DMP - Multiscale modelling of pome fruit morphogenesis

Project Name DMP - Multiscale modelling of pome fruit morphogenesis - DMP - Multiscale modelling of pome fruit morphogenesis

Grant Title 1189422N

Principal Investigator / Researcher Bart Dequeker

Project Data Contact +32 478 63 30 06, bart.dequeker@kuleuven.be

Description Pome fruit such as apple and pear constitutes a major export product from Flanders. Important visual quality factors in determining the commercial value and consumer acceptance of the fruit, are its size, shape and structure. However, the processes driving and controlling the development of these features during growth are poorly understood. In order to sustain the supply of healthy and high quality fruit while adapting to climate change and reducing greenhouse gas emissions, it is therefore crucial to gain insight into the development of growing fruit. Hereto in this PhD project, a mathematical model of water transport and mechanical deformation of growing apple fruit is developed, taking into account the 3D multiscale geometry and vascular formation, for investigation of the quality changes during growth.

Institution KU Leuven

1. General Information

Name applicant

Bart Dequeker

FWO Project Number & Title

1189422N Multiscale modelling of pome fruit morphogenesis

Affiliation

- KU Leuven

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data
- Reuse existing data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of data	Format	Volume	How created
Apple and pear (tissue) CT scans	.tiff (uncompressed)	25 TB	Acquisition from X-ray CT scans
CT image processing	.am, .m, .png, .stl, .csv		CT image processing files, including visual and textual results from outputs
Fruit measurements	.xlsx, .csv	250 MB	Manual measurements and reused data from literature and databases
model files	.m, .py, .mph, .cpp	250 MB	Created model files
Textual model outputs	.csv	500 MB	Automatic outputs of model simulations
Visual model outputs	.jpg, .png, .gif, .vtk	2 GB	Automatic outputs of model simulation

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will

use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- No

N/A

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- No

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

1. X-ray (micro-)CT images: automatically generated .txt files containing the scanning and reconstruction settings will be included to the folders containing the projection and reconstructed (micro-)CT images, respectively. Additionally, a template based Word document will be added for each scan, containing information on the scanned fruit material, such as orchard, picking date, storage conditions, storage period and cultivar. Also, additional information on sample preparation (e.g. contrast solution and incubation, sample location in the fruit, ...) will be added to this document.
2. Fruit measurements: the first sheet will be used to summarize the set-up of the experiment and to provide enough information to understand the raw and processed data. A summary of the experimental description and data structure for measurements in a separate PDF file will be available.
3. Model data: .txt ReadMe file with clear information about all the model files, required inputs, and instructions on how to run the model will be provided. All model files will be clearly commented in the code. Version control, documentation and joint management of model files will be implemented with Gitlab.
4. Model outputs: labels and units will be provided for all variables. .txt ReadMe file will be provided for each category explaining both outputs and simulation settings generating the data.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

No metadata standard is available for data used and generated. The main effort will be to align metadata and documentation with group common practices through a metadata portal RDIS designed by MeBioS group.

I will use guidelines for documentation of used programming languages (Matlab, Python, C++) with provided metadata infrastructure on Gitlab for version control and documentation. ReadMe files and documentation will be provided as explained in the previous entry.

5. Data storage and backup during the FWO project

Where will the data be stored?

The data will be stored on a network drive of KU Leuven. The folder is open for the group

members and is secured and backed-up by the ICTS service of KU Leuven. The model's code will be available at the university GitLab with restricted access to the research group. Published data will be available according to the publisher standards. All data are locally saved on one of the hard disks on the lab's personal desktop computer.

How is backup of the data provided?

The general network drive is maintained by the ICTS service of KU Leuven with automatic daily back-up and mirror procedures. Automatic backup of the drive is further secured on university One Drive cloud with regular automatic backup as per the group guidelines. The backup of code is secured by Gitlab versioning with every commit.

Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available then explain how this will be taken care of.

- Yes

Yes, KU Leuven provides sufficient storage and back-up capacity during and after the project. A dedicated storage share will be made for the project on which the collaborators will work jointly and store data files. For the expected large volume of 3D image data, we will rent separate large volume storage of the ICTS service.

What are the expected costs for data storage and back up during the project? How will these costs be covered?

Type 1 server back-end storage with mirror backup for the project share folder will cost 519 Euro per Tb per year. The large volume storage will cost 156.60 Euro per Tb per year. Costs will be covered by the project consumables budget.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The network drive of the shared folder is secured by the ICTS service of KU Leuven with a mirror copy. Only specific lab members will have access to the shared folder. Unauthorized persons do not have access to this system. GitLab repository can be managed only by its author and responsible ICTS personnel and is secured by restricted access only to the lab members.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data obtained during this FWO project will be retained for the expected 5 year period. Even after this period, the data will remain available for lab members of the MeBioS group.

Where will the data be archived (= stored for the longer term)?

Essential data will be stored on the network drive on university’s central server. Essential model files will remain in Gitlab.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

Expected data sharing costs are the same as specified in section 5.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?

All data will be published and made available after the end of the project. Data with valuable IP will be protected prior to publication. The MeBioS group is implementing a web-based platform for sharing of CT data which can be used to share the 3D image data.

Where/how will the data be made available for reuse?

- In a restricted access repository

- Upon request by mail
- Other (specify):

The data will be stored and be available for lab members using a shared network drive and large volume storage provided by the university.

Metadata of all data will be available on the MeBioS metadata portal with a possibility to request the complete dataset for research purposes.

Model files will be available with restricted access to the lab members on Gitlab.

When will the data be made available?

- Upon publication of the research results

Who will be able to access the data and under what conditions?

All data will be available without restrictions to all the lab members. Published data will be available for everybody with access to the publication as per publisher's rules. Metadata of all data will be available on the MeBioS metadata portal RDIS with a possibility to request the complete dataset for research purposes.

What are the expected costs for data sharing? How will the costs be covered?

Expected data sharing costs are minimal and covered by university services.

8. Responsibilities

Who will be responsible for data documentation & metadata?

Bart Dequeker - the FWO fellow.

Who will be responsible for data storage & back up during the project?

Bart Dequeker - the FWO fellow, together with promotors and ICTS.

Who will be responsible for ensuring data preservation and reuse ?

Promotors and ICTS.

Who bears the end responsibility for updating & implementing this DMP?

Bart Dequeker - the FWO fellow.