SimpleDIVA User Manual

Authors: Frank H. Guenther, Alfonso Nieto-Castañón, Riccardo Falsini, Elaine Kearney, & Hasini R. Weerathunge

Version 1.2, Released May 14 2020

Contents

SimpleDIVA User Manual
- Contents
- Quick Start Guide
- Introduction
- Installation
  - Windows
  - Mac
  - Linux
- Examples
  - Individual data or mean data across subjects
  - Data from multiple subjects
  - Data from multiple subjects across different experimental paradigms
- Advanced Options
- Uninstall SimpleDIVA
  - Windows
  - Mac
  - Linux
- FAQ
  - Can I fit blocked data as well as individual trial data?
  - When would I want to fix the model parameters? And what values do I use?
  - Can I fit data from an experiment that used masking noise?
  - What about missing data?
  - Where can I find the latest version of the app?
  - How do I cite the software?
- License

Quick Start Guide

- Step 1: **Install** the SimpleDIVA application (see below for details).
- Step 2: **Format your dataset** as a csv file with one row per trial/block of the experiment. Column 1 contains the perturbation value for each trial (as a decimal percentage; e.g., a 30%

downward shift is -0.3). Each subsequent column contains response data for an individual subject.
- Step 3: Start the **SimpleDIVA** app.
- Step 4: Enter your csv data file as an '**Input File**'.
- Step 5: Click on the '**Fit Data**' button.

Introduction

This app allows a user to apply a simplified version of the DIVA model to data collected from auditory sensorimotor adaptation experiments. Specifically, the app takes as input a spreadsheet (.csv) that has one row per trial/block of the experiment. The first column contains the perturbation value for each trial (as a decimal percentage; e.g., a 30% downward shift of the first formant (F1) is -0.3), and each subsequent column contains data for an individual subject (one data point per trial/block per subject). The software then finds optimal fit values for three DIVA model parameters: an auditory feedback control gain, a somatosensory feedback control gain, and a trial-to-trial learning (adaptation) rate. The model's fit to the data is then displayed and statistics for the fit are provided. The resulting parameter values provide a mechanistic interpretation of the experimental data.

The three DIVA model parameters capture the following aspects of speech motor control :

- Auditory feedback control gain (α_A): When an auditory feedback perturbation is applied, the auditory feedback controller detects a difference between the target sound and the actual production, and attempts to compensate for the perturbation. A higher α_A leads to a higher compensatory response to an auditory perturbation.
- Somatosensory feedback control gain (α_S): When an auditory feedback perturbation is applied, the somatosensory feedback controller attempts to keep the vocal tract in the normal somatosensory configuration for the sound, thus partially counteracting the compensatory response. A higher α_S leads to a decrease in the compensatory response to an auditory perturbation.
- Trial-to-trial feedforward learning rate (λ_FF): When an auditory feedback perturbation is sustained over a number of trials, the feedforward command for the next trial is updated by some fraction of the feedback-based corrective command for the current trial. A higher λ_FF leads to a larger fraction of the corrective command being added to the feedforward command for the next trial.

Installation

Installer files are provided for Windows, Mac, and Linux.

Windows

1. Download SimpleDIVA_WinInstaller, unzip, and click to run.
2. Click '**Yes**' to give SimpleDIVA permission to install.
3. Click '**Next >**'.

4. Choose installation folder for SimpleDIVA; the default is to save the app to C:\Program Files in a subfolder called 'GuentherLab'. You can also opt to add a shortcut to the desktop. Click '**Next >**'.
5. Choose installation folder for MATLAB Runtime 2017b; the default is C:\Program Files in the MATLAB subfolder. If Runtime 2017b already exists in this folder, you will not have to choose an installation folder and Runtime will not be installed again. Click '**Next >**'.
6. Click '**Install >**'. Note that if Runtime 2017b needs to be installed, the installation will take a few minutes. Otherwise installation should complete in less than 20 seconds.
7. Click '**Finish**'.
8. Go to the folder where you installed SimpleDIVA. Inside SimpleDIVA, there are four subfolders (appdata, application, sys, and uninstall).
9. Go to the '**application**' subfolder and move 'example_data.csv' and 'example_data2.csv' to another location on your computer outside of C:\Program Files (e.g., Desktop).

Mac

1. Download SimpleDIVA_MacInstaller, unzip, and click to run. You may get a warning that says the installer can't be opened because it is from an unidentified developer. If this happens, right-click on the installer, and click 'open' and then click 'open' again in the pop-up window.
2. Provide your user name and password to allow java to make changes.
3. Click '**Next >**'.
4. Choose installation folder for SimpleDIVA; the default is to save the app to the Applications folder in a subfolder called 'GuentherLab'. Click '**Next >**'.
5. Choose installation folder for MATLAB Runtime 2017b; the default is the Applications folder in the MATLAB subfolder. If Runtime 2017b already exists in this folder, it will not be installed again. Click '**Next >**'.
6. Click '**Install >**'. Note that if Runtime 2017b needs to be installed, the installation will take a few minutes. Otherwise installation should complete in less than 20 seconds.
7. Click '**Finish**'.

Linux

Note: The Linux installer has not been fully tested due to limited access to a Linux machine.

1. Download SimpleDIVA_Linux, go to the download folder, and unzip.
2. Run the following two commands:

   ```
   chmod +x SimpleDIVA_LinuxInstaller.install
   ```

   ```
   ./SimpleDIVA_LinuxInstaller.install
   ```

3. Follow the command line prompts to complete instllation.
4. If the automatic installation (steps 2-3) does not work, please refer to the manual installation instructions in the readme.txt file in the SimpleDIVA_Linux directory.

Examples

We have included two example data files to get started with the SimpleDIVA app. Both files contain simulated data:

- example_data.csv has two columns: the first contains the perturbation values and the second contains F1 data for an individual subject (or potentially, mean data across multiple subjects)
- example_data2.csv has 11 columns: the first contains the perturbation values and columns 2-11 contain F1 data for 10 different subjects

Here we demonstrate how to apply SimpleDIVA to these datasets.

Individual data or mean data across subjects

1. Go to the folder where you installed SimpleDIVA.
2. Go to the '**application**' subfolder and click on '**SimpleDIVA**' to launch the app.
3. Under '**Input file(s)**', click '**Browse**' and then select the 'example_data.csv' file. The file path and file name will be shown in the top box, and underneath you will see information about the data (number of subjects, number of time points, and the maximum perturbation magnitude). This information can be used as a sanity check that the correct file was selected and loaded as expected. You can edit the '**Output file**' name to be something more meaningful than 'Default_Out1'.
4. Under '**Model Parameters**', you have the option to **fix** or **fit** the three model parameters (auditory feedback control gain (α_{A}), somatosensory feedback control gain (α_{S}), and trial-to-trial feedforward learning rate (λ_{FF})). Fixing the parameters means setting them to a pre-determined value rather than having the model find an optimal value, whereas fitting the parameters means having the model find optimal values. Generally, we want to fit the three parameters so that the model will estimate the contribution of all three parameters to the data. Scenarios requiring fixed parameters will be explained in the FAQ section.

   For each fitting parameter, you can also set **lower (LB)** and **upper bounds (UB)**. These correspond to the minimum and maximum values that the model can solve for. As a default (**recommended**), they are set to 0 and 1 for lower and upper bounds, respectively, as this covers the expected range for a neurotypical population. Take the auditory feedback control gain (α_{A}) as an example. An α_{A} of less than 0 would mean that a subject responds in the direction of the perturbation instead of opposing it (exacerbating instead of correcting for auditory errors), while an α_{A} of greater that 1 would mean a subject overcompensates for auditory errors. Neither of these scenarios is likely true of a neurotypical subject over the long run, but session-to-session and trial-to-trial variability may lead to a dataset that would lead to optimal parameter values that are outside the 0 to 1 range. You may wish to adjust these ranges to fit clinical populations or to fit individual subjects with highly variable perfomance measures. However, in the latter case, we recommend instead using **regularization** (described under Advanced Options below), which will seek a solution that maintains reasonable values for parameters, even for subjects with noisy data.

The **model equations** are also provided. For concreteness, we have filled F1 into the equations (as if modeling a first formant adaptation study), but F1 can be substituted for any auditory dimension being studied (e.g., second formant, fundamental frequency).

5. Click '**Fit Data**'. The model results will appear in the lower section of the app.

   The left panel figure shows the experimental data (blue) and the model fit (red) as a time series over the number of trials in the experiment. The y-axis units are the same as the input data (F1 (Hz) for example_data.csv).

   The right panel figure shows the model fit parameters: a Pearson's correlation (*r*) describes the relationship between the data and model fit, and the optimized fit parameters are given for α_A, α_S, and λ_{FF}.

   The final results shown in the app are the quantitative values for the model fit parameters. The values for *r*, α_A , α_S , λ_{FF} correspond to those plotted in the right panel figure. Additional values are provided for Root Mean Square Error (RMSE; a normalized value that captures how well the model fits the data) and the target F1 ($F1_T$; calculated as the mean F1 during the baseline phase of the experimental data).

6. Under '**Options**', there are three further options that affect the fitting of the data and the left panel figure.

   When '**Fit mean data**' is selected, the model is fit to the mean of all subjects in the .csv file. When 'Fit mean data' is not selected, the model is fit to each subject separately. If there are only data from a single subject, selecting this option will have no effect on the model output, but it is best to keep 'Fit mean data' selected so that the quantitative values for the model fit parameters are shown.

   When '**Include full compensation plot**' is selected, the plot will include a green line showing what full compensation would look like; that is, if F1 was increased by 30%, full compensation would be a 30% decrease in F1. We rarely observe full compensation - formant adaptation experiments tend to show about 30% compensation, while pitch adaptation experiments show about 50% compensation.

   When '**Include SEM bounds**' is selected, the mean of the experimental data and the standard error across subjects will be shown in the left panel figure. If there are only data from a single subject, standard error cannot be calculated and selecting this option will have no effect on the figure.

   We advise against changing the parameters available under '**Advanced Options**' (with the exception of **regularization** when fitting noisy individual subject data) unless you are familiar with those aspects of the model. Advanced Options will be addressed below.

7. In addition to the display of results in the app, three **output files** are saved to the folder where you selected the input data from:

   - [Output file name]_stats.txt: Records the file path/ file name of the input file and the resulting optimized model parameters.
   - [Output file name].csv: Records time-series data of the perturbation magnitude and the model fit.

- [Output file name].fig: Saves the model figure as a MATLAB .fig file.

Data from multiple subjects

1. The above steps for using SimpleDIVA with 'Individual data or mean data across subjects' also apply to data from multiple subjects. The only differences are as follows:
    - Instead of selecting 'example_data.csv' as the '**Input file**', select 'example_data2.csv'.
    - To fit the model to the mean of these subjects (10 subjects in the case of 'example_data2.csv'), select '**Fit mean data**'.
        - Selecting '**Include SEM bounds**' under '**Plotting Options**' will now result in error bars being shown in the left-panel figure.
    - To fit the model to each subject separately, unselect '**Fit mean data**'. The plots for each individual subject will appear one by one in the lower section of the app. Each figure will be saved as before to the folder where you selected the input data from (the output file name will be appended with a subject #). See **regularization** under Advanced Options for tips on handling noisy individual subject data.

Data from multiple subjects across different experimental paradigms

Another possibility is to fit multiple data files simultaneously with one set of model parameters. You may want to take this approach if you have multiple subjects within a study who completed different protocols or if you want to fit data from multiple studies together (e.g., see Figure 12, Kearney et al., 2020).

The two example data files have different experimental paradigms. 'example_data.csv' has a short ramp phase and a total of 180 trials, whereas 'example_data2.csv' has a gradual ramp phase and a total of 160 trials. Note that each variation in experimental paradigm requires a different csv file.

1. Under '**Input file(s)**', click '**Browse**' and then select 'example_data.csv' and 'example_data2.csv' (there is no limit on the number of files that can be fit simultaneously, it will simply take a little longer).
2. You can click on the file path/name under 'Input file(s)' to display the info for a given file (number of subjects, number of time points, and the maximum perturbation magnitude).
3. To fit all files, click '**Select All**'. To fit a subset of the files, use ctrl + click (Windows) or command + click (Mac) to select the target files.
4. Select '**Fit mean data**' and then '**Fit data**'.

    The left panel figures includes a subplot of the data and model fit for each .csv file (the first file you entered will be on the bottom). The right panel figure shows the model parameters, optimized to fit across all of the files.

    Note that *r* is calculated across all trials in all files, and is influenced more heavily by datasets with more trials.

Advanced Options

Advanced options are available with the recommendation that they are best left at their default values.

Three additional parameters can be fitted in the model. When 'fixed', these parameters are set to default values.

- Auditory error nonlinearity (ϵ): Applies a scaling factor that weights smaller perturbations relatively more than larger perturbations. Default value = exp(-1), which is equivalent to no nonlinearity.
- Auditory target learning rate (λ_{AT}): Determines how much the auditory target is updated on each production. Default value = 0, corresponding to the assumption that the auditory target for a sound does not change appreciably over the course of an experiment.
- Somatosensory target learning rate (λ_{ST}): Determines how much the somatosensory target is updated on each production. Default value = 0, corresponding to the assumption that the somatosensory target for a sound does not change appreciably over the course of an experiment.

The **regularization** option is useful when fitting individual subject data with high subject-to-subject variability. An indication to use at least a small amount of regularization when fitting individual subjects is when no regularization results in one or more of the parameters' optimal values being at the end of its bounded range (by default, 0 or 1).

When choosing this option, a regularization term acts to keep parameter values from drifting too far from "central" values. By default (**recommended**), the model will use the optimized parameter values found for the mean of the dataset as the central values. It is also possible, however, to set your own custom central values.

To use the regularization function:

1. Check '**Enable regularization**'.
2. In the value box, input the regularization factor, which in effect determines how strongly the parameter values are "punished" for straying far away from the central values during optimization . This value is typically less than 0.1 and defaults to .001. A good regularization factor is one that (a) eliminates most or all cases in which the optimized parameter is at the end of its range, while (b) still allowing for substantial between-subject differences in optimized parameter values.
3. If you wish for the model to use the paramaters from the mean data fit as the central values for regularization, make sure that the '**Use parameters from mean data fit for regularization**' option is checked.
4. Otherwise, uncheck '**Use parameters from mean data fit for regularization**' and switch back to the '**Regular Options**'. The value boxes for each fitting parameter will now be editable. These values are the central values that will be used in regularization.

Uninstall SimpleDIVA

You can uninstall the SimpleDIVA app at any time. Follow the instructions below for the platform you are using.

Windows

1. Go to the folder where you installed SimpleDIVA.
2. Go to the '**uninstall\bin**' subfolder and click on '**uninstall.exe**' to uninstall the app.

Mac

1. Go to the folder where you installed SimpleDIVA and simply delete the folder.

Linux

1. Go to the folder where you installed SimpleDIVA
2. Run the following command:

```
rm SimpleDIVA
```

FAQ

Can I fit blocked data as well as individual trial data?

Yes! For some experiments, it makes sense to calculate the mean value over every *N* trials in order to reduce the variability in the experimental data. For example, if the stimuli consisted of 3 words that were randomly presented in blocks of 3, it would be appropriate to estimate the mean of every 3 trials. Similarly, if only one stimulus was used throughout the experiment, it would be appropriate to again estimate the mean of every 3 trials. If, however, multiple stimuli were used but were not presented in a blocked fashion, it would be best to retain the individual trial values. You would need to block the data (take the mean over *N* trials) when preparing the csv file for input into SimpleDIVA. No other changes are required when interacting with the SimpleDIVA app.

When would I want to fix the model parameters? And what values do I use?

It's possible to use optimized model parameters from one experimental condition (let's call it condition1) to predict the model fit to another experimental condition (condition2). To do this, first fit condition1.csv as described previously. When fitting condition2.csv (again as Input file 1), select 'Fixed' for all three model parameters. In the 'Value' boxes, insert the optimized model parameters from condition1.csv. Remember that these values have been saved to an output file called [Output file name]_stats.txt. The model will then use these values when fitting the data and they will remain unchanged in the output. The resulting *r* and *RMSE* values will reflect how well the parameters from condition1 predicted the model fit in condition2.

Can I fit data from an experiment that used masking noise?

Yes! Masking noise is sometimes used during the hold phase of experiments to assess adaptation in the absence of auditory feedback. To indicate which trials have masking noise, replace the perturbation value in column 1 of the csv file with 'NaN' (not a number). For the noise-masked trials, no error will be detected between the auditory target and auditory feedback but somatosensory feedback errors will still be detected and corrected.

What about missing data?

Missing data points can be handled in the model. Either leave the data point blank or replace it with 'NaN'.

Where can I find the latest version of the app?

http://sites.bu.edu/guentherlab/software/simplediva-app

How do I cite the software?

Please cite our paper in *Frontiers in Psychology*:

Kearney, E., Nieto-Castañón, A., Weerathunge, H. R., Falsini, R., Daliri, A., Abur, D., Ballard, K. J., Chang, S-E., Chao, S-C., Heller Murray, E. S., Scott, T. L., & Guenther, F. H. (2020). A simple 3-parameter model for examining adaptation in speech and voice production. *Frontiers in Psychology*, *10*, 2995. https://doi.org/10.3389/fpsyg.2019.02995

License