CoDo - COmbined DOsimetry model User Guide

Daina Romeo

May 19, 2021
Version: 1.2

Documentation

CoDo - COmbined DOsimetry model User Guide

Daina Romeo

Reviewer Dr. Peter Wick
Particles-Biology Interaction Lab
Empa - Switzerland

May 19, 2021

Acknowledgement

I would like to thank Owen Price for his support in enlightening the details of the MPPD model, and Cordula Hirsch for her insights about the DG model and its results. My sincere thanks go to Daniele Lain for his fundamental contribution to the automatising of MPPD calculations.

This work has been developed under the European project Nanorigo. The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No. 814530. This document reflects only the author's view and the Commission is not responsible for any use that may be made of the information it contains.

Contents

1 About CoDo

1.1 CoDo in a nutshell

The COmbined DOsimetry model - CoDo - is a model for the estimation of the exposure levels corresponding to the concentrations of nano- or micro-particles used in *in vitro* submerged systems (Figure 1.1). It is based on the simulation of particles sedimentation and diffusion dynamics *in vitro*, and the simulation of diffusion, sedimentation, impaction, and clearance processes in the respiratory system of humans.

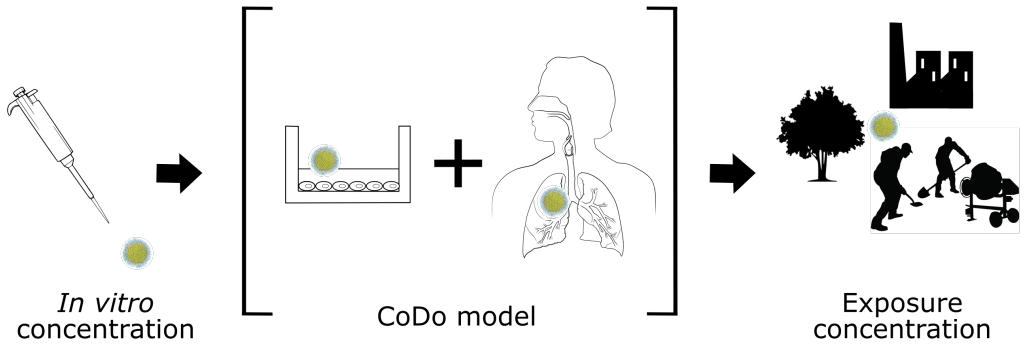

Fig. 1.1: CoDO is able to link *in vitro* concentrations to exposure levels thanks to the combined application of *in vitro* and *in vivo* simulation of particles dynamics.

CoDo is developed using Python programming language, and is supported by three additional files (more details in section 2.2): a template to provide input data, as xlsx file; a database of lung deposition fractions obtained using the MPPD software, as csv file; and a picture of the MPPD software icon, as png file. The results are saved as csv files, while optional plots can be saved as pdf files.

In a nutshell, this is how CoDo works (Figure 1.2):

1. CoDo loads the input template file;

2. CoDo checks if all the entries that the user wants to be simulated have a corresponding entry in the MPPD database;

3. If yes, jump to point 5;

4. If not, the user will be asked to open the MPPD program, with which CoDo will interact automatically. The results of the simulation will be saved in the MPPD database;

5. CoDo will simulate the deposition of particles *in vitro*;

6. CoDo will combine the results of the *in vitro* simulation with the lung deposition fractions from the MPPD database to obtain the exposure levels corresponding to the *in vitro* doses;

7. The results are saved in a csv file, including the deposition of particles *in vitro* and the corresponding exposure levels. Additional plots and detailed files of the *in vitro* simulation are saved if specified in the input template file.

A more detailed description of the step-by-step procedure is available in chapter 3.

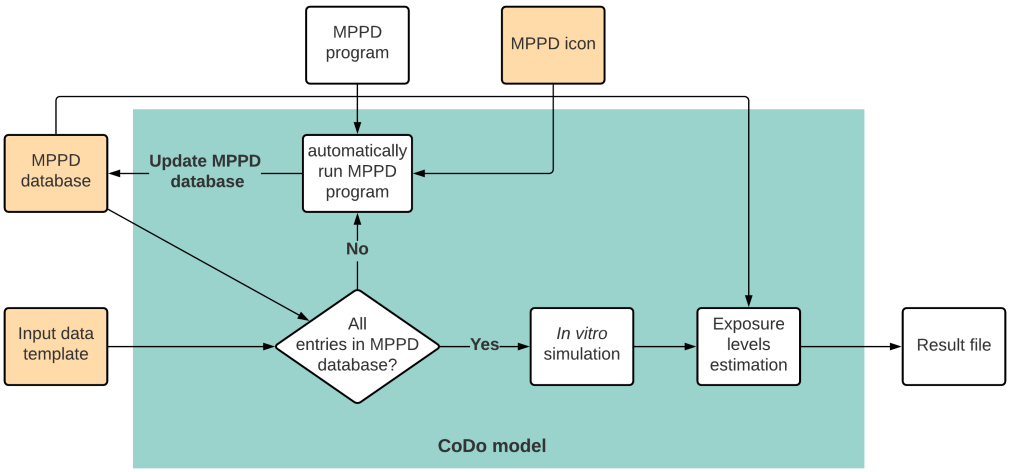

Fig. 1.2: A simplified scheme of CoDo operation and it's dependency on the supporting files (in yellow) and external software.

1.2 Why use CoDo

CoDo is an open source model that aims at facilitating linking *in vitro* studies to environmental exposure levels. It can be used, for example, to evaluate existing *in vitro* data to verify if the doses used represent realistic conditions, or to choose the doses to use in *in vitro* studies so that they represent specific exposure conditions. Moreover, by integrating *in vitro* data with dosimetry simulations, it can support the estimation of toxicological dose descriptors.

The easy-to-fill template input file allows to include multiple entries, and the user can specify a range of entries to evaluate in a single run. This is particularly efficient in case of big data sets or to test multiple options (e.g. consider both sexes for lung dosimetry).

1.3 Availability and reuse

The model, including the supporting files, will be downloadable from GitHub repository at `https://github.com/dainaromeo/CoDo`, but will also be linked in the Nanorigo Risk Governance Framework website (which will provide the option of generating input files). We release the model as Open Source, but we require the user to cite the following works:

- Daina Romeo et al. „Combined in vitro-in vivo dosimetry enables the comparison of in vitro doses and exposure levels: a proof of concept using titanium dioxide toxicity data.“ In: *Manuscript in preparation* (2020)

- Glen M. DeLoid et al. „Advanced computational modeling for in vitro nanomaterial dosimetry“. In: *Particle and Fibre Toxicology* 12.1 (Dec. 2015), p. 32

- Frederick J. Miller et al. „Improvements and additions to the Multiple Path Particle Dosimetry model“. In: *Journal of Aerosol Science* 99 (Sept. 2016), pp. 14–26

1.4 Reference dosimetry calculations

CoDo relies on two existing dosimetry models for the simulation of the particles dynamics in media and in air.

In the case of *in vitro* dosimetry, which is directly implemented into CoDo, we refer to the work by DeLoid et al. [2], which explains how to model the behavior of particles by subsequent time-discreet simulations of diffusion and sedimentation processes. The protocol published by the authors [4] can aid the user in the measurement of the parameters needed for the model.

For lung dosimetry, CoDo do not directly simulate the particles behavior, but is able to interact with the existing Multiple-Path Particle Dosimetry model (MPPD) to automatically use such software and save the results in a dedicated database (MPPD database). The MPPD model is a well-known and commonly-used lung dosimetry

model [5], available for free at `https://www.ara.com/mppd/`. A summary of the model and its uses can be found in Romeo et al. [6], while detailed information is available in Miller et al. [3], Asgharian et al. [7], and Asgharian et al. [8, 9]. The MPPD database provided as supporting file already includes more than 500 entries, to accelerate the simulation in case of entries that have already been evaluated.

2 Getting started

2.1 Prerequisites

Requirements:

- Python 3.7+
- MPPD model (available for free at `https://www.ara.com/mppd/`)
- Any software able to manage xlsx and csv files (e.g. Microsoft Office, Libre Office, Google Docs)
- The supporting files

The code has been tested on Windows 10.

2.2 Supporting files

In addition to the model itself (CoDo.py), we provide three supporting files that are necessary for the model.

2.2.1 Template input data - CoDo_template.xlsx

CoDo_template.xlsx is the template used to provide the input data to the model (Figure 2.1). It comprises a "Data" sheet where the user can input their data, and a "Legend" sheet that includes an explanation of each parameter, the available default values, and additional values. Alternatively to the template file, the user can input the data through any csv file in which the column names correspond to the parameters codes (as in row 3 of the template file).

A color code guides the user in the compilation of the template file. The parameters for which default values are available are indicated in green: in these cases, the user can leave the cell empty, and the default value will be used. The parameters

in pink/orange are parameters required for the simulation, but that, if missing, will be calculated by the model. In the case of the "Agglomerate effective density", the "fractal dimension", in light pink and green, is required, and a default value is available.

Two entries are included as examples; the data have been extracted from [10] and [11], and have been adapted to better illustrate the different ways in which the file can be filled, for example some parameters have been omitted to show that the default value will be used.

Hereafter the description of each parameter:

- **General info**
  - ID: a unique identifier for each entry.
  - Substance: The name of the particle, written in its chemical formula, e.g. TiO_2, Fe_2O_3. It is used as identifier to match the entries in the MPPD database file.
- **MPPD parameters (lung dosimetry)**
  - Sex of person: either "male" or "female" (not case-sensitive). Affects the dosimetry in the lung as the lung and breathing parameters are different for the average male and average female. It is used as identifier to match the entries in the MPPD database file.
  - Type of particle in air: either "pp" or "agg", respectively standing for "primary particle" and "agglomerate" (not case-sensitive). Defines the type of particle in air: either the model considers the primary particle, or an agglomerate corresponding to the agglomerate *in vitro*.
- **Media parameters (*in vitro* dosimetry)**
  - Media viscosity: the viscosity of the media in Pa·s.
  - Media density: the density of the media in g/cm^3.
  - Media temperature: the temperature of the media, in degrees Centigrade. Default value = 37°C.
- **Substance parameters**

- Primary particle diameter: in nm.

- Primary particle density: in g/cm^3.

- Primary particle specific surface area: in m^2/g, if not provided it is calculated from the diameter and the density of the primary particle according to the formula: $SSA = 0.0001 \cdot \frac{6}{\text{primary particle diameter} \cdot \text{primary particle density}}$

- Agglomerate diameter: the diameter of the agglomerate in the media in nm, e.g. the average hydrodynamic diameter. The model is not able to simulate agglomerates with diameters smaller than 26 nm.

- Agglomerate effective density: the effective density of the agglomerates in the media, in g/cm^3. If not provided by the user, it is calculated according to the Sterling equation:

$$\text{Agglomerate porosity} = 1 - \left(\frac{\text{agglomerate diameter}}{\text{primary particle diameter}}\right)^{\text{Fractal dimension}-3}$$

$$\begin{aligned}\text{Agglomerate effective density} &= \\ &= (1 - \text{Agglomerate porosity}) \cdot \text{primary particle density} + \\ &\quad + \text{Agglomerate porosity} \cdot \text{media density}\end{aligned}$$

- Fractal dimension: represents the fractal structure of the agglomerate, and varies between 1 and 3, with 3 being a perfect sphere with no fractal structure and no porosity. Default value = 2.1.

- **Experimental parameters (*in vitro* dosimetry)**

  - Column height: the height of the media in mm.

  - Initial particle concentration: concentration of the nanomaterial in the media, in mg/cm^3.

  - Simulation time: length of the exposure, in hours.

- **Threshold for deposition rate (*in vitro* dosimetry)**

  - Extend the simulation time until the defined deposition rate is reached: TRUE or FALSE, default is FALSE. Indicates whether to extend the simula-

tion until the increase in deposition per hour is equal or less than a set threshold (e.g. the change in deposition is very small over time, and we can consider to have reached a quite "stable" condition).

  - Threshold deposition rate as percentage per hour: arbitrary default value = 0.002, i.e. the simulation will be extended until the change in the deposited fraction per hour is equal or less to 0.002. It is calculated as:

$$\frac{\text{Deposited fraction}_{t1} - \text{Deposited fraction}_{t0}}{t_1 - t_0} < 0.002$$

    with t1 and t0 two consecutive time points.

- ***In vitro* simulation parameters**

  - Compartment height: the height of the theoretical subcompartments used for the simulation, default = 0.005 mm.

  - Simulation time interval: the interval time of the simulation, default = 0.5 s.

  - Subcompartment height: the height of the bottom compartment in which the particles are considered as deposited. It must be evenly divisible by the compartment height. Default = 0.01 mm.

- **Output parameters (*in vitro* dosimetry)**

  - Plot graphs: TRUE or FALSE, default = FALSE. If set to True, the deposited fraction over time, the deposited mass over time, and the deposited surface area over time will be plotted and saved together as a pdf file. The file name will be "Deposited_*Output file name*.pdf.

  - Save extensive results: TRUE or FALSE, default FALSE. If set to True, a csv file will be saved for each entry that has been simulated, reporting the results of the *in vitro* dosimetry simulation at each time point as defined by the output time interval. The file will be named *Output file name*.csv.

  - Output file name: the name to use to save the plots and the extensive results. The default name is *Substance_ID_date of simulation*.

  - Output time interval: the time interval in minutes at which the results are saved and reported in the extensive result file. Default = 30 min.

- ***In vitro* process parameters**

  - Sedimentation concentration dependence: indicates the dependence of the sedimentation coefficient on the concentration of the particles, and usually ranges between 0 and 0.1. Default = 0.

  - Diffusion concentration dependence: indicates the dependence of the diffusion coefficient on the concentration of the particles, and usually ranges between 0 and 0.1. Default = 0.

  - Initial dissolution fraction: fraction of the particles that is dissolved at time 0. Default = 0.

  - Dissolution rate type: behavior of dissolution after time 0. If no dissolution occurs, type = 0; if the particle dissolves at a constant rate, type = 1, if the dissolution rate changes over time, type = 2. Default = 0.

  - Dissolution rate if rate type = 1: fraction of particles dissolved per hour. Note that if dissolution rate $\cdot$ simulation time ≥ 1 all particles are dissolved at the end of the exposure time. Default = 0.

  - Dissolution time h if rate type = 2: time points in hours at which dissolution has been measured. Indicate the values separated by a comma (e.g. 0.5, 0.8, 2, 4). Default = 0.

  - Dissolution fraction if rate type = 2: dissolved fractions at the times reported in "Dissolution time". Indicate the values separated by a comma (e.g. 0.02, 0.06, 0.1, 0.18). Default = 0.

  - Sticky bottom: TRUE or FALSE, default = FALSE. Indicates whether the particles are adsorbed irreversibly on the bottom (if True) or if they can be resuspended (if False). In the first case, the particles will deposit completely at a certain point, while in the second an equilibrium different from 100% deposition might be reached. Generally, sticky conditions are appropriate when the particles are so big and dense that diffusion is negligible, or when there is a strong adsorption or uptake of the particles by the cells (which removes the particles from the system). More common condition instead is that the interactions between particles and cells are weak, which is mimicked by non-sticky conditions [2].

  - Adsorption dissociation constant: used to model the binding of agglomerates as a Langmuir isotherm adsorption process. Default = 10^{-9}.

1	General info		MPPD parameters		Media data			Substance parameters					
2	ID	Substance	Sex of person	type of particle in air	Media viscosity Pa s	Media density g/cm3	Media temperature degree C	primary particle diameter nm	Primary particle density g/cm2	Primary particle specific surface area	Agglomerate diameter nm	Agglomerate effective density g/cm3	Fractal dimension
3	ID	Substance	sex	air_type	media_vis	media_de	media_te	pp_diame	pp_density	ssa	agg_diameter	agg_effectiv	DF
4	1	TiO2	female	agg	0.00093	1.0211	37	17	4.2		900	1.4939	
5	2	Amorphous S	male	pp	0.000726	0.999		12	2.2	200	220		

Fig. 2.1: An example of the Data sheet in the CoDo_template.xlsx file. An example of the color code can be observed, with examples of parameters where default values are available (green), parameters that can be calculated by the model if missing (pink/orange), and parameters that are conditionally needed (light pink).

2.2.2 MPPD database of lung deposition fractions - MPPD_df_auto.csv

The MPPD_df_auto.csv file is a data set containing information about the total alveolar retention and the alveolar retention per cm^2 occurring after the exposure to a concentration of particles of 1 mg/m^3, at different exposure conditions, for particles of different sizes (Figure 2.2). It includes 532 entries including different particles and different exposure scenarios: short-term exposure equal to *in vitro* exposures, and one and 45 year(s) of exposure on the workplace (8h/day, 5 days/week). The parameters reported are general information (G), input data for the MPPD model (I), or results extracted after the analysis (R):

- **Substance** (G): the type of particle, e.g. TiO_2. It is used as identifier to match the entries in the template input file.

- **agg_diameter_nm** (I): The diameter of the particle in air, either as agglomerate or primary particle. CoDo gives the user the option to choose, for the particle diameter in air, either the primary particle diameter or the diameter of the agglomerate measured *in vitro* in media. It is used as identifier to match the entries in the template input file.

- **density_g/cm^2** (I): The density of the agglomerate or of the primary particle in air. It is used as identifier to match the entries in the template input file.

- **pp_diameter_nm** (G): The diameter of the primary particle. It is used as identifier to match the entries in the template input file.

- **sex** (G): Either male or female, the sex of the default person considered for the simulation. It influences the lung and breathing parameters: if male, the default values are kept, i.e. Functional Residual Capacity FRC = 3300 ml,

Upper Respiratory Tract Volume URT = 50 ml, Breathing Frequency BR = 12/min, Tidal Volume TV = 625 ml. For women, FRC = 2680 ml, URT = 40 ml, BF = 14/min, TV = 464 ml [12]. It is used as identifier to match the entries in the template input file.

- **Exposure** (I): Indicates if the exposure levels change over time. Set to "constant". Note: it does not indicate the duration of the exposure but only the change in exposure levels over time.

- **Exposure_time_h** (I): The length of the exposure in hours. It is used as identifier to match the entries in the template input file, based on the length of the exposure *in vitro*. To calculate the alveolar retention after one or 45 (work life) years of exposure on the workplace, it is set to 8 hours per day.

- **Exposure_time_d** (I): The number of days per week in which the exposure takes place. It is set to 1 when considering the exposure time as *in vitro*, e.g. Exposure_time_h = 48, Exposure_time_d = 1 instead of Exposure_time_h = 24, Exposure_time_d = 2 (the results do not change).When considering long-term exposure on the workplace, it is set to 5 days.

- **Exposure_time_w** (I): The number of weeks of exposure. It is set to 1 when considering the exposure time as *in vitro*. When considering long-term exposure on the workplace, it is set to 1820 days (45 years), which provides both the data for 5 days, 1 year (52 weeks) and 45 years of exposure.

- **Lung_Model** (I): The lung morphology model selected for the lung dosimetry simulation. In the provided data the Yeh/Schum Symmetric model was chosen, as it provides accurate results for regional deposition within a very short computation time [7].

- **Clearance** (G): Indicates that clearance processes have been accounted for in the simulation. Clearance exclusion is not supported.

- **Alveolar_retention_h_mg** (R): Result from the simulation, indicates the mass of the particles that are deposited in the pulmonary region after the exposure time "Exposure_time_h", accounting for clearance processes.

- **Alveolar_retention_five_mg** (R): Result from the simulation, indicates the mass of the particles that are deposited in the pulmonary region after five days of workplace exposure, accounting for clearance processes.

- **Alveolar_retention_y_mg** (R): Result from the simulation, indicates the mass of the particles that are deposited in the pulmonary region after one year of workplace exposure ("Exposure_time_y"), accounting for clearance processes.

- **Alveolar_retention_w_mg** (R): Result from the simulation, indicates the mass of the particles that are deposited in the pulmonary region after 35 years of workplace exposure ("Exposure_time_w"), accounting for clearance processes.

- **Alveolar_surface_cm^2** (G): The surface of the alveoli in cm^2. It's set to 792000 cm^2 for men, and 559000 cm^2 for women [13].

- **Deposited_per_area_h_mg/cm^2** (R): The amount of deposited particles per alveoli area after the exposure time "Exposure_time_h". It is calculated as the ratio between the alveolar retention and the surface area of the alveoli.

- **Deposited_per_area_five_mg/cm^2** (R): The amount of deposited particles per alveoli area after five days of exposure on the workplace. It is calculated as the ratio between the alveolar retention and the surface area of the alveoli.

- **Deposited_per_area_w_mg/cm^2** (R): The amount of deposited particles per alveoli area after one year of workplace exposure ("Exposure_time_y"). It is calculated as the ratio between the alveolar retention and the surface area of the alveoli.

- **Deposited_per_area_y_mg/cm^2** (R): The amount of deposited particles per alveoli area after 35 years of workplace exposure ("Exposure_time_w"). It is calculated as the ratio between the alveolar retention and the surface area of the alveoli.

Alveolar_retention_five_mg	Alveolar_retention_h_mg	Alveolar_retention_w_mg	Alveolar_retention_y_mg	Alveolar_surface_cm2	Clearance	Deposited_per_area_five_mg/cm2	Deposited_per_area_h_mg/cm2	Deposited_per_area_w_mg/cm2	Deposited_per_area_y_mg/cm2	Exposure	Exposure_time_d	Exposure_time_h	Exposure_time_w	Lung_Model	Substance	agg_diameter_nm	density_g/cm2	pp_diameter_nm	sex
0.088				792000	Yes		1.11E-07			constant	1	4	1	Yeh/Schun	MgO	852.2	0.168973	46.66	man
	0.8686	159.7	29.5	792000	Yes	1.10E-06		0.000202	3.72E-05	constant	5	8	1820	Yeh/Schun	MgO	852.2	0.168973	46.66	man
0.4952				792000	Yes		6.25E-07			constant	1	4	1	Yeh/Schun	MgO	46.66	3.14	46.66	man
	4.886	898.2	165.9	792000	Yes	6.17E-06		0.001134	0.000209	constant	5	8	1820	Yeh/Schun	MgO	46.66	3.14	46.66	man
0.5268				792000	Yes		6.65E-07			constant	1	24	1	Yeh/Schun	MgO	852.2	0.168973	46.66	man
2.963				792000	Yes		3.74E-06			constant	1	24	1	Yeh/Schun	MgO	46.66	3.14	46.66	man

Fig. 2.2: Screenshot of the MPPD_df_auto.csv file.

2.2.3 MPPD_icon.png

MPPD_icon.png is the image of the MPPD software logo at 2560 x 1440 resolution, 100% scaling (Figure 2.3). It is used if the user's entry is not already in the MPPD_df_auto database. In that case, the user will be asked to open the MPPD program, so that CoDo can automatically interact with it to calculate the alveolar

retention for the entry, and update the MPPD_df_auto database. The icon is needed to identify the position of the MPPD program on screen.

Fig. 2.3: The MPPD_icon.png file is used when the user's entry is not already included in the MPPD_df_auto database.

3 Running the model

After downloading CoDo and its supporting files, and filling the template input data with any number of entries to simulate, the user can open the CoDo.py file (i.e. CoDo model) to run the model. Figure 3.1 depicts the model steps and their connection to the supporting files (in light blue) and external software.

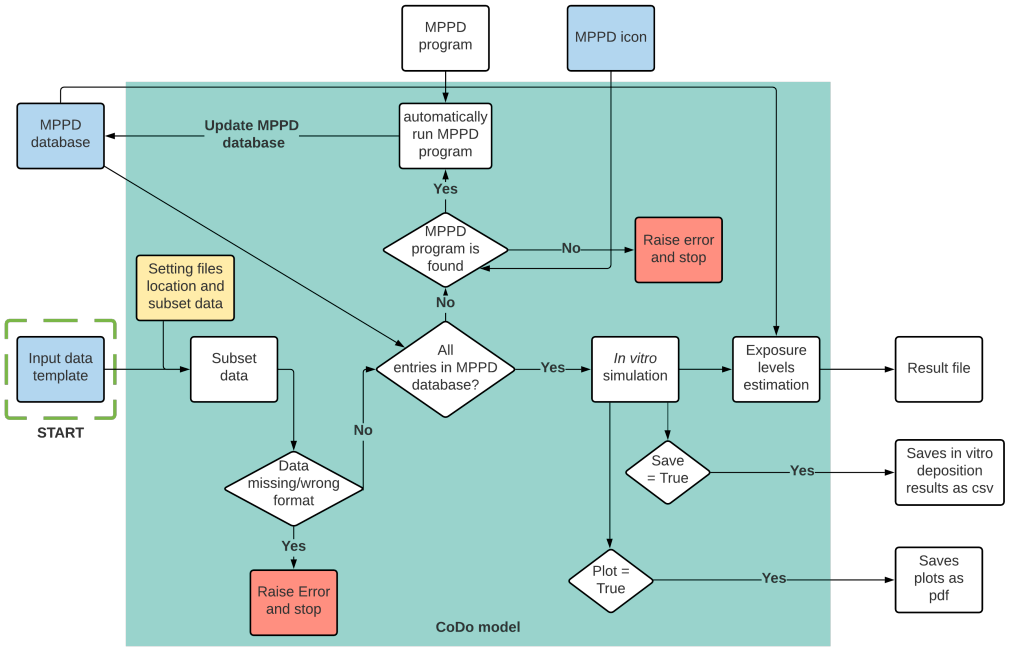

Fig. 3.1: The steps of CoDo's operation and their connection to the supporting files and external software. The input required to the user inside the CoDo.py file is highlighted in yellow, in light blue the supporting files, and in red the points in the simulations where an error could be raised and the simulation stopped if fundamental constraints are not satisfied (e.g. necessary input parameters are missing).

3.1 Setting files paths and row numbers in the code

To run CoDo, the user is required to provide a few inputs in the following section of the model code itself (CoDo.py), which is found in the beginning of the code:

```
1
2 # PLEASE NOT THAT THE FILEPATH USES "/" AND NOT "\"
```

```
# -------- INDICATE THE FULL PATH OF THE MPPD DATASET CSV FILE----

MPPD_dataset = 'C:/Users/roda/Documents/Python Scripts/MPPD_df_auto.csv'

# ------ INDICATE THE LOCATION OF THE MPPD SOFTWARE FOLDER - USUALLY FOUND IN DOCUMENTS

MPPD_results = 'C:/Users/roda/Documents/MPPD'

# ------- INDICATE THE FULL PATH OF THE TEMPLATE FILE---------
template_file = 'C:/Users/roda/Documents/Python Scripts/CoDo template.xlsx'

# ------- INDICATE THE FULL PATH OF THE MPPD ICON IMAGE------------

MPPD_icon = 'C:/Users/roda/Documents/Python Scripts/MPPD_icon.png'

# - INDICATE THE FIRST AND LAST LINE TO COMPUTE (IF ONLY ONE LINE, START_VALUE=END_VALUE)
start_value = 4
end_value = 4

# -----------------------------------------------------
#            END OF USER'S INPUTS - DO NOT MODIFY AFTER THIS LINE
# -----------------------------------------------------
```

The user should indicate the full path of the three supporting files, surrounded by either single quotation marks, or double quotation marks. Note that the notation should use slash "/" instead of backslash "\" (as used instead in Windows file paths), e.g.:

MPPD_dataset = 'C:/Users/username/Documents/CoDo/MPPD_df_auto.csv'

and not:

MPPD_dataset = 'C:\Users\username\Documents\CoDo\MPPD_df_auto.csv'

Moreover, the user is required to indicate the location of the folder where the results from the MPPD software are saved, to retrieve them during the simulation. This folder is automatically created when installing the MPPD software, and is usually located in the Documents folder. E.g.:

MPPD_results = 'C:/Users/username/Documents/MPPD'

Last, the user should indicate which entries from the input template file they want to run the simulation on. The user is required to indicate the first and last ROWS (not ID) of the template file corresponding to the desired entry/range of entries (Figure 3.2).

If start_value = 4 and end_value = 4, the entry in row 4 will be simulated (in our case, this corresponds to ID 1).

If start_value = 4 and end_value = 5, the entries in rows 4 and 5 will be simulated (in our case, this corresponds to ID 1 and 2).

If start_value = 4 and end_value = 6, the entries in rows 4, 5, and 6 will be simulated.

Note: the first entry in the template file is in ROW 4, as the first three are used for the parameters names. If using a csv input file, the first entry will be in ROW 2.

3.2 Run the model

The computation starts by running the model via the "Run" command in Python. The subsequent steps are all automatic, except if the entries are not available in the MPPD database, in which case the user's input will be required.

3.3 Loading the input template file and checking the input data

As a first step (Figure 3.3), CoDo will load the input template file, and keep the subset of entries to run the simulation on, as indicated at the beginning of the CoDo file (the subset identified by the start_value and end_value, as in section 3.1). Then, the model will check that the data provided is correct and no data is missing, and if not, it will raise one of the following errors and stop running:

```
'One or more required parameters are missing. Please check the
    template input file.'
'At least one agglomerate diameter is < 26nm. Please remove/modify
    entry.'
'Air type parameter is different from pp or agg. Please check input
    file.'
'Sex parameter is different from male or female. Please check input
    file.'
```

1	General info		MPPD parameters	
	ID	Substance	Sex of person	type of particle in air
2				
3	ID	Substance	sex	air_type
4	1	TiO2	female	agg
5	2	Amorphous S	male	pp

Fig. 3.2: The start_value and end_value refer to the row number of the entry.

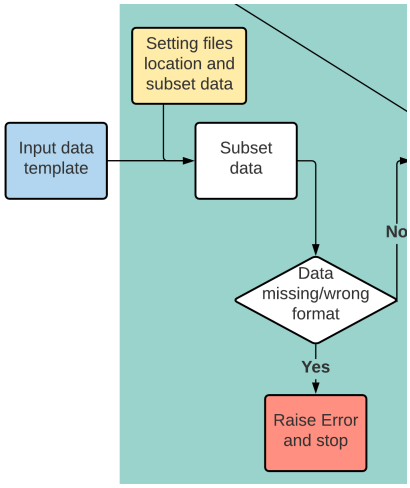

Fig. 3.3: As a first step, the template input file is loaded and the entries are checked for any missing data or wrong data format.

Each error refers to one of the following cases:

- An indispensable parameter is missing. The indispensable parameters are: 'ID', 'Substance', 'sex', 'air_type', 'media_viscosity', 'media_density', 'pp_density', 'agg_diameter', 'column_height', 'initial_concentration', 'simulation_time';
- At least one entry has an agglomerate diameter smaller than 26 nm, which cannot be simulated by the model;
- The "air_type" parameter of at least one entry is different from the accepted options, either "pp" or "agg", not case-sensitive ("aGg", "PP" are all valid options);
- The "sex" parameter of at least one entry is different from the accepted options, either "female" or "male", not case-sensitive ("MALE", "Female" are all valid options).

3.4 Availability of entries in MPPD database

In the next step, CoDo will check if corresponding entries are already available in the MPPD database, or if it necessary to simulate them via the MPPD model (Figure 3.4).

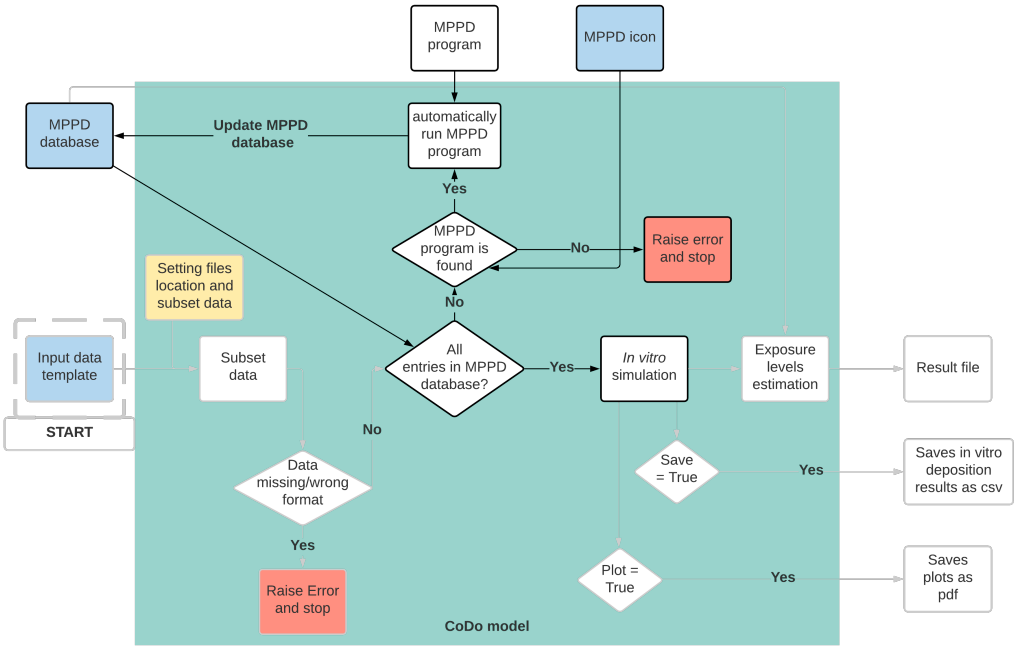

Fig. 3.4: CoDo checks if every entry is already covered in the MPPD database, otherwise it will automatically interact with the MPPD software.

3.4.1 At least one entry not in MPPD database

If at least one entry is not covered in the MPPD database, CoDo will have to interact with the MPPD software to calculate the missing data. To do so, the code will pause and indicate the user what are the next steps to be taken:

```
'Not all particles are available in the MPPD database. They will be
    calculated now. Please open the MPPD program, and make sure the
    resolution of your main screen is 2560 x 1440 (100% scaling). Keep
     the MPPD program on your main screen. To run the analysis using
    the MPPD software, you will need to have the MPPD software window
    on screen (not covered by any other program/window). You should
    not use or move the mouse, as the MPPD will be run by
    automatically moving the mouse to run the program multiple times.
    When you are ready, press ENTER; you will have 2 seconds to have
    the MPPD software window open and visible on screen (if not
    already), afterwards please do not move the mouse anymore.
Please, press ENTER to start running the calculation of MPPD results'
```

Respectively, the user is required to:

1. Verify that their main screen resolution is 2560 x 1440, and the scaling is 100%. Such settings are required to proceed;

2. Open the MPPD model on their main screen;

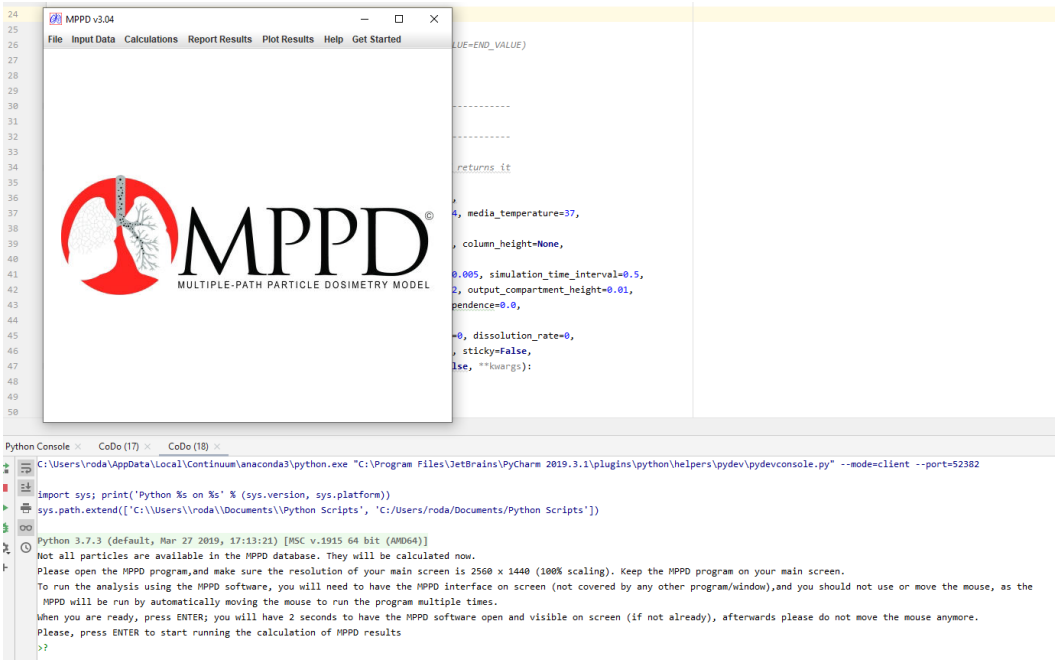

Fig. 3.5: To automatically interact with the MPPD interface, the MPPD software window must be visible on the main screen of the user.

3. Press "Enter" in the Python console;

4. After pressing "Enter", the user has two seconds to have the MPPD software window on their screen, as in Figure 3.5;

5. After that, the mouse should not be used until the interaction with the MPPD software is completed.

To automatically interact with the MPPD software, the model will first load the MPPD icon, and use it to identify the position on screen of the MPPD software window. If the position is not found, a warning will appear:

```
'MPPD window not located on screen, please make sure that the
    resolution is 2560 x 1440 (100% scaling), and the MPPD program is
    open and visible on your main screen.'
'If all points mentioned before have been fixed, press ENTER to retry
    . You will have 2 seconds to have the MPPD software window on
    screen and to stop moving the mouse'
```

After the user presses "Enter", if the MPPD window is again not found, CoDo will stop:

```
Impossible to find the MPPD window. The model will stop running.'
```

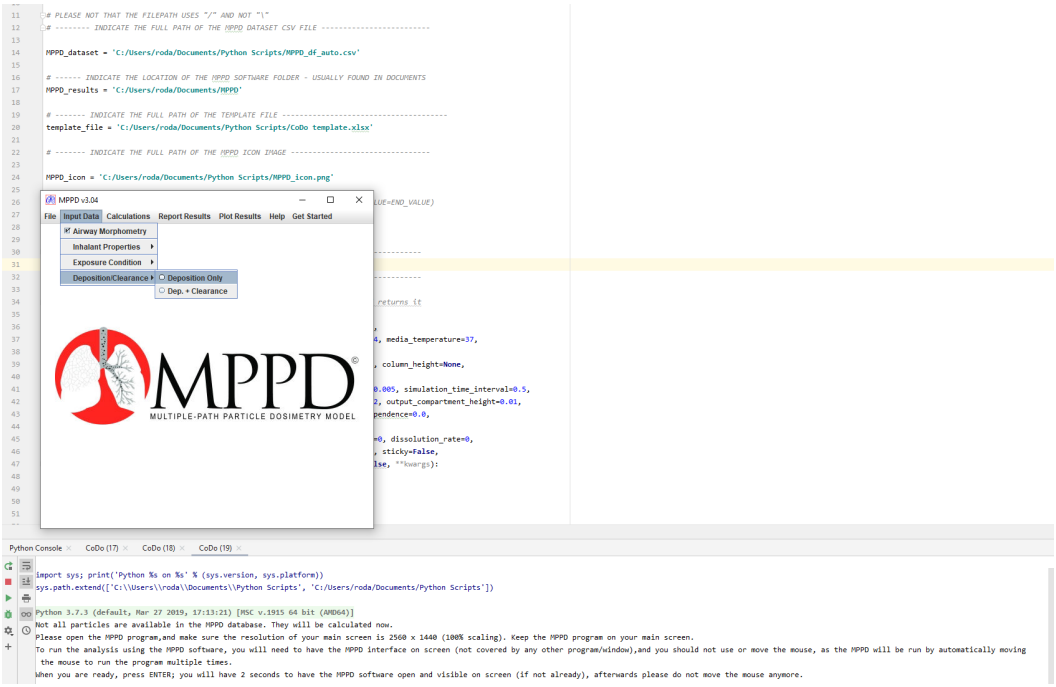

Fig. 3.6: A screenshot of CoDo automatically interacting with the MPPD interface.

If no problems occur, CoDo will automatically move the mouse to click on the MPPD interface and fill in all the input data (Figure 3.6). After the interaction with the MPPD model is concluded, this message will be displayed:

```
'MPPD results obtained: you can now close the MPPD program and use
    the mouse again.'
```

The MPPD database will be updated to include the new entries, and the model will continue running.

3.4.2 All entries in MPPD database

If/when all the entries are included in the MPPD database, the model will start the simulation of *in vitro* deposition.

3.5 *In vitro* dosimetry simulation

The simulation of *in vitro* deposition is performed through a loop of time-discreet rounds of sedimentation, diffusion, and, if the case, dissolution processes. A description of the formulae used can be found in the work by DeLoid et al. [2]. During the simulation, the Python console will display the ID number of the entry that is being simulated, the loop number, and the point that is being calculated. By

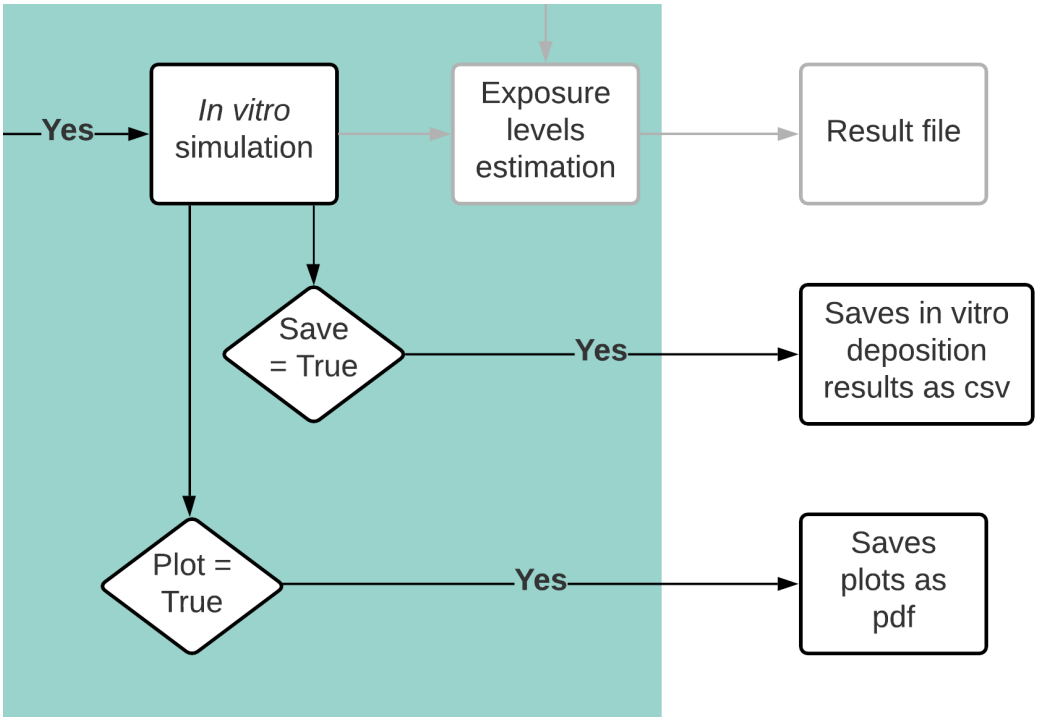

Fig. 3.7: During the *in vitro* simulation, depending on some input parameters, additional files can be saved.

default, a simulation round is performed at intervals of 0.5 seconds (total number of rounds/loops = exposure time in seconds/0.5), while a point is calculated at intervals of 30 minutes (total number of points = exposure time in hours/0.5). The end of the simulation for each entry is reported as well:

```
'ID 2, end loop 21600.0,calculating point 13'
'ID 2, end loop 21600.5,calculating point 13'
'ID 2: In vitro simulation complete'
```

Some parameters in the template input file refer to optional actions related to the *in vitro* simulation (Figure 3.7). These actions are specific for each entry in which the corresponding parameters are set to the activating option:

- If "Extend the simulation time until the defined deposition rate is reached" is True, the simulation will stop only when the change in deposition rate will be equal or smaller than the "Threshold deposition rate as percentage per hour". If this does not occur before the end of the exposure time ("Simulation time"), the simulation will be extended until the threshold rate is reached. This option can be used to identify at which point the deposition rate becomes so small that the deposited fraction can be considered quite stable over time (varying very slowly). The time at which the threshold is reached, and the deposited fraction, mass, surface area, and particle number are saved and included as well in the calculations of the exposure levels.

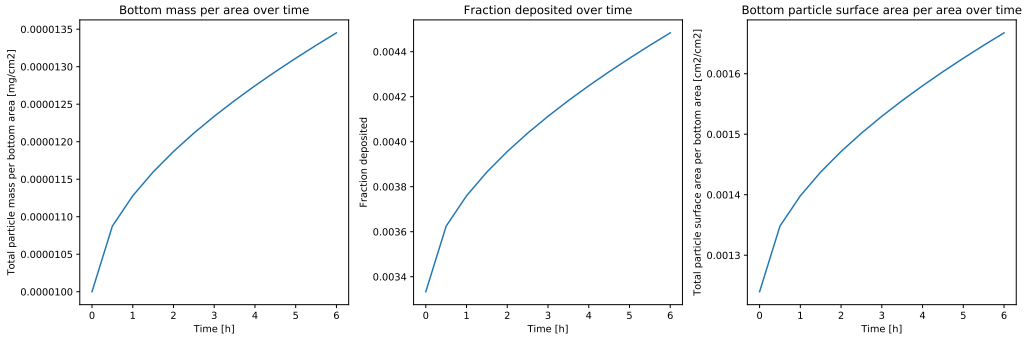

Fig. 3.8: The plots saved if the "Plot graphs True/False" parameter is set to True. Corresponds to the entry with ID 2 available as example in the template input file.

		Time: 0.0 hours	Time: 0.5 hours	Time: 1.0 hours
2	Time at which the marginal increase in deposited fraction per h is < 0.002 (h)	not defined	0	0
3	Deposited fraction (of total mass)	0.003333333	0.003624886	0.003759398
4	Mass deposited per area (mg cm^-2)	1.00E-05	1.09E-05	1.13E-05
5	Mass deposited and dissolved fraction per area (mg cm^-2)	1.00E-05	1.09E-05	1.13E-05
6	Number of particles per area (cm^-2)	11174496.03	12151881.29	12602813.28
7	Particles surface area per area (cm^2 cm^-2)	0.001239669	0.001348098	0.001398123
8	Concentration of dissolved fraction (mg cm^-3)	0	0	0

Fig. 3.9: A screenshot of the file saved if the "Save extensive results" is set to True. Corresponds to the entry with ID 2 available as example in the template input file.

- If "Plot graphs True/False" is set to True, the deposited mass per area over time, the deposited fraction over time, and the deposited surface area per area over time will be plotted and saved as a pdf file (Figure 3.8). The name of the file will be "Deposited_Output file name".

- If "Save extensive results" is set to True, the results of the *in vitro* simulation of the corresponding entry will be saved in an ad hoc csv file, named "Output file name.csv". The results will report the deposited amounts in different units at each "Output time interval minutes", e.g., at time 0, 30 min, 1 h, 1.5 h and so on (Figure 3.9).

3.6 Estimation of exposure levels

After the end of the *in vitro* simulation, the results will be combined with the lung deposition fractions available in the MPPD database (e.g. "Deposited_per_area_h_mg/cm2") to calculate the corresponding exposure levels (Figure 3.10). The exposure level is obtained according to formula:

$$\text{Exposure level} = \frac{\text{In vitro deposited mass per area}}{\text{Deposited fraction per area lung}}$$

The exposure levels will be calculated considering an *in vivo* exposure lasting:

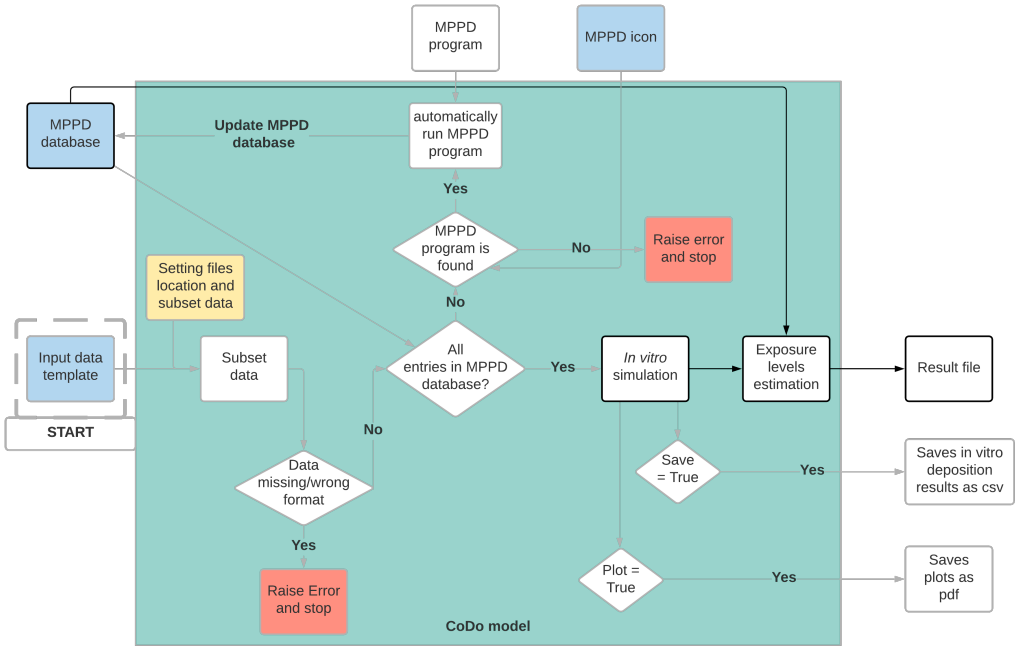

Fig. 3.10: The estimation of the exposure levels is obtained by the combination of the *in vitro* dosimetry simulation results and the lung deposition fractions saved in the MPPD database.

1. The same time as the exposure time *in vitro*;

2. One year of exposure on the workplace, i.e. 8 hours a day, 5 days a week;

3. Five days of exposure on the workplace, i.e. 8 hours a day;

4. A work life of exposure on the workplace, i.e. 8 hours a day, 5 days a week, for 35 years;

5. The same time needed *in vitro* to reach a "stable" deposition rate, if the "Extend the simulation time until the defined deposition rate is reached" parameter is set to True.

For each of these time points, a specific lung deposition fraction is available in the MPPD database, except for 5. In this case, the lung deposition fraction is estimated assuming a linear behavior over time between the exposure time and the five days exposure, according to the formulae:

$$\text{Change of lung deposition fraction per hour} = \frac{\text{Lung deposition fraction five} - \text{Lung deposition fraction at exposure time}}{120 - \text{Exposure time}}$$

$$\begin{aligned}\text{Lung deposition fraction at stable time} = \\ \text{Lung deposition fraction at exposure time} + \\ + (\text{Stable time} - \text{Exposure time}) * \text{Change of lung deposition fraction per hour}\end{aligned}$$

with "120" the number of hours in five days, and the "Stable time" the time at which the threshold *in vitro* deposition rate is reached.

Such calculation provides an indicative exposure level, as the change in lung deposition fraction is not linear. However, given that the "Stable" time is often similar to the Exposure time, we consider this an acceptable approximation. If a more precise estimation is needed, the user can create a new entry in the input template file where the Exposure time corresponds to the "Stable" time, and run the model for the new entry. In this way, a precise lung deposition fraction will be estimated via the MPPD software.

The results of the analysis are then saved in a csv file, with name "Codo_results_Year-Month-Day_hour_minute", where the date and time is the one of the simulation. The successful completion of the full simulation and its duration are print to screen:

```
"Task completed. Time for simulations: 4 seconds"
```

4 Results

The final results of the simulation are saved in a csv file; two additional files related to the *in vitro* dosimetry simulation are optionally saved depending on the indications on the input template file.

4.1 CoDo results file

The final results file, named "Codo_results_Year-Month-Day_hour_minute" (where the date and time is the one of the simulation) includes a summary of the input data, the results of the *in vitro* simulations, and the exposure levels (Figure 4.1). The file reports, for each entry, the following information (the parameters reported from the input template file are not explained):

- ID;
- Substance;
- Primary particle diameter nm;
- Agglomerate diameter nm;
- Initial concentration *in vitro* in mg/cm^3;
- End time: the exposure time *in vitro*, in hours;

1	ID	Substance	Primary_particle_diameter_nm	Agglomerate_diameter_nm	Initial concentration_mg/cm3	End_time_h	Equilibrium_time_h	Fraction_deposited_end
2	1	TiO2	17	900	0.256	24	8.5	0.909326
3	2	Amorphous SiO2	12	220	0.01	6	not define	0.004484

Fig. 4.1: An example of some of the parameters saved in the CoDo result file.

- Stable time: the time in hours needed to reach a "stable" deposition rate. If the "Extend the simulation time until the defined deposition rate is reached" parameter is set to False, the file will report "not defined";

- Fraction deposited end: The fraction of particles deposited at the "End time", i.e. the exposure time;

- Mass per area end: The mass deposited per square centimeter *in vitro* at the end of the exposure time, in mg/cm^2;

- Mass + Dissolved per area end: The sum of the mass deposited per square centimeter and the fraction of the dissolved particles that are in the bottom of the well *in vitro* at the end of the exposure time. Since the dissolved species are uniformly distributed in the media, the dissolved mass corresponds to the concentration of the dissolved species multiplied by the height of the subcompartment considered as the bottom of the well. In mg/cm^2;

- SA per area end: The surface area of the deposited particles per square centimeter *in vitro* at the end of the exposure time, in cm^2/cm^2;

- N per area end: The number of particles deposited per square centimeter *in vitro* at the end of the exposure time, in #/cm^2;

- Mass per area eq: The mass deposited per square centimeter *in vitro* at the Stable time, in mg/cm^2;

- SA per area eq: The surface area of the deposited particles per square centimeter *in vitro* at the Stable time, in cm^2/cm^2;

- N per area eq: The number of particles deposited per square centimeter *in vitro* at the Stable time, in #/cm^2;

- Agglomerate density air: The density of the particle or agglomerate in air, in g/cm^3. If the air particle type was set to "pp", i.e. the primary particle is considered, the density is the same as the primary particle's. If the agglomerate with the same size as the agglomerate *in vitro* is chosen, the density of the agglomerate in air is calculated as:

$$\text{Agglomerate density in air} = \\ = \frac{\text{agglomerate effective density} - \text{media density}}{\text{primary particle density} - \text{media density}} \cdot \text{primary particle density}$$

- Sticky bottom;

- SA per area end pp: The same as SA per area end, but the surface are is calculated not from the surface area of the agglomerate, but from the specific surface area (SSA) of the primary particle:

$$\text{SA per area end pp} = \text{Mass per area end} \cdot \text{SSA}$$

- Sex;

- Air particle type;

- Exposure concentration air end: The exposure concentration needed to obtain the same deposited dose as *in vitro*, considering the same exposure time also *in vivo*. In mg/m^3.

- Exposure concentration air eq: The exposure concentration needed to obtain the same dose deposited *in vitro* at the Stable time, considering the same exposure time also *in vivo*. In mg/m^3.

- Exposure concentration air five d: The exposure concentration needed to obtain the same deposited dose as *in vitro*, considering five days of exposure on the workplace *in vivo*. In mg/m^3.

- Exposure concentration air year: The exposure concentration needed to obtain the same deposited dose as *in vitro*, considering one year of exposure on the workplace *in vivo*. In mg/m^3.

- Exposure concentration air worklife: The exposure concentration needed to obtain the same deposited dose as *in vitro*, considering 35 years of exposure on the workplace *in vivo*. In mg/m^3.

4.2 Optional plots file

For each entry for which the parameter "Plot graphs True/False" has been set to True, three graphs will be plotted and saved as a single pdf file: the deposited mass per area over time, the deposited fraction over time, and the surface area of the particles deposited per area over time (Figure 4.2). The file name will be "Deposition_Output file name"; if the Output file name is not specified in the input template file, a default name will be created by combining the Substance name, the ID, and the date of the

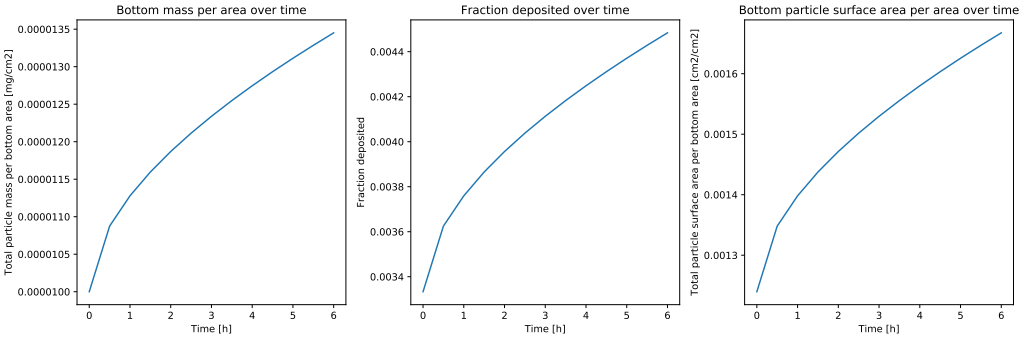

Fig. 4.2: The plots saved if the "Plot graphs True/False" parameter is set to True. Corresponds to the entry with ID 2 available as example in the template input file.

1		Time: 0.0 hours	Time: 0.5 hours	Time: 1.0 hours
2	Time at which the marginal increase in deposited fraction per h is < 0.002 (h)	not defined	0	0
3	Deposited fraction (of total mass)	0.003333333	0.003624886	0.003759398
4	Mass deposited per area (mg cm^-2)	1.00E-05	1.09E-05	1.13E-05
5	Mass deposited and dissolved fraction per area (mg cm^-2)	1.00E-05	1.09E-05	1.13E-05
6	Number of particles per area (cm^-2)	11174496.03	12151881.29	12602813.28
7	Particles surface area per area (cm^2 cm^-2)	0.001239669	0.001348098	0.001398123
8	Concentration of dissolved fraction (mg cm^-3)	0	0	0

Fig. 4.3: An example of the file saved if the "Save extensive results" is set to True. Corresponds to the entry with ID 2 available as example in the template input file.

simulation, and the file name will be "Deposited_Substance_ID_Day-Month-Year", e.g. "Deposited_TiO2_1_27-10-2020".

4.3 Optional individual *in vitro* result file(s)

For each entry for which the parameter "Save extensive results" has been set to True, a csv file will be saved reporting the information about particles *in vitro* deposition over time (Figure 4.3).

The file reports, for each time point defined by the "Output time interval minutes":

- The Stable time, i.e. the time at which the marginal increase in deposited fraction per h is < of the deposition rate threshold, in hours: if the option is not selected in the input template file, this will be reported as "not defined".
- The fraction of the total mass of particles that is deposited on the bottom/on the cells;
- The mass of the particles deposited per area;

- The total mass of the particles deposited per area, including the dissolved fraction.
- The number of particles deposited per area;
- The surface area of the particles deposited per area;
- The concentration of the dissolved fraction.

Bibliography

[1] Daina Romeo, Beatrice Salieri, Roland Hischier, Bernd Nowack, and Peter Wick. „Combined in vitro-in vivo dosimetry enables the comparison of in vitro doses and exposure levels: a proof of concept using titanium dioxide toxicity data.“ In: *Manuscript in preparation* (2020) (cit. on p. 3).

[2] Glen M. DeLoid, Joel M. Cohen, Georgios Pyrgiotakis, et al. „Advanced computational modeling for in vitro nanomaterial dosimetry“. In: *Particle and Fibre Toxicology* 12.1 (Dec. 2015), p. 32 (cit. on pp. 3, 9, 21).

[3] Frederick J. Miller, Bahman Asgharian, Jeffry D. Schroeter, and Owen Price. „Improvements and additions to the Multiple Path Particle Dosimetry model“. In: *Journal of Aerosol Science* 99 (Sept. 2016), pp. 14–26 (cit. on pp. 3, 4).

[4] Glen M. Deloid, Joel M Cohen, Georgios Pyrgiotakis, and Philip Demokritou. „Preparation, characterization, and in vitro dosimetry of dispersed, engineered nanomaterials“. In: *Nature Protocols* 12.2 (Feb. 2017), pp. 355–371 (cit. on p. 3).

[5] Satish Anjilvel and Bahman Asgharian. „A Multiple-Path Model of Particle Deposition in the Rat Lung“. In: *Fundamental and Applied Toxicology* 28.1 (Nov. 1995), pp. 41–50 (cit. on p. 4).

[6] Daina Romeo, Beatrice Salieri, Roland Hischier, Bernd Nowack, and Peter Wick. *An integrated pathway based on in vitro data for the human hazard assessment of nanomaterials*. Apr. 2020 (cit. on p. 4).

[7] B Asgharian, W Hofmann, and R Bergmann. „Particle Deposition in a Multiple-Path Model of the Human Lung“. In: *Aerosol Science & Technology* 34.4 (2001), pp. 332–339 (cit. on pp. 4, 11).

[8] Bahman Asgharian, Frederick J. Miller, Owen Price, et al. „Modeling particle deposition in the pig respiratory tract“. In: *Journal of Aerosol Science* 99 (Sept. 2016), pp. 107–124 (cit. on p. 4).

[9] Bahman Asgharian, Owen Price, Gene McClellan, et al. „Development of a rhesus monkey lung geometry model and application to particle deposition in comparison to humans“. In: *Inhalation Toxicology* 24.13 (Nov. 2012), pp. 869–899 (cit. on p. 4).

[10] Susanne Fritsch-Decker, Clarissa Marquardt, Tobias Stoeger, Silvia Diabaté, and Carsten Weiss. „Revisiting the stress paradigm for silica nanoparticles: decoupling of the antioxidative defense, pro-inflammatory response and cytotoxicity“. In: *Archives of Toxicology* 92.7 (July 2018), pp. 2163–2174 (cit. on p. 6).

[11] Sivakumar Murugadoss, Frederic Brassinne, Noham Sebaihi, et al. „Agglomeration of titanium dioxide nanoparticles increases toxicological responses in vitro and in vivo“. In: *Particle and Fibre Toxicology* 17.1 (Dec. 2020), p. 10 (cit. on p. 6).

[12] James S Brown, Terry Gordon, Owen Price, and Bahman Asgharian. „Thoracic and respirable particle definitions for human health risk assessment“. In: *Particle and Fibre Toxicology* 10.1 (Apr. 2013), p. 12 (cit. on p. 11).

[13] H. J.H. Colebatch and C. K.Y. Ng. „Estimating alveolar surface area during life“. In: *Respiration Physiology* 88.1-2 (Apr. 1992), pp. 163–170 (cit. on p. 12).

Glossary

entry Corresponds to one row in the input template file, and represents one specific case to simulate. 1, 6, 8, 12, 13, 17, 18, 36

stable time We define as "stable time" the time at which the change in deposition rate will be equal or smaller than the "Threshold deposition rate as percentage per hour".. 25, 28–30

List of Figures