**tcaCALC20190901Delta**

**Quick Start Document**

**J.R. Alger**

**November 26, 2024**

[**Jeffry.Alger@UTSouthwestern.edu**](mailto:Jeffry.Alger@UTSouthwestern.edu)

[**jeff@neurospectroscopics.com**](mailto:jeff@neurospectroscopics.com)

**New features of tcaCALC20190901Delta**

* There are now four versions of tcaCALC20190901. These are
  + tcaCALC20190901Alpha
  + tcaCALC20190901Bravo
  + tcaCALC20190901Charlie
  + tcaCALC20190901Delta
* Each includes bug fixes and improvements. In general, users should run the latest version (tcaCALC20190901Delta).
  + The tcaSIM version that performs the underlying metabolic simulation is unchanged through the above named revisions
* The primary motivation for tcaCALC20190901Delta was as follows. In Aug 2022 I received reports of errors that occurred when two or more substrate isotopomers of the same molecule were being fit. For instance when Lac oxo and Lac oox were both substrates and were unknown (ie were being fit). In this situation the search routine would lead to situations in which the sum of the fitted substrates is greater than 1.0 (eg Lac oox + Lac oxo > 1) which is an error. tcaCALC20190901Delta handles this problem in the same way that the problem of LDH < 0 is handled. It gives such instances a large error value so the search routine will prefer to go elsewhere.
* A second motivation was the need for the user to enter bounds on the searched parameters. Versions before tcaCALC20190901Delta used default bounds on the searched parameters.
* Addition of user-specified search bounds can also help with situations in which two substrate isotopomers of the same molecule were being fit. For instance if Lac oxo and Lac oox are both substrates are to be fit, the user can set the upper limits on Lac oxo and Lac oox so that the sum of all Lac isotopomers is less than 1.0 throughout the search.
* The general user interface (GUI) does not support situations in which two or more substrate isotopomers of the same molecule are being fit. Users who wish to perform such an analysis must create a ‘metabolic model’ spreadsheet (a .csv file) and select it as input. An example is provided in the distribution in InputTemplates\MetabolicModelInput\ MetabolicModels\_All.csv. This example illustrates a situation in which the following substrate isotopomers are to be fit: Lac oox, Lac oxo, FA xx.
* When spreadsheet entry of the metabolic model is used it is necessary for the user to specify the upper and lower bounds for all fitted parameters. See the example in InputTemplates\MetabolicModelInput\ MetabolicModels\_All.csv. All to be fit parameters are entered as Value:LowerBound:UpperBound. It is necessary to include the colons (‘:’) and for there to be no white spaces.
* tcaCALC requires installation of the MATLAB Statistics and Machine Learning Toolbox
* tcaCALC has been tested with MATLAB 2024a

**New features of tcaCALC20190901 (compared to tcaCALC20180427beta)**

* tcaCALC has been completely rewritten.
* The tcaCALC GUI has been completely redesigned.
* tcaCALC20190901 is the first version of tcaCALC that explicitly deals with the possibility that the metabolic network being studied may have some pathways that are turned off. Specifically, the software considers all feasible metabolic situations in which one or more of the four primary pathways (PDH, PK, Ypc, Ys) may be turned off. These combinations are shown schematically on the opening GUI (labeled Model01 – Model16)
* tcaCALC20190901 is the first tcaCALC version that does metabolic calculations using the 2018-19 tcaSIM metabolic simulation software.
* tcaCALC20190901 has been updated to support the following measured data inputs: isotopomers, 13C-NMR multiplets, fractional enrichments, isotopologues or combinations.
* tcaCALC spectral display module has been updated to display the frequencies of 13C NMR multiplet signals.
* tcaCALC 13C NMR spectrum reader has been updated to read from a .CSV file. The spectrum must be reconstructed, phased and baseline fixed using other software.

**General Warnings**

* The MATLAB Development Environment (DE) must be installed on the local system to run the source software.
* Win 10 installer and/or Win 10 executable can be provided for users who do not have access to the MATLAB DE. Send email to the author to request an executable version.
* Operation of the source code on MacOS and linux/unix has not been tested.
* Executables for MacOS and linux/unix cannot be built because of cost limitations. Users who operate MacOS and/or linux/unix systems must purchase and install the MATLAB DE to run the software.
* The software was developed with MATLAB 2017b. Performance with earlier or later MATLAB versions is not guaranteed.
* The GUI appearance is dependent on hardware settings. The GUI was set up using a laptop computer running Win 10 with the following settings: Resolution = 3200 x 1800. Magnification (change the size of text …) = 150%. These settings should be used for best appearance.
* The GUI does not trap for all potential human input errors. Some knowledge of the underlying metabolic principles is required.
* This software version has been developed with focus on fitting of the default set of metabolic models defined in the opening GUI. Turning on additional pathways (eg GK, YGln) or altering certain default values may produce unanticipated problems.

**To start the software**

Double left click on tcaCALC20190901.m to open MATLAB.

Left click on the MATLAB DE green arrow to start the software.

**GUI functions**

Models01 – Model16 checkboxes

These checkboxes permit the users to eliminate specific models from consideration. Checkboxes that have a checkmark define the metabolic models that will be fit.

All Models checkbox

Checking this checkbox will result in all default models being considered.

PK = 0 checkbox

Checking this checkbox will eliminate any model that has nonzero value of PK from consideration.

Ys = Ypc = 0 checkbox

Checking this checkbox will eliminate any model that has nonzero values of Ys and Ypc from consideration.

Ys = 0 checkbox

Checking this checkbox will eliminate any model that has nonzero value of Ys from consideration.

Load Metabolic Models CSV pushbutton

This pushbutton will load metabolic model information from a .csv formatted spreadsheet. This allows the user to quickly and reliably define complex metabolic models without tedious clicking and typing information into various GUI fields. InputTemplates\MetabolicModelInput\MetabolicModels\_All.csv defines a template for the metabolic models file that can be loaded. The user may edit this file to make adjustments to the metabolic models and the fitting parameters.

*Warning: Loading metabolic information from a .csv spreadsheet will override all GUI settings.*

*Warning: The metabolic information that is loaded from a .csv spreadsheet IS NOT reported in the GUI.*

*Warning: Output is keyed on having metabolic models identifiers ‘Model01 – Model16’. Use of different model identifiers may produce unanticipated problems.*

Display Complete Pathway Diagram pushbutton

This pushbutton provides a metabolic pathway diagram that illustrates key pathways that are known to tcaCALC.

*Warning: This pathway diagram is meant to provide information only. It is not a GUI and does not have active buttons.*

Display/Change Model(s) Parameters

This pushbutton opens a child GUI that allows the user to change metabolic model parameters. This GUI is meant to be self-explanatory.

*Warning: Fitting of parameters other than PDH, PK, Ypc and Ys has not been tested.*

13C enriched substrates checkboxes

Selecting these checkboxes allow the user to define which 13C enriched substrates were used in the metabolic tracing experiment that tcaCALC is evaluating. Each of these opens a child GUI that allows the user to exactly specify which isotopomers of the substrate were supplied to the metabolizing system in the experiment.

*Warning: The software expects that the user has exact knowledge of all substrate isotopomers for a particular substrate. For instance if the experiment used Lac xxx = 0.90, the user should do his/her best to enter the true enrichments of the remaining seven Lac isotopomers.*

*Warning: The substrate GUIs trap for input errors. If the user enters a negative value the software sets the value to zero. Values greater than 1 are set to 1. The sum of all isotopomers must be equal to 1. If this is not the case, the software resets the substrate to natural abundance*

13C enriched substrates Fit checkboxes

These checkboxes allow the user to define whether the substrate 13C enrichment is known exactly or should be fit as an unknown. A checkmark indicates the substrate 13C enrichment is should be fit as an unknown.

*Warning: tcaCALC is not able to fit multiple unknown isotopomers of the same substrate. For example tcaCALC can fit unknown values of Lac xxx or Lac oox but it cannot simultaneously fit both. If the user defines that multiple isotopomers of one substrate are present at levels above natural abundance, tcaCALC will fit only the largest of these as an unknown.*

*Warning: This version of tcaCALC does not support user definition of starting values, upper bounds or lower bounds for unknown substrate enrichment. If the Fit checkbox is selected, the starting value, lower bound and upper bound are set to 0.5, 0.0, and 1.0 respectively.*

Substrates exact natural abundance checkbox

Checking this checkbox indicates all substrates being consumed by the metabolic system have their true natural abundance. For instance,

Lac xoo = Lac oxo = Lac oox = 0.01

Lac xxo = Lac oxx = Lac xox = 0.01\*0.01,

Lac xxx = 0.01\*0.01\*0.01

Lac ooo = 1 – (Lac xoo + Lac oxo + Lac oox + Lac xxo +Lac oxx + Lac xox + Lac xxx).

An unchecked checkbox sets the uniformly unlabelled isotopomer of all substrates to 1.

*Warning: Checking this checkbox resets all substrates to their natural abundance.*

Load/Analyze 13C NMR Spectrum

This pushbutton opens a child GUI that helps the user determine the relative multiplet intensities in a measured 13C NMR spectrum. See below for further instructions.

Load Measured Data CSV

This pushbutton allows the user to load measured values of isotopomers, 13C-NMR multiplets, fractional enrichments, isotopologues or combinations. Four unique formats for the file that is read are allowed. Templates for these files can be found in the InputTemplates folder within the source code distribution. The most thoroughly tested format is the GeneralDataArrayInput.csv file within the GDAInput folder.

*Warning: tcaCALC will crash if the input CSV files have been incorrectly formatted. Users should follow the examples defined in the template exactly.*

Run tcaCALC pushbutton

This pushbutton initiates the processes of fitting all the user-defined metabolic models to the measured data. See below for a specification of the output files that are produced at the conclusion of the fitting process.

*Warning: tcaCALC will crash if metabolic models and measured data have not been loaded correctly.*

*Warning: This software version was primarily tested with 13C NMR multiplets. Fitting of metabolic models to measured isotopmers, isotopologues, fractional enrichments has not been fully tested. Fitting of models to combinations of 13C NMR multiplets, isotopmers, isotopologues, fractional enrichments has not been tested.*

Quit pushbutton

Self explanatory.

**13C NMR spectral fitting with tcaCALC**

REQUIRED file setup

Agilent FIDs must be stored in

XXX/Agilent/YYY/

where XXX and YYY are any legal folder names

The fid file can have any legal filename

A file named procpar must be present and must specify the fid acquisition parameter (see examples in source distribution folder)

Bruker TopSpin processed spectra must be stored in

XXX/BrukerTopSpin/YYY/pdata/1/1r

where XXX and YYY are any legal folder names

A proc and procs file must be present (see examples in source distribution folder)

processed spectra saved as .CSV files must be stored as

XXX\SpectrumCSV\YYY.csv

where XXX and YYY are any legal file/folder names

The CSV file must contain the following lines

SF,uuu,

ImagSign,vvv,

PPM,Real,Imag

Where

uuu is the spectrometer frequency

vvv provides the sign of the imaginary points (usually -1), but could be +1 for some NMR systems and analysis software.

Following these lines, are lines that provide the PPM, the real intensity and the imaginary intensity for each point in the spectrum. For example

221.864372,-5.946579,178.326019

The PPM scale should be set up so that Lac C3 S appears near 20.8 ppm

An example file can be found in the software distribution at

\InputTemplates\SpectrumCSV

Select Load FID, Reconstruct & Autophase

Agilent FIDs take a few minutes for autophasing. Bruker spectra are read as TopSpin output which are already phased and therefore load faster.

Click Lac C3 S radio button first (REQUIRED), then left click on top of Lac C3 S signal

Repeat above instructions to assign locations of all other singlets

Inspect the frequency positions of the signals of all multiplets as they appear in the GUI. Corrections can be made by reselecting specific singlet signals.

Note that tcaCALC uses known values for isotope shifts

Select Fit Multiplets (pull down menu).

About 5 minutes are necessary to fit all multiplets. Output is found in

XXX\AgilentFIDs\YYY\IMA1Results\

or

XXX\BrukerTopSpin\XXX\pdata\1\ IMA1Results\

When fit is completed return to tcaCALC window

*Warning: The tcaCALC spectral fitting module will produce fits for Lac multiplets. However Lac multiplets are not valid input for the tcaCALC metabolic fitting module. The presence of Lac in the MultipletResults.csv file will cause tcaCALC to crash. The Lac lines in the MultipletResults.csv file should be deleted. Alternatively the user can assume Lac/Ala transamination is in equilibrium and manually reassign all Lac lines in MultipletResults.csv file to Ala, which is a valid tcaCALC measured input.*

Select Load Measured Data CSV

Select the MultipletsResults.csv file

Select Run tcaCALC to perform the metabolic fits.

**Output**

On conclusion the software produces multiple results files in the same folder as the input data files.

*Warning: The user must have write permission on this folder.*

A .png formatted ‘ModelsResults’ file shows plots of all fitted parameters organized from best to worse according to the Akaike Information Criterion corrected for small sample sizes (AICc). The lowest AICc defines the mostly likely model that explains the data. In these plots open circles convey situations in which the model had a parameter value explicitly defined as zero. Solid circles and error bars convey best fit results for the particular parameter and model. Error bars are determined by error propagation from estimates of uncertainty for input measured parameters. The Ln(Residual) are the natural logarithms of the sum of squares residual. In these plots ADp is the p-value for the Anderson-Darling test for random residuals. Low values of ADp (ie < 0.05) suggest the residuals are not randomly distributed and the model may be ‘overfitting’ the data.

*Warning: The software fixes the standard deviation of all input measured parameters to be 0.1. The present version does not allow the user to alter these estimates.*

A .png formatted file that displays the results for each of the fitted metabolic models is produced. These are named according to the AICc from best to worst.

A .csv formatted spreadsheet named ‘ModelsResults.csv’ provides the detailed numerical values of fitted values, residuals, uncertainities etc. This can be used for further more detailed analyses.