



(19) 대한민국특허청(KR)  
(12) 등록특허공보(B1)

(45) 공고일자 2025년02월17일  
(11) 등록번호 10-2768733  
(24) 등록일자 2025년02월12일

- (51) 국제특허분류(Int. Cl.)  
G06N 20/20 (2019.01) G01N 15/00 (2025.01)  
G01N 15/0227 (2024.01) G06N 3/126 (2023.01)
- (52) CPC특허분류  
G06N 20/20 (2021.08)  
G01N 15/0227 (2024.01)
- (21) 출원번호 10-2024-0067600  
(22) 출원일자 2024년05월24일  
심사청구일자 2024년05월24일
- (56) 선행기술조사문헌  
KR102374002 B1\*  
US20040164235 A1\*  
KR102642421 B1\*  
\*는 심사관에 의하여 인용된 문헌
- (73) 특허권자  
세종대학교산학협력단  
서울특별시 광진구 능동로 209 (군자동, 세종대학교)
- (72) 발명자  
아불가셈  
서울시 광진구 능동로 209 대양AI센터, 433호  
최수미  
서울시 광진구 능동로 209 대양AI센터, 720호  
(뒷면에 계속)
- (74) 대리인  
양성보

전체 청구항 수 : 총 7 항

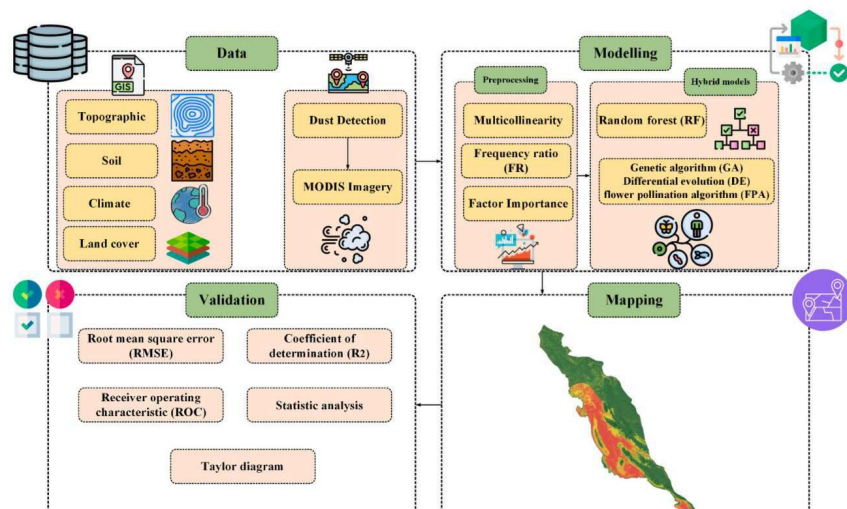
심사관 : 박승철

(54) 발명의 명칭 위성 이미지 시계열 및 앙상블 기계 학습과 진화 알고리즘의 통합을 지원하여 먼지 감지 및 민감성 매핑 방법 및 시스템

(57) 요약

위성 이미지 시계열 및 앙상블 기계 학습과 진화 알고리즘의 통합을 지원하여 먼지 감지 및 민감성 매핑 방법 및 시스템이 개시된다. 일 실시예에 따른 먼지 민감성 매핑 시스템에 의해 수행되는 먼지 민감성 매핑 방법은, 위성 이미지로부터 감지된 먼지 발생 지점과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비하는 단계; 기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하는 단계; 상기 구축된 먼지 민감성 매핑 모델을 통해 상기 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성하는 단계; 및 상기 생성된 먼지 민감성 지도에 기초하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 단계를 포함할 수 있다.

대표도 - 도1



(52) CPC특허분류

**G06N 3/126** (2023.01)

**G01N 2015/0096** (2013.01)

(72) 발명자

**바히드 라자비 데르메**

서울특별시 광진구 능동로 209, 광개토관, 1001호  
(군자동)

**리즈완 알리 나크비**

서울시 광진구 능동로 209 대양AI센터, 452호

이 발명을 지원한 국가연구개발사업

과제고유번호 1711179366

과제번호 RS-2022-00156354

부처명 과학기술정보통신부

과제관리(전문)기관명 정보통신기획평가원

연구사업명 대학ICT연구센터지원사업

연구과제명 실-가상 연계 메타버스를 위한 초실감 XR 기술 연구

과제수행기관명 세종대학교산학협력단

연구기간 2022.07.01 ~ 2029.12.31

공지예외적용 : 있음

---

## 명세서

### 청구범위

#### 청구항 1

먼지 민감성 매핑 시스템에 의해 수행되는 먼지 민감성 매핑 방법에 있어서,

위성 이미지로부터 감지된 먼지 발생 지점과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비하는 단계;

기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하는 단계;

상기 구축된 먼지 민감성 매핑 모델을 통해 상기 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성하는 단계; 및

상기 생성된 먼지 민감성 지도에 기초하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 단계를 포함하고,

상기 준비하는 단계는,

MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지를 이용한 먼지 모니터링을 통해 먼지 발생 지점을 감지하고, 먼지 발생에 영향을 미치는 지형, 기후, 토양 특성, 토지 속성을 포함하는 복수 개의 요인을 수집하고, 상기 수집된 복수 개의 요인과 상기 감지된 먼지 발생 지점을 포함하는 먼지 발생 데이터를 준비하는 단계

를 포함하고,

상기 구축하는 단계는,

상기 먼지 발생 데이터에 대해 주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도를 포함하는 데이터 전처리를 수행하고, 랜덤 포레스트(Random forest, RF) 모델에 유전 알고리즘(Genetic algorithm, GA), 차분 진화(Differential Evolution, DE) 알고리즘, 꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 중 어느 하나의 진화 알고리즘을 앙상블함에 따라 랜덤 포레스트 기반의 먼지 민감성 매핑 모델을 구축하고, 먼지 발생 데이터를 이용하여 상기 구축된 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정하는 단계

를 포함하는 먼지 민감성 매핑 방법.

#### 청구항 2

삭제

#### 청구항 3

삭제

#### 청구항 4

삭제

#### 청구항 5

삭제

#### 청구항 6

제1항에 있어서,

상기 구축하는 단계는,

상기 차분 진화 알고리즘을 통해 먼지에 영향을 미치는 요인의 중요도를 계산하는 단계를 포함하는 먼지 민감성 매핑 방법.

#### 청구항 7

제1항에 있어서,

상기 검증하는 단계는,

RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 단계

를 포함하는 먼지 민감성 매핑 방법.

#### 청구항 8

제1항, 제6항 내지 제7항 중 어느 한 항의 먼지 민감성 매핑 방법을 상기 먼지 민감성 매핑 시스템에 실행시키기 위해 비-일시적인 컴퓨터 판독가능한 기록 매체에 저장되는 컴퓨터 프로그램.

#### 청구항 9

먼지 민감성 매핑 시스템에 있어서,

위성 이미지로부터 감지된 먼지 발생 지점과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비하는 데이터 준비부;

기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하는 모델링부;

상기 구축된 먼지 민감성 매핑 모델을 통해 상기 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성하는 지도 생성부; 및

상기 생성된 먼지 민감성 지도에 기초하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 성능 검증부를 포함하고,

상기 데이터 준비부는,

MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지를 이용한 먼지 모니터링을 통해 먼지 발생 지점을 감지하고, 먼지 발생에 영향을 미치는 지형, 기후, 토양 특성, 토지 속성을 포함하는 복수 개의 요인을 수집하고, 상기 수집된 복수 개의 요인과 상기 감지된 먼지 발생 지점을 포함하는 먼지 발생 데이터를 준비하는 것을 포함하고,

상기 모델링부는,

상기 먼지 발생 데이터에 대해 주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도를 포함하는 데이터 전처리를 수행하고, 랜덤 포레스트(Random forest, RF) 모델에 유전 알고리즘(Genetic algorithm, GA), 차분 진화(Differential Evolution, DE) 알고리즘, 꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 중 어느 하나의 진화 알고리즘을 앙상블함에 따라 랜덤 포레스트 기반의 먼지 민감성 매핑 모델을 구축하고, 먼지 발생 데이터를 이용하여 상기 구축된 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정하는

먼지 민감성 매핑 시스템.

#### 청구항 10

삭제

#### 청구항 11

삭제

#### 청구항 12

삭제

### 청구항 13

삭제

### 청구항 14

제9항에 있어서,

상기 모델링부는,

상기 차분 진화 알고리즘을 통해 먼지에 영향을 미치는 요인의 중요도를 계산하는

것을 특징으로 하는 먼지 민감성 매핑 시스템.

### 청구항 15

제9항에 있어서,

상기 성능 검증부는,

RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는

것을 특징으로 하는 먼지 민감성 매핑 시스템.

## 발명의 설명

### 기술 분야

[0001] 아래의 설명은 위성 이미지를 이용하여 먼지 민감성을 매핑하는 기술에 관한 것이다.

### 배경 기술

[0003] 먼지 폭풍은 경제, 환경, 건강(급성 호흡기 질환), 농산물에 부정적인 영향을 미치는 치명적인 기상 현상이다. 사막 지역, 반건조 사막, 마른 호수 바닥은 먼지를 생성하는 지역 중 하나이다. 먼지 발생은 빈도, 위치, 크기, 먼지 방출 강도, 입자 모양 및 크기에 영향을 받는다. 먼지 폭풍은 건조하고 사막 지역의 풍속 (6.5~7m/s)이 풍식 임계값을 초과하여 공기 중에 부유하는 토양 입자의 수가 증가할 때 발생한다. 이 경우 가시성은 1km 미만으로 감소하며, 극단적인 경우 먼지 입자 농도가 6000g/m<sup>3</sup>를 초과한다. 먼지 폭풍은 지구 표면의 30% 이상을 차지하는 건조 및 반건조 지역에서 더욱 빈번해지고 심해지고 있다. 주요 먼지 발생원 중에는 이라크, 사우디아라비아, 이란, 시리아 등 중동 사막이 있다. 인간 활동과 기후 변화는 지난 수십 년 동안 중동 지역에서 심각한 환경 악화를 일으켰으며, 이 지역은 전 세계 먼지 입자의 약 25%를 차지한다. 이러한 추세는 중동과 이란에서 환경 문제로 대두되어 주민들의 삶의 질, 생태계, 가시성, 교통, 미기후 및 통신 시스템에 영향을 미치고 있다.

[0004] 한편, 대한민국공개특허 제 10-2374002호(2022.03.08. 등록일)에 위성이미지를 통한 미세먼지농도 측정 장치, 방법 및 프로그램이 개시되어 있다.

### 발명의 내용

#### 해결하려는 과제

[0006] 랜덤 포레스트와 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하고, 구축된 먼지 민감성 매핑 모델을 통해 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성할 수 있다.

#### 과제의 해결 수단

[0008] 먼지 민감성 매핑 시스템에 의해 수행되는 먼지 민감성 매핑 방법은, 위성 이미지로부터 감지된 먼지 발생 지점

과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비하는 단계; 기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하는 단계; 상기 구축된 먼지 민감성 매핑 모델을 통해 상기 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성하는 단계; 및 상기 생성된 먼지 민감성 지도에 기초하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 단계를 포함할 수 있다.

- [0009] 상기 준비하는 단계는, MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지를 이용한 먼지 모니터링을 통해 먼지 발생 지점을 감지하고, 먼지 발생에 영향을 미치는 지형, 기후, 토양 특성, 토지 속성을 포함하는 복수 개의 요인을 수집하고, 상기 수집된 복수 개의 요인과 상기 감지된 먼지 발생 지점을 포함하는 먼지 발생 데이터를 준비하는 단계를 포함할 수 있다.
- [0010] 상기 모델링하는 단계는, 상기 먼지 발생 데이터에 대해 주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도를 포함하는 데이터 전처리를 수행하는 단계를 포함할 수 있다.
- [0011] 상기 모델링하는 단계는, 랜덤 포레스트(Random forest, RF) 모델에 유전 알고리즘(Genetic algorithm, GA), 차분 진화(Differential Evolution, DE) 알고리즘, 꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 중 어느 하나의 진화 알고리즘을 앙상블함에 따라 랜덤 포레스트 기반의 먼지 민감성 매핑 모델을 구축하는 단계를 포함할 수 있다.
- [0012] 상기 모델링하는 단계는, 먼지 발생 데이터를 이용하여 상기 구축된 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정하는 단계를 포함할 수 있다.
- [0013] 상기 모델링 단계는, 상기 차분 진화 알고리즘을 통해 먼지에 영향을 미치는 요인의 중요도를 계산하는 단계를 포함할 수 있다.
- [0014] 상기 검증하는 단계는, RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 단계를 포함할 수 있다.
- [0015] 먼지 민감성 매핑 방법을 상기 먼지 민감성 매핑 시스템에 실행시키기 위해 비-일시적인 컴퓨터 판독가능한 기록 매체에 저장되는 컴퓨터 프로그램을 포함할 수 있다.
- [0016] 먼지 민감성 매핑 시스템은, 위성 이미지로부터 감지된 먼지 발생 지점과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비하는 데이터 준비부; 기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축하는 모델링부; 상기 구축된 먼지 민감성 매핑 모델을 통해 상기 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 생성하는 지도 생성부; 및 상기 생성된 먼지 민감성 지도에 기초하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증하는 성능 검증부를 포함할 수 있다.
- [0017] 상기 데이터 준비부는, MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지를 이용한 먼지 모니터링을 통해 먼지 발생 지점을 감지하고, 먼지 발생에 영향을 미치는 지형, 기후, 토양 특성, 토지 속성을 포함하는 복수 개의 요인을 수집하고, 상기 수집된 복수 개의 요인과 상기 감지된 먼지 발생 지점을 포함하는 먼지 발생 데이터를 준비할 수 있다.
- [0018] 상기 모델링부는, 상기 먼지 발생 데이터에 대해 주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도를 포함하는 데이터 전처리를 수행할 수 있다.
- [0019] 상기 모델링부는, 랜덤 포레스트(Random forest, RF) 모델에 유전 알고리즘(Genetic algorithm, GA), 차분 진화(Differential Evolution, DE) 알고리즘, 꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 중 어느 하나의 진화 알고리즘을 앙상블함에 따라 랜덤 포레스트 기반의 먼지 민감성 매핑 모델을 구축할 수 있다.
- [0020] 상기 모델링부는, 먼지 발생 데이터를 이용하여 상기 구축된 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정할 수 있다.
- [0021] 상기 모델링 단계는, 상기 차분 진화 알고리즘을 통해 먼지에 영향을 미치는 요인의 중요도를 계산할 수 있다.
- [0022] 상기 먼지 민감성 매핑 시스템은, 상기 성능 검증부는, RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 상기 구축된 먼지 민감성 매핑 모델의 성능을 검증할 수 있다.

## 발명의 효과

[0024] 랜덤 포레스트 모델과 진화 알고리즘을 앙상블함으로써 먼지 민감성 모델링의 정확성을 향상시킬 수 있다.

## 도면의 간단한 설명

[0026] 도 1은 일 실시예에 있어서, 먼지 민감성 매핑 동작을 설명하기 위한 도면이다.

도 2는 일 실시예에 있어서, 연구 지역의 지리적 위치를 설명하기 위한 예이다.

도 3은 일 실시예에 있어서, 연구 지역의 먼지 발생을 설명하기 위한 도면이다.

도 4는 일 실시예에 있어서, 먼지에 영향을 미치는 요인을 설명하기 위한 도면이다.

도 5는 일 실시예에 있어서, 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정하는 동작을 설명하기 위한 도면이다.

도 6은 일 실시예에 있어서, 먼지 민감성 매핑 시스템을 설명하기 위한 블록도이다.

도 7은 일 실시예에 있어서, 먼지 민감성 매핑 방법을 설명하기 위한 흐름도이다.

## 발명을 실시하기 위한 구체적인 내용

[0027] 이하, 실시예를 첨부한 도면을 참조하여 상세히 설명한다.

[0029] 먼지가 인간의 건강과 환경에 미치는 영향을 완화하려면 먼지에 취약한 지역을 식별하는 모델과 지도를 생성하는 것이 중요하다. 실시예에서는 MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지를 사용하여 특정 지역에서 먼지 발생을 식별한 후, 앙상블 머신러닝 모델을 개선하여 먼지 민감도 지도(Dust Susceptibility Map, DSM)를 준비하고자 한다. 여기서, RF(Random Forest) 앙상블 모델을 향상시키기 위해 DE(Differential Evolution), GA(Genetic Algorithm), FPA(Flower Pollination Algorithm)를 사용한다. 모델링을 위해 복수 개(예를 들면, 519개)의 먼지 발생 지점(MODIS 이미지에서 추출)과 먼지에 영향을 미치는 15개의 요인(경사, 용적 밀도, 측면, 점토, 고도, 모래, 강우, 암석학, 토양 순서, 강까지의 거리, 토양 질감, 정규화된 차이 식생 지수(NDVI), 토양 수분 함량, 토지 피복 및 풍속)을 포함하는 공간 데이터베이스가 생성될 수 있다. DE 알고리즘을 활용하여 먼지 발생에 영향을 미치는 이러한 요인의 중요성을 결정한 결과, 고도, 풍속, 토지피복이 가장 큰 영향을 미치는 요인인 반면, 하천까지의 거리, 용적밀도, 토양조직은 먼지 발생에 덜 영향을 미치는 것으로 나타났다. 데이터는 다중공선성 분석과 빈도비(FR) 접근법을 사용하여 전처리될 수 있다. 실시예에서는 먼지 민감도 지도를 위해 RF-FPA, RF-GA 및 RF-DE라는 세 가지 RF 기반 메타 휴리스틱 최적화 알고리즘이 생성될 수 있다. 구축된 모델의 유효성 예측은 RMSE(Root-Mean-Square-Error) 지수, AUC-ROC(Area Under the Receiver Operating Characteristic Curve) 및 최고에서 최악까지의 결정 계수( $R^2$ )를 통해 RF-DE(RMSE = 0.131, AUC-ROC = 0.988,  $R^2$  = 0.93), RF-GA(RMSE = 0.141, AUC-ROC = 0.986,  $R^2$  = 0.919), RF-FPA(RMSE = 0.157, AUC-ROC = 0.981,  $R^2$  = 0.9) 및 RF(RMSE = 0.173, AUCROC = 0.964,  $R^2$  = 0.878)이다. 이에, 진화 알고리즘과 RF 모델을 결합하면 먼지 민감성 모델링의 정확성이 향상된다는 것을 보여준다.

[0030] 도 1은 일 실시예에 있어서, 먼지 민감성 매핑 동작을 설명하기 위한 도면이다.

[0031] 먼지 민감성 매핑 시스템은 데이터 준비, 모델링, 먼지 민감성 지도(Dust Susceptibility Map, DMS) 준비 및 결과 평가를 포함하는 네 가지의 주요 단계를 수행할 수 있다.

[0032] 데이터 준비 단계에서, 먼지 민감성 매핑 시스템은 MODIS(Moderate-Resolution Imaging Spectroradiometer) 이미지와 지형, 기후, 토지 피복, 토양을 포함하는 네 가지 범주에 대한 유효 기준과 관련된 데이터를 이용하여 먼지 발생 데이터를 준비할 수 있다.

[0033] 모델링 단계에서, 먼지 민감성 매핑 시스템은 데이터 전처리(주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도)와 모델링(RF 모델 개발)을 수행할 수 있다.

[0034] 먼지 민감성 지도 준비 단계에서, 먼지 민감성 매핑 시스템은 RF, RF-FPA, RF-GA, RF-DE의 네 가지 모델을 사용하여 먼지 민감성 지도를 생성할 수 있다.

[0035] 결과 평가 단계에서, 먼지 민감성 매핑 시스템은 RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver



Operating Characteristic) 곡선 및 통계 지수를 사용하여 모델링 결과 및 먼지 민감성 지도를 평가할 수 있다.

[0036] 도 2는 일 실시예에 있어서, 연구 지역의 지리적 위치를 설명하기 위한 예이다.

[0037] Bushehr 주는 페르시아만 연안에 긴 해안선을 따라 남부 이란에 위치해 있다. 이 지역은 북위 27° 19'~30° 16', 동경 50° 1'~52° 59'에 위치해 있다. Bushehr 주의 면적은 약 27,653km<sup>2</sup>이며, 2016년 인구 조사에 따르면 인구는 321,826가구에서 약 1,163,400명이다. Bushehr 주는 연평균 기온이 24C인 해안 기후로 겨울에는 6° C, 여름에는 50° C이다. 이 지역에는 연간 평균 강우량이 217mm이며, 강수량은 11월과 겨울 3월 사이로 제한된다. Bushehr 주의 우세한 풍향은 북서쪽이다. 이라크와 시리아의 외부 먼지 발생원으로 인해 이란의 서쪽과 남서쪽에서 먼지가 더 자주 발생한다. 건조하거나 반건조한 환경과 광범위한 사막 지역이 특징인 이 인근 국가들은 중요한 먼지 발생원 역할을 한다. 일반적인 바람 패턴은 먼지 입자를 운반하는 데 중요한 역할을 하며, 바람이 이라크와 시리아에서 입자를 운반하기 때문에 이란의 서부와 남서부 지역은 먼지에 취약하다. 지난 10년 동안, 주로 여름에 발생하는 이 지역의 먼지 폭풍은 중요한 환경 문제가 되었다.

[0038] 도 3은 일 실시예에 있어서, 연구 지역의 먼지 발생을 설명하기 위한 도면이다.

[0039] 도 3(a)는 MODIS 이미지를 이용한 먼지 감지, 도 3(b)는 먼지 발생 지점을 나타낸 예이다. 먼지 민감성 매핑 시스템은 데이터를 생성할 수 있다. 먼지 민감성 매핑 시스템은 위성 이미지를 이용하여 먼지를 감지할 수 있다. 다시 말해서, 먼지 민감성 매핑 시스템은 위성 이미지를 이용하여 먼지 발생 지역을 감지할 수 있다. 먼지 민감성 매핑 시스템은 먼지를 모니터링하기 위해 MODIS Terra 및 Aqua 위성 데이터를 사용할 수 있다. 이러한 이미지들은 Google 어스 엔진(GEE) 시스템에서 20년(2002-2022) 동안 에어로졸 광학 깊이(AOD)를 모니터링하는 데 사용되었다. 이러한 목표를 달성하기 위해, 먼지 민감성 매핑 시스템은 MCD19A2 V6과 대기 보정(Multi-Angle Implementation of the Atmospheric Correction, MAIAC) 알고리즘의 다중 각도 구현을 결합하여 1 × 1 Km의 공간 해상도로 레벨 2 AOD를 모니터링할 수 있다. 구름 감지, 에어로졸 검색 및 대기 보정의 정밀도를 향상시키기 위해 MAIAC라는 새로운 최첨단 방법은 시계열 분석과 픽셀 및 이미지 기반 처리의 하이브리드를 사용한다. 먼지 민감성 매핑 시스템은 MODIS 이미지의 파란색 밴드(0.47μm)를 사용하여 에어로졸 농도를 결정하고 20년 동안 평균 농도를 준비할 수 있다(도 3(a)). 고위험 영역은 모델링을 위해 519 지점으로 변환되었으며, 이 중 70%(363점)는 훈련 지점으로, 30%(156점)는 평가 지점으로 사용될 수 있다(도 3(b)). 먼지가 감지된 519개 영역(1인 값) 외에도 모델링을 위해 먼지 확률이 낮은 영역(0인 값)에서 519 지점을 무작위로 선택할 수 있다.

[0040] 도 4는 일 실시예에 있어서, 먼지에 영향을 미치는 요인을 설명하기 위한 도면이다.

[0041] 도 4(a)는 고도, 도 4(b)는 NDVI, 4(c)는 강까지의 거리, 4(d)는 경사, 4(e)는 강우량, 4(f)는 풍속, 4(g)는 벌크 밀도, 4(h)는 모래, 4(i)는 토양 수분 함량, 도 4(j), 도 4(k)는 토양 질감, 4(l)는 토양목, 도 4(m)는 도 4(n)는 암석학, 도 4(o)는 방향을 나타낸다.

[0042] 먼지 민감성 매핑 시스템은 먼지 발생에 영향을 미치는 요인을 수집할 수 있다. 머신러닝 모델에는 종속적(먼지 발생 면적)과 독립적(먼지 발생에 영향을 미치는 기준)을 포함하는 두 가지 유형의 데이터가 필요하다. 먼지 민감성 매핑 시스템은 독립적 데이터를 결정하기 위해 먼지 발생에 대한 15가지 효과적인 기준(요인)을 선택하고 준비할 수 있다. 이러한 기준은 지형(고도, 방향 및 경사), 기후(강우량 및 풍속), 토양 특성(벌크 밀도, 모래, 점토, 토양 수분 함량, 토양 질감, 토양목), 토지 속성(토지 피복, 암석학, 하천까지의 거리, 정규화 식생 지수(NDVI))를 포함하는 네 가지로 분류될 수 있다(도 4). 데이터 처리 및 준비에는 ArcGIS 10.8 소프트웨어와 GEE 플랫폼이 사용될 수 있다. 모델링을 위해 모든 기준의 픽셀 크기는 동일한 것으로 간주될 수 있다. 초기에 WGS 84/UTM 구역 39 N 좌표계를 채택하기 위해 변환된 계층이다. 벡터 계층은 하천으로부터의 거리, 강우량, 풍속, 토양목, 암석학을 포함하여 래스터 형식으로 변환될 수 있다. 픽셀 크기는 토양 기준과 일치하도록 250 × 250 m로 설정될 수 있다. 이후, 모든 기준에 걸쳐 일관성을 보장하기 위해 ArcGIS 10.8 소프트웨어의 resample 명령을 사용하여 데이터를 재샘플링하여 추가 모델링 목적을 위해 250 × 250 m의 균일한 픽셀 크기를 생성할 수 있다.

[0043] ● 지형 기준

[0044] 먼지 민감성 매핑 시스템은 세 가지 지형적 요인, 즉 고도, 경사, 경사면을 고려할 수 있다. 높은 지역에서는 고도가 풍속을 감소시켜 먼지가 확산되는 것을 방지하는 방풍구 역할을 한다. 경사가 낮은 평지와 평지는 급경사지보다 바람 침식과 먼지 생성 가능성이 더 높다. 이들 지역의 바람은 바람 침식 임계값에 빠르게 도달하여 먼지 폭풍이 발생한다. 경사면은 직사광선을 받는 데 영향을 미치며, 이는 토양 건조 및 바람 침식을 유발할



수 있다. 위에서 언급한 세 가지 기준은 SRTM(Shuttle Radar Topography Mission) 이미지에서 추출한 DEM(Digital Elevation Model)에서 획득될 수 있다. 이때, 이미지의 픽셀 크기는 30 X 30 m였으며 GEE 플랫폼을 사용하여 처리될 수 있다.

[0045] ● 기후 기준

[0046] 강우는 식생 성장을 증가시킴으로써 증가된 바람의 영향을 감소시킨다. 또한, 비는 토양 수분을 증가시킴으로써 토양 응집력을 증가시킨다. 먼지 폭풍은 바람의 침식에 의해 발생하며, 침식이 발생할 가능성은 주로 풍속에 의해 결정된다. 먼지 민감성 매핑 시스템은 주요 기준에 대한 지도를 생성하기 위해 Bushehr 주의 10개 시읍면 관측소의 20년 평균 데이터를 사용할 수 있다. 이러한 작업을 수행하기 위해 역 거리 가중치(Inverse Distance Weighting, IDW) 보간 방법과 함께 ArcGIS 10.8 소프트웨어가 사용될 수 있다. 이 과정에서 사용된 픽셀 크기는 30 × 30m이다.

[0047] ● 토양 특성 기준

[0048] 토양의 물리적 특성은 먼지 폭풍의 발생과 먼지 방출의 크기를 조절하는 데 중요한 요인이 될 수 있다. 침식에 취약하고 식생이 없는 토양은 먼지를 생성하는 능력이 더 크다. 실시예에서는 6가지 토양 특성(벌크 밀도, 모래, 점토, 토양 수분 함량, 토양 질감, 토양목)을 고려하기로 한다. 2000년부터 2018년까지 GEE 플랫폼의 미국 농무부(USDA) 데이터를 사용하여 벌크 밀도, 모래, 점토, 토양 수분 함량, 토양 질감 기준이 생성될 수 있다. Bushehr 주 토양 지도는 토양목 지도를 만드는 데 사용될 수 있다. 이 과정에서 사용된 픽셀 크기 250 × 250m로 준비될 수 있다.

[0049] ● 토지 속성 기준

[0050] 토지 피복은 먼지 폭풍과 밀접한 관련이 있다. 식생이 거의 없거나 인간 또는 자연적 요인으로 인해 토양 지각이 파괴된 지역은 먼지 생성의 가능성이 더 높다. GEE 플랫폼 내에서 Sentinel-1 및 Sentinel-2 이미지의 조합을 사용하여 토지 피복 지도가 제작될 수 있다. 토지 피복 분류는 GEE 시스템에서 RF 알고리즘을 사용하여 수행되었으며 91.35%의 전체 정확도(OA)를 달성할 수 있다. 암석학은 토지 형성 및 지표면의 침식성으로 정의된다. 암석학에 민감한 지역은 저항성 지질층으로 구성된 지역보다 먼지 폭풍의 발생률이 더 높다. 이란 지질 지도의 데이터를 사용하여 암석학 지도가 생성될 수 있다(1:100,000 축척). 식생은 표토를 안정화시켜 입자 이동성과 먼지 방출을 억제한다. 식생은 바람에 대한 장벽으로 작용하여 바람에 의한 토양 침식을 방지하고 지표면의 거칠기에도 기여할 수 있다. NDVI는 식생 피복의 계절적 및 연간 변화를 보여줄 수 있는 강력한 지표이다. NDVI 지수는 GEE 플랫폼의 Landsat-8 이미지를 사용하여 작성될 수 있다(수학적 1).

$$NDVI = \frac{Band\ 5 - Band\ 4}{Band\ 5 + Band\ 4} \quad (1)$$

[0051]

[0052] "Band 4"는 Landsat-8 OLI(Operational Land Imager) 센서의 적색 스펙트럼 대역에 해당하는 반면, "Band 5"는 근적외선 스펙트럼 대역을 나타낸다. 건기의 주요 먼지원 중 하나는 계절에 따른 강과 범람원의 바닥이다. 홍수는 장마철에 이 지역을 덮고 이 계절에는 미세한 입자가 있는 물질이 퇴적되기 때문에 건기에는 바람에 노출된다. DEM은 강 계층을 얻기 위한 소스로 사용되었으며, 이후 유클리드 거리를 사용하여 래스터화될 수 있다.

[0053] 먼지 민감성 매핑 시스템은 먼지 발생 데이터에 대해 주파수 비율다중 공선성, (Frequency Ratio FR) 및 요인 중요도를 포함하는 데이터 전처리를 수행할 수 있다.

[0054] 데이터 세트에서 높은 수준의 상관 관계는 다중 공선성의 한 가지 잠재적 원인이며, 이는 머신러닝 모델이 사용하는 변수가 독립적이라고 가정할 때 발생한다. 선형 통계 방법을 사용할 때 데이터의 다중 공선성은 결과의 신뢰성과 해석을 방해한다. VIF(Variance Inflation Factor)은 다중 공선성을 감지하는데 사용되는 지표로, 독립 변수 간에 선형 관계가 있을 때 발생하는 현상이다. 변수 간의 다중 공선성이 오류를 일으키기 쉬운 경우 VIF 값은 10보다 크다.

[0055] FR(Frequency ratio)은 주어진 원인 요인이 있는 지역에서 발생한 사건을 해당 요인이 없는 지역에서 발생한 동일한 사건과 비교하는 것을 포함한다. 계산의 빈도 비율은 사건의 발생을 예측하는데 있어 다양한 원인의 상대적 중요도의 순위를 매기는 데 사용될 수 있다. 요인과 사건 사이의 양의 관련성은 빈도 비율이 1보다 큰 것으로 나타난다. 대조적으로, 이러한 관련성이 없는 경우, 비율이 1보다 작은 것이 특징이다. FR 지수는 수학적 2를 기반으로 계산될 수 있다.

$$FR = \frac{\text{Dust occurrence ratio}}{\text{Class area ratio}}$$

(2)

[0056]

[0057]

랜덤 포레스트(Random forest, RF) 모델은 분류 또는 회귀를 위한 의사 결정 트리 세트와 인기 있는 비모수 앙상블 학습자 알고리즘이다. 변수의 값을 예측하거나 특성화하기 위해, RF 모델은 여러 가지 다른 의사 결정 트리 기술을 사용하는 회귀 접근 방식을 사용한다. RF의 각 의사 결정 트리는 부트스트랩된 버전의 훈련 데이터 세트를 사용하여 구성된다. 반복 분할의 원리는 트리가 루트 노드에서 시작하여 일부 종료 기준에 도달할 때까지 반복적으로 노드를 분할하여 개발된다는 것이다. 예측 도구로서의 강점은 수많은 약한 학습자의 학습을 집계한다는 사실에서 비롯된다. RF는 의사 결정 트리의 포레스트를 사용하여 예측을 수행한다. RF 알고리즘 내에서 의사 결정 트리는 개별적으로 평가된다. 모든 계산 트리에서 과반수 투표 또는 평균을 차지하여 주어진 관찰에 대한 최종 예측을 생성한다. RF 모델은 부트스트랩 샘플을 사용하여 각 트리를 개발하기 때문에 의사 결정 트리를 구축할 때 여러 관찰을 생략한다. 이러한 접근 방식은 구성된 포레스트의 예측 오류를 계산하기 위해 정규 관찰과 동일한 방식으로 OOB(out-of-bag) 샘플을 사용한다.

[0058]

유전 알고리즘(Genetic algorithm, GA)의 아이디어는 유전적 과정과 생물의 자연적 선택에 기반을 두고 있으며, 최적화 문제와 학습 과정을 해결하기 위해 확장되었다. GA 알고리즘은 가장 효과적인 솔루션을 찾기 위해 평가 과정을 모방하여 솔루션 공간을 탐색한다. 이 과정은 솔루션 공간 내에서 초기 솔루션을 생성하고 배포하는 것으로 시작하여 최적의 솔루션을 식별하는 것으로 끝난다. 돌연변이, 교차 및 선택은 GA 알고리즘의 세 가지 주요 연산자이다. 선택 연산자는 특정 문자열이 번식 과정에 참여하는지 결정하려고 하며, GA 알고리즘에서 수렴 속도는 선택 압력에 따라 달라진다. 교차 연산자는 GA 알고리즘의 주요 연산자로 두 염색체에서 동시에 연산하며 두 염색체 특성을 모두 결합하여 자손을 생성한다. 교차 연산자는 선택된 부모 사이에서 문자열의 일부를 교환하며, 여기서 무작위로 자녀의 각 부분은 부모의 해당 부분과 동일한 기회로 선택된다. 돌연변이 연산자는 다른 염색체에서 자발적인 무작위 변경을 생성하고, GA 알고리즘이 로컬 최적화에 갇히는 것을 방지하기 위해 후보 솔루션을 증가시킨다. GA 알고리즘에서 초기 솔루션은 반복 프로세스에 의해 변경되고 정의된 연산자를 적용하여 최상의 솔루션이 획득될 때까지 각 세대에서 최적의 솔루션으로 수렴한다. 반복 프로세스는 미리 정의된 반복 횟수에 도달하거나 솔루션이 더 이상 개선되지 않을 때까지 계속된다.

[0059]

차분 진화(Differential Evolution, DE) 알고리즘은 비선형, 비분화 및 다변량 목적 함수를 처리할 수 있는 모 집단 기반 최적화 기법으로 도입되었다. DE 기법을 연속 공간에서 문제에 대한 솔루션을 효과적으로 최적화할 수 있는 진화 알고리즘이라고 말한다. 또한, 제약 조건을 사용하여 최적화 문제를 효과적으로 해결할 수 있다. DE 알고리즘에서는 모 집단이 무작위로 발생하는 진화 알고리즘과는 다르게 교란된다(예를 들면, GA 알고리즘). DE 알고리즘은 GA 알고리즘과 마찬가지로 교차, 돌연변이 및 선택 연산자로 구성된다. DE 알고리즘은 교차 및 돌연변이를 통해 초기 자손 모집단을 생성한다. 돌연변이 연산자는 검색 공간에서 검색 메커니즘 역할을 한다. 교차 연산자는 성공적인 조합에 대한 정보를 혼합하여 기존 모 집단 구성 요소를 사용하여 테스트 벡터를 구성하고 더 나은 솔루션 공간을 검색할 수 있도록 한다. 선택 단계에서는 부모 세대에 대한 유아 세대를 평가하기 위해 목적 함수를 사용한다. 선택 연산자는 잠재적인 영역을 검색 공간을 탐한다. DE 알고리즘에서는 모든 구성원이 동일하게 부모로 선택될 가능성이 높으며 부모보다 더 나은 성능을 발휘하는 구성원은 다음 단계로 진행된다. 이러한 프로세스는 종료 조건이 충족될 때까지 계속된다.

[0060]

꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 알고리즘은 꽃이 피는 식물의 흐름 수분 과정에서 영감을 얻어 다중 목적 최적화를 위해 개발되었다. FPA 알고리즘은 식물의 수분 움직임을 모방하여 효율적이고 최적의 시스템을 만들고자 한다. 수분은 두 가지 방식으로 발생할 수 있다. 첫 번째는 꽃의 자체 꽃가루가 스스로 수정하는 자가 수분이다. 두 번째는 한 식물에서 다른 식물로 꽃가루 알갱이를 옮기는 것이다. 꽃이 꽃가루를 퍼뜨리는 데 사용하는 방법 중에는 1) 바람이 꽃가루 전달에 필수적인 역할을 하는 비생명체 방법(상호)이 있다. 2) 다른 방법은 박쥐, 새, 곤충 및 기타 동물을 통한 생물학적 수분이다. 생물학적으로 수분의 목적은 꽃 수분, 수, 적자 및 최상의 생존 측면에서 식물의 번식이다. FPA 알고리즘을 개발하기 위해 네 가지 규칙이 명시되어 있다.

[0061]

● 규칙 1: 교차 수분과 생물학적 수분은 모두 보편적인 수분 과정으로 간주됩니다. 꽃가루를 운반하는 꽃가루 매개자는 Levy 비행을 따르고 장거리를 비행한다.

[0062]

● 규칙 2: 로컬 수분은 생물학적 수분과 자가 수분을 통해 발생한다.

[0063]

● 규칙 3: 곤충은 꽃가루 매개자로서 꽃의 지속성을 초래할 수 있다. 이는 번식 가능성이 두 꽃 사이의 유사

성과 관련이 있는 것과 같다.

[0064] ● 규칙 4: 스위치 확률은 로컬 및 글로벌 수분  $p \in [1]$ 의 변화를 제어한다.

[0065] 알고리즘을 단순화하기 위해 각 식물은 하나의 꽃가루 배우자를 생산하는 꽃만 가지고 있다고 가정하고 꽃가루 배우자와 식물, 문제의 솔루션을 구분할 필요가 없다. FPA 알고리즘의 개발은 글로벌 수분과 로컬 수분의 두 가지 주요 단계로 구성될 수 있다.

[0066] 도 5는 일 실시예에 있어서, 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정하는 동작을 설명하기 위한 도면이다.

[0067] 먼지 민감성 매핑 시스템은 기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 생성할 수 있다. 먼지 민감성 매핑 시스템은 세 가지 진화 알고리즘(FPA, GA, DE)을 사용하여 RF 모델의 다섯 가지 하이퍼파라미터( $n\_estimators$ ,  $max\_features$ ,  $max\_depth$ ,  $min\_samples\_split$ ,  $min\_samples\_leaf$ )를 결정하고 개선할 수 있다. 세 가지 결합된 알고리즘의 목적 함수는 정규화된 평균 제곱 오차(Normalized Root Mean Square Error, NRMSE) 지수(수학식 3)를 최소화하는 것이다. NRMSE를 사용하는 장점에는 스케일 불변 특성, 해석 가능성, 상대 비교에 대한 적합성 및 이상치에 대한 복원력이 포함된다.

$$NRMSE = \frac{RMSE}{\text{Mean}(y)} \quad (3)$$

[0068]

[0069] 여기서,  $y$ 는 데이터의 실제 값이며, RMSE는 예측값과 실제값의 차이를 나타내는 모델 성능을 평가하는 표준 통계 측정치이다. RMSE 지수는 수학식 4를 기반으로 계산된다.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - t_i)^2} \quad (4)$$

[0070]

[0071] 여기서,  $y_i$ 와  $t_i$ 는 각각 예측값과 목표값이고,  $n$ 은 총 샘플 수이다.

[0072] 먼지 민감성 매핑 시스템은 DE 알고리즘을 기반으로 특정 지역의 먼지에 영향을 미치는 요인의 중요도를 결정할 수 있다. DE 알고리즘은 다양한 반복에서 목적 함수를 기반으로 가장 중요한 요인을 계산한다. 목적 함수(수학식 5)는 최소화이며, 낮은 값은 더 높은 알고리즘 정확도를 나타낸다.

$$RMSE + w * n \quad (5)$$

[0073]

[0074] 여기서  $n$ 은 요인의 총 개수이고,  $w$ 는 0에서 1 사이의 가중치이다.

[0075] 먼지 민감성 매핑 시스템은 모델링된 먼지 민감성 매핑 모델의 성능을 검증할 수 있다. 먼지 민감성 매핑 시스템은 RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 모델링된 먼지 민감성 매핑 모델의 성능을 검증할 수 있다.

[0076] 먼지 민감성 매핑 시스템은 RMSE(수학식 4)와  $R^2$ 를 이용하여 먼지 민감성 매핑 모델의 효율성을 측정할 수 있다.  $R^2$  지수는 모델의 적합도를 평가하기 위한 단위가 없는 지수이다.  $R^2$  지수의 값은 0과 1 사이이며, 1에 가까울수록 모델의 성능이 향상된다.  $R^2$  지수는 수학식 6을 이용하여 계산될 수 있다.

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \bar{y})^2}{\sum_{i=1}^n (t_i - \bar{t})^2} \quad (6)$$

[0077]

[0078] 여기서,  $n$ 은 총 샘플 수,  $y_i$  및  $t_i$ 는 각각 샘플  $i$ 의 예측값 및 샘플  $i$ 의 목표값이고,  $\bar{y}$  및  $\bar{t}$ 는 예측값의 및

목표값의 평균값이다.

[0079] ROC 곡선은 모델 성능을 평가하기 위해 공간 모델링에 사용되는 그래픽 도구이다. y축의 true positive 비율(감도)과 x축의 false positive(1-특이성)의 두 가지 매개변수를 포함한다(수학식 7 및 수학식 8). ROC 곡선 아래의 면적(AUC)은 모델의 전반적인 성능을 요약하는 통계 지표로 모델의 효과를 정량적으로 평가하는 데 사용된다. AUC 값의 범위는 0.5와 1이며, 값이 높을수록 모델의 예측 능력이 우수함을 나타낸다. AUC 값이 1이면 우수한 모델 성능을 나타낸다. AUC는 수학식 9에 따라 계산될 수 있다.

$$x = 1 - \text{specificity} = 1 - \left[ \frac{\text{TN}}{\text{TN} + \text{FP}} \right] \quad (7)$$

$$y = \text{sensitivity} = \left[ \frac{\text{TP}}{\text{TP} + \text{FN}} \right] \quad (8)$$

$$\text{AUC} = \frac{\sum \text{TP} + \sum \text{TN}}{\text{P} + \text{N}} \quad (9)$$

[0080]

[0081] 모델의 성능을 평가할 때 P는 파지티브(positive) 샘플의 총 수, N은 네거티브(negative) 샘플의 총 수, TP는 참 파지티브(true positive) 샘플의 수, TN은 참 네거티브(true negative) 샘플의 수를 나타낸다.

[0082] 도 6은 일 실시예에 있어서, 먼지 민감성 매핑 시스템을 설명하기 위한 블록도이고, 도 7은 일 실시예에 있어서, 먼지 민감성 매핑 방법을 설명하기 위한 흐름도이다.

[0083] 먼지 민감성 매핑 시스템(100)의 프로세서는 데이터 준비부(610), 모델링부(620), 지도 생성부(630) 및 성능 평가부(640)를 포함할 수 있다. 이러한 프로세서의 구성요소들은 먼지 민감성 매핑 시스템에 저장된 프로그램 코드가 제공하는 제어 명령에 따라 프로세서에 의해 수행되는 서로 다른 기능들(different functions)의 표현들일 수 있다. 프로세서 및 프로세서의 구성요소들은 도 7의 먼지 민감성 매핑 방법이 포함하는 단계들(710 내지 740)을 수행하도록 먼지 민감성 매핑 시스템을 제어할 수 있다. 이때, 프로세서 및 프로세서의 구성요소들은 메모리가 포함하는 운영체제의 코드와 적어도 하나의 프로그램의 코드에 따른 명령(instruction)을 실행하도록 구현될 수 있다.

[0084] 프로세서는 먼지 민감성 매핑 방법을 위한 프로그램의 파일에 저장된 프로그램 코드를 메모리에 로딩할 수 있다. 예를 들면, 먼지 민감성 매핑 시스템에서 프로그램이 실행되면, 프로세서는 운영체제의 제어에 따라 프로그램의 파일로부터 프로그램 코드를 메모리에 로딩하도록 먼지 민감성 매핑 시스템을 제어할 수 있다. 이때, 데이터 준비부(610), 모델링부(620), 지도 생성부(630) 및 성능 평가부(640) 각각은 메모리에 로딩된 프로그램 코드 중 대응하는 부분의 명령을 실행하여 이후 단계들(710 내지 740)을 실행하기 위한 프로세서의 서로 다른 기능적 표현들일 수 있다.

[0085] 단계(710)에서 데이터 준비부(610)는 위성 이미지로부터 감지된 먼지 발생 지점과 복수 개의 발생에 영향을 미치는 요인을 포함하는 먼지 발생 데이터를 준비할 수 있다. 데이터 준비부(610)는 MODIS 이미지를 이용한 먼지 모니터링을 통해 먼지 발생 지점을 감지하고, 먼지 발생에 영향을 미치는 지형, 기후, 토양 특성, 토지 속성을 포함하는 복수 개의 요인을 수집하고, 수집된 복수 개의 요인과 감지된 먼지 발생 지점을 포함하는 먼지 발생 데이터를 준비할 수 있다.

[0086] 단계(720)에서 모델링부(620)는 기계학습과 진화 알고리즘을 앙상블하여 먼지 민감성 매핑 모델을 구축할 수 있다. 모델링부(620)는 먼지 발생 데이터에 대해 주파수 비율(Frequency Ratio FR), 다중 공선성 및 요인 중요도를 포함하는 데이터 전처리를 수행할 수 있다. 모델링부(620)는 랜덤 포레스트(Random forest, RF) 모델에 유전 알고리즘(Genetic algorithm, GA), 차분 진화(Differential Evolution, DE) 알고리즘, 꽃수분 알고리즘(Flower Pollination Algorithm, FPA) 중 어느 하나의 진화 알고리즘을 앙상블함에 따라 랜덤 포레스트 기반의 먼지 민감성 매핑 모델을 구축할 수 있다. 모델링부(620)는 먼지 발생 데이터를 이용하여 구축된 랜덤 포레스트 기반의 먼지 민감성 매핑 모델의 하이퍼파라미터를 조정할 수 있다. 모델링부(620)는 차분 진화 알고리즘을 통해 먼지에 영향을 미치는 요인의 중요도를 계산할 수 있다.



- [0087] 단계(730)에서 지도 생성부(630)는 구축된 먼지 민감성 매핑 모델을 통해 준비된 먼지 발생 데이터를 이용하여 먼지 민감성 지도를 구축할 수 있다.
- [0088] 단계(740)에서 성능 평가부(640)는 구축된 먼지 민감성 지도에 기초하여 구축된 먼지 민감성 매핑 모델의 성능을 검증할 수 있다. 성능 평가부(640)는 RMSE(Root Mean Square Error), 결정 계수( $R^2$ ), ROC(Receiver Operating Characteristic) 곡선 및 통계 지수를 사용하여 구축된 먼지 민감성 매핑 모델의 성능을 검증할 수 있다.
- [0089] 실시예에 따르면, MODIS 이미지(영상)와 공간 모델링을 결합하고, 앙상블 머신러닝 알고리즘을 진화 알고리즘으로 개선하여 먼지 민감성 지도를 생성할 수 있다. 따라서, 연구 지역의 먼지 발생 및 먼지 발생 지점에 영향을 미치는 15가지 요인을 이용하여 4가지 개발 모델(RF, RF-FPA, RF-GA, RFDE)로 먼지 민감성 지도를 작성하고 평가하였다. 실시예에서 제안된 방법의 결과는 다음과 같다. 결과는 낮은 고도와 경사, 낮은 NDVI 값, 하천에 가까운 거리, 낮은 강수량, 높은 풍속, 중간 부피 밀도 값, 높은 모래 값, 낮은 점토 및 토양수량, NW 경사 측면, 양토사토질, 암석학의 Qft2 단위, 토지피복의 점토, 토양순서의 불량지 등이 연구 지역의 먼지 발생에 가장 큰 영향을 미쳤다. 연구 장소의 먼지 발생은 고도, 풍속, 토지 피복 요인에 의해 크게 영향을 받았다. 결과에 따르면, 개발된 먼지 민감성 지도 모델은 높은 정확도(AUC>0.9)를 나타냈다. 또한, 진화적인 알고리즘은 독립형 RF 모델의 정확도를 향상시켰다. RF-DE, RF-GA, RF-FPA, RF 모델은 각각 모델링에서 가장 정확했다. 연구 지역의 서부 및 남서부 지역은 먼지에 더 취약한 것으로 나타났으며, 이는 이러한 지역에 대한 통제 및 완화 조치 구현의 필요성을 강조한다.
- [0090] 이상에서 설명된 장치는 하드웨어 구성요소, 소프트웨어 구성요소, 및/또는 하드웨어 구성요소 및 소프트웨어 구성요소의 조합으로 구현될 수 있다. 예를 들어, 실시예들에서 설명된 장치 및 구성요소는, 예를 들어, 프로세서, 컨트롤러, ALU(arithmetic logic unit), 디지털 신호 프로세서(digital signal processor), 마이크로컴퓨터, FPGA(field programmable gate array), PLU(programmable logic unit), 마이크로프로세서, 또는 명령(instruction)을 실행하고 응답할 수 있는 다른 어떠한 장치와 같이, 하나 이상의 범용 컴퓨터 또는 특수 목적 컴퓨터를 이용하여 구현될 수 있다. 처리 장치는 운영 체제(OS) 및 상기 운영 체제 상에서 수행되는 하나 이상의 소프트웨어 애플리케이션을 수행할 수 있다. 또한, 처리 장치는 소프트웨어의 실행에 응답하여, 데이터를 접근, 저장, 조작, 처리 및 생성할 수도 있다. 이해의 편의를 위하여, 처리 장치는 하나가 사용되는 것으로 설명된 경우도 있지만, 해당 기술분야에서 통상의 지식을 가진 자는, 처리 장치가 복수 개의 처리 요소(processing element) 및/또는 복수 유형의 처리 요소를 포함할 수 있음을 알 수 있다. 예를 들어, 처리 장치는 복수 개의 프로세서 또는 하나의 프로세서 및 하나의 컨트롤러를 포함할 수 있다. 또한, 병렬 프로세서(parallel processor)와 같은, 다른 처리 구성(processing configuration)도 가능하다.
- [0091] 소프트웨어는 컴퓨터 프로그램(computer program), 코드(code), 명령(instruction), 또는 이들 중 하나 이상의 조합을 포함할 수 있으며, 원하는 대로 동작하도록 처리 장치를 구성하거나 독립적으로 또는 결합적으로(collectively) 처리 장치를 명령할 수 있다. 소프트웨어 및/또는 데이터는, 처리 장치에 의하여 해석되거나 처리 장치에 명령 또는 데이터를 제공하기 위하여, 어떤 유형의 기계, 구성요소(component), 물리적 장치, 가상장치(virtual equipment), 컴퓨터 저장 매체 또는 장치에 구체화(embody)될 수 있다. 소프트웨어는 네트워크로 연결된 컴퓨터 시스템 상에 분산되어서, 분산된 방법으로 저장되거나 실행될 수도 있다. 소프트웨어 및 데이터는 하나 이상의 컴퓨터 판독 가능 기록 매체에 저장될 수 있다.
- [0092] 실시예에 따른 방법은 다양한 컴퓨터 수단을 통하여 수행될 수 있는 프로그램 명령 형태로 구현되어 컴퓨터 판독 가능 매체에 기록될 수 있다. 상기 컴퓨터 판독 가능 매체는 프로그램 명령, 데이터 파일, 데이터 구조 등을 단독으로 또는 조합하여 포함할 수 있다. 상기 매체에 기록되는 프로그램 명령은 실시예를 위하여 특별히 설계되고 구성된 것들이거나 컴퓨터 소프트웨어 당업자에게 공지되어 사용 가능한 것일 수도 있다. 컴퓨터 판독 가능 기록 매체의 예에는 하드 디스크, 플로피 디스크 및 자기 테이프와 같은 자기 매체(magnetic media), CD-ROM, DVD와 같은 광기록 매체(optical media), 플롭티컬 디스크(floptical disk)와 같은 자기-광 매체(magneto-optical media), 및 롬(ROM), 램(RAM), 플래시 메모리 등과 같은 프로그램 명령을 저장하고 수행하도록 특별히 구성된 하드웨어 장치가 포함된다. 프로그램 명령의 예에는 컴파일러에 의해 만들어지는 것과 같은 기계어 코드뿐만 아니라 인터프리터 등을 사용해서 컴퓨터에 의해서 실행될 수 있는 고급 언어 코드를 포함한다.
- [0093] 이상과 같이 실시예들이 비록 한정된 실시예와 도면에 의해 설명되었으나, 해당 기술분야에서 통상의 지식을 가진 자라면 상기의 기재로부터 다양한 수정 및 변형이 가능하다. 예를 들어, 설명된 기술들이 설명된 방법과 다

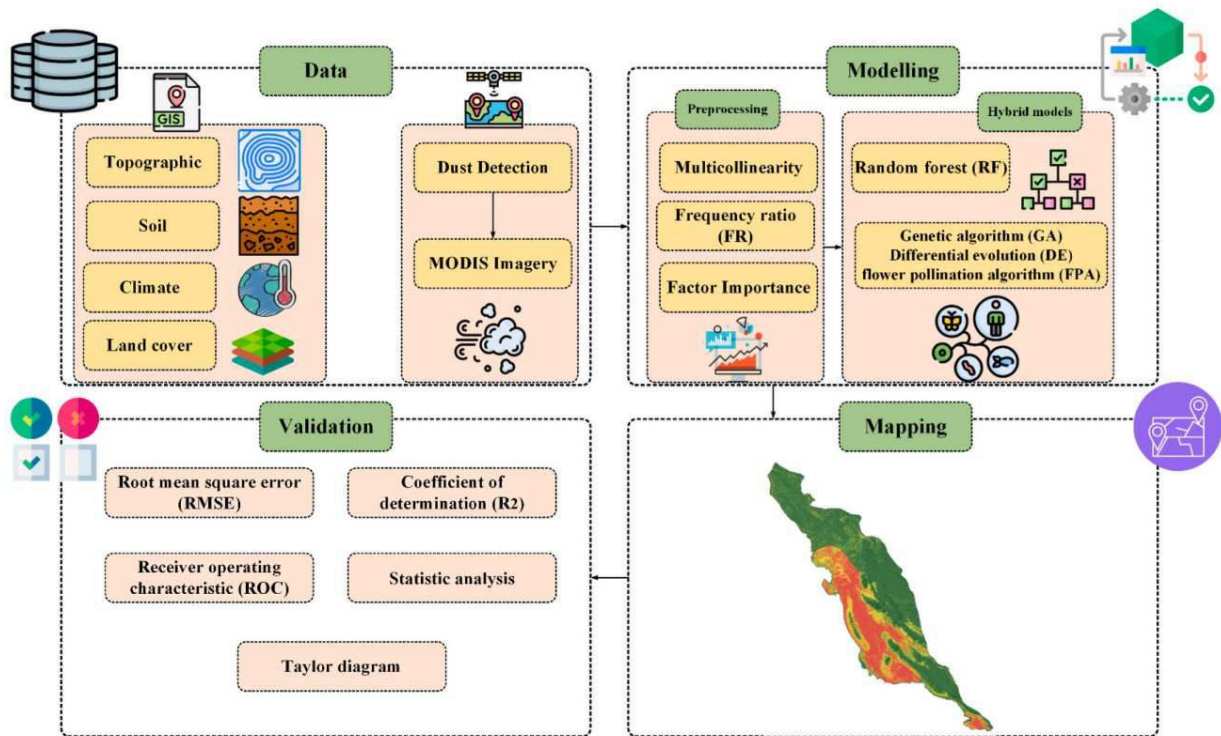
른 순서로 수행되거나, 및/또는 설명된 시스템, 구조, 장치, 회로 등의 구성요소들이 설명된 방법과 다른 형태로 결합 또는 조합되거나, 다른 구성요소 또는 균등물에 의하여 대치되거나 치환되더라도 적절한 결과가 달성될 수 있다.

[0094]

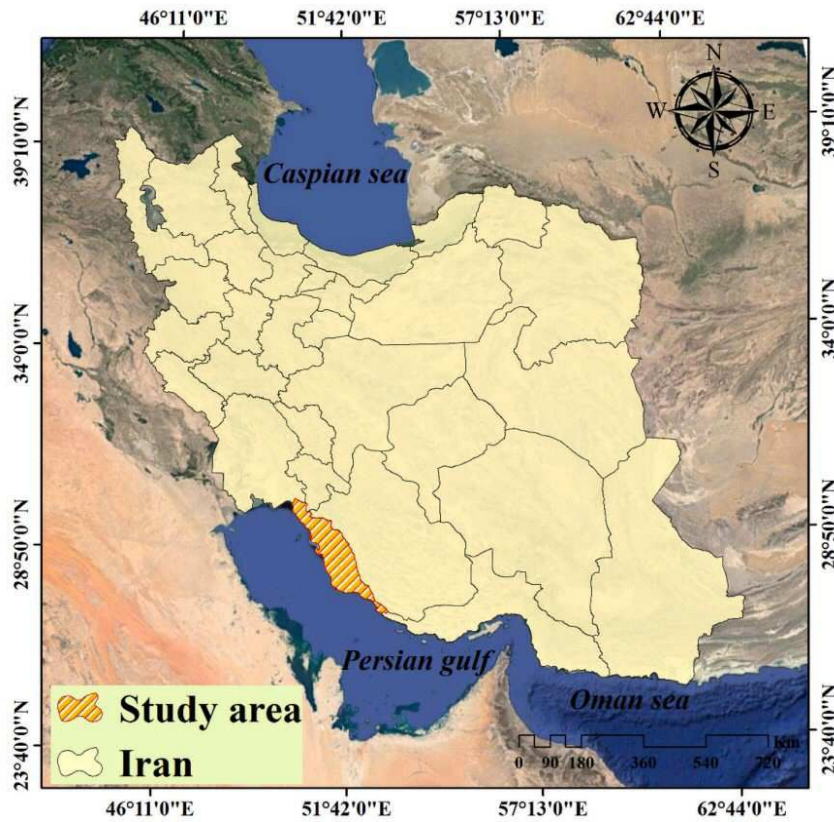
그러므로, 다른 구현들, 다른 실시예들 및 특허청구범위와 균등한 것들도 후술하는 특허청구범위의 범위에 속한다.

## 도면

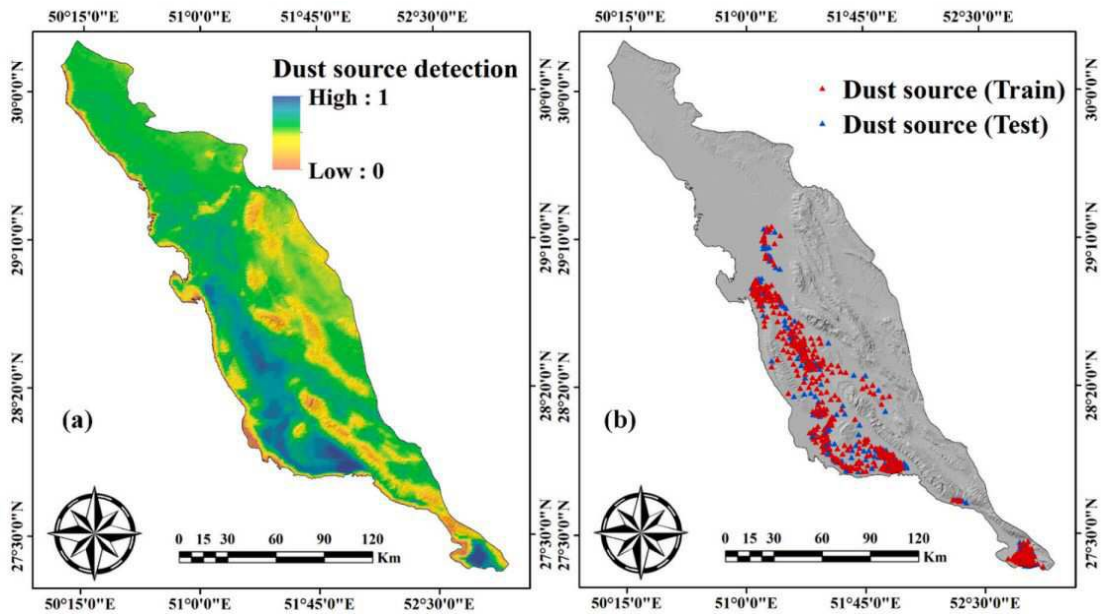
### 도면1



도면2

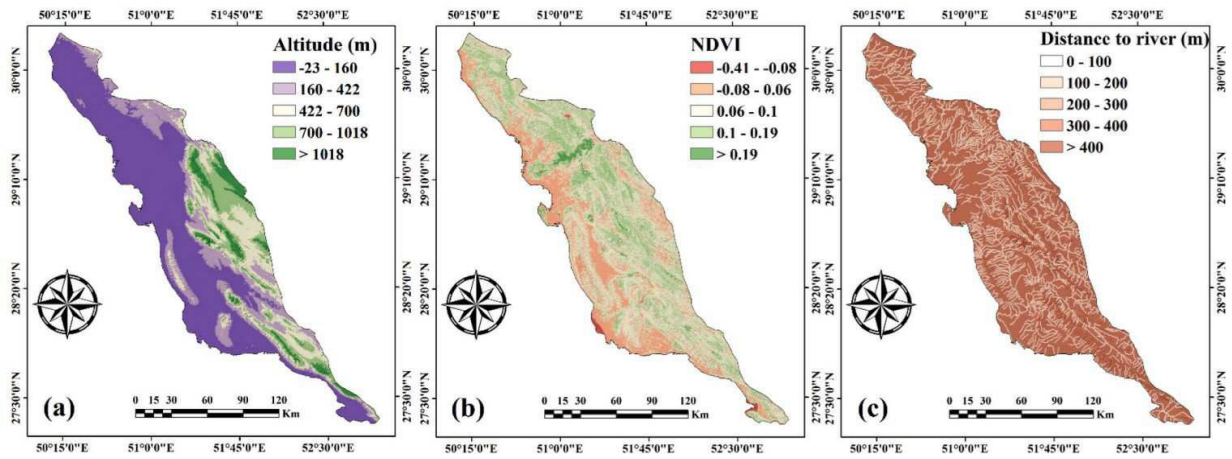


도면3

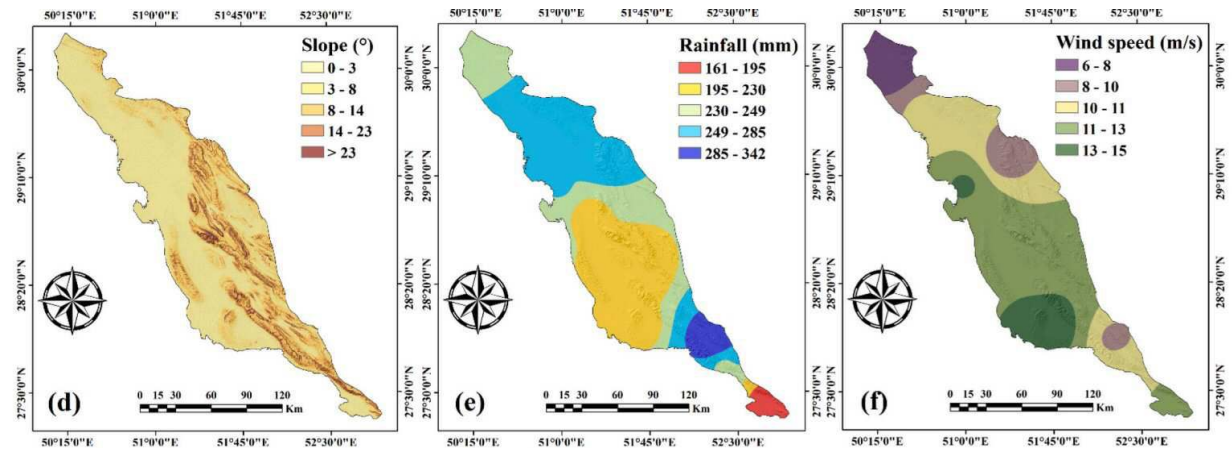




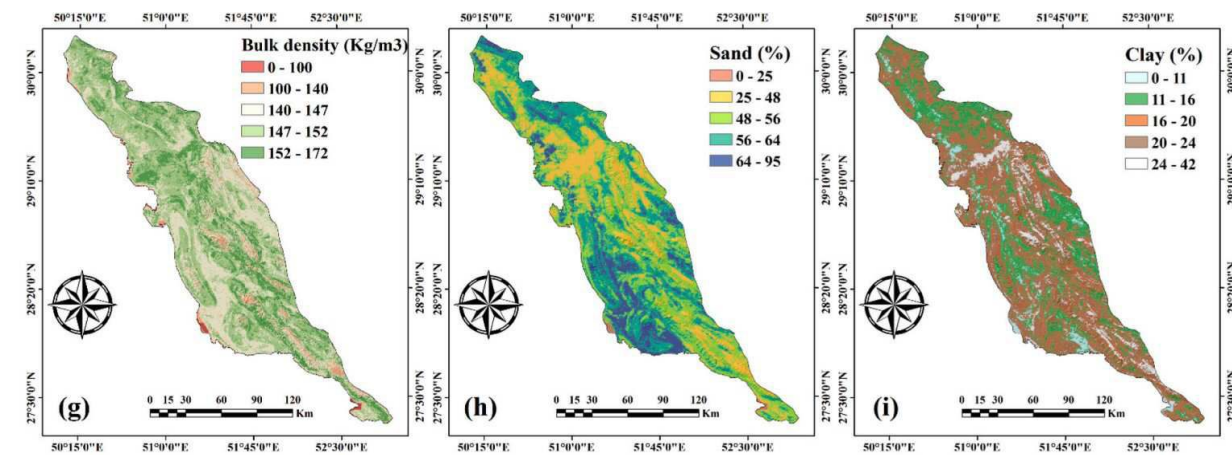
도면4a



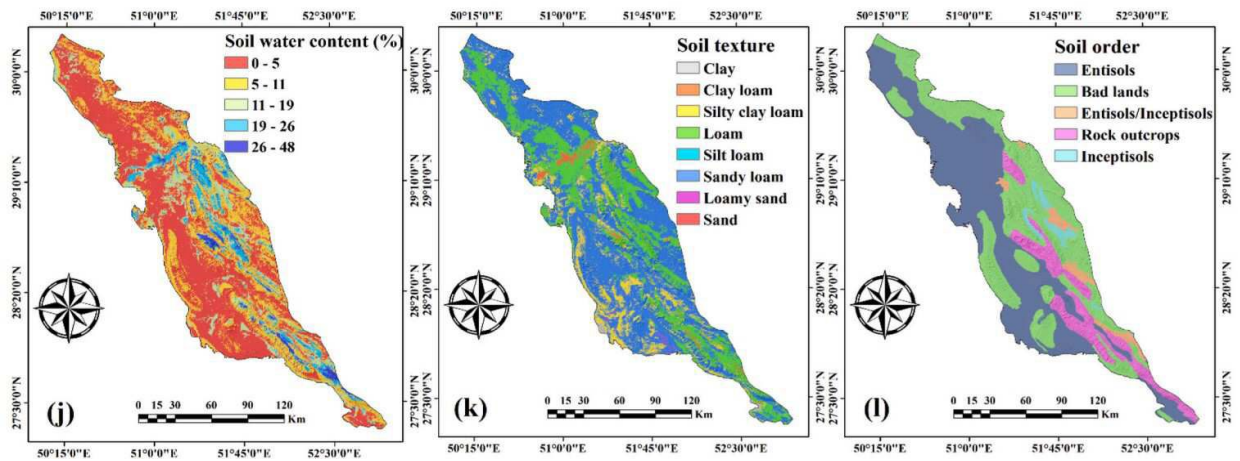
도면4b



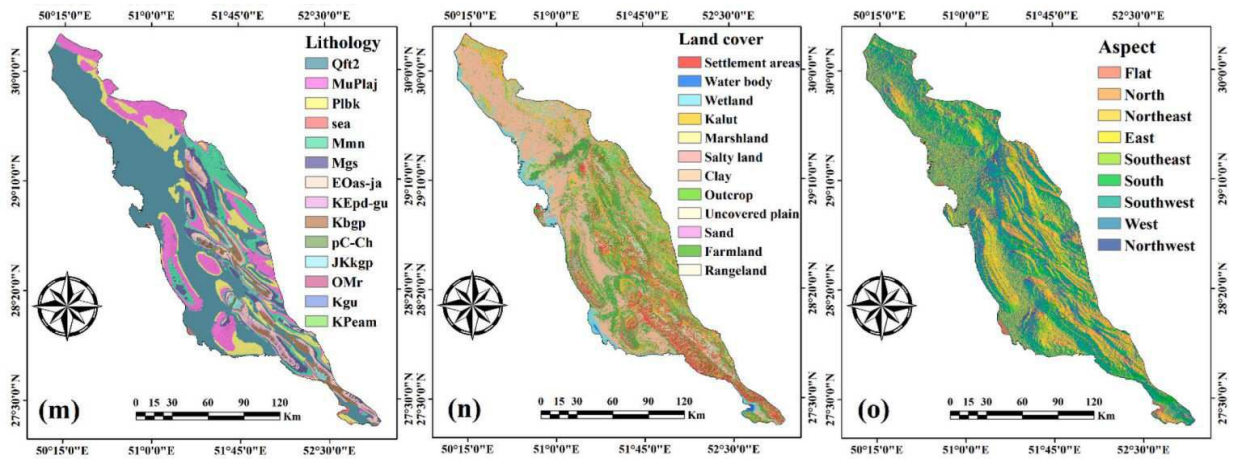
도면4c



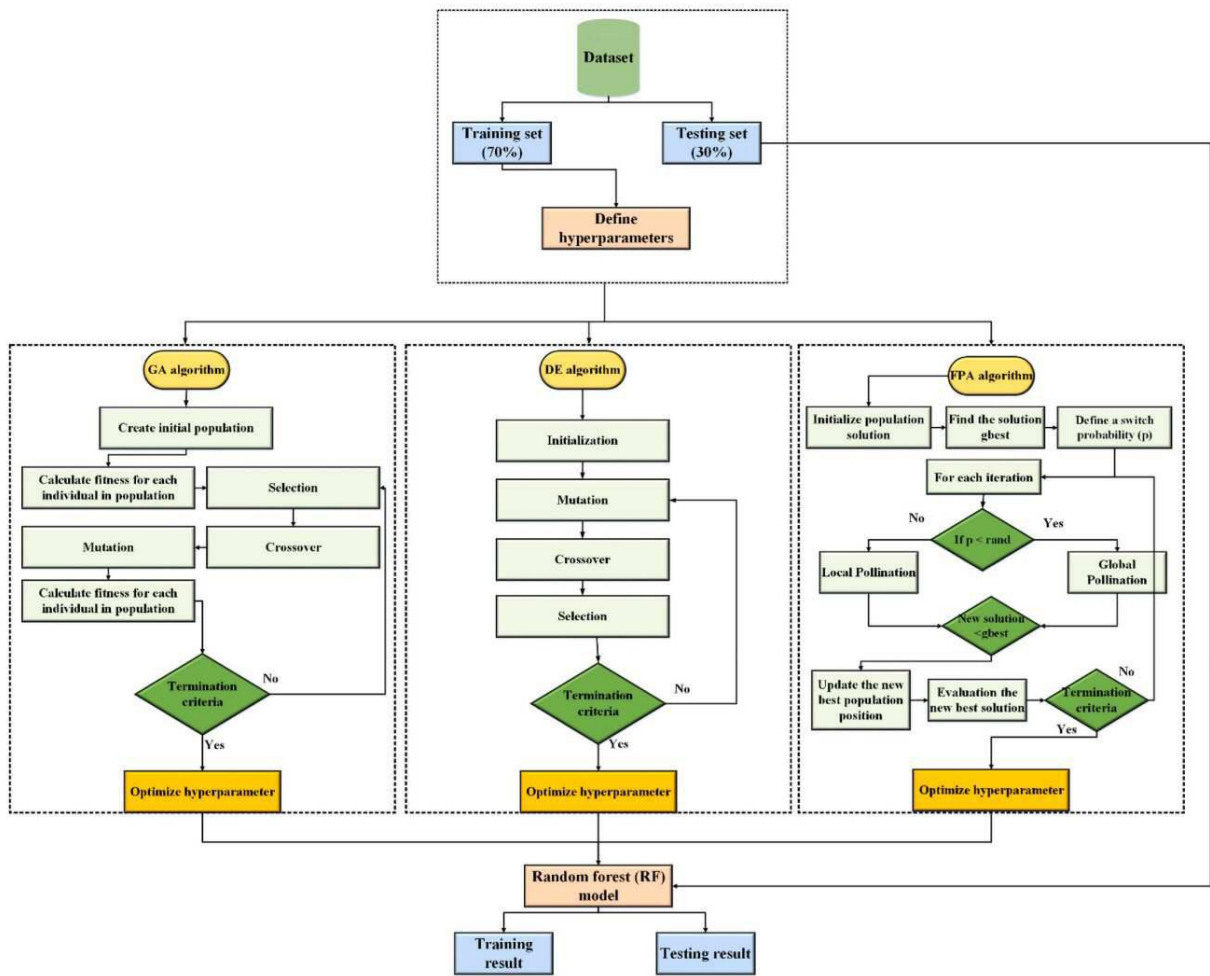
도면4d



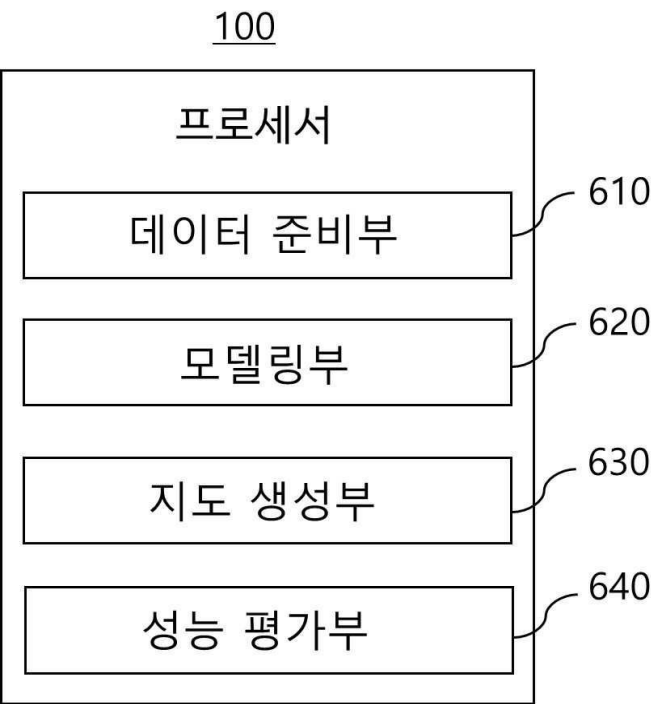
도면4e



도면5



도면6



도면7

