1S02523N - Machine learning for failure identification and reliable mitigation of rotor eccentricity in electrical drives

A Data Management Plan created using DMPonline.be

Creator: Philip Desenfans

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Philip Desenfans

Data Manager: Philip Desenfans

Project Administrator: Philip Desenfans

Grant number / **URL:** 1S02523N

ID: 199053

Start date: 01-11-2022

End date: 31-10-2026

Project abstract:

At this moment, more than 300 million electric motors are in use globally (Source: ABB 2021). Their replacement costs, upon failure, are usually dwarfed by the economic impact of the production losses. Consequently, increasing the dependability and fault tolerance of electric motors is a prominent endeavour. Many major causes of failure, e.g. bearing faults, induce rotor eccentricity. In turn, rotor eccentricity seriously jeopardises the health of other components. This project aims to realise rotor eccentricity monitoring and mitigation to prevent further damage
and increase the dependability of electrical drives through built-in control firmware. The proposed strategy does so without making structural changes to the electrical drive itself, increasing the marketability. The first research goal is to extract rotor eccentricity from stator voltages and currents, avoiding costly direct measurements. The second research goal is to build linearisation models of electrical drives using deep Koopman theory. Deep Koopman allows a complete linearisation of the unknown system. The third research goal is to combine the linearisation models with linear control methods to achieve rotor eccentricity control in the presence of mechanical faults. The proposed research aims to advance the state of the art from predictive maintenance to active solution strategies, leveraging the power of machine learning for failure identification and reliable mitigation of fault effects.

Last modified: 25-04-2023

1S02523N - Machine learning for failure identification and reliable mitigation of rotor eccentricity in electrical drives
FWO DMP (Flemish Standard DMP)

1. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data.

				Only for digital data	Only for digital data	Only for digital data	Only for physical data
Dataset Name	Description	New or reused	Digital or Physical	Digital Data Type	Digital Data format	Digital data volume (MB/GB/TB)	Physical volume
Case Western Reserve University bearing fault dataset	Faulty ball bearing acceleration dataset using motor setup of Case Western Reserve University	Reuse existing data	Digital	Experimental	.mat	<1GB	
Magnetic Equivalent Circuit model motor simulation dataset	Motor state data generated using the Magnetic Equivalent Circuit induction motor model and rotor suspension model Python scripts	Generate new data	Digital	Simulation data	.npy; Python dictionaries for metadata; .csv	<1TB	
KU Leuven Bruges FMEC lab 1.1 kW motor dataset	1.1 kW induction motor setup with various fault options. Current and voltage measurements, potentially added with acceleration, rotor positional measurements in the future	Generate new data	Digital	Experimental	.csv; .xlsx	<100GB	
Magnetic Equivalent Circuit model	Induction motor magnetic equivalent circuit model Python script	Generate new data	Digital	Software	Computational Script	<100MB	
Rotor suspension model	Induction motor rotor suspension model Python script	Generate new data	Digital	Software	Computational Script	<100MB	
(Embedded) UMP minimisation algorithm training script	UMP minimisation algorithm based on Deep Koopman Theory training script	Generate new data	Digital	Software	Computational Script	<100MB	
(Embedded) UMP minimisation algorithm inference script	UMP minimisation algorithm based on Deep Koopman Theory inference script	Generate new data	Digital	Software	Computational Script	<100MB	

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type:

The data named 'Case Western Reserve University bearing fault dataset' can be found in: https://engineering.case.edu/bearingdatacenter/welcome. (Smith WA and Randall RB (2015) Rolling element bearing diagnostics using the case western reserve university data: A benchmark study. Mechanical systems and signal processing 64: 100-131.)

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- No

Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)? If so, please comment per dataset or data type where appropriate.

- Yes

The continued development of the UMP minimisation algorithm, which uses '(Embedded) UMP minimisation algorithm training script' and '(Embedded) UMP minimisation algorithm inference script', has the potential for commercial exploitation. The contents of these scripts can be applied to the control logic of commercial motor controllers to add UMP minimisation as a control objective. The models named 'Magnetic Equivalent Circuit model' and 'Rotor suspension model' can accompany the previously mentioned scripts inside embedded controllers to aid model-based control strategies or to aid data generation.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements/ research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain in the comment section to what data they relate and which restrictions will be asserted.

- No

2. Documentation and Metadata

Clearly describe what approach will be followed to capture the accompanying information necessary to keep data understandable and usable, for yourself and others, now and in the future (e.g., in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded).

Measurement datasets will be accompanied by metadata which is done in-document for .xlsx-stored measurements or using linked guidance documents which accompany the raw data stored in .csv. Simulation datasets are stored as Python dictionaries for small datasets with (1) metadata added as key names or as metadata string values, (2) as .csv or (3) .npy files with accompanying text-based documents for metadata description. Scripts are explained thoroughly using in-script annotation. Furthermore, notebooks are accompanied by added text-segments which include LaTex symbols, linking implementation to theory. For finished model scripts, e.g. 'Magnetic Equivalent Circuit model', demo notebooks are created which show the model's functions and usage to assist future reuse.

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

For active data storage, the KU Leuven research storage platform ManGO will be used. There, metadata standards are defined per data type to ensure adequate data description. For measurement data this includes: setup specifications, measurement device specifications, fault state specifications, time of capture and person/people responsible. For simulation data, this includes: simulation setup specifications, hardware specifications, fault state specifications, time of simulation and person/people responsible. Lastly, for scripts and models, this includes: code annotation, theory referencing at the beginning of the script, accompanied by a demo notebook with further explanation.

3. Data storage & back-up during the research project

Where will the data be stored?

Active research data which is shared between researchers will be uploaded to ManGO, the active data management platform. Research data which is not (yet) shared is stored on the personal OneDrive for Business cloud storage. Manuscript-specific data such as data used to generate figures is additionaly stored in the KU Leuven Research Data Repository (RDR).

How will the data be backed up?

Local data, scripts and files in general are backed up using OneDrive for business' file synchronisation. Alternatively, the data stored in ManGO, RDR or Teams channels is kept on the uploader's personal hard drive and the uploader's synchronised OneDrive.

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

When overestimating the provided required storage per named dataset, it can be seen that 2 TB of storage should be sufficient for the project. The personal hard drive can only store 1 TB of data, which can be fully backed up by the OneDrive for Business storage, capable of storing 2 TB. ManGO will be used during the active stage of research, which provides capacity at increments of 1 TB. Currently, 1 TB is offered free of charge. For future expansion, this can be increased at the cost of 35 euros/TB, which can be paid for using the bench fee (appr. 2500 euros/year).

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The shared research platform ManGO is moderated by myself. At this point, only three people (including myself) have access to the data. Further access has to be given by myself. Inside ManGO, several groups will be defined such as read-only groups for students, where access to files can be given on a file-to-file basis or folder-to-folder basis. Additionaly, individual files are shared using OneDrive for business, where the generated links restrict access only to the intended receiver.

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

At this point in time, the used 1 TB of data storage on ManGO is free of charge. Should this be expanded or modified, the FWO bench fee will be used to fulfill the 35 euros/(TB*year) cost.

4. Data preservation after the end of the research project

Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...).

The large training datasets generated using simulations models, i.e. 'Magnetic Equivalent Circuit model motor simulation dataset', or physical motor setups, i.e. 'KU Leuven Bruges FMEC lab 1.1 kW motor dataset', along with the final control algorithm scripts '(Embedded) UMP minimisation algorithm training script' and '(Embedded) UMP minimisation algorithm inference script' will be stored.

Where will these data be archived (stored and curated for the long-term)?

Data will be stored in the KU Leuven Research Data Repository (RDR) for a minimum of ten years. This data will be linked to their corresponding publications.

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

Depending on the amount of data needed for long-term storage, the 50GB offered free of charge by KUL RDR may not be sufficient. In this case, the RDM helpdesk will be contacted to know the cost hereof. The cost of the earlier stated 2TB upper limit can be estimated from the cost of ManGO as 70 euros per year.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in a restricted access repository (after approval, institutional access only, …)

The formed scripts and datasets are reused in guided student projects. The reuse of data mainly concerns the 'Magnetic Equivalent Circuit model motor simulation dataset', the 'KU Leuven Bruges FMEC lab 1.1 kW motor dataset' and the Python scripts 'Magnetic Equivalent Circuit model' and 'Rotor suspension model'.

If access is restricted, please specify who will be able to access the data and under what conditions.

Access to the ManGO project file system is authorized by myself. This will be given to students who require the data or models, supervisors with whom I collaborate (mainly dr. ing. Dries Vanoost), and colleagues with whom I collaborate (e.g. M.Sc. Zifeng Gong, M.Sc. Chandrakanth Kancharla). When writing data is not required, the person will be granted read-only access, potentially limited to the subfolder that person requires.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain in the comment section per dataset or data type where appropriate.

- No

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Data will be made available to those with access using ManGO for all mentioned datasets and scripts. Here, 'Case Western Reserve University bearing fault dataset' is an exception as it is publicly available online.

When will the data be made available?

The ManGO project 'Electrical_Drive_Condition_Monitoring' is made available since 25/04/2023. In the near future, the transition will be made to this platform. Currently, a Microsoft Teams Channel is used as a placeholder.

Which data usage licenses are you going to provide? If none, please explain why.

At the end of the project, data will be made publicly available in RDR using Public Domain Mark (PD) for datasets and e.g. the Mozilla Public License 2.0 for the created software.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

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

The costs for data sharing in KUL RDR are free of charge up to 50GB, larger datafiles can be covered using the FWO bench fee.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Philip Desenfans. Those who upload new data are initially responsible for providing data documentation of their data.

Who will manage data storage and backup during the research project?

Philip Desenfans

Who will manage data preservation and sharing?

Philip Desenfans

Who will update and implement this DMP?

Philip Desenfans