kan *anses* rimlig och inte att vi har bevisat att den gäller.

Av särskilt intresse är relationer och mönster i patientdata för de 60% av patienter som påverkas väl av cancerbehandlingen, har de möjligtvis gemensamma karaktärer på sin tumörutveckling? Genom en ingående analys av patientdatan visar det sig att de väl påverkade patienternas tumörer har generellt skilda karaktärer. Tumörradien för patienterna skiljer sig markant; en

del har en stor tumör vid diagnos och en del har en liten tumör. Vidare varierar lutningen och våghastigheten av tumörutvecklingen från patient till patient. Det visar sig att detta gäller generellt. Genom att jämföra cancerkaraktärer från rådata av patienter med liknande förändringar av parametrar följer inga direkta samband. Detta kan tolkas som att cancers beteende är alldeles för komplicerad för att enkla, tydliga samband mellan patienter kan hittas under nuvarande modell, data och avgränsningar. Däremot finns det trender mellan $\hat{\alpha}$, D_v och dödsdag efter behandling för en andel av patienter. Dessutom finns det, för en andel av patienter, en trend för $\hat{\alpha}$, D_v och tumörradie vid diagnos. Dessa trender kommer att diskuteras härnäst.

5.2 Kopplingar mellan diverse patientdata och resultatet

Förändring i $\hat{\alpha}$ och D_v som resultat av behandling har analyserats gentemot tumörradiens storlek före behandling, där T1GD-känslighetsnivån har använts för att representera tumörradien. Att specifikt T1GD används som representant av tumörradien har ingen större påverkan på resultatet då samma trender hittades med T2GD/Flair-känslighetsnivån som radien.

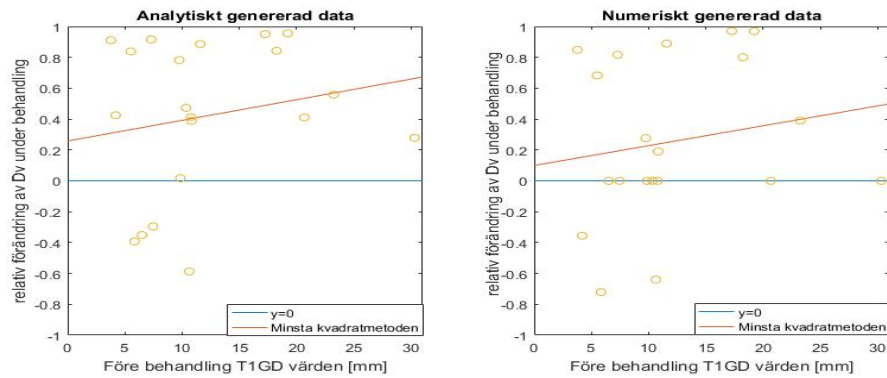


Figur 13: relativ $\hat{\alpha}$ minskning som resulterar behandling, gentemot tumörens T1GD-radie. Y-värdet av varje punkt är givet av ekvationerna (24).

$$\frac{\hat{\alpha}_{före} - \hat{\alpha}_{efter}}{\max[\hat{\alpha}_{före}, \hat{\alpha}_{efter}]} \quad (24)$$

Figur 13 illustrerar minskningen av tumörens $\hat{\alpha}$ värden. Förändringen av $\hat{\alpha}$, har dividerats med maximala $\hat{\alpha}$ före eller efter behandling, för att normera graferna och göra dem mer lättlästa⁵. Nollvärden innebär att ingen förändring skett under behandling, positiva värden innebär minskning av $\hat{\alpha}$ medan negativa värden innebär ökning. Minsta kvadratmetoden har applicerats för att analysera trender i datan. En linje längs origo har även lagts till.

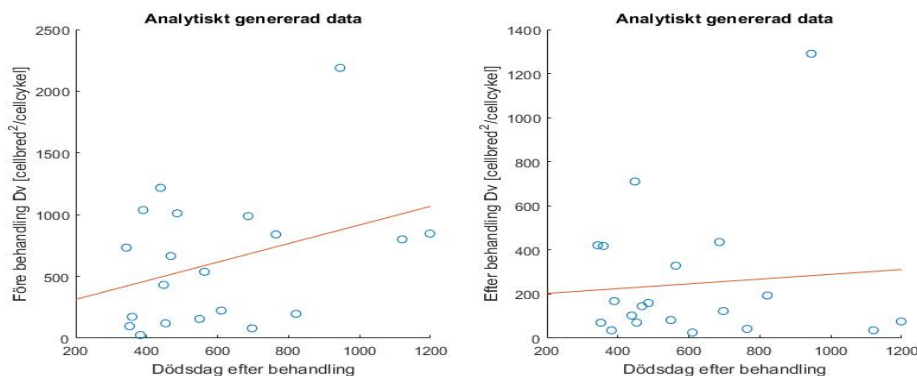
⁵relativa skillnaden med division av $\hat{\alpha}_{före}$ ger samma beteende och trender.



Figur 14: relativ D_v minskning som resulterar behandling. gentemot tumörens T1GD-radie. Y-värdet av varje punkt är givet av ekvation (25).

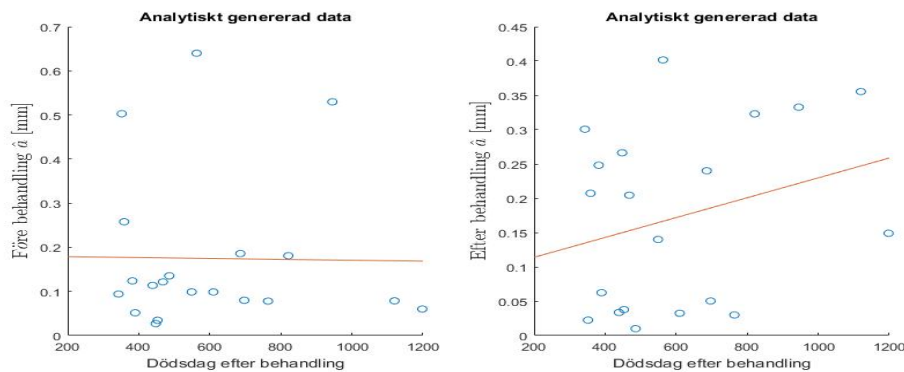
$$\frac{D_{v\text{före}} - D_{v\text{efter}}}{\max[D_{v\text{före}}, D_{v\text{efter}}]} \quad (25)$$

Motsvarande analys som redovisas ovan har även applicerats på D_v värdena, detta kan ses i figur 14. Det är värt att notera likheten mellan trenderna i den numeriska och analytiska datan för båda ovannämnda figurer. Vi kan se i figur 14 att cancerbehandlingen på mindre tumörer generellt ger ett dåligt resultat. Enbart små tumörer före behandling har ökande D_v värden, på resterande tumörer minskar värdena. Liknande trend kan noteras för \hat{a} värdena då större tumörer generellt resulterar i minskning av \hat{a} , medan mindre tumörer resulterar i motsatsen.



Figur 15: D_v relaterat till överlevnad

Det är även intressant att se hur utvalda parametrar är relaterade till överlevnad hos patienter. Då analytiska och numeriska metoder resulterar i samma trender har analytiska datan användas till figurerna. Som det kan ses i figur 15 har värdet på D_v före behandling större påverkan på patientens överlevnad än efter. Vilket kan indikera att behandlingen ger bättre resultat på tumören med högre motilitet.

Figur 16: \hat{a} relaterat till överlevnad

Gällande \hat{a} hittades en trend av högre värden efter behandling och längre överlevnad, vilket var något som vi inte hade förväntat oss. Tumörer med högre \hat{a} värden, borde vara mer aggressiva och därmed ha negativ korrelation på överlevnad. Men som tidigare nämnts har relativt lite data analyserats och en större undersökning måste göras för att kunna belysa om detta ytterligare.

5.3 Felkällor och avgränsningar

Vad gäller själva modelleringen (tillägg av en dödsterm som beskriver behandling) av Gerlees system, finns ett antal olika alternativ till det vi valt. Vi har valt att begränsa oss till systemet (3). Man kan tänka sig en olinjär apoptosterm på samma sätt som man kan tänka sig apoptosen som en funktion av tiden. Eftersom analysen varit nog komplicerad i det linjära fallet har vi valt att avgränsa oss från mer komplicerade alternativ.

Vidare har vi antagit en endimensionell modell med avseende på rummet. Detta motiveras utifrån att den flerdimensionella versionen bara betar sig annorlunda i början av cancers utveckling, alltså då vågansatsen inte heller är rimlig. Vad som händer är att diffusionstermerna i stil med $\frac{\partial^2 u}{\partial r^2}$ blir lika med Laplacianen av canceren Δu , som i sfäriska koordinater är lika med $\frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r}$ på grund av sfärisk symmetri. Vi noterar att då canceren breder ut sig (dvs. då $r \rightarrow \infty$) går $\Delta u = \frac{\partial^2 u}{\partial r^2} + \frac{2}{r} \frac{\partial u}{\partial r} \rightarrow \frac{\partial^2 u}{\partial r^2}$.

Det största antalet inskränkningar görs i kopplingen mellan modell och patientdata. Eftersom de numeriska beräkningarna tar en väldigt lång tid att utföra samtidigt som den analytiska målfunktionen (18) är icke-konvex (dvs. man hittar inte nödvändigtvis globala optima analytiskt) har vi valt att i båda fall låta $q_m = q_p = 20$, för att minska antalet frihetsgrader och antalet lokala lösningar. Vi motiverar detta genom att det främst är parametrarna D_v och $\alpha - \mu$ vi vill undersöka för att jämföra med motsvarande egenskaper hos Fisher. Detta är dock en stor inskränkning och möjligtvis en anledning till att resultaten blir svårtolkade. Det är närmast en självklarhet att patienternas verkliga parametrar har varierande q_m och q_p . Det som dessutom ledde Gerlees amerikanska kollega till Gerlees kopplade system var just det fenotypiska bytesmönstret, vars beteende karakteriseras av parametrarna q_m och q_p . Att låta q_m och q_p vara lika med samma konstant skulle möjligtvis kunna betyda att lösningen $d = p + m$ betar sig asymptotiskt så likt Fisher som möjligt, eftersom man långt i tiden kan se fenotypbytena som mindre betydande. Men en viktig skillnad som kvarstår är hur behandlingen bara påverkar de prolifererande cellerna i Gerlees modell.

Vi har dessutom bara tittat på en uppsättning mätningar efter behandling, av vilka det i många fall finns ett stort antal att välja mellan. Vi har valt datapunkter utifrån samma tumregel som användes för analysen med Fisher-ekvationen och den går ut på att välja den

första tidpunkten för uträkning av lutningen. Viktigt att nämna är att man skulle kunna tänka sig senare tidpunkter, speciellt tidpunkter då alla patienter genomgått samma behandling. I praktiken för den patientdata vi fick given är detta svårt och skakigt motiverat eftersom många patienter besökte olika sjukhus.

5.4 Avslutande kommentarer

Dataanalysen bygger på tillgången till rådata om endast 20 patienter, med helt skilda cancerkaraktärer vid diagnos. Det behövs en djupare och bredare dataanalys än den som har genomförts i detta projekt för att dra slutsatser om modellens riktighet. En sådan dataanalys kräver data från fler patienter. Speciellt krävs det data från flera patienter med liknande karaktärer på sin cancer vid diagnos, för att kunna göra relativa jämförelser av patienter.

Viktiga resultat från detta projekt är dock de numeriska metoderna och de analytiska uttrycken, samt associerade dataanalysmetoder. Dessa resultat kan användas i fortsatta studier om denna modell för hjärncancern glioblastom.

Bilagor

A Beskrivning & förenkling av ekvationer

A.1 Fisher's ekvation

Betrakta Fisher-ekvationen,

$$\frac{\partial u}{\partial t} = D\nabla^2 u + \rho u(1 - u).$$

Den beskriver hur cancercellernas densitet $u(x, t)$ varierar med tiden och rummet. D kallas diffusionskonstanten och är kopplad till Laplace-operatoren på u som i sin tur mäter en viss böjning av u med avseende på rumsvariabeln x . Med andra ord beskriver termen $D\nabla^2 u$ hur mycket cellmassan tenderar att fördelas ut i rummet från en viss given punkt genom ren diffusion.

Konstanten ρ är relaterad till tillväxthastigheten. Den multipliceras med faktorn $u(1 - u)$ som motsvarar just en tillväxt, eller proliferering, av cancers densitet. I detta fall är ekvationen normerad med en övre gräns så att $|u(1 - u)| \leq 1$. Det första u :et tar med antalet celler som kan dela sig i beräkningen, medan den andra faktorn $(1 - u)$ tar hänsyn till den maximala kapacitet som vävnaden kan uppbära. När cancers densitet är halva uppbärningskapaciteten anmärker vi att u har snabbast tillväxt. Tillväxten minskar däremot ju närmare u rör sig mot uppbärningskapaciteten eller mot noll. Detta modellerar väl det vi vet om naturlig populations-tillväxt.

A.2 Gerlee och Nelander's ekvationer

Vänd uppmärksamheten till de kopplade ekvationerna av Gerlee och Nelander,

$$\begin{cases} \frac{\partial p}{\partial t} = D_\alpha(1 - p - m)\frac{\partial^2 p}{\partial x^2} + \alpha p(1 - p - m) - (q_m + \mu)p + q_p m, \\ \frac{\partial m}{\partial t} = D_v((1 - p)\frac{\partial^2 m}{\partial x^2} + m\frac{\partial^2 p}{\partial x^2}) - (q_p + \mu)m + q_m p. \end{cases} \quad (2)$$

En första anmärkning är att summan av $p(x, t)$ och $m(x, t)$ beskriver cancers densitet. De prolifererande cellerna beskrivs av p medan de diffunderande eller migrerande cellerna beskrivs av m . Modellen utgår ifrån en stokastisk modell där cancercellerna befinner sig i ett rutnät och migrerar från ruta till ruta genom en Markov-process. Ekvationerna i systemet (2) är tidskontinuerliga varianter som beskriver sannolikheten att finna en cell av respektive karaktär i en viss ruta.

Notera vidare att vi den här gången har två diffusionskonstanter D_α och D_v , som hör till densiteterna p respektive m . Dessa konstanter bestäms enligt $D_\alpha = \alpha/2$ samt $D_v = v/2$, där α är cellens fördelningstakt och v motsvarar dess motilitet (rörlighet). Motsvarigheten till Fishers tillväxthastighet ρ är den ovannämnda konstanten α som då alltså bara infinner sig i ekvationen för p , det vill säga i ekvationen för de stillasittande prolifererande cellerna. Konstanterna q_m och q_p används i ekvationen för att representera cellövergång mellan de två fenotyperna. Dessa konstanter multipliceras sedan med densiteten för att få det totala flödet av celler mellan de två fenotyperna.

Till sist används dödskonstanten μ som svar för cancers apoptos, det vill säga takten som celler dör på. Det ska noteras att dödskonstanten antas vara samma för båda fenotyper av cellen. Dessutom sätter Gerlee $\mu = 0$ under motiveringen att det generellt gäller att $\mu \ll \alpha$ i sammanhang utan behandling [7]. Detta innebär att vi egentligen utgår ifrån det 2016

analyserade systemet,

$$\begin{cases} \frac{\partial p}{\partial t} = D_\alpha(1-p-m)\frac{\partial^2 p}{\partial x^2} + \alpha p(1-p-m) - q_m p + q_p m, \\ \frac{\partial m}{\partial t} = D_v((1-p)\frac{\partial^2 m}{\partial x^2} + m\frac{\partial^2 p}{\partial x^2}) - q_p m + q_m p. \end{cases} \quad (26)$$

Att byta variabel i ekvation (26) till $\tilde{x} = x/\sqrt{D_v}$ gör Gerlee m. fl. tidigt i sin analys [7]. Därifrån kan man genom vetskapen att $D_v \gg D_\alpha$ radera termen $\frac{D_\alpha}{D_v}(1-p-m)\frac{\partial^2 p}{\partial x^2}$. Detta lämnar oss, då vi i efterhand bytt tillbaka till $x = \tilde{x}\sqrt{D_v}$, med,

$$\begin{cases} \frac{\partial p}{\partial t} = \alpha p(1-p-m) - q_m p + q_p m \\ \frac{\partial m}{\partial t} = D_v((1-p)\frac{\partial^2 m}{\partial x^2} + m\frac{\partial^2 p}{\partial x^2}) - q_p m + q_m p \end{cases} .$$

Det visar sig att de olinjära diffusionstermerna $p\frac{\partial^2 m}{\partial x^2}$ och $m\frac{\partial^2 p}{\partial x^2}$ tillför väldigt lite inverkan på lösningen. Beloppet av differensen av lösningarna med eller utan dessa olinjära termer är uppåt begränsad av $8 \cdot 10^{-4}$. Vi sätter således $p\frac{\partial^2 m}{\partial x^2} = m\frac{\partial^2 p}{\partial x^2} = 0$, i den fortsatta analysen av systemet.

$$\begin{cases} \frac{\partial p}{\partial t} = \alpha p(1-p-m) - q_m p + q_p m, \\ \frac{\partial m}{\partial t} = D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p. \end{cases}$$

Slutligen gäller det att införa termer relaterade till behandling. Eftersom cellgifter samt strålningsbehandling angriper cancernas arvsmassa (och därigenom deras celledelning) är det rimligt att anta att deras påverkan på de migrerande cellerna är mild. Vi inleder således vår uppgift med följande system som utgångspunkt, som vi kallar det *analytiska systemet*,

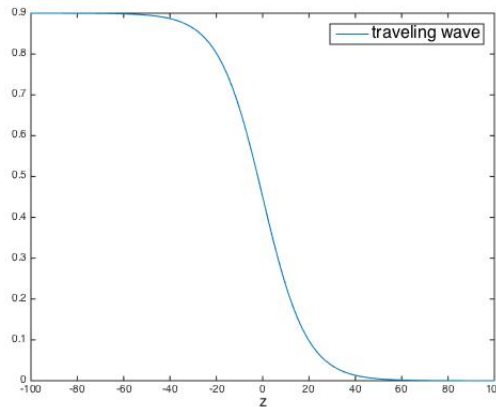
$$\begin{cases} \frac{\partial p}{\partial t} = \alpha p(1-p-m) - (q_m + \mu)p + q_p m, \\ \frac{\partial m}{\partial t} = D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p. \end{cases} \quad (3)$$

B Motivation för gräns- och randvärden

Vi vänder vår uppmärksamhet till det analytiska systemet (3),

$$\begin{cases} \frac{\partial p}{\partial t} &= \alpha p(1-p-m) - (q_m + \mu)p + q_p m, \\ \frac{\partial m}{\partial t} &= D_v \frac{\partial^2 m}{\partial x^2} - q_p m + q_m p. \end{cases}$$

Dessa ekvationer har färdande våglösningar som visats numeriskt (se sektion 3.1). Som läsaren kanske är medveten om ser en typisk färdande våg ut som Fig. 17, den rör sig i x-led utan att ändra form. Våra perturberade lösningar bör som första krav erhålla detta beteende.



Figur 17: Färdande våglösning.

Vidare undersöker vi beteendet av våra lösningar i randvärdena till grafen⁶. Dessa ligger i $z \rightarrow -\infty$, $z \rightarrow \infty$ och $z = 0$. I grafen ser vi att $p + m$ har horisontella asymptoter i både $-\infty$ och ∞ , vilket innebär att vi kan betrakta $p + m$ som konstanta om vi följer z ut långt nog åt båda hållen. Därav blir $\frac{\partial p}{\partial t} = -c \frac{\partial p}{\partial z} |_{z \rightarrow \pm \infty} = 0$ (och analogt för m samt om derivatan tas map. x) och därför kan vi räkna ut gränsvärdena av $p + m$ explicit från

$$\begin{cases} 0 &= \alpha p(1 - p - m) - (q_m + \mu)p + q_p m, \\ 0 &= -q_p m + q_m p. \end{cases}$$

Ur ekvationssystemet fås två par lösningar,

$$\begin{cases} p_1 = 0, & m_1 = 0 \\ p_2 = q_p \frac{1-\mu/\alpha}{q_p+q_m}, & m_2 = q_m \frac{1-\mu/\alpha}{q_p+q_m}. \end{cases}$$

Vi tar p_1 och m_1 som våra respektive randvärden i ∞ samt p_2 och m_2 som motsvarande i $-\infty$. Här ser vi dessutom från p_2 & m_2 ett krav på hur α och μ måste förhålla sig till varandra för att få rimliga resultat. Eftersom p_2 är en densitet har vi kravet $p_2 = q_p \frac{1-\mu/\alpha}{q_p+q_m} \geq 0$, som är ekvivalent med $\alpha - \mu \geq 0$. Dessutom är den totala andelen begränsad så att $p_2 + m_2 = q_p \frac{1-\mu/\alpha}{q_p+q_m} + q_m \frac{1-\mu/\alpha}{q_p+q_m} = 1 - \mu/\alpha \leq 1$. Samlat får vi,

$$0 \leq \mu \leq \alpha.$$

En viktig detalj vad gäller dataanalysen, är att den kräver en strängare övre begränsning på μ . Detta beskrivs även i kapitel 3.3. Begränsningen uppstår på grund av att en av patientfilernas två mätningar görs på 80%-nivån av andelen cancer per volymenhet. Om våra lösningar ligger under 0.8 blir de därför ej möjliga att upptäcka. Alltså för att, i skedet efter behandling haft sin inverkan på cancer, kunna analysera modellens passning på rådatan, låter vi

$$0 \leq \mu \leq 0.2\alpha.$$

Vidare kan vi nu motivera följande randvärden och egenskaper för den totala densiteten cancerceller $d(z) = p(z) + m(z)$,

$$\begin{cases} \lim_{z \rightarrow -\infty} d(z) &= 1 - \mu/\alpha = \frac{\hat{\alpha}}{\alpha}, \\ d(z)|_{z=0} &= \frac{1}{2} \frac{\hat{\alpha}}{\alpha}, \\ \lim_{z \rightarrow \infty} d(z) &= 0. \end{cases}$$

⁶I Gerlees rapport [7] sätter man speciella krav på f_0 och g_0 medan de andra blir noll, i till exempel $z = 0$. Vi har dessutom något annorlunda grundläggande krav.

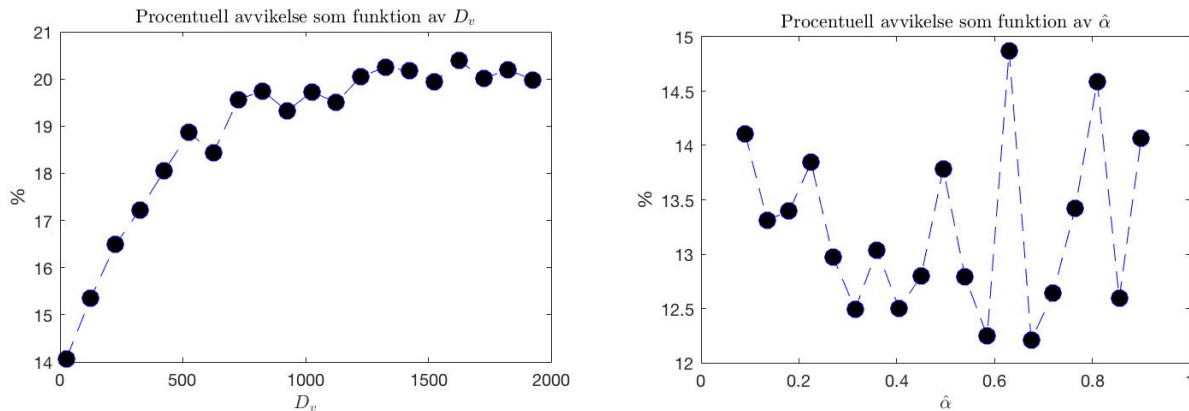
C Beskrivning av rådata

Rådatan vi använder är tagen från 20 patienter under behandling av glioblastom. Datan av en patient är en kronologiskt ordnad excell-fil som beskriver när en patient blir behandlad, vilken typ av behandling patienten genomgår, och som ger specifika information gällande behandlingen. Excell-filen inkluderar även information gällande radien av tumören. Radien är framtagen med hjälp av MR-röntgen. Tre sådana metoder är av särskilt intresse, nämligen s.k. T1GD, Flair och T2GD. Dessa metoder ger röntgenbilder av tumören och ger en kontrast mellan tumören över och under en given procenthalt av uppbärningskapaciteten, för hjärnan. T1GD ger kontrasten av tumören över 80% medan T2GD och Flair båda ger av tumören över 16%. Initialt tas alltid en röntgenbild med de tre känslighetsgraderna på patienten, efter detta är röntgenbilder sporadiskt tagna. Utifrån dessa bilder har sedan radien tagits ut.

Dessa data är givna till oss från en amerikansk kollega till vår handledare Philip Gerlee. Med hjälp av radien har cancers spridningshastighet räknats ut genom minsta kvadratmetoden. Vi har tagit friheten att använda de framtagna hastigheterna givna till oss av amerikanen.

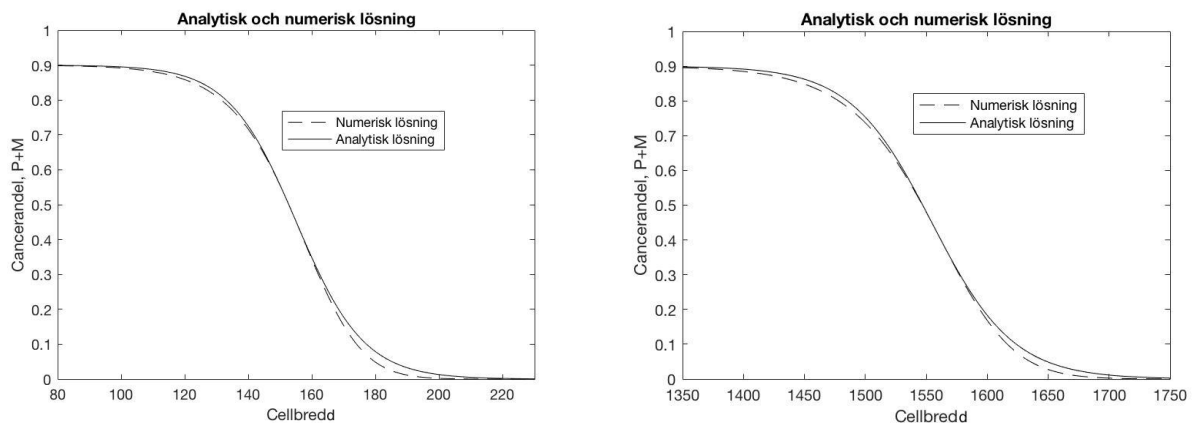
D Analytiska resultat i jämförelse med numeriska

En stor del av resultaten beror på de analytiska uttrycken i kapitel 4.1. Speciellt den analytiska perturbationslösningen (15) och den förenklade våghastigheten (11). Genom att jämföra de analytiska uttrycken mot numeriska simuleringar av lösningar kan en fingervisning erhållas på hur väl de analytiska uttrycken presterar. Detta är viktigt, speciellt eftersom de optimala $\hat{\alpha}$ och D_v för olika patienter redogjorda i kapitel 4.2 varierar från patient till patient. Om en större underliggande avvikelse faktiskt existerar har det en direkt inverkan på dataanalysens resultat.



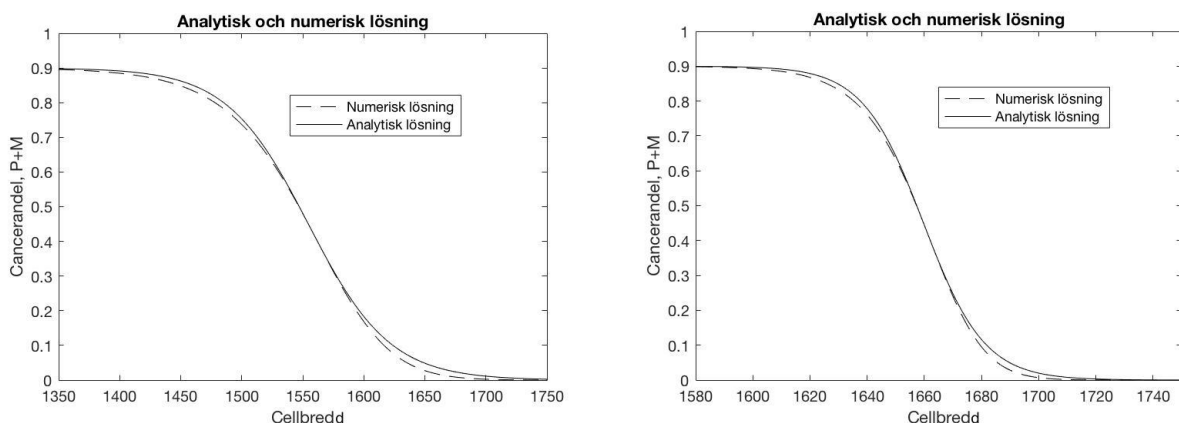
Figur 18: Maximal procentuell avvikelse mellan perturbationslösning och numerisk lösning, för $q = 20$. Till vänster: $\hat{\alpha} = 0.9$. Till höger: $D_v = 25$.

Den maximala procentuella avvikelsen mellan perturbationslösning och numerisk lösning i vågfronten, som funktion av ingående parametrar, illustreras i Fig. 18. Avvikelsen tenderar att öka då D_v ökar. En viktig anmärkning är att avvikelsen når över 20%, vilket är relativt högt. Dessutom finns en högre variation av avvikelsen för ökande $\hat{\alpha}$. Perturbationslösningen baseras huvudsakligen på perturbationsvariabeln $\hat{\alpha}$ och att den uppfyller $\hat{\alpha} \ll 1$, vilket när $\hat{\alpha}$ ökar blir mindre och mindre sant. Vilket kan förklara den till synes ökande variationen. För att vidare undersöka avvikelsen kan vi studera exempel på lösningar, för att åskådliggöra var den huvudsakliga avvikelsen befinner sig.



Figur 19: Numerisk och analytisk lösning för $\hat{\alpha} = 0.9$, $q = 20$. $D_v = 25$ till vänster, och $D_v = 2000$ till höger.

Den största avvikelserna existerar till synes i högra änden av vågfronten. Det verkar även existera en mindre avvikelse i övre, vänstra änden. Kom nu ihåg att rådata för patienter mäter tumörradien på känslighetsnivåerna 16% respektive 80% av den totala celldensiteten, och är även de känslighetsnivåerna som används i dataanalysmetoderna. Genom att betrakta ovanstående Fig. 19 existerar även de största felen vid dessa känslighetsnivåer, speciellt vid känslighetsnivån 16%. En naturlig konsekvens är att den analytiska lutningen (16) kan ge missvisande lutningsvärden.



Figur 20: Numerisk och analytisk lösning för $D_v = 25$, $q = 20$. $\hat{\alpha} = 0.09$ till vänster, och $\hat{\alpha} = 0.9$ till höger.

En likadan trend följer vid variation av $\hat{\alpha}$, vilket illustreras i Fig. 20. Generellt kan det uttryckas att den analytiska dataanalysmetoden resulterar i fel.

För att undvika fel är det till fördel att använda den numeriska dataanalysmetoden. Alternativt kan högre ordningens termer inkluderas i härledningen av perturbationslösningen, vilket medför mer komplicerade uträkningar och uttryck. Däremot, med dagens programvaror som Mathematica blir det mindre omständigt. Hur stort fel det är i praktiken att använda den analytiska dataanalysmetoden är en bra frågeställning. Av resultaten i 4.2 ger de nästintill identiska resultat, vilket kan indikera att avvikelserna inte har någon nämnvärd effekt i praktiken.

E Jämförelse av numeriskt kontra analytiskt minimeringsproblem

För en djupare analys av resultaten från minimeringsproblemen kan optimala $\hat{\alpha}$, D_v från båda metoderna tas i relation till varandra, för båda behandlingsfaserna. Skulle metoderna ge iden-

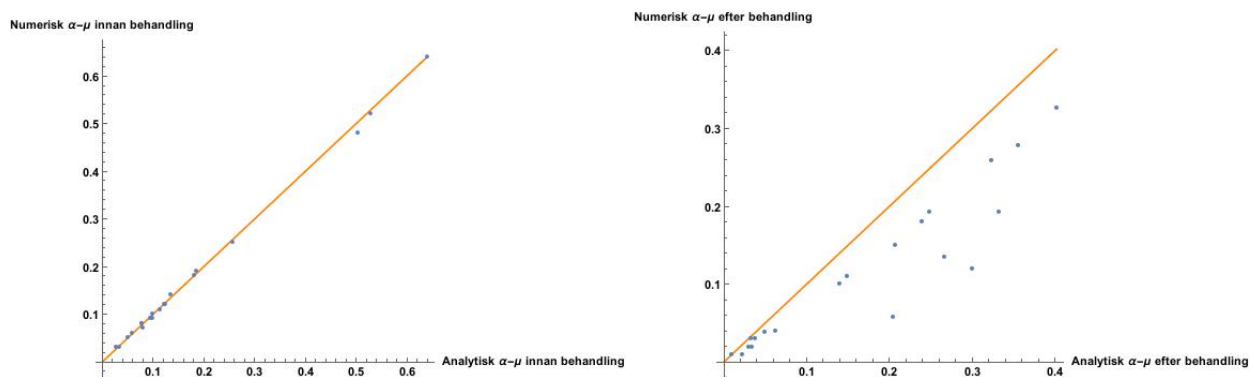
tiska resultat bör punkter av optimal parameter via numerisk metod relativt analytisk metod för patienter ligga utefter linjen $y = x$. Detta skulle indikera att metoderna är ekvivalenta, med avseende på dataanalysresultat. Innan vi går in på den djupare analysen påminns läsaren av dels det numeriska minimeringsproblemet

$$\underset{f_i \in \mathcal{F}}{\text{minimera}} \quad M_k(f_i), \quad (19)$$

och dels det analytiska minimeringsproblemet

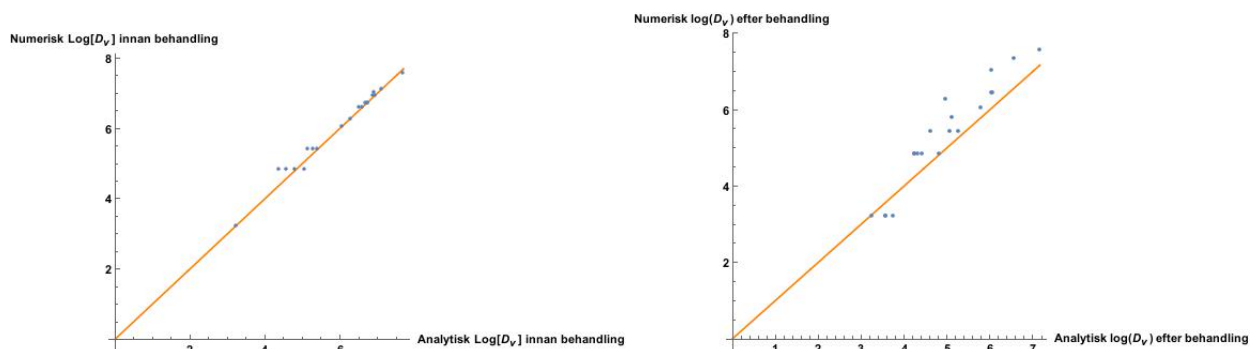
$$\begin{aligned} &\underset{(\alpha, D_v, \mu) \geq (0, 25, 0)}{\text{minimera}} && E(\alpha, D_v, \mu), \\ &\text{så att} && 5\mu < \alpha. \end{aligned} \quad (21)$$

E.1 Jämförelse av optimala parametrar för dataanalysmetoderna



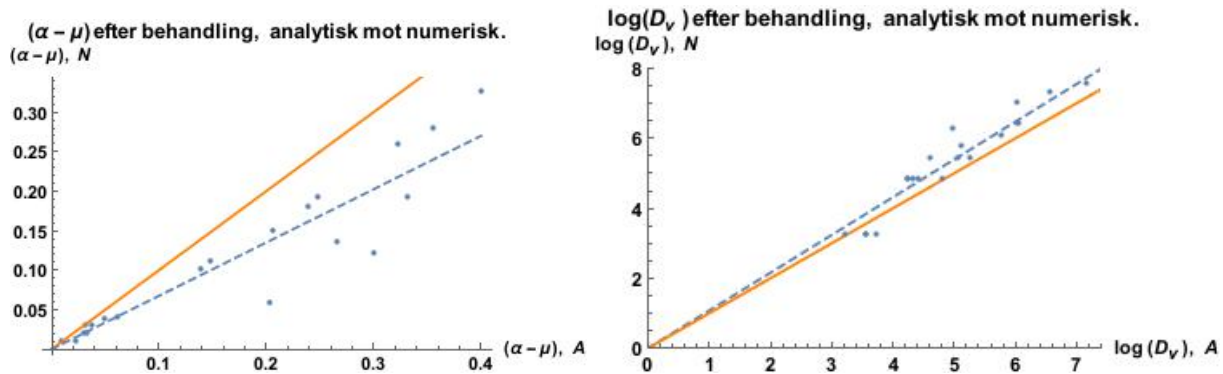
Figur 21: Jämförelse av optimala $\hat{\alpha}$ för numerisk och analytisk dataanalysmetod. Till vänster: före behandling. Till höger: efter behandling

Resultaten av de två dataanalysmetoderna före behandling har nästintill identiska värden, vilket illustreras till vänster i Fig. 21. Kom nu ihåg att i den analytiska dataanalysmetoden kunde de optimala parametrarna $\hat{\alpha}$, D_v före behandling räknas ut explicit, vilket visades i 3.3.2. Av denna anledning är likheten av $\hat{\alpha}$ före behandling inte oväntat. Det existerar dock en mindre avvikelse för en mindre andel patienter. Detta kan ses som en konsekvens av existerande avvikelse mellan numeriska och analytiska lösningar diskuterat i bilaga D. I detta fall är denna avvikelse till synes försumbar. En större avvikelse existerar dock efter behandling, vilket indikeras till höger i Fig. 21. Den procentuella avvikelsen varierar från patient till patient, men trenden att den analytiska metoden överestimerar $\hat{\alpha}$ är tydlig.



Figur 22: Jämförelse av optimala D_v för numerisk och analytisk dataanalysmetod. Till vänster: före behandling. Till höger: efter behandling

Beteendet av D_v är annorlunda än dess $\hat{\alpha}$ motpart, som illustreras i Fig. 22. I fasen före behandling är uttrycken nästintill identiska. Däremot tenderar D_v att underestimeras av den analytiska metoden. Denna avvikelse är däremot betydligt mindre än dess $\hat{\alpha}$ motpart. För vidare analys av avvikelsen i fasen efter behandling för $\hat{\alpha}$, D_v utförs en linjär dataanpassning. Detta är främst motiverat av att överestimationer och underestimationer i Fig. 21 respektive Fig. 22 har en till synes linjär form.



Figur 23: Optimala parametrar via analytisk metod (A) relativt numerisk metod (N) efter behandling. Streckad linje är linjär dataanpassning.

Den linjära anpassningen är bäst för D_v . Vidare följer den linjära anpassningen ekvationen $y = 1.0788x$, vilket är nästintill identiskt med $y = x$. Detta, återigen, indikerar prestationen av den analytiska dataanalysmetoden med avseende på att estimerar D_v . För $\hat{\alpha}$ har den linjära anpassningen en ekvation enligt $y = 0.6748x$, vilket tyder på ett sämre beteende än den för D_v . Avvikelsen av värden med avseende på $\hat{\alpha}$ mellan dataanalysmetoderna medför att dataanalysmetoderna inte är direkt ekvivalenta. Däremot är den relativa avvikelsen, totalt sett, på en god nivå.

Den största anledningen för avvikelse mellan resultatet från de olika minimeringsproblemen skulle vara den färdande vågansatsen som bara det analytiska minimeringsproblemet bygger på. Vågansatsen är ett tidsasymptotiskt antagande som vid användning antar att patienten redan haft tumören under en viss tid innan diagnos. Dessutom, vilket visats i bilaga D, medför det även avvikelser att använda analytiska uttryck kontra numeriska metoder. Det analytiska minimeringsproblemet baseras huvudsakligen på analytiska resultat, vilket i sin tur naturligt medför fel.

Efter behandling har vi ytterligare en faktor att ta hänsyn till. Vår lösare för det analytiska minimeringsproblemet hittar lokala minima på grund av målfunktionens icke-triviala karaktär. Alltså skulle vi, givet att vågansatsen är rimlig, ändå kunna observera diskrepans mellan metodernas lösningar. Dessutom kan det vara så att den valda diskretiseringen i implementationen av den numeriska dataanalysmetoden inte är tillräcklig. Utan att det istället krävs en finare partition för att ge bättre dataanpassningar. För att kringgå detta kan interpolationsteknik användas för att estimerar värden c_{f_i} , s_{f_i} som ligger mellan diskretiseringen.

E.2 Datorintensivitet och effektivitet av minimeringsproblemen

Anpassningsmetoden baserad på numeriska metoder enligt Ekv. (19) är en väldigt datorintensiv algoritm. För att se detta, ponera att $\alpha \in [0, 1]$ diskretiseras med en steglängd på 0.01, och att $D_v \in [25, 2000]$ diskretiseras med en steglängd på 100. Att simulera en godtycklig numerisk lösning och beräkna dess våghastighet och lutning tar ungefär 30 sekunder. Detta resulterar i en total simulationstid på ungefär 17 timmar. I fasen efter behandling, där μ inte nödvändigtvis är

noll, kan den totala simulationstiden sträcka över flera dagar, beroende på ingående bestämd noggrannhet. Fördelen är att metoden, av dess konstruktion, finner globala optimum, utan att behöva implementera avancerade optimeringsalgoritmer. Minimeringsproblemet (19) blir i praktiken ekvivalent med att hitta det minsta elementet i en matris, vilket är en trivial process.

Minimeringsproblemet (21) kräver ej samma tidsåtgång som för (19). Problemet (21) tar i praktiken mindre än ca. 10 sekunder att lösa, vilket är väldigt tidseffektivt. Däremot är lösandet av det analytiska minimeringsproblemet (21) en relativt komplicerad process. Dessutom, som konsekvens av (32) och (33) icke-triviala karaktärer, blir det ett problem att kunna avgöra lokal kontra global optimalitet av erhållna lösningar, vilket i praktiken inte behövs för (19) där global optimalitet är garanterad. Slutligen baseras metoden på vågansatsen, vilket (19) är helt oberoende av. Vilket potentiellt medför avvikelser om vågansatsen inte stämmer.

Det visar sig även att de optimala lösningarna till dataanalysmetoderna som funktion av optimala parametrar redogjorda i 4.2 ger utmärkt små målfunktionsvärden. Målfunktionerna (18) respektive (20) som funktion av den optimala lösningen för en godtycklig patient erhåller värden mindre än storleksordning 10^{-6} . Detta medför att de optimala parametrarna för varje patient både före och efter behandling är väldigt bra anpassade.

E.3 Optimala parametrar före behandling samt explicit målfunktion

Här vill vi först härleda uttrycken för optimala parametrar $(\alpha_k, D_{v,k})$ för en godtycklig patient med index k före behandling. Låt nu c_k och s_k denotera hastighet respektive lutning av tumörradien för en godtycklig patient med index k . Vi approximerar nu c och s med c_k respektive s_k . Antag vidare att $q_m = q_p = q$ har ett fixt värde, vilket medför att $\nu = 1$.

Av detta ges uttrycken för vidare analys av

$$c_k = \sqrt{\frac{D_v \hat{\alpha}}{1 - \hat{\alpha}/(3q)}}, \quad q = q_m = q_p, \quad (27)$$

$$s_k = -\frac{0.16 - 0.8}{x(0.16, t, \alpha, \mu, D_v) - x(0.8, t, \alpha, \mu, D_v)}, \quad (28)$$

$$x(d, t, \alpha, \mu, D_v) = ct + \frac{c(1 + \nu)}{\hat{\alpha}} \log\left(\frac{\hat{\alpha}}{d\alpha} - 1\right). \quad (29)$$

Av antagandet $\mu = 0$ i fasen före behandling noterar vi att med (29) i hand kan nämnaren i (28) ekvivalent skrivas

$$x(0.16, t, \alpha, 0, D_v) - x(0.8, t, \alpha, 0, D_v) = \frac{2c_k}{\alpha} \log(21), \quad (30)$$

vilket vidare insatt i (28) ger att α kan lösas ut explicit enligt

$$\alpha_k = \frac{2s_k c_k \log(21)}{0.64}, \quad (22)$$

där ett index för patient k har införts. Med liknande metodik, genom att kvadrera (11), lösa ut för D_v och slutligen sätta in (22) erhålles

$$D_{v,k} = \frac{c_k^2}{\alpha_k} \left(1 - \frac{\alpha_k}{3q}\right), \quad (23)$$

där, precis som förut, ett index för patient k har införts. Härledningen av de explicita uttrycken för α_k , $D_{v,k}$ är nu klart.

För att hitta ett explicit uttryck på målfunktionen i minimeringsproblemet för den analytiska dataanalysmetoden noteras det att

$$x(0.16, t, \alpha, \mu, D_v) - x(0.8, t, \alpha, \mu, D_v) = \frac{2c_k}{\hat{\alpha}} \log\left(\frac{0.8(\hat{\alpha} - 0.16\alpha)}{0.16(\hat{\alpha} - 0.8\alpha)}\right), \quad (31)$$

efter att dels använda Ekv. (29) och dels utföra diverse omskrivningar. Notera även att högerleden från (28) och (27) kan tolkas som funktioner av tre variabler; α , D_v och μ . Det vill säga, ekvationerna kan skrivas som

$$s(\alpha, D_v, \mu) = \frac{0.64}{2c_k} \cdot \frac{\hat{\alpha}}{\log\left(\frac{0.8(\hat{\alpha} - 0.16\alpha)}{0.16(\hat{\alpha} - 0.8\alpha)}\right)}, \quad (32)$$

$$c(\alpha, D_v, \mu) = \sqrt{\frac{D_v \hat{\alpha}}{1 - \hat{\alpha}/(3q)}}, \quad (33)$$

där uttrycket (31) har använts. Eftersom målfunktionen för minimeringsproblemet lyder enligt

$$E(\alpha, D_v, \mu) = \frac{(s_k - s(\alpha, D_v, \mu))^2}{s_k^2} + \frac{(c_k - c(\alpha, D_v, \mu))^2}{c_k^2}, \quad (20)$$

ser vi att med Ekv. (32) och Ekv. (33) följer det att målfunktionen (20) kan uttryckas explicit.

F Kod: Matlab och Mathematica

För en fullständig tillgång till Mathematica och Matlab kod underliggande numeriska metoder hänvisas läsaren till

https://www.dropbox.com/s/4hccs9xdxuj4hxp/Mathematica_kod_Glioblastom.zip?dl=0,
https://www.dropbox.com/s/j0hg7648h6ujk5b/Matlab_kod_Glioblastom.zip?dl=0.

Referenser

- [1] Cancer.gov, "Adult Central Nervous System Tumors Treatment (PDQ)–Health Professional Version", "<https://www.cancer.gov/types/brain/hp/adult-brain-treatment-pdq#section/all>", Februari 2017.
- [2] American Brain Tumor Association, "Glioblastoma (GBM)", "<http://www.abta.org/brain-tumor-information/types-of-tumors/glioblastoma.html?>", Februari 2017.
- [3] Tracqui P, Cruiwagen GC, Woodward DE, Bartoo GT, Murray JD, et al. (1995) A mathematical model of glioma growth: the effect of chemotherapy on spatio-temporal growth. *Cell Prolif* 28: 17–31.P. TracquiGC CruiwagenDE WoodwardGT BartooJD Murray1995A mathematical model of glioma growth: the effect of chemotherapy on spatio-temporal growth. *Cell Prolif*281731
- [4] Woodward DE, Cook J, Tracqui P, Cruiwagen GC, Murray JD, et al. (1996) A mathematical model of glioma growth: the effect of extent of surgical resection. *Cell Prolif* 29: 269–288.DE WoodwardJ. CookP. TracquiGC CruiwagenJD Murray1996A mathematical model of glioma growth: the effect of extent of surgical resection. *Cell Prolif*29269288
- [5] Swanson K, Bridge C, Murray J, Alvord E (2003) Virtual and real brain tumors: using mathematical modeling to quantify glioma growth and invasion. *J Neurol Sci* 216: 1–10.
- [6] Gerlee, Philip and Nelander, Sven, The Impact of Phenotypic Switching on Glioblastoma Growth and Invasion, *PLoS Computational Biology* (1553-7358). Vol. 8, issue 6 (2012).
- [7] Gerlee P, Nylander S, Travelling wave analysis of a mathematical model of glioblastoma growth *Mathematical Biosciences* (0025-5564). Vol. 276 (2016).
- [8] William H. Press and Saul A. Teukolsky and William T. Vetterling and Brian P. Flannery (2007), *Numerical Recipes 3rd Edition: The Art of Scientific Computing*, Cambridge University Press.
- [9] J.D. Murray (2011) *Mathematical Biology II: Spatial Models and Biomedical Applications (Interdisciplinary Applied Mathematics)* (v. 2), Springer.
- [10] Kolmogorov et. al. Articles on Partial Differential Equations: Studies of the Diffusion with the Increasing Quantity of the Substance; Its Application to a Biological Problem.