Data Analytics and Optimization Modelling based **Decision Support Tool** for Canadian Arctic Oil Spills Responses

PhD Defense by *Tanmoy Das*

Dept. of Industrial Engineering, Dalhousie University

Supervisor: Dr. *Floris Goerlandt*

Co-supervisor: Dr. *Ronald Pelot*

Committee members: Dr. *Ahmed Saif*, Dr. *Hassan Sarhadi*

External examiner: Dr *Marko Perkovič*

DALHOUSIE UNIVERSITY

MARS Maritime Risk and Safety Research Group

Nov 2023

Outline of the presentation

Introduction

Background
Research Objective and Questions
Structure of thesis

Method & Results

Estimating spill
Ranking Technologies
Optimizing spill coverage and cost
Quantifying Uncertainty

Discussion

Scientific contribution
Limitation and Future work

Introduction

Introduction

- **Oil spill in Arctic**
  - The most crucial threat to Arctic environment and marine life (Arctic Council 2009, Fu et al., 2021)
  - Increased Arctic shipping in recent years (Carter 2018, Lasserre & Bartenstein, 2022)
  - Arctic challenges includes harsh weather, limited resources, remote areas

Figure 1. Number of shipping voyages in Canadian Arctic doubled in recent years

- **Decision Support Tool (DST)**
  - Integrates key elements of spill responses
  - Captures Arctic conditions
  - Leverages data analytics for better response
  - Recommends facilities in optimal locations

DST: a computer-based tool equipped with modelling and analytical capabilities

Spill Response Technologies in Arctic

- **Mechanical Containment and Recovery (MCR)**
  Removing oil from water surface by physical barriers and mechanical devices e.g. skimmers

- **Chemical Dispersant Use (CDU)**
  Spraying chemical products into oil spill to disperse it into the water, and to accelerate natural dispersion

- **In-situ Burning (ISB)**
  Controlled-burning of oil in its original place of spill

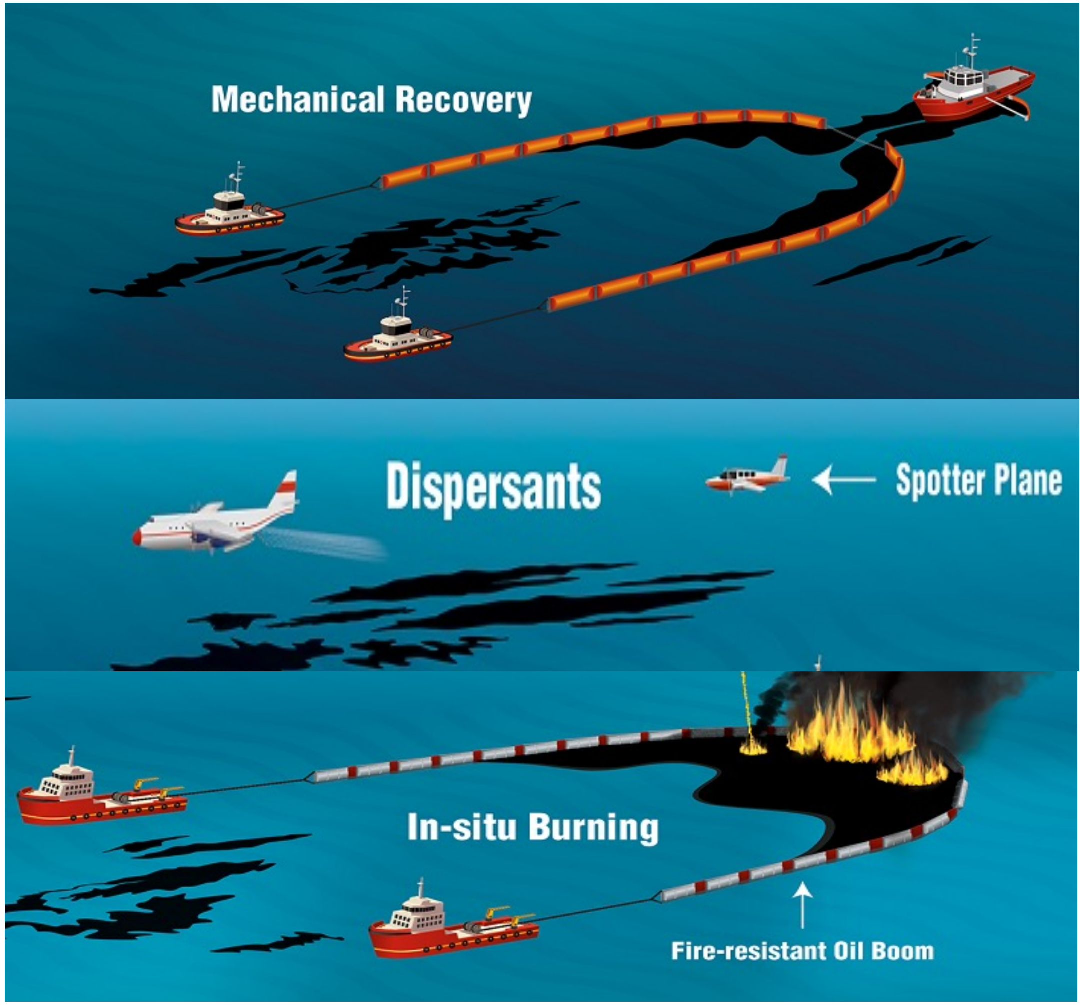


Figure 2. Oil Spill Response Technologies

Literature Review

Table 1. Literature review on DST for oil spill

Reference	Main Features	Key Findings	Opt	Str	Arctic	PPR
(Ye, 2022)	Resource allocation optimization under complex conditions	Enhance capacity of resource allocation	☒	☒	☐	☒
(Li, 2014)	Technology classification and simulation optimization	Integrated DST incorporating uncertainty in better response	☒	☒	☒	☐
(Garrett et al., 2017a)	Dynamic resource allocation in the Arctic spill	Minimum total weighted response time	☒	☐	☒	☐
(Božanić et al., 2023)	Multi criteria decision making to evaluate alternatives	Expert opinions can supplement incomplete information of onshore oil spill	☐	☒	☐	☐
(Pourvakhshouri et al., 2006)	Oil spill estimation, shoreline ranking, and resource prioritization	Priority ranking scale is designed by expert interview	☐	☒	☐	☐
(Amir-Heidari & Raie, 2019)	Spill response planning by deterministic & probabilistic model	Assessment of response drills & risk-based prioritization of coastal areas	☒	☒	☐	☒
Proposed DST (This thesis)	**Facility location and allocation for spill in Canadian Arctic**	**Improved Strategic planning of oil spill response**	☒	☒	☒	☒

Note: Opt: Optimization Model, Str: Strategic, PPR: Pollution Preparedness and Response

Research Gap

1. A fast and accurate model to predict spill size (slower engineering models are available)
2. A computationally efficient method to rank spill response technologies (subjective judgment and labor-intensive models are available)
3. Arctic facility location-allocation model considering environmental sensitivity and Arctic remoteness
4. Systematic selection of uncertain parameters and data-centric uncertainty quantification

A DST integrating individual models for estimating spill size, ranking technologies, locating facilities-allocating resources and quantifying uncertainty for harsh Arctic weather

Research Objectives (RO) and Questions (RQ)

Key RO: Develop an integrated DST for strategic oil spill response planning in Arctic

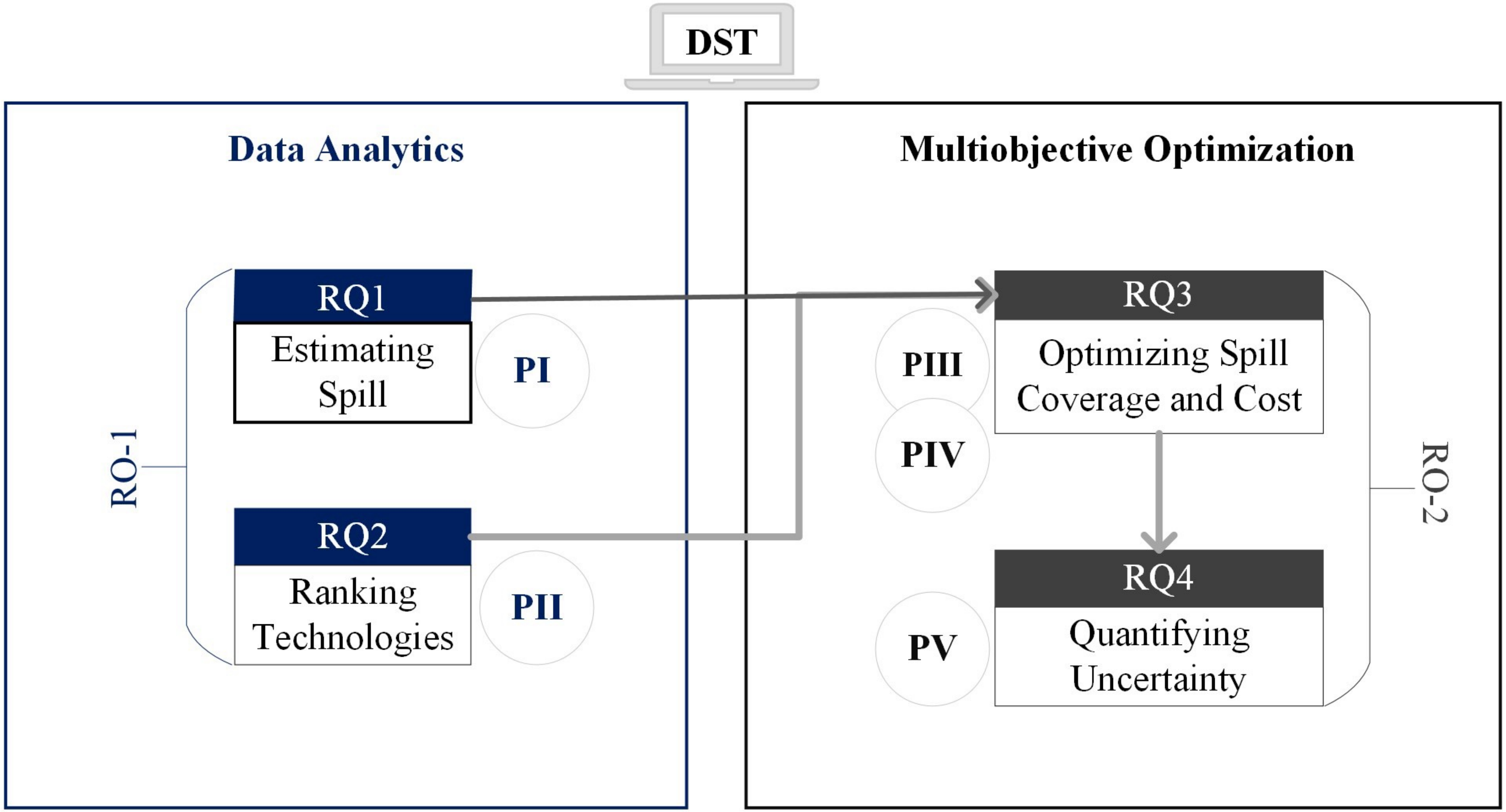
RO-1: Model oil spill problem to predict spill size and rank response technologies

- **RQ1:** What is the estimated volume of the oil spills in Arctic conditions?
- **RQ2:** What is a suitable method to accurately rank spill response technologies?

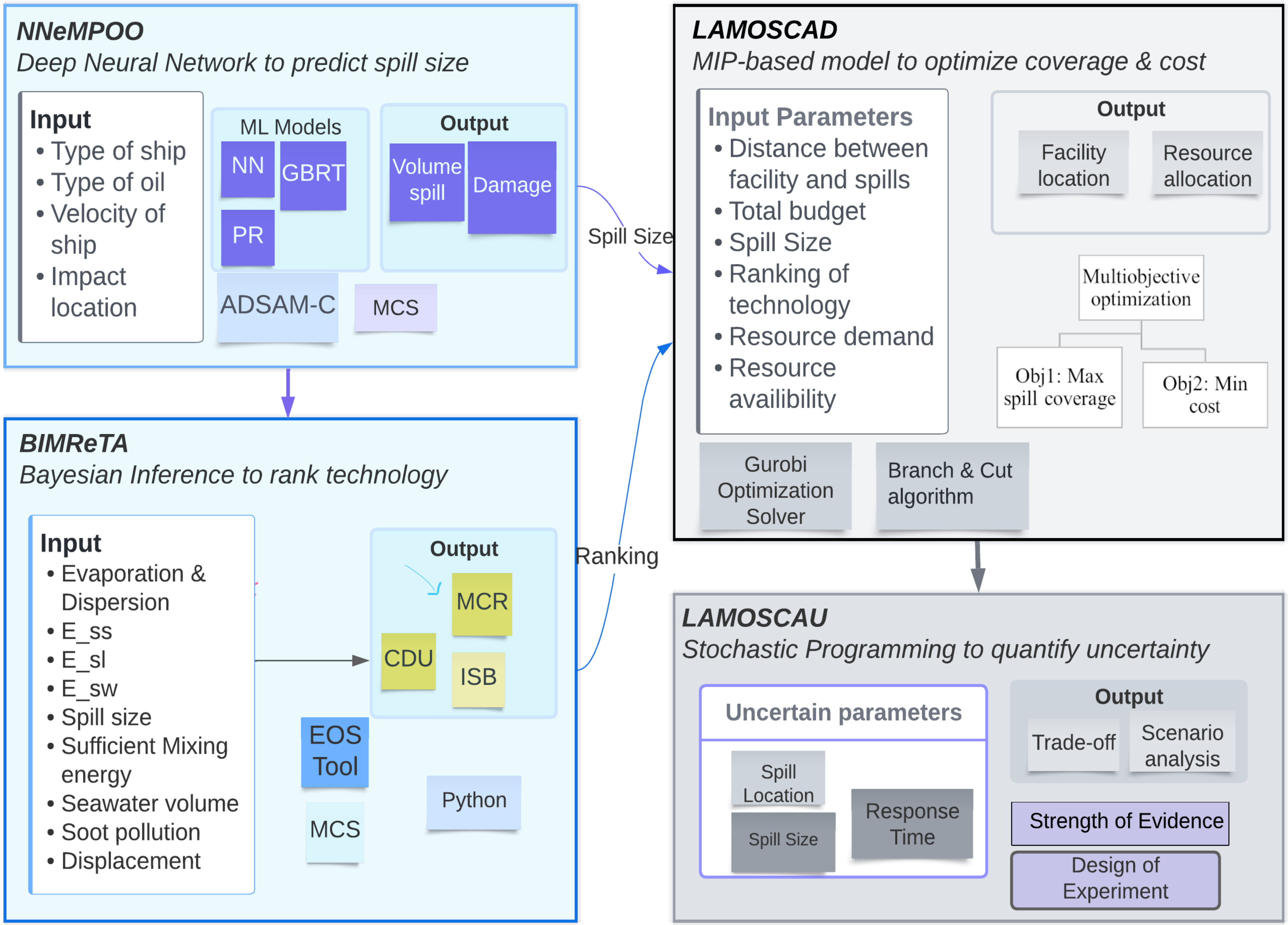
RO-2: Develop model to optimize response facility setup and resource allocation

- **RQ3:** Where and how many facilities are to be built in the Canadian Arctic, and how much resource to store?
- **RQ4:** How to quantify uncertainty in oil spill response?

Thesis Structure



DST & its models



Method & Result

Publications

This thesis is a summary document of the following five papers:

Publication No.	Title	Journal Name & Author	Author & Link
Paper I	An optimized meta**model for predicting** damage and **oil outflow** in tanker collision accidents	*PIME, Part M: Journal of Engineering for the Maritime Environment*	Das, T., Goerlandt, F., & Tabri, K. (2021) https://doi.org/10.1177/14750902211039659
Paper II	Bayesian inference modelling to **rank** response **technologies** in arctic marine oil spills	*Marine Pollution Bulletin*	Das, T., & Goerlandt, F. (2022) https://doi.org/10.1016/j.marpolbul.2022.114203
Paper III	A **Mixed Integer Programming** Approach to Improve Oil Spill Response Resource Allocation in the Canadian Arctic	*Multimodal Transportation*	Das, T., Goerlandt, F., & Pelot, R. (2023a) https://doi.org/10.1016/j.multra.2023.100110
Paper IV	**Multiobjective** Facility Location Model for Optimizing Oil-Spill Response: A case of Canadian Arctic region	*Annals of Operations Research (submitted)*	Das, T., Goerlandt, F., & Pelot, R. (2023b); Id: ANOR-D-23-02539
Paper V	Optimization under **Uncertainty**: Stochastic Facility Location Model for Oil Spill Response in Canadian Arctic	*Reliability Engineering and System Safety (Under review)*	Das, T., Goerlandt, F. (2023) Manuscript number: JRESS-D-23-02447

PI

An Optimized Metamodel for **Predicting** Damage and **Oil Outflow** in Tanker Collision Accidents

Tanmoy Das, Floris Goerlandt and Kristjan Tabri

Proceedings of the Institution of Mechanical Engineers, Part M:
Journal of Engineering for the Maritime Environment, 236(2), 412–426
2021

NNeMPOO model (PI)
Deep Neural Network to Predict Spill size

Dataset

Monte Carlo Simulation to generate data based on ADSAM-C model

Response variables: **Volume of Oil Spill**, Damage Extents

Metamodels

- Deep Neural Network
- Polynomial Regression
- Gradient Boosted Regression Tree

Performance Metric: R^2, RMSE, MAE, Runtime

Multiobjective Optimization

Improve performance of metamodels by tuning hyperparameters

Select Optimal metamodel using Pareto Frontier

NNeMPOO: Neural Network Model to **Predict Oil Outflow**

Findings from NNeMPOO model

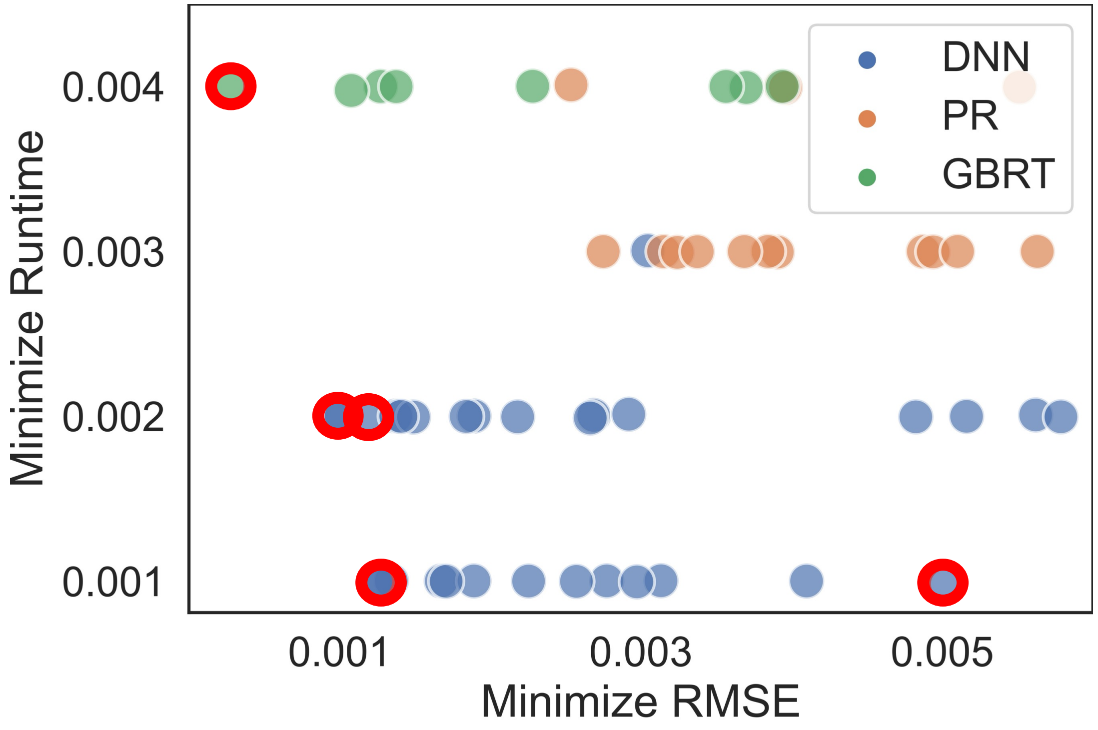


Figure 3. Trade-off between predictive accuracy vs. computational time

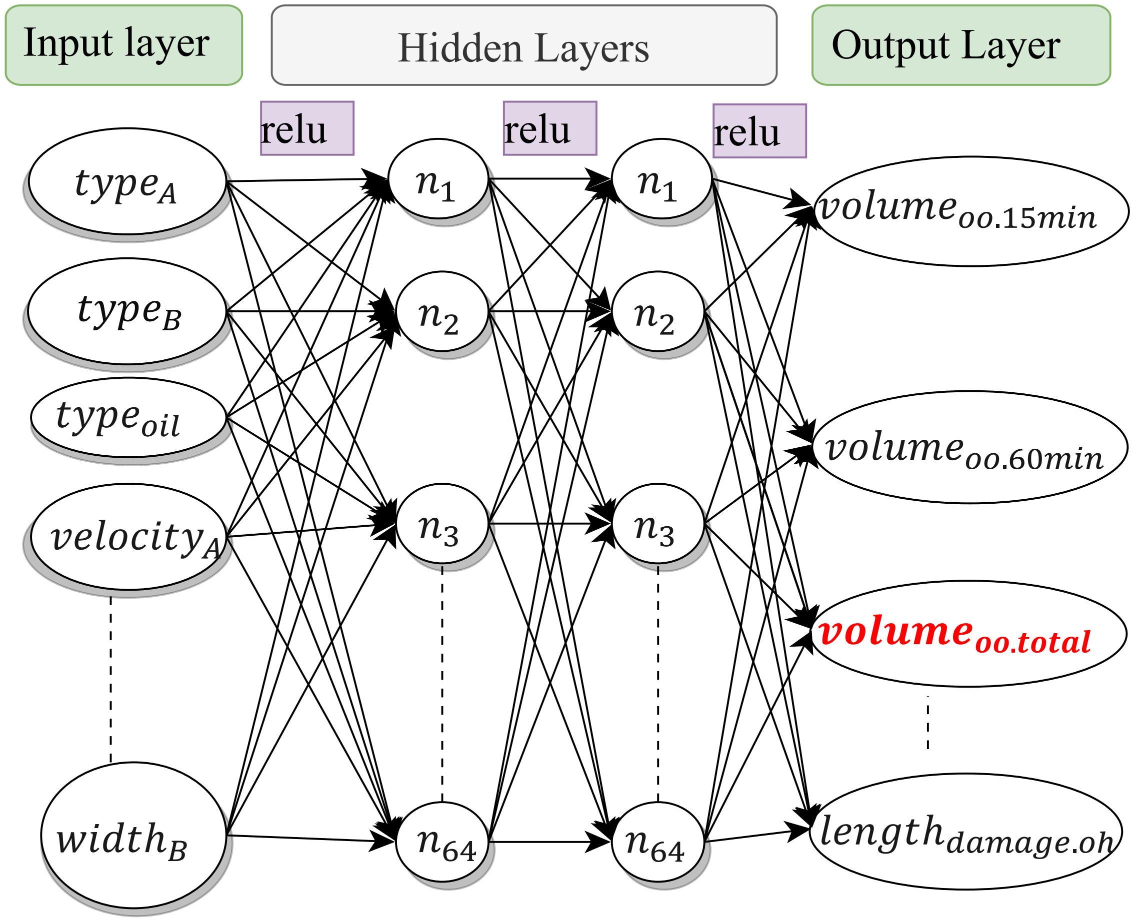


Figure 4. Structure of NNeMPOO model

PII

Bayesian Inference Modelling to **Rank Response Technologies** in Arctic Marine Oil Spills

Tanmoy Das, Floris Goerlandt

Marine Pollution Bulletin
2022

BIMReTA model (PII)
Bayesian Inference Model to Rank Response Technologies

Data

Key Factors for Cleanup Operations
Environmental: water temperature, ice presence, wind speed
Oil Properties: spill size, slick thickness
Other Arctic condition: Remoteness

Response Systems	Classes
MCR	Ok, Consider, Go next season, Unknown
CDU	Ok, Consider, Not recommended, Unknown
ISB	Ok, Consider, Unknown

BIMReTA: Bayesian Inference Model to Rank Response Technology in Arctic

Bayesian Inference Modeling

Formulation

$$\text{Probability}(\text{label} \mid x) = \frac{\text{Probability}(x \mid \text{label}).\text{Probability}(\text{label})}{\text{Probability}(x)}$$

$$\mathcal{P}(f) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2} f^T \Sigma^{-1} f\right)$$

$$\mathcal{P}_{ideal}\left(v_k \succ u_k \middle| f(v_k), f(u_k)\right) = \begin{cases} 1 & \text{if } f(v_k) \succ f(u_k) \\ 0 & \text{otherwise} \end{cases}$$

Statistical Evaluation

- Training vs testing performance using ROC-AUC, Label Ranking Average Precision, Confusion Matrix

Findings from BIMReTA model

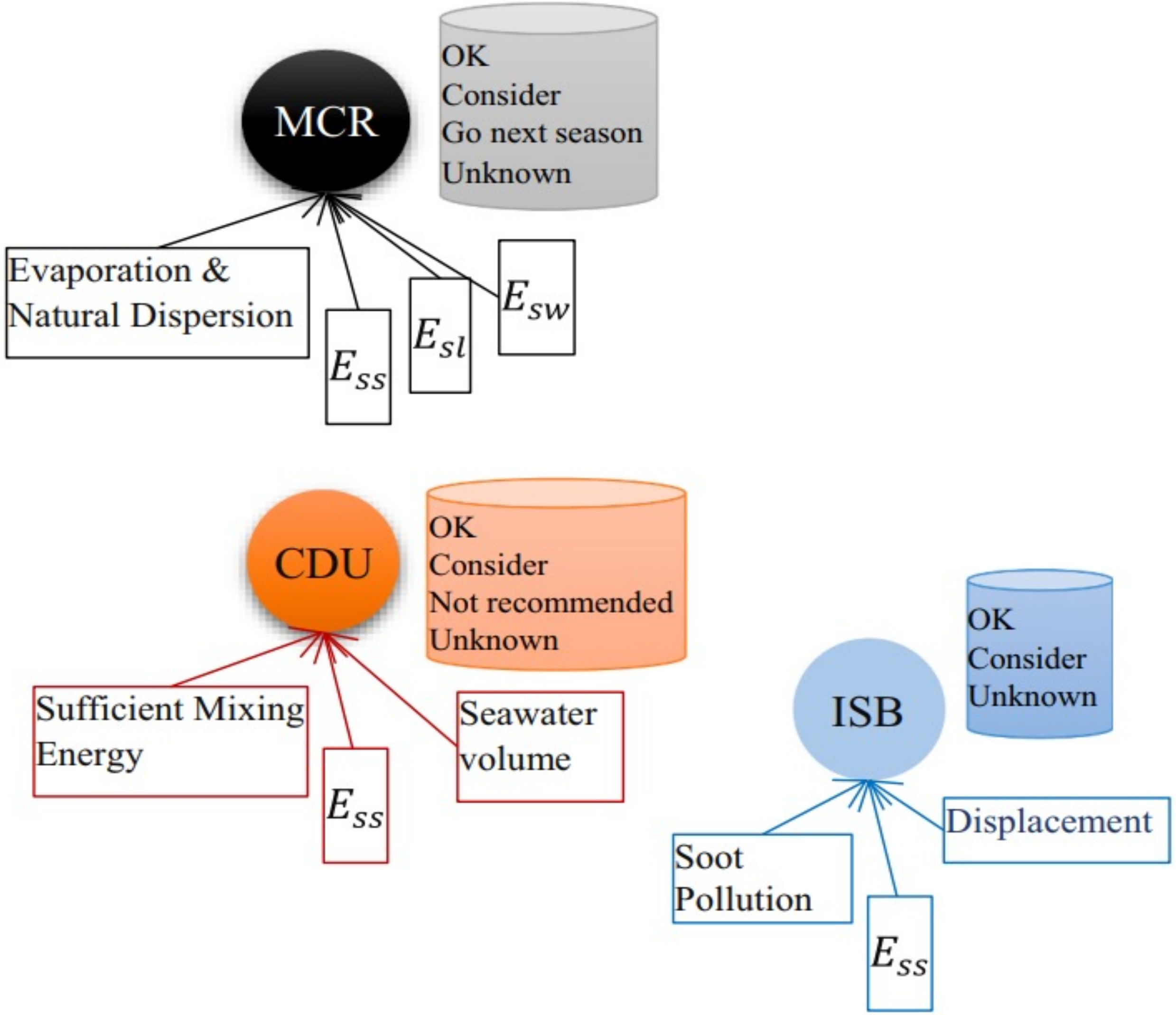


Figure 5. Structure of the BIMReTA model

Findings from BIMReTA model

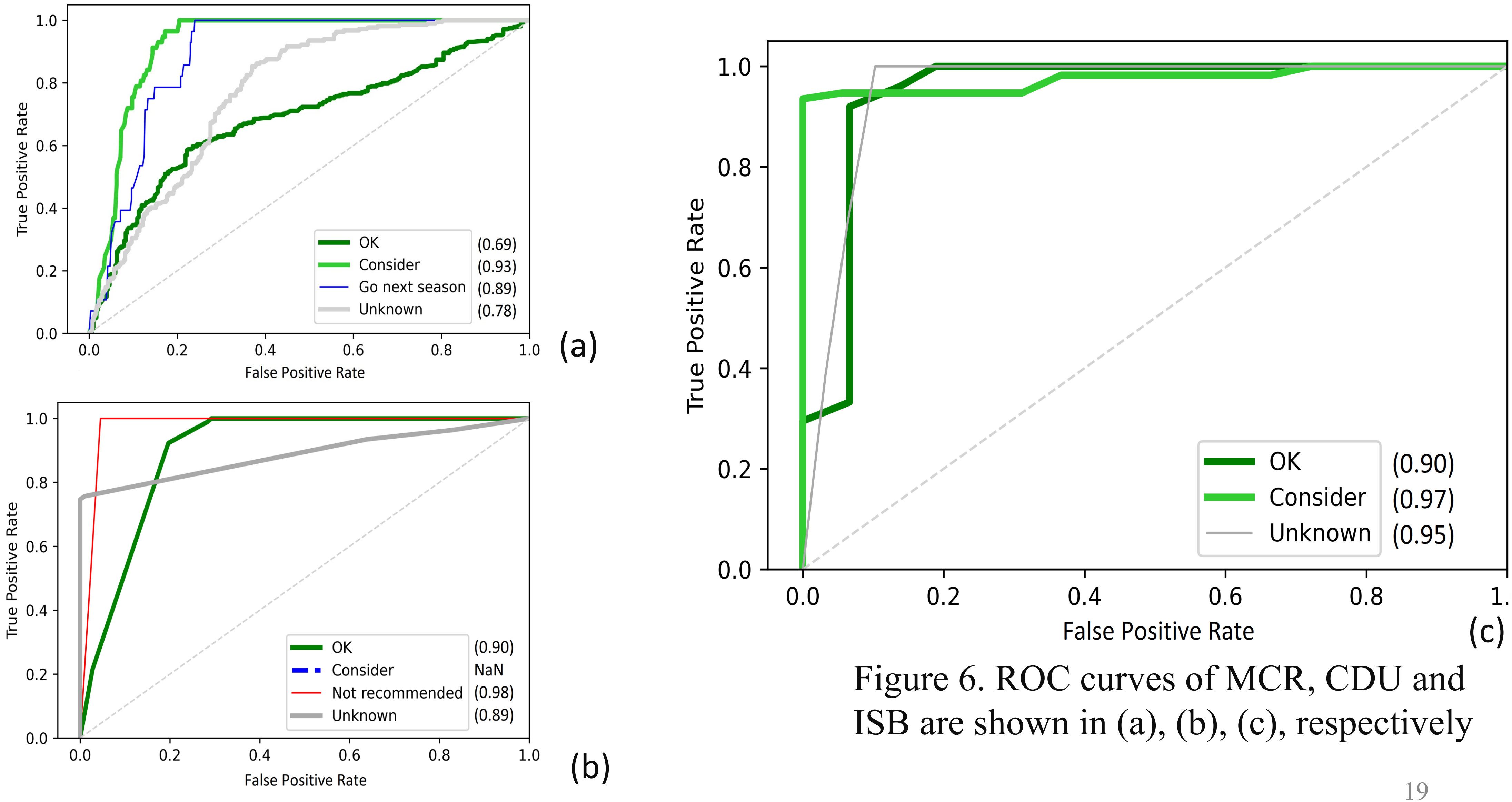


Figure 6. ROC curves of MCR, CDU and ISB are shown in (a), (b), (c), respectively

PIII

A **Mixed Integer Programming** Approach to Improve Oil Spill Response Resource Allocation in the Canadian Arctic

Tanmoy Das, Floris Goerlandt and Ronald Pelot

Multimodal Transportation
2023

LAMOSCAD-v1 Model (PIII)

A deterministic model to optimize spill coverage

$$Max \sum_{o \in O} (w_1 . \mathfrak{v}_o + w_2 . \mathfrak{y}_o - w_3 . t_{o,s}) . Y_{o,s}$$

Subject to,

$$X_{s:(o,s) \in P} \geq Y_{o,s}, \quad \forall o \in O$$

$$\sum_{s \in S} X_s = n_s$$

Classical MCLP constraint

$$\sum_{s:(o,s) \in P} Y_{o,s} \leq 1, \quad \forall o \in O$$

$$\sum_{s \in S} Z_{s,o} \leq r_o . X_s$$

$$\sum_{o \in O} Z_{s,o} \leq a_s . m . X_s, \forall s \in S$$

Demand and capacity constraint

$$\sum X_u \leq n_u, \quad X_u \subseteq X_s$$

Remoteness

$$Y_{o,s} \in [0,1], X_s \in [0,1], Z_{s,o} \geq 0$$

Set & Index

- $o \in O, s \in S$

Decision Variables

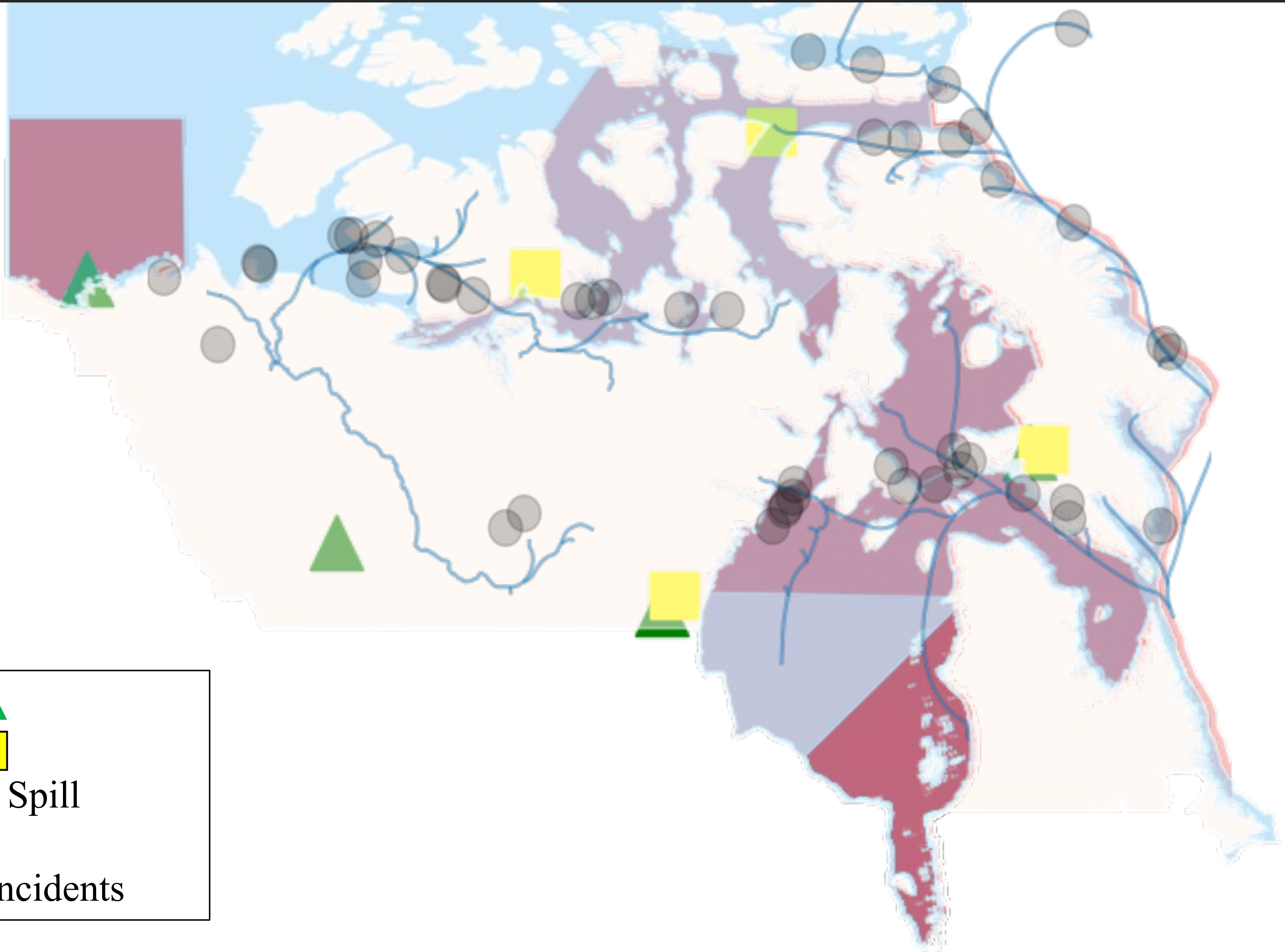
- $Y_{o,s} = 1$ if spill o is covered by station s
- $X_s = 1$ if station s is open
- $Z_{s,o}$ = the quantity of resource deployed from station s to oil spill zone o

Objective

Maximize weighted spill coverage

LAMOSCAD: Location-Allocation Modelling to Optimize Spill Coverage and Cost in the Canadian Arctic for Deterministic oil spills

Facilities, Oil Spill, Sensitivities in Canadian Arctic



Current setup: ▲
Proposed setup: ■
Facility/Supply: Oil Spill Response Stations
Demand: Oil Spill Incidents

Findings from LAMOSCAD-V1 model

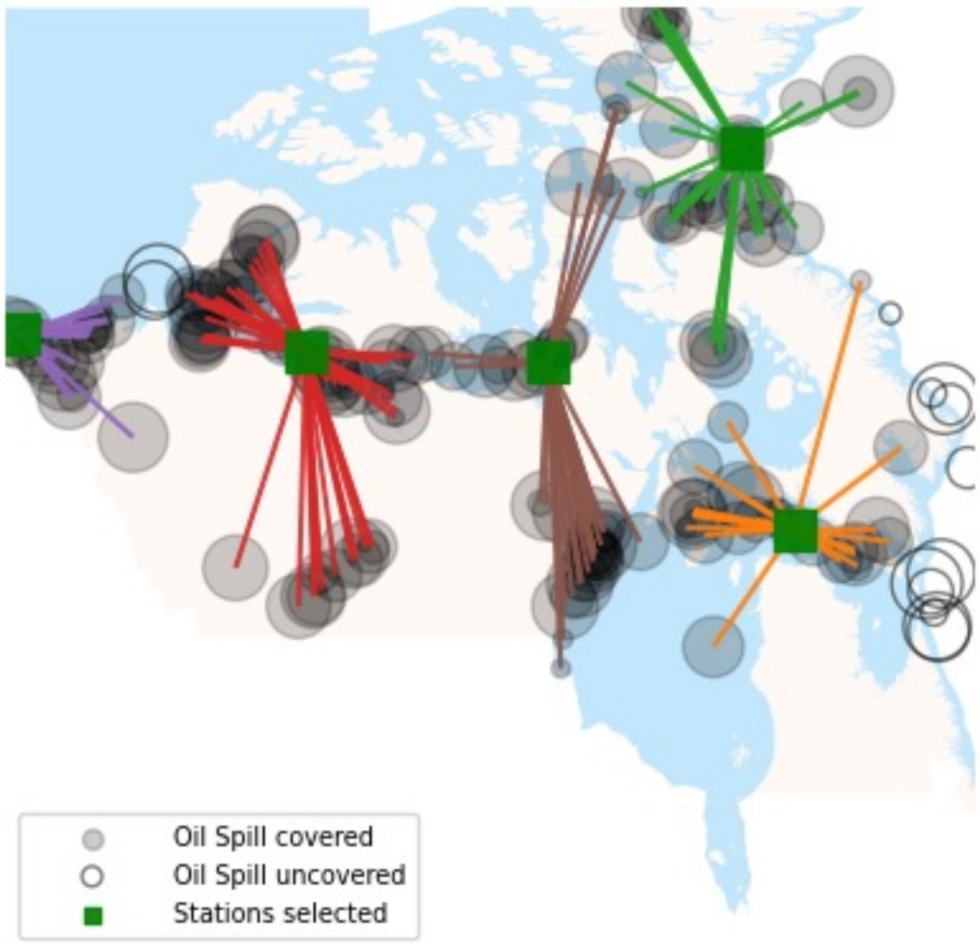


Figure 7. Network diagram of proposed response stations (green squares) and oil spills (circles)

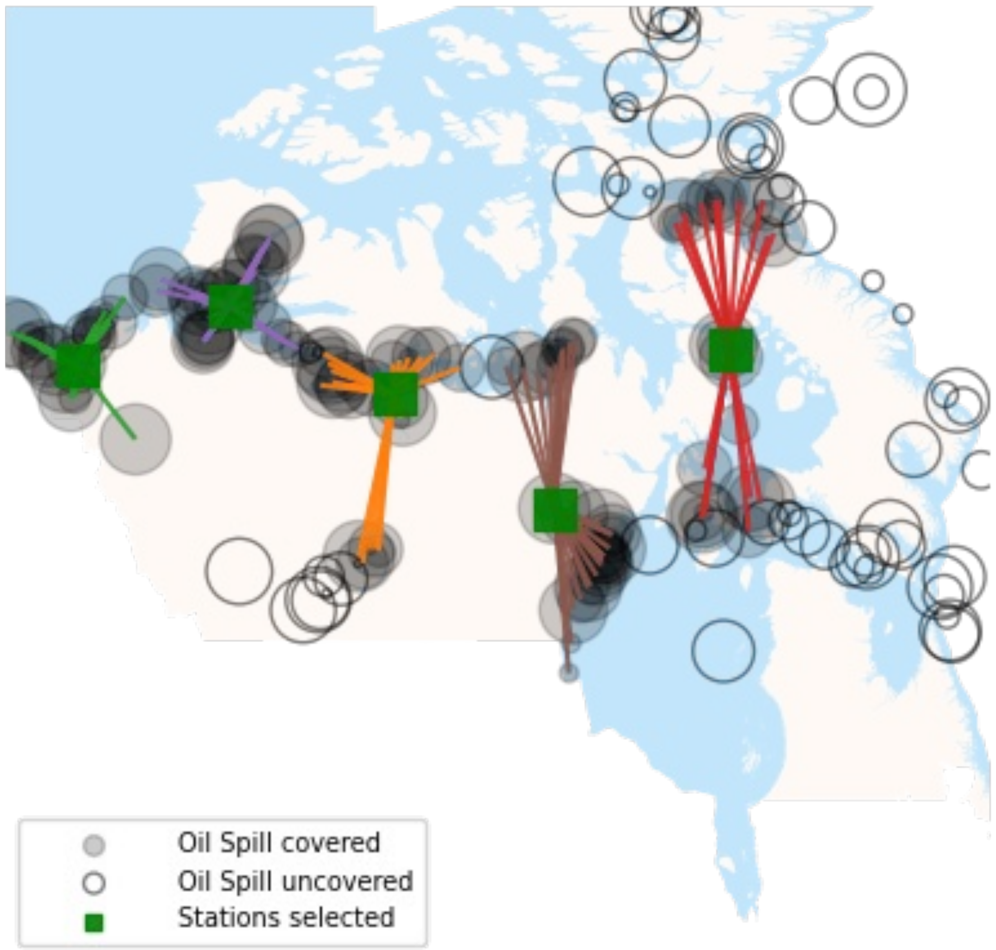


Figure 8. Network diagram obtained by applying a traditional $model_{mclp}$

PIV

Multiobjective Facility Location Model for **Optimizing Oil-Spill Response**: A case of Canadian Arctic region

Tanmoy Das, Floris Goerlandt and Ronald Pelot

Annals of Operations Research (submitted)
2023

LAMOSCAD Model (PIV)

A deterministic location-allocation model to optimize spill coverage and cost

$$\text{Max} \quad \sum_{o} \left(\underbrace{w_1 \mathfrak{v}_o}_{\text{Spill size}} + \underbrace{w_2 \mathfrak{y}_o}_{\text{Sensitivity}} - \underbrace{w_3 \sum_{s} t_{o,s}}_{\text{Response time}} \right) Y_{o,s}$$

$$\text{Min} \quad \underbrace{w_4 \sum_{s} cf_s X_s}_{\text{Fixed cost}} + \sum_{s} \left(\underbrace{w_5 \sum_{r} cu_s^r}_{\text{Variable cost}} - \underbrace{w_6 \sum_{o} \sum_{r} \sum_{t} ef_{s,o,t}^r}_{\text{Effectiveness term}} \right) Z_{s,o}^r + \sum_{s} \underbrace{\left(w_7 ct \sum_{o} d_{s,o} + w_8 \sum_{o} \sum_{r} pn_{o,s}^r \right) Vh_{s,o}^v}_{\text{Vessel related variable cost}}$$

MOO: Objective 1 = Maximize weighted oil spill coverage; Objective 2 = Minimize fixed and variable costs

Set & Index: $o \in O, s \in S, r \in R, v \in V$

Decision Variables: $Y_{o,s} = 1$ if spill o is covered by station s; $X_s = 1$ if station s is open

- $Z_{s,o}^r$ The quantity of resource r deployed from station s to oil spill zone o
- $Vh_{s,o}^v = 1$ if vessels v to send from s to o

Mathematical Programming of LAMOSCAD Model

s.t.

$$\sum_{s:(o,s)} Y_{o,s} = 1, \forall\, o \;\text{ if } \mathfrak{y}_o > snT \tag{23}$$

$$\sum_{s} \mathrm{cf}_s . X_s \leq bT \tag{25}$$

$$\sum X_H \geq n_H \;, where\; X_H \subseteq X_s \tag{26}$$

$$Y_{o,s} \leq X_s \quad, \quad \forall (o,s) \in P \tag{21}$$

$$\sum_{s:(o,s)} Y_{o,s} \leq n_{FOM}, \forall\, o \tag{22}$$

$$\sum_{s} X_s = n_s \tag{24}$$

$$\sum X_{uN} \leq n_{uN}, where\; X_{uN} \subseteq X_s \tag{27}$$

$$\sum_{o} Z_{s,o}^{r} \leq a_s^r . X_s . m, \qquad \forall r\, \forall s \tag{28}$$

$$\sum_{s}\sum_{r} Z_{s,o}^{r} \leq r_{o,r}, \qquad \forall o \tag{29}$$

$$\sum_{r} Z_{s,o}^{r} \geq \mathrm{nQ}, \qquad \forall \mathrm{s} \tag{30}$$

$$\sum_{s}\sum_{r} Z_{s,o}^{r} = 0, \forall o \;\; \text{if } d_{o,s} > dm \; or \; t_{s,o} \geq tm \tag{31}$$

$$\sum_{s} Z_{o,s}^{r=c} \leq lc_p, \qquad \forall o \tag{32}$$

$$\sum_{s} Z_{o,s}^{r=i} \leq li_p, \qquad \forall o \tag{33}$$

$$\sum_{o} \left(dr_{o,r} + vl_{o,r}\right) . Z_{s,o}^{r} \leq \sum_{v} av_v, \forall s \tag{34}$$

$$\sum_{o} Vh_{s,o}^{v} = Z_{s,o}^{r} \quad \text{and} \quad \sum_{s} Vh_{s,o}^{v} = Z_{s,o}^{r} \tag{35}$$

$$\sum_{v} Vh_{s,o}^{v} \leq n_{v,s}, \qquad \forall s \tag{36}$$

$$X_s = [0,1], Y_{s,o} = [0,1], Z_{s,o}^{r} \geq 0, Vh_{s,o}^{v} = [0,1] \tag{37}$$

Findings from LAMOSCAD model

Table 2. Different configurations of LAMOSCAD model

Model Config	Input Parameters			Outputs			
	Maximum no. of stations	Choice of facility set	Maximum distance (kms)	Obj 1	Obj 2	Coverage Percentage, $\mathfrak{C}$	$\tilde{\tau}$ (hours)
$\boldsymbol{model_C}$	4	Current	1200	963	1.5	83%	10.59
$model_2$	4	ACP	800	1006	525	78%	7.91
$model_3$	5	Current & ACP	800	600	2353	83%	6.96
$\mathbf{model_P}$	**5**	**ACP**	**800**	**1127**	**800**	**96%**	**7.33**
$model_5$	6	ACP	800	1149	4249	99%	6.64
$model_6$	8	ACP	800	1149	5057	99%	7.47

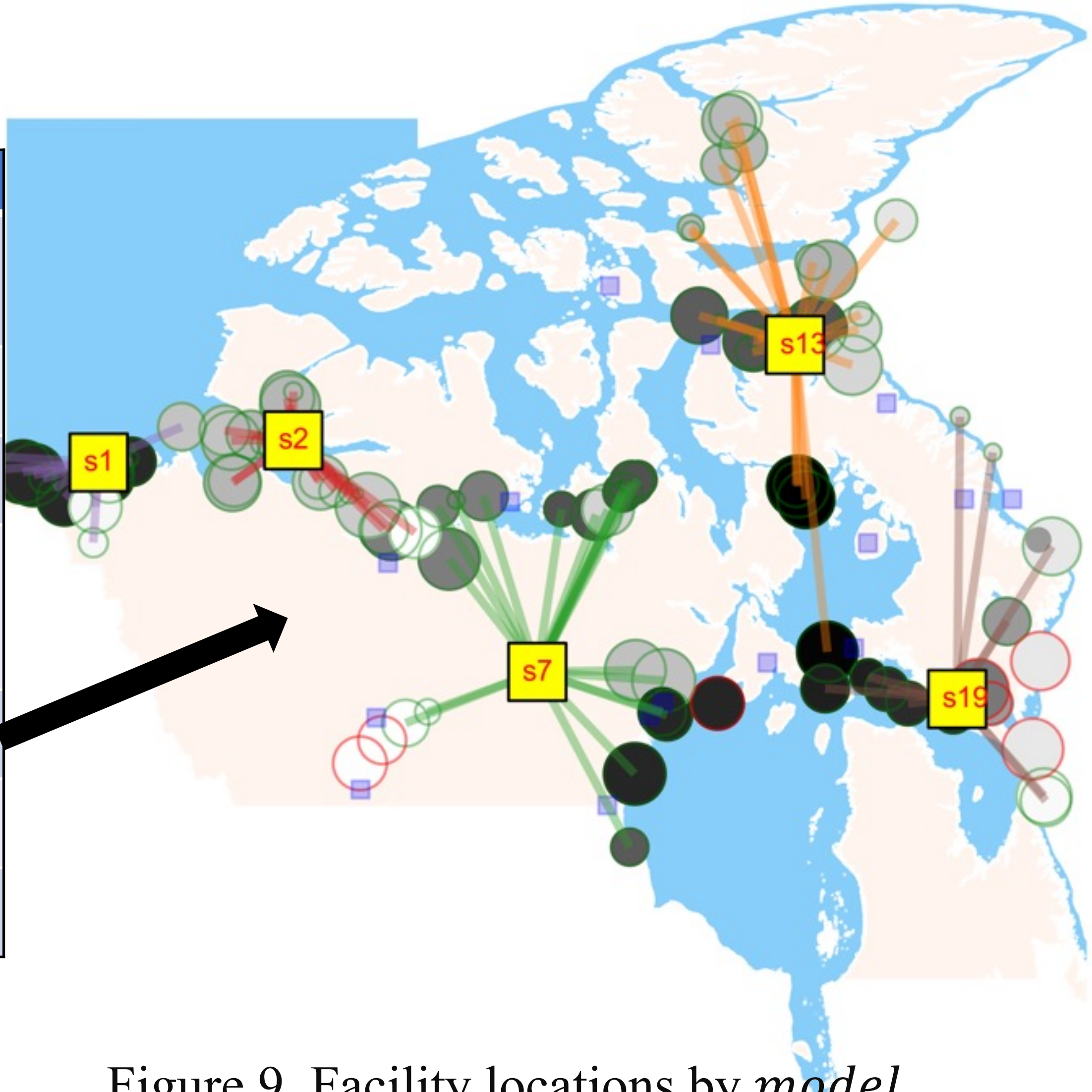


Figure 9. Facility locations by $model_p$ (proposed setup)

PV

Optimization under Uncertainty: Stochastic Facility Location Model for Oil Spill Response in Canadian Arctic

Tanmoy Das, Floris Goerlandt

Reliability Engineering and System Safety (Under review)
2023

LAMOSCAU Model (PV)
Stochastic Programming to Quantify Uncertainty

Uncertainty Parameters

Spill size, response time, location of spill

Data
- Deterministic reference model (LAMOSCAD)
- SoE to select uncertain parameters

Uncertainty based modeling (LAMOSCAU)

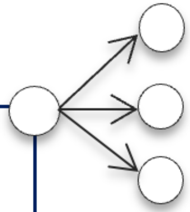
- Distribution of parameters, and MCS
- DOE & RSM for parametric estimate

Maximize spill coverage
Minimize cost

Stochastic Programming

$t_{o,s} \sim \text{Uniform}(lb, ub)$

$v_o \sim \text{Uniform}(lb, ub)$



LAMOSCAU: Location-Allocation Modelling to Optimize Spill Coverage and Cost in the Canadian Arctic for Uncertain oil spills

Model configurations

Model 1 – 5 based on different combinations of uncertainty

Findings from LAMOSCAU model

Table 3. Strength of Evidence to prioritize uncertain parameters

Uncertain Input parameters	*Data*		*Model*		*J*	*A*	*Justification of choice*	*SoE*
	Q	*Am*	*E*	*T*				
Spill size							(Das et al., 2021; Lu et al., 2019)	
Oil type							(Lu et al., 2019)	
Spill location							(Das et al., 2023a; Fu et al., 2016) cite	
Clean-up effectiveness							(Al Sharkawi, 2022; Das & Goerlandt, 2022)	
Response time							(Al Sharkawi, 2022; Das et al., 2023a)	
Response resource availability							(Das et al., 2023b)	
Weather conditions							(Al Sharkawi, 2022; Das et al., 2021; Das & Goerlandt, 2022)	
Economic factors							(Afenyo et al., 2022; DeCola et al., 2018)	
Ecological factors							(NRC, 2003)	
Note: Q = Quality, Am = Amount, E = Empirical, T = Theoretical, J = Judgement, A = Assumption								

Importance Score					
Not relevant	Low	Low-medium	Medium	Medium-high	High

Stochastic Programming of LAMOSCAU model

$$\boldsymbol{Max} \sum_{o\in O}\sum_{e\in E}\left(w_1.\mathfrak{v}_o + w_2.\mathfrak{y}_o - w_3.t_{o,s}\right).Y_{o,s}^{e} \tag{40}$$

$$\boldsymbol{Min} \sum_s \mathrm{w}_4.\mathrm{cf}_s.X_s + \sum_s w_5.cu_s^r.Z_{s,o}^{r,e} - \sum_o\sum_s\sum_r w_6.pn_{o,s}^r.\mathrm{ef}_{s,o,t}^r.Z_{s,o}^{r,e} + \sum_s w_9.ce_o.Z_{s,o}^{r,e} + \left(\sum_s\sum_o\sum_v w_7.ct_o.d_{s,o} + w_8.cp\sum_{s,v}\max\left(t_{s,o} - tm, 0\right)\right)Vh_{s,o}^{v,e} \tag{41}$$

Subject to,

$$X_{\mathrm{s:(o,e,s)\in P}} \geq Y_{\mathrm{o},s}^{e}, \forall \mathrm{o}\in\mathrm{O}\ \forall \mathrm{e}\in\mathrm{E} \tag{42}$$

$$\sum_{\mathrm{s}\in S}\sum_{\mathrm{e}\in E} X_s \leq n_s, \qquad \forall \mathrm{e}\in\mathrm{E} \tag{43}$$

$$Y_{\mathrm{o},s}^{e} \leq 1, \qquad \forall \mathrm{o}\in\mathrm{O}\ \forall \mathrm{e}\in\mathrm{E} \tag{44}$$

$$\sum_{s\in S}\sum_{e\in E} Z_{s,o}^{r,e} \leq \mathbb{d}_o^e.X_s \tag{45}$$

$$\sum_{s\in S}\sum_{e\in E} Z_{s,o}^{r,e} \leq a_s^r.X_s.mb \tag{46}$$

$$Y_{o,s}^{e}, X_s \in [0,1], Z_{\mathrm{s,o}}^{\mathrm{r},e} \geq 0, Vh_{\mathrm{s,o}}^{\mathrm{v},e} \geq 0 \tag{47}$$

Findings from LAMOSCAU model

Table 4. Model instances and their performance

Model instance	Input	n_o	Obj function	Method	Node	Incumbent	Best Bound	MIP Gap (%)	CPU Time
Model 1	[4, 0.08, 10]	100	$\mathcal{F}_1$	Heuristic	0	83	83	0	0.03
			$\mathcal{F}_2$	CP MIR1	0	7178.39	7195.9	0.24	0.00
		500	$\mathcal{F}_1$	Heuristic	0	81	81	0	0.002
			$\mathcal{F}_2$	CP MIR1	0	11232	11232	0	0.07
		3100	$\mathcal{F}_1$	Heuristic	0	82.7	82.7	0	0.03
			$\mathcal{F}_2$	CP MIR1	7	23109	23112	0.00	0.18
Model 2	[4, 0.08, 10]	100	$\mathcal{F}_1$	B&C	0	81.1	81.5	0	0.002
			$\mathcal{F}_2$	CP MIR1	0	6701	6701	0	0.00
		500	$\mathcal{F}_1$	B&C	0	72.6	73.2	-0.01	0.01
			$\mathcal{F}_2$	CP MIR1	0	12330	12337	0.000	0.00
		3100	$\mathcal{F}_1$	Heuristic	0	85.7	85.7	0	0.04
			$\mathcal{F}_2$	CP MIR1	7	19877	19877	0	0.18
Model 5	[4, -, 10]	100	$\mathcal{F}_1$	B&C	0	72	86.1	-0.195	0
			$\mathcal{F}_2$	CP MIR1	0	11200	11180	0.001	0
		500	$\mathcal{F}_1$	B&C	0	79	74.6	0.055	0.001
			$\mathcal{F}_2$	CP MIR1	0	12345	12340	0.000	0
		3100	$\mathcal{F}_1$	Heuristic	0	79	80	-0.012	0.02
			$\mathcal{F}_2$	CP MIR1	5	19883.5	19878	0.000	0.024

Note: F1 = objective function 1, F2 = objective function 2, CP: Cutting Plane

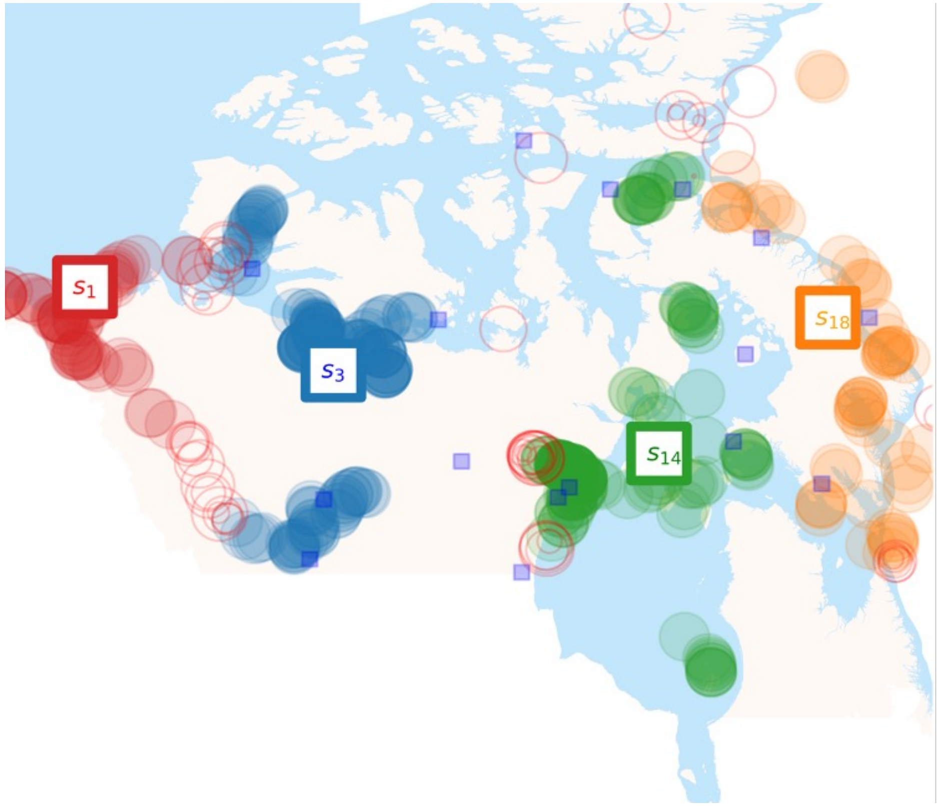
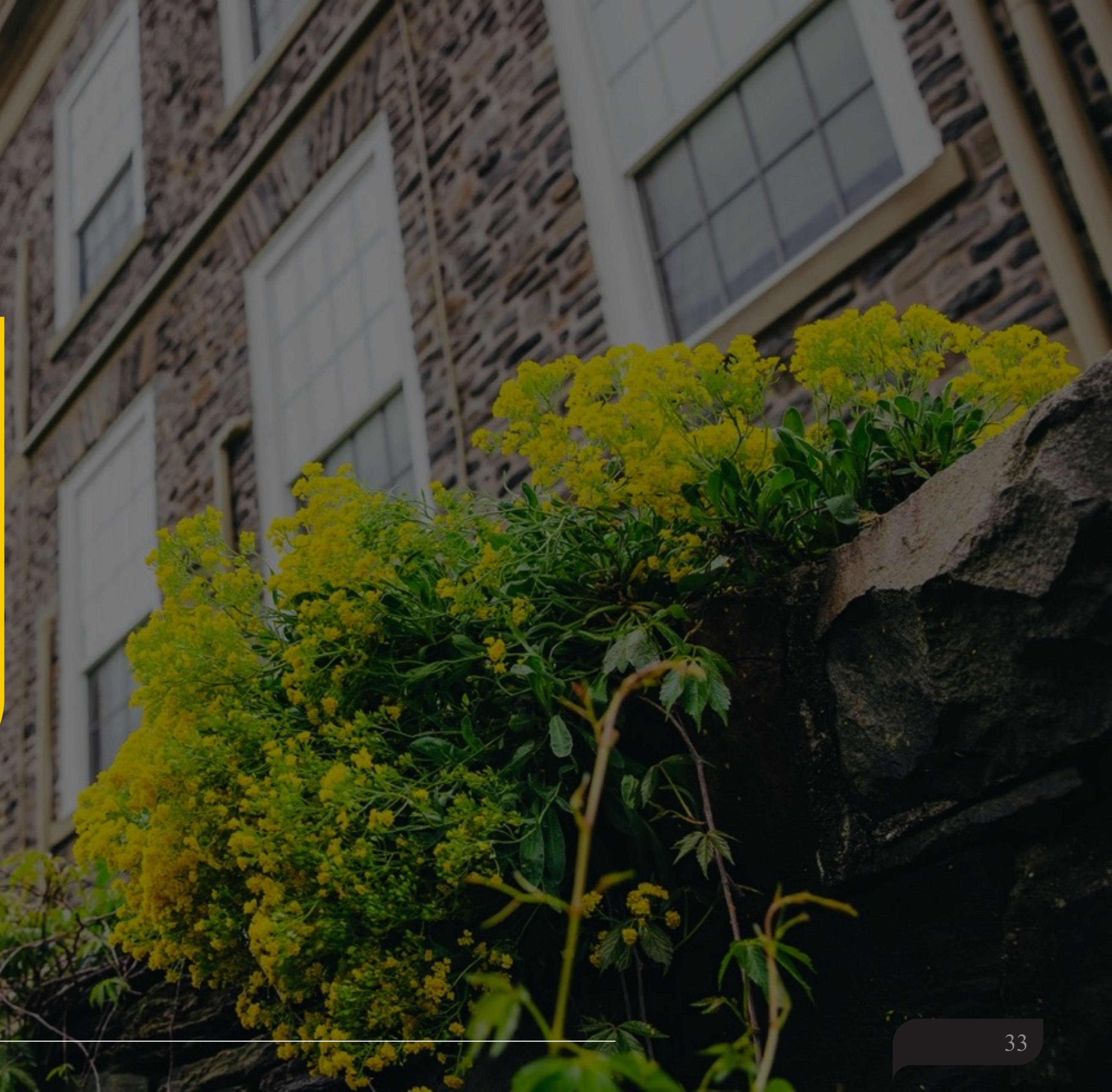


Figure 10. Cluster graph to display facilities and spills assigned to facilities

Discussion



Scientific contribution: DST Perspective

- **Decision Support Tool (DST)**
  - Development of a unified DST for oil spill response in the Canadian Arctic
  - Improved strategic spill response
- **Managerial insights**
  - Recommended response facility locations, ensuring long-term strategic planning and risk mitigation strategies
  - Enhanced understanding and management of oil spills in the challenging Canadian Arctic context
- **Data Analytics**
  - Machine Learning models (NNeMPOO & BIMReTA)
  - Application of Monte Carlo Simulation to generate representative oil spills
- **Optimization**
  - A multiobjective optimization model for transportation networks involving oil spill demand and response facilities
  - Uncertainty quantification by Stochastic Programming

Scientific contribution: Individual models

- **NNeMPOO** [PI]
  - The first article to develop an optimized model for spill estimation
  - 18 times faster than engineering models and is highly accurate
- **BIMReTA** [PII]
  - Employs data analytics and machine learning, trained on a large dataset using EOS Tool
  - Offers a data-driven alternative to expert-dependent ranking methods
- **LAMOSCAD** [PIII, PIV]
  - A novel multi-objective optimization of spill coverage and cost
  - Supports strategic preparedness and response planning in Arctic region
- **LAMOSCAU** [PV]
  - Stochastic Programming quantifies parametric uncertainty
  - Strength of Evidence approach selects uncertain parameters

Limitation and Future work

Limitations

- Difficulty in finding reliable, high-quality data for oil spills in the Canadian Arctic
- The NNeMPOO model's lack of consideration for expert judgment. The model's failure to explicitly address Black Swan events
- Non-intuitive ranking scaling and the lack of explicit consideration for seasonal and dynamic changes in BIMReTA model
- Distributional assumptions of LAMOSCAU model, the curse of dimensionality

Future Work

- Encompass ecosystem dynamics, socio-cultural dimensions, and collaboration with ongoing research efforts
- Developing predictive models that consider Arctic corridors and their impact on spill response
- Evaluating and incorporating emerging technologies like drones, autonomous surface vehicles