

NumKinFit is a python-based (Python 3.7.4) 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, were used to create this program. The numerical calculations involve integration using the odeint function of scipy to a system of ordinary differential equations. A .py script file (source code), an executable file, a user manual and some example data sets are provided in the [https://github.com/rc13564/num\\_kin\\_fit\\_git](https://github.com/rc13564/num_kin_fit_git) repository. The executable file can be opened in Windows 7 and 10 without installing Python. In order to best run the .py script file, a Python 3.7.4 install along with the modules listed are necessary.

Listed below are the various inputs that should be provided by the user for kinetic simulation 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', 'C0 min -Fit' and 'C0 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, C0' 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.
- 11) The text box and other entry boxes in the GUI are relatively small and inconvenient to use for large reaction schemes. The 'Box Model' button allows users to enter a text file containing reaction schemes and other parameters in such scenario. An example of the input file is provided in Example 3.

### **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, CH<sub>2</sub>OO, 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 CH<sub>2</sub>OO. 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<sup>-10</sup> cm s<sup>-1</sup>, d=7.6cm and σ<sub>355nm</sub> = 1.1×10<sup>-10</sup> cm<sup>2</sup> molecule<sup>-1</sup>. The datapoints for negative time and at time zero were deleted. The CH<sub>2</sub>OO loss processes are dominant only after time zero. The resultant time column and the corresponding CH<sub>2</sub>OO concentration values were saved in the github repository as example\_1.txt file.

The self-reaction and unimolecular reaction of CH<sub>2</sub>OO is described by the following reaction model:

Reaction Model =

CH<sub>2</sub>OO + CH<sub>2</sub>OO = Product1 : k1

CH<sub>2</sub>OO = Product2 : k2

Reaction Species = CH<sub>2</sub>OO: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 CH<sub>2</sub>OO is around 10<sup>-10</sup> cm<sup>3</sup> molecule<sup>-1</sup> s<sup>-1</sup> (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<sup>-1</sup> depending on experimental conditions.

Time Start = 0

Time Stop = 0.01

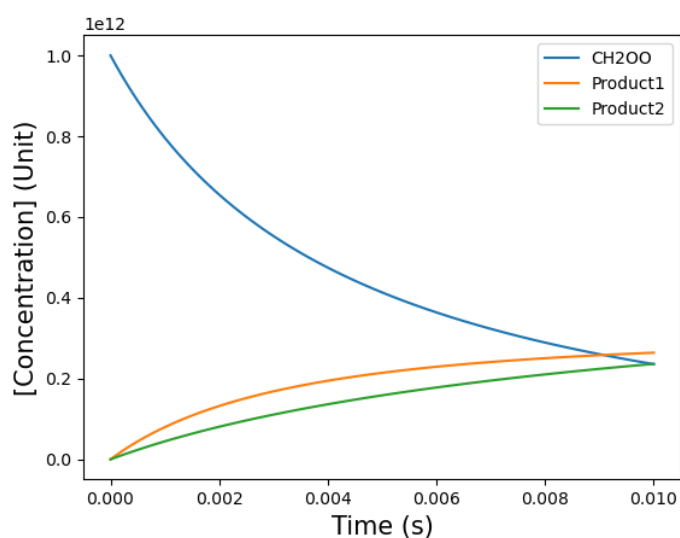
Number of Steps = 100

The time points for the numerical integration should be similar to the time window for the CH<sub>2</sub>OO decay measurement we are using for this example. The Number of steps of around 100 gives smooth traces.

Measured Species = None

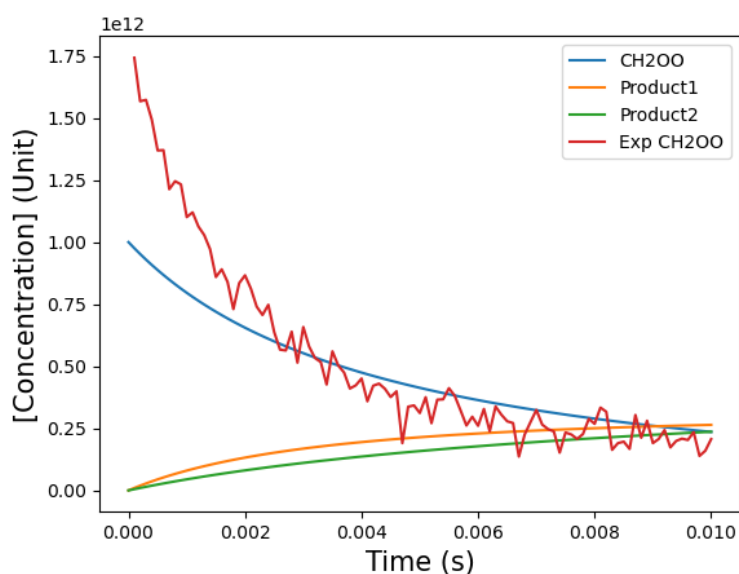
Species to Plot = CH<sub>2</sub>OO:Product1:Product2

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 = CH<sub>2</sub>OO

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:



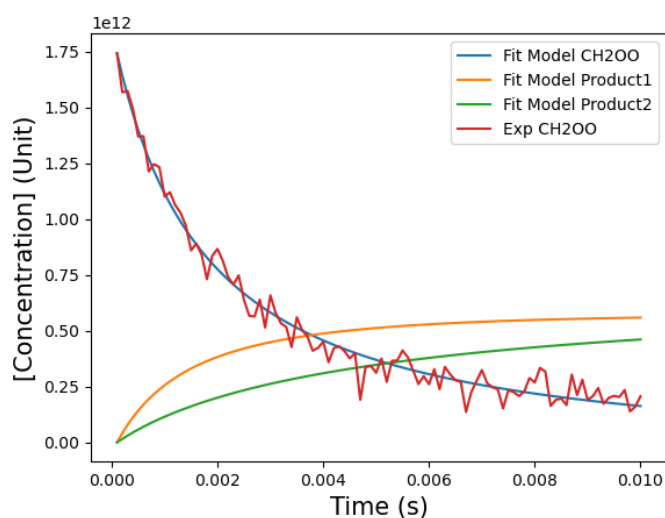
C0 min -Fit = 1E12:0:0

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 CH<sub>2</sub>OO is present at time zero. The lower and upper limit for CH<sub>2</sub>OO initial concentration values are provided as  $1 \times 10^{12}$  and  $1 \times 10^{13}$  molecule cm<sup>-3</sup>. 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  $1 \times 10^{-11}$  to  $1 \times 10^{-9}$  cm<sup>3</sup> molecule<sup>-1</sup> s<sup>-1</sup> and k2 value from 10 to 1000 s<sup>-1</sup>. 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, CH<sub>2</sub>OO<sub>0</sub> (initial concentration value for CH<sub>2</sub>OO), Product1<sub>0</sub> and Product2<sub>0</sub> were used and the first

three were varied in the fit. The error values for the fitted parameters are  $1\sigma$  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 CH<sub>2</sub>OO concentrations and other physical conditions to separate out the two reactions in the time domain.

#### References

1. [https://data.bris.ac.uk/datasets/3lflkl8guhjf2ain57w18p7ue/CRDS/CH2OO\\_CH3NH2/T\\_dep\\_50\\_Torr/311K/04-06-2016-001.txt](https://data.bris.ac.uk/datasets/3lflkl8guhjf2ain57w18p7ue/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 Sequential Reactions**

In example 1, we were able to model self-reaction and unimolecular reaction of CH<sub>2</sub>OO and then fit the model parameters to experimental measurements. Here we will perform global fitting for sequential reactions involving formation of CH<sub>2</sub>OO and its subsequent reaction. 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 CH<sub>2</sub>OO<sup>+</sup> and CH<sub>3</sub>N(H)CH<sub>2</sub><sup>+</sup> ions. The CH<sub>2</sub>OO<sup>+</sup> ion is produced form photoionization of CH<sub>2</sub>OO whereas CH<sub>3</sub>N(H)CH<sub>2</sub><sup>+</sup> ion is produced from photo-fragmentation of the adduct product from reaction of CH<sub>2</sub>OO with CH<sub>3</sub>NH<sub>2</sub>. 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  $\text{CH}_2\text{I}$  reacts rapidly with  $\text{O}_2$  to form  $\text{CH}_2\text{OO}$  which then reacts rapidly with  $\text{CH}_3\text{NH}_2$  to form Product1. Both reactions are in pseudo first order as excess concentrations of  $\text{O}_2$  and  $\text{CH}_3\text{NH}_2$  were used for these measurements. Thus, concentrations of  $\text{O}_2$  and  $\text{CH}_3\text{NH}_2$  do not need to be modelled explicitly here. Make sure there are not empty lines!

Reaction Species =  $\text{CH}_2\text{I}:\text{CH}_2\text{OO}:\text{Product1}$

Initial Concentrations,  $\text{C}_0 = 1:0:0$

Initial Rate Coefficients,  $k = 8000:500$

Time Start = 0

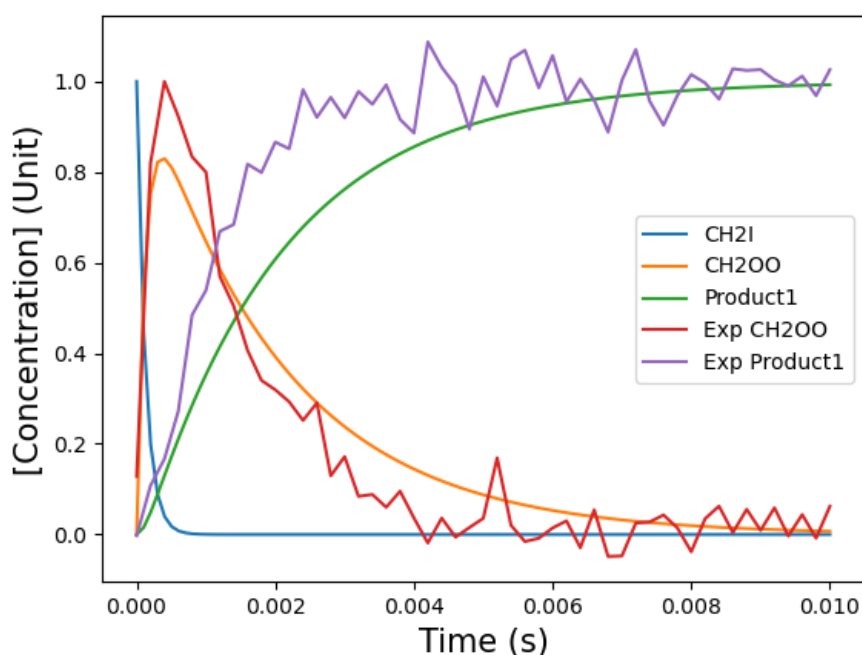
Time Stop = 0.01

Number of Steps = 100

Measured Species:  $\text{CH}_2\text{OO}:\text{Product1}$

Species to Plot =  $\text{CH}_2\text{I}:\text{CH}_2\text{OO}:\text{Product1}$

Press the Model button and open the example\_2I.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:

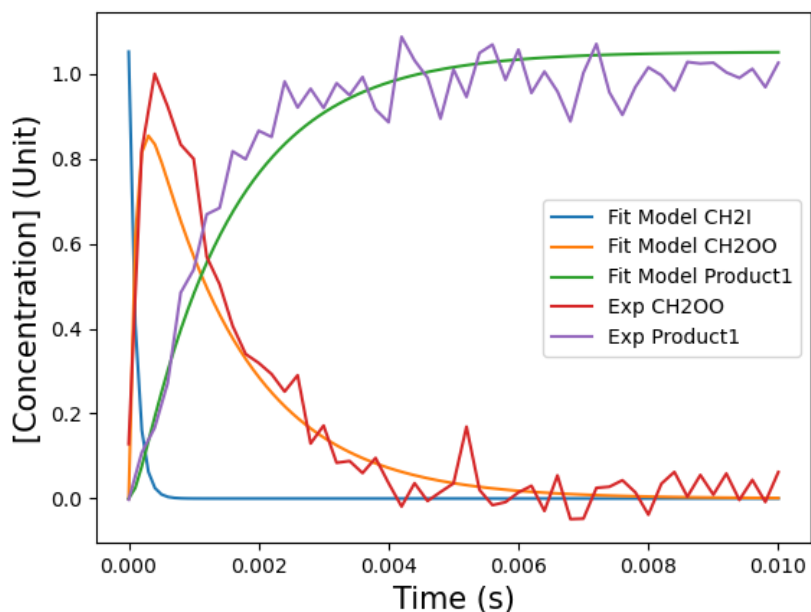
$\text{C}_0$  min -Fit =  $0.8:0:0$

$\text{C}_0$  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)

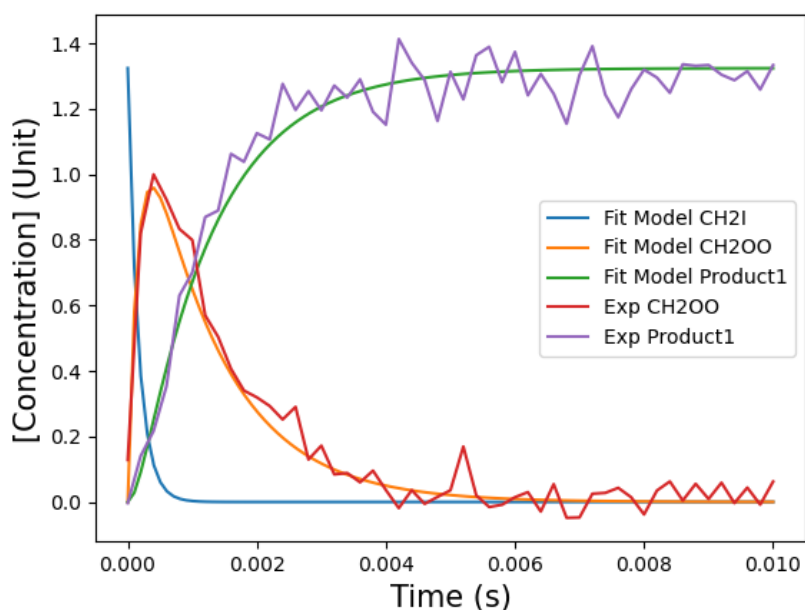
Initial\_Product1: 0 (fixed)

[[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 CH<sub>2</sub>OO signal is not reproduced well. This is likely because our model assumes that all the initial CH<sub>2</sub>I eventually produces Product 1 so the initial CH<sub>2</sub>I and final Product1 signal should be similar. However, we scaled CH<sub>2</sub>OO 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, Initial\_CH2I) = -0.188

C(k1, 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>

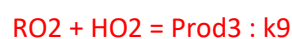
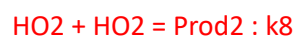
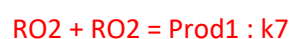
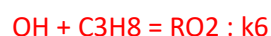
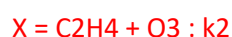
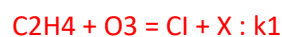
#### **Example 3: Box Modelling**

Under lot of realistic conditions, for e.g. in the atmosphere, the initial reagents are being constantly emitted. In such scenarios, the concentrations of reactants and the intermediate species created do not change rapidly, unlike pump probe experiments where a pulse of initial reagents are created by laser radiation which reacts away rapidly. A box modelling approach can be used to in which reactants



are being replenished. Shown here is a case in which an alkene and ozone react to form a Criegee intermediate, OH and HO<sub>2</sub> radicals in the troposphere. The main sink of the RO<sub>2</sub> and HO<sub>2</sub> radicals are self-reaction and reaction with each other in absence of NO. Similarly, main sink of the Criegee intermediate is self-reaction and unimolecular reaction in absence of H<sub>2</sub>O. Show below is the input file example\_3.txt which has the reaction input for the box model:

#### Reaction Model



#### Reactions Species

C<sub>2</sub>H<sub>4</sub>:O<sub>3</sub>:X:CI:CH<sub>2</sub>OO:

HO<sub>2</sub>:OH:C<sub>3</sub>H<sub>8</sub>:RO<sub>2</sub>:Prod1:

Prod2:Prod3:Prod4

#### Initial Concentrations

4.6E14:6E11:0:0:0:

0:0:8.6E16:0:0:

0:0:0

#### Rate Coefficient Values

1.6E-18:1E6:3.7E5:1.3E5:1.3E5:

7.6E-12:3E-13:2.5E-12:1E-11:7.3E-11:

1.1E-3

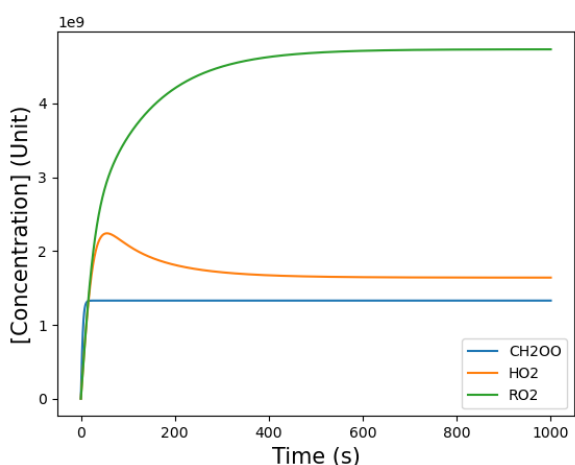
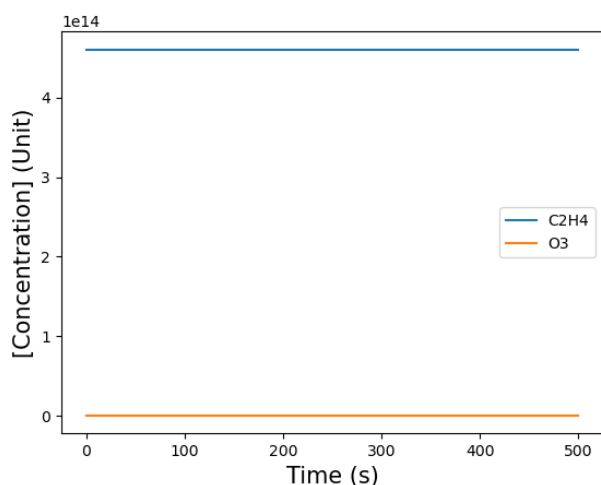
## Species to Plot

CH<sub>2</sub>OO:HO<sub>2</sub>:RO<sub>2</sub>

## Time

0:100:1000

There are six sections in the input file: Reaction Model, Reaction species, Initial Concentrations, Rate Coefficient values, Species to plot and Time. The formats for each entry are similar to the ones described in the previous examples. Each section should be separated by a blank line. The entries for reaction species, initial concentrations, rate coefficient values and Species to plot can be given in separate lines for better code readability. In the box model above concentrations of C<sub>2</sub>H<sub>4</sub> and O<sub>3</sub> are kept constant during the simulation to create a 'boxed condition'. Shown below are the results of the model simulation.



We can see that the concentration of C<sub>2</sub>H<sub>4</sub> and O<sub>3</sub> are constant during the simulation. The concentration of RO<sub>2</sub>, HO<sub>2</sub> and Criegee intermediate CH<sub>2</sub>OO increases and then reaches a steady state concentration within few minutes as the production and loss rates balance out.