

머신러닝 기반 한반도 기상 변수의 대기 굴절률 영향 분석

Machine Learning-Based Analysis of the Influence of Meteorological Variables on Atmospheric Refractive Index in the Korean Peninsula

안 민* · 손성환** · 조재연** · 남아리** · 박용배*^{***}

Min An* · Sung Hwan Sohn** · Jae Yeon Jo** · A Ri Nam** · Yong Bae Park^{***}

요약

본 논문에서는 한반도 7개 고층 기상 관측소의 2020~2023년 기상 데이터를 기반으로 기압, 기온, 수증기압의 고도별 변화량이 대기 굴절 계수 변화량에 미치는 영향을 분석하였다. 상관 계수 분석과 머신러닝 기법을 활용한 결과, 세 가지 기상 요소 중 수증기압 변화율이 대기 굴절 계수 변화에 가장 큰 영향력을 갖는 것으로 나타났다. 부분 의존도 도표 분석을 통해 세 기상 변수와 대기 굴절 계수 간의 관계를 수치화하였으며, 그 결과 수증기압 변화의 영향력이 기압과 기온 변화에 비해 현저히 크다는 것을 확인하였다. 이러한 결과는 특히 해안이나 산악 지역처럼 수증기압 변동이 큰 환경에서 레이더 운용 시 수증기압의 고도별 분포 특성을 중점적으로 고려할 필요가 있음을 의미한다.

Abstract

In this study, we analyzed the effects of the vertical gradients of pressure, temperature, and water vapor pressure on the atmospheric refractivity gradient using meteorological data from seven upper-air observation stations across the Korean Peninsula from 2020 to 2023. Through correlation coefficient analysis and machine learning techniques, we observed that, among the three meteorological factors, the water vapor pressure gradient had the most significant influence on the refractivity gradient changes. Using the partial dependence plot analysis, we quantified the relationships between the three meteorological variables and the refractivity gradient, confirming that the impact of water vapor pressure variations was substantially greater than that of pressure and temperature changes. These findings indicate that when operating radar systems in environments with significant water vapor fluctuations, such as coastal or mountainous regions, priority should be placed on the vertical distribution characteristics of the water vapor pressure.

Key words: Atmospheric Refractivity, Machine Learning, Radar Wave Propagation, Random Forest, Water Vapor Pressure, XGBoost

I. 서론

대기 중 전자파는 대기의 비균질함으로 인해 굴절되어

진행하는 특성을 갖는다. 이는 특히 대포병 레이더와 같이 표적에 대한 고도 오차가 포탄 궤적 추적에 큰 영향을 미치는 시스템에서 운용 성능 저하를 초래한다. 대기의

「이 논문은 2025년 정부(방위사업청)의 재원으로 국방기술진흥연구소의 지원을 받아 수행된 연구임(KRIT-CT-23-038-01).」

*아주대학교 AI융합네트워크학과(Department of AI Convergence Network, Ajou University)

**LIG넥스원(주) (LIG Nex1)

***아주대학교 전자공학과(Department of Electrical and Computer Engineering, Ajou University)

· Manuscript received June 23, 2025 ; Revised July 8, 2025 ; Accepted August 18, 2025. (ID No. 20250623-062)

· Corresponding Author: Yong Bae Park (e-mail: yong@ajou.ac.kr)

비굴절함은 대기 굴절률(n)으로 표현되며, 이를 이용해 간단한 지수식으로 대기 굴절 계수(N)를 도출한다. 대기 굴절 계수(N)는 기압(P), 기온(T), 수증기압(e) 세 가지 기상 변수에 의해 결정되는데, 장거리 전파 해석에서는 지구 곡률을 반영한 수정 대기 굴절률(m)과 수정 대기 굴절 계수(M)를 사용한다^[1]. 레이더 운용 및 전파 해석에서는 수정 대기 굴절 계수의 고도별 변화량(∇M) 범위에 따라서 전파 굴절 현상이 구분된다. 기존 대기 굴절률에 관한 연구들은 주로 대기 굴절률을 예측 또는 보간하거나^{[2],[3]}, 한국 기상 데이터를 활용한 한국형 지수 모델을 제안하는 등 모델링 관련 연구들이 진행되었다^[4]. 레이더 운용 측면에서는 대기 굴절률에 따른 전파 굴절 현상이 레이더 성능에 미치는 영향을 분석하였다^{[5]~[7]}. 이러한 연구들은 주로 대기 굴절률 자체의 모델링이나 레이더 성능 변화에 초점이 맞춰져 있었다. 그러나, 기압, 기온, 수증기압의 고도별 변화량(∇P , ∇T , ∇e)이 수정 대기 굴절 계수의 변화량(∇M)에 미치는 상대적 기여도와 가장 큰 영향을 미치는 변수에 대한 분석은 이루어지지 않았다. 이에 본 연구는 한반도의 지역적 기상 특성을 반영하여, 기상 변수의 고도별 변화량이 대기 굴절 계수의 변화량에 미치는 영향을 머신러닝을 이용하여 정량적으로 분석하였다. 대기 굴절 계수 수식을 이용한 분석도 가능하지만, 실제 대기 환경에서는 각 기상 변수가 서로 독립적이지 않고 상호 복잡한 영향을 주고받는다^{[8],[9]}. 또한 동서고저의 지형 특성과 삼면이 바다로 둘러싸인 한반도의 특성으로 인한 지역별 기상 차이를 일반화된 수식 기반 분석으로 완벽하게 반영하기 어렵다는 한계가 있다. 이에 한반도 전역의 실측 기상 데이터를 머신러닝 기법으로 분석함으로써 각 기상 변수가 대기 굴절 계수에 미치는 영향력을 직관적으로 비교하고, 어떤 변수가 굴절 계수 변화에 지배적인 역할을 하는지 정량적으로 확인하였다. 이를 통해 해안 지역이나 산악 지역과 같이 특정 기상 변수의 고도별 변화가 두드러지는 지역에서 레이더 운용 시 우선적으로 고려해야 할 기상 요소를 제시하고, 더 정확한 전파 해석을 위한 기초적인 분석 자료를 제공하는 것을 목표로 한다.

본 연구에서는 한반도 7개 고층 기상 관측소(백령도, 흑산도, 국가태풍센터(제주), 포항, 북강릉, 오산, 광주)에

서 2020~2023년 동안 하루 2~4회 관측을 통해 수집된 28,281개의 기상 데이터를 분석하였다. 각 데이터는 고도에 따른 기압, 기온, 이슬점 온도, 상대 습도 등으로 구성된다. 기상 변수의 고도별 변화량과 대기 굴절률 변화량 간의 상관관계를 분석하고, 머신러닝 모델인 random forest와 XGBoost을 활용하여 각 변수의 상대적 중요도를 평가하였다. 또한, 부분 의존도 도표(partial dependence plot)를 통해 random forest 모델에서 각 기상 변수의 기여도를 정량적으로 분석하였다.

II. 한반도 기상 데이터 및 수정 대기 굴절 계수

본 연구에서는 한반도 기상 변수의 대기 굴절률 영향 분석을 위해 University of Wyoming에서 제공하는 고층 기상 관측 데이터를 활용하였다^[10]. 국내 고층 기상 관측소는 그림 1과 같이 백령도, 흑산도, 국가태풍센터(제주), 포항, 북강릉, 오산, 광주 7개 지역에 위치하며, 하루 2~4회(00시, 06시, 12시, 18시) 약 100 m에서 30,000 m까지의 고도별 기상 데이터를 수집한다. 본 연구에서는 2020년부터 2023년까지 28,281개 데이터 파일(총 2,714,976행)을 분석하였다. 각 데이터 파일은 기압(P), 기온(T), 수증기압(e) 등으로 구성되며, 500~10,000 m 고도 범위를 100 m 간격으로 선형 보간하여 약 96행으로 표준화하였다. 관측소,

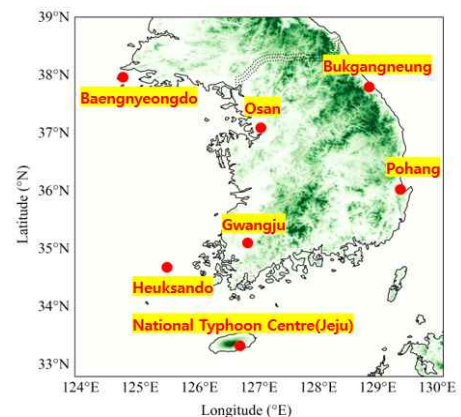


그림 1. 한반도 고층 기상 관측소

Fig. 1. Upper air radiosonde observation stations in the Korean Peninsula.

날짜, 시간에 따라 측정 고도가 달라질 수 있으므로, 최소 측정 고도를 확인하고 선형 보간을 적용해 일관된 고도 레벨(500~10,000 m, 100 m 간격)로 데이터를 정규화하였다. 분석 고도를 10,000 m 이내로 제한한 이유는 기압, 기온, 수증기압 등의 기상 변수가 이 범위에서 가장 활발히 변화하며, 그 이상에서는 상대적으로 일정한 경향을 보이기 때문이다. 또한, 군사용 레이더의 일반적인 운용 고도가 이 범위 내에 있어 실제 운용 환경에 적합한 분석을 수행하기 위함이다. 본 연구는 기상 변수의 고도별 변화량이 대기 굴절률에 미치는 영향을 분석하는 데 초점을 맞추었으며, 모든 측정 시간대의 데이터를 활용하였다.

University of Wyoming에서 제공하는 고층 기상 관측 데이터는 표 1과 같이 고도별 기압, 기온, 이슬점 온도, 상대 습도 등을 포함하고 있다. 본 연구에서는 대기 굴절률(n) 대신 수정 굴절 계수(M)을 사용하였다. 이는 대기 굴절률(n)은 거의 1에 가까운 값으로 해석 시 불편함이 있기 때문이며, 통상 식 (1)과 같이 대기 굴절률(n)에 1을 빼고 10⁶을 곱한 대기 굴절 계수(N)와 지구 곡률을 고려한 수정 대기 굴절 계수(M)을 주로 사용한다. 대기 굴절 계수(N)를 계산하기 위해서는 이들 중 기압(P), 기온(T), 그리고 이슬점 온도로부터 계산되는 수증기압(e)이 필요하다. 수증기압(e)은 다음 식 (2)와 같이 계산할 수 있다^[1].

표 1. 고층 기상 관측 데이터 예시(2024.07.24. 12시)
Table. 1. Example of upper air observation data (2024.07.24. 12 Z).

Height (m)	Pressure (hpa)	Temperature (°C)	Dewpoint temperature (°C)	Relative humidity (%)
158	1,000.0	23.8	23.3	97
382	970.0	25.0	25.0	100
458	960.0	24.8	23.7	94
707	928.0	24.2	19.5	75
731	925.0	24.0	19.4	75
1,078	889.0	22.3	18.8	81
⋮				
32,569	8.5	-41.3	-82.1	0
32,839	8.2	-40.0	-81.5	0

$$N = (n - 1) \times 10^6 = 77.6 \frac{P}{T} - 5.6 \frac{e}{T} + 3.75 \times 10^5 \frac{e}{T^2} \quad (1)$$

P: 기압[hpa], T: 기온[°K], e: 수증기압[hpa]

$$e = 6.112 \times \exp\left(\frac{17.67 T_d}{T_d + 243.5}\right) \quad (2)$$

T_d: 이슬점 온도[°C]

$$M = N + \frac{x}{a_e} \times 10^6 = N + 157x \quad (3)$$

a_e: 지구반지름 6,378[km], x: 고도[km]

계산된 수증기압과 기압, 기온을 이용하여 대기 굴절 계수(N)를 구하고, 이에 수정 대기 굴절 계수(M) 및 수정 대기 굴절 계수의 고도별 변화율(∇M)을 식 (3)을 통해 계산할 수 있다^[1]. 이러한 방법으로 총 28,281개, 2,714,976행의 기상 데이터를 처리하여 기압(P), 기온(T), 수증기압(e)을 고도에 대하여 미분한 변화율(∇P, ∇T, ∇e)과 수정 대기 굴절 계수의 고도별 변화율(∇M)을 산출하였다.

III. 기상 변수의 대기 굴절 계수 영향 분석

산출된 각 변수의 고도별 변화량 값과 수정 굴절 계수의 고도별 변화량 값을 이용하여 각 변수 간 피어슨 상관 계수를 구하였다. 그림 2는 변수 간 상관 계수를 나타낸 상관 행렬로, 피어슨 상관 계수는 변수 간 선형 관계의 강

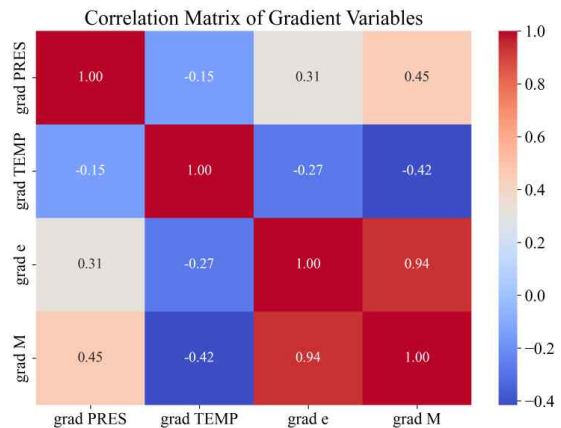


그림 2. 기상 변수 변화율 간 상관 계수 행렬
Fig. 2. Correlation matrix of gradient variables.

도를 나타낸 것이다^[11]. 즉, 변수 간 상관 관계를 직접적으로 나타낸 것으로, 각 변수의 고도별 변화율인 $\nabla P(\text{grad } P)$, $\nabla T(\text{grad } T)$, $\nabla e(\text{grad } e)$ 와 $\nabla M(\text{grad } M)$ 의 관계를 분석한 것이다.

그림 2에서 확인할 수 있듯이, 수증기압의 고도별 변화율(∇e)과 수정 대기 굴절 계수의 고도별 변화율(∇M) 사이에는 매우 강한 양의 상관관계($r=0.94$)가 존재한다. 이는 수증기압의 고도별 변화가 대기 굴절 계수의 고도별 변화를 설명하는 데 가장 중요한 요소임을 의미한다. 반면, 기압의 고도별 변화율(∇P)과 ∇M 사이에는 중간 정도의 양의 상관관계($r=0.45$)가 관찰되었다. 이는 기압의 변화가 대기 굴절 계수의 변화에 일정 부분 기여하지만, 그 영향력이 수증기압에 비해 상대적으로 제한적임을 나타낸다. 기온의 고도별 변화율(∇T)과 ∇M 사이에는 ∇P 와 유사한 정도의 음의 상관관계($r=-0.42$)가 존재하는데, 이는 기온이 증가하는 방향으로의 변화가 대기 굴절 계수를 감소시키는 방향으로 작용함을 의미한다.

이러한 상관관계의 물리적 의미를 더 명확히 하기 위해, 고도에 따른 기상 변수 변화율의 패턴을 분석하였다. 그림 3은 전체 데이터에 대해 고도에 따른 정규화된 변화율(∇M , ∇P , ∇T , ∇e) 평균값을 비교한 것이다. 모든 변화율은 0에서 1 사이의 값으로 정규화되어 있어, 상대적인 크기를 직접 비교할 수 있다.

저고도에서는 $\nabla e(\text{grad } e, \text{파란색 실선}), \nabla P(\text{grad } P, \text{초$

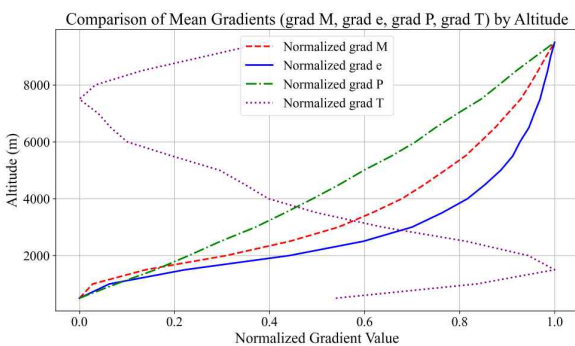


그림 3. 정규화된 고도에 따른 기상 변수 및 굴절 계수 변화율

Fig. 3. Normalized altitude profiles of meteorological variable and refractive index gradients.

록색 점선), $\nabla M(\text{grad } M, \text{빨간색 점선})$ 이 유사한 경향을 보이지만, 고도가 증가함에 따라 $\nabla e(\text{grad } e)$ 와 $\nabla M(\text{grad } M)$ 의 곡선이 점점 더 유사해지는 반면, $\nabla P(\text{grad } P)$ 의 영향은 상대적으로 감소한다. 반면, $\nabla T(\text{grad } T, \text{보라색 점선})$ 의 경우 고도에 따라 변화 패턴이 복잡하게 나타나지만, 전반적으로 ∇M 과는 음의 상관관계를 보이며 그 영향력은 제한적이다. 이는 앞선 상관관계 분석 결과와 일치한다. 그러나 상관관계 분석은 두 변수씩 쌍으로 비교하는 방식으로, 다른 변수들의 영향을 통제된 상태에서 분석하기 때문에 대기 환경에서 모든 변수가 동시에 상호 작용하는 실제 조건을 완전히 반영하지 못하는 한계가 있다. 특히 기압, 기온, 수증기압은 서로 밀접한 관계를 가지며 변화하는 특성을 갖기 때문에, 단순히 두 변수 간의 선형 상관 관계만으로는 이들이 대기 굴절 계수(M)에 미치는 영향을 정확히 평가하기 어렵다. 이러한 한계를 보완하기 위해 본 연구에서는 머신러닝 기법인 random forest와 XGBoost 알고리즘을 활용하였다. 이 알고리즘들은 여러 변수들이 동시에 존재하는 상황에서 각 변수의 상대적 중요도를 평가하는 데 활용될 수 있어 본 연구에서 복잡한 상호관계를 가진 기상 변수들이 수정 대기 굴절 계수 변화율에 미치는 영향을 분석하는 데 적용하였다. 이러한 모델들은 대기/기상 데이터 분석에서 복잡한 변수 상호작용과 중요도 추출에 강점이 있어 널리 사용된다^{[12]-[14]}. 모델 학습 시 하이퍼파라미터는 라이브러리 기본 세팅을 따랐으며, 별도의 튜닝 없이 높은 정확도를 확보하였다. 이는 수정 대기 굴절 계수(M) 자체가 기압, 기온, 수증기압으로부터 계산되고 이를 고도에 대하여 미분한 것이기 때문에, 이러한 기상 변수들의 변화율(∇P , ∇T , ∇e)들로 ∇M 을 예측할 때 모델의 정확도가 매우 높게 나타난다. 그러나, 본 연구의 목적은 이러한 높은 정확도의 예측 모델에서 각 기상 변수의 변화율이 굴절 계수 변화에 미치는 상대적 기여도를 정량적으로 평가하는 데 있다.

Random forest는 여러 개의 결정 트리로 구성된 앙상블 학습 방법으로, 복잡한 데이터 패턴을 효과적으로 식별할 수 있다^[15]. 이 모델은 다수의 결정 트리를 구성하는 과정에서 각 노드에서 데이터를 분할할 때 가장 유용한 변수를 선택한다. 이 과정에서 모델은 수많은 분기점에서 어

변 변수가 목표값(∇M)을 예측하는 데 가장 효과적으로 기여하는지 학습하게 된다. 각 변수가 노드 분할에 사용된 빈도와 예측 정확도 향상 정도를 측정함으로써 변수의 특성 중요도를 산출할 수 있다^[16]. 본 연구에서는 100개의 결정 트리를 사용하여 모델을 구성하였다. ∇M (grad M)을 예측하기 위해 ∇P (grad PRES), ∇T (grad TEMP), ∇e (grad e)를 입력 변수로 사용하였다. 전체 데이터셋의 80%를 훈련에 사용하고 20%를 테스트에 활용하여 높은 정확도($R^2=0.995$, $RMSE=1.376$)를 가지는 모델을 구축하였다. 그림 4는 이 random forest 모델의 각 기상 변수의 특성 중요도를 표로 나타낸 것이다. ∇e (grad e)의 특성 중요도가 0.8852로 매우 큰 비중을 차지했다. 이는 모델이 대기 굴절 계수의 변화를 예측하기 위한 결정 트리를 구성할 때, 약 89%의 경우에서 수증기압의 변화율을 주요 분할 기준으로 선택했음을 의미한다. 반면, ∇P (grad PRES)와 ∇T (grad TEMP)의 특성 중요도는 각각 0.0924와 0.0225로 상대적으로 매우 낮게 나타났다.

XGBoost는 random forest와 달리 이전 트리의 오차를 순차적으로 보완하는 부스팅(boosting) 방식을 사용한다. 이 알고리즘은 각 트리가 이전 트리들의 잔차(residual)를 학습하도록 설계되어 있으며, 손실 함수의 그래디언트(gradient)를 최소화하는 방향으로 모델을 최적화한다^{[17],[18]}. XGBoost에서 특성 중요도는 각 특성이 모델의 손실 함수를 감소시키는 데 기여한 정도를 기반으로 계산

된다. 본 연구에서도 random forest와 동일하게 100개의 트리를 사용하여 높은 정확도($R^2=0.968$, $RMSE=3.604$)를 가지는 모델을 구축하였다.

그림 5는 이 XGBoost 모델의 각 기상 변수의 특성 중요도를 표로 나타낸 것으로, random forest 모델과 마찬가지로 ∇e (grad e)의 중요도가 0.9188로 매우 큰 비중을 차지했다. 이는 수증기압의 고도별 변화가 XGBoost 모델에서 대기 굴절 계수의 변화를 예측하는 데 약 92%의 기여를 한다는 것을 의미한다. ∇P (grad PRES)와 ∇T (grad TEMP) 중요도는 각각 0.0476과 0.0336으로, random forest와 유사하게 낮은 값을 보였다. XGBoost가 random forest와 다른 학습 알고리즘을 가짐에도 불구하고 두 모델에서 도출된 특성 중요도가 매우 유사하다는 점은 ∇e (grad e)가 ∇M (grad M)에 가장 큰 영향을 미친다는 분석 결과의 신뢰성을 높인다.

각 기상 변수의 상대적 중요도 분석에서 확인된 변수 중요도를 더 정량적으로 탐색하기 위해, random forest 모델에서 PDP(partial dependence plot)를 생성하여 각 기상 변수가 예측 결과인 ∇M 에 미치는 영향을 정량적으로 분석하였다. 부분 의존도는 다른 모든 특성의 영향을 평균화했을 때, 특정 특성이 모델의 예측에 미치는 한계 효과(marginal effect)를 나타낸다^[18]. 이 방법은 특정 변수의 다양한 값 범위에 대해, 다른 모든 변수는 원래 값을 유지한 채 모델 예측을 수행하고, 각 예측값의 평균을 계산하

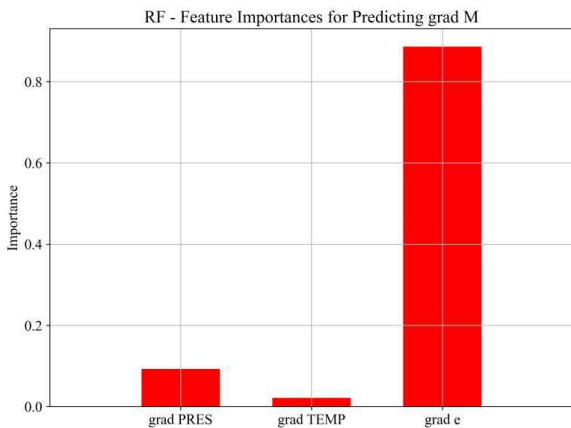


그림 4. Random Forest 특성 중요도
Fig. 4. Feature importance using random forest.

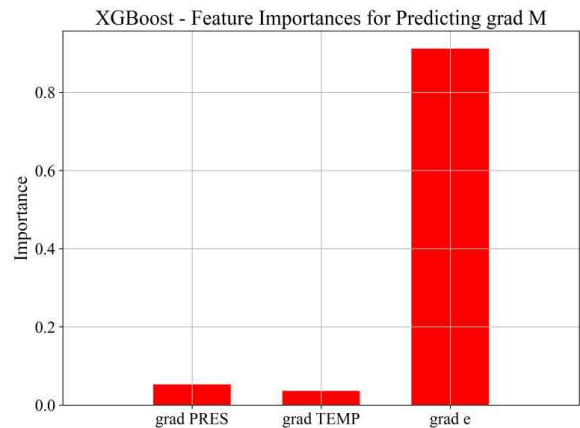


그림 5. XGBOOST 기법 중요도
Fig. 5. Feature importance using XGBOOST.

여 해당 변수가 목표 변수에 미치는 영향을 추정한다. 이러한 접근법은 각 기상 변수의 단위가 서로 다름에도 불구하고 이들이 굴절률 변화율에 미치는 영향을 객관적으로 비교할 수 있게 해준다.

그림 6은 $\nabla e(\text{grad } e)$ 에 대한 부분 의존도 도표이다. x축은 수증기압 변화율의 값, y축은 해당 값에서의 부분 의존도를 나타낸다. 여기서 부분 의존도는 분석 대상 변수(∇e)가 특정 값일 때, 다른 조건들($\nabla P, \nabla T$)은 데이터 셋에 존재하는 다양한 값들을 그대로 유지한 상태에서 모델이 예측하는 ∇M 의 평균값이다. 이 분석을 통해 ∇e 만의 독립적인 영향을 확인할 수 있다. 수증기압 변화율(∇e) 값이 $-7 \sim 0 \text{ hPa/km}$ 로 변화할 때, 부분 의존도는 약 $103 \sim 145 \text{ M-unit/km}$ 까지 증가하는 것을 볼 수 있다. 이는 ∇e 가 1 hPa/km 단위 증가할 때 ∇M 이 평균적으로 약 6 M-unit/km 증가함을 의미한다. 이러한 결과는 수증기압의 고도별 변화율이 증가할수록(즉, 수증기압이 고도에 따라 덜 감소할수록) 대기 굴절 계수의 고도별 변화율이 증가함을 보여주며, 전체 변화율 범위에서 거의 선형에 가까운 관계를 유지하고 있다.

그림 7의 $\nabla P(\text{grad PRES})$ 에 대한 부분 의존도 도표는 ∇e 보다 완만한 양의 기울기를 보인다. 기압 변화율(∇P) 값이 $-110 \sim -40 \text{ hPa/km}$ 로 변화할 때, 부분 의존도의 값은 약 $126 \sim 138 \text{ M-unit/km}$ 까지 증가한다. 이는 ∇P 가 10

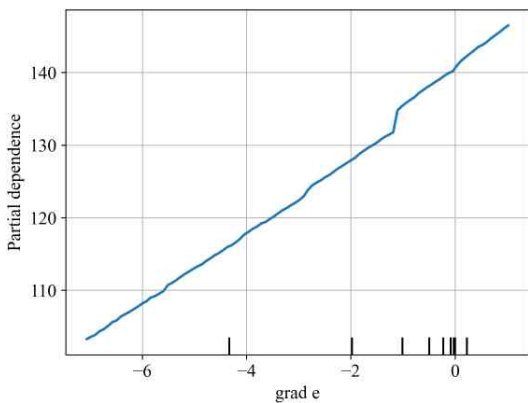


그림 6. 수증기압 고도별 변화율의 부분 의존도 도표
Fig. 6. Partial dependence plot of water vapor pressure gradient.

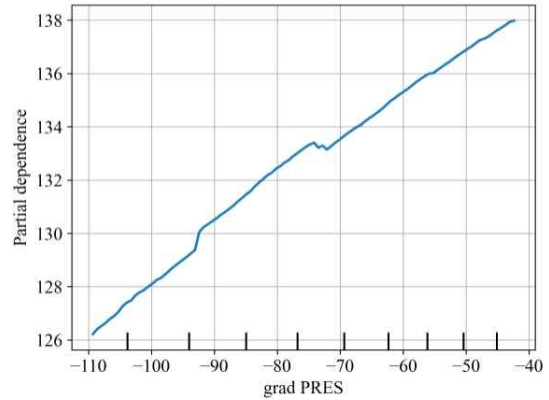


그림 7. 기압 고도별 변화율의 부분 의존도 도표
Fig. 7. Partial dependence plot of pressure gradient.

hPa/km 단위 증가할 때 ∇M 이 평균적으로 약 1.7 M-unit/km 증가함을 의미한다.

그림 8의 $\nabla T(\text{grad TEMP})$ 에 대한 부분 의존도 도표는 뚜렷한 음의 기울기를 보인다. 기온 변화율(∇T) 값이 $-9 \sim 0 \text{ }^\circ\text{C/km}$ 로 변화할 때, 부분 의존도 값은 약 $136 \sim 129 \text{ M-unit/km}$ 까지 감소한다. 이는 ∇T 가 $1 \text{ }^\circ\text{C/km}$ 단위 증가할 때 ∇M 이 평균적으로 약 0.8 M-unit/km 감소함을 의미한다. 세 기상 변수의 부분 의존도 도표 분석 결과를 종합적으로 비교하면, 표 2와 같다.

수증기압의 고도별 변화율이 수정 대기 굴절 계수의 변화에 미치는 영향이 기압 변화율 및 기온 변화율의 절

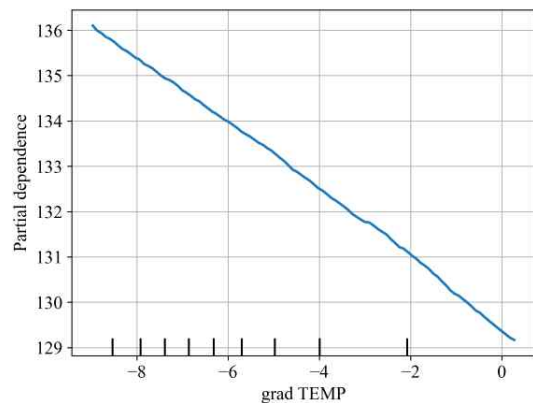


그림 8. 기온 고도별 변화율의 부분 의존도 도표
Fig. 8. Partial dependence plot of temperature gradient.

표 2. 기상 변수의 고도별 변화가 대기 굴절 계수 변화량에 미치는 영향

Table. 2. Effects of Vertical changes in meteorological variables on atmospheric refractivity gradient.

Variable	Unit change	Effects on ∇M
∇P	10 hPa/km	1.7 M-unit/km
∇T	1 °C/km	-0.8 M-unit/km
∇e	1 hPa/km	6 M-unit/km

댓값보다 더 크다는 것을 확인할 수 있다. 또한, 수증기압 변화에 의한 부분 의존도 값의 변동 범위(약 42 M-unit/km)가 기압 변화율(약 12 M-unit/km)과 기온 변화율(약 7 M-unit/km)에 비해 3~6배 더 넓게 나타나는 것은, 대기 굴절 계수의 변화가 수증기압 변화에 가장 민감하게 반응함을 의미한다. 이러한 PDP 분석 결과는 앞서 random forest와 XGBoost 모델에서 도출된 특성 중요도 결과와 일관된다. 두 모델 모두에서 grad e의 중요도가 약 90 %로 높게 나타났는데, 이는 PDP 분석에서 확인된 수증기압 변화율의 강한 영향력과 일치한다.

IV. 결 론

본 연구에서는 한반도 7개 고층 기상 관측소(백령도, 흑산도, 국가태풍센터(제주), 포항, 북강릉, 오산, 광주)에서 수집한 2020~2023년 동안의 기상 데이터(28,281일치, 2,714,976행)를 활용하여 기압, 기온, 수증기압의 고도별 변화량이 대기 굴절 계수 변화량에 미치는 영향을 머신러닝 기법을 통해 정량적으로 평가하였다. 사용된 학습 데이터는 기압(P), 기온(T), 수증기압(e)의 고도별 변화율(∇P , ∇T , ∇e)을 입력 feature로, ∇M 을 목표 변수로 구성되었으며, 이는 고도 500~10,000 m 범위의 관측값을 기반으로 한다. 상관관계 분석 결과, 수증기압 변화율과 수정 대기 굴절 계수 변화율 사이에 매우 강한 양의 상관관계($r=0.94$)가 확인되었으며, 기압 변화율과는 중간 정도의 양의 상관관계($r=0.45$), 기온 변화율과는 중간 정도의 음의 상관관계($r=-0.42$)가 관찰되었다. 머신러닝 모델 분석 결과, random forest에서는 수증기압 변화율의 중요도가 88.52 %로 매우 높게 나타났고, 기압 변화율은 9.24 %,

기온 변화율은 2.25 %에 불과했다. XGBoost 모델에서도 유사하게 수증기압 변화율이 91.88 %, 기압 변화율은 4.76 %, 기온 변화율은 3.36 %의 중요도를 보였다. 각 기상 변수의 영향을 정량적으로 분석하기 위해서 부분 의존도 도표(PDP)를 생성하고 이를 분석한 결과, 수증기압 변화율이 1 hPa/km 증가할 때 대기 굴절 계수 변화율이 약 6 M-unit/km 증가하는 것으로 나타났다. 반면, 기압 변화율은 10 hPa/km 증가할 때 약 1.7 M-unit/km 증가, 기온 변화율은 1°C/km 증가할 때 약 0.8 M-unit/km 감소하는 것으로 나타났다.

이러한 결과는 한반도, 특히 해안 지역이나 산악 지역과 같이 수증기압의 변화가 크고 불규칙한 지역에서 레이다 운용 시 대기 굴절 현상으로 인한 오차를 최소화하기 위해서는 수증기압의 고도별 분포와 변화를 우선적으로 고려한 보정이 필요함을 의미한다.

References

- [1] *The Radio Refractive Index: Its Formula and Refractivity Data*, ITU-R P.453-14, Aug. 2019.
- [2] J. Yang, J. Heo, J. H. Kim, Y. J. Park, H. Choo, and Y. B. Park, "Predicting atmospheric refractivity over the Korean Peninsula using deep learning techniques," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 34, no. 6, pp. 493-496, Jun. 2023.
- [3] J. Yang, J. Heo, J. Kim, Y. Park, H. Choo, and Y. B. Park, "Deep neural network-based time series atmospheric refractivity prediction model using meteorological observation data," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 34, no. 11, pp. 860-863, Nov. 2023.
- [4] M. H. Park, H. S. Kim, W. J. Jeon, S. K. You, S. W. Kwon, and H. W. Moon, et al., "Development of modeling & simulation tool for long range radar considering operational environment in time domain," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 30, no. 7, pp. 591-602,

- Jul. 2019.
- [5] M. H. Park, W. J. Jeon, H. S. Kim, S. W. Kwon, H. W. Moon, and K. W. Lee, "Radar performance variation according to the change of atmospheric refractive index," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 32, no. 8, pp. 743-750, Aug. 2021.
- [6] M. An, J. Heo, N. W. Choi, S. S. Kim, Y. S. Han, and Y. B. Park, "Analysis of radar detection range based on parabolic equation method considering atmospheric refractivity and actual terrain," *The Journal of Korean Institute of Electromagnetic Engineering and Science*, vol. 36, no. 3, pp. 258-261, Mar. 2025.
- [7] M. Kim, Y. Kim, S. Park, S. Cho, C. Hong, and H. Kim, "Error analysis of radar target measurement due to atmospheric refractivity," *The Journal of The Institute of Internet, Broadcasting and Communication*, vol. 24, no. 6, pp. 57-62, Dec. 2024.
- [8] V. Isaac, W. A. van Wijngaarden, "Surface water vapor pressure and temperature trends in North America during 1948-2010," *Journal of Climate*, vol. 25, pp. 3599-3609, May 2012.
- [9] E. Lu, "Understanding the effects of atmospheric circulation in the relationships between water vapor and temperature through theoretical analyses," *Geophysical Research Letters*, vol. 34, no. 14, Jul. 2007.
- [10] University of Wyoming Department of Atmospheric Science, "Upper air sounding data." Available: <http://weather.uwyo.edu/upperair/sounding.html>
- [11] P. Schober, C. Boer, and L. A. Schwarte, "Correlation coefficients: Appropriate use and interpretation," *Anesthesia & Analgesia*, vol. 126, no. 5, pp. 1763-1768, May 2018.
- [12] Y. Mai, Z. Sheng, H. Shi, and Q. Liao, "Using improved XGBoost algorithm to obtain modified atmospheric refractive index," *International Journal of Antennas and Propagation*, vol. 2021, no. 1, p. 5506599, Sep. 2021.
- [13] C. Yang, Y. Wang, A. Zhang, H. Fan, and L. Guo, "A random forest algorithm combined with Bayesian optimization for atmospheric duct estimation," *Remote Sensing*, vol. 15, no. 17, 4296, Aug. 2023.
- [14] J. A. Segovia, J. F. Toaquiza, J. R. Llanos, and D. R. Rivas, "Meteorological variables forecasting system using machine learning and open-source software," *Electronics*, vol. 12, no. 4, p. 1007, Feb. 2023.
- [15] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5-32, Oct. 2001.
- [16] T. T. Nguyen, J. Z. Huang, and T. T. Nguyen, "Unbiased feature selection in learning random forests for high-dimensional data," *The Scientific World Journal*, vol. 2015, no. 1, p. 471371, Mar. 2015.
- [17] J. H. Friedman, "Greedy function approximation: A gradient boosting machine," *The Annals of Statistics*, vol. 29, no. 5, pp. 1189-1232, Oct. 2001.
- [18] T. Chen, C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD '16)*, San Francisco, CA, Aug. 2016, pp. 785-794.

안 민 [아주대학교/석사과정]

<https://orcid.org/0009-0008-4302-5044>



2024년 2월: 아주대학교 전자공학과 (공학사)
2024년 3월~현재: 아주대학교 AI융합네트워크학과 석사과정
[주 관심분야] 장거리 Propagation 해석

남 아 리 [LIG넥스원/연구원]

<https://orcid.org/0009-0003-7292-7147>



2022년 2월: 숭실대학교 전자공학부 (공학사)
2024년 2월: 숭실대학교 지능형반도체공학 (공학석사)
2024년 1월~현재: LIG넥스원 레이더연구소 연구원
[주 관심분야] 레이더 신호처리, 레이더

성능분석

손 성 환 [LIG넥스원/수석연구원]

<https://orcid.org/0000-0002-9401-7182>



2004년 2월: 인하대학교 전자공학과 (공학사)
2006년 2월: 인하대학교 정보통신대학 (공학석사)
2010년 8월: 인하대학교 정보통신대학 (공학박사)
2011년 1월~현재: LIG넥스원 레이더연구소 수석연구원

[주 관심분야] 레이더 신호처리, 레이더 통제, 레이더 성능분석

박 용 배 [아주대학교/교수]

<https://orcid.org/0000-0002-7095-4614>



1998년 2월: 한국과학기술원 전기 및 전자공학과 (공학사)
2000년 2월: 한국과학기술원 전기 및 전자공학과 (공학석사)
2003년 2월: 한국과학기술원 전기 및 전자공학과 (공학박사)
2003년 2월~2006년 8월: KT인프라연구소 선임연구원

2006년 9월~현재: 아주대학교 전자공학과 중점 AI융합네트워크학과 교수

[주 관심분야] 전자파 Propagation 해석, 레이더, 레이돔, AI

조 재 연 [LIG넥스원/선임연구원]

<https://orcid.org/0009-0007-5664-9721>



2023년 8월: 한양대학교 전자공학부 (공학사)
2023년 5월~현재: LIG넥스원 레이더연구소 선임연구원
[주 관심분야] 레이더 시스템 등