

(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2025-0100232

(43) 공개일자 2025년07월03일

(51) 국제특허분류(Int. Cl.)

G01W 1/18 (2006.01) G06N 3/08 (2023.01)

(52) CPC특허분류

G01W 1/18 (2013.01)

G06N 3/08 (2023.01)

(21) 출원번호 10-2023-0191584

(22) 출원일자 2023년12월26일

심사청구일자 2023년12월26일

(71) 출원인

인하대학교 산학협력단

인천광역시 미추홀구 인하로 100(용현동, 인하대 학교)

(72) 발명자

우승범

경기도 성남시 분당구 대왕판교로 155 (금곡동) 108동 202호

임채욱

인천광역시 미추홀구 용정공원로83번길 49 (용현동) 1947호

서지명

서울특별시 중랑구신내로 128(신내동, 동성아파트) 6동 1301호

(74) 대리인

특허법인아이더스

전체 청구항 수 : 총 1 항

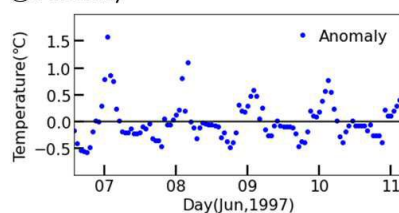
(54) 발명의 명칭 오토인코더를 이용한 해양 표층 정점 수온 관측자료 오측 탐지 방법

(57) 요약

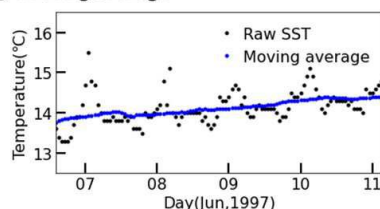
본 발명에 의한 오토인코더를 이용한 해양 표층 수온 관측자료 오측 탐지 방법은, 해양 표층 수온 관측자료를 이용하여 장주기 성분 오토인코더를 학습시키는 제1단계; 해양 표층 수온 관측자료를 이용하여 단주기 성분 오토인코더를 학습시키는 제2단계; 장주기 성분과 단주기 성분이 학습된 오토인코더에 표층 수온 관측자료를 입력하여 재구성값을 출력하는 제3단계; 해양 표층 수온 관측자료와 재구성값의 차이를 제공하여 이상점수의 값을 구하는 제4단계; 이상점수의 값이 관정 기준의 임계치값보다 큰 측정값들을 오측으로 판단하고 이상점수의 값이 관정 기준의 임계치값보다 작은 측정값은 정상 측정값으로 판단하는 제5단계;를 포함한다.

대표도 - 도6

① Anomaly



② Moving average



(52) CPC특허분류

G01W 2201/00 (2013.01)

이 발명을 지원한 국가연구개발사업

과제고유번호	1711179347
과제번호	00155915
부처명	과학기술정보통신부
과제관리(전문)기관명	정보통신기획평가원
연구사업명	인공지능융합혁신인재양성
연구과제명	인공지능융합혁신인재양성(인하대학교)
기 여 율	30/100
과제수행기관명	인하대학교 산학협력단
연구기간	2023.01.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1711198990
과제번호	00254371
부처명	과학기술정보통신부
과제관리(전문)기관명	(재단)과학기술일자리진흥원
연구사업명	실험실창업지원
연구과제명	2023년도 실험실 특화형 창업선도대학 사업(인하대학교)
기 여 율	30/100
과제수행기관명	인하대학교
연구기간	2023.04.01 ~ 2023.12.31

이 발명을 지원한 국가연구개발사업

과제고유번호	1525014994
과제번호	00256330
부처명	해양수산부
과제관리(전문)기관명	해양수산과학기술진흥원
연구사업명	쿠로시오 해류로 인한 한반도 해양위기 대응기술개발
연구과제명	쿠로시오 해류로 인한 한반도 해양위기 대응기술 개발
기 여 율	40/100
과제수행기관명	안양대학교 산학협력단
연구기간	2023.04.01 ~ 2027.12.31

명세서

청구범위

청구항 1

해양 표층 수온 관측자료를 이용하여 장주기 성분 오토인코더를 학습시키는 제1단계;

해양 표층 수온 관측자료를 이용하여 단주기 성분 오토인코더를 학습시키는 제2단계;

장주기 성분과 단주기 성분이 학습된 오토인코더에 표층 수온 관측자료를 입력하여 재구성값을 출력하는 제3단계;

해양 표층 수온 관측자료와 재구성값의 차이를 제공하여 이상점수의 값을 구하는 제4단계;

이상점수의 값이 판정 기준의 임계치값보다 큰 측정값들을 오측으로 판단하고 이상점수의 값이 판정 기준의 임계치값보다 작은 측정값은 정상 측정값으로 판단하는 제5단계;

를 포함하는 것을 특징으로 하는 오토인코더를 이용한 해양 표층 수온 관측자료 오측 탐지 방법.

발명의 설명

기술 분야

[0001] 본 발명은 오토인코더를 이용한 해양 표층 수온 관측자료 오측 탐지 방법에 관한 것으로서, 더욱 상세하게는, 비지도 기반의 학습 방법을 사용하는 오토인코더를 단주기·장주기 성분을 학습하는 준지도 학습 방법으로 수정하여 해양 표층 정점 수온 관측자료를 학습시킨 후 관측자료의 오측을 탐지하는 방법에 관한 것이다.

배경 기술

[0002] 해양 표층 수온은 해양-대기 상호작용, 기후 변화 등 해양과 대기의 현상을 이해하는 가장 중요한 변수 중 하나이다. 따라서 다양한 방식으로 해양 표층 수온을 관측하는데, 그 중 가장 중요한 자료 중 하나는 정점 관측 자료이다.

[0003] 도 1은 해양 표층 정점 수온 관측자료와 오측 자료 및 결측 자료의 예이다.

[0004] 도 1은 서해 덕적도의 해양 기상 부이 표층 수온을 1시간 단위로 1995년부터 2020년까지 측정하여 저장한 자료이다. 1996년~2003년의 자료를 검증자료, 2004년~2020년의 자료를 학습자료로 사용하였다.

[0005] 이러한 관측자료에는 관측상의 오차 또는 오류가 있을 수 있다.

[0006] 도 1의 아래쪽 그래프는 위쪽 그래프 중 1997년의 그래프를 확대한 그래프인데, 파란색 사각형 부분처럼 데이터가 없는 부분(결측된 부분이라고 말할 수 있고, 도 1에서는 "Missing"이라고 표시됨)도 있고, 빨간색 네모 부분처럼 데이터가 잘못된 부분(오측된 부분이라고 말할 수 있고, 도 1에서는 "Outlier"라고 표시됨)도 있다.

[0007] 이렇게 오측된 부분(잘못 측정된 부분)을 찾는 종래의 방법 중 하나는 비지도 학습 방법을 사용하는 오토인코더를 이용하여 오류를 찾는 방법이다. 정답 라벨이 없는 데이터를 비슷한 특징끼리 군집화 하여 새로운 데이터에 대한 결과를 예측하는 방법을 비지도학습 이라고 한다.

[0008] 도 2는 오토인코더 구조의 예이다.

[0009] 오토인코더(Autoencoder, AE)는 주어진 입력 데이터를 효과적으로 표현하고 압축하기 위해 설계된 인공 신경망의 한 유형이다. 주로 비지도 학습(Unsupervised Learning)으로 사용되며, 입력 데이터의 특징을 추출하거나 데이터의 차원을 줄이는 데 활용된다. 오토인코더는 크게 인코더(Encoder)와 디코더(Decoder) 두 부분으로 구성되어 있다.

[0010] 인코더는 입력 데이터를 저차원의 표현으로 매핑하고, 디코더는 이 저차원 표현을 다시 원래 차원으로 복원한다. 이러한 과정을 통해 모델은 입력 데이터의 중요한 특징을 학습하게 된다.

[0011] 오토인코더는 데이터에서 특징을 추출하고 효과적인 표현을 학습하는 데 사용되는 인공 신경망이다.

- [0012] 해양 표층 수온 관측자료 오측 탐지에 비지도 오토인코더를 적용한다면, 도 1과 같은 관측자료에 의해 학습된 비지도 오토인코더에, 도 1과 같은 관측자료가 입력되면, 학습된 비지도 오토인코더는 정측에 가깝게 재구성된 재구성값을 출력한다. 즉 무엇이 오측값인지 지정하지 않아도 정측에 가깝게 재구성된 재구성값을 출력한다.
- [0013] 도 3 (a) 은 학습이 완료된 비지도 오토인코더의 입력과 출력의 예이다.
- [0014] 검은색 점들의 값(Rawdata)이 입력되면, 빨간색 실선과 같은 값(재구성값)이 비지도 오토인코더(Unsupervised Autoencoder, AE)의 출력값으로 출력된다.
- [0015] 입력값과 출력값의 차이를 제공하여 이상점수(Outlier Score)를 산정할 수 있다.
- [0016] 도 3 (b) 는 도 3 (a)를 이용하여 이상점수 그래프를 그린 예이다.
- [0017] 판정 기준의 임계치값(Threshold)을 설정하고 이상점수(Outlier Score)의 값이 그 값보다 크면 오측으로 판단할 수 있다.
- [0018] 도 4는 비지도 오토인코더를 이용하여 오측을 탐지해낸 결과의 예이다.
- [0019] 도 4에서 검은색 점은 정상자료로 판정된 것이고, 붉은색 점들은 오측으로 판단된 것이다.
- [0020] 도 4의 좌측 상단부의 큰 사각형은 큰 사각형 오른쪽에 있는 작은 사각형(6월의 장기 오측 구간)을 확대한 것이다.
- [0021] 따라서 비지도 오토인코더는 다음의 단계를 포함한다.
- [0022] (1) 제1단계: 해양 표층 수온 관측자료에서 전체 평균을 뺀 아노말리(Anomaly) 성분을 이용하여 비지도 오토인코더를 학습시키는 단계
- [0023] (2) 제2단계: 해양 표층 수온 관측자료를 학습된 비지도 오토인코더에 입력하여 재구성값을 출력하는 단계
- [0024] (3) 제3단계: 해양 표층 수온 관측자료와 재구성값의 차이를 제공하여 이상점수(Outlier Score)의 값을 구하는 단계
- [0025] (4) 제4단계: 이상점수(Outlier Score)의 값이 판정 기준의 임계치값(Threshold)보다 큰 측정값들을 오측으로 판단하고 이상점수(Outlier Score)의 값이 판정 기준의 임계치값(Threshold)보다 작은 측정값은 정상 측정값으로 판단하는 단계
- [0026] 그런데 비지도 오토인코더의 경우 6월의 장기 오측 구간의 일부를 정상자료로 판단하여 검은색으로 표시하고 있는 등 여러 오류가 있다.
- [0027] 따라서 해양 표층 수온 관측자료의 오측을 좀 더 정확하게 탐지하는 방법이 필요하다.

선행기술문헌

특허문헌

- [0028] (특허문헌 0001) 한국 공개특허공보 KR 10-2020-0059085 A(공개일: 2020. 5. 28.)

발명의 내용

해결하려는 과제

- [0029] 본 발명의 목적은, 해양 표층 수온 관측자료의 오측을 효율적으로 정확하게 탐지해내는 것이다.

과제의 해결 수단

- [0030] 본 발명에 의한 오토인코더를 이용한 해양 표층 수온 관측자료 오측 탐지 방법은, 해양 표층 수온 관측자료를 이용하여 장주기 성분 오토인코더를 학습시키는 제1단계; 해양 표층 수온 관측자료를 이용하여 단주기 성분 오토인코더를 학습시키는 제2단계; 장주기 성분과 단주기 성분이 학습된 오토인코더에 표층 수온 관측자료를 입력하여 재구성값을 출력하는 제3단계; 해양 표층 수온 관측자료와 재구성값의 차이를 제공하여 이상점수의 값을 구하는 제4단계; 이상점수의 값이 판정 기준의 임계치값보다 큰 측정값들을 오측으로 판단하고 이상점수의 값이

판정 기준의 임계치값보다 작은 측정값은 정상 측정값으로 판단하는 제5단계;를 포함하는 것을 특징으로 한다.

발명의 효과

[0031] 본 발명에 의한 오토인코더를 이용한 해양 표층 정점 수온 관측자료 오측 탐지 방법은 해양 표층 수온 관측자료의 오측을 효율적으로 정확하게 탐지해낸다.

도면의 간단한 설명

[0032] 도 1은 해양 표층 수온 관측자료와 오측 자료 및 결측 자료의 예
 도 2는 오토인코더 구조의 예
 도 3 (a) 은 학습이 완료된 비지도 오토인코더의 입력과 출력의 예
 도 3 (b) 는 도 3 (a)를 이용하여 이상점수 그래프를 그린 예
 도 4는 비지도 오토인코더를 이용하여 오측을 탐지해낸 결과의 예
 도 5는 비지도 오토 인코더의 아노말리 성분과 오토 라벨링을 통해 학습시킨 정측 자료의 예
 도 6는 단주기 성분의 아노말리 성분과 장주기 성분의 이동 평균 성분의 예
 도 7은 본 발명이 제안한 방법 1에 의해 오측을 탐지해낸 결과의 예
 도 8는 본 발명에서 사용한 평가지표
 도 9은 본 발명에 의한 정량 평가 결과

발명을 실시하기 위한 구체적인 내용

[0033] 이하, 첨부한 도면을 참조하여 본 발명의 바람직한 실시예를 보다 상세하게 설명한다. 그리고, 본 발명은 다수의 상이한 형태로 구현될 수 있고, 기술된 실시 예에 한정되지 않는다.

[0034] 본 발명은 비지도 학습 기반의 오토인코더를 단주기 · 장주기 성분을 학습하는 준지도 기반의 오토인코더로 구성하여 해양 표층 정점 수온 관측자료 오측 탐지 방법이고, 다음의 방법1과 같다.

[0035] - 방법 1-

[0036] 방법 1은 다음의 단계를 포함한다.

[0037] (1) 제1단계: 해양 표층 수온 관측자료를 이용하여 장주기 성분 오토인코더를 학습시키는 단계

[0038] 이 단계에서는, 수온 관측자료를 이용하여 24시간 평균을 수행한 이동 평균(Moving Average) 성분으로 변환하여 학습한다.

[0039] (2) 제2단계: 해양 표층 수온 관측자료를 이용하여 단주기 성분 오토인코더를 학습시키는 단계

[0040] 이 단계에서는, 수온 관측자료에서 24시간 평균을 뺀 아노말리(Anomaly) 성분으로 변환하여 학습한다.

[0041] (3) 제3단계: 장주기 성분과 단주기 성분이 학습된 오토인코더에 해양 표층 수온 관측자료를 입력하여 재구성값을 출력하는 단계

[0042] (4) 제4단계: 해양 표층 수온 관측자료과 재구성값의 차이를 제공하여 이상점수(Outlier Score)의 값을 구하는 단계

[0043] (5) 제5단계: 이상점수(Outlier Score)의 값이 판정 기준의 임계치값(Threshold)보다 큰 측정값들을 오측으로 판단하고 이상점수(Outlier Score)의 값이 판정 기준의 임계치값(Threshold)보다 작은 측정값은 정상 측정값으로 판단하는 단계

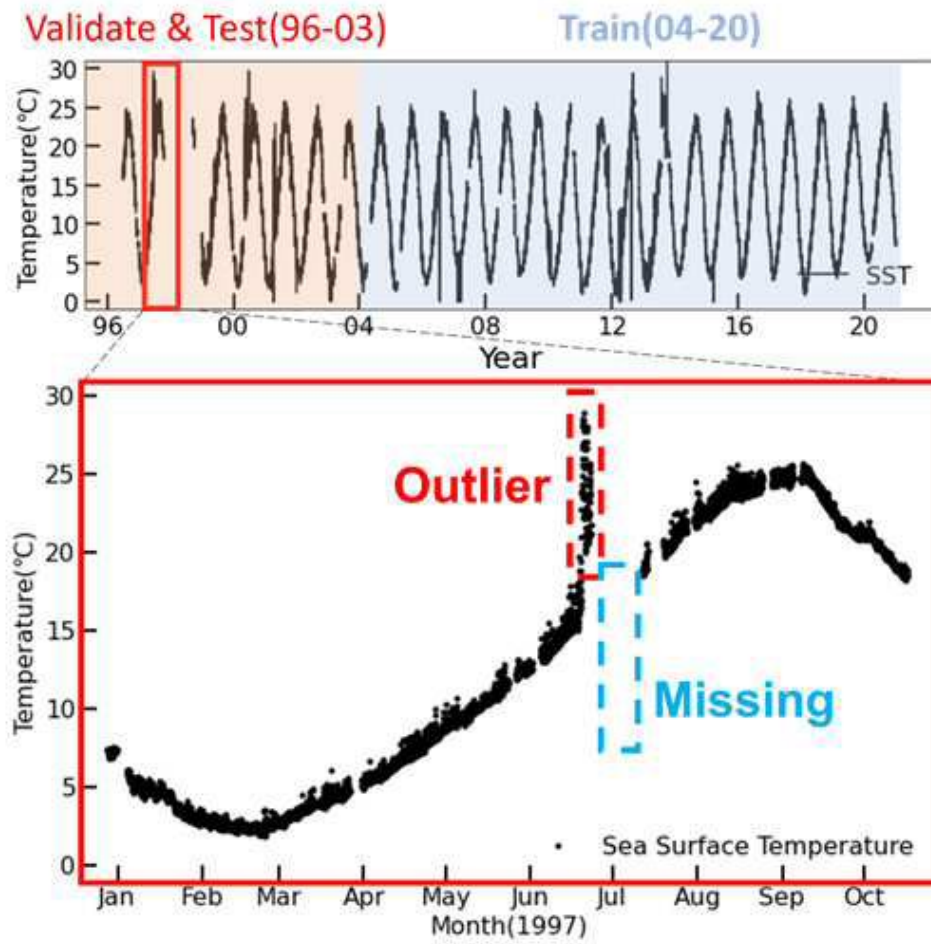
[0044] 도 5는 비지도 오토 인코더의 아노말리(Anomaly) 성분과 오토 라벨링(Auto Labeling)을 통해 학습시킨 정측 자료의 예이다.

[0045] 검은색 점은 아노말리(Anomaly) 자료를 의미하며, 붉은색 선은 정측자료로 판단을 위한 임계치값(Threshold)을 의미하고 푸른색 점은 임계치 값 내부에 존재하는 정측 값으로 판단된 아노말리(Anomaly) 자료를 의미한다.

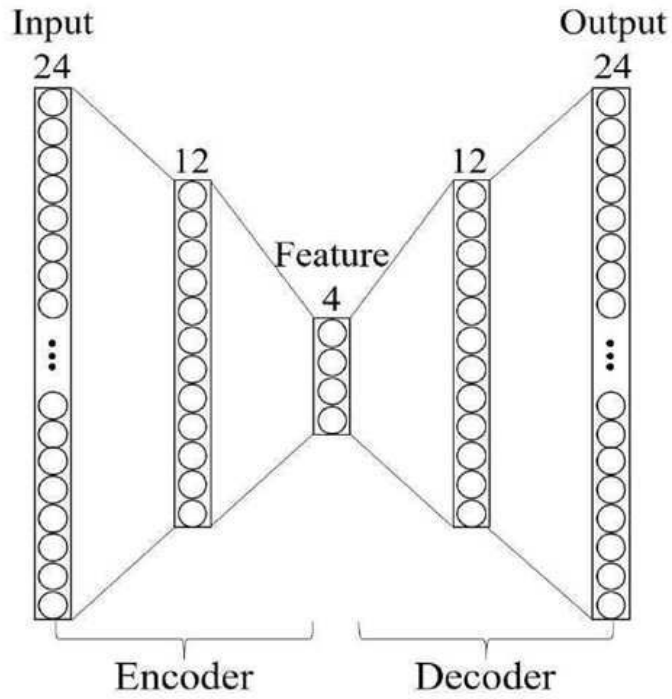
- [0046] 도 6는 단주기 성분의 아노말리 성분과 장주기 성분의 이동 평균 성분의 예이다.
- [0047] 도 6의 ①번 그림은 단주기 성분인 아노말리(Anomaly) 자료의 예이다.
- [0048] 장주기 성분과 단주기 성분이 모두 존재하는 수온 관측자료에서 장주기 성분만을 포함하는 Moving Average 성분을 빼서 제거하게 되면, 일일 이하의 주기를 가지는 성분만이 남게 되고 이를 단주기 성분이며 평균을 제거하였기 때문에 Anomaly 성분으로 명명하였다.
- [0049] 도 6의 ②번 그림은 장주기 성분인 이동 평균(Moving Average) 자료의 예이다.
- [0050] 장주기 성분을 푸른색 점으로 표시하였고, 원본 수온자료를 검은색 점으로 표시하였다.
- [0051] 수온 관측자료를 A라고 한다면 24시간 이동 평균의 경우 수온 자료를 24시간의 묶음으로 변환하여 24시간의 값을 평균한 자료를 생성한다. 이때 앞의 기준인 n번째를 기준으로 앞의 자료 n-11, n+12까지의 자료를 평균한다.
- [0052] 이때 n이 12보다 작거나 n+12의 값이 전체 수온 관측자료의 크기보다 클 경우 최대로 가용할 수 있는 자료만을 이용하여 평균 값을 산정한다.
- [0053] 예를 들어 12번째 자료의 경우 1번부터 24번까지의 수온 자료를 평균한 값을 가지게 된다.
- [0054] 반면 3번째 자료의 경우 1번부터 15번째 자료만을 이용하여 평균 값을 산출 하게 되고, 100의 크기를 가지는 자료에서 98번째 자료의 경우 86번째부터 100번째까지의 자료를 이용하여 산정한다.
- [0055] 이렇게 산정한 평균값을 24시간 평균을 수행한 이동 평균(Moving Average) 성분으로 명명하며 24시간 평균을 수행하여 일일 변동 성분을 제거하였기 때문에 변동 주기가 하루 이상의 성분만을 포함하게 된다.
- [0056] 일일 주기 성분과 장주기 성분이 모두 포함된 자료(Raw data)에서 장주기 성분인 이동 평균과의 차이를 통해 장주기 성분을 제거한 일일주기 성분의 아노말리(Anomaly) 성분을 단주기 성분을 학습하기 위한 입력자료로 구성할 수 있다.
- [0057] 도 7은 본 발명이 제안한 방법에 의해 오측을 탐지해낸 결과의 예이다.
- [0058] 도 7의 경우, 도 4의 결과에 비해, 빨간색 사각형(6월 14일에서 1월 17일 사이의 장기 오측구간)의 데이터를 모두 붉은 점으로 표시하여 모두 오측으로 판단하였다. 이것은 도 4와 도 7에서 탐지해내지 못한 오측을 정확히 탐지해낸 것으로 볼 수 있다.
- [0059] 도 8는 본 발명에서 사용한 평가지표이다.
- [0060] 도 8에서는 관측자료로부터 오측을 판단한 결과를 판단하는 지표로 정확도와 재현율을 제시하고, 그에 대한 정의를 내리고 있다.
- [0061] 도 9은 본 발명에 의한 정량 평가 결과이다.
- [0062] 도 9에서, unsupervised AE(비지도 오토인코더)를 이용한 탐지결과와 Semisupervised AE(본 발명의 방법 1에 의한 준지도학습 기반 오토인코더)를 이용한 오측 탐지결과를 보면, 정확도와 재현율이 향상되었음을 확인할 수 있다.

도면

도면1



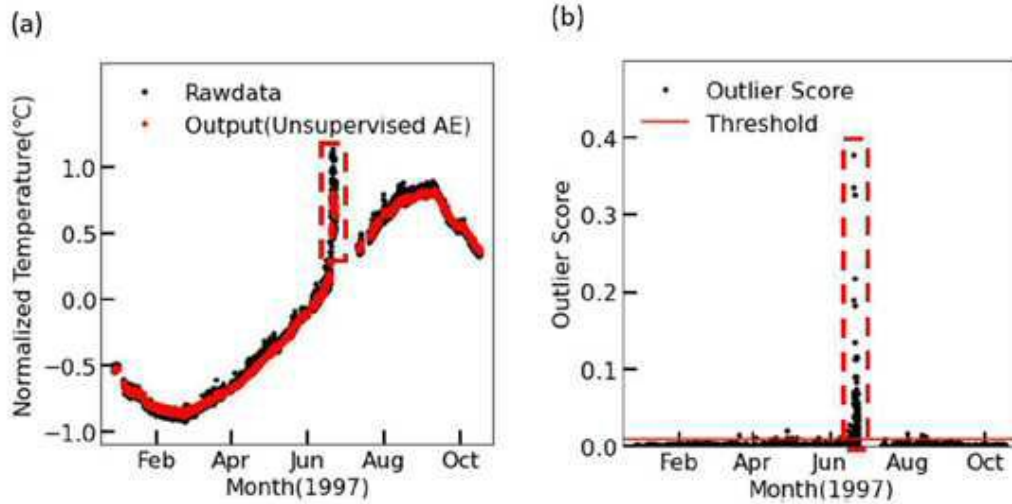
도면2



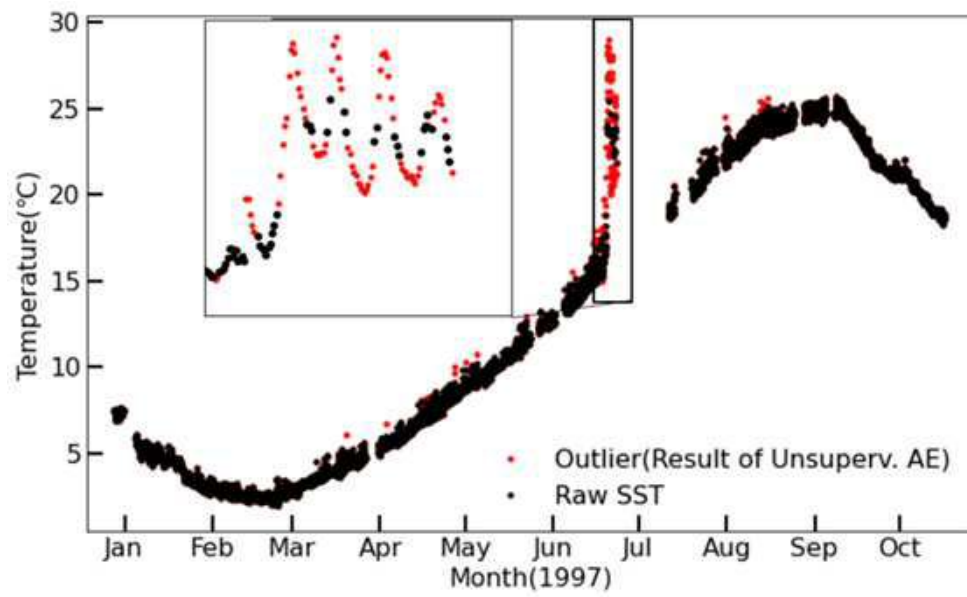
$$E(x) = W_2(F(W_1 * x + b_1)) + b_2$$

$$D(x) = W_4(F(W_3 * E(x) + b_3)) + b_4$$

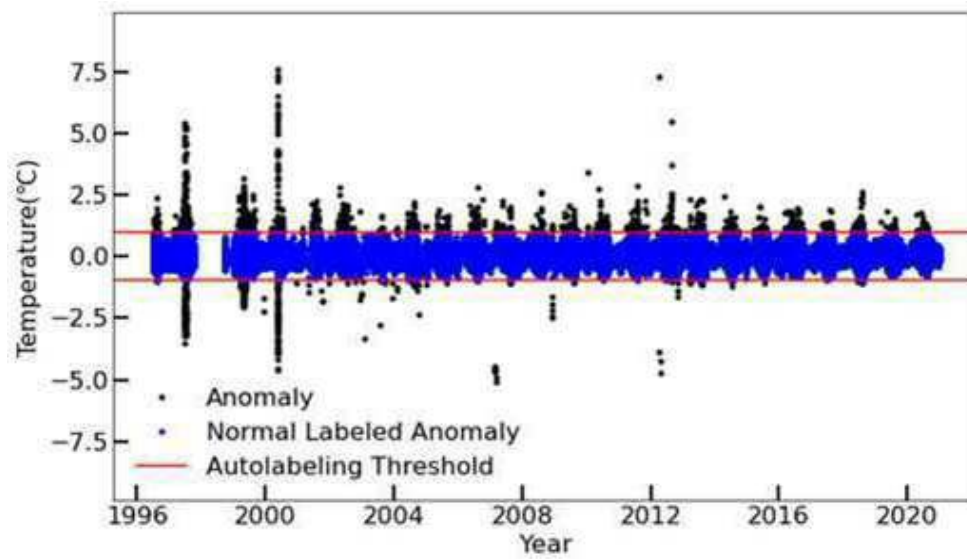
도면3



도면4

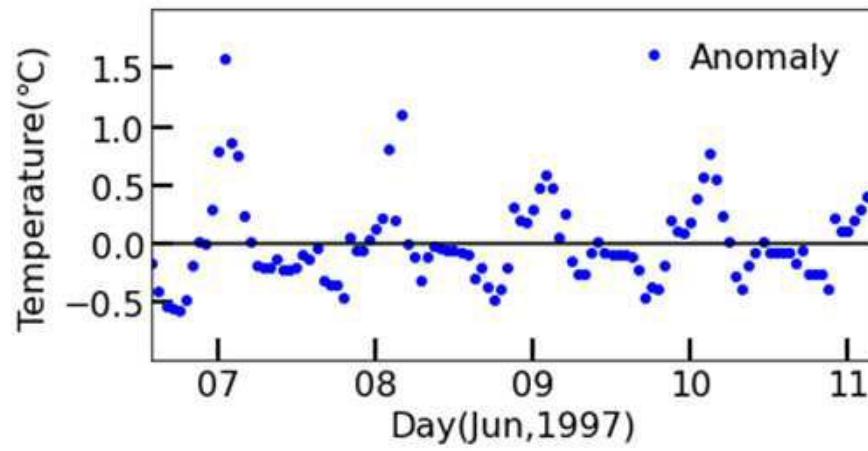


도면5

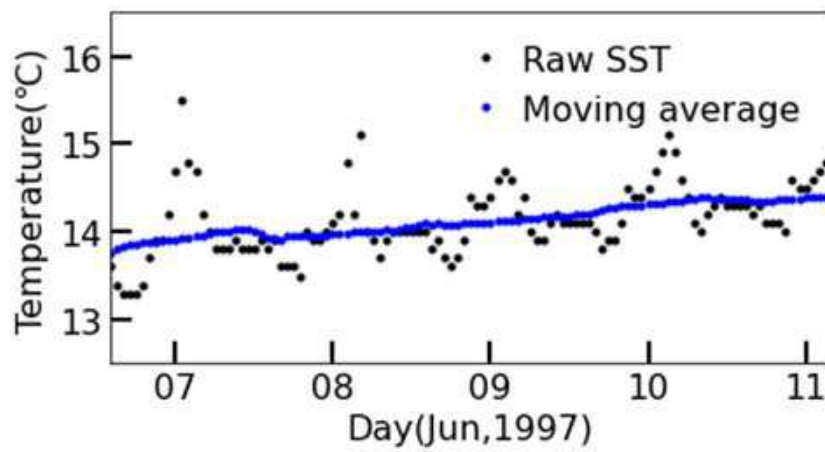


도면6

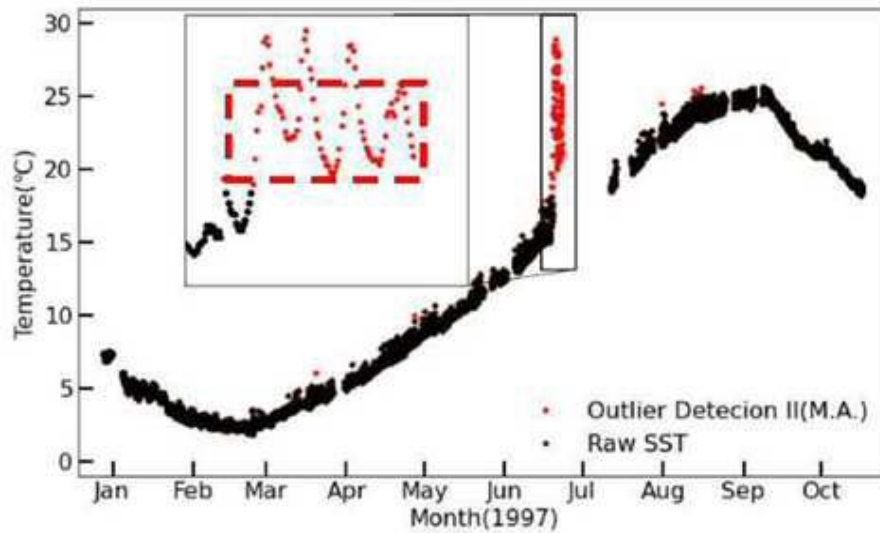
① Anomaly



② Moving average



도면7



도면8

지표		정측 자료 판단	오측 자료 판단
	실제 정측 자료	True Negative	False Positive
	실제 오측 자료	False Negative	True Positive

정확도: 전체 자료 중 분류 정보를 제대로 맞춘 비율

$$Accuracy = \frac{TP + TN}{TP + FN + FP + TN}$$

재현율: 실제 오측 자료 중 오측 자료로 판단된 비율

$$Recall = \frac{TP}{TP + FN}$$

도면9

Unsupervised AE	정측 자료 판단	오측 자료 판단	정확도: 0.98 재현율: 0.36
실제 정측 자료(3220)	3211	9	
실제 오측 자료(88)	56	32	

Semisupervised AE	정측 자료 판단	오측 자료 판단	정확도: 0.99 재현율: 0.97
실제 정측 자료(3220)	3216	4	
실제 오측 자료(88)	3	85	