

NumKinFit is a python-based (Python 3.7) program created to numerically calculate time resolved concentrations of reacting species. Moreover, it can also fit initial concentration and rate coefficient parameters of a reaction model to experimental measurements. Along with standard modules in Python 3.7 (default install with Anaconda) such as numpy, scipy, matplotlib and tkinter, a fitting module based on scipy, lmfit, are required to run this program. The python script, a manual and example data sets are provided in the https://github.com/rc13564/num_kin_fit_git repository. Listed below are the various inputs that should be provided by the user for kinetic modelling and fitting using the GUI version of this program:

- 1) The reactions should be entered in the 'Reaction Model' box. Each reaction is entered in a different line. An example of the reaction entry is provided as the default input. The reactants and products are separated by a '=' symbol. Each reaction needs a variable for rate coefficient separated from the reaction by a ':' symbol. All the letters and symbols are separated by a space key. Make sure that there are no empty lines!
- 2) All the molecules involved in the model should be entered in the 'Reaction Species' entry. The species labels should be consistent with the 'Reaction Model' entry and should be separated by a ':' symbol.
- 3) Concentration of all the species should be provided in the 'Initial Concentrations, CO' entry. All the values should be entered in the same order as the 'Reaction Species' entries and each value are separated by a ':' symbol. The unit of concentrations should match the unit for 'Initial Rate Coefficient, k' entry which is described later. [Be aware that the gas phase and condensed phase units are generally different]
- 4) The 'Time Start' and 'Time Stop' entries for the simulation should be in the units of seconds. The 'Number of Steps' specified determines the step size of the simulation time.
- 5) Any species involved in the model can be selected to be plotted after numerical integrations. These are specified in the 'Species to Plot' entry which should be consistent with the species label in 'Reaction Species' entry. Each species label should be separated by a ':' symbol.
- 6) Initial value of the rate coefficients for each reaction should be provided in the 'Initial Rate Coefficients, k' entry. The rate coefficients are entered in the order that the reaction appear in the model and separated by a ':' symbol.
- 7) Pressing the 'Model' button will result in numerical simulation of the 'Reaction Model' based on the values given for 'Initial Concentrations, CO', 'Time Start', 'Time Stop', 'Number of Steps' and the 'Initial Rate Coefficients, k' entries. A plot window will appear with concentrations of species specified in the 'Species to Plot' entry.
- 8) The default 'Measured Species' entry is 'None' which results in a plot of modelled concentrations. However, if any reaction species label is provided in the 'Measured Species' entry, hitting the 'Model' button will open a dialog box which allows user to enter a measurement text file. The measurement text file should be tab separated and the first column should be time. The order of species label in the 'Measured Species' entry should match the order of the measurement columns after the time column. The simulated and experimental concentrations are then plotted in the same graph and this feature can be useful in initial refinement of the reaction model.
- 9) The 'Fit' button will result in fitting of the model to the experimental concentration of the species provided in the 'Measured Species' Entry. The parameters varied during the fitting process depends on the values for 'k min -Fit', 'k max -Fit', 'CO min -Fit' and 'CO max -Fit' entries. The values for each of these entries should be provided in the same order as the 'Initial Rate Coefficient, k' and 'Initial Concentrations, CO' entries. If the minimum and maximum values are equal, these parameters are fixed during the fitting process.

- 10) Both Model and Fit processes produce traces which can be saved by the user by selecting the 'Yes' option for the Save? dialog box query. For the Model option, the time stamp parameters provided by the user and the modelled concentration at these time stamps will be saved. For the Fit option, the start and stop values are obtained from the measurement data and the number of steps in between will be taken from the entry provided in the GUI. A log file will also be created for both processes which lists all the entries used in the GUI for modelling and fitting.

Example 1: Modelling Competitive Reactions

The example_1.txt file was derived from a data set provided in a public data repository for a research paper (See Reference 1). This file contains data for decay of the simplest Criegee intermediate, CH₂OO, due to self-reaction and unimolecular reaction at 311 K. The first column is time delay between photolysis and probe lasers. The fourth column is the change in ring-down rate upon photolysis which is directly proportional to concentration of CH₂OO. The relationship between ringdown rate and concentration is provided by equation 16 of the paper listed in Reference 2 and using L=106 cm, c = 2.99×10⁻¹⁰ cm s⁻¹, d=7.6cm and σ_{355nm} = 1.1×10⁻¹⁰ cm² molecule⁻¹. The datapoints for negative time and at time zero were deleted. The CH₂OO loss processes are dominant only after time zero. The resultant time column and the corresponding CH₂OO concentration values were saved in the github repository as example_1.txt file.

The self-reaction and unimolecular reaction of CH₂OO is described by the following reaction model:

Reaction Model =

CH₂OO + CH₂OO = Product1 : k1

CH₂OO = Product2 : k2

Reaction Species = CH₂OO:Product1:Product2

Initial Concentrations, C0 = 1E12:0:0

For initial concentration, product concentrations should be zero. For fitting purpose, the best initial guess for the reactant initial concentration would be the first time point concentration value.

Initial Rate Coefficients, k = 1E-10:50

The rate coefficient for self-reaction of CH₂OO is around 10⁻¹⁰ cm³ molecule⁻¹ s⁻¹ (see Reference 2). The rate coefficient for the unimolecular loss processes which combines reactive and physical losses can range from 10 to few 100 s⁻¹ depending on experimental conditions.

Time Start = 0

Time Stop = 0.01

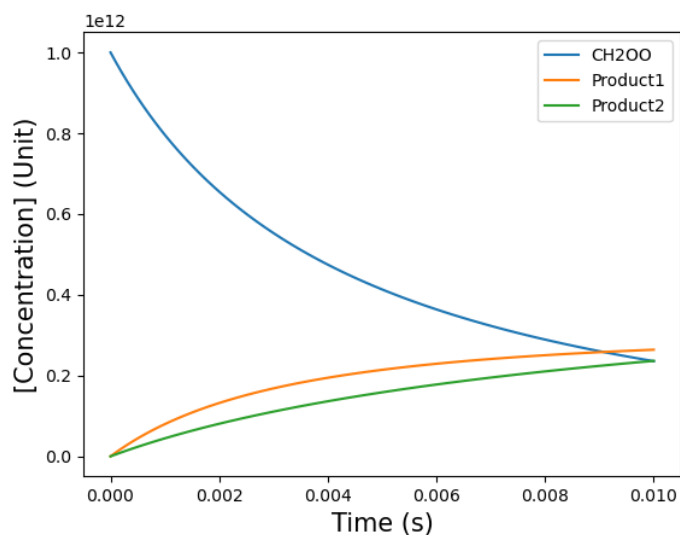
Number of Steps = 100

The time points for the numerical integration should be similar to the time window for the CH₂OO decay measurement we are using for this example. The Number of steps of around 100 gives smooth traces.

Measured Species = None

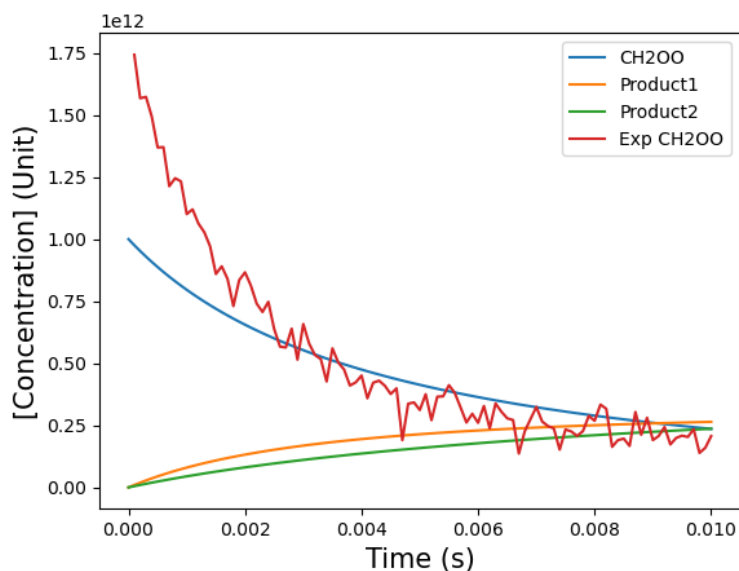
Species to Plot = CH2OO:Product1:Product2

The default value for Measured Species input is None. The Species to Plot labels should be consistent with Reaction Species input. If we press Model now, the GUI will perform numerical integration and show the traces of species which were input in the Species to Plot as shown below:



Measured Species = CH2OO

Change the input for Measured Species as above and press Fit button. A dialog box will appear and select the example_2.txt file. A plot window will appear showing modelled as well as the experimental data sets as shown below:



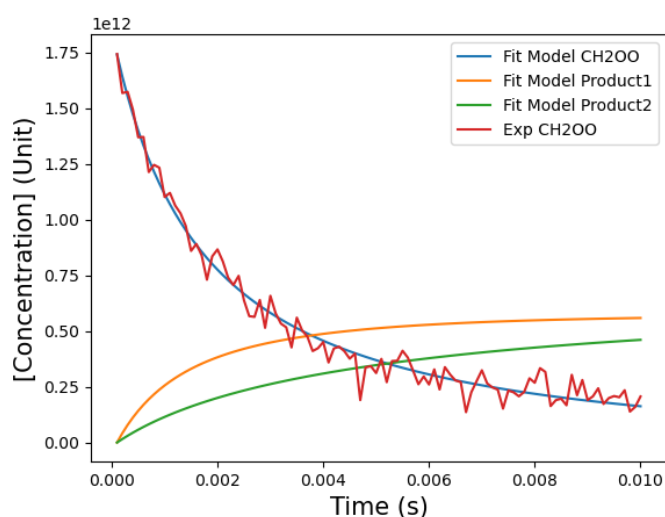
C0 min -Fit = 1E12:0:0

C0 max -Fit= 1E13:0:0

k min -Fit = 1E-11:10

k max -Fit = 1E-9:1000

The initial concentration of the various species can also be varied in the fit. Only CH₂OO is present at time zero. The lower and upper limit for CH₂OO initial concentration values are provided as 1×10^{12} and 1×10^{13} molecule cm⁻³. The minimum and maximum values for initial concentrations of Product1 and Product2 are same and thus these values will be fixed to values provided in the Initial Concentration, C0 input. The k1 value can vary from 1×10^{-11} to 1×10^{-9} cm³ molecule⁻¹ s⁻¹ and k2 value from 10 to 1000 s⁻¹. Pressing the Fit button should give the fitted trace along with the experimental trace as shown below:



The values of the fitted parameters are provided in the python console. Overall, five parameters k1, k2, CH₂OO₀ (initial concentration value for CH₂OO), Product1₀ and Product2₀ were used and the first three were varied in the fit. The error values for the fitted parameters are 1σ value. Correlation between various fitted parameters are also provided.

[[Variables]]

k1: 1.4576e-10 +/- 8.2092e-12 (5.63%) (init = 1e-10)

k2: 92.1173211 +/- 9.85983787 (10.70%) (init = 50)

Initial_CH2OO: 1.7408e+12 +/- 3.4867e+10 (2.00%) (init = 1e+12)

Initial_Product1: 0 (fixed)

Initial_Product2: 0 (fixed)

[[Correlations]] (unreported correlations are < 0.100)

C(k1, k2) = 0.839

C(k2, Initial_CH2OO) = 0.779

C(k1, Initial_CH2OO) = 0.727

! Although the fit result look good, correlations between the fit parameters are quite high. Ideally the fit parameters should have relatively low correlations (<0.5) which in this case may be achieved by systematic measurements for dependences on initial CH₂OO concentrations and other physical conditions to separate out the two reactions in the time domain.

References

1. https://data.bris.ac.uk/datasets/3lflkl8guhjfj2ain57w18p7ue/CRDS/CH2OO_CH3NH2/T_dep_50_Torr/311K/04-06-2016-001.txt
2. <https://pubs.rsc.org/en/content/articlepdf/2015/cp/c4cp04198d>

Example 2: Modelling Product Formation

In example 1, we were able to model self-reaction and unimolecular reaction of CH₂OO and then fit the model parameters to experimental measurements. Here we will perform global fitting for CH₂OO reactions and products. The data set were obtained from a public data repository for a research paper (see Reference 3). The first column shows the time delay (s) between photolysis and probe lasers. The second third column show signals for CH₂OO⁺ and CH₃N(H)CH₂⁺ ions. The CH₂OO⁺ ion is produced from photoionization of CH₂OO whereas CH₃N(H)CH₂⁺ ion is produced from photo-fragmentation of the adduct product from reaction of CH₂OO with CH₃NH₂. The ion signals are in arbitrary unit and directly proportional to the concentration of the neutral species. These values were scaled to 1 as numerical integration will not work properly for very small and large values as then saved to example_2l.txt file in the github repository. For reactions with multiple branching channels, different scaling factor may be needed for various products.

Reaction Model =

CH₂I = CH₂OO : k1

CH₂OO = Product1 : k2

Here CH₂I reacts rapidly with O₂ to form CH₂OO which then reacts rapidly with CH₃NH₂ to form Product1. Both reactions are in pseudo first order as excess concentrations of O₂ and CH₃NH₂ were used for these measurements. Thus, concentrations of O₂ and CH₃NH₂ do not need to be modelled explicitly here. Make sure there are not empty lines!

Reaction Species = CH₂I:CH₂OO:Product1

Initial Concentrations, C0 = 1:0:0

Initial Rate Coefficients, k = 8000:500

Time Start = 0

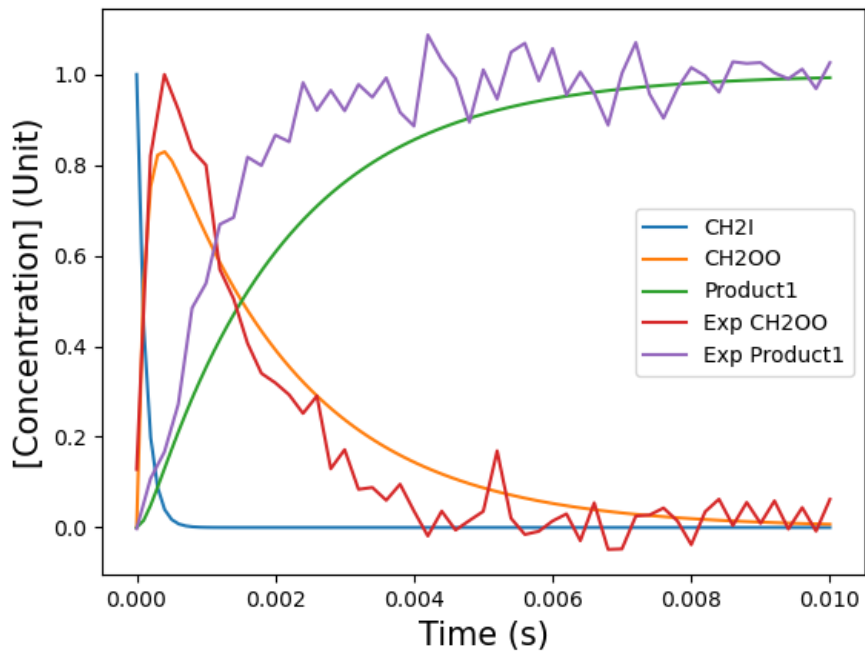
Time Stop = 0.01

Number of Steps = 100

Measured Species: CH₂OO:Product1

Species to Plot = CH₂I:CH₂OO:Product1

Press the Model button and open the example_2l.txt data file. Shown below is the plot you get showing comparison of the model to the experimental data.



The model does relatively good job and a fit was performed using the following bounds for the rate and initial concentration parameters:

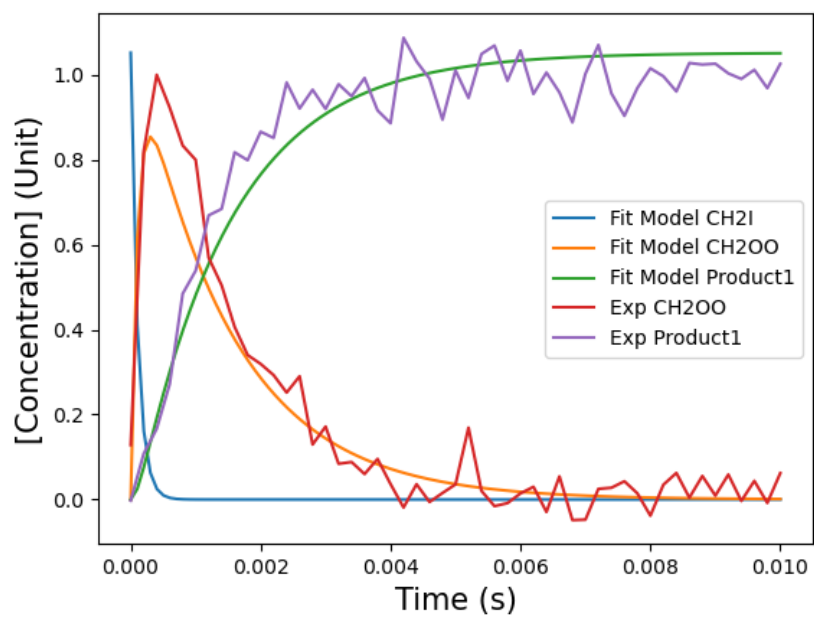
C0 min -Fit = 0.8:0:0

C0 max -Fit = 1.5:0:0

k min -Fit = 1000:100

k max -Fit = 10000:1000

Shown below are the fit results:



[[Variables]]

k1: 9296.37803 +/- 2299.11838 (24.73%) (init = 8000)

k2: 688.839935 +/- 27.9494270 (4.06%) (init = 500)

Initial_CH2I: 1.05190541 +/- 0.01151086 (1.09%) (init = 1)

Initial_CH2OO: 0 (fixed)

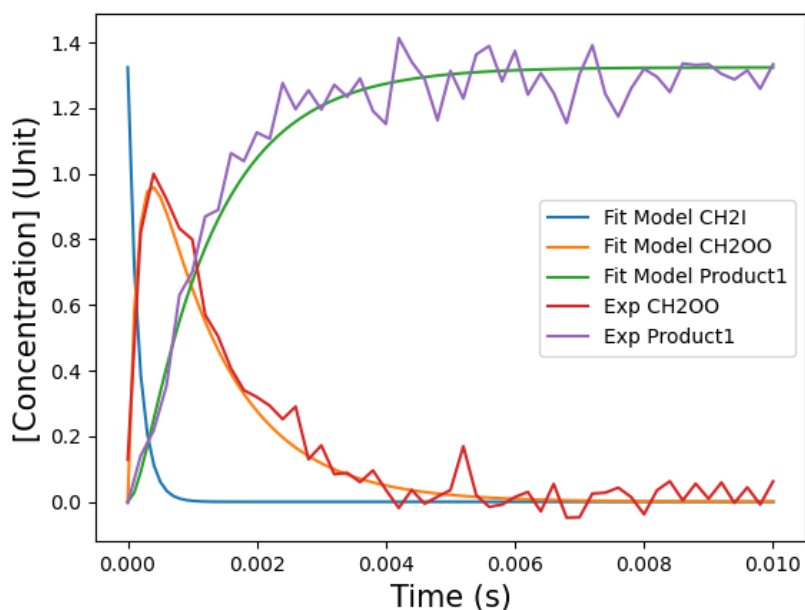
Initial_Product1: 0 (fixed)

[[Correlations]] (unreported correlations are < 0.100)

C(k1, k2) = -0.388

C(k2, Initial_CH2I) = -0.239

! This fit gives relatively good result overall. Only the amplitude of the CH₂OO signal is not reproduced well. This is likely because our model assumes that all the initial CH₂I eventually produces Product 1 so the initial CH₂I and final Product1 signal should be similar. However, we scaled CH₂OO and Product1 signal to 1. In example_2II.txt file, the Product1 signal is scaled to 1.3 and if we repeat the fit we get a much better fit result as shown below:



[[Variables]]

k1: 6111.09083 +/- 761.924384 (12.47%) (init = 8000)

k2: 858.116702 +/- 27.4637654 (3.20%) (init = 500)

Initial_CH2I: 1.32438347 +/- 0.00988239 (0.75%) (init = 1)

Initial_CH2OO: 0 (fixed)

Initial_Product1: 0 (fixed)

[[Correlations]] (unreported correlations are < 0.100)

$$C(k1, k2) = -0.432$$

$$C(k2, \text{Initial_CH2I}) = -0.188$$

$$C(k1, \text{Initial_CH2I}) = -0.106$$

Thus, in the case of fitting to multiple species, we need to know the accurate relative concentration of the measured species.

References

3. <https://data.bris.ac.uk/data/dataset/c16353ddad474649ef43fb77484f8def>