Plan Overview

A Data Management Plan created using DMPonline

Title: Mycotoxin contamination management in large-scale maize storage in China: prediction, mitigation, and economic analysis

Creator: Bing Xia

Principal Investigator: Bing Xia

Data Manager: Bing Xia

Project Administrator: Ine van der Fels-Klerx

Contributor: Cheng Liu

Affiliation: Wageningen University and Research (Netherlands)

Template: Data Management Plan | Wageningen University and Research

ORCID iD: 0000-0003-0186-6263

Project abstract:

Maize, a globally important staple crop, is predominantly stored in large-scale silos to ensure food security and reduce post-harvest losses. These storage systems frequently face a critical challenge, that is fungal infection and growth, and subsequent production of mycotoxins, particularly aflatoxins (AFBs), deoxynivalenol (DON), zearalenone (ZEN), and fumonisins (FBs), by fungi such as *Aspergillus* and *Fusarium* spp. The contamination is aggravated by the microclimates within silos, localized zones of high temperature, humidity, and poor airflow, conditions that are conducive to fungal growth and difficult to manage by current control strategies. This research project aims to develop an integrated framework to reduce mycotoxin contamination in large-scale maize silos scientifically and economically. The framework will combine environmental simulation-driven predictive modelling, with engineered enzymatic biocontrol, and cost-effectiveness analysis. The main underlying research question is how large-scale maize silo storage can be optimized in a cost-effective way to reduce mycotoxin contamination through the integration of predictive modelling, enzymatic mitigation, and economic assessment.

This research consists of four subjects and related chapters and will systematically address the silo structural layout and existing control methods (Chapter 1), develop spatial-temporal mycotoxin prediction models (Chapter 2), evaluate biocontrol strategies (Chapter 3), and analyze cost-effectiveness and scalability of each scenario, individually and in combination, in Chapter 2 and 3. (Chapter 4). Therefore, this research aims to evaluate the feasibility of different storage scenarios during grain storage and to propose a practical framework for improving food safety and storage resilience.

ID: 205737

Start date: 04-09-2024

End date: 04-09-2028

Last modified: 05-07-2026

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

Mycotoxin contamination management in large-scale maize storage in China: prediction, mitigation, and economic analysis

A. Describe the research project

1. Name researcher (please, add your full name):

Bing Xia

2. What is the name of your department(s)?

- Wageningen Food Safety Research

3. What is the name of your chair group(s) or business unit(s)? English name and abbreviation for chair groups from this page; business units from this page (expand to Wageningen Research and keep expanding to find your specific division / group). Examples: Bioprocess Engineering (BPE) or Contract Research Organization (CRO).

Busciness Economics group (BEC)

4. Describe the organisational context of your research project.

DMP version (or date last modified)	Version 11, 13 February 2026
Supervisor / (co-)promotors	Prof. Dr. HJ van der Fels-Klerx Dr. Cheng Liu
Graduate School (WU only)	WASS
Start date of project	04/09/2024
End date of project	04/09/2028
Project number	205737
Funding body	CAAS

5. Give a short description of your research project.

Title	Mycotoxin contamination management in large-scale maize storage in China: prediction, mitigation, and economic analysis
Summary	Maize, a globally important staple crop, is predominantly stored in large-scale silos to ensure food security and reduce post-harvest losses. These storage systems frequently face a critical challenge, that is fungal infection and growth, and subsequent production of mycotoxins, particularly aflatoxins (AFBs), deoxynivalenol (DON), zearalenone (ZEN), and fumonisins (FBs), by fungi such as *Aspergillus* and *Fusarium* spp. The contamination is aggravated by the microclimates within silos, localized zones of high temperature, humidity, and poor airflow, conditions that are conducive to fungal growth and difficult to manage by current control strategies. This research project aims to develop an integrated framework to reduce mycotoxin contamination in large-scale maize silos scientifically and economically. The framework will combine environmental simulation-driven predictive modelling, with engineered enzymatic biocontrol, and cost-effectiveness analysis. The main underlying research question is how large-scale maize silo storage can be optimized in a cost-effective way to reduce mycotoxin contamination through the integration of predictive modelling, enzymatic mitigation, and economic assessment. This research consists of four subjects and related chapters and will systematically address the silo structural layout and existing control methods (Chapter 1), develop spatial-temporal mycotoxin prediction models (Chapter 2), evaluate biocontrol strategies (Chapter 3), and analyze cost-effectiveness and scalability of each scenario, individually and in combination, in Chapter 2 and 3. (Chapter 4). Therefore, this research aims to evaluate the feasibility of different storage scenarios during grain storage and to propose a practical framework for improving food safety and storage resilience.

6. List the individuals responsible for the following data management tasks.

Data collection	Bing Xia, CAAS group
Data quality	Bing Xia, supervisor
Storage and backup	Bing Xia
Data archiving / publishing	Bing Xia
Data stewardship / support	Bing Xia, supervisor
Any other role [......]	

7. I have requested a review of this data management plan from:

- The (coordinating) data steward of my chair group / business unit.

8. Name of the data management support staff and / or data steward consulted during the preparation of this plan and date of consultation.

Dr. Cheng Liu

B. Describe the data to be collected, software used, file formats and data size.

9. Will you use existing data for this project?

- Yes. Please specify below which data (e.g. DOI, URL, or storage location) and the terms of use (e.g. licence).

Existing data will only be used as background information and for writing literature reviews. No existing datasets will be reused for new analysis. All sources will be properly cited, and terms of use or licences will be respected where applicable.

10. Will new data be produced?

- Yes.

11. Please describe the data you expect to generate and / or use in the table below. Include reused existing data as well (as these are files that you manage and store).

File contents	Data type	Software	(Open) file format	Estimated size of each file (range)	Estimated number of files (range)
(e.g. lab analysis, gene sequence, interviews, lesion scores, etc.)	(e.g. numerical)	(e.g. Excel)	(e.g. .csv)	(e.g. 20-50 Mb)	(e.g. 50-100)
Grain storage environmental monitoring data	Numerical / tabular	Excel	.csv, .xlsx	0.1–5 MB	20–100
Grain quality analysis data	Numerical / tabular	Excel	.csv, .xlsx	0.1–5 MB	20–100
Mycotoxin analysis results	Numerical/chromatographic	HPLC software, Excel	.csv, .txt, .xlsx	0.5–20 MB	50–200
GC-IMS volatile compound data	Numerical / spectral / tabular	GC-IMS software, Excel, R	csv, .txt, .xlsx	5–100 MB	50–200
Statistical analysis scripts	Code / textual	R, Python	.R, .py, .Rmd, .ipynb	0.01–10 MB	20–100

12. Estimate how much data storage you require in total (e.g. by using the information in the table at question 11).

- 10-100 GB

C. Storage of data and data documentation / metadata during research

13. Where will the data,code and accompanying documentation / metadata be stored and backed up during the research project (see the WUR Data Storage Finder)? Include platforms you use to share data, collect data on, or send data to for processing or analysis.

- W:drive Enterprise File Storage (WUR network drive).
- WUR OneDrive for Business - only when an up to date version of the research data is also safely stored on the W:drive or Yoda.

D. Structuring your data and information

14. Give a (visual) representation of the folder structure you intend to use.

The project data will be organised using the following folder structure:
PhD_Project_Mycotoxin_Storage/ │ ├── 00_Project_Administration/ │ ├── Data_Management_Plan/ │ ├── Ethical_Approval_and_Permissions/ │ └── Meeting_Notes/ │ ├── 01_Literature_Review/ │ ├── Search_Strategies/ │ ├── Reference_Library/ │ ├── Screening_Tables/ │ └── Extracted_Data/ │ ├── 02_Experimental_Protocols/ │ ├── Sampling_Protocols/ │ ├── Grain_Quality_Assays/ │ ├── Mycotoxin_Analysis/ │ ├── GC_IMS_Analysis/ │ ├── ITS_Sequencing/ │ └── Molecular_Biology/ │ ├── 03_Raw_Data/ │ ├── Silo_Environmental_Data/ │ │ ├── Temperature/ │ │ ├── Relative_Humidity/ │ │ └── CO2/ │ ├── Grain_Quality_Data/ │ ├── Mycotoxin_HPLC_Data/ │ ├── GC_IMS_Raw_Data/ │ ├── ITS_Sequencing_Raw_Data/ │ ├── qPCR_Raw_Data/ │ ├── Microbial_and_Phenotype_Data/ │ └── Image_Raw_Data/ │ ├── Colony_Images/ │ ├── Microscopy_Images/ │ └── Gel_Images/ │ ├── 04_Processed_Data/ │ ├── Cleaned_Tables/ │ ├── Normalised_Data/ │ ├── Peak_Tables/ │ ├── Sequencing_Feature_Tables/ │ └── Combined_Datasets/ │ ├── 05_Scripts_and_Code/ │ ├── R_Scripts/ │ ├── Python_Scripts/ │ ├── Statistical_Analysis/ │ ├── Machine_Learning/ │ └── Figure_Generation/ │ ├── 06_Results/ │ ├── Statistical_Outputs/ │ ├── Model_Outputs/ │ ├── Tables/ │ └── Figures/ │ ├── 07_Manuscripts_and_Reports/ │ ├── Manuscript_Drafts/ │ ├── Supplementary_Materials/ │ ├── Conference_Reports/ │ └── Presentations/ │ ├── 08_Shared_Data/ │ ├── Data_for_Collaborators/ │ ├── Data_for_Publication/ │ └── README_and_Metadata/ │ └── 09_Backup_and_Archive/ ├── Archived_Raw_Data/ ├── Archived_Processed_Data/ └── Final_Versions/

15. Describe the file naming conventions you intend to use. Please give one or multiple example(s).

A consistent file naming convention will be used throughout the project. File names will be descriptive and will include the project name, data type or experiment, sample/group information, date, and version number where applicable. Dates will be written using the international format YYYYMMDD, and version numbers will use leading zeros, such as v01, v02, and v03.

The general format will be:

[project]_[data/experiment]_[sample/group]_[date]_[version].[extension]

Examples:

MycotoxinStorage_SiloEnv_TempRHCO2_Silo13_20260525_v01.csv

MycotoxinStorage_GrainQuality_MoistureAcidValue_Silo14_20260630_v01.xlsx

Raw and processed files will be clearly distinguished by adding terms such as Raw, Cleaned, Processed, or Final in the file name when needed. Spaces and special characters will be avoided; underscores will be used to separate information.

16. How will you distinguish between versions of files (multiple answers possible)?

- The designation 'vRAW' is added to file names that contain raw unaltered data (before any processing and cleaning). Any alteration of RAW data is done on a copy of the RAW data and appended with a version number which increases with each file modification (e.g. v01, v02, v03 etc.).
- Dates within file names are updated when files are modified.

E. Data documentation and data quality

17. Describe below what data documentation and metadata will accompany the data to help make the data findable, understandable, and reproducible.

- The WUR readme file template (see template at https://doi.org/10.5281/zenodo.7701727).

18. Describe what data and analysis quality controls will be used?

- We will use discipline specific community standards for labelling and coding of data. Please specify the community standard used.
- We will use standardised coding and terms of data throughout all experiments so that data descriptions are equal throughout various datasets created.
- Supervisors or peers will review the data and results for any anomalies (e.g. unexpected inconsistencies, outliers, correct labeling of data and / or treatments, correct and consistent coding applied, etc.).
- We will consult statisticians.
- Statistical model assumptions are adhered to and assessed (e.g. (residual) distribution analysis, outlier analysis, (accounting for) independence, homogeneity of variance, etc.).
- We will use repeated measurements to validate results (e.g. duplicate or triplicate analysis, multiple observer agreement, measurements taken over time, etc.).
- We will use standard and validated protocols where appropriate.
- We will perform preliminary (pilot) experiments to validate intended experimental methods.

F. Working with sensitive data (personal data, ethics), data ownership, sharing and access

19. Who is the (rights)holder of the data (commonly known as the owner of the data)?

- WUR is not the (only) (rights)holder of the data and a WUR approved formal (consortium) agreement or contract between WUR and other parties is present.

20. What is the data classification for your project (for example as specified in SmartPIA) taking into account the (privacy) sensitivity of the data?

- Negligible.

21. Is this project registered in SmartPIA?

- No. Please register in SmartPIA in the case (privacy) sensitive data is collected (when applicable: via your supervisor, the project manager, see guidance).

22. Please specify the (sensitive) data and privacy protection measures. Note that any measures undertaken should be consulted with the Information Security Officer (ISO) and Privacy Officer (PO).

- Data is classified as negligible and standard WUR security measures are undertaken.

23. Are there other ethical issues that need to be taken into account which may include approval from ethical committees?

- No.

24. Will there be any intellectual property (IP) rights or alternative applications or routes to impact (such as commercial interests) associated with the data?

- No.

G. Data archiving and publishing

25. Are there reasons to restrict access to the data or limit which data will be made publicly available?

- No.

26. Describe what data from question 11 will be archived internally (e.g. WUR network drive / Yoda@WUR) and not published, for a minimum of 10 years? Include the exact name for the storage medium chosen (see the WUR Data Storage Finder).

- Not applicable as data will be published.

27. What data will be published and made available for reuse via a data repository?

- Data not underlying an article or report will also be published. Please specify below which data listed in question 11.

Cleaned and processed grain storage environmental data, including temperature, relative humidity and CO_2 data; cleaned grain quality analysis data; processed mycotoxin concentration data; GC-IMS peak tables or processed volatile compound data; ITS sequencing data and feature tables;statistical analysis scripts; machine learning scripts; and processed figures or tables.

28. When will the data be available for reuse, and for how long will the data be available?

- Data will be available for at least 10 years as soon as the article or report is published and not required for any other article publication.

29. Which data repository do you intend to use to make the data findable and accessible (see the WUR Repository Finder)?

- Zenodo.
- 4TU.ResearchData.
- DANS Data Stations

30. Which metadata standard will be used to describe the data during internal archiving and / or depositing in a data repository?

- Metadata standard from DANS Data Stations, 4TU.ResearchData and / or Zenodo (which often are the DublinCore or DataCite standard).

31. Which licence/terms of use will be applied to the data?

- Open access (Creative Commons Attribution licence (CC BY); anyone can access and reuse with attribution).

H. Data management costs

33. What resources (in time and / or money) will be dedicated to data management, data archiving or publication, and ensuring that data is reusable? Indicate as well how these costs will be covered.

- The PhD candidate and supervisor will spend at least 10% of their time on research data management to approach the FAIR principles as much as possible.
- All costs for 10 year data storage and access management to that data after journal publication or report are covered by the research group / project.