

STATISTICAL METHODS FOR MODELING THE SPATIAL STRUCTURE ON THE VISUAL
FIELD IN GLAUCOMA PROGRESSION RESEARCH

Samuel I. Berchuck

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biostatistics in the Gillings School of Global Public Health.

Chapel Hill
2018

Approved by:

Amy Herring

Joshua Warren

Jean-Claude Mwanza

Jason Fine

Joseph Ibrahim

©2018
Samuel I. Berchuck
ALL RIGHTS RESERVED

ABSTRACT

Samuel I. Berchuck: Statistical methods for modeling the spatial structure on the visual field in glaucoma progression research
(Under the direction of Amy Herring and Joshua Warren)

Diagnosing glaucoma progression is critical for limiting irreversible vision loss. A common method for assessing glaucoma progression uses a longitudinal series of visual fields (VF) acquired at regular intervals. VF data are characterized by a complex spatiotemporal structure due to the data generating process and ocular anatomy. Thus, advanced statistical methods are needed to make clinical determinations regarding progression status. In this dissertation, we introduce new modeling techniques that produce flexible spatial dependency structures within the framework of hierarchical Bayesian spatial models. The developed methodology is applied to VF data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry.

In chapter 2, we work within the framework of boundary detection and introduce a spatiotemporal boundary detection model that allows the underlying anatomy of the optic disc to dictate localized spatial structure on the VF. We show that our new method provides insight into vision loss that improves diagnosis of glaucoma progression in actual glaucoma patients. Simulations are presented, showing the proposed methodology is preferred over existing spatial methods for VF data. An R package `womblR` is provided that implements the method.

Chapter 3 aims to introduce the modeling framework from chapter 2 to the ophthalmology community. An optimal form of the metric is established and compared with standard methods for assessing glaucoma progression using a statistical diagnostic framework. In particular, we demonstrate the added value of using the novel metric in addition to established prediction models based on standard operating characteristics. Finally, we detail the procedure for implementing our novel technique in the clinical setting.

In chapter 4, we present a framework that brings together vital aspects of glaucoma management, i) prediction of future VF sensitivities, ii) predicting the timing and location of future vision loss, iii)

making clinical decisions regarding progression, and, iv) incorporation of anatomical information to create plausible data-generating models. We show that our method improves prediction and estimation of progression, and simulations are presented, showing the proposed methodology is preferred over existing models for VF data. An R package called `spCP` is provided that implements the method.

For Ania, who always gives me hope.

ACKNOWLEDGEMENTS

First and foremost, I want to thank my advisors Amy Herring and Josh Warren. Thanks to Amy, the first five years of my study were graciously supported by a National Institute of Environmental Health Sciences training grant (T32ES007018). Throughout my career at UNC Amy has provided unparalleled mentorship, always leading me in the right direction. I'm grateful for her expertise and guidance that have helped me navigate professional relationships, job applications, and the world of biostatistics over the past six years. Without her support and vision, this dissertation would never have been possible, especially because she introduced me to Josh. It has been an honor being Josh's first Ph.D. student. Somehow we are always on the same wavelength, and our interests are nearly perfectly aligned. I don't know what I would have done without his constant, day-to-day advising, theorizing, writing, and revising.

In addition to my advisors, I would like to thank my committee members Jean-Claude Mwanza, Jason Fine, and Joe Ibrahim. Jean-Claude has proven to be a fantastic colleague and I am thankful for the time he has invested in teaching me the basics of glaucoma. Before serving on my committee, both Drs. Fine and Ibrahim were fantastic teachers and I am appreciative of their efforts in the completion of this dissertation. Finally, I would like to thank my colleagues at the Duke Clinical Research Institute, in particular Karen Pieper for helping me develop professionally.

This dissertation would not have been possible without data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry, Perth, Western Australia. The data was generously provided by Brigid D. Betz-Stablein (School of Medical Sciences, University of New South Wales), William H. Morgan (Lions Eye Institute, University of Western Australia), Philip H. House (Lions Eye Institute, University of Western Australia), and Martin L. Hazelton (Institute of Fundamental Sciences, Massey University).

My time at UNC was made far more enjoyable in large part due to the many friends that became a part of my life. I am grateful for the time we spent at Weaver Street, dinners at IP3, and our memorable trips to the beach. I am indebted to them for their assistance, encouragement and

friendship in more ways than I knew were possible. I look forward to our continued relationships as colleagues.

I would like to thank my family for their love and encouragement throughout this process. To my parents, I am fortunate that you raised me with an eagerness towards learning that has not only provided me with aspirations towards academic pursuits, but also brightens life in times of darkness. To my older brothers, your work ethic and unconditional support have provided endless inspiration. Finally, to my wife. Together we have have learned resilience and an understanding that life has way more imagination than we do. Through trials and tribulations, we have become stronger. Any day spent with you is my favorite day. I love you.

TABLE OF CONTENTS

LIST OF TABLES	xii
LIST OF FIGURES	xv
LIST OF ABBREVIATIONS	xviii
CHAPTER 1: LITERATURE REVIEW	1
1.1 Introduction.....	1
1.2 Glaucoma Research	1
1.2.1 Diagnosing Glaucoma Progression	2
1.2.2 Visual Field Data	3
1.2.3 Statistical Methods for Visual Field Data	6
1.3 Modeling Spatially Referenced Data	9
1.3.1 Areal Data	10
1.3.2 Univariate Areal Data Models	10
1.3.3 Multivariate Areal Data Models	12
CHAPTER 2: DIAGNOSING GLAUCOMA PROGRESSION WITH VISUAL FIELD DATA USING A SPATIOTEMPORAL BOUNDARY DETECTION METHOD ...	14
2.1 Overview	14
2.2 Introduction.....	14
2.3 Visual Field Data.....	16
2.4 Spatial Boundary Detection.....	19
2.5 Methods	21
2.5.1 Observational Model	21
2.5.2 Neighborhood Model	23

2.5.3	Temporal Model	23
2.5.4	Specifying Hyperprior Distributions	24
2.5.5	Prediction	25
2.6	Analysis of Visual Field Data and Glaucoma Progression Risk	26
2.6.1	Study Data	26
2.6.2	Accounting for Zero-Truncation	26
2.6.3	Creating a Dissimilarity Metric	28
2.6.4	Model Estimation	28
2.6.5	Diagnosing Glaucoma Progression	29
2.6.6	Sensitivity Analyses	32
2.7	Simulation Study	32
2.7.1	Data Generation	32
2.7.2	Results	33
2.8	Discussion	35
2.9	Supplementary Material	37
CHAPTER 3: IMPROVED DETECTION OF VISUAL FIELD PROGRESSION USING A SPATIOTEMPORAL BOUNDARY DETECTION METHOD		38
3.1	Overview	38
3.1.1	Purpose	38
3.1.2	Methods	38
3.1.3	Results	38
3.1.4	Conclusions	39
3.2	Introduction	39
3.3	Methods	41
3.3.1	Data Description	41
3.3.2	Spatiotemporal Boundary Detection	41
3.3.3	Model Comparison	44
3.3.4	Assessing Early Stage Progression	46

3.4	Results	46
3.4.1	Diagnostic Capability of the Metrics	46
3.4.2	Correlation Between Methods	47
3.4.3	Combination of Methods.....	47
3.4.4	Determination of Progression Over Time	49
3.5	Discussion	54
CHAPTER 4: A SPATIALLY VARYING CHANGE POINT MODEL FOR AS-		
SESSING GLAUCOMA PROGRESSION FROM VISUAL FIELD DATA		56
4.1	Overview	56
4.2	Introduction	56
4.3	Change Point Theory	59
4.4	Methods	60
4.4.1	Spatial Change Point Model	61
4.4.2	Multivariate Dissimilarity Metric Spatial Process	62
4.4.3	Hyperpriors	63
4.5	Assessing Visual Field Progression Using Change Points	64
4.5.1	Study Data	64
4.5.2	Change Point Model for Visual Field Data	64
4.5.3	Model Estimation	66
4.5.4	Model Comparison	67
4.5.5	Diagnosing Progression Using Change Points	68
4.6	Simulation Study	72
4.6.1	Data Generation	72
4.6.2	Results	74
4.6.3	Estimation of the Observed Change Point	77
4.7	Discussion	80
4.8	Supplementary Materials	82
APPENDIX A: SUPPLEMENTARY MATERIALS FOR CHAPTER 2		83

A.1	Introduction.....	83
A.2	Implementation of the Model	83
A.3	Prediction.....	88
A.4	Sensitivity to the Mis-specification of ρ	89
A.5	Sensitivity Analysis in the Visual Field Analysis	89
A.6	Additional Figures	90
APPENDIX B: SUPPLEMENTARY MATERIALS FOR CHAPTER 4		92
B.1	Change Point Likelihood	92
B.2	Full Conditionals.....	93
BIBLIOGRAPHY		99

LIST OF TABLES

2.1	Assessing the diagnostic capability of VF metrics. Each metric is regressed against the clinical assessment of progression using a logistic regression model with the slope estimates being displayed (*significance level of 0.05). “Mean CV” is the CV of the VF wide mean of DLS at each visit, PLR is the minimum p-value of the location specific regression slopes, and “Space CV” and “ST CV” represent the mean posterior CV of α_{tGH} from the Lee and Mitchell (2011) and our spatiotemporal models, respectively. Each of the metrics are standardized.	31
2.2	Results from simulation study estimating the posterior mean of the CV of α_{tGH} in setting A: no temporal correlation, no cross-covariance, B: no temporal correlation, cross-covariance, C: temporal correlation, no cross-covariance, and D: temporal correlation, cross-covariance. Each setting is also implemented for the median (7) and maximum (21) number of VF visits. Each reported estimate is based on 1,000 simulated datasets.	34
3.3	Summary of criteria included in each diagnostic method. For each of the metrics, the covariates that are listed are included in a logistic regression model used to predict clinical assessment of progression. The resulting predicted probabilities of progression are used as the progression metric.	45
3.4	Assessing the diagnostic capability of the metrics. Each metric is regressed against the clinical assessment of progression, both with and without the global index (GI). AIC, AUC and pAUC. The pAUC is limited to the clinically relevant range of specificity, 85-100%. Also reported are p-values corresponding to hypothesis tests *) for each metric marginally, and †) in addition to GI. The bold cells indicate a significant improvement in AUC or pAUC over the GI model at the $\alpha = 0.05$ level.	47
3.5	Correlation matrix and hypothesis test results for the diagnostic metrics. The estimated Pearson correlations and p-values between the metrics are presented below and above the diagonal, respectively. Bold values indicate that the correlation does not statistically differ from zero at the $\alpha = 0.05$ level of significance.	48
3.6	Estimated coefficients from the logistic regression model combining the global index (GI) and STBound.	48
3.7	Estimates of AUC and pAUC presented at yearly intervals. These raw estimates correspond to the LOESS smoothed curves presented in Figure 3.10. Bold cells indicate a significant improvement in either AUC or pAUC over the corresponding values of the global index (GI) at the $\alpha = 0.05$ level.	49
3.8	Sensitivity (sens.) and specificity (spec.) for each metric using the clinical definition for thresholds.	54

4.9 Model fit and prediction diagnostics for model comparison. Model fit is determined using the deviance information criterion (DIC) and effective number of parameters, p_D . Prediction capability is determined using the mean squared prediction error (MSPE) and continuous ranked probability score (CRPS). The ability of each model to diagnose progression is assessed using a metric dependent on the latent CPs. For each model with a latent CP process, the metric $\max_i\{P[\theta(\mathbf{s}_i) < t_\nu]\}$ is calculated and regressed against the clinically determined progression status. Diagnostic capability is measured using AIC, area under the curve (AUC), and the p-value for each measure. 68

4.10 Model fit diagnostics for simulation study outlined in Section 4.6.1. The diagnostics included are DIC and p_D and are presented for each of the models introduced in Section 4.5.4. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets. 74

4.11 Prediction diagnostics for estimation of the VF visits at times 0.75 and 1. The diagnostics included are CRPS and MSPE and are presented for each of the models introduced in Section 4.5.4. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets. 75

4.12 Summary diagnostics for estimation of the observed, $\theta(\mathbf{s})$, and latent, $\eta(\mathbf{s})$, CPs. The diagnostics included are bias, MSE, and EC (95% nominal) and are presented for each of the models introduced in Section 4.5.4 that are capable of estimating either the observed or latent CP. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets. 76

4.13 Summary diagnostics for estimation of the observed CPs in the simulation detailed in Section 4.6.3. The diagnostics included are bias, MSE, and EC (95% nominal) and are presented for each of the models that are capable of estimating the continuous observed CP. Standard errors (SE) correspond to the variation across all VF locations. Each reported estimate is based on 1,000 simulated datasets. 79

A.1 The effect of mis-specifying the spatial correlation parameter ρ on model performance. The table displays MSE and EC for both mis-specified and concordant settings. Each reported estimate is based on 1,000 simulated datasets. 89

A.2 Sensitivity analysis to various assumptions in analysis of visual field data and glaucoma progression risk. For each assumption new values are presented with risk presented for comparison. Assumptions include i) upper bound for $\log(\alpha_{tGH}), v_1$, ii) correlation structure, $\Sigma(\phi)$, iii) mean hyperparameter for $[\delta]_1, [\mu_\delta]_1$, and iv) criterion for the bounds of ϕ 90

LIST OF FIGURES

1.1	Visual field output for a normal eye.	4
1.2	Retinal nerve fiber layer. The fibers can be seen extending from throughout the retina to the optic disc. (Note: the large “fibers” are blood vessels). Printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).	6
1.3	Example of stable (left) and progressing (right) VFs. Each cell represents the observed DLS at each location over visits, with the red line representing a linear regression line. The stable eye has constant trend over time, while the progressing eye has cells with large negative slopes over visits and big variation within DLS. In the southern hemisphere of the progressing eye there are cells with zero truncation.	7
2.4	Demonstrating the VF data generating mechanism, a) A patient (the dissertation author) participates in a VF examination. b) The HFA-II stimulus is absorbed by the retina and transmitted through the optic disc, and along the optic nerve to the brain. c) Each VF test point corresponds to a location on the underlying RNFL. d) The corresponding nerve fiber enters the optic disc at 77° . e) Each VF test produces a map that shows the intensity a stimulus is detected at each test location. The gray locations represent the blind spot. f) Over time the patient accrues VF tests. The eye in b) is printed with permission from the Lions Eye Institute and the RNFL in c) is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).	17
2.5	Top row: Example of stable (left) and progressing (right) VFs. The observed DLS across all visits for a VF location is shown in each cell. Bottom row: Posterior mean estimates of α_{tGH} plotted across time along with the average of the spatial adjacencies across different sectors of the VF for the same two patient eyes. Sector 1 (Inferior) is the superior VF, corresponding to the inferior sector of the optic disc, and Sector 2 (Superior) is the inferior VF, corresponding to the superior sector of the optic disc.	27
2.6	Demonstrating the Garway-Heath dissimilarity metric across the VF (right). The angles are defined as the degree at which the underlying nerve fiber enters the optic disc for each location. The edges and corners are shaded to represent the distance between bordering locations on the VF, where darker shading represents a larger distance. The RNFL is displayed to demonstrate the similarity of the pattern that appears on the VF by using the Garway-Heath angles (left). The RNFL (left) is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).	29

3.7	The metric STBound describes variability in the spatial structure of the VF across time. The statistical model incorporates the anatomy of the eye, where the distance at which the underlying nerve fibers enters the optic disc dictates spatial correlation between data from two VF locations. These angles, measured by Garway-Heath et al. (2000), are displayed across the VF (right). The edges and corners are shaded to represent the distance between bordering locations on the VF, where darker shading represents a larger distance. The RNFL is displayed to demonstrate the similarity of the pattern that appears on the VF by using the Garway-Heath angles (left). The RNFL (left) is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).	42
3.8	Longitudinal series of VFs for one progressing (left) and stable (right) eye are presented in the bottom row. Each cell represents the observed VF sensitivities over all visits. The progressing eye has more variability across the VF and in particular in the region between the superior and inferior sectors. In the top row, the estimated parameter that describes the spatial correlation in the VF data, α , is plotted over time and the corresponding values of STBound (i.e., predicted probability of progression) are given as 0.99 and 0.17 for the progressing and stable eye, respectively. For these two patients STBound indicates that eye a) is much more likely to be progressing.	43
3.9	Receiver operating characteristic curves for statistically significant diagnostic metrics.	50
3.10	Performance of diagnostic metrics in the initial years from baseline visit, AUC (a) and pAUC (b). Estimates are presented as smooth curves using the LOESS method. The horizontal dashed line indicates the value at which a metric does not contribute any diagnostic information.	51
3.11	Demonstrating the performance of sensitivity (a) and specificity (b) in the initial years from baseline visit. Both sensitivity and specificity are obtained by using a clinically motivated threshold defined to maximize sensitivity, while forcing the specificity to be no smaller than 85%. Estimates are presented as smooth curves using the LOESS method.	52
3.12	Density estimates of the time of first diagnosis for each model in days from baseline visit along with the mean number of days.	53
4.13	Example VF follow-up for a patient with glaucoma. At each location, the time series of the observed threshold sensitivities is plotted. The trends and clustering of the time series motivate the use of spatially varying CPs.	65

4.14	Demonstrating the ability of the Spatial CP model to identify CPs that have yet to occur. At each cell, the observed DLS are presented for an example patient with nine visits. Based on only the first seven visits, the posterior estimates of the CPs for the Spatial CP and Non-spatial (L) models are presented as density estimates in blue, and green lines, respectively. The estimated CP based on the Spatial CP model using all of the nine VF visits are presented as dashed vertical lines. The DLS and CPs are shown in red if they occur after the seventh visit and thus are not observable at the time the posterior densities were estimated.	70
4.15	Output generated from the Spatial CP model. The plot on the left presents the estimated mean (with 95% credible intervals) and CPs for the same example patient in Figure 4.14. The mean process is presented in red and the change points are vertical blue lines. The panel on the right presents the posterior probability of an observed CP at each location on the VF.	71
4.16	Simulation study comparing the performance of the spatial CP and non-spatial CP models in estimating the CPs across the VF. The true value of the CP is presented as a dotted black line and the threshold sensitivity values for an example simulated dataset are presented as gray dots. The posterior estimates of the CPs for the Spatial CP, Non-spatial CP (D), Non-spatial CP (C), and Non-spatial CP (L) models are presented as density estimates in red, blue, green, and purple lines, respectively.	78
A.1	Trajectories of the posterior mean estimates of α_{tGH} over time for all patients, subset by progression status.	91
A.2	Pairwise scatterplots between standardized diagnostic metrics. Included are estimated Pearson correlations, corresponding p-values from the hypothesis that the correlations are equal to zero, and regression trend lines.	91

LIST OF ABBREVIATIONS

AIC	Akaike information criterion
AR	Autoregressive
asb	Apostilbs
AUC	Area under the curve
CAR	Conditional autoregressive
CP	Change point
CV	Coefficient of variation
dB	Decibels
DLS	Differential light sensitivity
EC	Empirical coverage
GI	Global index
GLMM	Generalized linear mixed model
GMRF	Gaussian Markov random field
GPA	Guided progression analysis
HFA	Humphrey Field Analyzer
IOP	Intraocular pressure
LOESS	Local regression
MCAR	Multivariate conditionally autoregressive
MCMC	Markov chain Monte Carlo
MD	Mean deviation
MRF	Markov random field
MS	Mean slope
MSE	Mean squared error
OAG	Open-angle glaucoma
PLR	Point-wise linear regression
PPD	Posterior predictive distribution
RGC	Retinal ganglion cell
RNFL	Retinal nerve fiber layer

ROC	Receiver operating curve
SAP	Standard automated perimetry
SITA	Swedish Interactive Thresholding Algorithm
SPROG	Spatial PROGression
ST	Spatiotemporal
STBound	Spatiotemporal boundary detection
VF	Visual field
VFI	Visual field index

CHAPTER 1: LITERATURE REVIEW

1.1 Introduction

Throughout the course of this dissertation, we study patients diagnosed with glaucoma and attempt to describe their disease progression using visual field (VF) data. In Chapter 1, we begin by introducing glaucoma disease pathology and characterize progression. Then, we present VF data in detail and introduce statistical methods currently used for detecting glaucoma progression. We finish Chapter 1, with an introduction to spatial statistics, focusing on the study of areal data methods, both univariate and multivariate.

1.2 Glaucoma Research

Glaucoma is an optic neuropathy that results in damage to the optic nerve and subsequent deterioration of the VF (Foster et al., 2002). In 2012 there were approximately 60 million people with glaucomatous optic neuropathy worldwide and an estimated 8.4 million people blind as a result. These numbers are expected to increase to 80 million and 11.2 million, respectively, by 2020 (Cook and Foster, 2012). The most common form of glaucoma is primary open-angle glaucoma (OAG). The biological basis of the disease is not fully understood, however a recognized risk factor for primary OAG is increased intraocular pressure (IOP). IOP is the fluid pressure in the eye and can be treated with eye drops or surgery (AGIS Investigators, 2000). If IOP increases and goes untreated the optic nerve becomes damaged, resulting in permanent vision loss. Since the impairment caused by glaucoma is irreversible, and valid treatments exist for the most common forms of glaucoma, early detection is essential (Weinreb and Khaw, 2004).

1.2.1 Diagnosing Glaucoma Progression

Patients diagnosed with glaucoma are at risk for vision loss and must be monitored for disease progression. During progression, glaucoma increases in severity and leads to irreversible blindness. Early detection of progression is essential for clinicians to properly manage treatment and follow-up in order to minimize visual disability. It remains one of the most challenging aspects of glaucoma management, since it is difficult to distinguish progression and variability due to natural degradation (Vianna and Chauhan, 2015). There are numerous techniques to diagnose progression, all of which detect structural or functional changes in either the i) optic nerve head, ii) retinal nerve fiber layer (RNFL), iii) retinal ganglion cell (RGC) layer or iv) VF. Each of these techniques attempts to assess glaucoma in varying ways (Katz et al., 1997). In order to understand the relationship between these techniques it is important to be aware of basic ocular physiology.

In the human eye the retina acts as a lens, absorbing visual information and transmitting it to the brain through the optic nerve. The optic nerve consists of a bundle of fibers that extend from the retina to the optic nerve, creating the RNFL. The fibers from the retina converge at the optic disc, which is the starting point of the optic nerve. More specifically, these nerve fibers are axons of RGCs. The RGCs disperse across the retina and use photoreceptors to encode visual information and then transmit it to the brain along their axons (Davson, 2012). Glaucoma damages the optic disc, killing RGCs. Damage to specific regions of the optic disc corresponds to loss of RGCs, whose axons enter the damaged region. Thus, vision loss across the retina is a result of systemic ganglion cell loss, stemming from damage to the optic disc. This physiology motivates the three most common methods of glaucoma detection.

We know that a method to assess glaucoma progression must be capable of assessing damage to the optic disc (Heijl et al., 2002). The first method directly measures characteristics about the optic disc, while the other two explore adjacent pathways in the visual process. There are numerous properties of the optic disc that are measured. The most common metric is the cup-to-disc ratio, which measures the impact of increased IOP on the geometry of the optic disc (Klein et al., 1992). The second technique involves detecting structural changes in the RNFL, where methods are developed to correlate changes in the RNFL to damage on the optic disc (Medeiros et al., 2009). The third method involves assessing change in the VF. The VF represents the visualization of images by the

brain, transmitted by ganglion cells. Thus, detecting deterioration in the VF is a proxy for assessing optic disc damage. All of these methods are effective to varying degrees, with structural assessment often allowing earlier detection of glaucoma progression, while functional evaluations are generally preferred for slower progressing patients. In this dissertation, we focus on a diagnostic technique that determines progression from a longitudinal series of VFs.

1.2.2 Visual Field Data

The VF is the spatial array of visual sensations that your brain perceives as vision (Smythies, 1996). In order to approximate the VF in patients with glaucoma, standard automated perimetry (SAP) is used (Chauhan et al., 2008). SAP constructs a VF image by assessing a patient's response to varying levels of light sensitivities (Johnson et al., 1993). Patients are instructed to focus on a central fixation point as light is introduced randomly in a preceding manner over a grid on the patient's field of vision. When light is observed, the patient presses a trigger and the current light intensity is recorded. This process is repeated until the entire VF is imaged. An example printout of a VF examination is presented in Figure 1.1 for a normal eye. This VF examination was conducted using the most common form of SAP, the Humphrey Field Analyzer-II (HFA) (Carl Zeiss Meditec Inc., Dublin, CA). The 24-2 test pattern was used with the Swedish Interactive Thresholding Algorithm (SITA) which tests 54 locations. However, two of these locations correspond to a natural blind spot on the optic disc, resulting in 52 informative locations. The SITA standard strategy is preferred, since it offers high accuracy and relatively short test times (Heijl et al., 2012).

In the upper left corner of Figure 1.1 the threshold sensitivities are displayed in a gridded fashion. The two locations corresponding to the blind spot straddle the small triangle on the right. SAP measures the differential light sensitivity (DLS) across the VF. The measurement represents a contrast between the background of the machine, normally white, and the light stimulus. The intensity of the stimulus is initially similar to the background, but as the intensity increases the contrast grows and the probability of detecting the stimulus increases. These intensities are displayed in decibels (dB), where DLS values near 40 indicate good vision, while values near zero represent possible blindness. A grayscale map is also displayed that smooths the raw thresholds, where dark areas represent reduced sensitivity. In addition to the numerical threshold sensitivities, STATPAC, the built in software of the HFA, also outputs a selection of diagnostic measures.

Single Field Analysis

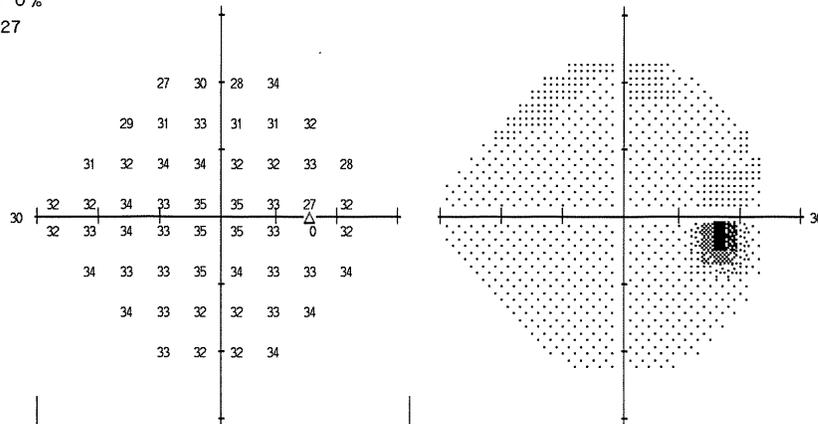
Eye: Right

Name: BERCHUCK SAMUEL	DOB: 09-05-1989
ID: T52081	

Central 24-2 Threshold Test

Fixation Monitor: Gaze/Blind Spot	Stimulus: III, White	Pupil Diameter: 4.6 mm	Date: 07-11-2016
Fixation Target: Central	Background: 31.5 ASB	Visual Acuity:	Time: 11:34 AM
Fixation Losses: 1/10	Strategy: SITA-Fast	RX: DS DC X	Age: 26
False POS Errors: 11 %			
False NEG Errors: 0 %			
Test Duration: 02:27			

Fovea: 39 dB



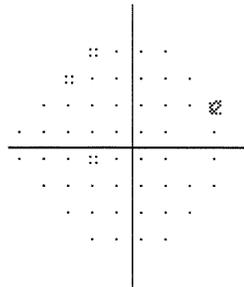
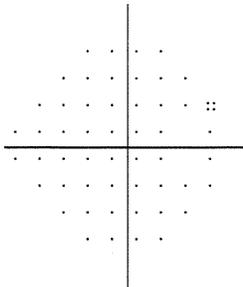
-2	0	-1	5
-2	-1	1	0
0	-1	1	-1
2	0	0	-1
3	1	1	-1
2	1	-1	1
2	1	-1	1
2	1	1	2

-4	-2	-3	3
-4	-3	-1	-2
-2	-3	-1	-1
0	-2	-2	-3
1	-1	-1	-3
0	-1	-3	-1
0	-2	-3	-1
0	-1	-1	0

GHT
Within Normal Limits
VFI 100%
MD +0.38 dB
PSD 1.30 dB

Total Deviation

Pattern Deviation



:: < 5%
⊗ < 2%
⊗ < 1%
■ < 0.5%



© 2010 Carl Zeiss Meditec
HFA II 750-42595-5.1.2/5.1.2

Figure 1.1: Visual field output for a normal eye.

Total deviation maps compare the observed thresholds to normal age matched versions and present both the raw deviations and probability map of any significant deviation. Pattern deviation is similar to total deviation, however, it corrects for changes due to the height of the hill of vision, such as caused by cataracts. In addition to these measures, STATPAC presents global indices of VF loss, mean deviation (MD), visual field index (VFI), and pattern standard deviation. MD is a weighted average of total deviations, with 0 indicating no deviation from normal and large negative values implying vision loss. VFI is a robust version of MD that is not affected by cataracts and is more sensitive to changes near the center of the field. VFI values of 100% are associated with normal vision. Finally, pattern standard deviation provides a measure of the unevenness of the field, where low values indicated a normal field (Heijl et al., 2012).

The VF test is a link between the structural damage on the optic disc and functional vision loss. In particular, the functional loss observed on the VF is a result of RGC axons death occurring at the optic disc. The pattern of VF deterioration follows the structure of the RNFL, seen in Figure 1.2. Because of this relationship, glaucomatous VF deterioration is generally characterized by localized trends. Since the VF does not map directly to the optic disc, the relationship has been studied extensively (Chauhan et al., 2001; Artes and Chauhan, 2005). Garway-Heath et al. (2000) quantified this relationship and for each VF location estimated the angle that the corresponding RGC axons enter the optic disc. The structure of the RNFL implies that adjoining locations on the VF do not necessarily map to the same location on the optic disc. For example, locations that straddle the hemisphere line have RGC axons that disperse to disparate positions on the optic disc. As such, the patterns of vision loss seen in patients correspond “to the anatomy of the nerve fiber layer of the retina and in its projections to the optic nerve” (Walsh, 2010). A detailed description of the VF data generating process is given in Section 2.3.

There are characteristics of the VF that are important to consider when studying glaucoma progression. Figure 1.3 presents a series of VFs for one patient diagnosed as stable and one diagnosed as progressing. In the figure, each cell represents the observed DLS at a location over visits and the red line is a regression line. There are many differences in these plots that differentiate the two patients. The stable patient has a consistent pattern of DLS over the entire VF, while the progressing patient has much more variability throughout, in particular within the inferior hemisphere. Within these cells there is a clear negative linear trend, along with an increase in variability in DLS. Furthermore, in the

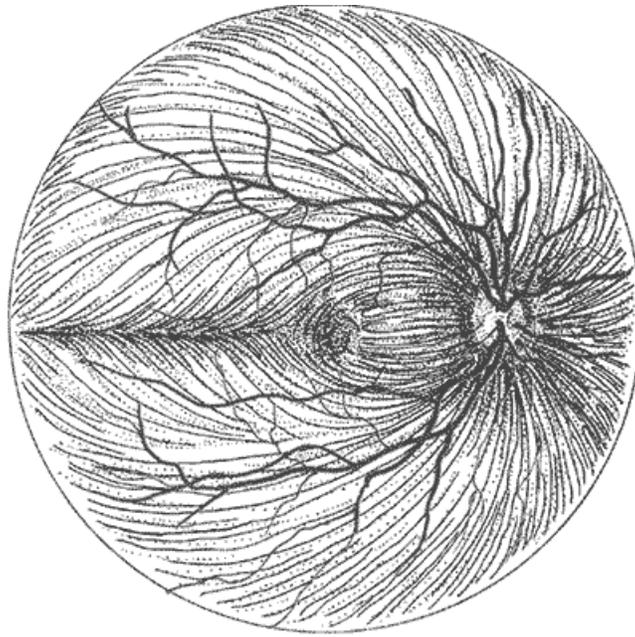


Figure 1.2: Retinal nerve fiber layer. The fibers can be seen extending from throughout the retina to the optic disc. (Note: the large “fibers” are blood vessels). Printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).

progressing patient the DLS become zero truncated, a characteristic of VF data due to the HFA-II testing process. Thus, to identify a progressing VF it is important to quantify the variability, both within and between locations, and any trends in the DLS over time. These characteristics are the basis for statistical methods that are developed to diagnose glaucoma progression using VF data.

1.2.3 Statistical Methods for Visual Field Data

VF testing is a common technique for assessing glaucoma progression, but there is currently no gold standard approach. Methods can be grouped in two categories, event-based and trend-based. In event-based analyses, new VF measurements are compared with a baseline, and change is considered significant if it exceeds expected deviation. Meanwhile, trend-based analyses model measurements over time, using some form of regression to estimate rates of change and statistical significance (Vianna and Chauhan, 2015). There are advantages and disadvantages to both classes of methods. Event-based methods possibly detect progression earlier, while trend-based methods are more robust since they incorporate all measurements (Vesti et al., 2003; Casas-Llera et al., 2009). The most well

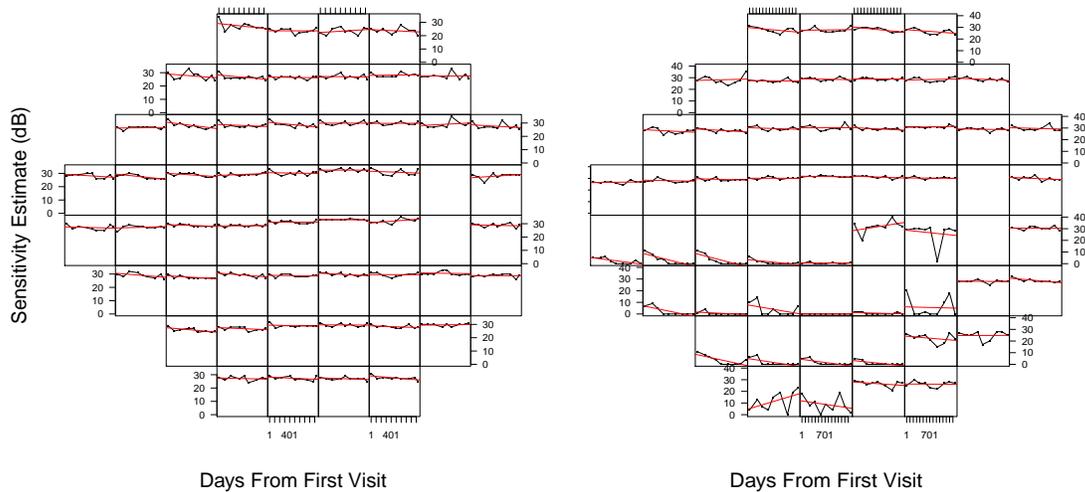


Figure 1.3: Example of stable (left) and progressing (right) VFs. Each cell represents the observed DLS at each location over visits, with the red line representing a linear regression line. The stable eye has constant trend over time, while the progressing eye has cells with large negative slopes over visits and big variation within DLS. In the southern hemisphere of the progressing eye there are cells with zero truncation.

known example of an event-based method is guided progression analysis (GPA) of the HFA. GPA compares each new test point to corresponding locations on two baseline tests, and significant change is assessed at each location using the criterion specified in the Early Manifest Glaucoma Trial (Heijl et al., 2003). Then, progression is determined if three or more locations have significant change at two consecutive visits. This method, characteristic of event-based methods, suffers from high false-positive results and is limited when the baseline measurements show variability (Wesselink et al., 2009; Nouri-Mahdavi and Caprioli, 2014).

Within the class of trend-based methods, modeling techniques can be classified as global or pointwise. Global methods summarize all measurements on the VF into a single metric at each visit, while pointwise methods analyze each individual location on the VF over time. The most common global trend-based method is MD (presented on the HFA-II printout), which calculates the average of pointwise differences in DLS between values from the test and a normal age matched VF. This method is advantageous due to its robust nature, however it can miss highly localized changes (Heijl et al., 1987). A commonly used trend-based pointwise method, PROGRESSOR, estimates trends at each VF location independently (Fitzke et al., 1996). This method is characteristic of a general class of point-wise linear regression models (PLR). Detection of progression in PLR

methods is dependent on p-values that correspond to a statistically significant slope of the location specific regression lines (Katz et al., 1997; Ernest et al., 2012). These methods can detect localized progression, however individual measurements are more variable than averages, causing difficulty in detecting true progression (Vianna and Chauhan, 2015). In recent years more advanced methods have been developed that introduce more flexible modeling frameworks.

The methods outlined above are simple, however they are widely used in practice because they are effective and easy to interpret. Throughout the past decade new methods have been developed using more advanced techniques. A class of machine learning classifiers have been used to identify progression based on variational Bayesian independent component analysis mixture models. These methods aim to distinguish between patterns of VF deterioration and provide meaningful clusters (Tucker et al., 2005; Goldbaum et al., 2005). Unfortunately, these methods do not improve diagnostic capabilities over standard techniques (Jampel et al., 2011). Another class of methods use spatial filters to smooth the spatial surface of the VF before assessing progression (Fitzke et al., 1995). These techniques estimate a correlation matrix based on the observed dependence structure and use it to smooth the VF surface. Initial attempts struggle to detect localized changes in DLS and are limited since they do not incorporate the anatomy of the eye in smoothing (Spry and Johnson, 2002). Gardiner et al. (2004) incorporated the anatomical structure of the VF introduced by Garway-Heath et al. (2000) and demonstrated improved predictive capabilities (Strouthidis et al., 2007). These methods are implemented in populations with relatively stable VFs and are untested in more advanced subjects. A fundamental issue with these methods is the lack of a formal inferential framework, due to the fixed specification of the spatial filter.

More recently, methods have been developed that incorporate the complex nature of VF data and provide a proper statistical framework for inference. Bryan et al. (2015) introduced a flexible two stage Bayesian hierarchical model that allows for location specific means and variances over visits and incorporated dependence within patients. This model is highly over-parameterized. However, it is based on assumptions that are reflective of the true data process. Unfortunately, this model ignores the spatial dependence on the VF. Zhu et al. (2014) introduced a method that incorporates spatial information of the VF and the underlying RNFL and proposed a likelihood that has a variance that increases inversely with the DLS. This model has many appealing attributes, however it uses a two-stage technique that ignores variability in likelihood parameter estimates. Finally, spatial

disease mapping techniques have been used to detect changes in DLS over time, controlling for spatial dependencies in the VF through the specification of a neighborhood structure (Betz-Stablein et al., 2013). This technique introduces a spatially varying intercept and slope over time, and is appealing since it accounts for spatial dependence in a fully inferential framework. However, the model assumes a constant variance, limiting model flexibility. These more recent advances represent an improvement in VF data modeling, but there is still a demand for newer advanced methods.

1.3 Modeling Spatially Referenced Data

In this dissertation, we develop statistical methods for diagnosing glaucoma progression taking into account the spatial nature of the VF. Throughout the work, the framework we use to account for correlations between VF locations is through the field of spatial statistics. In the context of spatial statistics, there are three commonly studied data types, geo-statistical (or point-referenced), areal (or lattice), and point-process (Gelfand et al., 2010; Cressie, 2015). These data types are characterized through the definition of their spatial location. Geo-statistical data are observed at specific locations across a smooth domain, areal data are defined over a set of finite regions, and locations are random in point-processes. Theory has been developed within each of these data types that leverages each definition of a location. The foundational assumption in spatial statistics states that dependence between observations weakens as the distance between locations increases (Banerjee et al., 2003). This assumption manifests differently in each of the data types. Geo-statistical models utilize proper covariance functions and model spatial decay, areal models often depend on Gaussian Markov random fields (GMRF) that induce neighborhood dependence across the region, and point-process models rely on clustering methods for generating spatial point patterns (Geman and Geman, 1984; Diggle et al., 2003; Illian et al., 2008). Throughout this dissertation we treat the spatial locations on the VF as areal units. The following section provides a detailed introduction to spatial analyses of areal referenced data using Bayesian methods. Bayesian methods are preferred, because they induce a fully inferential framework for modeling the spatial structure of the VF field.

1.3.1 Areal Data

Areal data is observed over regions, called areal units, and is sometimes referred to as discrete spatial modeling, since models are specified for a finite set of random variables. Areal data is spatially referenced, so that observations are defined as $Y(\mathbf{s})$, for location $\mathbf{s} \in D$, where D is a discrete space. The locations $\{\mathbf{s}_1, \dots, \mathbf{s}_m\}$ represent the m disjoint spatial regions that comprise the spatial domain, D . The complete set of observations is represented through the finite collection of random variables, $\{Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_m)\}$. Examples of areal data include various geographic data, such as states, counties and census tracts. In this dissertation, $Y(\mathbf{s})$ will represent the observed DLS at location \mathbf{s} on the VF.

The statistical machinery for areal data is distinct from classical geo-statistical modeling, since areal data has different underlying spatial assumptions. In particular, the concept of distance, which features heavily in geo-statistical processes, is meaningless in areal data. A new definition of spatial distance defines two areal units as similar if they are neighbors (i.e., share a border). It should be noted that the main goals for areal data analysis are to, i) smooth the data over the spatial surface and, ii) describe any spatial pattern. The idea of prediction in areal data is less obvious since presumably, data is recorded at every areal unit in the spatial domain.

1.3.2 Univariate Areal Data Models

We begin by presenting the most common class of areal data methods, conditional autoregressive (CAR) models. CAR models were first introduced by Besag (1974), where Brook's lemma is used to relate a set of conditional distributions to a joint distribution. Consider the joint distribution, $f(\mathbf{Y})$, where $\mathbf{Y} = \{Y(\mathbf{s}_1), \dots, Y(\mathbf{s}_m)\}^T$. It is clear that given this joint distribution the conditional distributions, $f(Y(\mathbf{s}_i)|Y(\mathbf{s}_j), j \neq i), i = 1, \dots, m$ are uniquely determined. Brook's lemma demonstrates the converse, allowing the joint distribution to be determined from the set of conditionals (Brook, 1964). Assume that we specify a set of conditional distributions, $f(Y(\mathbf{s}_i)|Y(\mathbf{s}_j), j \in \delta_i), i = 1, \dots, m$, where δ_i represents the set of neighbors of location i . We hope to construct a joint distribution from these local conditional distributions. This notion of using local conditional distributions to specify a joint distribution is known as a Markov random field (MRF). The Hammersley-Clifford Theorem demonstrates that if we have a MRF then the

joint distribution is given by a Gibbs distribution (Besag, 1974; Clifford, 1990) and the converse is supplied by Geman and Geman (1984). This theorem is the foundation for the CAR modeling framework.

The conditional distributions used in the CAR specification are defined as follows,

$$Y(\mathbf{s}_i)|Y(\mathbf{s}_j), j \in \delta_i \sim \mathbf{N}\left(\frac{\sum_{j \in \delta_i} Y(\mathbf{s}_j)}{n_i}, \frac{\tau^2}{n_i}\right), \quad i = 1, \dots, m,$$

where n_i is the number of neighbors for location i . This conditional specification is elegant, as the mean is a weighted average of its neighbors and the variance shrinks as the number of neighbors increases. This model can be rewritten in terms of adjacency weights,

$$Y(\mathbf{s}_i)|Y(\mathbf{s}_j), j \neq i \sim \mathbf{N}\left(\frac{\sum_{j=1}^m w_{ij} Y(\mathbf{s}_j)}{\sum_{j=1}^m w_{ij}}, \frac{\tau^2}{\sum_{j=1}^m w_{ij}}\right), \quad i = 1, \dots, m,$$

where $w_{ij} = 1(i \sim j)$, $1(\cdot)$ is an indicator function and $i \sim j$ is the event that locations i and j are neighbors. Using the Hammersley-Clifford Theorem a joint distribution can be written down in the form of a Gibbs distribution, $\mathbf{Y}|\tau^2 \propto \exp\{-\frac{1}{2\tau^2} \mathbf{Y}^T \mathbf{W}^* \mathbf{Y}\}$, where $[\mathbf{W}^*]_{ij} = -w_{ij}$ for $i \neq j$ and $[\mathbf{W}^*]_{ii} = \sum_{j=1}^m w_{ij}$ for all i . The random variable \mathbf{Y} is said to be an intrinsic CAR process with variance τ^2 (i.e., $\mathbf{Y} \sim CAR(\tau^2)$). This distribution is not proper, thus it is commonly specified as a prior for random effects to induce spatial structure into a model.

For normally distributed data, a standard areal data model with a spatially varying intercept can be specified as follows,

$$Y(\mathbf{s}) = \mu(\mathbf{s}) + \phi(\mathbf{s}) + \epsilon(\mathbf{s}), \quad \epsilon(\mathbf{s}) \stackrel{\text{iid}}{\sim} \mathbf{N}(0, \sigma^2),$$

where $\phi(\mathbf{s})$ is a spatial parameter at location \mathbf{s} , specified from a CAR model. Define the spatial effect, $\boldsymbol{\phi} = \{\phi(\mathbf{s}_1), \dots, \phi(\mathbf{s}_m)\}^T$. The mean structure is general, but a standard specification introduces location specific covariates, $\mu(\mathbf{s}) = x(\mathbf{s})\boldsymbol{\beta}$. Then, the posterior distribution of the full set of parameters, $\boldsymbol{\theta} = (\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2, \tau^2)$ is as follows,

$$f(\boldsymbol{\theta}|\mathbf{Y}) \propto f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2) \times f(\boldsymbol{\phi}|\tau^2) \times f(\boldsymbol{\beta}, \sigma^2, \tau^2),$$

where, $\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\phi}, \sigma^2 \sim \mathbf{N}(X\boldsymbol{\beta} + \boldsymbol{\phi}, \sigma^2\mathbf{I}_n)$ and $\boldsymbol{\phi} \sim \text{CAR}(\tau^2)$. Then, we set a prior such that $f(\boldsymbol{\beta}, \sigma^2, \tau^2) = f(\boldsymbol{\beta})f(\sigma^2)f(\tau^2)$. A weakly informative prior may be placed on $\boldsymbol{\beta}$, however careful consideration must be placed on the priors for σ^2 and τ^2 . It is generally desired to specify similar priors for σ^2 and τ^2 , since they represent the random and spatial error variances, however since one is specified conditionally and the other jointly, it can be difficult (Bernardinelli et al., 1995). The posterior distribution can be estimated using Markov chain Monte Carlo (MCMC) methods and the methods apply to the generalized linear mixed model (GLMM) setting.

1.3.3 Multivariate Areal Data Models

Multivariate areal data models are commonly developed for disease mapping in the presence of multiple diseases that may not be independent. Most multivariate CAR (MCAR) models are members of the family developed by Mardia (1988). Similar to the univariate case, the joint distribution is derived from the full conditional distributions. Under the MRF assumption these conditional distributions can be specified as follows,

$$f(\boldsymbol{\phi}(\mathbf{s}_i)|\boldsymbol{\phi}(-\mathbf{s}_i), \boldsymbol{\Gamma}_i) = \text{MVN}\left(\sum_{i \sim j} \mathbf{B}_{ij}\boldsymbol{\phi}(\mathbf{s}_j), \boldsymbol{\Gamma}_i\right), \quad i, j = 1, \dots, m$$

where $\boldsymbol{\phi}(\mathbf{s}) = \{\phi_1(\mathbf{s}), \phi_2(\mathbf{s}), \dots, \phi_p(\mathbf{s})\}^T$ is a p -dimensional vector, and $\boldsymbol{\Gamma}_i$ and \mathbf{B}_{ij} are $p \times p$ matrices. Mardia (1988) proved that the full conditional distributions uniquely determine the joint distribution, $f(\boldsymbol{\phi}|\{\boldsymbol{\Gamma}_i\}) \propto \exp\left\{-\frac{1}{2}\boldsymbol{\phi}^T\boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}})\boldsymbol{\phi}\right\}$, where $\boldsymbol{\phi} = \{\boldsymbol{\phi}(\mathbf{s}_1)^T, \boldsymbol{\phi}(\mathbf{s}_2)^T, \dots, \boldsymbol{\phi}(\mathbf{s}_m)^T\}^T$, $\boldsymbol{\Gamma}$ block diagonal with blocks $\boldsymbol{\Gamma}_i$ and $\tilde{\mathbf{B}}$ is $mp \times mp$ with (i, j) -th block \mathbf{B}_{ij} . From this formulation, we can choose different $\boldsymbol{\Gamma}$ and $\tilde{\mathbf{B}}$ to obtain different MCAR model structures. As in the univariate case, symmetry of $\boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}})$ is required. A convenient special case sets $\mathbf{B}_{ij} = b_{ij}\mathbf{I}_p$. The standard MCAR can be obtained by setting $b_{ij} = w_{ij}/\sum_{i \sim j} w_{ij}$ and $\boldsymbol{\Gamma}_i = \boldsymbol{\Sigma}/\sum_{i \sim j} w_{ij}$. Then the precision can be simplified using the Kronecker product, $\boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}}) = \mathbf{W}^* \otimes \boldsymbol{\Sigma}^{-1}$. This induces the following conditional distributions,

$$f(\boldsymbol{\phi}(\mathbf{s}_i)|\boldsymbol{\phi}(-\mathbf{s}_i), \boldsymbol{\Sigma}) = \text{MVN}\left(\frac{\sum_{j=1}^m w_{ij}\boldsymbol{\phi}(\mathbf{s}_j)}{\sum_{j=1}^m w_{ij}}, \frac{\boldsymbol{\Sigma}}{\sum_{j=1}^m w_{ij}}\right), \quad i, j = 1, \dots, m.$$

The moments of the standard MCAR have desirable interpretations. Similar to the univariate case the mean is a weighted average of its neighbors, although $\phi(\mathbf{s})$ is now a vector, and the covariance is a $p \times p$ matrix, Σ , that shrinks as the number of neighbors increases.

There have been numerous methods that have generalized the standard MCAR introduced by Mardia (1988). Carlin and Banerjee (2003) and Gelfand and Vounatsou (2003) developed MCAR models that permit unique spatial decay for each spatial effect using the Cholesky and spectral decompositions, respectively. Jin et al. (2005) introduced an approach for multivariate areal data to directly specify the joint distribution for a multivariate spatial process through the specification of simpler conditional and marginal forms. Additional methods to generalize the MCAR are through the linear model of coregionalization (Jin et al., 2007), and novel definitions of the neighborhood structure (Sain et al., 2011).

CHAPTER 2: DIAGNOSING GLAUCOMA PROGRESSION WITH VISUAL FIELD DATA USING A SPATIOTEMPORAL BOUNDARY DETECTION METHOD

2.1 Overview

Diagnosing glaucoma progression is critical for limiting irreversible vision loss. A common method for assessing glaucoma progression uses a longitudinal series of VFs acquired at regular intervals. VF data are characterized by a complex spatiotemporal structure due to the data generating process and ocular anatomy. Thus, advanced statistical methods are needed to make clinical determinations regarding progression status. We introduce a spatiotemporal boundary detection model that allows the underlying anatomy of the optic disc to dictate the spatial structure of the VF data across time. We show that our new method provides novel insight into vision loss that improves diagnosis of glaucoma progression using data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry. Simulations are presented, showing the proposed methodology is preferred over existing spatial methods for VF data. Supplementary materials for this article can be found in Appendix A and the method is implemented in the R package `womb1R`.

2.2 Introduction

Glaucoma is a leading cause of blindness worldwide, with a prevalence of 4% in the population aged 40-80 (Tham et al., 2014). The most common form of glaucoma is primary OAG. The biological basis of the disease is not fully defined, however the most significant risk factor for primary OAG is elevated IOP. Elevated IOP can be treated with eye drops, surgery, or laser. If this condition goes untreated the optic nerve may be damaged, resulting in permanent vision loss. Since visual impairment caused by glaucoma is irreversible and efficient treatments exist, early detection of the disease is essential. As such, patients diagnosed with glaucoma are monitored for disease progression even if they are receiving treatment, because the role of the treatment is to slow the progression. Determining if the disease is progressing remains one of the most challenging aspects of glaucoma

management, since it is difficult to distinguish true progression from variability due to natural degradation or noise (Vianna and Chauhan, 2015). Numerous techniques have been developed to monitor progression, but there is currently no consensus as to which method is best. In this study we focus on VF testing.

A VF test is a psychophysical procedure that assesses a patient's field of vision. The test results in a 2-dimensional map that represents the level of eyesight uniformly across the retina. Glaucoma patients normally receive bi-annual VF tests and have follow-up lasting numerous years (Chauhan et al., 2008). The collection of VF data results in a longitudinal series of spatially referenced measurements that exhibit a complex spatiotemporal structure. The spatial surface of the VF is observed on a lattice (i.e., uniform areal data), however it is a complex projection of the underlying optic disc and exhibits anatomically induced spatial dependencies. In particular, localized damage to the optic disc can result in clinically deterministic deterioration across the VF (Quigley et al., 1992). Incorporating this non-standard spatial dependence structure into our methodology is a priority for properly analyzing these data.

There are comparable methodological complexities that arise in Alzheimer's disease, attention deficit hyperactivity disorder, and multiple sclerosis, all related to white-matter connectivity of the brain (He et al., 2008, 2009; Konrad and Eickhoff, 2010). However, complex brain imaging data are generally analyzed using point-referenced statistical models as opposed to areal data models (Bowman et al., 2008). The point-referenced framework has a rich theory that accounts for non-standard spatial dependencies, mainly through the assumptions of non-stationarity (Sampson, 2010) and anisotropy (Ecker and Gelfand, 2003). Castruccio et al. (2016) model anatomical regions of interest of the brain in fMRI data using these assumptions in a study of stroke rehabilitation, while another study accounts for brain connectivity in Alzheimer's patients (Thompson et al., 2004).

The literature surrounding complex spatial dependencies is less developed in the areal data setting with spatiotemporal methods even less common. One reason for this is that stationarity and isotropy (i.e., correlation as a function of distance alone) cannot be defined for areal data due to the contrasting definition of spatial proximity between the two frameworks. When analyzing areal data, spatial similarity is typically dictated by a local neighborhood structure (Banerjee et al., 2003). Over time, modifications to the neighborhood structure have been proposed, and in this manuscript we work within the boundary detection framework. An extension of directional gradients from

point-referenced theory (Banerjee and Gelfand, 2006), boundary detection was originally developed to identify boundaries on geographical maps in the context of disease mapping (Ma and Carlin, 2007). The inherited method of modifying the local neighborhood structure provides an intricate framework for introducing complex spatial structure on the VF. We introduce a novel spatiotemporal boundary detection model that allows the underlying anatomy of the optic disc to dictate the spatial structure of the VF across time and show that it offers new and valuable information for improved glaucoma progression detection through analysis of data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry.

This chapter is outlined as follows. Section 2.3 details the data generating mechanism for VF data and the setting of glaucoma progression diagnostics. We briefly review spatial boundary detection methods in Section 2.4. In Section 2.5, our newly developed statistical methodology is described. We apply our method to a dataset of VF tests from glaucoma patients in Section 2.6 and compare its performance to an existing boundary detection method via simulation study in Section 2.7. We conclude in Section 2.8 with a discussion.

2.3 Visual Field Data

The VF is the spatial array of visual sensations that the brain perceives as vision (Smythies, 1996). The most common technique for testing the VF is SAP (Chauhan et al., 2008). In this study we analyze data acquired with the HFA-II (Carl Zeiss Meditec Inc., Dublin, CA). The VF data generating process is displayed in Figure 2.4. We follow a single observation throughout the figure, presented as a diamond. In Figure 2.4a, a patient (the dissertation author) is tested on a HFA-II. SAP constructs a VF map by assessing a patient's response to varying levels of light. Patients are instructed to focus on a central fixation point as light is introduced randomly in a preceding manner over a grid on the VF. As light is observed, the patient presses a button and the current light intensity is recorded (Johnson et al., 1993). The process is repeated until the entire VF is tested. In the figure, the author stares at the background of the machine, waiting until he observes the stimulus to press the buzzer.

The HFA-II measures 54 test points, however two of these correspond to a natural blind spot corresponding to the optic disc, resulting in 52 informative points. The points that correspond to

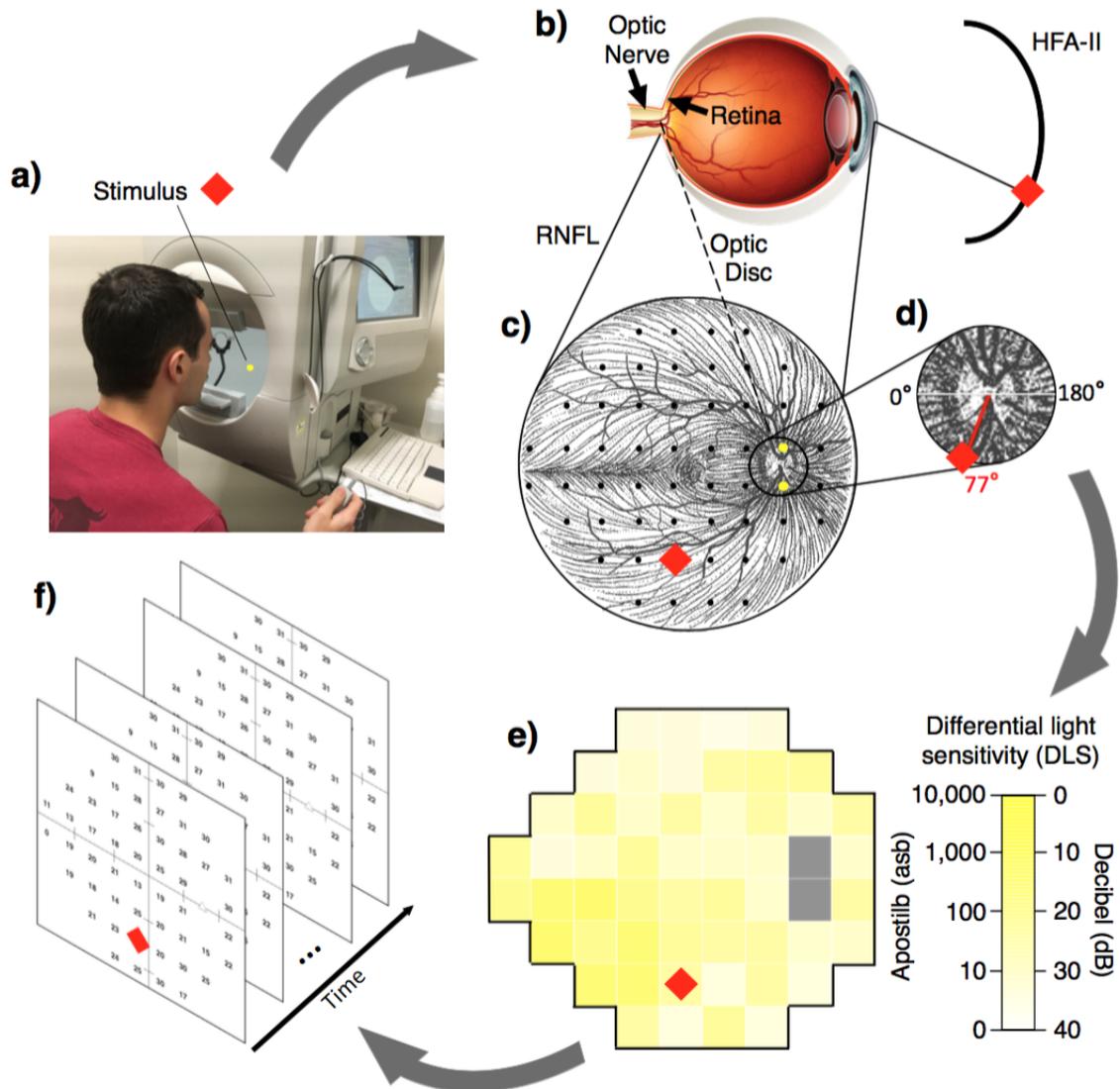


Figure 2.4: Demonstrating the VF data generating mechanism, **a)** A patient (the dissertation author) participates in a VF examination. **b)** The HFA-II stimulus is absorbed by the retina and transmitted through the optic disc, and along the optic nerve to the brain. **c)** Each VF test point corresponds to a location on the underlying RNFL. **d)** The corresponding nerve fiber enters the optic disc at 77° . **e)** Each VF test produces a map that shows the intensity a stimulus is detected at each test location. The gray locations represent the blind spot. **f)** Over time the patient accrues VF tests. The eye in **b)** is printed with permission from the Lions Eye Institute and the RNFL in **c)** is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).

the blind spot are highlighted yellow (Figure 2.4c). Figure 2.4b presents the anatomy of the eye. The retina is a light-sensitive layer at the back of the eye that absorbs stimulus from the HFA-II and transmits information to the brain through the optic disc along the optic nerve. The optic nerve is a bundle of more than one million fibers that carry visual information in the form of electrical signal from the retina to the brain. The RNFL, Figure 2.4c, is made of RGC axons. All the axons converge at the optic disc, which is the departure point of the optic nerve.

These RGC axons are responsible for encoding visual information. The ganglion cells disperse across the retina, but are mostly concentrated in the center of the retina, and transmit electrical visual signal, captured by photoreceptors, to the brain along their axons (Davson, 2012). Both the RGCs and their axons may die progressively as a result of elevated IOP. In particular, damage to specific regions of the optic disc corresponds to loss of RGCs whose axons enter the damaged region. Thus, vision loss across the VF and the corresponding damage to the optic disc are the result of the death of RGCs and their axons. Furthermore, correlation between two test points on the VF is dependent on the spatial proximity that their underlying nerve fibers enter the optic disc. This indicates that variability in the neighborhood structure of the VF is possibly indicative of progression. The nerve fibers (not to be confused with the thicker and more sparse blood vessels) are shown extending throughout the retina with the VF test points represented by black dots.

Each test location has underlying fibers that track across the RNFL and enter the optic disc at a particular angle. Garway-Heath et al. (2000) studied the relationship between the VF test points and the underlying RNFL. They quantified the relationship between the VF and optic disc by estimating the angle that each test location's underlying fiber enters the optic disc, measured in degrees ($^{\circ}$). The measure ranges from 0-360 $^{\circ}$, where 0 $^{\circ}$ is designated at the 9-o'clock position (right eye) and angles are counted counter clockwise. The Garway-Heath angle for the example location is 77 $^{\circ}$ (Figure 2.4d).

In the course of a VF test all 52 informative locations are assessed, resulting in a grid of test values (Figure 2.4e). The two locations corresponding to the blind spot are gray. SAP measures the DLS across the VF. The measurement represents a contrast between the background of the machine, normally white, and the light stimulus. The intensity of the stimulus is initially similar to the background, but as the intensity increases the contrast grows and the probability of detecting the stimulus increases. The intensity of the stimulus is measured in Apostilbs (asb), where larger

values represent a greater intensity. The HFA-II is capable of outputting intensities ranging from one (similar to the background) to 10,000 asb. All stimuli that are not detected by 10,000 asb are censored, due to the constraints of the machine. In practice, these intensities are converted to dB, where $\text{dB} = 40 - 10 \log_{10}(\text{asb})$. This inverts the scale, such that DLS values near 40 indicate good vision, while values near zero represent possible blindness. In the course of monitoring a patient with glaucoma, VF testing is performed on a regular basis and a longitudinal series of VFs is obtained (Figure 2.4f).

2.4 Spatial Boundary Detection

VF data exhibit a complex spatiotemporal structure that is characterized by localized spatial dependencies dictated by the underlying anatomy of the optic disc. These data are generated over a lattice, a subset of areal data in spatial literature. In spatial statistics, the foundational assumption states that dependence between observations weakens as the distance between locations increases (Gelfand et al., 2010; Cressie, 2015). In areal data models, this assumption often manifests through GMRFs that induce neighborhood dependence across the region (Geman and Geman, 1984). A common GMRF is the CAR process (Besag, 1974).

The CAR model achieves spatial smoothing through random effects, φ_i at location i for $i = 1, \dots, m$, with spatial structure defined through the set of neighborhood adjacencies, $\{w_{ij}\}$. These adjacencies are fixed, such that $w_{ij} = 1(i \sim j)$, where $1(\cdot)$ is the indicator function and $i \sim j$ is the event that locations i and j share a border ($w_{ii} = 0$ for all i). This specification induces an elegant conditional distribution for each random effect. We present a general form of the CAR process originally introduced by Leroux et al. (2000),

$$\varphi_i | \varphi_{-i}, \mu, \tau^2 \sim \text{N} \left(\frac{\rho \sum_{j=1}^m w_{ij} \varphi_j + (1 - \rho) \mu}{\rho \sum_{j=1}^m w_{ij} + 1 - \rho}, \frac{\tau^2}{\rho \sum_{j=1}^m w_{ij} + 1 - \rho} \right), \quad (2.1)$$

where $\varphi_{-i} = (\varphi_1, \dots, \varphi_{i-1}, \varphi_{i+1}, \dots, \varphi_m)^T$. Note that setting $\rho = 1$ reduces Equation 2.1 to the standard intrinsic CAR process. The mean is a weighted average of the neighbors with variance decreasing inversely with the number of neighbors. The standard CAR model provides an attractive representation and is flexible in its ability to model smooth spatial processes. However, it can be limited in settings where spatial structure is fragmented into localized regions due to $\{w_{ij}\}$ being

fixed. In the areal data setting, a flexible class of models called boundary detection can be used to remedy this issue (Banerjee et al., 2003).

Boundary detection was originally explored by Womble (1951), but has gained a niche in the context of disease mapping. The motivation for boundary detection is to identify regions on the spatial surface where there are sharp changes in the response value (Jacquez and Greiling, 2003). In disease mapping, these boundaries can take many forms, for example geographic obstacles such as mountain ranges or socioeconomic boundaries caused by pockets of increased poverty. Standard methods attempt to control for disjoint spatial regions by including covariates in the mean structure of the CAR process. This technique can be effective in producing highly variable spatial surfaces, but is limited in producing truly localized spatial smoothing (Banerjee et al., 2003). Boundary detection improves on this naive approach by carefully considering the form of the adjacencies.

Initially boundary detection methods were parameterized for use in disease mapping, defining boundaries as a function of the difference in standardized incidence ratios (Lu and Carlin, 2005). This method is limited, since it can be difficult to have knowledge of boundaries a priori and does not generalize outside of disease mapping. Numerous methods in the boundary detection literature treat the adjacencies as random variables and construct hierarchical models to estimate the adjacency matrix (Lu et al., 2007; Ma and Carlin, 2007; Ma et al., 2010), and even provide extensions to the spatiotemporal setting (Rushworth et al., 2017). However, inference from these models can be highly sensitive to prior specifications on certain parameters (Li et al., 2015). Furthermore, these methods introduce $\binom{m}{2}$ additional random variables leading to potential identifiability issues. Li et al. (2011) propose another class of methods that enumerate all possible permutations of the adjacencies in parallel models, using the Bayesian information criterion to choose between them. This class of methods was formalized when Lee et al. (2014) introduced a novel joint prior distribution for the spatial random effects and adjacency matrix. This prior has been extended to spatiotemporal models (Lee and Mitchell, 2014).

A final class of methods models the adjacencies using dissimilarity metrics. The method introduced in Lee and Mitchell (2011) generalizes the form of Equation 2.1 to allow for the adjacency weights to be modeled as a function of a small number of regression parameters, $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_q)^T$. According to Lee and Mitchell (2011), “boundaries in the risk surface are likely to occur between populations that are very different because homogeneous populations should have similar risk

profiles”. They define q non-negative dissimilarity metrics $\mathbf{z}_{ij} = (z_{ij1}, \dots, z_{ijq})^T$, where $z_{ijk} = |z_{ik} - z_{jk}|$ for $k = 1, \dots, q$. The q covariates, z_{ik} at location i , drive detection of boundaries and are characterized by their importance in defining the neighborhood structure. The choice of q is problem specific and based on the availability of useful explanatory information for describing the boundaries. The adjacencies are modeled as follows,

$$w_{ij}(\boldsymbol{\alpha}) = 1(i \sim j)1(\exp\{-\mathbf{z}_{ij}^T \boldsymbol{\alpha}\} \geq 0.5) \quad (2.2)$$

where each α_k is constrained to be non-negative so that a larger dissimilarity metric indicates a higher likelihood of a boundary (or zero weight). In this model, ρ is fixed at 0.99 to force the spatial correlation structure to be determined locally by $\{w_{ij}(\boldsymbol{\alpha})\}$, rather than globally by μ . This form has many appealing properties that make it amenable to boundary detection. In particular, if there are no adjacencies (i.e., $w_{ij}(\boldsymbol{\alpha}) = 0$ for all $i \neq j$) the conditional mean and variance are still defined. The method proposed by Lee and Mitchell (2011) provides a parsimonious framework for introducing localized smoothing.

2.5 Methods

Following the approach of Lee and Mitchell (2011), we propose modeling localized spatial correlation through a set of weights $\{w_{ij}(\boldsymbol{\alpha}_t)\}$ as a parsimonious function of dissimilarity metrics and their regression parameters. However, we propose extending the framework to account for spatiotemporal localized smoothing and therefore define $\boldsymbol{\alpha}_t = (\alpha_{t1}, \dots, \alpha_{tq})^T$, for $t = 1, \dots, \nu$. With appropriate temporal dependency structures in place, this specification allows for localized smoothing in instances of true temporal correlation. Inference for this model is based on MCMC simulation, and a description of the algorithm is given in Appendix A.2. Spatiotemporal models are computationally intensive, so the MCMC algorithm is implemented using Rcpp (Eddelbuettel et al., 2011) and is available from the R package `womb1R` (R Core Team, 2016).

2.5.1 Observational Model

We begin by describing our new methodology generally before applying it to VF data in Section 2.6. Let Y_{it} denote an observation from spatial location i at time t , $i = 1, \dots, m_t$, for $t = 1, \dots, \nu$.

The number of locations can vary over time. Define $\boldsymbol{\varphi}_t = (\varphi_{1t}, \dots, \varphi_{m_t t})^T$, with φ_{-it} missing the i^{th} entry. The observational model is given by,

$$\begin{aligned} Y_{it} | \vartheta_{it}, \boldsymbol{\zeta} &\stackrel{\text{ind}}{\sim} f(Y_{it} | \vartheta_{it}, \boldsymbol{\zeta}) \quad \text{for } i = 1, \dots, m_t, t = 1, \dots, \nu, \\ g(\vartheta_{it}) &= \varphi_{it}, \\ \varphi_{it} | \boldsymbol{\varphi}_{-it}, \mu_t, \tau_t^2, \boldsymbol{\alpha}_t &\sim \text{N} \left(\frac{\rho \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t) \varphi_{jt} + (1 - \rho) \mu_t}{\rho \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t) + 1 - \rho}, \frac{\tau_t^2}{\rho \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t) + 1 - \rho} \right). \end{aligned} \quad (2.3)$$

The parameter ϑ_{it} describes the distribution of Y_{it} and our novel spatiotemporal random effect, φ_{it} , is introduced as a linear predictor of $g(\vartheta_{it})$, with $g(\cdot)$ a link function. Finally, $\boldsymbol{\zeta}$ is a vector of variance (or nuisance) parameters, for example the over-dispersion parameter in the negative binomial distribution. This modeling framework is general and accommodates GLMM. The GLMM setting can be obtained by setting $\vartheta_{it} = \mathbb{E}[Y_{it} | \vartheta_{it}]$. Due to the general specification, our methodology can be used to induce spatiotemporal localized smoothing in a general areal data setting, such as disease mapping.

The random effect for φ_{it} represents an extension of the Lee and Mitchell (2011) specification with temporally referenced parameters, μ_t , τ_t^2 , and $\boldsymbol{\alpha}_t$ (referred to as observational level parameters). The ρ parameter acts as a gauge of present spatial variation, where $\rho = 0$ corresponds to global independence and $\rho \rightarrow 1$ implies strong spatial correlation. In the same vein as Lee and Mitchell (2011), we fix $\rho = 0.99$ to guarantee that spatial correlation can be determined locally by the set of weights $\{w_{ij}(\boldsymbol{\alpha}_t)\}$. We fully explore the impact of this decision through simulation as described in Section 2.6.6 and in Appendix A.4. Overall, we find that the results are robust to this choice.

The conditional distributions of the random effect can be written jointly using Brook's Lemma, $\boldsymbol{\varphi}_t | \mu_t, \tau_t^2, \boldsymbol{\alpha}_t \stackrel{\text{ind}}{\sim} \text{MVN} \left(\mu_t \mathbf{1}_{m_t}, \tau_t^2 \mathbf{Q}(\boldsymbol{\alpha}_t)^{-1} \right)$, $t = 1, \dots, \nu$, where the spatial matrix $\mathbf{Q}(\boldsymbol{\alpha}_t) = [\rho \mathbf{W}^*(\boldsymbol{\alpha}_t) + (1 - \rho) \mathbf{I}_{m_t}]$, and $\mathbf{1}_m$ and \mathbf{I}_m are an m dimensional column of ones and identity matrix, respectively. The matrix $\mathbf{W}^*(\boldsymbol{\alpha}_t)$ has diagonal elements $w_{ii}^*(\boldsymbol{\alpha}_t) = \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t)$ and off-diagonal elements $w_{ij}^*(\boldsymbol{\alpha}_t) = -w_{ij}(\boldsymbol{\alpha}_t)$.

2.5.2 Neighborhood Model

We use a similar framework as Lee and Mitchell (2011) to model the adjacency weights by writing them as a function of dissimilarity metrics, z_{ij} . The weights are defined as follows,

$$w_{ij}(\boldsymbol{\alpha}_t) = 1(i \sim j) \exp \left\{ -z_{ij}^T \boldsymbol{\alpha}_t \right\}. \quad (2.4)$$

Unlike the original Lee and Mitchell (2011) specification (Equation 2.2), the weights are not forced to be binary. We do specify the components of $\boldsymbol{\alpha}_t$ to be non-negative, forcing the adjacencies in the open unit interval. These constraints on α_{tk} yield intuitive interpretations at extreme values. As $\alpha_{tk} \rightarrow \infty$ the adjacencies become zero, resulting in an independence model, while $\alpha_{tk} \rightarrow 0$ reduces the adjacencies to a standard CAR process (Equation 2.1). This new specification changes how neighbors share information. It is best understood through the conditional mean from Equation 2.1, which becomes a weighted average of neighbors under our new definition of an adjacency, versus a simple average.

2.5.3 Temporal Model

From the joint specification of the random effects, we see that spatial structure is introduced through the covariance of φ_t at each time point and there is conditional independence ($\varphi_{t_1} \perp \varphi_{t_2} | \mu_t, \tau_t^2, \boldsymbol{\alpha}_t : t = t_1, t_2$). To induce temporal dependence between the φ_t , we specify a separable temporal structure on the observational level parameters. Define,

$$\boldsymbol{\theta} = [\boldsymbol{\theta}_{\cdot 1} \cdots \boldsymbol{\theta}_{\cdot \nu}] = \begin{bmatrix} \boldsymbol{\theta}_{1\cdot} \\ \boldsymbol{\theta}_{2\cdot} \\ \boldsymbol{\theta}_{2+1\cdot} \\ \vdots \\ \boldsymbol{\theta}_{2+q\cdot} \end{bmatrix} = \begin{bmatrix} \mu_1 & \cdots & \mu_\nu \\ \log(\tau_1) & \cdots & \log(\tau_\nu) \\ \log(\alpha_{11}) & \cdots & \log(\alpha_{\nu 1}) \\ \vdots & & \vdots \\ \log(\alpha_{1q}) & \cdots & \log(\alpha_{\nu q}) \end{bmatrix}.$$

Using properties of the vectorization function, $\text{vec}(\cdot)$, and the Kronecker product, \otimes , a separable process is specified such that

$$\text{vec}(\boldsymbol{\theta}) | \boldsymbol{\delta}, \mathbf{T}, \phi \sim \text{MVN}(\mathbf{1}_\nu \otimes \boldsymbol{\delta}, \boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}). \quad (2.5)$$

This process yields elegant interpretations for the row and column moments of $\boldsymbol{\theta}$. From the moments, we see that $\boldsymbol{\delta}$ is a constant that corresponds to the mean of the observational level parameters at time t . The matrix \mathbf{T} can be interpreted as the local covariance of the observational level parameters at each time t .

Finally, the correlation matrix, $\boldsymbol{\Sigma}(\phi)$, represents the temporal correlation of each observational level parameter over time. Due to the properties of the separable covariance, each of the observational level parameters has the same temporal structure dictated by the form of $\boldsymbol{\Sigma}(\phi)$. The form of $\boldsymbol{\Sigma}(\phi)$ is general such that any standard temporal correlation function may be specified (e.g., exponential or first order autoregressive). The parameter ϕ acts as a temporal tuning parameter describing the strength of correlation across time and can be interpreted within the context of each specific temporal structure.

2.5.4 Specifying Hyperprior Distributions

In order to complete the model specification, we define hyperprior distributions for the introduced parameters such that,

$$\boldsymbol{\delta} \sim \text{MVN}(\boldsymbol{\mu}_\delta, \boldsymbol{\Omega}_\delta), \quad \mathbf{T} \sim \text{Inverse-Wishart}(\xi, \boldsymbol{\Psi}), \quad \phi \sim \text{Uniform}(\mathbf{a}_\phi, \mathbf{b}_\phi).$$

The choice of the entries in $\boldsymbol{\mu}_\delta$ can be informative or non-informative depending on the context and user. It is important to judiciously consider the entries of $\boldsymbol{\Omega}_\delta$. For simplification, we detail a situation where $\boldsymbol{\Omega}_\delta$ is diagonal, $\boldsymbol{\Omega}_\delta = \text{Diag}(1000, 1000, v_1, \dots, v_q)$. The entries in $\boldsymbol{\Omega}_\delta$ that are of importance are those that correspond to $\log(\alpha_{tk})$, since the large variances for μ_t and $\log(\tau_t)$ induce approximately flat priors. More care is needed in specifying v_1, \dots, v_q . These hyperprior variances are chosen for purposes of regularization, in order to encourage $\log(\alpha_{tk})$ to be in a realistic range. Regularization is a common use of Bayesian priors (Gelman and Shalizi, 2013). In particular, these

variances are chosen so that α_{tk} do not become larger than α_k^* , such that $[\alpha_k^* : \exp\{-\alpha_k^* z_k\} = 0.5]$ with $z_k = \min_{i,j} \{z_{ijk}\}$. This condition comes from Equation 2.4, where we isolate each dissimilarity metric individually.

For our prior on \mathbf{T} , we use an inverse-Wishart distribution with degrees of freedom $\xi = (q+2)+1$ and scale matrix, $\Psi = \mathbf{I}_{q+2}$. This prior is appealing since it induces marginally uniform priors on the correlations of \mathbf{T} and allows for the diagonals to be weakly informative (Gelman et al., 2014). Finally, we specify the hyperprior distribution for the temporal tuning parameter ϕ for correlation structures with one parameter. The bounds for ϕ cannot be specified arbitrarily since it is important to account for the magnitude of time elapsed. We specify the following conditions for finding the bounds, $[a_\phi : [\Sigma(a_\phi)]_{t,t'} = 0.95, |t - t'| = x_{\max}]$ and $[b_\phi : [\Sigma(b_\phi)]_{t,t'} = 0.01, |t - t'| = x_{\min}]$, where x_{\min} and x_{\max} are the minimum and maximum temporal differences between visits. These conditions state that the lower bound of ϕ is small enough so that the greatest length of time between time points can yield a correlation of 0.95 and the upper bound is set so that the shortest length of time between time points can reach 0.01. These conditions were specified so that ϕ can dictate a temporal correlation close to independence ($\phi \rightarrow b_\phi$) or strong correlation ($\phi \rightarrow a_\phi$), resulting in a weakly informative prior distribution.

2.5.5 Prediction

Once posterior samples have been obtained, prediction is often a priority. In particular, obtaining samples from the posterior predictive distribution (PPD) is of interest, $f(\mathbf{Y}_{\nu+1}|\mathbf{Y})$, where $\mathbf{Y}_t = (Y_{1t}, \dots, Y_{mt})^T$ and $\mathbf{Y} = (\mathbf{Y}_1, \dots, \mathbf{Y}_\nu)^T$. We first express the PPD as an integral $\int_{\Omega} f(\mathbf{Y}_{\nu+1}|\Omega, \mathbf{Y}) f(\Omega|\mathbf{Y}) d\Omega$ and then further partition the integral,

$$\int_{\Omega} \underbrace{f(\mathbf{Y}_{\nu+1}|g^{-1}(\varphi_{\nu+1}), \zeta)}_1 \underbrace{f(\varphi_{\nu+1}|\theta_{\nu+1})}_2 \underbrace{f(\theta_{\nu+1}|\theta, \delta, \mathbf{T}, \phi)}_3 \underbrace{f(\zeta, \theta, \delta, \mathbf{T}, \phi|\mathbf{Y})}_4 d\Omega, \quad (2.6)$$

where $\Omega = (\varphi_{\nu+1}, \theta_{\nu+1}, \zeta, \theta, \delta, \mathbf{T}, \phi)$. The convenient form of Equation 2.6 is a function of four known densities that are defined as a consequence of the methodology introduced in Section 2.5. As such, the PPD can be obtained through composition sampling (Tanner, 1996). This theory is presented with one future time point, but is easily generalized to multiple. Full prediction theory details are presented in Appendix A.3.

2.6 Analysis of Visual Field Data and Glaucoma Progression Risk

2.6.1 Study Data

In this study, we source data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry, Perth, Western Australia. The dataset contains 1,448 VFs from 194 distinct eyes (98 patients in total). Three of the eyes had no clinical assessment of progression and are discarded, yielding 191 series of VFs for analysis. All of the subjects have some form of primary OAG. The mean follow-up time for participants is 934 days (2.5 years) with an average of 7.4 tests per subject. The progression status of each eye was determined by a group of expert clinicians. Although there is no consensus gold standard for diagnosing progression, there is precedent for treating clinician expertise as a gold standard when introducing new analytic models (Betz-Stablein et al., 2013; Warren et al., 2016). Every VF series is diagnosed as progressing based on the clinical judgment of two independent clinicians. In the case that the two clinicians disagree, a third clinician is consulted (occurred for only 13 VF series). In our study, we have 141 (74%) stable and 50 (26%) progressing patient eyes. For a detailed description of the data, please refer to (Betz-Stablein et al., 2013).

The longitudinal series of VFs for one stable (left) and one progressing (right) eye are presented in the top row of Figure 2.5. In the figure, each cell represents the observed DLS at a VF location over visits. The progressing eye has more variability across the VF and in particular within different VF regions, as seen in the superior and inferior sectors.

2.6.2 Accounting for Zero-Truncation

We apply our newly developed methodology to a longitudinal series of VFs. From Section 2.3, we know that VF testing machinery does not allow observations below 0 dB, and therefore any zero measurement represents a potentially censored observation. This motivates the use of a Tobit model (Tobin, 1958), in which there is precedent in glaucoma progression research (Betz-Stablein et al., 2013; Bryan et al., 2015).

Define the observed DLS, Y_{it} , at VF location i with $i = 1, \dots, 52$ and visit t with $t = 1, \dots, \nu$. There are 52 VF locations excluding the blind spot and ν is the number of visits a patient accrues

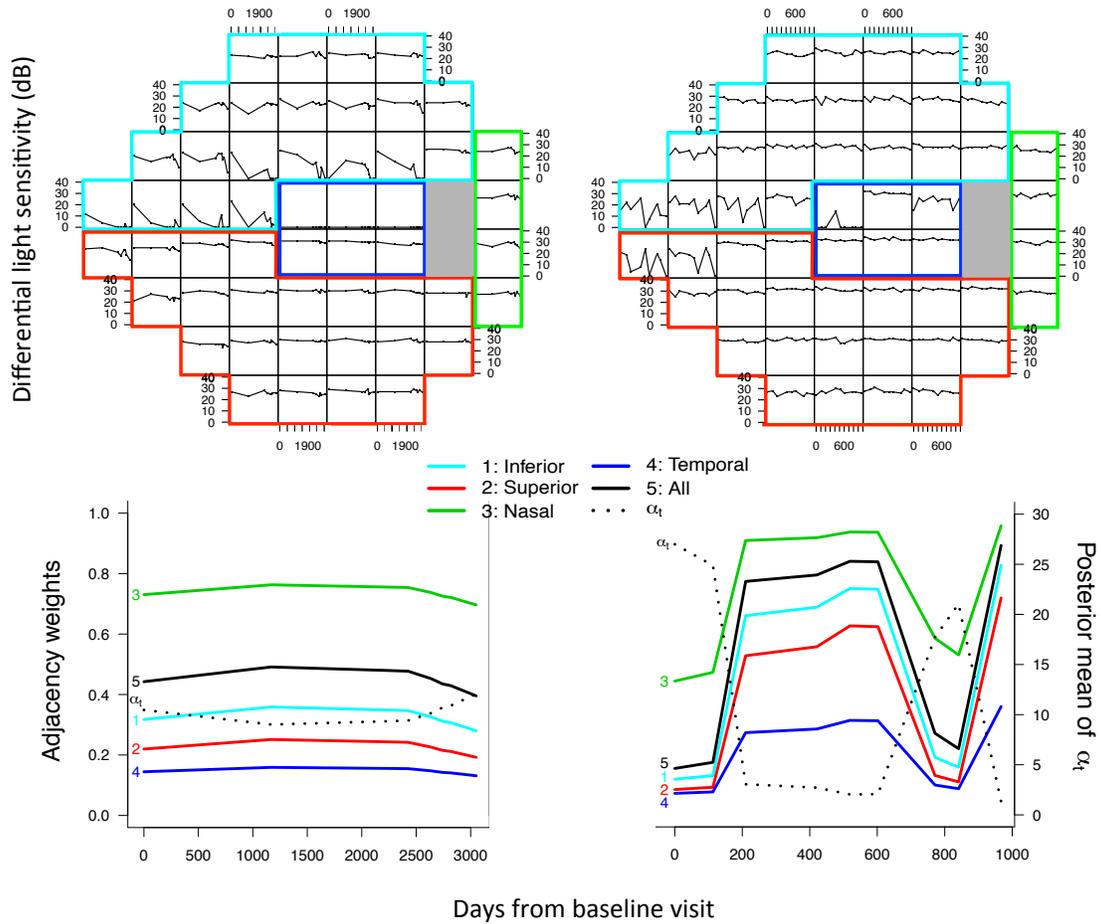


Figure 2.5: Top row: Example of stable (left) and progressing (right) VFs. The observed DLS across all visits for a VF location is shown in each cell. Bottom row: Posterior mean estimates of α_{tGH} plotted across time along with the average of the spatial adjacencies across different sectors of the VF for the same two patient eyes. Sector 1 (Inferior) is the superior VF, corresponding to the inferior sector of the optic disc, and Sector 2 (Superior) is the inferior VF, corresponding to the superior sector of the optic disc.

and is patient specific. To induce the Tobit model, define $g(\vartheta_{it}) = \varphi_{it}$ with identity link and specify,

$$f(Y_{it}; \vartheta_{it}) = P(Y_{it} = x | \vartheta_{it}) = 1(x = 0)1(\vartheta_{it} \leq 0) + 1(x = \vartheta_{it})1(\vartheta_{it} > 0), \quad x \geq 0.$$

This specification induces the standard Tobit model, $Y_{it} = \max\{0, \vartheta_{it}\}$, where ϑ_{it} is an underlying normally distributed latent process.

2.6.3 Creating a Dissimilarity Metric

We specify a dissimilarity metric based on the Garway-Heath angles defined in Section 2.3, since we know that correspondence between VF test locations and their underlying nerve fibers is important in determining local neighborhood structure. Figure 2.6 displays the dissimilarity metric. It shows that two locations on the VF may be neighbors, but can still be dissimilar in terms of the Garway-Heath angles. Interestingly, a pattern emerges across the VF that mirrors the RNFL (left in Figure 2.6). In particular, the locations that are separated by the superior and inferior regions are separated by nearly 180° and the flow of spatial dependency emulates the nerve fibers. This pattern is manifested in Figure 2.5, in the stable eye, as VF locations deteriorate regionally, further motivating the use of the Garway-Heath dissimilarity metric.

Formally, we define z_i as the Garway-Heath angle for location i . Then, the dissimilarity metric between locations i and j is $z_{ij} = ||z_i - z_j||$. We use the following distance metric, $||x - y|| = \min\{|x - y|, 360 - \max\{x, y\} + \min\{x, y\}\}$. This metric calculates the minimum difference in Garway-Heath angles on the arc of the circular optic disc. We define our dissimilarity parameter at time t as α_{tGH} , since we only use a single dissimilarity metric. We allow the event $1(i \sim j)$ to include both edges and corners (i.e., a queen specification).

2.6.4 Model Estimation

To finalize the model, we specify the temporal correlation structure and hyperparameters. We define an exponential correlation structure, such that $[\Sigma(\phi)]_{t,t'} = \exp\{-\phi|x_t - x_{t'}|\}$ where x_t is the number of days at visit t after the initial visit, with $x_1 = 0$ for each patient. Based on the criterion in Section 2.5.4 for the Garway-Heath dissimilarity metric, we specify $v_1 = 1$. Then we set

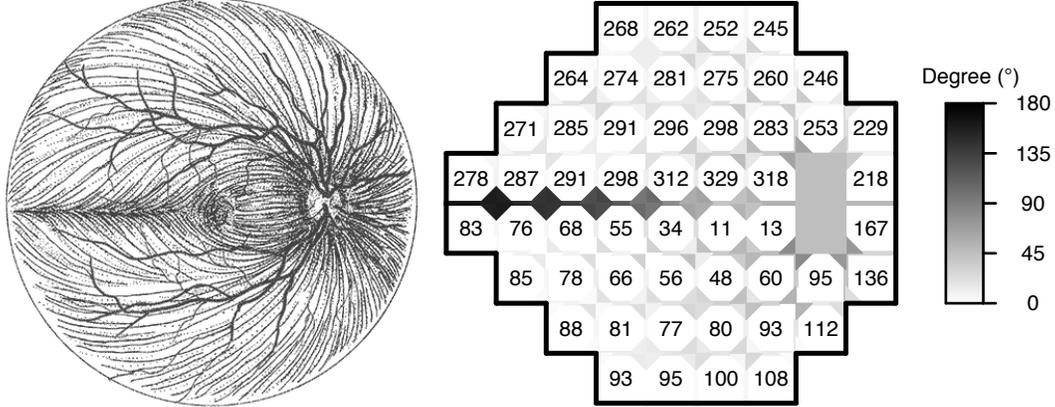


Figure 2.6: Demonstrating the Garway-Heath dissimilarity metric across the VF (right). The angles are defined as the degree at which the underlying nerve fiber enters the optic disc for each location. The edges and corners are shaded to represent the distance between bordering locations on the VF, where darker shading represents a larger distance. The RNFL is displayed to demonstrate the similarity of the pattern that appears on the VF by using the Garway-Heath angles (left). The RNFL (left) is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).

$\Omega_\delta = \text{Diag}(1000, 1000, 1)$ and $\mu_\delta = (3, 0, 0)$. For complete details on the implementation of the model, see Appendix A.2.

2.6.5 Diagnosing Glaucoma Progression

The methodology presented in Section 2.5 provides a novel framework for studying glaucoma progression. Since progression is characterized by worsening disease severity over time, we propose using a function of α_{tGH} that can quantify this variation. We suggest using the posterior mean of the coefficient of variation (CV) of α_{tGH} for $t = 1, \dots, \nu$. The CV is the ratio of the standard deviation and mean and is an ideal summary metric, because it accounts for the variability in a parameter while standardizing by its mean, allowing it to be comparable over populations (Brown, 1998). This metric is novel clinically as α_{tGH} does not describe mean trend in DLS over time, but rather is representative of optic disc damage and changes in the spatial covariance structure across time. We refer to our metric based on the spatiotemporal (ST) model as “ST CV”.

In the bottom row of Figure 2.5, the posterior mean estimates of α_{tGH} for the two patient eyes are presented over time along with the average of the spatial adjacency weights, $w_{ij}(\alpha_{tGH})$, across different sectors of the VF. If the α_{tGH} parameters were static across time we would expect to see

no variability in these values for each of the sectors, since α_{tGH} completely defines the adjacencies at each time point. In Figure A.1 of Appendix A, we display posterior mean estimates of α_{tGH} over time for all patients, subset by progression status. The figure suggests that there is more variability in the estimates across time for the progressing patients generally, and that there isn't a clear deterministic trend structure in the estimates across time. This further motivates a metric like CV to quantify general variability across time for the purposes of diagnosing progression status.

In order to assess the novelty of our method, we compare it to two metrics that aim to be representative of a class of standard VF trend-based methods. Trend-based VF diagnostic techniques can be grouped in two categories, global and point-wise, and study associations between VF changes and progression (Vianna and Chauhan, 2015). We define the global metric “Mean CV” as the CV of the VF wide mean of DLS at each visit. If the patient is stable, the means should be similar at each visit and the CV of the metric should be near zero. On the other hand, a progressing patient will have varying means at each visit and the CV should increase as visits accrue. We also compare our model to a commonly used PLR method. The progression metric is defined as the minimum p-value for the slope parameter from separately run simple linear regression analyses across all VF locations, where observed DLS is the dependent variable and time from first visit is the independent variable (Smith et al., 1996). Using both “Mean CV” and PLR as comparators is far-reaching, as global methods are generally more robust, while point-wise methods identify local changes in visual ability. In Figure 2.5, progression status for each patient eye was incorrectly diagnosed by “Mean CV” and PLR but correctly diagnosed by “ST CV”, indicating the potential of “ST CV” to detect variability that the other methods overlook. Although these trend metrics misclassified the eyes in Figure 2.5, they are easily implemented and their interpretations are simple and clinically useful. Thus, a novel metric must have clinical utility and provide supplemental diagnostic capability.

To assess the diagnostic capability of our method, we construct logistic regression models regressing various metrics on the clinical assessment of progression. We compare our method to “Mean CV”, PLR, and also the posterior mean CV of α_{tGH} from the Lee and Mitchell (2011) model, where α_{tGH} is estimated independently at each visit. We refer to this metric as “Space CV”. To obtain “Space CV”, we apply the Lee and Mitchell (2011) methodology with a Tobit likelihood. Thus, the appreciable differences between the two models that produce “Space CV” and “ST CV” are the

Table 2.1: Assessing the diagnostic capability of VF metrics. Each metric is regressed against the clinical assessment of progression using a logistic regression model with the slope estimates being displayed (*significance level of 0.05). “Mean CV” is the CV of the VF wide mean of DLS at each visit, PLR is the minimum p-value of the location specific regression slopes, and “Space CV” and “ST CV” represent the mean posterior CV of α_{tGH} from the Lee and Mitchell (2011) and our spatiotemporal models, respectively. Each of the metrics are standardized.

Metrics	Estimate	Std. Error	z value	Pr(> z)
Mean CV	0.40	0.16	2.55	0.011*
PLR	-0.59	0.25	-2.37	0.018*
Space CV	-0.07	0.17	-0.41	0.680
ST CV	0.39	0.16	2.44	0.015*

definition of the weights, cross-covariance, and temporal correlation structure for the observational level parameters.

In Table 2.1 we present the results from the logistic regression analyses using a significance level of 0.05. Each predictor is standardized in order to facilitate comparisons of the different metrics. We can see that as expected “Mean CV” and PLR are significantly associated with glaucoma progression, with p-values of 0.011 and 0.018, respectively. For “Mean CV”, the estimated slope coefficient of 0.40 indicates that as a patient’s “Mean CV” increases, their risk of progression increases. For PLR, a smaller minimum p-value suggests an increased risk of progression (estimated slope coefficient of -0.59). Based on the glaucoma literature, we would expect trend-based methods such as “Mean CV” and PLR to have good discriminatory capability between stable and progressing eyes. It was less clear for our new metrics dependent on α_{tGH} . We see that “Space CV” is not significantly associated with glaucoma progression with a p-value of 0.680, while “ST CV” is significantly associated with a p-value of 0.015. This result is encouraging, yet surprising, since the models are similar and illuminates their differences in the VF data setting. However, for this newly defined metric to be impactful we must verify that it is explaining a novel pathway in glaucoma progression, independent of existing methods.

In order to assess whether “ST CV” provides novel diagnostic capabilities, we explore the correlation between the metrics. In Figure A.2 of Appendix A, we show pairwise correlation plots. We also present Pearson correlation estimates (ρ) and p-values from the hypothesis test: $H_0 : \rho = 0$, $H_1 : \rho \neq 0$. “ST CV” is uncorrelated with both “Mean CV” and PLR with estimated correlations of 0.06 and 0.11, respectively, and large p-values. This result has important implications;

indicating that in addition to being highly predictive of progression, “ST CV” is uncorrelated with the standard trend-based metrics. This suggests that “ST CV” and the trend-based metrics can be used in conjunction in order to diagnose progression. Finally, the estimated correlation between “Space CV” and “ST CV” is 0.48 and the p-value is highly significant (< 0.001), which is not surprising since “Space CV” and “ST CV” are estimating the same quantity. It is, however, interesting that these two metrics have such different associations with glaucoma progression (see Table 2.1). When “Space CV” was calculated with continuous weights this association did not change (p-value: 0.326, not included in Table 2.1), indicating the importance of temporal correlation and cross-covariance for properly modeling VF data. Presumably due to our enhanced methodology, “ST CV” is more precisely estimating CV of α_{tGH} by smoothing the α_{tGH} and eliminating temporal noise and cross-covariance dependencies. To formalize this hypothesis a simulation study is designed in Section 2.7.

2.6.6 Sensitivity Analyses

In Appendix A, we present a number of sensitivity analyses and an additional simulation study to explore various modeling assumptions, including hyperparameter choices, the exponential correlation structure, the bounds of ϕ , and mis-specification of ρ . Overall, we find that the results are robust to these assumptions. For full details on these analyses, please see Appendix A.

2.7 Simulation Study

A simulation study is designed to assess the performance of the proposed model in the presence of temporal dependence and cross-covariance. We focus on the estimation of CV of α_{tGH} , since we propose using this posterior distribution in diagnosing progression.

2.7.1 Data Generation

In order to understand the gains of using our methodology in the presence of temporal correlation and cross-covariance dependence, we design a simulation study comparing the spatial model of Lee and Mitchell (2011) and our spatiotemporal method. We simulate data based on a set of known hyperparameters and then estimate the posterior mean of the CV of α_{tGH} using both models,

comparing the estimates to the known truth using bias, mean squared error (MSE), and empirical coverage (EC). The simulation is designed to explore model performance in a typical patient from our study data. As such, we fix the true hyperparameters in the simulation study to the posterior means obtained from Section 2.6 for an average patient. Here an average patient is one whose posterior mean estimates are average amongst all patients. The true hyperparameters used in the simulation study are as follows,

$$\boldsymbol{\delta} = \begin{bmatrix} 2.446 \\ 0.070 \\ 0.974 \end{bmatrix}, \quad \mathbf{T} = \begin{bmatrix} 0.820 & 0.004 & -0.028 \\ 0.004 & 0.380 & -0.191 \\ -0.028 & -0.191 & 0.840 \end{bmatrix}, \quad \phi = 0.163.$$

Simulation settings are developed to incrementally understand the impact of the cross-covariance, \mathbf{T} , and temporal correlation, $\Sigma(\phi)$. To facilitate this analysis, Equation 2.5 is utilized. Define $\mathbf{T}_{\text{Diag}} = \text{Diag}(\mathbf{T})$, such that \mathbf{T}_{Diag} has zeros on the off-diagonal and $\phi_I = 100$ (note that a large ϕ for the exponential correlation implies temporal independence). The simulation settings are as follows with the data generating covariance given in parentheses, A: no temporal correlation, no cross-covariance ($\Sigma(\phi_I) \otimes \mathbf{T}_{\text{Diag}}$), B: no temporal correlation, cross-covariance ($\Sigma(\phi_I) \otimes \mathbf{T}$), C: temporal correlation, no cross-covariance ($\Sigma(\phi) \otimes \mathbf{T}_{\text{Diag}}$), D: temporal correlation, cross-covariance ($\Sigma(\phi) \otimes \mathbf{T}$). Finally, we must specify the days and number of VF visits. To obtain the visit days, we sampled from a Poisson distribution with rate parameter equal to the average difference in days between VF visits (rate = 117.25 days). We present the simulation at the median (7) and maximum (21) number of visits in our study data.

2.7.2 Results

For each simulation setting, we generate 100 values of $\boldsymbol{\theta}$ and then use each $\boldsymbol{\theta}$ to simulate 10 datasets. This yields 1,000 simulated datasets for each of the simulation settings. In Table 2.2, the bias, MSE, and EC are presented for the two models across all settings and at the median and maximum number of visits. We begin by noting the average (across all settings) simulation standard errors of the bias (0.223, 0.063), MSE (0.039, 0.028) and EC (0.141, 0.013) for the space and

Table 2.2: Results from simulation study estimating the posterior mean of the CV of α_{tGH} in setting A: no temporal correlation, no cross-covariance, B: no temporal correlation, cross-covariance, C: temporal correlation, no cross-covariance, and D: temporal correlation, cross-covariance. Each setting is also implemented for the median (7) and maximum (21) number of VF visits. Each reported estimate is based on 1,000 simulated datasets.

Setting	Model	# Visits					
		7 (Median)			21 (Maximum)		
		Bias	MSE	EC	Bias	MSE	EC
A	ST	0.032	0.107	0.97	0.023	0.084	0.95
	S	0.102	0.111	0.87	0.174	0.098	0.81
B	ST	0.047	0.125	0.96	0.034	0.101	0.95
	S	0.119	0.120	0.80	0.250	0.176	0.68
C	ST	-0.113	0.088	0.97	0.002	0.037	0.98
	S	-0.306	0.172	0.51	-0.182	0.088	0.60
D	ST	-0.103	0.085	0.98	0.015	0.060	0.98
	S	-0.299	0.172	0.51	-0.169	0.087	0.59

spatiotemporal models, respectively. In general the spatiotemporal model has smaller standard errors than the space model, indicating less variability in estimation across the simulated datasets.

The bias results suggest that our spatiotemporal model produces an estimator of CV with bias closer to zero than the spatial model on average across all settings. In the settings with no temporal correlation (i.e., A and B) the MSE is quite similar between the models, except for Setting B where the maximum number of visits is considered. In this setting, the MSE is lower for the spatiotemporal model, revealing the importance of the cross-covariance for properly estimating the posterior mean of the CV of α_{tGH} . In all other settings, the estimated MSEs for the spatiotemporal model are smaller than for the spatial model, with the largest differences seen in Settings C and D for the median number of visits.

The EC is the proportion of the time that the estimated Bayesian credible intervals contain the true CV value. We define the nominal coverage as 95%. The EC results favor the spatiotemporal model across all settings. The spatial model has an EC of 0.87 and 0.81 in setting A for the median and maximum settings, respectively. This is the most ideal setting for the spatial model, because it was the setting in which the model was introduced (although within the context of disease mapping). The EC deteriorates in the spatial model as cross-covariance and temporal structure are introduced in the simulated data, falling as low as 0.51, while the spatiotemporal model performs consistently.

2.8 Discussion

In this chapter, we proposed a modeling framework for incorporating local neighborhood structure into complex spatiotemporal areal data. Based on this framework we developed an innovative and highly predictive diagnostic of glaucoma progression that outperformed the spatial-only model. Although motivated by VF data, the methodology was introduced in a general manner that permits the model to be applied in broad areal data settings, including disease mapping and more generally GLMM. The methodology is built upon theory in boundary detection literature, utilizing a Bayesian hierarchical modeling framework for inference. We extended the spatial-only method introduced by Lee and Mitchell (2011), that elegantly introduced a dissimilarity metric in a parsimonious framework. Our method allows for the local neighborhood structure to adapt over time as a function of changing dissimilarity metric parameters. Furthermore, the temporal correlation and cross-covariance are accounted for, eliminating known sources of excess variability. This parsimonious method induces non-standard local spatial surfaces in areal data that are capable of mirroring complex processes, such as the surface of the VF.

We have shown (Figure 2.6) that the spatiotemporal covariance structure specified in our methodology successfully induces local neighborhood structure across the VF. The novelty of employing the Garway-Heath angles in the form of a dissimilarity metric provides a connection between the VF and the underlying optic disc, resulting in a neighborhood structure on the VF that is representative of the RNFL. Other methods incorporated the Garway-Heath angles into statistical models, either collapsed into regions of anatomical interest (Betz-Stablein et al., 2013; Warren et al., 2016) or by VF location (Zhu et al., 2014), but none allowed the effect to change dynamically. The dissimilarity metric parameter, α_{tGH} , dictates the VF spatial surface at each visit, allowing the neighborhood structure to adapt alongside changes in DLS.

The results from applying our method to VF data (Section 2.6) demonstrated the added benefit of using our methodology, in both the clinical and statistical frameworks. We defined the diagnostic metric “ST CV” as the posterior mean of the CV of α_{tGH} , and showed that it is a significant predictor of clinically determined glaucoma progression while being uncorrelated with standard VF trend-based metrics (“Mean CV”, PLR). Since “ST CV” is independent of “Mean CV” and PLR, and each one is an effective predictor, there is clinical utility to combining their diagnostic capability (i.e., “ST

CV” is not meant to compete with “Mean CV” and PLR but should be used together). The “Space CV” metric was not significantly associated with progression.

Our simulation study indicated that the presence of temporal correlation and cross-covariance dependence impacted estimation of the posterior mean of the CV of α_{tGH} . The results (Table 2.2) illustrated superior bias, MSE, and EC for our proposed spatiotemporal model over the spatial model of Lee and Mitchell (2011). These results imply that our method provides a framework for more precisely estimating the CV of α_{tGH} by smoothing the α_{tGH} and eliminating temporal noise and cross-covariance dependencies. These results are consistent with studies that have shown when temporal correlation is ignored, estimators are often biased and variances are poorly estimated (West, 1996; Sherman, 2011).

The significant association observed between “ST CV” and glaucoma progression was statistically note-worthy, but taken in isolation had limited clinical implications. We showed the clinical utility of “ST CV” by demonstrating its independence from “Mean CV” and PLR, and established it as a novel predictor of progression. This fundamental finding shows that “ST CV” constitutes an alternate pathway for studying glaucoma progression using VF data. This pathway is facilitated by the dissimilarity metric framework introduced by the proposed methodology and consequently “ST CV” has a novel interpretation amongst glaucoma progression diagnostic metrics. In particular, “ST CV” quantifies damage to the optic disc over time as a function of stability on the spatial structure of the VF. It provides a method for discussing underlying damage to the optic disc using VF data. These clinical implications demonstrate the need for further exploration of the proposed metric.

In our implementation of the model for glaucoma data, we are not particularly interested in the specific location of the spatial boundaries across the VF, unlike traditional boundary detection applications. In isolation we do not believe those findings would be predictive of glaucoma progression, but would only inform about the current level of vision loss for a patient. Our results suggest that changes in these boundaries across time represent an informative, innovative, and unique metric for diagnosing glaucoma progression. Using the dissimilarity metric gives us the ability to directly quantify the boundary changes across time and therefore, to make clinical progression determinations for a patient. In future applications of the model outside of the VF setting, boundary locations across time may be of interest and can be explored using our method.

The need for spatiotemporal boundary detection with a dissimilarity metric is also driven by the unique features of our VF dataset. Using a more basic CAR structure with fixed neighborhood adjacencies would result in ignoring the fact that these neighborhood definitions are potentially changing across time due to disease progression. It would also require us to fix the adjacencies *a priori*. While the Garway-Heath angles provide a template for describing proximity on the VF, ultimately the appropriate amount of information shared across VF locations depends on the severity and location of current damage to the optic disc. On the other hand, using a point-referenced geostatistical model with a spatiotemporal covariance function would result in ignoring the inherent discreteness of the spatial domain, a grid overlaying the VF. Spatiotemporal boundary detection with a dissimilarity metric represents an ideal blend between the two methods by allowing for discreteness in the spatial domain and flexibility in defining neighbors through use of a continuous covariate such as the Garway-Heath angles.

Finally, this work opens up numerous avenues for future statistical research. If data are available, covariates can be incorporated in the observational model naturally via the link function, $g(\vartheta_{it}) = x_{it}\beta + \varphi_{it}$. Currently, the specification of the Leroux et al. (2000) likelihood includes a μ_t in the mean structure, thus in order to identify the μ_t and β_0 we must either apply a constraint, $\sum_{t=0}^{\nu} \mu_t = 0$, or re-parameterize the likelihood to be a function of only τ_t and α_t . The incorporation of covariates is of particular importance in disease mapping, where standard incidence rates are often mapped over varying risk factors (Lee and Mitchell, 2013). Another natural extension to this approach includes generalizing the separable temporal covariance to allow unique temporal decay parameters.

2.9 Supplementary Material

Appendix A contains details for implementing the MCMC sampler, including derivation of full conditionals, along with prediction theory, additional sensitivity analyses, and figures.

CHAPTER 3: IMPROVED DETECTION OF VISUAL FIELD PROGRESSION USING A SPATIOTEMPORAL BOUNDARY DETECTION METHOD

3.1 Overview

3.1.1 Purpose

To evaluate the diagnostic performance of a recently developed spatiotemporal boundary detection predictor (STBound) of glaucomatous VF progression in comparison with existing methods.

3.1.2 Methods

STBound, Spatial PROGgression (SPROG), and traditional trend-based progression methods (global index (GI) regression, mean regression slope (MS), PLR) were applied to longitudinal VF data from 191 eyes of 91 glaucoma patients with an average follow-up time of 2.6 years. The ability and speed of each method to correctly determine the progression status were compared using Akaike information criterion (AIC), full and partial area under the receiver operating characteristic curve (AUC), sensitivity, and specificity. Correlation between methods was explored to establish the novelty of STBound in describing VF progression.

3.1.3 Results

All trend-based methods successfully predicted VF progression (all p-values <0.05). GI (AIC: 204.11; AUC: 0.70; pAUC: 0.26) performed better than other trend-based methods (AIC: 210.46-217.55; AUC: 0.63-0.67; pAUC: 0.13-0.24). STBound offered improved diagnostic ability compared to the trend-based methods (AIC: 197.77; AUC: 0.74; pAUC: 0.31). Unlike the significant correlation between trend-based methods (all p-values \leq 0.05), STBound showed no correlation with the competing methods (r: -0.01 to 0.11; p-values: 0.11-0.93). STBound combined with GI provided a significantly improved performance over all individual trend-based metrics and compared to all

trend-based metrics combined with GI (all p-values <0.05). STBound combined with GI also had improved predictive ability during intermediate time points compared to competing methods.

3.1.4 Conclusions

STBound may be a valuable diagnostic tool for clinical use in detecting glaucomatous VF progression earlier and more accurately than existing methods. Given its lack of similarity with current methods, we recommend that it be used in conjunction with those methods to improve predictive performance.

3.2 Introduction

The chronic and progressive nature of OAG requires patients to be monitored over time in order to establish whether the disease is progressing or not. Determination of progression status is important for timely adjustment or escalation of treatment. However, detecting glaucomatous functional progression is the most challenging aspect of glaucoma management (Gardiner and Crabb, 2002; Artes et al., 2005; Vianna and Chauhan, 2015; Medeiros and Tatham, 2016). Despite the clinical benefit of detecting VF progression and the increased use of VF data as an endpoint in clinical trials, there is no standard method for determining VF progression (Gordon et al., 2002; Leske et al., 2007; Musch et al., 2009; Chauhan et al., 2010; De Moraes et al., 2012; Garway-Heath et al., 2015). One of the factors contributing substantially to this difficulty is the variability in test results. The variability due to measurement error and/or factors related to the psychophysics of the test, both within a given stimulus location and between consecutive VF examinations, can mimic VF change in the absence of true disease worsening or mask a true glaucoma-related functional change.

White-on-white SAP is the most commonly used functional testing method for VF assessment. SAP, which can be performed with the HFA (Carl Zeiss Meditec, Dublin, CA, USA), the Octopus (Haag-Streit, Koniz, Switzerland), or other similar devices, provides sensitivity estimates of each test point in the VF. These sensitivities are monitored over time to determine if glaucomatous progression is occurring for an eye. For decision-making about progression status, most clinicians use either their own subjective assessment of VF series or assessment of statistically significant progression determined by software with automated algorithms built into SAP devices.

Currently, statistical modeling for assessing VF progression mainly relies on event- or trend-based analyses (Spry and Johnson, 2002). Event-based analyses (e.g., HFA GPA) express progression status in a binary manner (progression or no progression) after comparing a recent VF test with baseline results, often without considering test results from all other visits (Heijl, 1991). These methods use pre-specified levels of deterioration in order to define progression. Trend-based analyses (e.g., HFA-compatible PROGRESSOR, Medisoft Inc, Leeds, UK) assess VF progression using rates of loss per year determined by regression modeling of global indices (e.g., MD) (Bengtsson and Heijl, 2008; Bengtsson et al., 2009; Artes et al., 2011), sectoral mean sensitivity, and/or sensitivity of individual test points over time (Smith et al., 1996). Advanced methods including spatial filters (Spry and Johnson, 2002; Gardiner et al., 2004; Strouthidis et al., 2007), statistical modeling of spatial correlation (Zhu et al., 2014, 2015; Warren et al., 2016), general hierarchical Bayesian modeling (Bryan et al., 2017), and machine learning (Yousefi et al., 2014) have also been introduced to properly differentiate true signal from random variability.

The majority of previous models in this setting focus mainly on changes in VF sensitivities over time in some form (e.g., changes from baseline; trends over time), but do not explicitly consider how changes in variability, and more specifically spatial correlation, over time may be an equally important indicator of progression. This idea is partially supported by the known inverse relationship between VF sensitivities and testing variability (Russell et al., 2012). In Chapter 2, we developed statistical theory for a novel spatiotemporal boundary detection statistical model that considers the anatomy of the optic disc when modeling longitudinal VF data and successfully quantifies changes in spatial correlation across time. This measure of variability in the spatial correlation over time, hereafter referred to as STBound, was shown to be an important predictor of glaucomatous VF progression. The purpose of this study is to extend STBound and validate it through comparison with existing VF progression statistical methods. Classification accuracy and the speed of the competing methods to correctly diagnose progression are both investigated.

3.3 Methods

3.3.1 Data Description

We source data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry, Perth, Western Australia, which contain 1,448 VFs from 194 distinct eyes (98 patients in total) (Betz-Stablein et al., 2013). All of the subjects have some form of OAG. Three of the eyes were removed from the final analysis because of missing information on the clinical determination of progression. The mean follow-up time for the remaining participants is 2.6 years (range: 0.2, 9.4) with an average of 7.4 tests (range: 2, 21) per eye. Glaucomatous VF progression status was determined at the final visit based on the expertise of two independent clinicians. A third clinician evaluated any discrepancies between the two primary assessors (only necessary for 13 VF series). The clinicians had access to GPA output and the VF printouts of the tests across time when making their determinations. In total, there are 141 (74%) stable and 50 (26%) progressing eyes. The trial adhered to the tenets of The Declaration of Helsinki and was approved by the University of Western Australia's ethics committee.

A description of the VF data collection has been previously provided (Betz-Stablein et al., 2013). Briefly, the VF data were obtained using the HFA 24-2 program with 15 patients using the full threshold SAP, 81 patients using SITA, and two patients with a combination of both. In order to remove the learning effect often seen in first time VF test takers, each patient took at least two VF tests before data collection began.

3.3.2 Spatiotemporal Boundary Detection

The statistical method developed in Chapter 2 provides a novel approach to diagnosing glaucomatous VF progression. Unlike previous methods that focus on estimating trends in changing VF sensitivities over time, it quantifies changes in the spatial correlation structure of VF data across time and uses the developed metric, STBound, to detect VF progression.

For each VF test, a statistical method known as spatial boundary detection with a dissimilarity metric is applied to the collected data and the spatial correlation structure of the data at the current visit is estimated. This correlation is partially determined by the anatomy of the optic disc and the

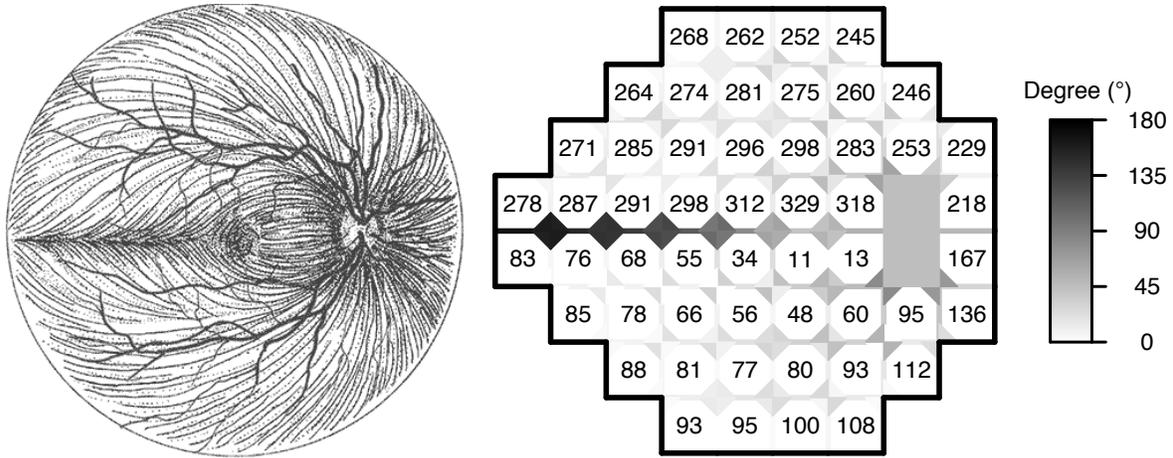


Figure 3.7: The metric STBound describes variability in the spatial structure of the VF across time. The statistical model incorporates the anatomy of the eye, where the distance at which the underlying nerve fibers enters the optic disc dictates spatial correlation between data from two VF locations. These angles, measured by Garway-Heath et al. (2000), are displayed across the VF (right). The edges and corners are shaded to represent the distance between bordering locations on the VF, where darker shading represents a larger distance. The RNFL is displayed to demonstrate the similarity of the pattern that appears on the VF by using the Garway-Heath angles (left). The RNFL (left) is printed with permission from Wallace L.M. Alward (Ophthalmology and Visual Sciences, University of Iowa Carver College of Medicine).

relationship between VF loss and damage to the optic disc with respect to the death of RGCs (Figure 3.7) (Garway-Heath et al., 2000). In practice, all data from each VF test for a particular eye are modeled simultaneously in order to account for similarities in estimation across time. If the estimated spatial correlation is changing substantially over time, we showed that this is highly predictive of progression (Chapter 2). If the estimates of spatial correlation are stable across time, the eye is more likely to be classified as non-progressing. In order to quantify the variability in spatial correlation across time and to create a standardized metric to facilitate comparisons across eyes, we used the CV (large value: high variability; small value: low variability).

While we previously only investigated the estimated CV value as a predictor of VF progression, in this chapter we explore the use of the estimated CV value, its standard error, and their interaction in combination, and refer to this predictor as STBound. Figure 3.8 shows an example of the variability in the estimated spatial correlation across time for two patients in the study, one progressing, one stable, along with the STBound metric (i.e., the predicted probability of progression). The full

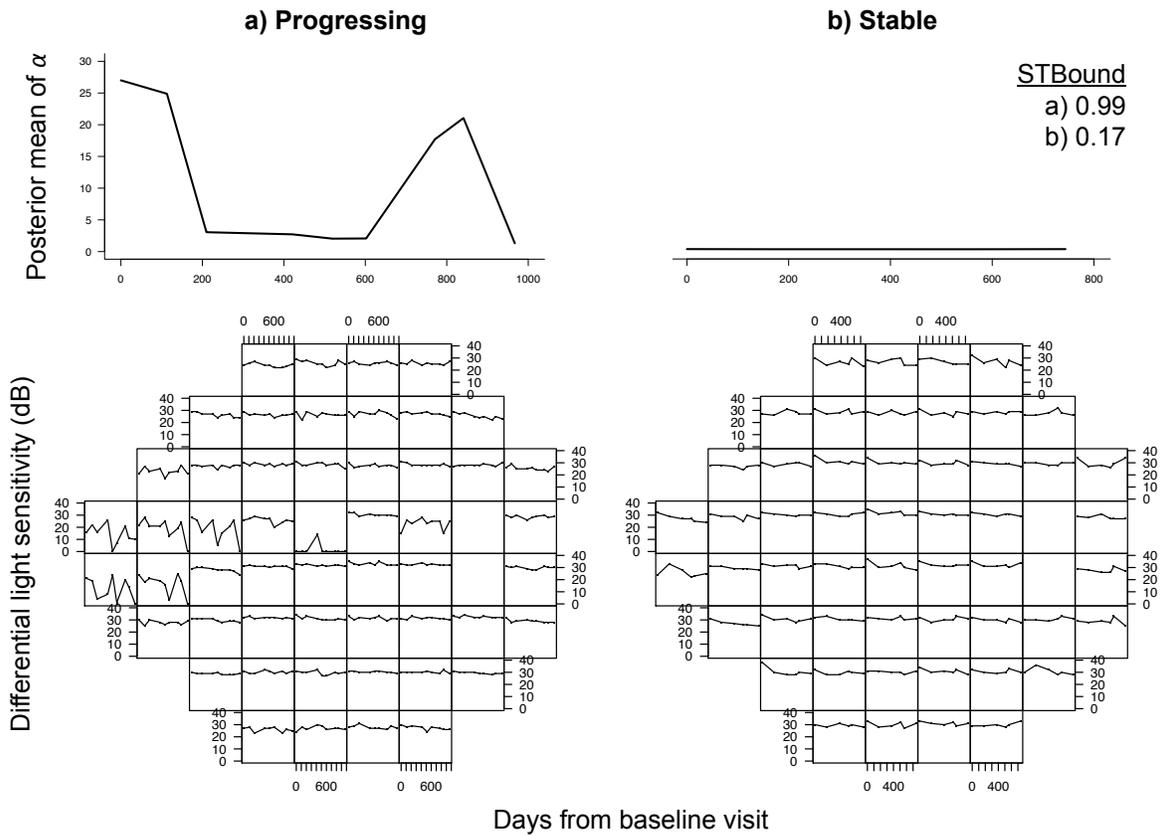


Figure 3.8: Longitudinal series of VFs for one progressing (left) and stable (right) eye are presented in the bottom row. Each cell represents the observed VF sensitivities over all visits. The progressing eye has more variability across the VF and in particular in the region between the superior and inferior sectors. In the top row, the estimated parameter that describes the spatial correlation in the VF data, α , is plotted over time and the corresponding values of STBound (i.e., predicted probability of progression) are given as 0.99 and 0.17 for the progressing and stable eye, respectively. For these two patients STBound indicates that eye a) is much more likely to be progressing.

details on the statistical model, including model fitting, can be found elsewhere. The model is freely available and implemented in the R package `womb1R` (Berchuck, 2017).

3.3.3 Model Comparison

In this chapter, we further explored the potential of STBound in diagnosing VF progression in comparison to existing methods and its ability to correctly diagnose earlier than these methods. To formally make these comparisons, we constructed logistic regression models, regressing all developed predictors of VF progression on the clinical assessment of progression and calculated the AIC, area under the receiver operating characteristic (ROC) curve (full and partial), sensitivity, and specificity. The final metrics used for diagnosing progression are the predicted probabilities of progression from each of the fitted logistic regression models.

We compared STBound to a GI regression method, the MS method (Bhandari et al., 1997), PLR (Smith et al., 1996), and SPROG (Betz-Stablein et al., 2013). In GI, we first averaged the VF sensitivities during an exam to produce a single VF value at that point in time. We then repeated this process separately for data from each VF test. Finally, we regressed these averages (one from each test) on time to determine the average rate of change in the VF across time. It is worth noting that any alternative global indices (e.g., MD) could be used in the regression and we chose the average VF sensitivities due to ease of interpretation and availability of information. We used the estimated regression slope and p-value, along with their interaction, as predictors of progression. Including the p-value when determining progression allows us to differentiate between statistically significant and non-significant slope estimates and should improve our ability to detect true progression.

GI will work well when large sections of the VF are degrading over time, but may struggle to detect highly localized damage (Chauhan et al., 1990; Birch et al., 1995; Katz et al., 1999). As a result, we also considered several location-specific analyses. PLR is a commonly used trend-based point-wise method, in which detection of progression corresponds to statistically significant slopes of the VF location-specific regression lines (sensitivities regressed on time) (Smith et al., 1996). We present results from four versions of PLR, where the progression criterion is based on the 1st, 2nd, 3rd, and 4th smallest slope p-values from all location-specific regressions being less than 0.001 (P1, P2, P3, P4) (Smith et al., 1996). Therefore, the 1st, 2nd, 3rd, and 4th smallest slope p-values are used as predictors of progression in the logistic regression analyses. In MS (Bhandari et al., 1997), we

Table 3.3: Summary of criteria included in each diagnostic method. For each of the metrics, the covariates that are listed are included in a logistic regression model used to predict clinical assessment of progression. The resulting predicted probabilities of progression are used as the progression metric.

Method	Progression Metric
GI	Slope and corresponding p-value from regressing the average (across all VF locations) DLS from each test across time, and their interaction.
MS	Overall mean slope of DLS over time for all location specific regression lines with a p-value <0.01 (Bhandari et al., 1997).
P1	1st order statistic of slope p-values from the location specific regression (Smith et al., 1996).
P2	2nd order statistic of slope p-values from the location specific regression (Smith et al., 1996).
P3	3rd order statistic of slope p-values from the location specific regression (Smith et al., 1996).
P4	4th order statistic of slope p-values from the location specific regression (Smith et al., 1996).
SPROG	The estimated overall slope that describes change in sensitivities over time across the entire VF after fitting a hierarchical Bayesian disease mapping statistical model (Betz-Stablein et al., 2013).
STBound	The posterior mean and standard deviation of the CV of the parameter that describes the spatial correlation of the VF data, and their interaction (Chapter 2).

first computed PLR at each VF location and then averaged the estimated slopes whose p-values were <0.01 . This average slope value was used as a predictor of progression in the logistic regression analysis.

We also include an advanced spatiotemporal statistical method for comparison (Betz-Stablein et al., 2013). SPROG simultaneously models sensitivities from each VF location linearly over time using hierarchical Bayesian disease mapping techniques while also accounting for the underlying anatomy of the eye. If the overall slope that describes change in sensitivities over time is statistically significantly less than zero, then SPROG classifies the eye as progressing. We included the posterior probability that the global slope is less than zero as a predictor of progression in the logistic regression analysis. Full details of SPROG have been previously detailed (Betz-Stablein et al., 2013). A full summary of the progression metrics is given in Table 3.3.

We compared metrics using AIC, AUC, partial AUC (pAUC), sensitivity, and specificity. A smaller value of AIC indicates an improved model fit, while larger values of AUC, pAUC, sensitivity, and specificity indicate superior discriminatory ability (Dodd and Pepe, 2003; Pepe, 2003). We

limited the pAUC to regions of clinically relevant specificity, 85-100% (Zhu et al., 2015). In addition, we used Pearson correlation and statistical hypothesis tests to investigate the similarity between each of the considered metrics. Methods that yield predictors with low correlation may be combined in order to improve the prediction of the progression outcome.

3.3.4 Assessing Early Stage Progression

Diagnosing glaucomatous VF progression early is critical for limiting irreversible vision loss, thus we study the performance of each metric in the early phases. To accomplish this, we study the longitudinal trends in AUC, pAUC, sensitivity, and specificity for each metric in the initial years from baseline with respect to the clinical determination of progression on the final testing date. We estimate these summaries at half-year increments up to 4.5 years of follow-up. Only visits that occurred on or before the cutoff time are included in each analysis. This imitates a clinical setting where each metric is calculated at every visit and progression is diagnosed.

3.4 Results

3.4.1 Diagnostic Capability of the Metrics

The results comparing the diagnostic capabilities of the different methods at the end of the study period are shown in Table 3.4. Assessment of the performance of each individual metric (“Without GI” columns of Table 3.4) indicates that each of the trend-based methods are highly predictive of progression, as noted by the significant marginal p-values that describe the overall utility of the predictor. The global metric, GI (AIC: 204.11; AUC: 0.70; pAUC: 0.26), is superior to the other existing trend-based methods (although not statistically significant, p-values: 0.05-0.99) in terms of AIC (210.46-217.55), AUC (0.63-0.67), and pAUC (0.13-0.26). The newly developed metric, STBound, has improved diagnostic ability compared to the competing trend-based methods, including GI, with respect to AIC (197.77), AUC (0.74), and pAUC (0.31); though the AUC and pAUC values are not statistically significantly larger than the values produced by GI (p-values: 0.51-0.60). However, this improved performance indicates that STBound may be useful for diagnosing progression on its own.

Table 3.4: Assessing the diagnostic capability of the metrics. Each metric is regressed against the clinical assessment of progression, both with and without the global index (GI). AIC, AUC and pAUC. The pAUC is limited to the clinically relevant range of specificity, 85-100%. Also reported are p-values corresponding to hypothesis tests *) for each metric marginally, and †) in addition to GI. The bold cells indicate a significant improvement in AUC or pAUC over the GI model at the $\alpha = 0.05$ level.

	Without GI				With GI			
	AIC	AUC	pAUC	p-value*	AIC	AUC	pAUC	p-value†
GI	204.11	0.70	0.26	<0.001	—	—	—	—
MS	210.45	0.64	0.24	<0.001	204.49	0.70	0.31	0.203
P1	214.16	0.64	0.13	0.002	201.50	0.74	0.28	0.032
P2	217.55	0.64	0.26	0.014	204.52	0.71	0.26	0.207
P3	215.13	0.63	0.19	0.004	203.43	0.72	0.27	0.102
P4	214.12	0.65	0.18	0.002	203.32	0.71	0.28	0.095
SPROG	214.84	0.67	0.15	0.003	206.11	0.70	0.26	0.953
STBound	197.77	0.74	0.31	<0.001	180.41	0.81	0.45	<0.001

3.4.2 Correlation Between Methods

The major strength of STBound is the fact that it doesn't rely on an underlying trend-based method, and is therefore providing unique predictive information to the decision-making process. This fact is shown clearly in Table 3.5 where the Pearson correlation estimates and corresponding p-values between each of the metrics from Table 3.4 are presented. The estimated correlations are presented below the diagonal with the p-values above the diagonal. All of the existing trend-based methods are highly correlated with each other (all p-values ≤ 0.05). STBound, however, is uncorrelated with the existing methods as indicated by the large bolded p-values (all p-values > 0.05). These correlations suggest that STBound explains a unique alternative pathway in understanding glaucomatous VF progression that is not incorporated into existing trend-based methods. As such, STBound can be used in conjunction with trend-based methods.

3.4.3 Combination of Methods

We tested the performance of combining GI with the other metrics, including STBound, relative to these metrics alone. GI was selected because it was shown to be the top performing trend-based method in the study. The "With GI" columns of Table 3.4 display the results when each of the metrics are combined with GI. Based on the correlation analysis in Table 3.5, we didn't expect large improvements in diagnostic ability when GI is added to the other trend-based methods due to their

Table 3.5: Correlation matrix and hypothesis test results for the diagnostic metrics. The estimated Pearson correlations and p-values between the metrics are presented below and above the diagonal, respectively. Bold values indicate that the correlation does not statistically differ from zero at the $\alpha = 0.05$ level of significance.

Metric	GI	MS	P1	P2	P3	P4	SPROG	STBound
GI	—	0.00	0.00	0.00	0.00	0.00	0.00	0.26
MS	0.55	—	0.00	0.00	0.00	0.00	0.00	0.93
P1	0.26	0.34	—	0.00	0.00	0.00	0.05	0.48
P2	0.31	0.25	0.83	—	0.00	0.00	0.00	0.47
P3	0.33	0.21	0.72	0.92	—	0.00	0.00	0.23
P4	0.38	0.24	0.68	0.85	0.92	—	0.00	0.18
SPROG	0.57	0.32	0.14	0.27	0.26	0.28	—	0.11
STBound	0.08	-0.01	0.05	0.05	0.09	0.10	0.11	—

large degree of similarity overall. This appears to be the case as p-values that describe the statistical significance of the added metrics are generally >0.05 , with the exception of P1 (p-value = 0.032) (though the improvements in AUC and pAUC are not statistically significant). STBound is the only metric that improves the predictive performance across all diagnostic indicators. When STBound is added to GI the increase in AUC and pAUC are statistically significant with p-values of 0.002 and 0.003, respectively. The model with GI and STBound (referred to as GI + STBound) has optimal characteristics with a minimum AIC of 180.41 vs. 197.77, and maximum AUC and pAUC of 0.81 vs. 0.74 and 0.45 vs. 0.31, respectively. The estimated regression coefficients, standard errors, and p-values for GI + STBound are presented in Table 3.6.

Table 3.6: Estimated coefficients from the logistic regression model combining the global index (GI) and STBound.

	Estimate (SE)	p-value
Intercept	-1.75 (0.97)	0.073
GI: P-value	-0.25 (0.75)	0.738
GI: Slope	-13.73 (4.2)	0.001
GI: P-value * Slope	18.7 (2.83)	0.008
STBound: Mean	7.47 (6.53)	0.172
STBound: SE	-8.92 (7.37)	0.011
STBound: Mean * SE	-13.37 (14.97)	0.372

In Figure 3.9, ROC curves are displayed for each of the metrics from Table 3.4 that were significant, GI, MS, P1, SPROG, STBound, GI + P1, and GI + STBound. We only present P1 to represent the PLR methods in order to simplify the figure. The ROC curve can be interpreted as sensitivity and (1-specificity) over various thresholds. Figure 3.9 confirms the results in Table 3.5, as

the ROC curve of GI + STBound provides significant improvement over the GI curve, in particular in the region of high specificity. Although P1 was shown to improve GI, it is clear from the ROC curves that the improvement is marginal at best.

3.4.4 Determination of Progression Over Time

Table 3.7: Estimates of AUC and pAUC presented at yearly intervals. These raw estimates correspond to the LOESS smoothed curves presented in Figure 3.10. Bold cells indicate a significant improvement in either AUC or pAUC over the corresponding values of the global index (GI) at the $\alpha = 0.05$ level.

Metric	Criteria	Years from baseline visit				
		1	2	3	4	End
GI	AUC	0.55	0.66	0.73	0.73	0.70
	pAUC	0.07	0.23	0.31	0.30	0.26
MS	AUC	0.52	0.58	0.64	0.66	0.64
	pAUC	0.07	0.22	0.29	0.29	0.24
P1	AUC	0.56	0.60	0.65	0.65	0.64
	pAUC	0.12	0.10	0.11	0.14	0.13
SPROG	AUC	0.58	0.63	0.71	0.71	0.67
	pAUC	0.11	0.14	0.27	0.26	0.15
STBound	AUC	0.59	0.70	0.74	0.75	0.74
	pAUC	0.16	0.22	0.27	0.32	0.31
GI + P1	AUC	0.57	0.69	0.76	0.75	0.74
	pAUC	0.20	0.21	0.32	0.32	0.28
GI + STBound	AUC	0.52	0.74	0.82	0.82	0.81
	pAUC	0.14	0.35	0.45	0.47	0.45

We also explored the performance of the metrics in the early years of the study with respect to progression determined at the end of the study. The presentation of the longitudinal trends of AUC and pAUC are displayed in Figure 3.10 for all metrics found to be significant in Table 3.4 (again only including P1 to represent the PLR methods). The figure presents smoothed local regression (LOESS) estimates of AUC and pAUC (Cleveland and Devlin, 1988). In Figure 3.10a, the AUC trends are presented, where GI + STBound has superior AUC over the entire time period. The pAUC trends are displayed in Figure 3.10b, with an even more pronounced separation of GI + STBound over the other metrics. Again, it can be seen that although P1 significantly improved upon GI, this improvement did not transfer to the operating characteristics of GI + P1. Table 3.7 presents the raw estimates of AUC and pAUC at the first four years of the study, corresponding to the smoothed lines presented in Figure 3.10. Bolded entries in the table coincide to values that significantly improve upon the

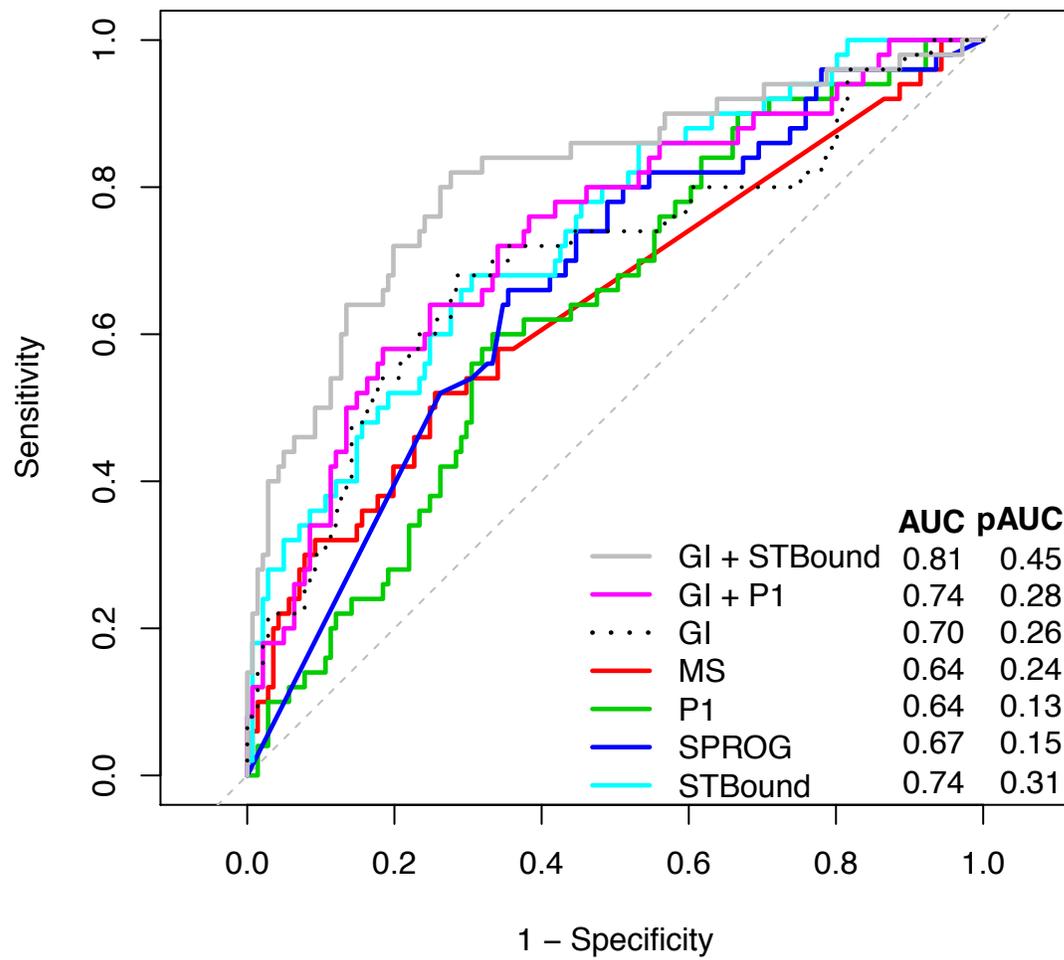


Figure 3.9: Receiver operating characteristic curves for statistically significant diagnostic metrics.

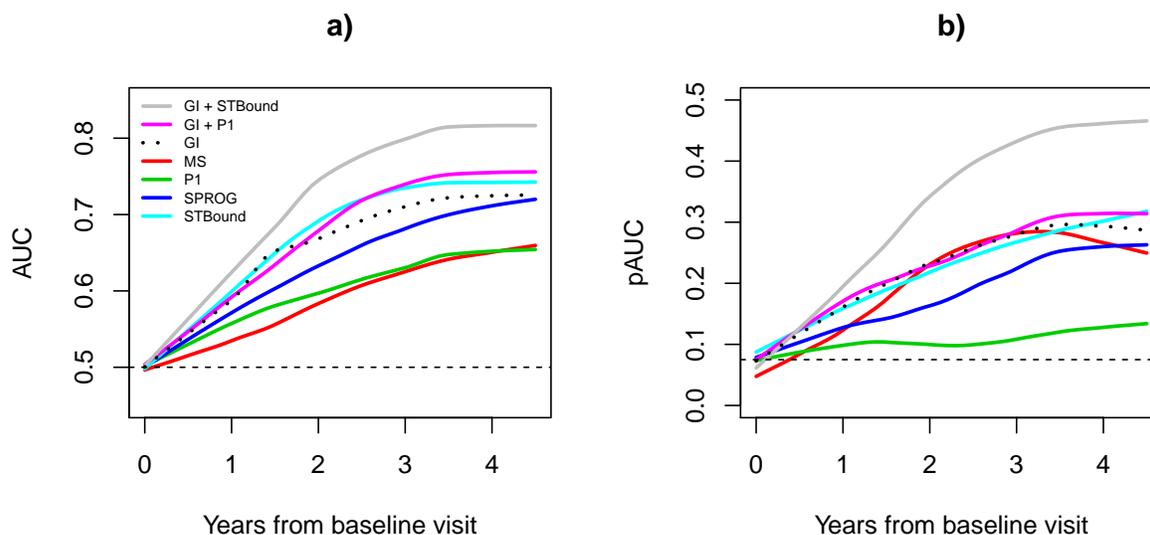


Figure 3.10: Performance of diagnostic metrics in the initial years from baseline visit, AUC (a) and pAUC (b). Estimates are presented as smooth curves using the LOESS method. The horizontal dashed line indicates the value at which a metric does not contribute any diagnostic information.

GI value at the same year, demonstrating that the gains seen in GI + STBound in Figure 3.10 are statistically significant. Overall, GI + STBound is shown to produce improved diagnostic metrics earlier than the competing methods.

Figure 3.11 and Table 3.8 present sensitivity and specificity estimates for each metric annually for four years after the baseline visit and at the end of the study. The sensitivity and specificity values are obtained through calculation of an optimal threshold. The threshold in use is motivated clinically, so that the specificity is guaranteed to be at least 85%, while maximizing sensitivity. The thresholds are calculated using the completed study and are given in Table 3.8. In Figure 3.11, it is clear that the sensitivity is superior in the GI + STBound model beginning early in follow-up (0.34-0.64 vs. 0.06-0.60). Due to the definition of the threshold, the specificity for all models is around the 85% level throughout follow-up. Specifically in Table 3.8, by the end of the study GI + STBound has optimal operating characteristics, with specificity in the clinical range, while showing statistically significant improvement in sensitivity over GI. In the earlier years, the specificity does not deteriorate, and the sensitivity only truly begins to drop-off significantly within two years from baseline, with sensitivity of 0.52.

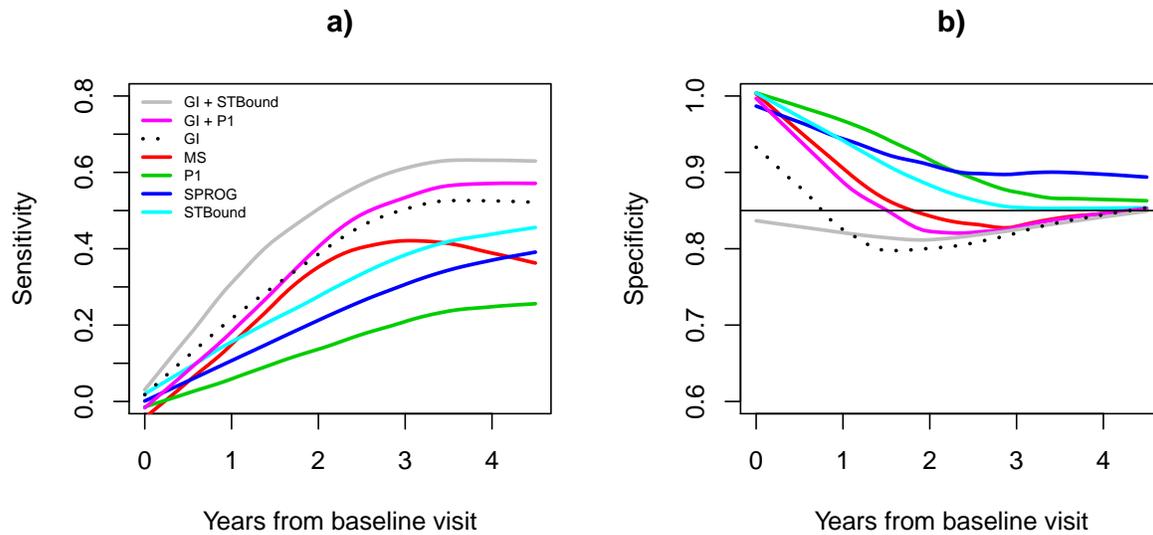


Figure 3.11: Demonstrating the performance of sensitivity (a) and specificity (b) in the initial years from baseline visit. Both sensitivity and specificity are obtained by using a clinically motivated threshold defined to maximize sensitivity, while forcing the specificity to be no smaller than 85%. Estimates are presented as smooth curves using the LOESS method.

Finally, we explore how much faster the GI + STBound metric is able to successfully diagnose progression in terms of days from baseline visit. For all 50 progressing patients, we determine which visit each metric would have identified the patient as progressing using the clinically motivated thresholds in Table 3.8. If a model was incapable of detecting progression at the end of the study time, they were assigned the date of the last visit. Figure 3.12 presents the density estimate of the time of first diagnosis for each model in days from baseline visit along with the mean number of days. The GI + STBound metric is capable of diagnosing progressing fastest, improving approximately two months over the second fastest, GI. Meanwhile, when P1 was added to GI there is no improvement, further emphasizing the importance of STBound for detecting progression earlier.

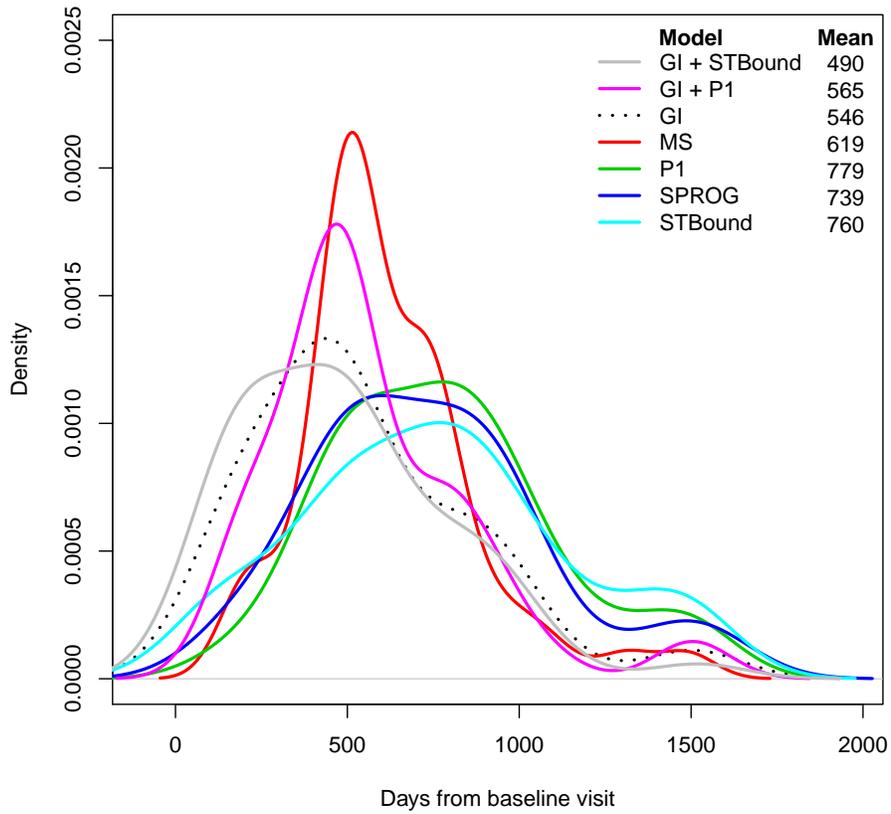


Figure 3.12: Density estimates of the time of first diagnosis for each model in days from baseline visit along with the mean number of days.

Table 3.8: Sensitivity (sens.) and specificity (spec.) for each metric using the clinical definition for thresholds.

Metric	Threshold		Years from baseline visit				
			1	2	3	4	End
GI	0.344	<i>Sens.</i>	0.22	0.38	0.54	0.52	0.48
		<i>Spec.</i>	0.78	0.76	0.82	0.84	0.85
MS	0.322	<i>Sens.</i>	0.12	0.38	0.42	0.42	0.34
		<i>Spec.</i>	0.88	0.79	0.83	0.85	0.85
P1	0.338	<i>Sens.</i>	0.06	0.16	0.22	0.26	0.24
		<i>Spec.</i>	0.90	0.91	0.86	0.87	0.85
SPROG	0.323	<i>Sens.</i>	0.10	0.22	0.36	0.38	0.52
		<i>Spec.</i>	0.94	0.88	0.91	0.89	0.74
STBound	0.302	<i>Sens.</i>	0.16	0.28	0.40	0.46	0.44
		<i>Spec.</i>	0.94	0.87	0.85	0.86	0.85
GI + P1	0.335	<i>Sens.</i>	0.18	0.42	0.60	0.58	0.52
		<i>Spec.</i>	0.81	0.79	0.83	0.84	0.85
GI + STBound	0.325	<i>Sens.</i>	0.34	0.52	0.64	0.64	0.64
		<i>Spec.</i>	0.77	0.81	0.83	0.84	0.85

3.5 Discussion

This chapter validates the method introduced in Chapter 2 as a modeling framework for diagnosing VF glaucoma progression. The method utilizes an innovative spatiotemporal boundary detection framework that accounts for the anatomy of the eye, resulting in a novel diagnostic metric that represents a departure from standard trend-based methods. Instead, STBound quantifies change in the spatial correlation of the VF data, representative of underlying damage to the optic disc across time.

When establishing a novel diagnostic, assessing its performance marginally is fundamental; however demonstrating the incremental performance over established metrics is necessary for introducing it for clinical utility. All of the metrics in the study were found to be marginally statistically significant predictors of progression (Table 3.4, “Without GI” columns). This was not surprising, since all but STBound were established diagnostics. However, only STBound provided improved performance in addition to GI (Table 3.4, “With GI” columns). Although P1 was also significant, it is clear from the results that this improvement was minimal. The established metrics GI, MS, PLR, and SPROG are derived from analyzing changes in the mean trend of VF sensitivities across time and therefore, explain similar components of progression. This is confirmed in Table

3.5, where the Pearson correlation estimates for all of the mean trend metrics are significant. On the other hand, the Pearson correlations between STBound and the mean trend metrics are close to zero, establishing their independence. As an independent diagnostic, STBound has clinical utility in its ability to assess an adjacent pathway in glaucoma progression.

Having demonstrated the utility of STBound clinically, we then established it as significantly superior in the important early stages from baseline visit as diagnosing progression early is critical for limiting irreversible vision loss. The AUC, pAUC, sensitivity, and specificity show GI + STBound is superior over the entire early stage (Figures 3.10-3.11; Tables 3.7-3.8). This indicates the accelerated efficacy of STBound in the early stages from baseline. It also indicates the independence of STBound from GI, as GI + STBound is able to improve over the GI curve, while the GI + P1 metric makes no notable improvements. Early detection of glaucoma progression is important clinically as interventions can be effective in reducing a patient's likelihood of vision loss (Heijl et al., 2002). Therefore, diagnostic metrics such as STBound are critical for providing better patient outcomes.

With further validation, we propose that the metric STBound be integrated into the clinical environment. The metric STBound requires a minimum of two visits to compute and from Figure 3.12, we know that it can be used to decrease the time until a correct progression diagnosis by two months on average over GI. In practice, clinicians can implement the method defined in this study based on the clinically defined thresholds in Table 3.8. We suggest using the GI + STBound model, as it has optimal operating characteristics across the entire study (Table 3.8). These characteristics are relatively conservative, and this is due to the definition of the threshold, which prioritizes high values of specificity. This argument is based on the idea that it is better to overestimate progression than underestimate it, since doing so could delay treatment and lead to vision loss. GI + STBound can be calculated using the regression coefficients presented in Table 3.6 along with the individual components of STBound that can be obtained using the `womb1R` R package (Berchuck, 2017).

CHAPTER 4: A SPATIALLY VARYING CHANGE POINT MODEL FOR ASSESSING GLAUCOMA PROGRESSION FROM VISUAL FIELD DATA

4.1 Overview

The assessment of glaucoma progression using VF data is defined by slow (or stable) deterioration, followed by a rapid decrease in visual ability. Determining the transition point of the disease trajectory to a more severe state is important clinically for disease management. We present a unified framework that brings together multiple vital aspects of glaucoma management, i) prediction of future VF sensitivities, ii) predicting the timing and spatial location of future vision loss, iii) making clinical decisions regarding progression, and, iv) incorporation of anatomical information to create plausible data-generating models. We accomplish this by introducing a spatially varying change point model to detect structural change in both the mean and variance process across the VF. The change point is modeled using a latent process that represents the biological change and multivariate spatial structure is introduced that allows the spatial structure of the VF to mimic the anatomy of the eye. Finally, we show that our method improves prediction and estimation of progression, and simulations are presented, showing the proposed methodology is preferred over existing models for VF data. An R package called `spCP` is provided that implements the method.

4.2 Introduction

Glaucoma is an optic neuropathy that is the leading cause of irreversible vision loss worldwide (Tham et al., 2014). Once a patient is diagnosed, they are at risk of disease progression and are closely monitored by clinicians who must balance the potential of further irreversible deterioration of vision with the costs (financial and physical) of medical and surgical interventions. The functional ability of a glaucoma patient is most often measured through SAP, an interactive machine that estimates the threshold sensitivity to light at each of 54 locations across a patient's field of vision. The light is generated with increasing brightness ranging from approximately 40 dB (excellent vision) to 0

dB (near blindness). The patient uses a hand-held trigger to indicate whether the light stimulus was detected and the machine reports the dimmest stimulus detected at each location. The resulting gridded spatial layout is an estimate of a patient's field of vision and is called the VF. These VF tests are performed during clinical visits over time and results are monitored for signs of glaucomatous VF progression (Weinreb and Khaw, 2004). Determining if the disease is progressing remains one of the most challenging aspects of glaucoma management, as it is difficult to differentiate true progression from random testing variability (Vianna and Chauhan, 2015).

Although VF testing is a common technique for assessing glaucoma progression, there is currently no reference standard approach for converting it into a clinical decision. The majority of past studies can be grouped into two categories, trend-based and global-based. Global-based methods provide a single summary of the VF data at each visit, thus aggregating and potentially ignoring key information across VF grid locations (Heijl et al., 1987). These global summaries are then monitored over time to investigate the potential of disease progression. Trend-based methods analyze the time series of sensitivities at each VF location (Fitzke et al., 1996). These methods often resort to separately modeling the sensitivities over time assuming a linear relationship (Katz et al., 1997; Ernest et al., 2012). The inference obtained from these methods are often poor diagnostics of progression or have inferior prediction accuracy, and suffer from low power because they generally ignore the spatial nature of VF data and aggregate over space (Vianna and Chauhan, 2015). To improve the prediction of future VF sensitivities for glaucoma patients, we develop a spatiotemporal model that relaxes questionable assumptions made in past modeling attempts.

In this chapter, we develop a model that utilizes change points (CP) to accurately model the sensitivities over time. Patients diagnosed with glaucoma are often monitored for years with only slow changes in visual functionality (Heijl et al., 2002). It is not until disease progression that notable vision loss occurs, and the deterioration is often swift (Jay and Murdoch, 1993). This disease course inspires a modeling framework that can identify a point of structural change in the course of follow-up. By utilizing the CP framework, the concept of disease progression becomes intrinsically parameterized into the model. Not only does the specification increase flexibility in modeling the trajectory of the VF time series, generalizing the linearity assumption, but the CP represents the point of structural change and can be used to make statements about progression. This framework is particularly rich, however, because in addition to there being a structural change

in the VF sensitivities upon the start of progression, there is also an increase in variance in areas of decreased sensitivity. This inverse variance relationship upon progression is well known in the glaucoma literature, although is often ignored (Russell et al., 2012). In this chapter, we use a CP to segment the variance into a stable process that has the flexibility to increase after the progression point. Finally, a focus will be on the joint modeling of all available VF data for a patient, accounting for the underlying spatial correlation in the data.

The VF exhibits a complex spatial structure that is best modeled considering the anatomy of the eye. Naive spatial structures that assume adjacent neighbors are uniformly correlated ignore this anatomy which can aide in modeling these adjacencies. In Chapter 2, we introduced a framework for inducing localized neighbors as a function of ocular anatomy using a spatiotemporal boundary detection method. This method utilized a measure developed by Garway-Heath et al. (2000) that estimates the angle that each VF test location's underlying retinal nerve fiber enters the optic disc. The Garway-Heath angles are used as a dissimilarity metric, a covariate that explains localized spatial structure. In the present chapter, the methodology from Chapter 2 is extended to the multivariate setting, such that the Garway-Heath angles dictate the neighborhood on the VF for the multiple spatial parameters, including the mean and variance processes and CPs. We believe this is the first extension of boundary detection using a dissimilarity metric to the multivariate setting.

In this chapter, we develop a spatiotemporal CP model to both improve prediction of future VF sensitivities and assess progression for a glaucoma patient. Having the ability to accurately predict future deterioration of the VF will improve a clinician's ability to make timely, data-driven treatment adjustments to preserve vision. By relaxing questionable assumptions from prior studies the predictive power and uncertainty in the predictions are increased. Furthermore, a latent CP process is introduced that allows us to spatially predict CPs at spatial locations that are currently censored. Previous statistical work has focused on one or some of these aspects typically, but none have considered each simultaneously. This chapter is outlined as follows. Section 4.3 details CP theory and the current literature for spatial methods for CPs. In Section 4.4, our newly developed spatial CP method is detailed and we develop a multivariate boundary detection spatial process for localized smoothing. We apply our method to a dataset of VF tests from glaucoma patients in Section 4.5 and present a simulation study in Section 4.6 that demonstrates an improved estimation of CPs using our model. We conclude in Section 4.7 with a discussion.

4.3 Change Point Theory

The detection of CPs refers to the problem of estimating the time at which properties of a time series change. It originated in the 1950s, with the foundational work of Page (1954), and was used as a quality control method to automatically detect faults in industrial processes. For any interval (x_1, x_ν) on the real line ($x_1 < x_\nu$), the standard CP linear regression model with at most one CP is given as,

$$Y_t = \begin{cases} \beta_0 + \beta_1 x_t + \epsilon_t & t = 1, \dots, \theta, \\ \beta_0^* + \beta_1^* x_t + \epsilon_t & t = \theta + 1, \dots, \nu, \end{cases} \quad (4.7)$$

where Y_t , $t = 1, \dots, \nu$ are observations of the dependent variable, x_t is a time varying regressor ($x_1 < \dots < x_\nu$), and $\epsilon_t \stackrel{\text{iid}}{\sim} \text{N}(0, \sigma^2)$. The parameters β_0 , β_1 , β_0^* and β_1^* represent the intercept and slope of the regression line, before and after the CP, $\theta \in \{1, 2, \dots, \nu - 1\}$, respectively.

The parameters are often constrained so that $\beta_0 + \beta_1 x_\theta = \beta_0^* + \beta_1^* x_\theta$, to induce a continuous process, although not differentiable at the CP (Julious, 2001). This constraint reduces the dimensionality of the problem such that,

$$Y_t = \begin{cases} \beta_0 + \beta_1 x_t + \epsilon_t & t = 1, \dots, \theta, \\ \beta_0 + \beta_1 x_\theta + \beta_1^* (x_t - x_\theta) + \epsilon_t & t = \theta + 1, \dots, \nu, \end{cases} \quad (4.8)$$

where the parameter space is reduced to $\Omega = (\beta_0, \beta_1, \beta_1^*, \sigma, \theta)$. Distribution theory for the maximum likelihood estimator of both the regression parameters and CPs is well discussed in the literature (Quandt, 1958; Hinkley, 1970; Menzefricke, 1981) and more recently Bayesian methodology has been explored (Carlin et al., 1992; Raftery, 1994; Perreault et al., 2000). Equations 4.7 and 4.8 present a model for detecting a change in the mean, however, CPs in the variance process have also been addressed (Menzefricke, 1981). The problem of estimation and hypothesis testing for a CP in either the mean or variance process at an unknown time point has been studied extensively in the literature for one CP (Quandt, 1960; Liu et al., 2008), and multiple CPs (Bai, 1997), frequently using likelihood ratio testing techniques. The model in Equation 4.8 can be written as a mixture

distribution,

$$f(\mathbf{Y}|\boldsymbol{\Omega}) = \prod_{t=1}^{\nu} f(Y_t|\omega_1)^{1(t \in \{1, \dots, \theta\})} f(Y_t|\omega_2)^{1(t \in \{\theta+1, \dots, \nu\})}, \quad (4.9)$$

where $\mathbf{Y} = (Y_1, Y_2, \dots, Y_\nu)^T$, $\omega_1 = (\beta_0, \beta_1, \sigma)^T$ and $\omega_2 = (\beta_0, \beta_1, \beta_1^*, \sigma)^T$. The likelihood form presented in Equation 4.9 enforces the interpretation of the CP as a parameter that segments the data into two distinct populations.

Although the CP literature is extensive, the problem of detecting an unknown change in spatially referenced data is relatively sparse. The majority of previously introduced CP regression models for spatial data assume that a single CP is shared across all responses,

$$Y_t(\mathbf{s}) = \begin{cases} \beta_0(\mathbf{s}) + \beta_1(\mathbf{s})x_t(\mathbf{s}) + \epsilon_t(\mathbf{s}) & t = 1, \dots, \theta, \\ \beta_0(\mathbf{s}) + \beta_1(\mathbf{s})x_\theta + \beta_1^*(\mathbf{s})\{x_t(\mathbf{s}) - x_\theta\} + \epsilon_t(\mathbf{s}) & t = \theta + 1, \dots, \nu, \end{cases} \quad (4.10)$$

where the observed data are now spatially referenced, $Y_t(\mathbf{s})$, along with the regression parameters and regressors, $x_t(\mathbf{s})$. The CP, θ , is shared across all locations. This assumption may be inappropriate if the observations are arising from different spatial locations. There have been various Bayesian CP formulations that account for spatial or spatiotemporal dependencies through adjustments in the intercept (Beckage et al., 2007; Yu et al., 2008; Cai et al., 2016), and extensions that allow for independent spatial processes before and after the CP (Majumdar et al., 2005). However, these methods did not consider modeling spatial variability in the CPs themselves. In a recent study, Wagner and Midway (2014) introduced a model with spatially varying intercepts, slopes, and CPs. Their method accounted for correlation among parameters at a specific location, but assumed independence across space. This method has been extended to account for spatial dependency using an MCAR process (Warren et al., 2017).

4.4 Methods

Following the approach of Warren et al. (2017), we propose modeling a set of spatially varying slopes, intercepts, and CPs with an MCAR process. However, we propose extending the framework to allow for a CP in the variance process as well. In addition, we introduce a latent process that

represents the true biological CP. This flexible specification permits spatial prediction of currently censored CPs. Finally, a novel MCAR process is developed that allows for a dissimilarity metric in the spatial covariance. Inference for this model is based on MCMC simulation, and a description of the algorithm is given in Appendix B. Spatial CP models are computationally intensive, so the MCMC algorithm is implemented using Rcpp (Eddelbuettel et al., 2011) and is available from the R package `spCP` (R Core Team, 2016).

4.4.1 Spatial Change Point Model

Let $Y_t(\mathbf{s})$ be an observation from location \mathbf{s} and time x_t , for $\{\mathbf{s}_i : i = 1, \dots, m\}$, $t = 1, \dots, \nu$, where the times are strictly increasing. We introduce a model that allows for location specific CPs in the mean and variance. We model the observed data as follows,

$$Y_t(\mathbf{s}) = \mu_t(\mathbf{s}) + \epsilon_t(\mathbf{s}), \quad \epsilon_t(\mathbf{s}) \stackrel{\text{ind}}{\sim} \mathbf{N}(0, \sigma_t^2(\mathbf{s})),$$

so that both the mean process, $\mu_t(\mathbf{s})$, and variance process, $\sigma_t^2(\mathbf{s})$, are location and time specific. An extension of Equation 4.10, the mean process is defined to be a function of spatially varying intercepts, slopes, and CPs such that,

$$\mu_t(\mathbf{s}) = \begin{cases} \beta_0(\mathbf{s}) + \beta_1(\mathbf{s}) x_t(\mathbf{s}) & \text{if } x_1(\mathbf{s}) \leq x_t(\mathbf{s}) \leq \theta(\mathbf{s}), \\ \beta_0(\mathbf{s}) + \beta_1(\mathbf{s}) \theta(\mathbf{s}) + \beta_1^*(\mathbf{s}) \{x_t(\mathbf{s}) - \theta(\mathbf{s})\} & \text{if } \theta(\mathbf{s}) \leq x_t(\mathbf{s}) \leq x_\nu(\mathbf{s}). \end{cases} \quad (4.11)$$

Furthermore, the regression line standard deviation is unique at each spatial location and changes over time,

$$\log\{\sigma_t(\mathbf{s})\} = \begin{cases} \lambda_0(\mathbf{s}) + \lambda_1(\mathbf{s}) x_t(\mathbf{s}) & \text{if } x_1(\mathbf{s}) \leq x_t(\mathbf{s}) \leq \theta(\mathbf{s}), \\ \lambda_0(\mathbf{s}) + \lambda_1(\mathbf{s}) \theta(\mathbf{s}) + \lambda_1^*(\mathbf{s}) \{x_t(\mathbf{s}) - \theta(\mathbf{s})\} & \text{if } \theta(\mathbf{s}) \leq x_t(\mathbf{s}) \leq x_\nu(\mathbf{s}). \end{cases}$$

To permit spatial prediction of the CP at locations where the CP is currently censored (i.e., $\theta(\mathbf{s}) = x_\nu(\mathbf{s})$ or $x_1(\mathbf{s})$), it is now specified as a continuous quantity. This is in contrast to the standard model that forces the CP to be an index of the time series. When $\theta(\mathbf{s}) = x_1(\mathbf{s})$ the CP has occurred before the start of follow-up, while $\theta(\mathbf{s}) = x_\nu(\mathbf{s})$ indicates the CP has not yet occurred. As such, both

the lower and upper bounds of the CP are censored. Therefore, we model the observed CP as a function of the underlying true biological CP process, $\theta(\mathbf{s}) = \max\{\min\{\eta(\mathbf{s}), x_\nu(\mathbf{s})\}, x_1(\mathbf{s})\}$. The true process, $\eta(\mathbf{s})$, has support on the real line and is allowed to vary spatially.

4.4.2 Multivariate Dissimilarity Metric Spatial Process

We develop a novel MCAR that allows for a dissimilarity metric within the framework introduced by Mardia (1988). Similar to the univariate case, the joint distribution is derived from the full conditional distributions. Under the MRF assumption these conditional distributions can be specified as follows,

$$f(\phi_i | \phi_{j \neq i}, \boldsymbol{\mu}, \boldsymbol{\Gamma}_i) = \text{MVN} \left(\boldsymbol{\mu}_i + \sum_{i \sim j} \mathbf{B}_{ij}(\phi_j - \boldsymbol{\mu}_j), \boldsymbol{\Gamma}_i \right), \quad i, j = 1, \dots, m$$

where $\phi_i = (\phi_{i1}, \phi_{i2}, \dots, \phi_{ip})^T$ and $\boldsymbol{\mu}_i = (\mu_{i1}, \mu_{i2}, \dots, \mu_{ip})^T$ are p -dimensional vectors, $\boldsymbol{\mu} = (\boldsymbol{\mu}_1^T, \boldsymbol{\mu}_2^T, \dots, \boldsymbol{\mu}_m^T)^T$, and $\boldsymbol{\Gamma}_i$ and \mathbf{B}_{ij} are $p \times p$ matrices. Using Brook's lemma, Mardia proved that the full conditional distributions uniquely determine the joint distribution given by, $f(\boldsymbol{\phi} | \boldsymbol{\mu}, \{\boldsymbol{\Gamma}_i\}) \propto \exp \left\{ -\frac{1}{2}(\boldsymbol{\phi} - \boldsymbol{\mu})^T \boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}})(\boldsymbol{\phi} - \boldsymbol{\mu}) \right\}$, where $\boldsymbol{\phi} = (\phi_1^T, \phi_2^T, \dots, \phi_m^T)^T$, $\boldsymbol{\Gamma}$ is a block diagonal matrix with blocks $\boldsymbol{\Gamma}_i$ and $\tilde{\mathbf{B}}$ is an $mp \times mp$ matrix with (i, j) -th block \mathbf{B}_{ij} . Different MCAR forms can be obtained through various specifications of $\boldsymbol{\Gamma}$ and $\tilde{\mathbf{B}}$.

In this case, we induce an MCAR structure that generalizes the univariate version of the Leroux CAR specification, thus allowing the dissimilarity metric to be used in a multivariate setting (Leroux et al., 2000; MacNab and Gustafson, 2007). As in the univariate case, symmetry of $\boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}})$ is required. A convenient special case sets $\mathbf{B}_{ij} = b_{ij}\mathbf{I}_p$. Under this condition, the multivariate version of the Leroux dissimilarity metric likelihood can be obtained with the following specifications,

$$b_{ij} = \frac{\rho w_{ij}(\alpha)}{\rho \sum_{j=1}^m w_{ij}(\alpha) + (1 - \rho)}, \quad \boldsymbol{\Gamma}_i = \frac{\boldsymbol{\Sigma}}{\rho \sum_{j=1}^m w_{ij}(\alpha) + (1 - \rho)}. \quad (4.12)$$

The definitions in Equation 4.12 introduce information about the spatial neighborhood structure through the adjacencies $\{w_{ij}(\alpha)\}$.

The adjacencies are defined similar to the specification in Chapter 2. In particular, they are a function of a dissimilarity metric and underlying parameter, such that $w_{ij}(\alpha) = 1(i \sim j) \exp\{-z_{ij}\alpha\}$,

where $1(\cdot)$ is the indicator function and $i \sim j$ is the event that locations i and j share an edge or corner ($w_{ii} = 0$ for all i). The quantity z_{ij} is a measure of (dis-)similarity between two locations and is defined as $z_{ij} = \|z_i - z_j\|$, where z_i is a covariate at location s_i . The parameter α is forced to be non-negative, so that the adjacencies are in the open unit interval, and this is in fact a condition for guaranteeing a proper covariance.

The specification given in Equation 4.12 induces a multivariate version of the conditional distributions that is analogous to the univariate Leroux likelihood,

$$f(\phi_i | \phi_{j \neq i}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha, \rho) = \text{MVN} \left(\frac{\rho \sum_{j=1}^m w_{ij}(\alpha) \phi_j + (1 - \rho) \boldsymbol{\delta}}{\rho \sum_{j=1}^m w_{ij}(\alpha) + (1 - \rho)}, \frac{\boldsymbol{\Sigma}}{\rho \sum_{j=1}^m w_{ij}(\alpha) + (1 - \rho)} \right)$$

where $\boldsymbol{\mu}_i = \boldsymbol{\delta}$, so that the mean is constant. The distribution for the joint specification can be established by noting that $\boldsymbol{\Gamma}^{-1}(\mathbf{I}_{mp} - \tilde{\mathbf{B}}) = \mathbf{Q}(\alpha, \rho) \otimes \boldsymbol{\Sigma}^{-1}$, with $\mathbf{Q}(\alpha, \rho) = \rho \mathbf{W}^*(\alpha) + (1 - \rho) \mathbf{I}_m$. The matrix $\mathbf{W}^*(\alpha)$ has diagonal elements $w_{ii}^*(\alpha) = \sum_{j=1}^m w_{ij}(\alpha)$ and off-diagonal elements $w_{ij}^*(\alpha) = -w_{ij}(\alpha)$. Finally, this form of the adjacencies yields the joint distribution, $f(\boldsymbol{\phi} | \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha, \rho) \sim \text{MVN}(\mathbf{1}_m \otimes \boldsymbol{\delta}, \mathbf{Q}(\alpha, \rho)^{-1} \otimes \boldsymbol{\Sigma})$. We call a random variable $\boldsymbol{\phi}$ with this distribution, $\boldsymbol{\phi} \sim \text{MCAR}(\alpha, \rho, \boldsymbol{\delta}, \boldsymbol{\Sigma})$.

4.4.3 Hyperpriors

To complete the model specification, we need to place priors on the hyperparameters. We define the mean of the MCAR process, $\boldsymbol{\delta} = (\beta_0, \beta_1, \beta_1^*, \lambda_0, \lambda_1, \lambda_1^*, \theta)^T$. The prior for the mean, $\boldsymbol{\delta}$, covariance $\boldsymbol{\Sigma}$, and boundary detection parameter α are as follows,

$$\boldsymbol{\delta} \sim \text{N}(\mathbf{0}, \kappa^2 \mathbf{I}_7), \quad \boldsymbol{\Sigma} \sim \text{Inverse-Wishart}(\xi, \boldsymbol{\Psi}), \quad \alpha \sim \text{Uniform}(a_\alpha, b_\alpha).$$

The mean $\boldsymbol{\delta}$ is zero centered with variance, κ^2 , that is chosen to be large in order to yield a weakly informative prior. For our prior on $\boldsymbol{\Sigma}$, we specify an inverse-Wishart distribution with degrees of freedom $\xi = p + 1$ and scale matrix, $\boldsymbol{\Psi} = \mathbf{I}_p$, where p is the dimension of the spatial parameter, ϕ_i . This prior is appealing since it induces marginally uniform priors on the correlations of $\boldsymbol{\Sigma}$ and allows for the diagonals to be weakly informative (Gelman et al., 2014). The dissimilarity metric parameter, α , is forced to be non-negative with a uniform distribution. The bounds are defined as $a_\alpha = 0$ and

$b_\alpha = -\log\{0.5\}/\min_{i,j}\{z_{ij}\}$. As in Lee and Mitchell (2011), the upper limit is set so the smallest dissimilarity metric between two locations can still obtain a spatial adjacency of 0.5.

We fix ρ at 0.99 to force spatial dependency, in which there is precedent from Chapter 2 and Lee and Mitchell (2011). In particular, Section A.4 demonstrates that ρ fixed at 0.99 does not impact estimation of α . Therefore, multivariate spatial structure is introduced through, $\phi_i = \{\beta_0(\mathbf{s}_i), \beta_1(\mathbf{s}_i), \beta_1^*(\mathbf{s}_i), \lambda_0(\mathbf{s}_i), \lambda_1(\mathbf{s}_i), \lambda_1^*(\mathbf{s}_i), \eta(\mathbf{s}_i)\}^T$, which are assigned an MCAR(α, δ, Σ) process. With $\rho = 0.99$ the joint distribution for the spatial effects is defined as,

$$\phi \sim \text{MVN}\left(\mathbf{1}_m \otimes \delta, \mathbf{Q}(\alpha)^{-1} \otimes \Sigma\right). \quad (4.13)$$

4.5 Assessing Visual Field Progression Using Change Points

4.5.1 Study Data

In this study, we source data from the Vein Pulsation Study Trial in Glaucoma and the Lions Eye Institute trial registry, Perth, Western Australia. The dataset contains 1,448 VFs from 194 distinct eyes (98 patients in total). All of the subjects have some form of primary OAG. The mean follow-up time for participants is 934 days (2.5 years) with an average of 7.4 tests per subject.

In Figure 4.13, the time series of VF observations is presented for a sample patient. Figure 4.13 provides a motivating example for the CP framework. The majority of past modeling attempts have assumed a linear relationship between sensitivities and time. However, from the patient in Figure 4.13 it is clear that a CP may provide an improved description of the trajectory of VF sensitivities over time, and in particular for areas of decreased sensitivities. Furthermore, the figure illuminates the inverse relationship between VF sensitivities and testing variability.

4.5.2 Change Point Model for Visual Field Data

Using notation developed in Section 4.4, we define $Y_t^*(\mathbf{s})$ as the VF observation at location \mathbf{s} and x_t days from baseline visit ($x_1 = 0$), for $\{\mathbf{s}_i : i = 1, \dots, 52\}, t = 1, \dots, \nu$, where the number of visits, ν , is patient specific. The number of locations, 52, represents the locations on the VF removing the two corresponding to the blind spot. The data generating mechanism for VF data is such that observations are censored at zero. Therefore, we use a Tobit model (Tobin, 1958), in which there is

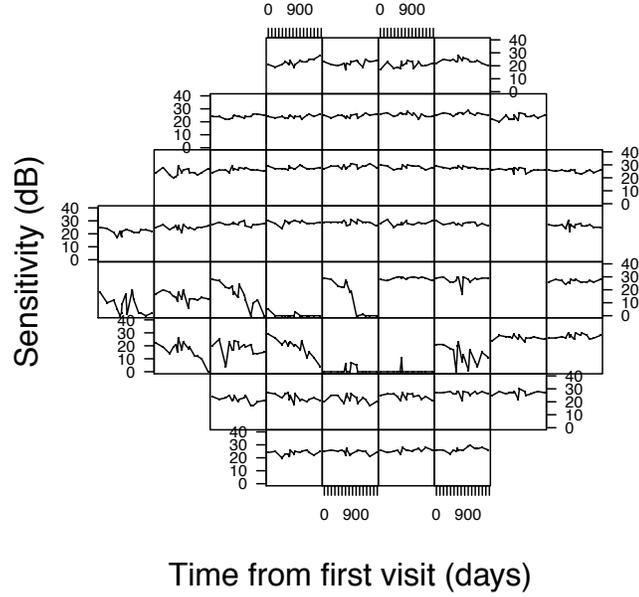


Figure 4.13: Example VF follow-up for a patient with glaucoma. At each location, the time series of the observed threshold sensitivities is plotted. The trends and clustering of the time series motivate the use of spatially varying CPs.

precedent in the glaucoma progression literature, including Chapter 2 (Betz-Stablein et al., 2013; Bryan et al., 2015). To induce the Tobit model, we define a latent process through the standard Tobit link, $Y_t^*(\mathbf{s}) = \max\{0, Y_t(\mathbf{s})\}$.

The latent process can then be modeled using the framework developed in Section 4.4. To appropriately model the course of glaucoma progression, we will slightly modify the mean and variance structures. Patients with glaucoma are typically stable until the point of progression when their VF time series begin to deteriorate. As such, for VF data we define the mean process,

$$\mu_t(\mathbf{s}) = \begin{cases} \beta_0(\mathbf{s}) & \text{if } 0 \leq x_t \leq \theta(\mathbf{s}), \\ \beta_0(\mathbf{s}) + \beta_1(\mathbf{s}) \{x_t - \theta(\mathbf{s})\} & \text{if } \theta(\mathbf{s}) \leq x_t \leq x_\nu, \end{cases} \quad (4.14)$$

so that before the CP there is a constant intercept at each location and after a potential deterioration. This form is a subset of Equation 4.11, with $\beta_1(\mathbf{s}) = 0$ and then $\beta_1^*(\mathbf{s}) \equiv \beta_1(\mathbf{s})$ redefined. Furthermore, since the time of visits is the same for all locations, $x_t(\mathbf{s}) = x_t$. The same specification is used

for the variance process, so that the variance can increase with the onset of deterioration,

$$\log\{\sigma_t(\mathbf{s})\} = \begin{cases} \lambda_0(\mathbf{s}) & \text{if } 0 \leq x_t \leq \theta(\mathbf{s}), \\ \lambda_0(\mathbf{s}) + \lambda_1(\mathbf{s}) \{x_t - \theta(\mathbf{s})\} & \text{if } \theta(\mathbf{s}) \leq x_t \leq x_\nu. \end{cases}$$

Spatial dependency is introduced using Section 4.4.2 theory, $\phi \sim \text{MCAR}(\alpha, \delta, \Sigma)$, where $\phi_i = \{\beta_0(\mathbf{s}_i), \beta_1(\mathbf{s}_i), \lambda_0(\mathbf{s}_i), \lambda_1(\mathbf{s}_i), \eta(\mathbf{s}_i)\}^T$. To finalize the model, we specify hyperpriors based on details in Section 4.4.3, where $\kappa^2 = 1000$ and $p = 5$. We define the dissimilarity metric, z_i , as the Garway-Heath angle at location i , so that the upper bound for α can be determined. This completes the model specification.

4.5.3 Model Estimation

For complete details on the implementation of the model, see Appendix B. We draw samples from the joint posterior using an MCMC sampler (Metropolis et al., 1953; Geman and Geman, 1984; Gelfand and Smith, 1990). The Tobit model specification is particularly amenable to MCMC, since the latent process has a closed form full conditional distribution. Many of the parameters have conjugate full conditionals. In particular, all of the parameters have closed form solutions, with the exception of α , $\lambda_0(\mathbf{s})$, $\lambda_1(\mathbf{s})$, and $\eta(\mathbf{s})$. A Gibbs sampler is used with Metropolis steps for parameters lacking conjugacy. For each eye, the MCMC sampler is run for 250,000 iterations after a 10,000 burn-in and thinned to a final sample size of 10,000.

To guarantee that Metropolis acceptance rates are within an acceptable range, pilot adaptation is used (Banerjee et al., 2003). Convergence is assessed through examination of traceplots and the Geweke diagnostic statistic (Geweke, 1992), with no signs of non-convergence observed for any of the 191 VF series. The observed VF data were scaled by 10 along with the visit days being scaled by 365 to improve the stability of the parameter estimation. Finally, we scale our dissimilarity metric by 100 to improve MCMC mixing.

4.5.4 Model Comparison

In order to establish the capability of the introduced spatially varying CP model to both predict and diagnose glaucoma progression, we compare it to various simpler models. In this analysis, we present five models defined as follows,

1. *Piece-wise linear regression (PLR)*: Tobit linear regression fit at each location.
2. *Non-spatial CP (D)*: Tobit linear regression with a discrete CP, based on Equation 4.8. Each of the following CP models specify a constant process before the CP, as specified in Equation 4.14.
3. *Non-spatial CP (C)*: Tobit linear regression with a continuous CP, where the CPs at each location are independent and uniformly distributed.
4. *Non-spatial CP (L)*: Tobit linear regression with a continuous CP with an underlying normally distributed latent process, as specified in Section 4.4.1.
5. *Spatial CP*: Spatial CP model from Section 4.5.2 using the MCAR distribution introduced in Section 4.4.2.

Each of the comparison models utilize the Tobit likelihood to illuminate other differences in the model. The motivation of fitting the simpler models, such as PLR and basic CP models, is to have a benchmark for comparison and emphasize the importance of the latent CP specification and spatial structure. Note that each of the models are subsets of Model 5, with the exception of the model with discrete CPs.

The models are compared using both model fit diagnostics and hold-out prediction. To assess model fit, the Deviance Information Criterion (DIC) is used. DIC is based on the deviance statistic and penalizes for the complexity of a model with an effective number of parameters estimate, p_D (Spiegelhalter et al., 2002). DIC is preferred for comparing explanatory performance and a smaller value indicates improved fit (Banerjee et al., 2003). For prediction, the final visit for each of the VF series is held out, and the remaining visits are used to predict the hold-out visit. Prediction accuracy is measured using both the mean squared prediction error (MSPE) and continuous ranked probability score (CRPS). MSPE measures the expected squared distance between the predictor and truth, while

CRPS compares the empirical cumulative distribution function for the prediction to the truth. CRPS is an appropriate measure for evaluating prediction in the Bayesian framework, since it accounts for the full distribution (Hersbach, 2000). Small values of DIC, MSPE, and CRPS are preferred.

Table 4.9: Model fit and prediction diagnostics for model comparison. Model fit is determined using the deviance information criterion (DIC) and effective number of parameters, p_D . Prediction capability is determined using the mean squared prediction error (MSPE) and continuous ranked probability score (CRPS). The ability of each model to diagnose progression is assessed using a metric dependent on the latent CPs. For each model with a latent CP process, the metric $\max_i \{P[\theta(s_i) < t_\nu]\}$ is calculated and regressed against the clinically determined progression status. Diagnostic capability is measured using AIC, area under the curve (AUC), and the p-value for each measure.

Model	Model Fit		Prediction		Diagnostic		
	DIC	p_D	MSPE	CRPS	AIC	AUC	p-value
PLR	31.11	120.30	0.20	0.21	—	—	—
Non-spatial CP (D)	-14.27	119.30	0.45	0.28	—	—	—
Non-spatial CP (C)	-18.76	112.13	0.32	0.28	—	—	—
Non-spatial CP (L)	-29.05	67.69	0.23	0.40	211.36	0.64	0.001
Spatial CP	-53.74	68.21	0.13	0.17	208.67	0.69	<0.001

The model fit and prediction diagnostics can be found in Table 4.9. The DIC value for the Spatial CP is superior to the rest of the models, with a minimal DIC of -53.74. In particular, it can be observed that there is a monotone decreasing trend as the models become more complex. This suggests that each of the models incrementally improve upon the prior model, indicating that the continuous CP with a latent process and spatial structure are beneficial. The number of effective parameters also trends smaller with added model complexity, indicating that the added model structure provides a more concise framework for modeling VF data. The prediction measures also favor the Spatial CP model, with optimal MSPE and CRPS of 0.13 and 0.17, respectively. Both MSPE and CRPS for the three non-spatial CP models are inferior to the PLR model, indicating that the inclusion of a CP does not guarantee improved prediction. Only when accounting for the spatial nature of VF data does a CP model offer improved prediction over the PLR model, again illuminating the importance of modeling spatial variability. Overall the results in Table 4.9 indicate that the Spatial CP has superior model fit and prediction ability for VF data.

4.5.5 Diagnosing Progression Using Change Points

To illustrate the superiority of the introduced methodology, in addition to model fit and prediction, we demonstrate the diagnostic capability of the CPs. As hypothesized in the introduction, the

location of the CPs can act as a proxy for disease progression. Based on this theory, we propose a diagnostic metric that is a function of the probability that a CP has occurred up to the last VF visit, $x_i = P[\theta(s_i) < t_\nu]$. Then, we define a simple metric as the maximum of all these probabilities, $\max_i\{x_i\}$. This is an intuitive metric, because any longitudinal series of VFs with a large CP is more likely to be progressing.

In order to assess the diagnostic ability of the CPs, we construct logistic regression models, regressing the defined metric on the clinical assessment of progression and present AIC, AUC and a p-value for the slope in the logistic regression. These measures are presented in Table 4.9 for models that have a latent CP process (the only models capable of estimating the CP probability). From Table 4.9, it is clear that the CPs arising from the Spatial CP model are highly associated with progression (p-value: < 0.001). The non-spatial model also yields a significant slope estimate, however the p-values are of a different magnitude (p-value: 0.001). Furthermore, the values of AIC and AUC are optimized under the Spatial CP model with minimal AIC, 208.67, and maximal AUC, 0.69. These results help validate the claim that the CPs in some way represent disease progression and can aid in clinical diagnoses.

To further illuminate the diagnostic ability granted by the latent CP framework, we present an example patient as a proof of practice. This patient is appropriate, because they have late occurring CPs that are truncated when removing visits near the end of follow-up. The patient has nine VF visits, however we only use the initial seven to predict the CPs that were estimated using the full data. The patient's VF data is presented in Figure 4.14, where at each location the first seven VF sensitivities are black dots and the last two are red. Furthermore, the mean CPs estimated by the Spatial CP model for the full data are presented as dashed vertical lines. Any CP that was estimated to have occurred after the seventh visit is colored red and are technically truncated for purposes of clinical use at visit seven. Finally, the posterior CP density estimates (based on the reduced data) for the Spatial CP, and Non-spatial CP (L) models are presented as density estimates in blue, and green, respectively.

From Figure 4.14 it is clear that the Spatial CP model is capable of predicting future CPs that have not been observed. This is remarkable, since at the seventh visit there was no clear sign of a CP at many of these locations. It is even more remarkable to compare the performance of the Spatial CP model with the non-spatial model. The Non-spatial CP (L) model performs poorly, indicating the

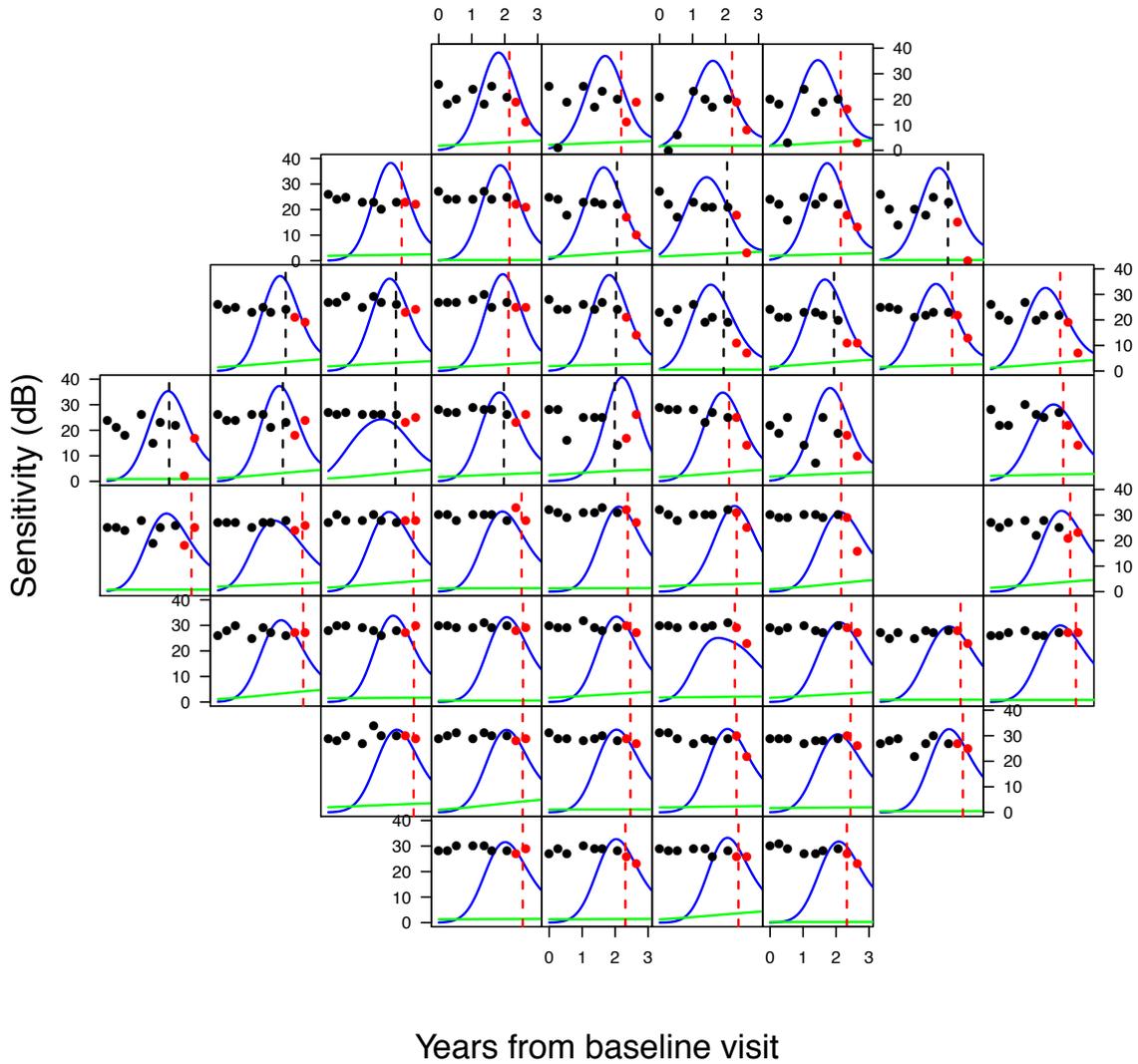


Figure 4.14: Demonstrating the ability of the Spatial CP model to identify CPs that have yet to occur. At each cell, the observed DLS are presented for an example patient with nine visits. Based on only the first seven visits, the posterior estimates of the CPs for the Spatial CP and Non-spatial (L) models are presented as density estimates in blue, and green lines, respectively. The estimated CP based on the Spatial CP model using all of the nine VF visits are presented as dashed vertical lines. The DLS and CPs are shown in red if they occur after the seventh visit and thus are not observable at the time the posterior densities were estimated.

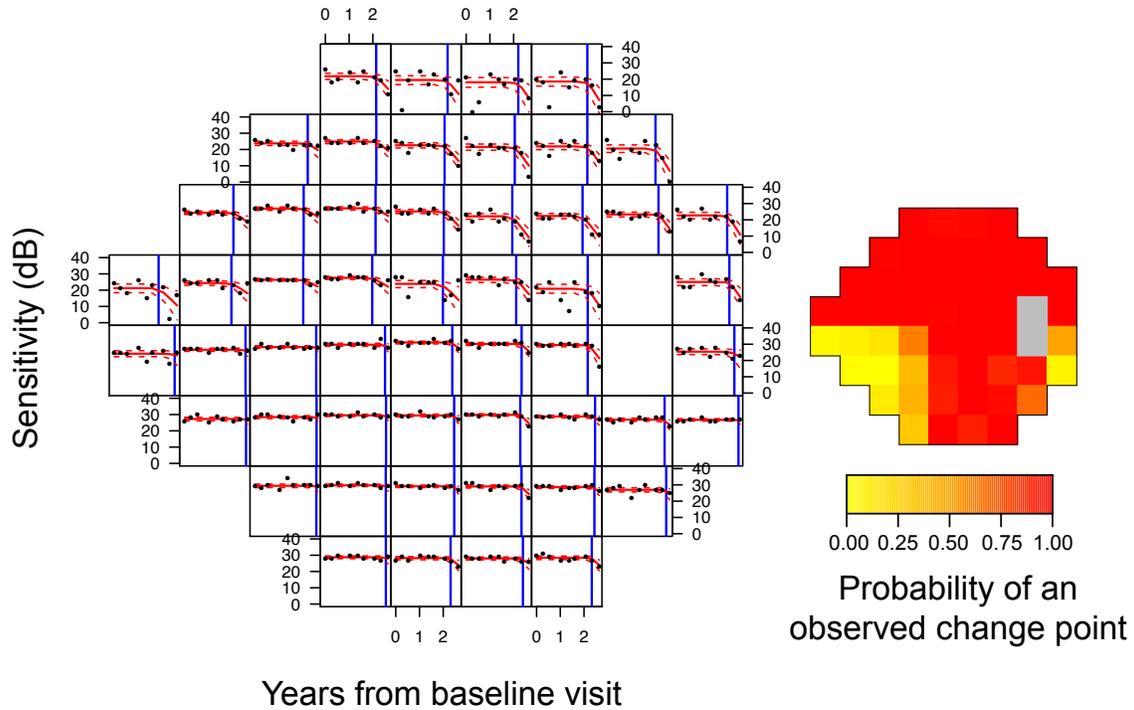


Figure 4.15: Output generated from the Spatial CP model. The plot on the left presents the estimated mean (with 95% credible intervals) and CPs for the same example patient in Figure 4.14. The mean process is presented in red and the change points are vertical blue lines. The panel on the right presents the posterior probability of an observed CP at each location on the VF.

importance of spatial structure. To quantify the conclusions drawn from Figure 4.14, for each model we present diagnostics for prediction of the estimated CPs. Each model is compared to the “true” CPs estimated from the same model. The Spatial CP and Non-spatial (L) models have MSPE/CRPS of 4.70/1.05 and 8,525.82/23.66, respectively. Clearly, the Spatial CP model has superior prediction of the unobserved CPs, with minimal values of both prediction criteria. These powerful results indicate that the introduced methodology has the capability of predicting future unobserved CPs, and therefore progression. This is possible due to both the latent CP specification and multivariate spatial structure using the Garway-Heath angles. As we have seen from the results in Table 4.9 and Figure 4.14, without either of these specifications the model suffers in terms of model fit, prediction and diagnostic capability.

Having established the potential for using the Spatial CP model in the clinical setting for prediction of VF sensitivities and CPs, we now present example output for use clinically. In Figure 4.15, on the left, the fitted trend is presented on the VF with the estimated posterior CPs. The plot

presents the estimated mean (with 95% credible intervals) and CPs for the same example patient in Figure 4.14. The mean process is presented in red and the CPs are vertical blue lines. This figure demonstrates the flexibility of the Spatial CP to produce non-linear trends, not two lines connected at a point. This type of trajectory is a result of the CPs being random variables. In addition to the predicted fit, it is useful to have a heatmap of the estimated CPs having occurred in the observed data (right of Figure 4.15). This presentation corresponds to the estimated CPs in the left frame, but further provides a useful map for clinicians to understand the location and severity of progression. Both of these plot in conjunction with the diagnostic potential from the Spatial CP model provide a useful framework for clinicians to assess progression. The combination of the latent CP process and novel spatial structure has produced a technique that goes beyond the simple CP models and PLR in terms of model fit, prediction, and progression diagnoses. To better understand these improvements the models are studied through simulation in Section 4.6.

4.6 Simulation Study

4.6.1 Data Generation

In order to assess the performance of the proposed model, a simulation study is designed that utilizes various data generating settings. The purpose of the simulation is to explore the performance of each of the models detailed in Section 4.5.4 to different types of VFs. The settings explored in this simulation are, 1) progressing (all CPs before the beginning of follow-up), 2) stable (all CPs after the end of follow-up), 3) non-spatial CPs, 4) spatial CPs with no covariance, and 5) full spatial CP model. The simulation settings are all subsets of the spatial covariance structure of the full spatial CP model given in Equation 4.13, $MVN(\mathbf{1}_m \otimes \boldsymbol{\delta}, \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma})$. For the full model (setting 5) these parameters are fixed as follows,

$$\boldsymbol{\delta} = \begin{bmatrix} 25 \\ -15 \\ -0.5 \\ 0.1 \\ 0.5 \end{bmatrix}, \quad \boldsymbol{\Sigma} = \begin{bmatrix} 0.025 & -0.500 & -0.500 & -0.500 & 0.500 \\ -0.500 & 0.025 & 0.500 & 0.500 & -0.500 \\ -0.500 & 0.500 & 0.025 & 0.250 & -0.500 \\ -0.500 & 0.500 & 0.250 & 0.025 & -0.500 \\ 0.500 & -0.500 & -0.500 & -0.500 & 0.025 \end{bmatrix}, \quad \alpha = 0.1. \quad (4.15)$$

Settings 1-4 are obtained by changing parts of $\boldsymbol{\delta}$, $\boldsymbol{\Sigma}$, and α . To obtain settings 1 and 2, α is fixed at an arbitrarily large value, 1000, and $\boldsymbol{\Sigma}$ is restricted to being diagonal, with the fourth entry zero. This specification yields the PLR model as long as the CPs are not observed. As such, setting 1 is obtained by fixing the fifth entry in $\boldsymbol{\delta}$ to an arbitrarily large negative value, -10, while a large positive value, 10, is used for setting 2. To guarantee no CP in the variance, the fourth entry in $\boldsymbol{\delta}$ is fixed at zero. The third simulation setting introduces a CP in the variance process by using the specification of $\boldsymbol{\delta}$ from Equation 4.15 and the fully diagonal $\boldsymbol{\Sigma}$. The value of α is still 1000, so that there is no spatial dependence. Finally, setting 4 is obtained by setting $\boldsymbol{\Sigma}$ to be diagonal and represents a spatial CP model with no covariance between spatial effects.

For each setting, we simulate data based on a set of known hyperparameters and then fit each model to assess model fit, prediction and estimation of CPs, both latent and observed. As in Section 4.5 for model fit, we utilize DIC and p_D and for prediction, CRPS and MSPE. The framework for assessing prediction is innately linked to how the data were simulated. Each dataset was simulated assuming there were 21 VF visits (the maximum number of visits in our study data). Then, the actual simulation was only conducted with the first 14 visits, and the 16th and final visit are used for prediction. The visit times were specified as the sequence of numbers between 0 and 1 by a 0.05 increment. Finally, estimation of both the observed CPs and latent CP process are assessed using bias, MSE, and EC. The EC is the proportion of the time that the estimated Bayesian credible intervals contain the true value. We define the nominal coverage as 95%.

4.6.2 Results

A new dataset is generated for each simulated value of ϕ to ensure that the results are not affected by a particular realization of ϕ . The results are displayed in Tables 4.10, 4.11, and 4.12 and relate to 250 simulated data sets for each simulation setting. In Table 4.10, the model fit diagnostics, DIC and p_D , are presented for the five models across all settings. As expected, the models all have generally similar DIC values for the first two simulation settings, which are just progressing and stable representations of PLR. As the models increase in model complexity the DIC value corresponding to the Spatial CP model separates from the non-spatial CP models and PLR. For example, in setting 5 the DIC for Spatial CP is -351.22, while the nearest competitor only reaches -318.36. This indicates that with increased data complexity, as we see in VF data, the Spatial CP becomes more important for model fit. The effective number of parameters, p_D , is similar for all models in setting 1, but is generally smaller for the Spatial CP and Non-spatial CP (L) models for the remaining settings. It is encouraging that the Spatial CP model has a smaller p_D in the more complex settings, indicating that the increased model structure maintains parsimony even in the presence of complex data.

Table 4.10: Model fit diagnostics for simulation study outlined in Section 4.6.1. The diagnostics included are DIC and p_D and are presented for each of the models introduced in Section 4.5.4. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets.

Metric	Model	Simulation Setting				
		1	2	3	4	5
DIC	PLR	-466.33	-409.93	-360.16	-401.16	-276.76
	Non-spatial CP (D)	-457.89	-434.85	-426.27	-455.50	-293.29
	Non-spatial CP (C)	-412.21	-442.09	-414.03	-448.02	-293.54
	Non-spatial CP (L)	-473.37	-476.06	-446.44	-446.77	-318.36
	Spatial CP	-479.61	-455.21	-457.57	-468.10	-351.22
p_D	PLR	118.23	118.39	118.59	121.08	118.72
	Non-spatial CP (D)	111.19	118.51	117.44	134.85	125.36
	Non-spatial CP (C)	133.70	105.29	120.37	135.56	115.00
	Non-spatial CP (L)	109.34	69.65	87.82	78.27	77.70
	Spatial CP	118.80	92.41	110.22	94.71	90.56

Next, we studied the capability of each model to predict VF sensitivities into the future. In this simulation, data was simulated across the time interval [0,1] by 0.05. This yields 21 visits, but only

the first 14 values were used, implying a final visit of $x_\nu = 0.65$. Based on this data generation, prediction ability was assessed for each model by predicting the simulated VF at future times 0.75 and 1 (i.e., visits 16 and 21). The prediction results can be found in Table 4.11, where both CRPS and MSPE are presented. For setting 1, both CRPS and MSPE show similar values for all models and at

Table 4.11: Prediction diagnostics for estimation of the VF visits at times 0.75 and 1. The diagnostics included are CRPS and MSPE and are presented for each of the models introduced in Section 4.5.4. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets.

Time	Metric	Model	Simulation Setting				
			1	2	3	4	5
0.75	CRPS	PLR	12.77	22.68	18.38	18.48	21.30
		Non-spatial CP (D)	12.77	22.48	18.34	18.55	21.31
		Non-spatial CP (C)	12.77	22.67	18.43	18.53	21.29
		Non-spatial CP (L)	12.72	22.64	18.35	18.43	21.21
		Spatial CP	11.80	22.75	18.05	18.20	21.13
	MSPE	PLR	215.14	577.70	450.40	358.81	531.48
		Non-spatial CP (D)	213.56	572.25	447.12	364.53	532.32
		Non-spatial CP (C)	215.35	577.26	451.93	362.37	531.95
		Non-spatial CP (L)	213.62	577.54	451.17	359.04	530.65
		Spatial CP	187.63	577.25	439.73	349.58	524.76
1	CRPS	PLR	9.44	22.58	17.09	15.10	20.72
		Non-spatial CP (D)	9.34	21.83	16.71	15.13	20.47
		Non-spatial CP (C)	9.41	22.41	17.01	15.13	20.59
		Non-spatial CP (L)	9.33	22.22	16.87	14.80	20.29
		Spatial CP	8.20	22.71	16.70	14.78	20.39
	MSPE	PLR	131.81	573.39	451.89	248.02	541.26
		Non-spatial CP (D)	129.24	558.73	443.05	258.77	539.04
		Non-spatial CP (C)	132.91	573.20	455.12	254.08	542.85
		Non-spatial CP (L)	131.47	574.85	452.77	247.87	541.05
		Spatial CP	109.62	575.05	443.39	239.76	527.46

both times, except for the Spatial CP model which is superior. This is surprising, since this setting should be equally accessible for each model, but possibly indicates that the Spatial CP model is preferred even in the absence of clear spatial structure. In setting 2, all of the models have comparable performance, while in settings 3 and 4 the Spatial CP model generally shows small improvements. Finally, in the most complex setting, the Spatial CP model has superior prediction at time 0.65 and this is confirmed at time 1, with a MSPE of 527.46 compared to 539.04 of the next best model. Interestingly, the CRPS for the Non-spatial CP (L) model is smaller than the Spatial CP model at

time 1, indicating the flexibility of the latent CP specification, even without spatial structure. Overall, however, it is clear the Spatial CP model has superior prediction capability over the other models.

We finish this simulation by studying the estimation ability of each model to identify both the observed, $\theta(s)$, and latent, $\eta(s)$, CPs. In Table 4.12, we present bias, MSE, and EC for models that are capable of estimating either the observed or latent CP. Note that the Non-spatial CP (D)

Table 4.12: Summary diagnostics for estimation of the observed, $\theta(s)$, and latent, $\eta(s)$, CPs. The diagnostics included are bias, MSE, and EC (95% nominal) and are presented for each of the models introduced in Section 4.5.4 that are capable of estimating either the observed or latent CP. The data generating settings for the simulation increase in model complexity and are as follows, 1) progressing, 2) stable, 3) non-spatial CP, 4) spatial CP with no covariance, and 5) full spatial CP model. Each reported estimate is based on 250 simulated datasets.

Estimand	Metric	Model	Simulation Setting				
			1	2	3	4	5
$\theta(s)$	Bias	Non-spatial CP (C)	-0.11	0.19	0.07	0.07	0.09
		Non-spatial CP (L)	-0.00	0.06	0.01	0.04	0.01
		Spatial CP	-0.00	0.00	0.00	0.01	0.00
	MSE	Non-spatial CP (C)	0.02	0.07	0.05	0.03	0.05
		Non-spatial CP (L)	0.00	0.04	0.03	0.04	0.04
		Spatial CP	0.00	0.00	0.00	0.01	0.00
	EC	Non-spatial CP (C)	0.00	0.00	0.16	0.58	0.16
		Non-spatial CP (L)	1.00	1.00	0.98	0.90	0.98
		Spatial CP	1.00	1.00	0.98	0.97	0.99
$\eta(s)$	Bias	Non-spatial CP (L)	16.00	-17.69	-3.82	-11.18	-18.99
		Spatial CP	-1.59	1.01	-0.00	-0.01	0.01
	MSE	Non-spatial CP (L)	4088.01	5399.53	4615.01	4453.89	5531.15
		Spatial CP	33.33	26.85	1.90	0.04	0.56
	EC	Non-spatial CP (L)	0.99	0.99	0.90	0.84	0.96
		Spatial CP	0.85	0.84	0.90	0.91	0.95

model was not included in the observed CP section, since it is estimating a different, discrete, CP. The results for estimating the observed CP demonstrate the importance of the latent CP process for properly modeling the true CP. The most important take away from these results are that the Non-spatial CP (C) model was incapable of estimating the observed CP, with inferior bias, MSE and EC. In particular, the EC is notably poor, with EC of zero for the first two settings. This is clearly due to the lack of a latent CP specification, as the two models with a latent CP specification have improved EC. The high EC values for these models are a result of the observed CP being a censored object in their model specifications. Finally, it is important to note that the Spatial CP model has superior bias and MSE in estimating the observed CP, indicating that in addition to the latent CP,

spatial structure is also crucial. This is further validated when looking at the estimation of the latent CPs.

In the simulation analysis for estimation of the latent CP, only the Spatial CP and Non-spatial CP (L) models are compared. It is immediately clear that the inclusion of spatial structure is crucial for estimating the latent CP. In all simulation settings the bias and MSE are improved in the Spatial CP model. The bias and MSE are larger in the first two settings, because there are no observed CPs in the generated data. The EC values for the Spatial CP model increase to the nominal coverage of 95%, while the Non-spatial CP (L) model has EC near one for the first two settings. This is most likely a result of the huge variation seen in the MSE, while the Spatial CP model underestimates the nominal coverage, presumably a result of the smoothing applied by the spatial structure. From Table 4.12, it is clear that the components of the Spatial CP model are needed for estimating both the latent and observed CPs. Having the ability to estimate the latent CP in the context of glaucoma progression is clinically important, because it represents the point of structural change in the progression follow-up.

4.6.3 Estimation of the Observed Change Point

In Section 4.5, we established the importance of estimating the observed CPs for identifying the moment of progression. Then in Section 4.6.2, we established in simulation the importance of the latent CP specification and spatial structure for estimating both the observed and latent CPs. As a result, it is important to better understand the behavior of each modeling feature in estimating CPs, especially at the censored times. To gain a better understanding, we develop a second simulation that is based on a fixed ϕ , so that the CPs are fixed for all simulations. The realization of ϕ was chosen to produce CPs that occur across all stages of follow-up. To generate the ϕ , we use the same data generating scheme as Section 4.6.1, with the same values of Σ and α , with $\delta = (25, -30, 1, 0.5, 0.5)$. Based on the value of ϕ , 1000 datasets are simulated. In this simulation, the VF visit times are defined as all 21 visits from Section 4.6.1.

Each of the models were fit to the simulated datasets yielding an estimate of the observed CP. In Figure 4.16, the density estimates of the distribution of estimated observed CPs are presented for each model across the VF (minus PLR). Furthermore, the true observed CP is presented as a vertical black dotted line and the threshold sensitivity values for an example simulated dataset are presented as gray dots. This figure is illuminative for many reasons. First, in locations where the true CP is near

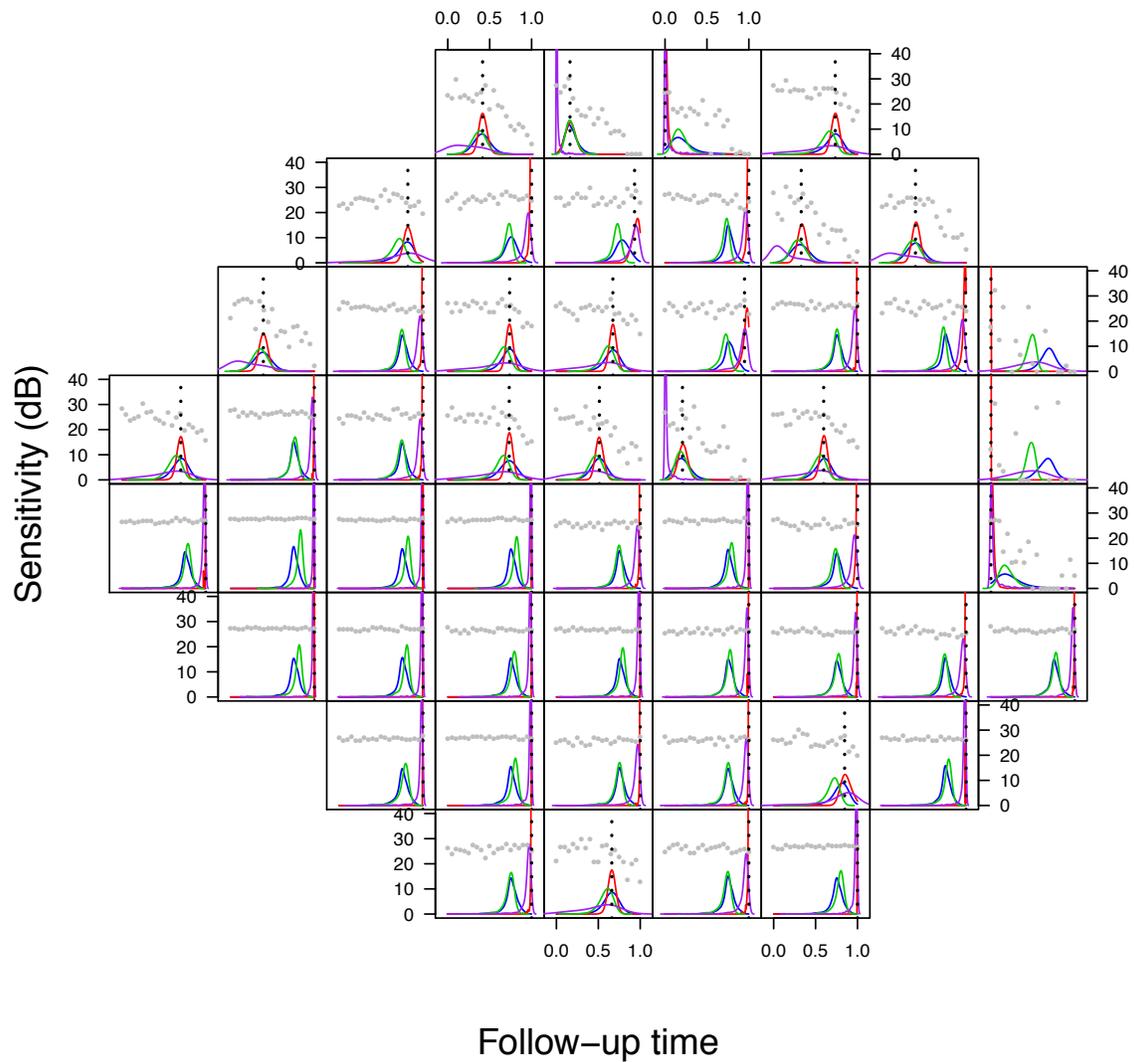


Figure 4.16: Simulation study comparing the performance of the spatial CP and non-spatial CP models in estimating the CPs across the VF. The true value of the CP is presented as a dotted black line and the threshold sensitivity values for an example simulated dataset are presented as gray dots. The posterior estimates of the CPs for the Spatial CP, Non-spatial CP (D), Non-spatial CP (C), and Non-spatial CP (L) models are presented as density estimates in red, blue, green, and purple lines, respectively.

the middle of the follow-up time (i.e., easy to identify), the Spatial CP model is unbiased and has the narrowest bounds. In these settings, the other models generally have small bias, however the density estimates are much wider. Then, in locations where the true CP is near or beyond the beginning or end of follow-up the Spatial CP is clearly superior. The two models that have no latent CP process are incapable of approaching the edges of the follow-up times, with apparent bias. The Non-spatial CP (L) model provides an improvement, allowing the density estimate to approach the bounds. However, it is important to note that this appears to be a trade-off with wider densities in locations with a true CP in the middle of follow-up. The only model that is capable of producing acceptable CP estimation in the middle and bounds of follow-up is the Spatial CP model. Interestingly, at certain locations at the bounds the Spatial CP model is capable of producing a point mass.

Table 4.13: Summary diagnostics for estimation of the observed CPs in the simulation detailed in Section 4.6.3. The diagnostics included are bias, MSE, and EC (95% nominal) and are presented for each of the models that are capable of estimating the continuous observed CP. Standard errors (SE) correspond to the variation across all VF locations. Each reported estimate is based on 1,000 simulated datasets.

Metric	Model	Estimate (SE)
Bias	Non-spatial CP (C)	0.147 (0.171)
	Non-spatial CP (L)	0.058 (0.123)
	Spatial CP	-0.001 (0.008)
MSE	Non-spatial CP (C)	0.102 (0.065)
	Non-spatial CP (L)	0.083 (0.087)
	Spatial CP	0.002 (0.002)
EC	Non-spatial CP (C)	0.33 (0.461)
	Non-spatial CP (L)	0.94 (0.146)
	Spatial CP	0.99 (0.018)

To formalize the results seen in Figure 4.16, summary diagnostics are presented in Table 4.13. The results for the non-spatial CP model with a discrete CP are excluded, since it is estimating a discrete CP. The results in Table 4.13 confirm the trends observed in Figure 4.16 as the bias and MSE are improved with the introduction of a latent process and then spatial structure. The Spatial CP has superior bias and MSE, -0.001 and 0.002, respectively, and has the smallest standard errors, indicating less variability in estimation across the simulated datasets. The EC is only acceptable with the inclusion of the latent CP process. The Non-spatial CP (C) model has an EC of 0.33, which is clearly a result of all of the true CPs near the bounds. The Spatial CP model slightly overestimates the nominal coverage, which corresponds to the prior simulation and makes sense since the observed CP

is a censored parameter. Overall, this additional simulation provides a closer look into the estimation of the observed CP, demonstrating the behavior of each model at various locations over follow-up. These results again illuminate the importance of the Spatial CP model for estimating the observed CP, especially in models with CPs near the bounds of follow-up, which are commonly the interesting patients.

4.7 Discussion

In this chapter, we present a unified framework that brings together multiple vital aspects of glaucoma management, i) prediction of future VF sensitivities, ii) predicting the timing and spatial location of future vision loss, iii) making clinical decisions regarding progression, and, iv) incorporation of anatomical information to create plausible data-generating models. Previous statistical work has focused on one or some of these aspects typically, but nobody has done all simultaneously, in particular the latent CP process. Although motivated by VF data, the methodology was introduced in a general manner that permits the model to be applied in broad areal data settings where a CP is appropriate. The methodology is built upon theory in the CP literature, utilizing a Bayesian hierarchical modeling framework for inference.

We extended previous attempts at modeling CPs across spatial units, that either ignore spatial dependencies (Wagner and Midway, 2014) or do not include a CP in the variance process (Warren et al., 2017). The inclusion of the variance added flexibility and was motivated by the known inverse relationship between decreases in VF sensitivity and variability (Russell et al., 2012). Finally, we extended the univariate spatial boundary detection framework for allowing a dissimilarity metric to dictate the local neighborhood structure on the VF. The novel multivariate extension allows the anatomy of the optic disc to dictate the neighborhood adjacencies for the mean, variance, and CP processes. This specification functions jointly and accounts for known dependencies between the separate processes through a cross-covariance. The combination of these model characteristics leads to a superior modeling technique for longitudinal series of VFs.

The results from applying our method to VF data (Section 4.5) demonstrated the benefit of using our methodology for both prediction of future VF sensitivities and model fit. The Spatial CP model was compared to non-spatial CP models and PLR and was shown to have superior DIC and

prediction diagnostics, CRPS and MSPE (Table 4.9); this was later confirmed in simulation (Tables 4.10 and 4.11). These results indicate the promise of the Spatial CP model clinically, as predicting future VF sensitivities is incredibly important for managing progression (Crabb et al., 1997). These improvements can be attributed to the inclusion of a spatial dependency structure and the flexible trajectories produced by the CPs across time. The need for spatiotemporal treatment of VF data is well established (Bryan et al., 2013), and is demonstrated here with a reduction in prediction errors and clinically meaningful results.

In addition to an increase in prediction precision, the CP framework allows for prediction of the timing and spatial location of future vision loss. This information is accessible due to both the modeling of spatial dependencies and the latent specification of the CP. The benefits of the latent CP are that it provides a biological interpretation of true disease progression and makes statements about VF sensitivity outside of the observed follow-up. The latent CP is modeled using the novel multivariate spatial process that allows the anatomy of the optic disc to dictate the local neighborhood structure. As such, the patterns of future vision loss based on the latent CP are anatomically driven. This can be seen in the estimated probabilities of a CP for an example patient (right of Figure 4.15), which display a clear differentiation in the inferior and superior sectors as defined by Garway-Heath et al. (2000). The practical utility of this latent CP specification is rooted in the biological motivation, but also in that it provides an alternate method for predicting severity on the VF.

The latent CP is clearly an improvement over standard CP specifications that limit the range of the CP to the observed follow-up period. Furthermore, it allows us to predict a CP for a spatial location that currently has none. This was exemplified with an example patient with many observed CPs near the end of follow-up. For this patient, we showed that when removing the last two visits (and thus truncating these CPs), the Spatial CP model was still capable of detecting the presence of an imminent CP. This served as a proof of practice for predicting unobserved CPs and was facilitated by spatial communication of VF locations, as the non-spatial CP model was incapable of detecting these truncated CPs (Figure 4.14). This was further confirmed in simulation, as the Spatial CP model estimated the latent CP with lower bias and efficiency (Table 4.12).

We hypothesized that the CP framework would provide a technique for parameterizing the rate of glaucoma progression intrinsically into a model. Through the implementation of our spatially varying CP model, it is clear that structural change on the VF can be properly characterized by the CP

framework. This contribution is important clinically, since a detectable change in a longitudinal series of VF series can be highly suggestive of progression, even if standard VF metrics may appear well within normal limits (Artes and Chauhan, 2005). We defined a diagnostic metric as the maximum probability across the VF of a CP occurring in the observed follow-up and showed that it is a significant predictor of clinically determined glaucoma progression. This metric provides a method for describing structural change in VF data and the clinical implications demonstrate the need for further exploration of the proposed metric.

This chapter introduced a framework for assessing glaucoma progression from numerous angles and demonstrates the clinical utility of CPs acting as a proxy for disease progression. To validate these findings it will be important to compare the diagnostic efficiency of the CPs to other established metrics in the glaucoma progression literature, in addition to the ones presented here. Furthermore, the estimation of progression can be improved when combining functional (e.g., VF) and structural (e.g., optic disc) testing techniques (Nicolela et al., 2003). As such, a synthesis of these two data sources could provide improvement in estimating the CPs. This would provide stronger clinical evidence of progression as damage to the optic disc could be corroborated with the VF and vice-versa (Artes and Chauhan, 2005).

Finally, this work opens up numerous avenues for future statistical research. Outside of the VF setting it is often of interest to have a trajectory with multiple CPs. This can easily be obtained using the framework introduced in this chapter by adapting the mean (and possibly variance) process to the setting of multiple CPs and then including the additional spatial CPs in the spatial parameter, ϕ_i . Another natural extension to this approach includes generalizing the separable specification of the multivariate spatial process. Currently the MCAR implies that each set of spatial parameters has the same spatial structure, dependent on the unique dissimilarity metric parameter α . In the future, it would be beneficial to allow each set to depend on their own parameter, α_p , which can be accomplished, for example, by using the linear model of coregionalization.

4.8 Supplementary Materials

Supplementary material for this chapter can be found in Appendix B, where the implementation of the model, referenced in Section 4.5.3, is detailed.

APPENDIX A: SUPPLEMENTARY MATERIALS FOR CHAPTER 2

A.1 Introduction

This document provides supplementary material to Chapter 2, ‘Diagnosing Glaucoma Progression with Visual Field Data Using a Spatiotemporal Boundary Detection Method’, and is divided into five sections. Section A.2 describes the implementation of the Bayesian model proposed in the paper using MCMC simulation. Section A.3 details the prediction capability of our model, first introduced in Section 2.5.5. Section A.4 presents an additional simulation study to assess the impact of mis-specifying ρ on our results, while Section A.5 presents sensitivity analyses to various modeling assumptions in the VF analysis of Section 2.6. Finally, Section A.6 includes additional figures from the analysis.

A.2 Implementation of the Model

We draw samples from the joint posterior using a MCMC sampler (Metropolis et al., 1953; Geman and Geman, 1984; Gelfand and Smith, 1990). The Tobit model specification is particularly amenable to MCMC, since the latent process has a closed form full conditional distribution. Many of the parameters have conjugate full conditionals. In particular, all of the hyperparameters have closed form solutions, with the exception of ϕ . None of the observational level parameters display obvious conjugacy. However, due to the separable specification of θ , the full conditional for μ_t , $t = 1 \dots \nu$ can be written in closed form. Due to the variety of full conditionals, a Gibbs sampler is used with Metropolis steps for parameters lacking conjugacy. For each eye, the MCMC sampler is run for 250,000 iterations after a 10,000 burn-in and thinned to a final sample size of 10,000.

To guarantee that Metropolis acceptance rates are within an acceptable range, pilot adaptation is used (Banerjee et al., 2003). Convergence is assessed through examination of traceplots and the Geweke diagnostic statistic (Geweke, 1992), with no signs of non-convergence observed for any of the 191 VF series. The observed VF data were scaled by 10 along with the visit days being scaled by 365 to improve the stability of the parameter estimation. Finally, we scale our dissimilarity metric by 100 to improve MCMC mixing.

We present the MCMC sampler for the general method. The parameters in the model consist of $[\varphi = (\varphi_1^T, \dots, \varphi_\nu^T)^T, \zeta, \theta, \delta, \mathbf{T}, \phi]$, and MCMC posterior sampling proceeds in the following manner. The MCMC sampler is implemented in the R package `womblR`.

1. Sample from the full conditional distribution of φ_{it} :

Begin by sampling φ_{it} from their full conditional distribution for the set $(\{i, t\} : i = 1, \dots, m_t, t = 1, \dots, \nu)$. It is clear that each φ_{it} will only depend on the other observations at time t . The full conditional for φ_{it} can be written as,

$$f(\varphi_{it} | Y_{it}, \zeta, \boldsymbol{\varphi}_{-it}, \mu_t, \tau_t^2, \boldsymbol{\alpha}_t) \propto f(Y_{it} | g^{-1}(\varphi_{it}), \zeta) \times f(\varphi_{it} | \boldsymbol{\varphi}_{-it}, \mu_t, \tau_t^2, \boldsymbol{\alpha}_t).$$

In general, the full conditional distribution has no closed form, requiring the Metropolis algorithm to be used for sampling. However, for VF data we use the Tobit link, $Y_{it} = \max\{0, \varphi_{it}\}$, which yields the following closed form full conditional density,

$$\begin{aligned} f(\varphi_{it} | Y_{it}, \boldsymbol{\varphi}_{-it}, \mu_t, \tau_t^2, \boldsymbol{\alpha}_t) &\propto 1(\varphi_{it} \leq 0) \times \mathbf{N}(\mathbb{E}_{it}, \mathbb{V}_{it}) \\ &= \begin{cases} \text{TN}(\mathbb{E}_{it}, \mathbb{V}_{it}, \leq 0) & Y_{it} = 0, \\ Y_{it} & Y_{it} > 0, \end{cases} \end{aligned}$$

where $\text{TN}(\mu, \sigma^2; \leq 0)$ specifies a truncated normal distribution with mean, μ , and variance, σ^2 , truncated above by zero. Recall the moments, $\mathbb{E}_{it} = [\rho \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t) \varphi_{jt} + (1 - \rho)\mu_t] / \Delta$ and $\mathbb{V}_{it} = \tau_t^2 / \Delta$ with $\Delta = \rho \sum_{j=1}^{m_t} w_{ij}(\boldsymbol{\alpha}_t) + (1 - \rho)$.

2. Sample from the full conditional distribution for ζ :

$$f(\zeta | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T}, \phi) \propto \left[\prod_{t=1}^{\nu} \prod_{i=1}^{m_t} f(Y_{it} | g^{-1}(\varphi_{it}), \zeta) \right] \times f(\zeta).$$

The full conditional for ζ depends on the observational model specification. In general, the full conditional distribution for ζ has no closed form. For VF data there are no nuisance parameters in the likelihood, so it is not necessary to specify this full conditional.

3. Sample from the full conditional distribution of $\boldsymbol{\theta}_{1\cdot}$: (i.e., μ_1, \dots, μ_ν)

It is not immediately obvious, but with some manipulation we can find a closed form full conditional distribution for $\boldsymbol{\theta}_{1\cdot}$. The key is that we can find an analytical form of the distribution $f(\boldsymbol{\theta}_{1\cdot} | \boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi)$. Using properties of the multivariate normal distribution we know that $\text{vec}(\boldsymbol{\theta}^T) \sim \text{MVN}(\text{vec}(\mathbf{M}^T), \mathbf{T} \otimes \boldsymbol{\Sigma}(\phi))$, where $\mathbf{M} = \boldsymbol{\delta} \mathbf{1}_\nu^T$. Define $s = \{2, 3, \dots, q + 2\}$,

so that we can write the covariance in block form,

$$\mathbf{T} = \begin{bmatrix} \mathbf{T}_{11} & \mathbf{T}_{1s} \\ \mathbf{T}_{s1} & \mathbf{T}_{ss} \end{bmatrix} \implies \mathbf{T} \otimes \boldsymbol{\Sigma}(\phi) = \begin{bmatrix} \mathbf{T}_{11} \otimes \boldsymbol{\Sigma}(\phi) & \mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi) \\ \mathbf{T}_{s1} \otimes \boldsymbol{\Sigma}(\phi) & \mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi) \end{bmatrix}.$$

Using the properties of the multivariate normal distribution we can compute the moments of $f(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_s, \boldsymbol{\delta}, \mathbf{T}, \phi) \sim \mathcal{N}(\mathbb{E}_{1|s}, \mathbb{C}_{1|s})$,

$$\begin{aligned} \mathbb{E}_{1|s} &= \mathbb{E}[\boldsymbol{\theta}_1 | \boldsymbol{\theta}_s] = \mathbf{M}_1 + [\mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi)] [\mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi)]^{-1} \text{vec}(\boldsymbol{\theta}_s - \mathbf{M}_s) \\ &= \mathbf{M}_1 + [\mathbf{T}_{1s} \mathbf{T}_{ss}^{-1} \otimes \mathbf{I}_\nu] \text{vec}(\boldsymbol{\theta}_s - \mathbf{M}_s), \end{aligned}$$

and

$$\begin{aligned} \mathbb{C}_{1|s} &= \mathbb{C}(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_s) = [\mathbf{T}_{11} \otimes \boldsymbol{\Sigma}(\phi)] - [\mathbf{T}_{1s} \otimes \boldsymbol{\Sigma}(\phi)] [\mathbf{T}_{ss} \otimes \boldsymbol{\Sigma}(\phi)]^{-1} [\mathbf{T}_{s1} \otimes \boldsymbol{\Sigma}(\phi)] \\ &= [\mathbf{T}_{1|s} \otimes \boldsymbol{\Sigma}(\phi)], \end{aligned}$$

where $\mathbf{T}_{1|s} = (\mathbf{T}_{11} - \mathbf{T}_{1s} \mathbf{T}_{ss}^{-1} \mathbf{T}_{s1})$. Now, we can derive the full conditional for $\boldsymbol{\theta}_1$. We define a matrix specification of the random effects,

$$f(\boldsymbol{\varphi} | \boldsymbol{\theta}) \sim \text{MVN}(\mathbf{Z}_\theta \boldsymbol{\theta}_1, \mathbf{Q}^{-1}),$$

where $\mathbf{Z}_\theta = (\mathbf{I}_\nu \otimes \mathbf{1}_{m_t})$ and $\mathbf{Q} = \text{BlockDiag}\{\mathbf{Q}(\alpha_t) / \tau_t^2\}$, for $t = 1, \dots, \nu$. Then, the full conditional is as follows,

$$\begin{aligned} f(\boldsymbol{\theta}_1 | \mathbf{Y}, \boldsymbol{\zeta}, \boldsymbol{\varphi}, \boldsymbol{\theta}_s, \boldsymbol{\delta}, \mathbf{T}, \phi) &\propto f(\boldsymbol{\varphi} | \boldsymbol{\theta}) \times f(\boldsymbol{\theta}_1 | \boldsymbol{\theta}_s, \boldsymbol{\delta}, \mathbf{T}, \phi) \\ &\propto \exp \left\{ -\frac{1}{2} \left[(\boldsymbol{\varphi} - \mathbf{Z}_\theta \boldsymbol{\theta}_1)^T \mathbf{Q} (\boldsymbol{\varphi} - \mathbf{Z}_\theta \boldsymbol{\theta}_1) \right. \right. \\ &\quad \left. \left. + (\boldsymbol{\theta}_1 - \mathbb{E}_{1|s})^T \mathbb{C}_{1|s}^{-1} (\boldsymbol{\theta}_1 - \mathbb{E}_{1|s}) \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\theta}_1^T (\mathbf{Z}_\theta^T \mathbf{Q} \mathbf{Z}_\theta + \mathbb{C}_{1|s}^{-1}) \boldsymbol{\theta}_1 \right. \right. \\ &\quad \left. \left. - 2\boldsymbol{\theta}_1^T (\mathbf{Z}_\theta^T \mathbf{Q} \boldsymbol{\varphi} + \mathbb{C}_{1|s}^{-1} \mathbb{E}_{1|s}) \right] \right\} \\ &\sim \text{MVN}(\mathbb{E}_\theta, \mathbb{V}_\theta), \end{aligned}$$

where, $\mathbb{V}_\theta = \left(\mathbf{Z}_\theta^T \mathbf{Q} \mathbf{Z}_\theta + \mathbb{C}_{1|s}^{-1} \right)^{-1}$ and $\mathbb{E}_\theta = \mathbb{V}_\theta \left(\mathbf{Z}_\theta^T \mathbf{Q} \boldsymbol{\varphi} + \mathbb{C}_{1|s}^{-1} \mathbb{E}_{1|s} \right)$.

4. Sample τ_t^2 using a Metropolis step:

Each τ_t^2 , $t = 1, \dots, \nu$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

$$f(\log(\tau_t) | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}_{-\tau_{tk}}, \boldsymbol{\delta}, \mathbf{T}, \phi) \propto f(\boldsymbol{\varphi}_t | \mu_t, \tau_t^2, \boldsymbol{\alpha}_t) \times f(\boldsymbol{\theta}_2 | \boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi),$$

where $s = \{1, 3, \dots, q + 2\}$.

5. Sample α_{tk} using a Metropolis step:

Each α_{tk} , $t = 1, \dots, \nu$ and $k = 1, \dots, q$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

$$f(\log(\alpha_{tk}) | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}_{-\alpha_{tk}}, \boldsymbol{\delta}, \mathbf{T}, \phi) \propto f(\boldsymbol{\varphi}_t | \mu_t, \tau_t^2, \boldsymbol{\alpha}_t) \times f(\boldsymbol{\theta}_{k+2} | \boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi),$$

where $s = \{1, 2, 3, \dots, q + 2\}$ with k removed.

6. Sample from the full conditional distribution for $\boldsymbol{\delta}$:

Define $\mathbf{Z}_\delta = \mathbf{1}_\nu \otimes \mathbf{I}_3$ and $\mathbb{C}_\theta = \boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}$. Then the full conditional distribution can be derived as follows,

$$\begin{aligned} f(\boldsymbol{\delta} | \mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \mathbf{T}, \phi) &\propto f(\text{vec}(\boldsymbol{\theta}) | \boldsymbol{\delta}, \mathbf{T}, \phi) \times f(\boldsymbol{\delta}) \\ &\propto \exp \left\{ -\frac{1}{2} \left[(\text{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta \boldsymbol{\delta})^T \mathbb{C}_\theta^{-1} (\text{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta \boldsymbol{\delta}) \right. \right. \\ &\quad \left. \left. + (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta)^T \boldsymbol{\Omega}_\delta^{-1} (\boldsymbol{\delta} - \boldsymbol{\mu}_\delta) \right] \right\} \\ &\propto \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\delta}^T (\mathbf{Z}_\delta^T \mathbb{C}_\theta^{-1} \mathbf{Z}_\delta + \boldsymbol{\Omega}_\delta^{-1}) \boldsymbol{\delta} \right. \right. \\ &\quad \left. \left. - 2\boldsymbol{\delta}^T (\mathbf{Z}_\delta^T \mathbb{C}_\theta^{-1} \text{vec}(\boldsymbol{\theta}) + \boldsymbol{\Omega}_\delta^{-1} \boldsymbol{\mu}_\delta) \right] \right\} \\ &\sim \text{N}(\mathbb{E}_\delta, \mathbb{V}_\delta), \end{aligned}$$

where, $\mathbb{V}_\delta = (\mathbf{Z}_\delta^T \mathbb{C}_\theta^{-1} \mathbf{Z}_\delta + \boldsymbol{\Omega}_\delta^{-1})^{-1}$ and $\mathbb{E}_\delta = \mathbb{V}_\delta (\mathbf{Z}_\delta^T \mathbb{C}_\theta^{-1} \text{vec}(\boldsymbol{\theta}) + \boldsymbol{\Omega}_\delta^{-1} \boldsymbol{\mu}_\delta)$.

7. Sample from the full conditional distribution of \mathbf{T} :

We begin by rewriting the quadratic form of $\boldsymbol{\theta}$,

$$\begin{aligned} f(\text{vec}(\boldsymbol{\theta})|\boldsymbol{\delta}, \mathbf{T}, \phi) &\propto \exp \left\{ -\frac{1}{2} (\text{vec}(\boldsymbol{\theta}) - \text{vec}(\mathbf{M}))^T (\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T})^{-1} (\text{vec}(\boldsymbol{\theta}) - \text{vec}(\mathbf{M})) \right\} \\ &= \exp \left\{ -\frac{1}{2} \text{tr} \left((\boldsymbol{\theta} - \mathbf{M}) \boldsymbol{\Sigma}(\phi)^{-1} (\boldsymbol{\theta} - \mathbf{M})^T \mathbf{T}^{-1} \right) \right\} \\ &= \exp \left\{ -\frac{1}{2} \text{tr} (\mathbf{S}_\theta \mathbf{T}^{-1}) \right\}, \end{aligned}$$

where $\mathbf{S}_\theta = (\boldsymbol{\theta} - \mathbf{M}) \boldsymbol{\Sigma}(\phi)^{-1} (\boldsymbol{\theta} - \mathbf{M})^T$. Now, the full conditional distribution for \mathbf{T} is straightforward,

$$\begin{aligned} f(\mathbf{T}|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \phi) &\propto f(\text{vec}(\boldsymbol{\theta})|\boldsymbol{\delta}, \mathbf{T}, \phi) \times f(\mathbf{T}) \\ &\propto |\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T}|^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} \text{tr} (\mathbf{S}_\theta \mathbf{T}^{-1}) \right\} \\ &\quad \times |\mathbf{T}|^{-\frac{(\xi+3+1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr} (\boldsymbol{\Psi} \mathbf{T}^{-1}) \right\} \\ &\propto |\mathbf{T}|^{-\frac{(\xi+\nu+3+1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr} ([\mathbf{S}_\theta + \boldsymbol{\Psi}] \mathbf{T}^{-1}) \right\} \\ &\sim \text{Inverse-Wishart} (\xi + \nu, \mathbf{S}_\theta + \boldsymbol{\Psi}). \end{aligned}$$

8. Sample ϕ using a Metropolis step:

We transform ϕ to the real line to facilitate sampling. Define a new parameter $\Delta = h(\phi) = \log \left(\frac{\phi - a_\phi}{b_\phi - \phi} \right)$, such that $h^{-1}(\Delta) = (b_\phi \exp\{\Delta\} + a_\phi) / (1 + \exp\{\Delta\})$ and $|\frac{\partial}{\partial \Delta} h^{-1}(\Delta)| \propto \exp\{\Delta\} / (1 + \exp\{\Delta\})^2$. Now we can sample from the transformed proposal distribution, $\Delta^* \sim \text{N}(\Delta, \delta)$, where δ is a tuning parameter. Then we can obtain a proposal of ϕ , $\phi^* = h^{-1}(\Delta^*)$. Finally, if we define $\mathbf{R} = \text{vec}(\boldsymbol{\theta}) - \mathbf{Z}_\delta \boldsymbol{\delta}$ the Metropolis ratio can be calculated as

$$\begin{aligned} r &= \frac{f(\phi^*|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T})}{f(\phi|\mathbf{Y}, \boldsymbol{\varphi}, \boldsymbol{\zeta}, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T})} \\ &\propto \frac{f(\boldsymbol{\theta}|\boldsymbol{\delta}, \mathbf{T}, h^{-1}(\Delta^*)) \times \left| \frac{\partial}{\partial \Delta^*} h^{-1}(\Delta^*) \right|}{f(\boldsymbol{\theta}|\boldsymbol{\delta}, \mathbf{T}, h^{-1}(\Delta)) \times \left| \frac{\partial}{\partial \Delta} h^{-1}(\Delta) \right|} \\ &\propto \frac{|\boldsymbol{\Sigma}(\phi^*)|^{-\frac{3}{2}} \exp \left\{ -\frac{1}{2} [\mathbf{R}^T (\boldsymbol{\Sigma}(\phi^*) \otimes \mathbf{T})^{-1} \mathbf{R}] \right\} \frac{\exp\{\Delta^*\}}{(1 + \exp\{\Delta^*\})^2}}{|\boldsymbol{\Sigma}(\phi)|^{-\frac{3}{2}} \exp \left\{ -\frac{1}{2} [\mathbf{R}^T (\boldsymbol{\Sigma}(\phi) \otimes \mathbf{T})^{-1} \mathbf{R}] \right\} \frac{\exp\{\Delta\}}{(1 + \exp\{\Delta\})^2}}. \end{aligned}$$

Now accept ϕ^* with probability $\min\{1, r\}$, otherwise keep ϕ .

9. Repeat steps 1-8 until convergence has been achieved and an adequate number of posterior samples have been obtained post-convergence.

A.3 Prediction

In Section 2.5.5, we present the PPD for a future set of VF responses as a function of four known densities that are defined as a consequence of our methodology,

$$\begin{aligned}
 f(\mathbf{Y}_{\nu+1}|\mathbf{Y}) &= \int_{\Omega} f(\mathbf{Y}_{\nu+1}|\Omega, \mathbf{Y}) f(\Omega|\mathbf{Y}) d\Omega \\
 &= \int_{\Omega} \underbrace{f(\mathbf{Y}_{\nu+1}|g^{-1}(\varphi_{\nu+1}), \zeta)}_1 \underbrace{f(\varphi_{\nu+1}|\boldsymbol{\theta}_{\cdot\nu+1})}_2 \underbrace{f(\boldsymbol{\theta}_{\cdot\nu+1}|\boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T}, \phi)}_3 \\
 &\quad \times \underbrace{(\zeta, \boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T}, \phi|\mathbf{Y})}_4 d\Omega.
 \end{aligned} \tag{A.1}$$

Equation A.1.1 represents the observed likelihood function written in vector form and is problem specific (or in scalar form: $\prod_{i=1}^{m_{\nu+1}} f(Y_{i(\nu+1)}|g^{-1}(\varphi_{i(\nu+1)}), \zeta)$). Equation A.1.2 is the joint specification of the random effects. Equation A.1.3 is the least trivial of the four distributions to obtain, but using the properties of the conditional multivariate normal distribution in a similar manner to $f(\boldsymbol{\theta}_{1\cdot}|\boldsymbol{\theta}_{s\cdot}, \boldsymbol{\delta}, \mathbf{T}, \phi)$, the following can be obtained,

$$f(\boldsymbol{\theta}_{\cdot\nu+1}|\boldsymbol{\theta}, \boldsymbol{\delta}, \mathbf{T}, \phi) \sim \text{MVN}(\mathbb{E}_{\boldsymbol{\theta}_{\cdot\nu+1}}, \mathbb{C}_{\boldsymbol{\theta}_{\cdot\nu+1}}).$$

The moments are $\mathbb{E}_{\boldsymbol{\theta}_{\cdot\nu+1}} = \boldsymbol{\delta} + (\boldsymbol{\Sigma}^+ \otimes \mathbf{T})(\text{vec}(\boldsymbol{\theta}) - \mathbf{1}_{\nu} \otimes \boldsymbol{\delta})$ and $\mathbb{C}_{\boldsymbol{\theta}_{\cdot\nu+1}} = \boldsymbol{\Sigma}^* \otimes \mathbf{T}$ with $\boldsymbol{\Sigma}^+ = [\boldsymbol{\Sigma}(\phi)]_{\nu+1,1:\nu} [\boldsymbol{\Sigma}(\phi)]_{1:\nu,1:\nu}^{-1}$ and $\boldsymbol{\Sigma}^* = [\boldsymbol{\Sigma}(\phi)]_{\nu+1,\nu+1} - \boldsymbol{\Sigma}^+ [\boldsymbol{\Sigma}(\phi)]_{1:\nu,\nu+1}$. Here $\boldsymbol{\Sigma}(\phi)$ is a temporal correlation matrix including the new time point $\nu + 1$, so that $[\boldsymbol{\Sigma}(\phi)]_{\nu+1,1:\nu}$ is a subset including the row $\nu + 1$ and columns 1 up to ν . Finally, Equation A.1.4 is the posterior distribution obtained in the MCMC sampler from the original model fit.

A.4 Sensitivity to the Mis-specification of ρ

In Section 2.7, we explored performance of the model in settings of changing temporal correlation and cross-covariance, while in this simulation we explore the sensitivity of the model to changing values of ρ . We simulate data from setting D with the maximum number of visits (i.e., 21), such that there is temporal correlation and cross-covariance. In this study, we generate data with $\rho = 0.4, 0.6, 0.9, 0.99$, and apply the model proposed in Chapter 2 with ρ fixed at 0.99 allowing us to quantify the effect on model performance of mis-specifying ρ . Additionally, we fit concordant models for each value of ρ to provide context for the mis-specified settings. The results from the simulation study are displayed in Table A.1, where each reported estimate is based on 1,000 simulated datasets. In Table

Table A.1: The effect of mis-specifying the spatial correlation parameter ρ on model performance. The table displays MSE and EC for both mis-specified and concordant settings. Each reported estimate is based on 1,000 simulated datasets.

True ρ	Mis-specified		Concordant	
	MSE	EC	MSE	EC
0.99	–	–	0.060	0.98
0.90	0.056	0.94	0.052	0.98
0.60	0.085	0.96	0.074	0.99
0.40	0.115	0.95	0.095	0.99

A.1 it is clear that when our model is mis-specifying ρ , estimation of the posterior mean of the CV of α_{tGH} (i.e., “ST CV”) is stable. This is consistent for all of the values of ρ in Table A.1, as the EC stays near the nominal coverage value of 95%. Furthermore, the nearly identical MSE values in the mis-specified and concordant columns indicate fixing ρ at 0.99 does not result in any loss in efficiency.

A.5 Sensitivity Analysis in the Visual Field Analysis

In Section 2.6, we applied the introduced methodology to VF data. We assess sensitivity of those results to four modeling assumptions: i) the regularization upper bound for $\log(\alpha_{tGH}), v_1$, ii) the temporal correlation structure, $\Sigma(\phi)$, iii) the mean hyperparameter for $[\delta]_1, [\mu_\delta]_1$, and iv) the criterion for the bounds of ϕ . The “ST CV” metric is re-calculated under each sensitivity setting and regressed against the clinical assessment of progression with the results displayed in Table A.2

Table A.2: Sensitivity analysis to various assumptions in analysis of visual field data and glaucoma progression risk. For each assumption new values are presented with risk presented for comparison. Assumptions include i) upper bound for $\log(\alpha_{tGH})$, v_1 , ii) correlation structure, $\Sigma(\phi)$, iii) mean hyperparameter for $[\delta]_1$, $[\mu_\delta]_1$, and iv) criterion for the bounds of ϕ .

Assumption	Value	Estimate	Std. Error	z value	Pr(> z)
v_1	0.25	1.70	0.78	2.17	0.030
	0.50	2.03	0.84	2.43	0.015
	1.00	2.08	0.85	2.44	0.015
	2.00	2.02	0.86	2.35	0.019
$\Sigma(\phi)$	EXP	2.08	0.85	2.44	0.015
	AR(1)	3.05	1.02	2.98	0.003
$[\mu_\delta]_1$	0	2.07	0.85	2.44	0.015
	3	2.08	0.85	2.44	0.015
ϕ	0.99	2.07	0.82	2.52	0.012
	0.95	2.08	0.85	2.44	0.015

In the manuscript, the regularization upper bound for $\log(\alpha_{tGH})$ is set at 1.00, while values of 0.25, 0.50, and 2.00 are used for comparison in the sensitivity analysis. The results are clearly robust to the choice of v_1 , as the p-values do not change. To demonstrate sensitivity to the exponential temporal correlation structure, we compare it to a continuous time autoregressive with lag one structure, AR(1), $[\Sigma(\phi)]_{t,t'} = \phi^{|x_t - x_{t'}|}$. When using the AR(1) temporal correlation structure, it can be seen that “ST CV” is highly predictive of glaucoma progression (p-value: 0.0029), indicating that the results are robust over correlation structures. Next, we assess sensitivity to the specification of choosing $[\mu_\delta]_1 = 3$, which was clinically motivated, comparing it to the objective choice of zero. The results are robust to the choice of $[\mu_\delta]_1$, which is reassuring, but not surprising since the corresponding variance was 1,000. Finally, we compare the original results to a version where the lower bound of ϕ was based on the following criterion, $[a_\phi : [\Sigma(a_\phi)]_{t,t'} = 0.99, |t - t'| = x_{\max}]$, changed from the original 0.95. This criterion is motivated to give wider and symmetric bounds. From Table A.2, it is clear that the results are robust to changing this criterion. Overall, our model specification is robust to the modeling assumptions and prior choices in the VF data setting.

A.6 Additional Figures

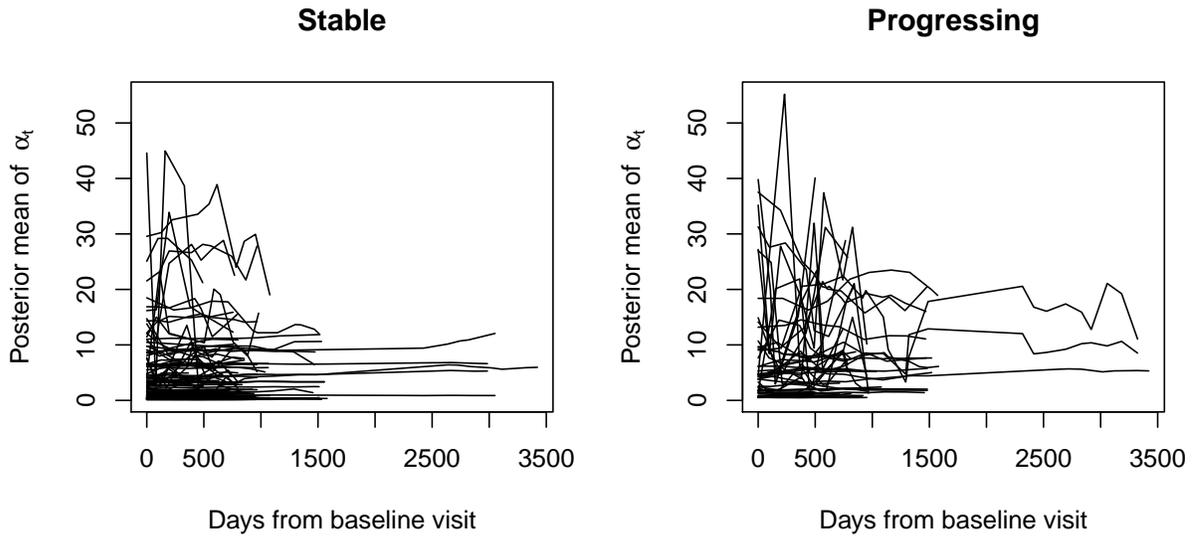


Figure A.1: Trajectories of the posterior mean estimates of α_{tGH} over time for all patients, subset by progression status.

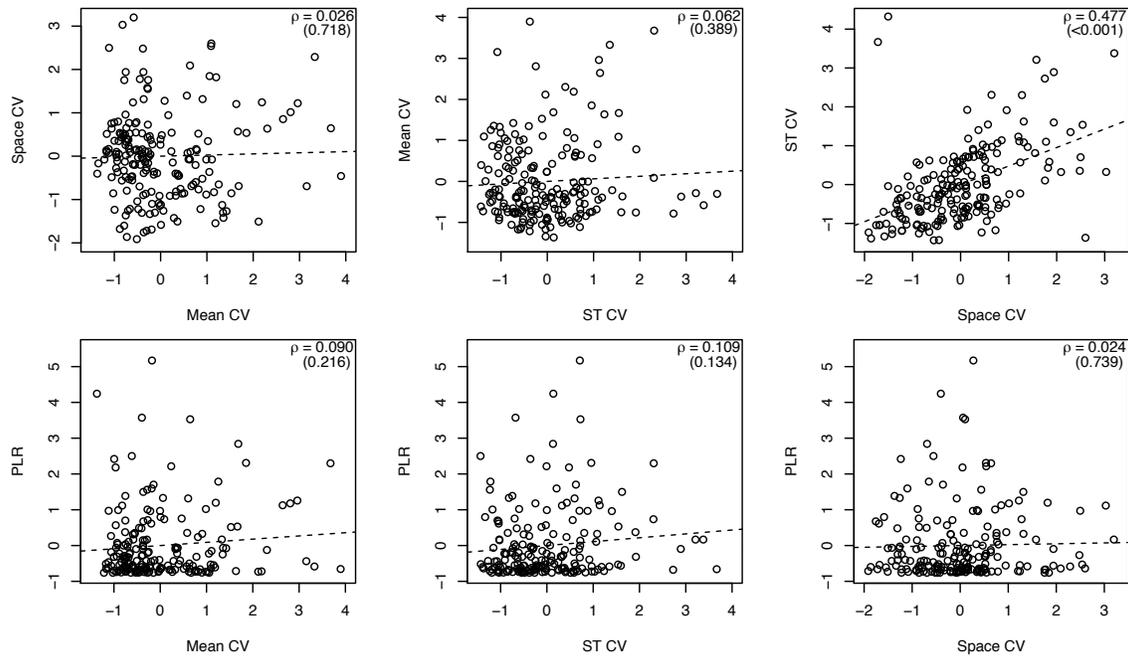


Figure A.2: Pairwise scatterplots between standardized diagnostic metrics. Included are estimated Pearson correlations, corresponding p-values from the hypothesis that the correlations are equal to zero, and regression trend lines.

APPENDIX B: SUPPLEMENTARY MATERIALS FOR CHAPTER 4

B.1 Change Point Likelihood

The change point likelihood can be written for the latent Tobit process,

$$f \left\{ Y_t(\mathbf{s}_i) \mid \beta_0(\mathbf{s}_i), \beta_1(\mathbf{s}_i), \lambda_0(\mathbf{s}_i), \lambda_1(\mathbf{s}_i), \theta(\mathbf{s}_i) \right\} = \prod_{t=1}^{\nu} \prod_{i=1}^m \left[\mathbf{N} \left(\beta_0(\mathbf{s}_i), e^{2\lambda_0(\mathbf{s}_i)} \right) \right]^{1(0 \leq x_t \leq \theta(\mathbf{s}_i))} \\ \times \left[\mathbf{N} \left(\beta_0(\mathbf{s}_i) + \beta_1(\mathbf{s}_i) \{x_t - \theta(\mathbf{s}_i)\}, e^{2[\lambda_0(\mathbf{s}_i) + \lambda_1(\mathbf{s}_i)\{x_t - \theta(\mathbf{s}_i)\}]} \right) \right]^{1(\theta(\mathbf{s}_i) \leq x_t \leq x_\nu)}.$$

This likelihood is informative, however for computational purposes we will write the likelihood in matrix form. The latent responses can be written more concisely using an indicator variable. Define the event $\Upsilon_t(\mathbf{s}_i) = 1\{\theta(\mathbf{s}_i) \leq x_t \leq x_\nu\}$. Then,

$$Y_t(\mathbf{s}_i) = \beta_0(\mathbf{s}_i) + \beta_1(\mathbf{s}_i) \{x_t - \theta(\mathbf{s}_i)\} \Upsilon_t(\mathbf{s}_i) + \epsilon_t(\mathbf{s}_i) \\ = x_t(\mathbf{s}_i; \theta) \boldsymbol{\beta}(\mathbf{s}_i) + \epsilon_t(\mathbf{s}_i),$$

where $\boldsymbol{\beta}(\mathbf{s}_i) = \{\beta_0(\mathbf{s}_i), \beta_1(\mathbf{s}_i)\}^T$ and $x_t(\mathbf{s}_i; \theta) = [1, \{x_t - \theta(\mathbf{s}_i)\} \Upsilon_t(\mathbf{s}_i)]$. Now, we determine the vector of latent observations at each time point, defining $\mathbf{Y}_t = \{Y_t(\mathbf{s}_1), Y_t(\mathbf{s}_2), \dots, Y_t(\mathbf{s}_m)\}^T$,

$$\mathbf{Y}_t = X_t(\theta) \boldsymbol{\beta} + \boldsymbol{\epsilon}_t,$$

where $\boldsymbol{\epsilon}_t = \{\epsilon_t(\mathbf{s}_1), \epsilon_t(\mathbf{s}_2), \dots, \epsilon_t(\mathbf{s}_m)\}^T$, $\boldsymbol{\beta} = \{\boldsymbol{\beta}(\mathbf{s}_1)^T, \boldsymbol{\beta}(\mathbf{s}_2)^T, \dots, \boldsymbol{\beta}(\mathbf{s}_m)^T\}^T$ and $X_t(\theta)$ is an $m \times 2m$ block diagonal matrix with blocks entries $x_t(\mathbf{s}_i; \theta)$. Finally, we write the full latent process, $\mathbf{Y} = \{\mathbf{Y}_1^T, \mathbf{Y}_2^T, \dots, \mathbf{Y}_\nu^T\}^T$,

$$\mathbf{Y} = \mathbf{X}(\theta) \boldsymbol{\beta} + \boldsymbol{\epsilon},$$

where $\boldsymbol{\epsilon} = \{\boldsymbol{\epsilon}_1^T, \boldsymbol{\epsilon}_2^T, \dots, \boldsymbol{\epsilon}_\nu^T\}^T$ and $\mathbf{X}(\theta)$ is an $m\nu \times 2m$ matrix stacking the $X_t(\theta)$ matrices.

Having written the mean process in matrix form, we turn our attention to the variance process. Define,

$\sigma_t^2 = \{\sigma_t^2(\mathbf{s}_1), \sigma_t^2(\mathbf{s}_2), \dots, \sigma_t^2(\mathbf{s}_m)\}^T$. Then $\sigma^2 = \{(\sigma_1^2)^T, (\sigma_2^2)^T, \dots, (\sigma_m^2)^T\}^T$. The variance process can be written using the same form as the mean process,

$$\log \{\sigma\} = \mathbf{X}(\theta)\boldsymbol{\lambda} \implies \sigma^2 = \exp\{2\mathbf{X}(\theta)\boldsymbol{\lambda}\},$$

where $\boldsymbol{\lambda} = \{\boldsymbol{\lambda}(\mathbf{s}_1)^T, \boldsymbol{\lambda}(\mathbf{s}_2)^T, \dots, \boldsymbol{\lambda}(\mathbf{s}_m)^T\}^T$, with $\boldsymbol{\lambda}(\mathbf{s}_i) = \{\lambda_0(\mathbf{s}_i), \lambda_1(\mathbf{s}_i)\}^T$. Define the full likelihood covariance as $\boldsymbol{\Omega} = \text{Diag}(\sigma^2)$. Then, we can write the matrix likelihood as follows,

$$f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\theta}) \sim \text{MVN}(\mathbf{X}(\boldsymbol{\theta})\boldsymbol{\beta}, \boldsymbol{\Omega})$$

$$\log \{\sigma\} = \mathbf{X}(\boldsymbol{\theta})\boldsymbol{\lambda},$$

where $\boldsymbol{\theta} = \{\theta(\mathbf{s}_1), \theta(\mathbf{s}_2), \dots, \theta(\mathbf{s}_m)\}^T$.

Having written the likelihood in a convenient matrix form, we now write the full data likelihood which will be used to find the full conditionals. Note that the likelihood does not include $\boldsymbol{\theta}$, but rather is parameterized to include the generating parameters,

$$\begin{aligned} f(\mathbf{Y}, \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) &\propto f(\mathbf{Y}|\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}) \\ &\times f(\{\boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}\}|\boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) f(\boldsymbol{\delta}) f(\boldsymbol{\Sigma}) f(\alpha), \end{aligned}$$

where $\boldsymbol{\eta} = \{\eta(\mathbf{s}_1), \eta(\mathbf{s}_2), \dots, \eta(\mathbf{s}_m)\}^T$.

B.2 Full Conditionals

1. Gibbs Step for $\boldsymbol{\delta}$:

Note that $(\mathbf{1}_m \otimes \boldsymbol{\delta}) = (\mathbf{1}_m \otimes \mathbf{I}_5)\boldsymbol{\delta} = \mathbf{Z}_\delta\boldsymbol{\delta}$. Then,

$$\begin{aligned} f(\boldsymbol{\delta}|\cdot) &\propto f(\boldsymbol{\phi}|\boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \times f(\boldsymbol{\delta}) \\ &\propto \exp\left\{-\frac{1}{2}\left[\{\boldsymbol{\phi} - \mathbf{Z}_\delta\boldsymbol{\delta}\}^T (\mathbf{Q}(\alpha) \otimes \boldsymbol{\Sigma}^{-1}) \{\boldsymbol{\phi} - \mathbf{Z}_\delta\boldsymbol{\delta}\} + \frac{\boldsymbol{\delta}^T \boldsymbol{\delta}}{\kappa^2}\right]\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\delta}^T \left\{\mathbf{Z}_\delta^T (\mathbf{Q}(\alpha) \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{Z}_\delta + \frac{\mathbf{I}_5}{\kappa^2}\right\} \boldsymbol{\delta} - 2\boldsymbol{\delta}^T \{\mathbf{Z}_\delta^T \boldsymbol{\Omega}^{-1} \boldsymbol{\phi}\}\right]\right\} \\ &\sim \text{MVN}(\mathbb{E}_\delta, \mathbb{C}_\delta), \end{aligned}$$

where $\mathbb{C}_\delta = \{\mathbf{Z}_\delta^T (\mathbf{Q}(\alpha) \otimes \boldsymbol{\Sigma}^{-1}) \mathbf{Z}_\delta + \frac{\mathbf{I}_5}{\kappa^2}\}^{-1}$, $\mathbb{E}_\delta = \mathbb{C}_\delta \{\mathbf{Z}_\delta^T \boldsymbol{\Omega}^{-1} \boldsymbol{\phi}\}$.

2. Gibbs Step for $\boldsymbol{\beta}$:

It is not immediately obvious, but with some manipulation we can find a closed form full conditional distribution for $\boldsymbol{\beta}$. The key is that we can find an analytical form of the distribution $f(\boldsymbol{\beta}|\boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha)$. Recall that, $\mathbb{C}(\boldsymbol{\phi}) = \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}$. Define, $k = \{1, 2\}$ and $j = \{3, 4, 5\}$ as the indices of the components of $\boldsymbol{\beta}$ in the vector of random effects $\boldsymbol{\phi}_i$. The mean can be partitioned as,

$$\mathbf{1}_m \otimes \boldsymbol{\delta} = \begin{bmatrix} \mathbf{1}_m \otimes \boldsymbol{\delta}_k \\ \mathbf{1}_m \otimes \boldsymbol{\delta}_j \end{bmatrix}. \quad (\text{B.1})$$

The covariance matrix can then be partitioned in block form as follows,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_{kk} & \boldsymbol{\Sigma}_{kj} \\ \boldsymbol{\Sigma}_{jk} & \boldsymbol{\Sigma}_{jj} \end{bmatrix} \implies \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma} = \begin{bmatrix} \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{kk} & \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{kj} \\ \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{jk} & \mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{jj} \end{bmatrix}.$$

Define $\boldsymbol{\phi}_{ij}$ as the original $\boldsymbol{\phi}_i$ vector, only keeping the indices included in j and $\boldsymbol{\phi}^j = \{\boldsymbol{\phi}_{1j}^T, \boldsymbol{\phi}_{2j}^T, \dots, \boldsymbol{\phi}_{mj}^T\}^T$. Using the properties of the multivariate normal distribution we can compute the moments of $f(\boldsymbol{\beta}|\boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \sim \text{MVN}(\mathbb{E}_{k|j}, \mathbb{C}_{k|j})$,

$$\begin{aligned} \mathbb{E}_{k|j} &= \mathbb{E}[\boldsymbol{\beta}|\boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha] = (\mathbf{1}_m \otimes \boldsymbol{\delta}_k) + \left[\mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{kj} \right] \left[\mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{jj} \right]^{-1} (\boldsymbol{\phi}^j - (\mathbf{1}_m \otimes \boldsymbol{\delta}_j)) \\ &= (\mathbf{1}_m \otimes \boldsymbol{\delta}_k) + \left[\mathbf{I}_n \otimes \boldsymbol{\Sigma}_{kj} \boldsymbol{\Sigma}_{jj}^{-1} \right] (\boldsymbol{\phi}^j - (\mathbf{1}_m \otimes \boldsymbol{\delta}_j)), \end{aligned}$$

and

$$\begin{aligned} \mathbb{C}_{k|j} &= \mathbb{C}(\boldsymbol{\beta}|\boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) = \left[\mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{kk} \right] - \left[\mathbf{I}_n \otimes \boldsymbol{\Sigma}_{kj} \boldsymbol{\Sigma}_{jj}^{-1} \right] \left[\mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{jk} \right] \\ &= \left[\mathbf{Q}(\alpha)^{-1} \otimes \boldsymbol{\Sigma}_{k|j} \right], \end{aligned}$$

where $\Sigma_{k|j} = (\Sigma_{kk} - \Sigma_{kj}\Sigma_{jj}^{-1}\Sigma_{jk})$. Now, we can derive the full conditional for β .

$$\begin{aligned}
f(\beta|\cdot) &\propto f(\mathbf{Y}|\beta, \boldsymbol{\lambda}, \boldsymbol{\eta}) \times f(\beta|\boldsymbol{\lambda}, \boldsymbol{\eta}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \\
&\propto \exp\left\{-\frac{1}{2}\left[\{\mathbf{Y} - \mathbf{X}(\boldsymbol{\theta})\boldsymbol{\beta}\}^T \boldsymbol{\Omega}^{-1}\{\mathbf{Y} - \mathbf{X}(\boldsymbol{\theta})\boldsymbol{\beta}\} + (\boldsymbol{\beta} - \mathbb{E}_{k|j})^T \mathbb{C}_{k|j}^{-1}(\boldsymbol{\beta} - \mathbb{E}_{k|j})\right]\right\} \\
&\propto \exp\left\{-\frac{1}{2}\left[\boldsymbol{\beta}^T \left\{\mathbf{X}(\boldsymbol{\theta})^T \boldsymbol{\Omega}^{-1} \mathbf{X}(\boldsymbol{\theta}) + \mathbb{C}_{k|j}^{-1}\right\} \boldsymbol{\beta} - 2\boldsymbol{\beta}^T \left\{\mathbf{X}(\boldsymbol{\theta})^T \boldsymbol{\Omega}^{-1} \mathbf{Y} + \mathbb{C}_{k|j}^{-1} \mathbb{E}_{k|j}\right\}\right]\right\} \\
&\sim \text{MVN}(\mathbb{E}_\beta, \mathbb{C}_\beta),
\end{aligned}$$

$$\text{where } \mathbb{C}_\beta = \left\{\mathbf{X}(\boldsymbol{\theta})^T \boldsymbol{\Omega}^{-1} \mathbf{X}(\boldsymbol{\theta}) + \mathbb{C}_{k|j}^{-1}\right\}^{-1}, \mathbb{E}_\beta = \mathbb{C}_\beta \left\{\mathbf{X}(\boldsymbol{\theta})^T \boldsymbol{\Omega}^{-1} \mathbf{Y} + \mathbb{C}_{k|j}^{-1} \mathbb{E}_{k|j}\right\}.$$

3. Metropolis Step for $\lambda_0(\mathbf{s}_i)$:

Each $\lambda_0(\mathbf{s}_i)$, $i = 1, \dots, m$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

$$\begin{aligned}
f\{\lambda_0(\mathbf{s}_i)|\cdot\} &\propto f(\mathbf{Y}|\beta, \boldsymbol{\lambda}, \boldsymbol{\eta}) \times f(\phi|\boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \\
&\propto \prod_{t=1}^{\nu} \left[\text{N}\left(\beta_0(\mathbf{s}_i), e^{2\lambda_0(\mathbf{s}_i)}\right) \right]^{1(0 \leq x_t \leq \theta(\mathbf{s}_i))} \\
&\quad \times \left[\text{N}\left(\beta_0(\mathbf{s}_i) + \beta_1(\mathbf{s}_i)\{x_t - \theta(\mathbf{s}_i)\}, e^{2[\lambda_0(\mathbf{s}_i) + \lambda_1(\mathbf{s}_i)\{x_t - \theta(\mathbf{s}_i)\}]}\right) \right]^{1(\theta(\mathbf{s}_i) \leq x_t \leq x_\nu)} \\
&\quad \times f(\lambda_0|\{\beta, \boldsymbol{\lambda}_1, \boldsymbol{\eta}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha),
\end{aligned}$$

where $f(\lambda_0|\{\beta, \boldsymbol{\lambda}_1, \boldsymbol{\eta}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \sim \text{MVN}(\mathbb{E}_{k|j}, \mathbb{C}_{k|j})$, with $k = \{3\}$ and $j = \{1, 2, 4, 5\}$.

Here $\boldsymbol{\lambda}_j = \{\lambda_j(\mathbf{s}_1), \lambda_j(\mathbf{s}_2), \dots, \lambda_j(\mathbf{s}_m)\}^T$, $j = 0, 1$.

4. Metropolis Step for $\lambda_1(\mathbf{s}_i)$:

Each $\lambda_1(\mathbf{s}_i)$, $i = 1, \dots, m$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

$$\begin{aligned} f\{\lambda_1(\mathbf{s}_i) | \cdot\} &\propto f(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}) \times f(\boldsymbol{\phi} | \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \\ &\propto \prod_{t=1}^{\nu} \left[\text{N}\left(\beta_0(\mathbf{s}_i), e^{2\lambda_0(\mathbf{s}_i)}\right) \right]^{1(0 \leq x_t \leq \theta(\mathbf{s}_i))} \\ &\quad \times \left[\text{N}\left(\beta_0(\mathbf{s}_i) + \beta_1(\mathbf{s}_i) \{x_t - \theta(\mathbf{s}_i)\}, e^{2[\lambda_0(\mathbf{s}_i) + \lambda_1(\mathbf{s}_i)\{x_t - \theta(\mathbf{s}_i)\}]} \right) \right]^{1(\theta(\mathbf{s}_i) \leq x_t \leq x_\nu)} \\ &\quad \times f(\boldsymbol{\lambda}_1 | \{\boldsymbol{\beta}, \boldsymbol{\lambda}_0, \boldsymbol{\eta}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha), \end{aligned}$$

where $f(\boldsymbol{\lambda}_1 | \{\boldsymbol{\beta}, \boldsymbol{\lambda}_0, \boldsymbol{\eta}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \sim \text{MVN}(\mathbb{E}_{k|j}, \mathbb{C}_{k|j})$, with $k = \{4\}$ and $j = \{1, 2, 3, 5\}$.

5. Metropolis Step for $\eta(\mathbf{s}_i)$:

Each $\eta(\mathbf{s}_i)$, $i = 1, \dots, m$, is sampled using a Metropolis step given the following quantity that is proportional to the full conditional density,

$$\begin{aligned} f\{\eta(\mathbf{s}_i) | \cdot\} &\propto f(\mathbf{Y} | \boldsymbol{\beta}, \boldsymbol{\lambda}, \boldsymbol{\eta}) \times f(\boldsymbol{\phi} | \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \\ &\propto \prod_{t=1}^{\nu} \left[\text{N}\left(\beta_0(\mathbf{s}_i), e^{2\lambda_0(\mathbf{s}_i)}\right) \right]^{1(0 \leq x_t \leq \theta(\mathbf{s}_i))} \\ &\quad \times \left[\text{N}\left(\beta_0(\mathbf{s}_i) + \beta_1(\mathbf{s}_i) \{x_t - \theta(\mathbf{s}_i)\}, e^{2[\lambda_0(\mathbf{s}_i) + \lambda_1(\mathbf{s}_i)\{x_t - \theta(\mathbf{s}_i)\}]} \right) \right]^{1(\theta(\mathbf{s}_i) \leq x_t \leq x_\nu)} \\ &\quad \times f(\boldsymbol{\eta} | \{\boldsymbol{\beta}, \boldsymbol{\lambda}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha), \end{aligned}$$

where $f(\boldsymbol{\eta} | \{\boldsymbol{\beta}, \boldsymbol{\lambda}\}, \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \sim \text{MVN}(\mathbb{E}_{k|j}, \mathbb{C}_{k|j})$, with $k = \{5\}$ and $j = \{1, 2, 3, 4\}$.

6. Metropolis Step for α :

We transform α to the real line to facilitate sampling. Define a new parameter $\Delta = h(\alpha) = \log\left(\frac{\alpha - a_\alpha}{b_\alpha - \alpha}\right)$, such that $h^{-1}(\Delta) = (b_\alpha \exp\{\Delta\} + a_\alpha) / (1 + \exp\{\Delta\})$ and $\left|\frac{\partial}{\partial \Delta} h^{-1}(\Delta)\right| \propto \exp\{\Delta\} / (1 + \exp\{\Delta\})^2$. Now we can sample from the transformed proposal distribution, $\Delta^* \sim \text{N}(\Delta, \delta)$, where δ is a tuning parameter. Then we can obtain a proposal of α , $\alpha^* = h^{-1}(\Delta^*)$.

$$f(\alpha | \cdot) \propto f(\boldsymbol{\phi} | \boldsymbol{\delta}, \boldsymbol{\Sigma}, \alpha) \times f(\alpha) \times \left| \frac{\partial}{\partial \Delta} h^{-1}(\Delta) \right|.$$

7. Gibbs Step for Σ :

Define $\Phi_{5 \times n} = \{\beta_0, \beta_1, \lambda_0, \lambda_1, \eta\}^T$. Note that $\phi = \text{vec}(\Phi)$. Therefore, using the following properties of the Kronecker product and trace, $(B^T \otimes A)\text{vec}(X) = \text{vec}(AXB)$ and $\text{tr}(AB) = \text{vec}(A)^T \text{vec}(B)$ and defining $M = \delta \mathbf{1}_m^T$, we may rewrite the MCAR prior as follows,

$$\begin{aligned}
f(\phi | \delta, \Sigma, \alpha) &\propto \exp \left\{ -\frac{1}{2} (\phi - (\mathbf{1}_m \otimes \delta))^T (\mathbf{Q}(\alpha) \otimes \Sigma^{-1}) (\phi - (\mathbf{1}_m \otimes \delta)) \right\} \\
&= \exp \left\{ -\frac{1}{2} (\text{vec}(\Phi) - \text{vec}(\mathbf{M}))^T (\mathbf{Q}(\alpha) \otimes \Sigma^{-1}) (\text{vec}(\Phi) - \text{vec}(\mathbf{M})) \right\} \\
&= \exp \left\{ -\frac{1}{2} \text{vec}(\Phi - \mathbf{M})^T (\mathbf{Q}(\alpha) \otimes \Sigma^{-1}) \text{vec}(\Phi - \mathbf{M}) \right\} \\
&= \exp \left\{ -\frac{1}{2} \text{vec}(\Phi - \mathbf{M})^T \text{vec}(\Sigma^{-1}(\Phi - \mathbf{M})\mathbf{Q}(\alpha)) \right\} \\
&= \exp \left\{ -\frac{1}{2} \text{tr}((\Phi - \mathbf{M})^T \Sigma^{-1}(\Phi - \mathbf{M})\mathbf{Q}(\alpha)) \right\} \\
&= \exp \left\{ -\frac{1}{2} \text{tr}((\Phi - \mathbf{M})\mathbf{Q}(\alpha)(\Phi - \mathbf{M})^T \Sigma^{-1}) \right\} \\
&= \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{S}_{\Phi, \alpha} \Sigma^{-1}) \right\},
\end{aligned}$$

where $\mathbf{S}_{\Phi, \alpha} = (\Phi - \mathbf{M})\mathbf{Q}(\alpha)(\Phi - \mathbf{M})^T$. Now, the full conditional for Σ is straight forward,

$$\begin{aligned}
f(\Sigma | \cdot) &\propto f(\phi | \delta, \Sigma, \alpha) \times f(\Sigma) \\
&\propto |\mathbf{Q}(\alpha) \otimes \Sigma^{-1}|^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{S}_{\Phi, \alpha} \Sigma^{-1}) \right\} |\Sigma|^{-\frac{(\xi+p+1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\Psi \Sigma^{-1}) \right\} \\
&\propto (|\Sigma^{-1}|^m |\mathbf{Q}(\alpha)|^p)^{\frac{1}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\mathbf{S}_{\Phi, \alpha} \Sigma^{-1}) \right\} |\Sigma|^{-\frac{(\xi+p+1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr}(\Psi \Sigma^{-1}) \right\} \\
&\propto |\Sigma^{-1}|^{\frac{m}{2}} \exp \left\{ -\frac{1}{2} \text{tr}([\mathbf{S}_{\Phi, \alpha} + \Psi] \Sigma^{-1}) \right\} |\Sigma|^{-\frac{(\xi+p+1)}{2}} \\
&\propto |\Sigma|^{-\frac{m}{2}} \exp \left\{ -\frac{1}{2} \text{tr}([\mathbf{S}_{\Phi, \alpha} + \Psi] \Sigma^{-1}) \right\} |\Sigma|^{-\frac{(\xi+p+1)}{2}} \\
&\propto |\Sigma|^{-\frac{(m+\xi+p+1)}{2}} \exp \left\{ -\frac{1}{2} \text{tr}([\mathbf{S}_{\Phi, \alpha} + \Psi] \Sigma^{-1}) \right\} \\
&\sim \text{Inverse-Wishart}(m + \xi, \mathbf{S}_{\Phi, \alpha} + \Psi).
\end{aligned}$$

8. Gibbs Step for $Y_t(\mathbf{s}_i)$:

The latent process $Y_t(\mathbf{s}_i)$ is sampled from their full conditional distribution of,

$$f\{Y_t(\mathbf{s}_i) | Y_t^*(\mathbf{s}_i), \boldsymbol{\Omega}\} \propto f\{Y_t^*(\mathbf{s}_i) | Y_t(\mathbf{s}_i)\} \times f\{Y_t(\mathbf{s}_i) | \boldsymbol{\Omega}\},$$

where $\boldsymbol{\Omega}$ is a vector of the model parameters. Then, we know that there are two possibilities, $Y_t^*(\mathbf{s}_i) = 0$ or $Y_t^*(\mathbf{s}_i) = Y_t(\mathbf{s}_i)$. If $Y_t^*(\mathbf{s}_i) = 0$, then,

$$f\{Y_t^*(\mathbf{s}_i) | Y_t(\mathbf{s}_i)\} = P[Y_t^*(\mathbf{s}_i) = 0 | Y_t(\mathbf{s}_i)] = \begin{cases} 1 & Y_t(\mathbf{s}_i) \leq 0, \\ 0 & Y_t(\mathbf{s}_i) > 0. \end{cases}$$

Therefore, $f\{Y_t(\mathbf{s}_i) | Y_t^*(\mathbf{s}_i), \boldsymbol{\Omega}\} \propto 1\{Y_t(\mathbf{s}_i) \leq 0\} f\{Y_t(\mathbf{s}_i) | \boldsymbol{\Omega}\}$. Furthermore, when $Y_t^*(\mathbf{s}_i) = Y_t(\mathbf{s}_i)$, $f\{Y_t(\mathbf{s}_i) | Y_t^*(\mathbf{s}_i), \boldsymbol{\Omega}\} = f\{Y_t(\mathbf{s}_i) | Y_t(\mathbf{s}_i), \boldsymbol{\Omega}\} = Y_t(\mathbf{s}_i)$. Then, we see that the full conditional for the latent variable is,

$$f\{Y_t(\mathbf{s}_i) | Y_t^*(\mathbf{s}_i), \boldsymbol{\Omega}\} \sim \begin{cases} \text{TN}(\mu_t(\mathbf{s}_i), \sigma_t^2(\mathbf{s}_i); \leq 0) & Y_t^*(\mathbf{s}_i) = 0, \\ Y_t^*(\mathbf{s}_i) & Y_t^*(\mathbf{s}_i) = Y_t(\mathbf{s}_i). \end{cases}$$

where, $\text{TN}(\mu, \sigma^2; \leq 0)$ specifies a truncated normal distribution (truncated above by zero).

9. Repeat steps 1-8 until convergence has been achieved and an adequate number of posterior samples have been obtained post-convergence.

BIBLIOGRAPHY

- AGIS Investigators (2000). The Advanced Glaucoma Intervention Study (AGIS): 7. The relationship between control of intraocular pressure and visual field deterioration. *American Journal of Ophthalmology*, 130(4):429–40.
- Artes, P. H. and Chauhan, B. C. (2005). Longitudinal changes in the visual field and optic disc in glaucoma. *Progress in Retinal and Eye Research*, 24(3):333–354.
- Artes, P. H., Nicolela, M. T., LeBlanc, R. P., and Chauhan, B. C. (2005). Visual field progression in glaucoma: total versus pattern deviation analyses. *Investigative Ophthalmology & Visual Science*, 46(12):4600–4606.
- Artes, P. H., O’Leary, N., Hutchison, D. M., Heckler, L., Sharpe, G. P., Nicolela, M. T., and Chauhan, B. C. (2011). Properties of the statpac visual field index. *Investigative Ophthalmology & Visual Science*, 52(7):4030–4038.
- Bai, J. (1997). Estimation of a change point in multiple regression models. *Review of Economics and Statistics*, 79(4):551–563.
- Banerjee, S. and Gelfand, A. E. (2006). Bayesian wombling: Curvilinear gradient assessment under spatial process models. *Journal of the American Statistical Association*, 101(476):1487–1501.
- Banerjee, S., Gelfand, A. E., and Carlin, B. P. (2003). *Hierarchical Modeling and Analysis for Spatial Data*. CRC Press.
- Beckage, B., Joseph, L., Belisle, P., Wolfson, D. B., and Platt, W. J. (2007). Bayesian change-point analyses in ecology. *New Phytologist*, 174(2):456–467.
- Bengtsson, B. and Heijl, A. (2008). A visual field index for calculation of glaucoma rate of progression. *American Journal of Ophthalmology*, 145(2):343–353.
- Bengtsson, B., Patella, V. M., and Heijl, A. (2009). Prediction of glaucomatous visual field loss by extrapolation of linear trends. *Archives of Ophthalmology*, 127(12):1610–1615.
- Berchuck, S. I. (2017). womblr: Spatiotemporal Boundary Detection Model for Areal Unit Data. *R package version 1.0.0*.
- Bernardinelli, L., Clayton, D., and Montomoli, C. (1995). Bayesian estimates of disease maps: how important are priors? *Statistics in Medicine*, 14(21-22):2411–2431.
- Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society. Series B*, pages 192–236.
- Betz-Stablein, B. D., Morgan, W. H., House, P. H., and Hazelton, M. L. (2013). Spatial modeling of visual field data for assessing glaucoma progression. *Investigative Ophthalmology & Visual Science*, 54(2):1544–1553.
- Bhandari, A., Crabb, D. P., Poinosawmy, D., Fitzke, F. W., Hitchings, R. A., and Nouredin, B. N. (1997). Effect of surgery on visual field progression in normal-tension glaucoma. *Ophthalmology*, 104(7):1131–1137.

- Birch, M. K., Wishart, P. K., and O'Donnell, N. P. (1995). Determining progressive visual field loss in serial humphrey visual fields. *Ophthalmology*, 102(8):1227–1235.
- Bowman, F. D., Caffo, B., Bassett, S. S., and Kilts, C. (2008). A Bayesian hierarchical framework for spatial modeling of fMRI data. *NeuroImage*, 39(1):146–156.
- Brook, D. (1964). On the distinction between the conditional probability and the joint probability approaches in the specification of nearest-neighbour systems. *Biometrika*, pages 481–483.
- Brown, C. E. (1998). Coefficient of variation. In *Applied Multivariate Statistics in Geohydrology and Related Sciences*, pages 155–157. Springer.
- Bryan, S. R., Eilers, P. H., Lesaffre, E. M., Lemij, H. G., and Vermeer, K. A. (2015). Global visit effects in point-wise longitudinal modeling of glaucomatous visual fields. *Investigative Ophthalmology & Visual Science*, 56(8):4283–4289.
- Bryan, S. R., Eilers, P. H., Rosmalen, J. v., Rizopoulos, D., Vermeer, K. A., Lemij, H. G., and Lesaffre, E. M. (2017). Bayesian hierarchical modeling of longitudinal glaucomatous visual fields using a two-stage approach. *Statistics in Medicine*, 36(11):1735–1753.
- Bryan, S. R., Vermeer, K. A., Eilers, P. H., Lemij, H. G., and Lesaffre, E. M. (2013). Robust and censored modeling and prediction of progression in glaucomatous visual fields. *Investigative Ophthalmology & Visual Science*, 54(10):6694–6700.
- Cai, B., Lawson, A. B., McDermott, S., and Aelion, C. M. (2016). A bayesian semiparametric approach with change points for spatial ordinal data. *Statistical Methods in Medical Research*, 25(2):644–658.
- Carlin, B. P. and Banerjee, S. (2003). Hierarchical multivariate car models for spatio-temporally correlated survival data. In Bernardo, J., Bayarri, M., Berger, J., Dawid, A., Heckerman, D., Smith, A., and West, M., editors, *Bayesian Statistics 7: Proceedings of the Seventh Valencia International Meeting*, page 45.
- Carlin, B. P., Gelfand, A. E., and Smith, A. F. (1992). Hierarchical Bayesian Analysis of Changepoint Problems. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 41(2):389–405.
- Casas-Llera, P., Rebolleda, G., Muñoz-Negrete, F. J., Arnalich-Montiel, F., Pérez-López, M., and Fernández-Buenaga, R. (2009). Visual field index rate and event-based glaucoma progression analysis: comparison in a glaucoma population. *British Journal of Ophthalmology*, 93(12):1576–1579.
- Castruccio, S., Ombao, H., and Genton, M. G. (2016). A multi-resolution spatio-temporal model for brain activation and connectivity in fMRI data. *arXiv preprint arXiv:1602.02435*.
- Chauhan, B. C., Drance, S. M., and Douglas, G. R. (1990). The use of visual field indices in detecting changes in the visual field in glaucoma. *Investigative Ophthalmology & Visual Science*, 31(3):512–520.
- Chauhan, B. C., Garway-Heath, D. F., Goñi, F. J., Rossetti, L., Bengtsson, B., Viswanathan, A. C., and Heijl, A. (2008). Practical recommendations for measuring rates of visual field change in glaucoma. *British Journal of Ophthalmology*, 92(4):569–573.

- Chauhan, B. C., McCormick, T. A., Nicolela, M. T., and LeBlanc, R. P. (2001). Optic disc and visual field changes in a prospective longitudinal study of patients with glaucoma: comparison of scanning laser tomography with conventional perimetry and optic disc photography. *Archives of Ophthalmology*, 119(10):1492–1499.
- Chauhan, B. C., Mikelberg, F. S., Artes, P. H., Balazsi, A. G., LeBlanc, R. P., Lesk, M. R., Nicolela, M. T., and Trope, G. E. (2010). Canadian glaucoma study: 3. impact of risk factors and intraocular pressure reduction on the rates of visual field change. *Archives of Ophthalmology*, 128(10):1249–1255.
- Cleveland, W. S. and Devlin, S. J. (1988). Locally weighted regression: an approach to regression analysis by local fitting. *Journal of the American Statistical Association*, 83(403):596–610.
- Clifford, P. (1990). Markov random fields in statistics. *Disorder in Physical Systems: A Volume in Honour of John M. Hammersley*, pages 19–32.
- Cook, C. and Foster, P. (2012). Epidemiology of glaucoma: what’s new? *Canadian Journal of Ophthalmology/Journal Canadien d’Ophtalmologie*, 47(3):223–226.
- Crabb, D. P., Fitzke, F. W., McNaught, A. I., Edgar, D. F., and Hitchings, R. A. (1997). Improving the prediction of visual field progression in glaucoma using spatial processing. *Ophthalmology*, 104(3):517–524.
- Cressie, N. (2015). *Statistics for Spatial Data*. John Wiley & Sons.
- Davson, H. (2012). *Physiology of the Eye*. Elsevier.
- De Moraes, C. G., Liebmann, J. M., Greenfield, D. S., Gardiner, S. K., Ritch, R., and Krupin, T. (2012). Risk factors for visual field progression in the low-pressure glaucoma treatment study. *American Journal of Ophthalmology*, 154(4):702–711.
- Diggle, P. J., Ribeiro Jr, P. J., and Christensen, O. F. (2003). An introduction to model-based geostatistics. In *Spatial Statistics and Computational Methods*, pages 43–86. Springer.
- Dodd, L. E. and Pepe, M. S. (2003). Partial auc estimation and regression. *Biometrics*, 59(3):614–623.
- Ecker, M. D. and Gelfand, A. E. (2003). Spatial modeling and prediction under stationary non-geometric range anisotropy. *Environmental and Ecological Statistics*, 10(2):165–178.
- Eddelbuettel, D., François, R., Allaire, J., Chambers, J., Bates, D., and Ushey, K. (2011). Rcpp: Seamless R and C++ integration. *Journal of Statistical Software*, 40(8):1–18.
- Ernest, P. J., Viechtbauer, W., Schouten, J. S., Beckers, H. J., Hendrikse, F., Prins, M. H., and Webers, C. A. (2012). The influence of the assessment method on the incidence of visual field progression in glaucoma: a network meta-analysis. *Archives of Ophthalmology*, 90(1):10–19.
- Fitzke, F. W., Crabb, D. P., McNaught, A. I., Edgar, D. F., and Hitchings, R. A. (1995). Image processing of computerised visual field data. *British Journal of Ophthalmology*, 79(3):207–212.
- Fitzke, F. W., Hitchings, R. A., Poinosawmy, D., McNaught, A. I., and Crabb, D. P. (1996). Analysis of visual field progression in glaucoma. *British Journal of Ophthalmology*, 80(1):40–48.
- Foster, P. J., Buhrmann, R., Quigley, H. A., and Johnson, G. J. (2002). The definition and classification of glaucoma in prevalence surveys. *British Journal of Ophthalmology*, 86(2):238–242.

- Gardiner, S. K. and Crabb, D. P. (2002). Examination of different pointwise linear regression methods for determining visual field progression. *Investigative Ophthalmology & Visual Science*, 43(5):1400–1407.
- Gardiner, S. K., Crabb, D. P., Fitzke, F. W., and Hitchings, R. A. (2004). Reducing noise in suspected glaucomatous visual fields by using a new spatial filter. *Vision Research*, 44(8):839–848.
- Garway-Heath, D. F., Crabb, D. P., Bunce, C., Lascaratos, G., Amalfitano, F., Anand, N., Azuara-Blanco, A., Bourne, R. R., Broadway, D. C., Cunliffe, I. A., et al. (2015). Latanoprost for open-angle glaucoma (ukgts): a randomised, multicentre, placebo-controlled trial. *The Lancet*, 385(9975):1295–1304.
- Garway-Heath, D. F., Poinoosawmy, D., Fitzke, F. W., and Hitchings, R. A. (2000). Mapping the visual field to the optic disc in normal tension glaucoma eyes. *Ophthalmology*, 107(10):1809–1815.
- Gelfand, A. E., Diggle, P., Guttorp, P., and Fuentes, M. (2010). *Handbook of Spatial Statistics*. CRC Press.
- Gelfand, A. E. and Smith, A. F. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85(410):398–409.
- Gelfand, A. E. and Vounatsou, P. (2003). Proper multivariate conditional autoregressive models for spatial data analysis. *Biostatistics*, 4(1):11–15.
- Gelman, A., Carlin, J. B., Stern, H. S., and Rubin, D. B. (2014). *Bayesian Data Analysis*, volume 2. Chapman & Hall/CRC Boca Raton, FL, USA.
- Gelman, A. and Shalizi, C. R. (2013). Philosophy and the practice of Bayesian statistics. *British Journal of Mathematical and Statistical Psychology*, 66(1):8–38.
- Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, PAMI-6(6):721–741.
- Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to calculating posterior moments. In Bernardo, J. M., Berger, J., Dawid, A. P., and Smith, J. F. M., editors, *Bayesian Statistics 4*, pages 169–193. Oxford University Press, Oxford.
- Goldbaum, M. H., Sample, P. A., Zhang, Z., Chan, K., Hao, J., Lee, T.-W., Boden, C., Bowd, C., Bourne, R., Zangwill, L., et al. (2005). Using unsupervised learning with independent component analysis to identify patterns of glaucomatous visual field defects. *Investigative Ophthalmology & Visual Science*, 46(10):3676.
- Gordon, M. O., Beiser, J. A., Brandt, J. D., Heuer, D. K., Higginbotham, E. J., Johnson, C. A., Keltner, J. L., Miller, J. P., Parrish, R. K., Wilson, M. R., et al. (2002). The ocular hypertension treatment study: baseline factors that predict the onset of primary open-angle glaucoma. *Archives of Ophthalmology*, 120(6):714–720.
- He, Y., Chen, Z., and Evans, A. (2008). Structural insights into aberrant topological patterns of large-scale cortical networks in Alzheimer’s disease. *Journal of Neuroscience*, 28(18):4756–4766.

- He, Y., Dagher, A., Chen, Z., Charil, A., Zijdenbos, A., Worsley, K., and Evans, A. (2009). Impaired small-world efficiency in structural cortical networks in multiple sclerosis associated with white matter lesion load. *Brain*, 132(12):3366–3379.
- Heijl, A. (1991). Extended empirical statistical package for evaluation of single and multiple fields in glaucoma: Statpac 2. *Proceedings of the Ninth International Perimetric Society Meeting Amsterdam, NY Kugler*, pages 303–315.
- Heijl, A., Leske, M. C., Bengtsson, B., Bengtsson, B., and Hussein, M. (2003). Measuring visual field progression in the early manifest glaucoma trial. *Acta Ophthalmologica Scandinavica*, 81(3):286–293.
- Heijl, A., Leske, M. C., Bengtsson, B., Hyman, L., Bengtsson, B., and Hussein, M. (2002). Reduction of intraocular pressure and glaucoma progression: results from the Early Manifest Glaucoma Trial. *Archives of Ophthalmology*, 120(10):1268–1279.
- Heijl, A., Lindgren, G., and Olsson, J. (1987). A package for the statistical analysis of visual fields. In *Seventh International Visual Field Symposium, Amsterdam, September 1986*, pages 153–168. Springer.
- Heijl, A., Patella, V. M., and Bengtsson, B. (2012). *The field analyzer primer: effective perimetry*. Carl Zeiss Meditec Incorporated.
- Hersbach, H. (2000). Decomposition of the continuous ranked probability score for ensemble prediction systems. *Weather and Forecasting*, 15(5):559–570.
- Hinkley, D. V. (1970). Inference about the change-point in a sequence of random variables. *Biometrika*, pages 1–17.
- Illian, J., Penttinen, A., Stoyan, H., and Stoyan, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*, volume 70. John Wiley & Sons.
- Jacquez, G. M. and Greiling, D. A. (2003). Geographic boundaries in breast, lung and colorectal cancers in relation to exposure to air toxics in Long Island, New York. *International Journal of Health Geographics*, 2:4.
- Jampel, H. D., Singh, K., Lin, S. C., Chen, T. C., Francis, B. A., Hodapp, E., Samples, J. R., and Smith, S. D. (2011). Assessment of visual function in glaucoma: a report by the american academy of ophthalmology. *Ophthalmology*, 118(5):986–1002.
- Jay, J. and Murdoch, J. (1993). The rate of visual field loss in untreated primary open angle glaucoma. *British Journal of Ophthalmology*, 77(3):176–178.
- Jin, X., Banerjee, S., and Carlin, B. P. (2007). Order-free co-regionalized areal data models with application to multiple-disease mapping. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69(5):817–838.
- Jin, X., Carlin, B. P., and Banerjee, S. (2005). Generalized hierarchical multivariate CAR models for areal data. *Biometrics*, 61(4):950–961.
- Johnson, C. A., Adams, A. J., Casson, E. J., and Brandt, J. D. (1993). Progression of early glaucomatous visual field loss as detected by blue-on-yellow and standard white-on-white automated perimetry. *Archives of Ophthalmology*, 111(5):651–656.

- Julious, S. A. (2001). Inference and estimation in a changepoint regression problem. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 50(1):51–61.
- Katz, J., Congdon, N., and Friedman, D. S. (1999). Methodological variations in estimating apparent progressive visual field loss in clinical trials of glaucoma treatment. *Archives of Ophthalmology*, 117(9):1137–1142.
- Katz, J., Gilbert, D., Quigley, H. A., and Sommer, A. (1997). Estimating progression of visual field loss in glaucoma. *Ophthalmology*, 104(6):1017–1025.
- Klein, B. E., Klein, R., Sponsel, W. E., Franke, T., Cantor, L. B., Martone, J., and Menage, M. J. (1992). Prevalence of glaucoma: the beaver dam eye study. *Ophthalmology*, 99(10):1499–1504.
- Konrad, K. and Eickhoff, S. B. (2010). Is the ADHD brain wired differently? A review on structural and functional connectivity in attention deficit hyperactivity disorder. *Human Brain Mapping*, 31(6):904–916.
- Lee, D. and Mitchell, R. (2011). Boundary detection in disease mapping studies. *Biostatistics*, 15(3):457–469.
- Lee, D. and Mitchell, R. (2013). Locally adaptive spatial smoothing using conditional auto-regressive models. *Journal of the Royal Statistical Society: Series C*, 62(4):593–608.
- Lee, D. and Mitchell, R. (2014). Controlling for localised spatio-temporal autocorrelation in long-term air pollution and health studies. *Statistical Methods in Medical Research*, 23(6):488–506.
- Lee, D., Rushworth, A., and Sahu, S. K. (2014). A Bayesian localized conditional autoregressive model for estimating the health effects of air pollution. *Biometrics*, 70(2):419–429.
- Leroux, B. G., Lei, X., and Breslow, N. (2000). Estimation of disease rates in small areas: A new mixed model for spatial dependence. In *Statistical Models in Epidemiology, the Environment, and Clinical Trials*, pages 179–191. Springer.
- Leske, M. C., Heijl, A., Hyman, L., Bengtsson, B., Dong, L., and Yang, Z. (2007). Predictors of long-term progression in the early manifest glaucoma trial. *Ophthalmology*, 114(11):1965–1972.
- Li, P., Banerjee, S., Hanson, T. A., and McBean, A. M. (2015). Bayesian models for detecting difference boundaries in areal data. *Statistica Sinica*, pages 385–402.
- Li, P., Banerjee, S., and McBean, A. M. (2011). Mining boundary effects in areally referenced spatial data using the Bayesian information criterion. *Geoinformatica*, 15(3):435–454.
- Liu, Y., Zou, C., and Zhang, R. (2008). Empirical likelihood ratio test for a change-point in linear regression model. *Communications in Statistics-Theory and Methods*, 37(16):2551–2563.
- Lu, H. and Carlin, B. P. (2005). Bayesian areal wombling for geographical boundary analysis. *Geographical Analysis*, 37(3):265–285.
- Lu, H., Reilly, C. S., Banerjee, S., and Carlin, B. P. (2007). Bayesian areal wombling via adjacency modeling. *Environmental and Ecological Statistics*, 14(4):433–452.
- Ma, H. and Carlin, B. P. (2007). Bayesian multivariate areal wombling for multiple disease boundary analysis. *Bayesian Analysis*, 2(2):281–302.

- Ma, H., Carlin, B. P., and Banerjee, S. (2010). Hierarchical and Joint Site-Edge Methods for Medicare Hospice Service Region Boundary Analysis. *Biometrics*, 66(2):355–364.
- MacNab, Y. C. and Gustafson, P. (2007). Regression b-spline smoothing in bayesian disease mapping: with an application to patient safety surveillance. *Statistics in Medicine*, 26(24):4455–4474.
- Majumdar, A., Gelfand, A. E., and Banerjee, S. (2005). Spatio-temporal change-point modeling. *Journal of Statistical Planning and Inference*, 130(1-2):149–166.
- Mardia, K. (1988). Multi-dimensional multivariate Gaussian Markov random fields with application to image processing. *Journal of Multivariate Analysis*, 24(2):265–284.
- Medeiros, F. A. and Tatham, A. J. (2016). Structure versus function in glaucoma: the debate that doesn't need to be. *Ophthalmology*, 123(6):1170–1172.
- Medeiros, F. A., Zangwill, L. M., Alencar, L. M., Bowd, C., Sample, P. A., Susanna Jr, R., and Weinreb, R. N. (2009). Detection of glaucoma progression with stratus oct retinal nerve fiber layer, optic nerve head, and macular thickness measurements. *Investigative Ophthalmology & Visual Science*, 50(12):5741.
- Menzefricke, U. (1981). A Bayesian Analysis of a Change in the Precision of a Sequence of Independent Normal Random Variables at an Unknown Time Point. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 30(2):141–146.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, 21(6):1087–1092.
- Musch, D. C., Gillespie, B. W., Lichter, P. R., Niziol, L. M., and Janz, N. K. (2009). Visual field progression in the collaborative initial glaucoma treatment study: the impact of treatment and other baseline factors. *Ophthalmology*, 116(2):200–207.
- Nicolela, M. T., McCormick, T. A., Drance, S. M., Ferrier, S. N., LeBlanc, R. P., and Chauhan, B. C. (2003). Visual field and optic disc progression in patients with different types of optic disc damage: a longitudinal prospective study. *Ophthalmology*, 110(11):2178–2184.
- Nouri-Mahdavi, K. and Caprioli, J. (2014). Measuring rates of structural and functional change in glaucoma. *British Journal of Ophthalmology*, pages bjophthalmol–2014.
- Page, E. S. (1954). Continuous Inspection Schemes. *Biometrika*, 41(1):100–115.
- Pepe, M. S. (2003). *The statistical evaluation of medical tests for classification and prediction*. Oxford University Press.
- Perreault, L., Bernier, J., Bobée, B., and Parent, E. (2000). Bayesian change-point analysis in hydrometeorological time series. Part 1. The normal model revisited. *Journal of Hydrology*, 235(3-4):221–241.
- Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association*, 53(284):873–880.
- Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association*, 55(290):324–330.

- Quigley, H. A., Katz, J., Derick, R. J., Gilbert, D., and Sommer, A. (1992). An evaluation of optic disc and nerve fiber layer examinations in monitoring progression of early glaucoma damage. *Ophthalmology*, 99(1):19–28.
- R Core Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Raftery, A. E. (1994). Change point and change curve modeling in stochastic processes and spatial statistics. *Journal of Applied Statistical Science*, 1(4):403–423.
- Rushworth, A., Lee, D., and Sarran, C. (2017). An adaptive spatiotemporal smoothing model for estimating trends and step changes in disease risk. *Journal of the Royal Statistical Society: Series C*, 66(1):141–157.
- Russell, R. A., Crabb, D. P., Malik, R., and Garway-Heath, D. F. (2012). The relationship between variability and sensitivity in large-scale longitudinal visual field data. *Investigative Ophthalmology & Visual Science*, 53(10):5985–5990.
- Sain, S. R., Furrer, R., and Cressie, N. (2011). A spatial analysis of multivariate output from regional climate models. *The Annals of Applied Statistics*, pages 150–175.
- Sampson, P. D. (2010). Constructions for nonstationary spatial processes. In Gelfand, A. E., Diggle, P. J., Fuentes, M., and Guttorp, P., editor, *The Handbook of Spatial Statistics*, chapter 9, pages 119–130. CRC Press, Boca Raton, FL.
- Sherman, M. (2011). *Spatial Statistics and Spatio-temporal Data: Covariance Functions and Directional Properties*. John Wiley & Sons.
- Smith, S. D., Katz, J., and Quigley, H. A. (1996). Analysis of progressive change in automated visual fields in glaucoma. *Investigative Ophthalmology & Visual Science*, 37(7):1419.
- Smythies, J. (1996). A note on the concept of the visual field in neurology, psychology, and visual neuroscience. *Perception*, 25:369–372.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P., and Van Der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B*, 64(4):583–639.
- Spry, P. G. and Johnson, C. A. (2002). Identification of progressive glaucomatous visual field loss. *Survey of Ophthalmology*, 47(2):158–173.
- Strouthidis, N. G., Scott, A., Viswanathan, A. C., Crabb, D. P., and Garway-Heath, D. F. (2007). Monitoring glaucomatous visual field progression: the effect of a novel spatial filter. *Investigative Ophthalmology & Visual Science*, 48(1):251–257.
- Tanner, M. (1996). Tools for statistical inference: methods for the exploration of posterior distributions and likelihood functions. Springer-Verlag. *New York*.
- Tham, Y.-C., Li, X., Wong, T. Y., Quigley, H. A., Aung, T., and Cheng, C.-Y. (2014). Global prevalence of glaucoma and projections of glaucoma burden through 2040: a systematic review and meta-analysis. *Ophthalmology*, 121(11):2081–2090.
- Thompson, P. M., Hayashi, K. M., Sowell, E. R., Gogtay, N., Giedd, J. N., Rapoport, J. L., de Zubicaray, G. I., Janke, A. L., Rose, S. E., Semple, J., et al. (2004). Mapping cortical change in Alzheimer’s disease, brain development, and schizophrenia. *NeuroImage*, 23:S2–S18.

- Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica*, 26(1):24–36.
- Tucker, A., Vinciotti, V., Liu, X., and Garway-Heath, D. (2005). A spatio-temporal bayesian network classifier for understanding visual field deterioration. *Artificial Intelligence in Medicine*, 34(2):163–177.
- Vesti, E., Johnson, C. A., and Chauhan, B. C. (2003). Comparison of different methods for detecting glaucomatous visual field progression. *Investigative Ophthalmology & Visual Science*, 44(9):3873–3879.
- Vianna, J. R. and Chauhan, B. C. (2015). How to detect progression in glaucoma. *Progress in Brain Research*, 221:135–158.
- Wagner, T. and Midway, S. R. (2014). Modeling spatially varying landscape change points in species occurrence thresholds. *Ecosphere*, 5(11):1–16.
- Walsh, T. (2010). *Visual fields: examination and interpretation*, volume 3. Oxford University Press.
- Warren, J. L., Mwanza, J.-C., Tanna, A. P., and Budenz, D. L. (2016). A Statistical Model to Analyze Clinician Expert Consensus on Glaucoma Progression using Spatially Correlated Visual Field Data. *Translational Vision Science & Technology*, 5(4):14–14.
- Warren, J. L., Pingali, S. C., and Weinberger, D. M. (2017). Spatial variability in the persistence of pneumococcal conjugate vaccine-targeted pneumococcal serotypes among adults. *Epidemiology*, 28(1):119.
- Weinreb, R. N. and Khaw, P. T. (2004). Primary open-angle glaucoma. *The Lancet*, 363(9422):1711–1720.
- Wesselink, C., Heeg, G. P., and Jansonius, N. M. (2009). Glaucoma monitoring in a clinical setting: glaucoma progression analysis vs nonparametric progression analysis in the groningen longitudinal glaucoma study. *Archives of Ophthalmology*, 127(3):270–274.
- West, M. (1996). *Bayesian Forecasting*. Wiley Online Library.
- Womble, W. H. (1951). Differential systematics. *Science*, 114(2961):315–322.
- Yousefi, S., Goldbaum, M. H., Balasubramanian, M., Jung, T.-P., Weinreb, R. N., Medeiros, F., Zangwill, L. M., Liebmann, J. M., Girkin, C., Bowd, C., et al. (2014). Glaucoma progression detection using structural retinal nerve fiber layer measurements and functional visual field points. *IEEE Transactions on Biomedical Engineering*, 61(4):1143–1154.
- Yu, Q., Scribner, R., Carlin, B., Theall, K., Simonsen, N., Ghosh-Dastidar, B., Cohen, D., and Mason, K. (2008). Multilevel spatio-temporal dual changepoint models for relating alcohol outlet destruction and changes in neighbourhood rates of assaultive violence. *Geospatial Health*, 2(2):161.
- Zhu, H., Crabb, D. P., Ho, T., and Garway-Heath, D. F. (2015). More accurate modeling of visual field progression in glaucoma: ANSWERS. *Investigative Ophthalmology & Visual Science*, 56(10):6077–6083.
- Zhu, H., Russell, R. A., Saunders, L. J., Ceccon, S., Garway-Heath, D. F., and Crabb, D. P. (2014). Detecting changes in retinal function: analysis with non-stationary weibull error regression and spatial enhancement (ANSWERS). *PloS One*, 9(1):1–11.