

Extrapolative neural operator for SOFAR-channel transmission loss

Jaewan Lee¹, Junghwan Kook^{1*}

¹ Gyeongnam Aerospace & Defense Institute of Science and Technology, Gyeongsang National University ,501 Jinju-daero, Jinju, Gyeongnam, Republic of Korea

Submitted to **Journal of the Acoustic Society of America**

December 20, 2025

*Corresponding author: jkook@gnu.ac.kr

Tel. +82 055-772-1631

Abstract

심해 SOFAR(Sound Fixing and Ranging) 채널에서의 전달손실(Transmission Loss, TL)은 장거리 수중 음향 시스템의 탐지 거리와 음영 영역을 결정하는 핵심 지표이다. 그러나 포물선 방정식 및 정상 모드 해석기는 높은 정확도에도 불구하고, 주파수, 음원 깊이, 환경 조건의 광범위한 변화를 대상으로 매개변수 연구를 수행하거나 운용 중 의사결정을 지원하기에는 계산 비용이 크다. 이러한 연산 비용의 한계를 완화하기 위해 BELLHOP 음선 궤적을 입력으로 고정 격자 TL 맵을 예측하는 합성곱 신경망(Convolutional Neural Network, CNN) 모델들이 도입되었으나, 이미지 회귀(image-to-image regression) 구조에 둑여 격자 해상도와 도메인에 종속되며, 학습 범위를 벗어난 외삽(extrapolation) 조건에서의 성능과 물리 일관성은 충분히 검증되지 않았다.

이러한 격자 종속성과 외삽 한계를 완화하기 위해, 본 연구에서는 SOFAR 채널 TL 예측을 환경 변수에서 공간적 TL 필드로의 연속적인 매핑 연산자(operator) 학습 문제로 정식화하고, 잔차형 DeepONet(R-DeepONet)에 기반한 신경 연산자(neural operator)를 제안한다. 제안된 모델은 BELLHOP로부터 얻은 음선 밀도(ray-density)와 주파수, 음원 깊이를 합성곱 브랜치 네트워크로 환경 임베딩으로 압축하고, 이를 좌표 기반 트렁크 네트워크와 결합하여 임의의 범위–수심 좌표에서 TL 값을 계산한다. 학습 시에는 TL 값 오차에 대해 상반성(reciprocity)과 공간적 매끄러움(smoothness)을 반영한 손실을 함께 최소화하였다. 제안 기법의 정량적 성능을 검증하기 위해 ray 기반 CNN 기준 모델과 비교한 결과, 제안된 R-DeepONet은 분포 내(in-distribution) 조건에서 유사한 TL 오차를 유지하면서, 학습 범위를 벗어난 외삽 조건에서는 수렴 및 음영 구역 등 핵심 음향 구조를 상대적으로 더 잘 보존하는 경향을 보였다. 또한 물리 일관성 손실을 포함한 학습은 예측 TL 필드의 상반성 위반과 비물리적 미소 스케일 잡음을 감소시켰으며, 이를 통해 신경 연산자 기반 TL 대리모델이 향후 연속 학습 및 불확실성 정량화 기법이 탑재될 수 있는 확장적 TL 백본(deterministic TL backbone)으로 활용될 수 있음을 보여준다.

Keywords: SOFAR channel, Transmission loss prediction, Neural operator learning, Extrapolation

1. Introduction

SOFAR(Sound Fixing and Ranging) 채널은 심해에서 음속(sound speed)이 깊이에 따라 변하면서 음파가 반복적으로 굴절(refraction)되어 장거리 전파가 유지되는 도파관(waveguide)이다.^[1] SOFAR 채널에서 전달손실(Transmission Loss, TL)은 특정 위치의 음압 레벨이 기준 대비 얼마나 감소했는지를 dB로 나타낸 값이며, 장거리 수중 음향 탐지와 통신 시스템에서 탐지 가능 범위와 수렴 구역(convergence zone), 음영 구역(shadow zone)의 형성을 판단하는 핵심 물리량이다.^[2] TL은 KRAKEN,^[3] SCOOTER와 같은 파동 기반 수치 해석기(wave based numerical solver)로 계산하며, BELLHOP과 같은 음선 추적(ray tracing)은 전파 경로 정보를 산출한다.^[4] 이러한 파동 기반 수치 해석기는 높은 신뢰도의 TL을 제공하지만, 주파수, 음원 수심, 음속 구조(Sound Speed Profile, SSP), 경계 조건이 바뀔 때마다 해석을 다시 수행해야 한다.^[5] 이 반복 해석은 매개변수 연구와 설계 검토처럼 조건을 대량으로 바꿔가며 TL을 평가하는 작업, 그리고 제한된 시간 안에 의사결정을 내려야 하는 운용 시나리오에서 계산 비용을 상승시키고, 병목현상을 유발한다. 이러한 계산비용 상승 및 병목현상을 제거하기 위해 파동 기반 해석기의 TL을 학습 기반 모델로 근사하여 추론 단계에서 빠르게 산출하려는 접근법이 등장했다.^[6]

Wang et al.^[7]은 BELLHOP 음선 정보를 입력으로 사용하고 파동 기반 해석기가 생성한 TL 맵을 출력으로 예측하는 CNN 기반 TL 예측 모델을 보고했으며, Deo et al.^[8]은 범위 존(range-dependent) 환경에서 조건부 CNN과 연속 학습(continual learning)을 결합하여 환경 변화에 대한 적응을 다루었다. 그러나 두 연구 모두 TL 예측을 고정 격자 이미지 회귀(image-to-image regression)로 구성한다. 이 구성은 출력이 특정 격자와 도메인 크기에 고정된다는 제약을 만들고, 임의의 범위–수심 좌표에서 TL을 직접 계산하는 기능과 해상도 또는 도메인 크기를 바꿔 적용하는 기능을 구조적으로 제공하지 못한다. 또한 학습 범위를 벗어난 주파수, 음원 수심, SSP 또는 환경 조건에서는 TL 회랑(corridor)과 수렴, 음영 구

역 같은 핵심 구조의 보존이 불안정해질 수 있으며, 연속 학습 접근도 추가 학습과 가중치 업데이트를 전제로 하는 경우가 많다. 따라서 추가 학습 없이 외삽 조건에서 구조 보존을 확보하면서 임의 좌표에서 TL을 계산 가능한 형태로 제공하는 모델링 틀이 필요하다.

본 연구는 이러한 공백을 메우기 위해 SOFAR 채널 TL 예측을 연산자 학습(operator learning)^[9] 문제로 다루고, 잔차형 DeepONet(Residual DeepONet, R-DeepONet)에 기반한 신경 연산자(neural operator)를 제안한다. R 브랜치 네트워크(branch network)는 음선 밀도(ray density) 맵과 주파수, 음원 수심을 입력으로 받아 환경의 잠재 표현(latent representation)을 생성하고, 트렁크 네트워크(trunk network)는 범위, 수심 좌표를 입력으로 받아 좌표 기저 함수를 생성한다. 모델은 두 출력을 결합하여 임의 좌표에서 TL 값을 계산 하므로 동일 환경에 대해 해상도에 독립적인 TL 계산을 제공한다. 학습은 TL 오차만 줄이지 않고, 상반성(reciprocity)^[10]과 공간적 매끄러움(spatial smoothness)을 반영하는 물리 일관성 손실(physics consistent loss)을 함께 사용하여 비상반성 패턴과 미소 스케일 노이즈 같은 비물리적 아티팩트를 억제하고 핵심 음향 구조의 안정성을 강화한다. 검증은 동일한 음선 기반 입력을 사용하는 고정 격자 CNN 기준 모델과의 비교로 수행하며, 분포 내 조건에서의 정확도와 함께 학습 범위를 벗어난 주파수, 음원 수심, SSP 교란 조건에서의 구조 보존, 물리 일관성, 계산 시간 측면에서 결과를 제시한다.

2. Problem formulation and data generation

II.A. Physical setting and TL solvers for the SOFAR channel

본 연구는 심해 SOFAR 채널에서의 전달손실(Transmission Loss, TL) 예측 문제를 2차원 범위-수심(range-depth) 평면에서 정의하고, 범위에 대해 균질한(range-independent) 원통 대칭 환경을 가정하였다. 좌표계는 원통 좌표계 (r, z) 를 사용하며, r 은 음원으로부터의 수평 거리, z 는 해수면으로부터의 수심을 나타낸다. 해석 도메인은 $r \in [0, 100]$ km, $z \in [0, 5000]$ m로 고정하였고, 이 도메인을 각각 256점으로 균일 샘플링하여 256×256 격자 위에서 음장과 TL 필드를 산출하

였다. (Fig 1.)

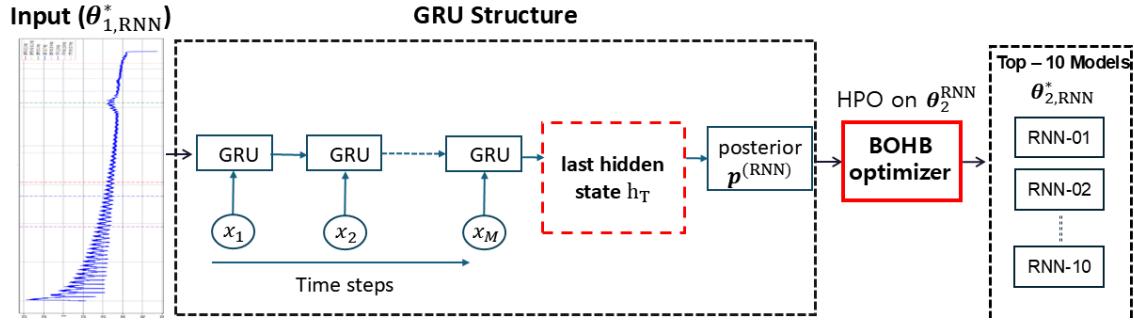


Figure 1. GRU Branch Architecture—Input $\theta_{1,RNN}^*$ → GRU stack → last hidden state h_T → posterior $p^{(RNN)} \in \mathbb{R}^c$ → BOHB (HPO) → Top–10 models.

심해 음속 구조는 Munk canonical 음속 프로파일로 설정하였다.^[11] 본 연구는 축심 수심(axis depth) z_a , 기준음속 c_0 , 무차원 프로파일 파라미터 ε 를 사용하여 수심 방향 음속 $c(z)$ 를 다음과 같이 정의하였다(Eq. (1)).

$$\eta = \frac{2(z - z_a)}{z_a}, \quad c(z) = c_0[1 + \varepsilon(e^{-\eta} + \eta - 1)] \quad (1)$$

해수면 경계는 압력 해방(pressure-release) 조건으로 두었고, 해저는 유체 반공간(acoustic half-space)으로 모델링하였다. 해저 매질은 압축파 속도 c_p , 전단파 속도 c_s , 밀도 ρ , 압축파 감쇠 α_p 로 기술하였다. 또한 장거리 전파에서 수치적 과감쇠로 인한 TL 포화(saturation)를 방지하기 위해 체적 흡수(volume attenuation)는 적용하지 않았다. 본 절에서 사용한 물리 도메인, SSP, 경계 조건 및 해석기 설정의 수치 파라미터는 (Table 1)에 요약하였다.

Table. 1 Numerical setup and dataset sweep for the deep-ocean SOFAR benchmark

Block	Settings (numerical values)
Domain & Grid	$r \in [0,100] \text{ km}$, $z \in [0,5000] \text{ m}$, Uniform 256×256 sampling
SSP(Munk)	$z_a = 1234.5 \text{ m}$, $c_0 = 1548.52 \text{ m/s}$, $\varepsilon = 0.00737$
Boundaries & Bottom	Surface: pressure-release; Bottom: acoustic half-space($c_p = 1600 \text{ m/s}$, $c_s = 0$, $\rho = 1.8 \text{ g/cm}^3$, $\alpha_p = 0.8 \text{ dB}/\lambda$), volume attenuation off

Solver roles	BELLHOP (ray tracing): input signature KRAKEN (+FIELD): TL ground truth
BELLHOP beams	$\theta \in [-30^\circ, 30^\circ]$ (BELLHOP beam launch angle), $N_{beam} = 361$
Sweep & Cases	$f = 20 - 10000 \text{ Hz}$ (log-spaced, 250 points), z_s (source depth): 20 values(50-4000 m) total 5000 cases

TL 데이터 생성은 “입력 서명(input signature)”과 “정답 TL(ground truth)”의 역할을 분리하여 수행하였다. 먼저 BELLHOP 음선 추적기를 사용해 동일 환경에서 음선 궤적을 계산하고, 이 결과로부터 학습 입력이 되는 음선 기반 서명(ray signature)을 생성하였다. BELLHOP 계산에서 는 발사 각 범위를 -30° 에서 30° 로 두고 총 361개의 빔을 발사하여 음선 궤적을 산출하였다. 다음으로 파동 기반 해석기(wave-based solver)인 KRAKEN을 사용해 동일한 물리 환경에서 복소 압력장 $p(r, z)$ 를 계산하고, 이를 TL 정답으로 사용하였다. KRAKEN 출력은 FIELD 후처리를 통해 지정된 256×256 격자에서의 압력장을 얻었으며, TL은 Acoustics Toolbox 관례에 따라 unit source로 정규화된 복소 압력장의 크기를 사용하여 (Eq. (2))로 정의하였다.

$$TL(r, z) = -20 \log_{10} |p(r, z)| \text{ [dB]} \quad (2)$$

데이터셋은 주파수와 음원 수심을 체계적으로 스윕하여 구성하였다. 주파수는 20–10000 Hz 범위에서 255점을 로그 간격으로 샘플링하였고, 음원 수심은 50–4000 m 구간에서 20개 이상 값으로 고정하였다. 따라서 총 $250 \times 20 = 5000$ 개의 환경–주파수 조건에 대해 BELLHOP 기반 입력 서명과 KRAKEN 기반 TL 정답을 동일 셋업에서 생성하였다.

II.B. Ray-density and TL representations for learning

본 연구는 II.A의 동일 물리 셋업에서 “입력 서명(input signature)”과 “정답 TL(ground truth)”을 짹(pair)으로 구성하였다. 입력 X 는 BELLHOP이 산출한 음선 궤적을 256×256 range-depth 격자에 누적하여 얻은 ray-density map이며, 주어진 SSP 및 경계 조건이 규정하는 전파 경로의 기하학적 서명(geometric descriptor)로서 파동 기반 TL 필드 예측의 입력으로 사용된다. 정답 Y 는 동

일 환경에서 KRAKEN(+FIELD)이 계산한 복소 압력장 $p(r, z)$ 로부터 (Eq. (2))의 정의에 따라 계산한 TL 필드이다. 또한 학습 안정성을 위해 ray-density map은 격자 내 누적 ray count를 최대값으로 나누어 $X \in [0,1]$ 로 정규화하였고, TL 타깃은 TL을 [40,120] dB로 클리핑한 뒤 선형 정규화를 적용하여 $Y \in [0,1]$ 로 구성하였다. 추론 시 출력의 역정규화 및 보고 범위는 하단 (Eq. (3))에 따른다. 각 샘플은 (X, Y) 와 함께 주파수 f , 음원 수심 z_s 를 메타데이터로 포함하는 HDF5 레코드로 저장하여, 동일 격자에서 입력 서명과 TL 정답의 대응을 유지하였다. 입력 ray-density map과 이에 대응하는 TL map의 예시는 Fig. 2에 제시하였다.

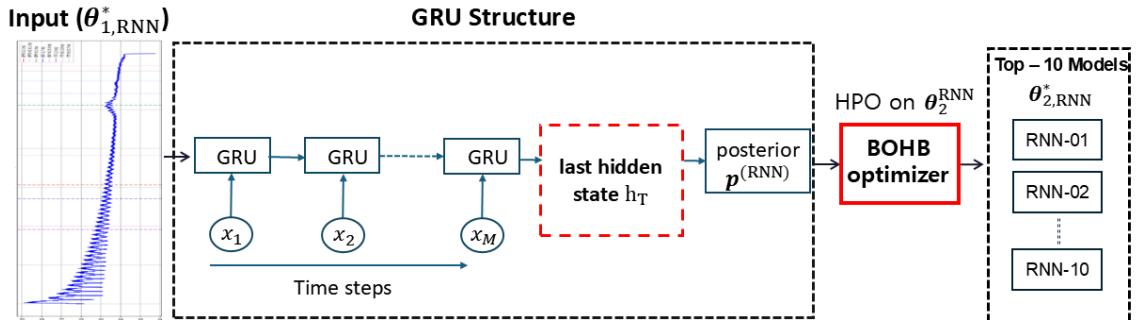


Figure 2. (left) X : ray-density map, (right) TL or Y : corresponding TL map.

$$Y = \frac{\text{clip}(TL, 40, 120) - 40}{80}, \quad TL_{dB} = \text{clip}(80\hat{Y} + 40, 40, 120) \quad (3)$$

II.C. Operator learning versus fixed-grid CNN surrogates

앞서 정의한 TL 예측 문제를 본 절에서는 오퍼레이터 학습 관점에서 재정식화한다. 오퍼레이터 관점에서 목표는 ray-density 입력 X 와 조건 (f, z_s) 가 주어졌을 때 범위-수심 평면의 TL 함수를 산출하는 연산자 \mathcal{G} 를 학습하는 것이다(Eq. (4)).

$$\mathcal{G} : (X, f, z_s) \mapsto \text{TL}(\cdot, \cdot), \quad \text{TL}(r, z) = \mathcal{G}(X, f, z_s)(r, z) \quad (4)$$

반면, 기존의 CNN surrogate는 고정 격자 기반으로써, 동일 입력으로부터 미리 정해진 격자 D 우의 TL 맵을 직접 회귀한 것이다(Eq. (5)).

$$\tilde{\mathcal{G}} : (X, f, z_s) \mapsto \widetilde{\text{TL}_D}, \quad D = \{(r_i, z_j)\}_{i,j}. \quad (5)$$

두 정의의 차이는 예측 결과를 “임의 좌표에서 평가 가능한 연속 함수”로 취급할지, “특정 해상도의 고정 격자 맵”으로 취급할지를 결정하며, 해상도 변경 가능성, 좌표 단위 평가, 그리고 학습 범위를 벗어난 조건에서의 구조 보존을 어떤 방식으로 검증할지까지 직접적으로 좌우한다. 본 연구는 이 오퍼레이터 정의를 기반으로 R-DeepONet을 구성하고, 학습 목적에 TL 오차와 물리 일관성 제약을 결합하여(III) 외삽 조건에서의 핵심 구조 보존을 정면으로 평가한다(IV).

3. Models and training methodology

III.A. R-DeepONet architecture for SOFAR TL

본 연구는 SOFAR 채널에서의 전달손실(Transmission Loss, TL) 예측을 연속 좌표에서의 스칼라 출력으로 다루기 위해 DeepONet 계열 신경 연산자를 구성하였다. 제안 모델은 입력 서명 X 와 조건 벡터 \mathbf{c} 로부터 환경 임베딩 \mathbf{b} 를 생성하는 브랜치 네트워크와, 정규화 좌표 $\tilde{\mathbf{x}} = (\tilde{r}, \tilde{z})$ 로부터 좌표 기저 $\mathbf{t}(\tilde{\mathbf{x}})$ 를 생성하는 trunk network로 이루어진다. 본 절에서 R-DeepONet의 R은 브랜치 네트워크에 사용한 ResNet-18의 residual block 구조를 의미하며,^[12] 출력 레벨에서 별도의 skip connection을 추가하지 않고 표준 DeepONet 결합(내적) 형태를 따른다.

모델 입력은 (i) ray-density map $X \in \mathbb{R}^{1 \times 256 \times 256}$, (ii) 조건 벡터 $\mathbf{c} = [\tilde{\mathbf{f}}, \tilde{z}_s] \in [0,1]^2$, (iii) 좌표 입력 $\tilde{\mathbf{x}} = (\tilde{r}, \tilde{z})$ 로 정의하였다. $\tilde{\mathbf{f}}, \tilde{z}_s, \tilde{r}, \tilde{z}$ 의 정규화 TL 타깃 정규화 및 역정규화 규칙은 II.B에서 정의한 설정을 따른다.

먼저 브랜치 네트워크는 두 갈래로 구성된다. 첫째, ray 서명 X 는 1채널 입력으로 수정된 ResNet-18 인코더 $E_\theta(\cdot)$ 를 통해 전역 특징으로 요약되며, adaptive average pooling 이후 512차원 특징 $\mathbf{h}_c \in \mathbb{R}^{64}$ 로 변환한다. 두 특징을 결합하여 환경 임베딩 $\mathbf{b} \in \mathbb{R}^K$ 를 생성하는 과정은 다음과 같이 정의한다(Eq. (6)).

$$\mathbf{h}_X = E_\theta(X) \in \mathbb{R}^{512} \quad (E_\theta: \text{ResNet-18}^{1\text{-channel}} + \text{GAP}) \quad (6-a)$$

$$\mathbf{g}_X = FC_X(Drop_p(\mathbf{h}_X)) \in \mathbb{R}^{512} \quad (FC_X: 512 \rightarrow 512, Drop_p: p = 0.06) \quad (6\text{-b})$$

$$\mathbf{h}_c = M_\phi(\mathbf{c}) \in \mathbb{R}^{64} \quad (M_\phi: 2 \rightarrow 128 \rightarrow 64, GELU) \quad (6\text{-c})$$

$$\mathbf{b} = FC_b([\mathbf{g}_X; \mathbf{h}_c]) \in \mathbb{R}^K \quad ([\cdot; \cdot]: concat, FC_b: 576 \rightarrow K, K = 384) \quad (6\text{-d})$$

여기서 $Drop_p(\cdot)$ 는 학습 시에만 적용되며 평가 시에는 항등으로 동작한다.^[13] 또한 본 연구에서 는 ResNet-18을 사전학습 없이 초기화하고, 초기 계층은 고정하여 학습 안정성을 확보하였다.

반면 Trunk Network는 좌표 $\tilde{\mathbf{x}}$ 에 대해 사인, 코사인 기반 positional encoding을 적용한 뒤 MLP 를 통해 K 차원 좌표 기저를 생성한다 (Eq. (7)).^[14]

$$\phi(\tilde{\mathbf{x}}) = [\tilde{\mathbf{x}}, \{\sin(2^k \cdot 2\pi\tilde{\mathbf{x}}), \cos(2^k \cdot 2\pi\tilde{\mathbf{x}})\}_{k=0}^{L-1}] \in \mathbb{R}^{2+4L} \quad (L = 9) \quad (7\text{-a})$$

$$\mathbf{t}(\tilde{\mathbf{x}}) = T_\psi(\phi(\tilde{\mathbf{x}})) \in \mathbb{R}^K \quad (T_\psi: 38 \rightarrow 320 \times 7 \rightarrow K, GELU, K = 384) \quad (7\text{-b})$$

최종 출력은 환경 임베딩과 좌표 기저의 내적으로 계산된다(Eq. (8)).

$$\hat{Y}(\tilde{\mathbf{x}}) = \mathbf{b}^\top \mathbf{t}(\tilde{\mathbf{x}}) \quad (8)$$

$\hat{Y}(\tilde{\mathbf{x}})$ 는 정규화된 TL 스칼라 예측값이며, dB 스케일 복원은 II.B에서 정의한 역정규화 및 클리 팅 규칙(Eq. (3))에 따라 평가 단계에서 적용한다. 이 구성은 동일한 (X, c)에 대해 $\tilde{\mathbf{x}}$ 만 변화시키 며 TL 값을 산출할 수 있도록 하여, 고정 격자 출력에 종속되는 이미지 회귀 모델과 구분되는 연 속 좌표 기반 예측을 제공한다.

III.B. Physics-consistent loss functions and training setup

R-DeepONet 모델을 구축할 시, SOFAR 채널 TL 예측에서 단순한 격자 일치(pointwise fit)만으 로는 수렴 구역(convergence zone)과 음영 구역(shadow zone) 등 구조적 패턴이 쉽게 흔들리고, 비 물리적 아티팩트가 발생할 수 있다는 점을 전제로 한다. 이에 따라 학습 목표를 값 정확도와 물 리 일관성을 동시에 만족해야하며, 손실 함수는 값 기반 회귀 항에 물리 일관성 항을 선택적으

로 결합하는 형태로 구성하였고 학습에서 최소화 하는 전체 손실은(Eq. (9))와 같다.

$$\mathcal{L} = \mathcal{L}_{Huber} + \lambda_{rec} \cdot \mathcal{L}_{rec} + \lambda_{tv} \cdot \mathcal{L}_{tv} + \lambda_{grad} \cdot \mathcal{L}_{grad} \quad (9)$$

여기서 \mathcal{L}_{Huber} 는 좌표 샘플에서의 TL 값 오차를 줄이는 회귀 항이며,^[15] \mathcal{L}_{rec} , \mathcal{L}_{tv} , \mathcal{L}_{grad} 는 각각 상반성(reciprocity) 및 공간적 매끄러움(smoothness)을 반영하기 위한 선택 항이다.

학습 프로토콜은 다음과 같이 구성하였다. 데이터는 파일 단위로 학습/검증 세트를 구성하고, 각 iteration에서 $B = 4$ 개의 환경을 샘플링한 뒤 환경마다 N_q 개의 좌표를 선택하여 좌표 기반으로 학습한다. 좌표 선택은 에지 집중(edge-focused) 샘플링을 적용하여 고기울기 영역을 우선 관측하도록 하였고, 최적화 안정화를 위해 gradient clipping을 적용하였다.^[16] 또한 EMA(exponential moving average)를 유지하여 평가 시 예측 변동을 줄였으며,^[17] 검증 지표를 기준으로 early stopping을 적용하고^[18] dB 스케일 MAE가 최소가 되는 체크포인트를 최종 모델로 선택하였다. dB 스케일 평가는 II.B의 역정규화 및 클리핑 규칙(Eq. (3))을 적용한 값으로 수행한다. 구체적인 학습 설정은 Table 2에 요약한다.

Table. 2 Training protocol summary

Item	Settings (numerical values)
Data split	Train/Val = 80/20
Batch	$B = 4$ environments per iteration
Coordinate sampling	$N_q = 2048$ coords per environment, edge-focused (edge_ratio=0.7)
value loss	Huber, $\delta = 1.0$
Physics-loss weights	$\lambda_{rec} = \lambda_{tv} = \lambda_{grad} =$
Optimizer	AdamW
Learning rate	$lr \approx 7 \times 10^{-4}$
Weight decay	$\approx 10^{-6}$
Stabilization	grad clip max-norm = 3.0 (L2), EMA decay = 0.999

III.C. Ray-based ResNet-18 CNN baseline (branch-only fixed-grid model)

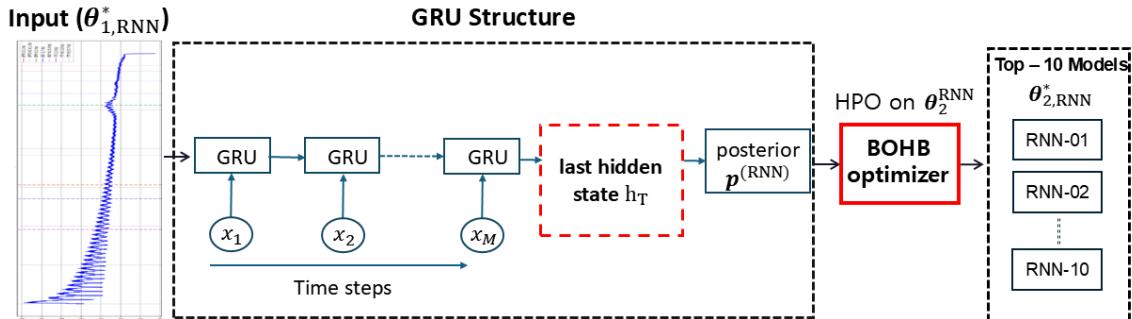


Figure 3. Output representation comparison: fixed-grid CNN vs R-DeepONet operator.

본 절은 R-DeepONet과, 그 브랜치에 사용한 ResNet-18을 그대로 사용한 고정 격자 CNN(ResNet-18 CNN)을 1:1로 비교하기 위한 통제된 기준선을 정의한다. 즉, ray-density map X 와 조건 벡터 $c = [\tilde{f}, \tilde{z}_s]$ 를 동일하게 사용하며, 브랜치 네트워크의 ResNet-18 인코더와 조건 MLP 또한 III.A의 설정을 그대로 유지한다. 이로써 두 모델의 차이점은 오직 출력 표현에서만 발생한다. 즉, R-DeepONet은 좌표 입력을 분리한 트렁크를 통해 $\widehat{TL}(r, z)$ 를 연속 좌표에서의 질의 가능한 형태로 출력하는 반면, ResNet-18 CNN 기준선은 디코더를 통해 256×256 고정 격자 TL 맵 \widehat{TL} 을 직접 회귀하며, 이러한 비교 설정은 Fig. 3에 도식화 하였다. 따라서 본 비교는 브랜치 네트워크 백본의 성능이 아니라 연산자 구조의 좌표 분리여부, 외삽, 질의 가능성의 효과만을 분리하여 검증한다. 끝으로, 두 모델의 입력, 출력 형태, 좌표 질의 가능성, 격자 및 도메인 변경 시 재학습 필요 여부는 Table 3에 요약한다.

Table. 3 Controlled comparison between R-DeepONet and the fixed-grid ResNet-18 CNN baseline.

Item	R-DeepONet (operator)	ResNet-18 CNN baseline (fixed-grid)
Input	$X(\text{ray-density map}), c = [\tilde{f}, \tilde{z}_s]$	$X(\text{ray-density map}), c = [\tilde{f}, \tilde{z}_s]$
Output	$\widehat{TL}(r, z)$ (continuous query)	\widehat{TL} map on 256×256 grid
Coordinate query	Yes (arbitrary (r, z))	No (grid-locked output)
Grid, domain change	Flexible via coordinate input	Requires retraining or redesign
Comparison focus	Operator structure (coord-separated)	Branch-only mapping (no trunk)

4. Numerical experiments and results

본 장에서는 심해 SOFAR 채널 전달손실(Transmission Loss, TL) 예측에서 제안한 R-DeepONet의 성능을 수치 실험으로 검증한다. 핵심 목표는 (i) 학습 분포 내(in-distribution)에서 고정 격자 기반 예측 모델과 비교해 값 정확도가 충분한지, (ii) 학습 범위를 벗어나는 조건에서 추가 학습 없이(Zero-shot) TL 회랑, 수렴 구역, 음영 구역 등 핵심 구조를 얼마나 안정적으로 보존하는지, (iii) 운용 관점에서 해석기 대비 계산 효율이 실질적으로 개선되는지를 동시에 확인하는 것이다.

모든 비교는 제안 모델(R-DeepONet)과 III.C에서 정의한 통제된 기준선(ResNet-18 CNN)을 중심으로 구성한다. 두 모델은 동일한 입력 서명(ray-density map)과 동일한 조건 변수(주파수, 음원 수심)를 사용하도록 정렬하여, 차이는 출력 표현(연속 좌표 질의 가능한 연산자 vs 고정 격자 TL 맵 회귀)에서 발생하도록 설계하였다.

IV.A. In-distribution accuracy on the SOFAR TL dataset

본 절에서는 외삽(IV.B) 논의에 앞서, 학습 분포 내(in-distribution) 조건에서 R-DeepONet의 baseline인 ResNet-18 CNN 대비 더 높은 정확도를 확보함을 보인다. 정량 평가는 dB 스케일 TL 기준으로 수행하며, 예측 출력은 II.B의 역정규화 및 클리핑 규칙(Eq. (3))을 적용한 뒤 MAE, RMSE(및 필요 시 MSSIM)로 비교하였다. 또한 수치 지표가 실제 회랑, 음영, 수렴 구역이 보존되는지 확인하기 위해, 동일 조건에서의 TL 맵을 나란히 제시한다(Fig. 4). 결과적으로 R-DeepONet은 홀드아웃 세트에서 MAE = "___" dB, RMSE = "___" dB (MSSIM = "___")를 달성했으며, ResNet-18 CNN은 MAE = "___" dB, RMSE = "___" dB (MSSIM = "___")를 보였다.

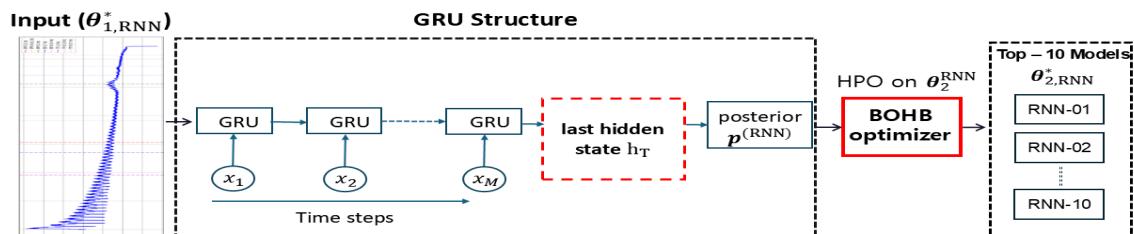


Figure 4. Ground Truth, R-DeepONet, CNN (a) TL map, (b) error map

IV.B. Zero-shot extrapolation in frequency, source depth, and SSP

본 절은 추가 fine-tuning이나 가중치 업데이트 없이(zero-shot) 학습 분포 밖 조건에서의 일반화 성능을 평가하여, ResNet-18 CNN과 R-DeepONet의 차이를 직접 비교한다. 외삽 축은 (i) 주파수 f , (ii) 음원 수심 z_s (iii) SSP이며, 각 실험에서 훈련 도메인과 테스트 도메인을 분리하여 동일한 입력 조건을 유지한 채 실행하였다. 여기서 SSP 외삽은 canonical Munk을 기준으로 한 perturbed Munk family를 테스트 도메인으로 구성하였으며, 하단 (Eq. (10))과 같다. 정량 결과는 (Table 4)에 요약하고, 구조 봉괴 여부는 대표 외삽 케이스의 TL 맵 비교로 확인한다(Fig. 5).

$$c_{test}(z; \alpha) = c_{Munk}(z) + \alpha \Delta c(z), \quad \alpha \in [0, \alpha_{max}] \quad (10)$$

여기서 $\Delta c(z)$ 는 수심 방향으로 매끄러운(smooth) 교란 함수이며, α 는 교란 강도를 조절하는 스칼라 파라미터이다.

(Fig. 5)는 외삽이 진행될수록 CNN 기준선이 TL 회랑(corridor)·수렴/음영 구역 구조를 빠르게 상실하는 반면, R-DeepONet은 구조를 상대적으로 안정적으로 유지함을 시각적으로 보여준다. 동일한 경향은 Table 4의 오차 지표에서도 확인되며, 주파수 외삽과 z_s 외삽에서 CNN은 훈련 경계 이후 오차가 급격히 증가하는 반면 R-DeepONet은 증가율이 완만하다. SSP 외삽에서는 Munk 기반의 작은 형상 변화에도 CNN의 공간 패턴이 쉽게 붕괴하는 경향이 두드러지며, R-DeepONet은 값 오차뿐 아니라 공간 구조 측면에서도 더 견고한 출력을 제공한다. 요약하면, 본 절의 결과는 고정 격자 이미지 회귀의 외삽 취약성과, 좌표 분리 연산자 표현의 구조 보존 이점을 zero-shot 조건에서 분리해 보여준다.

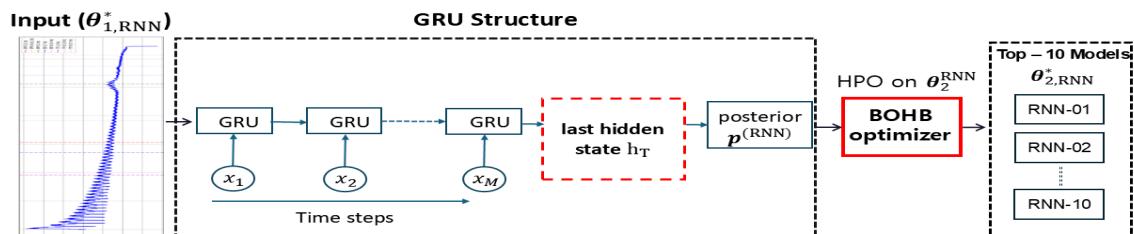


Figure 5. Zero-shot extrapolation results in (a) frequency, (b) source depth, and (c) SSP (perturbed Munk). Representative TL map comparisons illustrate the structural

degradation of the fixed-grid CNN baseline and the relative structure preservation of R-DeepONet under out-of-distribution conditions.

Table 4. Quantitative zero-shot extrapolation performance of R-DeepONet and ResNet-18 CNN baseline across frequency, source-depth, and SSP shifts (test domains held out from training).

IV.C. Mechanistic and practical validation: physics-consistency ablation and runtime trade-offs

본 절에서는 IV.B에서 관찰된 외삽 조건의 구조 보존 차이가 모델 아키텍처가 아닌 학습 목표에서 비롯되는 것인지 확인하기 위해 동일한 R-DeepONet 설정을 유지한 채 손실 구성만 바꾼 비교를 수행한다. 손실 항은 앞선 (Eq. (9))를 따르며, 본 절에서는 loss-guided와 value-only인 설정을 비교한다. 두 설정의 차이는 dB 스케일 오차 뿐만 아니라 reciprocity 위반 및 고주파 아티팩트 수준으로 나타나는지를 함께 확인한다.

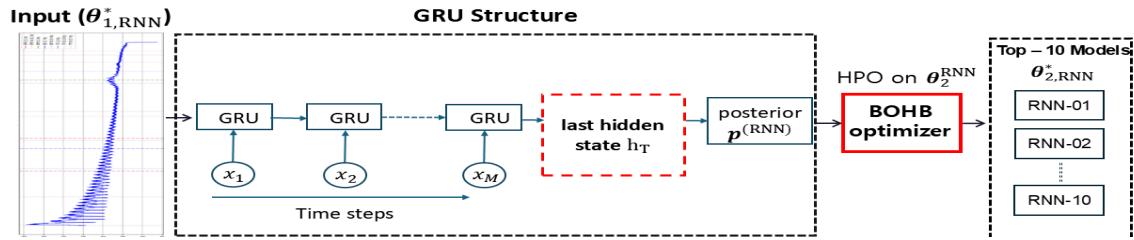


Figure 6. (a) GT TL map, (b) R-DeepONet value-only TL map, (c) R-DeepONet loss-guided TL map, (d) 구조 불괴를 드러내는 보조맵, ϵ Runtime bar chart: solver / CNN baseline / R-DeepONet, (f) R-DeepONet query scaling: N_q vs time.

Fig. 6(a-d)는 동일한 외삽 케이스에서 GT, value-only, loss-guided 예측을 동일 스케일로 비교하여, 값 오차보다 “구조 불괴 형태”의 차이를 직접 드러낸다. 특히 value-only에서는 corridor/shadow 경계의 흐트러짐이나 점상 고주파 노이즈가 TL 필드 전반에 확산되는 양상이 나타나는 반면, loss-guided에서는 주요 음향 구조의 연속성과 대비가 상대적으로 유지되는 경향을 보인다. 이때 구조적 차이는 TL 맵 자체의 비교와 함께, 보조 맵인 reciprocity residual과 $|\nabla TL|$ 을 함께 제시하여 정확도와 물리적 패턴 안정성을 분리하여 확인하였다.

Fig. 6(e-f)는 파동 기반 해석기 대비 학습 기반 모델의 런타임 이점을 정량화하고, R-DeepONet의 좌표 질의 수에 따른 비용 스케일링을 함께 제시하여 실시간 운용 시나리오에서의 실용성을

평가한다. 본 절에서 사용한 평가 지표는 Table 5에 통합 요약하여, 외삽에서의 구조 보존이 추가 계산비용을 과도하게 요구하지 않음을 함께 보여준다.

Table. 4 Summary of value-only vs loss-guided training under zero-shot extrapolation and runtime evaluation

Item	Value-only	Loss-guided
Training objective	Value loss only	Value loss + physics-consistency terms
Loss weights	$\lambda_{rec}, \lambda_{tv}, \lambda_{grad} = 0$	$\lambda_{rec} =, \lambda_{tv} =, \lambda_{grad} =$
MAE		
RMSE		
MSSIM		
Reciprocity violation metric		
Artifact, high-frequency metric		
Qualitative structure	More unstable	More stable
Wall-clock time per case (s)		

5. Conclusions

본 연구는 심해 SOFAR 채널에서 레이 기반 환경 서명과 운용 조건으로부터 전달손실 필드를 예측하는 문제를 연속 좌표에서 질의 가능한 신경 연산자 학습으로 정식화했다. 이를 위해 ResNet 기반 브랜치와 좌표 기반 트렁크를 결합한 R-DeepONet을 제안하고, 물리 일관성 손실을 포함한 학습 목표로 비물리적 아티팩트와 구조 붕괴를 억제하도록 설계했다. 비교의 공정성을 위해 동일한 입력 인코더와 조건 임베딩을 공유하되 출력 표현만 다른 고정 격자 CNN 기준선을 함께 구성하여, 성능 차이가 백본이 아니라 출력 표현에서만 발생하도록 통제했다. 실험 결과 제안 모델은 분포 내 조건에서 기준선과 동급의 정확도를 유지하면서도, 주파수와 음원 수심, 그리고 교란된 Munk SSP에 대한 추가 학습 없는 외삽에서 TL 회랑과 음영 구역의 구조를 더 안정적으로 보존했다. 향후에는 범위 의존 환경과 지형 변화를 포함한 데이터로 확장하고, 불확실

성 정량화를 연산자에 통합할 것이다.

Acknowledgements

Reference

- Ewing, M., Worzel, J. L., Worzel, J. L., Ewing, M., and Pekeris, C. L. (1948). "LONG-RANGE SOUND TRANSMISSION," in *Propagation of Sound in the Ocean* (Geological Society of America), p. 0.
- Cable, P. G. (1971). "Average Transmission Loss for Model SOFAR Channels."
- Porter, M. (1992). "The KRAKEN normal mode program."
- Porter, M. B., and Bucker, H. P. (1987). "Gaussian beam tracing for computing ocean acoustic fields," *The Journal of the Acoustical Society of America* **82**, 1349–1359.
- Langhirt, M., Martinelli, S., and Holland, C. W. (2018). "A study of underwater propagation model performance in complex environments with inherent uncertainty," *The Journal of the Acoustical Society of America* **144**, 1735–1735.
- Mallik, W., Jaiman, R. K., and Jelovica, J. (2022). "Predicting transmission loss in underwater acoustics using convolutional recurrent autoencoder network," *The Journal of the Acoustical Society of America* **152**, 1627–1638.
- Wang, H., Peng, S., He, Q., and Zeng, X. (2024). "Predicting underwater acoustic transmission loss in the SOFAR channel from ray trajectories via deep learning," *JASA Express Letters* **4**.
- Sun, Z., Wang, Y., and Liu, W. (2025). "End-to-end underwater acoustic transmission loss prediction with adaptive multi-scale dilated network," *The Journal of the Acoustical Society of America* **157**, 382–395.
- Lu, L., Jin, P., Pang, G., Zhang, Z., and Karniadakis, G. E. (2021). "Learning nonlinear operators via DeepONet based on the universal approximation theorem of operators," *Nature Machine Intelligence* **3**, 218–229.
- He, T., Xing, G., Li, Y., Li, Q., and Zhou, S. (2021). "A novel baseline-free defect detection and localization method of welded steel plate based on reciprocity loss," *Measurement Science and Technology* **32**, 025602.
- Scully-Power, P. D., Malay, J. T., and Browning, D. G. (1981). "Canonical eddies—An estimation of acoustical effects," *The Journal of the Acoustical Society of America* **69**, S33–S34.
- He, K., Zhang, X., Ren, S., and Sun, J. (2016). "Deep Residual Learning for Image Recognition," in *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 770–778.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). "Dropout: a simple way to prevent neural networks from overfitting," *J. Mach. Learn. Res.* **15**, 1929–1958.
- Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, L., and Polosukhin, I. (2017). "Attention is All you Need," in *Neural Information Processing Systems*.
- Meyer, G. P. (2021). "An Alternative Probabilistic Interpretation of the Huber Loss," in *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5257–5265.

- Pascanu, R., Mikolov, T., and Bengio, Y. (2012). "On the difficulty of training recurrent neural networks," in *International Conference on Machine Learning*.
- Polyak, B. T., and Juditsky, A. B. (1992). "Acceleration of Stochastic Approximation by Averaging," *SIAM Journal on Control and Optimization* **30**, 838–855.
- Vilares Ferro, M., Doval Mosquera, Y., Ribadas Pena, F. J., and Darriba Bilbao, V. M. (2023). "Early stopping by correlating online indicators in neural networks," *Neural Networks* **159**, 109–124.