A semi-“smart predict then optimize”(semi-SPO) method for efficient ship inspection

Presenter: YAN Ran (鄢然)

Yan, R., Wang, S., & Fagerholt, K. (2020). A semi-“smart predict then optimize”(semi-SPO) method for efficient ship inspection. *Transportation Research Part B: Methodological*, *142*, 100-125.

The Hong Kong Polytechnic University

22 January 2021

Contents

1. Introduction

1 Introduction

Example of maritime accidents

Source: https://www.google.com/search?q=maritime+accident&source=lnms&tbm=isch&sa=X&ved=0ahUKEwjP39-A_dPhAhWKPXAKHWqfAUYQ_AUIDigB&biw=1522&bih=652

1 Introduction

Objectives

- Reduce maritime risks
- Protect marine environment
- Guarantee good working and living conditions for the crew.

International rules and recommendations

❑ **SOLAS**:
Sets minimal safety standards in the construction, equipment and operation of merchant ships.

❑ **MARPOL**:
Aims to minimize pollution of the oceans and seas, including dumping, oil and air pollution.

❑ **STCW**:
Sets minimum qualification standards for masters, officers and watch personnel on seagoing merchant ships and large yachts.

maritime transportation

maritime safety

Port state Control

“second line of defense”

“last safety net”

Flag state

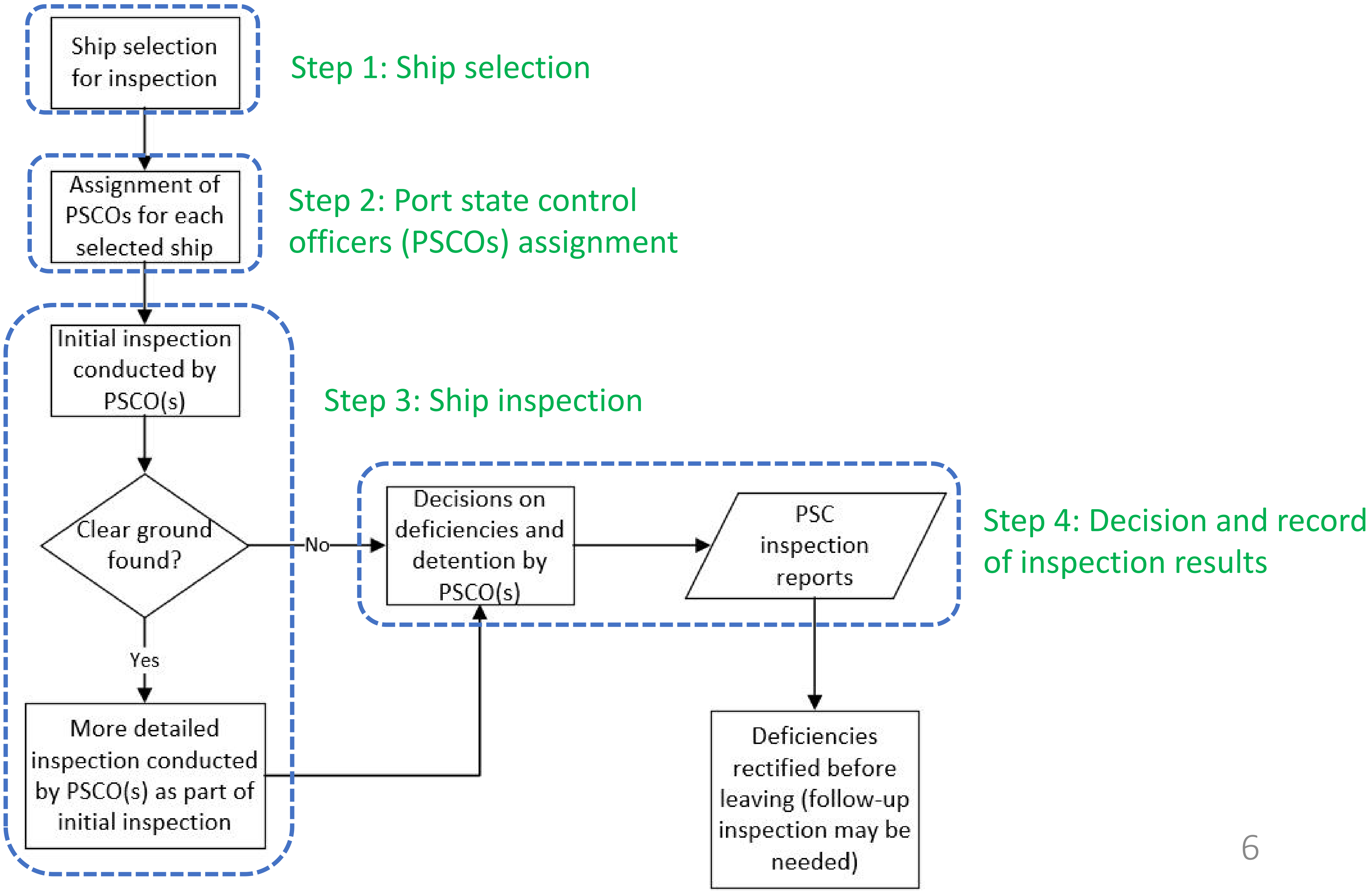
“first line of defense”,

unable to perform well

Substandard ships

Ships cannot comply with the conventions.

1 Introduction



Example of deficiencies detected by the PSC authority

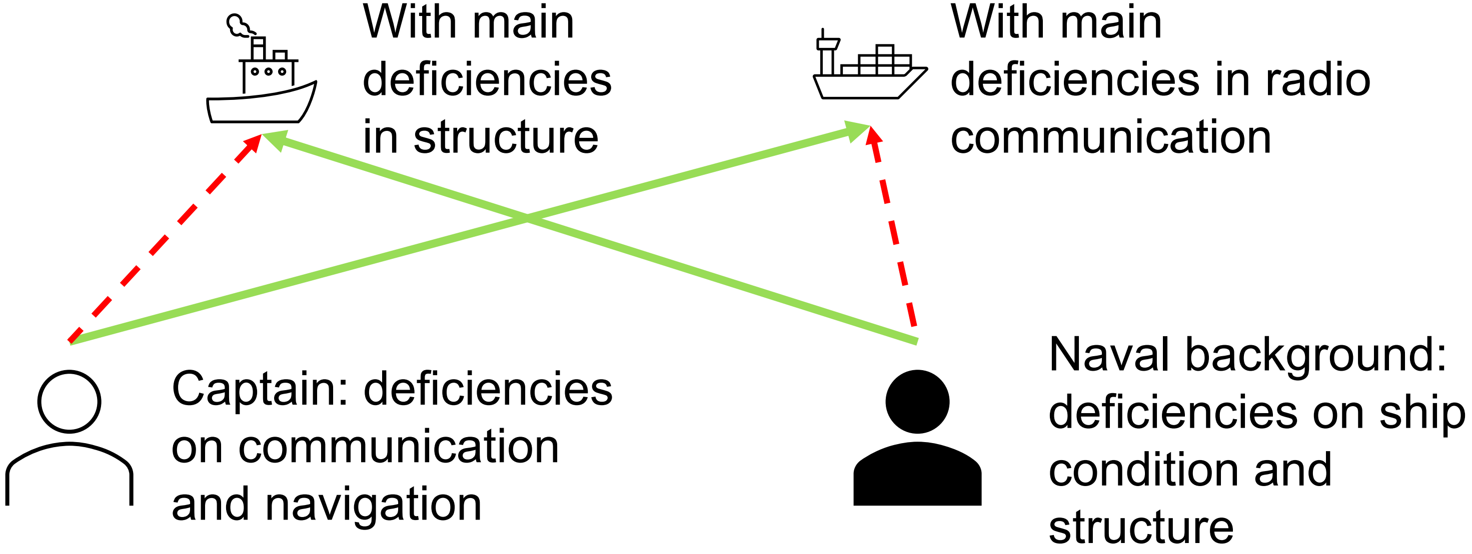
Source: Tokyo MoU, 2017; Tokyo MoU, 2018ba

Deficiency category and items

Deficiency category	Deficiency items
C1: ship safety	D4 Emergency system, D7 Fire safety, D11 Life saving appliances, D12 Dangerous goods
C2: ship management	D1 Certificates and documentation, D9 Working and living conditions, D14 Pollution prevention, D15 International Safety Management (ISM), D18 Labour conditions
C3: ship condition and structure	D2 Structural condition, D3 Water/Weathertight condition, D6 Cargo operations including equipment, D13 Propulsion and auxiliary machinery
C4: communication and navigation	D5 Radio communication, D8 Alarms, D10 Safety of navigation

Port state control officer (PSCO)

Requirement	An experienced person with both theoretical knowledge and seagoing
	erchant marine captains, d radio officers.
	subjectivity, individuality, nent, different backgrounds COs at the same port may rtise in identifying different encies.
	be ignored due to the lack kgrounds and knowledge.
	onditions and the expertise hed.



https://safety4sea.com/cm-port-state-control-a-flag-state-perspective/

1 Introduction

Contribution

Theoretically

- ✓ **3 sequential prediction and optimization approaches for the PSCO assignment problem**

Practically

- ✓ **Address a practical and meaningful problem in maritime transportation**
- ✓ **Matching the expertise of PSCOs with various conditions of ships**
- ✓ **4.70%, 4.55%, and 4.86% more deficiencies can be detected compared with random PSCO assignment**

Model	Prediction targets	Splitting criteria	Decision trees	Assignment model	Assignment decision
MTR-RF1	Number of deficiencies under each deficiency category	MSE	$f^{MTR}(\mathbf{x})$	M1	A1
MTR-RF2	Number of deficiencies identified by each PSCO	MSE	$f'^{MTR}(\mathbf{x})$	M2	A2
MTR-RF3	Number of deficiencies identified by each PSCO	MSO	$f''^{MTR}(\mathbf{x})$	M2	A3

*MTR-RF3: semi- smart predict then optimize (SPO) framework
Partially integrate optimization models into the prediction, improving the performance while incurring limited extra computational burden
**MSO: the mean squared difference regarding the overestimates (i.e., predicted value minus actual value) in the numbers of deficiencies that can be detected among the PSCOs for each ship

2. Data and models

2 Data and models

Description of input features

Feature name	Meaning	Min value	Max value	Average value
Age (year)	Difference between keel laid date and inspection date.	0	47	11.00
GT (100 cubic feet)	A measure of a ship's overall internal volume.	299.00	266,681.00	44,075.59
Length (meter)	The overall maximum length of a ship.	32.29	400.00	212.73
Depth (meter)	The vertical distance measured from the top of the keel to the underside of the upper deck at side.	3.70	36.02	17.60
Beam (meter)	The width of the hull.	7.38	60.05	31.64
Type	Bulk carrier (12.70%), container ship (57.05%), general cargo/multipurpose (10.95%), passenger ship (1.35%), tanker (11.50%), other (6.45%).	/	/	/
Number-of-times-of-changing-flag	The sum of the times the ship's flag has been changed after keel laid date.	0	8	0.69
Total-detention-times	The sum of the times the ship has been detained by all PSC authorities.	0	18	0.62
Casualties-in-last-five- years	1, if the ship is encountered with casualties in last five years; 0, otherwise.	0	1	0.09
Ship-flag-performance*	White (92.10%), grey (3.20%), black (4.05%), not listed (0.65%).	/	/	/
Ship-RO-performance*	High (95.85%), medium (2.30%), low (0.10%), very low (0), not listed (1.75%).	/	/	/
Ship-company- performance*	High (34.50%), medium (40.25%), low (15.70%), very low (9.10%), not listed (0.45%).	/	/	/
Last-inspection-time (month)	The time of last PSC inspection within Tokyo MoU.	0.03	180.70	10.12
Last-deficiency-number	The deficiency number of last PSC inspection within Tokyo MoU.	0	55	3.41
Follow-up-inspection- rate	The total number of follow-up inspections divided by total number of inspections within Tokyo MoU.	0.00	0.64	0.15

* Note: Ship flag performance, recognized organization (RO) performance, and company performance are calculated based on flag Black-Grey-White list, RO performance list, and company performance list provided by Tokyo MoU, respectively. The performance of the flags on white-list is better than those on grey-list, and much better than those on black-list. For RO and company, the performance gets worse in the sequence of "high", "medium", "low", and "very low". If the performance of the ROs and companies is not shown on the lists, the performance state is recorded as "not listed".

Source: http://www.tokyo-mou.org/inspections_detentions/psc_database.php

Prediction of natural targets– MTR-RF1

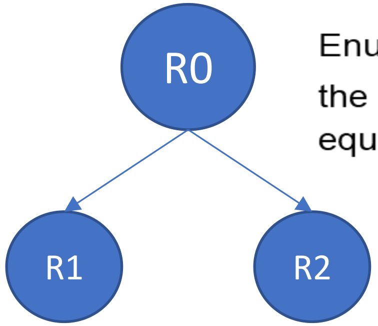
Prediction targets

The numbers of deficiencies in each deficiency category of a ship

Splitting criteria

Mean squared error (MSE)

Node split in an ordinary multi-target regression (MTR) tree (based on CART algorithm)



Enumerate of all features and their values to find the best split pair $(j^*, w^*{}_{j^*})$ by the following equation

Error of the prediction targets of all examples in R1

$$(j^*, w^*_{j*}) \in \underset{\substack{j\in(x_1,\dots,x_J)\\ w_j\in\Omega_j}}{\arg\min} \left[\sum_{e\in R_1(j,w_j)} \sum_{c=1}^{C} \left(\alpha^{ec} - \frac{1}{|R_1(j,w_j)|} \sum_{e_1\in R_1(j,w_j)} \alpha^{e_1c} \right)^2 + \sum_{e\in R_2(j,w_j)} \sum_{c=1}^{C} \left(\alpha^{ec} - \frac{1}{|R_2(j,w_j)|} \sum_{e_2\in R_2(j,w_j)} \alpha^{e_2c} \right)^2 \right] \quad (1)$$

Error of the prediction targets of one example

where $R_1(j,w_j) = \{e \in R_0 | x^{ej} \le w_j\}$ and $R_2(j,w_j) = \{e \in R_0 | x^{ej} > w_j\}$.

MTR-RF1 based on MTR trees

- **Two layers of randomness**
  - ✓ The training set for each MTR tree is generated by bootstrapping
  - ✓ A subset of all features is used to split each node
- **The predicted number of deficiencies in each category is the average of the outputs of all MTR trees in the forest**

2 Data and models

Optimization model using the output of MTR-RF1: M1

Notation	Meaning
S	The set of ships selected to be inspected for one day
P	The number of PSCOs that can be assigned for inspection
C	The number of deficiency categories, equals 4
u_{pc}	Expertise of PSCO p for inspecting deficiency category c, i.e. the percentage of deficiencies that can be inspected
Θ	The number of ships that can be inspected by a PSCO for each day

[M1] Output of MTR-RF1

$$\max \sum_{p=1}^{P} \sum_{s \in S} \sum_{c=1}^{C} \boxed{\hat{\alpha}^{sc}} u_{pc} z_{ps}$$

Maximizes the inspection expertise of the PSCOs

subject to

$$\sum_{s \in S} z_{ps} \le \Theta, \; p = 1, \ldots, P$$

Limit the number of ships that can be inspected by one PSCO in one day

$$\sum_{p=1}^{P} z_{ps} = 1, \; s \in S$$

Guarantee that one ship is inspected by one PSCO

$$z_{ps} \in \{0, 1\}, \; p = 1, \ldots P, \; s \in S.$$

Proposition 1

Model [M1] can be solved in polynomial time of the length of the input parameters

2 Data and models

Prediction of coefficients in the objective function of optimization model– MTR-RF2

$$\max \sum_{p=1}^{P}\sum_{s\in S}\sum_{c=1}^{C}\hat{\alpha}^{sc}u_{pc}z_{ps} \Rightarrow \beta^{sp}=\sum_{c=1}^{C}\alpha^{sc}u_{pc},\quad \boldsymbol{\beta}^{s}=\left(\beta^{s1},\ldots,\beta^{sp},\ldots,\beta^{sP}\right)$$

Prediction targets

The total numbers of deficiencies of a ship that can be detected by the PSCOs

Splitting criteria

Mean squared error (MSE)

Optimization model using the output of MTR-RF2: M2

[M2]

Output of MTR-RF2

$$\max \sum_{p=1}^{P}\sum_{s\in S}\boxed{\beta^{sp}}z_{ps}$$

Maximizes the inspection expertise of the PSCOs

Subject to constraints of M1.

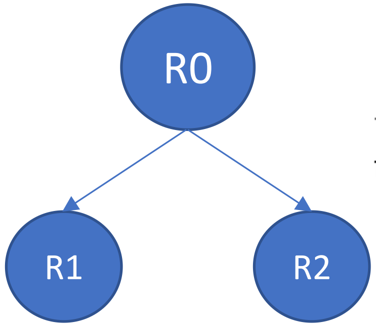
Error of the prediction targets of all examples in R1

$$(j^{*},w_{j*}^{*})\in \underset{\substack{j\in(x_1,\ldots,x_J)\\ w_j\in\Omega_j}}{\arg\min}\left[\sum_{e\in R_1(j,w_j)}\sum_{p=1}^{P}\left(\beta^{ep}-\frac{1}{|R_1(j,w_j)|}\sum_{e_1\in R_1(j,w_j)}\beta^{e_1p}\right)^2\right.$$

Error of the prediction targets of one example

$$\left.+\sum_{e\in R_2(j,w_j)}\sum_{p=1}^{P}\left(\beta^{ep}-\frac{1}{|R_2(j,w_j)|}\sum_{e_2\in R_2(j,w_j)}\beta^{e_2p}\right)^2\right]$$

where $R_1(j,w_j)=\{e\in R_0|x^{ej}\le w_j\}$ and $R_2(j,w_j)=\{e\in R_0|x^{ej}>w_j\}$.

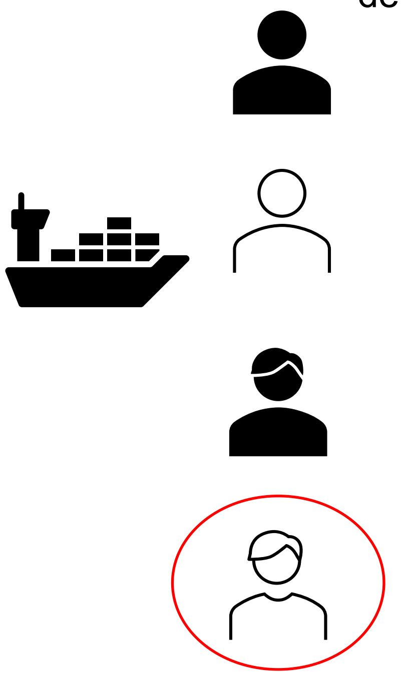


Enumerate of all features and their values to find the best split pair $(j^{*},w^{*}{}_{j^{*}})$ by the following equation

2 Data and models

Prediction of fundamental parameters that are fed into the optimization model– MTR-RF3

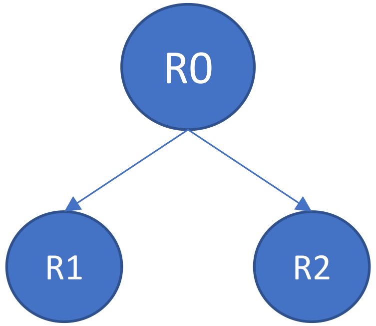
Example



Actual number of deficiencies identified	Predicted number of deficiencies identified
6	8 (+2)
7	9 (+2)
8	10 (+2)
9	11 (+2)

Prediction targets

The total numbers of deficiencies of a ship that can be detected by the PSCOs



Enumerate of all features and their values to find the best split pair $(j^*, {w^*}_{j^*})$ by the following equation

Splitting criteria

The mean squared difference regarding the overestimates (predicted value minus actual value) in the number of deficiencies that can be detected among the PSCOs for a ship (MSO))

Difference between the overestimates

$$(j^*, w^*_{j*}) \in \underset{\substack{j \in (x_1,\dots,x_J) \\ w_j \in \Omega_j}}{\arg\min} \left[\sum_{e \in R_1(j,s_j)} \sum_{p=1}^{P-1} \sum_{p'=p+1}^{P} \left(\left(\beta^{ep} - \frac{1}{|R_1(j,w_j)|} \sum_{e_1 \in R_1(j,w_j)} \beta^{e_1 p} \right) - \left(\beta^{ep'} - \frac{1}{|R_1(j,w_j)|} \sum_{e_1 \in R_1(j,w_j)} \beta^{e_1 p'} \right) \right)^2 + \sum_{e \in R_2(j,s_j)} \sum_{p=1}^{P-1} \sum_{p'=p+1}^{P} \left(\left(\beta^{ep} - \frac{1}{|R_2(j,w_j)|} \sum_{e_2 \in R_2(j,w_j)} \beta^{e_2 p} \right) - \left(\beta^{ep'} - \frac{1}{|R_2(j,w_j)|} \sum_{e_2 \in R_2(j,w_j)} \beta^{e_2 p'} \right) \right)^2 \right]$$

Sum of the differences between each two PSCOs

Motivation

If the predicted number of deficiencies t
can identify for a ship are overestimated or underestimated by the same value, the final optimal assignment decision will not be changed.

3. Computational experiments

3 Computational experiments

Construction of MTR-RFs

Expertise of each PSCO in each deficiency category

PSCO/deficiency category	C1	C2	C3	C4
PSCO1	0.8	0.5	0.7	0.6
PSCO2	0.7	0.9	0.4	0.5
PSCO3	0.7	0.6	0.8	0.7
PSCO4	0.4	0.7	0.6	0.7

1,400 examples	300 samples	300 samples
Initial training set	Validation set	Test set

Hyperparameters in the MTR-RF models

Hyperparameter	Explanation	Value/Search space	Optimal values		
			MTR-RF1	MTR-RF2	MTR-RF3
n_estimators	The total number of DTs contained in an RF model.	200	200	200	200
max_features	The number of features considered for each split.	{3, 4, 5, 6, 7}	4	4	6
max_depth	The maximum depth of each DT in the RF model	{4, 5, 6, 7, 8}	8	7	8
min_samples_leaf	The minimum number of examples required to be at a leaf node	{2, 3, 4, 5, 6, 7, 8}	5	3	4

Construction of MTR-RFs

Experient settings

- Combine the initial training set and validation set to form a new training set for final model construction
- Divide the test set into 30 groups with each group containing 10 ships
- The number of PSCOs is 4, the maximum number of ships that can be inspected by one PSCO is 3
- Run each model 10 times

Model performance

Model	Metric	Min	Mean	Max	Variance
MTR-RF1	MSE	3.9756	4.0173	4.0762	0.0009
MTR-RF2	MSE	15.4953	15.8342	16.1237	0.0437
MTR-RF3	MSE	16.7775	17.1684	17.5571	0.0443
MTR-RF3	MSO	3.0242	3.0513	3.0863	0.0002

- Model performance (MSE): MTR-RF2 < MTR-RF3
- Model variability: MTR-RF1 < MTR-RF2 < MTR-RF3

3 Computational experiments

Combination of MTR-RFs with PSCO assignment models

Mean inspection expertise of the three models

	Random PSCO assignment	A1 (MTR-RF1+M1)	A2 (MTR-RF2+M2)	A3 (MTR-RF3+M2)	Best in theory
Average	38.90	40.73	40.67	40.79	42.90
Ratio	90.68%	94.94%	94.80%	95.08%	100%

Randomness of model performance

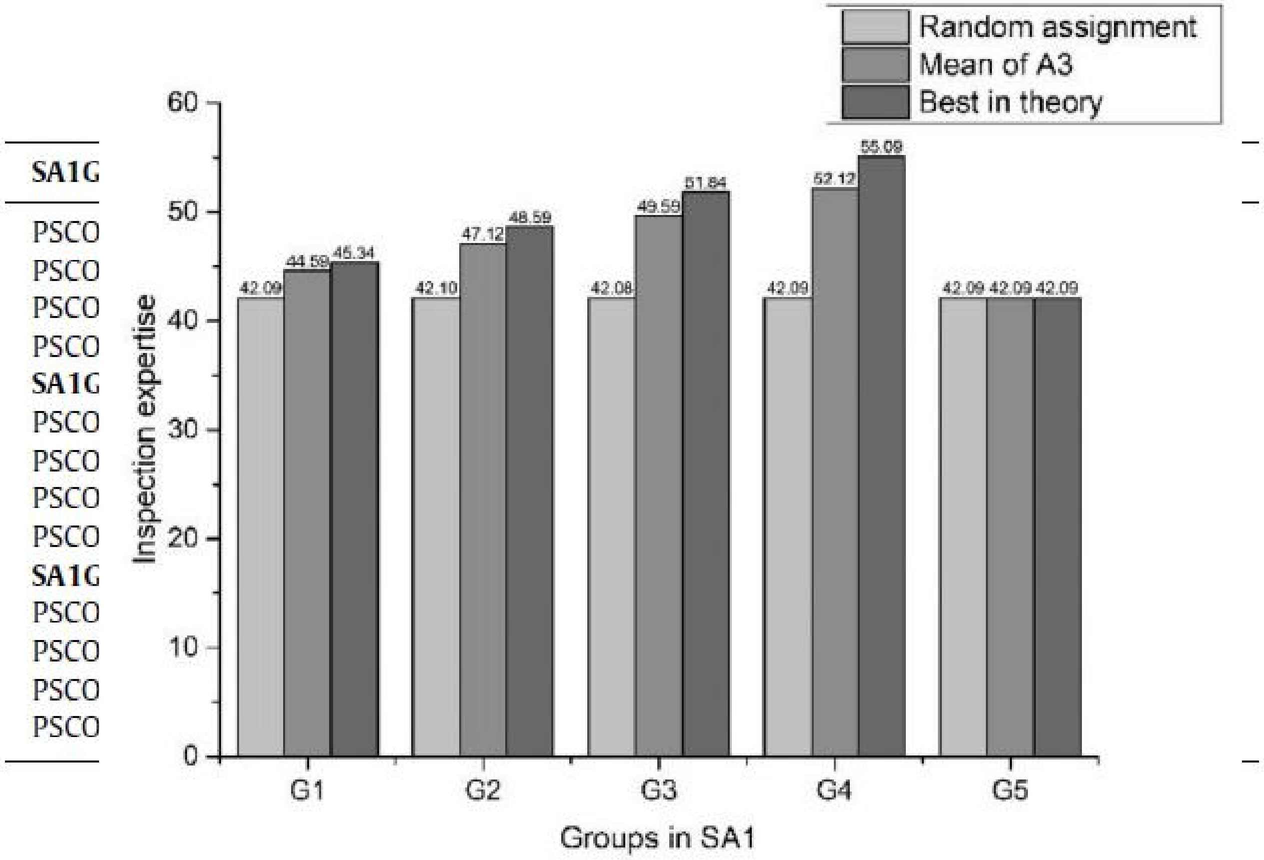
	Min inspection expertise			Max inspection expertise			Variance of inspection expertise		
	A1	A2	A3	A1	A2	A3	A1	A2	A3
Average	40.11	40.05	40.10	41.29	41.27	41.35	0.2924	0.2095	0.2411

- ✓ The performance of all the three PSCO assignment schemes is stable and is much better than the performance of random PSCO assignment
- ✓ Model performance: A3>A1>A2
- ✓ Model variance: A1>A3>A2

3 Computational experiments

Sensitivity analysis

SA1. composition of a group of PSCOs

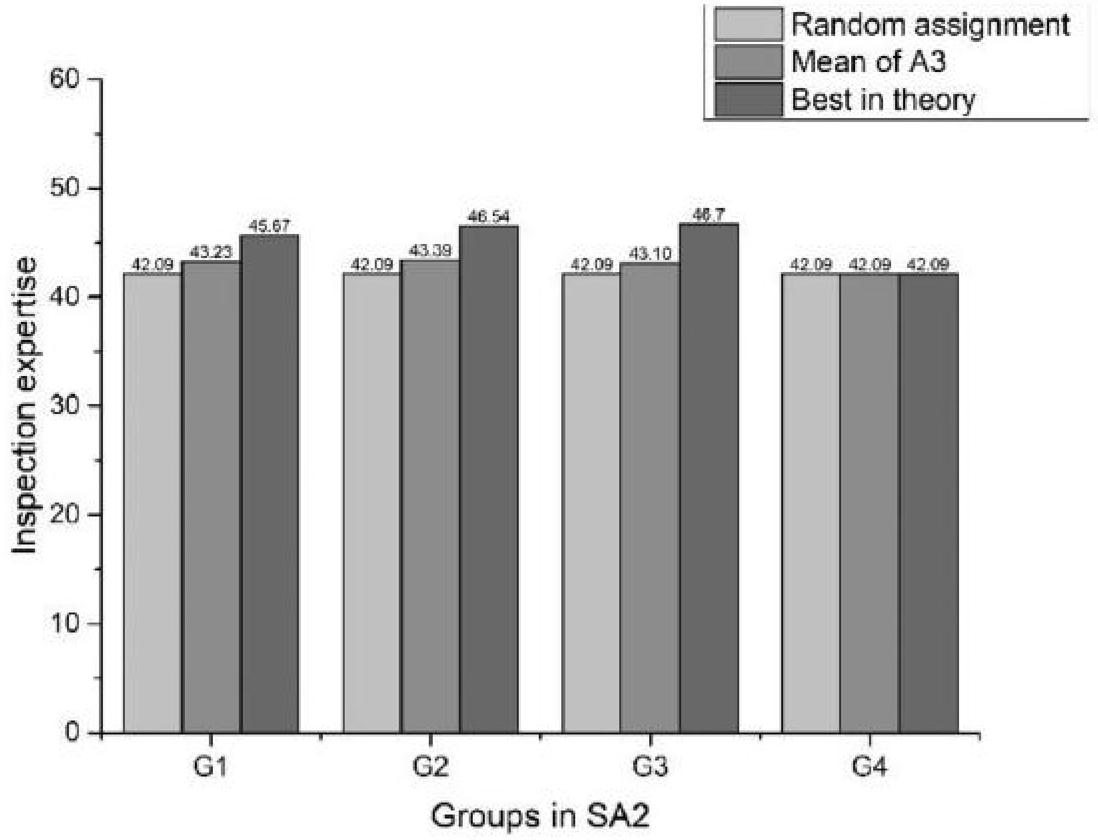


- ✓ As the divergence of expertise of the group of PSCOs become larger, both the best inspection expertise in theory and the mean inspection expertise achieved by using A3 increase
- ✓ The superiority of the PSCO assignment scheme generated by A3 over random PSCO assignment becomes more obvious when the inspection expertise of the PSCOs gets more diverse.
- ✓ The mean inspection expertise achieved by A3 is equal to the best inspection expertise in theory when all the PSCOs have the same expertise.

3 Computational experiments

Sensitivity analysis

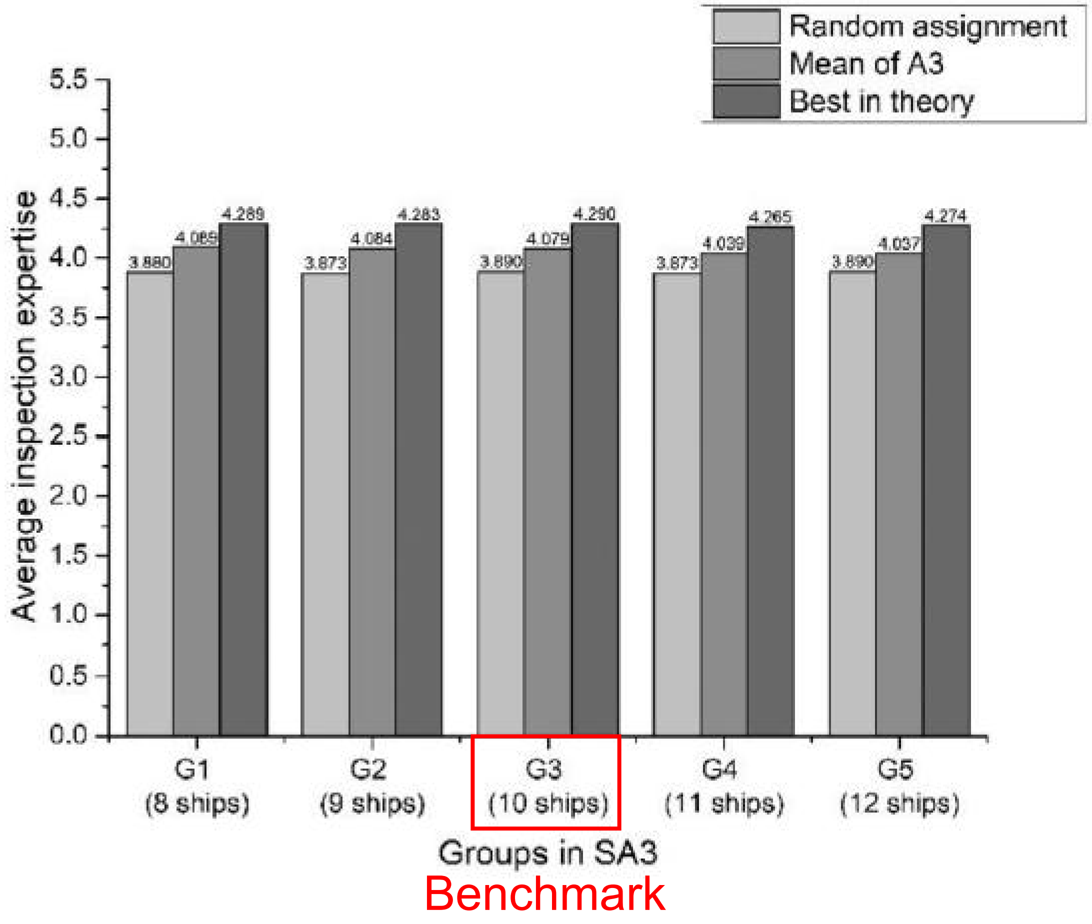
SA2. divergence in expertise of a PSCO



- ✓ The best inspection expertise in theory shows gentle increase when the divergence of the PSCOs' expertise increases.
- ✓ Due to randomness, when the variations in the expertise of the PSCOs increase, the predicted mean inspection expertise can either increase or decrease modestly.

Sensitivity analysis

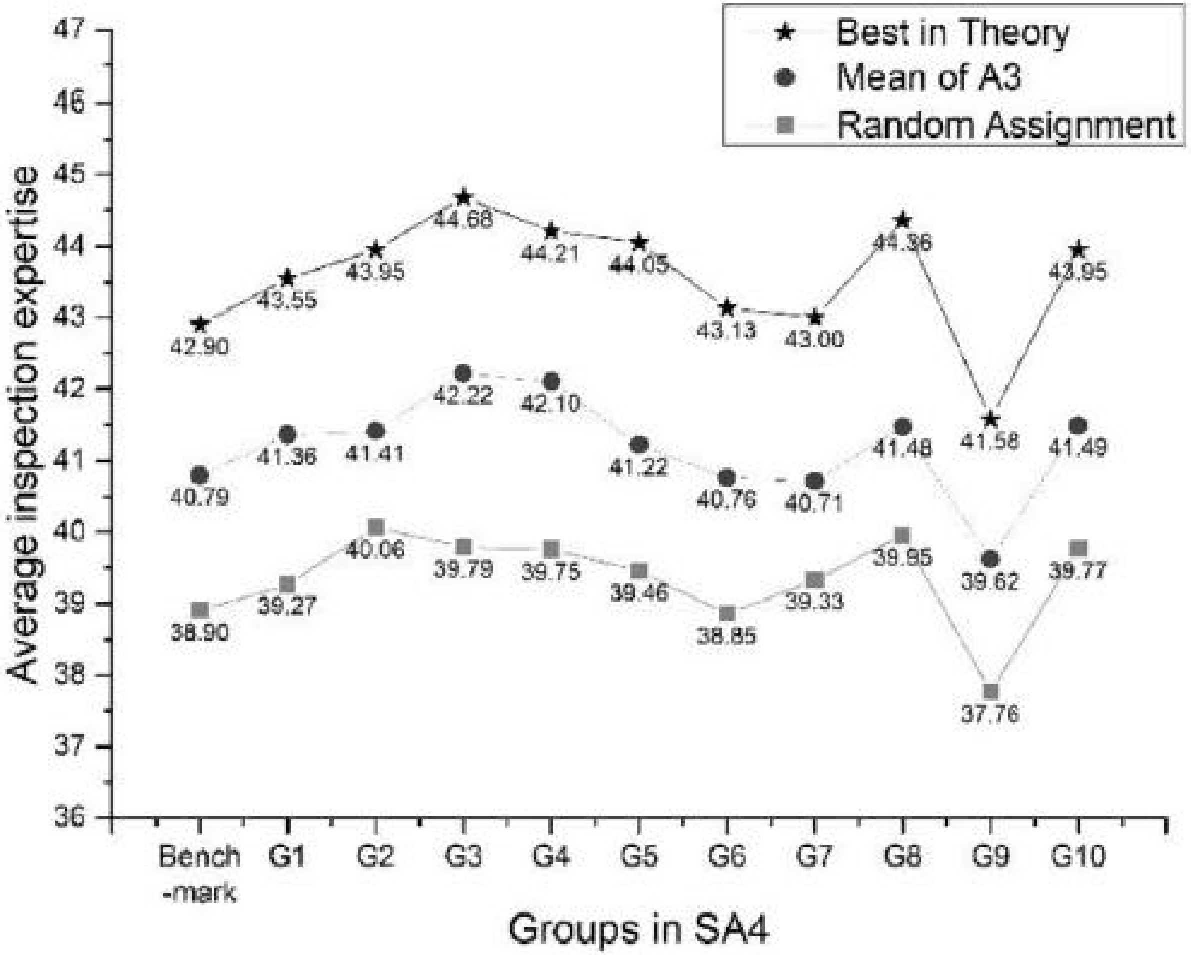
SA3. adequacy of PSCO resources



- ✓ The average inspection expertise of one ship remains stable when the number of ships in a group grows while the maximum number of ships can be inspected by one PSCO remain unchanged.

3 Computational experiments

Sensitivity analysis



- ✓ The differences of the best inspection expertise in theory compared to benchmark is between −3.08% and + 4.15% , which are much smaller than ±10%.
- ✓ The differences between the benchmark and the 10 groups range from −2.87% to + 3.51% , which are much smaller than ±10%.
- ✓ The proposed model can identify about 94.5% of the total deficiencies and is always better than random assignment.

4. Conclusion

4 Conclusion

1

Develop three machine learning models to **predict ship deficiency condition**: MTR-RF1, MTR-RF2, MTR-RF3.

Prediction model

2

Propose two PSCO assignment models M1 and M2 to **match the diverse ship deficiency conditions with the expertise of PSCOs**.

PSCO assignment

3

Compared with random assignment of PSCOs, the proposed models can help to detect **4.70%, 4.55%, and 4.86%** more deficiencies after inspecting the same groups of ships by using the same PSCO resources.

Model performance

4

The gap between the proposed prediction models and the best prediction model in theory is **about 5%**.

Model performance

Thank you for your attention!

Q&A

Deficiency number/PSCO	PSCO 1	PSCO 2	PSCO 3	PSCO 4
Real	6	7	8	9
Predicted (split 1)	8	8	10	11
Overestimate (split 1)	2	1	2	2
MSO (split 1)	$(2-1)^2+(2-2)^2+(2-2)^2+(1-2)^2+(1-2)^2+(2-2)^2=3$			
Predicted (split 2)	8	10	11	10
Overestimate (split 2)	2	3	3	1
MSO (split 2)	$(2-3)^2+(2-3)^2+(2-1)^2+(3-3)^2+(3-1)^2+(3-1)^2=11$			