Centre for Maintenance Optimization & Reliability Engineering

Director
Chi-Guhn Lee

Semi-annual report

June 13, 2019
Michael E. Charles Council Chamber

Mechanical & Industrial Engineering
UNIVERSITY OF TORONTO

cmore.mie.utoronto.ca | 416·946·7867

Table of contents

	Page number
Executive summary	3
Visits and interactions	9
C-MORE leadership activities	13
Overall project direction	17
Goals and retrospectives	17
Theoretical work	17
Collaboration with companies and site visits	20
Technical reports	23
Global study of current approaches to maintenance education and training	23
DSTL: long-term military asset procurement strategies	25
DND: propulsion diesel engine replacement models	47
TTC: track re-inspection schedule optimization	49
Kinross Gold: Caterpillar haul truck engine decision policy	53
Teck: project exploration	63
Development of digital twin for predictive maintenance	65
Digital twin of reheat furnace	81
Epsilon-BMC: Bayesian model combination approach to Epsilon-greedy exploration in model-free reinforcement learning	99
Ongoing research on the remaining useful life prediction for high voltage circuit breakers	107
Toronto Hydro: investment spike smoothing and asset risk growths	113

Executive summary
Chi-Guhn Lee, C-MORE Director

Introduction

The following report summarizes work undertaken between C-MORE and collaborating companies and notes the major changes at C-MORE since the meeting in December 2018.

As Director, I have continued to expand C-MORE's service and research portfolio by ensuring the engagement of consortium members, driving the research program in new directions, attracting new researchers, and expanding the Consortium. I am happy to say we have welcomed a new Consortium member: Department of National Defence.

This year, one of C-MORE's special activities included a professional development session co-hosted with PEMAC's GTA chapter. This was an opportunity to share our broad research area with maintenance and asset management professionals. Several contacts demonstrated interest in membership with C-MORE as a result of this event. Another important undertaking was the submission of an NSERC Collaborative Research and Development (CRD) grant, with Kinross and Titan as partners. This grant leverages industry funds with federal government funding. The title of this proposal is "Maintenance and reliability in the face of uncertain data."

Other special activities for C-MORE included the signing of a memorandum of understanding with TAMS, a new research group at Inha University in Korea, and hosting a three-day course in asset management, "Asset Management 4.0," immediately before the Progress Meeting.

At the same time, I have continued to lead a team of research graduate students doing research in reinforcement learning, learning in non-stationary environment, degradation modelling, and digital twin. Since the last consortium meeting in December, we have published at NeurIPS 2018 and in *European Journal of Operational Research*. I have been busy outside Toronto as well, speaking at a research retreat held by Fujitsu Co-Creation Lab in March 2019 and visiting LG CNS headquarter in Seoul in February 2019 and Teck in BC in May 2019.

In the following sections, I give more details about the work of the individual members of the C-MORE team and outline our projects with Consortium members.

The C-MORE team

Janet Lam, Assistant Director

In the first half of 2019, Janet continued working on various projects with all member companies through direct research as well as student supervision. In April, she delivered a two-day training course on EXAKT to member company Kinross. On May 9, she presented a talk, "A neural network approach to equipment health prediction using free-form comments," at the Reliability Conference in Seattle, WA. In June, she was one of five instructors in the Asset Management 4.0 course hosted by C-MORE.

Andrew K. S. Jardine, Professor Emeritus

Andrew Jardine has continued to make a valuable contribution to maintenance courses offered to professionals and students around the world, including a graduate course and an industry course, a guest lecture at the University of Toronto, and a training course at a Canadian manufacturing company. Andrew has also continued to work with companies. On April 5, he visited BrightOrder Inc. to discuss extensions to their Fleet Management software. Andrew is currently working with Dr. Albert Tsang of Hong Kong Polytechnic University and Dr. Sharareh Taghipour of Ryerson University to finalize the 3rd edition of the textbook *Maintenance, Replacement and Reliability: Theory and Applications*. Since December 2018, Andrew has been named to a number of committees and boards. He has been appointed as member of the International Editorial Advisory Committee of the West Indian Journal of Engineering, and an Advisory Board Member for the International Maintenance Association (IMA), presently working on "A Global Study of Current Approaches to Maintenance Education and Training." On March 12 and 19 and April 24, he attended meetings of the Awards Committee for the Plant Engineering and Maintenance Association of Canada (PEMAC). On March 30, Andrew was inducted as Fellow of the Engineering Institute of Canada (FEIC), the first member of Canadian Region of the Institute of Industrial and Systems Engineering (IISE) to be elected FEIC.

Dragan Banjevic, C-MORE Consultant

In his work with C-MORE, Dragan collaborated with Janet Lam on projects with Consortium members, notably with Kinross Gold and TTC, and to some extent with MOD and DND. He also provided help in some projects with C-MORE students.

Sharareh Taghipour, Ryerson, External Collaborator

Sharareh is currently supervising/co-supervising two postdoctoral fellows, seven PhD students, two Master's students, and two undergraduate. One of her PhD and two of her Master's students recently completed their programs. Sharareh finished a collaborative project entitled "Model recipes for incoming workload prediction" with Intelerad Medical Systems. She also received an Early Researcher Award from the Ministry of Economic Development, Job Creation and Trade for a project entitled "Intelligent predictive maintenance and production scheduling for Industry 4.0." Her John R. Evans Leaders Fund was approved by Canada Foundation for Innovation for "Industry 4.0 Smart Factory System." She is serving on a number of committees at Ryerson University, including the

Department Hiring Committee (DHC), Strategic Research Plan (SRP) Steering Committee, and Faculty Awards Committee.

Scott Sanner, University of Toronto
Scott has been involved in a range of applied projects covering network and power grid security, predictive modelling for residential HVAC, prediction of hospital readmission for Sunnybrook Hospital, uses of social media in financial applications, and a number of projects involving recommender systems for eCommerce applications. Scott also continues to engage in fundamental research on machine learning and data mining (paper accepted at SDM 2019), reinforcement learning (paper accepted at UAI 2019), deep learning (paper accepted at AAAI 2019), constraint satisfaction (a best paper award at CPAIOR 2018 and a new paper accepted to CPAIOR 2019), recommendation systems (two papers accepted to SIGIR 2019) adaptive user interfaces (paper accepted to CHIIR 2019), and sequential decision-making in support of the aforementioned applications.

Fae Azhari, University of Toronto
Two new people have joined Fae's research group, making a total of one post-doctoral fellow, four PhDs, four MASc students, and one undergraduate. Her projects include: complex naval asset management using sensor data, optimizing the fabrication and performance of multifunctional cementitious composites, the application of digital image correlation for effective non-contact strain measurements in SHM, the development of a sensing device for monitoring lateral soil pressure, the application of fibre optic sensors in torsional vibration monitoring, the development of a sensing system for gait analysis, bridge scour monitoring, and condition-based maintenance of bridges. Fae's student Scott Koshman attended RAMS and Niloofar Heirani attended IMABM, where both presented their work. Fae has been meeting with various people in the industry about research opportunities.

Ali Zuashkiani, Director of Educational Programs
Ali has been providing consulting services to various industries, such as oil and gas, power generation and distribution, mining, and petrochemical. He has been especially active working with a major utility company (Marafiq) to improve its Operation and Maintenance business processes and procedures. Since December, Ali has delivered courses in asset management, reliability centred maintenance, and spare parts management in Abu Dhabi, Kuwait, Dubai, and Toronto. He is also developing a five-day comprehensive spare parts management course with Don Barry and Steve Sinkoff. Ali was a session chair at the Maintenance 4.0 Digitalization Forum at the Reliability Conference and received CMRP and CRL designations.

C-MORE students and postdoctoral fellows

Li Yang, postdoctoral fellow
In the last six months, Li has made progress on the following three research topics. First, opportunistic maintenance for wind farms considering weather impacts; this work comprehensively investigates the impacts of wind conditions on system reliability, power generation, positive impact on maintenance and negative impact on maintenance (maintenance delays). Second, the development of a hybrid prognostic framework to

predict remaining useful lifetime; this framework innovatively incorporates a recurrent neutral network (RNN) and the Wiener-based degradation model to estimate model parameters of RUL. The project is in collaboration with Gaoyang. Third, a mission abort policy based on early-warning information; this work designs mission abort policies for typical mission-critical systems, such as UAV, submarines, nuclear devices and so on. The objective is to balance the trade-off between mission reliability and system survivability. Based on these, Li has authored six journal papers, including three accepted/published papers and three under review/revision

Danish Anis, PhD student
While Danish has been busy with coursework throughout this academic term, he has continued to make progress on digital twin for reliability. He has submitted an abstract to RAMS 2020 and one to the MIE graduate research symposium. He is currently working towards extending an LSTM-RNN technique to generate a probabilistic RUL prediction within a digital twin framework as a means of synchronization with changing operational states. In theory, an LSTM encoder-decoder (LSTM-ED) is used to train a neural network and reconstruct the sensor data input time-series corresponding to a healthy state. The resulting reconstruction error can be used to estimate health index (HI) training and testing sets. Using a time lag to record similarity between the HI curves, a weighted average of the final RUL estimation can thus be obtained and uncertainty can be quantified using Monte Carlo (MC) Simulation. This approach is being evaluated first on a publicly available engine degradation data set.

Jing Janice Cao, EngSci thesis student
Jing completed her undergraduate thesis on April 9. Her research project modelled the UKMOD procurement problem with an impairment factor, incorporating the investment amount and reward in the decision tree model with different impairment situations. She modelled the problem was modelled as a Markov decision process and developed a backward calculation for each decision and chance node. She simulated a 20-year period of impairment conditions were simulated and presented a distribution of optimal stopping year of investment

Kuilin Chen, PhD student
Kuilin is continuing his research on digital twin for reheat furnaces while a full-time employee at Arcelor Mittal Dofasco. He gave a talk at a PEMAC professional development event on March 21 and is presenting at the Progress Meeting.

Michael Gimelfarb, PhD student
Michael's recent paper, "Epsilon-BMC: a Bayesian model combination approach to Epsilon-greedy exploration in model-free reinforcement learning," was accepted for the UAI 2019 conference and will be presented in July. A more detailed review of this work is included in the report. He is currently working on knowledge transfer in reinforcement learning using graph-structured data, Bayesian approaches, and hierarchical RL.

Scott Koshman, PhD student
In January, Scott presented at the Reliability and Maintainability Symposium (RAMS) "Incorporating condition monitoring for multi-faceted decisions" (co-authored with

Professor Azhari). The paper will be published in the pending IEEE conference proceedings. At the same symposium, he participated in a breakout session with Department of Defense (USA) where they discussed developing the body of knowledge in reliability. In April, he completed his requirements for PhD candidacy by passing his qualifying examination where he presented his research proposal "Resolving irregular data scenarios in equipment health monitoring decision support systems for complex naval platforms."

Gaoyang Li, visiting PhD student
As Gaoyang wraps up his study abroad program at C-MORE, he is working on modelling degeneration through machine learning. He has developed a new integrating technique to fuse the stochastic process model with deep learning neural network. In this model, deep learning methods are used to learn the existing degeneration paths, and the stochastic process is employed to model the variability inside the data to get a confidence interval for the results provided by the deep learning model. The integrated method is expected to outperform the traditional stochastic process while maintaining the confidence interval of the prognosis results.

Avi Sokol, PhD student
As a flex-time PhD, Avi is developing a research project directly linked with his employment. His work is aimed at creating a decision support system to advise on the procurement of inventory; and pricing and sale. His initial project is an inventory problem for inventory valued at $3M in volume.

Gary Wang, EngSci thesis student
Gary completed his final year of undergraduate studies in Engineering Science: Mathematics, Statistics, and Finance Option. He has taken a position as a hedge fund analyst at Rosalind Advisers. He is also a co-founder of Team Flyhand at The Entrepreneurship Hatchery at the University of Toronto, where he is developing an application called Image-Cloud which uses its image-recognition capabilities to allow users to search relevant information through images rather than keywords.

C-MORE activities with consortium members

Defence Science and Technology Laboratory (DSTL)
In this term, DSTL and C-MORE worked on a decision-making tool for long-term projects using utility as the variable. The problem is based on making long-term decisions that may be expensive to adapt in the future, using limited information that is available in the present. The project formed a part of Jing Cao's thesis and will be presented at the June meeting.

Department of National Defence
DND is our newest member company, joining in March 2019. The initial project is a condition-based project for propulsion diesel engines on the city-class ships. We are currently identifying failure modes and definitions to set up the project.

Kinross
Having settled on a hazard model, we incorporated cost information for preventive and corrective maintenance actions to finalize the decision model. By applying the decision model retroactively, we were able to assess that had the decision model been applied, several failures would have been prevented, and many PM actions would have been advised to continue running.

Teck
In May, C-MORE visited Teck's Sparwood location to have a face-to-face meeting discussing potential areas for collaboration. A KPI-based physical availability model and a decision tool for mid-life evaluation were discussed, as well as possibilities for harnessing a hybrid degradation model with a Bayesian neural network.

Toronto Transit Commission
TTC wrapped up the NDT line-test project early in 2019. We are exploring options to create a scheduling tool that will implement the optimized line-test schedule. A new project was launched in May to discuss re-inspection intervals using the NDT inspection crew. This will be presented at the June meeting.

C-MORE educational programs

In a new venture, C-MORE offered a three-day course in asset management, Physical Asset Management 4.0, just before the June Progress meeting. Day 1 featured an introduction to asset management in the 21st century; day 2 focused on evidence-based asset management (EBAM), and day 3 dealt with machine learning.

Conclusion

As the executive summary makes clear, we have been busy at C-MORE, continuing some ongoing projects and starting new ones. I want to take this opportunity to thank C-MORE staff, students, collaborating faculty, and collaborating companies for their outstanding support of our mission.

Chi-Guhn Lee
June 2019

Visits and interactions with consortium members and others

Jan 23, 2019
Andrew Jardine was appointed a member of the International Editorial Advisory Committee of the West Indian Journal of Engineering for a three-year period (2019-2021).

Jan 28-31, 2019
Sharareh Taghipour attended the Annual Reliability and Maintainability Symposium (RAMS) in Orlando, Florida. At this conference she presented two papers: "Maintenance Effectiveness Estimation from Observable Covariate Data with Applications to Railway Industry" and "Energy-Efficient Optimization of Flexible Job shop Scheduling and Preventive Maintenance."

Feb 7, 2019 **Kinross**
Dragan and Janet visited the Kinross Toronto office to finalize the hazard model and discuss how to determine appropriate costs for the decision model

Feb 12, 2019
Andrew was appointed Advisory Board Member, International Maintenance Association (IMA).

Feb 19, 2019 **Inha University**
Professor Hosang Jung from Inha University Asia Pacific School of Logistics visited C-MORE to discuss a potential memorandum of understanding between the two research groups.

Feb 22, 2019 **Teck**
Alex Creagh, Kevin Hatch and Graeme Dillon had a call with C-MORE to discuss potential projects and our main areas of expertise.

Mar 7, 2019 **TTC**
Dragan and Janet visited TTC's Hillcrest office for a meeting with Aleks Urosevic, David Girodat, Mo Gaus, Jennifer Lu, Hossein Mohammadian, Mark Vella, and Horacio Werchow to wrap up the NDT line-test project and kick off the re-inspection project.

Mar 9-10 & 23-24, 2019
Andrew taught a graduate course at the University of the West Indies, Trinidad, titled: "Maintenance Analysis and Optimization."

Mar 12, 19, April 24, 2019
Andrew attended meetings of the Awards Committee for the Plant Engineering and Maintenance Association of Canada (PEMAC)

Mar 20, 2019 **Keolis**
Suzanne Manaigre and Chris White from Keolis visited C-MORE hiring undergraduate student interns for operations contracts for Kitchener-Waterloo light rail transit.

Mar 21, 2019 **PEMAC**
C-MORE and PEMAC co-hosted a professional development event at the University of Toronto campus. There were three presentations by C-MORE: Janet Lam, Chi-Guhn Lee, and PhD student Kuilin Chen.

Mar 22, 2019 **Shanghai BDO**
Chi-Guhn and Janet met with George Dai from the University of Toronto business development office to discuss strategies for making connections with businesses in Shanghai

Mar 26, 2019
Andrew co-taught University of the West Indies course for industry with Kishore Jaghroo titled: "Optimizing Maintenance and Reliability Systems in the Heavy Industries."

Mar 30, 2019
Andrew was inducted as Fellow of the Engineering Institute of Canada (FEIC), the first member of Canadian Region of the Institute of Industrial and Systems Engineering (IISE) to be elected FEIC.

Apr 4, 2019 **DND**
C-MORE and DND had a conference call to discuss our first project in their membership. We discussed initial data requirements and established a set of milestones.

April 4, 2019
Andrew gave a guest lecture on "Application of Data Analytics to Engineering Asset Management" to undergraduate students of MIE 469: Reliability and Maintainability Engineering, taught by Dr. Janet Lam

April 5, 2019
Andrew visited BrightOrder Inc. regarding extensions to their Fleet Management software.

April 9, 2019
Jing Janice Cao and Gary Wang gave their Engineering Science thesis final presentations.

Apr 12, 2019
Scott Koshman had his qualifying exam with committee members Professors Fae Azhari, Chi-Guhn Lee, and Scott Sanner. His thesis proposal was "Resolving Irregular Data Scenarios in Equipment Health Monitoring Decision Support Systems for Complex Naval Platforms."

Apr 17, 2019 **Toronto Hydro**
Chi-Guhn and Janet visited Sushma Narisetty at the Toronto Hydro College Street location to provide a summary of collaboration over the years and suggest possibilities of continued collaboration.

Apr 23, 2019 **Cerrejón**
C-MORE and Murray Wiseman had a phone call with David Velandia of Cerrejón to discuss their ISO55000 asset management process report and how to implement the recommendations.

Apr 30, 2019 **DND**
Janet had a conference call with DND to discuss details of the data and to select a promising candidate for further study. The propulsion diesel engines were chosen.

May 2, 2019
Gaoyang Li gave a lecture on Bayesian deep learning to Dynamic Optimization Laboratory's summer students.

May 6-10, 2019
Janet and Ali attended The Reliability Conference in Seattle. Ali served as a session chair in the Digitalization Forum track. Janet gave a talk entitled "A neural network approach to equipment health prediction using freeform comments."

May 7, 2019
Andrew taught day one of a three-day education and training course to a Canadian manufacturing company on "Weibull Analysis and Extensions."

May 13, 2019
Mike Gimelfarb's paper entitled "Epsilon-BMC: A Bayesian Model Combination Approach to Epsilon-Greedy Exploration in Model-Free Reinforcement Learning" was accepted to UAI-2019

May 14, 2019 **Teck**

Chi-Guhn and Janet visited Teck's Sparwood location to discuss a range of potential projects for collaboration.

May 16, 2019

Mike Gimelfarb submitted an abstract to NeurIPS-2019 for a paper on knowledge transfer in reinforcement learning using graph-structured data.

May 22, 2019 **Linde**

Chi-Guhn and Janet attended the Trade Commissioner's Service B2B meeting with Linde to demonstrate C-MORE's value proposition.

May 23, 2019 **TTC**

Dragan and Janet visited TTC's Dundas West location to present preliminary results on the re-inspection project and to probe the project's objectives and decision criteria

May 24-June 5, 2019

Andrew held meetings with Dr. Albert Tsang of Hong Kong Polytechnic University and Dr. Sharareh Taghipour of Ryerson University to finalize the 3rd edition of the textbook *Maintenance, Replacement and Reliability: Theory and Applications*.

May 28, 2019 **DND**

Janet had a conference call with DND to discuss details of the propulsion diesel engine data and possibilities for distinguishing engine failures from minor corrective maintenance actions.

May 29, 2019 **Titan**

Chi-Guhn and Li Yang visited Titan's Hangzhou offices to discuss the applications of machine learning for asset prognostics and condition-based maintenance.

June 3-7, 2019

Li Yang attended the 11th International Conference on Mathematical Methods in Reliability (MMR 2019) hold in Hong Kong. He gave a presentation titled "Operations and Maintenance of Wind Farms Incorporating Multiple Impacts of Wind Conditions" in the Preventive Maintenance track. The talk focused on an advanced opportunistic maintenance strategy incorporating both the positive and negative impacts of wind conditions.

June 10-12, 2019

C-MORE's team delivered a three-day course in asset management to participants from Colombia, Korea, and Canada

C-MORE leadership activities

Chi-Guhn Lee, Director

Since December, Chi-Guhn has continued to lead a team of research graduate students doing research in reinforcement learning, learning in non-stationary environment, degradation modelling, and digital twin. He was invited as a speaker at a research retreat held by Fujitsu Co-Creation Lab in March 2019. He and his team have published at NeurIPS 2018 and in *European Journal of Operational Research* since the last consortium meeting. He visited LG CNS headquarter in Seoul in February 2019 and Teck in BC in May 2019.

Janet Lam, Assistant Director

In the first half of 2019, Janet continued on various projects with all member companies through direct research as well as student supervision. In April, she delivered a two-day training course on EXAKT to member company Kinross. On May 9, she presented a talk titled “A Neural Network Approach to Equipment Health Prediction Using Free-Form Comments” at The Reliability Conference in Seattle, WA. In June, she was one of five instructors in the Asset Management 4.0 course hosted by C-MORE.

Andrew K. S. Jardine, Professor Emeritus

Andrew Jardine has continued to make a valuable contribution to maintenance courses offered to professionals and students around the world, including a graduate course and an industry course, a guest lecture at the University of Toronto, and a training course at a Canadian manufacturing company. Andrew has also continued to work with companies. On April 5, he visited BrightOrder Inc. to discuss extensions to their Fleet Management software. Andrew is currently working with Dr. Albert Tsang of Hong Kong Polytechnic University and Dr. Sharareh Taghipour of Ryerson University to finalize the 3rd edition of the textbook *Maintenance, Replacement and Reliability: Theory and Applications*.

Since December 2018, Andrew has been named to a number of committees and boards. He has been appointed as member of the International Editorial Advisory Committee of the West Indian Journal of Engineering, and an Advisory Board Member for the

International Maintenance Association (IMA), presently working on "A Global Study of Current Approaches to Maintenance Education and Training." On March 12 and 19 and April 24, he attended meetings of the Awards Committee for the Plant Engineering and Maintenance Association of Canada (PEMAC). On March 30, Andrew was inducted as Fellow of the Engineering Institute of Canada (FEIC), the first member of Canadian Region of the Institute of Industrial and Systems Engineering (IISE) to be elected FEIC.

Dragan Banjevic, C-MORE Consultant

In his work with C-MORE Dragan collaborated mostly with Janet Lam on projects with consortium members, notably with Kinross Gold, and TTC, and to some extent with MOD and DND. He also provided help in some projects with C-MORE students.

Sharareh Taghipour, Ryerson, External Collaborator

Sharareh is currently supervising/co-supervising 2 postdoctoral fellows, 7 PhD, 2 Masters, and 2 undergraduate students. One of her PhD and two of her masters' students recently completed their programs. Sharareh finished a collaborative project entitled "Model recipes for incoming workload prediction" with Intelerad Medical Systems. She also received an Early Researcher Awards from the Ministry of Economic Development, Job Creation and Trade for a project entitled "Intelligent predictive maintenance and production scheduling for Industry 4.0". Her John R. Evans Leaders Fund was approved by Canada Foundation for Innovation for "Industry 4.0 Smart Factory System". She is serving on a number of committees at Ryerson University, including the Department Hiring Committee (DHC), Strategic Research Plan (SRP) Steering Committee, and Faculty Awards Committee.

Scott Sanner, University of Toronto

Scott has been involved in a range of applied projects covering network and power grid security, predictive modeling for residential HVAC, prediction of hospital readmission for Sunnybrook hospital, uses of social media in financial applications, and a number of projects involving recommender systems for eCommerce applications. Scott also continues to engage in fundamental research on machine learning and data mining (paper accepted at SDM 2019), reinforcement learning (paper accepted at UAI 2019), deep learning (paper accepted at AAAI 2019), constraint satisfaction (with a best paper award at CPAIOR 2018 and a new paper accepted to CPAIOR 2019), recommendation systems (two papers accepted to SIGIR 2019) adaptive user interfaces (paper accepted to CHIIR 2019), and sequential decision-making in support of the aforementioned applications.

Fae Azhari, University of Toronto

Two new people have joined Fae's research group, making a total of 1 post-doctoral fellows, 4 PhDs, 4 MASc students, and 1 undergraduate. Her projects include: complex naval asset management using sensor data, optimizing the fabrication and performance of multifunctional cementitious composites, the application of digital image correlation for effective non-contact strain measurements in SHM, developing a sensing device for monitoring lateral soil pressure, the application of fibre optic sensors in torsional vibration monitoring, developing a sensing system for gait analysis, bridge scour monitoring, and condition-based maintenance of bridges. Fae's student Scott Koshman attended RAMS and Niloofar Heirani attended IMABM, and presented their work. Fae has been meeting with various people in the industry regarding research opportunities.

Ali Zuashkiani, Director of Educational Programs

Ali has been providing consulting services to various industries, such as oil and gas, power generation and distribution, mining, and petrochemical. He has been especially active working with a major utility company (Marafiq) to improve its Operation and Maintenance business processes and procedures. Since December, Ali has delivered courses in asset management, reliability centred maintenance, and spare parts management in Abu Dhabi, Kuwait, Dubai, and Toronto. He is also developing a five-day comprehensive spare parts management course with Don Barry and Steve Sinkoff. Ali was a session chair at the Maintenance 4.0 Digitalization Forum at The Reliability Conference and received CMRP and CRL designations.

Overall project direction

Janet Lam, Assistant Director

Goals and retrospectives

This section highlights the some of the main achievements in C-MORE for the period January 2019 – June 2019. This year, one of C-MORE's special activities included a professional development session co-hosted with PEMAC's GTA chapter. This was an opportunity to share our broad research area with maintenance and asset management professionals. Several contacts demonstrated interest in membership with C-MORE as a result of this event.

An NSERC Collaborative Research and Development (CRD) grant was submitted with Kinross and Titan as partners. This grant leverages industry funds with federal government funding. The title of this proposal is Maintenance and reliability in the face of uncertain data.

Other special activities for C-MORE includes the signing of a memorandum of understanding with TAMS, a new research group at Inha University in Korea, and hosting a three-day course in Asset Management 4.0 that immediately preceded the progress meeting.

Activities

Theoretical work

This section on theoretical work is oriented toward students' and postdoctoral fellows' research topics.

Name	Activity
Li Yang, postdoctoral fellow	In the last 6 months, Li has made progress on the following three research topics:

Name	Activity
	1) A hybrid prognostic framework to predict remaining useful lifetime. This framework innovatively incorporates recurrent neutral network (RNN) and Wiener-based degradation model to estimate model parameters of RUL. This project is in collaboration with Gaoyang; 2) Opportunistic maintenance for wind farms considering weather impacts. This work comprehensively investigates the impacts of wind conditions on (i) system reliability, (ii) power generation, (iii) positive impact on maintenance (opportunistic maintenance) and (iv) negative impact on maintenance (maintenance delays); 3) Mission abort policy based on early-warning information. This work designs mission abort policies for typical mission-critical systems, such as UAV, submarines, nuclear devices et al. The objective is to balance the trade-off between mission reliability and system survivability. Based on these, Li has authored six journal papers, including three accepted/published papers and three papers under review/revision.
Danish Anis, PhD student	While Danish has been busy with coursework throughout this academic term, he has continued to make progress on digital twin for reliability. In addition to submitting an abstract to the MIE graduate research symposium, his paper titled "Optimal RUL Estimation: A State-of-Art Digital Twin Application" has been accepted for the RAMS 2020 conference. He is currently working towards extending an LSTM-RNN technique to generate a probabilistic RUL prediction within a digital twin framework as a means of synchronization with changing operational states. In theory, an LSTM encoder-decoder (LSTM-ED) is used to train a neural network and reconstruct the sensor data input time-series corresponding to a healthy state. The resulting reconstruction error can be used to estimate health index (HI) training and testing sets. Using a time lag to record similarity between the HI curves, a weighted average of the final RUL estimation can thus be obtained and uncertainty can be quantified using Monte Carlo (MC) Simulation. This approach is being evaluated first on a publicly available engine degradation data set.
Jing Janice Cao, EngSci thesis student	Jing successfully completed her undergraduate thesis on April 9. Her research project modelled the UKMOD Procurement problem with an impairment factor, incorporated the investment amount and reward in the decision tree model with different impairment situations. She modelled the problem as

Name	Activity
	a Markov decision process and developed a backward calculation for each decision and chance node. She simulated a 20-year period of impairment conditions and presented a distribution of optimal stopping year of investment.
Kuilin Chen, PhD student	Kuilin is continuing his research on digital twin for reheat furnace while a full-time employee at Arcelor Mittal Dofasco. He gave a talk at a PEMAC professional development event on March 21. He will be presenting at the June progress meeting.
Michael Gimelfarb, PhD candidate	Michael began his PhD program in September 2018 under the co-supervision of Professor Scott Sanner and Professor Chi-Guhn Lee. His recent paper, "Epsilon-BMC: A Bayesian Model Combination Approach to Epsilon-Greedy Exploration in Model-Free Reinforcement Learning," was accepted for the UAI 2019 conference and will be presented in July. A more detailed review of this work is included in the report. He is currently working on knowledge transfer in reinforcement learning using graph-structured data, Bayesian approaches and hierarchical RL.
Scott Koshman, PhD student	Scott is a Flex-Time PhD Candidate in the Mechanical and Industrial Engineering department. He is supervised by Professor Fae Azhari and in his 2nd year of studies. His full time employer is the Royal Canadian Navy where he is a Naval Combat Systems Engineer (Officer) with a specialty in reliability engineering, supportability, and quality management. The focus of his research is Equipment Health Monitoring for Halifax Class Frigates. In January, he presented at the Reliability and Maintainability Symposium (RAMS) "Incorporating Condition Monitoring for Multi-Faceted Decisions" (co-authored with Prof Azhari) which will be published in the pending IEEE conference proceedings. At the same symposium, he participated in a breakout session with Department of Defense (USA) where they discussed developing the body of knowledge in reliability. In April, he completed his requirements for PhD candidacy by passing his Qualifying Examination where he presented his research proposal "Resolving irregular data scenarios in Equipment Health Monitoring Decision Support Systems for Complex Naval Platforms."
Gaoyang Li, Visiting PhD student	As Gaoyang wraps up his study abroad program at C-MORE, he is working on modelling degeneration through machine learning. He has developed a new integrating technique to fuse the stochastic process model with deep learning neural

Name	Activity
	network. In this model, deep learning methods are used to learn the existing degeneration paths, and the stochastic process is employed to model the variability inside the data, to get a confidence interval about the results provided by the deep learning model. The integrated method is expected to outperform the traditional stochastic process while maintaining the confidence interval of the prognosis results.
Avi Sokol, PhD student	As a flex-time PhD, Avi is developing a research project directly linked with his employment. His work is aimed at creating a decision support system to advise on the procurement of inventory; and pricing and sale. His initial project is an inventory problem for inventory valued at $3M in volume. His ongoing research interests are dynamic pricing and sales suggestions.
Gary Wang, EngSci thesis student	Gary completed his final year of undergraduate studies in Engineering Science: Mathematics, Statistics, and Finance Option. He has taken a position as a hedge fund analyst at Rosalind Advisers. He is also a co-founder of Team Flyhand at The Entrepreneurship Hatchery at the University of Toronto. They are currently developing an application called Image-Cloud, which utilizes its image-recognition capabilities to allow users to search relevant information through images rather than keywords.

Collaboration with companies and site visits

This section gives details on progress in research conducted with consortium members.

Member	Collaborations
Defence Science and Technology Laboratory	In this term, DSTL and C-MORE worked on a decision-making tool for long-term projects using utility as the variable. The problem is based on making long-term decisions that may be expensive to adapt in the future, using limited information that is available in the present. The project formed a part of Jing Cao's thesis and will be presented at the June meeting.
Department of National Defence	DND is our newest member company, joining in March 2019. The initial project is a condition-based project for propulsion diesel engines on the city-class ships. We are currently identifying failure modes and definitions to set up the project.
Kinross	Having settled on a hazard model, we incorporated cost information for preventive and corrective maintenance actions

Member	Collaborations
	to finalize the decision model. By applying the decision model retroactively, we were able to assess that had the decision model been applied, several failures would have been prevented, and many PM actions would have been advised to continue running.
Teck	In May, C-MORE visited Teck's Sparwood location to have a face-to-face meeting discussing potential areas for collaboration. A KPI-based physical availability model and a decision tool for mid-life evaluation were discussed, as well as possibilities for harnessing a hybrid degradation model with Bayesian neural network.
Toronto Transit Commission	TTC wrapped up the NDT line-test project early 2019. We are exploring options to create a scheduling tool that will implement the optimized line-test schedule. A new project was launched in May to discuss re-inspection intervals using the NDT inspection crew. This will be presented in the June meeting.

Global study of current approaches to maintenance education and training

Andrew Jardine

The International Maintenance Association (IMA) is a worldwide organization comprising members from 30 countries on five major continents, including North America, South America, Africa, Australia, the Middle East, the Far East, and Europe. Members include researchers, senior managers, and Subject Matter Experts in the maintenance industry, specialists, curriculum developers, and trainers in different fields. This cluster aims to enhance the management of physical asset maintenance. The IMA vision is “global platform for knowledge exchange of effective maintenance practices and management.”

A Committee was struck (I am one or the 11 members) to develop a questionnaire seeking insights into maintenance education and training.

The questionnaire is a part of a global study on “A Worldwide Overview of Current Practice of Maintenance Training and Education (MT&E)” dedicated to determining the shortcomings and gap analysis of ME&T to reach the best formula of ME&T globally and prepare guidelines at the international level to help local and regional organizations/ associations plan and strategically develop their maintenance education and training strategies.

A draft questionnaire has been prepared and subsequently reviewed by several “friends” of IMA. The questionnaire is now being finalized and will be circulated to an international group of training providers

The findings will be presented at "The 3rd IMA World Maintenance Forum" organized by the International Maintenance Association (IMA). The presentation will be under the conference theme of “First Global Meeting of Maintenance Societies / A Worldwide Overview of Current Practice of Maintenance Training & Educating.” The conference will be held in the Convention Centre, Lugano, Switzerland, 30-31 Oct 2019.

UK MoD: long-term military asset procurement strategies

Jing Cao

Background

Making economic decisions such as investment or procurement subject to various complex factors can be significant and risky to businesses, especially the development and research decisions for national military assets. These decisions can no longer sorely rely on the management's intuitive judgments and should be modelled in more technical decision analysis approaches. To support the decision-making process for management groups, the thesis focuses on developing a long-term investment strategy for military assets using decision tree models, Markov decision process and real options. Interactive programs are constructed with designed input variables to visualize different results and decision values based on different input variables from real-world estimates. The thesis adopts a real case raised by United Kingdom Ministry of Defence (UKMOD) on its aircraft procurement strategies and generalizes the problem to all similar military asset investments. It provides the entire process of formulating and addressing the investment decision problems based on external and internal factors and uncertainties.

Introduction

Running a business requires managers to make wise investment or operation decisions. Real-world projects are often complex and have substantial uncertainties in multiple phases. One decision may completely reverse the direction of a company and determine the success or failure of a business.

Personal judgement and subjective views are considered as one of the most significant parts in making decisions for companies. However, purely personal insights can be limited, biased and misleading. It is important to realize that human beings are imperfect in decision analysis as the decisions are not well supported by facts and are subject to risks and uncertainties.

As a result, the techniques of making wise decisions became an important topic of research. Conventional ways of decision making include statistical and economic methods, Return on Investment (ROI), Net Present Value (NPV) calculations, cost-benefit analysis, etc. These traditional ways of decision making ignore outcome uncertainty, the choice of investment timing and irreversibility of resource commitment [1]. Thus, there needs to be a more advanced and complex decision analysis framework that can evaluate and determine the proper investment or disinvestment.

The goal of the thesis project is to develop an optimizing process using decision analysis methods and real option theory to confirm role of decision analysis in business investment and procurement decisions. I will investigate a real-world business example for United Kingdom Ministry of Defence (UKMOD) and devise an optimization aircraft procurement model that will examine risks and uncertainties.

Literature review

Decision analysis

Decision analysis is a subject analysing and optimizing real world situations including forecasting return from new investments, understanding oil or gas market, managing research and development programs subject to uncertainties and risks. It helps management to address uncertainties and returns in a systematic, quantitatively and interactive way and reduces the risks involved with subjective judgement [3]. Decision analysis relies on seven steps: identify the decision situation and understand objectives, identify alternatives, decompose and model the problem, choose the best alternative, sensitivity analysis, further analysis and implementation [3].

Step 1: Identify the decision situation and understand objectives

The first thing to do in decision analysis is to explore every uncertainty and fact and understand the central value or ultimate goal of the decision. Talking to the stakeholders will be an essential way to establish the precise nature of the situation.

Step 2: Identify alternatives

With the well-established objectives and situations, we turned to identify and discover every possible alternatives of the decision. Sometimes the alternatives are hard to be explicitly discovered, and that requires experience and creativity to reveal the hidden solutions. Some techniques of creativity are useful too.

Step 3: Decompose and model the problem

This section involves decomposing the problem and understanding the structure of the problem. Dividing the problem into sub-pieces may be a good way to modelling each stage of the problem. Each part of the problem could be associated with different time stages or different uncertainty factors. Modelling of the problem is the most essential part of decision analysis. There are many different kinds of modelling, such as mathematical and graphic, hierarchical and network, or more specifically, decision trees, influence diagrams, and machine learning, etc.

Step 4: Choose the best alternative

After properly modelling the problem, choose or find one preferred alternative among all possible paths based on certain metrics, e.g. expected return, expected utility or certainty equivalent, etc.

Step 5: Sensitivity analysis

When the preferred alternative is chosen, the question that if we made a slight change to one of the factors, would the solution still be optimal will be raised. In order to answer that, sensitivity analysis is introduced to test the sensitivity of the decision with respect to small changes.

Step 6: Is further analysis needed?

Decision analysis is an iterative process. Once everything is settled down, we need to go back and check if further analysis is required according to the preferred solution. It might be necessary to refine some of the objectives or constraints due to various discoveries during the decision analysis process, so that all previous steps need to be re-performed. Such discoveries include the management's beliefs, likelihood of the uncertainties, preferences of the outcome, etc.

Step 7: Implement the chosen alternative

Once the preferred alternative is demonstrated to be the most efficient solution, implementation can be started.

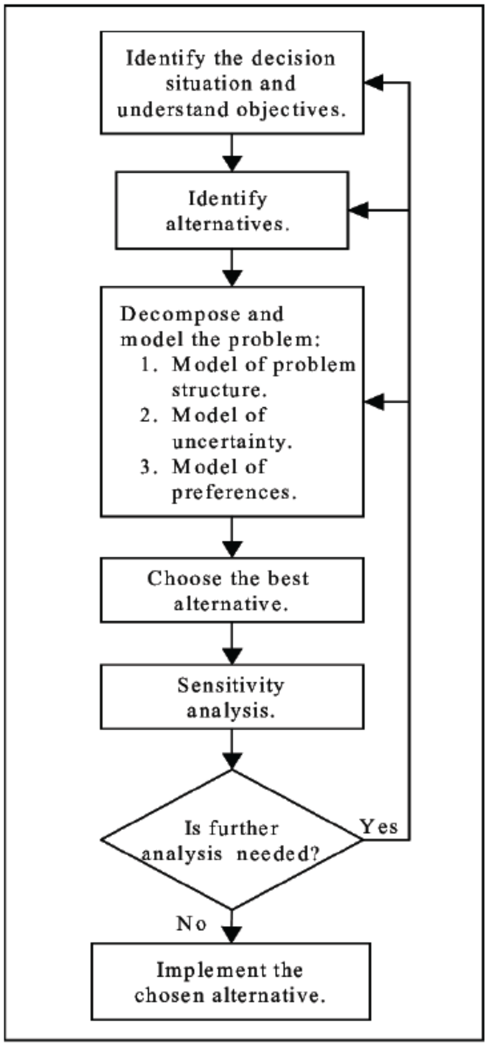


Figure 1: A decision analysis flow chart [4]

Decision tools

To perform the above analysis, one of the useful decision tools is the set of Excel add-on developed by Palisade's DecisionTools suite. The suite contains five programs: PrecisionTree, RISKview, BestFit, TopRank and @RISK. Each program is intended to be used for different stage of the decision analysis. PrecisionTree is used to structure decision tree models graphically and mathematically. It is also able to perform sensitivity analysis and model the decision variables with utility functions. TopRank is also used for sensitivity analysis and RISKview can be used to model uncertainties. @RISK is mainly for simulation purpose, one of the most commonly used examples is Monte Carlo simulations. These programs interact and support each other to provide a clear decision analysis interactive system.

Decision trees

Decision tree analysis is one of the most popular tools in decision modelling. In decision tree analysis, there are three essential elements: decisions, chances and consequences. Same as influence diagrams, decision tree uses squares to represent decisions and circles to represent chances. The branches followed by decisions are decision alternatives to make and the branches followed by the chances are possible outcomes with probabilities. Each branch has a consequence at the end.

Utility Functions

With certain probability distributions, we are able to calculate expected monetary value for decision alternatives. However, these monetary value does not capture risk attitude. For example, the negative impact of losing 100k to most people is larger than the positive impact of earning 100k. This is demonstrated as risk averse utility function [3]. Utility and wealth will display a concave shape for individuals who are risk averse a linear shape for the one that are risk neutral and convex for risk seeking. We assume in UKMOD example, our stakeholders are risk neutral as most of the people. Exponential utility functions capture risk attitude and has the general format of:

$$U(x) = 1 - e^{\frac{x}{R}}$$

The parameter R is called risk tolerance and determines how risk averse the utility function is. A larger value of R indicates a flatter utility curve and the individual is less risk averse.

Reinforcement learning

Markov Decision Process is one of the optimal control strategies within the scope of reinforcement learning. Reinforcement learning, similar to machine learning is simply learning what to do by setting a numerical reward signal and maximize it. In our decision analysis model, reinforcement learning helps to decide which actions to take by considering not only the immediate reward but also the following rewards in next series of the states. It composes two important features: trial-and-error search and delayed reward, which distinguish reinforcement learning from other methodologies.

One of the challenges in reinforcement learning is the exploration-exploitation dilemma, which demonstrate the trade-off between exploration and exploitation. Exploration is the process to explore actions that the agent has never selected before, and there is some possibility that the agent could get a larger reward. However, exploration consumes time and energy, while takes the risk of no better actions could be found, therefore exploration may not be always efficient. Exploitation, on the other hand, is just to use the actions that have already been experienced and get a known reward without looking for new actions to explore. Neither exploration nor exploitation can be used sorely in reinforcement learning process, so the trade-off between them is what we have to consider and balance. Reinforcement learning has four main elements which are essential for its environment: a policy, a reward signal, a value function and a model.

- A policy may be deterministic or stochastic. It is the core element of reinforcement learning as it can determine behaviour. It tells the mapping of following states or behaviours given the previous states and actions and this relationship can be as simple as a lookup take or as complex as some extensive computation and searching.
- A reward signal is the core element for the goal of reinforcement learning. As the ultimate objective that the agent is perusing is just to maximize the total expected reward, the reward signal defines the positive or negative impact that a certain action may bring to the whole process. It is also the primary reason for changing the policy: If an action results in a low reward, then the policy might be altered to choose a different action in the future.
- A value function specifies the value of a state considering all cumulative rewards in the future. It not only consists of the immediate reward, but also indicate the long-run desirability of a state by considering future rewards followed by the current state.
- A model is used for planning, by which is used to decide actions before actually experienced. It mimics the behaviour and make inferences of the environment. The ways to solve the reinforcement learning problem includes model-free and model-based methods.

One of the limitations of reinforcement learning is that it relies heavily on the state element by using it as the input and output variables of the policy, value function and the model. It does not define the state signals, but instead focuses on the value functions of the states what are already available in the environment.

Finite Markov decision process

Markov Decision Process is a reinforcement learning framing in order to achieve a goal. It has interactive relationship between an agent, which is a decision maker or learner and an environment as shown in below diagram. The environment responds with actions being taken and gives rise to rewards in numerical value form [5]. The ultimate goal in this process is for the agent to maximize the total rewards through choices of actions.

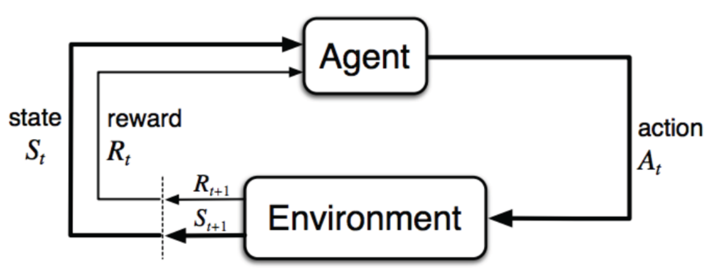


Figure 2: The agent-environment interaction in a Markov Decision Process

In finite MDPs, the agent receives the state of the environment $S_t \in S$ and select one of the actions $A_t \in A(s)$ at each time step, t = 0,1,2,3… by interacting with the environment. After the interaction, the agent receives a reward signal in numerical value form and finds itself in the next state S_{t+1}. Therefore, the agent follows a sequence that begins like:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, \ldots$$

The random variables R_t and S_t have discrete probability distributions depending on the previous state and the action taken because the MDP is discrete. Given the preceding state and an action, the probability of getting a reward r and jump to state s' is defined as:

$$\mathrm{p}(s', r|s, a) = \Pr\{S_t = s', R_t = r \mid S_{t-1} = s, A_{t-1} = a\}$$

for all $s', s \in \mathcal{S}, r \in \mathcal{R}, \text{and } a \in \mathcal{A}(s)$. The probability function p is an ordinary deterministic function based on the four variables.

In this process, we define objective as to maximize the total expected returns. In this case, at the current state S_t, we would like to sum up the sequence of returns and denote the sum as G_t, the expected return:

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + \cdots + R_T$$

where T is the final time step in the process. This approach works for processes with discrete episodes, and each episode has a terminal state followed by another round of starting state. The agent is learning how to choose the optimal A_t to maximize the total reward. Considering the discount factor $\gamma \in [0,1]$, the expected discount return G_t becomes:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots + R_T = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

If the discount rate is less than 1, then the equation is finite and could be maximized as long as $\{R_k\}$ is bounded. We can also write the expected return in a recursive form:

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \cdots \\ &= R_{t+1} + \gamma (R_{t+2} + \gamma R_{t+3} + \gamma^2 R_{t+4} + \cdots) \\ &= R_{t+1} + \gamma G_{t+1} \end{aligned}$$

This expected total return function works for all time steps before the terminal step. The other two elements value function and policy is defined by:

- Value function $v_\pi(s)$: the expected return when starting from a particular state s and follow by the policy π:

  $$v_\pi(s) = \mathbb{E}_\pi[G_t|S_t = s] = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1} | S_t = s\right]$$

  for all s ∈ $\mathcal{S}$.

- The action value function for policy π is defined as $q_\pi(s, a)$. It is the value of taking an action a at the state s considering all upcoming time steps:

$$q_\pi(s,a) = \mathbb{E}_\pi[G_t|S_t = s, A_t = a] = \mathbb{E}_\pi\left[\sum_{k=0}^{\infty}\gamma^k R_{t+k+1}|S_t = s, A_t = a\right]$$

- With the above definitions, we know one of the most important characteristics of reinforcement learning is that the value function satisfies the recursive relationship similar to the expected returns we defined before. Therefore, the Bellman Equation demonstrate the property by defining the value function in terms of the successor states:

$$\begin{aligned} v_\pi(s) &= \mathbb{E}_\pi[G_t|S_t = s] \\ &= \mathbb{E}_\pi[R_{t+1} + \gamma G_{t+1}|S_t = s] \\ &= \sum_a \pi(a|s) \sum_{s'}\sum_r \mathrm{p}(s', r|s, a)\,[r + \gamma \mathbb{E}_\pi[G_{t+1}|S_{t+1} = s']] \\ &= \sum_a \pi(a|s) \sum_{s',\mathrm{r}} \mathrm{p}(s', r|s, a)\,[r + \gamma v_\pi(s')] \end{aligned}$$

for all s ∈ $\mathcal{S}$. In this case, the value function is the unique solution to the Bellman equation and the Bellman equation provides a way to learn and approximate value function [5].

UKMOD: long-term procurement problem

To investigate and demonstrate the effectiveness of decision analysis, I worked with Centre for Maintenance Optimization and Reliability Engineering (C-MORE) on a military procurement problem raised by United Kingdom Military of Defence (UKMOD). C-MORE is a research organization under Department of Mechanical and Industrial Engineering at University of Toronto and UKMOD is one of the important clients of the organization. The thesis relies on the problem and conditions provided by UKMOD but can be generalized to the decision analysis of any military asset investment and procurement problem. The ultimate goal is to deliver a technologically and/or operationally complex end product using proper decision models.

Objectives and uncertainties

The objective of the analysis is to model the decision such that we can find an optimized process to invest in military aircrafts. Investing in military aircrafts is a long-term project. It requires multiple capital investments in consecutive years and each year is associated with the discrete decisions of continuing investing or stop the investment. The return from investing is unknown and subject to multiple uncertainties and risks, such as the source of procurement, military requirements, technology, etc. There will be a range of internal and external factors and uncertainties that can affect the efficiency and appropriateness of each potential approach in a given situation. It is desired to devise an optimization process that will examine the risks and uncertainties in a given situation and will calculate the range of potential outcomes for each approach and hence indicate range of potential approaches and which risks will need to be especially considered in each case.

The research gap is to develop a decision analysis strategy especially for the case of military assets for UKMOD by implementing various decision analysis models and tools. There are no specific data sets for this research and all the variables are designed to be input and can be modified to perform sensitivity analysis. The strategy should be concise and clear in terms of presentation and user interactions such that the management of military asset, e.g. UKMOD management, is able to modify the inputted variables and perform decision analysis to support their decision making.

Suggested modelling strategies

There are three main options of procurement strategies that could be adopted suggested by UKMOD: waterfall model, incremental model, and iterative model.

- Waterfall Model – In this approach the requirements for the final product are set, and then design, production and introduction into service activities are followed sequentially, delivering the final product in a single iteration. Decisions could only be made at the beginning of the process and cannot be modified once the investment is settled.
- Incremental Model – In this approach the requirements for the final product are set, together with the requirements for an interim level of performance to be delivered. The design work for the initial level of performance is undertaken, informed by the requirements that exist for the final level of performance, followed by the production and introduction into service of the initial capability. Once this has been achieved the design, production and introduction into service of the final version can be undertaken.
- Iterative Model – This approach is composed of a series of waterfalls, that iteratively work towards a final solution, but rather than setting the final requirements as the first activity, each iteration has a new set of requirements, informed by the results of the previous iterations. In this approach, decisions can be modified constantly at the beginning of each time step to adjust the direction of the investment to match the future expectations. However, some of the chances are lost due to delay in decision-making or diseconomies of scale.

Problem formulation

Using the utility theory as demonstrated in above literature review sections, we would like to find the stakeholders' utility for each possible scenario. The optimized or preferred alternative will have the highest utility value. We shall consider the below factors when measuring utility:

- Investment in the aircraft: The amount of investment on aircraft each year is a necessary factor in utility. If initial investment amount is too high and the return is expected to be unsatisfactory, it might decrease expected utility and hence affect the initial investment decision. Due to budget constraint, we assume lower investment with higher return will give more satisfaction and higher price with lower return will give less satisfaction.

- Maintenance and operating costs: The maintenance and operating costs or other related costs associated with the aircrafts, similarly, are subject to the budget constraint. Therefore, we assume the stakeholders would prefer less associated costs with respect to certain amount of return.
- Contribution to the mission: It is the core component of the reward and satisfaction to national defence because military aircrafts are mainly developed to meet certain military tasks or missions. The investments are always better to fit in the missions properly in order to give more satisfaction. Compared with the large value of reward brought by the proper contribution to the missions, the investment would be less concerned to the national defence management, and they are willing to spend large amount of investment to satisfy certain important missions.
- Impairment: The internal or external factors will cause the aircraft to reduce its value, either due to regular amortization or unexpected technology from enemies. Details to be followed in the one-factor decision model section.
- Unexpected events such as wars. If a war happens, the demand for military aircrafts will increase significantly and thus create a completely different utility function.

Each aspect has dynamic numerical values which can be stochastic variables over the time period. The relationship of each factor and the satisfaction may be linear or quadratic or exponential, which should be determined and modelled by the specific stakeholders or scenarios. In this project, we use exponential utility functions to model the relationship between wealth values and utilities. The relationship demonstrates a concave exponential curve, indicating the risk-averse nature of the investors.

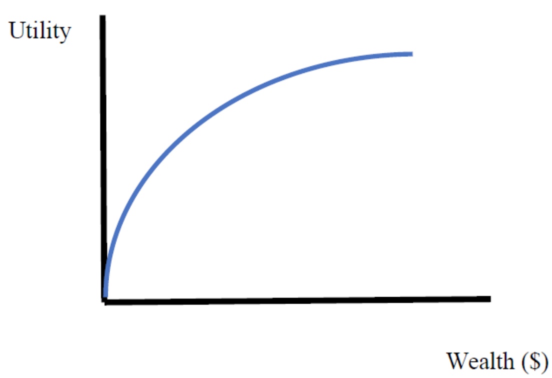


Figure 3: The exponential utility function for risk-averse investors

The concave curve turns flatter as the wealth value increases suggests that the risk-averse investors tend to get less satisfactions for the same unit of increase in wealth value when the wealth value gets very large. It is exactly opposite for the risk-seeking investors. For military asset holders, the nature of risk-averse investors is more appropriate because the national defence will not enjoy risky investments using the military budget.

The thesis assumes the military asset holders are able to provide numerical values of expected investment or reward values in terms of wealth value for each time step during the investment. The analysis process will also need to understand the risk-averseness of the investors and estimate proper variables to determine the utility function which could convert wealth value into satisfactions.

One factor model on impairment

Two different types of impairment

To better model the problem and understand what my stakeholders require, I decomposed the problem into several parts. Each part is associated with one uncertain factor illustrated above. We pick the factor "Impairment" to begin with. After back and forth communication with the military asset management group, we understand that for aircraft impairment, there are mainly two types of impairment described as the following:

- **Regular impairment:**
  Regular amortization of aircraft invested, usually moderate percentage (2-5%) of aircraft's value is impaired each year. An increase in impairment amount decreases the wealth value of the aircraft and hence decease the utility of the asset holders. In addition, we understand the fact that at early years of investment (1-3 years after initial investment), the impairment of aircraft can be negative, i.e. the value of the aircraft increases as UKMOD gets more familiar with the functions of aircrafts and the aircrafts become more useful. After typically 3 to 5 years, the regular positive impairment happens, which decrease the value of aircraft and hence decrease the utility.
- **Unexpected impairment:**
  This occurs when something unexpected happens which significantly reduces the value of our aircraft. For example, when enemies invent some kinds of witchcraft that completely counter the aircraft we invested and make our investment completely useless. Another example will be some unbelievable and advanced technology appears such that we no longer need the military aircraft we invested and therefore we could get a largely negative reward or a dramatic decrease in utility.

Interactive program for impairment analysis

Considering the one factor model for impairment, I decomposed the problem into several time periods, each period indicates 1 year. For each year, the management group can estimate and input the following variables:

- The amount of regular impairment in percentage (2-5%)
- The percentage decrease of the aircraft value when unexpected impairment happens (typically a large percentage of the aircraft value would be written off)
- The probability that the unexpected impairment happens (very unlikely at the beginning of the investment, small probability several years later)

- The utility coefficient (indicates the risk-averseness of the investors)
- An initial level of base utility

With these variables, a utility function can be constructed to convert the reward and investment into satisfaction. An expected utility is calculated by incorporating the estimated probabilities and returns for each scenario. This interactive program allows the management to analyse the expected satisfactions by inputting variables and constructs a visualized diagram as shown in the example below.

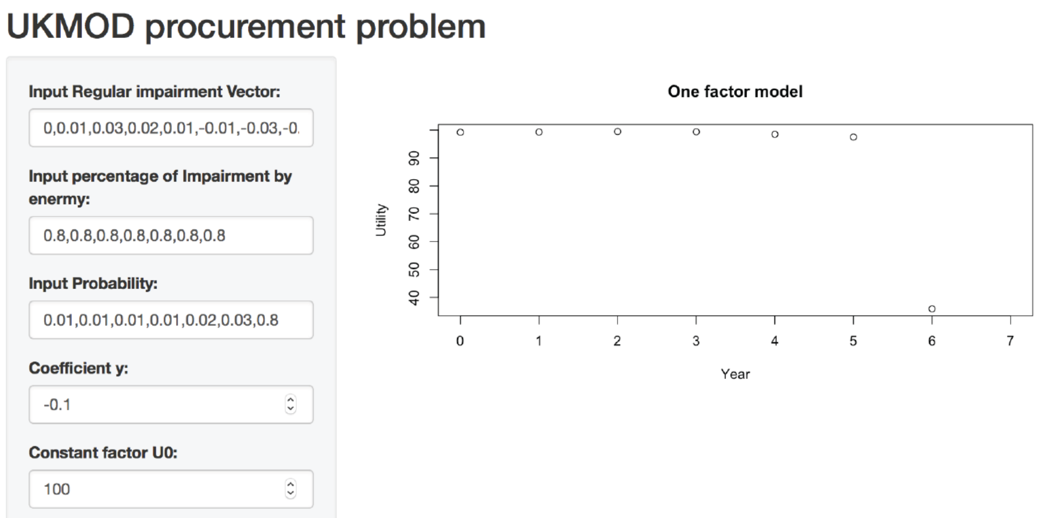


Figure 4: The impairment analysis interactive program example

Decision tree analysis

Introduction to decision tree analysis

On top of the one factor model, the next step is to add in other factors such as investment and contribution to the missions. Investment, in the context of military assets, is a long term research and development process and requires continuous capital all along the project. Contribution to mission creates high level of satisfaction to national defence and can be modelled as a numerical reward. To model the problem in decision tree analysis, we consider discrete time steps and model the investment as a negative utility value and the satisfaction from performing the missions successfully as a positive reward added to the utility value. The thesis assumes the rewards and investments can be estimated by UKMOD and all numerical values describe below are in terms of utility value.

In each time period, there is one decision to make and two chances followed. In the case of impairment, at the starting point of each time period, there are two decision branches: Continue investment and stop investment. At the end of this time period, there are two chance branches: Regular impairment or unexpected impairment. After several time periods, for each outcome there is one consequence at the very end.

We need the variables below in our decision tree analysis:

Variable name	Unit	Notes
Time periods	N/A	5 years in 1 time period
Probability of regular impairment	N/A	
Probability of unexpected impairment	N/A	=1- Probability of regular impairment
Investment amount of each period	dollars	
Reward from each period	dollars	
Regular impairment percentage	%	
Unexpected impairment percentage	%	

Table 1. Variables in Decision Tree Analysis

Implement decision trees

This decision tree model can be implemented with PrecisionTree, an Excel add-on to perform decision analysis. The initial decision at time 0, i.e. the beginning of the first time period would be to invest or not. The initial decision is represented by the green square in Figure 5 and the expected value "3283" is the higher of the values from two chance branches. The investment amount can be entered under the branch of the investment or link it to a cell with values that can be modified later.

After the initial decision, if we go with the investment branch, we can model two chance branches as regular impairment and unexpected impairment as shown in Figure 6. The probability of each chance is entered above the branch line and the return is below the branch line. Here the return incorporates both the reward brought by the aircraft and the amount of impairment during the period. Similarly, they can be lined to cells in this Excel file which can be modified or simulated later. The value under chance "3282" represents the expected wealth we could get under the modelling situation, i.e. when the probability of regular impairment is 99%.

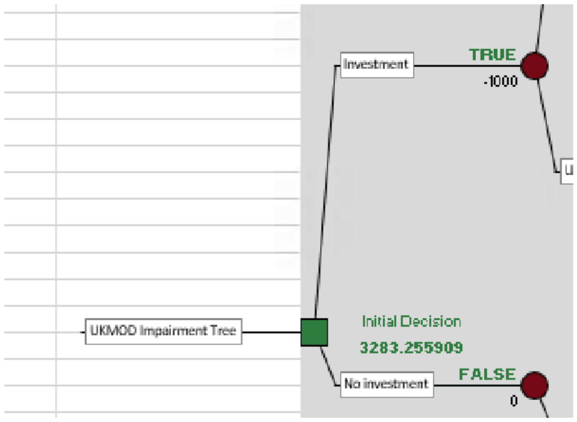


Figure 5. Initial Decision Node

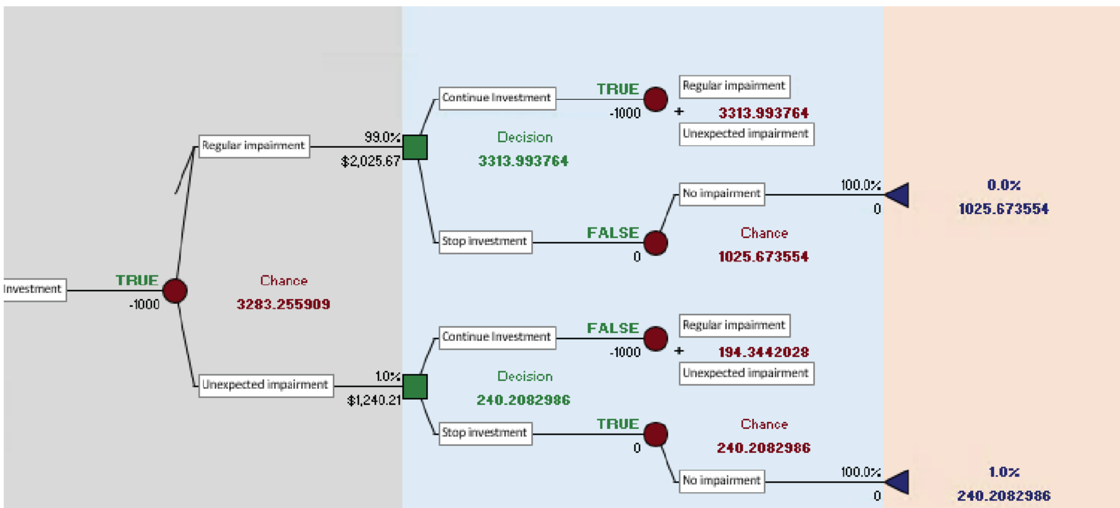


Figure 6. Investment Chance Node

After each chance, the second decision needs to be made: either continue investing or quit. Similar values and modelling follow until all five periods end. One the other end, if initially we choose not to invest, then like Figure 7, we have no investment amount and no impairment. Expected wealth equals to zero. The triangle at the very end represents a consequence.

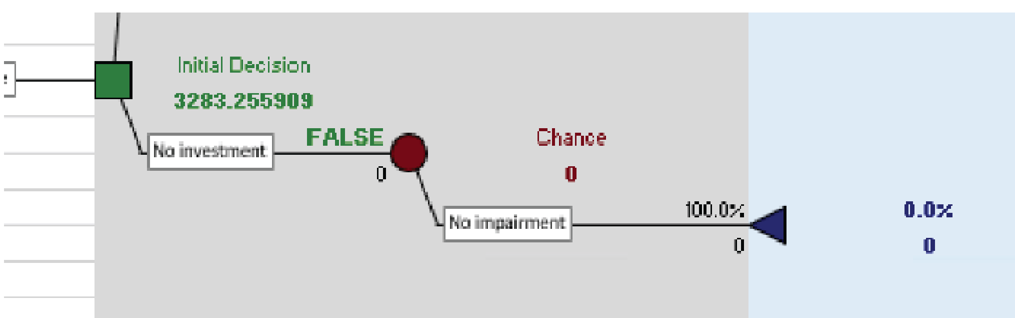


Figure 7. No investment Chance Node

All branches described above have an ending node which represents its consequence and indicates the expected value of this path. All chance nodes will have an expected value based on the following probability and all decision nodes will have one of the highest expected values from their following chance nodes.

Choose the optimal path

The optimal path is the one with highest expected value. PrecisionTree will automatically find the optimal path and indicate it by marking "True" on each of the decision. It allows utility function analysis and we can choose to display expected utility or certainty equivalent instead of expected values. (Shown in Figure 5) The optimal path will always be based on utility functions if you choose to use utility function in model settings.

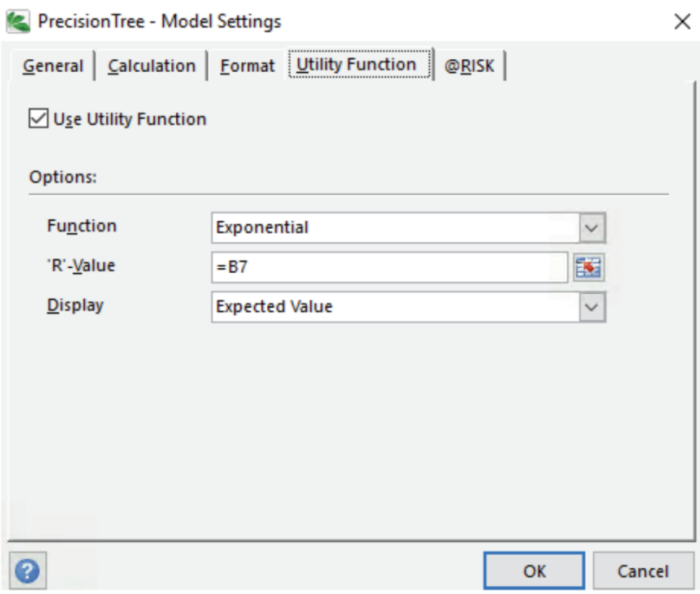


Figure 8. Utility Function Settings

Markov decision process model

Introduction to Markov decision process model

As mentioned in the literature review sections, Markov Decision Process is one of the decision analysis methods in reinforcement learning. Its goal is to select the proper actions at each state to maximize the expected total return considering current reward and all future following states.

Markov decision process formulation

Consider a time period of 20 years and we divide that into five time periods. Converting our decision tree model into a finite MDP, we have a series of states at each time period where

$$S_t \in \mathcal{S}$$

where we consider $\mathcal{S}$ as the market conditions, i.e. either regular impairment happens, or unexpected impairment happens, e.g. some kind of witchcraft is invented.

We have two possible actions in our model:

$$a \in \mathcal{A}(s)$$

where $\mathcal{A}(s)$ is the space of the actions we can perform if we can make decision at that state, i.e. either continue the investment or stop the investment.

In this finite MDP, the agent selects the action to invest or to stop invest at each time step, t = 0,1,2,3,4,5 by interacting with the environment. After the interaction, the agent receives a positive reward signal in numerical value from UKMOD by contributing to the missions in regular impairment market condition, or a negative reward signal if unexpected impairment happens. The agent then moves on to the next state S_{t+1}. Therefore, the agent follows a sequence that begins like:

$$S_0, A_0, R_1, S_1, A_1, R_2, S_2, A_2, R_3, \ldots$$

and $R_t > 0$ in regular impairment condition, $R_5 < 0$ in unexpected impairment condition.

Given the preceding state and an action, the probability of getting a reward r and jump to state s' $p(s', r|s, a)$ is deterministic and is inputted by the management as a variable. In this process, our objective is to maximize the total expected returns. At the current state S_t, the expected return is the sum of all following rewards after time period t.

$$G_t = R_{t+1} + R_{t+2} + R_{t+3} + \cdots + R_T$$

where T is the final time step in the process, which is 5. Considering the discount factor $\gamma \in [0,1]$, the expected discount return G_t becomes:

$$G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots + R_T = \sum_{k=0}^{\infty} \gamma^k R_{t+k+1}$$

The discount rate can also be an inputted variable. We can also write the expected return in a recursive form:

$$\begin{aligned} G_t &= R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \gamma^3 R_{t+4} + \cdots \\ &= R_{t+1} + \gamma (R_{t+2} + \gamma R_{t+3} + \gamma^2 R_{t+4} + \cdots) \\ &= R_{t+1} + \gamma G_{t+1} \end{aligned}$$

The process calculates the value function for each decision nodes by summing up the expected returns starting form a particular state s and follow by the policy π. The decision value of each node is the reward given by taking this particular action and the expected return from the next state. Therefore, the Bellman Equation demonstrate the property by defining the value function in terms of the successor states:

$$v_\pi(s) = R(s) + \gamma \sum_{S \in S'} p(S^{'}|S, a = \pi(s)) v_\pi(s')$$

This value equation provides a recursive relationship between the decision values and the successor values and defines a backward terminal value problem similar to calculating the value of financial options.

Markov decision process analysis

To compute the terminal values of each scenario, we need the input variables of investment, the positive reward from regular impairment scenarios and the negative reward for the unexpected impairment scenarios in terms of utility values. We take the data set below as an example, but the values can be modified by the decision makers.

Consider 5 time periods and each time period has four possible conditions:

- Continue investment and regular impairment happens
- Continue investment and unexpected impairment happens
- Stop investment and regular impairment happens
- Stop investment and unexpected impairment happens

Year	Investment	Regular Reward	Unexpected Reward
1	-1000	20000	-10000
2	-1000	20000	-15000
3	-1200	30000	-30000
4	-1500	40000	-60000
5	-1500	50000	-100000

Table 2. Example Data for Markov Decision Process Model

After 5 years, we have 1024 terminal values to be computed. The terminal values are just the summation of all previous returns, i.e. reward minus investment, in terms of utility, before whenever the first time we stop the investment.

Having the terminal values, we can compute the decision values for each decision nodes using the Bellman equation, which adds the current investment utility and the maximum value of the two following chance values. The same implementations are performed for all decision nodes in the five time periods.

Similarly, we can compute the chance values by adding the current reward to the expected values of the two following decision values using the Bellman equation.

Therefore, in our five-period model, for each decision, we have two decision values. Decision makers can choose the action with the maximum of the two values to perform because that is the optimal decision at that stage considering the given investment, reward and probability variables. The program returns values for each node in the process and helps decision makers to adjust investment decisions according to the constantly changing market variables.

Simulation of the optimal stopping year

Introduction

In this section, we address the same problem in another prospective. We would like to see the distribution of the optimal stopping time for 10000 simulations. To determine the optimal stopping year, we still take in the same inputted reward values and investment amount in in terms of utility as shown in the last chapter and choose the year with maximum total return. Instead of focusing on the decision to be made at each time step, we simulate the market condition and find the distribution of the optimal stopping year.

Find the optimal stopping year

For simulation purposes, we take in a distribution of estimated probabilities of unexpected impairment happens, and randomly generate a series of conditions for 20 years. We take in a positive reward for any regular impairment situation and a negative reward for unexpected situation as variables, calculating the total reward up to year t for each t = 1,2,3,4...20. Comparing the total rewards for each year, we found the year with the maximum reward value and indicate it as the optimal stopping year. We record the optimal stopping year in each simulation and repeat the simulation 10000 times. The following histogram is a typical representation of what we could get:

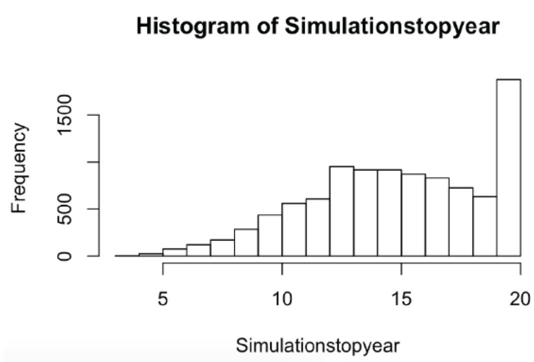


Figure 9. Optimal stopping year

Analysis of the results

In the histogram, we can see that a large frequency sits at year 20, which means as the probability of unexpected impairment is generally small, around 1500 simulations out of

10000 do not have unexpected impairment or have one at the very early stage, so it is optimal to keep investing until the project finishes. However, with the given data set, only 15% of the chances that the project could be kept until year 20, a large number of simulations are optimal to stop before year 20. As the initial probability of unexpected impairment is very low, and the impact of the impairment also gradually increases during the investment term, it is more likely that a serious impairment happens after year 10, and it would be optimal to stop before the impairment. The frequency reaches the maximum around year 12-13 and declines after year 13 because some simulations which were expected to have an unexpected impairment later already have one and it is already optimal to stop at the earlier stage.

The histogram does not provide direct suggestions for decision makers to make decisions but gives a general picture of optimal stopping year to guide the distribution of decisions.

Conclusion and reflection

In this thesis project, we addressed the long-term procurement strategy by first considering a one factor model, then modelling it in the decision tree model followed by the Markov decision process formulation. Finally, we simulated the process to find the distribution of optimal stopping time. By evaluating the decision values for each possible decision that we can make at each period of time, we tend to choose the decision with highest expected values to support our decision-making process.

As the thesis is a client-facing project, back-and-forth communications between C-MORE and UKMOD are essential to understand the problem definition more completely. Thus, the project requires a longer time to understand the client's requirements. To UKMOD, the project is on-going and needs to incorporate other specific and customized factors into the decision analysis process, such as the military technology, supply side constraints, etc.

The project can be generalized to any long-term asset investment and procurement considering different market conditions, the rewards and the investments. Considering any random market conditions similar to regular and unexpected impairment, the investment strategy can be modelled in decision tree models. Proper use of the decision values from Markov Decision Process Model and real options can be beneficial for decision-making. The simulation distributions can give a general picture of optimal stopping time for a risky investment.

Bibliography

[1] Tadeu, H. and Silva, J. (2014). Real Options Theory: Alternative Methodology Applicable to Investment Analyses in R & D Projects. Australian Journal of Basic and Applied Sciences, 8(6): 444-54.

[2] Howell, S., Stark, A., Newton, D., Paxson, D., Cavus, M., Pereira, J. and Patel, K. (2001). Real options. London: Financial Times Prentice Hall.

[3] Clemen, R. T., & Reilly, T. (2014). Making hard decisions with DecisionTools®. Mason, OH: South-Western, Cengage Learning.

[4] Developing Advanced Academic Degree Educational Profiles for Career Fields - Scientific Figure on ResearchGate. Available from: https://www.researchgate.net/figure/A-decision-analysis-process-flow-chart-Clemenand- Reilly-19966_fig1_235140604 [accessed 27 Jan, 2019]

[5] SUTTON, RICHARD S. BARTO, ANDREW G. (2018). REINFORCEMENT LEARNING: An introduction. Place of publication not identified: MIT Press.

DND: Propulsion diesel engine replacement models

Janet Lam

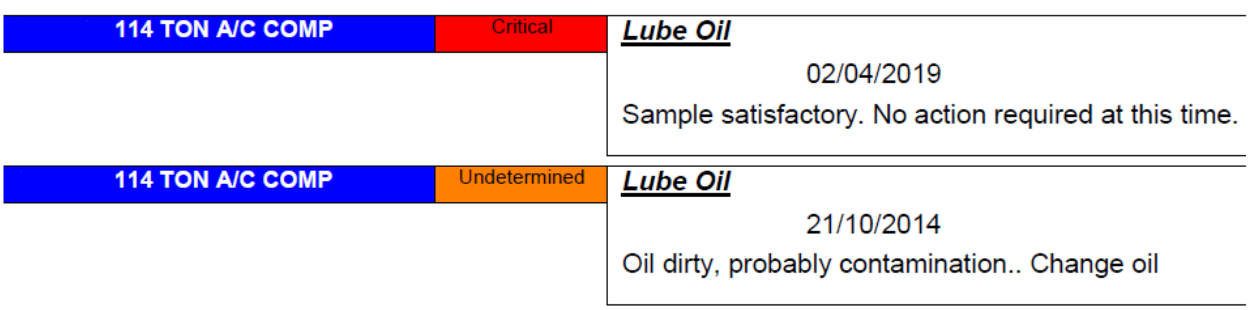
Background

In March 2019, the Canadian Department of National Defence became C-MORE's newest consortium member. As our initial project, we are looking at the replacement policies for propulsion diesel engines (PDE) in the Halifax Class ships.

Current procedures

The Halifax Class PDEs are currently under the Oil and Coolant Condition Analysis Program (OCCAP) wherein oil and coolant samples are taken every 30 days and analysed off-site. Once the samples are analysed, a report is produced and maintenance decisions are made by the fleet maintenance facility (FMF).

One of the challenges with the current procedures is what appears to be a disconnect between the OCCAP numerical results, and the analysed comments. In some cases, the OCCAP values will register as critical, and the technician comments will recommend no action, whereas in other cases the converse will occur.
Figure 1 demonstrates a situation in which there was a conflict between the OCCAP readings and the recommended action.



114 TON A/C COMP	Critical	Lube Oil 02/04/2019 Sample satisfactory. No action required at this time.
114 TON A/C COMP	Undetermined	Lube Oil 21/10/2014 Oil dirty, probably contamination.. Change oil

Figure 1 Sample excerpt of OCCAP results and recommendations

The objective of this project is to develop a model that expresses the relationship between OCCAP readings and the health of the PDEs in the Halifax Class fleet. By adding in cost information, we may also develop a decision policy that advises on performing preventive maintenance actions based on OCCAP's numerical results.

Preliminary analysis

Jamie Dreyer in collaboration with Nicolle Kilfoyle prepared three main data files. One file was the OCCAP readings, this is to be used as condition monitoring information. The second file was an excerpt from the Defense Resource Management Information System (DRMIS), and includes all of the work orders associated with the PDEs on the Halifax Class ships. This information will be used to obtain the event information associated with the engine uptime, downtime due to preventive maintenance and failure. The third file includes the engine running hours that is measured on a monthly basis.

The dataset ranges from 2013 to the present, and includes planned preventive maintenance actions and corrective maintenance actions. One of the challenges that face this project is determining failures of the engine from the work orders. The DRMIS data currently lists work order types as preventive or corrective (among others), however, a corrective maintenance action does not necessarily correlate to an engine failure. This is because the engine is a complex asset that may require repairs that are not critical to the function of the engine. For example, replacing a burned out lightbulb is considered corrective maintenance, but clearly should not be an indicator of engine failure.

One of the approaches that has been considered is using the cost of the replacement components as a proxy for whether the engine has failed or not. This is not unreasonable, as a component that is critical to engine functionality is more likely to be expensive compared to an auxiliary component. However, this hypothesis requires deeper consideration, as it is possible that an expensive component is not necessarily a critical component.

Our primary task in the immediate term is to distinguish failures, minor preventive maintenance and major preventive actions from the work orders so that an age-based analysis can be performed. The incorporation of OCCAP data for condition-based analysis will follow.

TTC: Track re-inspection schedule optimization

Janet Lam, Dragan Banjevic

Background

One of TTC's rail maintenance strategies includes the detection and resolution of track defects. Each defect is given a priority level, and a priority level is assigned a time limit in which the defect must be resolved, or repaired. In the event that maintenance resources are not available to repair defects in time, the non-destructive testing (NDT) team will re-inspect the defect, which restarts the clock on the time limit. Based on experience, TTC feels that a significant portion of the NDT team's resources are being spent on re-inspections, and would like to review the time limits associated with each priority level to be based on evidence.

Preliminary analysis

There are seven different priority levels, with differing time limits. The priorities are designated by colour, with red, yellow and purple being considered "high priority" defects, and blue, brown and gray not high priority defects.

Table 1 lists the priorities in increasing order and their time limits for resolution. Note that grey defects are not required to be updated; they are simply recorded for information in the event that they are upgraded to higher priority.

Table 1 Defect priorities and their associated time limits

Priority	Time limit (days)
Grey	None
Brown	365
Blue	45
Purple	21
Yellow	10
Red	1

In practice, red defects are analogous to track failures that require immediate resolution. Similarly, yellow defects are prioritized by the maintenance team, and thus are repaired relatively quickly. Thus, they may be excluded from the re-inspection project.

Each defect is given a unique identifier when it is first detected. Each follow-up inspection results in an entry labelled “updated”, and a final entry labelled “completed,” when the defect is repaired. Using data from 2015-2018, we counted the number of entries for each unique defect. Only considering the defects that had been completed, the summary is given in Table 2.

Table 2 Summary of defect updates

Initial defect colour	Number of unique defects	Mean # of entries per defect	Mean days to complete record	Avg inspection interval
Grey	131	2.8	319.5	112.5
Brown	6	2	110.8	55.4
Blue	17	4.9	167.7	33.9
Purple	330	9.1	139.6	15.3
Yellow	39	2.3	7.6	3.3
Red	4	2.8	2.25	0.8

The information summarized in Table 2 is limited to completed defects, so it does not include defects that are were still undergoing more re-inspections, and grey defects that did not have any follow-ups. The second column shows the average number of entries recorded for a unique defect. This is a proxy for the number of re-inspections, with the exception that these only include the completed defects. This means that the smallest number of entries will be two: one to record the new defect, and one to close the defect.

There is evidence that blue and purple defects are undergoing more re-inspections than strictly necessary, as blue defects have 5 records, and purple defects have 9 records on average. The defects are resolved in 168 and 140 days for blue and purple defects, respectively. The initial hypothesis is that the re-inspection time limit may be somewhat conservative.

Another point of interest is the average inspection interval. These values represent how many days lapse between each re-inspection, or new record. When these values are compared to the time limits listed in Table 1, it can be seen that the actual re-inspections are done much more frequently than the time limit allows. Particularly in the case of blue and purple defects with many re-inspections, this shortening of the time limit will result in additional re-inspection efforts that were not required to meet the current guidelines.

Objectives

This project has several areas that may be improved, or better understood. Throughout the project, we aim to maintain the current level of reliability of the tracks. That is, less frequent re-inspections should not result in a reduced reliability of tracks.

One of the objectives of this project is to re-evaluate the time limits associated with each priority level. From the preliminary analysis it can be seen that blue and purple defects are consuming quite a few inspection resources. Ideally, we will develop an evidence-based approach to generating the re-inspection time limits to reduce the number of inspections required.

Another objective is to determine the transition rate from one priority level to another. In the preliminary analysis, the defects were categorized by the initial priority level. However, over the duration of the defect, it can progress to a higher priority. In particular, once it reaches yellow or red priority levels, the resolution process may be quick. The transition to other priority levels was not captured in the preliminary analysis. Transition rates will definitely be necessary in order to determine the appropriate re-inspection guidelines for each priority level.

Another objective is to include some sensitivity analysis on the recommended re-inspection times. The NDT team consistently performs re-inspections before their required time limit. This may be due to operational issues that make it more convenient to schedule re-inspections at a certain time in their overall schedule. By performing sensitivity analysis, we will be able to provide a range of recommended time limits for re-inspection.

As we perform this analysis another domain of interest is to further categorize defects into the defect modes. Some defect modes such as bolt hole cracks are more likely to progress into full failures that need immediate repair, whereas others such as corrosion do not progress as quickly. The further categorization of defect modes is a stretch goal for this project.

Kinross Gold: Caterpillar haul truck engines decision policy

Dragan Banjevic, Janet Lam

Background

In April 2018, C-MORE began a new project with Kinross on the optimal replacement frequency of a fleet of haul truck engines. Kinross supplied event history and a record of inspections gathered at oil changes for its fleet of haul trucks. In the first stage (April – June 2018) the data preparation and analysis was conducted, including cleanup, identification of anomalies and potential errors. In the second stage, initial model selection was started, and some preliminary results were presented at the June C-MORE meeting. In July – December 2018, more data cleaning was performed, mostly to look for "outliers" (spurious measurements, sampling errors), as well as detailed model selection. The results were presented at the December 2019 meeting. Finally, a decision policy for truck engines was completed in January – February 2019. The results were presented to Kinross Gold in February. At the end of April, C-MORE gave a two half days course in EXAKT software to Kinross personnel (Emilio Sarno, Brian Wright, Alberto Van Oordt). An overview of the decision model and its analysis is given in this report.

A short summary of the data

- There are two different engine models: 793C and 793D. After discussion with Kinross engineers, it was concluded that they can be considered identical for the analysis.
- There are 15 units – 9 are model 793C and 6 are 793D.
- There are 34 histories (chronological data since installation/repair until failure/suspension) of which 8 ended in failure, 13 ended in suspension (due to high hours) and 13 were still in service.
- Inspections are made regularly throughout the lives of all units, 1323 inspections in total, every 300 hours (on average).
- The data range from July 2012 to February 2018.

- There are 37 columns of measurements, mostly metal and oil variables. Several of them (mostly oil variables) are with incomplete (missing) records. They were not included in the analysis.

What we did before December

- Initial data analysis – grouping events into failures, replacements, etc., data anomalies, errors, etc., checking variables, Integrate oil change and filter changes into events.
- Checking for outliers.
- Explored potentially critical measurements.
- Evaluation information appeared not really informative of failures.
- We tried different combinations of significant variables.
- Looking for costs – at failure and replacement.

Steps after our meeting in November: decision making modelling

- Estimating transition probabilities for variables in the model.
- Finding replacement and failure costs, including downtimes.
- Calculating decision replacement model.
- Evaluating decision model, how good and efficient it is.

Decision policy in short

- The decision policy is defined by selection of a threshold on hazard function (critical hazard level).
- At every inspection instant the hazard function is calculated and if its value is below the threshold, it is recommended to continue operation. If it is above the threshold, it is recommended to stop operation and perform repair/replacement.
- The hazard level is selected to make optimal balance between expected costs of preventive repairs and failure repairs.

Hazard and decisions

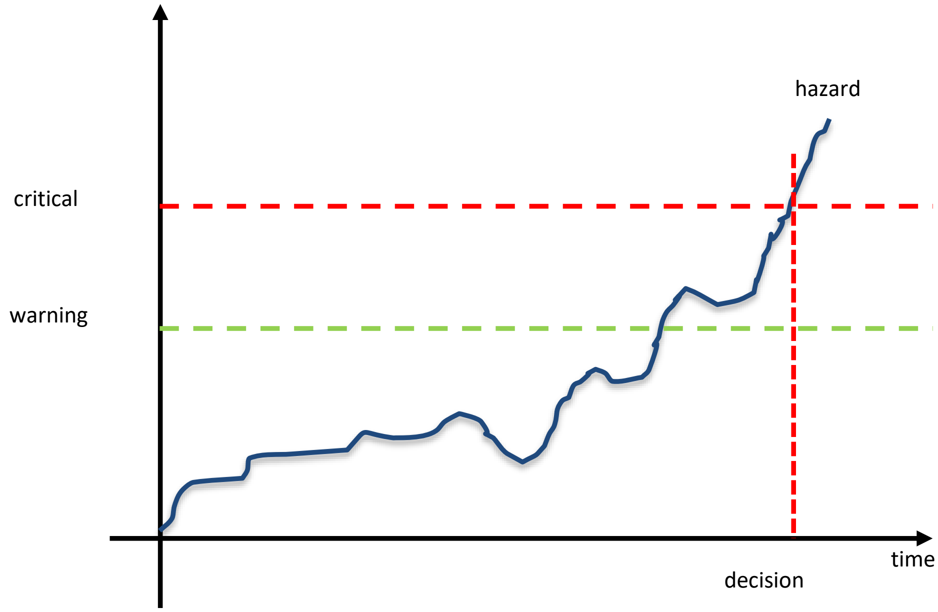


Figure 2 Decision policy and hazard

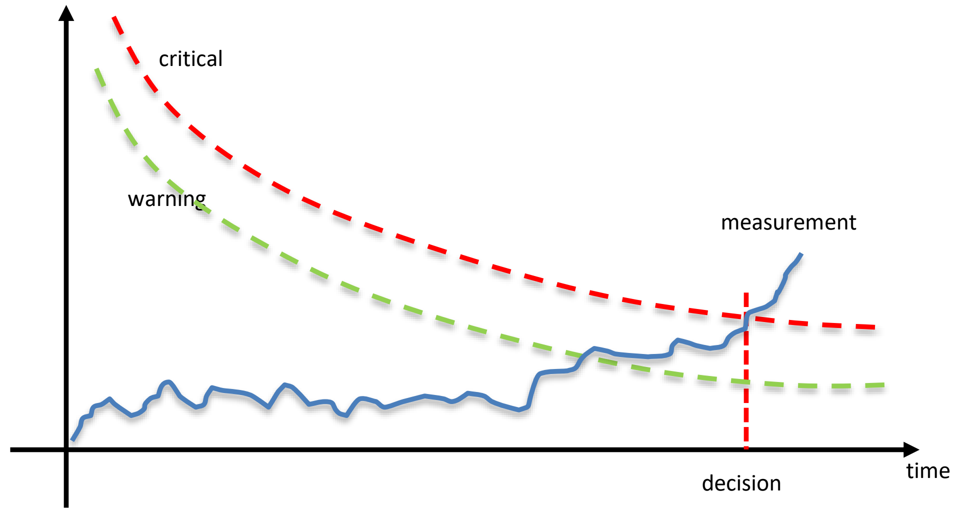


Figure 3 Transformed decision graph

Cost parameters

- Two cost parameters are required for decisions: cost of one preventive repair/replacement, and cost of one failure replacement (at least average costs).
- Brian Wright from Kinross estimated the preventive replacement cost of an engine to $380,000, and the failure replacement cost to $925,000 – including downtime. They can differ, depending on a site.
- Cost ratio of 925/380 = 2.43 is relatively moderate.

Two-parameter Weibull model

As an initial lifetime model for engines, without measurements, we estimated Weibull distribution from engines lifetimes.

Table 3 Weibull model parameters

Summary of Estimated Parameters (based on ML method)

Parameter	Scale	Shape	Mean Life	Med. Life	Char. Life	Std. Dev.
Estimate	23855	2.313	21134.9	20359.3	23855	9695.9

- Engines show ageing, shape parameter = 2.313. Mean life of an engine (if run until failure) is 21,135 hours.
- The Kinross policy is to replace an engine due to high hours at around 15,000-16,000 operating hours.
- From the Weibull model, a chance that an engine will survive to 16,000 hours replacement time is around 67%, not far from observed in the data. So, the chance to fail before replacement is 33%.

Time-based decision policy

Time-based replacement policy option is to replace an engine at planned preventive replacement time. We obtained the following summary of the cost analysis.

Table 4 Summary of time-based policy cost analysis

Summary of Cost Analysis

	Cost [$/h]	Preventive Repl. Cost [$/h]	Failure Repl. Cost [$/h]	Prev. Repl. [%]	Failure Repl. [%]	Expected Time Between Replacements
Optimal Policy	38.6349	13.3413 (34.5%)	25.2936 (65.5%)	56.2	43.8	16012
Replacement Only At Failure	43.7664	0 (0.0 %)	43.7664 (100 %)	0.0	100.0	21134.9
Saving	5.13145 (11.7 %)	-13.3413	18.4728	-56.2	56.2	-5122.95

- Optimal preventive replacement time - 18,793 operating hours. It would result in about 56% preventive replacements, and 44% failure replacements.
- Kinross policy is to replace due to high hours – at around 16,000 hours. Optimal replacement time is somewhat longer than for Kinross policy, which would result in more failures (44% to 33%), but in more utilized time and lower cost per time unit. With a cost ratio higher than now (of 2.43), the replacement time would be shorter.
- Expected saving compared to current cost ≈ 12.5% ($38.6/hour vs $44.2/hour).

Condition-based replacement policy

- Several hazard models with oil measurements had been tried for a reasonable decision policy.
- The models show, in theory, savings between 15-25% in comparison with the current policy.
- When different models were applied retroactively, saving were smaller, between 1-7%.

We selected model with metal variables Fe, Si, and Mo.

- When applied retroactively, it prevented 3 out of 8 failures. More details will follow.
- Expected saving – 15%
- More data (histories) are needed for validation

Optimal condition-based policy has the following form. It is presented as "composite covariate" Z on y-axes, versus working age on x-axis. Composite covariate is a linear combination (estimated from the data) of selected variables Fe, Si, and Mo.

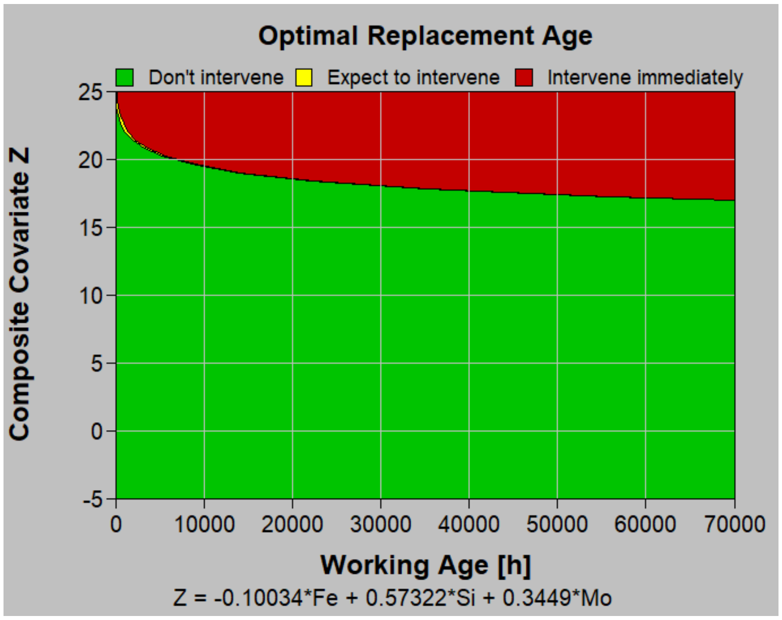


Figure 4 Decision policy "red-yellow-green" chart

Policy in work – an example of engine id 812 history

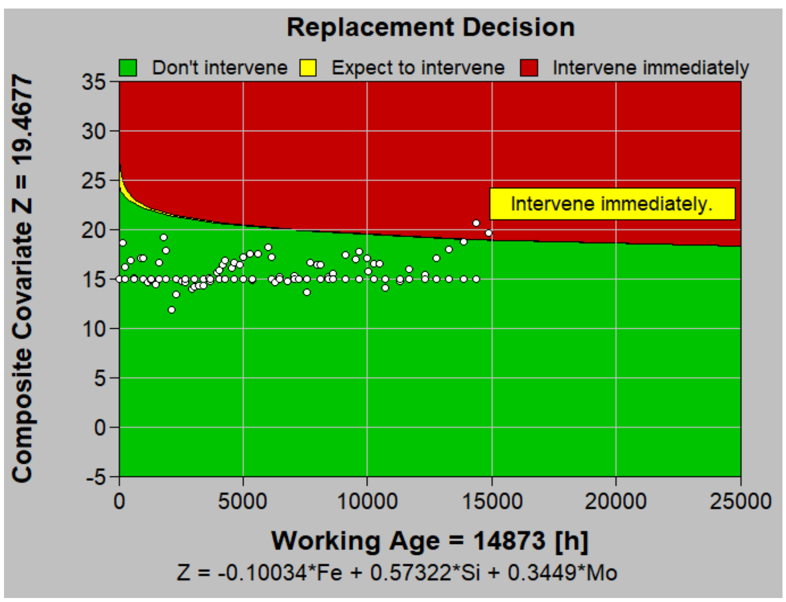


Figure 5 Decision policy applied to truck 812

Some comments on the history:

- Decision recommendation: Intervene immediately - points are in red zone
- Last inspection was at 14,873 hours
- Engine failed after 247 hours (in one month calendar time), at 14,873 + 247.

Caterpillar S.O.S service recommendations

Caterpillar provides decision recommendations in form of warning limits. The limits don't depend on engine's age.

Table 5 Caterpillar S.O.S service wear table

Element	No Action Required			Monitor			Action Required	
Copper	0	to	42	43	to	61	Over	61
Iron	0	to	26	27	to	33	Over	33
Chromium	0	to	1	1	to	1	Over	1
Aluminum	0	to	4	5	to	5	Over	5
Lead	0	to	3	4	to	5	Over	5
Silicon	0	to	7	8	to	9	Over	9
Tin	0	to	5	6	to	8	Over	8

It is not clear when an action is required. Is it when any of the elements is above the limit? There is no clear indication in the data whether any of those recommendations were applied with engines measurements. Mo (molybdenum) is not included in the S.O.S. list of critical metals. It was found significant in our analysis? Where does it come from?

Decision policy analysis

Decision policy was applied retroactively to truck histories. Out of 8 failures, 3 would have been prevented, with some reduction in operating hours, but saving in cost per hour, as mentioned above. In other 3 cases, recommendations to intervene would likely come close to failures.

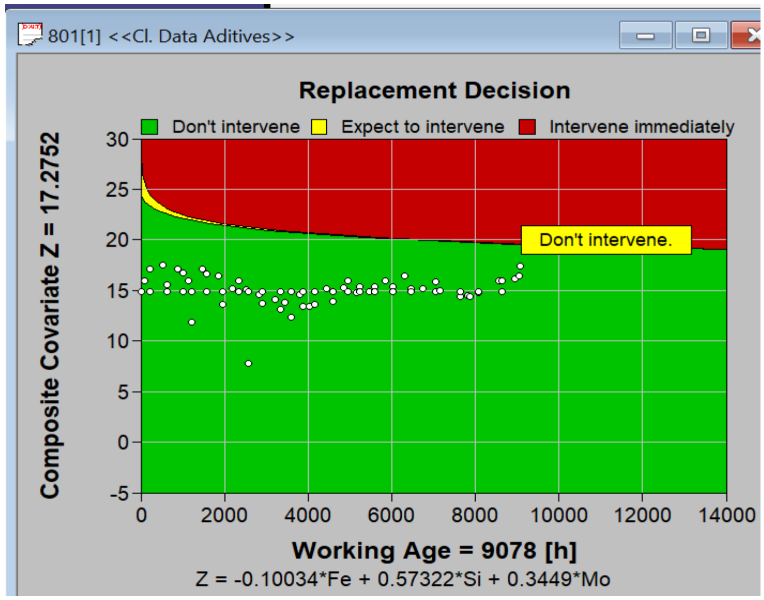


Figure 6 Engine failure before intervention, but approaching

A summary of all histories, failed and unfailed is presented in the following table.

Table 6 Summary of EXAKT policy applied to truck histories

Data	EXAKT	Count
Failed (and replaced, 8)	Saved	3
	Not saved, but close	3
	Not saved	2
Retired (2)	Continue	2
Replaced (11)	Intervene	2
	Continue	9
In operation (13)	Intervene	1
	Close to intervene	2
	Continue	11

The data show that the policy is, more or less, compatible with the data, e.g., for 26 unfailed histories, the policy would recommend for 23 to continue, and for 3 to intervene.

Conclusions

- Analysis show promising results for improvement of the replacement policy.
- The saving predicted is 15-20% of the current policy, but this requires more validation by more engine histories.
- Some failures may not be related to oil variables (to wear, or oil), so they cannot be accurately predicted by the policy, we don't have data to check it.
- The replacement only due to age (high hours), could be also improved, by somewhat extending the upper limit from 16,000 hours to more than 18,000 hours.

Next steps

- Preliminary implementation, learning and testing (we did it in form of short EXAKT course in April).
- EXAKT data base creation, EXAKT installation.
- Real data entering procedure.
- Decision policy application.
- Feed-back and updates.
- Finally – implementation on Kinross CMMS.

Teck: project exploration
Janet Lam

Background

In May 2019 Chi-Guhn and Janet visited the Teck Sparwood office in British Columbia to have a discussion on some potential areas of collaboration. Kevin Hatch and Alex Creagh were in attendance, with David Williams on a remote call.

Three main projects were identified. This report will discuss these projects in turn.

Physical availability prediction with KPIs

Teck is interested in improving its prediction of availability of their mobile assets. The availability of a truck is determined by many factors, including reliability, spare parts, labour, and operations. Rather than determining and using the direct inputs that may affect availability, we are interested in using key performance indicators (KPIs) that are currently measured and reported regularly.

Alex has already developed some preliminary models that express the relationship between some KPIs and the historical availability of their trucks. However, in some cases similar equipment with similar KPIs will have different availabilities. Thus, a model that better relates the inputs to outputs is of interest.

Economic decision tool

One of Teck's maintenance policies includes a late-life audit in which some of their equipment is assessed at some pre-determined life point in the later part of the equipment's life. Based on its previous performance and results of the assessment, a decision is made for the equipment on whether it is to be retired, or to be used for an additional number of years.

In Teck's experience, the performance of the equipment before and after the audit is not consistent, and equipment that was very reliable leading up to the audit may suddenly experience an increased failure rate following the audit.

We would like to further investigate the relationship between equipment health, age, the time of the audit, and how better evidence-based decisions may be made.

Holistic analysis of wheel motor maintenance

In the maintenance procedures of wheel motors, Teck currently uses historical records and makes maintenance decisions based on information from the whole life of a wheel motor. This method is known to be a good way to make decisions. However, the current method is a labour-intensive procedure.

Although using a basic proportional hazards model would be less work, the simplest application of the model only uses the currently health information, and doesn't incorporate the historical health records.

A method that incorporates or visualizes the whole life of the asset with minimal additional work is another project of interest for Teck.

Moving forward

There is also interest in applying one of the new developments in a hybrid remaining useful life model that incorporates a Bayesian neural network to model deterioration of equipment.

The plan is to select a small project that can be completed within this calendar year.

Development of digital twin for predictive maintenance

Mohamad Danish Anis, C-MORE PhD student

Executive summary

As the star concept behind the Industry 4.0 wave, a digital twin is a virtual simulation to mirror its physical counterpart's performance and serve the product lifecycle in a virtual space. Evidently, a digital twin can identify potential issues with its corresponding real twin. Thus, it is best suited for enabling a physics-based and data-driven model fusion to estimate component remaining useful life (RUL). A constraint with modern machinery is that sensor data is collected sporadically as per the requirements and policies of the data-provider. Traditional RUL prediction techniques assuming either a linear or exponential degradation curve may not be useful for such a scenario.

In our research so far, we are working towards extending an LSTM-RNN technique to generate a RUL prediction within a digital twin framework as a means of synchronization with changing operational states. In theory, an LSTM encoder-decoder (LSTM-ED) is used to train a neural network and reconstruct the sensor data input time series corresponding to a healthy state. The resulting reconstruction error can be used to estimate health index (HI) training and testing sets. Using a time lag to record similarity between the HI curves, a weighted average of the final RUL estimation can thus be obtained.

Currently, we are evaluating this approach first on a publicly available engine degradation data set. We later plan to extend this approach to tri-axial vibration data collected intermittently from a wind turbine. We expect the results to indicate a high RUL estimation accuracy with greater error reduction rate. Furthermore, the applicability of a probabilistic model shown within a digital twin framework can be an important contribution to the existing literature which is looking for solutions to real-time update as new sensor information is received.

Introduction

State of the art advancements and evolution in digital technologies are constantly challenging the traditional practices in many industries worldwide. The onset of what is seen by many as the fourth industrial revolution (Industry 4.0) finds its basis in a new generation of virtual reality and big-data driven models [1]. Studies such as [2]–[4] recognize Industry 4.0 as a paradigm shift in investment strategies of companies towards smart technology from an Industrial Internet of Things (IIoT) perspective. Digital twin technology has been recognized as a core component of Industry 4.0 with its ability to virtually represent the elements and dynamics of how an IIoT device operates throughout its life. The idea of a digital twin took shape with the vision of providing a comprehensive functional mirror of a component, product or system, aiming to replicate the physical asset's performance in virtual space with real-time synchronization [5].

As the number of wind farms grows to cope with the increasing energy demands, the reliability of wind turbines in addition to lowering their maintenance costs has been of an area of intensive research. A real world IIoT set up, such as a wind turbine drive train, has paved way for simultaneous monitoring of several sensors at their unique sampling rates. This has realized the need for artificial intelligence tools for robust data processing. However, the large size of input data requires real time monitoring and synchronization for online analysis. Even though sub-systems of such complex machinery, be it mechanical, electrical or hydraulic, are all independently designed and troubleshot, the system overall performance is only reflected during real running conditions. Owing to the system's requirement of continued availability and reduced manual dependency, precise self-prediction and self-assessment are a digital twin's main focus. As a part of its prognostic application, a digital twin can proactively identify propagating anomalies with its physical counterpart. Thus, RUL prediction of specific components can be facilitated by fusing pre-processed real time sensory data flow into a digital twin model.

Prognostics and Health Management (PHM) mainly aims at predicting the RUL of a system or component to plan optimal maintenance. Prognostic approaches can be broadly classified into data-driven, physics-based and hybrid. Given their simplicity, data-driven approaches are mostly preferred to make RUL predictions by directly analyzing equipment behaviour from condition-monitoring data [6], [7]. Machine learning and statistical learning are the two most widely used data-driven approaches. Increasingly complex sequential data today requires machine learning based deep learning techniques for accurate non-linear processing. DTs require offline computing resources to utilize deep learning models. They are trained using a combination of machine learning algorithms and data analytic techniques to create a model of a specific target, for instance a flight-critical component or a high speed rotor shaft that can derive an actionable outcome from the model. Upon availability of historical data, the training of such comprehensive data-driven models for RUL prediction has been heavily researched. Scholars have used supervised and unsupervised learning techniques like cluster analysis, back propagation neural networks, Support Vector Machine (SVM) algorithm etc. to diagnose and predict incipient and propagating faults. A discussion on one such deep learning technique, LSTM-RNN, follows a detailed literature survey in this report.

The following sections provide an update on the study of the underlying analytics and machine learning techniques behind developing a digital twin for machinery prognostics in the form of an optimal RUL estimation application.

Supporting literature review

Digital twins: concept and applications

Given the novelty of the DT concept, there exists little but a steadily growing literature on the different application specific definitions of a DT as understood by researchers from around the world. According to [8], [9], the DT concept is expected to have all data and information of a physical system replicated in software for simulation and analysis. This is consistent with some of the initial studies emphasizing on 3-D simulation as the core functionality of a DT in industries like manufacturing and design [10], [11]. However, in practice, this conceptual model can only embody certain essential elements like physical and virtual spaces and the linkage between the two [12], [13]. A heavier emphasis needs to be placed on the management of IIoT in a virtual, smart and manipulable space with better representation and communication capabilities [14], [15]. The first formal definition of a DT was proposed by NASA in its integrated technology road map, calling it a virtual equivalent of the physical system replicated for modelling, simulation and analysis purposes [16]. The broader aim was to have an integrated model to simulate, monitor, calculate, regulate and control system processes in a highly quasi-real fashion.

Based on the studies conducted over the last six years, Table 1 below summarizes the industrial application specific definitions of DT as they appear in the literature. Most of the existing and some very recent studies on DT, like [17] deal with a surface level introduction of the topic rather than going in depth of the underlying analytics behind the idea. This is understandable since most DTs' applications lie in military or other highly confidential and sensitive businesses. Furthermore, building a physics-based (1st principle or knowledge-based) DT can obviously be a complex task given the in-depth knowledge required to address any uncertainty due to the ignorance of the deep learning model parameters. The current gap in the literature can be filled by addressing uncertainty quantification and high fidelity real-time synchronization between the digital and physical counterparts.

Deep learning for prognostics

Deep learning (DL) is one of the most active sub-fields of machine learning research. Inspired by the biological brain architecture, DL refers to the supervised/unsupervised learning technique that can learn hierarchical patterns by stacking multiple layers of information processing modules in deep structures. Significant advantages that DL offers in predictive models are accuracy and increased processing power, thus, reducing the computational burden [26], [27]. Conversely, however, the deep architecture is also known to introduce diverse hyper-parameters during learning which can be challenging to optimize in the training process. Supervised DL techniques require large labeled training data sets in their training procedure, thus, the prediction accuracy is strongly reliant on the constructed run-to-failure labels. On the other hand, unsupervised techniques utilize a pre-training stage for degradation feature extraction from unlabelled data.

The literature today has come a long way from conventional data-driven DL techniques involving feature manual design, selection, extraction, model training etc. [28] in their extensive review of DL techniques on machine health monitoring study how deep neural networks can extract hierarchical representations from input data and perform a layer-wise non-linear transformation into outputs. This allows learning complex concepts from simpler information fed as inputs and extensive human labor is reduced. Furthermore, the model parameters including pattern classification/ regression modules don't have to be trained individually. Thus, studies like [29] have used DL techniques to train and re-train their models for both diagnostic and prognostic applications.

Table 7 Application specific literature review on digital twins

Ref. No.	Year	Application Description	Comments
[18]	2012	DT used to predict aircraft structural reliability over the course of the mission	The developed DT was a life-long model with configurations for sub-models in the entire inventory
[19]	2012	Used DT for confidence prediction and CBM decision making in an aircraft	The developed numerical simulation model was used for RUL prediction
[20]	2013	A DT was used to model the life of vehicle materials and structures	The developed high-fidelity model was able to predict cracks with on-board vehicle health management system
[11]	2015	DT was able to model the relationship between product manufacturing process and the operating environment through simulation	The current state and behaviour of the production system life-cycle was realistically modelled
[21]	2015	The developed DT was able to incorporate wind turbine blade structural fatigue damage in real-time	The high fidelity model focused on SHM to develop a FEM framework for fatigue damage prediction
[14]	2016	The study put to use DT for engineering and mechanical design integrity	The model makes use of CAD-based simulations along with Simulink
[22]	2016	The DT was put to use for virtual commissioning for robot control	Used a Virtual Environment and Robotic Simulation (VERS) for control optimization
[10]	2017	Integrated DT with cyber-physical system (CPS) and virtual factory for smart manufacturing	The aim of the study was to create an architecture reference model for a cloud based CPS
[23]	2017	Heuristic optimization algorithm so as to identify key parameters of the model	The fidelity of the model was increased by regularly comparing results against real data
[24]	2017	In a DT for additive manufacturing, analytical sub-models were used to predict spatial and temporal variations	The predicted variations of metallurgical parameters proved accurate with validation of experimental data
[25]	2018	A generic DT driven PHM is proposed for interaction mechanism and fused data in a Wind Turbine	The study efficiently explained the building blocks of DT in complex machinery and constructed its framework by taking into account cost, fidelity and large data size

Some of the most rapidly developing DL variants prominently used in RUL predictions include neural networks. Convolutional Neural Network (CNN), Deep Belief Network (DBN), Recurrent Neural Network (RNN) and LSTM are known to have outperformed the traditional prognostic algorithms for RUL prediction [28], [30], [31]. LSTM in particular has been widely explored for prognostics and has proven to be particularly efficient in supervised and semi-supervised studies. The following section provides an in-depth overview of the LSTM architecture and the supporting literature.

Long-short term memory (LSTM) network

A variant of the traditional RNN, the LSTM model was initially proposed to address the vanishing/exploding gradient problem in RNNs [32] and further improved to avoid any long-term dependency issues [33]. LSTM models have since achieved great success in sequence learning tasks such as speech recognition and machine translation where hidden patterns are to be discovered. The architecture of LSTM, as shown in Figure 1.1 below, uses a well-designed memory cell with four different gates to replace the activation function of hidden state in RNN. The cell state at the top, from C_{t-1} to C_t, converts information from end to end and stores in cell memory. Within the memory cell, the model has three gates composed out of a sigmoid neural net layer that provide the ability to add or remove information to the cell state. The output value is kept from 0 to 1 for each number in the cell state C_{t-1}, where 1 represents any information is allowed to pass and 0 presents nothing is allowed to pass. The forget gate decides which information has to be discarded from the cell state. Only conditional information is stored in the cell state and this is decided by the *sigmoid* layer (input gate layer) and the *tanh* layer (new candidate values). The old cell state, C_{t-1}, is thus updated into new cell state C_t by forget gate, input gate and new memory. Lastly, the output gate decides which information will be converted from the cell state into the current hidden layer data. The element wise multiplication is fostered by weights displayed in the mathematical representations.

As the flow of information is regulated in and out of the cell, the memory cell can preserve its state for longer duration (i.e. learning long term dependencies) and influence future predictions. Several impactful studies have been conducted in the field of RUL prognostics that confirm the same. [34] were able to reduce training time and improve performance using an GRU-LSTM(Gated RecurrentUnit). [35] later used LSTMto extract important degradation features from several operating conditions in a pre-training procedure and showed enhanced performance compared to RNN and GRU-LSTM. [36] proposed an LSTM-based deep learning method to capture degradation process and predict RUL. [27] tracked the raw historical machine degradation parameters and translated them in to health index (HI) for RUL estimation using a bi-directional LSTM model. Some of the more recent works on LSTM-based RUL estimation have been looking to build a HI instead of relying on degradation trends for RUL prediction. This is being done by using LSTM encoder-decoder.

LSTM encoder-decoder (ED)

LSTM-EDs have successfully been used for both anomaly detection [37], [38] and sequence-to-sequence learning tasks like natural language generation and reconstruction, parsing and image captioning [34], [39], [40]. Autoencoders or ED are artificial neural networks with symmetrical structures and at least one hidden layer that consists of less neurons than input and output layers. They are used to re-construct their inputs and learn a lower dimensional representation of input data in hidden layer. As seen in Figure 1, while the encoder maps a multivariate input sequence to a fixed dimensional vector representation, the decoder produces target sequences using this vector representation. Since a multi-variate time series has variables observed over multiple time stamps, the LSTM model here would require certain pre-processing steps to label negative data, reconstruct new samples and record prediction error. These steps shall be discussed in further detail in the Approach Overview section.

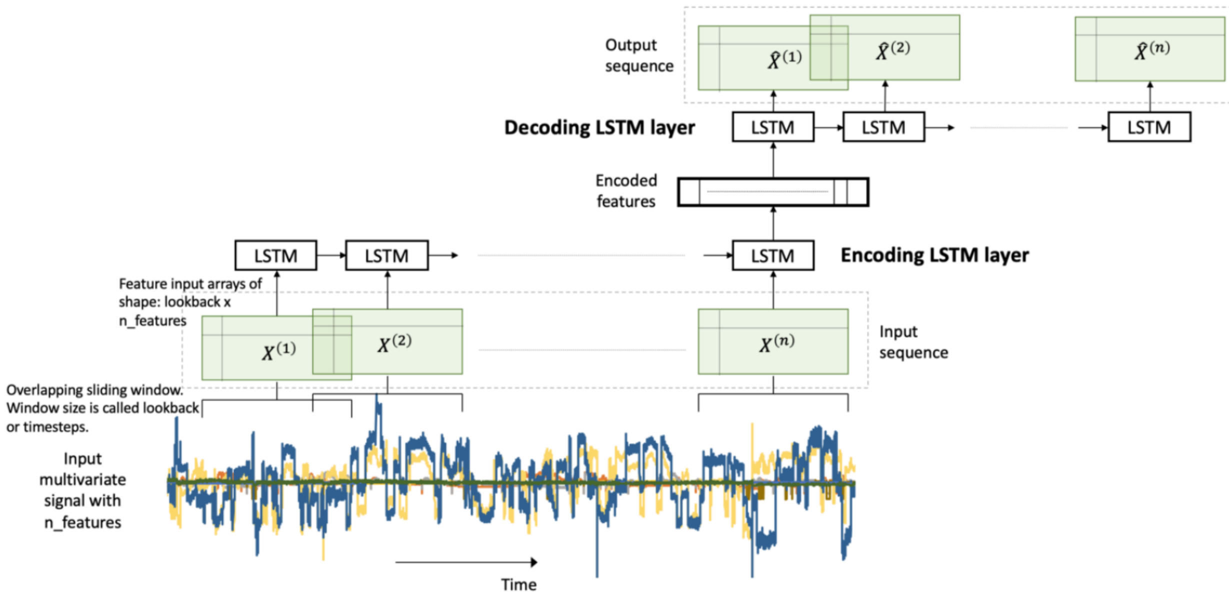


Figure 7 LSTM encoder-decoder

The main rationale of [38] and [41] behind using LSTM-ED was to reduce the dependency of the traditional prognostic approaches upon the assumption that the degradation trend must follow a specific curve. Some of the most famous case studies in data-driven prognostics like [42]–[46] assumed either an exponential or linear degradation trend to build HI prediction models. While this approach has its benefits when the data doesn't have explicit health parameters or labels, it doesn't take into account the time taken to reach the same level of degradation and the effect of noisy sensor data. Besides, standard DL-based anomaly detection approaches do not take all variables or failure modes from multiple sensors into account, leading to unreliable predictions.

The use of autoencoders captures system degradation without relying on domain knowledge or the assumption regarding degradation curve. With additional memory cells, LSTM-ED is able to reconstruct, and later store as embeddings, the multivariate 'normal-behaviour' input time series to detect anomalies, compute reconstruction error and an unsupervised HI for accurate RUL estimation. Given the novelty of this approach, not a lot of supporting literature exists that makes use of autoencoders to make accurate RUL predictions.

Approach overview

LSTM-ED is trained to reconstruct the input time series in the form of embeddings corresponding to the healthy state of the system. The rationale behind using autoencoders for this purpose is because of their proven ability to retain underlying patterns from time series representation while filtering out noise. As the learned model reconstructs subsequences that belong only to the healthy state, the resulting high reconstruction error for unhealthy subsequences gives the HI at those specific time instances. The difference in embeddings for normal and faulty machines tends to provide the degree of degradation and subsequently the RUL. We consider a scenario where RUL of a system is to be estimated at a certain operational instance with multi-sensor historical end-of-life data available. For this given scenario, let the multivariate input time series data be represented by the equation 1.1:

$$X_n^{(u)} = [x_1^{(u)}, x_2^{(u)}, \ldots x_{T_n}^{(u)}] \tag{1.1}$$

where $x^{(u)}_t \in R^p$ and p is the dimension of the sensor data (p no. of sensors at a time t), $u \in U$: is set of train instances of the system and T_n is the time of failure. The LSTM model takes each sequence of sensor measurements $X^{(u)}_n$ and learns how to model the whole sequence with respect to target RUL [31]. Let the Time Sequence of the i-th sensor measurement be given by the equation 1.2:

$$x_i^{(u)} = [x_i^{(u)1}, x_i^{(u)2}, \ldots x_i^{(u)T_n}] \tag{1.2}$$

where $i = (1, 2, \ldots, p)$, $t = (1, 2, \ldots, T_n)$. Let the sensor data be z-normalized over all U instances such that $x^{(u)}_{ti}$ is the sensor reading at time t for sensor i at instance u be transformed to $\frac{x_{ti}^{(u)} - \mu_i}{\sigma_i}$ where μ_i and σ_i are the mean and standard deviation for the i-th sensor data. Equation 1.3 below denotes a sub sequence of length l from the original input time series $X_n{}^{(u)}$ represented by:

$$X_n^{(u)}(t, l) = Z = [z_1, z_2, \ldots z_l] \tag{1.3}$$

such that $1 \leq t \leq L^{(u)} - l + 1$.

As seen in the many previous case studies like [46], [47], correlated multi-sensor data require an unsupervised pre-training stage for degradation feature extraction before the DL model can perform any supervised fine-tuning. Many times, the sensor data is inconsistent and requires a clustering algorithm for regime partitioning. This allows for a potentially more accurate RUL prediction. PCA has also previously been used in multi-sensor prognostics to reduce data dimensionality and preparing the validation set. Other studies like [48], [49] used PCA to observe the different operating condition patterns and build a HI for their case studies. The following sections break down the details of this approach in depth and explain the methodology of RUL estimation.

RUL estimation using embeddings

Although studies like [41], [42], [45] have estimated their machines' HI trend using a linear regression approach, thus assuming an exponential degradation, the drawbacks of such estimation have been stated in the previous section of this report. Referring to equation 1.2, every failed instance in $(u)_{Tn}$ corresponds to the total operating life and in equation 1.3, the input time series has been segmented into fixed length subsequences or windows. The encoder of our LSTM model can be trained in an unsupervised fashion to estimate health at the end of the fixed length $X_n^{(u)}(t, l)$. As seen in figure 2, the embeddings for these subsequences at the end of their fixed length retain important degradation related information and provides with the reconstruction error for healthy and faulty states. The following sections explain this process in detail.

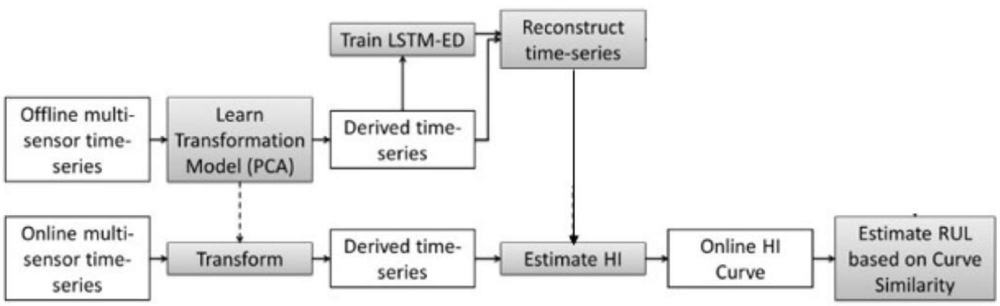


Figure 8 RUL Estimation steps using unsupervised HI based on LSTM-ED

Encoder-decoder based embeddings

We have discussed how LSTM-ED works in the ED section based on a sequence-to-sequence (seq2seq) framework, consisting of multi-layered DL models trained together. The RHS of equation 1.3 are the time series data points an LSTM encoder iterates through till they reach the final hidden state zl of a fixed dimension. The decoder on the other hand uses this z_l as its initial hidden state to reconstruct the input time series by performing gated transformation as seen in the equations in figure 1.1. The detailed mathematics behind this transformation can be found in [41]'s case study and the non-linear mapping of fixed dimensional time series between encoder and decoder is explained by figure 3 below. The representation of the encoder function as its final hidden state is basically the embedding to be learned. The reconstruction sequence (z_t') and error ($e_t^{(u)}$) obtained using LSTM-ED can mathematically be explained as equations 1.4 and 1.5 below respectively:

$$z_t^{'} = w^T a_t^{(D)} + b \tag{1.4}$$

$$e_t^{(u)} = ||z_t^{(u)} - z_t^{'(u)}|| \tag{1.5}$$

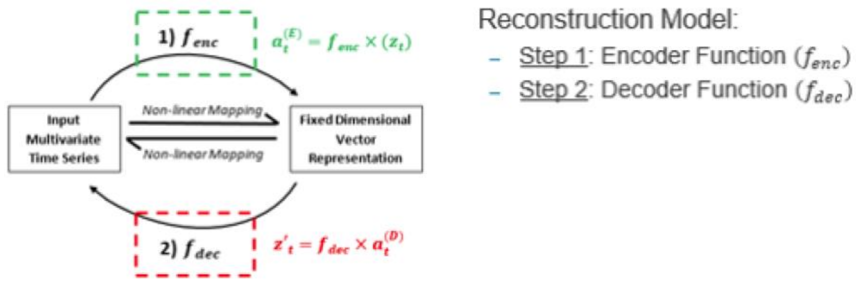


Figure 9 Encoder-decoder based reconstruction model

The LSTM-ED model is trained to minimize the loss function by squaring reconstruction error, provided by equation 1.6 below:

$$J = \sum_t ||z_t^{(u)} - z_t^{'(u)}||^2 \tag{1.6}$$

As no additional inputs are provided to the decoder following its initial hidden state from encoder, sufficient information required to reconstruct the time series is stored by the decoder. If the resulting output, the reconstructed time series is smooth, it means that the embeddings learned are able to retain the necessary degradation patterns from the input.

Obtaining HI and RUL prediction

The obtained encoder function or the embedding with important patterns can be used to highlight any differences between the healthy and faulty data. As the deterioration progresses over time, the healthy embeddings are used to train each subsequence of the original input series to provide prediction error. This difference between actual and predicted values is calculated to be the target normalized HI for that specific data point, represented by equation 1.7:

$$h_t^{(u)} = \frac{e_M^{(u)} - {e_t}^{(u)}}{e_M^{(u)} - {e_m}^{(u)}} \tag{1.7}$$

where $e_M{}^{(u)}$ and $e_m{}^{(u)}$ are the maximum and minimum values of reconstruction errors for u instances. The HI can have a value between 0 and 1, with 0 being very poor health and 1 being very healthy or normal. Most studies estimating HI use some scaling or normalizing procedure. The obtained HI curves for train and test instances are later matched for similarity (Equation 1.8) by varying the Euclidean distance between the two. The actual RUL prediction of the ground truth or test data is performed by the algorithm counting the number of train instances left after the last cycle of the test instance. If H^{u*} and H^u are the train and test instance HI curves, then their similarity can be quantified by:

$$s(u^*, u, t) = exp\frac{-d^2(u^*, u, t)}{\lambda} \tag{1.8}$$

where t is the time lag introduced to scale the inequality between the curves, d is the Euclidean distance and λ is an arbitrary parameter controlling similarity such that a smaller λ would mean less similarity and vice-versa. The weighted average RUL for the testing instance can be given by equation 1.9 below:

$$\hat{R}^{(u)^*} = \frac{\sum s(u^*, u, t).\hat{R}^{(u)^*}(u, t)}{\sum s(u^*, u, t)} \tag{1.9}$$

Experimental evaluation

Data description

The proposed model is used to calculate the RUL on NASA's famous C-MAPSS dataset for turbofan engine performance degradation and prognostics, generated from [42]. The data set is a multivariate time series representing turbofan engine usage from beginning till the end of its life. The data can be considered to be from a fleet of engines of the same type. Each entry (row) in the data set reflects an operational cycle of a specific engine identified by engine id and cycle time. There are multiple entries per engine to represent different reporting times. Other columns represents different features like 3 operational settings and 21 sensors.

The C-MAPSS data set is divided into 4 subsets each for training, test, and RUL (FD001, FD002, FD003, FD004) and each subset can have a 100 instances. A subset can have a different operational condition and consists of a different number of engines. In the training set, the last id, cycle entry is when the engine is declared unhealthy. For example if the first engine has 192 distinct time series events the cycle will go from 1 to 192, while the RUL will start with 192 and go down to 1. If the time series for an engine in the test data ends at 41, the model's goal is to identify the RUL at that point. Figure 4 below shows a random sample plot of 10 engines from sensors 7 to 12 from the training data. While there is some visible noise, we can say that not all engines display degradation at the same time and there is no necessary correlation between the data as the engine approach failure. Also, the data can be affected by incorrect sensor placement or other human error as might be the case in plots from sensor 10. There are a total of 20631 cycles for training engines and 13096 cycles for testing engines with each engine having a different degree of initial wear.

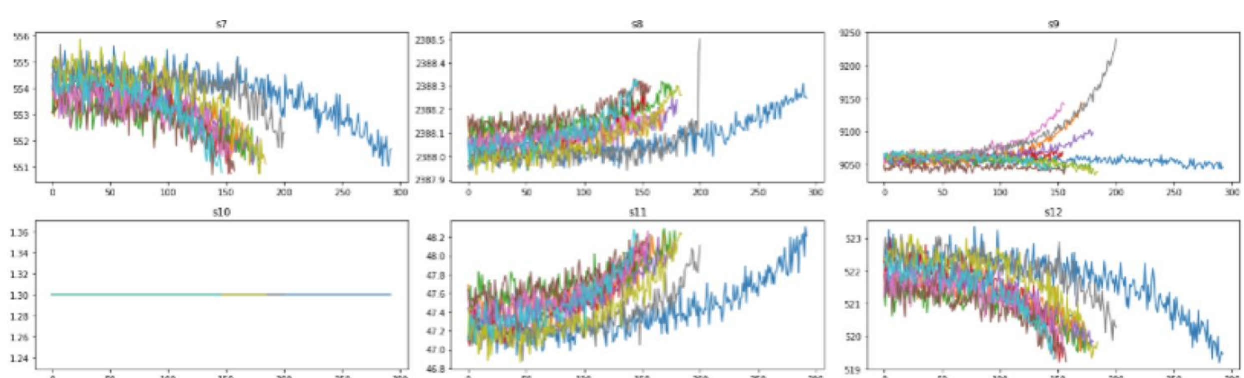


Figure 10 Data visualization sample of 10 random engines from sensors 7 – 12

Performance metrics

The LSTM network built for this experiment has 100 units in its first layer followed by another LSTM layer with 50 units. Dropout is also applied after each LSTM layer to control over fitting. The literature proposes several performance metrics to evaluate the performance of prognostic models. The proposed LSTM-ED model is evaluated on metrics such as Mean Absolute Error (MAE), Coefficient of Determination (R^2) and Model Loss [50]. Figure 5 provides the metric scores used.

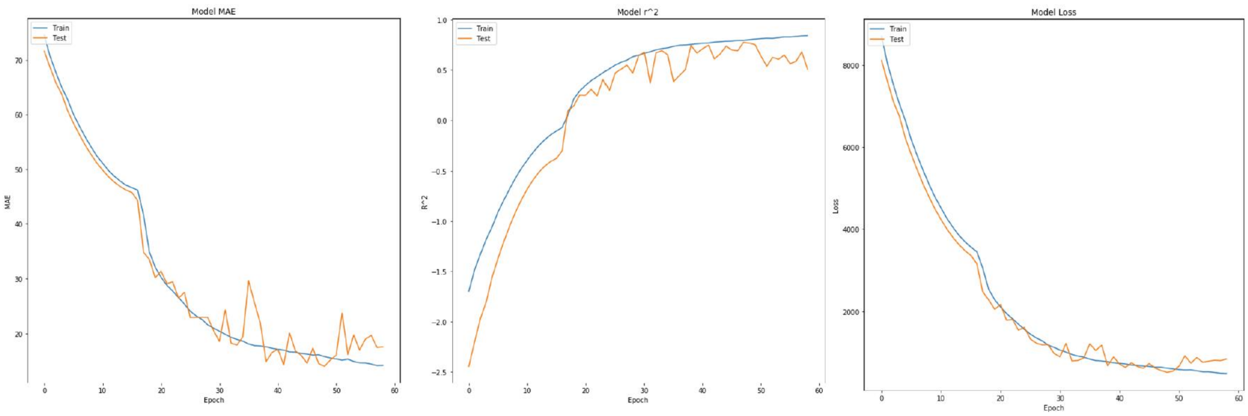


Figure 11 Robustness evaluation

Results: embedding RUL

Table 2 shows the performance of the proposed LSTM-ED model for Embedding RUL (HI similarity) approached as compared against benchmark models from previous studies like the Linear Regression based exponential curve assumption (LR-Exp) and reconstruction error based RUL (Recon-RUL) and their results as obtained. Figure 6 shows the final RUL of a random fleet of engines taken from each of the test files. The parameters chosen are the number of principal components p, the number of LSTM units in the hidden layers of encoder and decoder c, window/subsequence length l, maximum allowed time-lag τ, similarity threshold, maximum predicted RUL $R_{\max}$ and the λ parameter. The blue line reflects the actual RUL while the orange line is the predicted RUL. The actual RUL values are taken from the ground truth or actual RUL files. The earlier values seem to have monotonically increased from this point. We can say that to the best of our knowledge, the proposed approach scores higher on the performance metrics and the future work will be to further test the approach against other approaches from literature and also increase the number of metrics that this comparison will be based on.

Table 8 Performance of LSTED-ED model for embedding RUL

Model Approach	MAE	R^2	Model Loss
LR-Exp	9.8	NR	NR
Recon-RUL	8	NR	NR
Embed-RUL (Proposed)	9.9	NR	NR

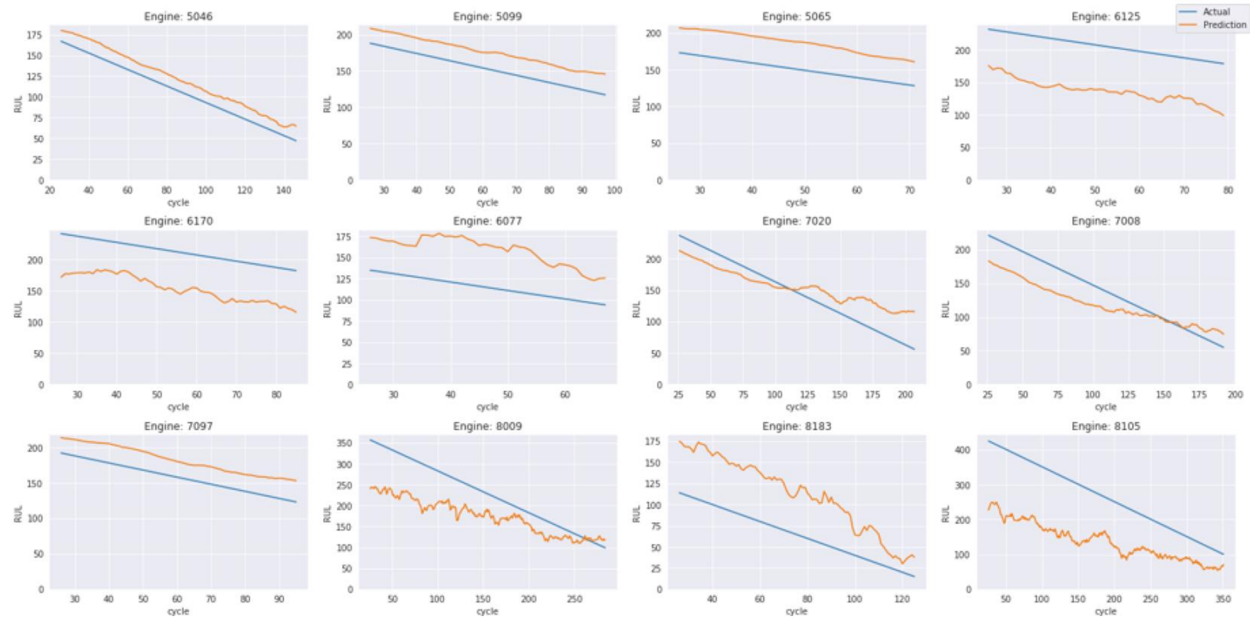


Figure 12 RUL of random set of engines from test files

References

[1] F. Tao, J. Cheng, Q. Qi, M. Zhang, H. Zhang, and F. Sui, “Digital twin-driven product design, manufacturing and service with big data,” *The International Journal of Advanced Manufacturing Technology*, vol. 94, no. 9-12, pp. 3563–3576, 2018.

[2] J. Lee, B. Bagheri, and H.-A. Kao, “A cyber-physical systems architecture for industry 4.0-based manufacturing systems,” *Manufacturing Letters*, vol. 3, pp. 18–23, 2015.

[3] K. Ashton, “That ‘internet of things’ thing.,” *RFID Journal*, vol. 11, no. 2, pp. 467–475, 2009.

[4] E.Negri, L. Fumagalli, and M. Macchi, “A reviewof the roles of digital twin in cps-based production systems,” *Procedia Manufacturing*, vol. 11, pp. 939–948, 2017.

[5] S. Boschert and R. Rosen, “Digital twin—the simulation aspect,” in *Mechatronic Futures*, Springer, 2016, pp. 59– 74.

[6] D. An, N. Kim, and J.-H. Choi, “Practical options for selecting data-driven or physics-based prognostics algorithms with reviews,” *Reliability Engineering and System Safety*, vol. 133, pp. 223–236, 2015.

[7] Z. Chen, S. Cao, and Z. Mao, “Remaining useful life estimation of aircraft engines using a modified similarity and supporting vector machine (svm) approach,” *Energies*, vol. 11, pp. 1–14, 2017.

[8] M. Grieves, “Virtually perfect,” 2011.

[9] M. Grieves and J. Vickers, “Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems,” in *Transdisciplinary perspectives on complex systems*, Springer, 2017, pp. 85–113.

[10] K. M. Alam and A. El Saddik, “C2ps: A digital twin architecture reference model for the cloud-based cyberphysical systems,” *IEEE Access*, vol. 5, pp. 2050–2062, 2017.

[11] R. Rosen, G. Von Wichert, G. Lo, and K. D. Bettenhausen, “About the importance of autonomy and digital twins for the future of manufacturing,” *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, 2015.

[12] A. Cerrone, J. Hochhalter, G. Heber, and A. Ingraffea, “On the effects of modeling as-manufactured geometry: Toward digital twin,” *International Journal of Aerospace Engineering*, vol. 2014, 2014.

[13] M. Bajaj, B. Cole, and D. Zwemer, "Architecture to geometry-integrating system models with mechanical design," in *AIAA SPACE 2016*, 2016, p. 5470.
[14] A. Canedo, "Industrial iot lifecycle via digital twins," in *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, ACM, 2016, p. 29.
[15] M. Schluse and J. Rossmann, "From simulation to experimentable digital twins: Simulation-based development and operation of complex technical systems," in *2016 IEEE International Symposium on Systems Engineering (ISSE)*, IEEE, 2016, pp. 1–6.
[16] E. Glaessgen and D. Stargel, "The digital twin paradigm for future nasa and us air force vehicles," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, 2012, p. 1818.
[17] K. Sivalingam, M. Sepulveda, M. Spring, and P. Davies, "A review and methodology development for remaining useful life prediction of offshore fixed and floating wind turbine power converter with digital twin technology perspective," in *2018 2nd International Conference on Green Energy and Applications (ICGEA)*, IEEE, 2018, pp. 197–204.
[18] E. Tuegel, "The airframe digital twin: Some challenges to realization," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, 1812, p. 1812.
[19] B. Smarslok, A. Culler, and S. Mahadevan, "Error quantification and confidence assessment of aerothermal model predictions for hypersonic aircraft," in *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, 2012, p. 1817.
[20] K. Reifsnider and P. Majumdar, "Multiphysics stimulated simulation digital twin methods for fleet management," in *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, 2013, p. 1578.
[21] Y. Bazilevs, X. Deng, A. Korobenko, F. L. di Scalea, M. Todd, and S. Taylor, "Isogeometric fatigue damage prediction in large-scale composite structures driven by dynamic sensor data," *Journal of Applied Mechanics*, vol. 82, no. 9, p. 091 008, 2015.
[22] G. Grinshpun, T. Cichon, D. Dipika, and J. Rossmann, "From virtual testbeds to real lightweight robots: Development and deployment of control algorithms for soft robots, with particular reference to," in *ISR 2016: 47st International Symposium on Robotics; Proceedings of*, VDE, 2016, pp. 1–7.
[23] B. R. Seshadri and T. Krishnamurthy, "Structural health management of damaged aircraft structures using digital twin concept," in *25th AIAA/AHS Adaptive Structures Conference*, 2017, p. 1675.
[24] T. DebRoy,W. Zhang, J. Turner, and S. Babu, "Building digital twins of 3d printing machines," *Scripta Materialia*, vol. 135, pp. 119–124, 2017.
[25] F. Tao, M. Zhang, Y. Liu, and A. Nee, "Digital twin driven prognostics and health management for complex equipment," *CIRP Annals*, 2018.
[26] X.-W. Chen and X. Lin, "Big data deep learning: Challenges and perspectives," *IEEE access*, vol. 2, pp. 514–525, 2014.
[27] J. Zhang, P. Wang, R. Yan, and R. X. Gao, "Long short-term memory for machine remaining life prediction," *Journal of manufacturing systems*, vol. 48, pp. 78–86, 2018.
[28] R. Zhao, R. Yan, Z. Chen, K. Mao, P. Wang, and R. X. Gao, "Deep learning and its applications to machine health monitoring," *Mechanical Systems and Signal Processing*, vol. 115, pp. 213–237, 2019.
[29] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on knowledge and data engineering*, vol. 22, no. 10, pp. 1345–1359, 2009.
[30] X. Li, Q. Ding, and J.-Q. Sun, "Remaining useful life estimation in prognostics using deep convolution neural networks," *Reliability Engineering & System Safety*, vol. 172, pp. 1–11, 2018.
[31] S. Zheng, K. Ristovski, A. Farahat, and C. Gupta, "Long short-term memory network for remaining useful life estimation," in *2017 IEEE International Conference on Prognostics and Health Management (ICPHM)*, IEEE, 2017, pp. 88–95.
[32] S. Hochreiter and J. Schmidhuber, "Long short-termmemory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
[33] A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional lstm and other neural network architectures," *Neural Networks*, vol. 18, no. 5-6, pp. 602–610, 2005.
[34] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using rnn encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.

[35] Y. Wu, M. Yuan, S. Dong, L. Lin, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla lstm neural networks," *Neurocomputing*, vol. 275, pp. 167–179, 2018.
[36] A. Z. Hinchi and M. Tkiouat, "Rolling element bearing remaining useful life estimation based on a convolutional long-short-term memory network," *Procedia Computer Science*, vol. 127, pp. 123–132, 2018.
[37] E. Marchi, F.Vesperini, F. Eyben, S. Squartini, and B. Schuller, "A novel approach for automatic acoustic novelty detection using a denoising autoencoder with bidirectional lstm neural networks," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, IEEE, 2015, pp. 1996–2000.
[38] P. Malhotra, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "Lstm-based encoder-decoder for multi-sensor anomaly detection," *arXiv preprint arXiv:1607.00148*, 2016.
[39] S. Bengio, O. Vinyals, N. Jaitly, and N. Shazeer, "Scheduled sampling for sequence prediction with recurrent neural networks," in *Advances in Neural Information Processing Systems*, 2015, pp. 1171–1179.
[40] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Advances in neural information processing systems*, 2014, pp. 3104–3112.
[41] P. Malhotra, V. TV, A. Ramakrishnan, G. Anand, L. Vig, P. Agarwal, and G. Shroff, "Multi-sensor prognostics using an unsupervised health index based on lstm encoder-decoder," *arXiv preprint arXiv:1608.06154*, 2016.
[42] A. Saxena, K. Goebel, D. Simon, and N. Eklund, "Damage propagation modeling for aircraft engine run-tofailure simulation," in *2008 international conference on prognostics and health management*, IEEE, 2008, pp. 1–9.
[43] A. Saxena and K. Goebel, "Turbofan engine degradation simulation data set," *NASA Ames Prognostics Data Repository*, 2008.
[44] T.Wang, J. Yu, D. Siegel, and J. Lee, "A similarity-based prognostics approach for remaining useful life estimation of engineered systems," in *2008 International Conference on Prognostics and Health Management*, IEEE, 2008, pp. 1–6.
[45] E. Ramasso, "Investigating computational geometry for failure prognostics.," *International Journal of Prognostics and Health Management*, vol. 5, no. 1, p. 005, 2014.
[46] A. Mosallam, K. Medjaher, and N. Zerhouni, "Data-driven prognostic method based on bayesian approaches for direct remaining useful life prediction," *Journal of Intelligent Manufacturing*, vol. 27, no. 5, pp. 1037–1048, 2016.
[47] A. L. Ellefsen, E. Bjørlykhaug, V. Æsøy, S. Ushakov, and H. Zhang, "Remaining useful life predictions for turbofan engine degradation using semi-supervised deep architecture," *Reliability Engineering & System Safety*, vol. 183, pp. 240–251, 2019.
[48] J. S. Bucknam, "Data analysis and processing techniques for remaining useful life estimations," 2017.
[49] Y. Liu, X. Hu, and W. Zhang, "Remaining useful life prediction based on health index similarity," *Reliability Engineering & System Safety*, vol. 185, pp. 502–510, 2019.
[50] A. Saxena, J. Celaya, E. Balaban, K. Goebel, B. Saha, S. Saha, and M. Schwabacher, "Metrics for evaluating performance of prognostic techniques," in *2008 International Conference on Prognostics and Health Management*, IEEE, 2008, pp. 1–17.

Digital twin of reheat furnace

Kuilin Chen

Introduction

Digital twin is one of the core concepts in Industry 4.0 with the integration of Internet of Things, artificial intelligence, and optimal decision making. The definition and application of digital twin is reviewed at the beginning of this article, followed by detailed design and development of digital twin for a reheat furnace through data-driven modelling. Finally, a dynamic generative network model is proposed to approximate the behavior of an industrial furnace

Literature review

In recent years, Industry 4.0 has become an emerging concept with the amazing growth and advancement in digital technologies that allow the integration of Internet of Things (IoT), cloud computing and artificial intelligence, etc. [Lasi et al., 2014, Jazdi, 2014]. In this fourth wave of industrial revolution, traditional manufacturing companies make huge investment to digitize their assets to take the advantage of smart factory, decentralized organization and flexibility to increase their efficiency and profitability in competition [Drath and Horch, 2014, Wan et al., 2016, Zhou et al., 2015]. Among several fundamental concepts within Industrial 4.0, digital twin (DT) plays a central role because it bridges the assets and processes in the physical world with the models and analytics at the virtual space in the real-time fashion [Babiceanu and Seker, 2016]. Essentially, DT can provide company with complete digital footprint of their products and equipment for the entire life cycle, leading to deep understanding of product development and equipment health conditions [Qi and Tao, 2018]. With the real-time synchronization of floor shop in physical world and integrated models and software in virtual space, manufacturing companies are able to detect and solve physical issues in product and equipment sooner, predict outcomes at high fidelity, design and build better products, and, ultimately, better serve their customers. Therefore, the DT concept becomes more and more prevalent in both industry and academia by boosting company's revenue through rapid product development, improved operation efficiency, reduced defect rate and optimized maintenance schedule.

Definition of digital twin

Although DT is a very new concept without a universal definition, researchers from different areas have attempted to define it in several different ways based on problems and issues they want to solve within the Industrial 4.0 framework in their area. The conceptual model of a virtual, digital equivalent to a physical product is proposed in 2003, with the expectation that all the data and information of a physical system could be replicated in software for simulation and analysis. [Grieves, 2011, Grieves and Vickers, 2017]. That conceptual model, though not being called as DT, but embodies all essential parts of DT, such as the physical space, the virtual space, and the linkage or interface between the two spaces [Cerrone et al., 2014, Bajaj et al., 2016]. DT is formally defined through a case study by NASA and US air force as an integrated simulation models of an as-built vehicle based on the best available information, and history data to reflect wear and tear of its corresponding physical twin while in use communicated by sensor updates [Glaessgen and Stargel, 2012, Tuegel et al., 2011]. Researchers in aerospace community develop various digital twins to mirror the health conditions of a aircraft or parts by integration with other aspects such as product life-cycle management (PLM), feedback from real world, and prognostics and diagnostics activities [Tuegel et al., 2011, Gockel et al., 2012, Reifsnider and Majumdar, 2013]. The concept of digital twin is expanded as a digital counterpart product to its physical instance, which could be used for products from any manufacturing industries, though their research focus is still on manufacturing of air vehicles [Ríos et al., 2015].

Thanks to the rapid development in data acquisition, cloud computing and big data technologies in manufacturing industries, DT is also developed and defined in manufacturing industries as a virtual counterpart of production resources based on coupled models to simulate the condition of the equipment and machines [Lee et al., 2015]. In parallel, the idea of DT is closely integrated with some ongoing research of cyber-physical system (CPS) and virtual factory (VF) in smart manufacturing area [Alam and El Saddik, 2017]. Simulation is a core functionality of DT in manufacturing, which provides seamless assistance and support for operation and service by means of direct linkage of operation data [Rosen et al., 2015]. Along with the view of simulation, DT is composed of very realistic models of the process current state and its behavior in interaction with the environment in the real world [Gabor et al., 2016]. On the other hand, DT also provides a new mechanism to manage IoT with virtual substitutes of real world objects consisting of virtual representations and communication capabilities making up smart objects acting as intelligent nodes inside the internet of things and services [Canedo, 2016, Schluse and Rossmann, 2016].

Application of digital twin

By definition, DT is designed to model sophisticated assets or processes with iteration to external environment, which are very difficult to quantify by traditional methodologies. As such, it is beneficial to review the current application of DT in a wide range of contexts to understand why and how DT is created and implemented before actually developing DT for any particular assets or processes.

DT was initially developed for aerospace vehicles because it's impossible to reproduce the extreme thermal, mechanical, and acoustical loadings in a laboratory environment to understand the health condition of the vehicle. However, health condition of the aerospace vehicle cannot be obtained from traditional simulation methods because they are not able to integrate the sub-models at different scale or handle the stochastic input data from external environment. Therefore, an integrated multiphysics, multiscale, probabilistic simulation of an as-built vehicle is developed as a digital twin to mirror the conditions of aerospace vehicle in actual flying [Glaessgen and Stargel, 2012]. The finite element method (FEM) model for local structural loads and computational fluid dynamics (CFD) model for artificially flying the aircraft, at different geometric scale, are closely coupled to predict the local structural damage and material state in repose to flight conditions, with the auxiliary of computer aided design (CAD) [Tuegel et al., 2011, Tuegel, 2012, Kraft, 2016]. Nevertheless, neither mathematical modelling details nor performance statistics of DT is provided in those early papers.

Recently, an FEM based DT with modelling details was presented to monitor structural health condition. Heuristic optimization algorithm is utilized to identify key parameters in the model, leaning to accurate monitoring results with comparison to real reference data [Seshadri and Krishnamurthy, 2017]. The aforementioned aircraft or airframe health condition monitoring is based on deterministic physical models within DT, which cannot handle epistemic uncertainty due to lack of knowledge. Consequently, machine learning algorithms such as particle filter driven dynamic Bayesian network are developed as an extra layer above the physical model based DT to get accurate prognostics and diagnostics results through stochastic input-output data from DT [Li et al., 2017].

Besides the aerospace field, DT is well received in smart manufacturing environment, especially with the wave of promoting Industry 4.0. DT has been endowed with new functionality and usage when applied to industrial processes in manufacturing area, beyond the scope of prognostics and diagnostics activities. Apart from acting as digital replica for physical assets at floor shop, DT of manufacturing processes is also developed to provide insightful guidance to improve production efficiency. For instance, a DT of an additive manufacturing process, using multiple analytical sub-models including heat transfer models and fluid dynamics models, is built to predict the spatial and temporal variations of metallurgical parameters, which is proved to be highly accurate with validation of experiment data [Knapp et al., 2017, DebRoy et al., 2017]. Meanwhile, a rapid DT design framework is proposed by adopting the idea of reference models, which refers to producing a copy of a production line and using it for reasoning about other instances of the similar production line. A DT based on such rapid development provides decoupling analytics for a hollow glass production line, as well as multi-objective optimization towards decision making [Zhang et al., 2017].

Gap in current research

Since the idea of DT was proposed in 2010s, a dozen of papers have been published around this topic. However, after thorough literature review, it is found that most of them are review and introduction papers about potential application of DT in different areas without any detailed examples. Table 1 lists all reviewed original research papers, which propose methodologies to build DT for a specific system or process.

Table 9 Digital twin examples

Ref	Methodology	Model Details	Performance Statistics
Tuegel et al. [2011]	FEM and CFD	No	No
Glaessgen and Stargel [2012]	3D FEM, quantum mechanics and molecular dynamics	No	No
Kraft [2016]	physics-based model and experiment	No	No
Li et al. [2017]	FEM and CFD	No	No
Seshadri and Krishnamurthy [2017]	FEM and genetic algorithm	Yes	Yes
Knapp et al. [2017]	analytical sub-models, heat transfer model and fluid dynamics model	Yes	Yes
DebRoy et al. [2017]	heat transfer model	No	Yes
Tao and Zhang [2017]	FEM and CFD	No	No
Zhang et al. [2017]	reference models	No	Yes

Although a mathematical model is an essential part of DT and almost every DT example listed in Table 1 uses different forms of mathematical models, the majority of current published papers do not provide any model details. No matter model details are provided or not, first-principle models or physical models are used in aforementioned examples. Identification of such models with nonlinear equations is a non-trial mathematical optimization problem. Apart from one paper explicitly using genetic algorithm, none of them illustrates how to identify the parameters in models. Such practice in current DT literature makes it extremely difficult to duplicate the methods proposed by those DT papers, though it is understood that some systems, products and processes contain highly confidential and sensitive information in business and military.

In addition, development of DT upon first-principle models requires in-depth knowledge of the system or process, and takes significant effort and time of experienced engineers. It is inevitable that some knowledge is not available for some very complex systems, leading to epistemic uncertainty in models due to lack of knowledge. Furthermore, less than half of the papers in Table 1 present the performance statistics of DT by comparison to its twin in physical world. DT itself is meaningless without high fidelity output compared with the physical system. Development of DT is not a one-time deal. Mismatch between the DT and physical system can happen due to equipment deterioration and external environmental change. How to develop an adaptive DT has not been studied yet.

It is an overwhelming task to develop DT for all processes and equipment in a manufacturing company all at once. A reasonable plan for digitalization of a tradition manufacturing company is to develop an initial DT for one chosen process along with its equipment, deliver value there, and continue to develop for other processes. The remaining of this article is organized as follows. Development of DT for a reheat furnace in a steel manufacturing company is discussed in details in the next section, followed by some preliminary results. Finally, we conclude our current work and propose some research ideas regarding development of DT for reheat furnace.

Design of digital twin for reheat furnace

The fundamental goal of the digital twin is to predict the output (slab drop-out-temperature and through-put) of reheat furnace with high fidelity. Due to the complexity of the reheat furnace system, the digital twin of reheat furnace can be decomposed into combustion system model, slab temperature model, control systems and simplified down stream model. In addition, the output of one subsystem can be the input for another subsystem. The mathematical details of subsystem models are presented in following subsections.

Combustion zone model

The reheat furnace consists of 12 combustion zones which control the zone temperature independently. In principle, the combustion zone temperature change Δz_t between current zone temperature z_t and last zone temperature z_{t-1} is drive by the air flow a_t and gas flow g_t through an unknown mapping function $f_p(\cdot)$ as follows

$$\begin{aligned} \Delta z_t &= f_p(a_t, g_t) \\ z_t &= z_{t-1} + \Delta z_t + \epsilon_{z_t} \end{aligned} \tag{1}$$

where ϵ_{zt} is the noise or disturbance. The air flow a_t and gas flow g_t are controlled by the zone temperature set point sp_t via a proportional-integral-derivative (PID) controller as follows

$$\begin{aligned} a_t &= K_{p_a} e_{z_t} + K_{i_a} \sum_{i=1}^{t} e_{z_i} + K_{d_a}(e_{z_t} - e_{z_{t-1}}) + \epsilon_{a_t} \\ g_t &= K_{p_g} e_{z_t} + K_{i_g} \sum_{i=1}^{t} e_{z_i} + K_{d_g}(e_{z_t} - e_{z_{t-1}}) + \epsilon_{g_t} \end{aligned} \tag{2}$$

where $e_{zt} = sp_t - z_t$ is the difference between the temperature set point and actual zone temperature, K_{pa}, K_{ia}, K_{da}, K_{pg}, K_{ig} and K_{dg} are the proportional, integral and derivative gains for air flow a_t and gas flow g_t respectively. It is not desirable to identify $f_p(\cdot)$ and predict zone temperature z_t using Eqs. 1 and 2 for several reasons. First, the model requires data from infinite past due to the integral part in Eq. 2. Second, the actual air and gas flows are subject to random disturbance, variation in air and gas composition and actuator failure which cannot be characterized by Eq. 2. The zone temperature prediction may not be accurate due to the propagation of errors. Third, it's very difficult or intractable to optimize the slab temperature performance or throughput based on this digital twin because the order of decision variable (temperature set point) is too large.

The best practice in control system optimization is to identify a transfer function $f_c(\cdot)$ between the input and output variables. Eq. 1 and Eq. 2 are equivalent to the transfer function as follows

$$z_t = f_c(z_{t-1}, z_{t-2}, ..., sp_{t-1}, sp_{t-2}, ...) + \epsilon_{z_t} \tag{3}$$

For a meaningful model, the orders of zone temperature and set points are truncated to r and s, respectively. Let vector $\mathbf{x}_t = [z_{t-1}, ..., z_{t-r}, sp_{t-1}, ..., sp_{t-s}]^T$ denotes the stacked past temperature and set points. The combustion zone model can be expressed as follows

$$z_t = f_c(\mathbf{x}_t, \omega) + \epsilon_{z_t} \tag{4}$$

where ω denotes the vector for model parameters and ϵ_{zt} is the model error term between measured zone temperature z_t and modelled zone temperature $\widehat{z_t} = f_c(x_t, \omega)$. The task of combustion zone model design is to find an appropriate model structure and model order (r and s) based on available data and knowledge about the combustion system to obtain minimized prediction error from the identified combustion zone model. The identification process is to find a $f_c(\cdot)$ with given structure and model orders by minimizing the error term as follows

$$\begin{aligned} J(\omega) &= \frac{1}{N}\sum_{t=1}^{N}(\hat{z}_t - z_t)^2 \\ \hat{\omega} &= \arg\max_{\omega} J(\omega) \end{aligned} \tag{5}$$

where N is the total available number of samples and ^! is the estimated model parameter vector.

Different from traditional simulation models, digital twin can be synchronized with the physical system. The frequency of synchronization depends on the purpose of digital twin. Due to the randomness in combustion system, the output from the digital twin can never be the same as the physical system. At each round of synchronization, the current predicted output can be calibrated by the real output data from physical system. The calibrated value can be used for future prediction. In addition, the dynamics in the transfer function is subject to change because the combustion system is a stochastic system. $\widehat{\omega}$ can be updated to capture the current dynamics in the system using the latest data from synchronization.

Slab temperature model

After a slab is charged into reheat furnace, it goes through different combustion zones when it travels from the charge side to discharge side. Due to the nature of digital computer control, the slab temperature model s_t is a discrete time model computed at a fixed time interval (30 seconds) as follows:

$$
\begin{aligned}
\Delta s_t &= f_s(\hat{s}_{t-1}, z_{t-1}, h) \\
\hat{s}_t &= \hat{s}_{t-1} + \Delta s_t \\
s_t &= \hat{s}_t + \epsilon_{s_t}
\end{aligned} \tag{6}
$$

where $\hat{s}_t$ is the predicted slab temperature, ϵ_{st} is the error term, h is the heat transfer coefficient based on thickness and chemistry composition of slab and $f_s(\cdot)$ is a known nonlinear heat transfer function available on reheat furnace control system. Note that only the initial charge temperature s_0 and final drop-out-temperature s_T can be measured. The future slab temperature is predicted based on current predicted slab temperature because no slab temperature measurement is available inside the furnace.

The design of slab temperature model in digital twin can be as simple as duplicating $f_s(\cdot)$ from reheat furnace control system. The input to the model is initial charge temperature s_0 and zone temperature z_t that can be calculated by model described in the combustion zone model.

Control system

The control system sets up the heating trajectory for each slab at charge time and computes the combustion zone set points for all zones and reheat furnace pacing dynamically at a fixed time interval (30 seconds). The set point of a combustion zone is calculated based on the temperature of all slabs in that zone with respect to their current temperature targets from the heating trajectory as follows:

$$
\begin{aligned}
e_t &= \sum_{j=1}^{K} (s_t^j - t_t^j) \\
sp_t &= g_s(e_t)
\end{aligned} \tag{7}
$$

where $j = \{1, \ldots, K\}$ denotes all slabs at one zone. Meanwhile, pace rate is calculated for each slab inside the furnace based on their current temperature and target. The furnace pacing is dominated by the slowest slab inside the furnace.

$$
\begin{aligned}
e_t^j &= s_t^j - t_t^j \\
p_t^j &= g_p(e_t^j) \\
p_t &= \min(e_t^1, e_t^2, \ldots, e_t^N)
\end{aligned} \tag{8}
$$

where $j = \{1, \ldots, N\}$ denotes all slabs inside the furnace, p^j_t represents the pace rate for an individual slab, p_t is the pace rate for the furnace. Eqs. 7 and 8 should be duplicated in the digital twin of reheat furnace to get the same control action behavior as the physical system.

Simplified downstream model

The reheat furnace requires downstream processing time. On the one hand, the slab temperature target relies on the predicted downstream processing time. On the other hand, the actual throughput is also affected by the downstream processing time because a slab cannot be extracted if the downstream processing is not complete. A simplified model can be developed to predict the downstream processing time.

$$t = g_d(d_1, d_2, c) \tag{9}$$

where d_1 is the original dimension (width, thickness and length) of the steel before downstream processing, d_2 is the final dimension of steel, c is a vector of slab characteristics such as chemistry composition, hardness, steel grade and special processing requirement. The input to $g_d(\cdot)$ includes both continuous and categorical data, and the relationship between input and output can be quite nonlinear. In this case, deep neural network could be a good candidate for simplified downstream model to predict the processing time.

Preliminary example for combustion zone model

A preliminary example of combustion zone model is demonstrated in this section. It can be treated as a toy example for digital twin of reheat furnace. Several transfer function identification methods are considered and compared in this section, including Box-Tiao transformation based ARMAX model, PEM based ARX model and gated recurrent unit (GRU) network. Diebold-Mariano test is used to compare the prediction accuracy between models. Finally, future research path is discussed based on pros and cons of all methods on the numerical performance on temperature control system identification.

System identification of combustion zone

System identification is the process of using appropriate mathematical models and optimization algorithms to determine a relationship between future outputs y_t and past observations of inputs u_{t-i} and outputs y_{t-i} by minimizing the error between measured outputs and model outputs. Without loss of generality, the input-output relationship can be expressed as follows:

$$\begin{aligned} &y_t = f(y_{t-1}, \ldots, y_{t-p}, u_{t-b}, \ldots, u_{t-b-s}, \epsilon_{t-1}, \ldots, \epsilon_{t-q}) + \epsilon_t \\ &r \geq 1, b \geq 0, s \geq 0, d \geq 0 \end{aligned} \tag{10}$$

where ϵ_t is the error between measured and predicted system output at time t, while b, p, r and d denote input delay, input lag, output lag and error lag, respectively.

ARMAX model

Liner transfer function can be an Auto-Regressive Moving-Average with exogenous input (ARMAX) model, which relates the current model output, past inputs, past outputs and past prediction errors [Box et al., 2015]. The ARMAX model can be expressed as follows:

$$y_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \sum_{i=b}^{s} \omega_i u_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} + \epsilon_t \tag{11}$$

where ϕ_i, ω_i and θ_i are the ARMAX model parameters. The ARMAX model can be identified by the Box-Tiao method in an iterative way [Box and Tiao, 1975]. A group of preliminary AR orders p and MA orders q can be selected on autocorrelation function (ACF) plot and partial autocorrelation function (PACF) plot. The final model order is determined by Akaike information criterion (AIC). The prediction residuals of ARMAX model identified by Box- Tiao method is guaranteed to be uncorrelated.

GRU network

Neural network has been widely used for time series forecasting since 1980s due to its strong capability to approximate nonlinear relationship between past and future data [Hunt et al., 1992, Narandra and Parathasarathy, 1990]. The recent emergence of deep learning and efficient learning algorithms have enhanced the modelling and prediction capability of NN, but traditional feedforward NN cannot use its reasoning about previous input-output data to inform future ones. Recurrent neural networks (RNN) are perfect for time-series forecasting because they are networks with loops inside them, allowing memory of past information to persist [Funahashi and Nakamura, 1993]. In Fig. 1, input $\mathbf{x}_t$ is fed to a neural network unit A with an output value h_t. The output of current step of the network is passed to the next step. The directed acyclic graph of RNN can be unrolled as Fig. 1 and the unrolled chain-like structure reveals that it's intimately related to time series.

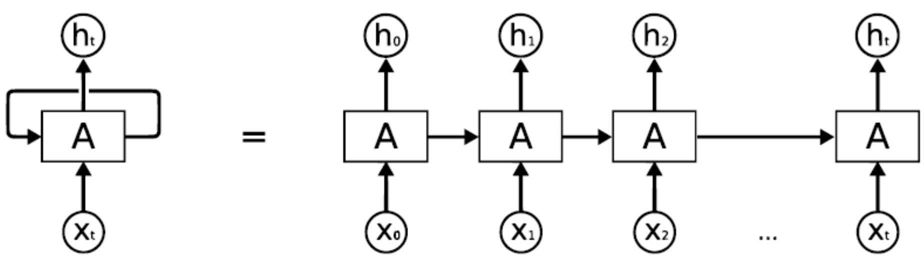


Figure 13 Unfolded recurrent neural network

In practice, RNN is not able to handle the long-term dependencies in time series well and it also suffers from gradient vanishing problem. The long short-term memory (LSTM) networks are a special kind of RNN, with recurrent gates called forget gates to avoid the long-term dependency problem [Hochreiter and Schmidhuber, 1997]. LSTM also prevents gradient vanishing or exploding in network training. More recently, GRU without an output gate is developed as one of the most notable LSTM variants [Cho et al., 2014]. GRU is chosen in this study because it has simpler structure and fewer parameters but leads to better performance than traditional LSTM [Greff et al., 2017, Jozefowicz et al., 2015]. The structure of a GRU is show in Fig. 2 and the mathematical details are presented below.

$$
\begin{aligned}
z_t &= \sigma(W_z \cdot [h_{t-1}, \mathbf{x}_t] + b_z) \\
r_t &= \sigma(W_r \cdot [h_{t-1}, \mathbf{x}_t] + b_r) \\
h_t &= z_t \circ h_{t-1} + (1 - z_t) \circ \tanh\left(W_h \cdot [r_t \circ h_{t-1}, \mathbf{x}_t] + b_h]\right)
\end{aligned} \tag{12}
$$

where $\circ$ denotes the Hadamard product, $\sigma(\cdot)$ is sigmoid function, f_t is the forget gate's activation, i_t is the input gate's activation, C_t is the cell state, $\tilde{C}_t$ is the new candidate cell state, o_t is the output gate's activation, W_* and b_* are weight and bias term for each gate to be identified through network training.

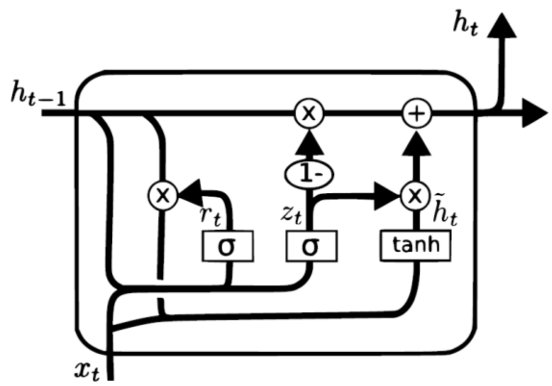


Figure 14 GRU

The neural network training process is to find the optimal network parameters (Φ = [**W**, **b**]) by minimizing a cost function, which can be the mean squared error between network's predicted output $\hat{y}_t = h_t$ and measured output y_t,

$$
\begin{aligned}
J(\Phi) &= \frac{1}{n}\sum_{t=1}^{n}(\hat{y}_t - y_t)^2 + \lambda\|\Phi\|_2 \\
\hat{\Phi}^T &= \arg\max_{\Phi} J(\Phi)
\end{aligned} \tag{13}
$$

where λ is a regularization term to prevent overfitting. Stochastic gradient descent (SGD) algorithm is utilized to update each parameter in the direction of negative gradient of the cost function with respect to the corresponding parameter as follows

$$
\Phi_{k+1} = \Phi_k - \alpha \nabla_{\Phi} J(\Phi) \tag{14}
$$

where α is the learning rate in SGD algorithm [Goodfellow et al., 2016]. A local minima can be found by SGD algorithm because Eq. 13 is a high-dimensional non-convex optimization problem. In order to find the best parameters of GRU, parameters are initialized at different values randomly in various training rounds.

Numerical results

The root mean squared error (RMSE) of predicted future temperature by three methods are shown in Tab. 2. RMSE of training and test data sets are calculated based on one-step ahead prediction. One-step ahead prediction can be used as input iteratively to get multistep ahead time-series generation. It's obvious that GRU performs much better in both one-step and multi-step ahead predictions.

Table 10 Prediction performance – RMSE

Case	ARMAX	GRU
Training	4.02	2.21
Test	4.77	2.26
Generation	16.57	10.91

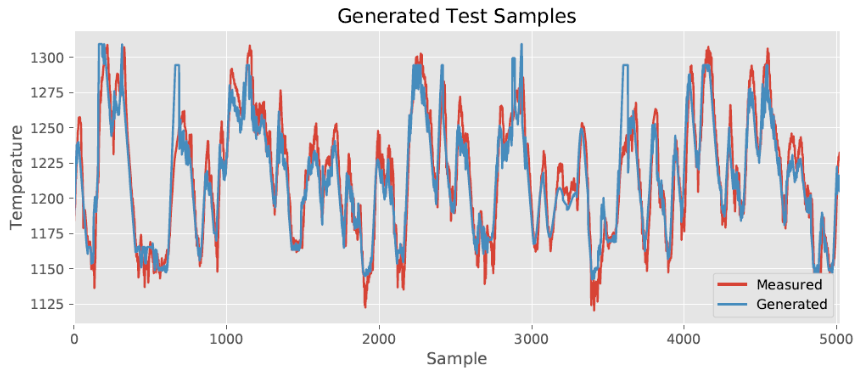


Figure 15 Generated time-series by ARMAX

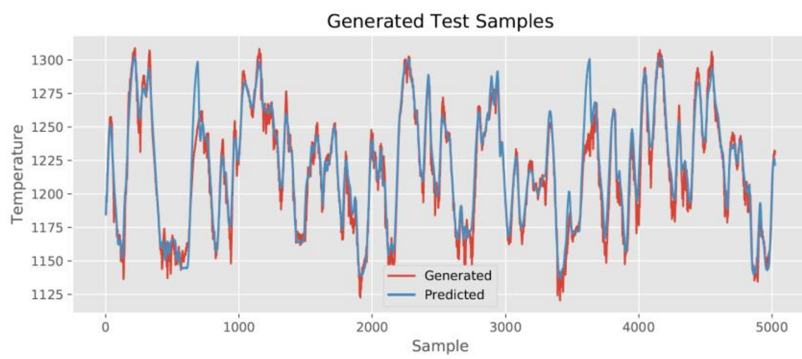


Figure 16 Generated time-series by GRU

Generative model

Introduction to generative model

The task of generative models is to generate new samples from the probabilistic distribution $p_D(\mathbf{x})$, defined over high-dimensional data $\mathbf{x}$ (image, text, audio, or other time-series sequences). Since the true probabilistic distribution $p_D(\mathbf{x})$ is unknown, a modeled probabilistic distribution $p_\theta(\mathbf{x})$ is proposed as a generative model. However, direction optimization over p_θ to approximate p_D is not feasible due to high dimensionality. Instead, a latent variable $\mathbf{z}$ with fixed prior distribution $p(\mathbf{z})$ is introduced and passed through a deep neural network to generate $\mathbf{x}$. Generative adversarial networks

(GANs) and variational autoencoders (VAEs) are two popular generative model approaches. GANs simultaneously train two models: a generative model that maps from a latent variable **z** to data **x**, and a discriminative model that discriminates between generated samples and training samples [Goodfellow et al., 2014]. On the other hand, VAEs are composed on of an probabilistic encoder that produces a simple distribution over latent variable **z** where training data point **x** could have been generated, and a probabilistic decoder that reconstruct **x** from **z** [Kingma and Welling, 2013].

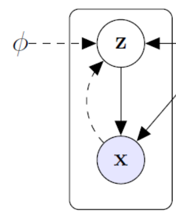


Figure 17 Solid lines denote the generative model $p_\theta(z)\ p_\theta(x|z)$, dashed lines denote the variational approximation $p_\phi(x|z)$. The variational parameters ϕ are learned jointly with the generative model parameters θ.

The generative process in GANs and VAEs cannot be explicitly controlled. Hence, conditional generative adversarial networks (CGANs) and conditional variational autoencoders (CVAEs) are developed to model latent variables **z** and data **x**, both conditioned on external input u [Mirza and Osindero, 2014, Sohn et al., 2015]. However, the external input in CGANs and CVAEs is limited to categorical data.

Dynamic generative model conditioned on external input

It is assumed that the observations (combustion zone temperature) $\mathbf{x} = [\mathbf{x}_1, \mathbf{x}_2, ..., \mathbf{x}_T]$ come from a latent state $\mathbf{z} = [\mathbf{z}_1, \mathbf{z}_2, ..., \mathbf{z}_T]$, whose evolvement over time is driven by external control input (temperature set point) $\mathbf{u} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_T]$. The initial latent state is $\mathbf{z}_0 = 0$. This system can be formulated as a general probabilistic state space model (SSM) as follows:

$$\begin{aligned} \mathrm{z}_t \sim & p_{\theta_z}(\mathrm{z}_t|\mathrm{u}_t, \mathrm{z}_{t-1}) \\ \mathrm{x}_t \sim & p_{\theta_x}(\mathrm{x}_t|\mathrm{z}_t) \end{aligned} \tag{15}$$

where $p_{\theta z}$ $(\mathbf{z}_t|\mathbf{u}_t, \mathbf{z}_{t-1})$ is the prior transition distribution and $p_{\theta x}(\mathbf{x}_t \mid \mathbf{z}_t)$ is the emission distribution. SSM can be used to generate temporally correlated sequence, but traditional SSM is not capable of modelling long sequence with complex structure. In parallel, this system can also be modelled by RNNs through a hidden state $\mathbf{h}_t$. Modern RNNs (e.g. LSTM and GRU) are capable of modelling long time-series through gated internal memory unit, but the deterministic hidden unit $\mathbf{h}_t$ can not be used to generate sequence.

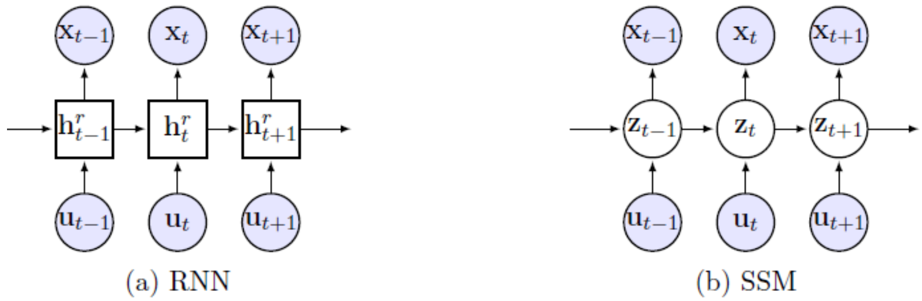


Figure 18 Graphical models to generate $\mathbf{x}_{1:T}$ with a recurrent neural network (RNN) and a state space model (SSM). Rectangle-shaped units are used for deterministic states, while circles are used for stochastic ones.

The joint probability of observations and latent states can be factorized as

$$
\begin{aligned}
p_\theta(\mathbf{x}_{1:t}, \mathbf{z}_{1:t}|\mathbf{u}_{1:t}) =& p_{\theta_x}(\mathbf{x}_{1:t}|\mathbf{z}_{1:t}) p_{\theta_z}(\mathbf{z}_{1:t}|\mathbf{u}_{1:t}) \\
=& \prod_{t=1}^{T} p_{\theta_x}(\mathbf{x}_t|\mathbf{z}_t) p_{\theta_z}(\mathbf{z}_t|\mathbf{u}_t, \mathbf{z}_{t-1})
\end{aligned}
\tag{16}
$$

Let the prior transition distribution $p_{\theta z} = N(\mu_{zt}\ ;\Sigma_{zt})$ be a Gaussian distribution whose mean and variance are parameterized by neural networks. In the emission distribution $p_{\theta x}(\mathbf{x}_t \mid \mathbf{z}_t)$, $\mathbf{x}_t$ depends on $\mathbf{z}_t$ through a neural network that is parameterized by θ_x. The generative model described by the joint density in Eq. 16 is parameterized by $\theta = [\theta_x, \theta_z]$. To generate samples that are merely like $\mathbf{x}$ through $p_{\theta x}(\mathbf{x}_t \mid \mathbf{z}_t)$, we have to sample values of $\mathbf{z}$ that are likely to produce $\mathbf{x}$. This means that we have to sample from the posterior distribution $p_\theta(\mathbf{z} \mid \mathbf{x}, \mathbf{u})$, which is computationally intractable. An approximate posterior distribution $q_\phi(\mathbf{z} \mid \mathbf{x}, \mathbf{u})$ is proposed to obtain the lower bound of the marginal likelihood

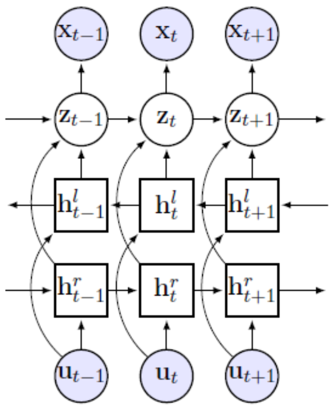
$$
\begin{aligned}
& \mathcal{D}_{KL}(q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) || p_\theta(\mathbf{z}|\mathbf{x}, \mathbf{u})) \\
=& \int q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) \log \frac{q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u})}{p_\theta(\mathbf{z}|\mathbf{x}, \mathbf{u})} dz \\
=& \mathbb{E}_{\mathbf{z}\sim q_\phi}[\log q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) - \log p_\theta(\mathbf{z}|\mathbf{x}, \mathbf{u})] \\
=& \mathbb{E}_{\mathbf{z}\sim q_\phi}[\log q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) - \log p_\theta(\mathbf{x}, \mathbf{z}|\mathbf{u})] + p_\theta(\mathbf{x}|\mathbf{u})
\end{aligned}
\tag{17}
$$

Negating both sides, rearranging, and contracting part of $\mathbf{E}_{\mathbf{z}\sim q\phi}$ into a KL-divergence terms yields:

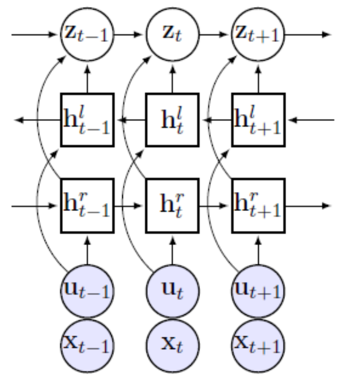
$$
\begin{aligned}
& \log p_\theta(\mathbf{x}|\mathbf{u}) - \mathcal{D}_{KL}(q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) || p_\theta(\mathbf{z}|\mathbf{x}, \mathbf{u})) \\
=& \mathbb{E}_{\mathbf{z}\sim q_\phi}[\log p_\theta(\mathbf{x}|\mathbf{z}, \mathbf{u})] - \mathcal{D}_{KL}[q_\phi(\mathbf{z}|\mathbf{x}, \mathbf{u}) || p_\theta(\mathbf{z}|\mathbf{u})] \\
=& \mathcal{L}(\theta, \phi)
\end{aligned}
\tag{18}
$$

where $\boldsymbol{L}(\theta, \phi)$ is the evidence lower bound (ELBO) because KL-divergence is non-negative. As such, maximizing the marginal likelihood $p_\theta(\mathbf{x}|\mathbf{u})$ is equivalent to maximizing ELBO $\boldsymbol{L}(\theta, \phi)$ w.r.t. variational parameters ϕ and generative parameters θ.

When $q_\phi(\mathbf{z}\,|\mathbf{x},\mathbf{u})$ and $p_\theta(\mathbf{z}|\mathbf{u})$ are Gaussian distributions, $\boldsymbol{D}_{KL}[q_\phi(\mathbf{z}|\mathbf{x},\mathbf{u})\;||\;\log p_\theta(\mathbf{z}|\mathbf{u})]$ can be computed analytically.



(a) Generative model p_θ



(b) Inference network q_ϕ

Figure 19 A dynamic generative model p_θ for a sequence $x_{1:T}$. Posterior inference of $z_{1:T}$ is done through an inference network q_ϕ, which uses a bidirectional-RNN to approximate the nonlinear dependence of z_t on future observations $x_{t:T}$ and states $u_{t:T}$

Algorithm 1 Dynamic generative model

Initialize parameters θ, ϕ
repeat
  Get random minibatch datapoints $\mathbf{x}, \mathbf{u}$
  Get Monte Carlo samples $\mathbf{z}^*$ from distribution $q_\phi(\mathbf{z}|\mathbf{x},\mathbf{u})$
  Evaluate $\mathbb{E}_{\mathbf{z}\sim q_\phi}[\log p_\theta(\mathbf{x}|\mathbf{z},\mathbf{u})]$ using $\mathbf{z}^*$
  Update parameters using gradients $\nabla_{\theta,\phi}\mathcal{L}$ (e.g. SGD)
until convergence of parameters θ, ϕ
return θ, ϕ

Numerical results

Some initial results from the proposed dynamic generative network are shown below.

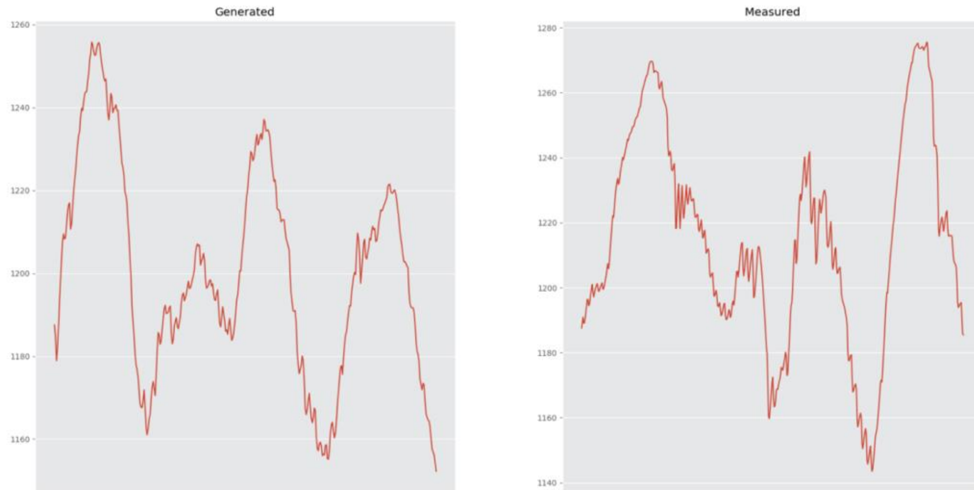


Figure 20 Generated time-series by dynamic generative model

References

K. Alam and A. El Saddik. C2ps: A digital twin architecture reference model for the cloudbased cyber-physical systems. *IEEE Access*, 5:2050–2062, 2017.

R. F. Babiceanu and R. Seker. Big data and virtualization for manufacturing cyber-physical systems: A survey of the current status and future outlook. *Computers in Industry*, 81: 128–137, 2016.

M. Bajaj, B. Cole, and D. Zwemer. Architecture to geometry-integrating system models with mechanical design. In *AIAA SPACE 2016*, page 5470. 2016.

G. E. Box and G. C. Tiao. Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association*, 70(349):70–79, 1975.

G. E. Box, G. M. Jenkins, G. C. Reinsel, and G. M. Ljung. *Time series analysis: forecasting and control*. John Wiley & Sons, 2015.

A. Canedo. Industrial iot lifecycle via digital twins. In *Proceedings of the Eleventh IEEE/ACM/IFIP International Conference on Hardware/Software Codesign and System Synthesis*, page 29. ACM, 2016.

A. Cerrone, J. Hochhalter, G. Heber, and A. Ingraffea. On the effects of modeling asmanufactured geometry: toward digital twin. *International Journal of Aerospace Engineering*, 2014, 2014.

K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, 1724–1734, 2014.

T. DebRoy, W. Zhang, J. Turner, and S. Babu. Building digital twins of 3d printing machines. *Scripta Materialia*, 135:119–124, 2017.

R. Drath and A. Horch. Industry 4.0: Hit or hype? *IEEE industrial electronics magazine*, 8(2):56–58, 2014.

K.-i. Funahashi and Y. Nakamura. Approximation of dynamical systems by continuous time recurrent neural networks. *Neural networks*, 6(6):801–806, 1993.

T. Gabor, L. Belzner, M. Kiermeier, M. T. Beck, and A. Neitz. A simulation-based architecture for smart cyber-physical systems. In *Autonomic Computing (ICAC), 2016 IEEE International Conference on*, pages 374–379. IEEE, 2016.

E. Glaessgen and D. Stargel. The digital twin paradigm for future nasa and us air force vehicles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, page 1818, 2012.

B. Gockel, A. Tudor, M. Brandyberry, R. Penmetsa, and E. Tuegel. Challenges with structural life forecasting using realistic mission profiles. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, page 1813, 2012.

I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, 2672–2680, 2014.

I. Goodfellow, Y. Bengio, A. Courville, and Y. Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016.

K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber. Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28(10): 2222–2232, 2017.

M. Grieves. *Virtually perfect: Driving innovative and lean products through product lifecycle management*. Space Coast Press, 2011.

M. Grieves and J. Vickers. Digital twin: Mitigating unpredictable, undesirable emergent behavior in complex systems. In *Transdisciplinary perspectives on complex systems*, pages 85–113. Springer, 2017.

S. Hochreiter and J. Schmidhuber. Long short-term memory. *Neural computation*, 9(8): 1735–1780, 1997.

K. J. Hunt, D. Sbarbaro, R. Zbikowski, and P. J. Gawthrop. Neural networks for control systems–a survey. *Automatica*, 28(6):1083–1112, 1992.

N. Jazdi. Cyber physical systems in the context of industry 4.0. In *Automation, Quality and Testing, Robotics, 2014 IEEE International Conference on*, pages 1–4. IEEE, 2014.

R. Jozefowicz, W. Zaremba, and I. Sutskever. An empirical exploration of recurrent network architectures. In *International Conference on Machine Learning*, pages 2342–2350, 2015.

D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

G. Knapp, T. Mukherjee, J. Zuback, H. Wei, T. Palmer, A. De, and T. DebRoy. Building blocks for a digital twin of additive manufacturing. *Acta Materialia*, 135:390–399, 2017.

E. M. Kraft. The air force digital thread/digital twin-life cycle integration and use of computational and experimental knowledge. In *54th AIAA Aerospace Sciences Meeting*, page 0897, 2016.

H. Lasi, P. Fettke, H. Kemper, T. Feld, and M. Hoffmann. Industry 4.0. *Business & Information Systems Engineering*, 6(4):239–242, 2014.

J. Lee, B. Bagheri, and H. Kao. A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters*, 3:18–23, 2015.

C. Li, S. Mahadevan, Y. Ling, S. Choze, and L. Wang. Dynamic bayesian network for aircraft wing health monitoring digital twin. *AIAA Journal*, 55(3):930–941, 2017.

M. Mirza and S. Osindero. Conditional generative adversarial nets. *arXiv preprint arXiv:1411.1784*, 2014.

K. Narandra and K. Parathasarathy. Identification and control of dynamic systems using neural networks. *IEEE Trans. on Neural Networks*, 1(1):4–27, 1990.

Q. Qi and F. Tao. Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison. *IEEE Access*, 6:3585–3593, 2018.

K. Reifsnider and P. Majumdar. Multiphysics stimulated simulation digital twin methods for fleet management. In *54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, page 1578, 2013.

J. Ríos, J. Hernández, M. Oliva, and F. Mas. Product avatar as digital counterpart of a physical individual product: Literature review and implications in an aircraft. In *ISPE CE*, pages 657–666, 2015.

R. Rosen, G. Von Wichert, G. Lo, and K. D. Bettenhausen. About the importance of autonomy and digital twins for the future of manufacturing. *IFAC-PapersOnLine*, 48(3): 567–572, 2015.

M. Schluse and J. Rossmann. From simulation to experimentable digital twins: simulationbased development and operation of complex technical systems. In *Systems Engineering (ISSE), 2016 IEEE International Symposium on*, pages 1–6. IEEE, 2016.

B. R. Seshadri and T. Krishnamurthy. Structural health management of damaged aircraft structures using digital twin concept. In *25th AIAA/AHS Adaptive Structures Conference*, page 1675, 2017.

K. Sohn, H. Lee, and X. Yan. Learning structured output representation using deep conditional generative models. In *Advances in neural information processing systems*, pages 3483–3491, 2015.

F. Tao and M. Zhang. Digital twin shop-floor: a new shop-floor paradigm towards smart manufacturing. *IEEE Access*, 5:20418–20427, 2017.

E. Tuegel. The airframe digital twin: some challenges to realization. In *53rd AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference 20th AIAA/ASME/AHS Adaptive Structures Conference 14th AIAA*, page 1812, 2012.

E. Tuegel, A. Ingraffea, T. Eason, and S. Spottswood. Reengineering aircraft structural life prediction using a digital twin. *International Journal of Aerospace Engineering*, 2011.

J. Wan, S. Tang, Z. Shu, D. Li, S. Wang, M. Imran, and A. V. Vasilakos. Software-defined industrial internet of things in the context of industry 4.0. *IEEE Sensors Journal*, 16(20): 7373–7380, 2016.

H. Zhang, Q. Liu, X. Chen, D. Zhang, and J. Leng. A digital twin-based approach for designing and multi-objective optimization of hollow glass production line. *IEEE Access*, 5:26901–26911, 2017.

K. Zhou, T. Liu, and L. Zhou. Industry 4.0: Towards future industrial opportunities and challenges. In *Fuzzy Systems and Knowledge Discovery (FSKD), 2015 12th International Conference on*, 2147–2152. IEEE, 2015.

Epsilon-BMC: a Bayesian model combination approach to Epsilon-greedy exploration in model-free reinforcement learning

Michael Gimelfarb

Introduction

Balancing exploration with exploitation is a well-known and important problem in reinforcement learning [Sutton and Barto, 2018]. If the behaviour policy focuses too much on exploration rather than exploitation, then this could hurt the performance in an on-line setting. Furthermore, on-policy algorithms such as SARSA or TD(λ) might not converge to a good policy. On the other hand, if the exploration policy focuses too much on exploitation rather than exploration, then the state space could not be explored sufficiently and an optimal policy would not be found.

Historically, numerous strategies have been proposed for addressing the exploration-exploitation trade-off specifically in model-free reinforcement learning, including Boltzmann exploration and epsilon-greedy [McFarlane,2018]. Epsilon greedy is formally defined as the following distribution over actions

$$\pi^{\varepsilon}(a|s) = \begin{cases} 1 - \varepsilon_t + \frac{\varepsilon_t}{|\mathcal{A}|} & \text{if } a = \arg\max_{a' \in \mathcal{A}} Q_t(s, a') \\ \frac{\varepsilon_t}{|\mathcal{A}|} & \text{otherwise} \end{cases}$$

The balance between exploration and exploitation is often controlled by one or more tuning parameters, such as ε in epsilon-greedy or the temperature in Boltzmann exploration. However, these parameters have to be tuned manually for each task in order to obtain good performance. This motivates the design of exploration algorithms that adapt according to some measure of the learning progress. For example, Tokic [2010] and dos Santos Mignon and da Rocha [2017] followed the simple epsilon-greedy policy, but the ε parameter was automatically tuned based on the Bellman error or similar quantification of learning progress. Tokic and Palm [2011] later built on this idea by combining it with Boltzmann exploration. Count-based methods were also introduced that used the state visit counts to guide exploration [Thrun, 1992, Bellemare et al., 2016, Ostrovski et al., 2017]. Dearden et al. [1998] followed a different approach to exploration by sampling from a posterior distribution over Q-values. However, despite recent developments in exploration strategies, epsilon-greedy is still often the exploration approach of choice [Vermorel and Mohri, 2005, Heidrich-Meisner, 2009, Mnih et al., 2015, Van Hasselt et al., 2016]. A benefit of epsilon-greedy exploration is that it can be easily combined with more sophisticated strategies, such as options [Bacon et al., 2017].

Unfortunately, the performance of epsilon-greedy in practice is highly sensitive to the choice of ε, and existing methods for adapting ε from data are ad-hoc and offer little theoretical justification. In this paper, we take a fully Bayesian perspective on adapting ε based on return data. Recent work has demonstrated the strong potential of a Bayesian approach for parameter tuning in model-free reinforcement learning [Downey and Sanner, 2010]. Another key advantage of a fully Bayesian approach over heuristics is the ability to specify priors on parameters, such as the predictive inverse variance of returns τ in this work, that are more robust to noise or temporary digressions in the learning process. In addition, our approach can be combined with other exploration policies such as Boltzmann exploration [Tokic and Palm, 2011].

A Bayesian interpretation of expected SARSA

The paper proceeds in the framework of Markov decision processes. In the paper, we first show that, for state $s' = s_{t+1}$, action $a^* = argmax_a Q_t(s', a)$ and reward $r' = r_{t+1}$, the expected return under the expected SARSA algorithm [Sutton and Barto, 2018] is

$$G_t = (1 - \varepsilon_t) G_t^Q + \varepsilon_t G_t^U \quad (1)$$

where G_t^Q is the standard Q-learning bootstrap, and

$$\tilde{G}_t^U = r_{t+1} + \gamma \frac{1}{|\mathcal{A}|} \sum_{a' \in \mathcal{A}} Q_t(s_{t+1}, a')$$

This leads to the following important observation. We can now view expected SARSA as a probability-weighted average of two models: the greedy model which trusts the current Q-value estimates and acts optimally with respect to them, and the uniform model which completely distrusts the current Q-value estimates and consequently places a uniform belief over them. Now, the problem of adapting ε from data in the expected SARSA framework can be addressed using Bayesian ensemble learning. Under this interpretation, ε and $1 - \varepsilon$ are the posterior beliefs assigned to the two aforementioned models, respectively.

Bayesian Q-learning

In order to facilitate tractable learning and inference, we assume that the return observation $q_{s,a}$ at time t, given the model, is independently normally distributed:

$$
\begin{aligned}
q_{s,a}|Q,\tau &\sim \mathcal{N}\left(\tilde{G}_t^Q, \tau^{-1}\right), \\
q_{s,a}|U,\tau &\sim \mathcal{N}\left(\tilde{G}_t^U, \tau^{-1}\right),
\end{aligned}
$$

In order to update τ, we considered a Normal-Gamma prior

$$
\begin{aligned}
q_{s,a}|\mu,\tau &\sim \mathcal{N}\left(\mu, \tau^{-1}\right), \\
\mu,\tau &\sim \text{NormalGamma}\left(\mu_0, \tau_0, a_0, b_0\right),
\end{aligned}
$$

where $q_{s,a}$ are i.i.d. given τ. Given data $\mathcal{D}$ of previously observed returns, the joint of τ and μ is also Normal-Gamma, and the marginal posterior of τ is

$$
\begin{aligned}
a_t &= a_0 + \frac{t}{2}, \\
b_t &= b_0 + \frac{t}{2}\left(\hat{\sigma}_t^2 + \frac{\tau_0}{\tau_0 + t}(\hat{\mu}_t - \mu_0)^2\right),
\end{aligned}
\tag{9}
$$

where $\hat{\mu}$ and $\hat{\sigma}$ are the sample mean and standard deviation of the return data [Bishop, 2006]. These quantities can be updated online after each new observation, in constant time [Welford, 1962]

$$
\begin{aligned}
\hat{\mu}_{t+1} &= \hat{\mu}_t + \frac{d' - \hat{\mu}_t}{t+1}, \\
M_{t+1} &= M_t + (d' - \hat{\mu}_t)(d' - \hat{\mu}_{t+1}), \\
\hat{\sigma}_{t+1}^2 &= \frac{M_{t+1}}{t},
\end{aligned}
\tag{10}
$$

Finally, by marginalizing out the (posterior) uncertainty in τ, we were able to show that the returns, given the model, are student T distributed

$$
\begin{aligned}
q_{s,a}|Q,\mathcal{D} &\sim \text{St}\left(2a_t, \tilde{G}_t^Q, \frac{a_t}{b_t}\right), \\
q_{s,a}|U,\mathcal{D} &\sim \text{St}\left(2a_t, \tilde{G}_t^U, \frac{a_t}{b_t}\right),
\end{aligned}
\tag{11}
$$

We now show how to use this likelihood function to adapt epsilon given the expected SARSA decomposition (1).

Epsilon-BMC: an adaptive epsilon greedy algorithm

The expected posterior return can be written as an average over all possible combinations of greedy and uniform model

$$\mathbb{E}[q_{s,a}|\mathcal{D}] = \int_0^1 \mathbb{E}[q_{s,a}|w]\, \mathbb{P}(w|\mathcal{D})\, \mathrm{d}w,$$

where w is the weight assigned to the uniform model (and so $1 - w$ is the weight assigned to the greedy model). By performing this integral exactly, and simplifying the algebra, we showed that this expectation was equivalent to

$$(1 - \mathbb{E}[w|\mathcal{D}])\, \tilde{G}_t^Q + \mathbb{E}[w|\mathcal{D}]\tilde{G}_t^U.$$

So we have shown that $\varepsilon_t = E[w|\mathcal{D}]$, e.g. the data-dependent epsilon parameter should be set to the expectation under the posterior of w.

Unfortunately, the posterior distribution of w is intractable. However, a good conjugate prior argued in the paper is the Beta distribution. To find the optimal values of the Beta distribution (α, β), we apply the Dirichlet moment-matching technique, first presented in [Hsu and Poupart, 2016], and further elaborated in [Gimelfarb et. al., 2018] to handle the multiple experts case. Applying the approach in this paper, in which the first two moments are matched, we obtain an analytic expression for the Beta parameters

$$m_t = \frac{\alpha_t}{\alpha_t + \beta_t + 1} \frac{e_t^U(\alpha_t + 1) + e_t^Q \beta_t}{e_t^U \alpha_t + e_t^Q \beta_t}, \tag{13}$$

$$v_t = \frac{\alpha_t}{\alpha_t + \beta_t + 1} \frac{\alpha_t + 1}{\alpha_t + \beta_t + 2} \frac{e_t^U(\alpha_t + 2) + e_t^Q \beta_t}{e_t^U \alpha_t + e_t^Q \beta_t}, \tag{14}$$

$$r_t = \frac{m_t - v_t}{v_t - m_t^2}, \tag{15}$$

$$\alpha_{t+1} = m_t r_t, \tag{16}$$

$$\beta_{t+1} = (1 - m_t) r_t, \tag{17}$$

Under this approximation, the optimal epsilon parameter value in a data-driven sense is just the expected value of a Beta distribution, which is

$$\varepsilon_t^{BMC} \approx \mathbb{E}_{\mathrm{Beta}(\alpha_t, \beta_t)}[w|\mathcal{D}] = \frac{\alpha_t}{\alpha_t + \beta_t}.$$

The complete algorithm is given in the following pseudocode as Algorithm 1.

Finally, we were able to prove the following theorem for our algorithm, which is still valid even with function approximation for the Q-value function (e.g. deep reinforcement learning).

Theorem 1: Suppose $0 < \alpha_0 \leq \beta_0 < \infty$. Then $\varepsilon_{t+1}^{BMC} \leq \varepsilon_t^{BMC}$ for all $t = 1,2, \ldots$ By the monotone convergence theorem [Rudin, 1976, page 56], ε_t^{BMC} converges as $t \to \infty$.

The short and simple proof is provided in the paper.

Algorithm 1 $\varepsilon-$BMC with Expected SARSA

* initialize $\mu_0, \tau_0, a_0, b_0, \hat{\mu}, M, \hat{\sigma}^2, \alpha, \beta$
for each episode **do**
  initialize s
  for each step in the episode **do**
    * $\varepsilon \leftarrow \frac{\alpha}{\alpha+\beta}$
    choose action a using ε-greedy policy π^{ε}
    take action a, observe r and s'
    $\tilde{G}^Q \leftarrow r + \gamma \max_{a'} Q(s', a')$
    $\tilde{G}^U \leftarrow r + \gamma \frac{1}{|\mathcal{A}|} \sum_{a'} Q(s', a')$
    $\tilde{G}^{ExpS} \leftarrow r + \gamma \sum_{a'} \pi^{\varepsilon}(a'|s') Q(s', a')$ {note $\tilde{G}^{ExpS} = (1 - \varepsilon)\tilde{G}^Q + \varepsilon \tilde{G}^U$}
    $Q(s, a) \leftarrow Q(s, a) + \eta[\tilde{G}^{ExpS} - Q(s, a)]$
    * update $\hat{\mu}$ and $\hat{\sigma}^2$ using (10)
    * compute a and b using (9)
    * compute e^Q and e^U using (11)
    * update α and β using (13)-(17)
    $s \leftarrow s'$
  end for
end for

Experiments

We applied this algorithm two three different problems: (1) a discrete grid-world environment with sub-goals [Ng et. al., 1999] the classical control problem called inverted pendulum/cart-pole, and a noisy supply chain/inventory management problem [Kemmer et al., 2018]. We applied this approach with two different reinforcement learning algorithms: (1) the tabular/exact expected SARSA algorithm (as presented in the pseudocode), and using the off-policy deep reinforcement learning algorithm with experience replay (DQN) [Mnih et. al., 2015]. In order to compare how well our algorithm adapts the epsilon parameter, we set the following families of epsilon adaptation schemes as base-lines

- **constant:** $\varepsilon_t = c$, where $c \in \{0.01, 0.05, 0.1, 0.25, 0.5\}$;
- **geometric:** $\varepsilon_t = \frac{1}{2}\rho^t$, where $\rho \in \{0.85, 0.9, 0.95, 0.975, 0.99\}$ and t is the episode number;
- **inverse:** $\varepsilon_t = \frac{1}{2}(t+1)^{-\beta}$, where $\beta \in \{0.25, 0.5, 1.0, 1.5\}$ and t is the episode number;
- **adaptive:** VDBE [Tokic, 2010] with $\varepsilon_0 = 0.5$, $\delta = 1/|\mathcal{A}|$, and $\sigma \in \{0.01, 0.05, 0.1, 0.5, 1.0, 10.0, 100.0\}$.

The remaining details are provided in the paper. The results are presented below for the supply chain problem. We have omitted the other experiments, but they will be published along with our paper later this year.

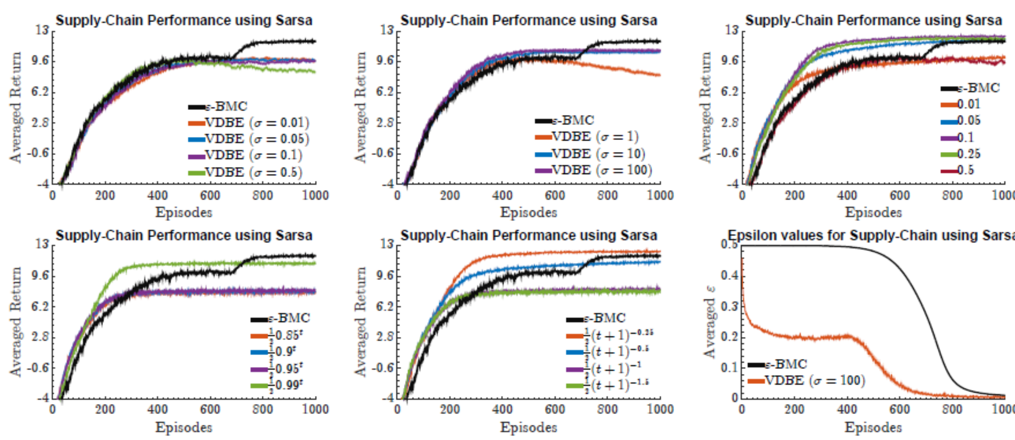


Figure 5: Average performance (return) on the supply-chain domain using expected SARSA.

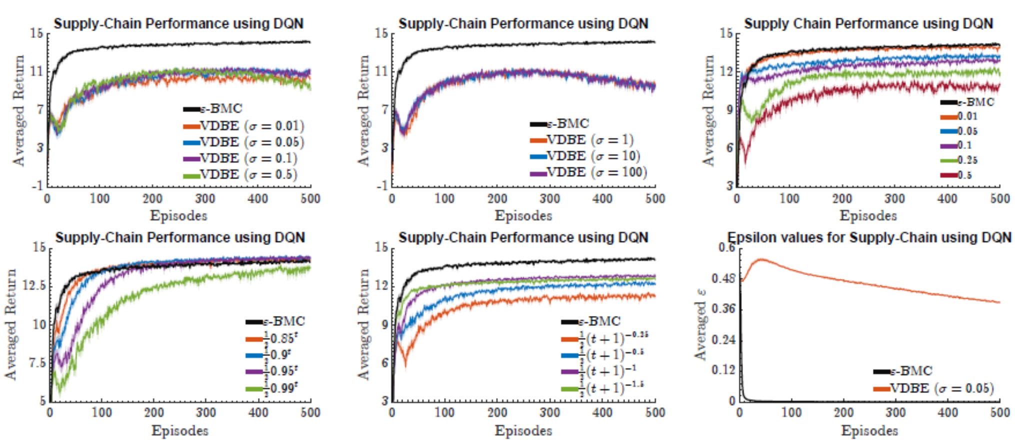


Figure 6: Average performance (return) on the supply-chain domain using deep Q-learning.

Overall, we see that epsilon-BMC consistently outperformed all other types of annealing strategies, including VDBE, or performed similarly. However, epsilon-BMC converged slightly later than VDBE on the grid-world domain and the fixed annealing

$\varepsilon_t = \frac{1}{2}(t+1)^{-0.25}$ on the supply chain problem, using tabular expected SARSA. However, in the former case, epsilon-BMC outperformed all fixed tuning strategies, and in the latter case, it outperformed VDBE by a large margin. These observations are related to the speed of convergence; asymptotically, epsilon-BMC approached the performance of the best policy that was attained (for grid-world this is indeed the optimal policy). While it performed well on the simple grid-world domain, VDBE performed considerably worse than epsilon-BMC on the more complex supply-chain problem. We believe that the Bayesian approach of epsilon-BMC smooths out the noise in the return signals better than VDBE and other ad-hoc approaches for adapting epsilon. This also suggests why our algorithm performed better with value function approximation than other methods.

Furthermore, we see that no single family of annealing strategies worked consistently well across all domains and algorithms. For instance, geometric decay strategies worked well on the grid-world domain, while performing poorly on the supply-chain problem using tabular SARSA. The inverse decay strategies worked well on the supply-chain problem using tabular SARSA, but failing to match the performance of other strategies when switching to DQN. Also, the performance of VDBE was highly sensitive to the choice of the sigma parameter. A lower value of sigma worked well for grid-world and cart-pole, but higher values of sigma worked better for supply-chain. The performance of epsilon-BMC was relatively insensitive to the choice of prior parameters for mu and tau, so we were able to use the same values in all our experiments. However, unsurprisingly, it was more sensitive to the strength of the prior on ε (α_0, β_0). Since we can always set them to be equal, this effectively reduces to the problem of selecting a single parameter that controls the strength of the prior on epsilon. This is considerably easier to do than to select both a good annealing strategy and the parameter(s) that control the speed of convergence.

Conclusion

In this paper, we proposed a novel Bayesian approach to solve the exploration-exploitation problem in general model-free reinforcement learning, in the form of an adaptive epsilon-greedy policy. Our novel algorithm, epsilon-BMC, is a novel approach for tuning the epsilon parameter automatically from return observations based on Bayesian model combination and approximate moment-matching based inference. It was argued to be general, efficient, robust, and theoretically grounded, and was shown empirically to outperform fixed annealing schedules for epsilon and even a state-of-the-art heuristic adaptation scheme proposed in the literature.

References

P.-L. Bacon, J. Harb, and D. Precup. The option-critic architecture. In Thirty-First AAAI Conference on Artificial Intelligence, 2017.

M. Bellemare, S. Srinivasan, G. Ostrovski, T. Schaul, D. Saxton, and R. Munos. Unifying count-based exploration and intrinsic motivation. In Advances in Neural Information Processing Systems, pages 1471–1479, 2016.

C. M. Bishop. Pattern Recognition and Machine Learning. Springer, 2006.

R. Dearden, N. Friedman, and S. Russell. Bayesian q-learning. In AAAI/IAAI, pages 761–768, 1998.

A. dos Santos Mignon and R. L. d. A. da Rocha. An adaptive implementation of epsilon-greedy in reinforcement learning. Procedia Computer Science, 109:1146–1151, 2017.

C. Downey and S. Sanner. Temporal difference Bayesian model averaging: A bayesian perspective on adapting lambda. In International Conference on Machine Learning, pages 311–318, 2010.
M. Gimelfarb, S. Sanner, and C.-G. Lee. Reinforcement learning with multiple experts: A bayesian model combination approach. In Advances in Neural Information Processing Systems, pages 9549–9559. Curran Associates, 2018.
V. Heidrich-Meisner. Interview with richard s. sutton. KI, 23(3):41–43, 2009.
W.-S. Hsu and P. Poupart. Online bayesian moment matching for topic modeling with unknown number of topics. In Advances In Neural Information Processing Systems, pages 4536–4544, 2016.
L. Kemmer, H. von Kleist, D. de Rochebouet, N. Tziortziotis, and J. Read. Reinforcement learning for supply chain optimization. In European Workshop on Reinforcement Learning 14, 10 2018.
R. McFarlane. A survey of exploration strategies in reinforcement learning. McGill University, http://www.cs.mcgill.ca/ cs526/roger.pdf, 2018. Unpublished.
T. P. Minka. Bayesian model averaging is not model combination. Technical report, 2002.
V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, et al. Human-level control through deep reinforcement learning. Nature, 518 (7540):529, 2015.
K. Monteith, J. L. Carroll, K. Seppi, and T. Martinez. Turning bayesian model averaging into Bayesian model combination. In Neural Networks (IJCNN), pages 2657–2663, 2011.
A. Y. Ng, D. Harada, and S. Russell. Policy invariance under reward transformations: Theory and application to reward shaping. In International Conference on Machine Learning, volume 99, pages 278–287, 1999.
G. Ostrovski, M. G. Bellemare, A. van den Oord, and R. Munos. Count-based exploration with neural density models. In Proceedings of the 34th International Conference on Machine Learning-Volume 70, pages 2721–2730, 2017.
M. L. Puterman. Markov decision processes: discrete stochastic dynamic programming. JohnWiley & Sons, 2014.
W. Rudin. Principles of mathematical analysis. McGraw-Hill Book Co., New York, third edition, 1976. ISBN 0-07-085613-3. International Series in Pure and Applied Mathematics.
R. S. Sutton and A. G. Barto. Reinforcement learning: An introduction. MIT press, 2018.
S. B. Thrun. Efficient exploration in reinforcement learning. Technical report, 1992.
A. D. Tijsma, M. M. Drugan, and M. A. Wiering. Comparing exploration strategies for q-learning in random stochastic mazes. In 2016 IEEE Symposium Series on Computational Intelligence (SSCI), pages 1–8, 2016.
M. Tokic. Adaptive "-greedy exploration in reinforcement learning based on value differences. In Annual Conference on Artificial Intelligence, pages 203–210, 2010.
M. Tokic and G. Palm. Value-difference based exploration: adaptive control between epsilon-greedy and softmax. In Annual Conference on Artificial Intelligence, pages 335–346, 2011.
H. Van Hasselt, A. Guez, and D. Silver. Deep reinforcement learning with double q-learning. In Thirtieth AAAI Conference on Artificial Intelligence, 2016.
H. Van Seijen, H. Van Hasselt, S. Whiteson, and M. Wiering. A theoretical and empirical analysis of expected sarsa. In Adaptive Dynamic Programming and Reinforcement Learning, 2009. ADPRL'09. IEEE Symposium on Adaptive Dynamic Programming and Reinforcement Learning, pages 177–184, 2009.
J. Vermorel and M. Mohri. Multi-armed bandit algorithms and empirical evaluation. In European conference on machine learning, pages 437–448. Springer, 2005.
B. Welford. Note on a method for calculating corrected sums of squares and products. Technometrics, 4(3): 419–420, 1962.

Ongoing research on the remaining useful life prediction for high voltage circuit breakers

Gaoyang Li

Introduction

Acting as the vital device to cut off the short-circuit currents and isolate the faulty parts, the high voltage circuit breaker plays an important role in the relaying protection system of the power grid against various faults. Nowadays, the long-lived mechanical maintenance policies adopting a scheduled maintenance scheme are clearly sub-optimal compared to the methodologies based on prognostics and health management (PHM), which relies on an estimation of the health status to trigger the maintenances to avoid needless interventions. In the PHM, it is a key step to quantify the remaining useful life (RUL) of the mechanical systems based on the condition monitoring data.

Reviewing the literature devoted to the RUL estimation, the existing prognostics research can generally be classified into three categories: 1) model-based, 2) data-driven, and 3) hybrid-based. The model-based approaches are dependent on the physical models of specified mathematical expressions to describe the degradation of the system, of which the Wiener process, Gamma process, and inverse Gaussian process model are popular choices. With the analytic methods or approximate solutions such as Particle filter and Kalman filter, it is possible to update the parameters' posterior distributions online with sequential Bayesian methods. The data-driven models, on the contrary, mostly explore the run-to-failure data from field or laboratory experimentation with some machine learning models. Successful applications can be found in literature with Neural Network, Support Vector Machine (SVM). Finally, the hybrid approaches attempt to combine model-based and data-driven techniques to maximizing the existing information considering the dilemma between the availability of data and the reliability of prior knowledge. Although the model-based methods can be quite accurate with a detailed knowledge of the failure mechanism, the data-driven approaches have gained a wide diffusion for the complex engineered systems where the analytical model is laborious or even impossible. In particular, the deep learning methods have attracted tremendous attention recently as an active sub-field for the data-driven estimation area owing to the highly complex non-linear structure. There are, but not limited to, several types of deep structures including Recurrent Neural Network(RNN), Restricted Boltzmann machine(RBM), convolutional neural networks(CNN) that have been successfully applied in the PHM domain and outperformed the traditional prognosis algorithms.

However, one major divergence among these methods is that, while it is quite natural to estimate the uncertainty and variability of the statistical-based or model-based methods, most of the deep neural network based approaches can only achieve point estimation of RUL and are assumed to be accurate. In fact, In order to provide a meaningful RUL-based decision-making auxiliary, it is important to analyze not only the deterministic but also the prediction interval due to the presence of unknown factors, indicating to what degree can we trust the results. Besides, another drawbacks of the application of deep learning in the RUL estimation, which is often ignored, is the challenges in utilizing censored data in the model training.

Bayesian neural networks (BNNs) recently emerged as a principled framework to estimate either the aleatoric uncertainty or epistemic uncertainty. The BNN replaces the deterministic parameters with some distributions and optimizes the posterior approximation of the model parameters. Therefore, the present paper focuses attention on reasoning about both the epistemic uncertainty and the heteroscedastic aleatoric uncertainty in the BNN-based RUL estimation methods. The run-to-failure data from the operating mechanism of high voltage circuit breakers are collected to validate the effectiveness of the proposed method.

Model description

The two kinds of uncertainties, epistemic uncertainty and aleatoric uncertainty, can be considered as either putting a prior distribution on the model parameters or on the outputs. Therefore, it quite straight to distinguish what kinds of uncertainties are included in the prognosis model based on the basic setups. Accordingly, this section presents an overview of how the model-based and data-driven methods treat the uncertainty differently, and what kinds of uncertainties have been considered in the existing models.

Uncertainty in the model-based and hybrid-based methods

In model-based methods, the physical mechanism of equipment degradation is often described in the form of some stochastic processes, such as the Wiener process, Gamma process, or inverse Gaussian process. These models usually assumed a process with certain independent increment distribution, and the RUL is defined based on the distance between the cumulative increments and a preset threshold. Therefore, only the aleatoric uncertainty is considered in such models indicated by the increment distributions. The recent progress of model-based methods also tried to incorporate the epistemic uncertainty as well by putting a prior distribution on the drift parameters, offering a more comprehensive uncertainty estimation. For example, Zhai utilizes another Brownian motion for the drift parameter. The similar idea is also shared by Si. Taking advantage of the relatively simplified structure, the noise degrees of both the drift parameter and output can be estimated by the Maximum Likelihood Estimation (MLE) method.

However, the assumption that the degradation follows some preset path restricts their applications. Therefore, the hybrid approaches attempt to combine the data-driven techniques with the model-based methods by either replace the degradation model or measurement model, or both, with data-driven methods. In this kind of sequential Bayesian framework, the noise parameters in the degradation model and the measurement model can be considered as epistemic uncertainty and aleatoric uncertainty, respectively. However, although the noise levels have a big influence on the updating results, they are usually considered as hyperparameters. Only a few work attempts to deal with the practical challenge of noise level estimation.

Deep learning in RUL

Recently, deep learning has emerged as a potential alternative to process the highly non-linear and varying sequential sensor data in RUL estimation. For instance, Liao proposed an enhanced RBM to automatically generate features in RUL prediction. Li extracted multi-scale features of bearing vibrations using CNN. Besides, Zhao proposed local feature-based GRU networks to realize automatic feature learning for machine health monitoring. Ellefse applied a semi-supervised LSTM in turbofan engine degradation analysis. However, uncertainty is not the primary concern in the mentioned methods above, which reflect the essential divergence between the generating models and discriminant models.

To fill this gap, some literature has tried to integrate some external remedies into deep learning based methods for uncertainty quantification. An ensemble method is presented to estimate the aleatoric uncertainty, while the epistemic uncertainty is gained from a separated dataset. However, the ensemble learning is quite computationally-expensive and the separation operation is a waste of data.

Bayesian neural network

Different from the external approaches such as resemble learning, the BNN is a fully Bayesian method by replacing the deterministic weights with distributions. BNN is easy to formulate but difficult to conduct inference. Early studies relay on some approximation approaches such as Laplace approximation, Monte Carlo Markov Chain (MCMC), and variational inference. In the deep learning era, Blundell introduced Bayes by Backprop to realize variational inference in DNN, based on which a lot of new inference techniques have been proposed.

Today, deep learning algorithms are able to extract high-level representations automatically and have had tremendous success in research fields such as biology, physics, and prognosis. Many of the applications achieved state-of-the-art performance, with most of them not able to output the confidence interval. However, in reliability analysis, the uncertainty of the prognosis is quite critical in resource management and maintenance planning. A meaningful RUL prediction with uncertainty will provide not only the expectation of the mean RUL but also the probability that the machine will fail immediately, which is quite important for certain research objects of high reliability requirements such as electrical system and medical research.

From the perspective of Bayesian theory, the uncertainty of the prediction model derives from the prior probability and the likelihood function from the statistical model for the observation data.

$$p(\theta \mid D) = \frac{p(D \mid \theta)\, p(\theta)}{p(D)}$$

where θ indicate the model parameters and D are the observed data.

The key idea to generate uncertainty is to specify the prior parameter of the model parameters, indicating our belief about the hypothesis space. Most of the prognostic models that are able to quantitate the model uncertainty are based on this basic theory. In addition to offering uncertainty, another advantage of Bayesian inferences is the robustness to overfitting owing to the prior knowledge.

If a deep leaning model $f_\omega(X)$ parameterized by the collection of model parameters ω is utilized in this framework, a prior distribution must be assigned on ω. For example, if a Gaussian prior is assumed arbitrarily,

$$\omega \sim N(0,1)$$

with the observations as $X = \{x_1, x_2, \ldots, x_n\}$ and their corresponding labels as $Y = \{y_1, y_2, \ldots, y_n\}$, the posterior distribution of over the model parameters ω can be inferred from Eq.

$$p(\omega \mid X, Y) = \frac{p(Y \mid f_\omega(X))\, p(\omega)}{\int_\omega p(Y \mid f_\omega(X))\, p(\omega)\, d\omega}$$

Given a new observation x^*, the new prediction distribution is obtained by marginalizing the posterior distribution of ω

$$p(y^* \mid x^*) = \int_\omega p(y^* \mid f_\omega(x^*))\, p(\omega \mid X, Y) d\omega$$

According to the law of total variance, if X and Y are random variables, the variance of the prediction Y can be decomposed as unexplained and explained components

$$Var(Y) = E\left[Var(Y \mid X)\right] + Var\left[E(Y \mid X)\right]$$

The first term reflects the ignorance of the model parameters, referred to as the model uncertainty, is also known as the Epistemic uncertainty. The second term, aleatoric uncertainty, comes from the noise in the likelihood function, which is expected to be explained away with more data.

Neural network structure

Different neural networks can be adjusted to fit for such a framework. In practice, we use RNN to transform the input x of multiple time steps, with its output split to predict the parameters of the parametric survival distribution. In principle, RNN is more suitable for capturing relations in sequential data and have been applied extensively in as speech recognition, natural language translation, and video recognition. The so-called plain RNN has a hidden state h_i that is recurrently updated by

$$h_i = \sigma\left(Wx_i + Uh_{i-1} + b\right)$$

where the hidden state $h_{i\text{-}1}$ is a d-dimensional vector, and b is the bias. W and U are the weights. Tanh is the activation function. However, one particular drawback of RNN is the difficulty to capture long-term dependencies due to the vanishing or exploding gradients. To alleviate the problem, multiple variants have been proposed, of which the LSTM and GRU are successful ones. Although both are based on gating mechanisms, the GRU is claimed to be computationally cheaper compared to LSTM while maintaining similar performance. The reason lies mainly in that the GRU reduces the three gates from the LSTM to two, retaining an update gate z_i and a reset r_i only, thus having fewer parameters and calling for fewer data to generalize. Generally speaking, the GRU is presented in the following form:

$$h_i = (1 - z_i) \cdot h_{i-1} + z_i \cdot \widetilde{h_\iota}$$
$$\widetilde{h_\iota} = g(W_h x_i + U_h(r_i \cdot h_{i-1}) + b_h)$$

with the update gate and reset gate presented as

$$z_i = \sigma\left(W_z x_i + U_z h_{i-1} + b_z\right)$$
$$r_i = \sigma\left(W_r x_i + U_r h_{i-1} + b_r\right)$$

where W_z W_z W_r U_r are the model weights, and b_z b_r are the bias. In this formulation, the reset gate allows the hidden state ht to drop any information that is irrelevant to the future task, and the update gate controls the how much the previous hidden state is integrated with current observation in forming the current hidden state. Intuitively, the reset gate is more relevant to capturing short-term dependencies, while the update gate is more active when longer-term dependencies show more significance. By realizing forgetting and selectively memorizing in one step, as indicated in (), GRU saves one gating signal and the associated parameters.

In this paper, the GRU is utilized for fusing the information of multiple steps when the continuous monitoring data are available. Specifically, the GRU is further transformed into a Bayesian GRU by integrating the dropout variational inference technique into the original GRU model.

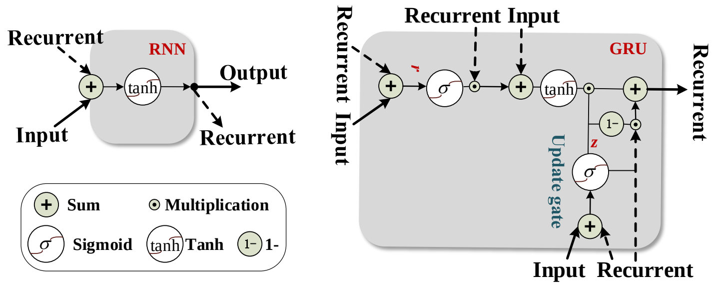


Figure 21 Architectures of a general RNN and GRU

Summary

We presented a novel Bayesian deep learning based framework to reason about the uncertainty in the RUL prediction. First, a dropout BNN is applied by assigning the prior distributions on every weight parameters of a GRU Neural Network to capture our ignorance about the model parameters, which is also known as the epistemic uncertainty.

The approach was assessed with the end-to-life data belonging to several hydraulic spring operating mechanisms in high voltage circuit breakers. The direct prediction results clearly show the superiority of the proposed methods, especially the Weibull BNN, over the traditional GRU with RMSE loss and several traditional probabilistic machine learning models with a higher point estimation accuracy and a better balance between the confidence interval width and PCIP. In the future, the ambition is to measure the uncertainty more comprehensively by incorporating the model structure uncertainty into the Bayesian framework as well.

References

Y. Liao, L. Zhang, and C. Liu, "Uncertainty prediction of remaining useful life using long short-term memory network based on bootstrap method," in *2018 IEEE International Conference on Prognostics and Health Management (ICPHM)*, 2018, pp. 1-8: IEEE.

X.-S. Si, W. Wang, C.-H. Hu, and D.-H. Zhou, "Remaining useful life estimation–A review on the statistical data driven approaches," *European journal of operational research,* vol. 213, no. 1, pp. 1-14, 2011.

Q. Zhai and Z.-S. Ye, "RUL prediction of deteriorating products using an adaptive Wiener process model," *IEEE Transactions on Industrial Informatics,* vol. 13, no. 6, pp. 2911-2921, 2017.

X.-S. Si, "An adaptive prognostic approach via nonlinear degradation modeling: Application to battery data," *IEEE Transactions on Industrial Electronics,* vol. 62, no. 8, pp. 5082-5096, 2015.

L. Liao and F. Köttig, "Review of hybrid prognostics approaches for remaining useful life prediction of engineered systems, and an application to battery life prediction," *IEEE Transactions on Reliability,* vol. 63, no. 1, pp. 191-207, 2014.

P. Baraldi, M. Compare, S. Sauco, and E. Zio, "Ensemble neural network-based particle filtering for prognostics," *Mechanical Systems and Signal Processing,* vol. 41, no. 1, pp. 288-300, 2013.

L. Liao, W. Jin, and R. Pavel, "Enhanced Restricted Boltzmann Machine with Prognosability Regularization for Prognostics and Health Assessment," vol. 63, no. 11, pp. 1-1, 2016.

Toronto Hydro Corporation: investment spike smoothing and asset risk growths

Gary Wang

Overview

In the C-MORE consulting project with Toronto Hydro, the objective is to address investment spike issues when a large proportion of assets come due for replacement simultaneously. Differences in asset risk growths are also major concerns when developing replacement schedules.

As an extension to our last report, this report will present the problem and objectives, and present a brief literature review. Three mathematical solutions to this problem will be presented in this report. Each of the three mathematical solutions have their unique characteristics, and are designed to solve the investment spike problems based on different investment targets, asset demographics and number of data entries. The results and recommendations on how to approach such problem are also included in this report.

Introduction: background, problem, and objectives

Toronto Hydro has a large variety of physical assets in the distribution system; each type of asset has a large and dynamic population. The company has a replacement program(s) for each category of assets. All assets are subject to replacements, reactively and/or proactively, due to failures, failure risks, aging, upgrades, legal obligations, etc. Assets of the same type often have similar useful-lives. Historically, large projects/initiatives have installed assets within a short period of time (or in a cyclical nature of a period of years), which can contribute to spikes in renewal expenditures in the future. Due to capital budget constraints, not all assets can be replaced at or before the optimal/expected replacement time, which further contributes to the spikes in renewal expenditures. To smooth out the investment spikes, assets have to be replaced earlier or later than the current expected replacement time. If asset is replaced early, asset may not have been fully depreciated, values within the asset has not been used to its fullest; if asset is replaced later than expected time, the asset carries additional risk, which may cause catastrophic failures in the distribution system, imposing additional costs to the company. Asset risk growths are also a major concern of this project. If two assets have different risk growths, even if the asset with the greater risk growth has a lower risk cost in the current period, it may be ideal to replace the asset with greater risk growth first for long-term outlooks.

This project will focus on creating an optimization approach to smooth out investment spikes with reasonable economic efficiency. The model should allow analysts and planners to setup constraints that are unique to a specific project/program. The types of constraints should include, but not limited to: capital budgeting constraints, replacement quantity constraints, and investment period constraints. For more details on the proposed constraints, please review section 6.

In this problem, we will ignore the effect of interest rates (which can be easily implemented in the mathematical approaches presented in later sections) to eliminate noises from other factors, and focus on investment spike smoothing and asset risk growths.

Hypothesis

Toronto Hydro's investment smoothing problem can be solved as an optimization problem with a set of constraints on the difference in budgeting within the investment horizon. Other constraints, such as, replacement amounts and risk levels can also be imposed in the optimization model, if needed. Asset risk growth plays a major role in long-term investment planning.

Literature review

Maintenance resource planning for utility poles in a power distribution network [1]

A recent report from C-MORE, "Maintenance resource planning for utility poles in a power distribution network," discusses prevention of unexpected increase in demand and resources, optimizing through the use of delayed renewal process [1]. The paper provides insights into preventative planning and risk management in the utility industry. Preventative planning takes on a similar concept as proactive replacement planning where certain proactive and pre-cautionary actions are performed to minimize the risk/cost of outages. The idea of risk management also coincides with the given research objective. The company does proactive replacement planning on a risk basis, prioritizing risky assets over assets that are unlikely to fail in a given period of time.

Equipment-replacement problem [2]

An article published by the University of Texas at Dallas (UTDallas) investigates an equipment replacement problem. The article breaks dynamic programming into four general stages [2]:

1) Definition of appropriate stages and states.

2) Definition of the optimal-value function.
3) Construction of a recurrence relation.
4) Recursive Computation.

For step 1, it is important to properly define stage variables and state variables. A stage variable should indicate the timeframe/timestamp of an equipment replacement problem. A state variable provides a snapshot of asset conditions at each timestamp [2]. In Toronto Hydro's investment problem, the stage variables are the years we plan for asset replacement; the state variables are the asset demographics of a given year defined as stage variables. In step 2, an appropriate optimization function must be defined [2]. In the Toronto Hydro investment spike smoothing problem defined earlier, the goal is to minimize risk cost through active replacements while constraining budget expenditures for each year. In step 3 and 4, a recursive function will be created to compute the optimized replacement schedule from the farthest replacement year included in the stage variables to the closest year (to now) [2].

Theoretical concerns

As a regulated utility company, Toronto Hydro faces many regulations and budgeting constraints. The model should allow analysts and planners to setup constraints that is unique to a specific project/program. The types of constraints should include, but not limited to: capital budgeting constraints, replacement quantity constraints, and investment period constraints.

The capital budgeting constraints set limits on annual capital expenditures. This constraint may vary from year to year. All values will be provided in present value.

The replacement quantity constraints set limits annual replacement quantities. This constraint may vary from year to year. The units of quantity may vary, for example, underground cables are measured in meters while transformers are counted per individual unit.

The investment period constraints set time limits to the replacement program. These constraints are a set of time variables; these variables indicate the desired time frame to achieve the set objectives.

Given Asset A and B in year i, A has a lower risk cost than B in year i, but if A has a higher risk cost growth, then in some year j, where j > i, the risk cost of A may potentially outgrow B's risk cost. Thus, it may be more efficient to replace A before B for better long-term outlooks, as A has a larger risk cost growth. The risk cost growth problem must be taken into consideration when formulating the mathematical optimization.

If an asset is replaced too early, the given asset may not be fully depreciated, the early replacement causes the company to lose value on the replaced asset; if an asset is replaced later, the given asset bear an increasing amount of outage risk due to age-related factors. If an outage occurs, the consequence cost of the outage may be significantly higher than the replacement cost or the asset value. The replacement timing trade-offs will need to be closely examined when developing investment smoothing algorithms.

Methods: proposed approaches

For the investment spike smoothing problem, "risk costs" carried by physical assets will be used to provide snapshots of asset demographics/asset health at each stage (current methodology used at Toronto Hydro). Toronto Hydro performs most maintenance activities on a fixed cycle, asset replacements will not affect routine maintenance activities. Therefore, maintenance costs will not be factored into this model. Toronto Hydro will provide a consequence cost for each individual asset based on its setup (e.g. carried load, location, outage duration, etc.). A mathematical program will be created to meet the given expenditure constraints while minimizing the risk costs. The model should provide an optimized investment schedule within the given constraints.

Non-linear programming approach and mathematical formulation

In this approach, the mathematical formulation will minimize the sum of annual risk cost over the investment planning period while setting constraints on annual investment budgets to smooth out the planned investment expenditures. This formulation considers the impact of risk cost growths within the investment period. Due to the non-linearity nature of this mathematical formulation and having two sets of decision variables, the program may have a long run-time when dataset becomes very large.

Variables and definitions

A_{ij}	Age, age of asset i in year j.
B_j	Annual Budget, the forecasted expenditure based on suggested replacement schedule in year j.
$Bmax_j$	Maximum Annual Budget, the maximum forecasted expenditure in year j.
$Bmin_j$	Minimum Annual Budget, the minimum forecasted expenditure in year j.
C_i	Consequence Cost, the cost to Toronto Hydro and customers if asset I fails or malfunctions.
$Favg_j$	Flexibility (Average), the allowed expenditure differences between the average annual expenditure and the annual budget in year j.
$Fcons_j$	Flexibility (Consecutive Years), the allowed expenditure differences between annual budgets of year j and year j + 1.
HZ_{ij}	Hazard Rates, the chance of failure of asset i in year j. Derived based on Weibull distributions.
$RepAmount_j$	Replacement Amount, the number of units replaced for a specific type of assets (major types, e.g. underground transformers) in year j.
$RepCost_i$	Replacement Cost, the cost to replace asset i.
$RiskCost_{ij}$	Risk Cost, is equivalent to Consequence Cost of asset i multiplied by the chance of failure (Hazard Rate) in year j.
S_i	Scale, a parameter used in the Weibull/Hazard Rate Function for asset i.
SF_i	Scale Factor, a parameter used in the Weibull/Hazard Rate Function for asset i. This parameter is used after assigning Shape to 3
Shape	Shape, a parameter used in the Weibull/Hazard Rate Function.

X_{ij} Decision Variable, X_{ij} takes on 0 or 1. If $X_{ij} = 1$, asset i will be replaced in year j. If $X_{ij} = 0$, asset i will remain in the system in year j.

Mathematical formulations and constraints

$$\min \sum_{i,j} \text{RiskCost}_{ij} = C_i \times HZ_{ij}$$

subject to

$$A_{i,j+1} = (A_{ij} + 1) \times (1 - X_{i,j})$$

$$HZ_{ij} = \left(\frac{\text{Shape}}{S_i}\right) \times \left(\frac{A_{ij}}{S_i}\right)^{(\text{Shape}-1)}$$

Most asset types have a Shape factor of 3, if all asset have a shape of 3, then HZ_{ij} can be simplified to:

$$\begin{aligned} HZ_{ij} &= \left(\frac{3}{S_i}\right)\left(\frac{A_{ij}}{S_i}\right)^{3-1} \\ &= \left(\frac{3}{s_i^3}\right) A_{ij}^2 \end{aligned}$$

$$\text{where } SF_i = \frac{3}{S_i^3}$$

$$\sum_i X_{ij} = \text{RepAmount}_j$$

$$B_j = \sum_i (X_{ij} \times \text{RepCost}_i)$$

The optimization must satisfy at least one of the following:

$$B\text{min}_j \leq B_j = \sum_i (X_{ij} \times \text{RepCost}_i) \leq B\text{max}_j$$

$$\left|\bar{B} - B_j\right| = \left|\bar{B} - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{avg}_j$$

$$|B_{j+1} - B_j| = \left|\sum_i (X_{i,j+1} \times \text{RepCost}_i) - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{cons}_j$$

The objective of this formulation is to minimize the overall system risk cost over the planning horizon. The budget constraints are imposed to smooth out the planned expenditures spikes.

This mathematical program contains two sets of decision variables, A_{ij} and X_{ij}. Although A_{ij} is a set of dependent variables of X_{ij}, in many mathematical optimization tools/solvers (in this problem, we use IBM CPLEX), any dynamic variable, or any variable that changes based on constraints is treated as a decision variable. Henceforth, the age variables A_{ij} act as a set of decision variables in the mathematical formulation. X_{ij} is a set of integer decision variables that determines whether or not to replace asset i in year j.

The first constraint is an age constraint. Because one is dealing with physical assets in this problem, aging is a significant contributor to assets' failure risks. If X_{ij} takes on the value of 0, meaning not to replace asset i in year j, then the constraint will set $A_{i,j+1}$ to $A_{i,j+1}$; if X_{ij} takes on the value of 1, meaning to replace asset i in year j, then the constraint will reset age $A_{i,j+1}$ to 0, reflecting that a new asset i has been put in to replace the old asset i in year j.

HZ_{ij} is the hazard rate of asset i in year j, given the age A_{ij}. This hazard rate provides the probability of asset i failing between the age of A_{ij} and the age of $A_{i,j+1}$, given that asset i has survived past age A_{ij}. The hazard rate function can be decomposed into two components, the probability density function f(t) and the survival function R(t), where $HZ_{ij} = f(t)=R(t)$. The probability density function, f(t), uses the Weibull distribution function assigned to each asset i based on its asset type. The Weibull distributions are often used in reliability engineering to predict the probability of failure of a given asset within a given time period. The survival function, R(t), provides the probability of an asset surviving past a certain age, in this case, surviving past A_{ij}.

$\Sigma_i X_{ij} = RepAmount_j$ is an optional constraint. This constraint gives engineers and analysts the option to limit the number of assets that are being replaced in year j.

B_j is the total budget proposed by the algorithm for year j. The main purpose of this research is to minimize the budgeting spikes within the investment period. This is done by imposing constraints on the annual budgets. The budget has to follow at least one of the budget constraint, but not limited to just one.

The first budget constraint has an upper and lower budget bound.

The second budget constraint provides a little more flexibility. It takes all proposed budgets over the planning horizon and finds the average budget. The proposed budget must fall within a determined range from the average budget.

The third budget constraint limits the budget change from year to year to a fixed value in each year.

Step-wise mathematical programming approach and mathematical formulation

In this approach, a step-wise mathematical formulation with a look-ahead period (k years in a look-ahead period) is introduced. At each investment year, the program will accumulate risk cost from current investment year to current investment year plus k years. In each investment year, the program will minimize the accumulated risk cost over the look-ahead period through asset replacement while staying within the budget constraints. A replacement schedule for the current year will be generated and the next year's asset demographics will be updated accordingly. This approach considers risk growth from the beginning for the investment year to the end of investment year plus the number of years in the look-ahead period. This approach has a looser constraint on risk growth than the previous approach. But due to its step-wise nature, breaking the problem into smaller sub-problems, this program has a short run-time compared to the previous approach.

Variables and definitions

$CumulativeRiskCost_{ij}$ CumulativeRiskCost, the sum of risk cost for asset i from year j to year (j + k) [k years in look-ahead Period].
A_{ij} Age, age of asset i in year j.
B_j Annual Budget, the forecasted expenditure based on suggested replacement schedule in year j.
$Bmax_j$ Maximum Annual Budget, the maximum forecasted expenditure in year j.
$Bmin_j$ Minimum Annual Budget, the minimum forecasted expenditure in year j.
C_i Consequence Cost, the cost to Toronto Hydro and customers if asset i fails or malfunctions.
$Favg_j$ Flexibility (Average), the allowed expenditure differences between the average annual expenditure and the annual budget in year j.
$Fcons_j$ Flexibility (Consecutive Years), the allowed expenditure differences between annual budgets of year j and year j + 1.
HZ_{ij} Hazard Rates, the chance of failure of asset i in year j. Derived based on Weibull distributions.
$RepAmount_j$ Replacement Amount, the number of units replaced for a specific type of assets (major types, e.g. underground transformers) in year j.
$RepCost_i$ Replacement Cost, the cost to replace asset i.
$RiskCost_{ij}$ Risk Cost, is equivalent to Consequence Cost of asset i multiplied by the chance of failure (Hazard Rate) in year j.
S_i Scale, a parameter used in the Weibull/Hazard Rate Function for asset i.
SF_i Scale Factor, a parameter used in the Weibull/Hazard Rate Function for asset i. This parameter is used after assigning Shape to 3. For more details, please see section 7.3.2.
Shape Shape, a parameter used in the Weibull/Hazard Rate Function.
X_{ij} Decision Variable, X_{ij} takes on 0 or 1. If $X_{ij} = 1$, asset i will be replaced in year j. If $X_{ij} = 0$, asset i will remain in the system in year j.

Mathematical formulations and constraints

Starting from the first year within the investment period, calculate the CumulativeRiskCost based on asset age and the look-ahead parameter, then determine an annual replacement schedule through integer programming optimization. After completing the optimization, update next year's asset demographics accordingly. Repeat for all years within the investment period.

$$\min \sum_{i} \text{CumulativeRiskCost}_{ij} \text{ for each } j \text{ in investment period}$$

subject to

$$\text{CumulativeRiskCost}_{ij} = \sum_{k=0}^{\text{lookAhead}} \text{RiskCost}_{ik}$$

$$\text{RiskCost}_{ij} = C_i \times HZ_{ij}$$

$$HZ_{ij} = \left(\frac{\text{Shape}}{S_i}\right) \times \left(\frac{A_{ij}}{S_i}\right)^{(\text{Shape}-1)}$$

Most asset types have a Shape factor of 3, if all asset have a shape of 3, then HZ_{ij} simplified to:

$$\begin{aligned} HZ_{ij} &= \left(\frac{3}{S_i}\right) \times \left(\frac{A_{ij}}{S_i}\right)^{3-1} \\ &= \left(\frac{3}{s_i^3}\right) A_{ij}^2 \end{aligned}$$

$$SF_i = \frac{3}{S_i^3}$$

$$\sum_i X_{ij} = \text{RepAmount}_j$$

$$B_j = \sum_i (X_{ij} \times \text{RepCost}_i)$$

The optimization must satisfy at least one of the following:

$$B\text{min}_j \leq B_j = \sum_i (X_{ij} \times \text{RepCost}_i) \leq B\text{max}_j$$

$$|\bar{B} - B_j| = \left|\bar{B} - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{avg}_j$$

$$|B_{j+1} - B_j| = \left|\sum_i (X_{i,j+1} \times \text{RepCost}_i) - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{cons}_j$$

The objective of this formulation is to minimize the overall system Risk Cost for year j and repeat for all years in the planning horizon. The budget constraints are imposed to smooth out the planned expenditures spikes.

This Step-Wise Mathematical Programming Approach introduces a look-ahead parameter, k, helps us to determine the CumulativeRiskCost measure, which accumulates the risk cost from year j to year j + k. This measure not only considers the risk cost in current years, but it also examines the impact of future risk costs (up to age j + k) if this asset remains in the distribution system.

X_{ij} is a set of integer decision variables that determines whether or not to replace asset i in year j.

The previous Non-Linear Programming Approach contains two sets of decision variables, A_{ij} and X_{ij}. Although A_{ij} is a set of dependent variables of X_{ij}, in many mathematical optimization tools/solvers (in this problem, we use IBM CPLEX), any dynamic variable, or any variable that changes based on constraints is treated as a decision variable. Henceforth, the age variables A_{ij} act as a set of decision variables in the mathematical formulation. This dramatically increases the complexity of the non-linear program as it requires the machine to solve two sets of decision variables, one of them being a set of integer decision variables, the runtime for the program is very long when processing large datasets.

The Step-Wise Mathematical Programming Approach removes the age constraint, thus re-moving one set of decision variables, A_{ij}. The A_{ij} used in this approach are fixed values during the optimization process and are only changed after each optimization cycle. The program utilizes the look-ahead variable and the mechanism of CumulativeRiskCost to evaluate the current and future impacts of risk cost. The program provides an optimized replacement schedule for the current year, then resets/increases the assets' age based on the replacement schedule. Once the ages are reset/increased, the program will move to the next year and loop this cycle until it reaches the end of the planning horizon.

Sorting approach and mathematical formulation

This Sorting Approach is built based on the foundation of the Step-Wise Mathematical Programming Approach. The Step-Wise Mathematical Programming Approach provides an optimal replacement schedule for minimizing the CumulativeRiskCost over the look-ahead period. It has a decent runtime when processing small and medium size datasets (less than 300 data entries). But due to the nature of integer programming, the complexity of the program grows significantly when the number of decision variables increases. Thus, this approach may take a very long time to process when the dataset is large. Given the large physical distribution asset population of Toronto Hydro, it is important that the algorithm is able to process large dataset efficiently. The Sorting Approach is introduced as a solution to this problem as it has significantly lower runtime compared to the Non-Linear Programming Approach and the Step-Wise Mathematical Programming Approach. The approach is motivated by the need of a faster investment planning algorithm for larger datasets. Instead of producing the replacement schedule using integer programming optimization, the approach utilizes sorting mechanisms to remove assets with the highest CumulativeRiskCost to ReplacementCost Ratios. The ratio provides a snapshot of the cost-efficiency for asset replacement given the CumulativeRiskCost from year j to year $j + k$. The ratio shows that, for every dollar spent in replacement cost, how much CumulativeRiskCost is removed from the system; the higher the ratio, the better the cost-efficiency.

In this approach, a sorting mechanism with a look-ahead period (k years in a look-ahead period) will be utilized. In each investment year, the program will accumulate risk cost from current investment year to current investment year plus k years. In each investment, the program will minimize the accumulated risk-cost over the look-ahead period through asset replacement while staying within the budget constraint. A replacement schedule for the current year will be generated and the next year's asset demographics will be updated accordingly. This approach considers risk growths from the beginning for the investment period to the end of investment period plus the number of years in the look-ahead period.

Mathematical formulations and constraints

The algorithm will start from the first year in the investment period; calculate the CumulativeRiskCost from year j to year j+k, then sort all assets by their CumulativeRiskCost to ReplacementCost Ratio. Determine a replacement schedule by removing the assets with the highest ratio while staying within budget constraints. After asset replacements, the algorithm updates next year's asset demographics accordingly and repeats for all years within the investment period.

Calculate the CumulativeRiskCost for asset i year j using the following formula:

$$\text{CumulativeRiskCost}_{ij} = \sum_{k=0}^{\text{lookAhead}} \text{RiskCost}_{ik}$$

$$\text{RiskCost}_{ij} = C_i \times HZ_{ij}$$

$$HZ_{ij} = \left(\frac{\text{Shape}}{S_i}\right) \times \left(\frac{A_{ij}}{S_i}\right)^{(\text{Shape}-1)}$$

Most asset types have a Shape factor of 3; if all asset have a shape of 3, then HZ_{ij} is simplified to:

$$\begin{aligned} HZ_{ij} &= \left(\frac{3}{S_i}\right) \times \left(\frac{A_{ij}}{S_i}\right)^{3-1} \\ &= \left(\frac{3}{s_i^3}\right) A_{ij}^2 \\ SF_i &= \frac{3}{S_i^3} \end{aligned}$$

Calculate the CumulativeRiskCost to ReplacementCost Ratio of each asset:

$$CumulativeRiskCost - to - ReplacementCost - Ratio_{i,j} = \frac{CumulativeRiskCost_{i,j}}{RepCost_i}$$

Sort the CumulativeRiskCost to ReplacementCost Ratio from highest to lowest. Replace assets with the highest CumulativeRiskCost to ReplacementCost Ratio until the replacement amount constraint and the budget constraints are met. After meeting constraint requirement, the program will continue to replace asset until no further replacements can be executed without violating the constraints. Please note that the list of asset is sorted by CumulativeRiskCost to Replacement Ratio without regards to the size of the replacement cost. Given the budget constraint, if an asset in the sorted list, say asset m, cannot not be replaced without violating the budget constraint, the algorithm will skip asset m and move to the next asset in the sorted list.

Please note that the replacement amount constraint may conflict with the budget constraints and make the solution less optimal under the Sorting Approach. If the analyst chooses to impose such constraint, it is recommended to have a very loose replacement amount constraint to minimize the impact of the conflict.

After determining the replacement schedule for the current year, the algorithm will reset/increase the age parameters according to the replacement schedule. The algorithm, then, will loop back to generating a new set of CumulativeRiskCost with year j + 1 as the current year, and repeat process until it reaches the end of the investment horizon.

The replacement amount constraint and the budget constraints are imposed as follows:

$$\sum_i X_{ij} = \text{RepAmount}_j$$

$$B_j = \sum_i (X_{ij} \times \text{RepCost}_i)$$

The optimization must satisfy at least one of the following:

$$B\text{min}_j \leq B_j = \sum_i (X_{ij} \times \text{RepCost}_i) \leq B\text{max}_j$$

$$|\bar{B} - B_j| = \left|\bar{B} - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{avg}_j$$

$$|B_{j+1} - B_j| = \left|\sum_i (X_{i,j+1} \times \text{RepCost}_i) - \sum_i (X_{ij} \times \text{RepCost}_i)\right| \leq F\text{cons}_j$$

Because this approach utilizes sorting mechanisms, the formulation does not have an optimization objective, instead, the formulation lists the steps within the algorithm.

The goal of this formulation is to reduce the overall system CumulativeRiskCost from year j to year j + k and repeat for all years in the planning horizon. The budget constraints are imposed to smooth out the planned expenditures spikes.

Similar to the Step-Wise Mathematical Programming Approach, the Sorting Approach uses a look-ahead variable, k, helps us to determine the CumulativeRiskCost measure, which accumulates the risk cost from year j to year j + k. This measure not only considers the risk cost in current years, but it also examines the impact of future risk costs (up to age j + k) if this asset remains in the distribution system.

The decision variables X_{ij} in this approach serves as a set of indicators to whether the replacement schedule has violated the constraints or not. A_{ij} will serve as parameters and will reset/increase according to the generated replacement schedule at the end of each year's replacement cycle.

As noted earlier, the previous Non-Linear Programming Approach contains two sets of decision variables, A_{ij} and X_{ij}. Although A_{ij} is a set of dependent variables of X_{ij}, in many mathematical optimization tools/solvers (in this problem, we use IBM CPLEX), any dynamic variable, or any variable that changes based on constraints is treated as a decision variable. Henceforth, the age variables A_{ij} act as a set of decision variables in the mathematical formulation. This dramatically increases the complexity of the non-linear program as it requires the machine to solve two sets of decision variables, one of them being integer decision variables, the runtime for the program is very long when processing large datasets.

The Sorting Approach moves away from integer programming optimizations. The program utilizes the look-ahead variable and the mechanism of CumulativeRiskCost to evaluate the current and future impacts of risk cost. The program provides a replacement schedule for the current year based on the CumulativeRiskCost to ReplacementCost ratio, then resets/increases the assets' age based on the replacement schedule. Once the ages are reset/increased, the program will move to the next year and loop this cycle until it reaches the end of the planning horizon.

Approaches in application, results, and discussions

In this section, we will discuss the problems that these approaches face in implementation and in application. The testing results will be presented and the takeaways of each approach will be discussed in details.

Non-linear programming approach in application

This approach is more costly in terms of mathematically complexity as it attempts to determine the optimal solution, without any loosening in constraints, over the planning horizon in one iteration. This formulation requires the server to possess a high level of processing power in order to complete the calculation.

For this reason, the implementation is done using IBM ILOG CPLEX, which supports high-level Optimization Programming Language (OPL) and allows the user to run the optimization on the IBM CPLEX Enterprise Server. IBM CPLEX Enterprise Server is a scalable server that provides mathematical optimization environments based on the size of the optimization. It is a powerful tool that is available for students for academic-use at zero cost.

The first obstacle encountered during implementation was within the age constraint. The age in current year (A_{i1}) is a fixed integer, but age in later years (A_{ij} where $j \geq 2$) becomes a variable based on the replacement schedule determined by X_{ij}. Though A_{ij} is dependent on the decision variables X_{ij}, in Optimization Programming Language (OPL), A_{ij} also becomes decision variables. The previously formulated age constraint:

$$A_{i,j+1} = (A_{ij} + 1) \times (1 - X_{i,j+1})$$

creates high mathematical complexity and renders the constraint non-convex. This makes the mathematical optimization very costly and IBM ILOG CPLEX is unable to handle non-convex constraints. To overcome this barrier, the age decision variables (A_{ij}) have been removed due to its dependency to X_{ij} and only asset age in the current year (AGEY 1[i]) is used in the formulation. All age parameters used in other constraints are derived directly using AGEY 1[i] and decision variables X_{ij}. The age parameter can be rewritten as:

$$AGEY1[i] + j + \sum_{k=1}^{\text{year j}} (AGEY1[i] + k) \times (X_{ik})$$

Unfortunately, this alternative formulation only holds true when limiting the replacement schedule to replace the same asset no more than once, thus:

$$\forall i, \sum_{j} X_{ij} \leq 1$$

This can be a reasonable assumption / constraint. Due to the long lifespan (average useful life) of utility assets, generally between 30 to 45 years, it is reasonable to assume/constrain that each asset may only be replaced once during the 5 to 20-year planning horizon.

Additionally, the original risk cost constraint:

$$RiskCost_{ij} = \sum_{i,j} C_i \times HZ_{ij}$$

is considered as a non-convex constraint. To make this constraint convex, the comparison operator is changed from equal "=" to greater than "≥"

$$RiskCost_{ij} \geq \sum_{i,j} C_i \times HZ_{ij}$$

This does not change the optimization results as the optimization objective minimizes $RiskCost_{ij}$.

This approach is able to provide a replacement schedule that achieves minimum total risk cost, over the planning horizon, given a set of budget constraints. But due to lack of relaxation in constraints, single-iteration optimization process, and costly integer programming approach, the optimization has an extremely long runtime. It is advised to use the Step-Wise Mathematical Programming Approach or the sorting Approach if relaxation in constraints is acceptable.

Non-linear programming approach: further formulation adjustments

The approach suggested previously solves the convexity issue but imposes significant runtime onto this problem. After carefully conducted analyses and tests, a new formulation has been proposed to replace the previous age constraint:

$$A_{i,j} + 1 <= A_{i,j+1} + X_{i,j} * M$$

where M is a number much larger than the average age of the physical asset population (an large M value of 10,000,000 is used in calculation).

To understand this formulation, we can break this constraint down to two different scenarios: when X_{ij} is 0 and when X_{ij} is 1. This constraint is imposed to a risk minimization optimization program. Because the age metrics (A_{ij} and $A_{i,j+1}$) are proportional to the risk measures, the minimization optimization will naturally force A_{ij} and $A_{i,j+1}$ to the smallest possible value without violating this or any other constraints.

When $X_{ij} = 0$, meaning asset i will not be replaced in year j, the $X_{ij} * M$ term has no effect in this inequality. The optimization will tempt to minimize the age and setting the constraint to $A_{ij} + 1 = A_{i,j+1}$

When $X_{ij} = 1$, meaning asset i will be replaced in year j, the $X_{ij} * M$ term becomes extremely large compared to the age metrics (A_{ij} and $A_{i,j+1}$). In this case, the right-hand side terms no longer requires $A_{i,j+1}$ to hold a large value to become larger than $A_{ij} + 1$. Due to minimization, $A_{i,j+1}$ will take on the smallest possible value, 0 (Age is set to be positive integers). This means that after the replacement of asset i in year j, the asset age will be reset to zero to reflect the state of the new asset i.

This constraint adjustment reduces the runtime for small to medium datasets by a significant margin (Approx. 30 percent). The inequality utilizes the nature of the optimization process to simplify the mathematical formulation.

Non-linear programming approach: concerns and obstacles

The advantage of this Non-Linear Programming Approach is that it finds the optimal solution that minimizes the total system risk over the given horizon while staying within the budget constraints.

The approach utilizes integer programming with two sets of decision variables, one set of them being integer decision variables, which has a very high mathematical complexity and can be very costly when analyzing large datasets. Though significant improvements were seen after the age constraint adjustment, the approach still has a very long run-time when processing large datasets. A 15,000 entry dataset would take more than 48 hours to execute. Because this approach can only process small datasets, due to limited data points, the sample results do not provide meaningful results.

Although an optimal solution is ideal, it may be preferable to provide a slightly sub-optimal solution that takes significantly less time to process. Given that is problem is defined based on a probabilistic concept, the optimal solution only outlines a replacement schedule given the likelihood of failures. The replacement schedule is subject to change in reality as it is derived based on a snapshot of the current system demographics. Thus, a close approximation of the optimal solution is sufficient enough in most scenarios given the circumstances. The Step-Wise Mathematical Programming Approach and the Sorting

Approach provide slightly sub-optimal solutions that process the problems with much shorter runtimes.

Empirical results and quantitative testing on small datasets have shown that the difference in optimality and cost-efficiency between the Non-Linear Programming Approach and the Step-Wise Mathematical Programming is very minimal.

Step-wise mathematical programming approach in application

In comparison with the Non-Linear Programming Approach, the Step-Wise Mathematical Programming Approach saves a significant amount of runtime when processing medium size dataset (50 to 300 data entries). Though the approach works very well for medium size dataset, it still has runtime issues when processing large datasets due the mathematical complexity of integer programming. Thus, a Sorting Approach has been introduced to target the optimization process of large asset groups. The sorting approach is slightly suboptimal in terms of finding the best replacement schedule with the highest cost-efficiency. But empirical and quantitative test results have shown that the difference in cost-efficiency is very minimal when the dataset becomes very large. In the sample code provided in the Appendix, the algorithm takes advantage of R Studio's built-in integer programming mechanism to further improve the algorithm's performances.

Step-wise mathematical programming approach results

The Step-Wise Mathematical Programming Approach is a great approach when processing small to medium datasets. The Step-Wise Mathematical Program has shown that when dealing with assets with same asset type but with different subtypes, the look-ahead parameter does not have too much impact on the optimization process as the assets with the same asset type tend to have similar risk growth (i.e. similar shape and scale). Because the differences in risk growth are reasonably small, sometimes programs with greater look-ahead parameters may under-perform compared to programs with smaller look-ahead parameters.

It is sometimes ideal to only consider the current annual risk cost and disregard the long-term effect of risk growths. To disregard the long-term effect of risk growths, set the look-ahead parameter to a small positive integer value (e.g. 1 or 2). If the look-ahead parameter is set to 1, then during optimization, the program will provide a replacement schedule based on the current year's annual risk only. If the look-ahead parameter is set to 2, then during optimization, the program will provide a replacement schedule based on the current year and next year's annual risk.

When the differences between the shape of each asset and the differences between the scale of each asset become significant, empirical and quantitative testing have shown that having considerations for the impact of risk growths could significantly benefit the company in the long-term. Please note that, due to limited access to Toronto Hydro's real dataset (for confidentiality reasons), these empirical and quantitative analysis are performed on designed datasets that simulate assets with large differences in shape and scale. The Step-Wise Approach test results of assets with large differences in shape and scale over 150 years have been shown in Figure 2. Each program has been given the same and constant annual budgets. The assets in the dataset used in testing has large differences in both shape and scale; the assets are, initially, at very young ages to emphasize the impact of large risk growths simulated in this example.

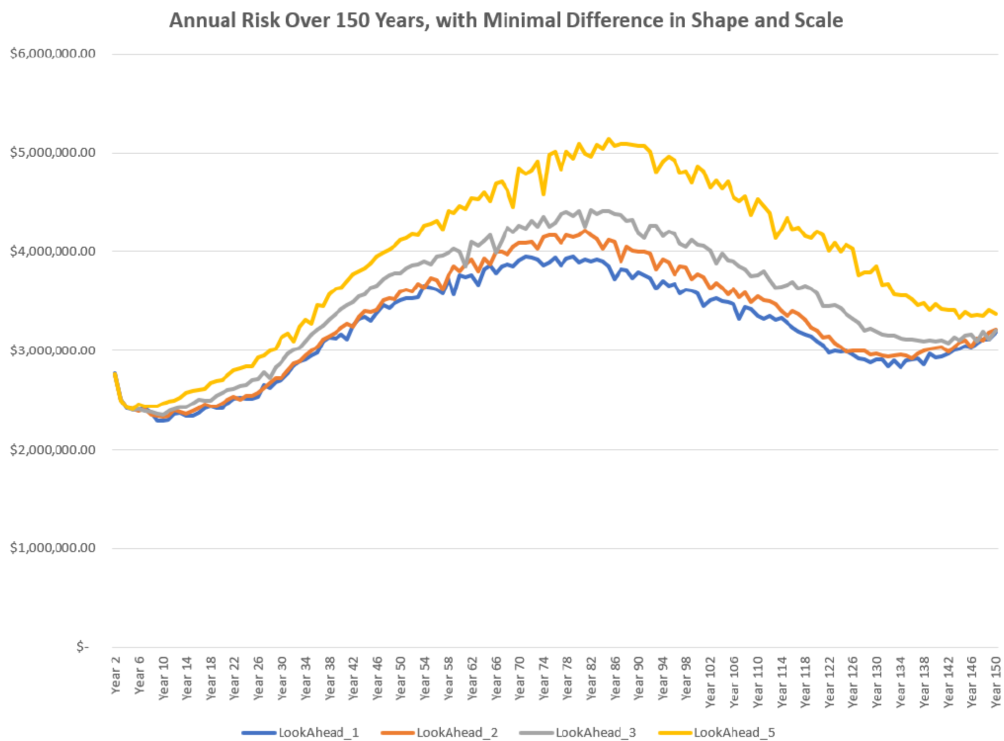


Figure 1: Step-Wise Approach: Annual Risk Over 150 Years, Minimal Differences in Shape and Scale

As we can see in Figure 2, the replacement program with a look-ahead parameter of 1 performs very well in the short-term, but begins to accumulate large risk costs in the long-term. As assets begin to age, the impact of risk growths becomes significant, and the more long-term oriented programs (i.e. programs with a greater look-ahead parameter) begin to outperform the "short-sighted" programs.

Please note that the look-ahead parameters require careful selection. Depending on asset demographics, each look-ahead parameter may perform differently. It is recommended to perform a similar quantitative test as shown in Figure 2, based on the quantitative results, select a look-ahead parameter that aligns with the company's investment strategies.

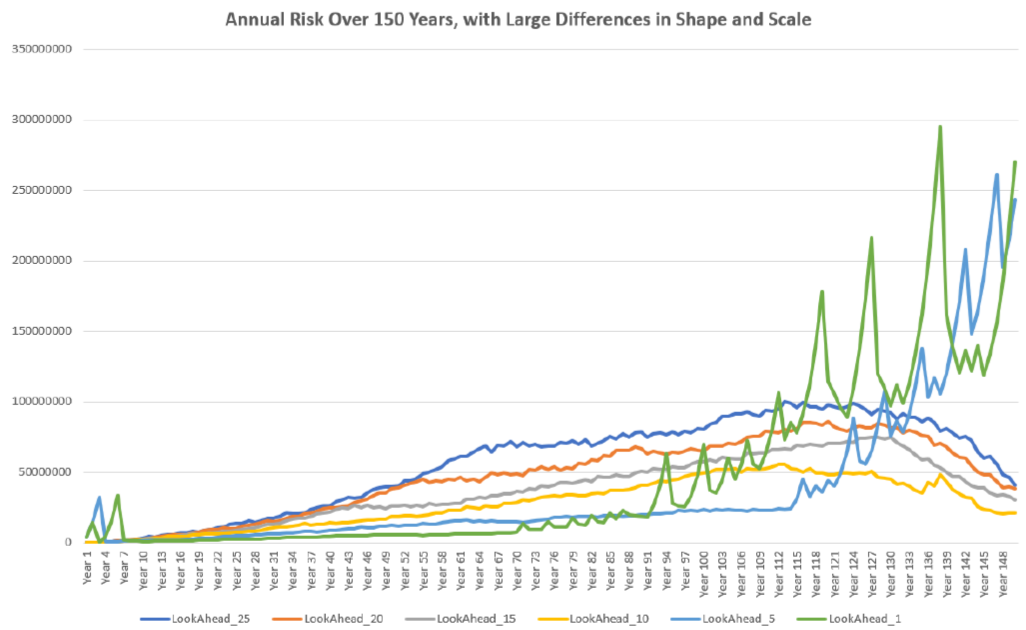


Figure 2: Step-Wise Approach: Annual Risk Over 150 Years, Large Differences in Shape and Scale

Sorting approach in application

The Sorting Approach is motivated by the need of reducing model's mathematical complexity and shorten the program runtime. In the sample code provided in the Appendix, the algorithm takes advantage of R Studio's built-in quick sorting mechanism to further improve the algorithm's performances. The runtime has been significantly reduced as the sorting mechanism only takes a fraction of the amount of time to run an integer optimization. Most of the runtime of the Sorting Approach is contributed by the calculations of CumulativeRiskCost.

Sorting approach results

The results are very promising as the deviations between the Sorting Approach and the Step-Wise Mathematical Programming Approach (which should produce a more optimal result in theory due to its usage of integer programming optimization) are very small.

The Sorting Approach is a great model to use when processing larger datasets. Similar with the Step-Wise Mathematical Programming Approach, the Sorting Approach has shown that when dealing with assets with same asset type but with different subtypes, the look-ahead parameter does not have too much impact on the optimization process as the assets with the same asset type tend to have similar risk growth (i.e. similar shape and scale). As we can see in Figure 2 and Figure 3 , the Sorting Approach results are very similar to the Step-Wise Approach results. It appears that the greater look-ahead parameter programs are likely to under-perform compared to smaller look-ahead parameter programs in the short-term. Due to the small differences in risk growth, a short-term oriented approach appears to be more appropriate for such scenarios.

Please note that the look-ahead parameters require careful selections. Depending on asset demographics, each look-ahead parameter may perform differently. It is recommended to perform a similar quantitative test as shown in Figure 3, based on the quantitative results, select a look-ahead parameter that aligns with the company's investment strategies.

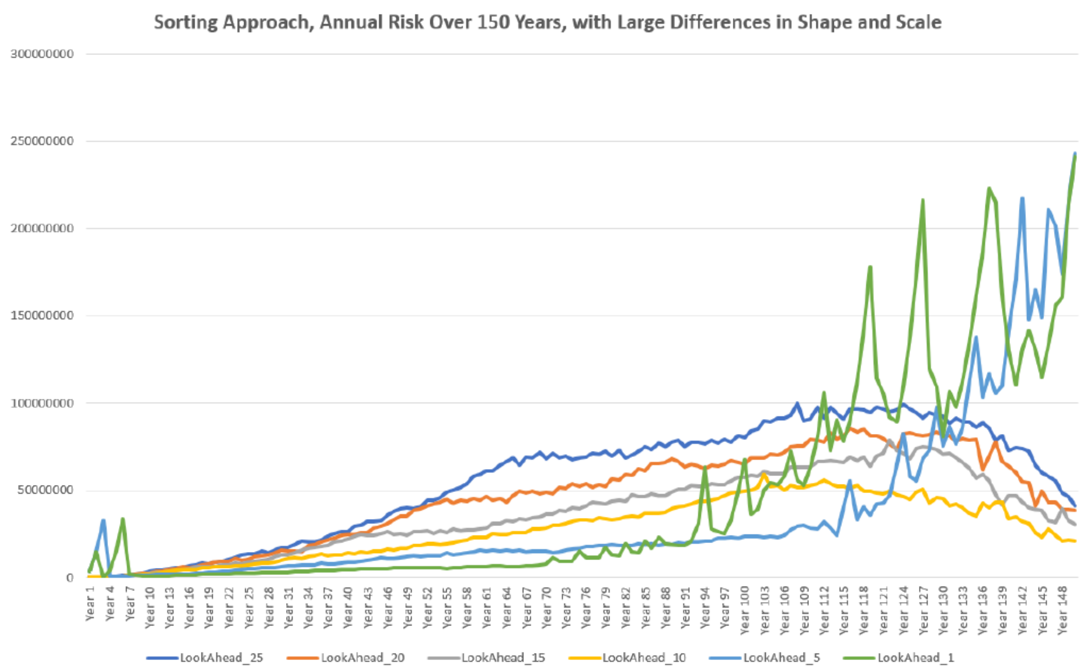


Figure 3: Sorting Approach: Annual Risk Over 150 Years, Large Differences in Shape and Scale

Conclusion

Toronto Hydro has been facing the problem of investment spiking when formulating investment plans. Assets of the same type often have similar useful-lives. Historically, large projects/initiatives have installed assets within a short period of time (or in a cyclical nature of a period of years), which can contribute to spikes in renewal expenditures in the future. Due to capital budget constraints, not all assets can be replaced at or before the optimal/expected replacement time, which further contributes to the spikes in renewal expenditures. Asset risk growths are also major concerns when developing investment plans.

This report has provided three mathematical approaches to aid Toronto Hydro's investment planning process, with a focus on smoothing out the investment spikes and reducing the impact of asset risk growths.

Though the Non-Linear Programming Approach provides an optimal replacement schedule for the investment smoothing problem while solving the asset risk growths problem within the investment horizon, the significant mathematical complexity of this approach makes the model difficult to use in real life settings. If time permits, we can use this approach to generate replacement schedules for small datasets.

The Step-Wise Mathematical Programming Approach is a viable solution to solving the investment spiking issue in a small to medium-size asset population. It removes one set of decision variable used in the Non-Linear Programming Approach and introduces the look-ahead parameters. It not only provides an optimized replacement schedule that smooths out the annual budgets, it also emphasizes solving the asset risk growth issues. The downside of this approach is that it takes a long time to execute large datasets due to the built-in integer programming optimization process.

The Sorting Approach is derived based on the foundation of the Step-Wise Mathematical Programming Approach. It utilizes the look-ahead parameters and the CumulativeRiskCost concept in the Step-Wise Mathematical Programming Approach. Instead of providing the replacement schedule through integer programming optimization, it utilizes sorting mechanisms to significantly reduce the program runtime. It is an ideal approach when optimizing investment plans with a very large asset population.

Each approach has advantages and downside trade-offs. It is important for engineers and analysts to recognize these trade-offs and select the appropriate approach based on the asset demographics and investment goals.

Future work

For future steps, we will continue to coordinate with Toronto Hydro Corporation and test out all approaches on real datasets (current implementation tests were done on smaller sample datasets).

The complete thesis work will be presented at the Centre for Maintenance Optimization and Reliability Engineering Progress Conference.