

---

# Voxel-wise nonlinear analysis toolbox for neurodegenerative diseases and aging

---

Authors

## Abstract

The abstract paragraph should be indented 1/2 inch (3 picas) on both the left- and right-hand margins. Use 10 point type, with a vertical spacing (leading) of 11 points. The word **Abstract** must be centered, bold, and in point size 12. Two line spaces precede the abstract. The abstract must be limited to one paragraph.

## 1 Introduction

Nowadays there is a vast armory of neuroimaging analysis tools available for the neuroscientific community whose ultimate goal is to conduct spatially extended statistical tests to identify regionally significant effects in the images without any a priori hypothesis on the location or extent of these effects. Some of these tools perform these tests at the voxel level<sup>1</sup>, whereas some others provide with specific environments for analyzing the images, based on 3D meshes<sup>2</sup> or boundaries<sup>3</sup>.

Irrespective of these differences, the vast majority of neuroimaging tools are based on different implementations of the General Linear Model (GLM). Even though the GLM is flexible enough for conducting most of the typical statistical analysis, its flexibility for modeling nonlinear effects is rather limited. To this regard, it is important to note that some relevant confounders in most analysis, such as the impact of age on regional brain volumes, have been described to be better described by nonlinear processes ([1], [2]).

In this event, the most widely used alternative is to refrain from performing a voxelwise analysis and quantify brain volumes based on regions of interest (ROIs) and then perform the nonlinear fitting of these trajectories by external software. This alternative has the obvious disadvantage that made voxel-wise analysis so popular on the first place: the need to identify a set of ROIs a priori.

In this work, we describe a new analysis toolbox that allows for the modeling of nonlinear effects at the voxel level that overcomes these limitations and illustrate its features using a relevant example in which distinct nonlinear trajectories were found as a function of a cerebrospinal fluid (CSF) related biomarker for participants in all stages of Alzheimer's disease.

## 2 The toolbox

The toolbox comprises an independent *fitting library*, made up of different *model fitting* and *fit evaluation* methods, a *processing* module that interacts with the fitting library providing the formatted data obtained from the *file system*, several *visualization* tools and a *command line interface* that allows the interaction between the user and the processing module, supported by a *configuration file*.

---

<sup>1</sup><http://www.fil.ion.ucl.ac.uk/spm/>

<sup>2</sup><http://freesurfer.net/>

<sup>3</sup><http://idealab.ucdavis.edu/software/bbsi.php>

## 2.1 Model fitting techniques

Model fitting consists in finding a parametric or a nonparametric function of some explanatory variables (**predictors**) and possibly some confound variables (**correctors**) that best fits the observations of the target variable in terms of a given quality metric or, conversely, that minimizes the loss between the prediction of the model and the actual observations. Implemented models are:

### General Linear Model (GLM)

The General Linear model is a generalization of multiple linear regression to the case of more than one dependent variable. Nonlinear relationships can be modeled with GLM using a polynomial basis expansion of degree  $d$ , that maps the input space into another feature space that includes the polynomial terms of the variables.

### Generalized Additive Model (GAM)

A Generalized Additive Model ([3]) is a Generalized Linear Model in which the observations of the target variable depend linearly on unknown smooth functions of some predictor variables:  $f(X) = \alpha + \sum_{i=1}^k f_i(X_i)$ . Here  $f_1, f_2, \dots, f_k$  are nonparametric smooth functions that are simultaneously estimated using scatterplot smoothers by means of the backfitting algorithm (REF). Several fitting methods can be accommodated in this framework by using different smoother operators, such as cubic splines, polynomial or Gaussian smoothers.

### Support Vector Regression (SVR)

In Support Vector Regression the goal is to find a function that has at most  $\epsilon$  deviation from the observations and, at the same time, is as flat as possible. However, the  $\epsilon$  deviation constraint is not feasible sometimes, and a hyperparameter  $C$  that controls the degree up to which deviations larger than  $\epsilon$  are tolerated is introduced. In the context of SVR nonlinearities are introduced with the "kernel trick", that is, a kernel function  $k(x_i, x_j) = \langle \Phi(x_i), \Phi(x_j) \rangle$  that implicitly maps the inputs from their original space into another high-dimensional space. The kernel function used in this toolbox is the Radial Basis Function (RBF) kernel, which is defined as  $k(x_i, x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ .

SVR methods rely on several hyperparameters, namely  $\epsilon$  and  $C$  in general and also  $\gamma$  when using a RBF kernel function. To address the search of these hyperparameters an automatic method based on grid search is included in this toolbox, which comprises the following steps: 1) sample the hyperparameters space in a grid using one of the several sampling methods provided in the toolbox; 2) fit a subset of the data with the combination of hyperparameters of each sample in the grid; 3) select the combination that minimizes the error function of choice.

## 2.2 Fit evaluation methods

### MSE, Coefficient of determination ( $R^2$ )

These metrics evaluate the predictive power of the model but don't penalize its complexity.

### Akaike Information Criterion (AIC)

Criterion based on information theory widely used for model comparison and selection. Unlike the previous ones, it penalizes the complexity of the model by requiring its number of parameters.

### F-test

The F-test evaluates whether the variance of the full model (correctors and predictors) is significantly lower — from a statistical point of view — than the variance of the restricted model (only correctors), that is, the inclusion of the predictors contributes to the explanation of the observations. This statistical test requires the degrees of freedom of both models, which are trivial to compute in GLM, but they aren't in GAM or SVR. The equivalent degrees of freedom for Support Vector Regression were introduced in the toolbox as defined in [4].

### Penalized Residual Sum of Squares (PRSS), Variance-Normalized PRSS

Penalized Residuals Sum of Squares is introduced in the toolbox in order to provide a fit evaluation metric that penalizes the complexity of the predicted curve without requiring the degrees of freedom. However, PRSS is not suitable enough in the context of morphometric analysis, as it always provides better scores for target variables with low-variance and flat trends than for target variables with

high-variance and nonflat trends, and that poses a problem as most of the voxels in the brain are 0-valued. For that reason a variance normalized version of the PRSS that weights the score with the inverse of the variance of the predicted curve was introduced, the Variance Normalized Penalized Residuals Sum of Squares.

### 2.3 Interactive visualization tool

An interactive visualization tool is included in the toolbox to provide added insight on the results: it allows to load a 3D statistical map — generated with the fit evaluation method of choice — and several fitted models, and then plots the predicted curves of all the models for the voxel selected with the cursor, hence easing the task of inspecting the curves in the significant regions. An example of the aforementioned tool is found in Figure 1.

## 3 Implementation details

The whole toolbox has been implemented in Python. Numpy and scipy have been used for the numerical and scientific computing, NiBabel for handling the morphometric data in NIfTI format, scikit-learn ([5]) for the machine learning algorithms and matplotlib and seaborn for plotting and for the visualization features.

## 4 Dataset

## 5 Experiments

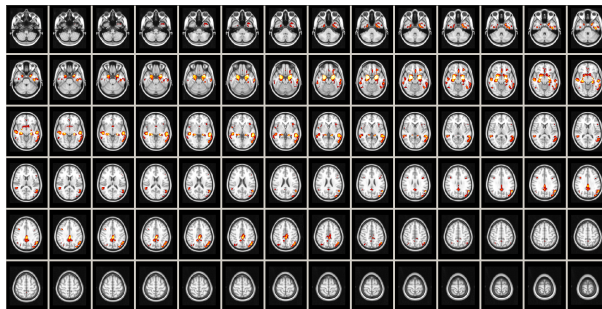


Figure 1: F-test statistical map for the polynomial SVR, with significance level ( $\alpha$ ) filtering at 0.001, minimum cluster size of 100 voxels and transformed into Z-scores for improved visualization.

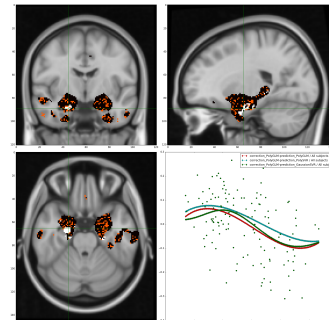


Figure 2: Comparison of polynomial GLM (black) vs polynomial SVR (orange) vs Gaussian SVR (white) with the corresponding curves for a voxel in which the best fitting score belongs to the Gaussian SVR model.

Sample figures.

## 6 Conclusions

As compared to ROI-based statistical analysis, the voxelwise approach has the key advantage to allow for spatially unbiased analysis of brain images. This strategy has become predominant in the last decades and typically relies on particular implementations of the GLM. However, this approach presents with severe limitations when it comes to the modeling of non-linear effects, hence current neuroimaging analysis tools are sub-optimal for the identification of such nonlinear patterns. As a consequence, the capacity of the neuroimaging community to detect spatially distributed critical points associated to brain maturation or pathology is as well limited.

In addition, the proliferation of neuroimaging repositories with longitudinal data provides with a unique opportunity for imaging researchers to analyze reference datasets with different tools, thus enabling the comparison and validation of analytical tools while gaining more insight on the data under analysis. Examples of these repositories are the OASIS, ADNI, and AIBL databases. In the case of ADNI, subjects have undergone several imaging assessments over time and, as for some imaging biomarkers, it is obvious that their temporal trajectories depart from linearity ([6]). Therefore, in the analysis of such longitudinal datasets, the availability of non-linear modeling tools is likewise critical to fully understand the interrelationships between the trajectories of different biomarkers that may be crucial for understanding downstream pathological effects.

## References

- [1] A. M. Fjell, L. T. Westlye, H. Grydeland, I. Amlien, T. Espeseth, I. Reinvang, N. Raz, D. Holland, A. M. Dale, and K. B. Walhovd. Critical ages in the life course of the adult brain: nonlinear subcortical aging. *Neurobiol. Aging*, 34(10):2239–2247, Oct 2013.
- [2] A. M. Fjell, L. T. Westlye, H. Grydeland, I. Amlien, T. Espeseth, I. Reinvang, N. Raz, A. M. Dale, and K. B. Walhovd. Accelerating cortical thinning: unique to dementia or universal in aging? *Cereb. Cortex*, 24(4):919–934, Apr 2014.
- [3] Trevor J Hastie and Robert J Tibshirani. *Generalized Additive Models*, volume 43 of *Monographs on Statistics and Applied Probability*. Chapman & Hall/CRC, June 1990.
- [4] Francesco Dinuzzo, Marta Neve, Giuseppe De Nicolao, and Ugo Pietro Gianazza. On the representer theorem and equivalent degrees of freedom of svr. *Journal of Machine Learning Research*, 8:2467–2495, December 2007.
- [5] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [6] P. S. Insel, N. Mattsson, M. C. Donohue, R. S. Mackin, P. S. Aisen, C. R. Jack, L. M. Shaw, J. Q. Trojanowski, and M. W. Weiner. The transitional association between  $\beta$ -amyloid pathology and regional brain atrophy. *Alzheimers Dement*, 11(10):1171–1179, Oct 2015.
- [7] Trevor Hastie, Robert Tibshirani, and Jerome Friedman. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction*. Springer, second edition edition, February 2009.
- [8] Alex J. Smola and Bernhard Schölkopf. A tutorial on support vector regression. *Statistics and Computing*, 14(3):199–222, August 2004.