



Hur betraktar vi en tavla?

En spatio-temporal matematisk modell för ögats attraktionsområden på en tavla

What do we see when we look at art?

A spatio-temporal mathematical model of the eye's regions of attraction on a painting

Kandidatarbete inom civilingenjörsutbildningen vid Chalmers

Examensarbete för kandidatexamen i matematisk statistik vid Göteborgs universitet

Simon Alacam

Nils Grimbeck

Måns Gustafsson

Simon Nordlund

Hur betraktar vi en tavla?

En spatio-temporal matematisk modell för ögats attraktionsområden på en tavla

Simon Alacam
Nils Grimbeck
Måns Gustafsson
Simon Nordlund

© Författarna, 2024

Institutionen för Matematiska vetenskaper, Chalmers Tekniska Högskola & Göteborgs Universitet
Göteborg, 2024

Omslagsbild: Claude Monet, *La terrasse à Sainte-Adresse*. Public domain. Källa: The Metropolitan Museum of Art.

Erkännanden, dedikationer och liknande personliga uttalanden återspeglar författarens egna åsikter.

Hur betraktar vi en tavla?

En spatio-temporal matematisk modell för ögats attraktionsområden på en tavla

Examensarbete för kandidatexamen i matematisk statistik vid Göteborgs universitet

Simon Alacam

Kandidatarbete i matematik inom civilingenjörsprogrammet Teknisk matematik vid Chalmers

Nils Grimbeck Simon Nordlund

Kandidatarbete i matematik inom civilingenjörsprogrammet Maskinteknik vid Chalmers

Måns Gustafsson

Handledare: Aila Särkkä Institutionen för Matematiska vetenskaper

Institutionen för Matematiska vetenskaper
CHALMERS TEKNISKA HÖGSKOLA
GÖTEBORGS UNIVERSITET
Göteborg, Sverige 2024

Förord

I detta kandidatarbete inom matematisk statistik ämnar vi att utforska hur man matematiskt kan modellera ögats rörelsemönster på en tavla. Arbetet har skrivits under handledning av Aila Särkkä, institutionen för matematiska vetenskaper vid Chalmers och Göteborgs universitet.

Vi vill först och främst uttrycka vår tacksamhet till vår handledare Aila Särkkä som kontinuerligt stöttat och väglett oss under arbetets gång.

Under arbetets gång har det förts en veckovis dagbok innehållande individuella och gemensamma bidrag. Ansvar för denna har skötts genom ett rullande schema. Varje medlem har även fört en tidslogg för antalet timmar som investerats i arbetet. Nedan finns även en tabell där huvudförfattare för respektive del står listad. Samtliga medlemmar har även reviderat hela rapporten. Det finns också en tabell för respektive medlems bidrag till modelleringsarbetet, vars källkod finns tillgänglig via <https://github.com/NilsGrimbeck/KandidatarbeteMVEX11-24-05>.

Avsnitt	Huvudförfattare
Förord Populärvetenskaplig presentation Sammandrag	Simon N & Nils Måns & Simon N Simon A
Inledning Syftesbeskrivning och avgränsningar Ögonrörelsedata	Nils Nils Nils
Data från universitetet i Jyväskylä Betraktarens blick förändras över tid Etiska och samhällsliga aspekter	Nils, Simon A Nils Måns
Metoder och inledande dataanalys Spatiala mönster av fixeringar Punktprocesser och intensitet	Nils Simon A Simon A
Kärnskattning Fixeringspunkters förskjutning relativt förväntade attraktionsområden Fördelningar för fixeringstid, sackadtid, och sackadlängd	Simon A Nils Simon A
Modellering Definition av ROI:er genom mean-shift-klustring Markovkedjor för att återskapa fokusförflyttningar mellan ROI:er	Måns & Simon N Nils (om mean-shift), Måns & Simon N (sista stycket & figurer) Nils & Simon A
Punktgenerering Statistikor Resultat och diskussion	Måns & Simon N Måns & Nils Nils
Slutsatser Appendix A Appendix B	Nils Nils & Simon A Simon A
Appendix C Appendix D Appendix E	Nils Simon N Nils

Modelldel	Implementerat av	Revision
Analys av fördelningar Mean-shift-klustring Markovkedjemodell	Simon A Nils Nils	Nils Simon A Simon A
Övergångssannolikheter Definition av ROI:er Punktgenerering	Simon A Måns Måns, Nils, Simon N	Nils Nils Simon A
Fixeringspunktsmönstermodell Intensitetsytor Statistikor	Nils Simon A, Måns, Simon N Nils, Måns, Simon N	Nils
Visualisering av ROI:er	Måns, Simon N	

Populärvetenskaplig presentation

När vi människor betraktar tavlor så rör sig blicken instinktivt och näst intill omedvetet över tavlans olika delar. Vissa områden och komponenter lockar blicken mer och vissa detaljer får oss att stanna upp för att titta lite extra. Hur våra ögon rör sig över tavlor kan säga mycket om vad den innehåller och ge insikter kring betraktarens kognitiva processer och egna preferenser. Frågan är vad som faktiskt lockar ögat och om det går att återfinna ett mönster hos olika personer. Går det att återskapa dessa rörelser med en modell genom att undersöka de processer som beskriver ögats rörelser? I detta arbete har ögats rörelser på en tavla undersökts och, med hjälp av matematisk statistik, modeller skapats för att återskapa hur ögat rör sig.

Ögats rörelser kan i huvudsak delas upp i fixeringar och sackader som sker i en alternerande sekvens. En fixering innebär att ögat är stilla och sackader är de korta ryck eller hopp som görs mellan fixeringar. För att undersöka ögonrörelser används idag avancerad teknik för att spåra ögats rörelser och position, så kallad eye-tracking. Eye-tracking utgår oftast från en kamera som spelar in ögats rörelser och använder algoritmer som översätter informationen från kameraflödet till hanterbara datapunkter. Datapunkterna är ofta koordinater, som anger var på det observerade objektet en person fäster sin blick och tidsstämplor som anger när fixeringarna ägde rum och hur länge de varade.

Detta arbete använder data från en tidigare studie där man jämfört ögonrörelser hos studenter från universitetet i Jyväskylä, Finland. I den studien betraktade studenterna sex olika tavlor i tre minuter vardera, medan eye-tracking-utrustningen spelade in ögonrörelserna. Vi undersöker ögonrörelsedata från en specifik tavla, nämligen Claude Monets *Terrassen vid Sainte Adresse* (1867) som syns i figuren nedan. Vi valde även att utgå ifrån data för de första trettio sekunderna för vår modell.

Grundidén för modellen utgår från premissen att ögat tenderar betrakta vissa delar av tavlan mer intensivt än andra. För Claude Monets *Terrassen vid Sainte Adresse* kan man ställa sig frågan, var kommer blicken dras? Är det på den franska flaggan, paret vid räcket, mannen och kvinnan som sitter ner eller kanske någon av båtarna vid horisonten? Ett sådant område med många fixeringspunkter benämns som attraktionsområde. När man fastställt tavlans attraktionsområden gäller det att bestämma exakt var inom detta område fixeringen bör hamna.



Claude Monets *Terrassen vid Sainte Adresse*. Public domain. Källa: The Metropolitan Museum of Art.

Sammantaget fungerar modellen som följande; först definieras alla attraktionsområden, sedan beräknas sannolikheten att blicken går från ett specifikt attraktionsområde till ett annat. Slutligen bestäms var fixeringen mest sannolikt bör hamna inom ett attraktionsområde eller i resterande del av tavlan. På så sätt kan vi efterlikna hur ögat rör sig på tavlan. Resultatet från modellen jämförs sedan med faktiska data med några statistiska mått.

Förhoppningen med modellen är att få en inblick i hur vi betraktar en tavla och utveckla nya verktyg för ögonrörelseanalys. Vår modell lyckas efterlikna ögats rörelser för denna tavla, men för att få en uppfattning om hur väl modellen efterliknar ögonrörelser generellt krävs att den valideras på andra tavlor och för ett större antal personer.

Sammandrag

I detta arbete undersöks ögonrörelser på tavlan *Terrassen vid Sainte Adresse* (1867) av Claude Monet, med hjälp av en stokastisk modell som simulerar hur betraktarens blick förflyttas på tavlan. Modellen baseras på ögonrörelsedata från 20 försökspersoner som observerade tavlan i tre minuter, där modelleringen begränsats till de första 30 sekunderna. I analysen används sackader och fixeringar för att beskriva ögonrörelser. Sackader refererar till snabba ögonrörelser mellan fixeringspunkter, medan fixeringar avser perioder när ögonen fokuserar på en specifik punkt.

Inledningsvis används mean-shift-klustring för att identifiera de intressanta områdena på tavlan. Därefter tillämpas Markovkedjor för att modellera övergångarna mellan dessa områden, och slutligen används en skattning av intensiteten av fixeringsmönstret och sackadlängderna för att positionera fixeringarna inom varje attraktionsområde. Huvudsyftet med arbetet är att utveckla en modell som återskapar fixeringspunktmönster med hänsyn till fokusförflyttningar mellan tavlans olika attraktionsområden och därmed introducera nya verktyg för ögonrörelseanalys. Analys av attraktionsområden genom klustring av fixeringspunkter är inget nytt inom konstvetenskaplig analys av betraktarens fokus men det har inte tidigare kombinerats med punktprocesser. För att bedöma hur väl modellen återskapar faktiska ögonrörelser, utvärderas modellen genom två statistikor, den totala sackadlängden och det konvexa höljet av fixeringspunkter. Resultaten visar att den utvecklade modellen framgångsrikt kan återskapa de grundläggande mönstren i fixeringspunktdata.

Abstract

This thesis examines eye movements on the painting *Terrace at Sainte Adresse* (1867) by Claude Monet using a stochastic model to simulate how an observer's gaze moves on the painting. The study is based on eye movement data from 20 participants who observed the painting for three minutes, where the modelling is limited to the first 30 seconds. The analysis utilizes saccades and fixations to describe eye movements. Saccades refer to rapid eye movements between fixation points, while fixations refer to periods when the eyes focus on a specific point.

Initially, mean-shift clustering is used to identify the regions of interest on the painting. Subsequently, Markov chains are applied to model the transitions between these regions, and finally an estimation of the intensity of the fixation pattern and the saccade lengths is used to position the fixations within each region of interest. The main purpose of the thesis is to develop a model that recreates the fixation patterns with regard to the focus transitions between different regions of interest on the painting and to introduce new tools for eye tracking analysis. Analysis of areas of attraction by clustering of fixation points is not new in art analysis of the observer's focus but it has not previously been combined with point processes. To assess how well the model recreates actual eye movements, the model is evaluated using two statistics, the total saccade length and the convex hull of fixation points. The results show that the developed model successfully recreates the fundamental patterns in fixation data.

Innehåll

1 Inledning	1
1.1 Syftesbeskrivning och avgränsningar	1
2 Ögonrörelsedata	3
2.1 Data från universitetet i Jyväskylä	3
2.2 Beträktarens blick förändras över tid	4
2.3 Etiska och samhällseliga aspekter	4
3 Metoder och inledande dataanalys	5
3.1 Spatiala mönster av fixeringar	5
3.1.1 Punktprocesser och intensitet	5
3.1.2 Kärnskattning	5
3.2 Fixeringspunkters förskjutning relativt förväntade attraktionsområden	6
3.3 Fördelningar för fixeringstid, sackattid, och sackadlängd	7
4 Modellering	11
4.1 Definition av ROI:er genom mean-shift-klustring	11
4.2 Markovkedjor för att återskapa fokusflyttningar mellan ROI:er	12
4.3 Punktgenerering	13
4.4 Statistikor	13
5 Resultat och diskussion	15
6 Slutsatser	18
Referenser	19
A Individuella fixeringspunktmönster	i
B Statistiska skattningar och diagnostiska plottar	iv
C Glasögons påverkan på fixeringspunktmönster	viii
D ROI:er för samtliga individer	ix
E Realisationer av modell B	xiii

1 Inledning

Redan i slutet av 1800-talet upptäckte Javal, Brown, och Hering separat [1] att ögat rör sig i korta ryck, kallade sackader, mellan vilopunkter, fixeringar, där ögat är relativt stilla [1, 2, 3] och ögonrörelser kan därför representeras som en alternerande sekvens av sackader och fixeringar [2, 3]. Vidare genereras tre till fyra sackader per sekund, under vilka synen hämmas, och de är nödvändiga för att förflytta centralgropen, ögats detaljcentrerande område på näthinnan [4], till relevanta områden i synfältet eftersom skärpan snabbt avtar mot periferin [3]. Dessa rörelser är ett resultat av en kognitiv process och har därför analyserats inom en rad områden så som psykologi, språk- och not-läsning, marknadsföring, och konstvetenskap [2]. I detta arbete är vi intresserade av ögonrörelser på tavlor.

Redan 1935 publicerades *How People Look at Pictures: a study of the psychology of perception in art*, en av de första systematiska eye-tracking studierna, av Guy Buswell där 200 försökspersoners ögonrörelser dokumenterades under betraktande av fotografier och målningar [5]. Buswell dokumenterade knappt 2000 fixeringspunktmönster på 55 fotografier vilka analyserades kvalitativt snarare än statistiskt. Enligt Nicholas J. Wade [5] drar Buswell slutsatserna att fixeringspunktmönster tenderar att klustra kring intresseområden i fotografiet, snarare än att drivas av slumpen, vilket även senare kvantitativa studier bekräftat [3, 6]. Vidare drar Buswell slutsatsen att olika individers fixeringsmönster tenderar att vara mer lika initialt än under observationens fortskridande samt att fixeringslängderna tenderar att öka med tiden och att skillnaden mellan personer med konstnärlig träning och noviser är liten [5].

På senare tid finns det en debatt kring vilka processer som i högst utsträckning definierar ögonrörelser under en estetisk upplevelse, delvis på grund av hur olika studier lägger olika vikt vid bottom-up och top-down processer. Massaro m.fl. [6] beskriver dessa processer som beroende av komposition, balans, färg, symmetri, etc. respektive åskådarens kulturella bakgrund, intresse, och erfarenhet. Ögonrörelserna visar därför även åskådarens selektion, och är en sammanställning av vad som observerats och åskådarens intresse [3]. Fixeringspunkternas klustring visar därmed vilka områden som skapar mening och eftersom detta delvis beror på uppgiften, så ändras även åskådarens ögonrörelser mellan olika uppgifter [3, 6, 7]. Sådana områden kommer fortsatt benämnas attraktionsområden och förkortas ROI från engelskans Region Of Interest.

Ögonrörelser kan avslöja mycket kring åskådarens visuella upplevelse, även om variationen är stor mellan olika tavlor, uppgifter, och individer. De senaste 10 åren har det gjorts försök att analysera detta som en realisering av en punktprocess eftersom fixeringspunkterna kan betraktas som ett spatio-temporalt punktmönster, se bland annat Barthelmé m.fl. [3] och Ylitalo m.fl. [2]. I den senare studien presenteras även en enkel spatio-temporal punktprocessmodell i vilken de slumpar en startpunkt från tavlans intensitetsyta genom att normalisera intensiteten till en bivariat sannolikhetsfördelning varpå en sackadlängd l slumpas från en gammafördelning. Därefter väljs nästa fixeringspunkt från den betingade sannolikhetsfördelningen av intensitetsytan på avstånd l från nuvarande punkt. Fixeringstiden dras sedan från en fördelning för denna och processen upprepas för önskat tidsintervall.

För att validera modellen använder de tre statistikor: det konvexa höljet, arean som täcks av diskar runt fixeringspunkterna, och den totala längden av sackaderna. De två första beskrivs väl av modellen medan den totala sackadlängden tenderas att underestimeras och de drar slutsatsen att de långa sackaderna som sker i samband med fokusförflyttning mellan attraktionsområden i tavlan inte tas hänsyn till i modellen.

1.1 Syftesbeskrivning och avgränsningar

Detta projektet syftar till att med avstamp i Ylitalo m.fl.s modell utveckla en stokastisk modell som även återskapar påverkan på fixeringsmönstret från fokusförflyttningar mellan tavlans ROI:er. Genom att förenklat anta att nästa ROI enbart beror på den föregående ämnar vi modellera fokusförflyttningarna med Markovkedjor genom att skatta kedjans övergångsmatris från ögonrörelsedata och i respektive ROI betrakta en realisering av en punktprocess för att generera ett fixeringspunktmönster vilket implementeras på två olika sätt. I modell A föreslås



Figur 1: Claude Monets Terrassen vid Sainte Adresse. Public domain. Källa: The Metropolitan Museum of Art.

nästa punkt likformigt i respektive ROI medan punktens placering i modell B föreslås på ett avstånd l som skattas från ögonrörelsedata om nästkommande punkt är i samma ROI. Vid övergång mellan ROI:er föreslås punkten så som i modell A.

Eftersom modellen endast baseras på fixeringspunktsdata, och inte direkt på tavlans komponenter, kan detta arbete inte besvara hur bottom-up respektive top-down processer samverkar. Fokus ligger därför på att återskapa de mekanismer som styr ögats rörelser på en populationsnivå och på så sätt skapa en modell ur vilken vidare undersökning av betraktande av konst kan baseras. Modellutvecklingsarbetet baseras på Claude Monets *Terrassen vid Sainte Adresse* (figur 1) och vi har valt att fokusera på en tavla eftersom det huvudsakliga syftet med detta arbete är att introducera nya verktyg för ögonrörelseanalys. Vidare är de data som detta arbete baseras på, vilka beskrivs i avsnitt 2.1, inte ett oberoende stickprov och resultaten kan inte utan validering med ögonrörelsedata från andra källor generaliseras till en hel population. Denna rapport bör därför ses som en förstudie där modellutvecklandet presenteras.



Figur 2: Fixeringspunkter över Claude Monets Terrassen vid Sainte Adresse, för samtliga försökspersoner. För individuella fixeringspunktmönster, se appendix A.

2 Ögonrörelsedata

Utifrån de grundantaganden som beskrivs i avsnitt 1.1 används ögonrörelsedata från en studie utförd vid universitetet i Jyväskylä, Finland, vilken även Ylitalo m.fl. [2] använt, för att skatta modellens parametrar. I detta avsnitt presenteras dessa data och avgränsningar samt etiska och samhälleliga aspekter av arbetet diskuteras.

2.1 Data från universitetet i Jyväskylä

I försöket vid Jyväskylä deltog 20 studenter som fick i uppgift att observera sex olika tavlor i tre minuter vardera medan de verbalt beskrev tavlornas stämning. Samtidigt registrerades deras ögonrörelser utifrån vilket fixeringspunkter, fixeringstider, och sackadtider beräknades, se Ylitalo m.fl. [2]. Modelleringen i detta arbete baseras på dessa tre mått och modellen avgränsas därmed såvida att tavlans komposition endast indirekt tas i beaktande genom fixeringspunkternas fördelning, vilka visualiseras i figur 2. Under försöket delades försökspersonerna upp i två grupper: experter, bestående av tio studenter med konstintresse, och noviser, vilket inkluderade tio studenter utan specifik konstbakgrund. Av de tjugo försökspersonerna var fem män och femton kvinnor varav tre män och sju kvinnor var noviser, vilket detta arbete inte tar hänsyn till.

Ylitalo m.fl. [2] hittade i deras artikel inte några större skillnader mellan grupperna och vi behandlar därför inte grupperna separat i modelleringsarbetet. Vissa individers data innehåller några fixeringspunkter som registrerats utanför tavlan och dessa filtreras bort för att enbart fånga det estetiska engagemanget utan störningar från omgivningen.

2.2 Beträktarens blick förändras över tid

I *The Vienna Integrated Model of Art Perception* som utvecklats av Palowski m.fl. [8] vid universitetet i Wien argumenterar författarna för hur en åskådare ser på konst förändras under betraktandet. Enligt dem dominerar de första 6-8 sekunderna av bottom-up-processer varpå top-down-processer blir allt mer dominerande och efter 30-60 sekunder påverkar betraktarens övergripande tankestruktur i allt högre grad. Empiriska analyser av de data som används i detta arbete stödjer denna hypotes genom att bland andra Berntsson m.fl. [9] och Ylitalo m.fl. [2] visat att fixeringstiderna tenderar att bli signifikant längre efter 30 sekunder än under de initiala 30 sekunderna. Vidare visade Smith m.fl. [10] att betraktarens engagemang med en tavla generellt avtar efter cirka 30 sekunder och Massaro m.fl. [6] visade att om en tavla innehöll människor så innehöll de två första ROI:erna deras ansikten bland 92.6% av betraktarna, och övriga delar av tavlan utforskades senare under betraktandet. Med denna bakgrund avgränsar vi därför modelleringen till de första 30 sekunderna av observationsfönstret på tre minuter från vilket data finns tillgängligt.

2.3 Etiska och samhälleliga aspekter

Vid forskning som innefattar mänskliga försökspersoner finns en rad bestämmelser kring etiska riktlinjer. De grundläggande etiska principerna för forskning kan kort sammanfattas till att forskningen bör göra nytta, undvika skada och respektera försökspersonernas integritet och autonomi [11]. Denna sammanfattning grundar sig i de villkor kring forskningsetik som beskrivs i Lag (2003:460) om etikprövning av forskning som avser människor [12]. Där beskrivs de utgångspunkter som finns vid etikprövning och vad som måste uppfyllas för att godkännas vid en prövning. För detta arbetets del, då endast data från Ylitalo m.fl. används, bör främst försökspersonernas integritet och rätt till anonymitet tas i beaktning.

Det finns stora framsteg inom ögonspårning och hur tekniken kan användas till användaridentifikation. Kröger m.fl. [13, Kap. 15, ss. 226–241] beskriver hur ögonspårningsdata kan användas för att på olika sätt identifiera en person. De skriver om studier som gjorts där man använt ögonspårningsdata från personer när de exempelvis betraktat stillbilder från romantiska filmer och under online-shopping, från vilket det var möjligt att avgöra könet på en person. Eftersom detta arbete inte använder data från specifik media som valts ut i ett särskilt syfte att identifiera en person bör möjligheten till identifikation av försökspersonerna vara försumbar. Risken att inskränka på försökspersonernas integritet eller anonymitet anses därför vara mycket liten.

En annan potentiell risk med arbetet är huruvida resultatet kan användas av någon utomstående för något som inte är i försökspersonernas intresse. Specifikt gäller detta om modellen som skapas kan användas som ett verktyg vid marknadsföring eller i andra kommersiella syften. Det skulle krävas att modellen i sådant fall är applicerbar på den typ av media som är intressant för en utomstående aktör och inte endast den tavla som arbetet är baserat på. Risken för detta anses också mycket låg eftersom modellen endast baseras på fixeringspunktmönster från försökspersoner och inte på tavlans komponenter, även om det finns en korrelation mellan dessa.

3 Metoder och inledande dataanalys

I modelleringen, vars grundläggande antaganden presenterats i avsnitt 1.1 och vars implementering presenteras i detalj i avsnitt 4 används sannolikhetsfördelningar och metoder från spatial statistik för att generera fixeringspunktmönster. Detta avsnitt syftar därför till att passa lämpliga sannolikhetsfördelningar för fixeringstid, sackaktid, och sackadlängd samt skatta intensiteten för data från Ylitalo m.fl. [2] som presenterats i avsnitt 2.1.

3.1 Spatiala mönster av fixeringar

I denna del av rapporten betraktar vi fixeringspunktmönstren, som observeras när individer betraktar konstverket, specifikt som realiseringar av spatiala punktprocesser. Detta möjliggör en kvantitativ analys och visualisering av de områden på konstverket som oftast fångar betraktarens uppmärksamhet.

3.1.1 Punktprocesser och intensitet

En punktprocess är en modell för en samling av punkter som är slumpmässigt fördelade över ett område och används för att beskriva spatiala punktmönster. I denna studie använder vi punktprocessmodellen för att beskriva och analysera fixeringar inom det visuella fältet av konstverket. Genom att modellera dessa fixeringar som en punktprocess kan vi kvantitativt undersöka och visualisera de områden på konstverket som mest frekvent drar till sig betraktarens uppmärksamhet [14, Kap. 6, ss. 128–129].

Intensitetsfunktionen $\lambda(x)$ av en punktprocess, anger det förväntade antalet punkter av processen per enhetsarea vid varje punkt x inom det undersökta området. För en homogen punktprocess, där intensiteten λ är konstant över hela området, kan detta uttryckas som:

$$\lambda = \frac{N(W)}{|W|}, \quad (1)$$

där $N(W)$ är det förväntade antalet punkter av processen inom området W , och $|W|$ är arean av W . För icke-homogena punktprocesser varierar $\lambda(x)$ över ytan. Det är denna variation i intensitet som vi avser att utforska och kvantifiera genom kärnskattning, en metod som är användbar för att skatta och visualisera variationen i intensitetsfunktionen [14, Kap. 6, ss. 157–160].

3.1.2 Kärnskattning

Kärnskattning, eller kernel density estimation (KDE) från engelskan, är en icke-parametrisk metod för att uppskatta intensitetsfunktionen baserat på ett punktmönster. Matematiskt kan kärnskattningen för en uppsättning punkter $x_1, x_2 \dots x_n \in W$ representeras som:

$$\hat{f}(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (2)$$

där $\hat{f}(x)$ är den uppskattade intensiteten vid punkten $x \in W$, K är kärnfunktionen som oftast är Gaussisk och h är bandbredden som avgör kärnans bredd och därmed hur slät intensitetsuppskattningen blir [15, Kap. 5, ss. 126–127].

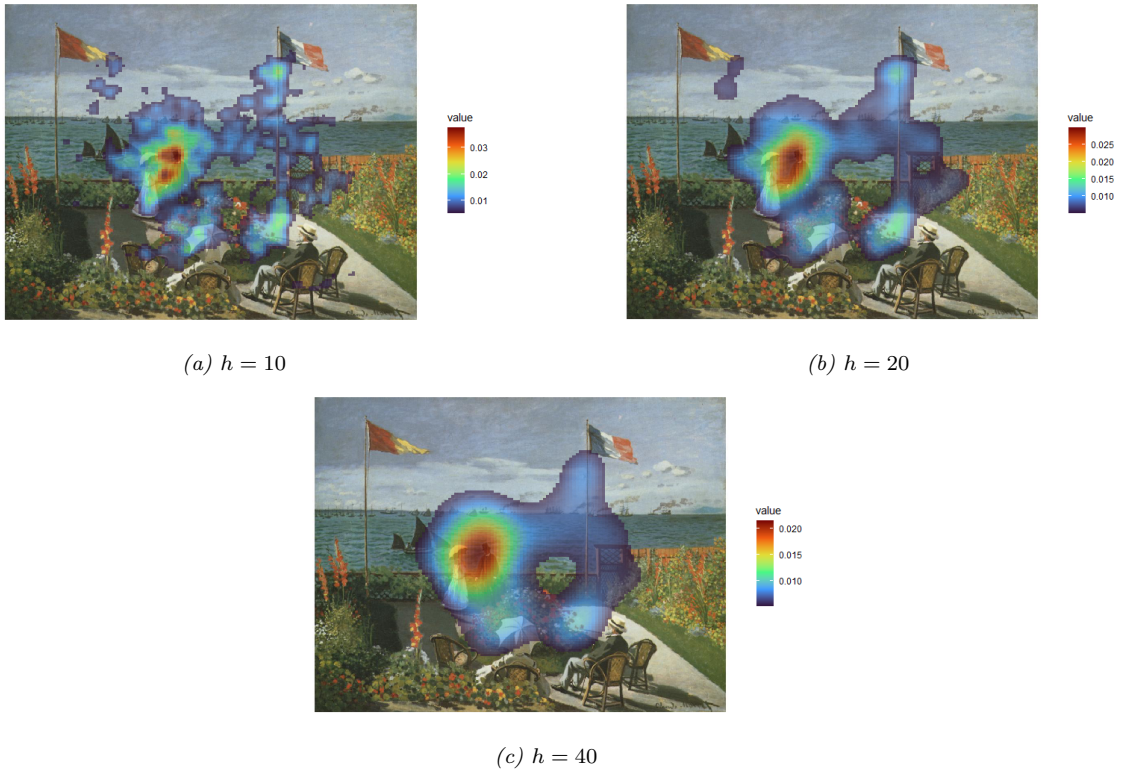
För att hantera komplexa eller multivariata datastrukturer kan bandbredden h ersättas med en bandbreddsmatrix H . Denna anpassning gör det möjligt att anpassa kärnskattningen för varje dimension. Formeln för kärnskattning i det multivariata fallet är:

$$\hat{f}(x) = \frac{1}{n|H|^{1/2}} \sum_{i=1}^n K\left(H^{-1/2}(x_i - x)\right), \quad (3)$$

där $|H|^{1/2}$ är kvadratroten av determinanten av bandbreddsmatrisen H , vilket normaliserar kärnfunktionen över dimensionerna av data [16].

Valet av bandbredd är avgörande, eftersom en för liten bandbredd kan leda till en brusig representation, medan en för stor bandbredd kan dölja strukturer hos data. I vår analys av ögonrörelser används kärnskattning för att skapa en intensitetsyta som visuellt representerar de områden på konstverket som drar mest uppmärksamhet. Ögonrörelsedata omvandlas till punktmönster i R genom att använda objektet `ppp` (planar point pattern) från `spatstat`-paketet, vilket anpassar data för spatial analys [14, Kap. 3, s. 54]. Därefter appliceras `density()`-funktionen för att utföra kärnskattningsanalysen, denna funktion tillämpar kärnfunktionen K och den valda bandbredden h för att generera en kontinuerlig intensitetsyta över det definierade området [14, Kap. 6, ss. 168–170].

Bandbredden h bestäms initialt i vårt fall genom funktionen `bw.diggle()`, som använder korsvalidering för att hitta en optimal bandbredd. Denna initiala skattning ger oss en grund att utgå ifrån, men ytterligare manuell justering av bandbredden har också utförts baserat på en noggrann visuell granskning av intensitetsytan för att finjustera resultatet [14, Kap. 6, s. 171]. Effekten av olika bandbredder h , inklusive valet av $h = 20$ som visade sig ge en optimal balans mellan detaljnivå och släthet, illustreras tydligt i figur 3.



Figur 3: Intensitetsytor över Claude Monets Terrassen vid Sainte Adresse, baserat på samtliga försökspersoner med olika val av bandbredd h .

3.2 Fixeringspunkters förskjutning relativt förväntade attraktionsområden

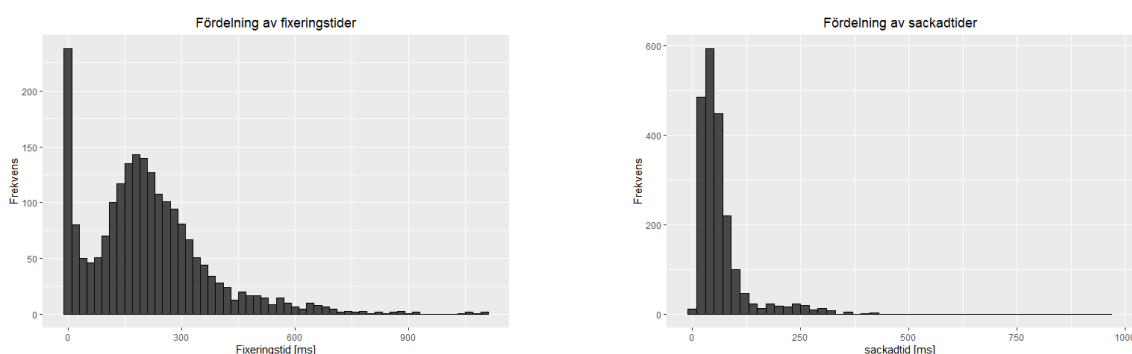
Enligt Massaro m.fl. [6] lokaliseras majoriteten av fixeringspunkterna kring mänskliga objekt, särskilt ansikten, under den inledande observationen. Eftersom figur 3b visar intensiteten för de första 30 sekunderna för samtliga individer överlagd på Claude Monets *Terrassen vid Sainte Adresse* är det därför att vänta att intensitetsytans maximipunkter bör sammanfalla med de fyra mänskliga objekten, men detta är inte fallet för kvinnan med parasollet och mannen i fätöljen. Fixeringspunkternas koordinater kan därför antas vara något förskjutna. Detta bör dock inte påverka modellen som presenteras i kommande avsnitt eftersom den endast baseras på fixeringspunktdata, och inte på tavlans komposition. Därför genomförs ingen korrigering för

denna misstänkta förskjutning, men med konsekvens att visualiseringar av faktiska och erhållna fixeringspunktmönster inte till fullo återger vad som betraktats.

Det framgår även av undersökningsprotokollet att fem av 20 försökspersoner använt glasögon under betraktandet av de sex tavlorna och att reflektioner i glaset kan påverka ögonrörelseinspelningen. För att undersöka detta jämförs intensitetsytorna för olika individer i appendix C. Inga större skillnader mellan intensitetsytorna kan visas och därför inkluderas samtliga individers fixeringspunktmönster i analyserna.

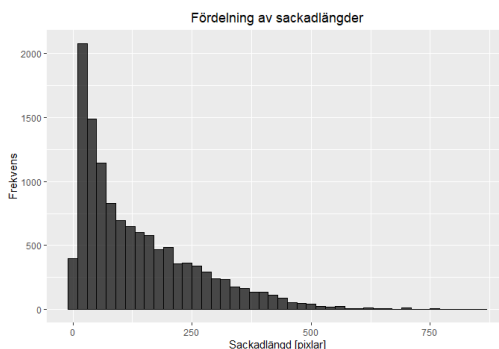
3.3 Fördelningar för fixeringstid, sackadtid, och sackadlängd

För att identifiera lämpliga fördelningar för fixeringstider, sackadtider och sackadlängder under de första 30 sekunderna av observationsfönstret på tre minuter, genomförs en inledande dataanalys. Denna analys inkluderar alla fixeringar som registrerats på tavlan och de visualiseras med hjälp av histogram i figur 4.



(a) Histogram över fixeringstider i ms för alla individer

(b) Histogram över sackadtider i ms för alla individer



(c) Histogram över sackadlängder i pixlar för alla individer

Figur 4: Histogram för fixeringstid och sackadtid i ms, och sackadlängder i pixlar.

Vi jämför olika statistiska fördelningar, nämligen gamma-, weibull-, lognormal- och exponentialfördelningar, för att bestämma vilken som bäst representerar de observerade datamönstren. För att ge en grundlig förståelse för dessa fördelningar, presenterar vi här definitionerna för de förekommande fördelningarna i vår analys.

Definition 1 (Exponentialfördelning) En kontinuerlig stokastisk variabel X sägs vara exponentialfördelad med intensitetsparameter $\lambda > 0$ om dess täthetsfunktion ges av

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & \text{för } x \geq 0, \\ 0 & \text{för } x < 0, \end{cases} \quad (4)$$

där väntevärdet och variansen för en exponentialfördelad stokastisk variabel ges av $E(X) = \frac{1}{\lambda}$ respektive $\text{Var}(X) = \frac{1}{\lambda^2}$ [17, Kap. 3, ss. 98–99].

Definition 2 (Lognormalfördelning) En positiv kontinuerlig stokastisk variabel Y sägs vara lognormalfördelad om dess logaritm $\ln(Y)$ är normalfördelad, alltså om $\ln(Y) \sim N(\mu, \sigma^2)$. Detta innebär att täthetsfunktionen för Y ges av

$$f_Y(y) = \frac{1}{y\sigma\sqrt{2\pi}} e^{-\frac{(\ln y - \mu)^2}{2\sigma^2}} \quad \text{för } y > 0, \quad (5)$$

där μ är medelvärdet och σ är standardavvikelsen för $\ln(Y)$ som bestämmer fördelningens form [17, Kap. 3, ss. 111–112].

Definition 3 (Gammafördelning) En kontinuerlig stokastisk variabel X sägs vara gammafördelad för $x \geq 0$ om dess täthetsfunktion ges av

$$f_X(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} \quad \text{för } x \geq 0, \quad (6)$$

där $\alpha > 0$ är formparametern, $\beta > 0$ dess intensitetsparameter och $\Gamma(\alpha)$ är gammafunktionen [17, Kap. 3, ss. 113–114].

Definition 4 (Weibullfördelning) En positiv kontinuerlig stokastisk variabel X sägs vara weibullfördelad om den har täthetsfunktionen

$$f_X(x) = \frac{\alpha}{\lambda} \left(\frac{x}{\lambda}\right)^{\alpha-1} e^{-\left(\frac{x}{\lambda}\right)^\alpha} \quad \text{för } x \geq 0, \quad (7)$$

där $\alpha > 0$ är formparametern och $\lambda > 0$ är skalfaktorn [18].

För att bestämma de parametervärden som mest sannolikt har genererat observerade data, tillämpas Maximum Likelihood-metoden för parameterskattning. Detta uppnås genom att maximera loglikelihoodfunktionen,

$$l(\theta) = \sum_{i=1}^n \log(f(x_i|\theta)), \quad (8)$$

där $f(x_i|\theta)$ är täthetsfunktionen för observation x_i givet parametern θ . Man maximerar med avseende på θ , med syftet att identifiera det värde på θ som maximerar sannolikheten för all observerad data, vilket indikerar bäst anpassning mellan modellen och data [17, Kap. 7, ss. 280–281].

Vi kompletterar denna metod med bootstrap-analyser för att generera konfidensintervall för de skattade parametrarna. Bootstrap är en statistisk teknik där man upprepat drar stickprov från den empiriska fördelningen med återläggning, så att varje stickprov innehåller samma antal observationer som originaldata, för att beräkna skattningar och konfidensintervall [17, Kap. 8, ss. 409–410]. Resultaten från dessa analyser presenteras i appendix B.

Med konfidensintervallen från bootstrap-analyserna som bekräftar parametrarnas stabilitet, tillämpas även Kolmogorov-Smirnovs-test för att ytterligare validera fördelningens anpassning. Testet kvantifierar avvikelsen mellan de teoretiska och empiriska fördelningarna, vilket ger en kvantitativ bedömning av fördelningens anpassning. För att formellt definiera detta test, presenteras följande:

Definition 5 (Kolmogorov-Smirnovs-test) Kolmogorov-Smirnov goodness of fit-testet är en icke-parametrisk metod som används för att bestämma om en samling oberoende observationer x_1, \dots, x_n överensstämmer med en specifik sannolikhetsfördelning. Testet jämför den empiriska kumulativa fördelningsfunktionen $F_n(x)$ med den teoretiska kumulativa fördelningsfunktionen $F(x)$.

Teststatistikan D definieras som:

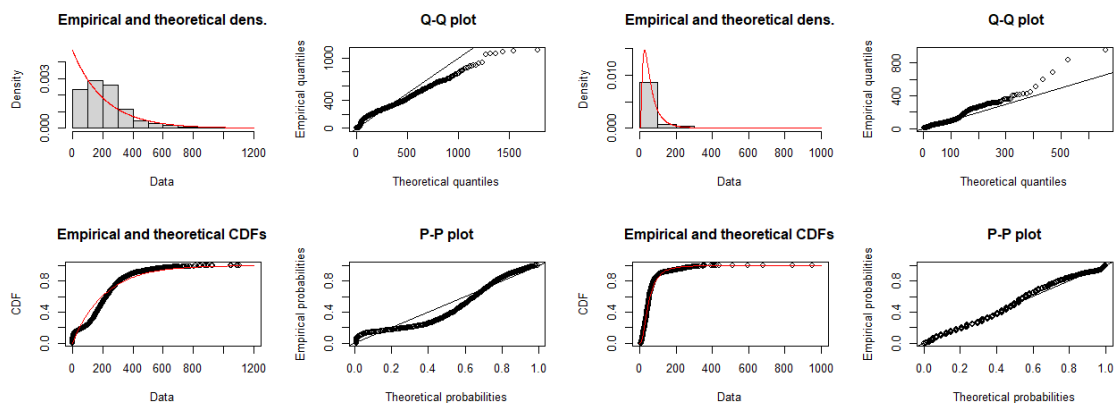
$$D = \sup_x |F_n(x) - F(x)|, \quad (9)$$

vilket är det maximala avståndet mellan $F_n(x)$ och $F(x)$, där \sup betecknar supremum. Ett högt värde på D indikerar en större avvikelse mellan de empiriska och teoretiska fördelningarna, vilket kan tyda på att den föreslagna modellen inte passar data väl. För att avgöra om nollhypotesen om att $F_n(x)$ kommer från $F(x)$ bör förkastas, kan kritiska värden från Kolmogorov-Smirnov-tabeller användas [19][14, Kap. 10, ss. 381–382].

Teststatistikans värden visas i tabell 1. Kolmogorov-Smirnov-testet ger även p -värden, men för alla fördelningar gällande fixerings-, sackaktider och sackadlängder erhåller vi p -värden < 0.05 , vilket vanligtvis indikerar en dålig passform eftersom att nollhypotesen förkastas för p -värden under signifikansnivån. Denna tolkning kan dock vara missvisande i stora datamängder, där även minimala avvikelser från den teoretiska fördelningen kan leda till mycket små p -värden. Detta beror på att testets känslighet ökar med antalet datapunkter och därför bör dessa p -värden inte användas isolerat för att bedöma fördelningens passform.

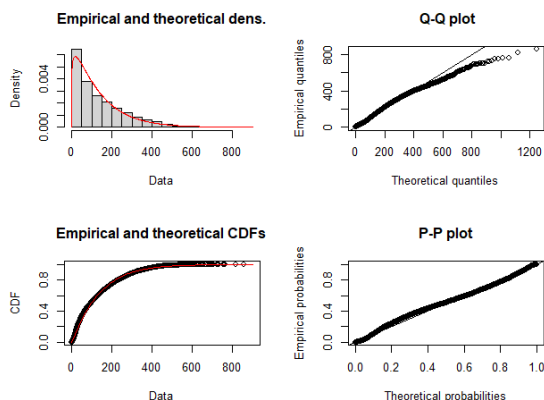
Tabell 1: Kolmogorov-Smirnov goodness of fit-statistik för de olika fördelningarna för fixeringstider, sackaktider och sackadlängder.

Kolmogorov-Smirnov	Gamma	Weibull	Lognormal	Exponential
Fixeringstider	0.1626	0.1392	0.2318	0.1589
Sackaktider	0.1257	0.1284	0.0641	0.1566
Sackadlängder	0.0459	0.0463	0.0664	0.0395



(a) Diagnostiska plottar för exponentialfördelning av fixeringstider

(b) Diagnostiska plottar för lognormalfördelning av sackaktider



(c) Diagnostiska plottar för gammafördelning av sackadlängder

Figur 5: Diagnostiska plottar inkluderande Q-Q-plot, P-P-plot samt jämförelser av empiriska och teoretiska fördelningsfunktioner.

För att komplettera den kvantitativa analysen, används diagnostiska plottar för att visualisera hur väl de teoretiska fördelningarna överensstämmer med observerade data. Från de diagnostiska plottarna i figur 5 presenteras de valda fördelningarna som lämpar sig ganska väl för dessa data. För fixeringstider väljer vi exponentialfördelningen, även om både weibull- och gammafördelningen också visade goda passformer. Valet av fördelning baseras inte på signifikanta skillnader mellan modellerna, utan snarare på att de presterade likvärdigt.

För sackadtider framstår lognormalfördelningen som den mest lämpliga modellen, även om exponentialfördelningen också skulle kunna vara ett lämpligt alternativ. När det gäller sackadlängder, visar gamma-, weibull- och exponentialfördelning goda passformer. Vi väljer gammafördelningen baserat på en helhetsbedömning av dess egenskaper i de diagnostiska plottarna, trots att skillnaderna mellan fördelningarna är marginella. Dessa val grundar sig på en noggrann analys med fokus på visuella bedömningar från Q-Q- och P-P-plottar. I varje fall har en av de mest passande fördelningarna valts baserat på dessa bedömningar. Detaljerna för dessa analyser, inklusive alla relevanta diagnostiska plottar och resultat, finns tillgängliga för fördjupad granskning i appendix B.

4 Modellering

Modellen består i huvudsak av tre komponenter: klustring, Markovkedjor och punktgenerering. Först utgår modellen från ett slumpmässigt valt tillstånd utifrån en Markovkedja, där tillstånden består av ROI:er och bakgrunden, vilka definierats genom en klustringsalgoritm. Därefter genereras en punkt slumpmässigt inom det givna området med avseende på den skattade intensitetsytan. Två olika algoritmer för punktgenerering presenteras. Dessa resulterar i en mängd modellbaserade fixeringspunkter på tavlan som jämförs med faktiska data med hjälp av två statistikor, vilka presenteras i avsnitt 4.4.

4.1 Definition av ROI:er genom mean-shift-klustring

För att definiera ROI:erna i tavlan utgår vi från den klustringsmetod, mean-shift (MS), som föreslås i *Robust Clustering of Eye Movement Recordings for Quantification of Visual Interest* av Santella m.fl. [7].

MS-klustring är en icke-parametrisk metod som identifierar kluster av punkter, i detta fall fixeringspunkter, genom att iterativt flytta punkter mot områden med högre densitet. Detta uppnås genom att beräkna ett lokalt masscentrum, där punkterna viktas med en kärnfunktion $K(x)$. De punkter som konvergerar mot lokala maxima, där punkttätheten är som högst, anses tillhöra samma kluster. Denna metod är beräkningsmässigt kostsam men effektiv för att hantera avvikande data och brus, vilket gör den väl lämpad för klustring av ögonrörelsedata [7]. Vi implementerar detta genom `ms`-funktionen i LPCM paketet i R, vilken baseras på Ameijerias-Alonso m.fl. [20, 16] vars resonemang återges här:

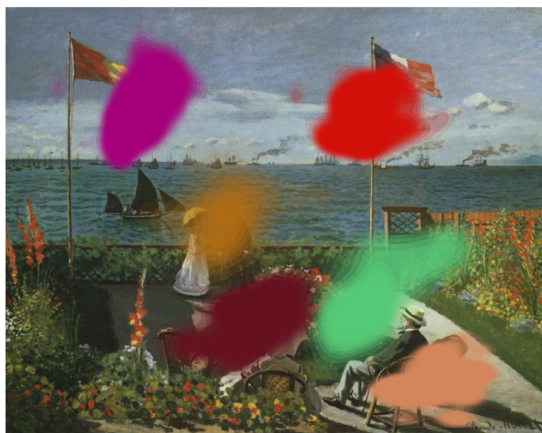
Definition 6 (Mean-shift-klustring) *Betrakta en stokastisk vektor X med täthetsfunktion $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ och medelvärde $\mu = (\mu_1, \mu_2)$ från vilken vi har n realisationer x_1, x_2, \dots, x_n . Låt $H = \text{diag}(h_1^2, h_2^2)$ vara en diagonal matris med positiva bandbredder h_i och låt $K(x)$ vara en kärnfunktion. I de lokala maximipunkterna är gradienten av intensitetsytan \hat{f} av f , som ges av ekvation (3), lika med noll. Genom att lösa ut x ur gradienten med en Gaussisk kärnfunktion $K(x) = \frac{1}{\sqrt{2\pi}} e^{-\|x\|^2/2}$ erhålles*

$$x = \frac{\sum_{i \in N} K(H^{-1/2}(x_i - x))x_i}{\sum_{i \in N} K(H^{-1/2}(x_i - x))} := \mu(x). \quad (10)$$

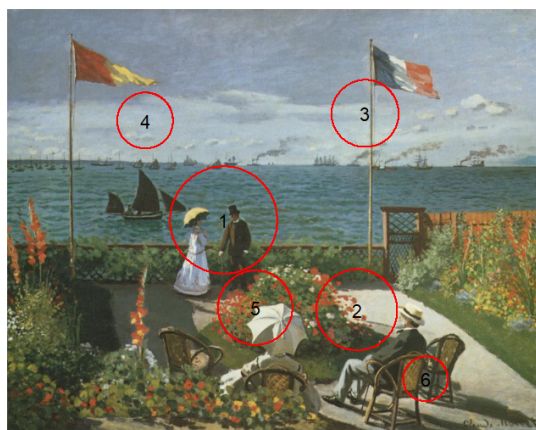
Det framgår av (10) att x är lika med det viktade medelvärdet av punkter i omgivningen av x . Genom att iterativt beräkna $x^{(n)} = \mu(x^{(n-1)})$ nås lösningen av (10) och alla punkter i dessa data som tillhör samma maxima sägs tillhöra samma kluster [16].

Valet av bandbredder H är därför en avgörande faktor i klusterformationen eftersom det avgör hur stor omgivning som tas hänsyn till i varje iteration och därmed även klusterseparationen [7, 16]. En vanlig utgångspunkt för att välja bandbredden H är att använda standardvärde om 5% av observationsfönstret. Detta erbjuder en praktisk början för bandbreddsval, som sedan kan anpassas genom kvalitativa eller empiriska tester för att bättre passa de specifika egenskaperna hos datamängden [7][14, s. 168–174]. Genom en kvalitativ analys av erhållna kluster valdes h_1 och h_2 till 4.5% av tavlans dimensioner och eftersom *Terrassen vid Sainte Adresse* ej är kvadratisk är $h_1 = 45.6 \neq h_2 = 34.6$. Detta skiljer sig från kärnskattningen av intensitetsytan i avsnitt 3.1.2 där en bandbredd, $h = 20$, används i båda dimensioner. h är mindre än h_1 och h_2 eftersom intensitetsytan i figur 3b används som en sannolikhetsfördelning i avsnitt 4.3 vilket kräver en högre upplösning. MS-klustringen används istället för att finna övergripande områden med många fixeringar vilket kräver en större bandbredd som även tar hänsyn till tavlans dimensioner för att inte generera kluster med ett fåtal fixeringar.

De ROI:er som erhålls från MS-klustringen är geometriskt irreguljära, se figur 6a, och för att förenkla beräkningarna antas varje ROI kunna modelleras som en cirkel med mittpunkt i klustrets centrum. Cirkelns radie anpassas så att de inte överlappar varandra och radien, definierad av en basradie r_b och en skalningsfaktor r_s , justeras dynamiskt beroende på antalet fixeringspunkter n i ett givet kluster, enligt formeln $r = r_b + r_s \cdot \log(n)^2$. För Claude Monets *Terrassen vid Sainte Adresse* valdes $r_b = 15$ och $r_s = 1.4$ och dessa visas på tavlan i figur 6b.



(a) ROI för samtliga individers data



(b) Cirklar innehållande tavlans ROI:er

Figur 6: Jämförelse mellan ROI-kluster och de ROI-cirklar som används i modellen.

4.2 Markovkedjor för att återskapa fokusförflyttningar mellan ROI:er

För att återskapa fokusförflyttningen mellan olika ROI:er i tavlan används Markovkedjor, och vi antar därför att Markovegenskapen kan betraktas som uppfylld under de 30 första sekunderna i denna process. Denna egenskap definierar vi enligt:

Definition 7 (Markovkedjor) Låt S vara en diskret mängd vilken vi benämner tillståndsrum. En Markovkedja är en sekvens av stokastiska variabler $X_0, X_1, \dots \in S$ med markovegenskapen

$$P(X_{n+1} = j | X_0 = x_0, \dots, X_{n-1} = x_{n-1}, X_n = i) = P(X_{n+1} = j | X_n = i), \quad (11)$$

för alla $x_0, x_{n-1}, i, j \in S$ och $n \geq 0$ [21, s. 41].

Övergångssannolikheten mellan två tillstånd i och j och definieras därmed som

$$p_{ij} = P(X_{n+1} = j | X_n = i). \quad (12)$$

Dessa sannolikheter sammanställs i övergångsmatrisen P , som är en $m \times m$ dimensionell matris med m antal unika tillstånd där varje element $p_{ij} \geq 0$ och uppfyller

$$\sum_{j=1}^m p_{ij} = 1 \quad \forall i = 1, 2, \dots, m, \quad (13)$$

vilket säkerställer att varje rad i P utgör en diskret sannolikhetsfördelning [21, 22].

Vi låter ROI:erna utgöra tillståndsrummet S och antar att den Markovkedja som genereras utifrån detta antagande återskapar betraktarens fokusförflyttningar mellan ROI:er. För att generera ett antal fixeringar som återspeglar individdata slumpas fixerings- och sackadtider från fördelningarna för dessa som valts i avsnitt 3.3 för varje fixeringspunkt så att den totala tiden blir 30 sekunder.

Den yta på tavlan som inte omsluts av en cirkel benämns som bakgrund och definieras som ett eget tillstånd i Markovkedjan och här inkluderas även de fixeringspunkter som tillhör kluster med färre än elva fixeringspunkter. Utifrån detta beräknas övergångsmatrisen där varje övergångssannolikhet p_{ij} är proportionell mot antalet övergångar från tillstånd j till i .

4.3 Punktgenerering

Punktgenereringen implementeras på två olika sätt, metod A och metod B. Båda metoder avser att generera punkter utifrån intensitetsytan genom att använda rejection sampling, se [23, Kap. 5, ss. 132–133], men implementationen skiljer sig åt.

Metod A implementeras genom att först slumpmässigt generera en fixeringspunkt likformigt inom det tillstånd som valts från övergångsmatrisen. Sedan normaliseras intensitetsytan med det största intensitetsvärdet λ_{max} i området som definieras av tillståndet. Därefter beräknas intensitetsvärdet $\lambda(x)$ för positionen. Om $\lambda(x) \geq X \sim unif[0,1]$ accepteras punkten, annars förkastas den och processen börjar då om tills en punkt har antagits inom det valda tillståndet. Ett större intensitetsvärde innebär alltså en större sannolikhet att fixeringspunkten antas och är kriteriet som rejection samplingen utgår från.

Metod B bygger på samma principer kring rejection sampling som metod A, men tar den föregående punktens position i beaktande likt metoden i Ylitalo m.fl. [2]. Om det aktuella tillståndet är samma som för den tidigare genererade punkten eller bakgrundstillståndet genereras punkten på en radie ifrån den förra punkten. Radien beräknas genom att slumpmässigt välja en längd, l , från gammafördelningen av sackadlängderna från kapitel 3.3. Om den genererade fixeringspunkten skulle hamna utanför det valda tillståndet genereras en ny fixeringspunkt. Fixeringspunkten accepteras eller förkastas sedan enligt samma kriterier för intensitetsytan som i metod A. Metod A ger sedan modell A och metod B ger modell B, där algoritmen för respektive modell beskrivs översiktligt i tabell 2 och 3.

Tabell 2: Översikt över punktgenereringsalgoritm A.

Punktgenereringsalgoritm A	
1	Generera en sekvens ROI:er, samt fixerings- och sackadtider med Markovkedjan enligt 4.2
2	Föreslå en punkt x likformigt inom en ROI
3	Acceptera x om $unif(0,1) \leq \lambda(x)$, återgå annars till steg 2
4	Upprepa steg 2 & 3 för alla ROI:er i sekvensen

Tabell 3: Översikt över punktgenereringsalgoritm B.

Punktgenereringsalgoritm B	
1	Generera en sekvens ROI:er, samt fixerings- och sackadtider med Markovkedjan enligt 4.2
2	Föreslå en punkt x likformigt inom en ROI vid byte av tillstånd. Vid fortsatt samma tillstånd eller övergång till bakgrund, föreslå en punkt x på avstånd l , slumpat från sackadlängdsfördelningen
3	Acceptera x om $unif(0,1) \leq \lambda(x)$, återgå annars till steg 2
4	Upprepa steg 2 & 3 för alla ROI:er i sekvensen

4.4 Statistikor

För att jämföra modellens resultat med faktiska data använder vi två olika statistikor, den totala sackadlängden och det konvexa höljet av fixeringspunkter. En sackadlängd definieras som det euklidiska avståndet mellan två efterföljande fixeringspunkter och mäter alltså hur långt ögat färdas mellan två fixeringspunkter. Den kumulativa summan av sackadlängder för en tid t_i definieras för en ordnad mängd $x = \{x_1, \dots, x_{n_t}\}$ av n_t fixeringspunkter av Ylitalo m.fl. [2] som

$$L(t) = \sum_{i=1}^{n_t-1} l_i \mathbf{1}(t_{i+1} \leq t), \quad (14)$$

där l_i är den euklidiska längden mellan fixeringspunkterna x_{i+1} och x_i .

Det konvexa höljet är den totala arean av den polygon som omsluter ett givet antal punkter och definieras av Ylitalo m.fl. [2] som antalet pixlar i mängden

$$C_x(t) = \left\{ \sum_{i=1}^{n_t} \alpha_i x_i : \alpha_i \geq 0 \quad \forall i \text{ \& } \sum_{i=1}^{n_t} \alpha_i = 1 \right\}. \quad (15)$$

För att beräkna detta användes `cxhull`-paketet i R. Det konvexa höljet mäter alltså hur stor area av tavlan som observeras och beräknas kumulativt för varje ny fixeringspunkt. Jämförelse av modellens resultat med data sker visuellt genom att plotta statistikorna och jämföra hur väl statistikornas värden för modellen och data sammanfaller.

5 Resultat och diskussion

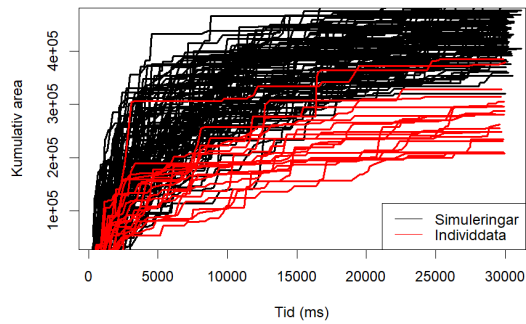
I detta avsnitt jämförs hur väl de två varianterna av modellen som presenteras i avsnitt 4 kan återskapa försökspersonernas fixeringspunktmönster genom att applicera de statistikor som presenteras i avsnitt 4.4.

Tabell 4: Medelantal fixeringar och medelandel av observationstiden som spenderas i varje ROI, där ROI 0 utgör bakgrundstillståndet.

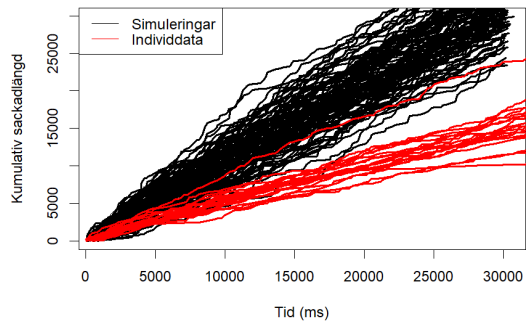
ROI	#fixeringar simulerad	#fixeringar data	andel fixeringstid simulerad	andel fixeringstid data
0	55.73	55.20	0.519	0.514
1	27.30	25.50	0.247	0.254
2	8.67	8.35	0.076	0.072
3	5.99	5.70	0.050	0.052
4	2.15	2.20	0.023	0.022
5	8.52	7.85	0.074	0.073
6	1.00	0.95	0.010	0.012

I figur 6b visas de sex erhållna ROI:erna och tabell 4 presenterar medelvärdet av den relativa vikt som försökspersonerna samt 100 realisationer av modellen lägger i respektive ROI, vilket mäts i antalet fixeringar och i andelen av den totala fixeringstiden. Det är tydligt att detta återskapas mycket väl av Markovkedjan, vars övergångssannolikheter skattas från antalet övergångar mellan ROI:er, och fixeringstidsfördelningen från avsnitt 3.3. Detta indikerar även att fixeringstidsfördelningen är någorlunda densamma i samtliga ROI:er och att betraktarna spenderar majoriteten av tiden i bakgrunden (ROI 0) och i ROI 1, vilket är paret i mitten av tavlån.

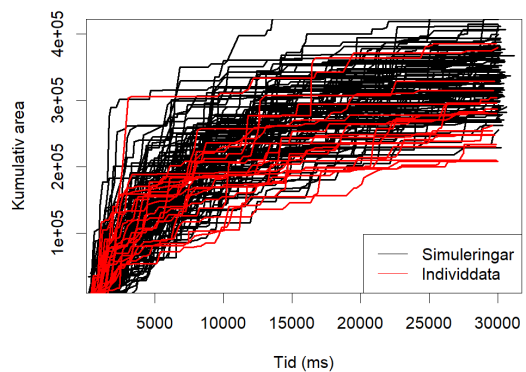
Valet av fördelningar för fixeringstid, sackadtid, och sackadlängd bör dock påverka resultatet. Det framgår från figur 4a att det förekommer ett antal korta fixeringar med en fixeringstid på under 40 millisekunder som ger upphov till två tydliga toppar. Denna observation tyder på att en bimodal fördelning skulle kunna vara mer lämplig än den använda exponentialfördelningen för att beskriva fixeringstider. Det är även troligt att dessa korta fixeringar kan betraktas som artefakter snarare än faktiska fixeringar på grund av deras kortare varaktighet än övriga fixeringar i data och i litteraturen [6]. Vi har i detta arbete valt att inkludera dem i analysen för att fånga den fulla dynamiken i ögonrörelsemönstren, men detta skiljer sig från Ylitalo m.fl. [2], där sådana korta fixeringar exkluderats med argumentet att de är artefakter. Den valda fördelningen återskapar dock den relativa fixeringstiden i respektive ROI väl och anses därför lämplig för att modellera fixeringstidsdata där korta fixeringar ej exkluderats.



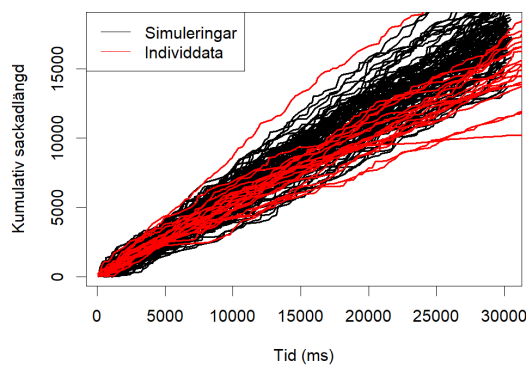
(a) Det konvexa höljet, modell A



(b) Kumulativ sacklängd, modell A



(c) Det konvexa höljet, modell B



(d) Kumulativ sacklängd, modell B

Figur 7: Sammanfattande statistikor för modell A och B för individdata i rött och 100 realiseringar av modellen i svart.

Utifrån den sekvens av ROI:er som produceras av Markovkedjan används punktgenereringsalgoritm A och B som presenteras i avsnitt 4.3 för att generera fixeringspunkter. Utifrån modellrealisationens fixeringspunktmönster beräknas det kumulativa konvexa höljet och den kumulativa sackadlängden, vars definitioner presenteras i avsnitt 4.4. I figur 7 presenteras dessa för individdata i rött och för 100 realisationer av respektive modell i svart. Det framgår att modell B bättre återskapar fixeringspunktmönstren än modell A vars realiseringar ger mycket högre värden på båda statistikor. Detta är att vänta eftersom i punktgenereringsalgoritm A föreslås en punkt med likformig sannolikhet i hela ROI:n vilket leder till långa sackader, speciellt i ROI 0 som täcker hela bakgrunden. Algoritm B tar med inspiration från Ylitalo m.fl. [2] också hänsyn till sackadlängdsfördelningen hos individdata när nästa punkt föreslås vilket begränsar den totala sackadlängden och individdata i rött återskapas mycket väl av modellen under de första 15 sekunderna. För $t \in [15,30]$ överestimerar modellen båda statistikor något, men modellen återskapar båda statistikor någorlunda väl i detta tidsintervall.

Notera att i figur 7d rör sig en individs öga ett mycket större avstånd än övriga försökspersoners efter ca. 10 sekunder vilket kan påverka parameterskattningen för sackadlängdsfördelningen. Det är därför tänkbart att en exklusion av denna individ skulle resultera i en bättre modellpassning för övriga individer när $t \in [15,30]$. Det är även tänkbart att processerna som styr ögats rörelse förändras efter 10-15 sekunder och att denna modell endast återskapar mekanismerna bakom den första. För detta finns det visst stöd i litteraturen där Pelowski m.fl. [8] argumenterar i sin modell *The Vienna Integrated Model of Art Perception* för att bottom-up processer dominerar under de första 6-8 sekunderna varvid top-down processer och individens övergripande tankestruktur tar över i större utsträckning.



(a) Försökspersoner

(b) 20 realiseringar av modell B

Figur 8: Intensitetsytor av fixeringspunktmönster.

Utöver att återskapa den relativa vikten i respektive ROI, det konvexa höljet, och hur långt ögat rör sig på tavlan är även fixeringspunkternas placering av intresse. I figur 8 visas därför intensitetsytan för samtliga individer och för 20 realiseringar av modell B. Det framgår att de områden med hög intensitet, vilka visas i rött, gult och grönt, överensstämmer med ROI-cirklarnas placering och återskapas väl av modellen. De områden med låg intensitet, vilka visas i blått samt utan färgöverläggning, återskapas inte till fullo. Exempelvis ser vi att modellen genererat fler fixeringspunkter i området till höger om paret än försökspersonerna.

Det är inte oväntat att modellen återskapar dessa perifera områden med lägre intensitet sämre eftersom modellen endast baseras på försökspersonernas fixeringspunktmönster och inte på tavlans komposition. Bakgrundstadiet (ROI 0) är stort och det finns många möjliga punkter att välja. I modell B föreslås punkter i bakgrundstadiet på ett avstånd l från fördelningen för sackadlängder oavsett om tidigare punkt tillhör bakgrunden eller en ROI och riktningen i vilken ögonrörelsen övergår från en ROI till bakgrunden avgörs via rejection sampling. Eftersom intensiteten är nollskild på nästan hela tavlan kan fixeringspunkter antas i områden med låg intensitet vilket resulterar i fler fixeringspunkter i vissa områden, och färre i andra, än vad försökspersonerna genererat. Överlag återskapar modellen dock fixeringspunktmönstret väl sett till de högintensiva områdena och intensitetsytan i figur 8b bör se annorlunda ut om den baserats på 20 andra realiseringar av modellen. Några realisationer av modell B presenteras i appendix E.

6 Slutsatser

Vi har i detta arbete utvecklat en modell för ögonrörelser vid betraktande av Claude Monets *Terrassen vid Sainte Adresse* som utifrån ögonrörelsedata från 20 individer återskapar fixeringspunktmönstret genom att kombinera den modell för fixeringspunktgenerering som presenteras i Ylitalo m.fl. [2] med fokusförflyttningar mellan attraktionsområden (ROI) på tavlan. Analys av ROI:er genom klustring av fixeringspunkter är inget nytt inom konstvetenskaplig analys av betraktarens fokus [6, 7] men det har inte tidigare kombinerats med punktprocesser.

Ögonrörelsedata innehåller stor variabilitet mellan individer samtidigt som individens ögonrörelser påverkas av uppgiften och olika fixeringsmönster från samma individ skiljer sig därför åt [6]. Fixeringsmönstret är även beroende av andra aspekter så som tavlans komposition och individens känslor [5, 6, 8] vilket gör det svårt att modellera, men den modell som presenteras här lyckas återskapa både det konvexa höljet och sackadlängden för *Terrassen vid Sainte Adresse* mycket väl under de första 15 sekunderna och väl under de första 30.

I modellen görs flera val av bland annat storleken på cirkelarna som omsluter ROI:erna och fördelningar samt parametrar för fixeringstid, sackadtid, och sackadlängd och hur dessa väljs kan påverka resultatet. Mindre förändringar i valet av cirkelarnas storlek bör dock inte ha någon större påverkan på resultaten eftersom när modellen går från en ROI till bakgrundsstadiet väljs första punkten i bakgrundsstadiet på ett avstånd l från sackadlängdsfördelningen från föregående fixeringspunkt. Därför är sannolikheten hög att punkter föreslås i omgivningen till en ROI, där även intensiteten är hög, till skillnad från längre ifrån dem, men modellen bör vara känslig för stora variationer i detta val. Vid implementering av modellen är det därför av vikt att cirkelarnas storlek väljs så att de omsluter den del av de erhållna klustret med hög intensitet. Dessutom beskriver flera fördelningar ögonrörelsedata någorlunda väl och valet av dessa bör inte påverka modellens förmåga att återskapa fixeringspunktmönstren nämnvärt. Det skulle även vara möjligt att använda de empiriska fördelningarna men eftersom flera fördelningar passar data väl är påverkan på resultatet troligtvis inte särskilt stor.

Valet av ROI:er kan även påverka hur väl modellen återskapar ögonrörelserna, men eftersom fixeringspunkternas koordinater väljs beroende av intensitetsytan i respektive ROI bör detta inte ha en signifikant påverkan på resultaten. Dock, om fler ROI:er med färre än 11 (0.55 per individ) fixeringar skulle tas hänsyn till skulle övergångsmatrisen innehålla fler tillstånd, men eftersom dessa har få fixeringar skulle sannolikheten att hamna i dessa vara låg och därmed är påverkan på total sackadlängd minimal. Vidare används ROI begreppet främst i studier av realistiska konstverk, framför allt med landskaps- eller människomotiv [6], och det är därför oklart hur ROI begreppet generaliserar till abstrakt konst men vår implementation som baseras på fixeringspunktmönster bör generalisera väl.

Generellt är det oklart hur väl de modeller som presenteras i denna studie skulle återskapa fixeringspunktmönster för andra tavlor. För att validera modellen bör en analys där datan delas upp i två grupper genomföras där en grupp används för att anpassa fördelningarna och intensiteten, medan den andra gruppen används för att testa hur väl modellen passar för detta och andra konstverk. Det är dock troligt att de grundläggande idéer som presenterats kan tillämpas på andra konstverk, men intensitetsytan och valet av fördelningar kan skilja sig avsevärt mellan olika verk eftersom varje tavla har en unik komposition som påverkar fixeringspunktmönstret.

Gällande generaliserbarhet av vår studie bör det noteras att urvalet av 20 personer inte nödvändigtvis är representativt för en större population, vilket begränsar möjligheten att generalisera resultaten till andra grupper eller situationer. Trots dessa begränsningar kan denna studie ses som en värdefull förstudie och konceptbevis som visar potentialen i att kombinera Markovkedjor och punktprocesser för att återskapa ögonrörelser. Med detta tillvägagångssätt kan detta arbete tjäna som en utgångspunkt och ett verktyg för vidare konstvetenskapliga studier där ögonrörelser analyseras.

Referenser

- [1] N. J. Wade. “Pioneers of Eye Movement Research”. I: *i-Perception* 1 (2010), s. 33–68. DOI: 10.1068/i0389.
- [2] A-K. Ylitalo, A. Särkkä och P. Guttorp. “What We Look at in Paintings: A Comparison Between Experienced and Inexperienced Art Viewers”. I: *The Annals of Applied Statistics* 10(2) (2016), s. 549–574. DOI: 10.1214/16-AOAS921.
- [3] S. Barthelmé m. fl. “Modeling fixation locations using spatial point processes”. I: *Journal of Vision* 13(12):1 (2013), s. 1–34. DOI: 10.1167/13.12.1.
- [4] A. Bringmann m. fl. “The primate fovea: Structure, function and development”. I: *Progress in Retinal and Eye Research* 66 (2018), s. 49–84. ISSN: 1350-9462. DOI: 10.1016/j.preteyeres.2018.03.006.
- [5] N. J. Wade. “Looking at Buswell’s pictures”. I: *Journal of Eye Movement Research* 13(2):4 (2020). DOI: 10.16910/jemr.13.2.4.
- [6] D. Massaro m. fl. “When Art Moves the Eyes: A Behavioral and Eye-Tracking Study”. I: *PLoS ONE* 7(5) (2012), e37285. DOI: 10.1371/journal.pone.0037285.
- [7] A. Santella och D. DeCarlo. “Robust clustering of eye movement recordings for quantification of visual interest”. I: *Proceedings of the 2004 Symposium on Eye Tracking Research & Applications*. 2004, s. 27–34. DOI: 10.1145/968363.968368.
- [8] M. Pelowski m. fl. “Move me, astonish me. . . delight my eyes and brain: The Vienna Integrated Model of top-down and bottom-up processes in Art Perception (VIMAP) and corresponding affective, evaluative, and neurophysiological correlates”. I: *Physics of Life Reviews* 21 (2017), s. 80–125. DOI: 10.1016/j.plrev.2017.02.003.
- [9] F. Berntsson m. fl. “Läs upp konstens hemligheter: En jämförelse av intensitetsytor från olika tidsintervall med hjälp av blickspårning och spatiala punktprocesser”. Kandidatarbete. Göteborg: Chalmers tekniska högskola och Göteborgs universitet, 2023.
- [10] L. Smith, J. Smith och P. Tinio. “Time Spent Viewing Art and Reading Labels”. I: *Psychology of Aesthetics, Creativity, and the Arts* 11(1) (2016), s. 77–85. DOI: 10.1037/aca0000049.
- [11] Chalmers. *Genomföra kandidatarbete*. 2024. URL: <https://www.chalmers.se/utbildning/dina-studier/kandidat-och-examensarbete/kandidatarbete/genomfora-kandidatarbete/#samhalleliga-och-etiska-aspekter>. Hämtad: 2024-02-09.
- [12] *Lag (2003:460) om etikprövning av forskning som avser människor*. Svensk författningssamling. <https://www.riksdagen.se/svenskforfattningssamling/2003:460>, SFS nr: 2003:460, Utfärdad: 2003-06-05, Ändrad t.o.m. SFS 2023:39. 2003.
- [13] J. L. Kröger, O. H-M. Lutz och F. Müller. “What Does Your Gaze Reveal About You? On the Privacy Implications of Eye Tracking”. I: *Privacy and Identity Management. Data for Better Living: AI and Privacy*. Utg. av M. Friedewald m. fl. Vol. 576. IFIP Advances in Information and Communication Technology. Springer, 2020. Kap. 15, s. 226–241. DOI: 10.1007/978-3-030-42504-3_15.
- [14] A. Baddeley, E. Rubak och R. Turner. *Spatial Point Patterns Methodology and Applications with R*. 1 uppl. London: Chapman Hall, 2015.
- [15] W.N. Venables och B.D. Ripley. *Modern Applied Statistics with S*. 4 uppl. New York: Springer, 2002.
- [16] J. Ameijeiras-Alonso och J. Einbeck. “A fresh look at mean-shift based modal clustering”. I: *Advances in Data Analysis and Classification* (2023). DOI: 10.1007/s11634-023-00575-1.
- [17] Alm S.E. och Britton T. *Stokastik : Sannolikhets teori och statistikteori med tillämpningar*. 1 uppl. Stockholm: Liber, 2008.
- [18] A. Kızılersü, M. Kreer och A. W. Thomas. “The Weibull Distribution”. I: *Significance* 15.2 (2018), s. 10–11. DOI: 10.1111/j.1740-9713.2018.01123.x.
- [19] M. L. Delignette-Muller och C. Dutang. “fitdistrplus: An R Package for Fitting Distributions”. I: *Journal of Statistical Software* 64.4 (2015), s. 1–34. DOI: 10.18637/jss.v064.i04.

- [20] J. Einbeck. *Mean shift clustering, R documentation*. URL: <https://search.r-project.org/CRAN/refmans/LPCM/html/ms.html>.
- [21] R. P. Dobrow. *Introduction to Stochastic Processes with R*. Hoboken: Wiley, 2016.
- [22] P. H Hsu. *Schaum's Outline of Probability, Random Variables, and Random Processes*. 3 uppl. New York: McGraw-Hill Education, 2014, s. 212–215.
- [23] P. Clifford. “Methods in Experimental Physics”. I: *Statistical Methods for Physical Science*. Utg. av J. L. Stanford och S. B Vardeman. Vol. 28. Academic Press, 1994. Kap. 5, s. 132–133. ISBN: 0-12-475973-4.

A Individuella fixeringspunktmönster



(a) Fixeringspunkter för individ 1



(b) Fixeringspunkter för individ 2



(c) Fixeringspunkter för individ 3



(d) Fixeringspunkter för individ 4



(e) Fixeringspunkter för individ 5



(f) Fixeringspunkter för individ 6



(g) Fixeringspunkter för individ 7



(h) Fixeringspunkter för individ 8



(i) Fixeringspunkter för individ 9



(j) Fixeringspunkter för individ 10



(k) Fixeringspunkter för individ 11



(l) Fixeringspunkter för individ 12



(m) Fixeringspunkter för individ 13



(n) Fixeringspunkter för individ 14



(o) Fixeringspunkter för individ 15



(p) Fixeringspunkter för individ 16



(q) Fixeringspunkter för individ 17



(r) Fixeringspunkter för individ 18



(s) Fixeringspunkter för individ 19



(t) Fixeringspunkter för individ 20

B Statistiska skattningar och diagnostiska plottar

Tabell 5: Parameterskattningar och bootstrap 95% konfidensintervall för parametrar för de olika fördelningarna för fixeringstider med 1000 iterationer.

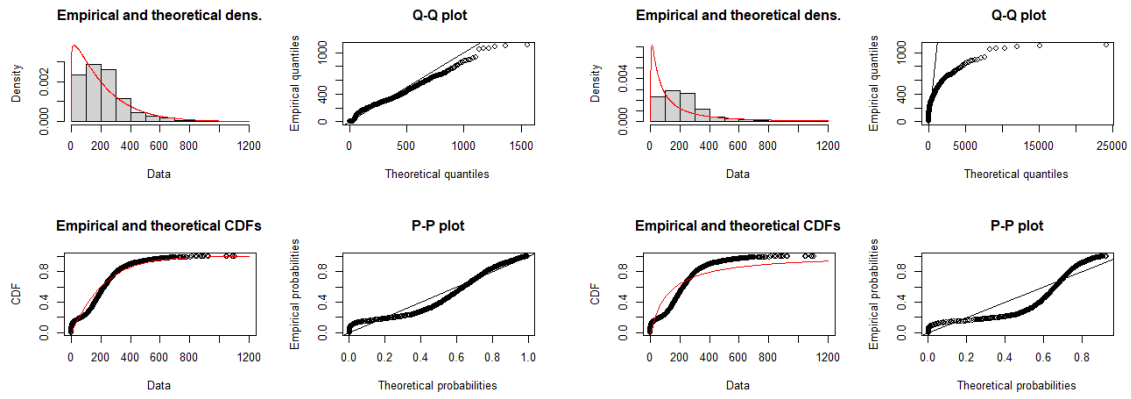
Fördelning	Parameter	ML uppskattning	95% Konfidensintervall
Exponential	Intensitet	0.0047	[0.0045, 0.0050]
Weibull	Form	1.0790	[1.0444, 1.1162]
	Skal	217.6841	[208.8630, 226.7961]
Gamma	Form	0.9754	[0.9252, 1.0287]
	Intensitet	0.0046	[0.0043, 0.0049]
Lognormal	Medellog	4.7661	[4.7031, 4.8314]
	Sdlog	1.5241	[1.4798, 1.5677]

Tabell 6: Parameterskattningar och bootstrap 95% konfidensintervall för parametrar för de olika fördelningarna för sackadtider med 1000 iterationer.

Fördelning	Parameter	ML uppskattning	95% Konfidensintervall
Exponential	Intensitet	0.0152	[0.0146, 0.0159]
Weibull	Form	1.2118	[1.1755, 1.2545]
	Skal	70.8245	[68.1720, 73.3194]
Gamma	Form	1.7471	[1.6544, 1.8525]
	Intensitet	0.0266	[0.0250, 0.0284]
Lognormal	Medellog	3.8723	[3.8409, 3.9060]
	Sdlog	0.7504	[0.7275, 0.7726]

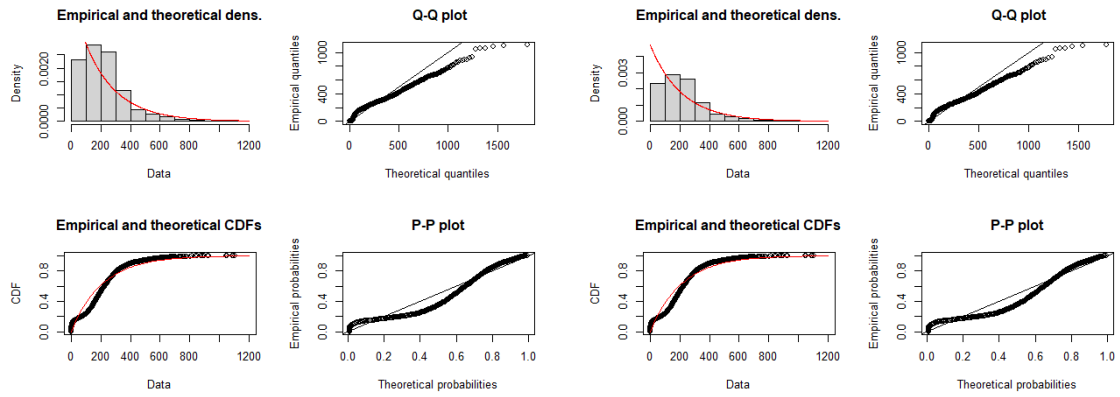
Tabell 7: Parameterskattningar och bootstrap 95% konfidensintervall för parametrar för de olika fördelningarna för sackadlängder med 1000 iterationer.

Fördelning	Parameter	ML uppskattning	95% Konfidensintervall
Exponential	Intensitet	0.0073	[0.0073, 0.0075]
Weibull	Form	1.0877	[1.0711, 1.1021]
	Skal	140.8287	[138.3803, 143.3929]
Gamma	Form	1.1556	[1.1307, 1.1821]
	Intensitet	0.0085	[0.0083, 0.0087]
Lognormal	Medellog	4.4234	[4.4036, 4.4427]
	Sdlog	1.1035	[1.0907, 1.1170]



(a) Diagnostiska plottar för weibullfördelning av fixeringstider

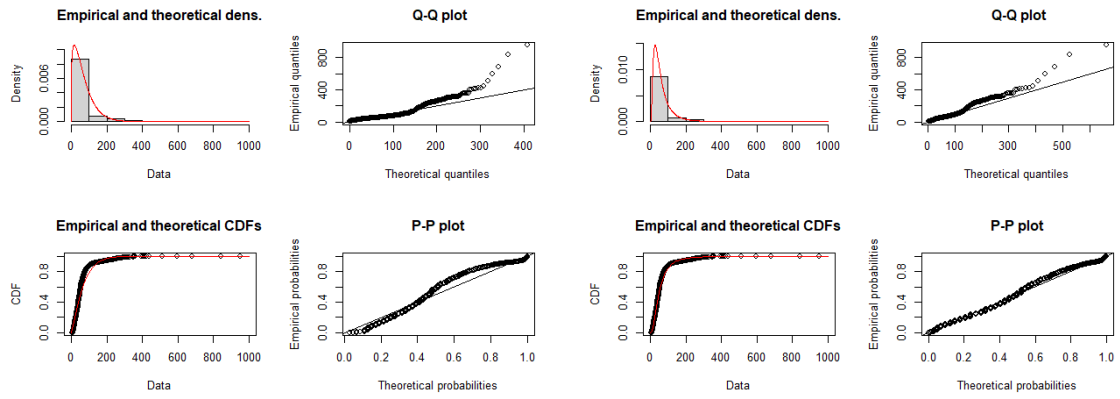
(b) Diagnostiska plottar för lognormalfördelning av fixeringstider



(c) Diagnostiska plottar för gammalfördelning av fixeringstider

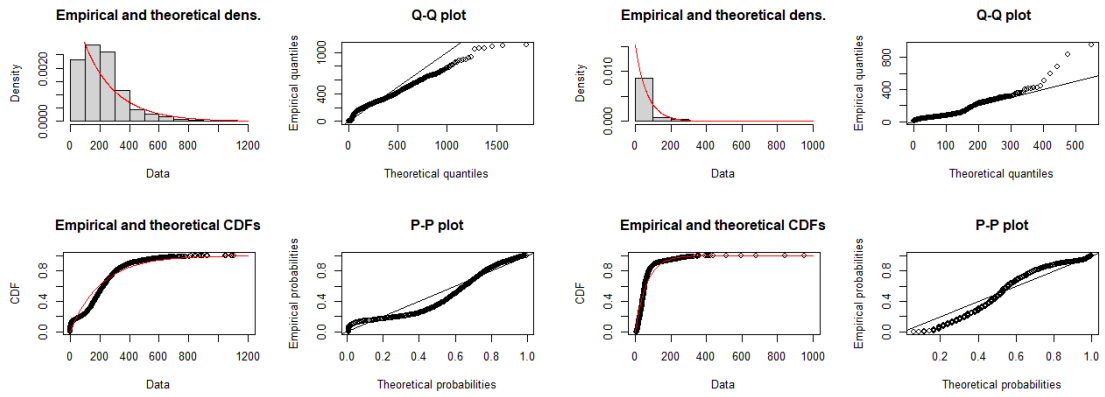
(d) Diagnostiska plottar för exponentialfördelning av fixeringstider

Figur 10: Diagnostiska plottar av fixeringstider, inklusive Q-Q-plot, P-P-plot samt empiriska fördelningsfunktionen mot den teoretiska fördelningsfunktionen.



(a) Diagnostiska plottar för weibullfördelning av sackadtider

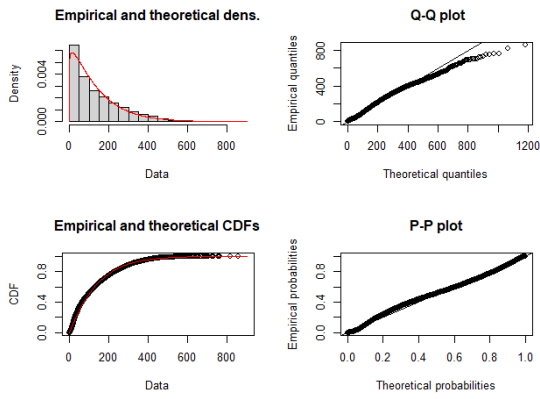
(b) Diagnostiska plottar för lognormalfördelning av sackadtider



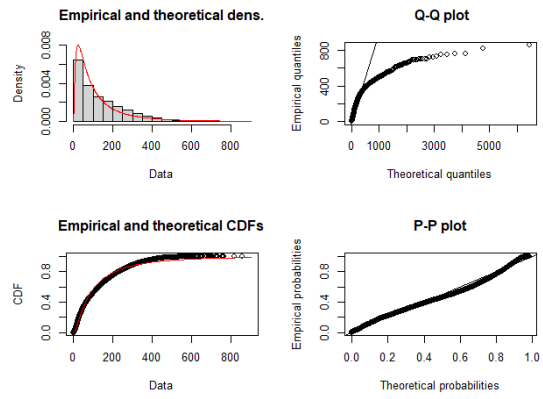
(c) Diagnostiska plottar för gammafördelning av sackadtider

(d) Diagnostiska plottar för exponentialfördelning av sackadtider

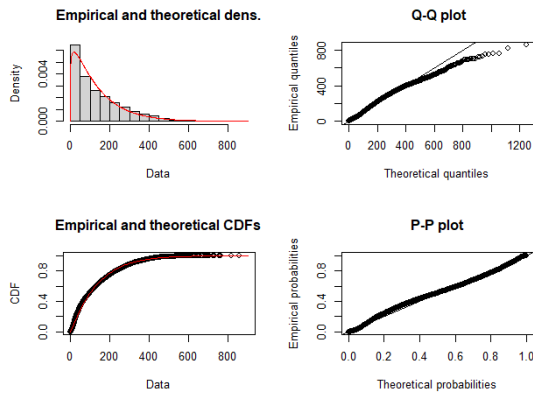
Figur 11: Diagnostiska plottar av sackadtider, inklusive Q-Q-plott, P-P-plott samt empiriska fördelningsfunktionen mot den teoretiska fördelningsfunktionen.



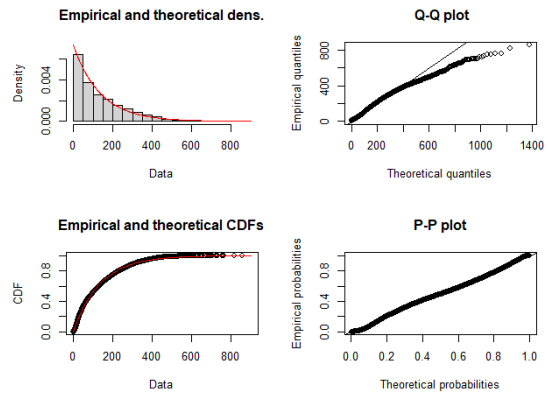
(a) Diagnostiska plottar för weibullfördelning av sackadlängder



(b) Diagnostiska plottar för lognormalfördelning av sackadlängder



(c) Diagnostiska plottar för gammafördelning av sackadlängder



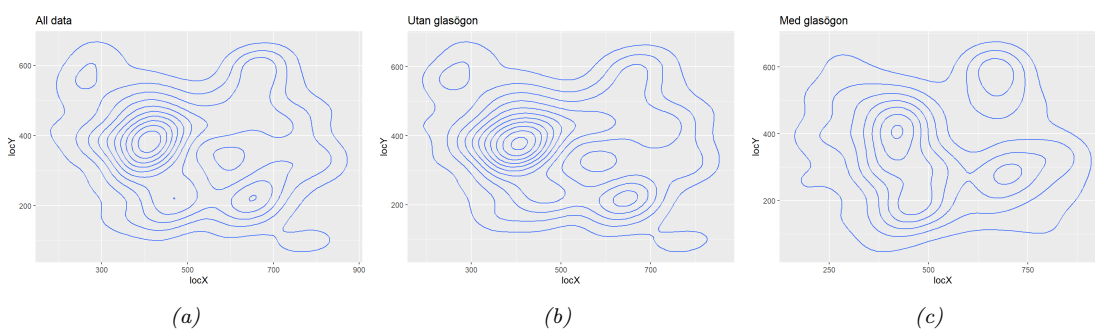
(d) Diagnostiska plottar för exponentialfördelning av sackadlängder

Figur 12: Diagnostiska plottar av sackadlängder, inklusive Q-Q-plot, P-P-plot samt empiriska fördelningsfunktionen mot den teoretiska fördelningsfunktionen.

C Glasögons påverkan på fixeringspunktmönster

I studieprotokollet framkommer det att fem (5) av 20 försökspersoner använt glasögon under undersökningen och att reflektioner i glasen kan ha påverkat ögoninspelningen genom att reflektioner uppstår i glasen. Vidare använde tre (3) av 20 försökspersoner linser, men inga effekter av detta har noterats i försöksprotokollet. Vi jämför här därför kvalitativt hur glasögonbärarna och potentiella mätfel på grund av deras glasögon påverkar fixpunktens placering i tavlan, men betraktar deltagarna med linser i samma grupp som de utan både glasögon och linser.

I figur 13 visas intensitetsytor av fixeringspunkter baserat på all data och de som inte använde respektive använde glasögon vid försökstillfället. Av en kvalitativ visuell analys framgår det att skillnaden mellan intensitetsytorna är små och modelleringen baseras därför på samtliga individers fixeringspunktmönster.



Figur 13: Intensitetsytor för alla individer, utan glasögon och med glasögon

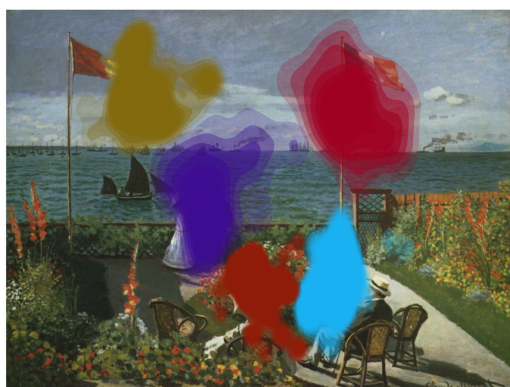
D ROI:er för samtliga individer



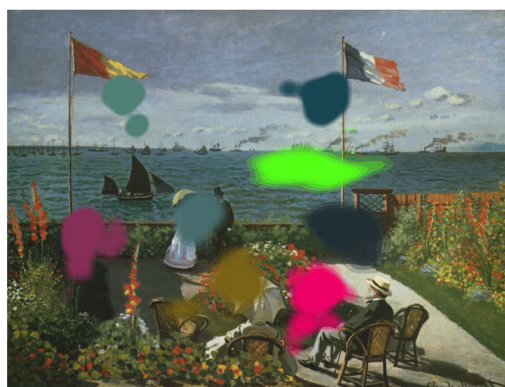
(a) ROI för individ 1



(b) ROI för individ 2



(c) ROI för individ 3



(d) ROI för individ 4



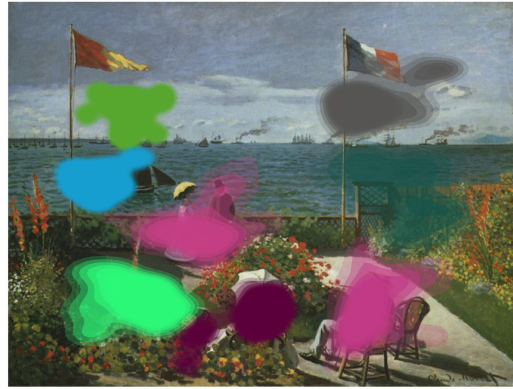
(e) ROI för individ 5



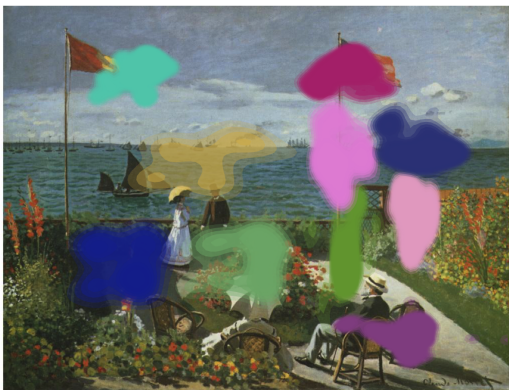
(f) ROI för individ 6



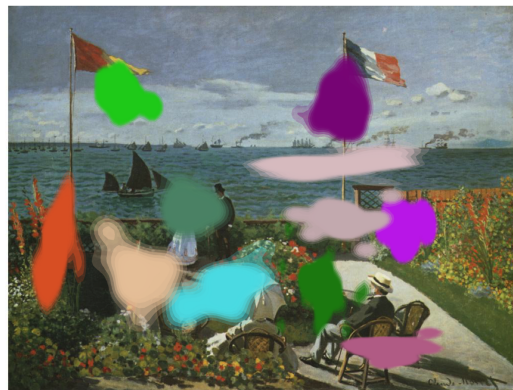
(g) ROI för individ 7



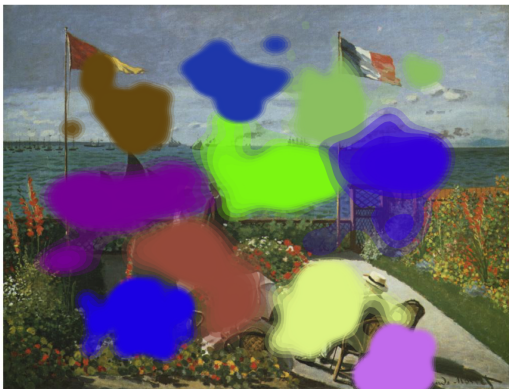
(h) ROI för individ 8



(i) ROI för individ 9



(j) ROI för individ 10



(k) ROI för individ 11



(l) ROI för individ 12



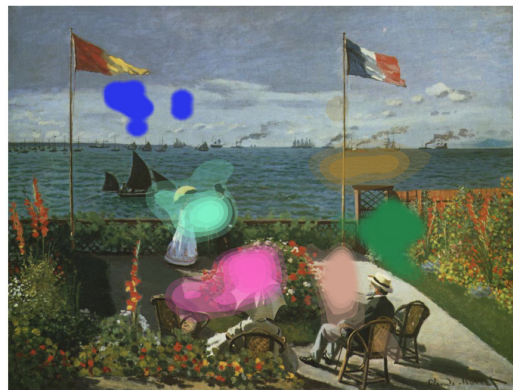
(m) ROI för individ 13



(n) ROI för individ 14



(o) ROI för individ 15



(p) ROI för individ 16



(q) ROI för individ 17



(r) ROI för individ 18



(s) ROI för individ 19



(t) ROI för individ 20

E Realisationer av modell B

