SUBJECT TITLE: PHYSICS-INFORMED DATA-DRIVEN APPROACHES TO NON-LINEAR SYSTEM-LEVEL ANALYSIS OF BATTERY STATE OF HEALTH

Select the Strategic Axes	
AI & Data	X
Autonomous Driving	
Carbon neutrality / clean energy	X
Connectivity	
Digitalization	
Electronics & Software	
Factory 4.0	
Materials	
Others (to be specified)	X

Others strategic axes: high voltage battery, durability and reliability

THE STELLANTIS LEADER OF THE DOCTORAL STUDENT:

Last Name, FirstName, Email. : LEFEBVRE, Nicolas, nicolas.lefebvre@stellantis.com, PLANTIER, Matthieu, matthieu.plantier@stelllantis.com
Direction, Département, Service : ENG/PRS/PCMP/PCBP/PSDR

DEVELOPEMENT OF THE RESEARCH PHD THESIS PROJECT (Context/ Objective/Planning/ Bibliography)

Desired start date: October 2025 **Duration : 3 years**

Context

High voltage battery lifetime prediction is a major issue in automative industry due to regulations on warranty contracts. For example, in Europe, the warranty period of the high voltage battery is imposed by completion to 160 000 km or 8 years, and the cost of replacement of the battery makes violation of this warranty period highly expensive compared to the selling price of the car (~20%). It is then of the most importance to have a robust and reliable prediction models of the State Of Health (SOH, i.e. the lifetime) of high voltage battery.

Battery State Of Health prediction models commonly rely on the evolution of the total amount of electricity generated due to electrochemical reactions in the battery against time in operation. Such models are mainly physics-based, with randomization of some of the parameters to account for the variability observed in lab tests. The scientific innovation of this project is to propose a switch from physics-based approaches to physics-informed data-driven approaches. The expected contributions are three folds:

- First, developing a joint modeling framework that - is useable for - and - takes benefit from - physicians and data-scientists,

- Second, providing a better understanding of predictions variability,
- Third, assessing the effects of prediction uncertainties on the downstream decisions making for battery design and operations management.

Objective

According to the state of the art, high voltage battery State Of Health models relies on ageing and degradation mechanism. A review of SOH models is provided in [1]:

- models with direct measurements techniques -continuous measurements recorded by the Battery Management System (BMS) such as time series of currents, voltages and temperatures- or measurements collected during experiments or procedures (Coulomb counting technique and controlled charging/discharging procedure, Hybrid pulse power characterization and electrochemical impedance spectroscopy, incremental capacity analysis and differential voltage analysis, …
- state-space models with observers, where sensor data are used to estimate parameters corresponding to underlying unobservable states using Kalman filter or particle filters (equivalent circuit models, electromechanical models),
- regression-based models including linear and machine learning regression,
- time-series models,
- survival type models that requires run to failure data,
- cumulative damage models,
- empirical/analytical models that fit empirical models to measurement data in order to capture relationships between SOH and stress factors (operation time, temperature, operational loads, …).

The main objective of this PhD is to focus on empirical/analytical models and the way they are challenged by validations test. Currently those tests are performed at module and pack level during roughly 1 year in accelerated conditions with high temperature. Some are based on standard constant current back to back cycling and others more on realistic customer usage. That kind of test, due to some specificities of “serial and parallel” cells connection, introduce some modeling and inference challenge. The research hypothesis of this PhD project is to assume that this category of models is still relevant and the objective is to challenge and improve models and methodologies in this category. Then the sub-objectives project are:

- Challenge/benchmark existing tools and algorithms to assess models’ parameters distributions through, typically, “Bayesian inference” methodologies since analytical solutions to assess this by the likelihood maximization might be tricky due to potential nonlinearity of considered models [2].
- Challenge the models’ structure by crossing data-scientist and physician expertise and questioning all the current limitations: e.g. correlated and/or “time” dependent parameters might be considered, investigating the use of causal inference, module weakest cell extreme value distribution Bayesian fitting, as well as extreme value theory to exploit correctly the fitted distribution [3].

- Exploit models to assess the high voltage battery durability and quantify risk associated to commercialization.

A second step in this PhD project could be to challenge the use of the lab tests by the improvement of validation plan and even Design of Experiments (DoE) to reduce the number of required tests, that are time consuming and expensive. This step may rely on recent works on DoE selections but potentially also on Active Learning and/or Transfer Learning [4, 5]. At last, a third step may be to investigate the use of Over The Air (OTA) data to validate the accuracy of SOH models and/or potentially improve them by their adaptations.

Organization of research work

The following table proposes planning for this research work where 2 major phases (Phases 3 and 4) are identified (between months 0 to 21 and months 15 to 33). The transition from phase 3 to phase 4 should allow the thesis to be refocused on objectives that may be redefined in view of the progress achieved and the difficulties encountered.

Phase 1: Months: 0-3	State of the art on Bayesian inference, extreme value theory, improvements and uncertainty propagation in risk assessment. *Note: the bibliographic activity will remain active throughout the whole duration of the project.*
Phase 2: Months: 0-6	Exploration of datasets used in the process, understanding of physical phenomena with an impact on high voltage battery durability. This work will be caried out in collaboration with specialists and fellows of battery ageing.
Phase 3: Months: 0-21	Benchmark of "Bayesian inference" methodologies and tools, as well as algorithms. This important phase may rely on datasets from providers or battery test center as well as simulated datasets.
Phase 4: Months: 15-33	Improvement of methodologies and algorithms associated to sub-objectives previously listed. Again, in this second important phase, these developments might be tested on practical datasets and on simulated ones. Depending on progress 2nd and/or 3rd steps might be processed. Internal (and external) communications have to be achieved.
Phase 5: Months: 30-36	PhD writing and developments sharing, within Stellantis, through software packages.

Bibliography

[1] Erik Vanem, Clara Bertinelli Salucci, Azzeddine Bakdi, Øystein Å sheim Alnes, Data-driven state of health modelling—A review of state of the art and reflections on applications for maritime battery systems, Journal of Energy Storage, Volume 43, 2021, 103158, ISSN 2352-152X.
https://doi.org/10.1016/j.est.2021.103158
[2] Zihao Zhou, David A. Howey, Bayesian hierarchical modelling for battery lifetime early prediction, IFAC-PapersOnLine, Volume 56, Issue 2, 2023, Pages 6117-6123, ISSN 2405-8963.
https://doi.org/10.1016/j.ifacol.2023.10.708

[3] Banfi, F., Cazzaniga, G., De Michele, C. Nonparametric extrapolation of extreme quantiles: a comparison study. Stoch Environ Res Risk Assess 36, 1579–1596 (2022).
https://doi.org/10.1007/s00477-021-02102-0
[4] Toth, C., Lorch, L., Knoll, C., Krause, A., Pernkopf, F., Peharz, R., & von Kügelgen, J. (2023). Active Bayesian Causal Inference. Advances in Neural Information Processing Systems: NeurIPS 2022 (Vol.35).
https://arxiv.org/pdf/2206.02063.pdf
[5] Jingxiang, L., Guan-Yu, H., Weiming, S., Junghui, C., A supervised functional Bayesian inference model with transfer-learning for performance enhancement of monitoring target batches with limited data, Process Safety and Environmental Protection, Volume 170, 2023, ISSN 0957-5820.
https://doi.org/10.1016/j.psep.2022.12.004

CANDIDATE :

% of time spent in the company during the PHD THESIS *: around 80%.

PHD thesis status: Stellantis hired

The candidate must have an engineering or a master degree with skills in applied mathematics and computer science. Moreover, knowledges in statistics, probabilities, statistical learning, reliability and safety are mandatory.

Good listening and synthesis capabilities are also expected: high voltage battery ageing is a complex system with many factors… It will be essential to interact effectively with all people who have in-house Stellantis knowledges on this.