Linköping Studies in Science and Technology
Dissertations No. 2244

On Machine Learning-Based Control for Energy Management in Construction Machines

Henrique Raduenz

Division of Fluid and Mechatronic Systems
Department of Management and Engineering
Linköping University, SE-581 83 Linköping, Sweden

Linköping 2022

On Machine Learning-Based Control for Energy Management in Construction Machines

Cover: Henrique Raduenz, 2022

ISBN 978-91-7929-415-1 (print)
ISBN 978-91-7929-416-8 (PDF)
https://doi.org/10.3384/9789179294168
ISSN 0345-7524

Printed in Sweden by LiU-Tryck, Linköping 2022.

UNIVERSIDADE FEDERAL DE SANTA CATARINA
CENTRO TECNOLÓGICO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA

Henrique Raduenz

On Machine Learning-Based Control for Energy Management in Construction Machines

Linköping
2022

Henrique Raduenz

Sobre Controle Baseado em Aprendizado de Máquina para Gerenciamento de Energia em Máquinas de Construção

Tese submetida ao Programa de Pós-Graduação em Engenharia Mecânica da Universidade Federal de Santa Catarina e pela Universidade de Linköping em regime de cotutela para a obtenção do título de Doutor em Engenharia Mecânica.
Orientadores: Prof. Victor Juliano De Negri, Dr. (UFSC) e
Prof. Petter Krus, Dr. (LiU)

Linköping / Florianópolis
2022

Henrique Raduenz

On Machine Learning-Based Control for Energy Management in Construction Machines

Thesis submitted to the Graduate Program in Mechanical Engineering of the Federal University of Santa Catarina and to Linköping University in a double degree format to obtain the Doctoral Degree in Mechanical Engineering.
Advisors: Prof. Victor Juliano De Negri, Dr. (UFSC) and
Prof. Petter Krus, Dr. (LiU)

Linköping / Florianópolis
2022

Ficha de identificação da obra elaborada pelo autor,
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Raduenz, Henrique
On Machine Learning-Based Control for Energy Management in Construction Machines / Henrique Raduenz ; orientador, Victor Juliano De Negri, orientador, Petter Krus, 2022.
140 p.

Tese (doutorado) - Universidade Federal de Santa Catarina, Centro Tecnológico, Programa de Pós-Graduação em Engenharia Mecânica, Florianópolis, 2022.

Inclui referências.

Trabalho elaborado em regime de co-tutela.

1. Engenharia Mecânica. 2. Machine Learning. 3. Energy Management. 4. Construction Machines. I. De Negri, Victor Juliano. II. Krus, Petter III. Universidade Federal de Santa Catarina. Programa de Pós-Graduação em Engenharia Mecânica. IV. Título.

Henrique Raduenz

On Machine Learning-Based Control for Energy Management in Construction Machines

O presente trabalho em nível de doutorado foi avaliado e aprovado por banca examinadora composta pelos seguintes membros:

Prof. Jonny Carlos da Silva, Dr. Eng.
Universidade Federal de Santa Catarina, Brasil

Prof. Bengt J. H. Jacobson, Ph.D.
Chalmers University of Technology, Suécia

Prof. Erik Frisk, Ph.D.
Linköping University, Suécia

Prof. Mats Björkman, Ph.D.
Linköping University, Suécia (Suplente)

e tendo como oponente:

Prof. Jouni Mattila, Dr. Tech.
Tampere University, Finlândia

Certificamos que esta é a **versão original e final** do trabalho de conclusão que foi julgado adequado para obtenção do título de Doutor em Engenharia Mecânica.

Documento assinado digitalmente
Paulo de Tarso Rocha de Mendonca
Data: 06/10/2022 10:30:52-0300
CPF: ***.822.691-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Paulo de Tarso Rocha de Mendonça, Dr.
Coordenador do Programa de Pós-Graduação

Documento assinado digitalmente
Victor Juliano de Negri
Data: 04/10/2022 17:57:25-0300
CPF: ***.505.630-**
Verifique as assinaturas em https://v.ufsc.br

Prof. Victor Juliano De Negri, Dr. Eng.
Orientador

Prof. Petter Krus, Dr.
Orientador

Linköping / Florianópolis
2022

To Shara

The important thing is not to stop questioning.
Curiosity has its own reason for existing.
Albert Einstein

Abstract

High energy efficiency is a key requirement for modern construction machinery. This is because of stricter environmental targets, electrification, and reduction of operation costs. To meet this requirement, the powertrain architectures of the machines are becoming increasingly complex, for example through hybridisation of drivetrain and work functions, or with improved hydraulic systems. However, the more complex the architecture is, the harder the management of splitting power between different sources and consumers. The number of work functions, operating environments, and tasks these machines engage in, along with the added degrees of freedom with respect to how energy can be recovered, exchanged, and reused, makes them unique. Therefore, the development of control strategies for energy management in such machines requires specific research and development with their architecture and application in focus. This doctoral thesis presents an analysis of two methods for the development of machine learning-based energy management strategies for construction machines. One is based on supervised learning and the other on reinforcement learning. The methods use optimisation to find optimised solutions for the control problem of the systems and machine learning for learning and implementing the control decisions. In both methods, models of the physical systems are used for the learning and training. The thesis highlights and confirms, with experimental results, the potential of such methods to derive control strategies for these machines. The studied methods can learn and implement improved control decisions in the real systems that result in the potential for increased efficiency. At the same time, their robustness is shown in the application to unseen scenarios during training, although that does not eliminate the need for further training in the real systems after deployment. The thesis also increases the comprehensiveness on energy management for construction machines. The thesis was completed in a double-degree format between the Federal University of Santa Catarina, Florianópolis, Brazil, and Linköping University, Linköping, Sweden.

Keywords: Machine Learning, Energy Management, Construction Machines

Populärvetenskaplig sammanfattning

Hög effektivitet är ett nyckelkrav för moderna byggmaskiner. Detta på grund av strängare miljömål, elektrifiering och sänkta driftskostnader. För att möta detta krav blir maskinernas arkitektur mer komplex, till exempel genom hybridisering av drivlina och arbetsfunktioner. Men ju mer komplex arkitekturen är desto svårare blir hanteringen av maktdelning mellan olika källor och konsumenter. Antalet arbetsfunktioner, driftsmiljö och uppgifter de engagerar tillsammans med de ökade frihetsgraderna med avseende på hur kraft kan återvinnas, bytas ut och återanvändas, gör dem unika. Därför kräver utvecklingen av styrstrategier specifik utveckling med deras arkitektur och tillämpning i fokus. Denna doktorsavhandling presenterar en analys av två metoder för utveckling av optimerade och intelligenta energihanteringsstrategier i realtid för delsystem av komplexa entreprenadmaskiner. De utvärderade metoderna använder optimering för att hitta optimala lösningar för systemens kontrollproblem, och använder maskininlärning som ett sätt att lära sig och implementera de optimerade besluten. I båda metoderna används modeller för lärandet och träningen. Avhandlingen belyser och bekräftar experimentellt potentialen hos sådana metoder för att härleda kontrollstrategier för dessa maskiner. De studerade metoderna kan lära sig och implementera optimerade styrbeslut i de verkliga systemen vilket leder till ökad effektivitet. Samtidigt visar det sig att de är robusta mot osynliga scenarier under träning, även om det inte eliminerar behovet av vidareutbildning i de verkliga systemen efter utplacering. Examensarbetet ökar också heltäckningen om energihantering för entreprenadmaskiner. Avhandlingen har utvecklats i ett dubbelgradersformat med Federal University of Santa Catarina, Florianópolis, Brasilien och Linköpings Universitet, Linköping, Sverige.

Resumo

Alta eficiência energética é um requisito para máquinas de construção modernas, sendo este uma consequência das metas ambientais, eletrificação e redução de custos. Para atender este requisito as arquiteturas dos trens de potência das máquinas têm se tornado mais complexas, por exemplo, através da hibridização do sistema de tração e funções de trabalho, ou sistemas hidráulicos melhorados. Entretanto, quanto mais complexa a arquitetura, mais difícil se torna o gerenciamento da divisão de potência entre as fontes e consumidores. O número de funções de trabalho, ambientes de operação e tarefas que elas executam, juntamente com os graus de liberdade relacionados à como energia pode ser recuperada, trocada e reutilizada, às tornam únicas. Dessa maneira, o desenvolvimento de estratégias de controle requer pesquisa e desenvolvimento específicos com as suas arquiteturas e aplicações em foco. Esta tese de doutorado apresenta uma análise de dois métodos para o desenvolvimento de estratégias baseadas em aprendizado de máquina para o gerenciamento de energia em máquinas de construção. Um é baseado em aprendizado supervisionado e outro em aprendizado por reforço. Os métodos avaliados usam otimização para encontrar soluções otimizadas para o problema de controle dos sistemas, e usam aprendizado de máquina como meio para aprender e implementar as decisões de controle. Em ambos os métodos, modelos dos sistemas físicos são utilizados para o aprendizado e treinamento. A tese destaca e confirma através de resultados experimentais, o potencial destes métodos em obter estratégias de controle para estas máquinas. Os métodos estudados são capazes de aprender e implementar decisões de controle melhores nos sistemas reais resultando em potencial aumento de eficiência energética. Ao mesmo tempo, é mostrado a sua robustez na prática a cenários não vistos durante o treinamento, apesar de isso não eliminar a necessidade de continuar o treinamento depois de implementadas no sistema real. A tese também aumenta a compressão sobre gerenciamento de energia em máquinas de construção. A tese foi desenvolvida em formato de cotutela com a Universidade Federal de Santa Catarina, Florianópolis, Brasil, e a Universidade de Linköping, Linköping, Suécia.

Palavras-Chave: Aprendizado de Máquina, Gerenciamento de Energia, Máquinas de Construção

Resumo Expandido

Introdução

Esta tese aborda métodos para a obtenção de estratégias de genrenciamento de energia baseadas em aprendizado de máquina para máquinas de construção.

O movimento global para atender metas para conter alterações climáticas impulssiona o desenvolvimento de sistemas mais eficientes. Isso também é válido para máquinas de construção.

Máquinas de construção como escavadeiras, carregadeiras de rodas e caminhões articulados são caracterizadas por um trem de potência não apenas dedicado ao movimento translacional mas também às funções de trabalho responsáveis por mover cargas pesadas. Sendo assim, para este tipo de máquina, o desenvolvimento de novos sistemas visando o aumento de eficiência, também envolve subsistemas que permitem a recuperação de energia cinética e potencial, ou que permitem modos de operação mais eficientes que os sistemas atuais.

Máquinas com a capacidade de recuperar, armazenar e reutilizar essas energias disponíveis têm graus de liberdade adicionais que, também precisam ser controlados, sendo estes controladores chamados de estratégias de gerenciamento de energia. Eles definem quando, onde e como a energia é gerada, recuperada, armazenada e reutilizada. Eles gerenciam o compartilhamento de energia entre múltiplas fontes e consumidores. É preciso ter uma visão holística da máquina, dos subsistemas e das tarefas de trabalho para otimizar esse gerenciamento de energia. O ganho em eficiência é, então, consequência da arquitetura da máquina e da estratégia de gerenciamento de energia.

Máquinas de construção com sistemas que permitem a recuperação de energia ou que possuam sistemas hidráulicos mais complexos que necessitem de gerenciamento de energia, têm uma maior capacidade de aumento de eficiência, mas ao mesmo tempo são mais desafiadoras sob o ponto de vista de desenvolvimento de estratégias de gerenciamento de energia. Por limitações relacionadas à sub-optimização, requisitos de computação, facilidade de implementação e adaptabilidade a novos cenários de operação, métodos já existentes para o desenvolvimento de estratégias de gerenciamento de energia aparentemente não atendem os requisitos de máquinas de construção. E em sua maiora, eles têm sido avaliados para veículos de rua e não para máquinas de construção.

Por outro lado, há indicativos de que o aprendizado de máquina possa atender aos requisitos e desafios relacionados à esse tipo máquina. Entretando, não existe grande quantidade de estudos demonstrando o seu potencial nessa aplicação. Ao mesmo tempo, as estratégias já estudadas não foram amplamente avaliadas em operação em tempo real em protótipos. Desta maneira, existe espaço para pesquisa específica e abrangente sobre métodos para obter estratégias otimizadas e baseadas em aprendizado de máquina para o gerenciamento de energia em máquinas de contrução.

Objetivos e Perguntas de Pesquisa

O objetivo desta tese é avaliar dois métodos para a obtenção de estratégias de gerenciamento de energia para máquinas de construção. Os métodos são baseados em otimização e aprendizado de máquina. É esperado que a combinação desses dois tipos de técnicas resulte em um aumento de eficiência das máquinas. Os dois métodos são avaliados com base nas estratégias de gerenciamento de energia quando do cumprimento da funcionalidade esperada, do aumento da eficiência da máquina e do atendimento de critérios de robustez e segurança.

Para alcançar este objetivo, esta tese é guiada pelas seguintes perguntas de pesquisa:

RQ1. Como estratégias de gerenciamento de energia baseadas em aprendizado de máquina podem ser obtidas para máquinas de construção?

RQ2. Que melhoria de eficiência pode ser esperada para máquinas de construção, quando operando com estratégias de gerenciamento de energia baseadas em aprendizado de máquina?

RQ3. Estratégias baseadas em aprendizado supervisionado e aprendizado por reforço, usando redes neurais como representação de funções, podem superar os desafios relacionados à arquiteturas de sistemas e operação de máquinas de construção?

RQ4. Quais vantagens e desvantagens se pode esperar de métodos baseados em aprendizado de máquina para o gerenciamento de energia em máquinas de construção?

Metodologia

O primeiro método avaliado é baseado em aprendizado supervisionado envolvendo a combinação de programação dinâmica e redes neurais.

Esta abordagem é aplicada para o controle de uma carregadeira híbrida de rodas. O segundo método é baseado em aprendizado por reforço também utilizando redes neurais. Esta abordagem é aplicada para o controle do braço de uma escavadeira. Cada método é avaliado em estudo de caso específico.

Os dois métodos têm uma fase inicial de aprendizado em simulação utilizando modelos para representar o comportamento dos sistemas físicos. Após o treinamento, os controladores são implementados diretamente nas máquinas para avaliação experimental.

O desenvolvimento desta tese envolve modelagem de sistemas físicos, simulação, otimização, aprendizado de máquina e avaliação experimental.

A robustez das estratégias de gerenciamento de energia é avaliada por meio de comparação entre os domínios de desenvolvimento e de aplicação.

Estudos de caso e resultados

Dois estudos de caso foram elaborados para permitir a avaliação dos métodos.

O primeiro estudo de caso consiste na avaliação do método que combina programação dinâmica e redes neurais. Ele é aplicado para o controle do sistema híbrido de uma carregadeira de rodas. O controle desse sistema permite a recuperação de energia cinética durante a frenagem e posterior uso para propulsão. Dessa maneira, o consumo de combustível pode ser reduzido.

Com base em um modelo suficientemente representativo do comportamento da máquina, vários ciclos de trabalho são utilizados como entrada para a otimização do controle do sistema híbrido. Programação dinâmica é a técnica utilizada para encontrar a solução ótima de como este sistema deveria ser controlado para minimizar o consumo de combustível durante os ciclos de trabalho. A otimização é restringida por regras que garantem a segurança da operação. Por exemplo, a ação de frenagem não pode ser passível de erros, e por conta disso, é implementada através de regras.

Uma rede neural é treinada para encontrar a função que mapeia o estado do sistema para a variável ótima de controle. A rede treinada, junto com regras adicionais, implementa a estratégia de controle.

Esse controle é implementado na máquina para avaliação experimental. Nestes experimentos, um operador profissional opera a máquina em ciclos de trabalho similares ao que a rede foi treinada. Resultados mostram uma eficiência superior à uma estratégia baseada em regras desenvolvida por engenheiros da empresa parceira. Dessa maneira, foi confirmada a capacidade desse tipo de método de encontrar automaticamente estratégias de controle otimizadas e de implementá-las nas máquinas. Ao mesmo tempo, foi percebida a necessidade de continuar o treinamento da rede após aplicada ao sistema real para corrigir diferenças entre os domínios de desenvolvimento e aplicação.

O segundo estudo de caso, consiste na avaliação do método baseado em aprendizado por reforço. A motivação para estudar esse tipo de método surge da capacidade deles de interagirem com o sistema enquanto aprendem a controlá-lo. Isso permite a continuidade do treinamento com o sistema real para corrigir diferenças entre os domínios de desenvolvimento e aplicação.

Este método foi aplicado para a seleção de modos de operação de um atuador hidráulico multicâmaras que faz parte do sistema de atuação do braço de uma escavadeira. A escolha de diferentes modos de operação permite a redução de perdas energéticas no conjunto de válvulas que controlam os atuadores. Dessa forma, o objetivo é, para cada estado do sistema, encontrar o modo que resulta em menores perdas energéticas.

Assim como no caso anterior, o agente é treinado em um ambiente de simulação onde, interagindo com o modelo do sistema, aprende automaticamente como controlá-lo segundo o objetivo desejado. Neste caso, também são adicionadas regras que garantem a operação segura do sistema.

Após treinado, o agente é aplicado para controlar o sistema real. Os experimentos mostram a capacidade do método de encontrar uma solução otimizada e de implementá-la diretamente no sistema real. Também é percebido uma elevada robustez à situações não necessariamente vistas durante o processo de treinamento. Também é visto a necessidade de continuidade do treinamento após aplicado.

Contribuições

As principais contribuições desta tese são:

- Demonstração da performance de máquinas de construção operando com estratégias de gerenciamento de energia baseadas em aprendizado de máquina;
- Avaliação de um método de aprendizado supervisionado e de um método de aprendizado por reforço;
- Uma análise da segurança e robustez de controladores baseados em aprendizado de máquina para gerenciamento de energia em máquinas de construção;
- Um estudo mostrando a importância de considerar partes da estrutura de controle da aplicação real já no processo de geração de dados; e

- Demonstração de que métodos baseados em aprendizado de máquina, usando dados de modelos de simulação para treinamento, podem resultar em controladores que operam bem na prática.

Conclusões

Esta tese avaliou o uso de aprendizado de máquina como o meio para aprender e implementar estratégias de controle para gerenciamento de energia em máquinas de construção. Dois métodos foram utilizados, um baseado em aprendizado supervisionado e outro baseado em aprendizado por reforço.

Os dois métodos foram avaliados desde o desenvolvimento até a implementação em experimentos, sendo um em uma máquina real e outro em uma bancada de testes. Tópicos relacionados à robustez, confiabilidade e performance foram abordados. Os dois estudos de caso forneceram as informações para responder as perguntas de pesquisa:

RQ1. Como estratégias de gerenciamento de energia baseadas em aprendizado de máquina podem ser obtidas para máquinas de construção?

Redes neurais, treinadas sob um abordagem de aprendizado supervisionado ou sob uma abordagem de aprendizado por reforço, possuem a capacidade de automaticamente aprender estratégias de controle optimizadas e de implementar elas diretamente nas máquinas como a estratégia de gerenciamento de energia. Entretanto, devido ao fato de o aprendizado inicial ser baseado em modelos, pode existir a necessidade de continuar o treinamento depois de implementadas para adaptar a estratégia ao sistema real.

RQ2. Que melhoria de eficiência pode ser esperada para máquinas de construção, quando operando com estratégias de gerenciamento de energia baseadas em aprendizado de máquina?

Os resultados dessa tese não evidenciaram quais estratégias de controle baseadas em aprendizado de máquina possuem limitações para aprender e implementar estratégias de controle. Dessa maneira, esses tipos de controladores podem fazer com que as máquinas operem mais próximas à sua eficiência teórica se comparados com estratégias baseadas em regras.

RQ3. Estratégias baseadas em aprendizado supervisionado e aprendizado por reforço, usando redes neurais como representação de funções, podem superar os desafios relacionados a arquiteturas de sistemas e operação de máquinas de construção?

Os métodos estudados nesta tese mostraram-se robustos à aplicação em máquinas de construção quando operando em condições similares às de treinamento. Isso significa que, eles são robustos a diferenças nos modelos e ainda assim capazes de operar em ambientes levemente diferentes daqueles que foram treinados. Entretanto, foi identificada a necessidade de continuar o trainamento após a implementação para aumentar a robustez. Também foi mostrado que eles são capazes de aprender as relações complexas entre variáveis do sistema e as decisões de controle para máquinas de construção.

RQ4. Quais vantagens e desvantagens se pode esperar de métodos baseados em aprendizado de máquina para gerenciamento de energia em máquinas de construção?

Adaptabilidade e segurança são os principais pontos de preocupação quanto à aplicabilidade deste tipo de método para gerar estratégias de gerenciamento de energia para máquinas de construção. As estratégias precisam ser acompanhadas de regras para garantir a segurança em todos os cenários de operação. Por outro lado, elas são capazes de encontrar automaticamente estratégias de gerenciamento de energia para máquinas de construção e aplicá-las diretamente nos sistemas com um nível considerável de robustez a diferenças entre os ambientes de desenvolvimento e de aplicação.

Palavras-Chave: Aprendizado de Máquina, Gerenciamento de Energia, Máquinas de Construção

Acknowledgements

The work presented in this thesis was undertaken at the Laboratory of Hydraulic and Pneumatic Systems – Federal University of Santa Catarina, and at the Division of Fluid and Mechatronic Systems – Linköping University. I am thankful to both institutions for providing this opportunity, and I recognise the financial support from the Brazilian Coordination for the Improvement of Higher Education Personnel (CAPES), the Brazilian National Council for Scientific and Technological Development (CNPq), and the Swedish Energy Agency (Energimyndigheten).

I am thankful to Prof. Victor J. De Negri for leading me to lands abroad to discover a whole world of possibilities. I am also thankful to Prof. Petter Krus for opening the doors of Sweden and for the trust in this partnership. Thank you very much Dr. Liselott Ericson for all the enriching discussions, guidance, and support during my studies. Thank you, Rita, for the all the support.

I am thankful to Volvo Construction Equipment, Eskilstuna, and especially to the people I met there. Karl Uebel, Kim Heybroek, Martin Rohdin, and Taerang Jung, your support and friendship along this thesis were invaluable.

I am extremely happy to have shared my daily work life with my colleagues at LASHIP since even before the start of this thesis. I am also happy to have shared the many fruitful conversations and friendship with my colleagues at FLUMES. A special thanks to Marcos, Dimitri, Vinícius, Artur, Christopher, Felix, Samuel, Viktor, Robert, Thomas, Ludvig, Heitor, and Jorge.

Esta tese não teria sido finalizada sem o suporte e motivação da minha família. Pai, Mãe, Vó Sonja, Tia Mica, Kika, Huf, Memi, Camila, Anna Laura, Benjamin, Theodoro, Maria, Seu Luiz, Dona Iva, José e Grazi, Muito Obrigado!

More important than anything is to share the beauty of life with the ones we love. Thank you so much, Shara, my beloved wife, for all the help and memorable moments we live together!

Thank you, God, for all the support provided along this journey.

Linköping, August 2022
Henrique Raduenz

Abbreviations

AI	Artificial Intelligence
AUX	Auxiliary Functions
AD	Application Domain
C	Controller
CN	Network trained on data from constrained optimisation
CONV	Conventional
CRS	Complementary Energy Recuperation and Storage System
DP	Dynamic Programming
DQN	Deep Q-Network
DT	Drivetrain
ECMS	Equivalent Consumption Minimisation Strategy
EMS	Energy Management Strategy
ES	Energy Storage
FWD	Forward
LiU	Linköping University
ML	Machine Learning
MPC	Model Predictive Control
NN	Neural Network
O	Operator
PM	Prime Mover
ReLU	Rectified Linear Unit
REV	Reverse
RB	Rule-Based
RL	Reinforcement Learning
SDP	Stochastic Dynamic Programming
SLC	Short Loading Cycle
SOC	Accumulator State of Charge
TD	Training Domain
UFSC	Federal University of Santa Catarina
UN	Network trained on data from unconstrained optimisation
WH	Work Functions/Hydraulics

Nomenclature

Variable	Description	Unit
a	Action	-
A	Sampled points from a distribution	-
A	Area	m^2
b	Bias	-
B	Sampled points form a distribution	-
D_{PM}	Pump-motor displacement	m^3/rad
E	Error	-
E	Energy	J
$\bar{E}$	Average error	-
f	Function	-
g	Function	-
g	Cost function of the optimal control problem	-
h	Function describing states to optimal control decision	-
F	External load	N
F	Traction force	N
i	Gear ratio	-
k	Discrete time step index	-
K	Constant	-
KD	Kick-Down	-
SOC	State of charge	-
ϵ_{PM}	Pump-Motor displacement setting	-
n	Number of points in a bin	-
n	Rotational Speed	rad/s
p	Probability	-
p	Pressure	Pa
P	Power	W
P	Population	-
Q	Flow rate	m^3/s
Q	Value function	-
r	Reward	-
R	Uniformly random variable	[0, 1]
s	Data set	-
$\boldsymbol{s}$	Observations	-
t	Time	s
T	Torque	Nm

u	Control input	-
$\boldsymbol{u}$	Vector of control actions	-
v	Speed	m/s
x	Position of the multi-chamber actuator	m
x	State variable	-
x	Input feature	-
$\boldsymbol{x}$	Vector of state variables or input features	-
x_{Brake}	Brake pedal position	m
x_i	Feature input	-
x_{NN}	Neural network input features	-
y	Output	-
ω_i	Neuron weights	-
Δp	Pressure difference	Pa
Δp	Difference in probability of occurrence	-

Subscripts	**Description**	**Unit**
$0, 1, 2, 3, \ldots$	Variable/Parameter/Load 0, 1, 2, 3, ...	-
A, B, C, D	Chambers of multi-chamber actuator	-
avg	Average	-
DV	Digital Valves	-
i	Index of weights and biases	-
ICE	Internal combustion engine	-
$loss$	Energy loss	-
LS	Load Sensing	-
M	Model	-
max	Maximum	-
min	Minimum	-
n	Final index of weights and biases	-
N	Final time step	-
NN	Neural Network	-
$Norm$	Normalisation	-
$Opti$	Optimal	-
PM	Pump/Motor	-
$Power$	Power	-
$pred$	Predicted	-
R	Real machine	-
R	Return line	-
ref	Reference	-
s	Supply	-
SLC	Short Loading Cycle	-

$Switch$	Switch counter for modes	-
t	Tank	-
$Test$	Relative to the test	-
tra	Transmission	-
$Train$	Relative to training	-
$Velocity$	Velocity	-
WH	Work hydraulics	-
$wheel$	Wheel	-

Superscripts	**Description**	**Unit**
o	Optimised	-
$*$	Ground truth	-
$'$	Next state or next observation	-

Papers

The following publications are included in the thesis and can be regarded as its foundation. They are referred by their Roman numerals throughout the text. Apart from formatting changes and minor errata, they are reproduced in their original form. At the time of publication of this thesis, Paper III and IV were submitted for publication.

I. H. Raduenz, L. Ericson, K. Uebel, K. Heybroek, P. Krus, and V. J. De Negri, "Energy management based on neural networks for a hydraulic hybrid wheel loader," in *2020 IEEE Global Fluid Power Soc. PhD Symp.* [Online], 2020.

II. H. Raduenz, K. Uebel, K. Heybroek, L. Ericson, V. J. De Negri, and P. Krus, "Rule- and neural network-based energy management for a hydraulic hybrid wheel loader," in *The 13th Int. Fluid Power Conf.*, Aachen, Germany, Jun. 13-15, 2022.

III. H. Raduenz, K. Uebel, K. Heybroek, L. Ericson, P. Krus, and V. J. De Negri, "Performance evaluation of neural network-based energy management for a hybrid wheel loader," in *IEEE Trans. Veh. Technol.*, 2022. (Submitted).

IV. H. Raduenz, L. Ericson, V. J. De Negri, and P. Krus, "Multi-chamber actuator mode selection through reinforcement learning - simulations and experiments," in *Energies*, 2022. (Submitted)

V. K. Uebel, H. Raduenz, P. Krus, and V. J. De Negri, "Design optimisation strategies for a hydraulic hybrid wheel loader," in *ASME/BATH 2018 Symp. on Fluid Power and Motion Control*, Bath, UK, Sept. 12-14, 2018, doi: 10.115/FPMC2018-8802.

VI. H. Raduenz, L. Ericson, K. Heybroek, V. J. De Negri, and P. Krus, "Extended analysis of a valve-controlled system with multi-chamber actuator," in *Int. J. of Fluid Power*, vol. 23, no. 1, 2021, pp. 79-108, doi: 10.13052/ijfp1439-9776.2314.

VII. T. Jung, H. Raduenz, P. Krus, V. J. De Negri, and J. Lee, "Boom energy recuperation system and control strategy for hydraulic hybrid excavators," in *Automat. in Const.*, vol. 135, 2022, doi: 10.1016/j.autcon.2021.104046.

The author of this thesis is the main author for all appended papers and has been responsible for the development of the studies and writing, except for papers V and VII. The co-authors have had a supervisory role. In papers V and VII the author provided support with simulations, analysis of results, and writing.

Additional Publications

The publications below are not central for the thesis but are connected to the topic.

VIII. H. Raduenz, L. Ericson, K. Heybroek, V. J. De Negri, and P. Krus, "Improving the efficiency of valve-controlled systems by using multi-chamber actuators," in *The 17th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, 2021, ISBN: 978-91-7929-013-9.

IX. M. P. Nostrani, H. Raduenz, A. Dell'Amico, P. Krus, and V. J. De Negri, "Multi-Chamber Actuator Using Pump for Position and Velocity Control Applied in Aircraft," in *2020 IEEE Global Fluid Power Soc. PhD Symp.* [Online], 2020, ISBN: 978-1-7281-4138-1.

X. H. Raduenz and V. J. De Negri, "Speed Compensation in Hydraulic Wind Turbine Control," in *2018 IEEE Global Fluid Power Soc. PhD Symp.*, Samara, Russia, 2018.

XI. H. Raduenz, Y. E. A. Mendoza, D. Ferronatto, F. J. Souza, P. P. Da Cunha Bastos, J. M. C. Soares, V. J. De Negri, "Online fault detection system for proportional hydraulic valves," in *J. Brazilian Soc. of Mech. Sci. and Eng.*, vol. 40, 2018.

XII. M. P. Nostrani, A. Galloni, H. Raduenz, and V. J. De Negri, "Design and Optimization of a Fast Switching Hydraulic Step-Down Converter for Position and Speed Control," in *The 15th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, 2017.

XIII. H. Raduenz, F. J. Souza, P. P. C. Bastos, D. Ferronatto, V. J. De Negri, and J. M. C. Soares, "Evaluation of an on-line fault detection method for proportional hydraulic valves," in *9th PhD Symp. on Fluid Power, FPNI2016,* Florianópolis, Brazil, 2016.

XIV. M. Nostrani, A. Galloni, H. Raduenz, and V. J. De Negri, "Theoretical and experimental analysis of a hydraulic step-down switching converter for position and speed control," in *The 8th Workshop on Digital Fluid Power*, *DFP2016,* Tampere, Finland, 2016.

XV. D. Ferronatto, Y. E. A. Mendoza, P. P. C. Bastos, F. J. Souza, H. Raduenz, V. J. De Negri, J. M. C. Soares, "Proportional hydraulic valve condition monitoring method for on-line fault detection," in *23rd ABCM Int. Congr. of Mech. Eng., COBEM2015*, Rio de Janeiro, Brazil, 2015.

XVI. E. Flesch, H. Raduenz, and V. J. De Negri, "Analysis of a hydrostatic transmission system for horizontal axis wind turbine," in *The 14th Scand. Int. Conf. in Fluid Power*, Tampere, Finland, 2015.

Contents

1

Introduction

This thesis is about the evaluation of methods based on machine learning to obtain energy management strategies for construction machines. More specifically, it addresses the development and evaluation of strategies based on two methods to learn and implement control decisions in real machines.

The global move towards meeting climate change goals is pushing the development of greener and more efficient systems. This is no different for mobile working machines since they are considered a fundamental part of the problem to be solved to meet those goals.

Mobile working machines, such as excavators, wheel loaders, haulers, and forwarders, are characterised by a powertrain architecture not only dedicated to the translational motion but also the control of work functions to handle heavy loads. Due to the weights of the machines and handled materials, they have high kinetic and potential energy. Hydraulics has been the primary technological solution for motion control in such machines, and therefore, it is one of the major power consumers. Typically, hydraulic systems of mobile machines have low efficiencies. Part of the solution to reducing the environmental impact of their operation involves the development of improved hydraulic systems that might include the capability to recover the available kinetic and potential energies.

The capacity to recover, store, and/or reuse energy results in added control degrees of freedom that need to be controlled to perform so-called energy management. Energy management strategies define when, how much, and where energy is generated, recuperated, stored, and reused. It manages the energy shared between multiple sources to multiple consumers. Therefore, a holistic view of the machine, its subsystems, and operation tasks is needed to improve the energy management. The energy efficiency gains are, therefore, a consequence of the system architecture and energy management strategy (EMS).

There are already several methods to generate EMSs for on-road hybrid vehicles, for example: Dynamic Programming, Equivalent Consumption Minimisation Strategy, Rule-based, Fuzzy Logic, Model Based Control, and Machine Learning. The EMSs obtained through such techniques differ in terms of their capacity to be implemented online, optimality, required computational power, capability to include various operational conditions and scenarios, adaptability, and easiness of realisation for complex systems.

The most widely used method for obtaining EMSs for construction machines is the rule-based one due to its high reliability and easy realisation. However, rule-based EMSs suffer from sub-optimality and are not easily developed for more complex machines, usually resulting in high development time and increased sub-optimality.

In this sense, construction machines with improved hydraulic systems and/or energy recuperation capabilities, have a higher capacity for efficiency improvement but are the most challenging ones regarding the design of EMSs. Thus, the maximum energy efficiency of construction machines will most likely not be achieved with rule-based EMSs, but with other techniques that can handle the complexity of their architectures, tasks, and working environments while maintaining their performance. Currently used methods to generate EMSs for on-road hybrid vehicles might be applicable to construction machines but it is not certain that they will extract their highest performance.

One category of methods, explored for on-road vehicles, uses machine learning techniques to generate the EMS or as an additional technique to improve their performance, for example, through prediction. Not only in the realm of control machine learning has proven capable of automatically learning complex functions from a set of data. When using machine learning, one is interested in their capability to generalise the learned knowledge for scenarios not seen during training.

Methods based on machine learning seem to be up to the challenges and requirements posed by energy management in construction machines. It is expected that they can learn and implement control decisions that result in higher machine efficiency. At the same time, it is expected that they can generalise the learned control strategies for other similar scenarios. This would be beneficial for energy management in construction machines because the tasks and load scenarios they experience during operation are vast. If the strategies yield increased machine efficiency and are robust when deployed to the real system, an additional advantage would be the capability to automatically learn a control strategy, which could possibly reduce the development time of such machines.

Despite the potential of this type of method, there are no extensive studies demonstrating, in prototypes, their applicability. In this way, a specific and comprehensive research on methods based on machine learning to obtain energy management strategies for construction machines is missing.

1.1 Aim and Research Questions

The aim of this thesis is to evaluate two methods based on machine learning to generate controllers for energy management in construction machines. The focus is on achieving energy efficiency improvements in the operation of the machines and on automating part of the development process of such controllers. Machine learning is selected due to its capability to automatically learn complex knowledge representations from a set of data and due to its generalisation capability.

One of the methods is supervised learning, where the learning target is the output of control optimisations with dynamic programming. The other method is reinforcement learning, where the energy management strategy is learnt while the agent interacts with the system.

The methods are evaluated based on the capacity of the resultant energy management strategies to fulfil system functionality and increase machine efficiency, as well as aspects related to robustness and safety. The last two points for evaluation are included because the methods are data-driven approaches that might suffer drawbacks in prediction performance when deployed to domains for which they were not trained.

To achieve the above aim, this thesis is guided by the following research questions:

RQ1. How can machine learning-based energy management strategies be obtained for construction machines?

RQ2. What efficiency improvements can be expected from construction machines when operating with machine learning-based energy management strategies?

RQ3. Can supervised learning and reinforcement learning-based methods, using neural networks as function representation, overcome the challenges related to system architecture and operation of construction machines?

RQ4. What advantages and drawbacks can be expected from machine learning-based methods for energy management in construction machines?

1.2 Methodology

To demonstrate that machine learning can be used for energy management in construction machines, the two described methods are used aiming to generalise two possible approaches to the goal.

With supervised learning the case addressed concerns a situation where the developer has the information on how the system should be controlled but does not have the means to implement it. It also covers the case where the information exists but other methods (e.g., through rules) result in poor machine performance. The information on how to control the system could, for example, come from control optimisations. This method is applied to a hydraulic hybrid wheel loader in the first case study.

With respect to reinforcement learning the case addressed involves a scenario where the developer does not have the information on how the system should be controlled. This method can learn how to control the system by interacting with it. This approach is applied to an excavator arm in the second case study.

Neural networks are used as function representation in both methods. In the supervised learning case, the network predicts the control action, while in the reinforcement learning case, it predicts the expected value for taking a control action, which allows for the selection of the best action. Due to their high capacity to map complex representations from input to output, it is expected that neural networks, trained under these two methods, will overcome the challenges of operation and system architectures of construction machines. In other terms, it is expected that they can handle the high nonlinearity and dimensionality associated with the control problem. Consequently, it is expected that these controllers can yield improved machine efficiency.

Both methods have an initial learning process performed offline in simulation, having models as the representation of the physical systems' behaviour. The models are used to generate the information of how the systems should be controlled. After training on this data, these trained controllers are directly deployed to the machines for experimental evaluation.

The development of this thesis involves modelling the physical systems, simulation, optimisation, machine learning, and experimental evaluation.

The robustness of the resultant EMSs is assessed by comparing the shift between the training domain and application domain.

1.3 Delimitations

Although generalising for other types of construction machines with different subsystems, like electric driven subsystems, this thesis uses as test cases hydraulic systems of combustion engine-driven construction machines. It is understood that the same techniques could be applied to other types of subsystems like battery and fuel cell hybrids.

The thesis is also limited to evaluating the two mentioned approaches for the development of energy management strategies. Therefore, the outcome of the thesis is not an argument concerning what is the best method for the development of energy management strategies for construction machines. Rather, it presents an extended analysis of two possible methods and presents arguments in favour or against them, mostly with focus on their application to real systems.

This thesis does not aim to find what the best machine learning model is for this type of application. Instead, it evaluates, on a higher level, the applicability of the two methods with networks of a size sufficient for the tasks. In this sense, it is understood that potential improvements can be achieved by a dedicated study focused on the machine learning part.

The training of the strategies does not continue after they are deployed to control the real systems. Even though continued training could possibly increase their performance and robustness, this is not done in this thesis.

1.4 Contribution

The main contributions of this thesis are summarised as follows:

- Demonstration of the performance of construction machines operating with machine learning-based energy management strategies;
- Evaluation of a supervised learning method and of a reinforcement learning method;
- Demonstration that machine learning-based methods, using data from simulation models for training, can result in controllers that perform well in practice;
- An analysis of the safety and robustness of machine learning-based controllers for energy management in construction machines in real application;
- A study showing the importance of considering parts of the control structure from the real application already in the data-generating process;
- Increased comprehensiveness on the problem of energy management for construction machines.

The structure of the content of the thesis and their connection with the papers is represented in Figure 1.1.

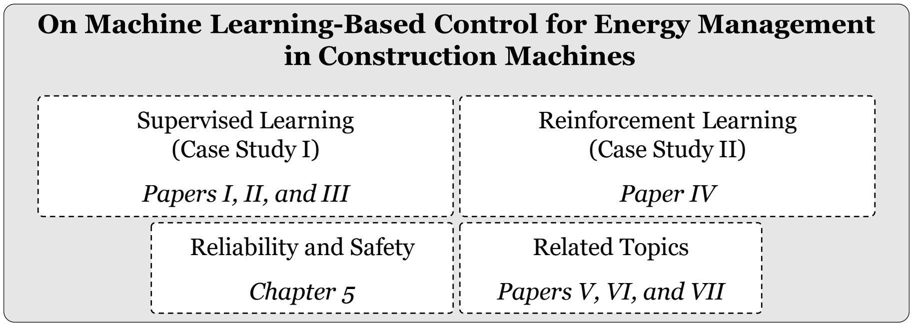

Figure 1.1 *Content of the thesis.*

1.5 Thesis Outline

The thesis was developed in a double degree and co-supervision format between Linköping University (LiU), Linköping, Sweden, and the Federal University of Santa Catarina (UFSC), Florianópolis, Brazil. The initial parts, including part of the literature review and the thesis proposal that led to the doctorate qualification, were developed at UFSC. The remaining phases were developed at LiU. This thesis had Volvo Construction Equipment AB in Eskilstuna, Sweden, as an industrial partner.

Chapter 2 presents a discussion on the complexity of construction machines and the problem of generating energy management strategies for them. Chapter 3 describes methods for energy management with a deeper focus on the techniques used in this thesis, while Chapter 4 describes the study cases performed to allow the assessment of the methods. Chapter 5 presents an analysis of the reliability of the generated EMSs, with Chapter 6 presenting the summary of the papers written on the topic. Finally, Chapter 7 presents discussions and future studies, followed by conclusions in Chapter 8.

2

Construction Machines

Construction machines are versatile equipment used in construction, mining, farming, and several other segments. Their scalability, ranging from few to hundreds of kilowatts, is a characteristic allowing them to meet a great number of needs. The tasks they perform can be succinctly described, but not limited to, moving a load from one point to another in an environment. Nevertheless, it is not a simple task to be accomplished. It involves the interaction between machine subsystems under operator control, handling them against variable and uncertain external loads. Therefore, they are challenging tasks that can be executed in different forms. At the same time, there are several possible subsystem architectures that result in machines capable of performing those tasks. Efficient operation has thus become a consequence of task, machine, and control.

2.1 Power Flow Between Machine Subsystems

Differently from on-road vehicles, where the power has one main path, in construction machines the power can have multiple paths. Power consuming subsystems can be divided into work functions (WH), also called work hydraulics, used to move the loads; drivetrain (DT), used for propulsion; and auxiliary functions (AUX), used for cooling, etc.

In conventional machines, the power to those functions is supplied by the internal combustion engine (ICE) with a fuel tank as energy storage (ES). An architecture of a conventional machine is shown in Figure 2.1.

Conventional machines do not usually have an installed capability to recover the kinetic energy from the drivetrain or the potential energy from the work functions. This results in comparatively simpler architectures with respect to how power can flow between subsystems. Additionally, the functions are usually mechanically coupled to the engine.

Hybrid machines, on the other hand, are characterised by two or more energy storage devices. Hybridisation aims to reduce energy consumption and/or reduce emissions through the combined use of several sources of energy to meet the power demanded from consumers [1]. Typically, one of the energy sources can perform energy recuperation. Therefore, lower energy consumption from the primary source is a result of the recuperation of kinetic and/or potential energy and decoupling between different power sources and consumers for more efficient individual operation, in addition to the maximum utilisation of installed power capacity.

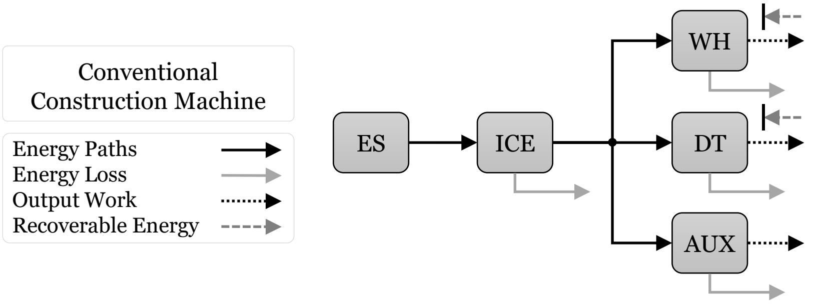

Figure 2.1 *Energy paths in a conventional construction machine.*

In construction machines, hybridisation targets either the drivetrain only or the work functions only, or both. The matrix representation in [2] provides an overview of possible architectures of hybrid construction machines. Figure 2.2 shows a hypothetical concept of a hybrid construction machine.

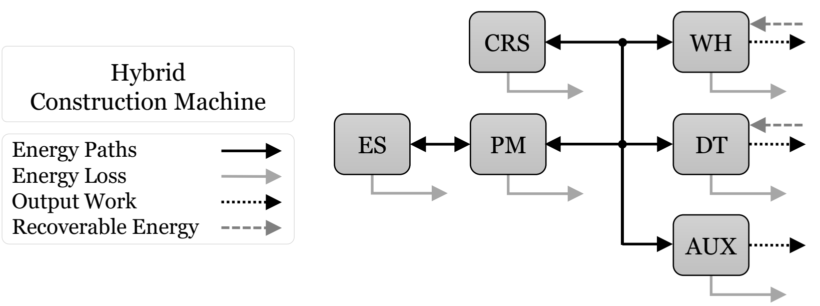

Figure 2.2 *Energy paths in a hypothetical hybrid construction machine.*

In this concept, the complementary energy recuperation and storage system (CRS), also called the secondary energy source, is connected to both the drivetrain and work functions. It could be a hydraulic accumulator, for example. To have a broader coverage, the engine is called prime mover (PM). This allows electric driven concepts to fit into this representation, which in this case, allows the main energy storage system to also recuperate energy. The represented paths for energy exchange between subsystems can be mechanical, hydraulic, electric, or a mix of them. This concept allows for the following additional energy exchange between subsystems:

- Kinetic energy from drivetrain to drive the work functions,
- Kinetic energy from drivetrain to charge the storage systems,

- Potential energy from work functions to charge the storage systems, and
- Potential energy from work functions to drive the drivetrain.

This is an example of the higher complexity level that can be involved in terms of system architecture and energy exchange possibilities. This is especially so if compared to on-road vehicles. In general terms, the degrees of freedom concerning how power can flow between subsystems are increased. As a result, the potential for energy efficiency improvement in such machines is high.

Hybridisation, or at least the capability to recover available energy, is a trend in the development of construction machines as can be noticed from review papers on the topic [3-8]. However, the architecture complexity depicted in Figure 2.2 is uncommon. What is more common are: architectures where one of the subsystems (WH or DT) is hybridised; architectures designed from the beginning to allow for the recovery of available energy (not necessarily hybrid); and more efficient subsystems, with more degrees of freedom, to drive the loads.

2.2 Examples of Machine Architectures

Series hybrid drivetrains allow a full decoupling of the engine from the wheels. Consequently, there is certain freedom to combine engine speed and torque to meet the demanded power. Therefore, high efficiency operation points can be set for the engine [9]. At the same time, the energy stored in the secondary source can be used at specific moments to reduce engine power requirements while meeting the necessary power demand [10]. In series hybrid wheel loaders, the decoupling between drivetrain and work functions is a major contribution for fuel savings [11]. This is because the power to each function can be controlled according to the respective demand. Examples of series hybrid mobile machines can be found in [12-16].

Parallel hybrid drivetrains do not necessarily decouple the engine from the drivetrain. They allow for the recovery of kinetic energy when braking and return it to the drivetrain in specific moments to reduce engine power demand. This configuration only allows engine operation point management when the secondary power source is delivering power. Hydraulic hybrid parallel systems are attractive for applications with frequent start-stop, which is the case for wheel loaders operating in short loading cycles. At least at the research level, parallel hybrid wheel loaders were studied in [10, 14, 17-19]. A parallel hybrid system that can recover energy from work functions and drivetrain is presented in [20].

Power-split hybrid drivetrains for wheel loaders are studied in [21] and [22]. In [21], the authors state that power-split drivetrains allow for the management of the engine's operation point and regenerative braking; they can also reduce interferences between drivetrain and work functions, which can improve fuel efficiency.

A considerable amount of hybrid architectures was also proposed for the recovery of potential energy from the linear actuation in work functions. In this case, the potential energy is recovered from overrunning loads instead of being throttled in control valves, for example.

Hydraulic transformers to recover the energy from loads and deliver it back to the same subsystems are also a possibility for energy efficiency improvement [23-27]. They can be seen as 'add-on' systems where a hydraulic machine is coupled at the meter-out port of the actuator. This machine, working as a motor, drives another hydraulic machine, working as pump, to charge an accumulator. The reverse process occurs when returning the energy to the system. An advantage of transformers is that they allow the decoupling of pressure levels between storage and load, whereas a drawback is the added pump/motor energy losses.

There are concepts where the recovered energy from the work functions is used to assist the engine. Excavators are the most common application for recovering potential energy from the boom and kinetic energy from the swing motion. Examples of such architectures are found in [28-31]. Such concepts reduce the interaction between the storage and control of the load when returning the recovered energy to the system. Essentially, the control of the actuation systems remains the same, except for the overrunning loads when recovering energy.

Figures 2.1 and 2.2 are high-level simplified representations of possible architectures for such machines, and the concepts presented up until this point mainly accomplish energy efficiency improvement by recovering available energy. However, energy efficiency improvement does not come only from hybridisation.

As indicated in [32], and also discussed in [33], energy efficiency improvement can also come from more efficient actuation systems, which also contain means of recovering energy but without an 'add-on' system. This usually requires a redesign of the powertrain with that requirement from beginning.

For the work functions, it is common to tackle throttling losses of valve-controlled systems with throttle-less actuation systems. A concept with this approach is presented in [34] and shown in Figure 2.3. A similar approach, but for an excavator, is presented in [35]. The linear actuation system is based on the secondary control concept presented in [36] and also studied in [37] for the linear actuation system of an excavator arm.

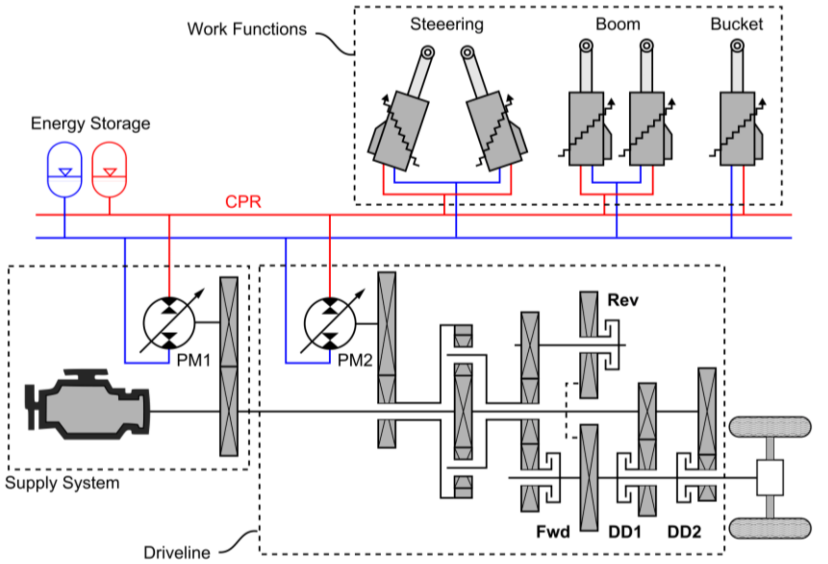

Figure 2.3 *Architecture of a wheel loader concept. Reproduced with permission from [34], Karl Pettersson, A Novel Hydromechanical Hybrid Motion System for Construction Machines, Copyright Karl Pettersson, 2016.*

The advantage of secondary controlled systems is the reduced throttling losses in comparison to valve-controlled systems, like load sensing. It also allows for the decoupling between loads and independent control of each port of the actuator.

Independent meter-in systems, based on the concept of common pressure rail, are presented in [38] and [39]. In similar way, they allow for the independent control of each actuator port to reduce throttling losses.

Another way of reducing throttling losses is by means of displacement-/pump-controlled systems. As the name indicates, the linear actuator is directly controlled by a variable displacement/speed pump/motor. An architecture with displacement-controlled actuators is shown in [40]. The system presented in [40] is still driven by the engine, thus there is a strong coupling between the actuation of each function. However, there are recently proposed concepts that take advantage of machine electrification and have independently driven actuation systems [41].

These concepts of displacement-/speed-controlled actuators are similar to the ones presented in [34] and [35] in the sense of having a common power distribution medium that is mechanic, hydraulic, or electric. In these cases, the energy from overrunning loads can be supplied to the power distribution medium to be used by other actuators or to charge the energy storage.

2.3 Architecture vs. Energy Management

In [42], the high variability of construction machines operation efficiency due to variability among operators is shown. In [43] and [44], it is shown that the optimisation of how the tasks are performed has also a significant contribution to the energy efficiency.

It is also known that efficiency is a consequence of the sizing of components and that optimisation of components sizes and control should be carried together in early development stages [1; 10; 11; 45-46]. In simple terms, the highest efficiency is only found with an optimal EMS and optimal component sizing. However, an optimal component sizing will only be found with an optimal EMS. The problem is coupled and requires a combined EMS and component sizing optimisation.

Therefore, energy efficiency improvement or the optimisation of a construction machine energy consumption is a consequence of the interaction between operator, task, system architecture, component sizing, and control. It is a multi-level task that ideally should be solved in one general optimisation problem with many degrees of freedom. However, this is a difficult problem that so far has been solved in part by constraining or locking some of the degrees of freedom. In this thesis, the focus is on the part of energy management with constrained system architecture, task, operator, and sizing of components. Despite the constraints it is still not a trivial problem.

The previous sections highlighted that there are many possible system architectures for construction machines and their subsystems. At the same time, it was highlighted how complex the energy exchange in, and/or between, the subsystems can be. It is also noticed that complex architectures and subsystems seem to inherently have the higher potential for energy efficiency gain. However, despite their installed potential for efficiency improvement, this is not achieved without a proper energy management. The strategy for energy management is vital to achieve the expected high efficiency improvement [5], [47].

An important point is observed from the assessment of possible machine architectures. The efficiency improvement can be dealt with at a machine level with 'add-on' systems or in the subsystem level, with, for example, displacement-controlled subsystems. In the first case, the control of the 'add-on' system is more focused on increasing the machine efficiency than on meeting the actuation control performance. In the second case, the control of the subsystem is more focused on meeting an actuation control performance than efficiency. Therefore, there is a duality in the control strategies. This means there must be a focus on control actuation performance and the energy efficiency aspect. How clear the division is between them seems to vary between one architecture to another.

As it can be noticed, the design of the energy management strategy for such machines and/or subsystems is not a simple task. This is due to the complexity of the task, the architecture, and the integration with the actual actuation control. This drives the focus of the thesis towards evaluating two methods to generate energy management strategies. In principle, it is assumed that they can handle the complexity around the operation and control of construction machines and their subsystems.

3
Methods for Energy Management

Energy management is a fundamental part of obtaining a high efficiency operation from the machines described in the previous chapter. The EMS defines where, how much, and when the power from different sources is split to fulfil the power requested by different functions of the machine. This chapter aims to describe methods to generate and implement energy management strategies. A classification is used to separate methods that can be used for offline or online operation. Despite the classification and description of several methods, more attention is given to the methods that are the central topic of this thesis.

3.1 Classification of Methods

Most studies on energy management address on-road vehicles. A common way to classify methods to generate EMSs is into two main categories, optimisation-based or rule-based methods [47-49]. Optimisation-based methods run an optimisation to select the control decision that minimises the cost function. Rule-based methods, as the name suggests, use rules to build the strategy, e.g., deterministic or fuzzy rules.

On the other hand, methods can be classified according to their causality. They are classified as methods that can generate online strategies if the final strategy is causal or if they rely on predictions of the future drive cycle conditions. They are classified as methods that generate offline strategies if the final strategy is non-causal, e.g., when performing a global optimisation where it is necessary to have prior knowledge of the drive cycle. This classification of methods is also used in [50] and [51].

To have the online or offline categories as the highest class, instead of rule- or optimisation-based, is more suitable for this thesis. This is because one of the methods studied aims at using machine learning to overcome the non-causality of the global optimisation method.

An overview of some of the methods to generate EMSs is presented in Figure 3.1 following the adopted classification structure. It is not a list of all methods but of those that, according to the literature review, seem to have more application cases to construction machines.

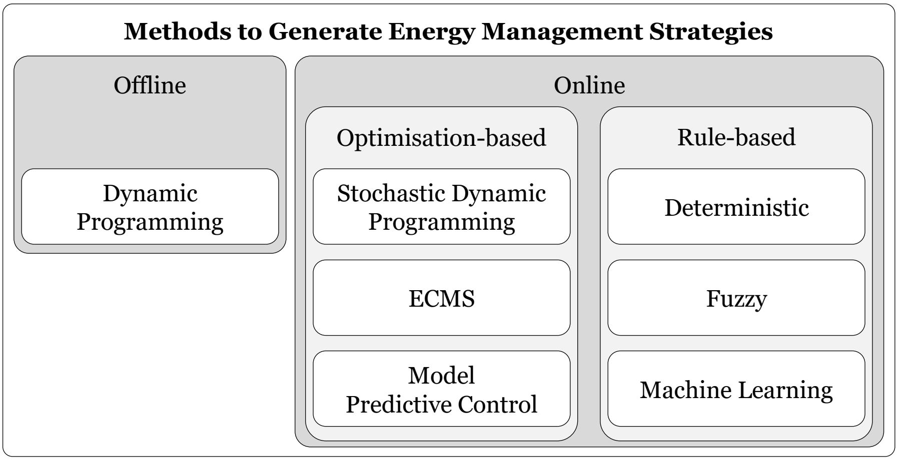

Figure 3.1 *Classification of methods to generate energy management strategies, inspired on [50].*

Although called offline or online, the development of the strategies is carried offline. In other terms, the strategies are not generated while in operation inside the systems they are controlling. In this sense, considering the methods addressed in this thesis, reinforcement learning would be the only fully online method because it can learn the strategy while interacting with the real system.

The methods evaluated in this thesis are a combination of dynamic programming and neural networks, and reinforcement learning (RL). Both are classified as machine learning-based methods, under the online rule-based category. Here it is considered that neural networks are in essence a rule-based method. The main difference is that neural networks encode the knowledge in the weights, biases, and activation functions. In a deterministic rule-based method this knowledge is constructed using "if-then-else" rules.

In one of the methods evaluated in this thesis, the neural network, learns from the results of dynamic programming optimisations and does not update while in the application/online. Thus, it is classified as an rule-based method and not as optimisation-based. In fact, the neural network is the means of implementing the energy management strategy and not what generates the knowledge.

In a similar way, RL is a machine learning-based method, and since in this thesis it does not continue to learn after being deployed to control the real system it is not classified as an optimisation-based method. In principle it could continue the learning after deployment, which, in this case, would make it fall into the category of optimisation-based methods as well.

3.1.1 Offline Methods

Some of the optimisation-based strategies are capable of finding the global optimal solution but require full a priori knowledge of the drive cycle and therefore not implementable online. Additionally, the optimal result is specific to the drive cycle they were optimised for. Nevertheless, they usually serve as benchmark to analyse, tune, evaluate, and propose other sub-optimal methods that can be applied online [49, 50].

One such method is dynamic programming (DP). A description of DP can be extracted from [52] and [53]. DP is a mathematical technique used for the optimisation of multi-stage decision-making processes where the cost of the present decision must consider the cost of future decisions. DP at each stage ranks decisions based on the sum of the present cost and the expected future cost, assuming optimal decisions for subsequent stages. It requires a discrete dynamic model of the system and an additive cost function. The dynamic system describes the behaviour of state variables influenced by decisions made at discrete time instances. It is capable of finding the optimal decision-making policy for a number of stages; this is the policy that minimises the total cost function.

Translating this explanation for the EMS case, the multi-stage decision-making process concerns, for example, how the power is split (control decision) between various power sources to meet power consumers demands over a drive cycle (decision-making process). DP can determine the optimal split of power to each time instance in a drive cycle, which minimises, e.g., the total fuel/energy consumption for the whole drive cycle. It is fundamental to have a model of the machine that describes its behaviour with the desired level of representativeness. Usually, backwards-facing quasi-static models are used for this purpose [10, 21].

DP have a high computational cost and requires a priori knowledge of the power demand for the whole drive cycle, thus it is not implementable online, as it results in a non-causal policy [54]. At the same time, the output EMS from DP is not straightforwardly interpreted, which makes the process for deriving an implementable EMS from them a time-consuming process, which also leads to sub-optimality.

In its most common formulation, it does not improve the performance of the drive cycle but finds a more efficient way by taking optimised control decisions to meet that specific drive cycle. It can, however, be formulated in different ways, like in [43], where the drive cycle is not prescribed but rather a consequence of the DP optimisation decisions.

The greatest advantage is that it guarantees global optimality for the given problem. Due to this advantage and constraints to online implementation, it commonly serves as a benchmark for the development of EMSs that can be implemented online, other causal controllers [54].

Examples of uses for such purposes are found in [10], [14], [21], [55], and [56].

DP is used for concept evaluation to ensure that the same cycle is met in the most efficient way possible by the different concepts. This should reduce the bias in concept evaluation during early development stages. Examples of such applications are found in [14] for diesel-electric hybrid wheel loaders and in [57] for hydraulic hybrid passenger vehicles.

Specifically for application in construction machines, an EMS for a wheel-loader with hybrid power-split transmission was developed in [21]. The authors adopted DP and made comparisons with a Rule-based (RB) approach. A DP approach to determining global optimal steering, lifting and tilting trajectories for a series electric hybrid wheel loader is studied in [43], and in [10], the problem of combined control and EMS optimisation for a parallel hybrid wheel loader is accessed. While optimising the size of components, DP ensures that the best possible efficiency was extracted from each design. DP and RB strategies were formulated to evaluate the combined optimisation approaches in terms of optimality and computational load.

3.1.2 Online Methods

For operation in the actual machines, the so-called online EMSs are required. It must be highlighted that forward-facing simulations of dynamic systems are also considered ‘online’ operation. However, the focus here is on actual real machine operation. According to Figure 3.1, they are divided into optimisation-based and rule-based categories.

Deterministic Rule-Based

Deterministic Rule-Based (RB) strategies are characterised by lower computational requirements, simplicity, and higher applicability online, at the cost of sub-optimality [49]. They are heuristic strategies implemented as “if-then-else” rules. Human expertise, intuition, operation boundaries, mathematical models, and safety considerations determine such rules [49, 58].

According to [50], there is the possibility that the control rules can be made detailed enough to address any special event that affects the vehicle. In real driving, the rules must be adapted to cover every driving condition. This can be valid to some extent for less complex passenger vehicles but could be a significantly harder task for the type of machines under consideration in this thesis.

In [5], it is stated that for hybrid construction machines RB-EMSs are the most widely used because of their high reliability and easy realisation. Examples of rule-based EMSs for construction machines are found in [18], where an RB-EMS is implemented for a parallel hybrid wheel loader. In [21], DP is used as a benchmark to guide the development of an RB-EMS for a power-split hybrid wheel loader. Similarly, [10] used DP to derive an RB strategy for a parallel hybrid wheel loader. A two-pressure level threshold rule to determine when the stored energy should be used to drive is presented in [28] and [59]. In [60] it is presented a parametric RB-EMS where the limits of saturation functions are the controller parameters that change according to predefined rules, therefore determining when to recover or reuse energy. In [61], an RB-EMS for a hybrid excavator is derived from the results of control optimisations with DP.

Fuzzy Rule-Based

Rather than decisions from deterministic RB strategies, Fuzzy Logic controllers can obtain a proportional output and continuous EMS. They can assume partially true values between true and false [51]. One example of fuzzy logic values could be: “Very low”, “Low”, “Medium”, “High”, and “Very high”. A description of the parts of a fuzzy logic controller can be found in [62].

Fuzzy logic when implemented in an EMS, yields output proportionality to different operating conditions, easy fuzzy rules tuning, and robustness to modelling and measurements errors are advantages [49].

The parameters of fuzzy logic can be tuned through optimisation to improve their optimality. A fuzzy logic EMS for a parallel hybrid hydraulic excavator is presented by [62], where genetic algorithms are used to optimise the membership functions parameters. An approach to deriving fuzzy logic EMSs for different hybrid electric vehicle architectures is presented in [63].

According to [49], fuzzy controllers can be further improved through adaptation and prediction. The parameters of the fuzzy logic can be adapted based on past, current and predicted vehicle operation information. Machine learning techniques such as neural networks can be developed to perform these adaptations and predictions.

One noticed characteristic of rule-based EMSs, either deterministic or fuzzy, is that usually they have only few monitored variables as inputs where the control decisions are based. Better decisions seem to emerge when more machine and environment variables are considered. When more variables are included for monitoring, rule-based controllers become even more difficult to be constructed.

The implementation of RB-EMSs for complex systems is a challenging task due to the difficulty to control continuous dynamic processes with a set of rules. They naturally result in sub-optimality, and this is a reason why great attention is given to optimisation-based EMSs. Naturally, optimisation-based methods show better fuel economy performance than do RB strategies [58].

Optimisation-Based

According to [50], online optimisation-based strategies simplify global optimisation problems into local optimisation problems, resulting in lower computational effort. This reduces the need for future drive cycle information, thus making them implementable online. According to the author, such techniques result in marginally sub-optimal results in comparison to offline optimisation strategies.

The following are examples of optimisation-based EMSs: Equivalent Consumption Minimisation Strategy (ECMS), Model Predictive Control (MPC), and Stochastic Dynamic Programming (SDP).

ECMS

According to [50], ECMS was first developed based on the concept that the energy used in a vehicle comes from the engine, as such, the hybrid system serves as an energy buffer.

ECMS is an instantaneous optimisation-based method for energy management where an equivalency factor is established to indicate the cost of using energy from the hybrid system. In a combustion engine hybrid vehicle, this equivalency translates the hybrid power to a fuel consumption. The total equivalent fuel consumption is determined by summing the engine fuel consumption and the equivalent fuel consumption. According to [15], depending on the direction of the hybrid energy, the equivalent fuel consumption can be higher or lower than the actual fuel consumption. At every time instance the equivalent fuel consumption is minimised, this leads to an optimised EMS [51]. Examples of ECMS application are found in [64] and [65].

A drawback of ECMS is the determination of the equivalency factor. It changes between tasks and states, which makes it a possible source of sub-optimality in the controller. One way of improving it is to make it adaptable and a parameter that can be tuned while in operation. An Adaptive-ECMS is developed for hydraulic hybrid off-road machines in [66].

MPC

According to [51], model predictive control (MPC) is a method for calculating a system control input to optimise the system's future output. This calculation is performed with a model of the system used to predict the system output for a given future time-horizon [50].

An optimisation problem is formulated based on the cost of the predicted control action. For hybrid vehicles the optimisation objective function can be the minimisation of fuel consumption as a function of the split of power between the energy sources as control decision.

The computational load of the controller is fundamental for its application in machines. Therefore, it usually relies on simplified models of the system to allow a fast computation, as can be noticed in [67] and [68]. An application of MPC for the energy management of construction machines is also found in [69].

SDP

Another technique that involves prediction of future behaviour is stochastic dynamic programming (SDP). EMSs derived through SDP are close to optimal and possible to implement online, which is an advantage over the deterministic DP [48].

According to [50], stochastic control is a framework developed to model and solve optimisation problems involving probabilities. For EMSs, the objective is to determine an optimal control action based on probabilities of transition from the current operation state to a future operation state. According to the author, an infinite-horizon stochastic dynamic optimisation problem is formulated. The vehicle power demand is modelled as a stationary Markov process, where the future state depends only on the current state, and the transition probabilities. According to [70], this removes the time-dependency of the problem. The future power demand is predicted based on current transition probabilities, and the optimal EMS to meet the future power demand is then obtained using SDP.

The optimisation using SDP is performed offline and results in a stationary lookup table that relates system states with optimal control decisions. The EMS is saved in look-up tables for interpolation of the control signals based on current states during operation [51; 70; 71].

According to [72], SDP-based EMSs are optimal if the driving behaviour matches the assumed Markov chain model. The authors also say that the major impact of SDP-EMSs is on the design of new vehicles because the method can automatically generate EMS faster than a person could do manually. However, it requires a collection of representative driving data to perform the optimisation and determine the Markov processes' probabilities of occurrence [71]. An example of the use of SDP for the energy management of a diesel-electric wheel loader is found in [73].

Section Summary

Although the above cited studies show the potential of the methods to address the energy management problem it is noticed that each of them has its drawbacks and advantages, which is probably why no consensus seems to exist for the application of more advanced methods instead of rule-based strategies in construction machines. This gives room for the evaluation of other methods, for example the ones based on machine learning.

3.2 Machine Learning-Based Methods

Machine learning, a branch of artificial intelligence (AI), encompasses algorithms that are able to improve their performance on a given task from experience [74]. In other terms, machine learning algorithms have the capability to acquire knowledge on their own by extracting patterns from raw data [75]. According to [75], it is the only viable approach for AI that can operate in real-world environments. Typical problems solved successfully by machine learning algorithms, for example, are image recognition, pattern identification, classification, and regression.

A distinction between machine learning and optimisation is that in machine learning one wants the generalisation/test error to be low. The generalisation error is the expected error on a new input [75]. In other terms, the goal is for the algorithm to also perform well in what it was not trained for.

Since in this thesis only neural networks are used as a means to encode the EMSs, it is the only machine learning technique described here.

Artificial neural networks (NN) are one of the techniques from machine learning. They provide a general, practical method for learning from examples [74]. They are formed by the interconnection of simple mathematical units (neurons). With a network composed by several layers of parallel units, it is possible to approximate complex functions to the point of being recognised as universal function approximators [76, 77]. The complexity of the functions they can approximate increases with the number of layers and neurons [78].

A neuron is illustrated in Figure 3.2. According to [74], each neuron takes a number of real-valued inputs x_i (possibly the outputs of other units). It performs a weighted sum by multiplying each input x_i by a weight ω_i, and a bias b is also added to each neuron. The output of the neuron (y) (which may become the input to many other units in subsequent layers) is determined by an activation function with the weighted sum value as input. The connection of neurons in layers forms a network. An example of the structure of a feed-forward network is shown in Figure 3.3.

During the training process of a network, the weights and biases are adjusted to minimise the error between the predicted value and target value. It is out of the scope of this thesis to describe training algorithms; however, a description of the backpropagation algorithm, used to train networks, can be found in [74].

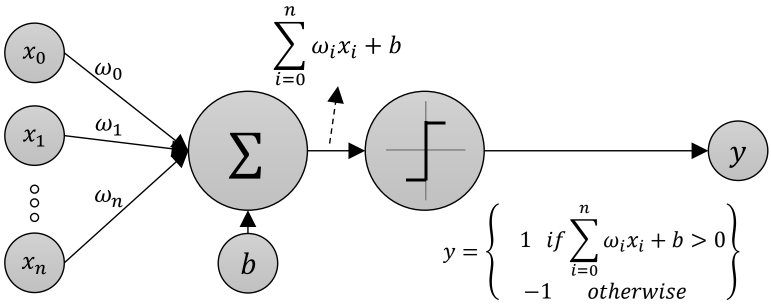

Figure 3.2 *A neuron, inspired on [74].*

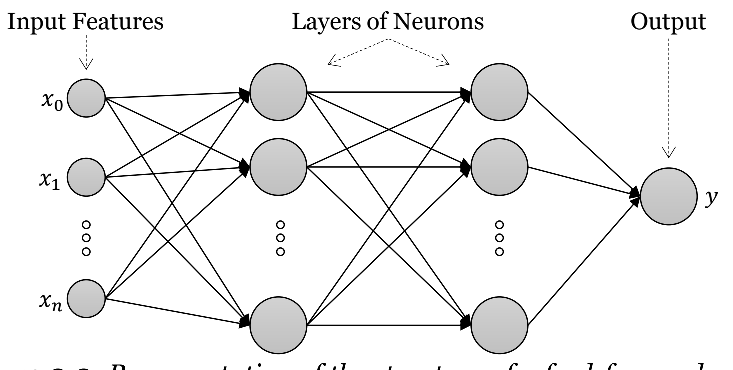

Figure 3.3 *Representation of the structure of a feed-forward neural network.*

According to [79], nonlinear activation functions, like hyperbolic tangent, when used in multilayer networks, help in creating compositions that increase the network's modelling power. The author continues, saying that more layers increase the depth of the network, and due to similar reasons, also increase the modelling power. However, the Occam's Razor statement provides guidance when defining the size of a network [78]. According to [75], simpler functions are more likely to generalise (reduce the test error), which is what is desired.

A few practical hints to consider when building and training a neural network are found in [76], [79], and [80]. It may be better to go to multiple hidden layers, preferring long and narrow networks to short and wide networks. When the number of hidden units is large, the generalisation accuracy tends to deteriorate, as it increases the chance to overfit the data. Therefore, learning should be stopped before overtraining occurs. Long training leads to weights and biases being tuned to fit a particular shape of the underlying structure in the dataset that might not represent the general dataset of the application target.

In this thesis, two methods based on machine learning to generate EMSs are evaluated: one belonging to supervised learning and another based on reinforcement learning. One characteristic that distinguishes them and drives the selection of one or the other is the availability of reference/target data, meaning that the choice for one or the other is guided whether or not the developer has the information about how the system should be controlled.

If this information exists, an approach based on supervised learning could be implemented and lead to a more straightforward development process aiming to train the algorithm to predict the target. This means one might know or have the information on how to control the system, but it is not possible to implement this knowledge by means of heuristic rules or in a way that does not result in significant deviation from the desired control behaviour. It could also be that the information concerning how to control the system is implicitly encoded in a dataset and not readily available to the developer.

If the knowledge on how to control the system does not exist, then a reinforcement learning-based approach could be implemented since in this type of algorithm the knowledge of how to control the system is generated while interacting with it.

3.2.1 Supervised Learning

One way of obtaining an EMS based on supervised learning is by using results from deterministic DP as inputs for supervised learning algorithms to train, for example, a neural network (NN).

NNs have been used in the control of hybrid electric vehicles due to its function-approximating ability [58], where the advantage is on learning and generalisation for mapping input features to a target output.

In [81], an approach to predict the driving condition in the near future and use this information to determine the optimal energy management is proposed. NNs are trained over a number of drive-cycles to perform that prediction. Results from the use of NNs to control the system are provided in [82].

In [9], a power management controller based on DP and NN for a hydraulic series hybrid on-road vehicle is proposed and investigated. If the optimal accumulator pressure trajectory is known, then an implementable control scheme can achieve a nearly global optimal fuel efficiency. An NN was then trained to control the optimal pressure trajectory based on the vehicle's velocity. In this way, the NN generalised the relationship between vehicle velocity and accumulator pressure.

Other similar approaches combining NN learning from optimisation results from DP or other control optimisation techniques were used by other authors as well for the control of on-road vehicles [83-92]. What is observed from this collection of studies on the topic is the ability of the networks to learn complex decision-making processes from system states and other measured variables and implement them in both simulation domains and in real systems. In some cases, the optimisations are multi-objective, like in [88]. Their performance seems to be independent from the type of systems and architectures, like series, parallel, busses, and light vehicles. It is also seen as a large distinction in types of architectures and sizes of networks used. However, it is a method not extensively studied for construction machines.

An investigation using the same type of method but for a hybrid wheel loader is presented in [93]. This simulation study was part of the studies presented in this thesis with the goal of evaluating the performance of the EMSs based on NNs along with the controllers of the machine and for a slightly different machine configuration than the one studied here.

The supervised training of the networks requires a large amount of drive cycles to be optimised with DP. This is to increase their capacity to find generalisation rules that are applicable to unseen cases. This could reduce the problem of optimisation-based EMSs highlighted by [11]. The author says that a problem is that they will optimise a certain known drive cycle, and that a single cycle does not guarantee optimality for the whole machine operation tasks. By learning from several examples, it is expected that NNs could find generalised rules. They would not operate optimally but could partially overcome the problem faced by some of the optimisation-based EMSs, like the adaptation to changing working conditions.

Positive aspects of this supervised learning approach are that the control developer knows what should be learnt and can work towards making sure that the network learns it while training. The availability of a reference to compare to is also a positive aspect. This means the learned control strategy can be directly compared to the target strategy.

A negative aspect is that usually the target reference is obtained from control optimisations based on simplified models of the system and of its interaction with the environment, like when using DP. This means that the target is an approximate solution of the real behaviour and not the global optimal solution for the real system. In some cases, the lack of ability to generate a target reference while the machine is operating prevents it from further training while deployed to the real system. They might suffer from adaptability to the actual machine and working conditions. This is a motivation to also look for algorithms that can continue the learning in practice.

3.2.2 Reinforcement Learning

"*Reinforcement learning is learning what to do - how to map situations to actions - so as to maximise a numerical reward signal. The learner is not told which actions to take, but instead must discover which actions yield the most reward by trying them,*" [94].

A representation of a reinforcement learning set-up, in comparison to a traditional control set-up, is shown in Figure 3.4.

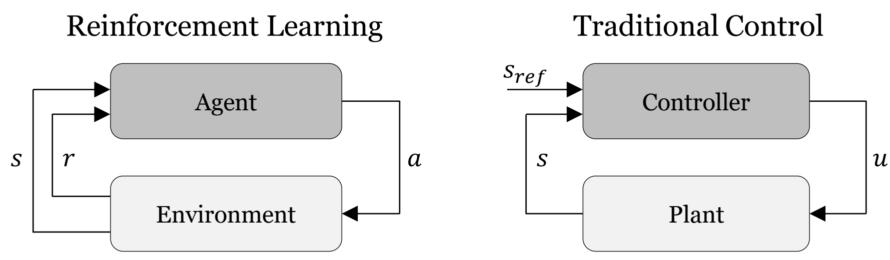

Figure 3.4 *Reinforcement learning framework vs. traditional control.*

To some extent the agent in RL is equivalent to the controller and the environment is equivalent to the plant. In traditional control there is an external reference (s_{ref}) generated, for example, by an operator, that is external to the plant being controlled. In RL the operator generating the reference is inside the environment with whom the agent is interacting with to maximise the reward (r). While in traditional control the controller is defined by the control engineer, in RL the agent learns automatically how to perform the task.

In RL, based on the current observations (s), the agent applies an action (a) to the environment, observes the resultant observations (s') and the reward (r) collected from the transition between $s \rightarrow s'$. As it progresses in the training, by testing different control actions and observing their outcome through the set of observations, it learns to take actions that maximise the reward. The phase where the agent is interacting with the system to find out the best way of controlling it is called exploration.

The action (a) is the control decision. Observations (s) are, for example, measurements from sensors that define a state of the system. The reward (r) is a function that defines how well it is performing the task. It could be, for example, the system's efficiency.

A more mathematically centred and detailed description of an RL algorithm can be found in [94] and [95]. However, here a brief description of an RL algorithm is performed based on these two references.

Usually, one is not interested in taking only one best control decision but in the best sequence of control decisions to have a strategy, also called a policy, that maximises the total collected reward over this sequence. This could be, for example, to maximise the collected points in a game. The sum of collected rewards over a sequence of decisions is called return (R). Since when performing the sequence of decisions there is no knowledge of future rewards, the return must be estimated. Since it is an estimation, it only provides an expectation of what could be the reward. Then, what one is after, for the current system state (s), is to find the action (a) that maximises the expected return $Q^o(s,a)$.

The optimal action-value function ($Q^o(s,a)$) gives the maximum expected return for being in state s and taking the control action a and after that following the policy π. Thus, based on the expected return, the agent can construct the action-value function (map from states to expected value) and apply a policy to select actions based on their values. For an agent, a policy could be to always choose the action that maximises it.

The action-value function can be approximated by a neural network, which is trained to predict the value (*Q-learning*) based on the observations (s) and action (a). Thus, the training, in this case, is to have a good representation of the action-value function to guide the decisions. A description of this method is found in [95]. Therefore, the function representation capability of NNs is used to tackle complex problems where the action-value function is highly nonlinear.

The fact that the agent in a RL framework can learn while interacting with the system allows their continuous improvement even after being deployed to the real system. In theory this is possible; however, there are practical and safety constraints that might undermine this approach, or at least not allow it to happen in a reasonable time window to cope with the varying nature of the tasks performed by mobile machines.

A disadvantage of such methods is the fact that the engineer does not know what solution to the optimal control problem it will find. It could be that it finds clever and optimised solutions that the engineer was unable to guess by her/himself, but it could also be that the solution found is far from the optimal one. This requires the assessment of the trained agent controlling the system to evaluate whether the solution is acceptable.

There are already examples of the application of reinforcement learning with a greater focus on energy management. In [96], it is shown the use of RL for a hybrid excavator, comparing it to a rule-based approach and an ECMS approach. Similarly, in [97], RL based on *Q-learning* is developed and evaluated for the real time energy management of a hybrid wheel loader. An approach based on *Q-learning* with demonstrated online learning capability is shown in [98]. In [99], the use of DQN (Deep Q-Network) for the mode selection of a multi-chamber actuator in the control of an excavator arm is shown. This was also a study that is part of the development of this thesis.

Another set of applications of RL for mobile machines are found in [100] where it is used, based on camera, lidar, and motion and force sensors, to perform the bucket loading in underground mine applications in a multi-objective target, including the maximisation of bucket loading; in [101], it is trained to control the motion of a forestry crane while considering the minimisation for energy consumption; in [102], it is used for the trajectory tracking control of the motion of an excavator arm aiming for autonomous application, where the controller generates the valve control signals directly; and in [103], RL is used to adapt for the real world conditions of a network trained to control the motion of the actuators of a wheel loader during bucket filling.

What is observed from these references is the capability of reinforcement learning to learn how to control systems of construction machines based on a variety of inputs, which allows the mapping of a system state to perform the best actions. Also noticed is their capability to deal with multi-objective goals.

It is also observed in these papers that the development of RL controllers usually starts with a pre-training of the agent in a simulation environment. Advantages of this approach are avoiding undesirable real-world consequences; that it can generally be performed at less cost than needed to obtain real experience; and that simulations typically run faster than real time [94].

What is not addressed in depth, for both supervised and reinforcement learning approaches are their application to real construction machines, the reliability, and safety of such controllers, which is of great concern especially in automated applications.

3.3 Review of the Reliability of Machine Learning-Based Methods

What motivates this review is the fact that machine learning-based controllers applied to real physical systems might underperform. In the worst case, this leads to unintended accidents. According to [104], accidents may emerge when one specifies the wrong objective function, is not careful about the learning process, or commits other machine learning-related implementation errors.

According to [75], the machine learning model must perform well on previously unseen inputs, not just on those it was trained for. This ability is called generalisation and allows them to be deployed to situations that they were not exactly trained for. However, a neural network will only be as good as the data used to train it [78]. As the author states: "*They are a technology that is at the mercy of the data*".

The concept of ground truth is fundamental for machine learning. It is the ideal expected result [105], and the result one would like the model to learn and predict. The accuracy of the trained model will depend on how close the defined target is to the actual ground truth. A machine learning model with high accuracy and high generalisation in the application would behave as intended.

Therefore, two fundamental conditions for supervised machine learning algorithms to work properly is to have access to the ground truth and that the datasets (instances/inputs and labels/targets) used for training match with the application. Therefore, since simulation models of physical systems are only approximations of the real behaviour, the methods used in this thesis result in the absence of ground truth and dataset shift. This leads to uncertainty in the predictions and in the performance of the deployed learned model.

Dataset shift, also called distributional shift, is a field of study in machine learning. According to [106], dataset shift happens where the joint distribution of inputs (features) and outputs (targets) differs between training and application. It is also called by [104] “robustness to distributional shift", and it discusses how to avoid having machine learning systems making bad/different decisions when given inputs that are different than which was seen during training. Meaning that one still wants to work with the learned model, but there is imperfect and incomplete information around it.

According to [106], the conditions in which the system operate differs from the conditions in which they were developed. The author says that environments are nonstationary, and sometimes the difficulties of matching the development scenario to the application are too great or too costly.

However, the application should still be similar to the training, which means that the training data must span the full range of the input feature space and also have similar distribution to the application for which the algorithm will be used. Therefore, these concepts of dataset shift and absence of ground truth are fundamental and must be addressed for machine learning-based methods to generate EMSs.

In this sense, RL has an advantage over supervised learning due to its capability to continue the learning after being deployed to the application. This allows it to adapt to differences between simulation and application. However, it is not completely free of these issues because it was initially trained on a simulation model. Therefore, it can still take wrong decisions during the exploration phase after deployment.

4

Case Studies

Machine learning-based methods to generate EMSs, using neural networks, seem to have a great potential to address the complexity of systems, operation scenarios, and tasks of construction machines. However, what is noticed from the literature review is a lack of studies focusing on the evaluation of the application of such EMSs to real systems. The two case studies developed in this thesis aim towards obtaining experimental results for the evaluation of the potential of machine learning-based methods for the generation of EMSs for construction machines.

Case Study I addresses a supervised learning approach. A neural network is trained on control optimisation results from dynamic programming to control the parallel hydraulic hybrid system of a wheel loader. The goal is to reduce fuel consumption. It is developed offline from simulations and applied to the real machine for evaluation. Papers I, II, III, and V contain the results and content of this case study.

Case Study II addresses a reinforcement learning approach. A neural network is trained to perform the mode selection of a multi-chamber actuator to reduce the energy losses of a valve-controlled hydraulic system for the actuation of an excavator arm. It is trained in simulation and deployed to the real system for evaluation. Paper IV is related to this case study.

The two case studies also address the potential of applying EMSs developed in a simulation environment directly in the real machines.

4.1 Case Study I - Energy Management for a Parallel Hydraulic Hybrid Wheel Loader

The goal of this case study is to demonstrate the expected performance of a hybrid construction machine when operating with an EMS developed based on supervised learning. The selected machine is a parallel hydraulic hybrid wheel loader. Its behaviour is simulated with a backwards-facing model. A fuel consumption optimisation problem is formulated and implemented in a DP tool for optimisation of the EMS of the secondary energy source. An NN is trained on the optimal results for several drive cycles to predict the optimised control variable. The NN along with additional rules form the EMS (NN-EMS). To increase the robustness and generalisation of the network, safety-critical rules from the real machine controller are implemented as constraints to the optimisation problem. Additionally, a method hereafter called 'state sweep' is implemented in the optimisation process to increase the robustness of the network to unseen scenarios that might occur in practice. Fuel consumption tests, with a professional operator, were done to evaluate the performance of the NN-EMS. The same field tests are performed with the conventional machine (non-hybrid mode) and with a rule-based EMS (RB-EMS).

Papers I, II, III, and V are connected to this case study. In Paper I, a simulation study shows that the NN-EMS does work in a simulation environment and has the potential for fuel consumption reduction when compared to a RB-EMS. Paper II shows that to generate an NN-EMS that is robust in the real application, it is necessary to include safety/control-critical rules as constraints in the optimal control problem in order for the solutions to resemble the operation of the real machine. Paper III presents the experimental results of the NN-EMS implementation in the machine, while Paper V presents the machine model used for the control optimisations and simulations.

The EMS is developed for the control of the hybrid system while the machine is operating in short loading cycles, which has its phases described in detail in [107]. Figure 4.1 shows a representation of the cycle.

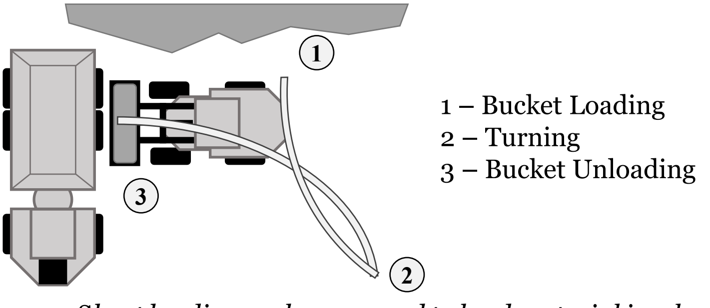

Figure 4.1 *Short loading cycle, e.g., used to load material in a hauler; [III].*

4.1.1 Machine, Model, and Optimisation

The machine used to demonstrate and evaluate the NN-EMS is a parallel hydraulic hybrid wheel loader. This concept allows kinetic energy recuperation when braking with storage in the hydraulic accumulator. The hydraulic pump/motor and accumulator form the complementary energy storage device (CRS). With proper management of the energy recovery and usage, fuel consumption can be reduced.

The machine is modelled in a backwards-facing way to allow the use of DP to solve the optimal control problem: the optimal split of power between engine and hybrid system to minimise the fuel consumption along the drive cycles. This is to generate the information on how this system should be controlled. It is expected that if the model description of the machine behaviour is sufficiently accurate, then the results from the control optimisation are a reference for how the real machine should be controlled to reduce fuel consumption in this case.

Figure 4.2 presents the machine concept diagram along with the model inputs and outputs, that are used to evaluate the agreement between model and measurements. As mentioned, the model is described in Paper V.

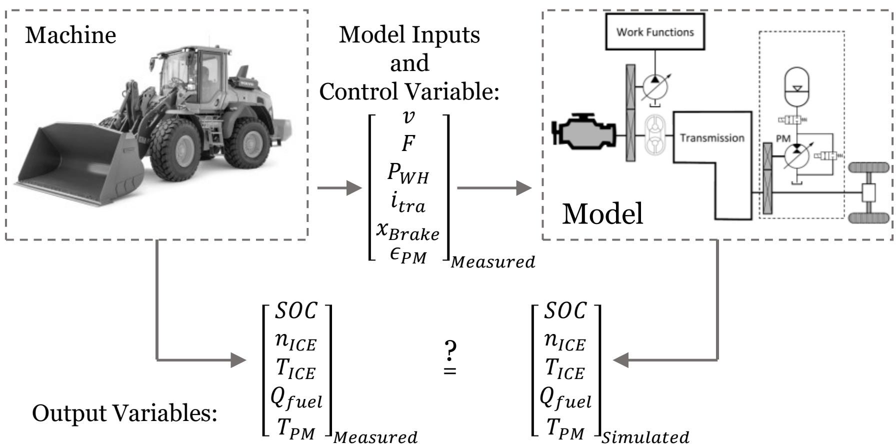

Figure 4.2 *Machine concept and model verification procedure; [III].*

The goal of the model is to describe the accumulator state of charge (*SOC*); engine speed (n_{ICE}) and torque (T_{ICE}); fuel consumption (Q_{fuel}); and pump/motor mechanical torque (T_{PM}), as functions of model inputs and the control action (ϵ_{PM}). The model inputs are machine speed (v), traction force (F), transmission gear (i_{tra}), work functions power (P_{WH}), and brake pedal position (x_{Brake}).

The backwards-facing model has, essentially, the drive cycle power consumption as input. This means that the cycle productivity (t/h) is prescribed. In other terms, the amount of work performed and the time taken is fixed for each cycle. What is obtained with the control optimisations with this type of model is information of how this same cycle could be performed with less power consumed, meaning less fuel consumed in this case. This also means that the productivity of the machine operating under the developed NN-EMS should have similar productivity to the recorded cycles.

The model agreement is performed as shown in Figure 4.2. Measured cycle inputs $(v, F, P_{WH}, i_{tra}, x_{Brake})$ and control action (ϵ_{PM}) are applied to the model, and the measured and simulated variables are compared. Figure 4.3 shows their agreement for one short loading cycle.

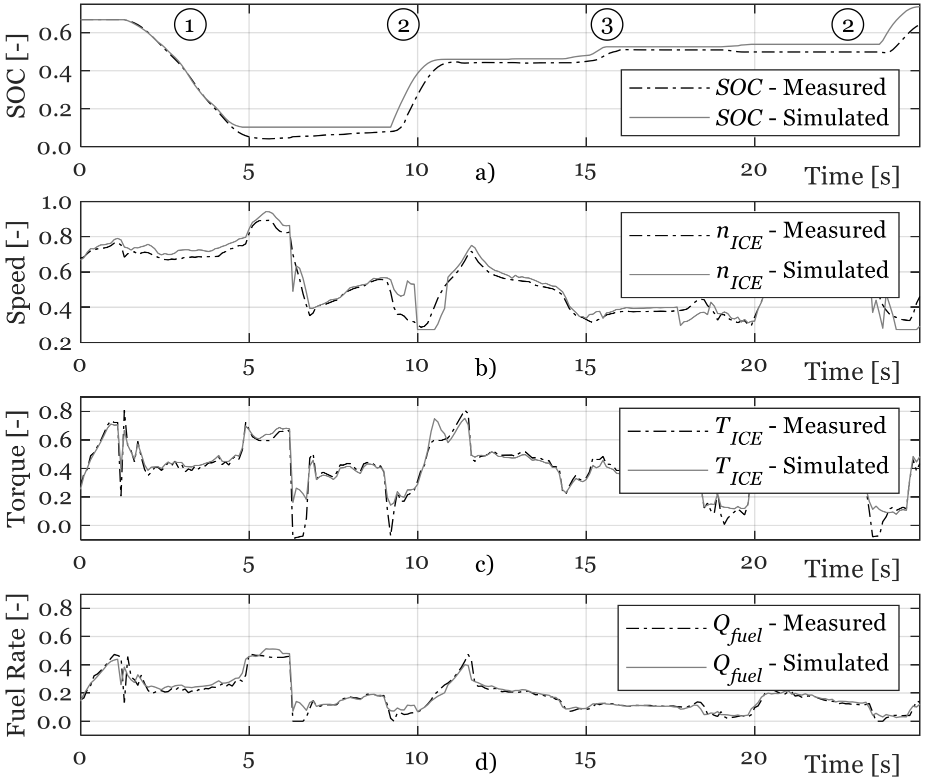

Figure 4.3 *Simulations and measurements agreement. a) Accumulator state of charge; b) Engine speed; c) Engine torque; d) Engine fuel rate; [III].*

Although the model is deliberately simplified to minimise the computational cost when performing the optimisations, the main characteristics of the real system are described. For this reason, the model is considered to provide a sufficiently accurate description of the machine behaviour for the intended purpose of generating information of how the real machine should be controlled. A better model would have a more accurate description of the machine behaviour. This would probably result in a better NN-EMS for the real machine, however the goal here is to demonstrate the potential of the method to generate EMSs.

Given the model agreement, it is possible to use this model to perform control optimisations to gather information on how this system should be controlled in practice. The optimal control problem, adapted from the formulations provided in [54], [108], and [109], can be formulated as:

$$min\left\{g_N(x_N) + \sum_{0}^{N-1} g_k(x_k, u_k)\right\}, k = 0,1, \dots, N \tag{4.1}$$

$$x_{k+1} = f_k(x_k, u_k)\,, \tag{4.2}$$

$$u_k = \begin{cases} u_k \; if \; u_{k,rule} = 0 \\ u_{k,rule} \; if \; u_{k,rule} \neq 0 \end{cases} \tag{4.3}$$

$$x_k \in X_k \subseteq \Re^n \tag{4.4}$$

$$x_0 = x_{IC} \tag{4.5}$$

$$x_N \in T \subseteq \Re^n \tag{4.6}$$

$$u_k \in U_k \subseteq \Re^m \tag{4.7}$$

where, g_N is the final cost, g_k is the additive cost function, f_k is the function describing the dynamic system, and u_k is the control action that can be any value from U_k, although it is $u_{k,rule}$ when a control action from a deterministic control rule takes place. x_k is the set of state variables, N is the final time step for the discrete-time problem, and T a target set as final state constraint. Therefore, the optimisation process runs as described in the referred papers, except when a deterministic rule applies, at which moment only the control variable $u_{k,rule}$, instead of every possible control action, is applied/tested from every discretized-state. It forces the optimisation algorithm to find the best solution for the whole problem given that at some parts of the decision-making process the control decisions were constrained. This process can be viewed as an control optimisation with intermediated control constraints. Further details on the effect of that in the optimisation results are provided in Section 4.1.3.

4.1.2 Method

Figure 4.4 presents a diagram with the activities performed along the method, from the development of the NN-EMS to the tests.

NNs can only output certain results if they were trained for it. If not enough examples are provided, or if they do not cover the whole operation envelope of the machine, it is likely that the network will behave in an unexpected way when it receives an unknown input.

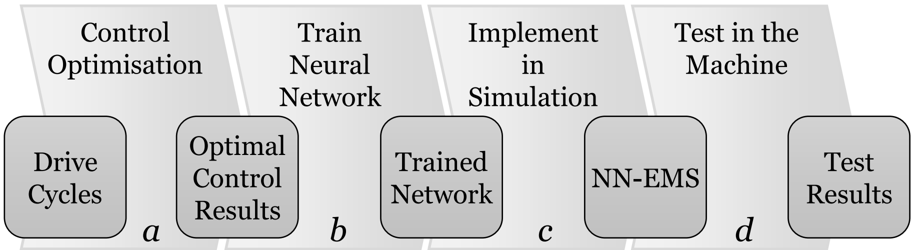

Figure 4.4 *Representation of the process to obtain an NN-EMS from dynamic programming and neural networks.*

Loading cycles in construction machines, despite being usually repetitive, have great variation, which can be observed in the measured trajectories in [44], and on machine and actuator speed profiles in [110]. Therefore, for the NN to learn how to control the system a significant number of examples, also called instances, is required. One instance is the tuple of attributes/input features and target. That is why the main input to the process is a large number of representative drive cycles. A brief description of each step in the process is provided in the sequence.

Step a: With the machine model to hand, drive cycles are simulated using DP to find the optimal control decisions that minimise the objective function while completing them under the set of control and state constraints. The optimisation results will contain the optimal trajectories of machine states and optimal control decisions.

Step b: Variables that affect the decision-making process are selected from the optimisations results to be the network input features to map the target. The target is the desired control variable to be predicted by the network. The dataset containing input features and target is used for the network training. When training, it learns the nonlinear relationship between features and target. The trained network is the causal controller that predicts the optimised control decisions in the machine.

Step c: The network is implemented in simulation alongside additional control rules to form the NN-EMS, to evaluate its performance, and to confirm its correct operation.

Step d: The NN-EMS is implemented and tested in the machine.

4.1.3 Control Optimisation (*Step a*)

The DP tool presented in [54] is used to solve the discrete-time optimal control problem with final state constraints. The optimisation is control-constrained by deterministic rules, as shown in equations (4.1) to (4.7). In the controller of a real machine, there will be safety-critical functions (e.g., braking), and possibly other simpler control actions, that cannot be prone to neural network miss-predictions, or there is no need to train the network. These cases can be addressed with deterministic control rules. Therefore, the final controller is a combination of rules and a neural network.

If the deterministic control rules are not added as constraints to the control optimisation problem the optimisation results will not resemble what the network finds in practice. The outcome of the control decisions in practice leads to a different machine behaviour than what the network was trained to achieve. As a result, the network will underperform because it encounters situations it was not trained for.

This is addressed in Paper II where it is shown, with simulation results, the influence on the control optimisation results between constraining it or not with the braking rule. The constrained and unconstrained optimisation results for the trajectory of the accumulator state of charge (*SOC*) for several drive cycles are shown in Figure 4.5. It allows a qualitative assessment of the deviation between the optimal control problem solutions for each case.

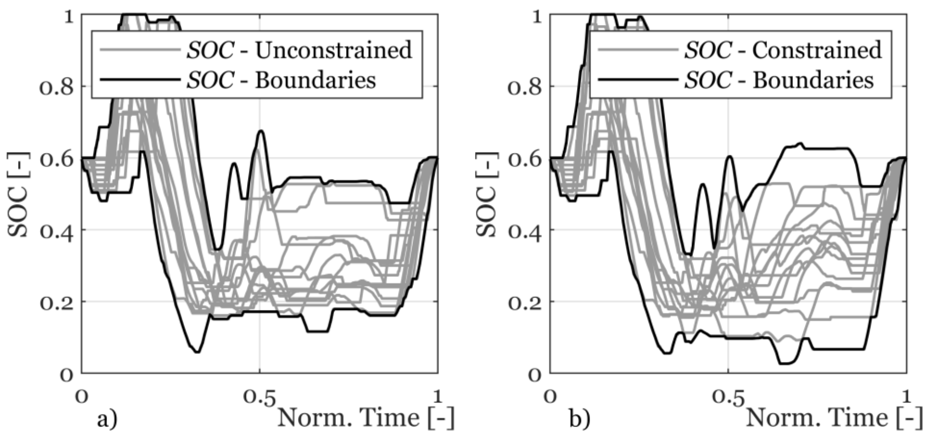

Figure 4.5 *Distribution of optimal SOC for the constrained and unconstrained control optimisation problem for individual short loading cycles. a) Unconstrained; b) Constrained; [II].*

In general, the solutions seem similar, and the differences are mostly located towards the final half of the cycle, which happens to be where most of the braking actions occur because the machine is approaching the hauler for the bucket unloading phase. However, if analysing in more detail it is noticed that differences are significant.

Since the braking rule overwrites the control decisions from the network, the braking parts are removed from the training set. Figure 4.6 shows a comparison between the SOC for the constrained and unconstrained cases for a number of cycles. If there is no significant difference between them, when plotted against each other they should lie close to the diagonal.

Even after removing the braking parts, there is still a significant deviation from the diagonal (Figure 4.6b), meaning that there are significant differences in the solution to the optimal control problem, even outside the braking actions. The presence of control constrains at some parts of the cycle affect the optimal solution for the whole drive cycle.

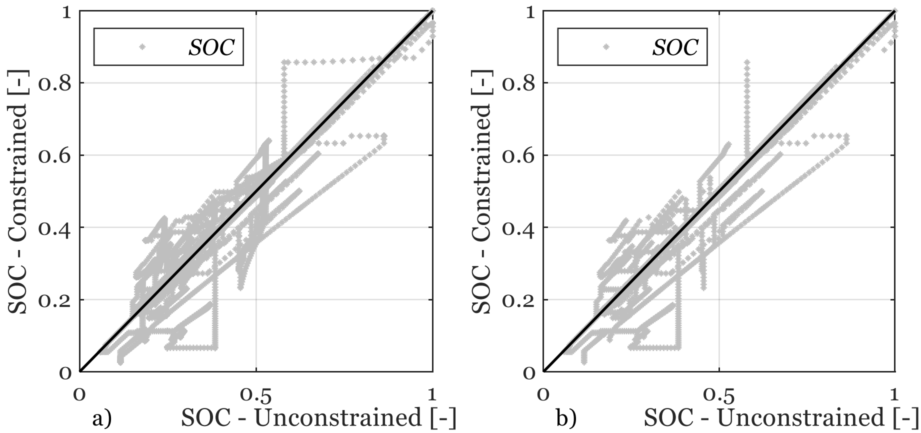

Figure 4.6 *Optimal SOC for the constrained vs. unconstrained control optimisation problem for a number of short loading cycles. a) Including the constrained parts; b) Excluding the constrained parts; [II].*

In essence, not considering the deterministic rules in the optimisations would result in a network trained for something different than what it will be interacting with in the application. In other terms, the inclusion of the rules reduces the dataset shift between training and application. This is confirmed when networks trained on data from each case are placed to control the machine.

Figure 4.7 shows the error (RMSE) between the control variable prediction (ϵ_{PM}) and optimal control variable, for two networks controlling the machine in simulation. One network is trained on a dataset from unconstrained control optimisations (UN) and the other on a dataset from constrained control optimisations (CN). The RMSE is also calculated for the *SOC*, allowing the assessment of the error in the state variable as a consequence of wrong control variable predictions.

The average error on the control variable (ϵ_{PM}) and state (SOC) are smaller for the NN-EMS trained on the dataset from constrained control optimisation (CN). These simulation results indicate that the hypothesis that one should consider required control rules to constrain the optimal control problem to achieve a better representation of what happens in reality holds true, otherwise the NN-EMS will underperform.

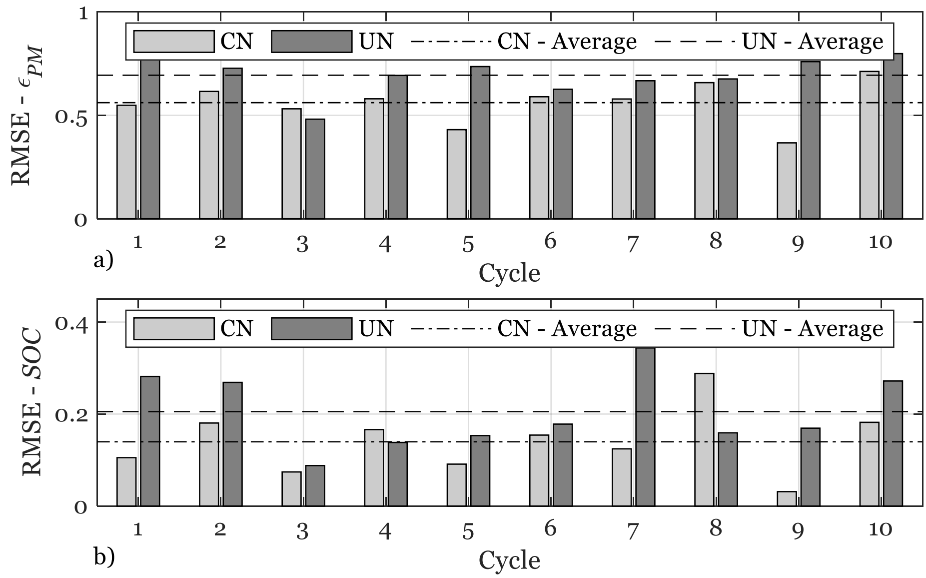

Figure 4.7 *ϵ_{PM} and SOC RMSE comparison between optimal solution and EMSs for the constrained (CN) and unconstrained (UN) problems. a) ϵ_{PM} RMSE; b) SOC RMSE; [II].*

The optimised control decisions from the optimisations must be analysed to make sure they make sense and to understand what type of control action must be learned by the network and applied to the real system. The optimisation results also provide information regarding the efficiency of that machine concept and how it should be controlled. Commonly, it provides non-intuitive ways to control the system to maximise its efficiency. The engineer can learn about the control of the system under development.

One expected advantage of the method to generate the EMS studied here is that the engineer does not need to design the whole controller for the machine since the network should learn that from the optimisation results.

It must be emphasized that the goal is for the network to learn when a control action must take place and its intensity. For the present system, this means when and how much to charge the accumulator and when, and how much to use the stored energy to drive the machine. Optimisation results for one short loading cycle are presented in Figure 4.8.

At around 5 seconds in Figure 4.8a, at point 1, the machine is loading the material from the pile. Therefore, the stored energy is discharged to avoid high torque converter losses. It also shows the recovery of kinetic energy and its reuse at point 2, when changing direction from the hauler to the pile of material, or vice-versa. It also shows, at around 8 and 20 seconds, the use of stored energy when leaving: the pile, point 1; and the hauler, point 3. This is mainly to overcome the torque converter losses at high torque and low-speed situations. Figure 4.8b shows the relationship between pump/motor operation and the hybrid system state, which means pumping to brake and recovering kinetic energy (Brake reverse – REV, or forward FWD), motoring to drive and reuse the stored energy (Drive reverse – REV, or forward FWD). These actions are confirmed in Figure 4.8c through the exchange of hydraulic potential energy and kinetic energy.

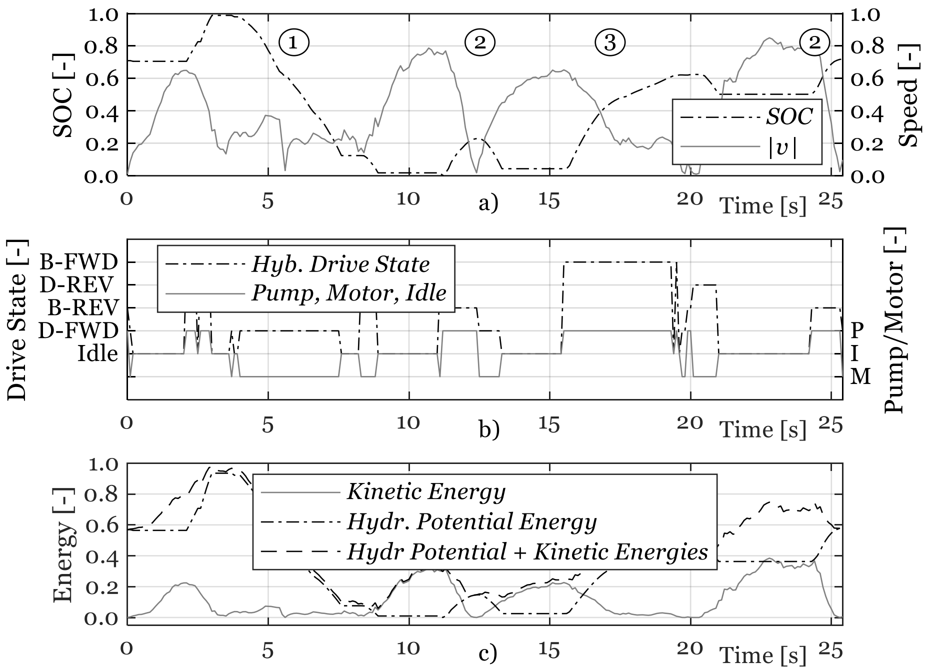

Figure 4.8 *Optimisation result for one short loading cycle. a) Machine speed and accumulator SOC; b) Hybrid system drive state and pump/motor operation; c) Kinetic and hydraulic potential energy; [III].*

It is the underlying information in the optimisation results, about how much and when to charge or discharge the accumulator, that is the objective for the neural network to map from the input features. Naturally, the drive styles of each operator, types and conditions of materials, distances between loading, turning and unloading points, and randomness in the interactions with the ground and loaded material will result in a high variation in the load cycles. Therefore, as many operation scenarios as possible must be covered to have an EMS that is robust to this variation in the application. Even though, the network generalises the predictions when training, it is expected that it can learn general optimised control solutions that perform well in the covered scenarios and are robust to unseen scenarios.

To achieve this robustness, "state sweeping" was implemented. This process is described in Paper III, but it essentially consists of also accounting for the solutions for short optimal control subproblems from within each optimised drive cycle. This process generates more information on how to control the system when it deviates from the optimal trajectory, increasing its robustness when being in such situations.

One example of SOC trajectories obtained with this process is shown in Figure 4.9, where it is indicated the starting points $(t, SOC(t))$ of the subproblems and the respective optimal state trajectory solutions.

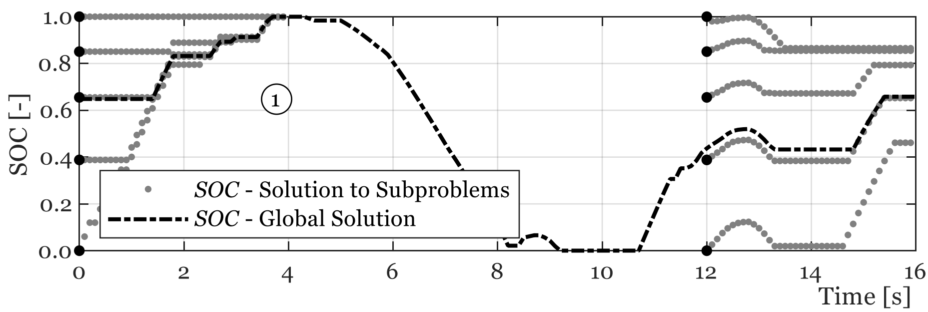

Figure 4.9 *Example for a short loading cycle of the accumulator SOC behaviour when applying state sweep in the optimisations; [III].*

To optimise several drive cycles, including deterministic rules as constraints, and applying state sweep should result in a dataset that contains sufficient, representative, and useful information to train a network that is expected to perform well and be robust in practice.

4.1.4 Dataset, Network, and Training (*Step b*)

About 100 short loading cycles were optimised individually to force the accumulator charge sustaining constraint at the end of each cycle. These cycles were recorded from several operators, loading different types of material, and in different machine modes (hybrid and non-hybrid). This makes the trained NN more robust for the scenarios found in practice.

No sensitivity analysis was performed to evaluate the impact of the number of cycles used in the development on the performance of the EMS in the application. However, it is not just the number of cycles that matters but also how well they cover the possible operation scenarios. This means the drive cycles used in the process should result in training points that cover the whole range of the input features used by the networks. At the same time, rare events should not have too few points, in comparison to more likely events, for the network to neglect them during training. In other words, a similar spread of points wouldn't make the network biased towards the more likely operation scenarios. This topic is addressed in Chapter 5.

The network used in this case study is a regression model that learns to map the input features to the target output – in this case, the optimal control action. The network architecture was determined iteratively aiming to reduce the number of neurons and hidden layers while maintaining high accuracy. A feed-forward multi-layer network, with tangent sigmoid neurons in the hidden layers, and a linear neuron in the output layer was selected since they are a standard architecture for regression problems [78].

Several training runs were performed for networks with different numbers of layers and neurons. Although not being the purpose of the thesis to find the best machine learning model, the objective of this iterative process was to find a small but sufficiently capable network model.

An important aspect that drives the use of a smaller network is the fact that they need to be implemented in the machine computer that usually has limited computational power. The hybrid system in the demonstrator used in this case study was controlled from a rapid control prototyping PC. Therefore, the size of the network wasn't a big constraint. However, this is a topic to be further considered in the future industrialisation of this type of EMS.

The selected structure was a network with three hidden layers, 25 neurons in each layer, and tangent sigmoid activation functions. The output has one neuron with a linear activation function to predict the normalised value of the pump/motor torque (T_{PM}). Although in the optimisations the control variable is the displacement setting (ϵ_{PM}), the network is trained to predict the torque which is transformed in a displacement setting, as shown in Figure 4.11.

The objective is to provide the minimum number of features to the network. This is still a task for the engineer to perform. It requires knowledge of the machine operation to identify which variables affect the system efficiency, can be measured or generated in the machine, and seem to have a strong correlation to the target control variable. The input features (x_{NN}) are

$$x_{NN}(t) = [v, i_{tra}, KD, SOC, x_{Accel}, n_{ICE}], \tag{4.8}$$

and the target ($T_{PM,NN}$) is

$$T_{PM,NN}(t) = T^{o}_{PM}(t), \tag{4.9}$$

where v is the machine speed, i_{tra} transmission gear, KD is the kick-down signal, x_{Accel} the accelerator position, n_{ICE} the engine speed, SOC

the accumulator state of charge, and T^o_{PM} the optimal pump/motor torque.

The input features and target are shifted by one time step. That is because in simulation, or in the real application, the current measured variables are used to predict the control action for the next time step.

As presented in Paper I, when simulating the system under control of the network, drive cycle data (c_i), state variables from the previous time-step (x_{i-1}) and control signals (u_i) are provided as inputs. The model computes the new state variables (x_i) and dependent variables (y_i). From these outputs, the features are calculated for the neural network (NN), which then calculates the new control signal that is used in the next time-step. This process is shown in Figure 4.10.

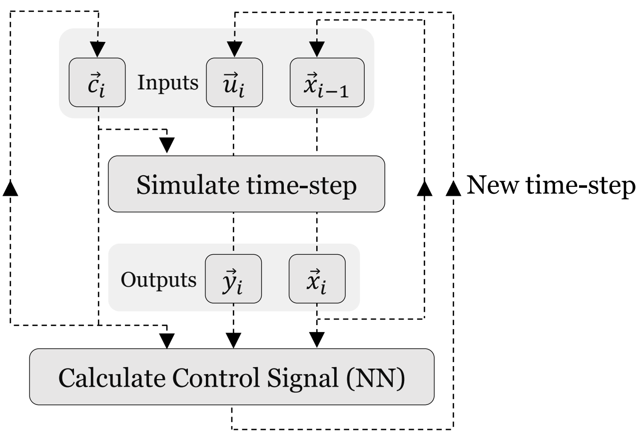

Figure 4.10 *Implementation of the neural network as the controller in simulation; [I].*

The network was trained on optimisation results from about 90 drive cycles and tested on the remaining drive cycles. The state sweep, described in Section 4.1.3, was applied to all cycles. As discussed in Paper II, the parts of the cycle where the control action is determined by a deterministic rule were removed from the training set. The reader is referred to Paper III, for further results on the training of the network.

4.1.5 Implementation (*Step c*)

In the implementation of NN-based EMSs, all safety-critical functionality should be kept as deterministic rules while the network should only predict the optimised control reference in non-safety-critical situations.

For the present case study, the additional rules added to ensure a correct and safe operation of the NN-EMS are summarised as follows:

- Predictions are bounded to the limits of the control variable;
- Predictions that would result in accumulator depletion or overcharging are not allowed to reach the system;
- The control signal is not applied to the hybrid system if the machine is not in a driving mode; and
- The torque reference is calculated by a map when the operator requests a braking action ($T_{PM,Brake}$).

The simplified block diagram of the NN-EMS is shown in Figure 4.11 where D_{PM} is the pump/motor displacement, Δp the pressure drop over it, and $T_{PM,NN}$ the predicted torque.

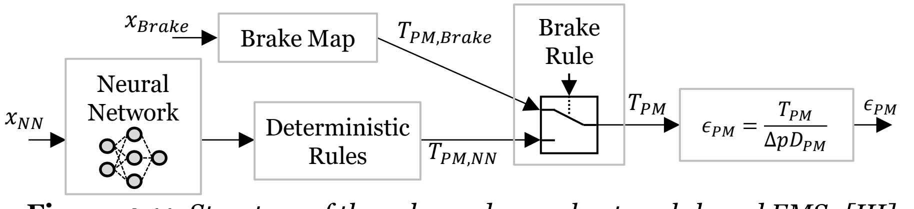

Figure 4.11 *Structure of the rule- and neural network-based EMS; [III].*

4.1.6 Test Results (*Step d*)

In practice, these machines perform a number of different drive cycles, however only short loading drive cycles (Figure 4.1) were used to train the network and are the only type of drive cycle tested.

Despite that the network was trained with data that originated from the recordings of drive cycles from several operators, only one professional operator drove the machine in all tests. This prevented the evaluation of the NN-EMS performance under different drive styles.

Two EMSs were tested: rule-based (RB-EMS) and neural network (NN-EMS). RB-EMS is a rule-based strategy developed by the industrial partner when developing and testing the demonstrator. NN-EMS is the one developed with the method presented in this case study. To have a baseline for comparisons, tests with a conventional machine (Conv) were also recorded. The conventional mode is run with the same machine but with the hybrid system in idle mode.

Four tests were performed for each of the three modes, where each test consists of one hauler loaded with four buckets – therefore, four short loading cycles per test. Productivity (t/h) and fuel efficiency (t/L) were the chosen parameters for evaluation. For that, loaded material, fuel consumption, and cycle time were recorded to assess the machine's performance. A picture of the tests is shown in Figure 4.12.

Figure 4.12 *Picture from the tests; [III].*

Operation of the NN-EMS in the Application

Figure 4.13 shows the operation of the machine for one short loading cycle while controlled by the NN-EMS.

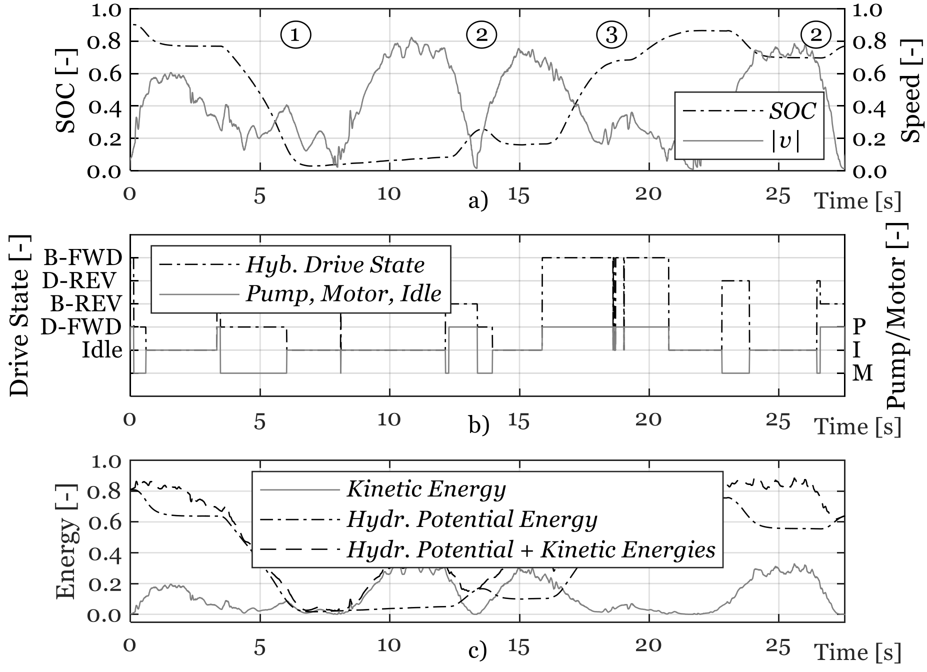

Figure 4.13 *Operation of the NN-EMS in one short loading cycle. a) Machine speed and accumulator SOC; b) Hybrid system drive state and pump/motor operation; c) Kinetic and hydraulic potential energy; [III].*

Figure 4.13a shows the discharge of the stored energy while loading the material (point 1). It also shows the recovery of kinetic energy and reuse of the stored energy when changing direction at point 2 and the use of stored energy when leaving the hauler after point 3.

This behaviour is close to the one obtained from the optimisations, as shown in Figure 4.8. This shows that the network learned and implemented in practice the optimised control decisions. These results provide evidence to confirm the expectation that the networks can implement directly in practice what they learned from the optimisation results from simulation.

Performance Comparison Between EMSs

As mentioned in Section 4.1.1, the productivity of the machine is not affected by the optimisation control decision. However, when controlled by the NN-EMS the machine might have different productivity, which could make it operate outside the conditions from which the network was trained. Therefore, it is important to evaluate how much the drive cycle is changed due to the use of the NN-EMS. If the NN-EMS significantly changes the machine behaviour, this should be seen in the drive cycle. In other words, the dataset shift between training and application would make the controller not operate as intended. Figure 4.14 presents the machine speed, SOC and variation in normalised cycle time and distance travelled for the RB-EMS and NN-EMS.

The cycles with the NN-EMS were similar to the cycles with the RB-EMS. The drive cycle is not significantly affected by the proposed EMS; it had a similar cycle performance. Therefore, at least concerning vehicle speed, cycle time, and state of charge, the NN-EMS is operating in similar conditions to what it was trained for.

Productivity and fuel efficiency are presented for the three operation modes in Figure 4.15. Each point represents one loaded hauler.

Lower productivity was observed for the hybrid modes. The reasons are explained in Paper III. However, the efficiency of the hybrid modes (RB and NN-EMS) is better than that of the conventional machine due to the hybrid system's capability to recover kinetic energy and reuse it. The NN-EMS resulted in better fuel efficiency for similar productivity if compared to the RB-EMS.

The results demonstrate that the NN-EMS performs similarly to the DP-EMS, which is the optimal one in simulation. Therefore, the neural network was able to learn the DP decision-making process and implement it in practice. The implementation of optimised control decisions leads to higher fuel efficiency in practice if compared to the RB-EMS. The reader is referred to Paper III for more results.

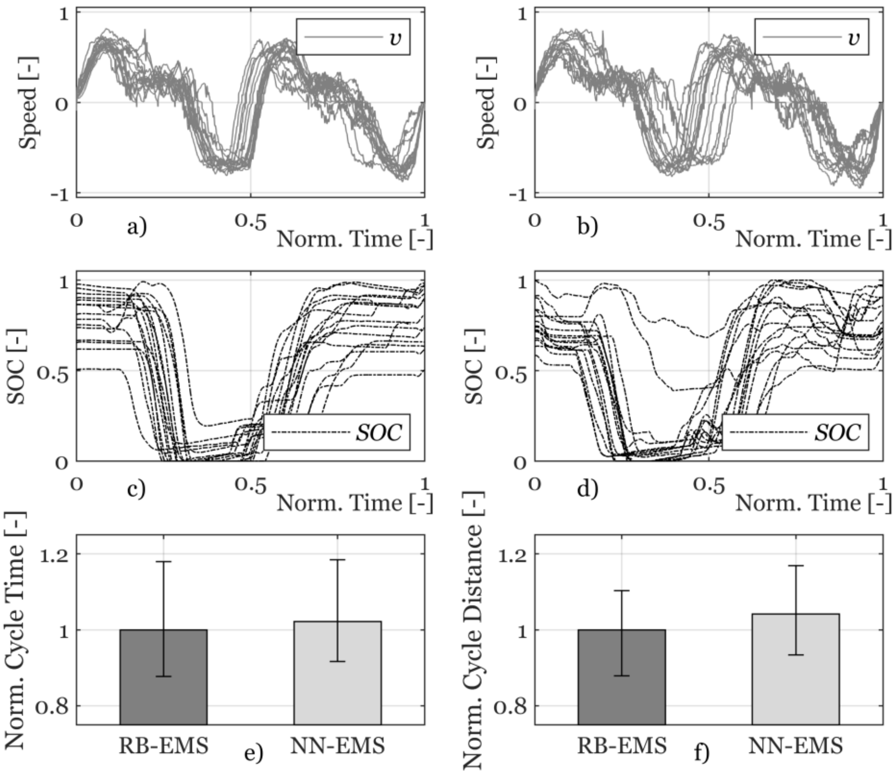

Figure 4.14 *Comparison between operations of the RB-EMS and the NN-EMS. a) and c) Machine speed and accumulator SOC for the RB-EMS; b) and d) Machine speed and accumulator SOC for the NN-EMS; c) Average, maximum and minimum cycle time; f) Average, maximum and minimum cycle distance; [III].*

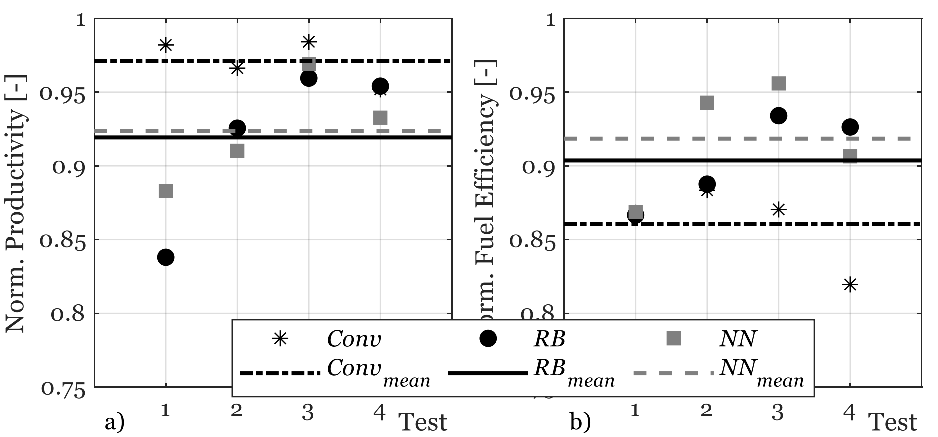

Figure 4.15 *Performance parameters measured from the tests. a) Productivity; b) Fuel efficiency; [III].*

4.1.7 Discussion

The objective of this case study was to demonstrate the implementation of an energy management strategy based on neural networks. More specifically, demonstrate that a controller can be learned from control optimisation results, be directly implemented in the real machine, and yield better machine performance than rule-based EMSs. Despite showing that, there are a few considerations to be made regarding the NN-EMS and its implementation.

For this case, the RB-EMS and the NN-EMS structures are relatively simple. Each strategy could be further developed to yield even better results. However, considering a more complex hybrid construction machine, the RB-EMS strategy might become significantly more difficult to be derived, implemented, and tuned. The NN-EMS would have the advantage to handle more input features due to its function representation capability.

It is expected that RB strategies would require a significant amount of development time for more complex systems. At the same time, this RB-EMS will potentially deviate even more from an optimal EMS. On the other hand, it is expected that NN-based strategies would not require as much development time and would have a greater potential to not deviate as much from optimality.

At the same time, another goal is to reduce the burden on the engineer to interpret the results from DP when deriving RB-EMSs, since the network can learn that, this burden is reduced. The downside is that even if the NN learns the DP decision-making process it would still be difficult to interpret the decisions taken. A shift in time dedicated to each activity is also noticed. When developing an NN-EMS, less time is dedicated to the development of the control knowledge, since that is extracted by the network; instead, time is dedicated to evaluating the outcomes of the learned strategy.

The results also indicated a robustness of the network to differences between the training domain and application domain. This is in part due to the inclusion of control rules into the optimisation process, the variety of cycles used to train it, and in part due to the state sweep. However, it is necessary to extend the covered range of operation scenarios to evaluate if similar robustness can be achieved. A deeper analysis of this topic is provided in Chapter 5.

There can be a discussion if the number of operators used in the tests, the number of drive cycles, the variation in types of loaded material, and so on, were sufficient for a broad qualitative assessment of the method. With that in mind, it is understood that the obtained test results serve as a proof of concept regarding the feasibility and potential of NN-based controllers for energy management in construction machines. However, future tests covering a broader range of operational conditions would provide additional resources for an assessment regarding the potential and feasibility for such controllers to be implemented in machines reaching the market.

4.2 Case Study II – Multi-Chamber Actuator Mode Selection Through Reinforcement Learning

In this case study, it is demonstrated through simulation and experimental results the training and implementation of a reinforcement learning-based (RL) controller for the mode selection of a multi-chamber hydraulic actuator. The actuator is part of a multi-actuator load sensing architecture driven by a single pump. This system is used to drive the boom and stick functions of an excavator arm.

The selection of different modes on the multi-chamber actuator allows the reduction of resistive losses in the control valves caused by unmatched pressure levels between actuators. A Deep Q-Learning (DQN - [112]) agent is created to learn how to select the modes to minimise the system energy losses.

In comparison to the first case study, presented in Section 4.1, this case study evaluates a technique that performs the optimisation of the control decisions while it learns the strategy. It also has the inherent conditions for continuing the training after deployment, with minor changes to the training algorithm.

This case study is related to papers IV and VI. The concept of the studied system is presented and theoretically evaluated in Paper VI, while Paper IV presents the study and results regarding the development of the RL agent and its performance in simulation and experiments.

4.2.1 System Description

In a multiple hydraulic actuator system, the load on each actuator results in different pressure levels. If the actuators belong to a load sensing architecture with pressure compensation valves, the supply pressure is regulated to the highest pressure required by the actuators. This causes a mismatch between pressures on the other actuators and the pump pressure, which is compensated by throttling in the valves.

In Paper VI it was shown that the reduction of such losses can be achieved by using a multi-chamber actuator in one of the functions. The hydraulic circuit diagram of the suggested system, where Load 2 is driven by a multi-chamber actuator, is show in Figure 4.16. An extended discussion on this system architecture is given in Paper VI.

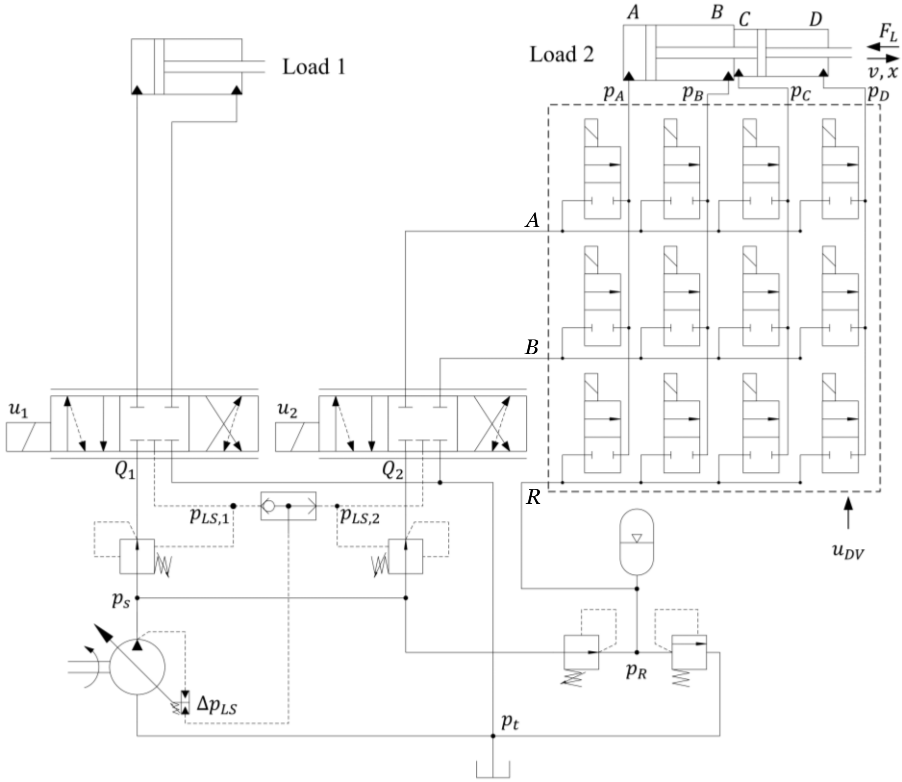

Figure 4.16 *Hydraulic system architecture of the valve-controlled pressure-compensated load sensing system with a multi-chamber actuator.*

The combination of different chambers defines possible actuator modes to be used. The selection of different actuator modes allows the modulation of the resultant pressure from the load to make $p_{LS,2}$ similar to $p_{LS,1}$. This reduces the throttling losses in the system. This pressure modulation is illustrated in the flow pressure diagrams in Figure 4.17.

The multi-chamber actuator adds another variable that must be controlled. In this concept, the mode selection acts as a motion enabler, which is still controlled by the proportional valve. Therefore, the control goal for the multi-chamber actuator is to select a mode that enables the requested motion to be completed while minimising the resistive control losses.

Although not dealing with energy recuperation or energy storage, the mode selection is performed to minimise the energy losses in the system. Therefore, this control problem can still be interpreted as being an energy management problem. At the same time, the focus here is on the evaluation of the method to generate an EMS and not on the system concept. It must be emphasized that this method could also be applied to other control problems, including the one presented in Case Study I.

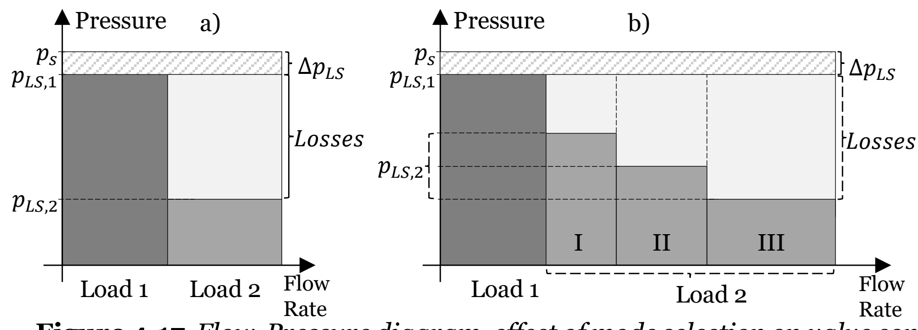

Figure 4.17 *Flow-Pressure diagram, effect of mode selection on valve control losses in a load sensing system with two actuators. a) The case with two conventional actuators; b) The case where one is a multi-chamber actuator.*

The loads and state conditions that such systems operate under are highly variable and cover a broad range. An analytical solution, or a solution based on rules, to determine the control action that results in reduced energy losses is not trivial and can be difficult to find. Therefore, the objective of this case study is to demonstrate that RL can also be used to find the EMS to control hydraulic systems of construction machines. This case study is about the selection of modes of multi-chamber actuators to reduce resistive control losses.

RL can be applied directly to the real system for training. However, this case study takes a similar approach as the previous one. A model of the real system is used for the agent to learn the EMS by interacting with it. The study is limited to only training the agent in the simulation environment and deploying it to the real system for experimental evaluation.

4.2.2 Modes, Model, and Control Structure

The multi-chamber actuator is connected to three supply lines (A/B/R). Most of the possible combinations of chamber areas are ruled out. The reasons for that are described in Paper VI. The modes used in this study are given in Table 4.1, Only mode 1 is not an agent's decision, as it is implemented as a rule for safety reasons. Modes are ordered in a decreasing force capacity.

The model of the physical system is developed in the multi-domain system simulation tool HOPSAN [113]. The model describes the motion of the boom and stick as functions of the motion of the hydraulic actuators. It also models the behaviour of pressures and flow rates in the hydraulic system. The verification of model agreement with experiments is described in [99]. Both the controller and training algorithm were implemented in Simulink [112].

Table 4.1 *Multi-chamber actuator modes used in the case study.*

Mode	A	B	R	Mode	A	B	R
1	-	-	-	4	A	BD	C
2	AC	BD	-	5	AB	D	C
3	ABC	D	-	6	C	BD	A

It is a forwards-facing model, which makes the performance of the load cycle not be prescribed as in Case Study I. Therefore, the control decisions of the RL agent affect the power losses of the system and how fast or slow the motion is performed. This allows the RL agent to find a trade-off between selecting a mode that is efficient but results in a slow motion or finding a mode that is less efficient but is faster.

Aside from the agent, the controller also contains other control and safety-critical rules. The motivation to include rules along with the agent, which is a neural network, are similar to the ones for the previous case study. Some situations cannot be prone to miss-predictions of the network and must be handled separately. Since the best control solution must account for these situations, these rules must be placed along with the agent when it is learning how to control the system.

In this case study, the training of the agent was only performed in simulation. However, in the case of letting the agent to train in the real system, there would be additional rules to ensure that the exploration phase, when the agent is interacting with the system to learn how to control it, is performed safely.

An overview of the control structure is shown in Figure 4.18. The reward branch of the agent is only used during the training phase.

Paraphrasing [95] for the present control problem: *The agent interacts with the environment, in a sequence of actions, observations, and rewards. At each time-step (t) the agent observes the observations* $s(t)$*, selects an action* $a(t)$ *from the set of modes (Table 4.1) and applies it to the environment. It observes the new observations* $s(t+1)$ *and the reward* $r(t)$. Based on the reward it updates the parameters of the network to improve the mode selection the next time it encounters a similar load and state condition. The agent tries to maximise the collected reward. A description of the network and its function is presented in Section 4.2.3.

In an RL framework, everything that is not the agent is the environment the agent is interacting with. Therefore, it is fundamental to make the environment in simulation as close as possible to the real environment. This includes the additional rules for safety or other situations and a sufficiently representative model of the physical system. This should reduce the differences between the development environment and application environment. If the differences are small, it is expected that the agent can control the real system despite training in simulation.

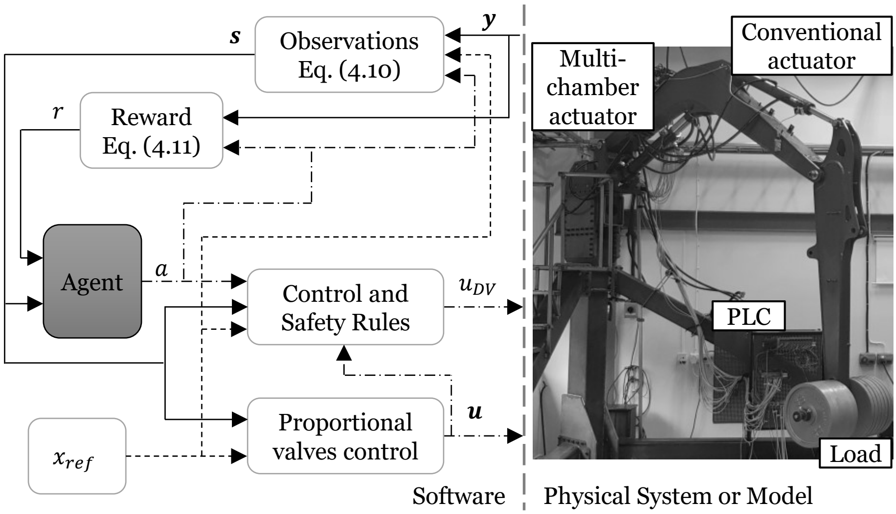

Figure 4.18 *Structure of the proportional valve and mode selector controller and picture of the test bench.*

A detailed description of the added rules for this case can be found in Paper IV. They mainly prevent modes that cannot drive the load to be applied in the system, handle mode 1, and the boom lowering motion, which is performed with a predefined mode.

As mentioned, the RL agent is responsible for the mode selection to enable the motion, another P controller 'mimics' the machine operator controlling the proportional valves. The P controller is what controls the boom motion.

4.2.3 Learning Setup for the Agent

A Deep Q-Learning (DQN) agent was selected because it allows the use of continuous variables as input features (measurements from the system) and discrete variables as output (modes).

The agent has a neural network that estimates the value of taking a certain action given the current observations (input features). The value is an estimation of the sum of the reward that the agent can collect over a future time horizon. At each agent's control time step, the value for taking an action is calculated and a greedy function selects and implements in the system the action with the highest value. In other words, the action that would lead to the best performance, according to the reward function.

The training algorithm makes sure that some exploration happens to ensure a search for decisions that yield a higher reward. If exploration doesn't happen, the agent would only exploit the 'best' decision found. In other terms, the exploration acts as an incentive for the agent to search for the global optimal solution and not converge to a locally optimised solution. The reader is referred to [95] and [112] for a detailed description of the type of agent and training algorithm.

This agent (neural network and greedy function) is a non-linear map between the multi-dimension space of input features (pressures, speed, position, ...) and the optimised action (modes). Each agent action corresponds to one mode (Table 4.1). The structure and parameters of the network are presented in Table 4.2.

Table 4.2 *Architecture of the action-value network.*

Layer	Size
Feature inputs (observations)	9 features
Fully connected with ReLU activation function	70 neurons
Fully connected with ReLU activation function	35 neurons
Fully connected with linear activation function (actions' value)	5 neurons

It is also a multi-layer feed-forward neural network. Despite the fact that the name of the agent refers to a deep network (several hidden layers), in this study, a shallow network with only two hidden layers was sufficient to encode the knowledge on how to control the system.

Only variables that can be measured in the system or extracted from the controller structure were used as observations (s), which are the input features for the network. They are,

$$\boldsymbol{s}(t) = [a, u_2, p_A, p_B, p_C, p_s, p_{LS,1}, v, x], \tag{4.10}$$

where a is the previous action, u_2 is the proportional valve control signal, p_{A-B-C} are the chambers' pressures, p_s is the supply pressure, $p_{LS,1}$ is the load sensing pressure for the conventional actuator, v the actuator speed, and x the actuator position.

The reward function is composed of three terms,

$$r = K_1 r_{Velocity} + K_2 r_{Power} + K_3 r_{Switch} \tag{4.11}$$

The velocity term ($r_{Velocity}$) penalises the agent if boom doesn't move with a minimum velocity. This encourages the agent to learn to meet a minimum control performance requirement. The power loss term (r_{Power}) is a penalty based on the hydraulic system pressure compensation losses to encourage the agent to find a mode that makes the motion be performed more efficiently. The switch term (r_{Switch}) penalises the agent for frequent mode switching. A detailed formulation of the terms r_i is presented in Paper IV.

What this reward function highlight is the fact that the optimisation objective can be composed of several goals. This does not mean that the task of formulating the reward function is simple. The weights between the terms (K_i) also play a guiding role in what the agent learns. They can be tuned to achieve the desired control behaviour.

Table 4.3 presents the twelve load cases used during training and testing. Other training parameters can be found in Paper VI.

Table 4.3 *Load cases and task used for training the agent.*

Load cases and task	Value
Initial position [m]	$0.10 + 0.02R$
Final position [m]	$0.40 + 0.05R$
External load [kg]	[40 80 120 160 200 240]
Load 1 load sensing pressure $p_{LS,1}$ [bar]	$[60\ 100] \pm 3R$
Load 1 flow rate Q_1 [lpm]	10

The agent is trained to lift the boom, under twelve different load cases, from a bottom position to an upper position. The model is not a perfect representation of the real system. Therefore, to increase the agent's robustness to the real application, randomisation (R – between 0 and 1) was added to the load cases and initial and final positions. Noise was also added to all the measured variables in the simulation.

In an excavator application, the external load would not be constant, there wouldn't be an initial and final position, there would be actuation in the four quadrants of force-speed directions, and more actuators are used simultaneously. As a consequence of the used test bench, the control problem that the agent faces is considerably different from that of an excavator operation. However, it is still a challenging control problem, it has the same implementation challenges as there would be for a construction machine, it has similar safety requirements, and the hydraulic system is similar to that of a mobile machine. Therefore, the objective of evaluating the capability of the RL-based method can still be achieved with this test bench and load cases.

4.2.4 Results

The trained agent was tested in simulation for all twelve load cases to evaluate its performance. First, to see if it learned to move the load from start to finish, and second, to see if it could learn to do so with modes that result in lower energy losses. The energy loss (E_{loss}) is calculated with equations (4.12) and (4.13).

$$P_{loss} = P_{loss,1} + P_{loss,2} = |Q_1(p_s - p_{LS,1})| + |Q_2(p_s - p_{LS,2})| \quad (4.12)$$

$$E_{loss} = E_{loss,1} + E_{loss,2} = \int_{t_i}^{t_f} (P_{loss,1} + P_{loss,2})dt \quad (4.13)$$

Simulation Results

An estimation of what mode results in the lowest power losses is made by evaluating Equation (4.12) for each mode and steady state conditions of the speed and position of the multi-chamber actuator. This indicates what mode selection to expect from the agent's learning. The solution for two load cases is shown in Figure 4.19. The safety boundary condition limiting the number of modes that can be applied to the system is also shown. This safety boundary is part of the additional rules mentioned earlier.

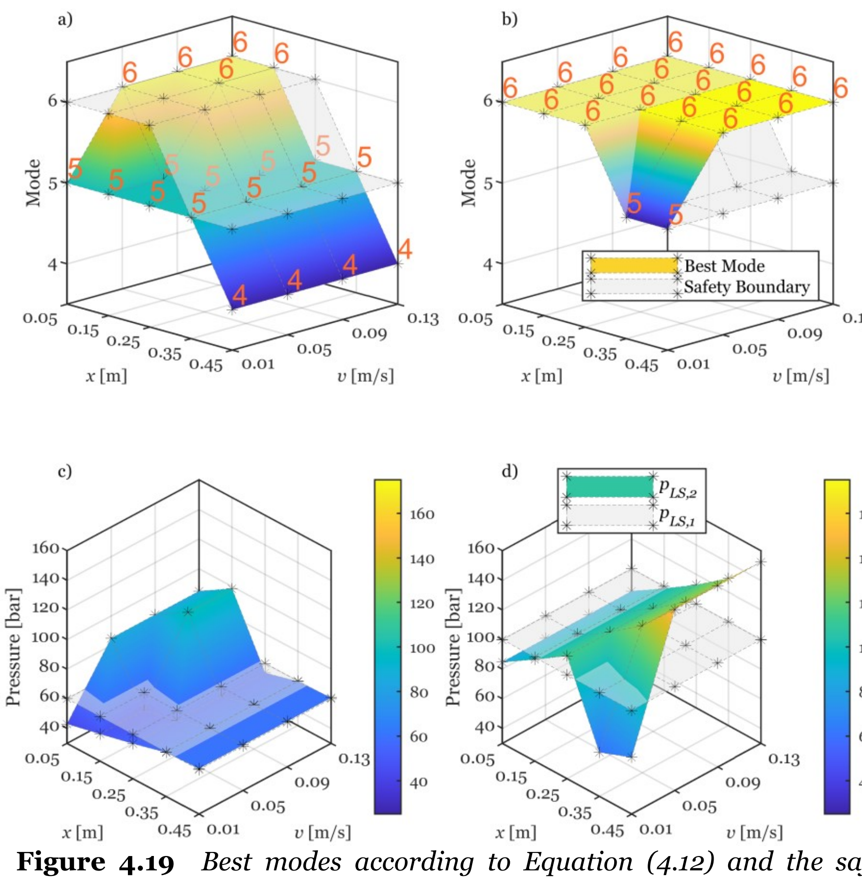

Figure 4.19 *Best modes according to Equation (4.12) and the safety boundary, 120 kg external load. a) For $p_{LS,1} = 60\ bar$; b) For $p_{LS,1} = 100\ bar$; Resultant $p_{LS,2}$: c) For $p_{LS,1} = 60\ bar$; d) For $p_{LS,1} = 100\ bar$.*

What Figure 4.19a shows is that modes that can exert less force are better when the actuator is retracted, while modes that can exert higher forces are better when it is extended due to the geometry of the boom arm. This results in a closer match of $p_{LS,1}$ and $p_{LS,2}$. However, the mode that results in the lowest power losses depends on the flow rate as well, Equation (4.12). That is the reason why, at higher speeds, weaker modes are chosen, because they result in a lower required flow rate. For a larger $p_{LS,1}$ pressure, the weaker modes result in a better match of pressures.

The theoretical assessment of the variables that compose the reward function, like pressures presented in Figure 4.19, provides insights into what the agent should learn. They provide guidance for tunning the reward function and for the evaluation of the agent's performance after training. However, as shown in Equation (4.11), reward functions usually have multiple terms which can be difficult to evaluate, especially when some of the terms are conflicting.

Figure 4.20 presents simulation results, for the trained agent performing the mode selection for the one load case presented in Figure 4.19. For comparison, it is also shown the system performance on the same tests but with a conventional actuator instead of the multi-chamber actuator for Load 2. Conv and RL are the abbreviations for the conventional actuator and the multi-chamber actuator, respectively.

Figure 4.20a shows the task completion, while Figure 4.20b shows the selected modes. If comparing Figure 4.20b with Figure 4.19b, it is noticed that the agent does learn to select modes that minimise energy losses, and that, in turn, results in a better match of pressures than the conventional case, as shown in Figure 4.20c. Consequently, the power losses are reduced significantly, as shown in Figure 4.20d.

All load scenarios were simulated, and Equation (4.13) is used to compare the energy losses. Results for some of the load cases, normalised by the highest total energy losses for all test cases of each level of $p_{LS,1}$, are presented in Figure 4.21. More results are presented in Paper IV.

The simulation results showed that the agent learned to select modes that result in lower energy losses. Exceptions occurred for some of the load cases where the agent chose to run the system as a conventional actuator (mode 2) but does not choose something worse than that, e.g., a mode that cannot drive the load.

Overall, what these results show is the capability of this type of method to find an optimised solution for the control problem with multiple objectives. The agent was trained several times and similar solutions were found, however, it is not known if it is the global optimal solution or a local optimum. This topic is further discussed in Section 4.2.5.

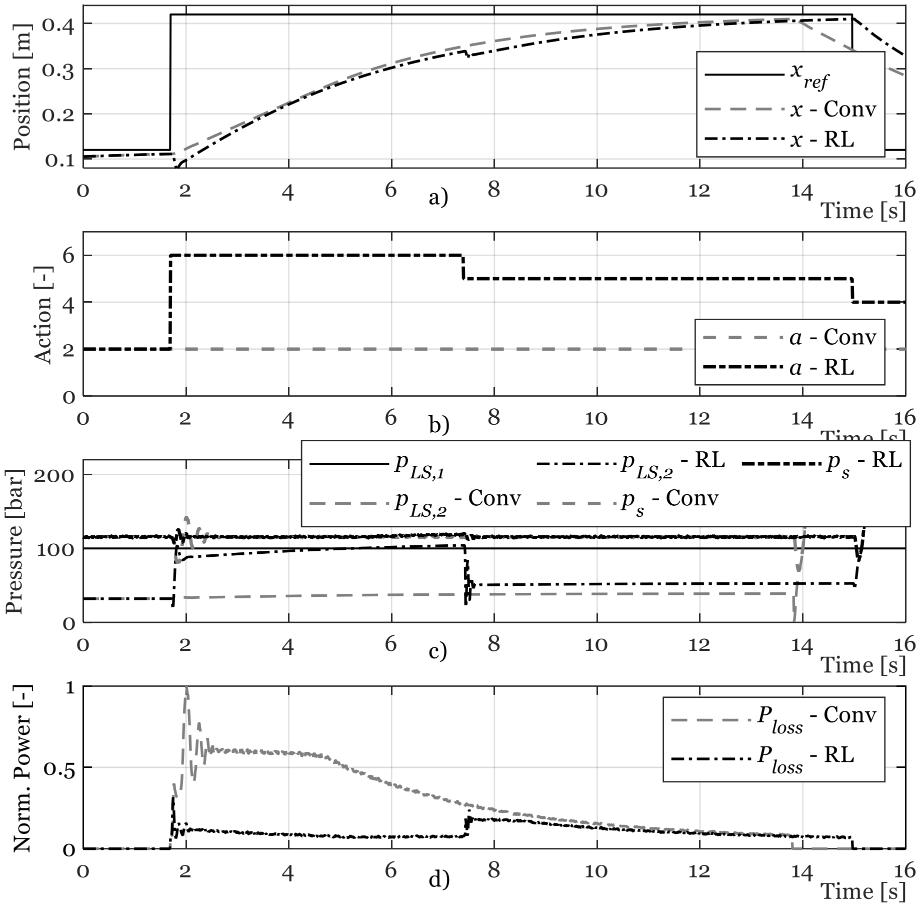

Figure 4.20 *Simulation result, 40kg external load and $p_{LS,1} = 100\ bar$. a) Actuator position; b) Agent action; c) System and load sensing pressures; d) Power loss.*

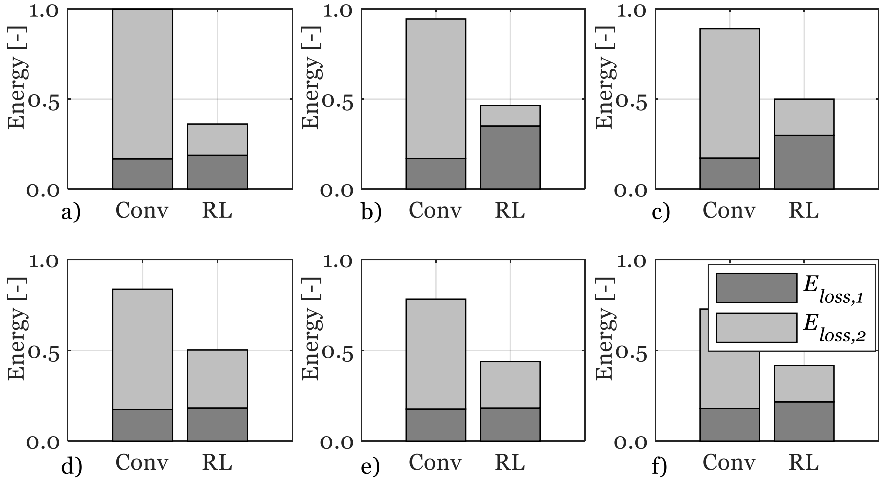

Figure 4.21 *Energy comparison between conventional system and the system with multi-chamber actuator with mode selection,* $p_{LS,1} = 100$ *bar. Loads: a) 40 kg; b) 80 kg; c) 120 kg; d) 160 kg; e) 200 kg; f) 240 kg.*

Experimental test performance

The trained agent was deployed to perform the mode selection on the real system. It is evaluated on the same load scenarios used in the training. The experimental test results, for the same load case of Figure 4.20 is shown in Figure 4.22.

The agent was able to apply in practice the same modes it applied in the simulation (Figure 4.20b – Figure 4.22b) even though there are differences in the input features. This can be noticed by comparing the actuator position (Figure 4.20a – Figure 4.22b) and system pressure (Figure 4.20c – Figure 4.22c).

Collective results for the agent's decisions in simulation are compared to the decisions in the experiments. They are shown in Figure 4.23. The abbreviations Sim and Test are used for the simulation and experimental test results respectively.

In most of the cases, the agent was able to implement in the real system the decisions also taken in simulation. Examples of load cases that failed are shown in figures 4.23b and c. Even though these cases show some deviation, they were not wrong during the whole test. This means that only a few decisions were different. However, they still follow the logic of choosing weaker modes at the start of the motion and stronger modes at the end of the motion. It cannot be said that the decisions in the experiments are wrong because the input features are different from the ones in the simulations. Different input features should lead to different control decisions depending on the generalisation achieved during training.

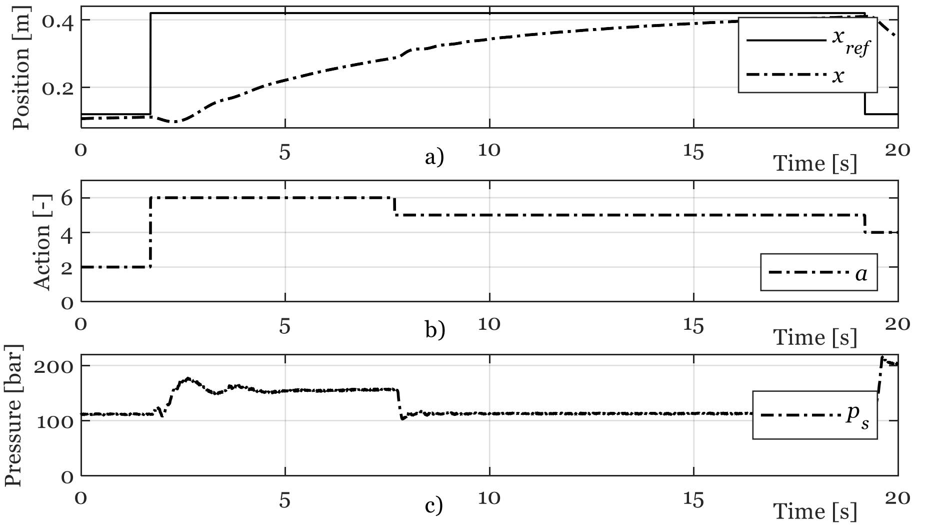

Figure 4.22 *Experimental results for the system with multi-chamber actuator and mode selection, 40kg load and* $p_{LS,1} = 100\ bar$*. a) Position; b) Action; c) System pressure.*

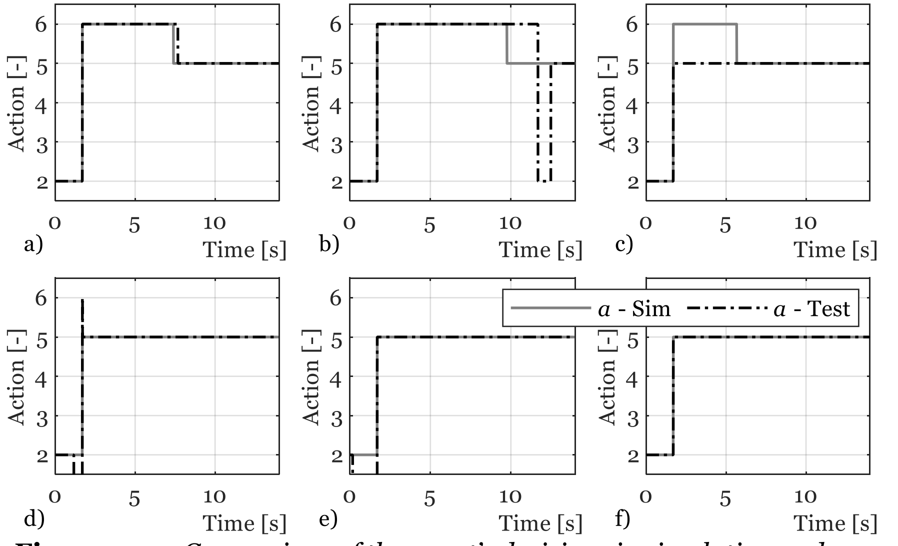

Figure 4.23 *Comparison of the agent's decisions in simulation and experiment,* $p_{LS,1} = 100\ bar$*. Loads: a) 40 kg; b) 80 kg; c) 120 kg; d) 160 kg; e) 200 kg; f) 240 kg.*

In this case study, only the selected mode was recorded. However, a more detailed assessment of the learning could be made by evaluating the output of the network which is predicting the value of taking each action. Changes in those values could be compared to changes in the input features to detect which inputs affect more the output. This could be used to understand the high sensitivity observed in the test cases.

The high sensitivity to the variation in input features seemed to result in low repeatability in the selection of modes in some of the load cases. The worst situation was observed for the load case of 120 kg and $p_{LS,1}$ 60 bar. An example is shown in Figure 4.24.

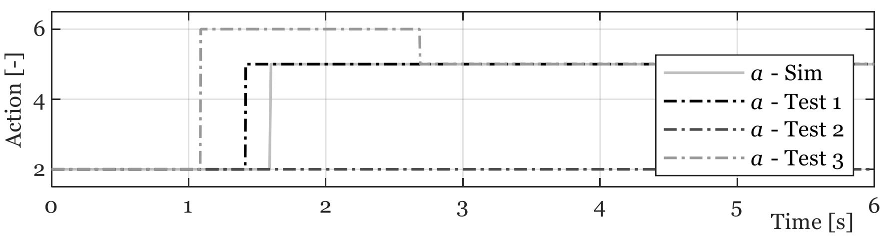

Figure 4.24 *Repeatability issues for 120 kg load and* $p_{LS,1} = 60$ *bar*.

Results of Figure 4.24 represent the fact that data-driven controllers, as expected, might make different predictions due to operating in different conditions that they were trained for. However, for the trained agent presented in this study, the decisions still made the agent complete the task in an acceptable way.

4.2.5 Discussion

The energy analysis based on the simulation results showed the reduction in energy losses caused by the selection of suitable modes by the agent. However, that does not mean that the training of the agent converged to the global optimal solution.

The model of the system was judged sufficiently accurate to describe the main characteristic and behaviour of the system. Still, there are deviations from the actual behaviour of the system. This supports the fact that RL-based methods can make use of simplified models to automatically obtain controllers that perform well in practice.

Although the use of simplified models was sufficient to pre-train the agent, its performance could likely be further improved by letting it, in a safe manner, interact with the real system and train from those experiences. It is also expected a higher performance if a more accurate model is used. In this sense, there seems to exist a trade-off to be found between using simpler models for pre-training with more training in the real system or using more accurate models for the pre-training and less training in the real system. The first approach with simpler models should lead to faster training of the agent, less cost with model verification, and more costs with the training in the real system. The second approach with more representative models should result in slower training, higher costs with model verification, and fewer costs with the training in the real system.

One limitation of this study was to evaluate the agent under similar conditions for which it was trained. In the simulation, random noise was introduced to all measurements from the system to increase its robustness in the experiments. However, there are still questions to be answered, and an evaluation to be made, regarding the robustness and generalisation capability of the agent.

5

Reliability of Machine Learning-Based EMSs

Despite the focus of the thesis being on the energy efficiency aspect, safety aspects cannot be neglected. For machine learning-based EMSs there are characteristics of the methods that inherently make them prone to result in an unsafe operation.

This section aims at assessing the reliability aspects of the EMSs developed in Chapter 4. The presence of dataset shift and absence of ground truth hinders the performance and reliability of the EMSs. Therefore, one would like to know if, when, and where the learned strategies can be trusted. In this chapter, they are assessed by comparing the training and application datasets.

5.1 Safety vs. System Architecture

As discussed previously, the systems for whom the EMSs based on machine learning were developed are sub-systems belonging to a machine. Therefore, the impact of the developed EMS on the machine safety is a consequence of what system it is controlling. In certain cases, the subsystem being controlled lies in parallel with the remaining sub-systems of the machine, as represented in Figure 5.1. It is a generalised representation of systems like the one studied in Section 4.1.

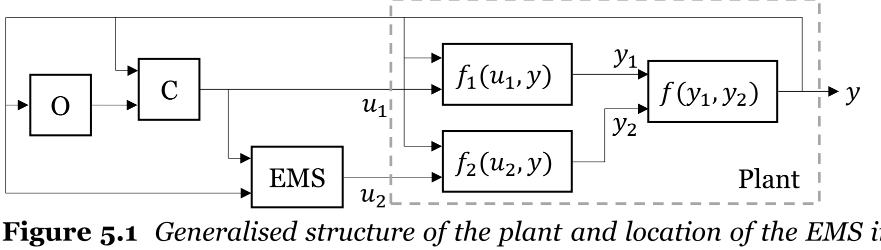

Figure 5.1 *Generalised structure of the plant and location of the EMS in the control chain, parallel case.*

In the figure: O represents the operator; C another lower level controller for other subsystems; the EMS is the one based on ML; u_1 and u_2 are control signals; f, f_1, f_2 are functions representing subsystems; y_1

and y_2 are the states directly affected by each subsystem; and y is the machine states, which might contain y_1 and y_2.

The relation between the output (y) and the control actions (u_1, u_2) for this type of system can be expressed as

$$y = f(y_1, y_2) = f\big(f_1(u_1, y), f_2(u_2, y)\big). \tag{5.1}$$

The parallel nature of the system results in that the output of the system is not only a function of the subsystem controlled by the EMS. If the subsystem f_2 has a strong impact on the output (y) than the safety requirements around this EMS would be strong as well. If there is a weak impact on the output, than the safety requirements are also weak. This is an argument based on the physical coupling between the subsystems regarding safety of implementing and EMS based on ML. For example, if the power delivered from f_2 is equivalent to the one delivered from f_1 it might be and indication of strong coupling and strong requirements on safety.

At the same time, the coupling between f_2 and the output can be made weak by control means. For example, make the f_2 transparent to the operator, in the sense that the output (y) can only be affected as much as the operator requires it to be affect. This means that the EMS cannot drive the output on its own. This control approach does not affect the strength of physical coupling and can be interpreted as a safety constraint.

However, the subsystem being controlled by the EMS might be in series with other subsystems, as represented in Figure 5.2. It is a general representation of the system studied in Section 4.2.

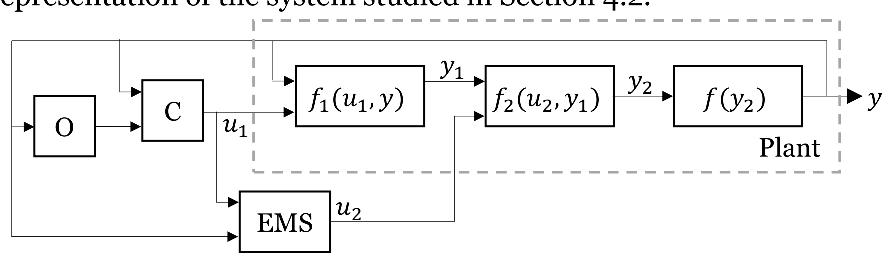

Figure 5.2 *Generalised structure of the plant and location of the EMS in the control chain, series case.*

The relation between the output (y) and the control actions (u_1, u_2) for this type of system can be expressed as

$$y = f(y_2) = f\big(f_2(u_2, y_1)\big) = f\left(f_2\big(u_2, f_1(u_1, y_1)\big)\right). \tag{5.2}$$

The output is a direct function of the subsystem controlled by the EMS, which means that the strength of this coupling is strong. Every undesired control action might negatively affect the output of the system, and the presence of this subsystem is difficult to make transparent to the operator. Therefore, this type of system results in stronger safety requirements to the development of EMSs based on machine learning.

Although important for both types of systems, the safety requirements will depend on the type of system being controlled.

5.2 Ground Truth and Dataset Shift

This section applies the method to generated EMSs presented in Section 4.1 as an example, but the concept applies to the method presented in Section 4.2 as well.

In [75], an important question is posed: "how can we affect (ensure) performance on the application set when we only get to observe the training set?" This question is also asked here but in the context of developing EMSs based on machine learning. This question leads to the fundamental assumptions under which these methods to generate EMSs are expected to work. They are as follows:

- By sampling drive cycles from real machines one can represent the population of possible working scenarios;
- By applying control optimisation to a sufficiently accurate system model, one can gather information on how to control the real system in an optimised way;
- There is a function between system states and optimal control decisions that can be learned and implemented in practice with machine learning;
- The machine performance will not be significantly affected when controlled by the learned EMS.

It is clear to see that these assumptions can be significantly affected by the uncertainties and differences between training and application that belong to the process of generating the EMSs.

Regarding the first assumption, in theory one could use a very large sample of the population as input to the method. However, construction machines are versatile and perform many tasks that might be impractical to collect an extremely large sample to cover the population. Therefore, it is known from beginning that the EMS might operate in conditions that it was not exactly trained for.

Regarding the second assumption, the problem is that there are model deviations; the discretisation of control actions and state space is not infinite in the optimisations; there can be inconsistencies in the drive cycle definition used as inputs for the model; and so on. Therefore, the solution obtained from the optimisation might not represent the optimal solution for the real system.

Despite having powerful function representation capability, machine learning models will hardly acquire a perfect learning of the task. At the same time, it is desired that they generalise for unseen cases, which means, one wants them to have a high performance in this inference from sample to population. Therefore, despite having an accurate model and optimisation, still the EMS might not apply the best decisions in practice. This is a point against the third assumption.

The EMS obviously affects the performance of the machine, which might make it operate in a different performance than initially trained for. In this case, it will operate outside its training conditions. This is a point against the fourth assumption.

The problems with sampling the population, model deviations, differences in optimisation results from model to reality, improper learning of the control strategy, and change in performance might negatively affect the performance of machine learning-based EMSs. These problems can be summarised as the methods suffering from absence of ground truth and presence of dataset shift. In this chapter, the absence of ground truth and dataset shift are evaluated by comparing the datasets generated along the EMS development process, shown in Figure 5.3.

Three datasets are generated: training, testing, and application. The testing dataset is used to evaluate the training of the network while the application dataset is from the recorded operation of the EMS when deployed to the system as the controller.

As a consequence of issues mentioned earlier, training and testing datasets are a representation of the application domain. However, it is expected that if the differences are small enough the assumptions made might be true and the EMSs should work reasonably well. In other terms, if there is a small dataset shift, and if the estimation of the ground truth is reasonable, the methods should yield reasonable results.

5.3 Methodology

The goal is to make an assessment of how well and safe the EMS is operating in a different domain. The reliability and performance of the resultant EMS are quantitatively assessed in the following ways between the training and application datasets:

- Calculating the range mismatch of input features;
- Calculating the difference in probability of occurrence (density) of input features;

- Calculating the performance of the network on the training dataset as the average absolute error along the covered space of input features, and using it to estimate the performance of the network on the application dataset as the error along the covered space of input features.

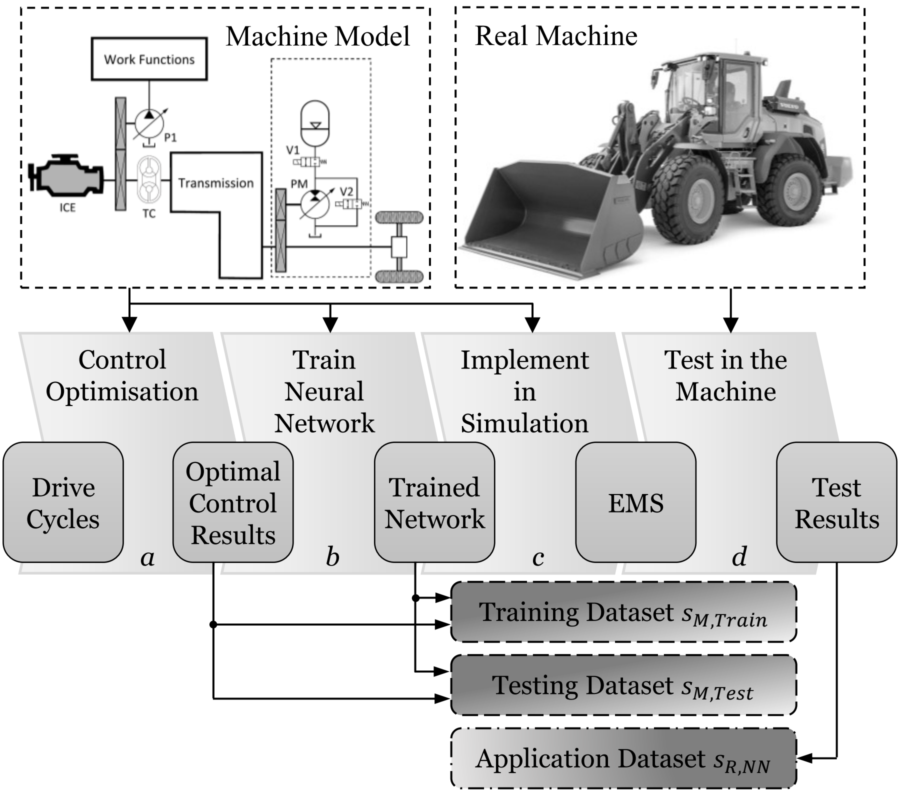

Figure 5.3 *Method to generate EMSs based on dynamic programming and neural networks and the generated datasets during the process. It is also shown what stages use models of the machine and what stage the real machine is used; [Machine image Volvo CE].*

The following equations are used to calculate the data mentioned above. If two large and representative samples, A and B, are extracted from the same population, and a neural network would be trained on the A sample, it is reasonable to assume the performance (e – error) of the network in the two samples to be approximately the same. This is even more certain if the network has a smooth interpolation between the training points. In this sense, the average performance of the network ($\bar{E}$

– average error) across close points of the samples would be similar. Therefore,

$$\bar{E}(A) = \bar{E}(B). \tag{5.3}$$

The average error is calculated by dividing the input space of features into small bins. The sum of the network prediction error on points inside each bin is calculated and divided by the number of points in the bin. An approximation to the total error (E) in each bin can be calculated with

$$E(A) = \bar{E}(A)n(A), \tag{5.4}$$

$$E(B) = \bar{E}(B)n(B), \tag{5.5}$$

where, n is the number of points in a certain bin. Combining equations (5.3) and (5.5) leads to

$$E(B) = \bar{E}(A)n(B) \tag{5.6}$$

which is an estimation of the network total prediction error on sample B given the average error on sample A. This equation, of course with all the assumptions around it, allows for the estimation of the network performance for situations where the ground truth or target is not known for the application.

The probability of occurrence is defined as the ratio of points in a certain bin divided by the total number of points

$$p(A) = \frac{n(A)}{N_A}, p(B) = \frac{n(B)}{N_B} \tag{5.7}$$

The difference between distributions is calculated between the two probabilities of occurrence

$$\Delta p = p(A) - p(B) \tag{5.8}$$

where regions with a negative value of Δp represents a lack of data in the A sample and positive values represent an excess of data. If A represents the training data set and B the application dataset, this means respectively, that the network had no sufficient information for the regions where it operates the most (B), and that it is likely biased towards regions that do not frequently occur.

It is important to make the right comparisons, since there is a significant distinction between comparing the network itself or comparing the EMS. If the generalisation of the EMS is to be accessed, then the rules and limits of operation that define the EMS must be accounted for. This means that the input features space of the network must be evaluated, but only for the situations where the network predictions reach the system.

Here, dataset shift is viewed in a more macro way, and the different types and sources of dataset shift are not classified. The reader is referred to the approaches of [105], [106], and [111] to interpret and identify dataset shift in a more detailed way.

5.4 Reliability and Performance

In practice, for the present method, the effects of the sources of dataset shift and absence of ground truth are combined. Here, no effort is placed in trying to isolate them. However, the effect of this combined problem is explored in depth by distinguishing the datasets and representing them graphically, see Figure 5.4.

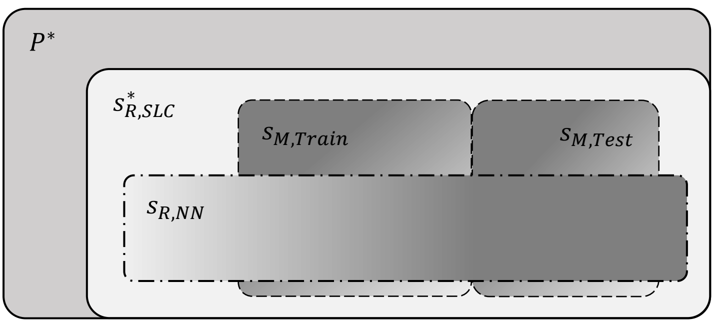

Figure 5.4 *Representation of the relationship between population, sample, and datasets generated for the method based on dynamic programming and neural networks.*

The multi-dimension state space vector of input features ($\boldsymbol{x}_R$) and global optimal decisions (u_R^*) for the real machine (ground truth) defines the population (P^*). It is what one would ideally like to cover with the EMS development method. The superscript * represents the global optimal solution (ground truth). The dataset is mathematically represented as

$$P^* = [\boldsymbol{x}_R, u_R^*]. \tag{5.9}$$

This population is defined by all possible applications and tasks that the machine (architecture + controller + driver + environment) can engage into, for example, short loading cycle, load and carry cycle, log stacking, and so on. Note that these types of application may or may not intersect or share similar states.

Extracting only, yet all short loading cycles from P^* means that only a sample of the population is being considered in the dataset. As described by [111], this type of action during the design phase introduces a bias to the training dataset. This sample of all short loading cycles is named $s^*_{R,SLC}$ and contains all possible states ($\boldsymbol{x}_{R,SLC}$) with their respective optimal control decisions ($u^*_{R,SLC}$). It is described as

$$s^*_{R,SLC} = \left[\boldsymbol{x}_{R,SLC}, u^*_{R,SLC}\right]. \tag{5.10}$$

To simplify the notation, in the continuation of this chapter the subscript SLC is removed from the notation of the datasets belonging to this group.

Up until this point perfect knowledge of the system and optimal decisions have been assumed. In practice, the sample of short loading cycles used in the development of the EMS is not composed by all possible scenarios in short loading applications. The sample aims at extracting a representative sample of short loading cycles but might fail to do so. Additionally, in practice there is no available ground truth for the real machine, and the optimal solution is only an estimation. Instead, for the present method, an estimation of the optimal control decision is obtained with the use of models and optimisation.

All these effects combined define a new dataset (s_M) composed by the states of the machine obtained from the model and recorded inputs ($\boldsymbol{x}_M$), and the estimation of the optimised control decisions (u^o_M). This is represented as

$$\begin{aligned} s_M &= s_{M,Train} \cup s_{M,Test} \\ s_M &= [\boldsymbol{x}_M, u^o_M]_{train} \cup [\boldsymbol{x}_M, u^o_M]_{test}, \end{aligned} \tag{5.11}$$

where the sample is split in a dataset for the training of the network $s_{M,Train}$ and the testing of the network $s_{M,Test}$. Here, the superscript ^{o} refers to the optimisation results, not the ground truth. The subscript M shows that the sample is generated from a model. It has a similarity to the original population P^* and sample $s^*_{R,SLC}$, but it is somewhat distorted by the dataset shift and absence of ground truth.

After deploying the trained network to the real machine for testing, another dataset is generated that is composed by states of the application ($\boldsymbol{x}_{R,NN}$) and by predicted control decisions from the network ($u_{R,NN}$)

$$S_{R,NN} = \left[\boldsymbol{x}_{R,NN}, u_{R,NN}\right]. \tag{5.12}$$

The subscript NN is introduced to indicate that it is a neural network decision or states affected by it. The subscript R means that the sample is generated from the real machine, not the model.

The only datasets that one has access to during the development of the EMS with the method are the $s_{M,Train}$, $s_{M,Test}$, and $s_{R,NN}$. Therefore, one can only give considerations and perform calculations regarding reliability and performance for the network on those datasets. It is not possible to fully characterise the dataset shift to P^* or $s^*_{R,SLC}$. The characteristics of the three datasets around the method are described in Table 5.1.

Table 5.1 *Description of the datasets.*

	Datasets		
	Training	**Testing**	**Application**
Symbol	$s_{M,Train}$	$s_{M,Test}$	$s_{R,NN}$
Domain	Model	Model	Real Machine
Target	From Optim.	From Optim.	No Target
Ground Truth	-	-	-
Features	Meas. + Model	Meas. + Model	Meas. only
Trained for it?	Yes	No	No

This study is limited to compare $s_{M,Train}$ and $s_{R,NN}$ since the analysis of $s_{M,Test}$ would give more insights into the training performance of the network instead of its reliability and performance on the application.

Based on figures 5.3 and 5.4 this chapter concerns the comparison of the network performance as the output of stage c as the EMS (dataset $s_{M,Train}$), with the results obtained from its operation in the application at the end of stage d (dataset $s_{R,NN}$). It also presents an estimation of the performance of the network in the application using

$$\bar{E}(s_{R,NN}) = \bar{E}(s_{M,Train})n(s_{R,NN}). \tag{5.13}$$

The two datasets in focus can be graphically represented on their own and for two generic feature variables (x_1, x_2), as shown in Figure 5.5. The region defined by the sample $s_{R,NN}$ is called application domain (AD), whereas, the region defined by the sample $s_{M,Train}$ is called training domain (TD).

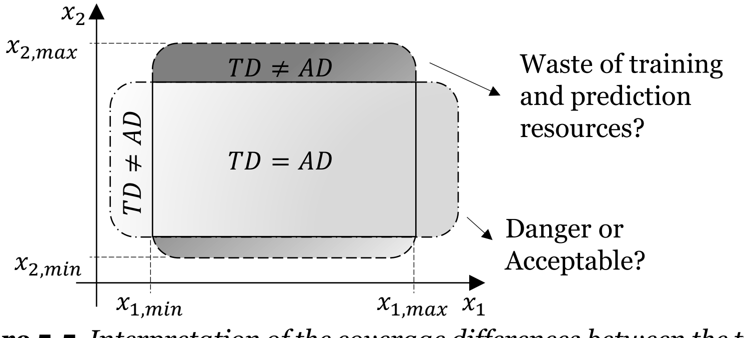

Figure 5.5 *Interpretation of the coverage differences between the training domain/dataset and application domain/dataset.*

Figure 5.5 shows what to expect when assessing reliability and performance by comparing the two datasets as described. There can be a mismatch in the coverage, both in range and probability of occurrence. The difference in range can lead to an area trained for that does not occur in the application; or, to an area not trained for that does occur in the application. The probability of occurrence concerning the two datasets will show if a certain combination of variables is more likely to occur in one over the other (this can be evaluated for the intersecting area).

It must be noticed that the training dataset is composed by many more cycles than the application dataset. The application tests with the implemented EMS were carried to characterise the operation of the controller and not to evaluate its reliability. Since extensive tests were not performed, one can only obtain an estimation of the reliability of the EMS in the application. Furthermore, only one operator drove the machine in the application. Thus, the collected results from the application are also just a sample from that domain. Nevertheless, it does allow the estimation of how the EMS performed in the application, which is supported by the machine performance results.

5.5 Results

The reliability assessment is based on the estimation of dataset shift, while the performance assessment is based on the estimation of the prediction performance based on the training accuracy.

Reliability Assessment

Figure 5.6 shows, for the training and application datasets, the plots for the probability of occurrence and coverage for the multi-dimensional space defined by the pair of input features engine speed and transmission output speed, which is proportional to the machine speed.

It must be remembered that figures 5.6a and 5.6b represent the input feature space for the situations where the output from the network was allowed to reach the system. The whole input feature space would likely have a wider range than the one shown in the figures, but assessing it would be irrelevant.

Figure 5.7 shows, for the training and application datasets, the plots for probability of occurrence and space coverage for the space defined by the pair of input features accelerator pedal and state of charge.

Figures 5.6f and 5.7f show the difference between the probability of occurrence (Δp). They indicate where there is a lack or excess of training data in comparison to the application. This information could guide the selection of representative drive cycles to compose the training dataset. They represent the dataset shift in terms of difference in the distribution.

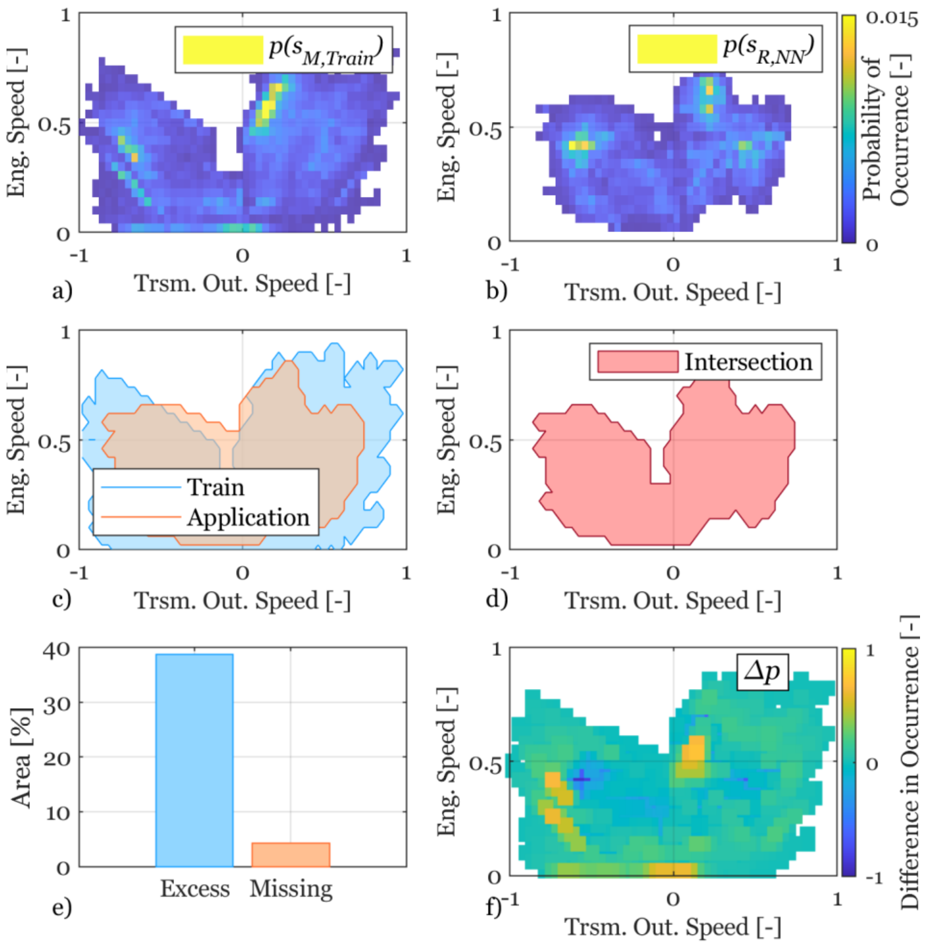

Figure 5.6 *Training dataset vs. application dataset for: Transmission Output Speed and Engine Speed. a) Training dataset probability of occurrence; b) Application dataset probability of occurrence; c) Space coverage comparison; d) Intersection area; e) Excess and missing coverage areas; f) Difference between datasets in probability of occurrence.*

Figures 5.6c and 5.7c show the difference in coverage of the two datasets where the excess area covered by the training dataset and possible missing areas not covered by the training dataset are seen. Figures 5.6d and 5.7d show the overlapping area between the two datasets, which is used to calculate the excess and missing areas. The excess area and the missing area are quantified and displayed in figures 5.6e and 5.7e. They represent the dataset shift in terms of the range of coverage.

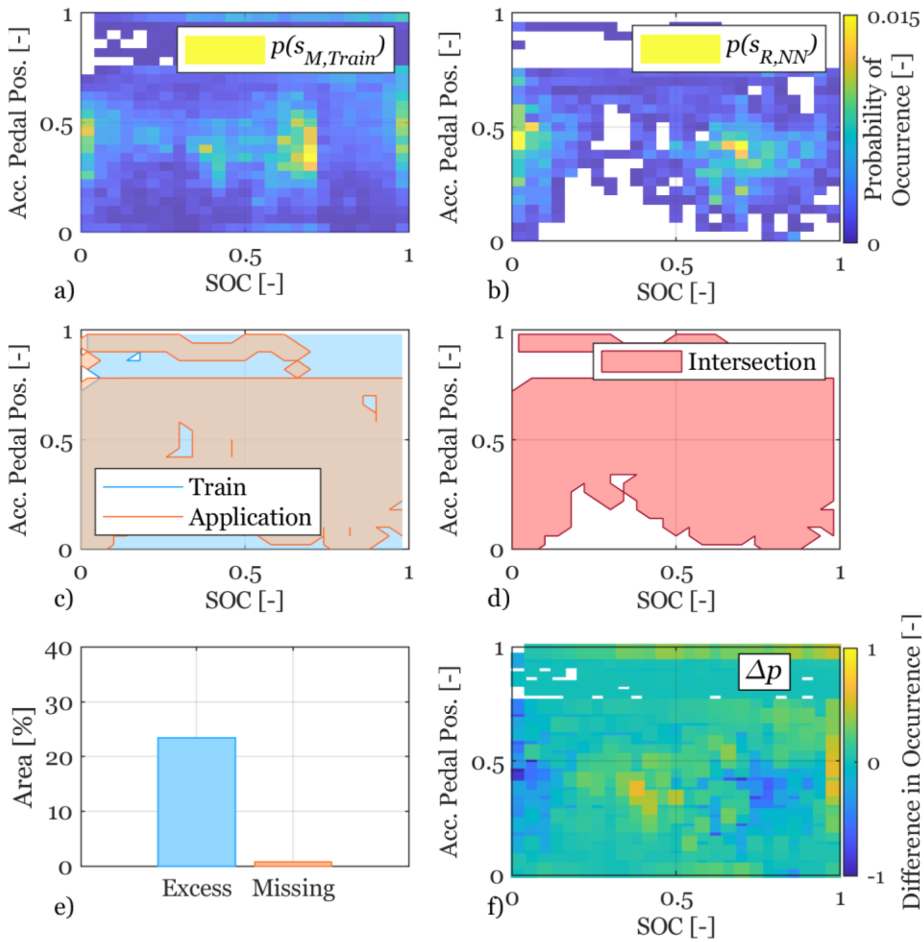

Figure 5.7 *Training dataset vs. application data set for: State of Charge and Accelerator Pedal Position. a) Training dataset probability of occurrence; b) Application dataset probability of occurrence; c) Space coverage comparison; d) Intersection area; e) Excess and missing coverage areas; f) Difference between datasets in probability of occurrence.*

Performance Assessment

To evaluate the performance of the EMS on its training domain the average absolute error ($\bar{E}$) is calculated for each pair of input features, by splitting the space in small bins; thus the input feature space is discretised. It must be remembered that the average absolute error can only be calculated for the training dataset, and it must be estimated for the application dataset according to Equation (5.13).

By plotting this average absolute error for the training dataset on top of the probability of occurrence for the training dataset allows one to see where, in the input feature space, the network has worst accuracy. Most likely, the regions of high probability of occurrence should see a small average error because the network had more instances to train for in comparison to regions with low probability of occurrence.

Plotting this average absolute error for the training data set on top of the probability of occurrence for the application dataset allows one to see where, in the input feature space, the network operated. By inference from training to application, one can see if the network operated in the application in a region of supposedly higher accuracy or in a region of lower accuracy, at least with respect to training.

Figures 5.8 and 5.9 present this kind of plot for the same two pairs of network input features used in figures 5.6 and 5.7. In the plots, the rings represent the average absolute error, where a larger ring means a larger error. Results are normalised with respect to the maximum average absolute error.

As expected for the training dataset, the higher average absolute error coincides with the regions of lower probability of occurrence, figures 5.8a and 5.9a. The opposite is also true; the lower average error coincides with the regions of higher probability of occurrence. This can be due to imbalanced data bias and/or because certain parts of the underlying function between inputs to target might be harder to map.

Now, assessing figures 5.8b and 5.9b, in the application the networks operated mostly in the parts of lower training error. In this sense, one could infer that it was operating in the regions of higher training accuracy and most likely made reasonable decisions, at least according to what it was trained for. However, it is difficult to evaluate the accuracy of the network in the application due to the lack of ground truth.

In figures 5.8c,d and 5.9c,d, the total error calculated with Equations (5.4) and (5.6), respectively is shown. They show, with the larger rings, where the EMS 'collected' more error. They indicate that more focus should be given to those areas in training to achieve a lower prediction error there because the multiplication of the number of points and average error is large.

Here this quantitative analysis of probability of occurrence vs. absolute average error is only shown for four of the input features, but in all other features and combinations of features, similar outcomes are observed.

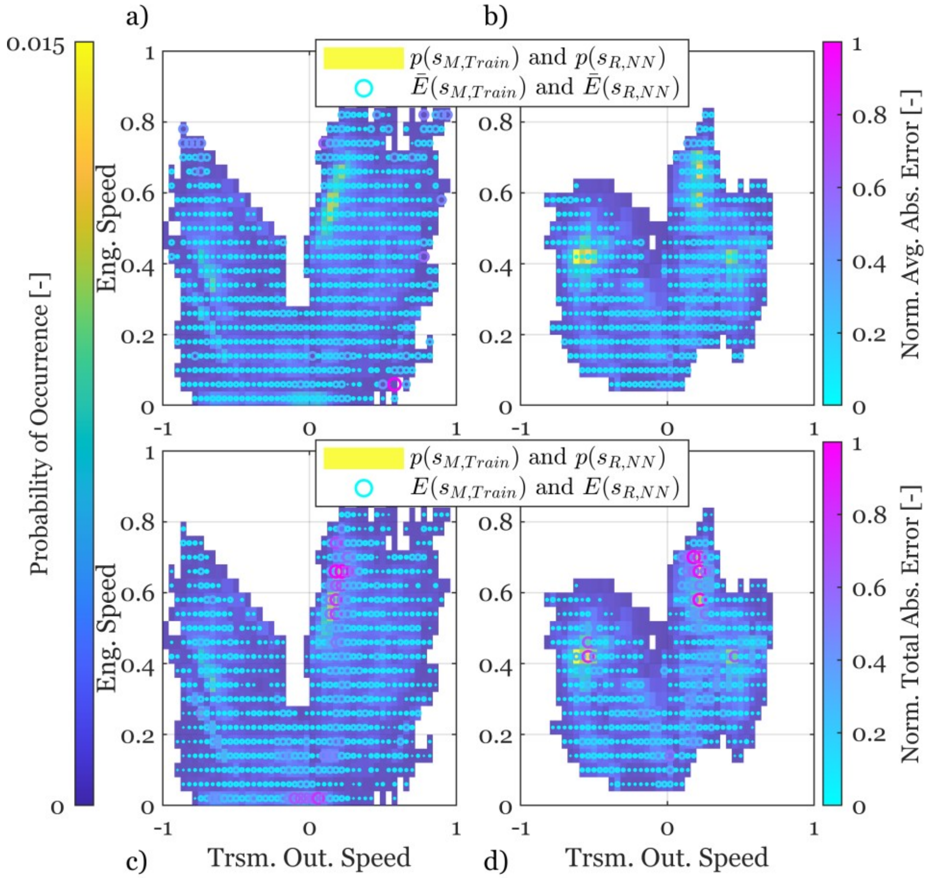

Figure 5.8 *Performance of the network in each domain for: Transmission Output Speed and Engine Speed. a) Training domain probability of occurrence and average error; b) Application domain probability of occurrence and average error; c) Training domain probability of occurrence and total error; d) Application domain probability of occurrence and total error.*

It is worth remembering two important points around the estimation of the absolute average error ($\bar{E}$). The first is that the level of confidence in them is not the same across the distribution. The calculated average error is more representative in the regions where there are more points. At the same time, it is averaging across all the remaining dimensions of the input feature space. This means the performance of the network can vary considerably even inside a bin.

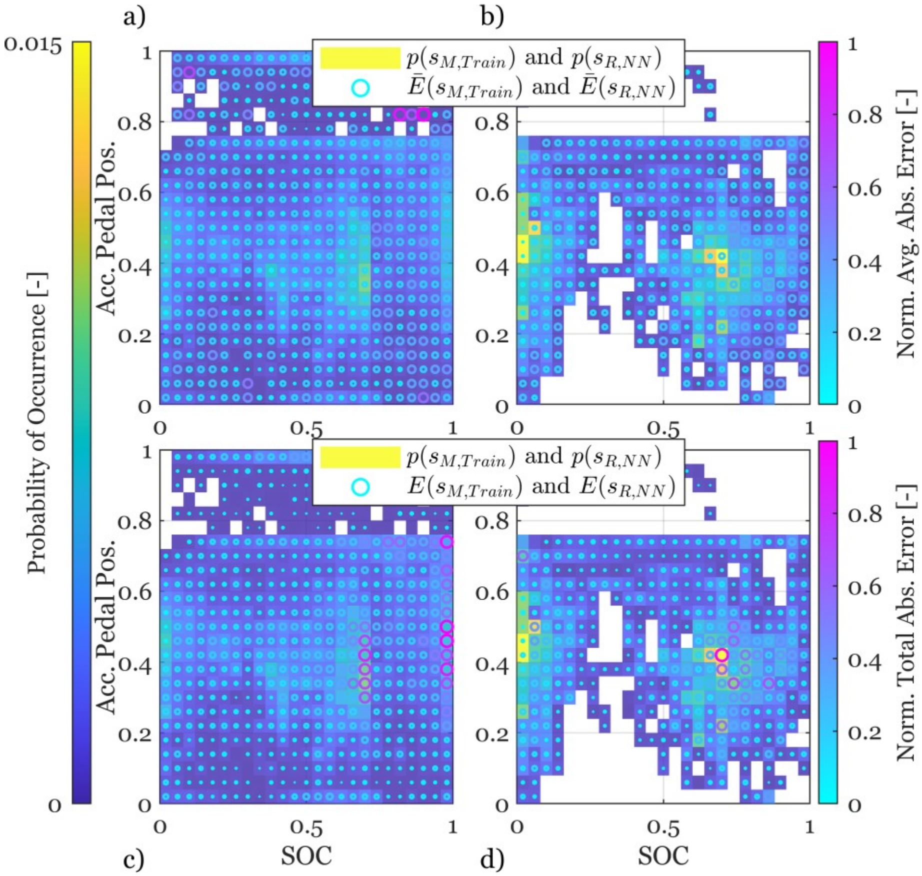

Figure 5.9 *Performance of the network in each domain for: Accelerator Pedal Position and State of Charge. a) Training domain probability of occurrence and average error; b) Application domain probability of occurrence and average error; c) Training domain probability of occurrence and total error; d) Application domain probability of occurrence and total error.*

5.6 Discussion

This chapter presented a few aspects on the assessment of the reliability of machine learning-based methods. However, there are additional topics that are worth to be discussed.

An argument for not calculating the accuracy of the network in the application through the model is that one would become similar to the accuracy of the network in the test dataset. This means that a dataset the network was not trained for and contains a dataset shift from the application. Therefore, if calculated that way, the estimated accuracy is not representative of the application.

The results demonstrate that the variability in the machine operation performance due to different operators can be significant, as indicated in [42]. Although a professional operator drove the machine during the tests, it is expected that high variability of productivity was present in the tests. However, this variability might not have been sufficient to fully characterise the application domain of the EMS, which means more operators would be necessary to conduct a thorough assessment of the reliability and performance of the EMS. However, the method presented here still is a valid initial point.

In figures 5.6e and 5.7e, it is indicated that there is a significant excess in terms of the coverage of the input features space. In [78], it is stressed that one of the negative side effects of a training space with broader coverage than the application space is that it is inefficient to fit the network outside the range of its use. This is especially true when the input dimension is large. On the other hand, the missing area is seen in the same figures. According to [78] and [104], there are techniques that indicate when a network is being used outside the range of the data for which it was trained. This will not improve the network performance, but it will prevent its use in situations where it is unreliable. In the case of being outside, one could trigger a rule-based controller to take over.

According to [78], a network trained to generalise will perform as well in new situations as it does on the data for which it was trained. For the scope of this study, this means that the generalisation of the network could be one way of reducing the effects of dataset shift. According to the author, there are at least five approaches used to obtain simple networks: growing, pruning, global searches, regularisation, and early training stopping.

In [78], it is also mentioned that it is not possible to guarantee network performance when the inputs to the network are outside the range of the training set, because they do not extrapolate well. For the present method, it may not be possible to define the active region of the input space because the operation region of these machines is vast. However, one can often collect the operation data of the machine and evaluate the coverage of the input space.

How can one reduce the EMS uncertainty caused by the dataset shift and the absence of ground truth in the application? One way is to improve the system model to gain a better representation of the machine behaviour, increase the discretisation of the state space and decrease the time step as well. This would make the solution of the optimal control problem a better estimation of the ground truth for the application. However, there is an increased computational cost in this case.

Another possibility is to increase the performance of the machine learning model itself to achieve better training accuracy and generalisation. This could be done by several means and the reader is referred to common literature in machine learning, like [75], [78], and [79].

Another possibility is to continue with the training of the network while in the application. However, the problem is how to find the optimal decisions. For that, a reinforcement learning algorithm could be applied in a similar way as suggested in [103].

Yet another possibility is to follow the recommendations of [105], [106], and [111] to account for and counteract dataset shift during the development and training of the machine learning model. The reduction of the effects of dataset shift should take the system closer to the ground truth by having a better alignment with the actual application.

Indirectly, a few procedures that were applied during the development of the EMS seem to have resulted in a reduced dataset shift. One of them was to find a better neural network structure that is capable of achieving higher prediction accuracy in the training and testing datasets. Nevertheless, the number of cycles used as inputs seemed to have achieved a reasonable coverage of the selected task. Moreover, some fundamental/mandatory rules from the rule-based architecture were included in the optimisation process as constraints, this also has an effect of reducing dataset shift.

In this section, the effects of the presence of dataset shift and absence of ground truth were collectively assessed by comparing the distributions of features of the datasets for training and application. The results show that there are clear differences in the datasets, indicating the presence of dataset shift. However, the performance results of the machine tests shown in Section 4.1 showed that these differences were not significant to the point of causing a bad performance of the EMS. Even though the optimised decisions are only an estimation of the ground truth it was still sufficient to achieve good performance in the machine. Therefore, the initial assumptions made at the start of this section were valid.

6

Review of Appended Papers

Paper I

Energy Management Based on Neural Networks for a Hydraulic Hybrid Wheel Loader

The paper presents the supervised learning method to obtain optimised energy management strategies based on dynamic programming and neural networks. Dynamic programming is used to obtain optimal offline energy management strategies for a series of drive cycles. The results are used as examples to train a neural network that implements the energy management strategy. Through simulation, the neural network's ability to learn the dynamic programming decision-making process is shown, resulting in the machine operating with fuel consumption close to that of the offline optimal energy management strategy. Aspects of modelling these machines for dynamic programming optimisation, the data necessary to train the network, the training process, variables used to learn the dynamic programming decision-making process, and the robustness of the network when facing unseen operational conditions were discussed. A limitation of the study was the use of only one base cycle that was randomised to generate similar drive cycles. Therefore, a question that remained was if the method would also be able to produce a similar performance for several recorded drive cycles, which represents a more complicated learning task closer to reality.

Paper II

Rule- and Neural Network-Based Energy Management for a Hydraulic Hybrid Wheel Loader

The paper highlighted the importance of considering required deterministic control rules from the machine control structure already in offline optimal control optimisations used to generate the data/knowledge to be learned by the neural network. The control rules are constraints to the optimal control problem. If not considered, the control optimisation results do not represent the reality and the EMS will have poor performance. In other words, to consider these types of rules increases the robustness of the network to the real application by reducing the differences between the control structure used in the development process and the control structure of the real machine. It reduces the dataset shift between the training/development domain and the application domain. Results showed that a better performance of the EMS is achieved if the rules from the application are considered in the optimal control problem. It was also shown that the implemented EMS is a combination of deterministic control rules and a neural network. The rules ensure a safe operation of the machine by reducing possible mispredictions of the network to be applied to the system during critical control actions requested by the operator.

Paper III

Performance Evaluation of Neural Network-Based Energy Management for a Hybrid Wheel Loader

This paper presented the development and evaluation of the supervised learning method to generate EMSs based on dynamic programming and neural networks. The developed EMS was implemented in the real machine for experimental evaluation in real working conditions. By comparing it to a rule-based strategy it was shown that the proposed approach can lead to, at minimum, equivalent fuel efficiency, which is a consequence of the networks' capability to implement optimised control decisions learned from dynamic programming results. It was then confirmed that simplified models are sufficient to generate a controller that works well in practice from beginning. However, issues related to adaptability and operator feeling were found to be the major goals for future development. This paper also drove the study regarding the robustness and reliability of machine learning-based methods presented in Chapter 5. Where it is shown that the dataset shift between training and application was not large, therefore not hindering the performance of the network to a significant extent. The lack of ability to continue the training after deployment to correct model and drive cycle deviations was the main driver to explore the use of reinforcement learning in a different test case.

Paper IV

Multi-Chamber Actuator Mode Selection Through Reinforcement Learning – Simulation and Experiments

The paper presented the development and implementation of a reinforcement learning agent as the mode selection controller for a multi-chamber actuator. The goal was to evaluate the capability of such a method to find and implement optimised control decisions for complex systems of construction machines. The reinforcement learning agent was trained to select the mode of the actuator to minimise system energy losses. The agent was trained in a simulated environment and afterwards deployed to the real system. As for the supervised learning case, the final controller is also a combination of network and rules to ensure safe operation. In the same way, these rules are included in the training process to minimise the dataset shift from training domain to application domain. In this study, the simulation results indicate the capability of the agent to minimise energy losses while maintaining the actuation performance. Experimental results show the capacity of the agent to perform the optimised mode selection in the real system. This study also shows that simplified, but representative models, are sufficient to learn the EMS and give the networks enough robustness to allow their direct application to control the systems. A limitation of the study was to not continue the investigation of the training after deployment to the real system.

Paper V

Design Optimisation Strategies for a Hydraulic Hybrid Wheel Loader

This paper contains the model used in papers I, II, and III. It presents a study on combined control and design optimisation. This is an important part of the conceptual design phase of construction machines, where concepts must be compared to one another while performing the same task. In this sense, it requires the design and the control to be optimised to allow an unbiased judgement between different concepts. Complexity of a system is often used as a factor when judging machine concepts. A machine concept that is complex in structure usually leads to an increased burden in the development of controllers. That was the starting point of thinking about solving this problem with machine learning methods. Therefore, to some extent, this thesis addresses this problem by evaluating methods that are capable of finding an optimised control strategy and able to implement them in practice.

Paper VI

Extended Analysis of a Valve-Controlled System with Multi-Chamber actuator

This paper presents a theoretical study on the potential of having a multi-chamber actuator in a load sensing architecture to reduce valve-controlled losses. It provided the basis for the selection of modes to be used in Paper IV. It also shows how complex the control problem for these architectures can be, which served as motivation to apply reinforcement learning to solve it. At the same time, it shows how a complex system can be simplified to make the development of controllers an easier task.

Paper VII

Boom Energy Recuperation System and Control Strategy for Hydraulic Hybrid Excavators

This paper presented the development and assessment of a rule-based EMS for the control of the boom potential energy recuperation of an excavator. In connection to this thesis, and together with the rule-based controller used as comparison in the first case study, it provided insights into how well rule-based controllers perform and how easy/difficult their implementation is. It also allowed for the understanding of a complex hydraulic system for construction machines, which is one of the motivations to evaluate the methods studied in this thesis. Another important lesson learned from this paper regards the performance assessment of construction machines. This is because the variability in the work cycles is large, which emphasises the importance of robustness in the developed control strategies.

7

Discussions and Future Studies

The studies performed along the development of this thesis showed a number of major points of concern and importance around the development and application of energy management strategies based on machine learning for construction machines.

Representativeness of the Models

At some point in the EMS development, the supervised machine learning and reinforcement learning approaches rely on models of the machines to obtain the information of how to control the systems, either from optimisations or by interaction. It was shown that sufficiently representative models were enough to obtain this information at a reasonable computational cost, i.e., they were also sufficient to present and evaluate the methods.

More accurate models should result in strategies more suitable for the real application due to their smaller deviation from the actual behaviour of the machines. However, models are always an estimation of real behaviour and there will always be deviations and compromises.

It is believed that important contributions to this topic could come from using even more representative and validated models than the ones used here. For example, the impact of different model accuracies on the resultant strategy and on the amount of continued training after deployment could be explored.

On the other hand, a balance between time to run the models and representativeness should be found. The reason is that several simulations must be performed to generate the data for the learning. The approach of this thesis was to work with simplified and less computationally heavy models. However, they were still representative enough as shown in the results of Chapter 4.

Optimisation

The methods used here are based on models; therefore the control solutions found do not represent the global optimal solution for the real machines. This leads to the presence of domain shift when applied to the real systems. It also results in absence of ground truth for the machine learning models. Therefore, it is more correct to say that the methods evaluated here resulted in optimised strategies rather than optimal ones. Future studies should focus on evaluating how close the optimised solution from these methods is from the global optimal solution.

The studies presented here focused mainly on the energy efficiency aspect of the control. However, in the control of construction machines it is highly desirable to have a smooth operation to reduce the workload of the operator while increasing its comfort. This could be extended to say that the optimisation goal is in practice never a single one but is a multi-objective problem. It is believed that, as long as they can be introduced in the optimisation generating the targets for learning, the networks should be able to learn the trade-off between the objectives. This is shown in the reinforcement learning case study, where the objective function had three terms.

Learning and Implementation of the Optimal Solution

The methods evaluated here were proven capable of finding optimised solutions and implementing them in practice. However, it is not known how close the controller implemented in practice is to the optimal solution. One problem is model deviations, and another could be the generalisation that happens during the training of the networks. Generalisation makes the learned strategy only a close prediction of the optimal one. On the other hand, the capability of machine learning to generalise to unseen cases is one of its advantages, and the actual end goal for their application.

It is believed that a trade-off between the representation of the optimal solution and generalisation can be adjusted in the machine learning model and training parameters.

There is also a deviation from the optimal caused by the necessity to use deterministic rules in the control architecture. This could be due to safety constraints for example. One could view them as constraints to the optimisation process, thus limiting the performance of the solution. At the same time, the final control structure implemented in the machine is possibly different from that used in the development. This is a source of dataset shift and a deviation from optimality since the optimal found in the development is not the one in practice.

Although important, to include all rules into the optimisation process in practice it is not an easy task, and sometimes even infeasible depending on the method. In this sense, it might be that the capability of making optimal decisions does not take place in practice all the time, although they should according to the optimisation. This can cause a significant dataset shift that affects the performance of the network. One could view the inclusion or exclusion of rules, as something similar to a reduced model representation where the model is not a true representation of reality.

Domain Shift and Reliability

The final control structures are in essence black-boxes, which does not allow, at least in a transparent way, their internal assessment to make sure they always operates in a reliable, safe, and stable manner. However, the studies of this thesis indicate that it is possible to compare the training and application domains to ensure that they operate inside the regions for which they were trained. On top of that, there must be rules, external to the networks, that always ensure a safe operation. Thus, the learned control strategy makes the system operate efficiently according to its training while rules ensure a safe operation. To some extent, this is also accompanying any type of controller. It does not necessarily result in additional development work for the machine learning-based controllers.

The difference in domain could be tackled by increasing the number of scenarios used for the learning; improved model; integration of all remaining parts of the machine controller in the optimisations; and continued learning after deployment.

This was briefly assessed in this thesis, but it remains a task for a continuation study to focus on the robustness of such methods in the real application.

Continuation of Training

Since their training is based on models of the machines and systems, it is natural to think of using this information only as a first training step to improve afterwards with actual results from machine operation.

The supervised learning approach tested here relies on results from dynamic programming, and it cannot further train in practice under the same approach due to the impossibility of applying dynamic programming in the machine. However, the same network could, in theory, be placed under a reinforcement learning approach to continue its training in practice. This would give it a degree of adaptability to improve over model imperfections or adapt to conditions it was not trained for.

Naturally, the controller trained under the reinforcement learning approach can directly be updated under the same structure while interacting/controlling the system.

The continuation of training after deployment was not explored in this thesis, but surely it would need to be accompanied by rules that result in stable convergence and safe operation. Thus, the evaluation of continued training after this initial learning from models remains for future study.

Development of EMSs Based on Machine Learning

In the end, it seems that there can be a shift in the type of activity the control engineer performs when using these methods, mostly so with regards to developing the controller or evaluating it.

Methods based on machine learning seem to shift the time to evaluation rather than development because the training of the networks is performed by the algorithms and that is where knowledge is extracted from data and built into the networks. This cannot, in principle, be affected by the engineer after the training by tweaking/tuning individual network parameters.

Although assessed for relatively less-complex systems, for more-complex systems a reduced control development time in comparison to rule-based methods is expected. This is mostly because of the methods' capability to find and learn the strategies automatically.

The two EMSs derived in the thesis cannot be easily tuned or adjusted after training to fix an unintended control behaviour. For example, it would be required to remove/fix that problem in the dataset and retrain it, or to implement the continued training after deployed, as discussed before. The tunning and adjustment of such EMSs after training should be a topic for further research.

One comparison is made between the different EMSs concerning the time required for development. Some might take a considerable amount of development time to be formulated, while others are more straightforward and demand less knowledge from the control engineer.

This thesis used simple but sufficient machine learning models to learn and apply control strategies in practice. No large emphasis was placed on comparing different network architectures. Therefore, there is the potential for further performance increase by using other machine learning models, which remains a future study.

Independence to System Concept

The methods evaluated in this thesis are concept independent. This means they can be implemented to, in principle, any type of system to perform energy management. This makes them also relevant for the trend in electrification of mobile machines where new machine architectures with a greater focus on efficiency are seen. Due to the lack of energy density of current batteries, for plug-in battery electric machines, it became a necessity to improve the machine efficiency to reduce battery pack size and reduce operation downtime due to charging.

Although the study cases were applied for hydraulic systems, it is believed that the same techniques could be applied to other domains such as mechanical and electrical ones.

Interaction with the Machine Operator

Despite the trend in the development of autonomous and automated functions, also called operator assist functions, most construction machines are still driven by an operator. According to [11], throughout the working life of a wheel loader the operator is the main influencer on efficiency and productivity. The same view is supported by [44] when the author mentions that operator decisions about control of the engine and vehicle speed, load lifting, and steering have direct effects on fuel consumption and drive cycle duration. In an empirical study, [42] shows that in a drive cycle, differences due to operator behaviour can be as much as 150% with respect to fuel efficiency and 300% regarding productivity among experienced operators.

The optimal operation is commonly generated during machine concept development to access its theoretical efficiency [11; 44; 114]. However, according to [62], the influence of the human operator is an aspect that is traditionally neglected in dynamic simulations. The author says that a "human element", introduced into dynamic simulations of working machines, provide more relevant answers with respect to operator-influenced complete-machine properties such as productivity and energy efficiency.

In the first case study, the operator reported the 'non-transparent' operation of the neural network controlling the sub-system. This is a potential source of performance degradation because of the increased workload due to the unexpected control action. This problem was not considered in the development of the controllers. However, operator and EMS interaction aspects should also be considered when developing optimised EMSs based on machine learning, and if possible, such aspects should be considered already in the modelling/optimisation phases.

Comparison to Rule-Based Strategies

The main difference between the RB-EMS and an NN-EMS is in how the knowledge is obtained and constructed. In the case of a neural network, the knowledge is constructed in the weights, biases, and activation functions. In the rule-based approach, the knowledge is in the heuristic rules. In the neural networks, the knowledge is obtained directly while training on the optimisation results, while in the rule-based approach it comes from the engineer. According to [76], the parameters learned by the neural network are difficult to be interpreted. Learned parameters of a neural network are not as easily communicated as deterministic rules.

As known since the beginning of this thesis, rule-based methods are preferred because of their transparency, predictability, and simplicity. However, the simplicity aspect is affected the more complex the systems become, which also contributes to the lower optimality of such methods. This could be one point where methods based on machine learning have an advantage over rule-based ones, for example, when the optimal control solutions are not known and/or too difficult to be implemented by hand. In such cases, the capacity of machine learning methods to find and implement strategies in the machines might have a significant advantage. However, the issues around safety and robustness must be studied in depth before their large-scale deployment to machines.

8

Conclusions

This thesis evaluated the use of machine learning as the means of learning and implementing control strategies for energy management in construction machines. Two methods were used: One based on a supervised learning and another one based on reinforcement learning.

Both methods were evaluated from the development to the application in experiments. Topics related to robustness, reliability, and performance were addressed. The two study cases presented in this thesis provided information to answer each research question:

RQ1. How can machine learning-based energy management strategies be obtained for construction machines?

This question is addressed in papers I, III and IV. Neural networks, trained under a supervised learning approach or a reinforcement learning approach, give the ability to automatically learn optimised control strategies and implement them as the energy management strategy in the machines. However, since the initial learning is based on models, there might be the necessity for a continuation of the learning process after deployment to adapt to the real system.

RQ2. What efficiency improvements can be expected from construction machines when operating with machine learning-based energy management strategies?

This question is addressed in papers I, III, and IV. The results of this thesis showed that control strategies based on machine learning do not seem to have a limitation on the learning and implementation of optimised control strategies. Therefore, these types of controllers have the potential to make the machine operate closer to their maximum efficiency than rule-based controllers.

RQ3. Can supervised learning and reinforcement learning-based methods, using neural networks as function representation, overcome the challenges related to system architecture and operation of construction machines?

This question is addressed in papers I, II, III and IV. The energy management strategies developed in this thesis were proven to be robust to the operation of construction machines when operating in conditions similar to those they were trained for. It means they are robust to model and control structure differences and still able to operate in a slightly different operation environment. However, the need for reducing these differences to increase their robustness is identified. It was also shown that they can learn the complex relationships between system variables and control decisions for such systems.

RQ4. What advantages and drawbacks can be expected from machine learning-based methods for energy management in construction machines?

Papers I, II, III, IV address this question. Adaptability and safety are the main points of concern for the applicability of this type of method to generate energy management for construction machines. The strategies must be accompanied by rules to ensure safety across all operation scenarios. On the other hand, they are able to automatically find control strategies for construction machines and implement them in the machines with a considerable level of robustness to the differences between the development domain and application domain.

References

[1] K. Baer, “Simulation-based optimization of a series hydraulic hybrid vehicle,” Ph.D. dissertation, Dept. of Man. and Eng., Linköping Univ., Linköping, Sweden, 2018. ISBN 978-91-7685-344-3.

[2] R. Filla, “Hybrid power systems for construction machinery: Aspects of system design and operability of wheel loaders,” in *Proc. of ASME IMECE* 2009, vol. 13, Nov. 13-19, 2009, pp 611-620, doi: 10.1115/IMECE2009-10458.

[3] A. Somà, “Trends and hybridization factor for heavy-duty working vehicles,” in *Intech*, 2017, doi: 10.5772/intechopen.68296.

[4] I. Naskoudakis and K. Petroutsatou, “A thematic review of the main research on construction equipment over recent years,” in *Periodica Polytechnica Architecture,* vol. 47, no. 2, 2016, pp. 110-118, doi: 10.3311/PPar.10384.

[5] J. Wang, Z. Yang, S. Liu, Q. Zhang, and Y. Han, “A comprehensive overview of hybrid construction machinery,” in *Advances in Mech. Eng.*, vol. 8, issue 3, Mar. 2016, doi: 10.1177/1687814016636809.

[6] T. Lin, Q. Wang, B. Hu, and W. Gong, “Development of hybrid powered hydraulic construction machinery,” in *Automat. in Constr.*, vol. 19, issue 1, Jan. 2010, pp. 11-19, doi: 10.1016/j.autcon.2009.09.005.

[7] X. He, and Y. Jiang, “Review of hybrid electric systems for construction machinery,” in *Automat. in Constr.*, vol. 92, Aug. 2018, pp. 286-296, doi: 10.1016/j.autcon.2018.04.005.

[8] A. Lajunen, P. Sainio, L. Laurila, J. Pippuri-Mäkeläinen, and K. Tammi, “Overview of powertrain electrification and future scenarios for non-road mobile machinery,” in *Energies*, vol. 11, no. 5, May 2018, doi: 10.3390/en11051184.

[9] M. Sprengel, and M. Ivantysynova, “Neural network based power management of hydraulic hybrid vehicles,” in *Int. J. of Fluid Power*, vol. 18, no. 2, 2017, pp. 79-91 doi: 10.1080/14399776.2016.1232117.

[10] K. Uebel, H. Raduenz, P. Krus, and V. J. De Negri, “Design optimisation strategies for a hydraulic hybrid wheel loader,” in *ASME/BATH 2018 Symp. on Fluid Power and Motion Control*, Bath, UK, Sept. 12-14, 2018, doi: 10.115/FPMC2018-8802.

[11] B. Frank, “On optimal control for concept evaluation and system development in construction machines,” Ph.D. dissertation, Fac. of

Eng., Lund Univ., Lund, Sweden, 2018. ISBN 978-91-88934-91-8.

[12] Q. Wen, F. Wang, B. Xu, and Z. Sun, "Improving the fuel efficiency of compact wheel loader with a series hydraulic hybrid powertrain," in *IEEE Trans. Veh. Technol.*, vol. 69, no. 10, Oct. 2020, pp. 10700-10709, doi: 10.1109/TVT.2020.3006155.

[13] Q. Zhang, F. Wang, B. Xu, and Z. Sun, "Online optimization of Pontryagin's minimum principle for a series hydraulic hybrid wheel loader," in *Proc. IMechE Part D: J. Automobile Eng.*, vol. 236, issue 7, Aug. 2021, pp. 1487-1499, doi: 10.1177/09544070211039819.

[14] B. Frank, "Using optimal control in concept evaluation and system optimization of diesel-electric hybrid construction machines," in *4th Int. Conf. on Elect. Syst. for Aircraft, Railway, Ship propulsion and Road Vehicles & Int. Transp. Electrific. Conf. (ESARS-ITEC)*, Toulouse, France, Nov. 2016, doi: 10.1109/ESARS-ITEC.2016.7841323.

[15] L. Serrao, G. Ornella, L. Balboni, C. M. G. Bort, C. Dousy, and F. Zendri, "A telehandler vehicle as mobile laboratory for hydraulic-hybrid powertrain technology development," in *10th Int. Fluid Power Conf.,* Dresden, Germany, Mar. 8-10, 2016, pp. 413-424.

[16] P. A. J. Achten, "A serial hydraulic hybrid drive train for off-road vehicles," in *IFPE 2008 Tech. Conf.*, Las Vegas, NV, USA, Mar. 12-14, 2008, pp 515-521.

[17] E. Lindzus, "HRB – Hydrostatic Regenerative Braking System: the hydraulic hybrid drive from Bosch Rexroth," Bosch Rexroth AG, 2010.

[18] S. Hui, and J. Junqing, "Research on the system configuration and energy control strategy for parallel hydraulic hybrid loader," in *Automat. in Constr.*, vol. 19, issue 2, Mar. 2010, pp. 213-220, doi: 10.1016/j.autcon.2009.10.006.

[19] X. Zeng, N. Yang, Y. Peng, Y. Zhang, and J. Wang, "Research on energy saving control strategy of parallel hybrid loader," in *Automat. in Constr.*, vol. 38, Mar. 2014, pp. 100–108, doi: 10.1016/j.autcon.2013.11.007.

[20] M. Erkkilä, F. Bauer, and D. Feld, "Universal energy storage and recovery system – a novel approach for hydraulic hybrid," in *The 13th Scand. Int. Conf. on Fluid Power,* Linköping, Sweden, Jun. 3-5, 2013.

[21] F. Wang, M. A. M. Zulkefli, Z. Sun, and K. A. Stelson, "Energy management strategy for a power-split hydraulic hybrid wheel loader,"

in *Proc. IMechE, Part D: J. Automobile* Eng., vol 230, issue 8, Sep. 2015, pp. 1105-1120, doi: 10.1177/0954407015600899.

[22] M. Schneider, O. Koch, and J. Weber, "Green Wheel Loader – Improving fuel economy through energy efficient drive and control concepts," in *10th Int. Fluid Power Conf.,* Dresden, Germany, Mar. 8-10, 2016, pp. 63-78.

[23] P. A. J. Achten, G. E. M. Vael, and K. Heybroek, "Efficient hydraulic pumps, motors and transformers for hydraulic hybrid systems in mobile machinery," in *Proc. of Wissensforum VDI*, Freidrichshafen, Germany, 2011.

[24] K. Heybroek, G. E. M. Vael, and J.-O. Palmberg, "Towards resistance-free hydraulics in construction machinery," in *8th Int. Fluid Power Conf.,* Dresden, Germany, Mar. 26-26, 2012.

[25] A. Hugo, K. Pettersson, K. Heybroek, and P. Krus, "Modelling and control of a complementary energy recuperation system for mobile working machines," in *The 13th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, June 3-5, 2013.

[26] K. Pettersson, K. Heybroek, A. Klintemyr, and P. Krus, "Analysis and control of complementary energy recuperation system," in *8th Int. Fluid Power Conf.,* Dresden, Germany. Mar. 26-28, 2012.

[27] M. Heikkilä, M. Huova, J. Tammisto, M. Linjama, and J. Tervonen, "Fuel efficiency optimization of a baseline wheel loader and its hydraulic hybrid variants using dynamic programming," in *ASME/BATH 2018 Symp. on Fluid Power and Motion Control*, Bath, UK, Sept. 12-14, 2018, doi: 10.1115/FPMC2018-8853.

[28] T. Lin, Q. Wang, B. Hu, and W. Gong, "Research on the energy regeneration systems for hybrid hydraulic excavators," in *Automat. in Constr.*, vol. 19, issue 8, Dec. 2010, pp. 1016-1026, doi: 10.1016/j.autcon.2010.08.002.

[29] P. Casoli, A Gambarotta, N. Pompini, and L. Riccò, "Hybridization methodology based on DP algorithm for hydraulic mobile machinery – Application to a middle size excavator," in *Automat. in Constr.*, vol. 61, pp. 42-57, Jan. 2016, doi: 10.1016/j.autcon.2015.09.012.

[30] T. Jung, H. Raduenz, P. Krus, V. J. De Negri, and J. Lee, "Boom energy recuperation system and control strategy for hydraulic hybrid excavators," in: *Automat. in Constr.*, vol. 135, Mar. 2022, doi: 10.1016/j.autcon.2021.104046.

[31] G. Jun, Z. Daqing, L. Changsheng, Z. Yuming, H. Peng, and Q. Weicai, "Optimization of electro-hydraulic energy savings in

mobile machinery," in *Automat. in Constr.*, vol. 98, Feb. 2019, pp. 132-145, doi: 10.1016/j.autcon.2018.08.011.

[32] H. Murrenhoff, S. Sgro, and M. Vukovic, "An overview of energy saving architectures for mobile applications," in *9th Int. Fluid Power Conf.* Aachen, Germany, Mar. 24-26, 2014.

[33] M. Vukovic, R. Leifeld, and H. Murrenhoff, "Reducing fuel consumption in hydraulic excavators – A comprehensive analysis," in *Energies*, vol. 10, no. 5, May. 2017, doi:10.3390/en10050687.

[34] K. Pettersson, K. Heybroek, P. Mattsson, and P. Krus, "A novel hydromechanical hybrid motion system for construction machines," in *Int. J. of Fluid Power*, vol. 18, no. 1, 2016, pp. 17-28, doi: 10.1080/14399776.2016.1210423.

[35] K. Heybroek and M. Sahlman, "A hydraulic hybrid excavator based on multi-chamber cylinders and secondary control – design and experimental validation," in *Int. J. of Fluid Power*, 2018, pp. 91-105, doi: 10.1080/14399776.2018.1447065.

[36] M Linjama, H-P Vihtanen, A Sipola, and M Vilenius, "Secondary controlled multi-chamber hydraulic cylinder," in *The 11th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, Jun. 2-4, 2009.

[37] A. Dell'Amico, M. Carlsson, E. Norlin, and M. Sethson, "Investigation of a digital hydraulic actuation system on an excavator arm," in *The 13th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, Jun. 3-5, 2013, pp. 505-511, doi: 10.3384/ecp1392a50.

[38] M. Vukovic, R. Leifeld, and H. Murrenhoff, "STEAM – a hydraulic hybrid architecture for excavators," in *10th Int. Fluid Power Conf.*, Dresden, Germany, Mar. 8-10, 2016.

[39] M. Ketonen and M. Linjama, "Simulation study of a digital hydraulic independent metering valve system on an excavator," in *The 15th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, Jun. 7-9, 2017.

[40] E. Busquets and M. Ivantysynova, "A multi-actuator displacement-controlled system with pump switching – A study of the architecture and actuator-level control," in *Int. J. of Fluid Power System*, vol. 8, issue 2, 2015, pp. 66-75, doi: 10.5739/jfpsij.8.66.

[41] P. Casoli, F. Scolari, C. M. Vescovini, and M. Rundo, "Energy comparison between a load sensing system and electro-hydraulic solutions applied to a 9-ton excavator," in *Energies*, vol. 15, no. 7, 2022, doi: 10.3390/en15072583.

[42] B. Frank, L. Skogh, and M. Alaküla, "On wheel loader fuel

efficiency difference due to operator behaviour distribution," in *2nd Commercial Vehicle Technol. Symp.*, Kaiserslautern, Germany, Mar. 13-15, 2012, pp.329-346.

[43] B. Frank, J. Kleinert, and R. Filla, "Optimal control of wheel loader actuators in gravel applications," in *Automat. in Constr.*, vol 91, Jul. 2018, pp. 1-14, doi: 10.1016/j.autcon.2018.03.005.

[44] V. Nezhadali, B. Frank, and L. Eriksson, "Wheel loader operation – optimal control compared to real drive experience," in *Control Eng. Practice*, vol. 48, Mar. 2016, pp. 1-9, doi: 10.1016/j.conengprac.2015.12.015.

[45] Z. Filipi *et al.*, "Combined optimisation of design and power management of the hydraulic hybrid propulsion system for the 6x6 medium truck," in *Int. J. of Heavy Vehicle Syst.*, vol. 11, nos. 3/4, 2004, pp. 372-402.

[46] K. Baer, L. Ericson, and P. Krus, "Framework for simulation-based simultaneous system optimization for a series hydraulic hybrid vehicle," in *Int. J. of Fluid Power*, 2018, doi: 10.1080/14399776.2018.1527122.

[47] W. Zhang, J. Wang, S. Du, H. Ma, W. Zhao, and H. Li, "Energy management strategies for hybrid construction machinery: Evolution, classification, comparison and future trends," in *Energies,* vol. 12, no. 10, 2019, doi: 10.3390/en12102024.

[48] V. Krithika and C. Subramani, "A comprehensive review on choice of hybrid vehicles and power converters, control strategies for hybrid electric vehicles," in *Int. J. of Energy Res.*, vol. 42, issue 5, Apr. 2018, doi: 10.1002/er.3952.

[49] A. M. Ali and D. Söffker, "Towards optimal power management of hybrid electric vehicles in real-time: A review on methods, challenges, and state-of-the-art solutions," in *Energies*, vol. 11, no. 3, Feb. 2018, doi: 10.3390/en11030476.

[50] W. Enang and C. Bannister, "Modelling and control of hybrid electric vehicles (A comprehensive review)," in *Renewable and Sustain. Energy Reviews,* vol. 74, Jul. 2017, pp. 1210-1239, doi: 10.1016/j.rser.2017.01.075.

[51] M. A. Karbaschian and D. Söffker, "Review and comparison of power management approaches for hybrid vehicles with focus on hydraulic drives," in *Energies*, vol. 7, no. 6, 2014, pp. 3512-3536; doi: 10.3390/en7063512.

[52] R. Bellman, *Dynamic Programming*, USA: Princeton Univ. Press. Princeton, 1957. ISBN 0-691-07951-X.

[53] D. P. Bertsekas, *Dynamic Programming and Optimal Control - Vol I*, 3rd ed. Delmont MA, USA: Athena Scientific, 2005. ISBN 1-886529-26-4

[54] O. Sundström, and L. Guzzella, "A generic dynamic programming Matlab function," in *2009 IEEE Control App., (CCA) & Intell. Control (ISIC)*, St. Petersburg, Russia, Jul. 8-10, 2009, pp. 1625–1630, doi: 10.1109/CCA.2009.5281131.

[55] F. Wang, M. A. M. Zulkefli, Z. Sun, and K. A. Stelson, "Investigation on the energy management strategy for hydraulic hybrid wheel loaders," in *Proc. ASME 2013 Dynamic Syst. and Control Conf.*, Palo Alto, California, USA, Oct. 21-23, 2013, pp. 1-10, doi: 10.1115/DSCC2013-3949.

[56] H. T. Do, "Energy management of parallel hydraulic hybrid wheel loader with focus on fuel consumption minimization," M.S. thesis, Sch. of El. Eng. and Comp. Sci., KTH, Stockholm, Sweden, 2018.

[57] K. A. Stelson, J. J. Meyer, A. G. Alleyne, and B. Hencey, "Optimization of a passenger hydraulic hybrid vehicle to improve fuel economy," in *Proc. 7th JFPS Int. Symp. on Fluid Power*, Toyama, Sep. 15-18, 2008, pp. 143-148, doi: 10.5739/isfp.2008.143.

[58] H. Lee, C. Kang, Y. Park, and S. W. Char, "A study on power management strategy of HEV using dynamic programming," *World Electr. Veh. J.*, vol. 8, no. 1, Mar. 2016, pp. 274-280, doi: 10.3390/wevj8010274.

[59] T. Lin, W. Huang, H. Ren, S. Fu, and Q. Liu, "New compound energy regeneration system and control strategy for hybrid hydraulic excavators," in *Automat. in Constr.*, vol. 68, Aug. 2016, pp. 11-20, doi: 10.1016/j.autcon.2016.03.016.

[60] G. Jun, Z. Daqing, G. Yong, L. Changsheng, Z. Yuming, H. Peng, and Q. Weicai, "Power control strategy and performance evaluation of a novel electro-hydraulic energy-saving system," in *App. Energy*, vol. 233-234, Jan. 2019, pp. 724-734, doi: 10.1016/j.apenergy.2018.10.066.

[61] W. Shen, J. Jiang, X. Su, and H. R. Karimi, "Control strategy analysis of the hydraulic hybrid excavator," in *J. Franklin Inst.*, vol. 352, issue 2, Feb. 2015, pp. 541-561, doi: 10.1016/j.jfranklin.2014.04.007.

[62] D. Wang, X. Lin, and Y. Zhang, "Fuzzy logic control for a parallel hybrid hydraulic excavator using genetic algorithm," in *Automat. in Constr.*, vol 20, issue 5, Aug. 2011, pp. 581-587, doi: 10.1016/j.autcon.2010.11.024.

[63] C. Cheng, A. McGordon, R. P. Jones, and P. A. Jennings, "Generalised fuzzy-logic-based power management strategy for various hybrid electric vehicle powertrain architectures," in *UKACC Int. Conf. on Control,* Coventry, UK, Sep. 7-10, 2010, pp. 197-202, doi: 10.1049/ic.2010.0280.

[64] S. Onori, L. Serrao, and G. Rizzoni, "Adaptive equivalent consumption minimization strategy for hybrid electric vehicles," in *Proc. ASME 2010 Dynamic Syst. and Control Conf., DSCC2010*, Cambridge, MA, USA, Sep. 12-15, 2010, pp. 499-505, doi: 10.1115/DSCC2010-4211.

[65] C. Musardo, G. Rizzoni, Y. Guezennec, and B. Staccia, "A-ECMS: An adaptive algorithm for hybrid electric vehicle energy management," in *Eur. J. of Control*, vol. 11, issues 4-5, 2005, pp. 509-524, doi: 10.3166/ejc.11.509-524.

[66] Q. Wen, F. Wang, M. Cheng, and B. Xu, "Adaptive equivalent consumption minimization strategy for off-road hydraulic hybrid vehicles: A cycle-to-cycle optimization approach," in *IEEE Trans. Veh. Technol.*, vol. 71, no. 3, Mar. 2022, pp. 2346-2357, doi: 10.1109/TVT.2021.3136658.

[67] T. O. Deppen, A. G. Alleyne, K. Stelson, and J. Meyer, "A model predictive control approach for a parallel hydraulic hybrid powertrain," in *Amer. Control Conf.,* San Francisco, CA, USA, Jun. 29 - Jul. 1, 2011, pp. 2713-2718, doi: 10.1109/ACC.2011.5991128.

[68] K. Heybroek and J. Sjöberg, "Model predictive control of a hydraulic multichamber actuator: A feasibility study," in *IEEE/ASME Trans. Mech.*, vol. 23, no. 3, Jun. 2018, pp. 1393-1403, doi: 10.1109/TMECH.2018.2823695.

[69] H. Wang, Y. Huang, A. Khajepour, H. He, and D. Cao, "A novel energy management for hybrid off-road vehicles without future cycles as a priori," in *Energy*, vol. 133, Aug. 2017, pp. 929-940, doi: 10.1016/j.energy.2017.05.172.

[70] T. Leroy, F. Vidal-Naquet, and P. Tona, "Stochastic dynamic programming based energy management of HEV's: an experimental validation," in *Proc. 19th World Congr., IFAC,* Cape Town, South Africa, Aug. 24-29, 2014, pp. 4813-4818, doi: 10.3182/20140824-6-ZA-1003.01868.

[71] C. Vagg, S. Akehurst, C. J. Brace, and L. Ash, "Stochastic dynamic programming in the real-world control of hybrid electric vehicles," in *IEEE Trans. Control Syst. Technol.*, vol. 24, no.3, May 2016, doi: 10.1109/TCST.2015.2498141.

[72] D. F. Opila, X. Wang, R. McGee, R. B. Gillespie, J. A. Cook, and J. W. Grizzle, "An energy management controller to optimally trade off fuel economy and drivability for hybrid vehicles," in *IEEE Trans. Control Syst. Technol.*, vol. 20, no.6, Nov. 2012, pp. 1490-1505, doi: 10.1109/TCST.2011.2168820.

[73] T. Nilsson, A. Fröberg, and J. Åslund, "Predictive control of a diesel electric wheel loader powertrain," in *Control Eng. Practice*, vol. 41, Aug. 2015, pp. 47-56, doi: 10.1016/j.conengprac.2015.04.008.

[74] T. M. Mitchell, *Machine Learning*: MacGraw-Hill, 1997. ISBN 978-0070428072.

[75] I. Goodfellow, Y. Bengio, and A. Courville, *Deep Learning,* Cambridge, MA, USA: The MIT Press, 2016. ISBN 9780262035613

[76] K. M. Hornik, M. Stinchcombe, and H. White, "Multilayer feedforward networks are universal approximators," in *Neural Networks*, vol.2, issue. 5, 1989, pp. 359-366, doi: 10.1016/0893-6080(89)90020-8.

[77] M. T. Hagan and H. B. Demuth, "Neural networks for control," in *Proc. Amer. Control Conf.*, San Diego, CA, USA, Jun. 1999.

[78] M. T. Hagan and H. B. Demuth, M. H. Beale, O. De Jesús, *Neural Network Design,* 2nd ed. ISBN-10: 0-9717321-1-6

[79] C. C. Aggarwal, *Neural Networks and Deep Learning – A Textbook*: Springer, 2018, ISBN 978-3-319-94462-3

[80] E. Alpaydin, *Introduction to Machine Learning*. 3rd ed. Cambridge, MA, USA: The MIT Press, 2014. ISBN 978-0-262-01243-0

[81] Y. L. Murphey, J. Park, Z. Chen, M. L. Kuang, M. A. Masrur, and A. M. Phillips, "Intelligent hybrid vehicle power control – part I: Machine learning of optimal vehicle power," in *IEEE Trans. Veh. Technol.*, vol. 61, no. 8, Oct. 2012, pp. 3519-3530, doi: 10.1109/TVT.2012.2206064.

[82] Y. L. Murphey, J. Park, L. Kiliaris, M. L. Kuang, M. A. Masrur, A. M. Phillips, and Q. Wang, "Intelligent hybrid vehicle power control – part II: Online intelligent energy management," in *IEEE Trans Veh. Technol.*, vol. 62, no. 1. Jan. 2013, pp. 69-79, doi: 10.1109/TVT.2012.2217362.

[83] C. Majed, S. H. Karaki, and R. Jabr, "Neural network technique for hybrid electric vehicle optimization," in *Proc. 18th Mediterranean Electrotechnical Conf.,* Cyprus, Apr. 18-20, 2016, doi: 10.1109/MELCON.2016.7495327.

[84] J. Park, Z. Chen, and Y. L. Murphey, "Intelligent vehicle power

management through neural learning," in *The 2010 Int. Joint Conf. on Neural Networks (JCNN),* Jul. 18-23, 2010, pp 1-7, doi: 10.1109/IJCNN.2010.5596725.

[85] Z. Chen, C. C. Mi, J. Xu, X. Gong, and C. You, "Energy management for a power-split plug-in hybrid electric vehicle based on dynamic programming and neural networks," in *IEEE Trans. Veh. Technol.*, vol. 63, no. 4, May 2014, pp. 1567-1580, doi: 10.1109/TVT.2013.2287102.

[86] C.-W. Yen, and M. L. Nagurka, "Design of predictive controllers by dynamic programming and neural networks," in *Proc. Amer. Control Conf.,* Denver, Colorado, USA, Jun. 4-6, 2003, doi: 10.1109/ACC.2003.1239766.

[87] J. Moreno, M. E. Ortúzar, and J. W. Dixon, "Energy-management system for a hybrid electric vehicle, using ultracapacitors and neural networks," in *IEEE Trans. Ind. Electronics*, vol. 53, no. 2, Apr. 2006, pp. 614-623, doi: 10.1109/TIE.2006.870880.

[88] J. Shen, and A. Khaligh, "A supervisory energy management control strategy in a battery/ultracapacitor hybrid energy storage system," in *IEEE Trans. Transp. Electrific.*, vol. 1, no. 3, Oct. 2015, pp. 223-231, doi: 10.1109/TTE.2015.2464690.

[89] G. Li, and D. Görges, "Fuel-efficient gear shift and power split strategy for parallel HEVs based on heuristic dynamic programming and neural networks," in *IEEE Trans. Veh. Technol.*, vol. 68, no. 10, Oct. 2019, pp. 9519-9528, doi: 10.1109/TVT.2019.2927751.

[90] S. Xie, S. Qi, and K. Lang, "A data-driven power management strategy for plug-in hybrid electric vehicles including optimal battery depth of discharging," in *IEEE Trans. Ind. Inform.*, vol. 16, no. 5, May 2020, pp. 3387-3396, doi: 10.1109/TII.2019.2917468.

[91] X. Tian, R. He, and Y. Xu, "Design of an energy management strategy for a parallel hybrid electric bus based on an IDP-ANFIS scheme," in *IEEE Access*, vol. 6, Apr. 2018, pp. 23806-23819. 10.1109/ACCESS.2018.2829701.

[92] C. K. D. Harold, S. Prakash, and T. Hofman, "Powertrain control for hybrid-electric vehicles using supervised machine learning," in *Vehicles*, vol. 2, no. 2, 2020, pp. 267–286, doi: 10.3390/vehicles2020015.

[93] M. R. Batli and S. S. Danaraddi, "Control of hydraulic hybrid wheel loader using machine learning," M.S. thesis, Dept. of El. Eng., Chalmers Univ. of Technol., Gothenburg, Sweden, 2020.

[94] R. S. Sutton and A. G. Barto, *Reinforcement Learning – An*

Introduction. 2nd ed. Cambridge, MA, USA: The MIT Press, 2018. ISBN 9780262039246

[95] V. Mnih, K. Kavukcuoglu, D. Silver, A. Graves, I. Antonoglou, D. Wierstra, and M. Riedmiller, "Playing Atari with deep reinforcement learning," 2013. arXiv:1312.5602v1

[96] Q. Zhu and Q. Wang, "Real-time energy management controller design for a hybrid excavator using reinforcement learning," in *J. Zheijang Univ.-SCIENCE A (Appl Phys & Eng)*, vol. 18, no. 11, 2017, pp. 855-870, doi: 10.1631/jzus.A1600650.

[97] W. Zhang, J. Wang, Y. Liu, G. Gao, S. Liang, and H. Ma, "Reinforcement learning-based intelligent energy management architecture for hybrid construction machinery," in *App. Energy*, vol. 275, Oct. 2020, doi: 10.1016/j.apenergy.2020.115401.

[98] G. Du, Y. Zou, X. Zhang, Z. Kong, J. Wu, and D. He, "Intelligent energy management for hybrid electric tracked vehicles using online reinforcement learning," in *App. Energy*, vol. 251, Oct. 2019, doi: 10.1016/j.apenergy.2019.113388.

[99] D. Berglund and N. Larsson, "Controlling a hydraulic system using reinforcement learning: Implementation and validation of a DQN-agent on a hydraulic multi-chamber cylinder system," M.S. thesis, Dept. of Man. and Eng., Linköping Univ., Linköping, Sweden, 2021.

[100] S. Backman, D. Lindmark, K. Bodin, M. Servin, J. Mörk, and H. Löfgren, "Continuous control of an underground loader using deep reinforcement learning," 2021, arXiv:2103.01283v2.

[101] J. Andersson, K. Bodin, D. Lindmark, M. Servin, and E. Wallin, "Reinforcement learning control of a forestry crane manipulator," 2021, arXiv:2103.02315v1.

[102] P. Egli and M. Hutter, "A general approach for the automation of hydraulic excavator arms using reinforcement learning," in *IEEE Robot. Automat. Lett.*, vol. 7, issue 2, Apr. 2022, pp. 5679-5686, doi: 10.1109/LRA.2022.3152865.

[103] S. Dadhich, F. Sandin, U. Bodin, U. Andersson, and T. Martinsson, "Adaptation of a wheel loader automatic bucket filling neural network using reinforcement learning," in *Int. Conf. on Neural Networks*, Glasgow, UK, Jul. 19-24, 2020, doi: 10.1109/IJCNN48605.2020.9206849.

[104] D. Amodei, C. Olah, J. Steinhardt, P. Christiano, J. Schulman, and D. Mané, "Concrete problems in AI safety," 2016, arXiv:1606.06565v2.

[105] M. Kull and P. Flach, "Patterns of dataset shift," 2014.

[106] J. Quiñonero-Candela, M. Sugiyama, A. Schwaighofer, and N. D. Lawrence, "Dataset Shift in Machine Learning," The MIT Press, 2009. ISBN 978-0-262-17005-5

[107] R. Filla, "An event-driven operator model for dynamic simulation of construction machinery," in *The 9th Scand. Int. Conf. on Fluid Power*, Linköping, Sweden, 2005, doi: 10.48550/arXiv.cs/0506033.

[108] O. Sundström, D. Ambühl, and L. Guzzella, "On implementation of dynamic programming for optimal control problems with final state constraints," in *Oil Gas Sci. Technol. – Rev. l'Institut Français du Pétrole*, vol. 65, no. 1, 2010, pp. 91–102.

[109] P. Elbert, S. Ebbesen, and L. Guzzella, "Implementation of dynamic programming for n-dimensional optimal control problems with final state constraints," in *IEEE Trans. Control Syst. Technol.*, vol. 21, no. 3, May. 2013, pp. 924-931, doi: 10.1109/TCST.2012.2190935.

[110] T. Samuelsson, R. Rilla, B. Frank, and L. Skogh, "Selecting representative working cycles from large measurement data sets," in *CVT 2016 Commercial Vehicle Technol. Symp.,* Kaiserslautern, Germany, Mar. 2016.

[111] A. J. Storkey, "When training and test sets are different: Characterizing learning transfer," in *Dataset Shift in Machine Learning*, Cambridge, MA, USA: The MIT Press, 2009, pp. 3-28. ISBN 978-0-262-17005-5

[112] DQN Agent, Mathworks Reinforcement Learning Toolbox, https://mathworks.com/help/reinforcement-learning/ug/dqn-agents.html, (accessed Apr. 1, 2022).

[113] HOPSAN multi-domain system simulation tool, Division of Fluid and Mechatronic System, Department of Management and Eng., Linköping Univ., Linköping, Sweden. https://liu.se/en/research/hopsan (accessed Apr. 1, 2022).

[114] K. Uebel, "Conceptual design of complex hydromechanical transmissions," Ph.D. dissertation, Dept. of Man. and Eng., Linköping Univ., Linköping, Sweden, 2017. ISBN 978-91-7685-447-1