Supplementary material: Variograms

The variogram is a technique usually used in geostatistics for the characterization of spatial processes. From the empirical variogram given by the data, the aim is to find the continuous function that represents the regional variation. Without going into much detail (for which we remand to the specialized literature\textsuperscript{1,2}), the most common fitting models for the variograms are the bounded models. In these models, the variance has a maximum, known as the \textit{sill variance}, which may be reached either at a finite distance, known as the \textit{range}, or asymptotically.

One of the most commonly used bounded models is the exponential model, which is described by the equation

\[ \gamma(r) = c \left\{ 1 - \exp \left( -\frac{r}{R} \right) \right\} \]

with sill $c$, distance between the pair of points $r$ and a distance parameter, $R$, that defines the spatial extent of the model. Since this model approaches the sill asymptotically, an \textit{effective range} is defined for practical purposes as the distance at which $\gamma$ equals 95% of the sill variance. This effective range is approximately $3R$ and the slope at the origin is $c/R$. This function is the representation of random processes, such as first-order autoregressive and Markov processes, and it has been the subject of many theoretical studies of the efficiency of sampling design.\textsuperscript{3-6}

Since in this work we calculated a variogram for every image, we had a total of 212 curves. A generally accepted procedure for fitting a theoretical model to an experimental variogram is still the object of debate. Some authors have suggested using an approach that includes both visual inspection and statistical fitting.\textsuperscript{1,2} Since the amount of data we analysed was considerable, we decided to analyse the linear fit
of the variograms at short distances, which can be easily replicated and makes no theoretical assumptions on the spatial distribution of cone reflectances.

In order to evaluate the overall trend of the semivariance in all study population, we inspected the variogram curve obtained taking the mean of all the curves (Figure S1). The resulting curve was well fitted by an exponential model, which can be explained by the fact that averaging the different features of all the subjects can be described by a random process. The exponential fit on the mean curve has as a distance parameter $R = 22\mu m$. Considering that this distance represents the average spatial extent of the individual variogram curves, we decided to use $R = 20\mu m$ as the maximum distance for all the linear fits.

We included in this appendix some experimental variograms, which we chose to be representative of the variety of spatial distributions that we encountered in our study population (Figures S2 and S3). The variograms are beside the corresponding images, to allow a visual comparison between the distribution of cone intensities and the shape of the curves. In all cases, the shape of the curves was constant with time. This allowed us to determine that the process at the basis of the spatial distribution of the intensity values in the cone layer remained the same through time, even if the reflectance of the individual cones changed.
References (Supplementary material)


2. Oliver MAaWR. Basic steps in geostatistics: the variogram and kriging: Springer; 2015.


Figure Legends

**Figure S1.** Mean of all the experimental variogram curves calculated in this study. The blue squares are the experimental data, the orange line is the exponential fitting function

**Figure S2.** Variograms of two control subjects. The cones in the images show no apparent pattern and this is confirmed by the variograms, which reach the maximum variance at short distances and then remain flat, showing no long-distance correlation. Size bar is 100μm

**Figure S3.** Variograms of two NPDR subjects. The shape of the curves allowed us to quantify what was only suggested by visual inspection, i.e. the presence of bright and dark patches in the images. The presence of peaks and valleys shows how cone reflectances are not independent from each other. NPDR2 shows a first peak around 60μm, while NPDR3 shows a large peak approximately at 100μm, which reflects the size of the bright clusters in the images. Especially in the latter case, the visual assessment is made easier by the comparison between the size of the bright cluster and the size bar (100μm), which are approximately the same.
Figure S1
Figure S2
Figure S3