

변수의 물리적 특성과 결측 패턴을 고려한 파랑 관측자료 결측보간 딥러닝 모델의 선택적 적용성 평가

Selective Applicability of Deep Learning Models for Imputation of Missing Wave Observation Data: Effects of Variable Physical Characteristics and Missing-Data Patterns

마경림* · 김종범** · 이승호*** · 최홍배****

Kyoungrim Ma*, Jongbeom Kim**, Seungho Lee*** and Heung-Bae Choi****

요지 : 파랑 관측자료의 결측보간에 가장 적합한 인공지능 모델은 대상 변수의 물리적 특성과 결측 패턴에 따라 달라질 수 있다. 본 연구에서는 남동해안 5개 관측소의 30분 간격 파랑 관측자료를 대상으로, 시간 중심 모델인 SAITS와 시공간 모델인 GRIN, SPIN, ImputeFormer를 동일한 자료, 데이터 분할, 시퀀스 및 마스크 조건에서 비교하였다. 모델 간 성능 차이와 통계적 동등성은 윈도우 단위 paired bootstrap 95% 신뢰구간을 통해 판정하였다. 복원 대상 관측소의 마지막 시점을 인위적으로 결측 처리한 평가에서, 유의파고와 유의파주기는 SAITS가 가장 낮은 오차를 보였으며, 순환 변수인 파향은 SPIN이 가장 낮은 각도오차를 나타냈다. 결측 패턴 민감도 실험에서는 무작위 결측의 경우 결측률 변화에 따른 모델 간 성능 차이가 전반적으로 작았다. 반면 윈도우 내 연속 결측에서는 결측률이 20% 이상일 때 유의파고와 유의파주기 모두에서 GRIN이 가장 우수한 성능을 보여, 무작위 결측과는 다른 성능 양상을 보였다. 특히 50% 연속 결측 조건에서 선형보간보다 낮은 오차를 보인 딥러닝 모델은 유의파고와 유의파주기 모두에서 GRIN이 유일하였다. 이러한 결과는 결측보간 성능이 전체 결측률보다 결측의 시간적 연속성에 더 민감할 수 있음을 보여준다. 특히 연속 결측률이 20% 이상인 조건에서 GRIN이 유의파고와 유의파주기 보간에 우수한 성능을 보인 결과는, 연속적인 누락 구간을 복원할 때 시공간 의존성을 활용하는 것이 효과적일 수 있음을 시사한다.

핵심용어 : 파랑 관측자료, 결측보간, 딥러닝, 변수 물리특성, 결측 패턴

Abstract : The optimal artificial intelligence model for imputing missing wave observation data may vary according to the physical characteristics of the target variable and the missingness pattern. This study compared a temporal model, SAITS, with spatiotemporal models, GRIN, SPIN, and ImputeFormer, using 30-min interval wave observations from five stations along the southeastern coast of Korea. All models were evaluated under identical data, data-splitting, sequence, and masking conditions. Performance differences and statistical ties between model pairs were assessed using window-level paired bootstrap 95% confidence intervals. In an evaluation in which the final time step at the target station was artificially masked, SAITS yielded the lowest errors for significant wave height and significant wave period, whereas SPIN achieved the lowest angular error for wave direction, a circular variable. In the missingness-pattern sensitivity experiments, performance differences among models were generally small across random missingness rates. In contrast, under contiguous missingness within a window, GRIN achieved the best performance for both significant wave height and significant wave period when the missingness rate was at least 20%, exhibiting a pattern distinct from that observed under random missingness. Notably, under 50% contiguous missingness, GRIN was the only deep learning model that produced lower errors than linear interpolation for both significant wave height and significant wave period. These findings indicate that imputation performance may be more sensitive to the temporal continuity of missing values than to the overall missingness rate. They further suggest that exploiting spatiotemporal dependencies can be effective for reconstructing consecutive missing segments, particularly when the contiguous missingness rate exceeds 20%.

Keywords : wave observation data, missing data imputation, deep learning, physical characteristics of variables, missing data pattern

*(주)지오시스템리서치 전임연구원(Associate Research Engineer, GeoSystem Research Corp.)

***(주)지오시스템리서치 책임연구원(Principal Research Engineer, GeoSystem Research Corp.)

***국립해양조사원 주무관(Administrative Officer, Korea Hydrographic and Oceanographic Agency)

****(주)지오시스템리서치 책임연구원(Corresponding author: Heung-Bae Choi, Principal Research Engineer, GeoSystem Research Corp., 306, Hanlim Human Tower, 172 LS-ro, Gunpo-si, Gyeonggi-do 15807, Korea, Tel: +82-31-5180-5862, hbchoi@geosr.com)

1. 서 론

파랑 관측자료는 연안 재해 대응과 해상 안전을 위한 예측 및 의사결정의 핵심 기초자료이다. 유의파고, 유의파주기, 파향과 같은 파랑 변수는 항만 운영, 연안 구조물 설계, 해수욕장 이안류(rip current) 감시 및 경보 산정에 직접 활용된다. 특히 이안류는 파랑 조건 변화에 민감하게 발생하며, 국내에서도 해운대해수욕장을 중심으로 발생기구 규명과 관측 기반 탐지 연구가 축적되어 왔다(Kim et al., 2011; Oh et al., 2019). 또한 파랑 예보를 기반으로 이안류 발생 가능성을 진단하는 연구도 이루어지고 있다(Choi, 2023). 이때 예보 및 진단의 신뢰도는 입력자료인 파랑 관측자료의 연속성과 완결성에 크게 좌우된다. 그러나 실제 관측망에서는 센서 고장, 통신 장애, 정기 점검 등으로 결측이 상시 발생한다. 결측이 포함되면 시계열 모델의 입력 시퀀스가 성립하지 않아 예측·경보가 중단될 수 있으므로, 결측 구간을 물리적으로 타당하게 복원하는 결측보간(imputation)은 연안 감시 체계의 전제 조건이 된다. 최근 관측망 확충과 계산 자원 발전으로 국내 해안·해양 분야에서도 자료 기반 인공지능(AI) 모델의 적용이 빠르게 확대되고 있으며(Lee et al., 2020; Kim and Lee, 2022), 파랑 변수에 대한 딥러닝 예측 연구도 보고되고 있다(Kim, 2020; Park and Kang, 2024).

다변량 시계열 결측보간을 위한 딥러닝 방법은 순환신경망 계열에서 출발하였다. BRITS는 양방향 순환구조로 결측값 자체를 학습변수로 두어 갱신하는 방식을 제시하였고(Cao et al., 2018), GAIN은 생성적 적대신경망을 결측보간에 적용하였다(Yoon et al., 2018). 이후 불규칙 표본을 다루는 연속시간 어텐션(mTAN; Shukla and Marlin, 2021)과 확률적 확산 모델 기반 보간(CSDI; Tashiro et al., 2021)으로 방법론이 확장되었다. 이 흐름에서 SAITS는 순환구조 없이 순수 self-attention만으로 시간 의존성과 특징 상관을 포착하여 결측보간에서 높은 정확도를 보고하였다(Du et al., 2023). SAITS는 시간축 중심 모델로, 관측소 간의 공간 구조를 명시적으로 입력하지 않는다는 특성을 가진다. 한편 센서망처럼 관측 지점이 공간적으로 연결된 자료에서는, 지점 간 관계를 그래프로 표현하여 이웃 관측을 활용하는 시공간 그래프 신경망(spatiotemporal graph neural network)이 제안되었다. GRIN은 양방향 순환 그래프 신경망으로 이웃 노드의 관측을 메시지 전파하여 결측을 복원하며(Cini et al., 2022), SPIN은 유효 관측만을 참조하는 희소 시공간 어텐션으로 오차 전파 문제를 완화한다(Marisca et al., 2022). 또한 ImputeFormer는 사전 정의된 그래프 없이 관측소 임베딩으로 공간적 상관관계를 학습하는 트랜스포머로 시간·공간 어텐션을 결합한다(Nie et al., 2024). 이들은 공통적으로 공간 정보의 활용이 결측 복원에 기여할 수 있음을 시사한다. 그러나 국내 해안·해양 분야의 파랑 관련 AI 연구는 대체로 단일 지점·단일 변수(주로 유의파고)에 대해 특정 모델의 우수성을 보이는 데 집중해 왔

으며(Kim, 2020; Park and Kang, 2024), 예측(forecasting)이 중심이어서 다지점 관측의 공간 정보를 활용한 결측보간, 그리고 복원 대상 변수의 물리적 특성이나 결측이 나타나는 패턴에 따라 어떤 모델이 적합한지를 체계적으로 비교한 연구는 상대적으로 드물다. 즉, 서로 다른 귀납적 편향을 가진 시간 중심 모델(SAITS)과 공간·그래프 모델(GRIN, SPIN, ImputeFormer)이 파랑 변수별로, 또 결측 형태별로 일관되게 우열이 유지되는지는 충분히 검증되지 않았다.

본 연구는 남동해안 5개소 파랑 관측자료(임랑·송정·해운대·울산·거제, 복원 대상 임랑·송정·해운대)를 대상으로, 동일한 자료·분할·시퀀스·평가 조건에서 시간 중심 모델(SAITS)과 시공간 그래프·어텐션 모델(GRIN, SPIN, ImputeFormer) 네 가지를 결측보간 과제로 학습·비교한다. 각 모델은 원저자가 공개한 소스코드를 바탕으로, 본 연구의 자료 구성과 평가 조건에 맞게 적용하였다(Du et al., 2023; Cini et al., 2022; Marisca et al., 2022; Nie et al., 2024).

본 연구의 가설은 다음과 같다. 파랑 결측보간에서 최적 모델은 단일하게 고정되지 않으며, (1) 복원 대상 변수의 물리적 특성과 (2) 결측이 나타나는 패턴에 따라 선택적으로 달라진다. 이를 검증하기 위해 두 축의 비교를 수행한다. 첫째, 변수 물리특성 축에서는 스칼라 크기 변수(유의파고·유의파주기)와 순환 변수(파향)를 구분하여 비교한다. 둘째, 결측 패턴 축에서는 무작위형(random) 결측과 원도 내 연속형(contiguous, 최대 12시간) 결측을 결측률별로 비교한다. 두 축의 조합에 따라 모델 간 우열이 어떻게 달라지고 역전되는지가 본 연구의 핵심 검증 대상이며, 구체적 수치와 반전 양상은 4장에, 종합 결과는 5장에 제시한다(자료의 공간상관 구조는 2.2절 참조).

따라서 본 연구의 차별성은 “어떤 모델이 항상 최적인가”를 제시하는 데 있지 않고, “어떤 물리적 특성을 가진 변수를 어떤 결측 패턴에서 복원할 때 어떤 모델을 선택해야 하는가”에 대한 정량적 선택 기준을 제시하는 데 있다. 이는 단일 변수 또는 단일 상황에 초점을 둔 기존 국내 파랑 AI 연구(Kim, 2020; Park and Kang, 2024)를 변수 물리특성과 결측 패턴의 교차 비교로 확장한다는 점에서 의의가 있다. 다만 본 연구는 임랑, 송정, 해운대 관측소와 2026년 1~7월의 시험 기간을 대상으로 하며, 5개 노드 완결 시퀀스만을 사용하는 자료 구성 규칙과 시험 구간 결측 편중으로 인해 완결 시험 원도 수는 변수별 106~123개이다. 이에 따라 1~2위 모델 간 성능 차이가 근소한 경우에는 통계적 동률로 해석하였으며, 결과의 일반화는 본 자료가 뒷받침하는 범위로 한정하였다. 자료 특성과 한계는 2장에, 결과 해석은 4장에 제시하였다.

2. 대상 자료 및 물리적 특성

2.1 대상 지역 및 자료 개요

본 연구는 부산·울산 연안의 파랑 관측망 5개소(임랑, 송정,

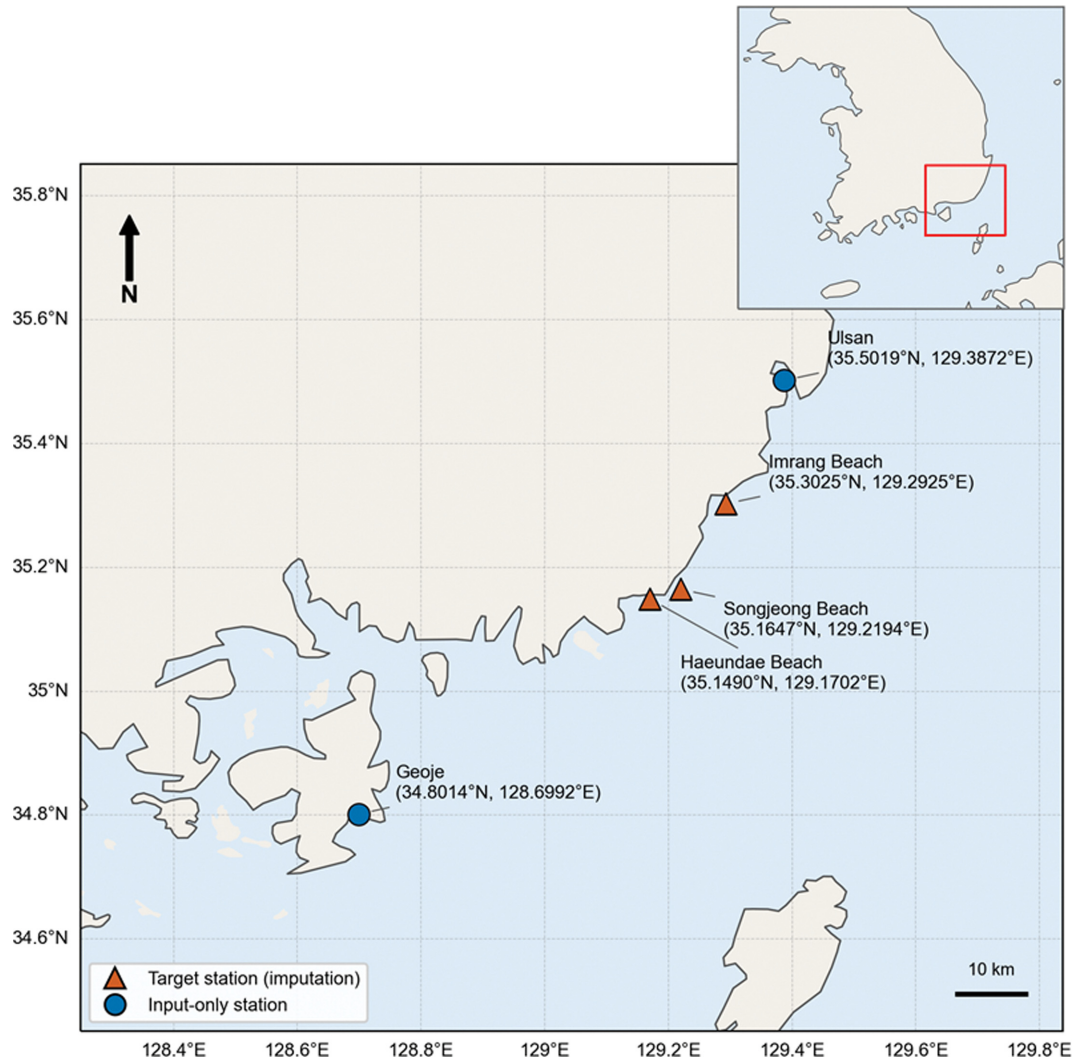


Fig. 1. Location map of the five wave observation stations along the southeastern coast of Korea (triangles: imputation targets — Imrang Beach, Songjeong Beach, Haeundae Beach; circles: input-only stations — Ulsan, Geoje).

해운대, 울산, 거제)를 대상으로 수행하였다. 이 중 임랑, 송정, 해운대 3개소를 결측 복원 대상(target)으로 설정하고, 울산과 거제를 포함한 5개 관측소의 자료를 모델 입력 노드로 사용하였다. 파랑 관측자료는 국립해양조사원(KHOA)의 바다누리 해양정보서비스(<https://www.khoa.go.kr/oceangrid/>)와 기상청(KMA)의 기상자료개방포털(<https://data.kma.go.kr/>)에서 제공하는 연안 파고부이 관측자료를 활용하였으며, 분석 기간은 2021년 1월 1일 00:00부터 2026년 7월 31일 23:30까지이다. 두 기관의 자료를 동일한 시간 해상도로 활용하기 위해 30분 간격으로 자료를 구성하였다. 또한 원자료에는 임의의 보간이나 추정을 수행하지 않았으며, 실제 결측을 포함하는 시퀀스는 학습 및 평가에서 제외하였다. 모델의 결측 복원 성능은 인위적인 마스킹(artificial masking)을 적용하여 평가하였다. 학습자료의 품질 확보를 위해 물리적 범위 기반 품질관리(QC)를 수행하였으며, 유의파고는 0~8 m, 유의파주기는 0~20 s, 파향은 0~360° 범위를 벗어나는 자료를 이상치로 판단하여 제거하였다. 관측소별 위경도와 역할은 다음과

같다. 복원 대상(target)은 임랑(35.3025°N, 129.2925°E), 송정(35.1647°N, 129.2194°E), 해운대(35.1490°N, 129.1702°E)이고, 입력 보조 관측소는 울산(35.5019°N, 129.3872°E)과 거제(34.8014°N, 128.6992°E)이다. 5개 관측소의 위치는 Fig. 1에 제시하였다.

2.2 대상 변수와 물리적 특성

복원 대상 변수는 유의파고, 유의파주기, 파향의 세 가지이며, 물리적 성격이 서로 다르다. 이 물리적 성격의 차이가 4장의 모델 성능 해석의 기준이 된다. 변수별 물리·통계 특성은 다음과 같다. 유의파고와 유의파주기는 방향성이 없는 양의 스칼라 연속량으로 노드당 1채널로 표현하며, 값 범위는 각각 0.01~8.00 m와 1.70~18.60 s이고, 파향은 0~360°의 순환량으로 노드당 $u = \cos(\theta)$, $v = \sin(\theta)$ 의 2채널로 분해하여 표현하며, 값 범위는 0~360°이다.

파향은 0°와 360°가 동일한 순환 변수이므로 각도 값을 그대로 회귀 대상으로 쓰면 경계 불연속이 발생한다. 이를 피

하기 위해 파향을 단위벡터 $u = \cos(\theta)$, $v = \sin(\theta)$ 로 분해해 노드당 2채널로 학습하고, 복원 후 다시 각도로 환원하여 순환 각도오차 $\min(|a - b|, 360 - |a - b|)$ 로 평가한다.

공간 구조 측면에서, 결측이 관측소 간에 부분적으로만 동시에 발생한다는 점이 중요하다. target 3개소(임랑·송정·해운대) 유의파고 결측의 공간 동시성을 관측 시각을 합집합으로 구성한 공통 시간격자(97,824 시각)에서 집계하면, 전 시각의 87.75%는 세 관측소가 모두 관측 중이고, 8.87%는 한 개소만 결측(나머지 두 개소는 관측), 1.23%는 두 개소 결측, 2.15%만이 세 개소 동시 결측이었다. 관측소 쌍별 결측 동시성(Jaccard)은 임랑-송정 0.267, 임랑-해운대 0.231, 송정-해운대 0.343으로 0.23~0.34 범위였다. 즉 한 관측소가 결측일 때 인접 관측소가 관측 중일 확률이 높은 “부분 공간상관” 구조로, 공간정보를 활용하는 그래프 기반 모델이 순수 시간 모델 대비 이득을 볼 구조적 여지가 존재한다. 이 가능성은 4.2절·4.3절에서 실측 결과와 함께 검토한다.

결측률은 전 기간 기준 3.8~6.9%로 낮고 변수 간 균질하나, 시험 기간(2026-01-01~07-31)에는 관측소별 편차가 크다. 시험 데이터 세트에서 결측률은 임랑 약 3.1%, 해운대 약 5.5%인 데 비해 송정은 약 11.5%로 다소 높으며(유의파고·유의파주기·파향 공통), 입력 보조노드 울산은 유의파주기에 한해 시험 기간의 12.48%가 결측이나, 다른 입력 보조 노드인 거제의 결측률은 0.6% 미만이다. 5개 노드가 모두 완결되어야 시퀀스가 성립하는 규칙(2.3절) 때문에 관측소별 결측률의 편차는 시험에 활용 가능한 완결 윈도 수를 좌우한다. 이는 모델 간 우열 및 통계적 동률 판정 시 유효 시험 표본수와 검정력의 제약으로 작용할 수 있으므로, 관련 한계는 4장에서 논의하였다.

2.3 자료 전처리 및 시퀀스 구성

학습(Train)·검정(Validation) 및 시험(Test) 자료는 시간 순서를 보존하여 기간 기준으로 분할하였다. 2021년 1월부터 2025년 12월까지의 약 5년 자료를 학습·검정에, 자료 취득 종료 시점 기준 최근 7개월인 2026년 1월 1일부터 7월 31일까지의 자료를 시험에 사용하였다. 이를 통해 미래 정보의 유입을 방지하고, 과거 자료로 학습한 모델의 최신 미관측 기간에 대한 일반화 성능을 평가하였다. 시퀀스는 시간 공백을 무시하고 이어붙이지 않는 규칙으로 구성하였다. 먼저 연속 관측 구간에서 $n_steps = 48(30분 \times 48 = 24시간)$ 윈도를 생성하였다. 학습·검정 자료는 자료 활용도를 높이기 위해 $stride = 24$ 를 적용하였고, 시험 자료는 윈도 단위 paired bootstrap으로 95% 신뢰구간을 추정할 때 윈도 간 의존성과 동일 관측값의 중복 집계를 방지하기 위해 비중첩 윈도로 구성하였다. 이후 결측을 포함한 윈도를 제외하고, 남은 학습·검정 윈도 중 30%를 무작위로 검정 자료로 분리하였으며(seed 42), 검정 윈도에는 10% 랜덤 마스킹(seed 43)을 적용하였다.

정규화는 신규 비교 모델의 경우 어댑터 내부에서 per-

feature 표준화를 수행하되, 성능 집계는 전 모델 모두 raw 정답 스케일에서 계산하여 공정성을 확보하였으며, 파향 자료는 각도의 주기성을 고려하여 표준화 이전에 $u = \cos(\theta)$, $v = \sin(\theta)$ 형태의 직교 성분으로 변환하였다. 물리적 범위 검토를 통과한 학습 윈도는 유의파고 986개, 유의파주기 986개, 파향 856개이고, 검정 윈도는 유의파고 423개, 유의파주기 422개, 파향 367개였다. 변수별 비중첩 시험 완결 윈도는 유의파고 123개, 유의파주기 116개, 파향 106개였다. 송정(11.5%)과 해운대(5.5%)의 결측, 그리고 유의파주기에서 울산(12.5%)의 결측으로 일부 구간은 평가에서 제외되었다. 다만 모든 모델을 동일한 윈도와 결측 마스크 조건에서 평가하여 상대적 성능 비교의 공정성은 확보하였으며, 제한된 윈도 수에 따른 해석상 한계는 4.4절에서 논의한다.

3. AI 모델 및 실험 설계

3.1 비교 모델

동일한 데이터·분할·시퀀스·마스크 조건에서 결측 복원 성능을 비교하기 위해, 시간 의존성만을 주로 활용하는 모델부터 관측소 간 공간 관계를 명시적 또는 학습형 구조로 반영하는 시공간 모델까지 포괄하도록 4종의 모델을 선정하였다. 모델 선정의 기준은 첫째, 결측보간을 목적으로 제안된 대표적인 딥러닝 기반 모델일 것, 둘째, 시간 정보와 공간 정보를 활용하는 방식이 서로 구분될 것, 셋째, 동일한 다변량 시계열 입력 조건에서 비교 가능할 것, 넷째, 원저자 공개 구현을 통해 재현 가능한 실험 구성이 가능할 것이었다. 이에 따라 시간축 self-attention 기반 모델인 SAITS, 그래프 메시지 전달 기반 시공간 모델인 GRIN, 시공간 점별 attention 기반 모델인 SPIN, 학습형 노드 임베딩을 활용하는 Transformer 계열 모델인 ImputeFormer를 비교 대상으로 설정하였다. 각 모델은 임의의 재구현 없이 원저자(Du et al., 2023; Cini et al., 2022; Marisca et al., 2022; Nie et al., 2024)의 공개 저장소의 코드를 기반으로 입출력 텐서 규약만 공통 어댑터를 통해 통일하였다. 이는 모델 구조 자체의 차이를 비교 대상으로 유지하면서, 자료 구성이나 평가 절차의 차이로 인한 성능 편향을 줄이기 위한 것이다.

기준 모델인 SAITS(Du et al., 2023)는 시간축 self-attention 기반 결측보간 모델이다. 본 연구에서는 5개 관측소의 특징을 하나의 다변량 시계열로 입력하여 시간 의존성과 특징, 즉 관측소 간 상관을 attention으로 학습하도록 하였다. 다만 SAITS는 사전 정의된 그래프나 명시적 공간 구조를 사용하지 않으므로, 공간 관계를 구조적으로 모델링하지 않는 시간 중심 모델로 분류하였다. 따라서 SAITS는 관측소 간 공간 전파 효과를 직접 사용하지 않는 기준 모델로서, 시공간 모델의 추가 이점을 평가하기 위한 비교 기준 역할을 한다.

GRIN(Cini et al., 2022)은 양방향 recurrent 구조와 graph message passing을 결합한 시공간 그래프 모델이다. 시간축

으로는 forward 및 backward 방향의 순환 구조를 통해 양측 시간 문맥을 반영하고, 공간축으로는 그래프를 따라 이웃 관측소의 정보를 전달받아 결측값 복원에 활용한다. 따라서 GRIN은 연속 결측과 같이 단일 관측소의 시간 정보만으로 복원이 어려운 조건에서, 주변 관측소 정보를 통해 복원 안정성을 높일 수 있는 모델로 선정하였다.

SPIN(Marisca et al., 2022)은 시공간 점별(point-wise) attention 기반 모델이다. 관측소를 그래프 노드로 보되, 결측 위치를 복원 대상 target point로 두고 해당 위치와 관련된 시점 및 이웃 관측소 정보를 attention으로 참조한다. GRIN이 그래프를 따라 정보를 순차적으로 전파하는 구조라면, SPIN은 결측 위치를 중심으로 어느 관측소의 어느 시점 정보가 중요한지를 선택적으로 학습하는 구조에 가깝다. 따라서 SPIN은 파향과 같이 단일 시간 흐름이나 단순 거리 기반 전파만으로 설명하기 어려운 변수에서 attention 기반 시공간 참조의 효과를 검토하기 위해 포함하였다.

ImputeFormer(Nie et al., 2024)는 Transformer 기반 결측보간 모델로, 명시적 거리 그래프 대신 학습형 노드 임베딩을 사용하여 관측소별 잠재 특성과 공간 관계를 반영한다. 또한 시간축 interaction과 공간축 interaction을 결합하여 각 관측소-시점 위치의 hidden representation을 갱신하고, 최종 출력층을 통해 결측값을 복원한다. 이에 따라 ImputeFormer는 고정된 거리 기반 그래프가 실제 관측소 간 관계를 충분히 설명하지 못할 가능성을 고려하여, 데이터로부터 관측소 간 잠재 관계를 학습하는 모델로 선정하였다.

이처럼 네 모델은 결측보간에 활용하는 정보와 귀납적 편향이 서로 다르다. SAITS는 시간축 attention을 중심으로 다변량 시계열 내 의존성을 학습하고, GRIN은 위경도 거리 기반의 명시적 그래프를 따라 주변 관측소 정보를 전파한다. SPIN은 그래프 구조 안에서 결측 위치별로 중요한 시공간 관측값을 attention으로 선택하며, ImputeFormer는 학습형 노드 임베딩을 통해 관측소 간 잠재 관계를 반영한다. 따라서 본 연구의 모델 구성은 단순한 성능 순위 비교가 아니라, 시간 중심 구조, 명시적 그래프 기반 구조, attention 기반 시공간 구조, 학습형 공간표현 구조가 파랑 관측자료 결측보간에서 어떻게 다르게 작용하는지를 비교하기 위한 설계이다. 특히 2.2절에서 제시한 변수의 물리적 특성과 결측 공간상관 특성을 고려할 때, 스칼라 크기량인 유의파고·유의파주기와 순환 방향량인 파향에서 각 모델의 적합성이 다르게 나타날 것으로 예상된다.

그래프 기반 모델인 GRIN과 SPIN에 필요한 인접행렬은 관측소 위경도로부터 결정론적으로 구성하였다. 먼저 5개 관측소의 쌍별 haversine 거리(km)를 계산하고, 거리 기반 Gaussian kernel을 적용하여 관측소 간 가중치를 산정하였다. 이때 Gaussian kernel의 폭을 결정하는 σ 는 관측소 간 거리의 표준편차로 설정하였다. 이후 자기연결을 나타내는 대각 성분은 1.0으로 부여하고, 대칭 정규화(symmetric normalization)

를 적용하였다. 관측소 수가 5개로 적어 임계값 기반 희소화를 적용할 경우 일부 노드가 고립될 가능성이 있으므로, 임계값은 0으로 두어 완전연결 그래프를 유지하였다. 이때 사전 대칭 정규화와 GRIN 내부 diffusion 정규화가 함께 적용되므로, 그래프 가중치의 상대적 차이가 완화될 수 있다. 구성된 인접행렬은 모든 실험에서 고정하였으며, 그래프 기반 모델 간 비교의 공정성을 위해 GRIN과 SPIN에 동일하게 적용하였다.

3.2 학습·검정 방법

전 모델의 공정한 비교를 위해 데이터, 분할 방식, 시퀀스 구성, 검정 마스킹 조건은 모두 동일하게 설정하였다. 공통 학습 조건으로 최대 300 epoch, patience 15의 조기 종료를 사용하였고, 조기 종료는 검정 성능을 기준으로 적용하였다. 또한, 재현성 확보를 위해 모든 실험에서 seed를 고정하고 동일한 전처리 절차와 동일한 그래프 구조를 사용하였다. 학습률과 배치 크기는 모델별 원저장소의 기본값을 그대로 사용하였다. 이는 특정 모델에 유리하도록 하이퍼파라미터를 임의로 조정함으로써 발생할 수 있는 비교 편향을 배제하기 위한 것이다.

모델별 학습률·배치 크기와 변수별(유의파고/유의파주기/파향) best epoch는 다음과 같다. SAITS는 학습률 5×10^{-4} , 배치 64, 변수별 best epoch는 149/112/13. GRIN은 학습률 1×10^{-3} , 배치 32, 변수별 best epoch는 98/108/82, SPIN은 학습률 8×10^{-4} , 배치 8, 변수별 best epoch는 56/58/32, ImputeFormer는 학습률 1×10^{-3} , 배치 16, 변수별 best epoch는 29/75/48이었다.

SPIN에 입력되는 시각 공변량은 각 원도의 시작 시각을 기준으로 일주기와 연주기의 sin/cos 값을 결정론적으로 생성하여 주입하였다. 또한 고정된 시험 원도 정렬을 보장하기 위해 각 원도의 시작 시각을 시험 세트에 함께 저장하였다. 정규화 방식은 모델별 구현에 따라 차이가 있으나, 성능 평가는 모든 모델에 대해 동일한 시험 마스크와 원자료(raw) 정답 스케일에서 수행하였다. 따라서 최종 성능 지표는 모델 내부의 정규화 방식과 무관하게 동일한 기준에서 산출되었다.

3.3 성능평가 지표 및 실험 설계

성능평가 지표는 대상 변수의 특성에 따라 구분하였다. 유의파고와 유의파주기는 R^2 , RMSE, MAE를 사용하였고, 파향은 순환 변수의 특성을 고려하여 두 각도 차이의 최솟값인 $\min(|a - b|, 360 - |a - b|)(\text{deg})$ 를 각도오차로 사용하였다. 모든 지표는 3개 대상 관측소를 통합한 전체 결과(ALL)로 산출하였다. 최적 모델 선정은 비중첩 시퀀스에서 복원대상 관측소의 마지막 시점을 인위적으로 결측 처리(기준 평가) 결과를 따랐으며, 유의파고와 유의파주기는 RMSE 최소, 파향은 각도오차 최소를 기준으로 하였다. 기준 평가는 실시간 관측·경보 체계에서 최신 시점 관측값이 수신되지 않는 상황을

모사하기 위한 설계로, 전후 관측이 모두 존재하는 통상적 결측보간 벤치마크의 마스킹 방식과 구별된다. 이는 비중첩 시퀀스 중 결측이 포함된 원도를 제외한 완결 원도만을 사용하였으며, 각 **target** 관측소의 마지막 1개 시점을 인위적으로 결측 처리하였다. 이후 해당 위치의 예측값만을 이용하여 성능 지표를 산출하였다. 모든 모델은 동일한 평가 원도, 동일한 결측 마스크, 동일한 원자료(raw) 정답 스케일에서 평가하였다.

모델 간 우열과 통계적 동등 여부를 원도 단위 **paired bootstrap**을 통해 판정하였다. 완결 원도를 재표집 단위로 설정하고, 10,000회 재표집을 수행하였으며(seed 42), 비교 대상인 두 모델에는 동일한 재표집 인덱스를 적용하였다. 이를 통해 모델쌍 간 지표 차이의 95% 신뢰구간을 산출하였다. 신뢰구간의 구체적 산출 절차는 다음과 같다. 먼저 각 모델의 예측에 대해 마스킹된 지점별 오차를 구하고, 이를 소속 원도별로 합과 개수로 집계한다. 다음으로 완결 원도 집합에서 원도 단위로 복원추출(with replacement)하여 원래 원도 수와 같은 크기의 재표집 표본을 만들고, 재표집된 원도들의 오차 합과 개수를 다시 합산해 지표를 재계산한다. 이때 유의파고와 유의파주기는 $\sqrt{(\Sigma 오차/\Sigma 개수)}$ 로 RMSE를 다시 계산하고, 파향은 $\Sigma 오차/\Sigma 개수$ 로 평균 각도오차를 다시 계산하여, 원도별 표본 수가 다른 경우에도 지표의 정의가 그대로 유지될 수 있게 하였다. 비교하는 두 모델에는 매 재표집마다 동일한 원도 인덱스를 적용하여 (paired) 원도 구성의 변동이 두 모델에 같은 방향으로 작용하도록 하였고, 이렇게 얻은 재표집별 지표 차이 10,000개의 2.5백분위수와 97.5백분위수를 취해 95% 신뢰구간으로 삼았다(percentile 방식). 원도를 재표집 단위로 둔 것은 같은 원도 안의 시점들이 서로 상관되어 있어 지점 단위로 재표집하면 신뢰구간이 실제보다 좁아지기 때문이다. 신뢰구간이 0을 포함하는 경우에는 ‘통계적 동등 (statistical tie)’로 판정하였고, 신뢰구간이 0을 포함하지 않는 경우에는 통계적으로 유의하다고 판단하여 ‘유의한 차이 (significant difference)’로 판정하였다. 이하 모든 표와 본문의 판정표기는 ‘통계적 동등’, ‘유의한 차이’ 두 가지로 통일하여 유의파고와 유의파주기의 동등 해석, 그리고 파향의 우월성 해석에 동일하게 적용하였다.

딥러닝 모델의 성능을 실무적 기준과 비교하기 위해 시간축 선형보간(linear interpolation)을 기준선으로 설정하였다. 각 원도 내 **target** 채널의 관측값만을 이용하여 결측 시점을 보간하였으며, 관측 구간 밖에 위치한 결측값은 최근접 관측값으로 대체하였다. 따라서 마지막 시점 결측과 같이 후속 관측값이 없는 경우에는 선형보간에 필요한 오른쪽 기준점이 존재하지 않으므로, 직전 관측값을 그대로 유지하는 방식으로 처리된다. 파향은 $u \cdot v$ 성분을 각각 선형보간한 뒤 각도로 복원하였다. 이 기준선은 별도의 학습 과정이 없는 방법으로, 운영 현장에서 일반적으로 사용되는 기본 보간 성능을 나타낸다. 추가적으로 결측률과 결측 패턴에 따른 모델 성능을 평가하였다. 결측 패턴에 따른 실험에서는 학습이 완료된 모델

의 가중치를 고정하고 재학습 없이 평가만 수행하였다. 성능 지표는 마스킹된 위치에서만 산출하였다. 결측률은 기준 평가와 동일한 마지막 1스텝 조건을 참조 조건으로 두고, 추가로 10%, 20%, 30%, 50%를 설정하였다. 원도 길이가 48스텝이므로 각 결측률에 대응하는 결측 스텝 수 $L = \text{round}(48 \cdot r)$ 는 각각 5, 10, 14, 24이다. 따라서 50% 연속 결측 조건은 24스텝, 즉 12시간의 연속 결측을 의미한다. 결측 패턴은 무작위형(random)과 연속형(contiguous)의 두 가지로 구성하였다. 무작위형 결측은 48개 원도 내에서 L 개 시점을 무작위로 결측 처리하는 방식이며, 연속형 결측은 길이 L 의 연속 구간 1개를 원도 내부에 완전히 포함되도록 무작위 위치에 배치하는 방식이다. 파향의 경우 $u \cdot v$ 두 성분을 동일한 시점에서 함께 결측 처리하였다.

실험 환경은 모델 구현에 따라 구분하였다. SAITS, GRIN, ImputeFormer는 Python 3.8, PyTorch 1.13.0+cu117 기반의 conda 환경에서 실행하였고, SPIN은 PyTorch 2.4.1+cu121 및 tsl 0.1.1 기반의 conda 환경에서 실행하였다. 실험 절차는 자료 전처리, 시험 세트 생성, 모델 학습 및 예측, 성능 지표 산출, 결측률 민감도 평가, paired bootstrap 신뢰구간 산출, 선형보간 기준선 산출의 순서로 수행하였다.

4. 결과 및 고찰

4.1 변수별 모델 성능 비교

임량, 송정, 해운대 통합 기준 평가(1step) 결과를 Table 1에 정리하였다. 모델 순위는 유의파고와 유의파주기는 RMSE가 낮은 순으로, 파향은 각도오차가 작은 순으로 산정하였으며, 비교 기준으로 시간축 선형보간 결과도 함께 제시하였다. Table 1에서는 변수별 최적 모델과 통계적 동등 모델을 굵은 글씨로 표시하였다. 통계적 동등 및 유의 판정은 3.3절에서 선언한 원도 단위 paired bootstrap 95% CI(재표집 10,000회, seed 42)를 단일 기준으로 적용하며, 그 결과를 Table 2에 제시한다.

유의파고의 경우 SAITS와 GRIN의 RMSE 차이는 0.0068로 작았고, paired bootstrap 95% 신뢰구간이 0을 포함하여 통계적 동등으로 해석되었다(Table 2). 유의파주기에서도 SAITS와 ImputeFormer의 RMSE 차이가 0.0081로 작았으며, 신뢰구간이 0을 포함하여 통계적 동등으로 판정되었다. 반면 파향에서는 SPIN과 ImputeFormer가 SAITS보다 낮은 각도오차를 나타냈다. SPIN과 SAITS의 차이는 약 3.7°로, 95% 신뢰구간이 0을 포함하지 않아 통계적으로 유의한 차이를 보였다. 다만 신뢰구간의 상한이 -0.03°로 0에 근접하므로 이 유의성은 경계에 가까운 수준으로 해석하는 것이 타당하다. ImputeFormer는 SAITS보다 약 1.0° 낮았으나 신뢰구간이 0을 포함하여 통계적 동등으로 판정되었다. 따라서 파향에서는 SPIN의 우세가 통계적으로 확인되며, ImputeFormer와 SAITS의 차이는 확정적이라고 보기는 어렵다. 선형보간 기준선과 비교하면, 각 변

Table 1. Performance by variable and model (reference evaluation, pooled over three targets, with the linear-interpolation baseline). Samples are complete windows/pooled points: significant wave height 123/369, significant wave period 116/348, wave direction 106/318

Variable	Performance Metric	SAITS	GRIN	SPIN	ImputeFormer	Linear Interpolation
Significant Wave Height	RMSE (m)	0.088	0.095	0.098	0.103	0.103
	MAE (m)	0.061	0.063	0.066	0.068	0.069
	R ²	0.943	0.934	0.930	0.922	0.923
Significant Wave Periods	RMSE (s)	0.468	0.520	0.495	0.476	0.542
	MAE (s)	0.343	0.388	0.355	0.351	0.397
	R ²	0.892	0.866	0.879	0.888	0.855
Wave Direction	Angular Error (°)	36.78	39.25	33.04	35.80	40.05

Table 2. Paired bootstrap 95% confidence intervals of metric differences between model pairs (diff = former minus latter; negative favors the former)

Variable	Model Pair	Difference	95% CI	Verdict
Significant Wave Height	SAITS – GRIN	−0.0068	[−0.0172, +0.0039]	statistical tie
Significant Wave Periods	SAITS – ImputeFormer	−0.0081	[−0.0283, +0.0120]	statistical tie
Wave Direction	SPIN – SAITS	−3.74°	[−7.70, −0.03]	significant difference
Wave Direction	ImputeFormer – SAITS	−0.98°	[−3.53, +1.59]	statistical tie

수의 최상위 딥러닝 모델은 모두 선형보간보다 낮은 오차를 보여 딥러닝 기반 접근의 이점이 비교적 일관되게 확인되었다.

4.2 물리적 특성과 모델 성능의 대응

2.2절에서 정리한 변수의 물리적 특성과 Table 1·2의 성능 결과를 교차 비교하면, 변수의 성격에 따라 적합한 모델 구조가 달라지는 대응 관계가 확인된다(변수·결측 패턴을 종합한 모델 선택 기준은 4.4절의 Table 5에 제시한다). 스칼라 크기량인 유의파고와 유의파주기에서는 시간 self-attention 기반 모델인 SAITS가 가장 낮은 RMSE를 보였다. 이는 마지막 시점 결측 복원에서 파고와 주기와 같은 스칼라 변수는 명시적 그래프 구조 없이 시간·특징 어텐션만으로도 상당한 수

준의 복원이 가능함을 시사한다. 반면 순환 방향량인 파향에서는 SPIN이 가장 낮은 각도 오차를 보였다. 이러한 결과는 파향과 같은 순환 변수의 복원에서 명시적 공간 구조 없는 시간·특징 어텐션만으로는 충분하지 않을 수 있으며, 관측소 간 관계를 함께 반영하는 구조가 유리하게 작용할 수 있음을 보여준다.

추가로, SAITS의 파향 best epoch가 13로 조기 수렴 또는 과소적합 가능성도 일부 포함될 수 있어, 이를 확인하기 위해 동일 설정으로 SAITS 파향 모델을 재학습하며 학습·검정 손실을 기록하였다(Fig. 2). patience를 120으로 완화해 133(13+120) epoch까지 학습을 확장한 재실험에서도 검정 손실이 그 이상 개선되지 않았으며 이는 단순한 학습 부족(과

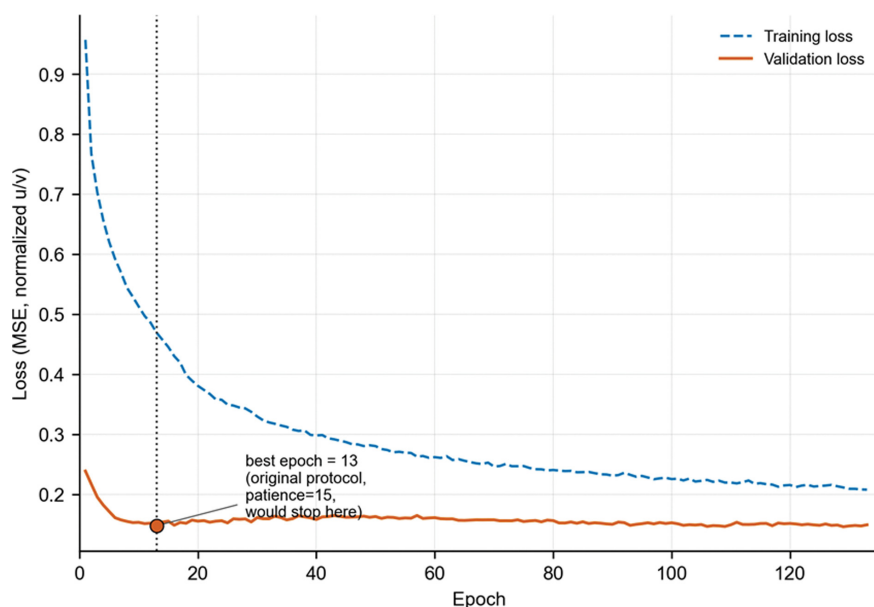


Fig. 2. SAITS wave-direction training and validation loss by epoch under the original protocol (seed 42, patience relaxed to 120 for this diagnostic re-run; the dotted line marks the epoch that would have triggered early stopping under the original patience = 15 setting).

소적합)보다는 이른 과적합에 가깝다고 볼 수 있다.

4.3 결측 패턴·결측률 민감도

동일 시험 세트에 결측 마스크만 바꿔 재평가한 결과, 결측 패턴(random 무작위 vs contiguous 연속)이 모델 우열을 뒤집는다. 전 결측률(1step, 10·20·30·50%)과 전체 변수의 결과를 Table 3에 전량 제시하고(선형보간 기준선 L 포함), 결측률에 따른 성능 변화를 Fig. 3에 도시한다(윗줄 (a)~(c) random, 아랫줄 (d)~(f) contiguous). 연속 50% 결측에서의 모델쌍 우열은 Table 4의 paired bootstrap 95% CI로 판정하였다.

Table 3의 최적 모델 판정은 3.3절과 동일한 window-level paired bootstrap 95% CI를 적용하여 확인하였고 변수별 최적 모델과 통계적 동물 모델을 굵은 글씨로 표시하였다. Table 3의 최적열에 병기한 동물 표기는 이 전체 비교 결과를 따른 것이며, 전 조건의 신뢰구간은 부록 A에 수록하였다. 무작위형(random) 결측에서는 4.1절의 기준 평가 결과가 대체로 유지되었다. 유의파고와 유의파주기에서는 SAITS가 대부분의 결측률 조건에서 최상위 또는 통계적 동물에 가까운 성능을 보였다. 일부 조건에서는 GRIN 또는 SPIN이 큰

소하게 낮은 RMSE를 보였으나, 그 차이는 기준 평가에서 동물이 판정된 차이보다 작아 실질적인 우열로 해석하기 어렵다. 파향에서는 모든 결측률 조건에서 SPIN이 가장 낮은 각도오차를 보였다. 또한 무작위형 결측에서는 결측률 증가에 따른 성능 저하가 전반적으로 완만하게 나타났다. 반면 연속형(contiguous) 결측에서는 유의파고와 유의파주기의 최적 모델이 GRIN으로 전환되었다. 유의파고와 유의파주기 모두 결측률 20% 이상에서 GRIN의 우위가 통계적으로 명확해졌다. 특히 50% 연속 결측 조건에서는 유의파고와 유의파주기 모두에서 GRIN이 SAITS보다 유의하게 낮은 RMSE를 보였으며, paired bootstrap 95% 신뢰구간도 0을 포함하지 않았다(Table 4). 이는 긴 연속 결측 구간에서 GRIN의 양방향 순환 그래프 전과 구조가 이웃 관측소 정보와 양측 시간 정보를 활용하는 데 유리하게 작용했을 가능성을 시사한다. 선형보간 기준선과 비교하면, 유의파고와 유의파주기 모두 30% 및 50% 연속 결측 조건에서 선형보간보다 뚜렷하게 낮은 오차를 보인 딥러닝 모델은 GRIN뿐이었다. 반면 SAITS, SPIN, ImputeFormer는 일부 연속 결측률이 높은 조건에서 선형보간과 유사하거나 오히려 높은 오차를 보였다. 이는 스칼라 변수의 긴 연속 결측 복원에서는 단순히 딥러닝 모델을 적용하

Table 3. Performance by missing pattern and missing rate (ALL). RMSE (lower is better) for significant wave height and significant wave period, angular error (deg) for wave direction. Model order: S (SAITS), G (GRIN), SP (SPIN), IF (ImputeFormer), L (linear interpolation). The 1step row uses the same mask as the reference evaluation

Variable	Missing Rate	Random		Ccontiguous	
		Best	S·G·SP·IF·L	Best	S·G·SP·IF·L
Significant Wave Height	1step	SAITS (≈GRIN)	0.088·0.095 ·0.099·0.103·0.103	←	same as left
	10%	SAITS (≈GRIN, SPIN)	0.080·0.082·0.081 ·0.086·0.082	GRIN (≈SAITS)	0.095·0.095 ·0.103·0.111·0.103
	20%	GRIN (≈4 Models)	0.086·0.086·0.088·0.089 ·0.090	GRIN	0.122· 0.108 ·0.132·0.188·0.122
	30%	SAITS (≈4 Models)	0.085·0.087·0.086·0.087 ·0.089	GRIN	0.156· 0.111 ·0.146·0.256·0.128
	50%	SAITS (≈GRIN)	0.091·0.092 ·0.094·0.099·0.093	GRIN	0.224· 0.133 ·0.180·0.321·0.159
Significant Wave Periods	1step	SAITS (≈IF)	0.468 ·0.520·0.495· 0.476 ·0.542	←	same as left
	10%	SPIN (≈4 Models)	0.472·0.469·0.459·0.462 ·0.493	SAITS (≈4 Models)	0.483·0.485·0.502·0.493 ·0.522
	20%	SAITS (≈SPIN, IF)	0.442 ·0.450· 0.445·0.443 ·0.486	GRIN	0.537· 0.508 ·0.573·0.711·0.560
	30%	ImputeFormer (≈SAITS, GRIN)	0.451·0.456 ·0.455· 0.449 ·0.489	GRIN	0.649· 0.537 ·0.607·0.926·0.598
	50%	SAITS (≈GRIN, SPIN)	0.470·0.472·0.473 ·0.482·0.501	GRIN	0.986· 0.567 ·0.755·1.143·0.727
Wave Direction	1step	SPIN (≈IF)	36.8·39.2· 33.0·35.8 ·40.0	←	same as left
	10%	SPIN	32.4·34.1· 30.9 ·32.7·36.8	SPIN	36.9·36.1· 33.9 ·36.3·40.8
	20%	SPIN	33.8·34.3· 32.4 ·33.6·37.7	ImputeFormer (≈4 Models)	36.2·36.2·35.9·35.5 ·41.0
	30%	SPIN	34.4·34.8· 32.4 ·33.8·36.9	SAITS (≈GRIN, IF)	34.7·35.2 ·36.9· 35.2 ·42.3
	50%	SPIN	34.8·35.0· 33.1 ·34.7·37.9	SAITS (≈GRIN)	35.5·35.6 ·39.4·37.3·42.9

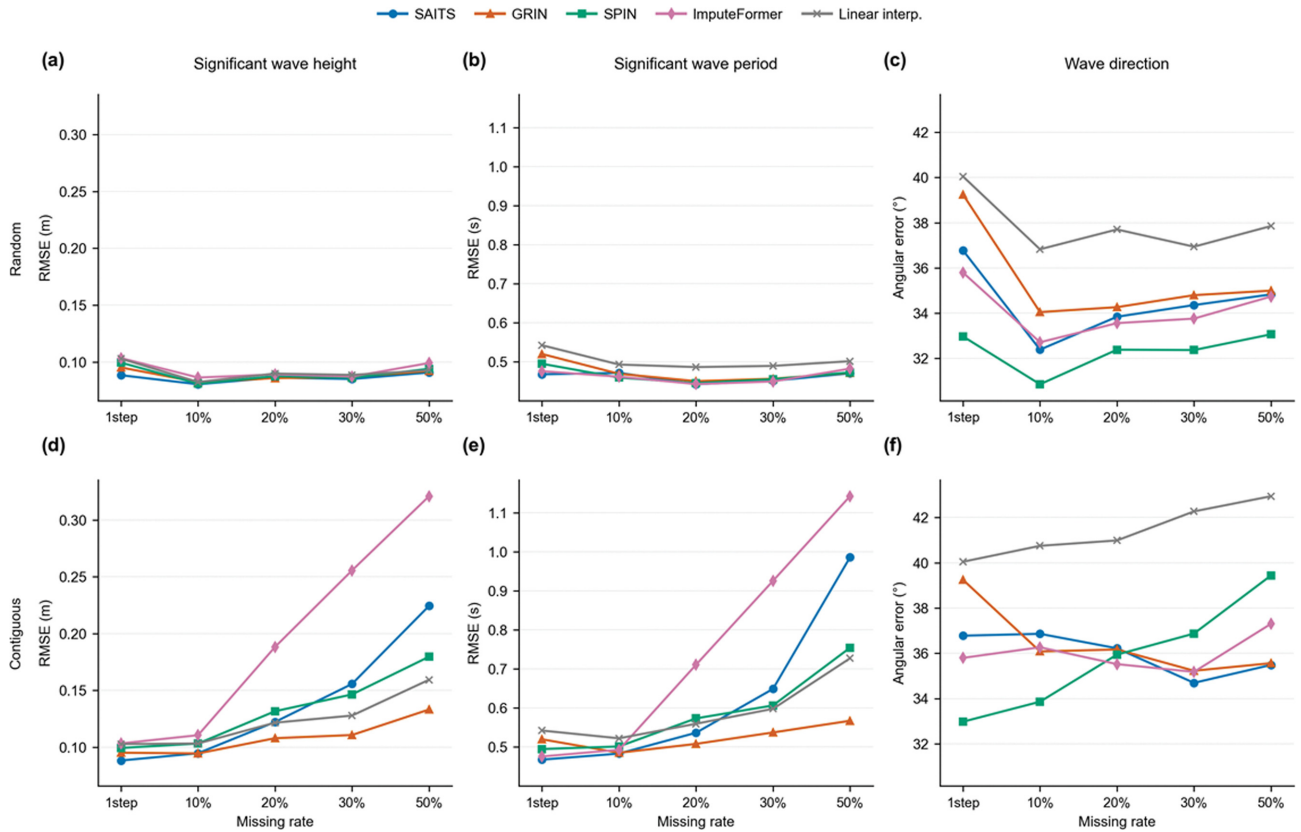


Fig. 3. Model performance versus missing rate with the linear-interpolation baseline (top row (a)~(c): random missing pattern; bottom row (d)~(f): contiguous missing pattern): (a, d) significant wave height, (b, e) significant wave period, and (c, f) wave direction.

Table 4. Paired bootstrap 95% confidence intervals of metric differences between model pairs under 50% contiguous missingness (diff = former minus latter; negative favors the former)

Missing	Variable	Model Pair	Difference	95% CI	Verdict
contiguous 50%	Significant Wave Height	GRIN – SAITS	-0.0911	[-0.1054, -0.0768]	significant difference
contiguous 50%	Significant Wave Periods	GRIN – SAITS	-0.4189	[-0.4487, -0.3886]	significant difference
contiguous 50%	Wave Direction	SAITS – SPIN	-3.96°	[-6.07, -1.95]	significant difference
contiguous 50%	Wave Direction	SAITS – GRIN	-0.0779	[-0.8195, 0.6804]	statistical tie

는 것만으로 충분하지 않으며, 결측 구간을 주변 관측소와 시간축 정보를 통해 안정적으로 전파할 수 있는 구조가 중요함을 의미한다.

파향은 스칼라 변수와 다른 양상을 보였다. 무작위형 결측에서는 SPIN이 모든 결측률 조건에서 가장 낮은 각도오차를 보였으나, 연속형 결측률이 높은 조건에서는 성능이 크게 저하되었다. 특히 30%와 50% 연속 결측 조건에서는 SPIN이 최하위로 떨어졌고, 50% 조건에서는 SAITS가 가장 우수한 성능을 보였다.

파향에서는 모든 결측 패턴과 결측률 조건에서 4개 딥러닝 모델이 모두 선형보간보다 낮은 각도오차를 보였다. 이는 파향과 같은 순환 변수에서는 단순 선형보간보다 딥러닝 기반 복원이 일관되게 유리할 수 있음을 보여준다. 반대로 유의파고의 무작위형 결측에서는 선형보간 기준선이 딥러닝 최상위 모델과 매우 근접한 성능을 보였다. 이는 양측 관측값이 남아 있는 내부 무작위 결측의 경우, 단순 보간만으로도 상당

한 수준의 복원이 가능함을 의미한다.

4.4 종합 고찰 및 한계

종합하면, 본 연구는 파랑 관측자료의 결측보간에서 대상 변수의 물리적 특성과 결측 패턴을 함께 고려한 모델 선택 기준을 제시하였다. 동일한 자료, 데이터 분할, 시퀀스 및 마스크 조건에서 비교한 결과, 스칼라 크기량인 유의파고와 유의파주기의 무작위 결측에는 시간축 attention 기반 SAITS가 전반적으로 우수한 복원 성능을 보였다. 순환 방향량인 파향의 무작위 결측에서는 SPIN이 가장 낮은 각도오차를 나타냈으며, 일부 조건에서는 ImputeFormer도 SPIN과 통계적 동률을 보였다. 반면 원도 내 최대 24스텝, 즉 12시간 이내의 연속 결측 구간에서는 결측률이 증가할수록 그래프 기반 정보 전파를 활용하는 GRIN의 상대적 강점이 뚜렷하게 나타났다. 이에 따라 무작위 결측 또는 짧은 연속 결손이 주로 발생하는 환경에서는 SAITS를 우선 검토하고, 파향과 같은 방향 변수

Table 5. Favorable models and supporting statistical evidence by variable physical characteristics and missing-data condition

Variable	Physical Properties	Missingness Condition	Appropriate Model	Justification
Significant Wave Height	Scalar quantity	Final time step·Random	SAITS (≈GRIN)	Under the baseline evaluation, the RMSE was 0.088, showing a statistical tie with GRIN (0.095) (Tables 1 and 2)
		Contiguous 20% or more	GRIN	Under 50% consecutive missingness, the RMSE was 0.133, showing a statistically significant difference from SAITS (0.224) (Tables 3 and 4)
Significant Wave Periods	Scalar quantity	Final time step·Random	SAITS (≈ImputeFormer)	Under the baseline evaluation, the RMSE was 0.468, showing a statistical tie with ImputeFormer (0.476) (Tables 1 and 2)
		Contiguous 20% or more	GRIN	Under 50% consecutive missingness, the RMSE was 0.567, showing a statistically significant difference from SAITS (0.986) (Tables 3 and 4)
Wave Direction	Directional quantity	Final time step·Random	SPIN (≈ImputeFormer)	The angular error was 33.0°, showing a statistically significant difference from SAITS (36.8°) and a statistical tie with ImputeFormer (35.8°) (Tables 1 and 2)
		Contiguous 20% or more	SAITS (≈GRIN)	Under 50% consecutive missingness, the angular error was 35.5°, showing a statistical tie with GRIN (35.6°) (Table 3)

에는 SPIN을 적용하는 기준을 제안한다. 약 5시간에 해당하는 20% 이상의 연속 결손이 빈번하고 최대 12시간 이내의 결손이 예상되는 환경에서는 GRIN을 우선적으로 검토할 수 있다. 특히 50% 연속 결측 조건에서 유의파고와 유의파주기 모두 선형보간보다 낮은 오차를 보인 딥러닝 모델은 GRIN이 유일하였다(4.3절). 이는 장시간의 연속 결손 상황에서 주변 관측소의 정보를 활용하는 시공간 접근이 실질적인 복원 대안이 될 수 있음을 보여준다. 변수 및 결측 패턴별 적합 모델과 통계적 근거는 Table 5에 종합하였다.

다만 본 연구는 조기 종료와 최종 모형 선택은 관측값의 10%를 무작위로 결측 처리한 검정 자료의 복원오차를 기준으로 수행한 반면, 핵심 평가는 복원 대상 관측소의 마지막 시점 결측 및 윈도 내 연속 결측 조건에서 수행하였다. 이러한 검정-평가 마스크 간 불일치는 연속 결측에 특화된 성능이 모형 선택 단계에 충분히 반영되지 못할 수 있다는 구조적 한계가 된다. 실제로 무작위 결측 조건에서 GRIN은 유의파고와 유의파주기에서 다른 모델보다 일관되게 우세하지 않았으나, 연속 결측률이 20% 이상인 조건에서는 두 변수 모두에서 가장 낮은 오차를 나타냈다. 따라서 GRIN의 최종 모형은 연속 결측 성능이 아니라 무작위 10% 결측 복원오차를 기준으로 선택되었으며, 이 선택 기준이 GRIN의 연속 결측 성능을 충분히 반영하지 못했을 가능성이 있다. 본 연구 결과만으로 해당 영향의 크기나 방향을 정량적으로 단정할 수는 없지만, 보고된 GRIN 성능은 연속 결측에 최적화된 검정 기준이 아닌 무작위 결측 기준으로 선택된 단일 모형의 성능으로 해석할 필요가 있다. 그럼에도 이러한 조건에서 GRIN이 연속 결측에 대해 가장 우수한 성능을 보였다는 결과는, 연속적인 누락 구간 복원에서 시공간 정보 활용의 가능성을 뒷받침한다. 향후에는 무작위 결측, 단일 시점 결측 및 연속 결측 등 예상되는 운영 환경의 결측 패턴별로 검정 자료를 구성하고, 각 조건에서 선택된 모형을 동일한 평가 조건에서 비

교함으로써 조기 종료 기준이 최종 성능과 모델 순위에 미치는 영향을 정량적으로 검토할 필요가 있다.

5. 결 론

본 연구는 파랑 관측자료 결측보간에서 대상 변수의 물리적 특성과 결측 패턴에 따라 적합한 모델이 어떻게 달라지는지를 동일한 비교 조건에서 검정하고, 실무적 모델 선택 기준을 제시하고자 하였다. 이를 위해 부산·울산 연안 5개 관측소 자료를 이용하여 3개 관측소를 복원 대상으로 설정하고, SAITS, GRIN, SPIN, ImputeFormer 및 시간축 선형보간을 동일한 자료 분할, 입력 시퀀스 및 결측 마스크 조건에서 비교하였다. 유의파고, 유의파주기 및 파향에 대하여 무작위형과 연속형 결측 조건을 구분하여 성능 변화를 평가하였으며, 모델 간 차이와 통계적 동등성은 윈도 단위 paired bootstrap 95% 신뢰구간을 기준으로 판정하였다.

분석 결과, 파랑 변수의 특성과 결측 양상에 따라 모델별 강점이 구분되었다. 스칼라 크기량인 유의파고와 유의파주기에서는 시간 self-attention 기반 SAITS가 가장 낮은 오차를 보였고, 유의파고의 GRIN 및 유의파주기의 ImputeFormer는 SAITS와 통계적 동등로 확인되었다. 순환 방향량인 파향에서는 SPIN과 ImputeFormer가 낮은 각도오차를 보였으며, SPIN은 SAITS보다 통계적으로 유의하게 우수하였다. 다만 파향에서의 성능 차이가 특정 구조적 요소에 의해 발생한 것인지는 본 연구의 비교만으로 분리하여 확인하지 않았으므로, 이에 대한 해석은 제한적으로 이루어져야 한다.

결측 패턴에 따른 차이는 더욱 뚜렷하였다. 무작위 결측 조건에서는 기준 평가에서 확인된 모델 순위가 대체로 유지되었으나, 윈도 내 연속 결측에서는 결측률이 증가함에 따라 유의파고와 유의파주기의 우수 모델이 GRIN으로 전환되었다. 특히 50% 연속 결측 조건에서 GRIN은 유의파고와 유의파

주기 모두에서 선형보간 기준선을 상회한 유일한 딥러닝 모델이었다. 이는 장시간 연속 결측 상황에서 인접 관측소의 정보를 활용하는 그래프 전파 방식이 시간축 정보만으로는 복원하기 어려운 구간에 유효하게 작동할 수 있음을 보여준다. 반면 파향은 고결측 연속 조건에서 SAITS가 우수한 성능을 보여, 방향 변수의 결측보간에는 연속 결측의 길이와 자료 구조를 함께 고려할 필요가 있음을 확인하였다.

이상의 결과를 바탕으로, 본 연구는 파랑 관측자료의 결측 보간에서 변수 특성과 결측 패턴을 함께 고려하는 모델 선택 기준을 제안한다. 무작위형 결측 또는 짧은 연속 결측이 주로 발생하는 환경에서는 SAITS를 우선 검토할 수 있으며, 파향과 같은 방향 변수에는 SPIN을 적용하는 방안을 고려할 수 있다. 반면 약 5시간(결측률 20%) 이상의 연속 결측이 빈번하고 최대 12시간 이내의 결측이 예상되는 경우에는 GRIN을 우선 검토하는 것이 적절하다. 즉, 관측 변수와 예상 결측 양상에 따라 모델을 선택적으로 운영하는 것이 단일 모델을 일률적으로 적용하는 방식보다 합리적인 접근이 될 수 있다. 향후에는 실제 운영 환경의 결측 패턴을 반영한 검정 기준을 구성하고, 관측 기간·관측망 확대, 외해 보조노드 누락 민감도, 수일 규모의 장기 결측, 실시간 자기회귀 다중 스텝 복원 조건까지 평가 범위를 확장할 필요가 있다. 이러한 후속 검정은 본 연구에서 제시한 변수별·결측 패턴별 모델 선택 기준을 보다 다양한 파랑 관측 환경으로 확장하는 기반이 될 것이다.

감사의 글

본 연구는 국립해양조사원의 「2026년 실시간 이안류 감시체계 운영」 사업의 지원을 받아 수행되었습니다(계약번호: R26TA0170412700).

References

- Cao, W., Wang, D., Li, J., Zhou, H., Li, L. and Li, Y. (2018). BRITS: Bidirectional recurrent imputation for time series. In *Advances in Neural Information Processing Systems* 31 (NeurIPS 2018), 6776-6786.
- Choi, J. (2023). Application of rip current likelihood distributions on rip current forecast system. *Journal of Korea Water Resources Association*, 56(8), 521-528 (in Korean).
- Cini, A., Marisca, I. and Alippi, C. (2022). Filling the gaps: Multivariate time series imputation by graph neural networks. In *International Conference on Learning Representations (ICLR 2022)*.
- Du, W., Côté, D. and Liu, Y. (2023). SAITS: Self-attention-based imputation for time series. *Expert Systems with Applications*, 219, 119619.
- Kim, I.C., Lee, J.Y. and Lee, J.L. (2011). Generation mechanism and numerical simulation of rip current at Haeundae Beach. *Journal of Korean Society of Coastal and Ocean Engineers*, 23(1), 70-78 (in Korean).
- Kim, T. (2020). A study on the prediction technique for wind and wave using deep learning. *Journal of the Korean Society for Marine Environment & Energy*, 23(3), 142-147 (in Korean).
- Kim, T. and Lee, W.D. (2022). Review on applications of machine learning in coastal and ocean engineering. *Journal of Ocean Engineering and Technology*, 36(3), 194-210 (in Korean).
- Lee, K.H., Kim, T.G. and Kim, D.S. (2020). Prediction of wave breaking using machine learning open source platform. *Journal of Korean Society of Coastal and Ocean Engineers*, 32(4), 262-272 (in Korean).
- Marisca, I., Cini, A. and Alippi, C. (2022). Learning to reconstruct missing data from spatiotemporal graphs with sparse observations. In *Advances in Neural Information Processing Systems* 35 (NeurIPS 2022), 32069-32082.
- Nie, T., Qin, G., Ma, W., Mei, Y. and Sun, J. (2024). ImputeFormer: Low rankness-induced transformers for generalizable spatiotemporal imputation. In *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (KDD 2024)*.
- Oh, C., Ahn, K. and Cheon, S.H. (2019). Automatic detection and analysis of rip currents at Haeundae Beach using X-band marine radar. *Journal of Korean Society of Coastal and Ocean Engineers*, 31(6), 485-492 (in Korean).
- Park, H.J. and Kang, Y.J. (2024). Comparison of wave prediction and performance evaluation in Korea waters based on machine learning. *Journal of Ocean Engineering and Technology*, 38(1), 18-29 (in Korean).
- Shukla, S.N. and Marlin, B.M. (2021). Multi-time attention networks for irregularly sampled time series. In *International Conference on Learning Representations (ICLR 2021)*.
- Tashiro, Y., Song, J., Song, Y. and Ermon, S. (2021). CSDI: Conditional score-based diffusion models for probabilistic time series imputation. In *Advances in Neural Information Processing Systems* 34 (NeurIPS 2021).
- Yoon, J., Jordon, J. and van der Schaar, M. (2018). GAIN: Missing data imputation using generative adversarial nets. In *Proceedings of the 35th International Conference on Machine Learning (ICML 2018)*, PMLR 80, 5689-5698.

Received 10 July, 2026

Revised 31 August, 2026

Accepted 31 August, 2026

부록 A

본 부록에서는 변수별·결측 패턴별 세부 성능값과 모델 간 paired bootstrap 95% 신뢰구간 결과를 제시한다.

Table A1. Paired bootstrap 95% confidence intervals for model differences in significant wave height across missingness patterns

	Missing Rate	Model A	Model B	Random	Contiguous
Significant Wave Height	1step	SAITS	GRIN	statistical tie	statistical tie
	1step	SAITS	SPIN	model A better	model A better
	1step	SAITS	Imputeformer	model A better	model A better
	1step	GRIN	SPIN	statistical tie	statistical tie
	1step	GRIN	Imputeformer	statistical tie	statistical tie
	1step	SPIN	Imputeformer	statistical tie	statistical tie
	10	SAITS	GRIN	statistical tie	statistical tie
	10	SAITS	SPIN	statistical tie	model A better
	10	SAITS	Imputeformer	model A better	model A better
	10	GRIN	SPIN	statistical tie	model A better
	10	GRIN	Imputeformer	statistical tie	model A better
	10	SPIN	Imputeformer	model A better	model A better
	20	SAITS	GRIN	statistical tie	model B better
	20	SAITS	SPIN	statistical tie	statistical tie
	20	SAITS	Imputeformer	model A better	model A better
	20	GRIN	SPIN	statistical tie	model A better
	20	GRIN	Imputeformer	statistical tie	model A better
	20	SPIN	Imputeformer	statistical tie	model A better
	30	SAITS	GRIN	statistical tie	model B better
	30	SAITS	SPIN	statistical tie	statistical tie
	30	SAITS	Imputeformer	statistical tie	model A better
	30	GRIN	SPIN	statistical tie	model A better
	30	GRIN	Imputeformer	statistical tie	model A better
	30	SPIN	Imputeformer	statistical tie	model A better
	50	SAITS	GRIN	statistical tie	model B better
	50	SAITS	SPIN	model A better	model B better
	50	SAITS	Imputeformer	model A better	model A better
	50	GRIN	SPIN	statistical tie	model A better
	50	GRIN	Imputeformer	model A better	model A better
	50	SPIN	Imputeformer	model A better	model A better

Table A2. Paired bootstrap 95% confidence intervals for model differences in significant wave period across missingness patterns

	Missing Rate	Model A	Model B	Random	Contiguous
Significant Wave Periods	1step	SAITS	GRIN	model A better	model A better
	1step	SAITS	SPIN	model A better	model A better
	1step	SAITS	Imputeformer	statistical tie	statistical tie
	1step	GRIN	SPIN	statistical tie	statistical tie
	1step	GRIN	Imputeformer	model B better	model B better
	1step	SPIN	Imputeformer	statistical tie	statistical tie
	10	SAITS	GRIN	statistical tie	statistical tie
	10	SAITS	SPIN	statistical tie	statistical tie
	10	SAITS	Imputeformer	statistical tie	statistical tie
	10	GRIN	SPIN	statistical tie	statistical tie
	10	GRIN	Imputeformer	statistical tie	statistical tie
	10	SPIN	Imputeformer	statistical tie	statistical tie
	20	SAITS	GRIN	model A better	model B better
	20	SAITS	SPIN	statistical tie	model A better
	20	SAITS	Imputeformer	statistical tie	model A better
	20	GRIN	SPIN	statistical tie	model A better
	20	GRIN	Imputeformer	statistical tie	model A better
	20	SPIN	Imputeformer	statistical tie	model A better
	30	SAITS	GRIN	statistical tie	model B better
	30	SAITS	SPIN	statistical tie	model B better
	30	SAITS	Imputeformer	statistical tie	model A better
	30	GRIN	SPIN	statistical tie	model A better
	30	GRIN	Imputeformer	statistical tie	model A better
	30	SPIN	Imputeformer	model B better	model A better
	50	SAITS	GRIN	statistical tie	model B better
	50	SAITS	SPIN	statistical tie	model B better
	50	SAITS	Imputeformer	model A better	model A better
	50	GRIN	SPIN	statistical tie	model A better
	50	GRIN	Imputeformer	model A better	model A better
	50	SPIN	Imputeformer	model A better	model A better

Table A3. Paired bootstrap 95% confidence intervals for model differences in wave direction across missingness patterns

	Missing Rate	Model A	Model B	Random	Contiguous
Wave Direction	1step	SAITS	GRIN	statistical tie	statistical tie
	1step	SAITS	SPIN	model B better	model B better
	1step	SAITS	Imputeformer	statistical tie	statistical tie
	1step	GRIN	SPIN	model B better	model B better
	1step	GRIN	Imputeformer	statistical tie	statistical tie
	1step	SPIN	Imputeformer	statistical tie	statistical tie
	10	SAITS	GRIN	model A better	statistical tie
	10	SAITS	SPIN	model B better	model B better
	10	SAITS	Imputeformer	statistical tie	statistical tie
	10	GRIN	SPIN	model B better	model B better
	10	GRIN	Imputeformer	statistical tie	statistical tie
	10	SPIN	Imputeformer	model A better	model A better
	20	SAITS	GRIN	statistical tie	statistical tie
	20	SAITS	SPIN	model B better	statistical tie
	20	SAITS	Imputeformer	statistical tie	statistical tie
	20	GRIN	SPIN	model B better	statistical tie
	20	GRIN	Imputeformer	statistical tie	statistical tie
	20	SPIN	Imputeformer	model A better	statistical tie
	30	SAITS	GRIN	statistical tie	statistical tie
	30	SAITS	SPIN	model B better	model A better
	30	SAITS	Imputeformer	model B better	statistical tie
	30	GRIN	SPIN	model B better	statistical tie
	30	GRIN	Imputeformer	model B better	statistical tie
	30	SPIN	Imputeformer	model A better	model B better
	50	SAITS	GRIN	statistical tie	statistical tie
	50	SAITS	SPIN	model B better	model A better
	50	SAITS	Imputeformer	statistical tie	model A better
	50	GRIN	SPIN	model B better	model A better
	50	GRIN	Imputeformer	statistical tie	model A better
	50	SPIN	Imputeformer	model A better	model B better

부록 B

본 부록에서는 무작위 및 연속 결측 조건에서의 보간 결과를 시각적으로 비교하기 위해, 송정·임랑·해운대 관측소의 복원 시계열을 제시한다. 전체 시계열의 길이를 고려하여 가독성을 확보할 수 있는 일부 구간을 임의로 선정하였다.

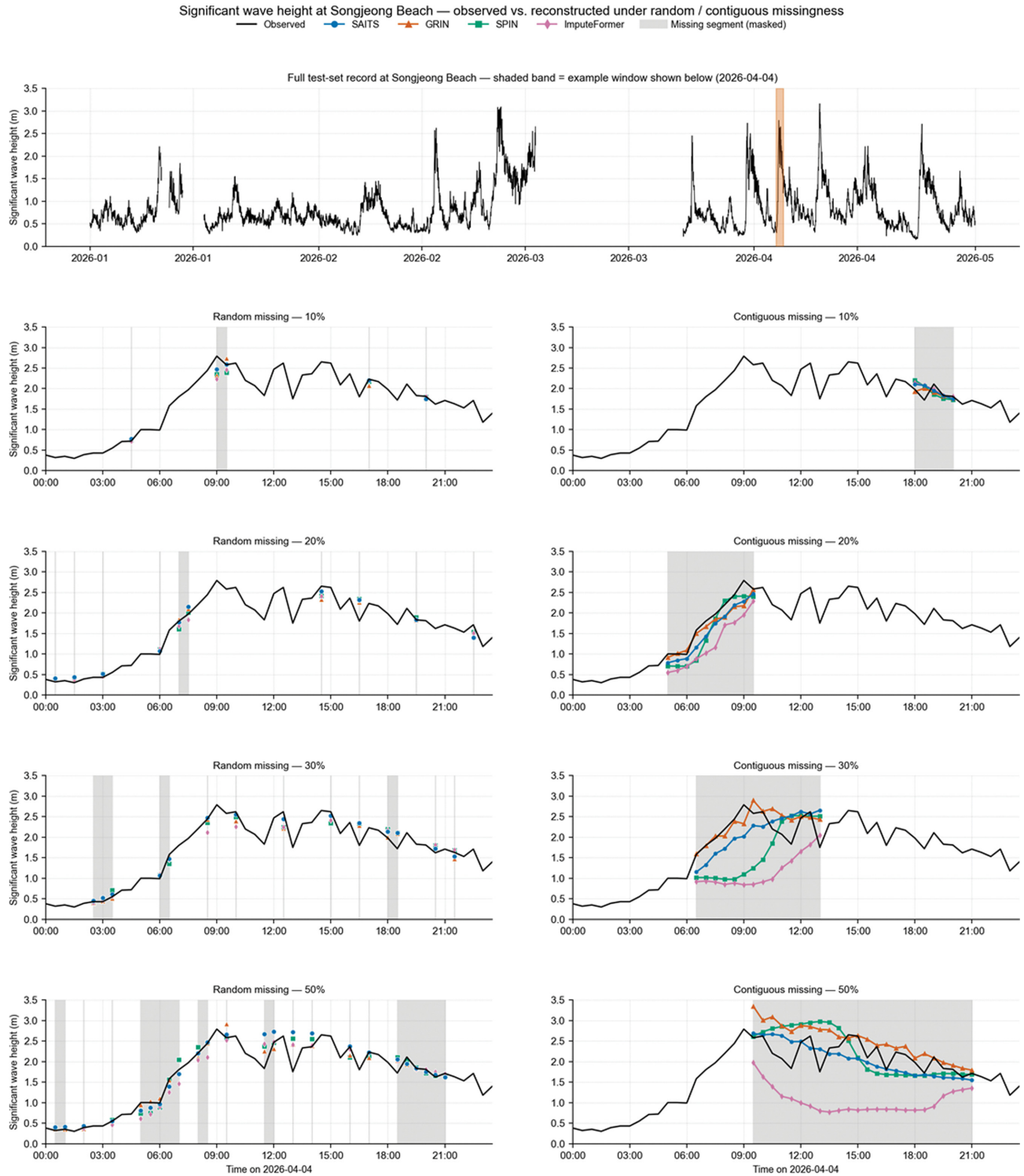


Fig. B1. Comparison of observed and reconstructed significant wave height at Songjeong Beach under random and contiguous missingness conditions.

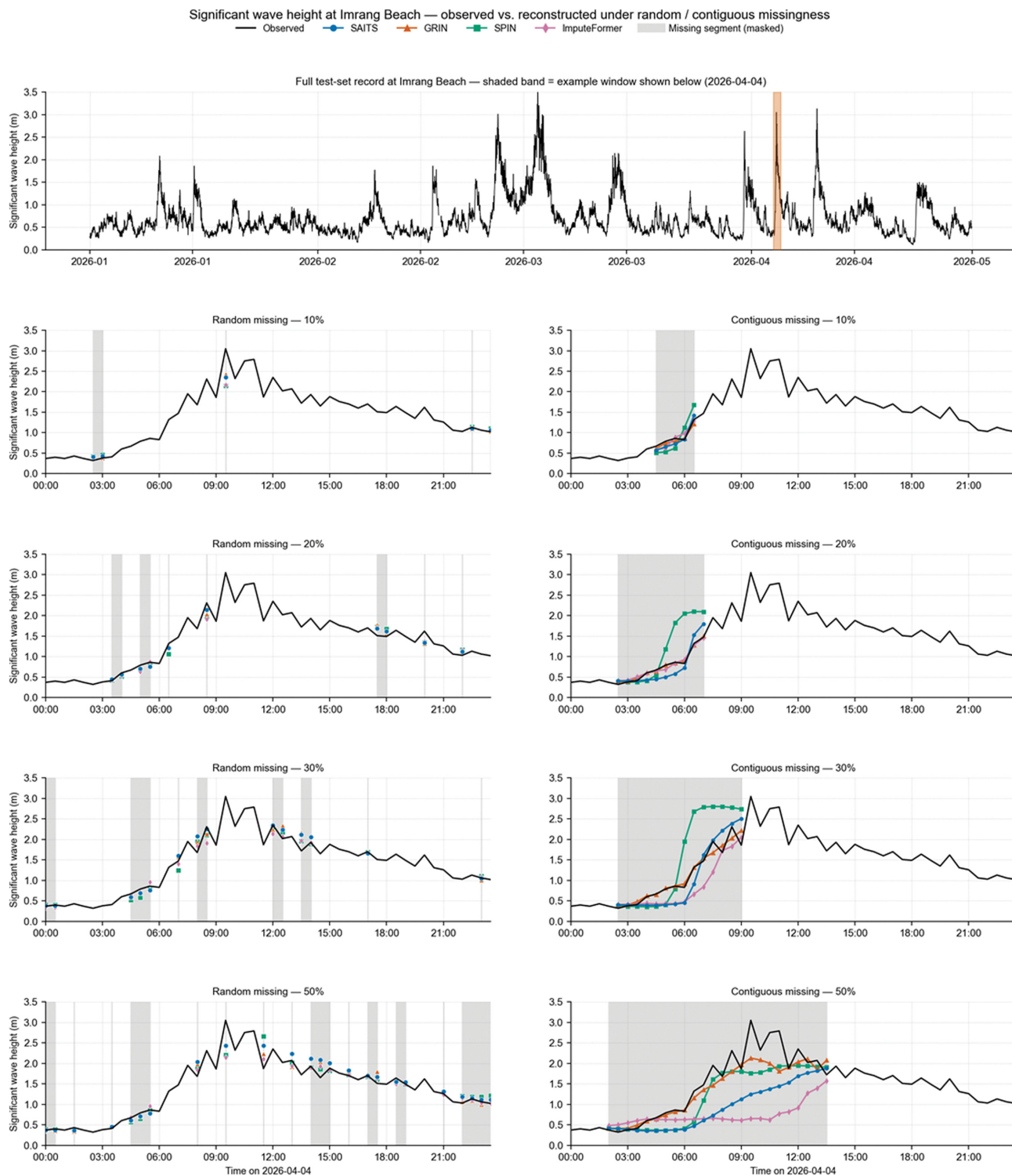


Fig. B2. Comparison of observed and reconstructed significant wave height at Imrang Beach under random and contiguous missingness conditions.

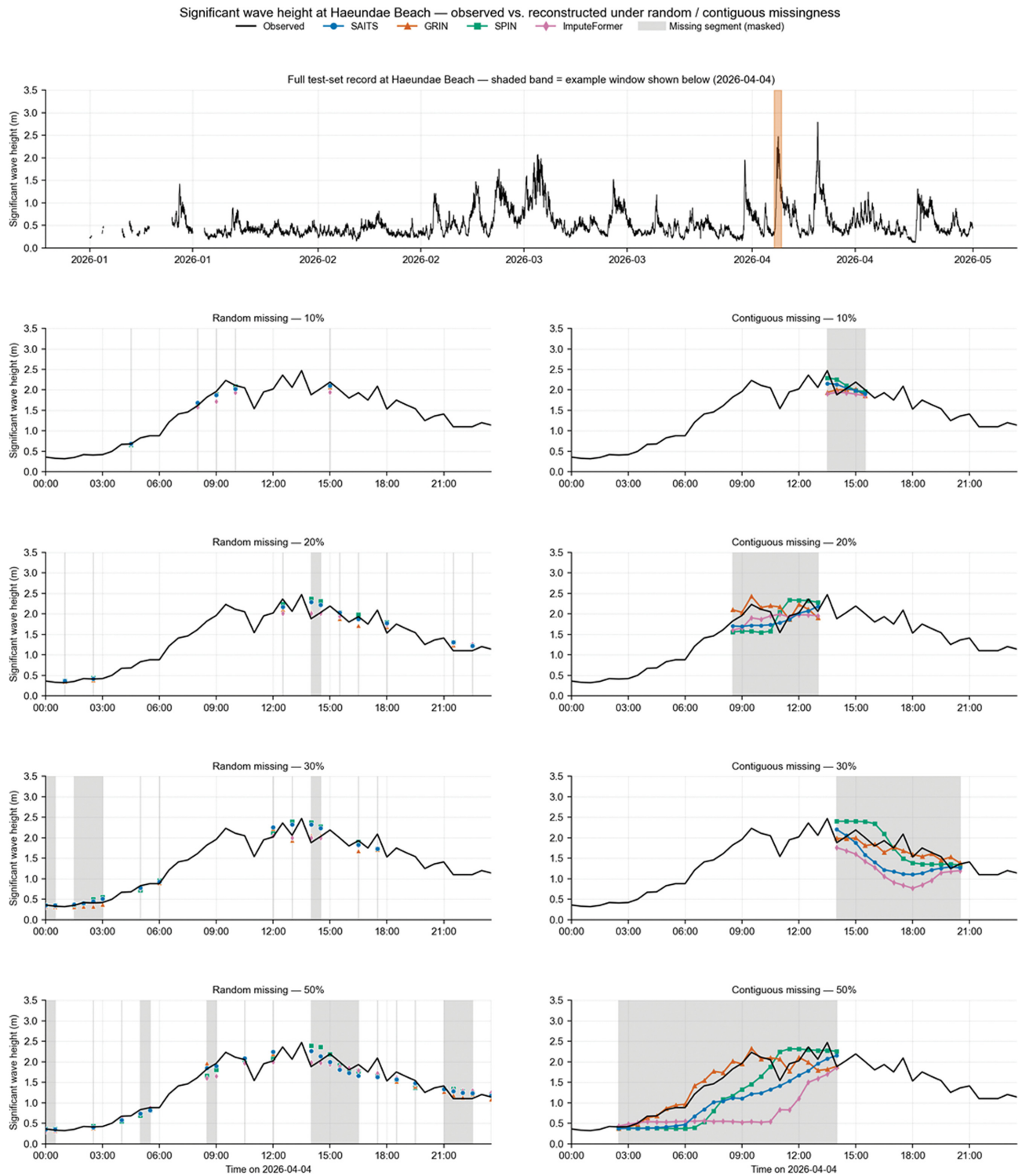


Fig. B3. Comparison of observed and reconstructed significant wave height at Haeundae Beach under random and contiguous missingness conditions.

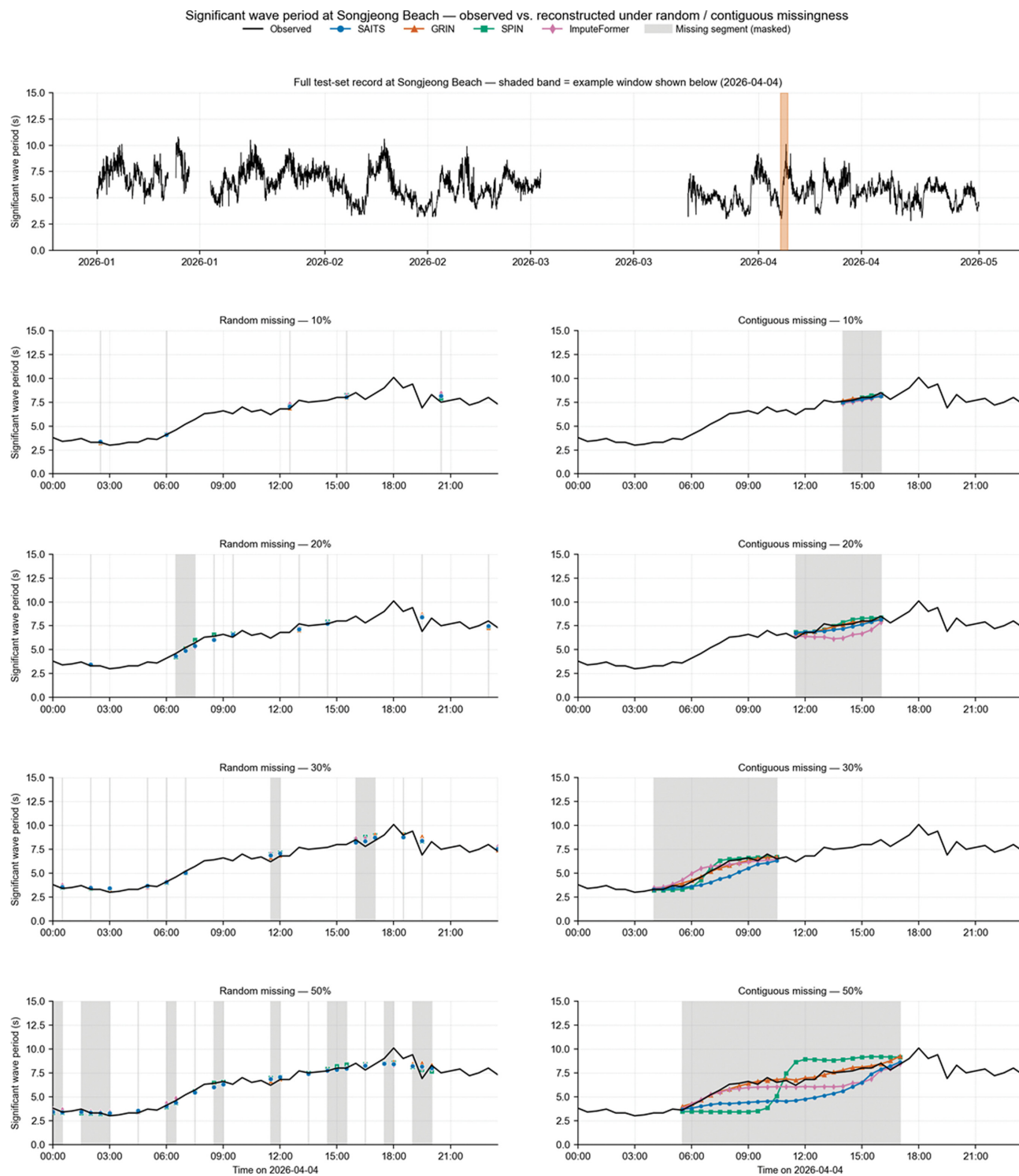


Fig. B4. Comparison of observed and reconstructed significant wave periods at Songjeong Beach under random and contiguous missingness conditions.

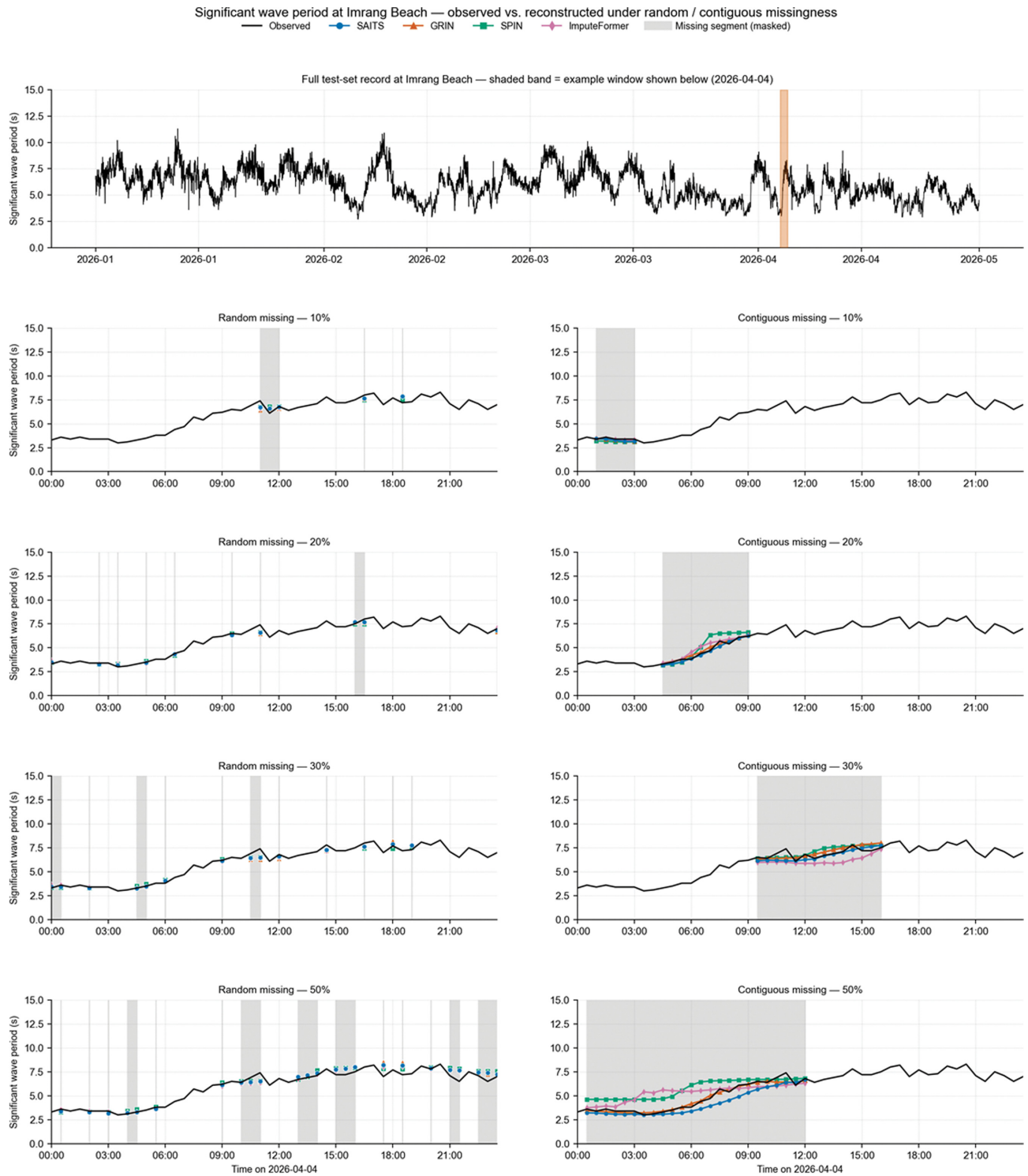


Fig. B5. Comparison of observed and reconstructed significant wave periods at Imrang Beach under random and contiguous missingness conditions.

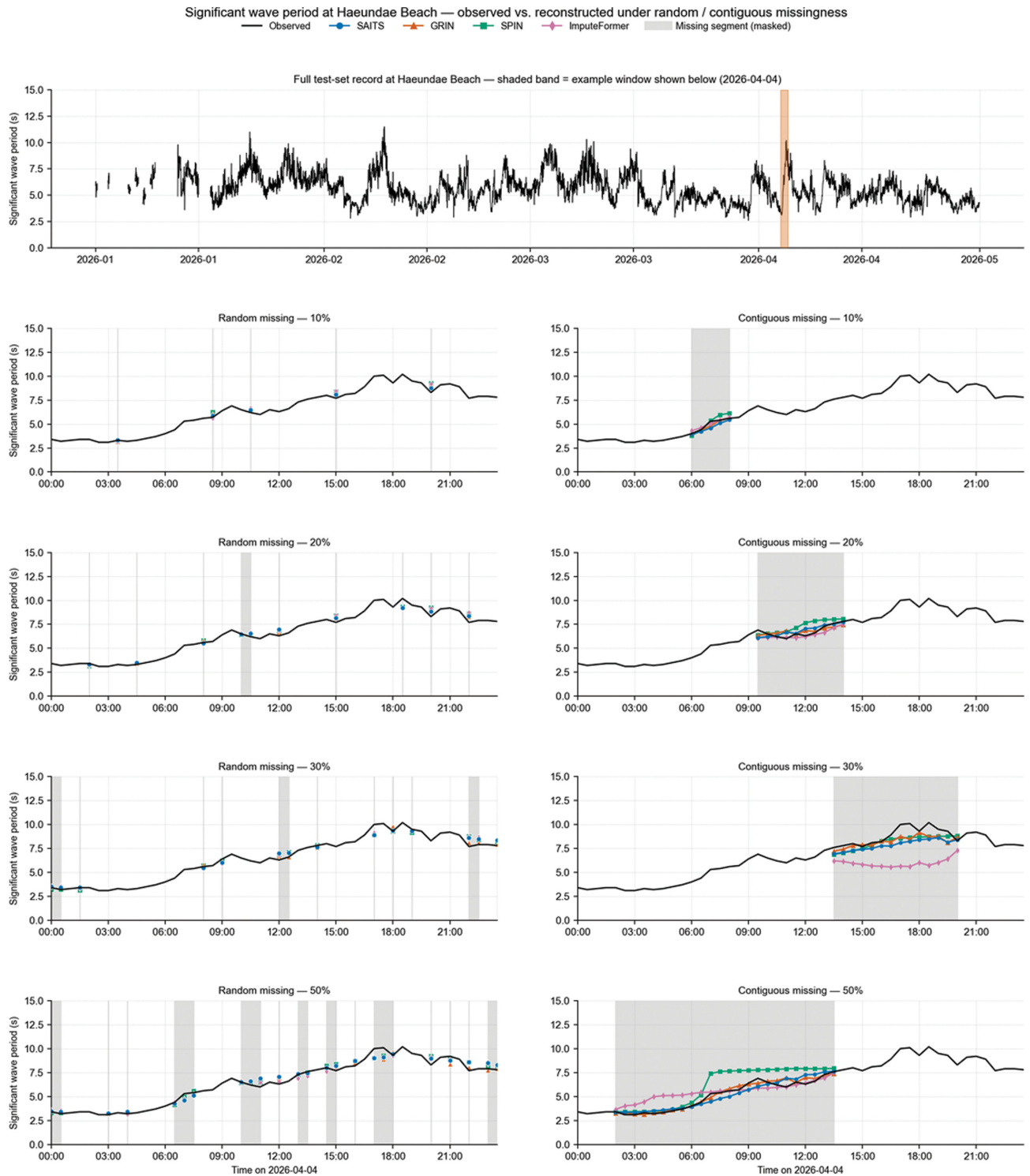


Fig. B6. Comparison of observed and reconstructed significant wave periods at Haeundae Beach under random and contiguous missingness conditions.

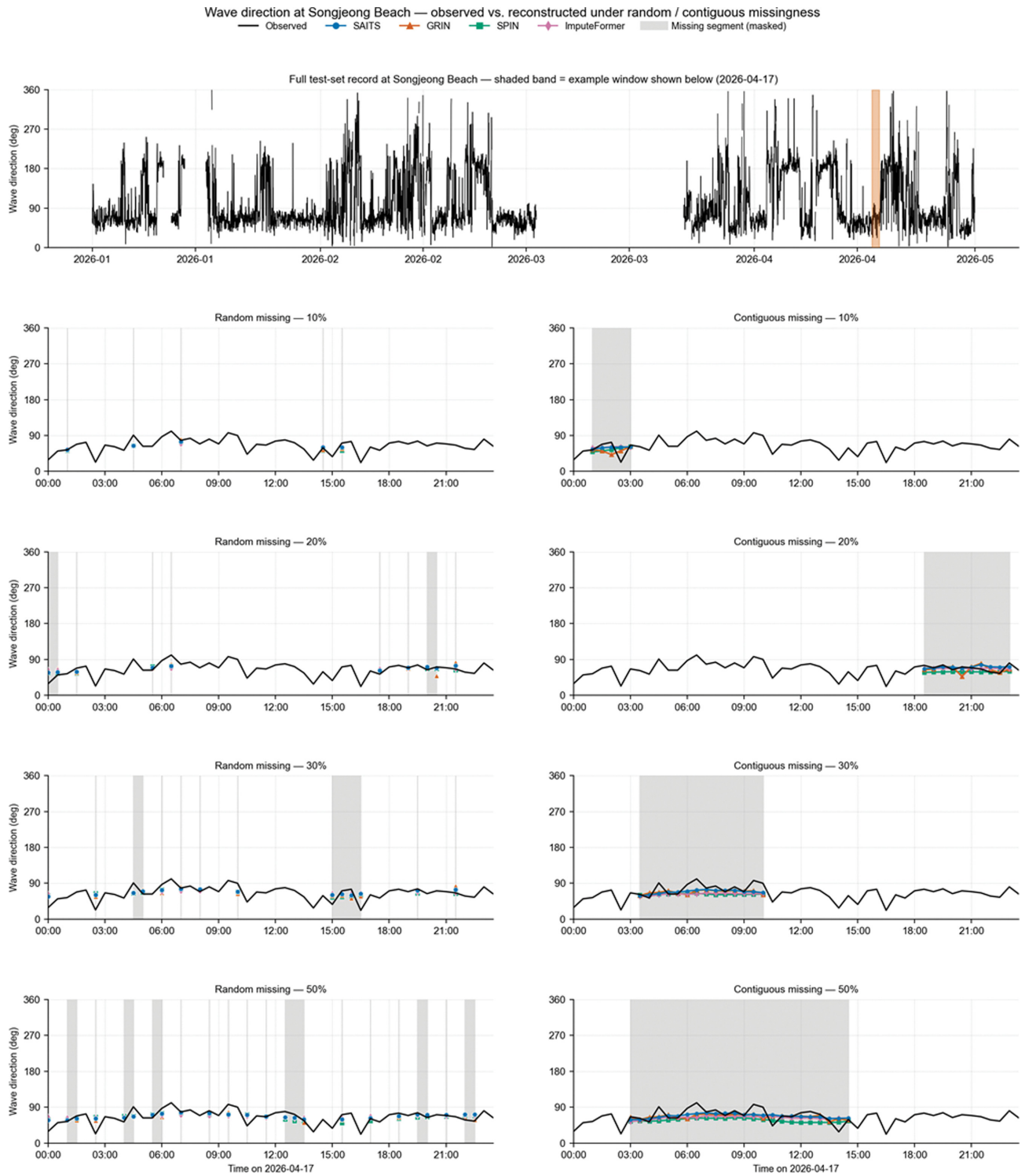


Fig. B7. Comparison of observed and reconstructed wave direction at Songjeong Beach under random and contiguous missingness conditions.

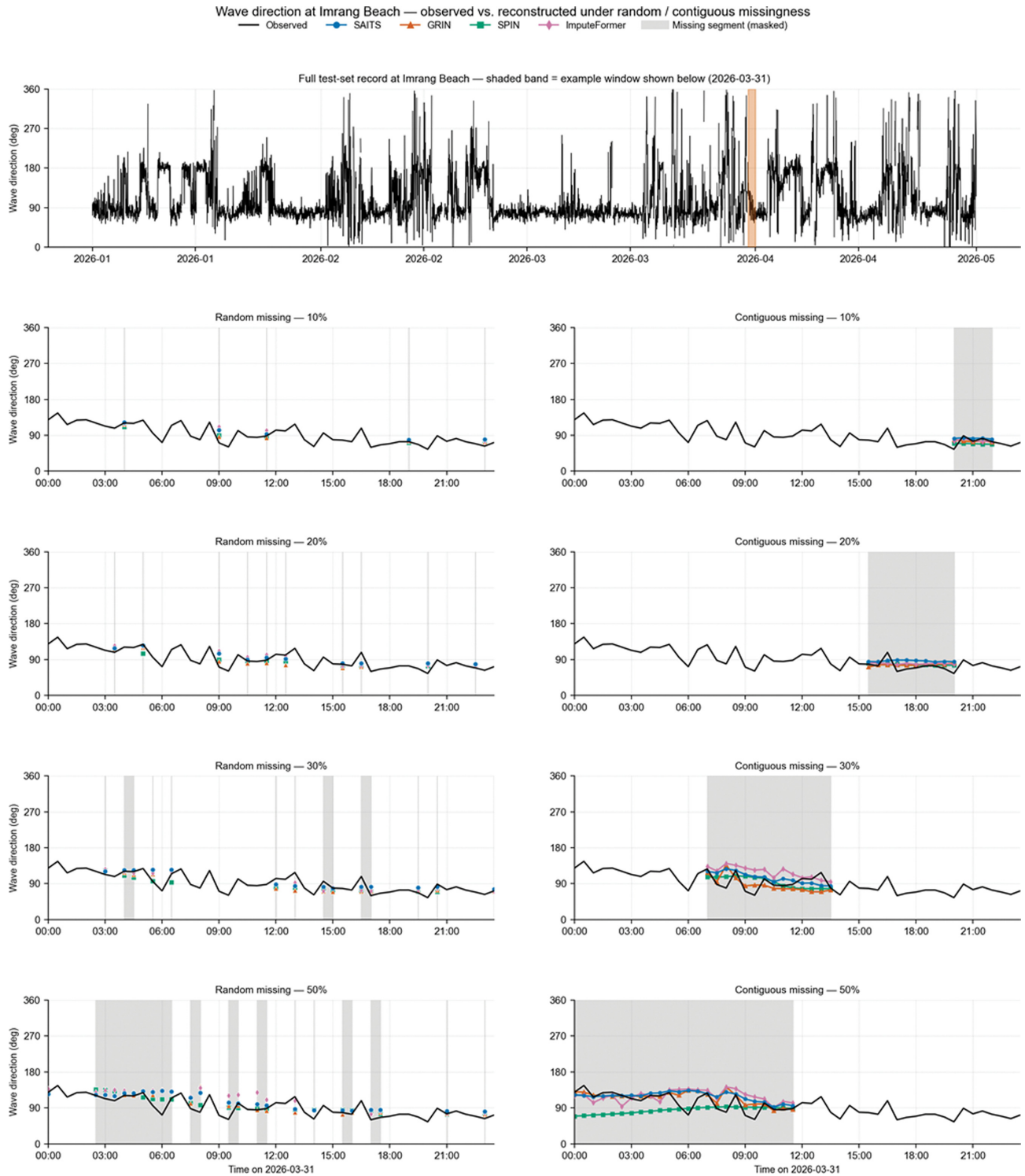


Fig. B8. Comparison of observed and reconstructed wave direction at Imrang Beach under random and contiguous missingness conditions.

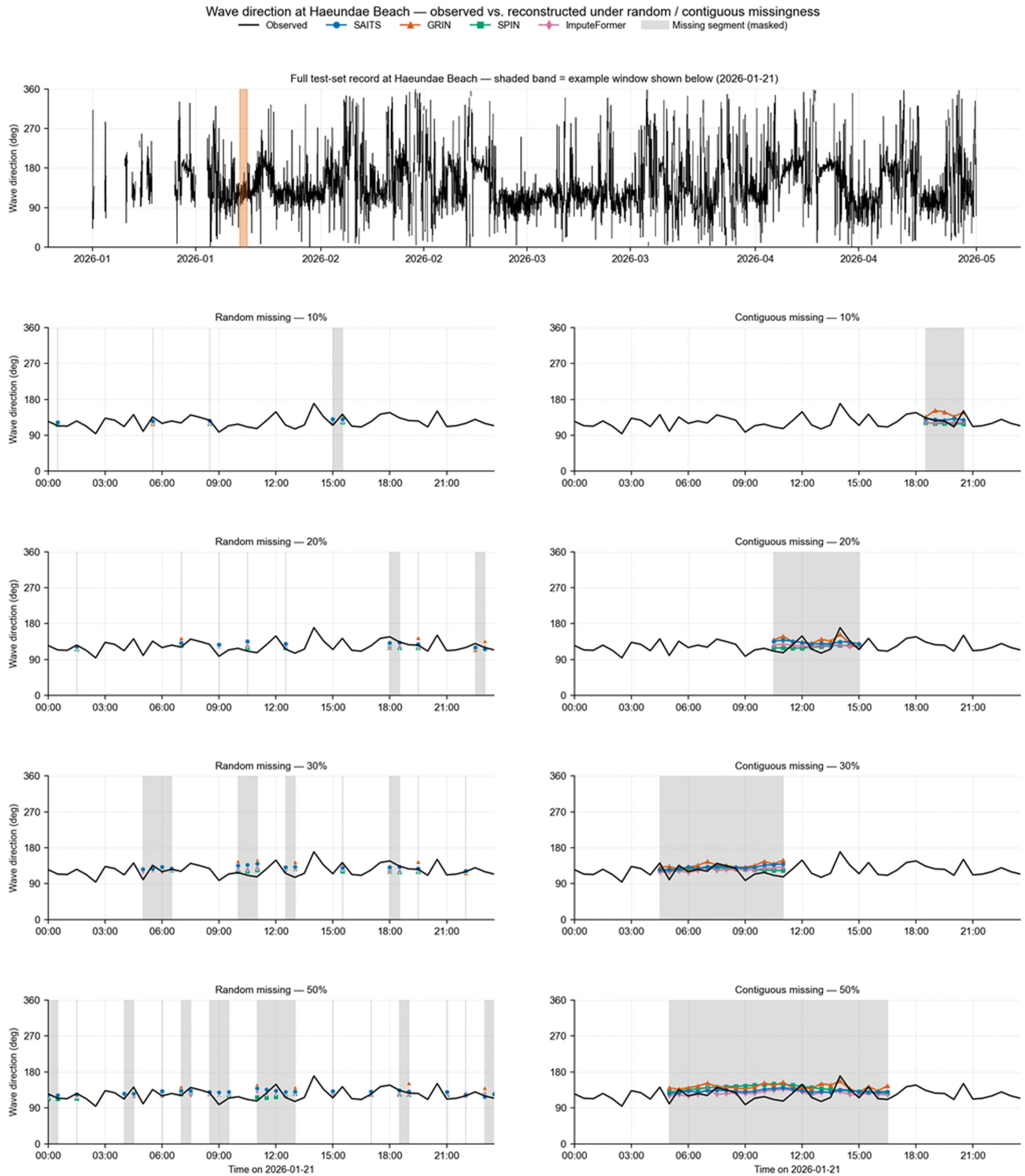


Fig. B9. Comparison of observed and reconstructed wave direction at Haeundae Beach under random and contiguous missingness conditions.