**README – CODES AND DATA - DC’s LARGE MEDIA**

**J.A. Duran**

**Feb 2021**

Here you find the codes used to compute numerical and experimental decorrelations and generate the DC figures, as in the paper of A.Duran et al. (2020). The simulations were run using a large media (33.6km x33.6km) and 20 Vmodels.

To be able to use the codes, one should first look to the previous folder A\_Tau\_Kernels\_Project/CODES for the SPECFEM2D version used, and the main codes contained in the folders *1\_Vmodels* and *2\_Elastic\_m.f.p.*

Three locations where tested (rp=3, 8, 9).

Here are the locations of all receivers used. Those are the unique locations where I have any data.



FOLDERS AND THEIR CONTENTS

DCnum:

* Contain samples of the output data (only for the first model) after the SPECFEM2D simulations (For Sim1 and Sim2).
* Here you find the codes used to compute the ensemble average intensities and the numerical Kernels.
* It contains the result of such operations: **Mean\_Intensities\_20Models.mat** and **Kernels\_20Models.mat**

DCexp:

Evaluates the decorrelation among waveforms before and after a localized vel anomaly is placed. The results here are always for a dv/v=+7% single point velocity anomaly. Always performed using x-displacement as my recordings.

* Code used to compute the ‘experimental’ decorrelations.
* Output of such code for three cases: i)when only a dvp is placed, ii) only a dvs , iii) Both dvp and dvs are placed at the same time wit.

The DCexp outputs are computed and saved.

DCplot\_LargeMed:

Here are the codes used to plot the results.

*A1\_DC\_plot\_Components.m* -> Given at specific rprime location (rp=3, 8 or 9), it computes DCbullk and loads the results of DCnum and DCexp.

For DCexp it considers the standard deviation, standard error and the interquartile range (IQR).

For DCnum I still used the errors computed in the model of 16.8km x 16.8m.

It saves all this information as “DCplot\_LargeMedDC\_EL\_plot\_LargeMed\_rp3.mat” so I can load it and plot it.

What I did in the past to compute DCnum error was to pick randomly 26Vmodels (from a total of 40) and compute the ensemble average intensities. Then, I compute DCnum using such average intensities and this gives me the first DCnum\_1 (this is step1, i.e. n=1) .

I repeat the procedure with another 26 models, compute the average intensities and then obtain DCnum\_2 (this is step 2, i.e. n=2). In this way I keep sampling the value of DCnum and I can make statistics. Then from such a set I can get the error.

DCnum\_n for n=1,..,40 -> I repeat the procedure 40 times.

Thus, it makes sense to me that the values of this error\_DCnum < error\_DCexp

But I’d need to generate enough data with these new large models to repeat the entire procedure.

*A2\_DC\_plots\_LargeMed.m* -> This is the code which load all the information, make a smoothing on the data and plot the final figure.

**Some words about the errors:**

From a Normal distribution (random variable), it is easier to take the error by studying mean and the std (avg distance from the values xi to the mean). Error is an indication of the level of confidence with the data, it is in a way an indication how spread the data is.

In various cases, the data is assumed to satisfy a normal distribution, like in T-student.

However, after plotting in the past the histograms of DCexp(r,to) for various sample times, I couldn’t tell if the distribution is normal. To me it seems to vary even between time steps, perhaps is a bimodal distribution or simply no normal.

When the distribution is not normal the interpretation of what standard deviation tells about the data is not clear. In such case, we could use non parametric approaches, in which there is no assumption about the type of distribution the data would fit in.

IQR is one of them. And it is the one I’m using for DCexp. However, I also observed that I found a better fit by simply using the mean over all the models as my result (i.e. DCexp=mean(DCexp)\_models) instead of using the median (DCexp).

In other words: mean(DCexp) fits better the DCbulk than median(DCexp).