

DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

PUBLICATION

EXAM-TM50G-PRE



PRE-TEST — SL 5G: ADVANCED ORSA

Maven Smart System (MSS) — USAREUR-AF

HEADQUARTERS
UNITED STATES ARMY EUROPE AND AFRICA
(USAREUR-AF)
Wiesbaden, Germany

DRAFT — NOT FOR OFFICIAL USE. FOR TRAINING PLANNING PURPOSES ONLY.

26 MARCH 2026

DRAFT — UNOFFICIAL — NOT FOR OPERATIONAL USE

PRE-TEST — SL 5G: ADVANCED ORSA

MAVEN SMART SYSTEM (MSS) — USAREUR-AF

Field	Detail
Course	SL 5G: Advanced ORSA
Form	Pre-Test
Level	SL 5G (Advanced Specialist)
Audience	FA49 / senior ORSA analysts; prerequisite: SL 4G + 18 months ORSA experience + graduate quantitative methods
Time Allowed	30 minutes
Passing Score	N/A — diagnostic only

INSTRUCTIONS

This diagnostic assessment establishes your baseline knowledge before training. Your score does not affect course eligibility. Answer honestly — results help the instructor tailor instruction to gaps.

SECTION 1 — MULTIPLE CHOICE

Circle the letter of the best answer. (2 points each)

1. In nonlinear programming, a key challenge that does not exist in linear programming is:

A. Defining an objective function
 B. The feasible region may be non-convex, and local optima may not be global optima
 C. Sensitivity analysis is not possible
 D. The solution requires integer decision variables

2. A "Pareto frontier" in multi-objective optimization represents:

- A. The single optimal solution that maximizes all objectives simultaneously
- B. The feasible region boundary defined by all binding constraints
- C. The set of solutions where no objective can be improved without degrading another
- D. The optimal solution under the most pessimistic assumptions

3. In agent-based modeling (ABM), "emergent behavior" refers to:

- A. Complex system-level patterns that arise from simple agent interactions, not explicitly programmed
- B. Behavior that was programmed into individual agents as a decision rule
- C. Behaviors that only appear when the model is run for more than 1,000 time steps
- D. The calibration result when agent rules are fitted to historical data

4. "Calibration" of an agent-based model involves:

- A. Setting the number of agents to match the population size
- B. Reducing computational complexity by simplifying agent decision rules
- C. Validating the model against a theoretical analytical solution
- D. Adjusting agent rules and parameters until the model reproduces observed historical patterns

5. In Bayesian analysis, "prior probability" represents:

- A. The probability calculated from the data before the model is fitted
- B. The analyst's initial belief about a parameter before observing the current data
- C. The probability assigned to the null hypothesis in a frequentist test
- D. The marginal distribution of the observed data

6. "Causal inference" differs from correlation analysis because it:

- A. Identifies variables that cause an outcome, not just those associated with it, enabling prediction of intervention effects
- B. Uses larger datasets to achieve statistical significance
- C. Requires an experimental design with randomized control and treatment groups
- D. Applies Bayesian methods rather than frequentist methods

7. A "DAG" (Directed Acyclic Graph) in causal modeling is used to:

- A. Represent the assumed causal structure between variables, identifying confounders and mediators
- B. Visualize the network of agent interactions in an ABM
- C. Display the dependency graph of a data pipeline
- D. Plot the posterior distribution of Bayesian model parameters

8. "Decision analysis under deep uncertainty" (DAUX) is appropriate when:

- A. The analyst has high-quality data and a well-specified probability model
- B. The commander requires a single recommended course of action
- C. The decision requires Monte Carlo simulation with more than 10,000 trials
- D. Key uncertainties cannot be meaningfully quantified with a single probability distribution

9. "Robustness analysis" in a DAUX context means:

- A. Testing the model's code against adversarial inputs
- B. Identifying strategies that perform acceptably across a wide range of scenarios, not just the best-case or expected-case scenario
- C. Measuring the model's sensitivity to changes in individual parameter values
- D. Validating that the model reproduces historical data within acceptable error bounds

10. An "ensemble forecasting" approach combines multiple forecasts by:

A. Selecting the single best model from a set of candidate models
B. Averaging or otherwise aggregating predictions from multiple models to reduce variance and improve overall accuracy
C. Running the same model multiple times with different random seeds
D. Stacking models sequentially where each model corrects the errors of the previous

11. A "hierarchical model" in forecasting is used when:

A. The data has multiple classification levels that require separate models
B. The forecast horizon exceeds 12 months and requires decomposition
C. Data has a nested structure (e.g., units within brigades within divisions) and estimates at each level should be informed by data at other levels
D. The model has more parameters than observations

12. Briefing a 3-star (LTG) or GO/SES audience requires that analytical findings be:

A. Presented with all technical details so the audience can validate the methodology
B. Limited to a single chart with no more than three data points
C. Reviewed and approved by the G3 before the ORSA presents directly to the GO
D. Presented at the strategic implications level — key finding, risk characterization, and decision required — with technical details in backup slides

13. In Army multi-echelon wargame support, an ORSA contributes most directly by:

A. Providing analytical products that quantify COA outcomes, adjudicate ambiguous results with modeling, and assess campaign-level risk
B. Running the wargame adjudication process as the primary controller
C. Briefing the commander on the wargame results after all moves are adjudicated
D. Serving as the opposing force (OPFOR) commander during the wargame

14. An "analytical product governance" requirement for a SL 5G level product means:

A. Only products with more than 1,000 data records require governance review
B. Products are classified at the SECRET level by default
C. All products must be reviewed by a peer ORSA before delivery to the commander
D. Products must be archived in GFEBS for audit purposes

15. The most significant difference between SL 4G and SL 5G ORSA work is:

A. SL 5G uses Python; SL 4G uses R
B. SL 5G requires a higher random seed value in Monte Carlo simulations
C. SL 5G addresses problems with deep uncertainty, system-level complexity, and GO/SES briefings where SL 4G techniques are insufficient
D. SL 5G involves a larger dataset size

SECTION 2 — SHORT ANSWER

Answer in 2–5 sentences. (6 points each)

SA-1. Explain the "backdoor criterion" in causal inference. Give a military example where failing to adjust for a backdoor path would lead to an incorrect analytical conclusion.

SA-2. A division G3 asks for a COA analysis under conditions where the probability of a key intelligence variable cannot be meaningfully estimated. Describe why standard expected-value analysis is inappropriate and what DAUX approach you would use instead.

SA-3. Describe the calibration challenge in building an ABM of unit logistics behavior, including what data you would use to calibrate agent rules and how you would validate the calibrated model.

SA-4. A GO asks for a "best estimate" of adversary force reconstitution time. You have significant uncertainty in the key parameters. Describe how you would construct and present the estimate to meet SL 5G GO/SES briefing standards.

SA-5. Describe the "peer review requirement" for SL 5G analytical products and explain why this requirement exists at the advanced ORSA level but is not formalized at SL 4G.

SCORING SUMMARY

Section	Questions	Points Each	Total Points
Multiple Choice	15	2	30
Short Answer	5	6	30
Total	—	—	60

Passing: N/A — Pre-test is diagnostic only.

ANSWER KEY — INSTRUCTOR USE ONLY

Do not distribute to students.

Multiple Choice: 1. B — Non-convex feasible regions mean local optima may not be global optima — key NLP challenge. 2. C — Pareto frontier = solutions where no objective improves without degrading another. 3. A — Emergent behavior = complex patterns arising from simple agent interactions, not programmed. 4. D — Calibration = adjusting agent parameters to reproduce historical observed patterns. 5. B — Prior probability = analyst's initial belief before observing current data. 6. A — Causal inference identifies what causes an outcome, enabling prediction of intervention effects. 7. A — DAG represents assumed causal structure, identifying confounders and mediators. 8. D — DAUX is for situations where uncertainty cannot be meaningfully quantified with a single probability model. 9. B — Robustness analysis finds strategies that perform acceptably across a wide range of scenarios. 10. B — Ensemble forecasting aggregates multiple model predictions to reduce variance. 11. C — Hierarchical models handle nested data structures where estimates at each level inform other levels. 12. D — GO/SES briefings focus on strategic implications, key finding, and decision required. 13. A — ORSA quantifies COA outcomes, adjudicates ambiguous results, and assesses campaign-level risk. 14. C — SL 5G products require peer review by another ORSA before commander delivery. 15. C — SL 5G addresses deep uncertainty, system complexity, and GO/SES communication gaps that SL 4G methods cannot handle.

Short Answer Guidance:

SA-1. Full credit: backdoor criterion — a backdoor path is an open non-causal path from treatment to outcome that passes through a common cause (confounder); failing to block it produces a confounded estimate; military example: training hours → unit readiness, but experienced units both train more and perform better — without adjusting for unit experience level, the estimated effect of training is confounded by baseline capability. Partial credit (3 pts) for correct definition without a military example.

SA-2. Full credit: standard expected-value analysis requires a probability distribution over the uncertain variable — if that distribution cannot be meaningfully specified, any expected-value result is arbitrary; use a scenario-based robustness approach: define a set of plausible scenarios (not a probability-weighted distribution), evaluate each COA across all scenarios, and identify which COA performs acceptably across the most scenarios (robustness) rather than which maximizes expected value; brief the GO on the scenario space and which COA is most robust. Partial credit (3 pts) for identifying expected value as inappropriate without the DAUX alternative.

SA-3. Full credit: calibration challenge — logistics behavior involves human decision-making (prioritization under uncertainty), which is hard to parameterize; data sources: historical unit supply request logs, distribution timing records, maintenance scheduling data; calibration method: adjust agent priority rules and response time parameters until the model reproduces observed logistics cycle times and demand patterns within a defined error tolerance; validation: test model predictions against a held-out period not used for calibration. Partial credit (3 pts) for calibration approach without data source or validation method.

SA-4. Full credit: construct the estimate using a range of assumptions for the uncertain parameters; present as a distribution or scenario range (e.g., 15–45 days depending on key assumptions); identify the most sensitive parameter (sensitivity analysis); present to GO as: "Under assumption A: 15 days. Under assumption B: 45 days. The key uncertainty is [parameter]. Our recommended planning range is [X–Y] days." Do not present a single point estimate; state assumptions explicitly; provide backup slides with methodology. Full credit requires scenario range + key uncertainty identification + GO briefing format.

SA-5. Full credit: peer review requirement = all SL 5G products must be reviewed by at least one other SL 5G qualified ORSA before delivery to the GO/SES; this exists at SL 5G because products influence strategic and operational decisions at the highest command levels — an undetected methodological error could produce catastrophically incorrect analytical conclusions; at SL 4G the stakes are lower (battalion/brigade level) and peer review is a best practice; at SL 5G the analytical complexity (nonlinear optimization, causal inference, ABMS) creates greater potential for subtle errors that only a peer can detect. Partial credit (3 pts) for the requirement without explaining why it escalates at SL 5G.

USAREUR-AF Operational Data Team TM-50G Pre-Test | Version 1.0 | March 2026