Analysis of Real Data

This is a step by step guide on how to generate all the results shown in the paper. You can run the default analysis and it will generate the plots shown in the paper or you can run the analysis on the pseudo data and generate the plots in the end of Step 4 of this report. All analyses and plots are done using the Matlab scripts described below. To fit the functional mixed models described in the manuscript, we call R-functions from Matlab. These functions are only available for Windows computers. An R-package that will work in any operational systems is under development.

1. Step 1 - Download and install the required software

   (a) Follow the instructions on wfmm_install.pdf file.

   (b) Open your uptodate version of R and install the packages: R.matlab, coda, and nlme.

   (c) EYE_Toolbox: Contains all the scripts and functions to run the analysis.

2. Step 2 - Run MainScript.m

   Open your Matlab and make sure that your current working directory is EYE_toolbox. Then open MainScript.m. Edit your R path in the first line of the code since the file calls an R script to perform the variable selection heuristics and also to run the MCMC for the basis model. If you wish to run the pseudo data, please comment lines 14 and 25 and un-comment lines 15 and 26. Finally, running MainScript.m will perform the following tasks:

   (a) From the data matrix Y_outlier_removed (or Y_simulated), obtain the wavelet basis transform and perform compression

   (b) Use the compressed basis to perform the model selection heuristics

   (c) Based on the model selection results and the basis obtained in (a), run the Bayesian model and obtain the MCMC samples for the fixed effects and variance components in the basis space. The priors on the variance components and the smoothing parameters are automatically estimated using the empirical Bayes method explained in the paper.

   The script will generate the following files and folders.

   - The file DCompressed.mat, containing the compressed basis and the settings information used to obtain the basis
   - The files B.txt, Omega.txt containing the B spline basis evaluation and penalty matrix for computing the degrees of freedom (DF)
   - The files Zg_spline.txt, and Zdg_spline.txt consisting of the random effects design matrices for evaluation of the nonparametric function and its derivative

- The file model_selection_output.txt containing the results for the model selection heuristics
- A folder named eye_filtered_wave_fmm that contains the MCMC results.

Note: The variable selection heuristics takes 20-25 minutes to complete. After the eye_filtered_wave_fmm folder is created, you can follow the MCMC progress by looking at the log file in the folder.

3. Step 3 - Run Postprocessing.m

   This script is divided in 4 main parts as follows.

   (a) Part 1: Computes MCMC samples of the random effects corresponding to the Demmler-Reinsch spline bases in the basis space and in the data space (U^* and U), MCMC samples of the fixed effect coefficients in the data space, MCMC samples of the nonparametric age effect in the basis space and data space, data space projection of the posterior mean of the variance components. The script will generate the following files.

   - MCMC_beta.mat: MCMC samples of the fixed effect coefficients in the basis space
   - MCMC_theta.mat: MCMC samples of the variance components in the basis space
   - MCMC_U_all.mat: MCMC samples of U^* and U as described above
   - MCMC_g_low_pass.mat: MCMC samples of the fixed effect coefficients in the data space
   - nonparametric_age_MCMC.mat: MCMC samples of the nonparametric age effect in the data space
   - MCMC_fstar.mat: MCMC samples of the nonparametric age effect in the basis space
   - MCMC_sig_mean.mat: data space projection of the posterior mean of the variance components

   (b) Part 2: Compute the degrees of freedom of the nonparametric age fit, the fitted MPS at various values of age and IOP, the AUC of the fitted MPS aggregated over all IOP values using the trapezoidal rule, and the derivative of the AUC with respect to age.

   - df_mean.mat: the degrees of freedom of the nonparametric age fit
   - fhat_summary_all.mat: this includes (1) the fitted MPS at various values of age and IOP and (2) the AUC of the fitted MPS
   - fdhat_summary_all.mat: the derivative of the AUC with respect to age

   (c) Part 3: Compute aggregated summaries of the fitted MPS and AUC over various functional regions and derivatives of the aggregated summaries

   - fhat_circum_mean.mat: aggregated summaries of the fitted MPS and AUC over circumferential regions

- fhat_summary_all_region_section.mat: aggregated summaries of the fitted MPS and AUC for each region [peripapillary(PP) region or mid-peripheral(MP) region] and each of 8 sections in the spherical domain
- fdhat_summary_all_region_section.mat: the derivative of the aggregated summaries for each region and each section
- region_section_results.mat: aggregated summaries of the raw MPS data for each region and each section

(d) Part 4: Compute the intrafunctional covariance structures induced by tensor wavelet bases

- InducedCovariances.mat: This includes (1) the diagonal elements of the induced covariance matrices in the data space and (2) variance components in the basis space

Note: The script might take 5 hours or more to run. You can follow the progress of each part as the files are saved. Commented in the script, you can see the length of the longest runs. Part 1 and 2 take most of the running time.

4. Step 4 - Run Plotting.m

This script uses the quantities computed in Step 3 to generate all the plots presented in the paper. It automatically creates a directory called 'Plots' with 7 sub-directories described below. An example of the plots contained in each directory is shown in the next page.

(a) age_by_pressure_with_df: contains the projection of the fitted MPS function, the posterior mean of degrees of freedom of the nonparametric MPS fit of age, and the estimated nonparametric MPS fit of age for all nine IOP levels as described in Figure 2 in the manuscript. Plots, one for each scleral location, are combined to generate the movie MPSbyAge.mp4.

(b) Combo_plots: contains the key summaries of the fitted model as described in Figure 3 in the manuscript. Plots, one for each scleral location, are combined to generate the movie combo_plot.mp4

(c) auc_circumferential: contains the posterior mean AUC (of MPS) as a function of age and distance from ONH as described in Figure 4 of the manuscript.

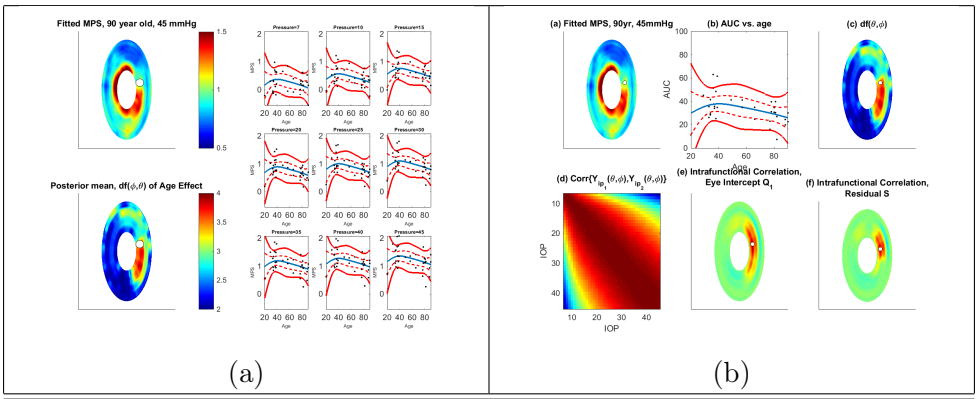
(d) AUC_dAUC_by_age_by_region: contains the aggregated AUC summaries as described in Figure 5 of the manuscript.

(e) Intra_corr: contains the intrafunctional correlation surface induced by our model and choice of tensor basis for the eye-to-eye random intercept and residual error as described in Figure 3 of the manuscript (panels (e) and (f)). Plots, one for each scleral location, are combined to generate the movie intrafunctional_correlation.mp4

(f) Intra_IOP: contains the serial correlation across IOP induced by the model. Plots, one for each scleral location, are combined to generate the movie Intra_IOP_corr.mp4.

(g) AUCvsAGE: contains the nonparametric MPS vs. age curve using the AUC to integrate over IOP as described in Figure 3 panel (b) of the manuscript. Plots, one for each scleral location, are combined to generate the movie AUCvsAGE_wave.mp4.

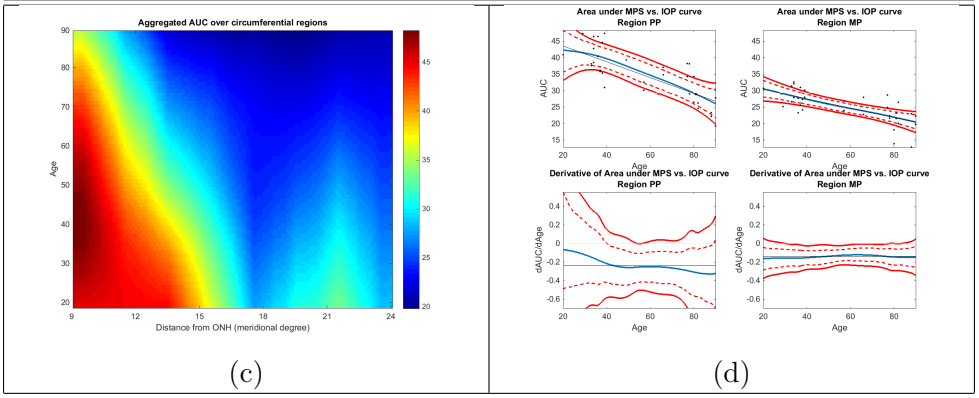
Note: For any of the plots generated for multiple scleral positions, we created the .mp4 movie files by starting a Microsoft Movie Maker session, dragging in the set of .png files from the appropriate directory, choosing "edit" to make the time for each slide 0.2 seconds, and then saving as .mp4 file "for computer". These steps can be followed to obtain the movie files for the folders "age_by_pressure_with_df", "AUCvsAG", "Combo_plots", "Intra_cor", and "Intra_IOP", if desired.


Fitted MPS, 90 year old, 45 mmHg
Posterior mean, df(ϕ,θ) of Age Effect
Pressure=7
Pressure=10
Pressure=15
Pressure=20
Pressure=25
Pressure=30
Pressure=35
Pressure=40
Pressure=45
MPS
Age

(a)

(a) Fitted MPS, 90yr, 45mmHg
(b) AUC vs. age
(c) df(θ,ϕ)
(d) Corr{$Y_{IP_1}(\theta,\phi), Y_{IP_2}(\theta,\phi)$}
(e) Intrafunctional Correlation, Eye Intercept Q_1
(f) Intrafunctional Correlation, Residual S
AUC
Age
IOP

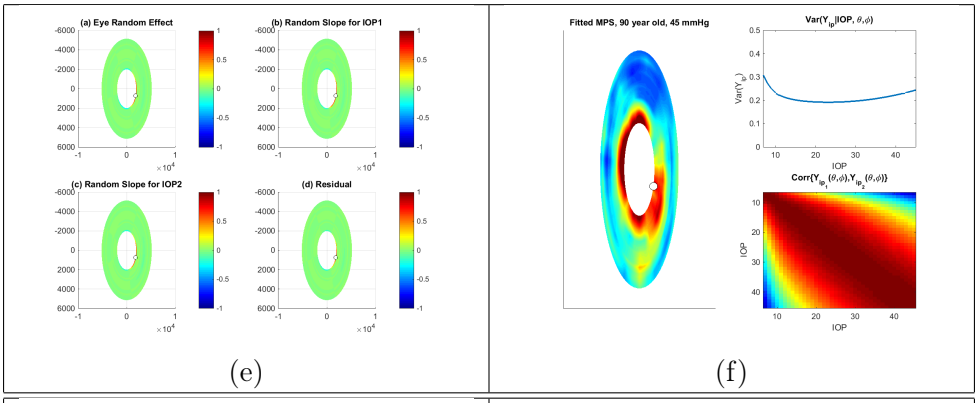
(b)


Aggregated AUC over circumferential regions
Age
Distance from ONH (meridional degree)

(c)

Area under MPS vs. IOP curve Region PP
Area under MPS vs. IOP curve Region MP
Derivative of Area under MPS vs. IOP curve Region PP
Derivative of Area under MPS vs. IOP curve Region MP
AUC
dAUC/dAge
Age

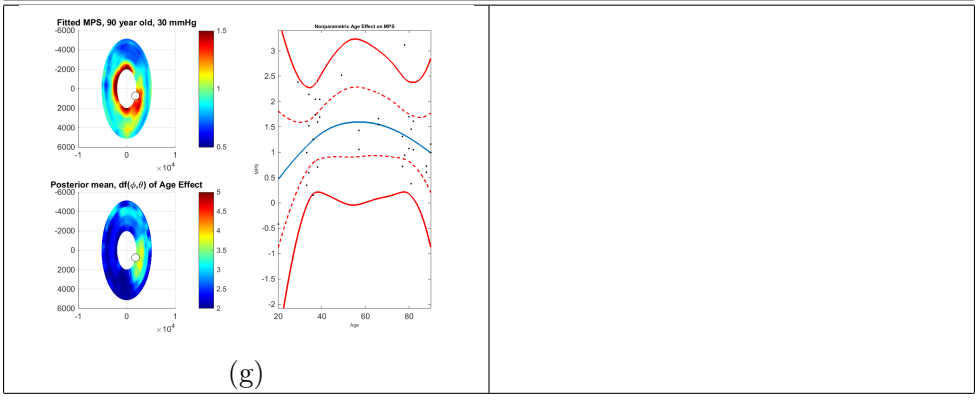
(d)


(a) Eye Random Effect
(b) Random Slope for IOP1
(c) Random Slope for IOP2
(d) Residual

(e)

Fitted MPS, 90 year old, 45 mmHg
Var(Y_{IP}|IOP, θ,ϕ)
Corr{$Y_{IP_1}(\theta,\phi), Y_{IP_2}(\theta,\phi)$}
IOP

(f)


Fitted MPS, 90 year old, 30 mmHg
Posterior mean, df(ϕ,θ) of Age Effect
Nonparametric Age Effect on MPS
MPS
Age

(g)

5. Step 5 - Run Diagnostics.m

   Similar to Step 2, you will need to edit your R path in the first line of the code. This script calls R and uses the package 'coda' to compute the Geweke statistics. Next, it obtains the results shown in the table below. It generates the file DiagnosticsSummary.mat, containing the variable TableSummary, a 14×4 matrix, in exactly the same order as shown in the blue part of the table. It also contains the variable MH_AcceptanceProbabilities, a vector in the same order as shown in the pink part of the table below.

Pseudo Data		Fixed Effects wo/ age	Age nonparametric	Variance Components	Combined
		Basis Space	Basis Space	Basis Space	Basis Space
P-value	Mean	0.5283	0.5095	0.4811	0.5053
	Q025	0.0202	0.0120	0.0119	0.0123
	Q05	0.0507	0.0238	0.0290	0.0263
	Median	0.5580	0.5147	0.4771	0.5089
	Q95	0.9607	0.9599	0.9350	0.9562
	Q975	0.9757	0.9778	0.9724	0.9771
Geweke	Mean	0.0171	0.0024	0.0728	0.0174
	Q025	-1.8195	-2.2085	-2.0355	-2.1315
	Q05	-1.6087	-1.7405	-1.7585	-1.7165
	Median	0.0081	-0.0345	0.0885	-0.0076
	Q95	1.7277	1.9217	1.9070	1.9069
	Q975	2.0453	2.2650	2.2997	2.2636
Effective sample size	Median	536.5004	876.3286	256.9461	771.4930
Proportion of Rejections		0.0249	0.0837	0.0605	0.0710
Metropolis-Hastings	Mean			0.9291	
	Q025			0.9070	
	Q05			0.9130	
	Median			0.9290	
	Q95			0.9460	
	Q975			0.9480	