Defence Research and Development Canada | Recherche et développement pour la défense Canada

Use of Agent-based modeling and data farming for the army ISR capability assessment

Maude Amyot-Bourgeois, Lynne Serré, Peter Dobias

email: maude.amyot-bourgeois@forces.gc.ca

NATO OR&A Conference,

October 2020

DRDC | RDDC

Canada NATO UNCLASSIFIED

Presentation outline

- **Intelligence, Surveillance, and Reconnaissance (ISR) Modernization (MOD) 2020:** project description
- **Proposed study orientation:** using data farming and advanced data analytics/machine learning to assess Land ISR options
- **Methodology**
  - Scenario context and model development in MANA
  - Design of the experiment
  - Metrics for assessment: measures of effectiveness and performance (MOEs/MOPs)
- **Analysis**
  - Description of machine learning methods used
    - *k*-nearest neighbour (KNN) and random forests (RF)
  - Performance of the methods and variable importance
- **Conclusion**

ISR MOD 2020: Project description

- **Land ISR MOD objectives:**
  - Digitized C2 System: connected network, sensor-to-effector linkage
  - Upgraded existing sensors: keep fleet baseline if meets requirement
  - Procurement of new sensors: incorporate new assets
  - Modernized sensor fleet
- **Sensor fleet purpose and goal:**
  - Intelligence, Surveillance and Reconnaissance in support of CA missions
  - Capability to detect, recognize, identify, track, locate targets
  - Near-real time situational awareness (SA)
  - Flexible, mobile, scalable fleet

Proposed study orientation

Using data farming and data analytics methods to assess Land ISR options

What is Data Farming?

- Generation of a large volume of data
- Broad exploration of the parametric space
- Simple and abstract model
- Data points generated represent various outcomes that can be analyzed to:
  - Predict trends
  - Identify, explain outliers
  - Measure impact from different factors
- Methodology is divided into six steps

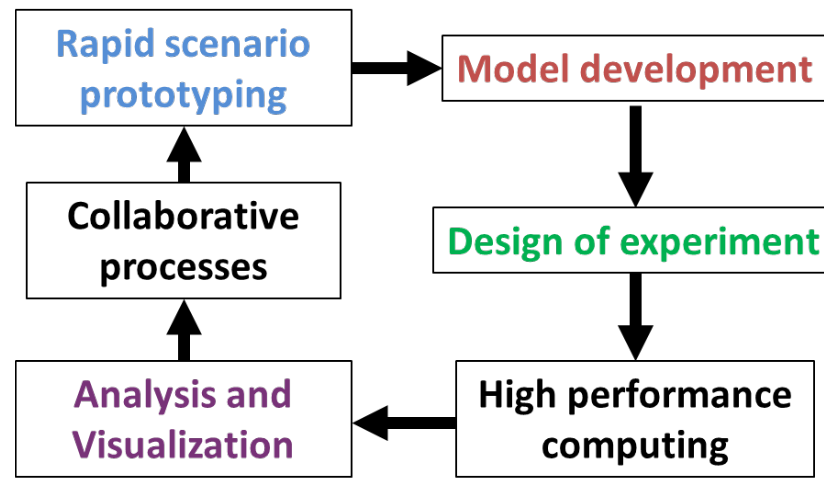

Figure 1. Six steps of the data farming process

Proposed study orientation

Using data farming and data analytics methods to assess Land ISR options

- **Scenario**: taken from the Canadian Army scenario vignettes
- **Model development**: Using agent-based model (MANA)
  - Simplify the scenario (terrain, players, players' goals)
  - Build the ISR architecture(s)
  - Select the metrics for assessment (MOEs/MOPs)
- **Design of experiment:**
  - Set the scope of the parametric option space to explore
- **Analysis and Visualization:**
  - Assess two methods (KNN and RF) for predicting MOEs/MOPs
  - Identify the key sensor parameters

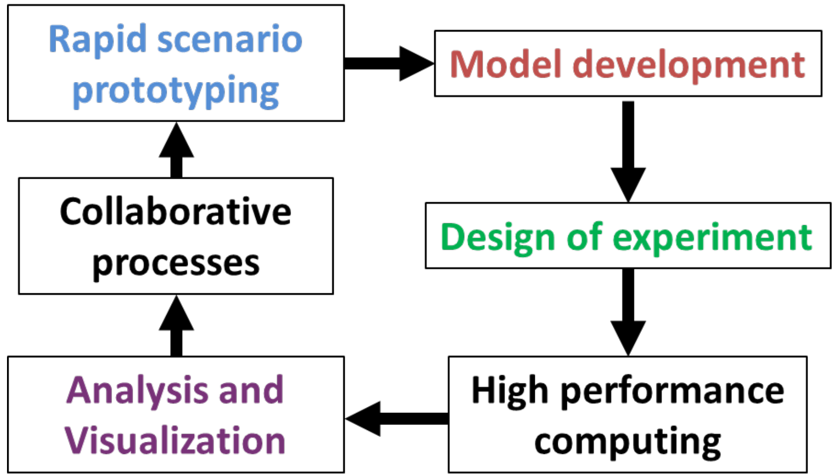

Figure 2. Six steps of the data farming process

Methodology: Scenario context and model development in MANA

- Fourth core mission in Canada's 2017 Defence Policy: lead or contribute to international peace operations and stabilization missions
- Shield at the border against invading neighbour country
- Engage if border crossing of Red team detected

Table 1: Entities description

Squad ID	Description	Population
1	Blue HQ	1
2	Red tanks company	10
3	Blue ground sensor border	1-2
4	Blue rotary wing aircraft	2
5	Blue tanks company	2-4

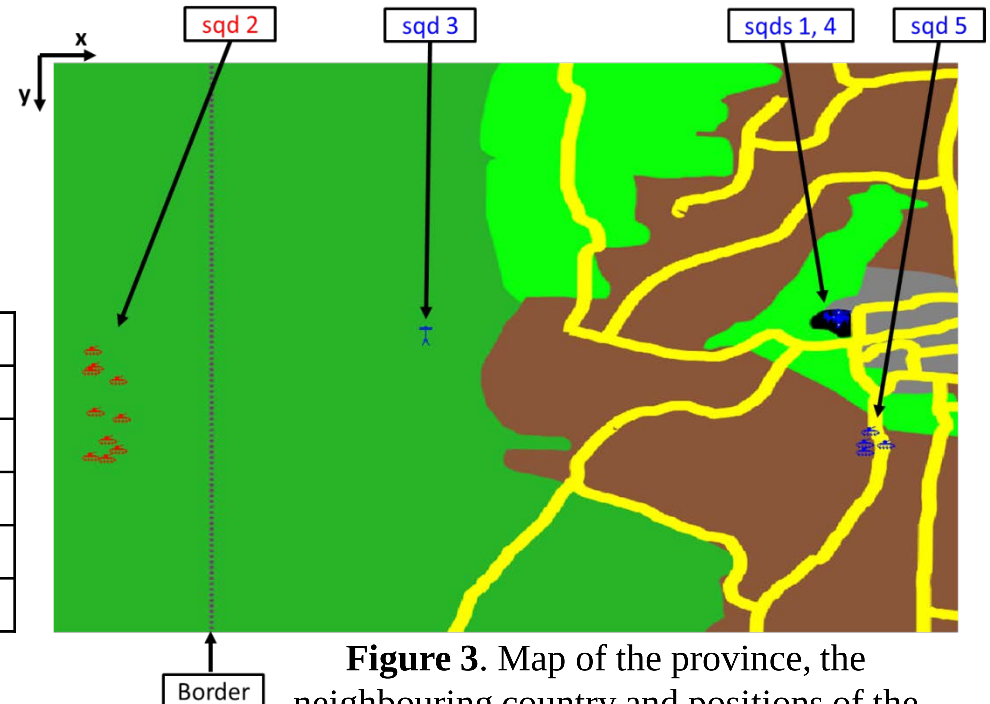

Figure 3. Map of the province, the neighbouring country and positions of the different squads

Methodology: Design of the experiment

- **Classification range and classification probability**: sensor performance metrics
- **Sensor Stealth**: source protection for continuous operability and protection of sensitive technology
- **Number of sensors**: continual and complementary coverage, redundancy
- **Sample size**: 10 iterations (threshold between sample size and simulation run time)

Table 2: Scope of the parameterizing of the variables of interest

Parameter	Unit	Range	Increment
Classification range	grids	10-200	10
Classification probability	-	0.1-1	0.1
Stealth	%	0-100	10
Number of sensors	-	1-2	1

Methodology: Metrics for assessment

Table 3: Description of the MOEs and MOPs selected for the assessment

MOEs/MOPs (Continuous)	**Units**	
First detection step (time until first Red tank detected)	Steps (time units)	
Mean first detection step (across all Red tanks)	Steps (time units)	
Mean detection range (across all Red tanks)	Grids (distance unit)	
Percentage of detected Red tanks	Ratio (detected/Red tanks initial number)	
MOEs (Categorical)	**Outcome at the end of the simulation**	**Score**
Blue mission success	Red didn't reach waypoint and all Red incapacitated	3
	Red reached waypoint and 70% or more Red incapacitated	2
	Red reached waypoint and less than 70% Red incapacitated	1
Blue RCS	Blue has 70% or more combat effectiveness	1
	Blue has less than 70% combat effectiveness	0

Analysis: Description of machine learning methods

- ***k*-Nearest-Neighbour (KNN):**
  - Memory-based method
  - Classify or predict test observations based on the majority vote or average value of the *k* closest observations in a training set
- **Random Forest (RF):**
  - A large set of de-correlated trees built by bootstrapping the training data
    - Tree split: only a random subset of the feature variables is considered
  - Classify or predict test observations by majority vote or averaging the predictions of the trees
- **Feature variables are the sensor parameters**
  - Classification range, classification probability, sensor stealth, number of sensors
- **Target (predicted) variables are the MOEs/MOPs**
  - Continuous target = regression problem
  - Categorical target = classification problem

Analysis: Performance of RF and KNN methods for predicting the MOEs/MOPs

- Regressors:
  - RF gives smaller MSE than KNN, indicating that RF performs better at predicting the metrics
  - Regressors with R^2>0.95: strong ability to predict unseen observations
  - R^2 = 0.68: weaker ability, but still outperforms a constant model predicting the expected value (R^2 = 0)
- Classifiers:
  - Accuracy > Balanced accuracy: class imbalance
  - Balanced accuracy > 1/2 or 1/3: mildly correctly classifies test observations
  - 0 < Cohen's Kappa < 0.1: slight agreement between predicted and true classes

Table 4: Performance of the RF and KNN methods

MOE/MOP	Metric	RF	KNN
First detection step (time until first Red tank detected)	MSE	66765.83	72154.24
	R^2	0.97	0.97
Blue mission success	Accuracy	58%	57%
	Balanced accuracy	41%	40%
	Cohen's kappa	0.07	0.06

Analysis: Variable importance for RF and KNN

- Classification range is the most important variable for regressors and Blue mission success
- Classification range and probability are the most important variables for the Blue RCS

Table 5: Variable importance for RF and KNN

	First detection step*	Blue mission success**
RF		
Classification range	1.9363 (0.0113)	0.0748 (0.0023)
Classification probability	0.0026 (0.0001)	0.0045 (0.0018)
Sensor stealth	0.0012 (0.0001)	0.0013 (0.0021)
Number of sensors	0.0011 (0.0001)	-0.0015 (0.0011)
KNN		
Classification range	1.9022 (0.0181)	0.0611 (0.0028)
Classification probability	0.0023 (0.0001)	0.0046 (0.0031)
Sensor stealth	0.0011 (0.0002)	0.004 (0.0039)
Number of sensors	0.0009 (0.0002)	0.001 (0.0023)

*Variable importance shown as mean decrease in R^2 (standard deviation)

**Variable importance shown as mean decrease in balanced accuracy score (standard deviation)

Analysis: Variable importance for RF and KNN

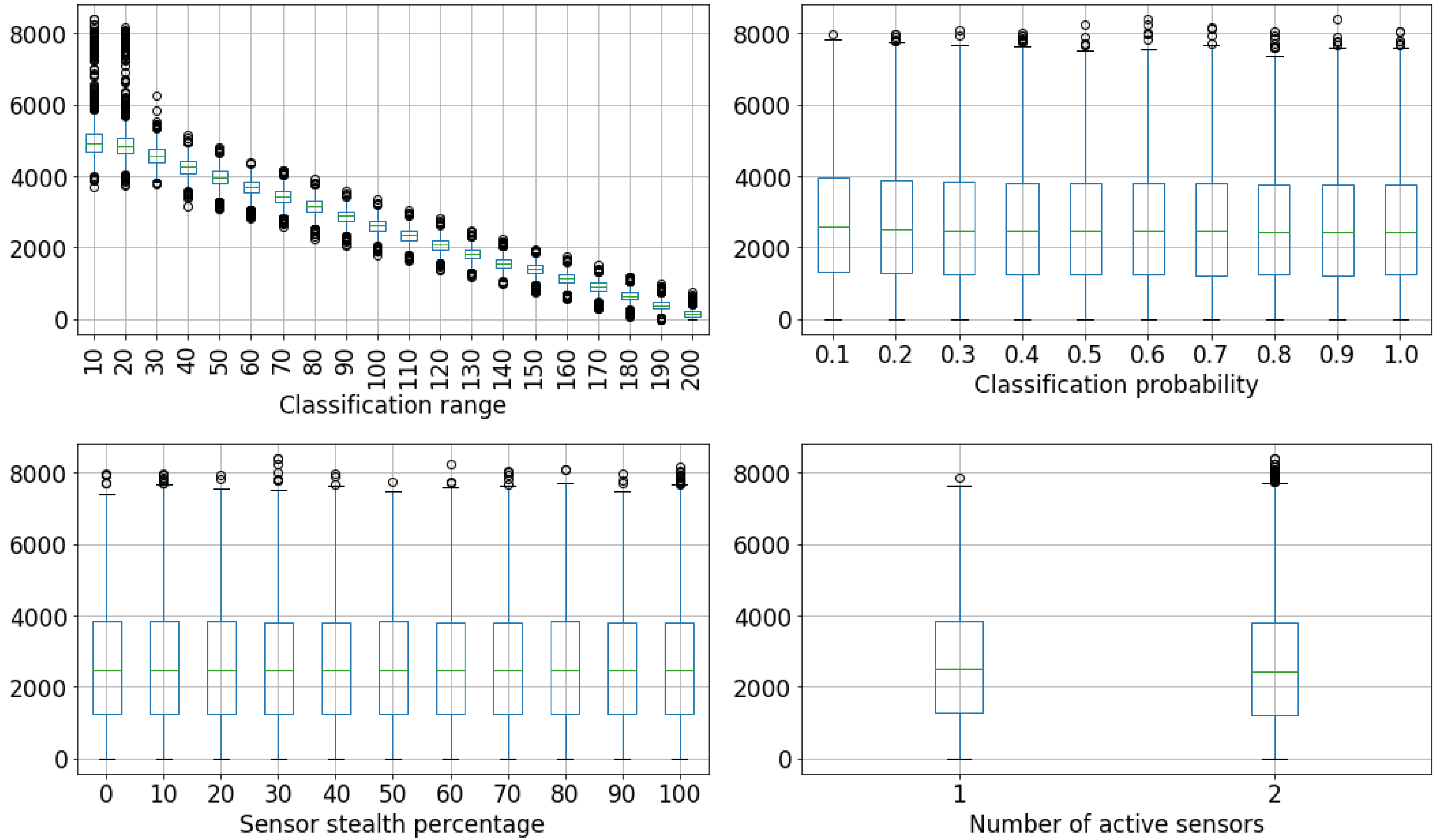

Figure 4. Distribution of the first detection step for each sensor parameter

Analysis: Variable importance for RF and KNN

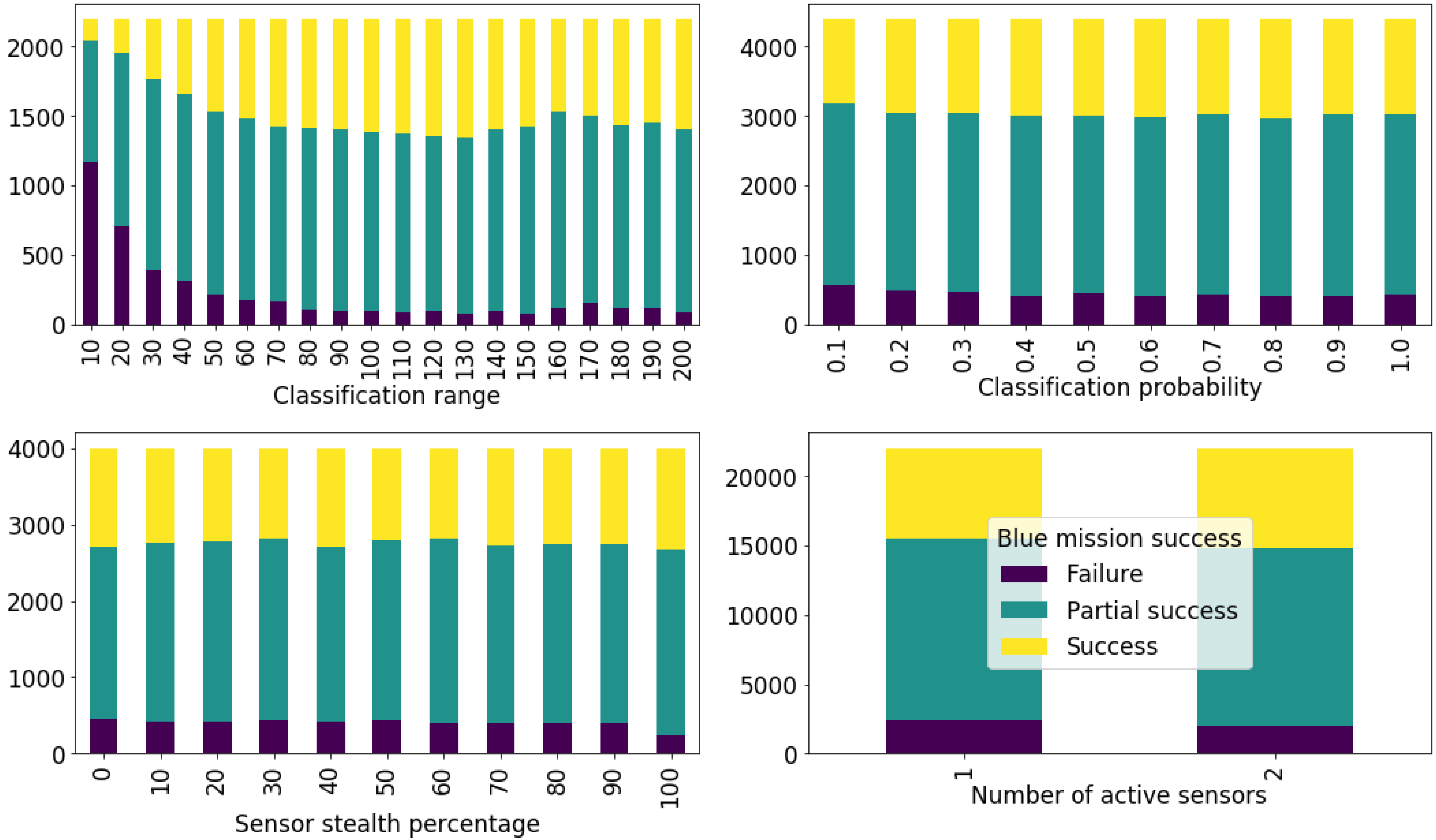

Figure 5. Distribution of the Blue mission success for each sensor parameter

Conclusion

- MANA successfully replicated the battlefield dynamics of an ISR task
- Promising results were found when using machine learning methods (KNN and RF) to establish patterns in the farmed data
- The process of data farming using MANA in combination with machine learning can potentially support ISR capability mix analysis
- Future work:
  - More comprehensive scenarios
  - More complex Red force
  - More complex Blue sensor mixes
  - Diversified scenario types
  - Use Cora HPC to shorten simulation time

Thank you for your time

email: maude.amyot-bourgeois@forces.gc.ca

DRDC | RDDC

SCIENCE, TECHNOLOGY AND KNOWLEDGE
FOR CANADA'S DEFENCE AND SECURITY

SCIENCE, TECHNOLOGIE ET SAVOIR
POUR LA DÉFENSE ET LA SÉCURITÉ DU CANADA