



Received: 2026/09/04
Revised: 2026/09/12
Accepted: 2026/09/29
Published: 2026/09/30

***Corresponding Author:**

Kyungwon Oh
E-mail: kwoh@hnu.kr

Abstract

본 연구는 전영역 인식과 JADC2 환경에서 지휘관에게 미래 임무 여건의 확률과 근거를 제공하는 AI 의사결정 지원모델 QFSVDM을 제안한다. 밀도행렬과 비가환 갱신으로 순차 정보를 표현하되 물리적 양자계를 가정하지 않으며, 대각 층분조건에서 부호 자유 고전 필터로 정확히 환원된다. 원천·수신·이용 가능시각을 구분하고 지휘관의 수용·수정·보류를 기록한다. 세 합성자료 생성과정의 독립 시드 153개에서 비교한 결과, 비대각 확장은 해당 구조가 있을 때만 고전 필터보다 낮은 표본외 로그손실을 보였으며, 시차 입력 로짓도 이득의 일부를 재현했다. 결과는 합성자료 검증에 한정되며 전장자료 학습이나 작전효과를 입증하지 않는다. 자료 품질·확률 예측·인간-AI 협업·통신 저하 복원성의 단계별 검증체계를 제시한다.

This study proposes the Quantum-like Future-State Vector Decision Model (QFSVDM), an AI decision-support model providing commanders with future mission-condition probabilities and evidence based on all-domain awareness in JADC2 environments. Density matrices and non-commutative updates represent sequential information without assuming a physical quantum battlefield. The model extends a signed classical filter, reduces exactly to it under sufficient diagonal conditions, distinguishes source, receipt, and availability times, and records commanders' acceptance, revision, or deferral. Across three synthetic generating processes and 153 independent seeds, the off-diagonal extension reduced out-of-sample log loss relative to the classical filter only when off-diagonal structure existed. A lagged-input logit reproduced part of this gain. These findings do not establish battlefield-data learning or operational effectiveness. A staged validation framework separates data quality, probabilistic forecasting, human-AI collaboration, and resilience under degraded communications.

Keywords

전영역 인식(All-Domain Awareness),
합동 전영역 지휘통제(JADC2, Joint All-Domain Command and Control),
지휘관 의사결정 지원(Command Decision Support),
인간-AI 협업(Human-AI Collaboration),
양자유사 미래상태벡터 의사결정 모델(QFSVDM,
Quantum-like Future-State Vector Decision Model)

Acknowledgement

이 논문은 2026학년도 한남대학교 학술연구비 지원에 의하여 연구되었음

전영역 인식 기반 지휘관 AI 의사결정 지원모델 QFSVDM: JADC2 환경의 비가환 정보 갱신, 인간 중심 검증체계 및 합성자료 반복 실험

An AI Decision-Support Model for Commanders Based on All-Domain Awareness (QFSVDM): Non-Commutative Information Updates, Human-Centered Validation, and Repeated Synthetic-Data Experiments in a JADC2 Environment

오경원*

한남대학교 기계공학과 교수

Kyungwon Oh*

Professor, Department of Mechanical Engineering, Hannam University

1. 서론

1.1 연구 배경과 문제

미래 전장의 지휘관은 지상·해양·공중·우주·사이버 영역에서 생산되는 정보와 전자기 환경, 기상, 통신 및 지원 여건을 함께 고려해야 한다. 그러나 여러 영역의 데이터를 한 화면에 모았다는 사실만으로 공통된 이해가 형성되지는 않는다. 보고의 기준시각이 다르거나, 서로 다른 체계가 동일한 사건을 독립 정보처럼 전송하거나, 핵심 자료의 부재가 정상 상태로 표시되면 통합 화면은 오히려 잘못된 확신을 만들 수 있다. 본 연구는 이러한 상황에서 AI가 무엇을 계산하고 무엇을 지휘관에게 남겨야 하는지에 주목한다.

미 국방부의 JADC2 전략 요약은 합동 지휘통제의 기능을 감지(sense), 이해(make sense), 행동(act)으로 제시하고, 데이터·인간·기술 기업체계와 임무 파트너 정보공유를 별도의 노력선으로 둔다[1]. NATO의 다영역작전 개념도 모든 작전영역·환경의 군사

활동을 비군사 활동 및 외부 이해관계자와 동기화하는 관점을 취한다[2]. 본 논문은 이러한 체계를 대체하는 설계가 아니라, 그 안에서 불확실한 정보와 지휘관의 판단을 연결하는 연구 모형을 다룬다.

핵심 연구 질문은 네 가지다. 상이한 시각과 품질의 정보를 같은 판단에 반영하려면 어떤 상태와 기록이 필요한가. 정보의 순서와 상충을 보존하는 모형이 더 단순한 고전 필터보다 추가 설명력을 갖는가, 그리고 그것을 어떻게 공정하게 검사하는가. AI의 확률 출력이 지휘관의 적절한 신뢰와 판단 유보에 기여하는지 어떻게 검증하는가. 통신 저하·정정·모형 갱신이 발생해도 과거 판단을 재현할 수 있는가.

1.2 목적, 범위와 기여

본 연구의 목적은 전영역 인식에 기반한 지휘관 AI 의사결정 지원모델의 개념 구조, 수학적 상태 갱신, 자료 생산 및 검증 절차를 제시하고, 그 가운데 계산으로 검사할 수 있는 부분을 실제로 검사하는 것이다. 분석 단위는 지휘소의 상황평가와 판단 지원이며, 예시는 정보·통신·지원의 가용성 질문에 한정한다. 표적 선정, 무기 할당, 교전 실행 알고리즘은 검증 대상이 아니다. 제시하는 확률과 합성 임계값은 작전

기준이나 실제 성능이 아니다.

기여는 다섯 가지다. 첫째, 전영역 정보·상황인식·미래 여건 판단·지휘관 선택을 구분한 구조와 QFSVDM의 명시적 정의(3장). 둘째, 부호 자유 고전 기준모형을 내포하는 비가환 확률 필터, 대각 충분조건 아래의 정확한 환원, 회전 게이지와 켈레 대칭에 따른 식별 가능 불변량(5장). 셋째, 이용 가능시각·정정 이력·접근 권한을 포함하는 재현 가능한 판단 기록(7장). 넷째, 다섯 추상 입력과 세 상태를 갖는 합성자료의 독립 시드 153개에서 고전 필터·비대각 확장·고전 대안을 같은 조건으로 적합한 반복 실험(9장). 다섯째, 수학적 정합성·예측 성능·인간-AI 협업·체계 복원성을 구별하는 단계별 검증 설계(10장).

여기서 저자가 제시하는 QFSVDM(quantum-like future-state vector decision model, 양자유사 미래 상태벡터 의사결정 모델)은 서로 다른 시각·품질로 순차 도착하는 정보를 밀도행렬로 표현된 내부 상태에 비가환적으로 갱신하여, 지정된 임무 질문에 대한 미래 여건의 확률과 근거를 지휘관에게 제공하는 의사결정 지원 모델이다. 부호 자유 고전 확률 필터를 특수 경우로 포함하며, 전장을 물리적 양자계로 가정하지 않고 연산자 형식만 차용한다. 최종 판단 권한은 지휘관에게 있다.



Fig. 1. Overview of all-domain awareness (author's illustration, AI-assisted)

Table 1. Sources of all-domain information and the context to be recorded with them (research classification example)

Category	Abstract input example	Context to record
Land	Field status and support reports	Observation coverage, report time, source identifier
Maritime	Sea state, weather, facility status	Spatial reference, observation gaps, data source
Air	Airspace, environment, observation availability	Update cycle, validity period, latency
Space	Observation, positioning, timing service availability	Service dependency, accessible range
Cyber	Information-system and network status	Integrity, access rights, service impact
Cross-domain	Electromagnetic, communications, weather context	Extent of effect, uncertainty, validity time

2. 전영역 인식과 관련 연구

2.1 전영역 인식의 정의

본 논문에서 전영역 인식은 여러 영역의 관측을 임무와 시간의 맥락에서 연결하여 현재 여건을 이해하고 미래 여건의 불확실성을 판단하는 과정이다. 이 용어는 특정 국가의 단일 사업명이 아니다. 상황인식을 지각·이해·미래 투영으로 구분한 Endsley의 이론은 데이터 확보, 의미 이해, 미래 판단을 구별하는 출발점을 제공한다[3]. 다만 AI의 계산 상태가 지휘관의 실제 인지 상태를 직접 측정한다고 가정하지 않는다.

육·해·공·우주·사이버는 정보의 출처와 맥락을 분류하는 축이고, 전자기·정보 환경은 여러 영역에 걸친다. 이를 곧바로 서로 배타적인 확률 상태로 만들지 않는다. 공통작전상황도에 어떤 객체가 표시되었다는 사실, 그 정보가 최신이라는 판단, 그 정보가 특정 임무에 충분하다는 판단은 서로 다른 명제이며, 본 연구는 세 번째 명제의 불확실성을 다루기 전에 앞의 두 명제의 출처와 품질을 확인하도록 설계한다.

한 원천이 여러 행에 관련될 수 있으므로 영역별 집계 수를 독립 정보원 수로 해석하지 않는다.

2.2 JADC2, 책임 있는 AI와 판단 지원

JADC2는 합동·연합의 지휘통제와 정보공유를 연결하는 상위 맥락이다[1]. QFSVDM은 그 안에서 확보된 관측을 받아 특정 임무 질문의 확률 판단을 제공하는 선택 가능한 구성요소이며, JADC2 연결 자체가 입력의 진실성이나 모형의 정확성을 보증하지 않는다. 미 국방부 CDAO의 2023년 데이터·분석·AI 도입 전략은 양질의 데이터, 분석, 책임 있는 AI를 연결하고 성과 기반의 지속적 평가와 피드백을 강조한다[4]. NATO의 2024년 개정 AI 전략은 적법성, 책임,

책무성, 설명가능성·추적가능성, 신뢰성, 통제가능성, 편향 완화의 여섯 원칙을 재확인하며[5], NIST AI RMF 1.0은 타당성·신뢰성·설명가능성과 위험 관리를 사용 맥락에 맞게 평가하도록 제시한다[6]. 본 논문은 이를 원천 연결, 모형 버전, 판단 이력, 인간의 수정 사유, 기능 중단 가능성이라는 연구 요건으로 구체화한다. 이러한 요건을 갖추는 것과 실제 체계의 인증을 받는 것은 구분한다.

2.3 인간-AI 협업과 신뢰

AI 지원의 목적은 지휘관이 AI를 최대한 따르게 만드는 것이 아니다. 자동화에 대한 신뢰는 체계의 실제 능력에 맞아야 한다[7]. 설명의 유창성, 화면의 완성도, 단일 신뢰도 수치가 적절한 의존을 보장한다고 가정하지 않으며, 낮은 품질의 보고를 AI가 설득력 있게 요약할 때 오판 가능성이 커지는지 평가해야 한다.

2.4 고전 모형과 비가환 모형

불완전 관측 아래의 상태 추정과 판단에는 Bayesian 필터, 잠재상태 모형, 부분관측 Markov의 사결정과정(POMDP)이 있다[8]. QFSVDM은 이들을 배제하는 이론이 아니며, 완전한 행동·관측·보상 생성모형을 갖춘 POMDP와 달리 주어진 정보 경로에서 확률을 갱신하는 조건부 표현에 초점을 둔다. 양자유사 인지 모형은 질문·판단의 순서효과를 비가환 연산자로 나타내지만[9,10], 기억과 맥락을 가진 고전 모형도 순서효과를 만들 수 있다. 따라서 순서에 따라 출력이 달라진다는 관찰만으로 비가환 형식의 고유한 우위를 주장하지 않으며, 비교 모형에는 같은 입력의 부호 자유 고전 필터, 시차 입력 로짓, 유사한 복잡도의 잠재상태 모형을 포함한다.

3. QFSVDM의 개념 구조와 정의

3.1 판단 질문을 먼저 정의하는 구조

모든 영역의 정보를 하나의 확률로 합치기 전에 지휘관이 답하려는 질문을 정의한다. 예를 들면 ‘정해진 검토 지평에서 상황평가에 필요한 정보·지원 여건이 충족될 것인가’이다. 임무 질문은 대상 범위, 판단 시각, 목표 시각, 관측 가능한 결과 기준을 포함하며, 같은 자료라도 다른 질문에는 다른 상태 정의와 부하가 필요할 수 있다.

Fig. 2는 제안 구조다. 감지 단계는 원천 관측을 확보하고, 자료 생산 단계는 출처·시각·중복·정정·접근 권한을 확인한다. 이해 단계에서는 수치 모형의 확률 판단과 생성형 AI의 근거 설명을 분리하여 생성하고, 지휘관의 검토가 수용·수정·보류를 결정한다. 판단과 이후 관측은 별도의 평가 기록으로 축적된다. 이 구조는 공식 JADC2 체계도나 실제 배치 구조가 아니다.

3.2 생성형 AI와 수치 모형의 역할 분담

생성형 AI는 보고에서 사건·시간·주장·출처를 추출하고, 서로 다른 용어를 정렬하며, 근거와 상충을 자연어로 설명하는 데 사용할 수 있다. 추출 결과는 원문 위치와 연결하고 원문에 없는 내용은 확인되지 않은 추론으로 표시하며, 외부 보고에 포함된 명령문은 실행 지시가 아닌 자료 내용으로 취급한다. 보안 등급과 공유 권한은 모델이 바꾸는 대상이 아니다. 확률 계산은 버전이 고정된 수치 모형이 수행하고, 언어모형이 만든 숫자를 필터의 산출값으로 표시하지 않는다. 설명에는 입력 마감시각, 사용·제외된 근거, 이전 판단과의 차이, 남은 정보 공백을 포함하며, 언어모형이 중단되어도 수치 결과와 원천 목록은 확인할 수 있도록 기능을 분리한다.

3.3 지휘관의 권한과 평가 대상

지휘관은 임무 질문과 검토 지평을 확인하고 AI의

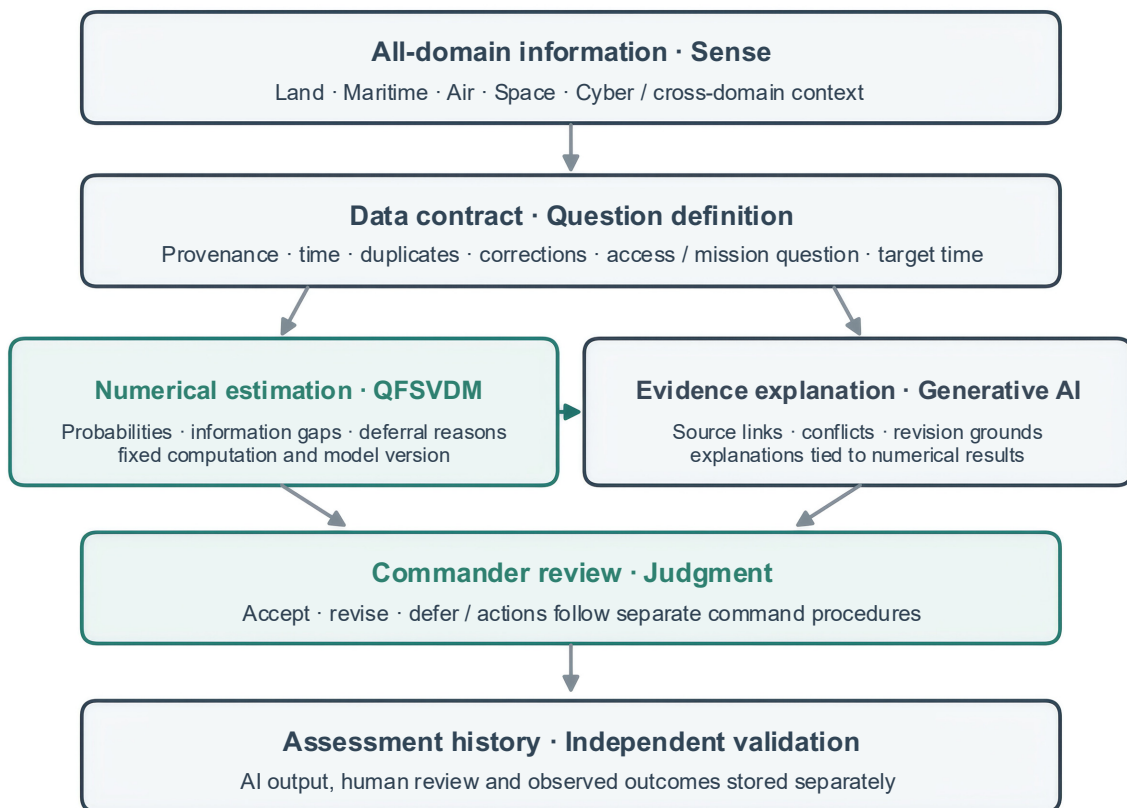


Fig. 2. Proposed structure of all-domain awareness and QFSVDM. Separately computed probabilities and evidence explanations converge on the commander's review; assessment and outcome records feed the validation system.

확률·근거·제약을 검토하여 판단한다. 모형의 확률은 특정 선택의 자동 실행 허가가 아니다. 인간의 결정은 수치 출력과 구분된 식별자로 저장하며, 수정·보류 사유를 기록할 수 있고, 인간이 AI에 동의했다는 사실을 정답 라벨로 사용하지 않는다. 판단 지원의 평가에는 정확한 사실의 파악, 중요한 상충의 발견, 정보 부족의 인식, 과도한 확신의 감소가 포함된다. 결정 시간의 단축만을 목표로 삼지 않으며, 더 빠른 잘못된 판단과 더 느리지만 필요한 확인을 거친 판단을 구분한다.

3.4 QFSVDM의 정의

위 구조를 하나의 모형으로 정의한다. QFSVDM은 임무 질문 q 에 대해 다음 여덟 요소의 튜플로 정의되는 확률 판단 지원 모형이다.

$$QFSVDM(q) = (K, F_t, \rho_t, \{U_{i,e}\}, T_d, \Pi, D, R).$$

K 는 임무 여건의 상호 배타적 결과 집합(기본 예시에서 충족·제한·미충족)이고, F_t 는 판단 시각 t 에 실제로 이용 가능하고 접근이 허용된 정보의 집합이다(4.1절). ρ_t 는 $|K|$ 차원 Hilbert 공간의 밀도행렬로 표현한 내부 상태이며 지휘관의 정신 상태가 아니다(4.3절). $U_{i,e}$ 는 정보원 i 의 점수 e 에 대한 정규화된 양의 행렬 갱신이고(5.2절), T_d 는 보고가 없는 시간 Δ 동안의 상태전이다(5.3절). Π 는 방향기저 사영자에 의한 확률 읽기 규칙 $p_k = \text{Tr}(\Pi_k \rho)$ 이다. D 는 확률·근거·제약을 받아 수용·수정·보류를 기록하는 지휘관 판단 인터페이스이고(6장), R 은 원천·판단·검토·정답을 변경 불가능한 식별자로 연결하는 기록 체계다(7장).

이 정의에서 수학적 대상은 (ρ_t, U, T, Π) 뿐이며, 나머지는 자료와 절차의 명세다. 모든 부하와 전이가 방향기저에서 대각이면 (ρ_t, U, T, Π) 는 확률벡터의 지수 점수 필터로 정확히 환원되므로(정리 1), QFSVDM은 그 고전 필터를 특수한 경우로 포함한다. 비대각 확장이 정당화되는지는 이름이 아니라 9장과 같은 비교 실험으로 판단한다.

4. 상태·시간·정보의 정의

4.1 임무 질문과 정보집합

평가 하나를 임무 질문 q , 생성시각 t , 목표시각 T , 정보 마감시각 c_t 로 식별한다. 원천 사건시각, 수신시각, 정규화와 접근 확인을 마친 이용 가능시각을 구분한다.

$$F_t = \sigma\{x_j : t_j^{available} \leq c_t, access_j(q, t) = 1\}, \quad (1)$$

$$c_t \leq t$$

이용 가능시각은 수신·처리완료시각보다 이르지 않다. 통신 지연으로 뒤늦게 도착한 보고는 원천 사건시각이 과거라는 이유로 이전 판단에 포함하지 않으며, 과거 판단의 재현은 당시 접근 가능했던 자료 목록과 버전을 사용한다. 동시 도착 보고의 순서는 수신 순번과 식별자로 결정하고 기록한다. 정정은 새 버전으로 남기고 이전 판단을 덮어쓰지 않는다. 새 판단에 정정을 반영할 때에는 정해진 체크포인트부터 원래 기록을 대체하여 재생하거나 정정 전용 관측모형을 적용하는 방법 중 하나를 명세하며, 정정값을 두 번째 독립 관측처럼 더하지 않는다.

4.2 임무 여건 상태와 정보부재의 구분

최소 예시에서 임무 여건의 상태는 충족(S), 제한(L), 미충족(I)의 세 범주이며, 하나의 사전 정의된 평가 질문에 대한 상호 배타적 결과다. 출처 영역을 이 범주에 대응시키지 않는다. ‘알 수 없음’은 자료와 판단의 한계일 수 있으므로 별도의 결측·유보 상태로 관리한다. 관측 가능한 결과 지표 $V_{q,T} \in [0,1]$ 와 고정 경계 $0 < b_1 < b_2 < 1$ 에 대한 연구용 라벨은 다음과 같다.

$$Y_{q,T} = \begin{cases} S, & V_{q,T} > b_2 \\ L, & b_1 \leq V_{q,T} \leq b_2 \\ I, & V_{q,T} < b_1 \end{cases} \quad (2)$$

V 는 군사적 가치를 대표하는 만능 점수가 아니다. 실제 적용에서는 한정된 가용성 질문에 대해 측정 항목, 정답 확보 절차, 결측 규칙을 정해야 한다. 본 논문의 합성 실험에서만 $b_1 = 0.35$, $b_2 = 0.65$ 를 사용하며,

이는 편의상의 수학적 구간이다. 정답을 관측하지 못한 사례는 미충족으로 바꾸지 않고 평가 가능 여부를 따로 남긴다.

4.3 밀도행렬과 확률 출력

세 상태의 정규직교기저를 갖는 $H = \mathbb{C}^3$ 에서

$$\begin{aligned} \rho_t &= \rho_t^\dagger, \quad \rho_t \succcurlyeq 0, \\ \text{Tr} \rho_t &= 1, \\ p_{k,t} &= \text{Tr}(\Pi_k \rho_t), \\ \Pi_k &= |k\rangle \langle k|, \quad k \in \{S, L, I\} \end{aligned} \quad (3)$$

p 는 지정한 질문·지평에 대한 조건부 상태 확률이며, 정확한 표기는 $\rho_{q,T|t}$ 이지만 고정된 질문에 대해 첨자를 줄여 쓴다. 갱신 시각 t 와 목표시각 T 는 다르다. 일반적인 3차원 밀도행렬은 8개의 실수 자유도를 가지므로 확률벡터보다 넓은 내부 표현을 허용한다는 비용도 함께 평가한다. 임무 질문마다 상태와 모수를 별도로 정의하고, 서로 다른 지평의 확률을 단순 평균하지 않으며, 지휘소 간 확률 공유는 질문 정의·정보 범위·모형 버전이 같을 때에만 직접 비교한다.

4.4 확률 집중도와 판단 신뢰의 구분

상태 엔트로피, 출력 확률의 엔트로피, 지정 기저의 상대엔트로피 결맞음을 구분한다[11].

$$\begin{aligned} S(\rho) &= -\text{Tr}(\rho \log \rho), \\ H(p) &= -\sum_k p_k \log p_k, \\ C_{\text{rel}} &= H(p) - S(\rho) \end{aligned} \quad (4)$$

$1 - S(\rho)/\log 3$ 은 상태의 혼합 정도를 나타내는 지표일 뿐 지휘관에게 제공할 ‘판단 정확도’가 아니다. 같은 출력 확률에도 다른 내부 상태가 가능하고, 확률이 한 상태에 집중되어도 핵심 자료가 누락되면 잘못된 확신일 수 있다. 경험적 신뢰는 독립된 예측·정답 기록의 보정 관계로 평가하며, 자료 품질·출력 확률·표본외 보정·판단 유보를 하나의 점수로 합치지 않는다.

5. 비가환 정보 갱신 모형

5.1 정보 점수와 출처 의존성

정보원 i 의 유한 실수 점수 $o_{i,t}$ 는 훈련 구간에서 고정된 특징 변환 규칙에서 얻는다. 품질 계수 $r_{i,t} \in [0, 1]$, 정보 연령 $a_{i,t} \geq 0$, 노후화율 $\eta_i \geq 0$ 로 유효 점수를 정의한다.

$$e_{i,t} = r_{i,t} o_{i,t} \exp(-\eta_i a_{i,t}) \quad (5)$$

품질 계수는 원천의 무결성·최신성·완결성에 대한 정의된 지표이며 보고가 참일 확률이 아니다. 마지막 관측값의 반복 전송은 새로운 독립 증거가 아니고, 결측 입력을 0이라는 정상 관측으로 대체하지 않으며, 관측이 없을 때는 결측 표시와 시간전이만 적용한다. 여러 보고가 같은 최초 원천에서 파생되면 원천 계보를 공유하며, 같은 사건·버전의 재전송은 한 번만 처리하고 서로 다른 원천의 상충은 보존한다. 같은 잠재량을 추정하는 불편 추정치들과 양정치 오차 공분산 Σ 가 있을 때에만 합이 1인 최소분산 가중

$$w = \frac{\Sigma^{-1} \mathbf{1}}{\mathbf{1}^\top \Sigma^{-1} \mathbf{1}} \quad (6)$$

을 사용하고, 공분산은 훈련 자료에서만 추정하며 조건수와 규제량(수축 추정[12] 등)을 기록한다. 보고수의 증가를 독립 증거의 증가로 판단하지 않는다.

5.2 정규화된 점수 필터와 공정한 기준모형

정보 점수에 대한 양의 행렬 필터를 다음과 같이 정의한다.

$$\begin{aligned} A_i(e) &= \exp(eL_i/2), \\ U_{i,e}(\rho) &= \frac{A_i(e)\rho A_i(e)^\dagger}{\text{Tr}[A_i(e)\rho A_i(e)^\dagger]} \end{aligned} \quad (7)$$

명제 1. Hermitian L , 유한 실수 e , 밀도행렬 ρ 에 대해 식 (7)은 밀도행렬을 반환한다. 그러나 일반적인

로 $A^\dagger A \leq I$ 는 성립하지 않는다.

증명. $A = \exp(eL/2)$ 는 양정치·가역이므로 $A\rho A^\dagger$ 는 Hermitian 양의 준정부호이고 대각합은 양이다. 이 값으로 나누면 대각합 1의 양의 준정부호 행렬이다. $L = G, e = I$ 에서 $A^\dagger A$ 의 고유값은 $e^1, 1, e^{-1}$ 이다.

따라서 식 (7)은 입력 점수가 주어진 조건부 필터이며, 모든 결과에 대해 $\Sigma_Z A_Z^\dagger A_Z = I$ 를 만족하는 완비 양자도구[13]나 입력 발생확률까지 정의한 생성모형이 아니다. 정규화 분모를 보고의 우도로 해석하지 않는다. 수치 구현은 $eL_i/2$ 의 최대 고유값을 뺀 뒤 지수화하며, 이 공통 스칼라는 정규화에서 소거된다.

기본 부하는 $G = \text{diag}(1, 0, -1)$, $J = \text{diag}(1, -2, 1)$, $X = |S\rangle \langle I| + |I\rangle \langle S|$, $Y = -i|S\rangle \langle I| + i|I\rangle \langle S|$ 의 추적 0 최소족으로 제한한다.

$$L_i = d_i G + n_i J + x_i X + y_i Y, \quad (8)$$

$$(d_i, n_i, x_i, y_i) \in \mathbb{R}^4$$

고전 기준모형 M_D 는 $x_i = y_i = 0$ 이며 d_i, n_i 의 부호를 제한하지 않는다. 확장모형 M_Q 는 비대각 계수를 추가한다. 부정적 보고를 표현하지 못하도록 고전 모형의 부호를 제한하면 확장의 효과를 과대평가하므로 같은 부호 자유도를 보장한다. 이 최소족은 상태 L 과 다른 상태 사이의 결맞음을 포함하지 않으며, 일반적인 세 상태 결맞음은 모수와 검증 부담이 더 큰 다른 모형이다.

5.3 시간전이와 정보 공백

새 보고가 없는 동안의 내부 전이는 재설정과 dephasing을 포함한 유한차원 생성자로 정의한다 [14,15].

$$(7\text{페이지 하단 참조}) \quad (9)$$

$\rho_\pi = \text{diag}(\pi)$ 는 합이 1인 비음 기준분포이고, γ 는 기준분포로 돌아가는 비율, λ 는 추가 비대각 감쇠율이다. 두 값은 상황이 안전해지는 속도나 지휘관의 확

신 감소율로 직접 해석하지 않으며, 기준분포는 혼련·검증으로 정한다. $H = \text{diag}(h)$ 이면 재설정 사상 $R(X) = \rho_\pi \text{Tr} X (R^2 = R)$ 의 고정점이 ρ_π 이고 대각 $H \cdot G$ 가 가환하므로 세 생성자가 가환하며, 정확한 시간전이는

$$T_A(\rho) = D_A \odot [e^{-\gamma A} U_A \rho U_A^\dagger + (1 - e^{-\gamma A}) \rho_\pi],$$

$$U_A = \text{diag}(e^{-i h_A A}), \quad (10)$$

$$(D_A)_{jk} = e^{-\lambda(g_j - g_k)^2 A/2}$$

이다. 일반 비대각 H 에서는 개별 채널의 합성식을 정확해라고 하지 않고 9×9 생성자의 행렬 지수 $\exp(\Delta L)$ 를 계산한다. 기본 비교 실험은 $H = 0$ 으로 고정하여 부하의 비대각 효과와 전이 효과를 혼동하지 않는다.

Hamiltonian과 재설정이 없는 순수 dephasing에서는 성분별로 $\rho_{jk}(t) = \rho_{jk}(0) \exp[-\lambda(g_j - g_k)^2 t/2]$ 이므로 S-I와 S-L 성분의 감쇠인자 로그 비는 4이다. 원래 진폭의 로그 비에는 초기값이 포함되므로 일반적으로 4가 아니며, 재설정을 포함하면 두 감쇠율은 $\gamma + 2\lambda$ 와 $\gamma + \lambda/2$ 가 되어 같은 비율이 유지되지 않는다. 최소 S-I 확장은 대각 초기상태에서 S-L 결맞음을 만들지 않으므로 이 비율을 출력 확률만으로 검증할 수 있다고 주장하지 않는다. 입력 노후화율 η_i , 재설정률 γ , 감쇠율 λ 는 서로 다른 역할이며, 통신 단절 중 필터가 기준분포에 가까워지더라도 관측이 복구되었다는 뜻은 아니므로 마지막 수신시각과 관측 공백을 출력과 함께 표시한다.

5.4 고전 환원과 식별 가능한 불변량

정리 1 (충분조건 환원). 초기상태와 ρ_π 가 상태기저에서 대각이고, 전이가 식 (10)의 대각 H 를 사용하며, 모든 L_i 가 대각이면 모든 시점의 상태는 대각이고 확률 갱신은 다음 고전 필터와 정확히 일치한다.

$$(8\text{페이지 상단 참조}) \quad (11)$$

$$L(\rho) = -i[H, \rho] + \gamma(\rho_\pi \text{Tr} \rho - \rho) + \lambda \left(G \rho G - \frac{1}{2} \{G^2, \rho\} \right), \quad \gamma, \lambda \geq 0 \quad (9)$$

$$p_k^- = e^{-r\Delta} p_k + (1 - e^{-r\Delta}) \pi_k, \quad p_k^+ = \frac{p_k^- \exp(e_i l_{ik})}{\sum_j p_j^- \exp(e_i l_{ij})}, \quad l_{ik} = (L_i)_{kk} \quad (11)$$

증명. 대각 상태의 유니타리 공액변환은 대각을 바꾸지 않고 $(D_{\Delta})_{kk} = 1$ 이므로 전이는 첫 식이다. 대각 부하에서 A_i 의 대각성분은 $\exp(e_i l_{ik}/2)$ 이므로 공액 변환 후 대각은 $p_k^- \exp(e_i l_{ik})$, 비대각은 0이고 대각합으로 나누면 두 번째 식이다. 관측 횟수에 대한 귀납으로 결론을 얻는다.

이는 충분조건이며, 조건 하나가 깨졌다고 모든 입력에서 출력 차이가 생긴다는 역명제는 성립하지 않는다. 대각 상태에서 λ 는 출력에 나타나지 않으므로 기준모형의 식별 가능 모수에서 제외한다.

M_Q 의 명목 모수 수를 식별 차원으로 보고하지 않는다. 두 종류의 관측동등성이 있기 때문이다. 첫째, 대각 유니타리 $\text{diag}(e^{i\phi_k})$ 로 상태와 모든 부하를 함께 변환하면 출력이 불변이며, 복소 연결계수 $z_i = x_i + iy_i$ 에 대해 이는 공통 위상 회전이다. 둘째, 기본 설정($H = 0$, 실수 대각 초기·재설정 상태, 실수 입력)에는 **컬레 대칭**이 있다. 모든 정보원에 대해 동시에 $y_i \rightarrow -y_i$ 로 바꾸면 $L_i \rightarrow \bar{L}_i$ 이고, 식 (7)·(10)이 실수 계수 연산이므로 모든 시점에서 $\rho_i \rightarrow \bar{\rho}_i$ 가 되어 대각 성분, 곧 출력 확률이 완전히 같다. 따라서 정보원이 n 개일 때 식별되는 것은 연결 크기 $|z_i|n$ 개와 상대 위상의 코사인 $\cos(\arg z_1 - \arg z_i)n-1$ 개이며, 상대 위상의 부호는 식별되지 않는다. 회전 게이지는 명목 자유도 $4n + 2$ 를 $4n + 1$ 로 줄이고 컬레 대칭은 최적해를 항상 쌍으로 만든다. 이 대칭은 H 에 복소 비대각 성분이 있거나 초기상태에 복소 결맞음이 있으면 깨진다. 귀무모형 $x_i = y_i = 0$ 에서는 위상과 λ 가 식별되지 않으므로 단순한 추가 모수 수에 따른 카이제곱 근사를 적용하기 어렵다 [16]. 실험에서는 Cartesian 계수로 최적화되 보고는 $|z_i|$, $\cos(\arg z_1 - \arg z_i)$, 정규화 교환자 $X_{iG} = \| [L_i, G] \|_F / (\| L_i \|_F \| G \|_F)$ 로 하고, 다중 시작점의 해는 두 동등성으로 정규화한 뒤 일치 여부를 판단한다.

5.5 정보 순서, 민감도와 절차 차이의 상한

두 정보원의 갱신은 일반적으로 순서에 따라 다르

다. 실제 수신 경로를 기본 경로로 고정하고 순서를 바꾼 계산은 진단용 비교로 구분한다.

$$\Delta_{ij}^{\text{order}} = \frac{1}{2} \sum_k |p_k^{(i \rightarrow j)} - p_k^{(j \rightarrow i)}|, \quad (12)$$

$$x_{ij} = \| [L_i, L_j] \|_F$$

Δ^{order} 는 같은 점수와 초기상태에서 비교한다. 시간적으로 동시이며 의미상 교환 가능한 보고에 대한 과도한 순서 민감도는 성능이 아니라 취약성일 수 있다. 회전각이 커질수록 비가환성이 단조 증가한다고 가정하지 않는다. S-I 평면의 회전

$$R_\theta = \begin{pmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{pmatrix} \text{에 대해 } L_\theta = R_\theta G R_\theta^\dagger \text{이면}$$

$\| [G, L_\theta] \|_F = 2\sqrt{2} |\sin 2\theta|$ 로 $\theta = \pi/4$ 에서 최대이고 $\theta = \pi/2$ 에서 다시 가환한다. 식 (8)의 부하에 대해서는 $\| [L_i, G] \|_F = 2\sqrt{2(x_i^2 + y_i^2)}$ 이다.

서로 다른 관측 절차의 확률 차이도 유계다. rank-1 기저 사영족 Q_m 의 비선택 측정 $D_Q(\rho) = \sum_m Q_m \rho Q_m$ 과 임의의 효과 $0 \leq \Pi \leq I$ 에 대해

$$|\text{Tr}[\Pi(\rho - D_Q \rho)]| \leq \frac{1}{2} \| \rho - D_Q \rho \|_1 \quad (13)$$

이다. $Z = \rho - D_Q \rho$ 가 Hermitian이고 대각합이 0이므로 양·음 부분의 대각합이 각각 $\| Z \|_1/2$ 이기 때문이다. 이는 서로 다른 절차의 분포 비교이며 하나의 조건부 확률공간에서 총확률법칙이 틀렸다는 뜻이 아니고, 비가환이라는 이유만으로 최대 불확실성이 보장되지도 않는다.

5.6 영역 간 영향과 공동 원천

영역 간 영향은 정보원 개수의 증가와 다르다. 확장 모형에서 시차 입력 $z_t = e_t + R z_{t-1}$ 를 사용한다면 R 은 검증할 재귀 구조이며 인과효과로 자동 해석하지 않는다. 상수 입력의 정상상태에 한해 $z = (I - R)^{-1}e$ 이고, 고

정된 유도 노름에서 $\|R\| \leq q < 1$ 이면 Neumann 급수에서

$$\|(I-R)^{-1}\| \leq \frac{1}{1-q} \quad (14)$$

이다. 스펙트럼 반경만으로 같은 상한을 주장하지 않는다. $R = \begin{pmatrix} 0 & 10 \\ 0 & 0 \end{pmatrix}$ 은 반경이 0이지만 역행렬의 2-노름이 약 10.099이다. 기본 참조 실험은 $R = 0$ 으로 두어 확인되지 않은 영역 간 전파를 추가하지 않으며, 여러 지휘소의 판단을 결합할 때 같은 원천을 다시 받아 자기 확신을 강화하는 순환을 막도록 근거 계보를 유지한다.

5.7 측정 순서 항등식과 적용 범위

양자유사 인지 문헌의 QQ 등식은 별도로 정의된 이전 판단 실험의 성질이며, 임무 상태를 예측하는 연속 점수 필터의 검정식이 아니다. 임의의 유한 차원에서 사영자 P , Q 와 $X = 2P - I$, $Y = 2Q - I$, 대칭 비예리 효과 $E_{A,\pm} = (I \pm r_A X)/2$, $E_{B,\pm} = (I \pm r_B Y)/2$ ($0 \leq r_A, r_B \leq 1$)의 양의 제공근을 Lüders 연산자로 사용한다.

정리 2. A 다음 B를 측정하여 같은 부호를 얻는 확률합은 $S_{AB} = 1/2 + r_A r_B / 4 \text{Tr}[\rho(XY + YX)]$ 이며 순서 교환에 불변이다.

증명. $A_{\pm}$ 를 $aP + b(I - P)$, $a(I - P) + bP$ ($a^2 - b^2 = r_A$)로 쓰면 $A_+ \rho A_+ - A_- \rho A_- = r_A / 2 (X\rho + \rho X)$ 이다. 같은 부호 확률합에 효과를 대입하면 비선택적 합에서 $1/2$, 차이항과 대각합의 순환성에서 나머지 항을 얻고, 우변은 $X \leftrightarrow Y$ 교환에 불변이다.

결과 의존 유니타리, 비대칭 잡음, 임의의 POVM으로 범위를 넓히면 같은 등식이 보장되지 않으며[17], QQ 성립만으로 고전적 인지 모형이 배제되지도 않는다[10,18]. 본 원고는 이 등식을 지휘관에게 정보를 특정 순서로 제시하여 판단을 유도하는 규칙으로 사용하지 않는다.

6. 확률 판단에서 지휘관의 선택으로

6.1 추천, 보류와 추가 확인

상태 확률은 ‘무엇이 예상되는가’의 출력이고 지휘관의 선택은 ‘현재 권한과 제약 아래 무엇을 할 것인가’의 판단이다. 같은 확률이라도 검토 시간, 근거의 결측, 임무 의도와 판단 권한이 다르면 선택이 달라지므로 최대 확률 상태를 곧바로 하나의 행동에 대응시키지 않는다. 개념 예시의 선택지는 현 판단 유지, 추가 정보 확인, 상급 검토 요청 같은 정보·검토 절차이며, 전장 행동의 허가는 해당 지휘체계의 별도 절차에 속한다.

자료가 부족하면 보류할 수 있지만 최대 확률이 낮다는 한 가지 이유로 모든 경우를 처리하지 않는다. 핵심 입력의 누락, 허용 연령 초과, 분포 이탈 입력, 원천 무결성 문제, 정의되지 않은 질문, 모형 오류를 별도의 상태로 표시하고, 보류 기준은 검증 전에 정하며 보류 사례를 성능 분모에서 조용히 제거하지 않는다. 추가 확인은 불확실성 감소 가능성과 소요 시간을 함께 고려하지만, 추가 관측의 비용·결과 생성모형이 없다면 ‘최적 정보 획득’이 계산되었다고 주장하지 않는다. 이것이 완전한 행동 모형을 포함한 POMDP[8]와 기본 QFSVDM의 범위를 가르는 지점이다.

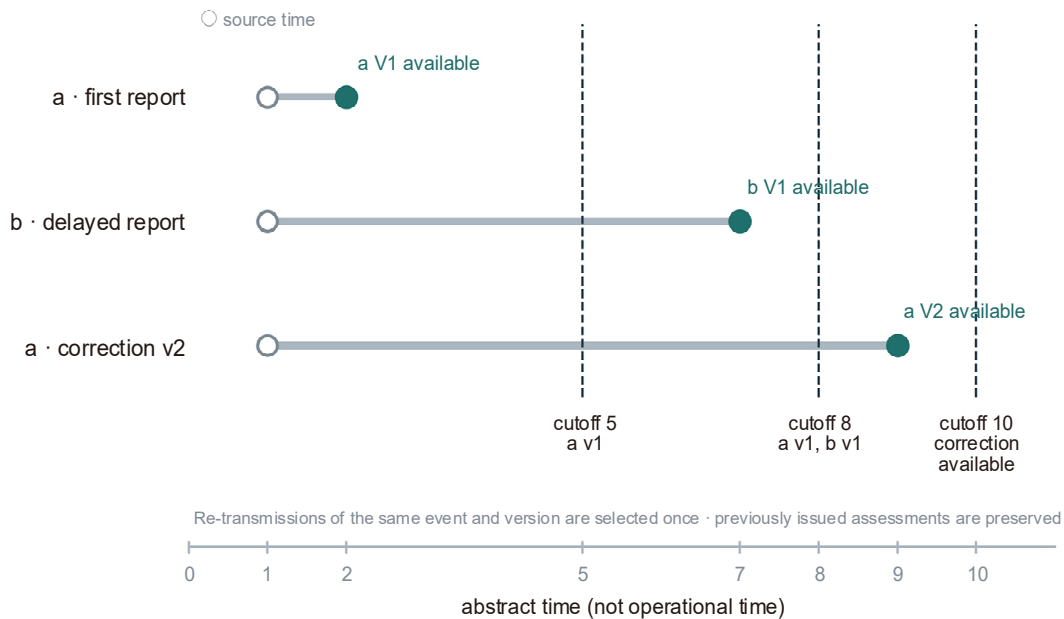
6.2 판단 근거의 설명과 인간의 피드백

설명은 관측 사실, 모형의 추정, 규칙에 따른 제한, 지휘관의 판단을 구분한다. 판단의 변화에는 새로 이용 가능해진 자료, 정정된 원천, 바뀐 질문이나 모형 버전을 함께 표시하며 ‘AI가 학습하여 판단이 바뀌었다’고 설명하지 않는다. 정보원 하나를 제외하고 재계산한 확률 차이는 민감도 설명에 쓸 수 있으나, 비선형·비가환 갱신에서는 기여의 합이 전체 변화와 같지 않으므로 이를 인과효과로 부르지 않는다. 자연어 설명은 재계산 결과와 근거 식별자에서 생성하고 사실 일치도를 별도로 평가한다.

지휘관의 수정은 유용한 연구 자료이지만 곧바로 정답이 아니다. 수정 사유를 임무 맥락 추가, 원천 오류 발견, 규칙 적용, 불확실성 수용으로 분리하고, 이후 확인된 사실과 AI-인간의 판단을 비교하여 해석한다. AI의 설명을 보고 내려진 판단을 다시 학습 정답

Table 2. Record types and minimum key fields

Record	Key fields	Condition to verify on reproduction
Source event	event_id, version, provenance	Re-transmissions and corrections distinguished
Time and access	source, received, available, access	Only data available at the time selected
AI assessment	assessment_id, question_id, cutoff, target	Target, information scope, and model version match
Computation basis	event_refs, input_hash, model_version	Same input reproduces the computation
Human review	review_id, assessment_id, reason	AI output and human choice separated
Observed outcome	outcome_id, available_at, evidence	Later corrections and unobserved status preserved

**Fig. 3.** Research example of event timing. At cut-off 5 only the first version of event a is available; at cut-off 8 the delayed b is added; at cut-off 10 the correction of a becomes available. Previously issued assessments are preserved.

으로 넣으면 출력이 되먹임되는 편향이 생긴다. 판단 이후의 행동이 미래 여건을 바꿀 수 있으므로 관측 결과만으로 선택하지 않은 대안의 결과를 안다고 가정하지 않으며, 초기 검증은 행동 개입을 통제한 정보평가 과제와 고정된 상황기록의 재생에서 시작한다.

7. 자료 생산과 운용 구조

7.1 원천 기록과 판단 기록

원천 사건은 변경 불가능한 식별자와 버전을 가지며 사건·수신·이용 가능시각, 원천 계보, 자료 종류, 품질 항목, 접근 범위를 기록한다. 판단은 임무 질문·범위·목표시각·정보 마감시각, 모형과 특징 변환의 버전, 사용한 사건 목록, 출력 확률·보류 사유, 이전 판단 연결을 포함한다. 해시는 동일 파일·동일 입력을

확인하는 수단이지 원천의 진실성을 입증하는 수단이 아니다. 정답은 목표시각 이후의 확인 가능한 관측 또는 독립 판정 절차로 확보하고, 최초 공개시각·판정 근거·정정 버전·평가 가능 여부를 따로 기록하며, 지휘관의 수용 여부와 실제 정답을 같은 열에 저장하지 않는다. 같은 목표에 대한 수백 번의 갱신은 수백 개의 독립 표본이 아니다.

7.2 연결 저하, 복구와 연합 공유

중앙과 현장의 연결이 항상 유지된다고 가정하지 않는다. 현장에서는 마지막으로 확인한 모형·자료 버전을 표시하고 공유 지연과 로컬 판단 여부를 구분하며, 마지막 값을 현재 관측처럼 표시하거나 수신이 없다는 이유로 정상 상태를 생성하지 않는다. 적용 범위를 벗어나면 제한 기능 또는 판단 보류로 전환한다.

연결 복구 시에는 재전송과 새 관측을 식별하고, 원래 발행된 판단은 보존한 채 새 판단을 생성하며, 정정이나 순서 변경이 있으면 체크포인트와 사건 이력에 따라 재생하고 복구 상태와 누락 범위를 표시한다.

연합 공유에서는 공유 가능한 자료와 존재하지만 공유할 수 없는 자료를 구분하고, AI가 처리한다는 이유로 원래 접근 권한을 확장하지 않는다. 서로 다른 정보집합에서 생성된 확률은 그 차이를 숨기지 않고 비교한다. JADC2·다영역 상호운용성은 메시지 전달 뿐 아니라 질문·시간·단위·상태 정의의 호환성을 요구하는 연구 문제이며, 본 논문의 기록 예시는 공식 메시지 표준이나 보안 승인된 인터페이스가 아니다.

7.3 모형 갱신과 감사 가능성

모형 변경은 새 버전으로 발행하고 이전 버전의 예측을 소급 수정하지 않는다. 자료 변환, 보정, 임계값, 부하계수, 계산 환경을 함께 기록하고, 갱신 전후 성능은 같은 정보 마감시각과 목표에 대해 비교한다. 업데이트가 지휘관의 화면이나 해석을 바꾸면 기능 평가와 인간-AI 협업 평가를 함께 수행한다. 책임 있는 AI 원칙의 체계 수준 평가는 수치 모형 외에 데이터 처리, 인터페이스, 권한 관리, 운영 절차를 포함하며 [5,6], 원장에 기록이 남는다는 사실만으로 설명가능성이나 통제가 완성되었다고 선언하지 않는다.

8. 수치검사와 합성자료

8.1 실제 수행한 구현·자료 생산 검사

연구용 사례는 여러 영역의 보고를 받는 가상의 합동 상황평가 셀을 가정하며, 지휘관의 질문은 ‘다음 검토 지평에 필요한 정보·지원 여건이 충족되는가’이다. 예시 입력은 다섯 개의 추상 신호이고 실제 센서·부대·작전자료가 아니다. Fig. 3의 흐름에서는 기존 보고, 지연된 보고, 중복 전달, 정정이 서로 다른 시각에 도착하며, 같은 과거 마감시각으로 재생한 판단에는 이후 도착한 정보가 포함되지 않아야 한다.

참조 구현은 Python 3.12, NumPy 2.3.5, SciPy 1.17.0에서 실행했고 수학 검사 시드는 20260907이다. Table 3는 그 결과다.

사건 선택 검사는 마감시각별 이용 가능한 버전 목

록을 확인한 최소 검사이며, 정정 후 전체 필터 경로를 재생하는 운영 서비스의 검증이 아니다. Table 3는 수학적 명세와 참조 코드의 일치를 지지할 뿐 지휘관 판단의 정확도나 작전효과를 측정하지 않는다.

Table 3. Implementation and data-contract checks actually performed

Check	Scope	Observed
Signed classical reduction (Theorem 1)	400 random settings, 5 updates each	max probability difference 3.89×10^{-16}
Density-matrix properties preserved	400 random complex states, 5 updates each	minimum eigenvalue $\geq$ on all paths (observed minimum 7.20×10^{-7}), max trace error 2.22×10^{-16}
Exact evolution (eq. 10)	fixed state with off-diagonal entries	difference from full generator exponential 4.83×10^{-17}
Event-selection contract	4 checks: delay, duplicate, correction, future data	4 passed

8.2 합성자료 생성과정

세 종류의 합성자료 A·B·C를 정의한다. A는 부호 자유 대각 필터, B는 같은 대각 부하에 비대각 계수를 더한 필터, C는 고전 잠재 체제와 로짓의 결합이다. 모든 자료는 실제 날짜·부대·센서와 연결되지 않은 추상 단위시간 $t = 0, \dots, T-1$ 을 쓴다. 다섯 입력의 잠재벡터는 $v_{-1} = 0$ 에서 시작하는 AR(1)이다.

$$\begin{aligned}
 v_t &= 0.65v_{t-1} + \sqrt{1 - 0.65^2} \epsilon_t, \\
 \epsilon_t &\sim N(0, \Sigma_x), \\
 x_{i,t} &= \tanh(v_{i,t})
 \end{aligned} \tag{15}$$

Σ_x 의 대각은 1, 모든 비대각은 0.15이다. 품질 계수는 1, 연령은 0, 네트워크 계수는 0이며, 따라서 이 자료로 노후화·지연·상충 처리의 성능을 평가하지 않는다.

A는 Table 4의 대각 계수만, B는 네 계수를 모두 사용하며, 두 경우 모두 $\rho_{\text{init}} = \rho_\pi = I/3$, $H = 0$, $\gamma = 0.12$, $\lambda = 0.08$, $\Delta = 1$ 이다. 매 시점 식 (10)의 전이를 먼저 적용하고 입력 1부터 5까지 순차 갱신하며 상태 확률은 대각성분으로 읽는다. C는 $Z_{-1} = 1$ 에서 시작하는 Markov

Table 4. Loading coefficients of the five synthetic sources

Input	d_i	n_i	x_i (B)	y_i (B)
1	0.80	0.10	0.55	0.00
2	-0.60	-0.12	0.20	0.50
3	0.40	0.05	-0.25	0.15
4	0.30	-0.08	0.10	-0.20
5	-0.35	0.07	0.15	0.10

체제 $Z_t \in \{0,1,2\}$ 와 Table 4의 대각 계수로 만든 $s_t = \sum_i d_i x_{i,t}$, $n_t = \sum_i n_i x_{i,t}$ 를 사용한다.

$$P_Z = \begin{pmatrix} .85 & .10 & .05 \\ .075 & .85 & .075 \\ .05 & .10 & .85 \end{pmatrix},$$

$$p_t = \text{softmax} \left[\begin{pmatrix} s_t + n_t \\ -2n_t \\ -s_t + n_t \end{pmatrix} + (1 - Z_t) \begin{pmatrix} .55 \\ 0 \\ -.55 \end{pmatrix} \right] \quad (16)$$

세 경우 모두 $Y_t \sim \text{Categorical}(p_t)$ 를 먼저 뽑고, S이면 (0.65,1), L이면 (0.35,0.65), I이면 (0,0.35)의 균등분포에서 다음 기간의 가용성 지표를 뽑은 뒤 식 (2)로 라벨을 다시 계산하여 일치를 확인한다. 이 검사는 구성상 항상 통과하며 생성 코드의 정합성 검사이지 모형의 증거가 아니다. 잠재 체제 Z와 결과 라벨 Y는 다른 변수이고 Z는 특징으로 제공하지 않는다. 동봉 파일 세 개($T = 1,000$, 시드 202609071-202609072-202609073)의 라벨 구성은 A 457/168/375, B 506/84/410, C 360/279/361(S/L/I)이며, 0-599행 훈련, 600-799행 내부 검증, 800-999행 시험으로 표시

한다. 생성확률과 미래 정답 열은 감사용이며 입력 특징에 포함하지 않는다.

이 자료의 두 성질을 명시한다. 첫째, 생성확률 경로의 1차 자기상관은 A 0.90, B 0.86, C 0.61이고 입력의 자기상관은 0.57~0.63이므로 1,000행의 독립 정보량은 대략 50~240 관측에 해당한다. 한 파일의 200행 시험 구간만으로는 모형 간 손실차를 판정할 수 없고 여러 독립 시드가 필요하다. 둘째, 생성확률로 채점한 오라클 로그손실은 A 0.629, B 0.711, C 0.916으로 균등 예보 $\ln 3 = 1.099$ 보다 각각 0.47, 0.39, 0.18 nat 낮다. 이 값이 어떤 모형도 넘을 수 없는 하한이다. 다섯 입력을 사용했다는 이유로 실제 다섯 작전영역의 특성을 재현했다고 해석하지 않는다.

9. 합성자료 반복 적합 실험

9.1 설계

실험 설정은 첫 실행 전에 설정 파일로 고정했고 실험 중 바꾸지 않았다.

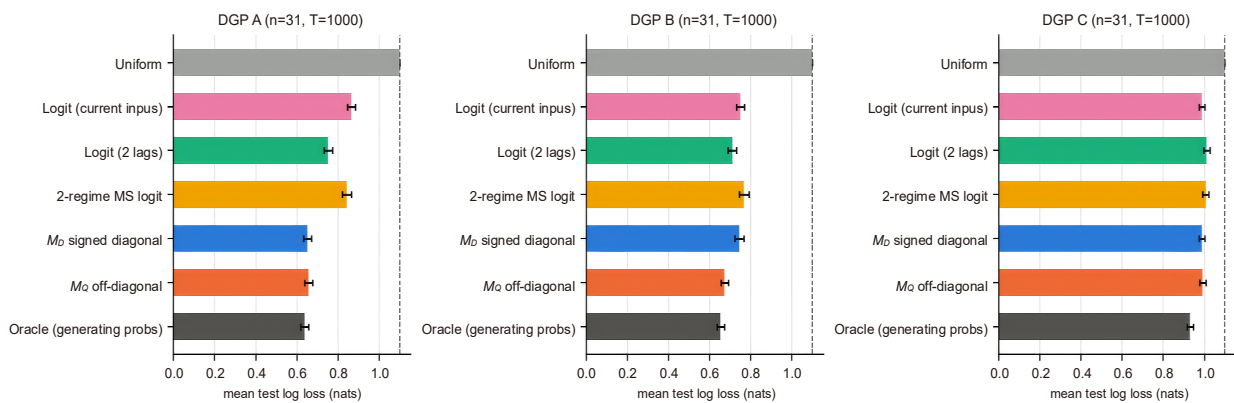
여덟 모형은 같은 입력과 라벨을 사용한다. 필터와 로짓은 입력만 사용하고 과거 정답을 갱신에 쓰지 않는다. 순차 빈도와 Markov 전환 로짓은 정의상 공개된 과거 라벨로 체제확률을 갱신하므로 필터보다 많은 정보를 사용한다. 시차 2 입력 로짓(모수 32개)은 필터의 기억을 흉내 내는 고전 대안이고, 2체제 전환 로짓(모수 26개)은 2.4절이 요구한 유사 복잡도의 고전 잠재상태 모형이다. 필터 상태는 입력만으로 자료

Table 5. Configuration of the repeated fitting experiment

Item	Setting
Processes and seeds	A, B, C: one enclosed file each plus 30 independent seeds at $T = 1,000$; 10 seeds each at $T = 500$ and $T = 2,000$ (153 datasets)
Split	time-ordered 60/20/20 (train / internal validation / test), identical for all models
Inputs and target	five inputs - $x_1 - x_5$ and the next-period state label; generating probabilities and future outcomes excluded
Models	(1) uniform, (2) training frequency, (3) sequentially updated frequency, (4) logit on current inputs, (5) logit with two lags, (6) two-regime Markov-switching logit, (7) M_D , (8) M_Q ; reference: oracle
M_D / M_Q parameters	(d_i, n_i) per source + $\gamma = 11$ / (d_i, n_i, x_i, y_i) per source + $\gamma, \lambda = 22$
Regularisation	L2 on loadings, $\omega \in \{10^{-4}, 10^{-3}, 10^{-2}\}$ chosen by internal-validation log loss; same grid for all fitted models
Optimisation	L-BFGS-B with box constraints ($ \text{loading} \leq 5$, $\gamma \in [0,3]$, $\lambda \in [0,5]$), at most 400 iterations, five starts for non-convex models
Scoring	mean test log loss (primary) and summed Brier score; paired and Wilcoxon signed-rank tests across seeds
Ledger	objective, function evaluations, convergence flag, time, and coefficients per experiment, model, regularisation value, and start

Table 6. Mean test losses ($T=1,000$, 31 datasets per process; standard errors across seeds in parentheses)

Model	A log loss	A Brier	B log loss	B Brier	C log loss	C Brier
Uniform	1.0986 (0.0000)	0.6667	1.0986 (0.0000)	0.6667	1.0986 (0.0000)	0.6667
Training frequency	1.0356 (0.0070)	0.6292	0.9242 (0.0081)	0.5755	1.0949 (0.0021)	0.6644
Sequential frequency	1.0330 (0.0068)	0.6271	0.9217 (0.0076)	0.5736	1.0932 (0.0018)	0.6632
Logit (current inputs)	0.8637 (0.0100)	0.5050	0.7498 (0.0100)	0.4375	0.9877 (0.0070)	0.5908
Logit (2 lags)	0.7519 (0.0101)	0.4318	0.7099 (0.0106)	0.4051	1.0098 (0.0076)	0.6047
2-regime MS logit	0.8414 (0.0109)	0.4886	0.7676 (0.0115)	0.4465	1.0059 (0.0076)	0.6004
M_D signed diagonal	0.6506 (0.0100)	0.3712	0.7444 (0.0112)	0.4071	0.9867 (0.0069)	0.5907
M_Q off-diagonal	0.6557 (0.0095)	0.3734	0.6733 (0.0095)	0.3824	0.9911 (0.0072)	0.5920
Oracle (generating probabilities)	0.6375 (0.0099)	0.3633	0.6539 (0.0097)	0.3744	0.9296 (0.0077)	0.5549

**Fig. 4.** Mean test log loss by generating process ($T=1,000$, 31 datasets). Error bars are 95% intervals of the mean across seeds; the dashed line is the uniform forecast $\ln 3$.

의 처음부터 순차 계산하므로 시험 구간에서도 누출이 없다. 적합 코드는 참조 구현과 별개로 작성한 U-I 블록 닫힌형 지수 구현이며, 동봉 자료 세 개의 입력-라벨-지표를 생성확률 2.2×10^{-15} 이내로 완전히 재생성함을 먼저 확인했다.

9.2 결과

모든 적합이 종료 조건 안에서 수렴했다(수렴 실패 0). M_D 는 5개 시작점이 모두 같은 최적값에 도달했고(시작점 간 목적함수 폭 중앙값 3.1×10^{-7}), M_Q 는 시작점의 96.7%가 최적값의 10^{-3} 상대오차 안에 들었으며, 2체제 전환 로짓은 52.9%만 최적값에 도달하여 다봉성이 뚜렷했다. 규제계수는 세 모형 모두 10^{-2} 가 가장 자주 선택되었다. Table 6는 $T=1,000$ 의 31개 자료에서 시험 구간 평균 로그손실과 Brier 점수의 시드 평균(괄호는 표준오차)이다.

Table 7은 같은 자료에 적합한 두 모형의 시험 로

그손실을 자료별로 뺀 대응 차이이며 음수는 앞 모형이 낮은 손실을 뜻한다. 대비 7개는 사전에 고정했고 다중 비교 보정을 적용하지 않은 원 p값을 보고한다.

Table 8은 주력 대비 $M_Q - M_D$ 를 세 길이에서 비교한 것이고 Fig. 5는 자료별 차이를 모두 표시한 것이다.

Table 9은 M_Q 의 게이지 불변 추정량과 전이 모수의 시드 중앙값을 생성 참값(B)과 나란히 놓은 것이다. A의 참값은 $|z_i| = 0$ 이고 C에는 참값이 없다.

9.3 해석

A(부호 자유 대각 생성과정). M_D 는 오라클과 0.013 nat 차이($T=2,000$ 에서는 0.004 nat)로 생성확률을 사실상 복원했다. M_Q 는 M_D 보다 0.005 nat 나뉘었고 이 차이는 통계적으로 유의하며, $T=500$ 에서 0.019, $T=2,000$ 에서 0.004 nat으로 표본이 커질수록 줄어든다. 이는 다섯 정보원에 추가된 자유 모수 11개의 대가이며, $|z_i|$ 중앙값이 0.07~0.13으로 0

Table 7. Paired differences in test log loss ($T = 1,000$, 31 datasets)

Contrast	DGP	Mean difference	95% CI	t-test p	Wilcoxon p	Fraction first model better
$M_Q - M_D$	A	+0.0052	[+0.0022, +0.0081]	0.002	< 0.001	0.26
$M_Q - M_D$	B	-0.0711	[-0.0846, -0.0576]	< 0.001	< 0.001	0.97
$M_Q - M_D$	C	+0.0044	[+0.0016, +0.0072]	0.004	0.007	0.32
M_D - Logit (2 lags)	A	-0.1014	[-0.1108, -0.0920]	< 0.001	< 0.001	1.00
M_D - Logit (2 lags)	B	+0.0345	[+0.0235, +0.0455]	< 0.001	< 0.001	0.13
M_D - Logit (2 lags)	C	-0.0231	[-0.0286, -0.0177]	< 0.001	< 0.001	0.97
M_Q - Logit (2 lags)	A	-0.0962	[-0.1053, -0.0871]	< 0.001	< 0.001	1.00
M_Q - Logit (2 lags)	B	-0.0366	[-0.0470, -0.0262]	< 0.001	< 0.001	0.87
M_Q - Logit (2 lags)	C	-0.0187	[-0.0248, -0.0126]	< 0.001	< 0.001	0.90
M_D - MS logit	A	-0.1909	[-0.2048, -0.1770]	< 0.001	< 0.001	1.00
M_D - MS logit	B	-0.0232	[-0.0380, -0.0084]	0.004	0.005	0.71
M_D - MS logit	C	-0.0192	[-0.0263, -0.0120]	< 0.001	< 0.001	0.84
M_Q - MS logit	A	-0.1857	[-0.1998, -0.1716]	< 0.001	< 0.001	1.00
M_Q - MS logit	B	-0.0943	[-0.1058, -0.0829]	< 0.001	< 0.001	1.00
M_Q - MS logit	C	-0.0148	[-0.0222, -0.0073]	< 0.001	< 0.001	0.77
M_D - Oracle	A	+0.0131	[+0.0087, +0.0174]	< 0.001	< 0.001	0.10
M_D - Oracle	B	+0.0905	[+0.0789, +0.1021]	< 0.001	< 0.001	0.00
M_D - Oracle	C	+0.0572	[+0.0481, +0.0662]	< 0.001	< 0.001	0.00
M_Q - Oracle	A	+0.0182	[+0.0123, +0.0241]	< 0.001	< 0.001	0.10
M_Q - Oracle	B	+0.0194	[+0.0132, +0.0257]	< 0.001	< 0.001	0.10
M_Q - Oracle	C	+0.0616	[+0.0524, +0.0707]	< 0.001	< 0.001	0.00

Table 8. The contrast $M_Q - M_D$ by series length

DGP	T	Datasets	Mean $M_Q - M_D$	95% CI	Wilcoxon p	Fraction M_Q better
A	500	10	+0.0186	[+0.0059, +0.0313]	0.020	0.10
A	1000	31	+0.0052	[+0.0022, +0.0081]	< 0.001	0.26
A	2000	10	+0.0040	[+0.0022, +0.0059]	0.004	0.10
B	500	10	-0.0874	[-0.1054, -0.0695]	0.002	1.00
B	1000	31	-0.0711	[-0.0846, -0.0576]	< 0.001	0.97
B	2000	10	-0.0688	[-0.0791, -0.0585]	0.002	1.00
C	500	10	+0.0065	[-0.0039, +0.0169]	0.375	0.50
C	1000	31	+0.0044	[+0.0016, +0.0072]	0.007	0.32
C	2000	10	+0.0030	[-0.0002, +0.0062]	0.131	0.30

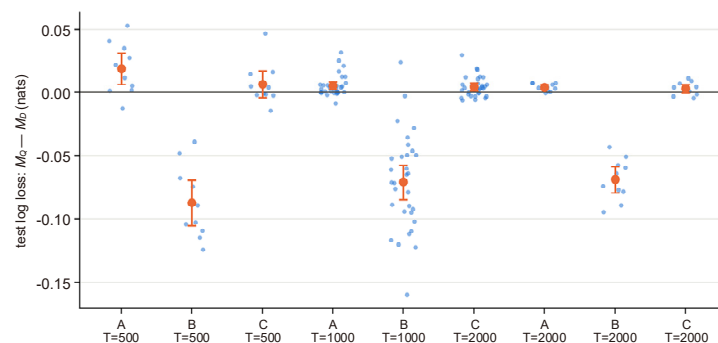
**Fig. 5.** Per-dataset differences $M_Q - M_D$ in test log loss. Small points are individual datasets; large points and bars are means with 95% intervals.

Table 9. Seed medians of the M_Q gauge invariants and evolution parameters ($T=1,000$)

DGP	Quantity	1	2	3	4	5
A	median $ z_i $	0.100	0.126	0.086	0.080	0.068
A	median $\cos(\arg z_1 - \arg z_i)$	—	-0.592	+0.905	+0.699	-0.497
A	median $\gamma (M_Q / M_D), \lambda$	0.106 / 0.114	0.226	(true value 0.12, 0.08)		
B	median $ z_i $	0.562	0.499	0.324	0.231	0.276
B	true	0.550	0.539	0.292	0.224	0.180
B	median $\cos(\arg z_1 - \arg z_i)$	—	+0.386	-0.817	+0.408	+0.504
B	true	—	+0.371	-0.857	+0.447	+0.832
B	median $\gamma (M_Q / M_D), \lambda$	0.101 / 0.247	0.011	(true value 0.12, 0.08)		
C	median $ z_i $	0.291	0.298	0.396	0.353	0.290
C	median $\cos(\arg z_1 - \arg z_i)$	—	-0.650	-0.210	+0.429	-0.251
C	median $\gamma (M_Q / M_D), \lambda$	3.000 / 3.000	0.000	—		

에 가깝고 규제 10^{-2} 가 가장 자주 선택된 사실이 이를 뒷받침한다. 두 필터는 시차 2 입력 로짓보다 0.10 nat, 2체제 전환 로짓보다 0.19 nat 낮은 손실을 얻었다. 부호 자유 대각 필터가 기억을 갖는 고전 모형이고, 같은 다섯 입력의 무기억 또는 짧은 기억 로짓이 그 기억을 대체하지 못한다는 뜻이다. 가설 H1의 귀무가설이 참일 때 비대각 확장은 허위 이득을 만들지 않았다.

B(비대각 생성과정). M_Q 는 M_D 보다 0.071 nat 낮은 손실을 얻었고(95% 구간, 31개 자료 중 30개에서 우위), 오라클과의 차이는 0.019 nat이었다. 세 길이 모두 같은 방향이며 $T=500$ 에서도 10개 자료 모두에서 M_Q 가 나왔다. 게이지 불변량은 큰 연결계수에서 참값을 복원했다. $|z_1|$ 0.56(참 0.55), $|z_2|$ 0.50(0.54), $|z_3|$ 0.32(0.29), $|z_4|$ 0.23(0.22), 상대 위상 코사인 +0.39(참 +0.37), -0.82(-0.86), +0.41(+0.45). 가장 작은 연결계수 $|z_5|$ (참 0.18)는 0.28로 과대추정되고 그 상대 위상(+0.50, 참 +0.83)도 부정확했는데, 이는 약한 신호의 위상이 유효 표본 50~240 관측으로는 식별되기 어렵다는 5.4절의 논의와 일치한다. λ 는 0.011로 참값 0.08보다 작게 추정되었고 γ 는 0.10(참 0.12)이었다. 반면 M_D 는 γ 를 0.25로 과대추정하여 표현할 수 없는 비대각 기억을 빠른 망각으로 대체했다.

이득의 해석에는 두 단서가 붙는다. 첫째, 시차 2 입력 로짓은 M_D 보다 0.035 nat 낮은 손실을 얻어 M_Q 의 이득 0.071 nat의 약 절반을 고전적 시차 향으로 재현했고, M_Q 는 그 로짓보다 0.037 nat 낮았다(87%의 자료에서 우위). 따라서 순수한 비대각 구조의 순이익은 0.04 nat 수준이다. 둘째, 2체제 전환 로짓은 M_D 보다

도 0.023 nat 나뉘고 시작점의 절반만 최적값에 도달했다. 잠재 체제 모형 하나의 실패가 모든 고전 잠재상태 모형의 실패를 뜻하지는 않으며, 11.1절에서 논의하듯 밀도행렬을 실수 좌표로 펼친 고전 상태공간 모형은 정의상 같은 경로를 표현한다. B의 결과는 “생성과정이 비대각 구조를 가질 때 그 구조를 가정한 추정기가 그것을 복원한다”는 정합성 검사이지 실제 자료에 그러한 구조가 있다는 증거가 아니다.

C(고전 잠재 체제). 모든 모형이 균등 예보보다 0.09~0.11 nat 낮은 데 그쳤고 오라클과의 격차 0.06 nat은 관측되지 않는 체제를 입력만으로 복원할 수 없음을 반영한다. 정답 라벨로 체제확률을 갱신하는 2체제 전환 로짓도 M_D 보다 0.019 nat 나뉘었다. M_Q 는 M_D 보다 0.004 nat 나뉘고(A와 같은 크기의 추가 보수 비용) 어떤 길이에서도 허위 이득을 만들지 않았다. 두 필터 모두 γ 가 상한 3에 도달하여 사실상 무기억 로짓으로 수렴했는데, 이는 생성과정의 기억이 입력이 아닌 잠재 체제에만 있기 때문이며 현재 입력 로짓과 M_D 의 손실이 0.001 nat 안에서 같다는 사실이 이를 확인한다.

가설과의 대응. H1은 B에서만 지지되고 A·C에서는 기각되지 않았으며 이는 설계된 대로다. 시드 간 대응 검정은 독립 반복에 대한 것이므로 10.2절의 조건을 충족한다. 실제 자료에 대한 H1은 미검증이며 이 절의 어떤 수치도 전장 판단의 증거로 사용하지 않는다. 실험이 보여주는 것은 세 가지다. 추정기는 자신이 가정한 구조를 복원하되 약한 연결계수의 위상은 식별하지 못한다. 구조가 없을 때 정보원당 두 개의 추가 자유도가 치르는 비용은 $T=1,000$ 에서 0.005

Table 10. Staged validation framework and current status

Stage	Object	Evidence required	Status
1	Mathematical and numerical consistency	Proofs, independent computation, numerical error	Checked in this paper
2	Data contract and recovery	Replay of event, correction, and access history	Minimal selection checks only
3	Probabilistic forecasting (synthetic)	Recovery of known structure, absence of spurious gain	Performed in Section 9
4	Probabilistic forecasting (real)	Independent test forecasts and observed outcomes	Not performed
5	Human-AI collaboration	Controlled situation assessment, participant studies	Not performed
6	System resilience and operational applicability	Tests under delay, disconnection, and error; staged evaluation in accredited environments	Not performed

nat 수준이고 표본이 커질수록 줄어든다. 그리고 시차 입력을 가진 단순 고전 모형이 비대각 이득의 절반을 대체하므로, 실제 자료의 비교에서는 10.2절의 기준선에 시차 입력 로짓을 반드시 포함해야 한다.

10. 지휘관 의사결정 지원의 검증 설계

10.1 단계별 검증과 주장 범위

검증은 계산이 작동한다는 사실과 판단에 도움이 된다는 사실을 구분한다. Table 10은 단계별로 필요한 증거와 현 상태다.

10.2 확률 예측과 기준모형

동일 임무 질문·목표시각·정보 마감시각의 예측을 비교한다. 최소 기준선은 균등확률, 훈련 빈도, 순차 갱신 빈도, 같은 입력의 규제 다항 로짓, 시차 입력 로짓, 부호 자유 대각 필터이며, 9장의 결과에 따라 시차 입력 로짓과 잠재상태 모형을 반드시 포함한다. 주요 손실은 자연로그 로그손실과 합산형 Brier 점수다[19].

$$L_{\log}(p, y) = -\log p_y,$$

$$L_B(p, y) = \sum_{k \in \{S, L, I\}} (p_k - \mathbf{1}\{y = k\})^2. \quad (17)$$

정확도 하나로 확률의 품질을 대신하지 않고 보정 곡선, 구간별 표본 수, 손실의 불확실성, 보류율, 조건별 실패를 함께 보고한다. 시간순 분할에서는 목표 종료시각과 정답 공개시각이 다음 예측보다 이전인지 확인하고, 중첩 목표나 같은 상황의 반복 갱신을 독립 표본으로 세지 않으며, 표준화·공분산·특징 선택·확률 보정은 훈련 또는 내부 검증에서만 결정한다. 실제

시계열의 손실 비교에는 장기분산 추정이나 블록 재표본화가 필요하다[20].

10.3 반증 가능한 연구 가설

H1 비대각 확장의 추가 예측력; 동일 입력과 계산 예산에서 M_Q 가 M_D 와 시차 입력 로짓보다 독립 시험 로그손실을 개선하는지 검증한다. 합성자료에서는 9장의 결과가 이를 조건부로 지지한다. 실제 자료에서 개선이 확인되지 않으면 단순 모형을 우선 후보로 유지하며, 유의하지 않다는 결과를 전장이 고전적이라는 증명으로 해석하지 않는다.

H2 정보 품질 관리의 기여; 중복 제거, 원천 계보, 이용 가능시각, 노후화, 결측 처리를 각각 제거한 절제실험을 수행한다.

H3 인간-AI 협업의 기여; 동일 과제에서 AI 미사용, 근거 목록만 제공, 확률과 근거 제공 조건을 비교한다. 목표는 사실 이해, 불확실성 판단, 적절한 수정·보류의 개선이며 AI 동의율은 성공 지표가 아니다.

H4 순서 민감도의 유용성과 취약성; 시간적으로 의미 있는 순서와 의미상 교환 가능한 보고의 표시 순서를 구분하고, 후자에서 불필요하게 큰 출력 차이가 발생하면 안정성 문제로 평가한다.

H5 연결 저하 시 재현 가능성; 지연·중복·단절·정정 조건에서 과거 판단 보존과 새 판단의 재현을 검사하고, 복구 성공률·복구 시간·미처리 사건 수·오류의 가시성을 측정한다.

10.4 참여자 연구와 체계 평가

초기 참여자 연구는 공개 가능한 가상 상황평가 과제를 사용하고, 정답·이용 가능한 보고·제한 시간·채

점 기준을 실험 전에 고정한다. 평가 항목은 핵심 사실의 파악, 상충 보고 탐지, 정보 공백 인식, 자신감과 정확도의 일치, 적절한 수정·보류, 소요시간과 주관적 부담이며, 틀린 AI 제안이 제시된 조건에서 이를 발견하고 수정할 수 있는지도 측정한다[7]. 참여자 수는 사전 효과크기 가정과 분석 단위에 따른 검정력으로 정하고, 사전 계획과 분석 변경 이력을 남기되 파일의 날짜와 시드만으로 외부 사전등록이 완료되었다고 표현하지 않는다. 체계 평가에서는 보고에서 추출한 사건·시간·출처의 정확도, 상충 보고의 누락, 정정 연결, 모형 오류 탐지, 설명의 원문 일치, 접근 제한 적용을 따로 측정하며, 통신 저하와 무결성 문제의 시험은 승인된 분리 환경에서 고정된 오류 조건으로 수행한다.

11. 논의와 결론

11.1 연구 모형의 의미

QFSVDM의 핵심은 다양한 자료를 하나의 큰 점수로 만드는 데 있지 않다. 특정 판단 질문을 정의하고, 당시 이용 가능한 근거를 고정하며, 확률과 상태 갱신의 규칙을 공개하는 데 있다. 비대각 상태는 순차 정보의 영향을 더 넓은 내부 공간에 보존하는 표현이지만, 그 복잡도는 자료가 뒷받침할 때에만 정당화된다. 밀도행렬을 실수 좌표로 펼치면 유한차원 비선형 상태공간 모형이므로 일반적인 고전 잠재상태 모형으로 같은 현상을 나타낼 수 없다고 주장하지 않는다. 9장의 결과는 이 점을 실제로 보여 준다. 비대각 구조가 있을 때 M_Q 는 그 구조를 복원하지만, 시차 입력 로짓도 그 이득의 일부를 재현하므로 실질적 질문은 특정 연산자 제약이 해석·규제·표본 효율성·재현에 도움이 되는가이다.

11.2 JADC2 적용의 함의

전영역 연결은 지휘관의 정보 접근을 넓히지만 서로 다른 시간과 의미를 가진 보고를 함께 이해하는 문제를 자동으로 해결하지 않는다. 본 논문은 감지에서 이해로 넘어오는 자료의 제약과, 이해에서 인간의 판단으로 넘어가는 확률·근거·제약의 연결을 명시했다. 지휘관이 마지막 판단 권한을 가진다는 원칙을 수정·

보류·감사 기록의 구조로 나타내는 것이 목표다. 향후 적용은 기존 영역인식·지휘통제 체계의 성능과 책임구조를 확인하면서 단계적으로 이루어져야 하며, 개정 필터는 기존 알고리즘이나 지휘절차를 즉시 대체하도록 설계된 체계가 아니다. 연합 파트너 간에는 자료의 공유 가능성과 같은 질문에 대한 의미적 적합성을 먼저 확인해야 한다.

11.3 한계

현재의 검증은 명시된 수식, 참조 구현, 추상 합성자료의 반복 적용에 한정된다. 합성 실험은 추정기가 자신이 가정한 구조를 복원하는지를 검사할 뿐 실제 군사자료에 그 구조가 있는지에 대해서는 아무것도 말하지 않는다. 생성과정은 다섯 입력·단일 지평·정지모수의 단순한 설계이고, 비교한 고전 잠재상태 모형은 2체제 전환 로짓 하나이므로 더 큰 고전 모형군이 같은 이득을 내지 못한다는 결론은 도출되지 않는다. 실제 군사자료의 적합, 지휘관 참여 실험, JADC2 인터페이스 연동, 연결 저하 환경의 전체 복구 서비스, 작전효과 검증은 수행하지 않았다. 임무 여건의 정답은 쉽게 정의되지 않을 수 있고 지휘관의 행동이 이후 결과를 바꿀 수 있으므로, 개입·관측 선택·검열을 무시하면 예측과 의사결정 효과를 혼동할 수 있다.

11.4 결론

본 연구는 전영역 인식과 JADC2 환경에서 지휘관의 AI 기반 의사결정을 지원하기 위한 QFSVDM을 정의하고, 정보 출처와 이용 가능시간, 임무 질문, 내부 상태, 출력 확률, 근거 설명, 인간의 판단을 구분했다. 부호 자유 고전 필터와 비대각 확장의 관계, 대각 충분조건 아래의 정확한 환원, 회전 게이지와 켈레 대칭에 따른 식별 가능 불변량을 수학적으로 명시했다. 실제 수행한 계산은 항등식과 참조 구현의 일치, 최소 사건 선택 제약, 그리고 합성자료 153개에서의 반복 적합이며, 비대각 확장은 생성과정에 비대각 구조가 있을 때만 이득을 내고 구조가 없을 때 허위 이득을 만들지 않았으며 시차 입력 로짓이 그 이득의 일부를 재현했다. 이러한 결과는 전장 판단의 우월성을 입증하지 않는다. 제안 모형의 가치는 동일한 정보 제약에서 더 신뢰할 수 있는 확률 판단을 제공하고 지휘관이

상충과 공백을 이해하며 적절히 수정·보류하도록 돕는지에 달려 있으며, 향후 검증은 예측 정확도뿐 아니라 인간의 이해, 책임 있는 사용, 연결 저하 시의 재현 가능성을 함께 평가해야 한다.

참고문헌

- [1] U.S. Department of Defense, Summary of the Joint All-Domain Command and Control Strategy, 2022. 공식 원문.
- [2] NATO Allied Command Transformation, Multi-Domain Operations. 공개 설명, 2026-09-06 확인. 공식 원문.
- [3] M. R. Endsley, 'Toward a Theory of Situation Awareness in Dynamic Systems,' *Human Factors*, VOL. 37, NO. 1, 1995, pp. 32-64. <https://doi.org/10.1518/001872095779049543>
- [4] U.S. Department of Defense, Chief Digital and Artificial Intelligence Office, 2023 Data, Analytics, and Artificial Intelligence Adoption Strategy — Fact Sheet (2023-11-02), 2023. 공식 원문.
- [5] NATO, Summary of NATO's Revised Artificial Intelligence (AI) Strategy, July 10, 2024. 공식 원문.
- [6] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0), NIST AI 100-1, 2023.
- [7] John D. Lee & Katrina A. See, 'Trust in Automation: Designing for Appropriate Reliance,' *Human Factors*, VOL. 46, NO. 1, 2024, pp. 50-80. https://doi.org/10.1518/hfes.46.1.50_30392
- [8] Leslie Pack Kaelbling, Michael L. Littman, & Anthony R. Cassandra, 'Planning and Acting in Partially Observable Stochastic Domains,' *Artificial Intelligence*, VOL. 101, NO. 1-2, 1998, pp. 99-134. [https://doi.org/10.1016/S0004-3702\(98\)00023-X](https://doi.org/10.1016/S0004-3702(98)00023-X)
- [9] Jerome R. Busemeyer & Peter David Bruza, *Quantum Models of Cognition and Decision* (2nd ed.), Cambridge University Press, 2025.
- [10] Zheng Wang, Tyler Solloway, Richard M. Shiffrin, & Jerome R. Busemeyer, 'Context Effects Produced by Question Orders Reveal Quantum Nature of Human Judgments,' *Proceedings of the National Academy of Sciences*, VOL. 111, NO. 26, 2014, pp. 9431-9436. <https://doi.org/10.1073/pnas.1407756111>
- [11] T. Baumgratz, M. Cramer, & M. B. Plenio, 'Quantifying Coherence,' *Physical Review Letters*, VOL. 113, 2014, Article 140401. <https://doi.org/10.1103/PhysRevLett.113.140401>
- [12] Olivier Ledoit & Michael Wolf, 'A Well-Conditioned Estimator for Large-Dimensional Covariance Matrices,' *Journal of Multivariate Analysis*, VOL. 88, NO. 2, 2024, pp. 365-411. [https://doi.org/10.1016/S0047-259X\(03\)00096-4](https://doi.org/10.1016/S0047-259X(03)00096-4)
- [13] Masanao Ozawa, 'Quantum Measuring Processes of Continuous Observables,' *Journal of Mathematical Physics*, VOL. 25, NO. 1, 1984, pp. 79-87. <https://doi.org/10.1063/1.526000>
- [14] G. Lindblad, 'On the Generators of Quantum Dynamical Semigroups,' *Communications in Mathematical Physics*, VOL. 48, 1976, pp. 119-130. <https://doi.org/10.1007/BF01608499>
- [15] Vittorio Gorini, Andrzej Kossakowski, & E. C. G. Sudarshan, 'Completely Positive Dynamical Semigroups of N-Level Systems,' *Journal of Mathematical Physics*, VOL. 17, NO. 5, 1976, pp. 821-825. <https://doi.org/10.1063/1.522979>
- [16] R. B. Davies, 'Hypothesis Testing When a Nuisance Parameter Is Present Only Under the Alternative,' *Biometrika*, VOL. 64, NO. 2, 1977, pp. 247-254. <https://doi.org/10.1093/biomet/64.2.247>
- [17] A. Lebedev & A. Khrennikov, 'Quantum-Like Modeling of the Order Effect in Decision Making: POVM Viewpoint on the Wang-Busemeyer QQ-Equality,' *In Infinite Dimensional Analysis, Quantum Probability and Related Topics (QP-PQ: Quantum Probability and White Noise Analysis)*. World Scientific. preprint v2: arXiv:1811.00045v2. https://doi.org/10.1142/9789811275999_0010
- [18] Thomas Boyer-Kassem, Sébastien Duchêne, & Eric Guerci, 'Testing Quantum-Like Models of Judgment for Question Order Effect,' *Mathematical Social Sciences*, VOL. 80, 2016, pp. 33-46. <https://doi.org/10.1016/j.mathsocsci.2016.01.001>
- [19] Tilmann Gneiting & Adrian E. Raftery, 'Strictly Proper Scoring Rules, Prediction, and Estimation,' *Journal of the American Statistical Association*, VOL. 102, NO. 477, 2007, pp. 359-378. <https://doi.org/10.1198/016214506000001437>
- [20] Dimitris Politis & Joseph P. Romano, 'The Stationary Bootstrap,' *Journal of the American Statistical Association*, VOL. 89, NO. 428, 1994, pp. 1303-1313. <https://doi.org/10.1080/01621459.1994.10476870>