Human Brain Mapping 2009

SurfStat: A Matlab toolbox for the statistical analysis of univariate and multivariate surface and volumetric data using linear mixed effects models and random field theory

KJ Worsley[1,2], JE Taylor[3], F Carbonell[2], MK Chung[4], E Duerden[5], B Bernhardt [2], O Lyttelton[2], M Boucher[2], AC Evans[2]

[1]University of Chicago, Chicago, IL, United States/[2]McGill University, Montreal, QC, Canada/[3]Stanford University, Stanford, CA, United States/[4]University of Wisconsin-Madison, Madison, WI, United States/[5]Université de Montréal, Montréal, QC, Canada

Introduction:
SurfStat does statistical analysis of univariate and multivariate surface data using linear mixed effects models (fitted by ReML) and random field theory (RFT). SurfStat is intended for cortical thickness data on triangular meshes, but it will handle any triangulated surface data; the only requirement is that the triangulation scheme must be the same for all surfaces, i.e. the data must be registered to a common surface.

Methods: Inference uses RFT for T, F, Hotelling's T^2 and Roy's maximum root statistics. Model formulae rather than design matrices specify the linear model. It is fast, because everything is loaded into memory, permitting a truly interactive analysis. Finally, off-the-shelf Matlab graphics are ready to publish.

It will also analyse volumetric data, e.g. VBM, DBM and PET data. To cope with increased memory demands, while maintaining fast execution, memory mapping is implemented.

Results: SurfStat will handle registered FreeSurfer data, but here we shall analyse 147 subjects from the ICBM data base. The scientific question is if cortical thickness is related to gender or age. Figure 1 shows data on a single subject.
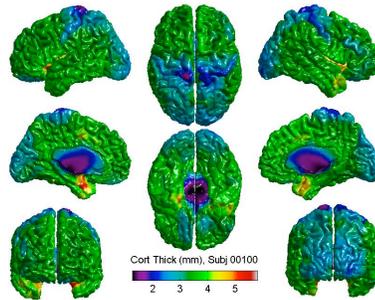
Cortical thickness has been read into a single subjects(147) × vertices(81920) matrix **Y**, and gender and age have been read into subjects(147) × 1 vectors **gender** (values **'male'** and **'female'**) and **age** (in years). We now convert these to *terms* that can be combined into a linear model **M** as follows:

Gender = *term*(**gender**);
Age = *term*(**age**);
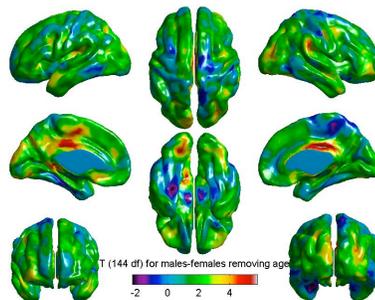**M** = 1 + **Gender** + **Age**;

The linear model is fitted by

slm = *SurfStatLinMod*(**Y**, **M**, **avsurf**);

where avsurf is a pre-computed structure that contains the triangulation scheme and an average surface (purely for display). To specify a gender contrast, calculate the T-statistic, and view it within a pre-computed mask:

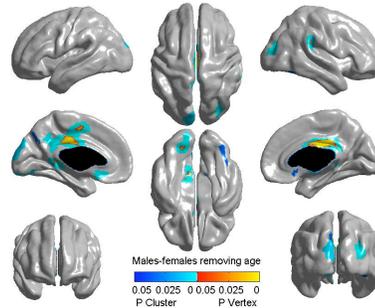**contrast** = **Gender.male** - **Gender.female**;
slm = *SurfStatT*(**slm**, **contrast**);
SurfStatView(**slm.t.*mask**, **avsurf**, 'T (144 df) for males-females removing age');

Finally, to see RFT corrected P-values inside the mask for peaks and clusters above **P**=0.001 (uncorrected):

SurfStatView(*SurfStatP*(**slm**, **mask**), **avsurf**, 'Males-females removing age');


Interactions can be specified as follows:

M = 1 + **Gender** + **Age** + **Age*Gender**;

F-statistics can be calculated by comparing nested models, e.g. to test for an interaction between age and gender:

slm0 = *SurfStatLinMod*(**Y**, 1 + **Gender** + **Age**, **avsurf**);
slm1 = *SurfStatLinMod*(**Y**, 1 + **Gender** + **Age** + **Age*Gender**, **avsurf**);
slm = *SurfStatF*(**slm1**, **slm0**);
SurfStatView(**slm.t.*mask**, **avsurf**, 'F statistic, 1,143 df');

Mixed effects models, fitted by ReML, can be specified as follows. Suppose some of the rows of **Y** are in fact repetitions on the same subject, specified by a variable **Subj** (35 "subjects"). Then a simple mixed effects model can be specified by:

M = 1 + **Gender** + **Age** + *random*(**Subj**) + **I**;
image(**M**);
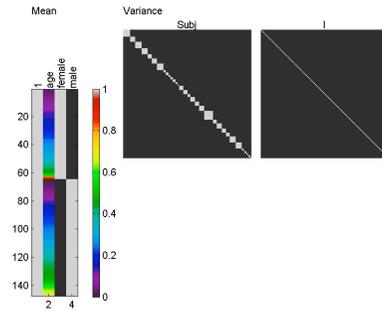
Multivariate data can be analysed simply by adding an extra dimension to Y, so that it becomes a subjects × vertices × variate array. T- and F-statistics become Hotelling's T, and Roy's maximum root.

Finally, 3D data can be analysed in exactly the same way, though it is slower and uses more memory.

Conclusions: SurfStat is a very flexible tool for the statistical analysis of surface or volumetric data via linear mixed effects models.

References:

Adler, R (2007), '', *Random fields and geometry*, vol. , no. , pp. .

Taylor, J (2008), 'Random fields of multivariate test statistics, with applications to shape analysis', *Annals of Statistics,*, vol. 36, no. , pp. 1-27.

Taylor, J (2007), 'Detecting sparse signal in random fields, with an application to brain mapping', *Journal of the American Statistical Association*, vol. 102, no. , pp. 913-928.

Worsley, K (1999), 'Detecting changes in non-isotropic images', *Human Brain Mapping*, vol. 8, no. , pp. 98-101.