

Computational modeling

1 ISETBio computational model

1.1 Overview

To connect the measured spatial transfer functions (STFs) to the receptive field (RF) organization of the underlying retinal ganglion cells, we employed a computational model which simulates optical, spectral, spatial, and temporal components of the AO stimulation apparatus, as well as the monkey’s optics and cone mosaic structure. The model computes a STF, assuming that cone signals are pooled by the center and surround mechanisms of an RGC according to a difference of Gaussians (DoG) spatial profile model. The parameters of the DoG model and therefore, the actual cone pooling weights, are estimated by minimizing the error between the model- and fluorescence-based- STFs. A schematic overview of this approach is depicted in Figure 1.

1.2 AO-based stimulus delivery modeling

The drifting monochromatic sinusoidal gratings used to measure RGC STFs are modeled as temporal sequences of ISETBio spatial-spectral radiance scenes, where each scene models one frame of the displayed stimulus. The spectral characteristics, spatial extent and the temporal properties of the AO display subsystem are all taken into account in generating the ISTEBio scenes. The spectral profile of the monochromatic beam is modeled as Gaussian shaped with a peak at 561 nm and a full-width-half-max bandwidth of 5 nm, and the mean irradiance on the retina is 1.29 mW/cm^2 . The visual stimulus has a pixel size of $1.03 \mu\text{m}$ and its spatial extent is 0.7×0.7 degrees. All sinusoidal gratings are presented with a nominal contrast of 1.0, and they are drifting at 6.0 Hz with a refresh rate of 25.3 Hz for a duration of 666 msec (4 cycles).

1.3 Retinal stimulus modeling

The monkey’s optics during AO-based stimulation are modeled as a diffraction-limited optical system model together with a small amount of residual blur. The diffraction limit is obtained using the 6.7 mm pupil diameter employed in the experiment assuming a conversion between degrees of visual angle and retinal extent of $199.26 \mu\text{m}/\text{deg}$. Residual blur in the AO apparatus, which might occur due to a slight defocus of the stimulus with respect to the plane of cone inner segments, is modeled by adjusting the Zernike defocus coefficient used to compute the optical point spread function. The amount of residual blur is not known a-priori and is estimated as part of model fitting, as described below. The ISETBio spatial-spectral radiance scenes modeling the AO stimulus are passed via this optical system to generate corresponding spatial-spectral retinal irradiance images, which are processed by the cone mosaic model as described in the next section.

1.4 Cone mosaic modeling

An ISETBio model of the monkey’s cone mosaic is generated from cone density maps measured during AO imaging, using an iterative approach as described by [Cottaris et al, 2019]. The spatial extent of the model cone mosaic is $1.3 \text{ deg} \times 1.3 \text{ deg}$, with a maximal cone density of $270,200 \text{ cones/mm}^2$ with relative L:M:S cone densities of 0.48:0.48:0.04. Cones are modeled with Gaussian entrance apertures with a characteristic radius equal to $0.204 \times \sqrt{2} \times D$, where D is the inner segment diameter as measured during AO-imaging. In this cone mosaic model,

visual stimulation modeling	
monochromatic stimulation (peak) (nm) :	561.0
monochromatic stimulation (FWHM) (nm) :	5.0
mean power ($mW \times cm^{-2}$) :	1.29
retinal pixel size (μm) :	1.03
spatial extent (degs) :	0.7×0.7
contrast :	1.0
drift rate (Hz) :	6.0
refresh rate (Hz) :	25.3
stimulus duration (seconds) :	0.666
optics modeling	
pupil diameter (mm) :	6.7
retinal magnification factor ($\mu m \times deg^{-1}$) :	199.26
AO-residual blur (diopters) :	0.067 or variable
cone mosaic modeling	
size (degs) :	1.3×1.3
cone density :	location dependent, matching AO-measurements
peak cone density ($10^3 cones \times mm^{-2}$) :	270.20
L-cone density :	0.48
M-cone density :	0.48
S-cone density :	0.04
cone aperture profile :	Gaussian
cone characteristic radius, c_{r_c} :	$0.204 \times \sqrt{2} \times$ inner segment diam
foveal cone characteristic radius (arc.min.) :	0.17
RGC modeling (spatial pooling of cone signals)	
peak sensitivity of the center (k_c) :	1
peak sensitivity of the surround (k_s) :	variable
characteristic radius of the center (r_c) :	c_{r_c} (single cone) or variable (multi-cone)
characteristic radius of the surround (r_s) :	variable

Table 1: Parameter values for modeling the AO-based stimulation paradigm, optics, cone mosaic, and RGC spatial pooling.

cone outer-segment lengths and macular pigment all vary with eccentricity as described in [Cottaris et al], and the cone quantal efficiencies are based on the Stockman-Sharpe (2000) normalized absorbance measurements.

1.5 Cone excitation response modeling

To compute a cone's excitation response, the spatial-spectral irradiance impinging on the retina is first spectrally weighted by the product of macular pigment transmittance and by each cone's spectral quantal efficiency, subsequently integrated over wavelength, and spatially integrated over each cone's Gaussian aperture. This excitation response is then integrated over the temporal duration of each stimulus frame (39.5 msec), to estimate the expected excitation events count, $E^j(\omega, t)$, for the j -th cone in the mosaic, at time t , in response to a drifting grating of spatial frequency ω . In these simulations, we do not include Poisson noise in $E^j(\omega, t)$, nor do we introduce jitter between the cone mosaic and the retinal stimulus due to eye movements.

1.6 Converting cone excitation responses to cone contrast responses

Assuming that cones are adapted to the mean background irradiance, the excitations response $E^j(\omega, t)$, is converted to contrast response, $R^j(\omega, t)$, by first subtracting the excitation of the cone to the background stimulus, E_o^j , and

then dividing by it, separately for each cone- j , i.e.:

$$R^j(\omega, t) = \frac{E^j(\omega, t) - E_o^j}{E_o^j} \quad (1)$$

where This operation captures an important effect of the photocurrent generation process which converts cone absorption events in the inner segment into ionic currents flowing through the cone outer segment, and which in effect down-regulates the stimulus-induced cone excitation rate with respect to the background cone excitation rate.

1.7 Computing model ganglion cell responses

Model ganglion cell responses, $RGC(\omega, t)$, are computed from the cone contrast responses assuming linear spatial pooling of cone contrast signals by the antagonistic center and surround mechanisms, as follows:

$$RGC(\omega, t) = RGC_c(\omega, t) - RGC_s(\omega, t) \quad (2)$$

$$= \sum_j W_c^j \times R^j(\omega, t) - \sum_j W_s^j \times R^j(\omega, t) \quad (3)$$

where W_c^j and W_s^j , are the factors with which the center and surround mechanisms, respectively, weigh responses $R^j(\omega, t)$. We are not modeling the temporal filtering and the delay between the center, $RGC_c(\omega, t)$, and surround $RGC_s(\omega, t)$ responses. Although real RGC responses may be affected both by temporal filtering and a center-surround delay, we have only response amplitude (not phase) measurements at only one temporal frequency and thus cannot meaningfully estimate these RGC parameters.

1.8 Estimating model RGC cone pooling weights from fluorescence-based STF measurements

To make computation of cone weights more tractable, we assume that the spatial distribution of cone weights to the center and surround mechanisms have concentric, Gaussian-shaped profiles:

$$W_c^j = \begin{cases} k_c \times \exp \left[-(d_j/r_c)^2 \right], & \text{for the multi-cone RF center model.} \\ k_c, & \text{for the single-cone RF center model.} \end{cases} \quad (4)$$

$$W_s^j = k_s \times \exp \left[-(d_j/r_s)^2 \right] \quad (5)$$

where d_j is the distance between cone- j and the spatial position of the center mechanism of the model RGC (which is the spatial position of the cone driving the center mechanism). The parameters k_c , k_s , r_c , and r_s , which represent the center and surround peak sensitivities and characteristic radii, respectively, of the Difference of Gaussians (DoG) RF model, are determined by minimizing the square root of the weighted error between the model-predicted STF, $STF^m(\omega)$, and the measured STF, $STF^{\Delta F/F}(\omega)$, accumulated over all spatial frequencies, ω :

$$RMSE = \sqrt{\sum_{\omega} \frac{1}{\epsilon(\omega)} \times [STF^m(\omega) - STF^{\Delta F/F}(\omega)]^2} \quad (6)$$

where $\epsilon(\omega)$ is the standard error of the mean of the $STF^{\Delta F/F}(\omega)$ measurement. To minimize the chance of getting stuck to local minima of the error function, we employ a multi-start minimizer which is ran 512 times, keeping the results from the run which results in the minimum RMSE.

The model-predicted STF, $STF^m(\omega)$, is computed by fitting a sinusoidal function,

$$RGC^m(\omega, t) = A(\omega) \times \sin [2\pi f t - \theta] \quad (7)$$

to the model response, $RGC(\omega, t)$, where the f is set to the temporal frequency of the drifting gratings, and $\theta, A(\omega)$ are free parameters. At each spatial frequency, the $A(\omega)$ value is taken as $STF^m(\omega)$.

In some cells, the measured $\text{STF}^{\Delta F/F}(\omega)$ can reach negative values for certain spatial frequencies ω . Since the fluorescence-based response to all grating stimuli is computed by subtracting the fluorescence-based response to a null (spatially uniform) stimulus, we interpret such negative values as a failure of the null stimulus to fully account for the baseline fluorescence response. In terms of modeling, there are 2 ways of handling this: (a) zeroing all negative values of the $\text{STF}^{\Delta F/F}(\omega)$ to be fitted, or (b) adding a nuisance parameter to the model, $\beta \leq 0$, such that $\text{STF}^m(\omega) = \beta + A(\omega)$. Here, we chose the latter option.

1.9 Model selection

To interpret the measured responses in terms of the underlying retinal wiring, we evaluated four different model scenarios, which differed in (a) the RGC receptive field center organization (having either a single or a variable number of input cones), and (b) the residual optical blur in the AO apparatus (being either zero or a variable amount), as seen in Table 3.

model scenario	variable params	fixed params
single cone center, zero residual defocus	k_c, k_s, r_s	$r_c = c_{r_c}, z_d = 0.000D$
single cone center, variable residual defocus	k_c, k_s, r_s, z_d	$r_c = c_{r_c}$
multiple cone center, zero residual defocus	k_c, k_s, r_c, r_s	$r_c = c_{r_c}, z_d = 0.000D$
multiple cone center, variable residual defocus	k_c, k_s, r_c, r_s, z_d	—

Table 2: Examined model scenarios

To select the most likely model, we employ a cross-validation approach, in which we train RF models based on STF data recorded at one session and evaluate their performance on STF data recorded at another session. This cross-validation procedure ensures that the most likely model has the best ability to generalize, and is therefore not picked because of overfitting the data.

1.9.1 Training the models

For each training session, s_{train} , we determine the cone weights to the center and surround receptive field subregions, W_c^j , and W_s^j , respectively, by minimizing the error between the predicted STF_{train}^m and the measured fluorescence-based STF, measured at that session, $\text{STF}_{train}^{\Delta F/F}$. Since the receptive field position of the recorded RGCs is known to lie within the central $40 \mu\text{m}$, but not known exactly, we derive W_c^j , and W_s^j estimates for a number of different mosaic positions, corresponding to different possible locations of the analyzed RGC. Therefore, for each training session, we obtain a number of possible RF models which may vary from one mosaic position to another due to local inhomogeneities of the local structure and local variations in cone density. This analysis is repeated for the 4 examined model scenarios (Table 3).

Figure 2 depicts the variability in the derived cone weights to the RF center and surround by fitting the measured data to the model. There are 3 main sources of variability.

C. Variability due to different model scenarios. Cone weights derived by the four examined model scenarios all fitted at the same position of the cone mosaic, using the same fluorescence data (session 3). In the single-cone center scenario, when we assume zero residual defocus, the model predicts pooling of cone signals from a wide area by the surround mechanism. When a residual defocus of 0.067D is assumed, the model predicts a much more focal pooling of cone signals by the surround. In the multiple cone center scenarios, when the residual defocus is zero, the model assigns several cones to the RF center and a somewhat similar distribution of cones to the RF surround. When the residual defocus is set to 0.067D , the model assigns only 2 cones to the RF center (with an orientation that is parallel to the vertically-oriented grating), so in effect a single cone along the x-axis, and a surround that is similar in size to when that derived for the the single cone center scenario.

aspects of the analysis for one RGC. Figure 2A depicts models trained at different mosaic positions. Shown in separate panels (left to right) are: the derived cone weights for the model RGC RF center (Fig. 2Aa), the derived cone weights for the model RGC RF surround (Fig. 2Ab), the horizontal profiles of the RF center and surround (Fig.

3Ac, red and blue shaded areas), computed by integrating cone weights across the y-dimension, and the corresponding model $\text{STF}^m(\omega)$ and measured STF data, $\text{STF}^{\Delta F/F}(\omega)$ (Fig. 3Ad, black line and red disks, respectively). The tree rows of Figure 3A depict data for 3 examined positions within the model cone mosaic, all derived from the same session and the same model scenario. Note that the derived weights for the surround differ somewhat from one position to another. This may be due to local inhomogeneities in the cone mosaic possibly combined with local minima in the error function being minimized. The model trained in the first position, has a slightly more focal surround than the models trained in the other two positions, which results in a little worse fit at the lowest spatial frequencies.

Figure 2B depicts the same analysis for models trained at identical mosaic positions and for the same model scenario, but using fluorescence data from different recording sessions. Note that the measured STFs differ somewhat from session to session, most notable here, the STF from session 2 appears to be a vertically shifted compared to the STFs measured in sessions 1 and 3, whereas the STFs measured in sessions 2 and 3 appear to be somewhat more narrow than the STF measured in session 1. As a result, the derived surround cone weights are more focal in sessions 2 and 3, than in session 1.

Figures 2C, D, E, and F depict results for the four examined model scenarios. These analyses are all derived from the STF measured at session 1, and for the same mosaic position.

1.9.2 Cross-validating models

To select the best model we assess their performance at generalizing across recording sessions. To do so we compute the error between STF_{strain}^m and $\text{STF}_{stest}^{\Delta F/F}$, with $stest \neq strain$, i.e. how by assessing the degree to which the model trained as session $strain$ can predict the data recorded at session $stest$. In this, cross-session comparison, we allow for an arbitrary scaling of STF_{strain}^m by a factor γ , which is selected so as to minimize the cross-validated error between the model STF computed using data from the $strain$ session, and the fluorescence-based STF measured at the $stest$ session:

$$\text{cv-RMSE}_{strain}^{stest} = \min_{\gamma} \left(\sqrt{\sum_{\omega} \left[\frac{\gamma \times \text{STF}_{strain}^m(\omega) - \text{STF}_{stest}^{\Delta F/F}(\omega)}{IQR} \right]^2} \right) \quad (8)$$

where IQR is the interquartile range of $\text{STF}_{stest}^{\Delta F/F}(\omega)$. Division with the IQR of the measured data standardizes the error with respect to the scale of the data, which would be different across different $stest$ sessions.

Averaging $\text{cv-RMSE}_{strain}^{stest}$ over all test sessions, we obtain the overall cross-validated RMSE for each training session, cv-RMSE_{strain} , and repeating this for all training sessions, we find the training session, s_{train}^{best} , with the minimal cross-validation error, $\text{cv-RMSE}_{s_{train}^{best}}$. The model computed at the s_{train}^{best} session is considered to have the best generalizing performance. Finally, we compute $\text{cv-RMSE}_{s_{train}^{best}}^j$ for each of examined j -cone positions within the central $40 \mu m$, and searching for the $\min_j (\text{cv-RMSE}_{s_{train}^{best}}^j)$ over all examined cone positions, we determine the best-generalizing model with the lowest cross-validation error amongst all examined RF center positions.

Figure 4 depicts the cross-validation analysis for the L-cone center RGC cell examined in Figure 3. Figure 4A, depicts the cross-validation of session-1 derived models, 4B, depicts the cross-validation of session-2 derived models, and 4C, depicts the cross-validation of session-3 derived models.

In the left panels, the gray lines depict the STF of the model RF that was trained in a particular session and which resulted in the minimum cross-validation error across the remaining 2 sessions, across all examined cone positions. The green and purple lines, depict those optimal cross-validated fits along with the scaling factor (arrows) that had to be applied to derive those fits from the STF of the model RF. The middle panels, plot the cross-validated errors against the training errors for all examined j -cone positions. The connected stars indicate where the minimum combined cross-validated error occurred, and the position of the cone at which this minimum occurred is denoted by the red circle in the right panels, which depict the spatial map of cross-validated errors.

The best cross-validation error across all positions (20.48) occurred with a model trained in session 1. The RF profile, corresponding cone weights, and the types of cones feeding into this best-generalizing model RF for the examined L-cone center RGC are depicted in Figure 5.

1.10 Assessing performance across different models

This cross-validation model selection approach can also be used to compare performance in models with different number of parameters. Without cross-validation, model with more parameters are more likely to overfit the data and have seemingly better performance. With cross-validation, overfitting becomes less likely, and selecting the best model becomes more meaningful.

We examined how different models can explain the measured STF responses (Table 3).

model description	variable params	fixed params	result
single cone center, zero residual defocus	k_c, k_s, r_s	$r_c = c_{r_c}, z_d = 0.000D$	worse cv-RMSE
single cone center, single residual defocus	k_c, k_s, r_s	$r_c = c_{r_c}, z_d = 0.067D$	pending
single cone center, variable residual defocus	k_c, k_s, r_s, z_d	$r_c = c_{r_c}$	pending
multiple cone center, zero residual defocus	k_c, k_s, r_c, r_s	$r_c = c_{r_c}, z_d = 0.000D$	pending
multiple cone center, variable residual defocus	k_c, k_s, r_c, r_s, z_d	—	pending

Table 3: Examined models

2 Results

3 Figures

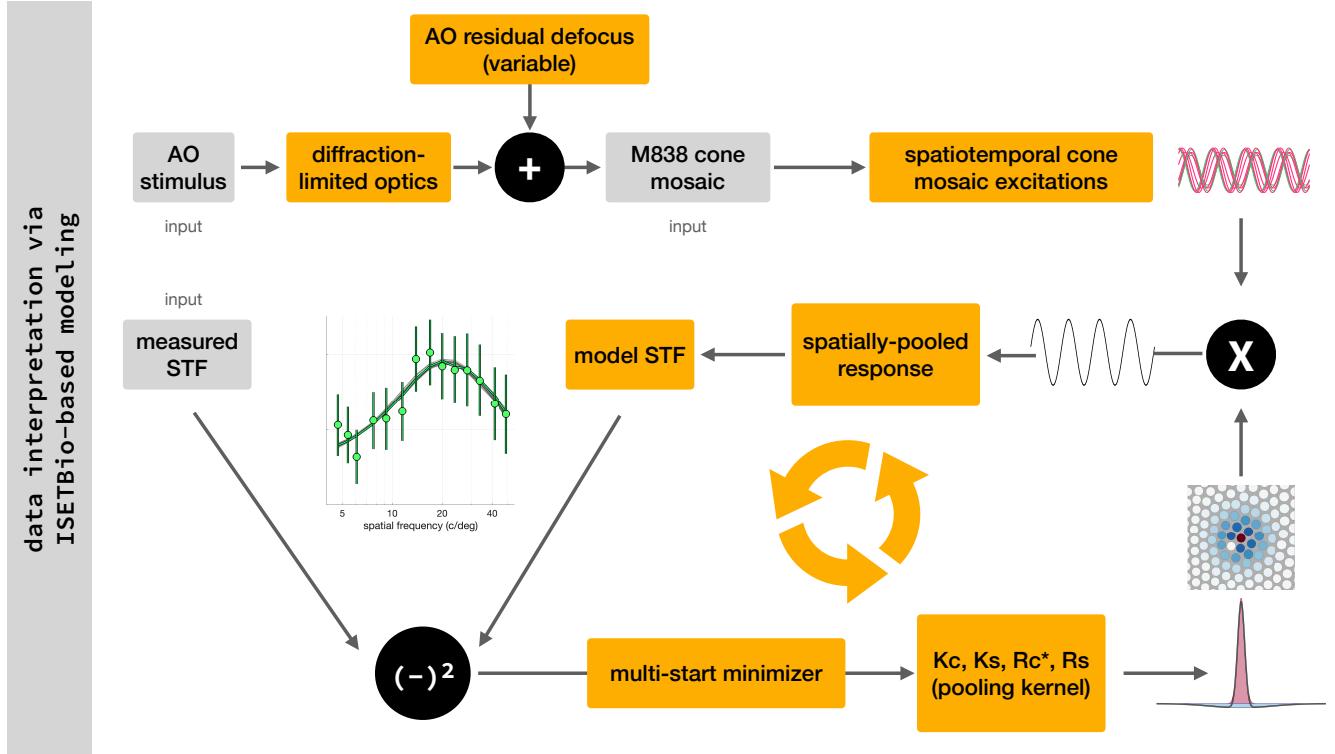
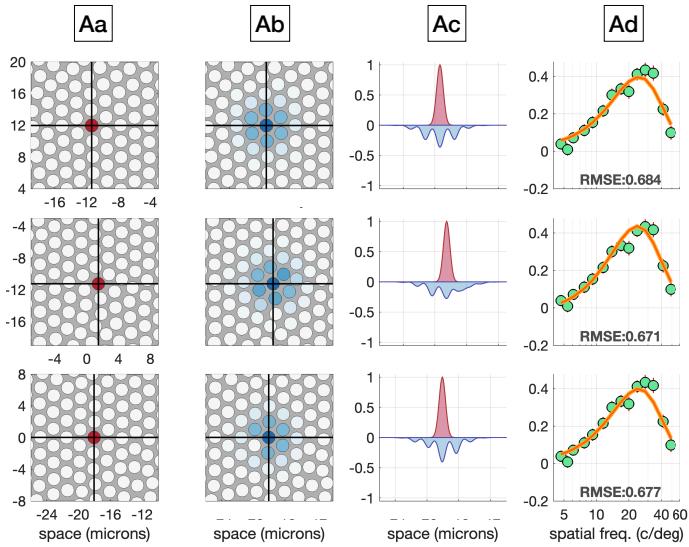
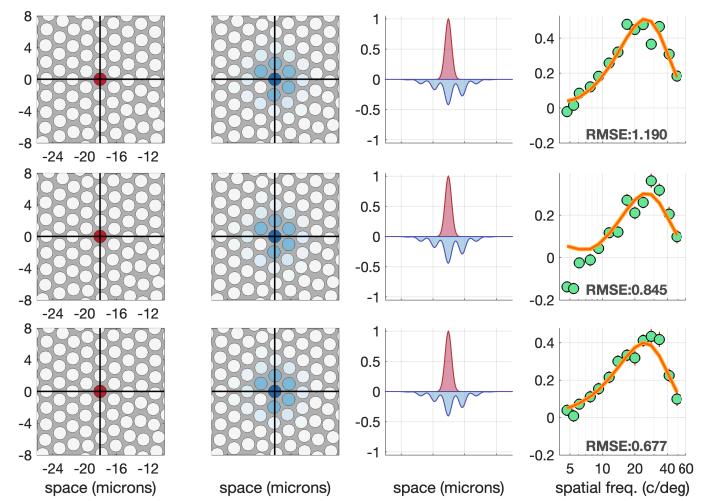


Figure 1: Schematic overview of the ISETBio computational model.

A. variability due to cone mosaic position



B. variability due to different recording sessions



C. variability due to different model scenarios

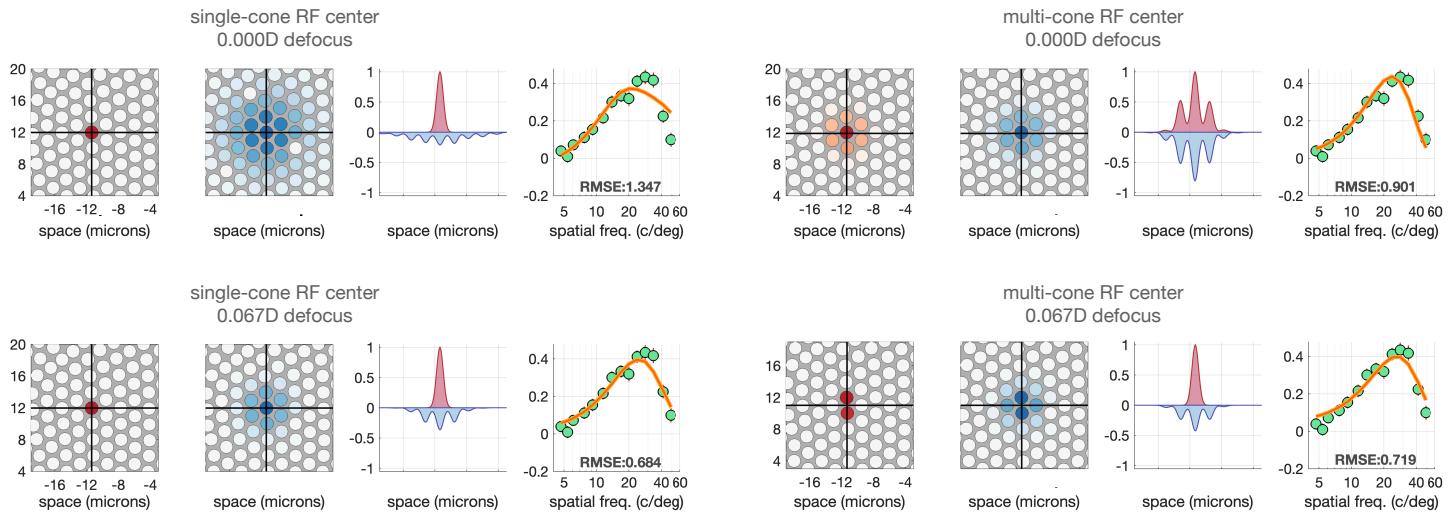
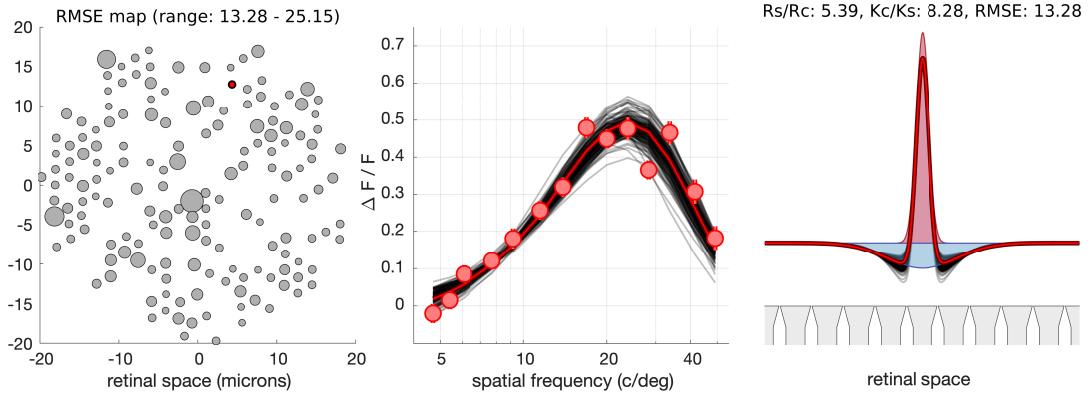
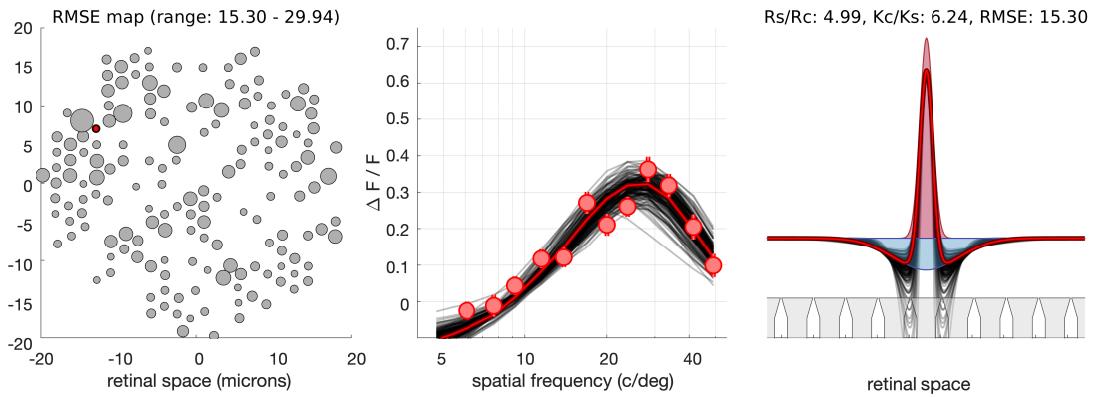


Figure 2: Variability in estimating spatial distribution of cone weights from fitting models to the measured STF. **A.** Variability due to assumed position of the recorded RGC receptive field (RF). **B.** Variability due to different recording sessions. **C.** Variability due to the different examined model scenarios. Figure format for all panels is as follows. The weights of cones feeding into the RF center are color coded in red in **Aa**. The weights of cones feeding into the RF surround are color-coded in blue in **Ab**. Horizontal slices through the RF center (red) and the RF surround (blue), computed by integrating over the y-dimension, the spatial profile on the input cones (cone aperture profile multiplied by the corresponding cone weight to the RF) are depicted in **Ac**. The measured fluorescence-based spatial transfer functions, which the computational model is fit to, $\text{STF}_{\text{stest}}^{\Delta F/F}(\omega)$ (green disks), and the STF of the best-fit model, $\text{STF}_{\text{strain}}^m(\omega)$ (orange line) are depicted in **Ad**.

(a) session 1- based models



(b) session 2- based models



(c) session 3- based models

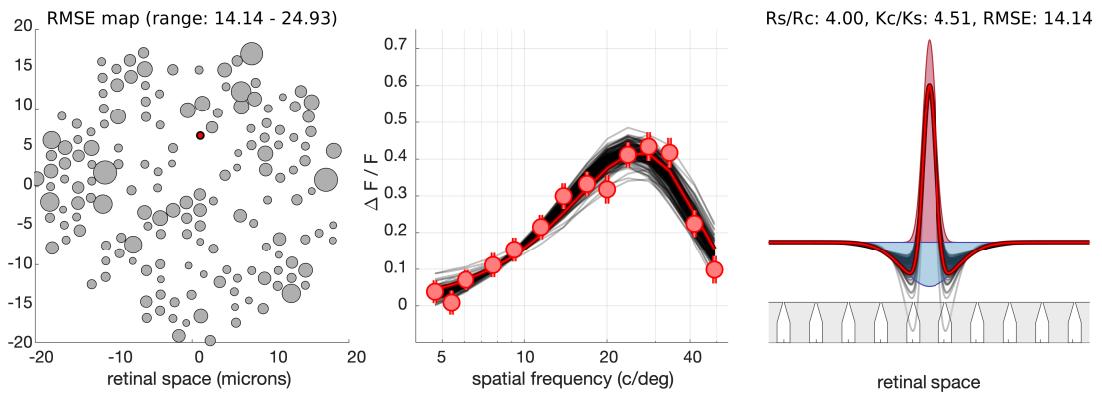
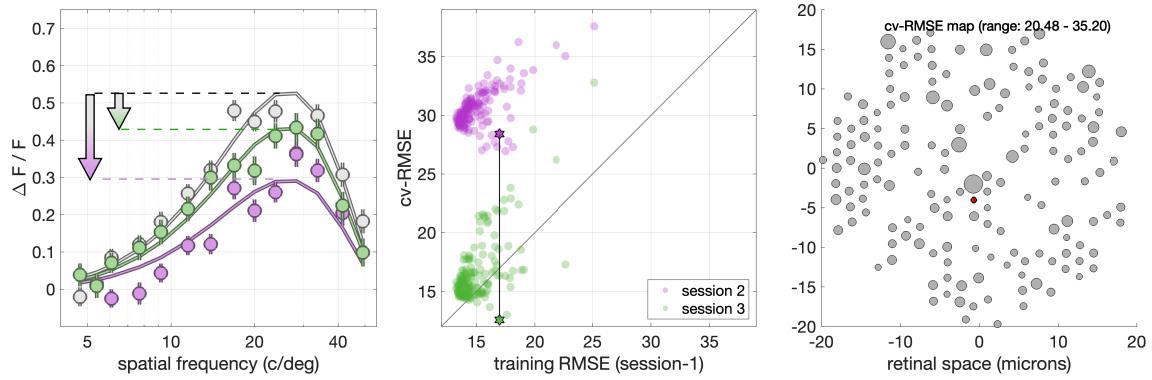
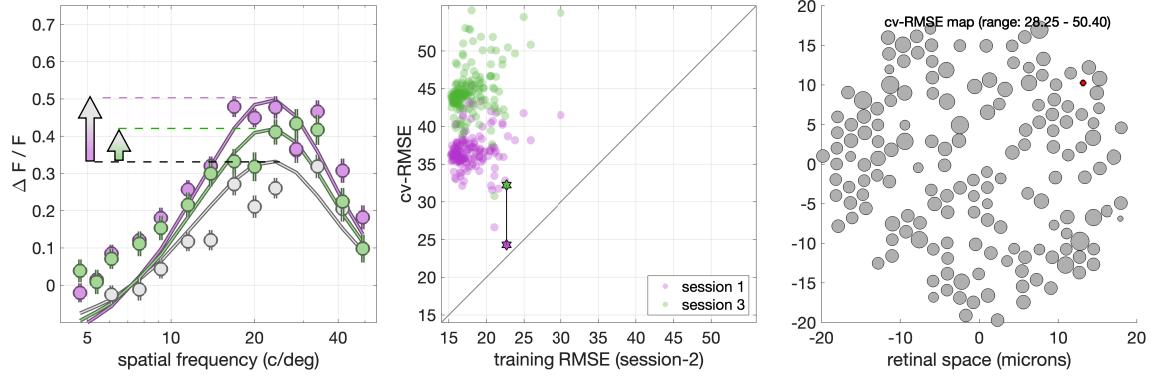


Figure 3: Training RF models for an L-cone center RGC based on STF data recorded during the first (**A**), second (**B**) and third (**C**) recording session. **Left panels:** spatial map of RMSE between model STF and measured STF, as a function of center cone position in the examined model. There are a total of 161 models, each centered on one of 161 cones within the mosaic. Bubble size denotes magnitude of RMSE for each model. **Middle panels:** Fits of model STF (black lines) to the measured STF data (red disks) for all 161 examined models. The red line represents the best fit. **Right panels:** Black thin traces depict the RF profiles of the 161 examined models. The red thick line depicts the RF profile of the best fit model, whereas the pink and blue shaded areas depict the center and surround mechanism profiles of the best fit model. The apertures of cones are depicted schematically in gray.

(a) cross-validating session-1 models



(a) cross-validating session-2 models



(a) cross-validating session-3 models

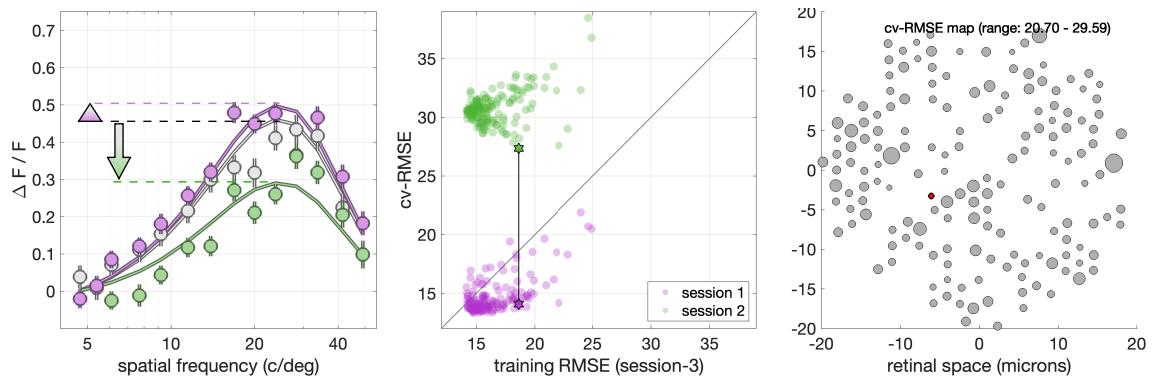


Figure 4: Cross-validating RF models for the same L-cone center RGC cell. The model RF is trained on STF data recorded during the first recording session (**A**), the second session (**B**) and third session (**C**). **Left panels:** Gray lines depict the STF of the model RF trained in the current session that resulted in the minimum combined cross-validation error over the remaining 2 sessions, across all examined cone positions. Green and purple lines, depict the optimal cross-validated fits along with the scaling factor (arrows) that had to be applied to derive them from the STF of the model RF. **Middle panels:** The cross-validated errors are plotted against the training errors for all examined cone positions. The connected stars indicate where the minimum combined cross-validated error occurred. **Right panels:** Spatial map of cross-validated errors, with bubble size indicated the error magnitude. The position of the cone at which this minimum occurred is denoted by the red circle.

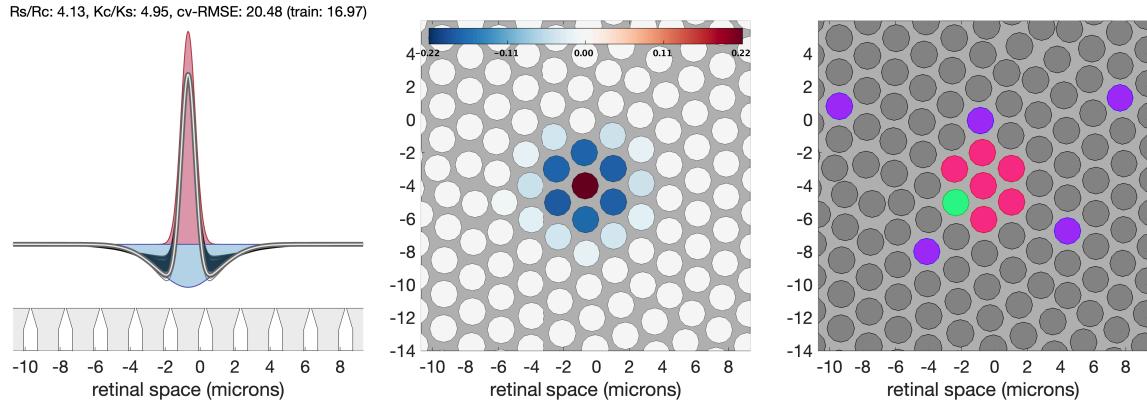


Figure 5: Best cross-validated RF model. **Left panel :** The RF profile of the best fit model, is depicted by the white thick line, whereas the pink and blue shaded areas depict the profiles of the center and surround mechanisms, respectively. Cone apertures are depicted schematically in gray below. **Middle panel:** Weights of cones feeding into the RF center (red) and the RF surround (blue) mechanisms. **Right panel:** Labeling of L/M- (red/green) cones whose weights to the best cross-validated RF model are greater than 5% of the center cone weight, or are S-cones (blue).