

논문

물리정보신경망을 이용한 파동방정식 모델링 전략 분석

조상인¹ · 최우창¹ · 지준² · 편석준^{1*}

¹인하대학교 에너지자원공학과

²한성대학교 AI응용학과

Analysis on Strategies for Modeling the Wave Equation with Physics-Informed Neural Networks

Sangin Cho¹, Woochang Choi¹, Jun Ji², and Sukjoon Pyun^{1*}

¹Department of Energy Resources Engineering, Inha University

²Department of AI Application, Hansung University

요약

편미분방정식의 해를 구하기 위한 여러 수치해법들의 한계와 순수 데이터 기반 기계학습의 단점을 극복하기 위해 물리정보신경망 (physics-informed neural network, PINN)이 제안되었다. 물리정보신경망은 편미분방정식을 손실함수 구성에 직접 활용하여 기계학습 훈련에 물리적 제약을 주는 기법으로 파동방정식 모델링에도 활용될 수 있다. 그러나 물리정보신경망을 이용하여 파동방정식을 풀기 위해서는 신경망 훈련 시 입력에 대한 2차 미분이 수행되어야 하고, 그 결과로 출력되는 파동장은 복잡한 역학적 현상들을 포함하고 있어 섬세한 전략이 필요하다. 이 해설 논문에서는 물리정보신경망의 기본 개념을 설명하고 파동방정식 모델링에 활용하기 위한 고려사항들에 대해 고찰하였다. 이러한 고려사항에는 공간좌표 정규화, 활성화함수 선정, 물리손실 추가 전략이 포함된다. 훈련자료의 공간좌표를 정규화한 후 사용하면 파동방정식 모델링을 위한 신경망 훈련에서 초기 조건이 더 정확하게 반영되는 것을 수치 실험을 통해 보였다. 또한 신경망을 통한 파동장 예측에 가장 적절한 활성화함수를 선정하기 위해 여러 함수들의 특성을 비교했다. 특성 비교는 각 활성화함수들의 입력자료에 대한 미분과 수렴성을 중심으로 이루어졌다. 마지막으로 신경망 훈련 중 손실함수에 물리손실을 추가하는 두가지 시나리오의 결과를 비교하였다. 수치 실험을 통해 훈련 초기부터 물리손실을 활용하는 전략보다 초기 훈련단계 이후부터 물리손실을 적용하는 커리큘럼 기반 학습전략이 효과적이라는 결과를 도출했다. 추가로 이 결과를 물리손실을 전혀 사용하지 않은 훈련 결과와 비교하여 PINN기법의 효과를 확인하였다.

주요어

물리정보신경망(PINN), 모델링, 파동방정식

ABSTRACT

The physics-informed neural network (PINN) has been proposed to overcome the limitations of various numerical methods used to solve partial differential equations (PDEs) and the drawbacks of purely data-driven machine learning. The PINN directly applies PDEs to the construction of the loss function, introducing physical constraints to machine learning training. This technique can also be applied to wave equation modeling. However, to solve the wave equation using the PINN, second-order differentiations with respect to input data must be performed during neural network training, and the resulting wavefields contain complex dynamical phenomena, requiring careful strategies. This tutorial elucidates the fundamental concepts of the PINN and discusses considerations for wave equation modeling using the PINN approach. These considerations include spatial coordinate normalization, the selection of activation functions, and strategies for incorporating physics loss. Our experimental results demonstrated that normalizing the spatial coordinates of the training data leads to a more accurate reflection of initial conditions in neural network training for wave equation modeling. Furthermore, the characteristics of various functions were compared to select an appropriate activation function for wavefield prediction using neural networks. These comparisons focused on their differentiation with respect to input data and their convergence properties.

Received: 3 July 2023; Revised: 18 July 2023; Accepted: 20 July 2023

*Corresponding author

E-mail: pyunsj@inha.ac.kr

Address: Department of Energy Resources Engineering, Inha University, 100 Inharo, Michuhol-gu, Incheon 22212, Korea

©2023, Korean Society of Earth and Exploration Geophysicists

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Finally, the results of two scenarios for incorporating physics loss into the loss function during neural network training were compared. Through numerical experiments, a curriculum-based learning strategy, applying physics loss after the initial training steps, was more effective than utilizing physics loss from the early training steps. In addition, the effectiveness of the PINN technique was confirmed by comparing these results with those of training without any use of physics loss.

KEYWORDS

Physics-Informed Neural Networks (PINN), modeling, wave equation

서론

편미분방정식(partial differentiation equation, PDE)은 미지의 함수가 두 개 이상의 독립 변수에 종속적일 때 일어나는 여러가지 물리학적, 기하학적 문제를 수학적으로 모델링(mathematical modeling)한 것이다(Kreyszig *et al.*, 2011). 물리 시스템을 지배하는 편미분방정식을 정의하고 초기값 및 경계값 문제를 푸는 것은 임의의 조건에서 물리적 현상을 이해하는 데 도움이 된다. 편미분방정식의 해를 구하는 방법은 크게 해석적인 접근과 수치적인 접근으로 구분될 수 있다. 미적분학을 활용하여 구해지는 해석해(analytic solution)는 정확한 값을 얻을 수 있지만 대형 연립방정식, 비선형성, 복잡한 기하학적 문제들을 다루는 데 한계가 있다. 실제로 해석적으로 정확한 해를 구할 수 없는 수학적 모델이 많기 때문에, 그 대안으로 해석해를 근사하는 수치해(numerical solution)를 얻는다(Chapra, 2012). 편미분방정식의 수치해를 구하는 기법으로는 유한차분법(finite difference method, FDM), 유한요소법(finite element method, FEM), 경계요소법(boundary element method, BEM) 등이 있다. 이러한 기법들을 통해 해석적으로 해를 구할 수 없는 문제에 대해 근사적으로 접근할 수 있다는 장점이 있지만, 수치해를 구할 때는 높은 계산 비용과 해의 유일성 및 수렴성과 관련된 문제 등이 발생할 수 있다. 특히 경계조건으로서 사용되는 정보에 누락(missing), 갭(gappy), 또는 잡음(noise)이 존재한다면, 현재로서는 기존의 방법으로 물리 시스템 문제를 해결하는 것이 불가능하다(Karniadakis *et al.*, 2021). 수치해법이 처음 제안된 이후로 많은 발전이 있었지만, 수치해법에 대한 일부 문제들은 아직 해결되지 않고 남아있다(Blechs Schmidt and Ernst, 2021).

전통적인 수치해법들의 대안으로서, 기계학습을 이용하여 편미분방정식의 해를 구하고자 하는 연구들이 수행되었다(Paganini *et al.*, 2018; Rasp *et al.*, 2018; Kasim *et al.*, 2020; Siahkoobi *et al.*, 2019). 기계학습은 일반적으로 대량의 자료를 사용해 신경망을 훈련시켜 분류, 패턴 인식, 회귀 등의 작업을 수행하는 기법이다. 특히, 비선형 활성화함수를 갖는 은닉층(hidden layer)을 하나 이상 포함한 신경망을 이용하여 훈련하면 보편적 근사 정리(universal approximation theorem)에 의해 이론적으로 어떠한 함수라도 근사할 수 있기 때문에 편미분방정식의 해를 구하는 과정에 기계학습 적용이 가능하다(Hornik *et al.*, 1989; Garcia Ferrero, 2018). 그러나 순수한 데이터 기

반(data-driven) 기계학습은 사용된 훈련 자료의 특성 분포 외의 정보는 학습할 수 없다는 문제가 있다(Moseley *et al.*, 2020a). 특히, 자료가 너무 적거나 편향된 특성을 가지고 있을 때 신경망을 통해 도출된 해가 물리적으로 타당하지 않거나 수렴하지 않을 수 있다. 이에 따라 전통적인 모델링 기법과 기계학습이 가지는 단점을 극복하기 위해 물리적 법칙을 기계학습 모델에 통합하는 접근에 대한 필요성이 등장하였다(Karniadakis *et al.*, 2021).

물리 법칙을 기계학습 기법에 적용하는 개념은 이미 1990년대에 제안되었다(Psichogios and Ungar, 1992; Lagaris *et al.*, 1998; Lee and Kang, 1990). 특히 Dissanayake and Phan-Thien (1994)는 기계학습 알고리즘에 사전지식(prior knowledge)을 통합하는 개념을 처음 제안한 문헌으로 여겨지며, 넓은 의미에서 최초의 물리정보신경망(physics-informed neural network, PINN)으로 간주될 수 있다(Cuomo *et al.*, 2022). 또한, Lagaris *et al.* (1998)는 미분방정식의 해를 매개변수화 하여 인공 신경망(artificial neural network, ANN)을 통해 최적의 매개변수를 결정하는 방법을 제시했다. 이러한 방법론은 역전파 알고리즘(Rumelhart *et al.*, 1985)을 사용해 네트워크의 입력에 대한 결과의 미분값을 효율적으로 계산할 수 있게 되면서 그 활용이 용이해졌다(Baydin *et al.*, 2018; Pytorch, 2018; Tensorflow, 2015). Mallat (2016), Arridge *et al.* (2019), Kondor and Trivedi (2018) 등의 연구들을 기반으로 Raissi *et al.* (2019)는 미분 방정식이 지배하는 물리적 시스템을 푸는 새로운 종류의 신경망인 PINN을 제안하였다.

PINN은 편미분방정식을 신경망의 손실함수 구성에 활용한다(Karniadakis *et al.*, 2021). 기존의 데이터 기반 심층 신경망 기법들은 라벨(label)과 신경망 출력 사이의 오차를 정의하여 손실함수로 사용하지만, PINN은 여기에 편미분방정식을 변형한 오차항을 더하여 물리 정보를 네트워크가 학습할 수 있도록 설계한다. 이 때 편미분방정식의 오차항은 모델링 영역 내의 임의의 지점에서 계산된다(Cuomo *et al.*, 2022). 따라서 PINN은 불완전한 자료가 있는 경우에 대해 민감도가 낮고, 물리학적 의미가 존재하는 예측을 수행할 수 있는 기계학습 기법이라고 할 수 있다.

한편, PINN 기법의 대표적인 문제는 그 수렴 특성이 완전히 파악되지 않아 원하는 해에 수렴하지 못하거나 발산할 가능성이 있다는 것이다. Raissi *et al.* (2019)는 적용된 편미분방정식이 타당(well-posed)하고 그 해가 유일할 때 PINN으로 도

출된 결과가 높은 예측 정확도를 가짐을 경험적으로는 확인하였으나, 수렴성에 대한 이론적인 근거는 부족하다. 이후에 Wang *et al.* (2022), Shin *et al.* (2020), Mishra and Molinaro (2022) 등은 PINN이 편미분방정식의 해에 수렴한다는 이론적 타당성을 제시하였으나, 여전히 이론적으로 PINN을 이해하는 과정의 초기 단계이며, 수렴성의 보장은 PINN의 주요 문제 중 하나로 남아있다(Moseley *et al.*, 2021).

PINN은 편미분방정식의 해를 추정하는 정문제(forward problem)와 관측된 자료로부터 모델 매개변수를 구하는 역문제(inverse problem)를 모두 다룰 수 있다. Raissi *et al.* (2019)가 PINN을 제안한 이후, 이를 활용하여 물리 시스템을 지배하는 편미분방정식의 정문제 및 역문제를 해결하기 위해 많은 연구가 이루어졌다. 특히 유체 역학 분야에서는 나비에-스토크스 방정식(Navier-Stokes equation)을 활용한 연구들이 많이 수행되었다(Erichson *et al.*, 2019; Raissi *et al.*, 2020; Sun *et al.*, 2020; Zhu *et al.*, 2019; Cai *et al.*, 2021). 또한 Sahli Costabal *et al.* (2020)은 심방세동(atrial fibrillation)을 진단하기 위한 파동 전파를 설명하기 위해 PINN을 사용하였고, Chen *et al.* (2020)은 미세 광섬유에서 빛의 산란에 대한 역문제에 PINN을 적용하였다. Haghghat *et al.* (2021)은 고체역학에서의 정문제 및 역문제에 PINN이 활용될 수 있는 가능성을 확인했다.

탄성과 탐사 분야에서도 파동이 전파할 때 나타나는 물리적 현상을 설명하기 위해 PINN을 적용하는 연구들이 진행되고 있다. 주시(traveltime)를 계산하기 위해 사용하는 아이코날 방정식(Eikonal equation)을 PINN에 적용하여, 파동장의 초동 주시를 예측하고 초동 주시 토모그래피 역산을 수행할 수 있음이 확인되었다(Smith *et al.*, 2020; Waheed *et al.*, 2020; Waheed *et al.*, 2021). Moseley *et al.* (2020b)은 2차원 음향 파동방정식의 해를 구할 때 PINN의 활용에 대한 연구를 수행했고, Song *et al.* (2021)은 이방성 매질에서 2차원 음향 파동방정식의 해를 PINN으로 예측하였다. PINN은 탄성과 자료로부터 물성을 예측하기 위한 역문제 풀이에 관한 연구에도 활발히 이용되고 있다. 대표적으로 음향 파동방정식을 기반으로 한 속도 모델 역산(Xu *et al.*, 2019), 탄성 파동방정식(elastic wave equation)을 이용한 1차원 속도와 밀도 역산(Karimpouli and Tahmasebi, 2020), 전파형 역산(full waveform inversion, FWI) 결과 개선(Rasht-Behesht *et al.*, 2022), 탄성 파동방정식을 이용한 석유탄성(petroelastic) 특성 추정(Junior *et al.*, 2019)에 관한 연구 등이 그 예이다. 또한 자료처리 분야에서는 국부 평면파 미분 방정식(local plane wave differential equation)을 PINN에 적용하여 탄성과 자료의 보간(interpolation)을 수행하고자 하는 연구가 이루어지고 있다(Brandolin *et al.*, 2022).

PINN을 사용하여 파동 방정식을 모델링하는 경우 2차 편미분방정식을 자주 사용한다. 따라서 시간과 공간에 대한 2차 미분항까지 고려해야 하기 때문에 1차 편미분방정식을 사용하는

시스템에 비해 복잡도가 높아지게 된다. 또한 지질 구조 모델은 역진층이나 단층을 포함하거나 이방성, 불균질성 등의 특성을 가질 수 있으므로 이러한 특성들이 파동장에 반영되어 복잡한 결과가 나타날 수 있다(Yilmaz, 2001). 이와 관련하여 PINN을 사용한 파동방정식 모델링을 수행할 때 몇 가지 문제가 발생할 수 있다. 신경망의 출력이 파동의 여러 역학적 현상들을 모두 표현할 수 있는지에 대한 불분명함(Moseley *et al.*, 2020b), 유일해를 구하기 위한 경계조건이 충분한지에 대한 모호함(Alkhadhr and Almekkawy, 2023), 1차 미분에 비해 늘어난 계산 비용 등이 그 문제에 해당한다. Moseley *et al.* (2020b)은 이러한 문제들을 고려하여 PINN을 훈련하는 동안 커리큘럼 훈련 전략을 기본으로 여러 시간 단계에서의 파동장 스냅샷(snapshot)을 경계조건 훈련자료로 사용하는 등 다양한 전략들을 활용하였다.

이 해설 논문에서는 물리정보신경망의 기본 개념을 설명하고, Moseley *et al.* (2020b)의 훈련 전략을 기초로 하여 2차원 음향 파동방정식의 해를 모사하는 심층 신경망 훈련 과정에 대해 상세히 서술한다. 이론 1에서는 일반적인 물리정보신경망의 개념을 설명하고, 이론 2에서는 파동방정식의 해를 물리정보신경망을 활용하여 도출할 때 고려해야 할 사항들에 대해 살펴본다. 수치 예제에서는 2차원 음향 파동 방정식 모델링을 물리기반신경망으로 구현할 때 각 고려사항에 대한 테스트를 수행하고, 이를 기반으로 최적화된 훈련 방식을 제안한다.

이론 1. PINN의 기본 개념

임의의 변수로 매개화된 비선형 편미분방정식의 일반적인 표현은 다음과 같다(Raissi *et al.*, 2019).

$$u_t + \mathcal{N}[u; \lambda] = 0 \quad (1)$$

이 시스템의 잠재적 해(latent solution)는 u 로 표현되며, 이 때 비선형 미분 연산자 \mathcal{N} 은 공간에 대한 편미분을, u_t 는 시간에 대한 편미분을 의미한다. Raissi *et al.* (2019)은 PINN 기법을 데이터 기반 솔루션(data-driven solution)과 데이터 기반 발견(data-driven discovery) 문제에 대한 접근으로 구분하여 설명하는데, 식 (1)에서 신경망을 통해 u 를 구하는 문제(정문제)는 데이터 기반 솔루션, λ 를 구하는 문제(역문제)는 데이터 기반 발견에 해당한다. 여기서는 정문제인 파동방정식 모델링을 위한 접근에 집중하여 데이터 기반 솔루션 문제에 관해서만 다룬다.

PINN 이론에 따르면 편미분방정식의 해를 출력하는 미분연산자는 기계학습의 관점에서 신경망(N)으로 대체할 수 있다. 해를 구하고자 하는 점의 시공간적 위치 정보(x, t)를 입력으로 하여 그 위치에서 예측된 해(\hat{u})를 출력하는 신경망 N 은 식 (2)와 같이 정의할 수 있고, 이 신경망은 가중치 θ 를 업데이트하는 방식으로 훈련된다.

$$\hat{u}(x, t) = N[(x, t); \theta] \quad (2)$$

PINN 기법의 훈련에서 입력 변수에 대한 미분 연산을 통해 편미분방정식을 활용하는 물리정보신경망을 f 로 표기하면, 이는 식 (1)의 좌변으로 정의될 수 있다($f := u_t + \mathcal{N}[u]$). 방정식 $f=0$ 을 만족하는 해는 물리 시스템의 지배방정식을 따른다는 것을 의미하기 때문에 식 (2)의 신경망 N 이 훈련될 때 f 가 0에 가까워지는 제약조건을 손실함수에 추가하는 방식으로 PINN이 구현된다. 즉, 신경망 N 의 출력($N[(x, t); \theta]$)을 물리정보신경망(f)에 적용한 결과($f(N[(x, t); \theta])$)가 0에 가까워지는 방향이 θ 를 업데이트 하는 데 반영된다고 할 수 있다. 이러한 제약조건은 식 (3)과 같이 기존의 데이터 기반 손실함수인 MSE_u 에 MSE_f 를 추가하여 손실함수로 사용하는 것으로 고려될 수 있다. 두 손실함수는 각각 데이터손실(data loss)과 잔차손실(residual loss)(Heldmann *et al.*, 2023), MSE손실(mean-square error loss)과 PINN손실(PINN loss)(Lim *et al.*, 2022) 등의 용어로 표현되며 이 논문에서는 2차원 파동 방정식 모델링을 수행한 Moseley *et al.* (2020b)에 따라 경계손실(boundary loss)과 물리손실(physics loss)로 각 항을 명명한다.

$$L = MSE_u + kMSE_f$$

$$= \frac{1}{N_u} \sum_{i=1}^{N_u} |u(x_u^i, t_u^i) - N[x_u^i, t_u^i; \theta]|^2 + \frac{k}{N_f} \sum_{j=1}^{N_f} |f(N[x_f^j, t_f^j; \theta])|^2 \quad (3)$$

식 (3)에서 아래 첨자 u 와 f 는 각각 경계손실과 물리손실을 계산하기 위한 자료임을 의미하고, 위 첨자 i, j 는 각 자료의 일련번호, N_u 와 N_f 는 각 손실이 계산되는 데 사용되는 자료의 개수를 나타낸다. 경계손실에 해당하는 첫번째 항은 경계 조건과 초기 조건에 해당하는 점(x_u^i, t_u^i)에서의 라벨과 동일 위치에서 신경망 출력 사이의 차이를 계산한다. 두번째 항은 물리손

실에 해당하며, 네트워크의 출력이 물리 시스템을 얼마나 잘 만족하는지에 대한 정보이다. 물리 손실은 모델링 영역 내에서 임의의 점을 추출하여 계산하며, 그 점들을 콜로케이션 포인트(collocation point)라고 부른다. 두 손실 항의 분포 범위가 다르기 때문에, 가중치 k 를 손실균형(loss balancing)으로 정의하여 사용하였다. 이 때, 최적화된 k 는 수치 예제를 통해 경험적으로 결정할 수 있다. 이 방식은 일반적인 지도학습의 손실함수에 해당하는 경계 손실에 물리 손실이라는 비지도 정규화(regularization)를 추가한 것이라고도 볼 수 있다.

이론 2. 파동방정식 모델링을 위한 PINN

2차원 공간에서 음향 파동 방정식은 매질의 밀도가 일정하다고 가정하고 송신원 항을 무시하면 다음과 같이 표현할 수 있다.

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial z^2} - \frac{1}{v^2} \frac{\partial^2 u}{\partial t^2} = 0 \quad (4)$$

이론 1절의 설명에 따라 파동방정식을 모델링 할 때 물리정보신경망(f)은 식 (4)의 좌변으로 정의되는 미분항의 조합을 계산하는 신경망이 된다. 파동방정식 모델링을 위한 PINN 훈련 과정을 2개의 은닉층 구조의 신경망으로 가정하여 간략하게 도식화하면 Fig. 1과 같다.

시공간적 위치 정보(x, t)가 심층 신경망의 입력으로 주어졌을 때, 신경망은 파동장을 묘사하는 결과값(\hat{u})을 출력하고, 이 출력 값의 입력에 대한 미분을 수행해 물리손실(MSE_f)을 계산한다. 손실함수는 역전파(back propagation) 기법을 통해 은닉층 사이의 가중치(θ)를 업데이트하는 데 활용된다.

이론 1에서 설명한 PINN의 훈련과정에 파동방정식을 적용하면 다음과 같은 손실함수가 도출된다.

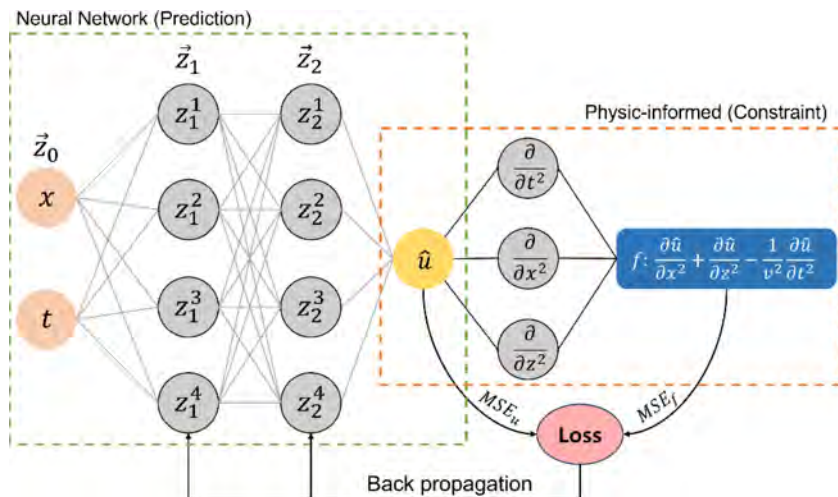


Fig. 1. Schematic diagram of PINN to solve the forward problem of the 2D acoustic wave equation.

$$L = \frac{1}{N_u} \sum_{i=1}^{N_u} |u_{FD}(x^i, t^i) - N(x^i, t^i)|^2 + \frac{k}{N_f} \sum_{j=1}^{N_f} |f[N(x^j, t^j); \theta; v(x^j)]|^2 \quad (5)$$

여기서 f 는 물리 기반 신경망, N 은 모델링 결과를 예측하는 심층신경망이며, 아래첨자 FD 는 유한차분법을 사용해 모델링한 결과임을 의미한다. 식 (5)에서 $v(x)$ 는 식 (1)에서 λ 로 나타낸 편미분방정식의 모델 매개변수를 의미하며 여기서는 파동방정식을 다루고 있으므로 모델링에 사용된 속도 모델이다.

경계손실(MSE_u)을 계산하기 위해 필요한 라벨 자료는 여러 수치해법을 통해 생성될 수 있지만, 여기서는 유한차분법을 사용하였다. 파동장 u_{FD} 은 속도, 밀도, 송신원 특성 등의 모델 매개변수를 고정한 후 처음 몇 개 시간 단계만큼 모델링하여 생성되며, 이는 파동방정식을 수치적으로 풀 때의 초기조건에 해당하지만, PINN에서는 경계손실을 계산하기 위해 사용한다는 의미에서 경계조건이라는 용어를 사용하여 설명한다. 식 (5)에서의 파동장 $u_{FD}(x_i, t_i)$ 은 $u_{FD} \in [0, X_{max}]$, $t_i \in [0, T_1]$ 의 범위 내에 있다고 표현할 수 있는데, X_{max} 는 모델링 영역의 상한, T_1 은 유한 차분법으로 모델링한 시간 단계 중 최댓값을 의미한다. 경계 손실은 이산화된 경계조건 집합에서 무작위로 지점들을 추출하여 계산되며 이는 훈련 중 가중치 업데이트 단계마다 수행된다. 이와 유사하게 물리 손실도 임의의 (x_j, t_j) 를 샘플링 하여 계산하며 이때의 자료 범위는 $x_j \in [0, X_{max}]$, $t_j \in [0, T_{max}]$ 으로 이는 모델링 하고자 하는 영역의 범위를 의미한다.

Moseley *et al.* (2020b)은 파동방정식을 PINN에 적용할 때 신경망으로 예측한 해의 정확도와 수렴성 개선을 위해 여러 전략들을 도입하였다. 이러한 전략에는 수렴성 개선을 위한 커리큘럼 훈련, 초기 해의 시간과 공간에서의 제약을 위한 여러 시간 단계에서의 경계조건 훈련자료 사용, 입력에 대한 그래디언트 계산을 추적하기 위해 비교적 작은 10 개 층의 완전 연결 신경망(fully connected network) 구성, 그리고 Softplus 활성화 함수 사용 등이 포함된다. 이 해설 논문에서는 이러한 전략 중 커리큘럼 훈련과 적절한 활성화 함수 선정에 대해 수치 예제

를 통해 분석하고자 한다. 이를 위해 여기서는 그에 해당하는 이론을 먼저 설명한다.

커리큘럼 훈련은 네트워크가 점진적으로 학습을 진행할 수 있도록 하는 전략이다. 훈련 초반에는 경계 손실만을 고려하여 가중치가 업데이트 되며 훈련 도중에 물리 손실을 추가하고, 점진적으로 시간 단계를 증가시키는 방식의 훈련 시나리오를 적용한다. 구체적으로 처음 10 개에서 20 개의 시간 샘플 동안은 경계손실만을 손실함수로 사용하여 신경망의 가중치를 업데이트하고, 그 이후 시간 단계를 늘려가며 경계손실과 물리 손실을 모두 손실함수 계산에 사용하는 방식으로 구현한다. 이러한 방식을 통해 신경망이 초기 파동장의 정보를 먼저 학습함으로써 본격적인 PINN 훈련 전에 좋은 초기 조건을 제공할 수 있다. 그 후, 시간 단계에 따라 학습을 진행해 나감으로써 해의 수렴성을 높일 수 있다.

비선형 활성화함수는 선형 함수라고 볼 수 있는 심층 신경망에 비선형성을 부여하기 위해 사용된다. 2 개 이상의 은닉층으로 구성된 심층 신경망 구조에는 기울기 소실(vanishing gradient) 문제를 방지하고 빠른 수렴을 위해 ReLU (rectified linear unit) 활성화함수가 주로 사용된다(Géron, 2022). Fig. 1에서 번째 층, j 번째 노드의 출력(z_i^j)과 신경망의 출력(\hat{u})은 다음과 같이 계산된다.

$$\begin{aligned} z_i^j &= \sigma_i(\vec{z}_{i-1} \cdot \vec{\theta}_{i-1}^j) \\ \hat{u} &= \sigma_{output}(\vec{z}_2 \cdot \vec{\theta}_2) \end{aligned} \quad (6)$$

여기서 σ_i 는 i 번째 층의 활성화함수를 의미한다. 파동방정식은 신경망의 출력(\hat{u})을 신경망의 입력 x 와 t 에 대하여 2차 미분하여 표현할 수 있다. 그러므로 활성화함수의 시간 및 공간에 대한 2차 미분을 계산할 수 없다면 파동방정식을 활용한 PINN 훈련에 문제가 될 수 있다. Table 1에 나타난 바와 같이 ReLU 활성화 함수는 2차 미분을 계산했을 때 그 값이 0이 되기 때문에 시간과 공간에 대한 2차 미분으로 표현되는 파동방정식의 특성 상 ReLU 함수를 훈련에 사용하는 것은 불가능하다.

따라서 Tanh (hyperbolic tangent)나 Softplus와 같은 활성화 함수들이 그 대안으로 사용될 수 있다. 이 두 활성화함수는 고차

Table 1. Definition of the activation functions and their 1st and 2nd derivative.

Name	Function	1st derivative	2nd derivative
ReLU	$\begin{cases} x & (x > 0) \\ 0 & (x \leq 0) \end{cases}$	$\begin{cases} 0 & (x < 0) \\ 1 & (x > 0) \\ \text{undefined} & (x = 0) \end{cases}$	0
Tanh	$\frac{e^x - e^{-x}}{e^x + e^{-x}}$	$1 - \left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right)^2$	$-2 \left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right) \left(1 - \left(\frac{e^x - e^{-x}}{e^x + e^{-x}}\right)^2\right)$
Softplus	$\ln(1 + e^x)$	$\frac{e^x}{(1 + e^x)}$	$\frac{e^{-x}}{(1 + e^{-x})^2}$
Siren	$\sin(x)$	$\cos(x)$	$-\sin(x)$

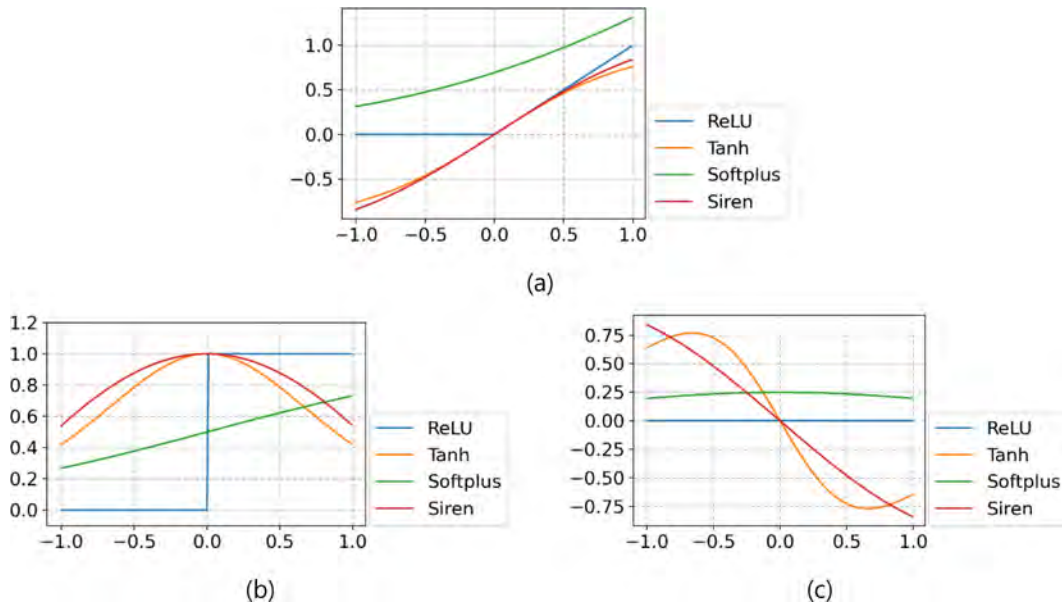


Fig. 2. Comparison between (a) four activation functions (ReLU, Tanh, Softplus and Siren), their (b) 1st derivative, and (c) 2nd derivative.

미분항이 계산될 수 있어 초기 문헌에서 주로 사용되었으나, 경계 자료를 이용하여 훈련하는 과정에서 수렴하지 못하는 문제가 발생하였다. 이를 해결하기 위해 Jagtap *et al.* (2019)는 기존의 활성화 함수에 하이퍼 파라미터(hyper-parameter)를 도입하여 PINN의 수렴성을 개선할 수 있는 적응 활성화 함수(adaptive activation function)을 제안하였다. Sitzmann *et al.* (2020)은 주기성이 있는 자료의 훈련에 적합한 siren 활성화 함수를 제안하였고, 이는 특히 탄성파 자료에 적합하며, PINN 훈련에도 성공적으로 적용될 수 있다. 각 활성화 함수의 파동방정식 모델링을 위한 PINN 훈련에 대한 적절성을 단순히 비교하기 위해 여기서는 하이퍼 파라미터 도입은 제외하고 분석한다. Fig. 2는 언급된 모든 활성화 함수들과 그 함수들의 1차, 2차 도함수를 그래프로 나타낸 것이다. 이를 통해 ReLU 보다는 Tanh, Softplus, Siren와 같은 함수들이 2차 미분항을 포함한 계산에 적절함을 확인하였다. 각 활성화 함수들의 수렴성에 대한 고찰은 수치 예제를 통해 다루인다.

수치 예제

파동방정식 모델링을 위한 물리 기반 신경망 구현을 위해 Moseley *et al.* (2020b)가 제안한 신경망 모델을 참고하였다 (Fig. 3). 그들은 1024 개 노드를 포함한 10 개의 완전연결층(fully connected layer, FCL)으로 구성된 신경망을 예측을 위해 사용했다. 또한 시간 지점, 공간 지점, 송신원 위치를 입력으로 하여 송신원 위치에 대해서는 재훈련이 필요하지 않도록 하였으나 여기서는 간단한 구현을 위해 시간과 공간 지점만을 입력으로 하는 신경망을 훈련하였다. 훈련을 위한 하이퍼 파라미터들은 경험적으로 결정하였다. 학습률(learning rate)은 10^{-5}

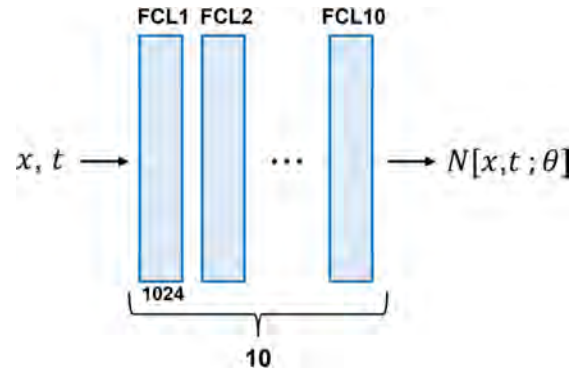


Fig. 3. Structure of the neural network used to solve the wave equation (modified after Moseley *et al.* (2020b)).

로 하여 ADAM (adaptive moment assessment) 최적화 기법을 사용하였고, 충분한 훈련을 위해 총 100,000번의 훈련 단계(training step)동안 훈련하였다. 물리손실과 경계손실 사이의 가중치를 나타내는 k 값은 10^{-6} 으로 결정하여 실험을 진행하였다.

수치 예제를 통해 Moseley *et al.* (2020b)가 제안한 커리큘럼 훈련과 여러가지 활성화 함수를 평가하고, 추가적으로 입력 자료의 공간 좌표 정규화를 분석한다. 공간정규화와 활성화 함수에 대한 테스트는 경계손실만을 손실함수로 하여 훈련한 결과를 도시하였다. 이를 통해 초기 조건이 신경망 훈련에 잘 반영되었는지 여부를 확인하고, 이러한 경계손실이 어느 경우에 잘 수렴하는지 파악할 수 있다. 또한, 물리손실 추가 전략들을 비교하여 손실함수 구성 방식에 따라 달라지는 결과를 확인하였다.

테스트를 위해 $1.5 \text{ km} \times 1.5 \text{ km}$ 크기의 영역에서 속도가

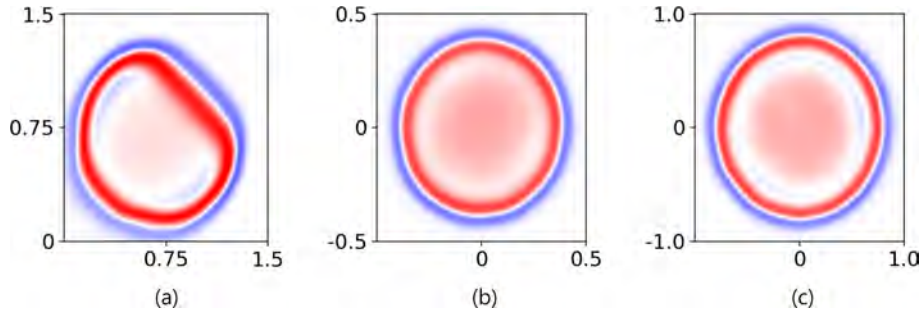


Fig. 4. Wavefield prediction from neural network training (a) without coordinate normalization and with normalization that results in coordinate ranging (b) $-0.5 \sim 0.5$, and (c) $-1 \sim 1$.

2500 m/s인 균질한 속도모델을 가정하였다. 송신원은 모델 중앙에 위치시켜 파동이 전파하는 형상을 직관적으로 확인하고자 하며, 중심 주파수가 20 Hz인 Ricker 송신파형을 사용하였다. 파동이 송신원으로부터 발생하여 전파하는 초기 시간대에는 경계조건(초기조건)으로 사용하는 파동장에 송신파형 정보가 불완전하게 반영되어 있으므로 파동방정식에서 송신원 항을 생략할 수 없다. 따라서 신경망이 초기 송신파형의 영향 없이 파동전파 매커니즘을 학습할 수 있도록 파형이 완전히 나타나는 0.1 초를 0 초로 가정하여 기준으로 사용했다. 총 0.3 초 동안의 파동장을 유한차분법으로 생성하고 0.1 초부터 0.3 초까지의 합성 파동장을 0 초에서 0.2 초까지의 파동장이라 명시했다. 새로운 시간 기준에서 0.04 초까지의 합성 파동장을 경계손실을 계산하기 위한 라벨자료로 사용했고, 나머지 시간대의 파동장은 결과 비교를 위해 사용했다. 모델링을 위한 공간 격자 간격은 분산분석을 통해 60 Hz 주파수에 대한 수치분산을 방지하기 위해 2 m로 설정하였고, 시간 격자 간격은 안정성 분석을 통해 수치해가 수렴할 수 있도록 2 ms로 설정하였다.

공간좌표 정규화

파동장의 모델링은 사전에 설정한 임의의 공간적 영역에서 수행된다. 따라서 공간 좌표 값의 범위는 모델링 매개변수 선정에 따라 달라진다. 그러나 파동이 송신원 위치로부터 퍼져 나가는 파동 전파 과정을 모사하는 파동방정식 모델링의 경우 절대적인 공간 좌표는 중요한 정보가 아니다. 예를 들어, 0 m 위치에 송신원을 두고 100 m 위치에서의 파동장을 확인하는 것과 100 m 위치에 송신원을 두고 200 m 위치에서의 파동장을 그려보는 것은 모델을 구성하는 매개변수(속도, 밀도 등)가 동일하다면 완전히 동일한 결과를 얻게 된다. 즉, 공간좌표 값이 아닌 송신 지점과의 상대적인 거리가 모델링 결과를 직접적으로 결정한다고 할 수 있다. 또한, 송신파형의 영향에서 자유롭다고 가정하면, 0 m로부터 100 m까지 100 m/s로 전파하는 파동장과 1000 m까지 1000 m/s로 전파하는 파동장의 모양도 완전히 동일하다. 이는 속도 스케일이 거리 스케일에 맞춰 조절된다고 가정하면, 거리를 정규화해도 파동장에는 영향을

주지 않음을 의미한다.

이를 고려하였을 때 훈련에 사용되는 모든 입력자료의 공간 좌표를 정규화 하는 것이 신경망 훈련에 과다한 정보가 주입되는 것을 방지할 수 있을 것이라 판단하였고 이에 대한 테스트를 진행하였다. Fig. 4a는 공간좌표를 그대로 사용하여 훈련한 결과이고, Figs. 4b, 4c는 공간좌표를 각각 $-0.5 \sim 0.5$, $-1 \sim 1$ 범위 내에 존재하도록 정규화 하여 훈련한 결과이다. Fig. 4에 도시한 결과는 경계손실만을 손실함수 계산에 사용하여 얻은 결과이다. 공간좌표를 정규화 하지 않은 경우 경계손실을 최적화해도 초기 시간대의 파동장이 부정확하기 때문에 그 이후의 물리손실 최적화 과정도 초기 조건을 잘 반영하지 못하게 된다. 이를 통해 PINN을 이용한 파동방정식 모델링은 공간좌표를 정규화 하는 과정이 필수적임을 확인할 수 있다.

활성함수

신경망에 사용되는 여러 활성화함수 중 ReLU, Softplus, Tanh, Siren 함수를 선택하여 파동방정식 모델링에 최적인 활성화함수를 분석해보았다. PINN을 이용한 파동방정식 모델링의 성패는 커리큘럼 훈련의 초기 단계인 경계손실 훈련에 매우 민감하므로 여기서는 경계손실만을 훈련의 손실함수로 사용했을 때의 결과를 비교하였다. 각 활성화함수의 성능 비교는 훈련 진행에 따른 경계손실(Fig. 5)과 0.04 초에서의 예측 결과(Fig. 6)로 이루어졌다.

훈련 단계에 따른 경계손실을 관찰해 보면 파란색으로 표현된 ReLU 함수와 빨간색으로 표현된 Siren 함수를 제외하고는 훈련이 진행되어도 계산된 손실 값이 줄어들지 않았다(Fig. 5). 손실함수 그래프가 이러한 형태를 보이는 이유는 Fig. 6에서 확인할 수 있다. 유한차분법으로 생성한 파동장(Fig. 6a)을 정답자료로 하여 0.04 초에서의 파동장을 비교했을 때, ReLU(Fig. 6b)와 Siren(Fig. 6e)을 사용한 경우에는 정답과 유사한 결과를 얻었으나 Softplus(Fig. 6c)와 Tanh(Fig. 6d)를 사용한 경우에는 훈련이 전혀 수행되지 못한 것을 확인할 수 있다. 이론전에서 언급한 바와 같이 물리손실이 추가되면 ReLU는 사용이 불가능하므로 파동방정식을 위한 PINN 훈련에는 Siren 함수가 활성화함수로 가장 적절하다.

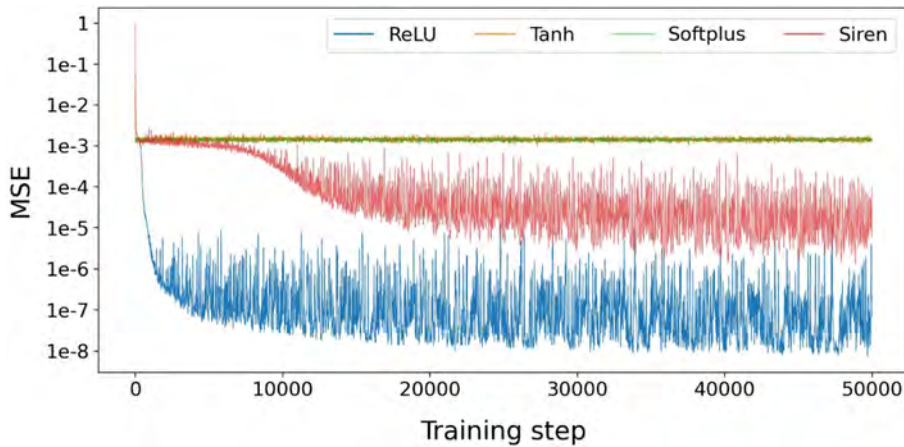


Fig. 5. Comparison of loss functions between four activation functions (ReLU, Tanh, Softplus, and Siren).

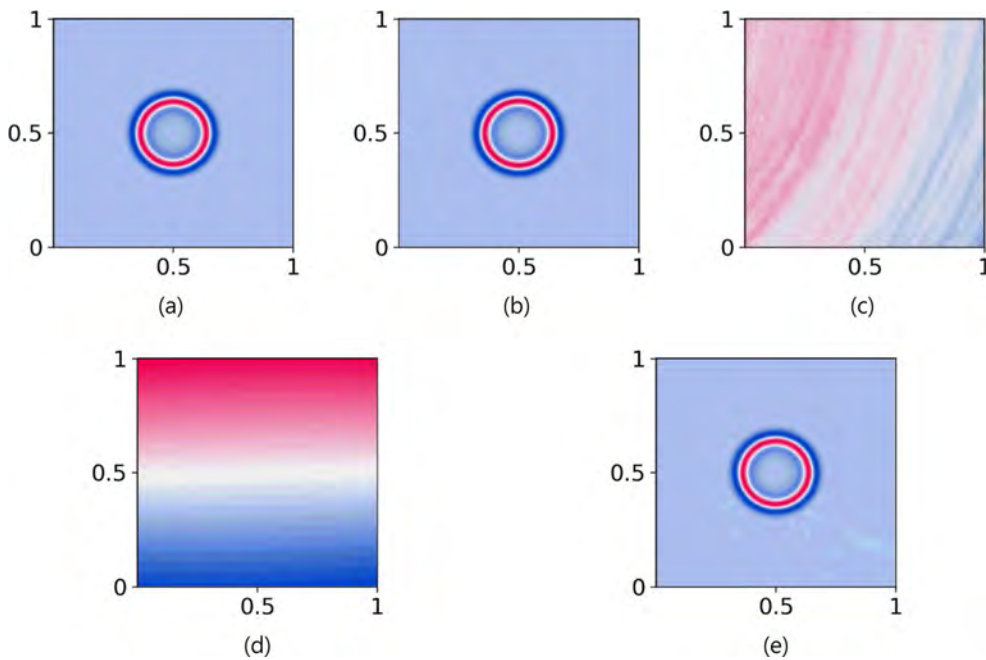


Fig. 6. Comparison between (a) label data generated with the FDM method and the results of the wavefield prediction at 0.04 s by neural network training with (b) ReLU, (c) Softplus, (d) Tanh, and (e) Siren activation functions.

물리손실 추가 전략

커리큘럼 훈련의 효과를 확인하기 위해 먼저 첫 훈련단계부터 물리손실과 경계손실을 모두 계산하여 훈련을 수행하였다. 또한, 처음에는 경계손실만으로 훈련을 진행하다가 50,000 번째 훈련단계부터 두 손실을 모두 반영하는 커리큘럼 기반 훈련도 수행하였다. 두 훈련 결과를 도시하면 Fig. 7과 같다. Fig. 7a는 유한차분법을 이용하여 모델링한 파동 전파 과정을 0.04 초 간격으로 나열하였다. PINN 훈련을 통해 예측한 파동장을 같은 시간 간격으로 나열하였고(Figs. 7b, 7d), 비교를 위해 유한차분법으로 얻은 파동장과의 차이를 도시하였다(Figs. 7c, 7e). PINN을 통해 출력된 두 결과 중 커리큘럼 학습 전략을 적

용한 학습 결과(Fig. 7b)가 그렇지 않은 학습 결과(Fig. 7d)보다 유한차분법을 통해 얻어진 해와 더 유사함을 확인할 수 있다. 각 훈련의 결과와 유한차분법으로 생성한 파동장의 차이를 보면 커리큘럼 전략을 사용했을 때(Fig. 7c)의 결과가 사용하지 않았을 때(Fig. 7e)보다 작게 나타났다. 이 테스트를 통해 커리큘럼 학습 전략을 도입하면 훈련의 수렴성과 성능이 개선된다는 결론을 도출할 수 있다.

마지막으로 물리 손실을 전혀 사용하지 않은 훈련 결과는 Fig.8과 같이 나타난다. 커리큘럼 훈련 전략을 사용한 경우(Fig. 7b)에는 Fig. 7c에서 볼 수 있듯이, 신경망 훈련을 위해 주어지지 않은 0.04 초 이후 시간대의 파동장 예측에서도 약

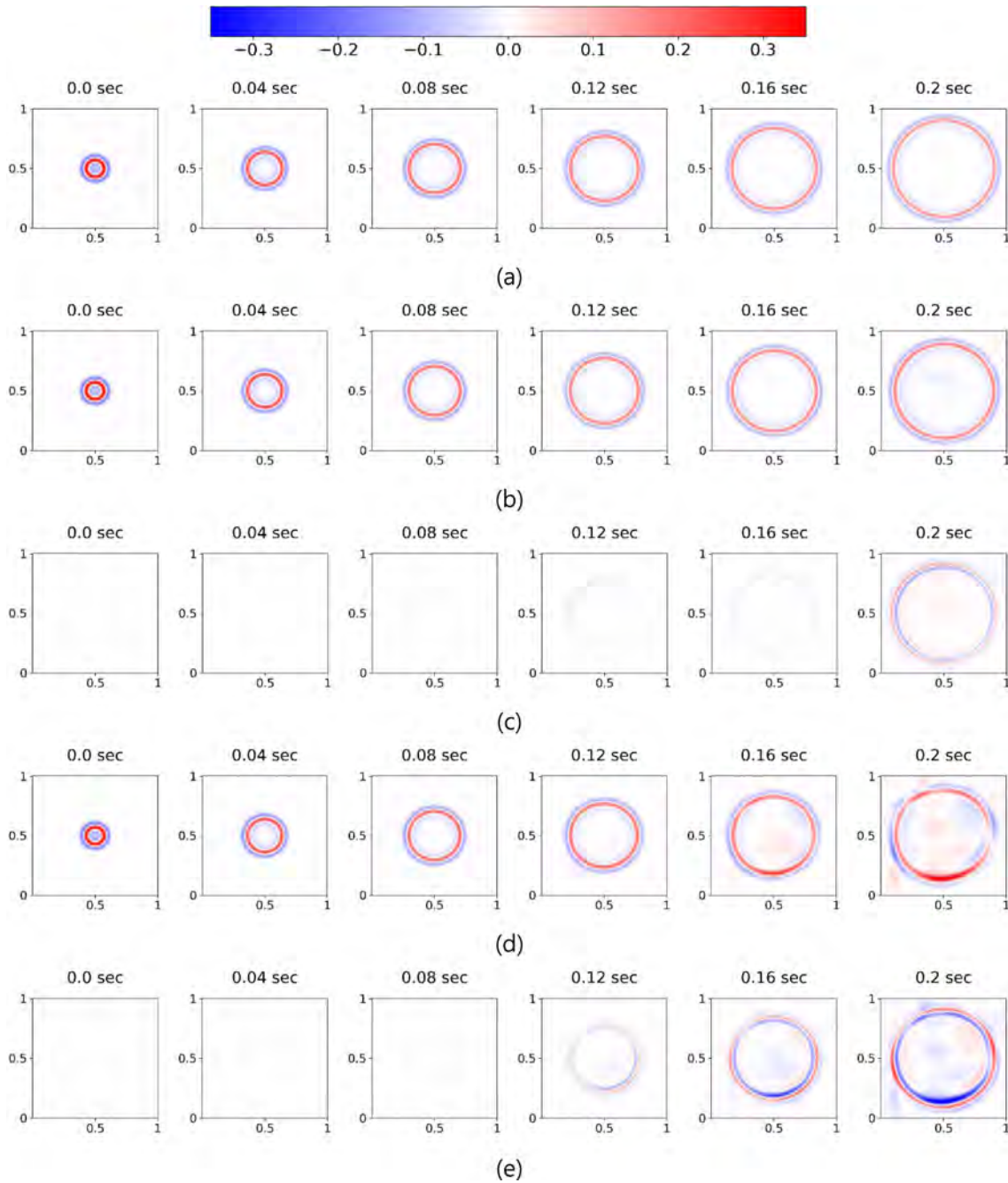


Fig. 7. Comparison between (a) label data generated with the FDM and the results of the wavefield prediction by neural network training (b) with and (d) without the curriculum-based learning strategy. (c) and (e) show the differences between (a) and (b), and (a) and (d), respectively.

간의 진폭 차이만 존재하고 거의 정확한 결과를 얻을 수 있다. 반면, 경계손실만을 사용하여 예측한 파동장의 경우(Fig. 8a), 초기 시간에서는 유한차분법 결과와 유사하지만, 라벨자료에 포함되지 않은 나중 시간의 파동장 예측이 거의 불가능한 것으로 나타났다(Fig. 8b). 이를 통해 PINN 기법이 기존 데이터 기반 기계학습보다 파동방정식 모델링에 우수한 성능을 가지고 있음을 확인할 수 있다.

결론 및 토의

이 연구에서는 편미분방정식을 기계학습에 직접적으로 반영하는 PINN 기법의 이론을 정리하고 이를 파동방정식 모델링에 적용하는 과정을 설명하였다. 특히 여러 수치 예제를 통해 2차원 파동방정식 모델링을 PINN으로 수행할 때 성공적인 훈련을 위한 전략에 대해 살펴보았다. 우선 경계손실을 계산하는

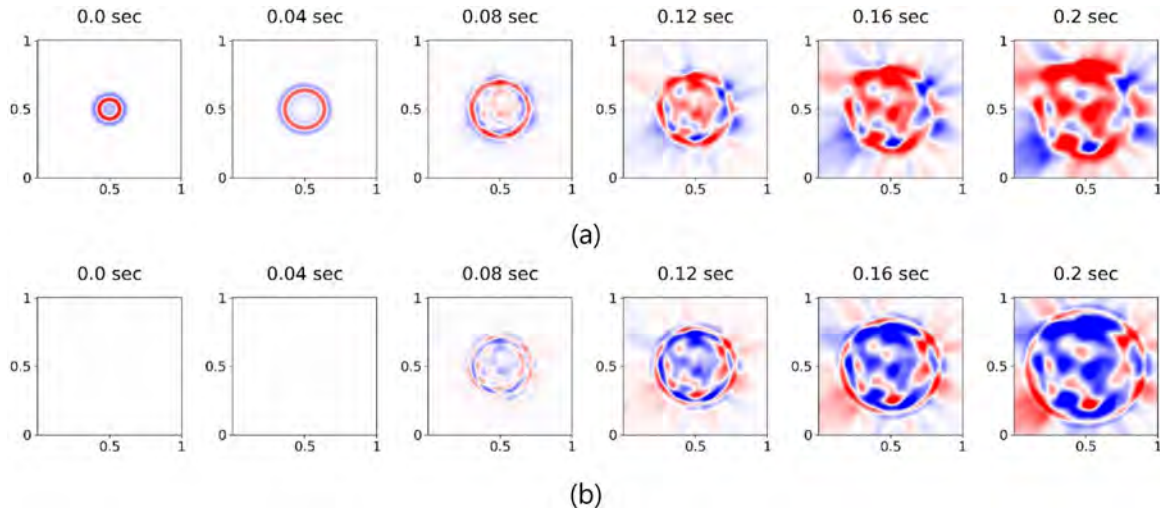


Fig. 8. (a) Results of the wavefield prediction achieved through neural network training without physics loss and (b) the differences observed compared to the wavefield generated by the FDM (Fig. 8a).

과정에서 공간좌표 값의 정규화를 통해 초기 조건이 더욱 정확하게 적용될 수 있음을 확인하였다. 활성화함수에 따른 경계손실의 수렴성 비교 결과를 보면 시간과 공간에 대한 2차 미분항을 포함한 파동방정식의 특성을 고려할 때 Siren 함수를 사용하는 것이 적절하다. 수치 예제를 통해 훈련 중 손실 함수에 물리 손실을 추가하는 것이 결과 개선에 효과적이라는 것을 확인할 수 있었다. 따라서 기존 문헌에서 제안한 커리큘럼 훈련 기반 학습을 사용하면 더 안정적이고 수렴성이 높은 훈련이 가능하다는 결론을 내릴 수 있다. PINN 기법은 적절한 전략을 통해 공간적인 모델링 경계 내에서는 기존 수치 해법과 유사한 파동장을 출력함을 확인하였다. 그러나 파동장이 공간적인 모델링 경계에 도달한 이후에는 경계면에서 반사가 일어나지 않고 무한 공간으로 전파하는 것과 같은 결과가 나타난다. 그 이유는 모델링 영역 외부에서의 파동장이 훈련에 사용되지 않았으며, 파동방정식을 통해서도 경계면에 대한 정보가 부여되지 않았기 때문이다. 경계면에서의 반사 현상을 모사하기 위해서는 단순히 물리손실에 경계조건 손실항을 추가하면 된다.

이 연구에서는 PINN의 수렴 특성을 파악하기 위해 균질한 속도모델을 물리손실 계산에 사용하였다. 하지만 실제 지질 구조를 고려한 모델링을 수행하려면 불균질한 속도모델을 사용해야 한다. 이를 위해 속도모델을 격자 구조로 이산화하고, 속도모델의 공간격자에 맞는 콜로케이션 포인트를 선정하여 물리손실을 계산함으로써 PINN에 불균질한 속도모델을 적용할 수 있다. 현재까지 제안된 PINN을 사용하여 파동방정식을 모델링하는 과정에 대한 문헌에서는 속도모델에 대한 일반화, 복잡한 구조에 대한 적용, 3차원 파동방정식 활용 등에 대한 고찰이 부족하다. 따라서 더욱 효율적이고 일반화된 활용을 위해 추후 연구가 필요하며, CNN이나 RNN 등 발전된 네트워크에

적용하여 효율성 및 정확도를 보완할 수 있을 것으로 기대한다. 그러나 PINN 기법을 파동방정식에 활용하면 기존의 수치 해법을 통해 전체 파동장을 계산하지 않아도 이산화 간격과 관계없이 특정 위치에서의 파동장 값을 알 수 있다는 장점이 있고, 기존 수치 해법과 비교했을 때 그 성능면에서도 충분히 활용될 수 있는 대안이라 생각된다.

감사의 글

이 논문은 2023년도 정부(산업통상자원부)의 재원으로 한국 에너지기술평가원의 지원을 받아 수행된 연구임 (20226A10100030, 고성능 해양 CO2 모니터링 기술개발).

References

- Alkhadr, S., and Almekkawy, M., 2023, Wave Equation Modeling via Physics-Informed Neural Networks: Models of Soft and Hard Constraints for Initial and Boundary Conditions, *Sensors*, **23**(5), 2792, doi: 10.3390/s23052792
- Arridge, S., Maass, P., Öktem, O., and Schönlieb, C. B., 2019, Solving inverse problems using data-driven models, *Acta Numer.*, **28**, 1-174, doi: 10.1017/S0962492919000059
- Baydin, A. G., Pearlmutter, B. A., Radul, A. A., and Siskind, J. M., 2018, Automatic differentiation in machine learning: a survey, *J. Mach. Learn. Res.*, **18**, 1-43. <https://www.jmlr.org/papers/volume18/17-468/17-468.pdf>
- Blechschmidt, J., and Ernst, O. G., 2021, Three ways to solve partial differential equations with neural networks—A review. *GAMM-Mitteilungen*, **44**(2), e202100006, doi: 10.1002/gamm.202100006
- Brandolin, F., Ravasi, M., and Alkhalifah, T., 2022, PWD-

- PINN: Slope-assisted seismic interpolation with physics-informed neural networks, In *Second International Meeting for Applied Geoscience & Energy* (pp. 2646-2650). Society of Exploration Geophysicists and American Association of Petroleum Geologists. doi: 10.1190/image2022-3742422.1
- Cai, S., Mao, Z., Wang, Z., Yin, M., and Karniadakis, G. E., 2021, Physics-informed neural networks (PINNs) for fluid mechanics: A review, *Acta Mech. Sin.*, **37**(12), 1727-1738, doi: 10.48550/arXiv.2105.09506
- Chapra, S., 2011, *EBOOK: Applied Numerical Methods with MATLAB for Engineers and Scientists*, McGraw Hill.
- Chen, Y., Lu, L., Karniadakis, G. E., and Dal Negro, L., 2020, Physics-informed neural networks for inverse problems in nano-optics and metamaterials, *Opt. Express*, **28**(8), 11618-11633, doi: 10.1364/OE.384875
- Cuomo, S., Di Cola, V. S., Giampaolo, F., Rozza, G., Raissi, M., and Piccialli, F., 2022, Scientific machine learning through physics-informed neural networks: where we are and what's next, *J. Sci. Comput.*, **92**(3), 88, doi: 10.1007/s10915-022-01939-z
- Dissanayake, M. W. M. G., and Phan-Thien, N., 1994, Neural-network-based approximations for solving partial differential equations, *Commun. Numer. Methods Eng.*, **10**(3), 195-201, doi: 10.1002/cnm.1640100303
- Erichson, N. B., Muehlebach, M., and Mahoney, M. W., 2019, Physics-informed autoencoders for Lyapunov-stable fluid flow prediction, *arXiv preprint arXiv:1905.10866*, doi: 10.48550/arXiv.1905.10866
- Garcia Ferrero, M. Á., 2018, *Global approximation theorems for partial differential equations and applications*, Ph.D. thesis, Universidad Complutense de Madrid. https://www.icmat.es/Thesis/2018/Tesis_Ma_Angeles_Garcia_Ferrero.pdf
- Géron, A., 2022, *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow*, O'Reilly Media, Inc.
- Haghighat, E., Raissi, M., Moure, A., Gomez, H., and Juanes, R., 2021, A physics-informed deep learning framework for inversion and surrogate modeling in solid mechanics, *Comput. Methods Appl. Mech. Eng.*, **379**, 113741, doi: 10.1016/j.cma.2021.113741
- Heldmann, F., Berkahn, S., Ehrhardt, M., and Klamroth, K., 2023, PINN training using biobjective optimization: The trade-off between data loss and residual loss, *J. Comput. Phys.*, **488**, 112211, doi: 10.1016/j.jcp.2023.112211
- Hornik, K., Stinchcombe, M., and White, H., 1989, Multilayer feedforward networks are universal approximators, *Neural Netw.*, **2**(5), 359-366, doi: 10.1016/0893-6080(89)90020-8
- Jagtap, A. D., Kawaguchi, K., and Karniadakis, G. E., 2020, Adaptive activation functions accelerate convergence in deep and physics-informed neural networks, *J. Comput. Phys.*, **404**, 109136, doi: 10.1016/j.jcp.2019.109136
- Junior, A. C. N., Almeida, J. D. S., Quiñones, M. P., and de Albuquerque Martins, L. S., 2019, Physics-based machine learning inversion of subsurface elastic properties, In *81st EAGE Conference and Exhibition 2019* (Vol. 2019, No. 1, pp. 1-5). European Association of Geoscientists & Engineers, doi: 10.3997/2214-4609.201901147
- Karimpouli, S., and Tahmasebi, P., 2020, Physics informed machine learning: Seismic wave equation, *Geosci. Front.*, **11**(6), 1993-2001, doi: 10.1016/j.gsf.2020.07.007
- Karniadakis, G. E., Kevrekidis, I. G., Lu, L., Perdikaris, P., Wang, S., and Yang, L., 2021, Physics-informed machine learning, *Nat. Rev. Phys.*, **3**(6), 422-440, doi: 10.1038/s42254-021-00314-5
- Kasim, M., Watson-Parris, D., Deaconu, L., Oliver, S., Hatfield, P., Froula, D., Gregori, G., Jarvis, M., Khatiwala, S., Korenaga, J., Topp-Muggleston, J., Viezzer, E., and Vinko, S., 2020, Building high accuracy emulators for scientific simulations with deep neural architecture search, arXiv preprint arXiv:2001.08055, doi: 10.48550/arXiv.2001.08055
- Kreyszig, E., 2011, *Advanced Engineering Mathematics*, 10th Ed, Wiley.
- Kondor, R., and Trivedi, S., 2018, On the generalization of equivariance and convolution in neural networks to the action of compact groups, In *International Conference on Machine Learning* (pp. 2747-2755). <https://proceedings.mlr.press/v80/kondor18a.html>
- Lagaris, I. E., Likas, A., and Fotiadis, D. I., 1998, Artificial neural networks for solving ordinary and partial differential equations, *IEEE Trans. Neural Netw.*, **9**(5), 987-1000, doi: 10.1109/72.712178
- Lee, H., and Kang, I. S., 1990, Neural algorithm for solving differential equations, *J. Comput. Phys.*, **91**(1), 110-131, doi: 10.1016/0021-9991(90)90007-N
- Lim, K. L., Dutta, R., and Rotaru, M., 2022, Physics Informed Neural Network using Finite Difference Method, In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)* (pp. 1828-1833), doi: 10.1109/SMC53654.2022.9945171
- Mallat, S., 2016, Understanding deep convolutional networks, *Philos. Trans. Royal Soc. A*, **374**(2065), 20150203, doi: 10.1098/rsta.2015.0203
- Mishra, S., and Molinaro, R., 2022, Estimates on the generalization error of physics-informed neural networks for approximating a class of inverse problems for PDEs, *IMA J. Numer. Anal.*, **42**(2), 981-1022, doi: 10.1093/imanum/drab032
- Moseley, B., Nissen-Meyer, T., and Markham, A., 2020a, Deep learning for fast simulation of seismic waves in complex media, *Solid Earth*, **11**(4), 1527-1549, doi: 10.5194/se-11-1527-2020
- Moseley, B., Markham, A., and Nissen-Meyer, T., 2020b, Solving the wave equation with physics-informed deep learning, *arXiv preprint arXiv:2006.11894*, doi: 10.48550/arXiv.2006.11894
- Moseley, B., Markham, A., and Nissen-Meyer, T., 2021, Finite Basis Physics-Informed Neural Networks (FBPINNs): a scalable domain decomposition approach for solving differential equations, *arXiv preprint arXiv:2107.07871*, doi: 10.48550/arXiv.2107.07871

- Paganini, M., de Oliveira, L., and Nachman, B., 2018, Accelerating science with generative adversarial networks: an application to 3D particle showers in multilayer calorimeters, *Phys. Rev. Lett.*, **120**(4), 042003, doi: [10.1103/PhysRevLett.120.042003](https://doi.org/10.1103/PhysRevLett.120.042003)
- Psichogios, D. C., and Ungar, L. H., 1992, A hybrid neural network-first principles approach to process modeling, *AIChE J.*, **38**(10), 1499-1511, doi: [10.1002/aic.690381003](https://doi.org/10.1002/aic.690381003)
- PyTorch, 2018, URL <https://www.pytorch.org>
- Raissi, M., Perdikaris, P., and Karniadakis, G. E., 2019, Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations, *J. Comput. Phys.*, **378**, 686-707, doi: [10.1016/j.jcp.2018.10.045](https://doi.org/10.1016/j.jcp.2018.10.045)
- Raissi, M., Yazdani, A., and Karniadakis, G. E., 2020, Hidden fluid mechanics: Learning velocity and pressure fields from flow visualizations, *Sci.*, **367**(6481), 1026-1030, doi: [10.1126/science.aaw4741](https://doi.org/10.1126/science.aaw4741)
- Rasht-Behesht, M., Huber, C., Shukla, K., and Karniadakis, G. E., 2022, Physics-Informed Neural Networks (PINNs) for Wave Propagation and Full Waveform Inversions. *J. Geophys. Res. Solid Earth*, **127**(5), e2021JB023120, doi: [10.1029/2021JB023120](https://doi.org/10.1029/2021JB023120)
- Rasp, S., Pritchard, M. S., and Gentile, P., 2018, Deep learning to represent subgrid processes in climate models, *Proceedings of the National Academy of Sciences*, **115**(39), 9684-9689, doi: [10.1073/pnas.1810286115](https://doi.org/10.1073/pnas.1810286115)
- Rumelhart, D. E., Hinton, G. E., and Williams, R. J., 1985, *Learning internal representations by error propagation*, California Univ San Diego La Jolla Inst for Cognitive Science. <https://apps.dtic.mil/sti/citations/ADA164453>
- Sahli Costabal, F., Yang, Y., Perdikaris, P., Hurtado, D. E., and Kuhl, E., 2020, Physics-informed neural networks for cardiac activation mapping, *Front. Phys.*, **8**, 42, doi: [10.3389/fphy.2020.00042](https://doi.org/10.3389/fphy.2020.00042)
- Shin, Y., Darbon, J., and Karniadakis, G. E., 2020, On the Convergence of Physics Informed Neural Networks for Linear Second-Order Elliptic and Parabolic Type PDEs, *Commun. Comput. Phys.*, **28**(5), 2042-2074, doi: [10.4208/cicp.OA-2020-0193](https://doi.org/10.4208/cicp.OA-2020-0193)
- Siahkoobi, A., Louboutin, M., and Herrmann, F. J., 2019, The importance of transfer learning in seismic modeling and imaging, *Geophysics*, **84**(6), A47-A52, doi: [10.1190/geo2019-0056.1](https://doi.org/10.1190/geo2019-0056.1)
- Sitzmann, V., Martel, J., Bergman, A., Lindell, D., and Wetzstein, G., 2020, Implicit neural representations with periodic activation functions, *Adv. Neural Inf. Process. Syst.*, **33**, 7462-7473, <https://proceedings.neurips.cc/paper/2020/hash/53c04118df112c13a8c34b38343b9c10-Abstract.html>
- Smith, J. D., Aizzadenesheli, K., and Ross, Z. E., 2020, Eikonet: Solving the eikonal equation with deep neural networks, *IEEE Trans. Geosci. Remote Sens.*, **59**(12), 10685-10696, doi: [10.1109/TGRS.2020.3039165](https://doi.org/10.1109/TGRS.2020.3039165)
- Song, C., Alkhalifah, T., and Waheed, U. B., 2021, Solving the frequency-domain acoustic VTI wave equation using physics-informed neural networks, *Geophys. J. Int.*, **225**(2), 846-859, doi: [10.1093/gji/ggab010](https://doi.org/10.1093/gji/ggab010)
- Sun, L., Gao, H., Pan, S., and Wang, J. X., 2020, Surrogate modeling for fluid flows based on physics-constrained deep learning without simulation data, *Comput. Methods Appl. Mech. Eng.*, **361**, 112732, doi: [10.1016/j.cma.2019.112732](https://doi.org/10.1016/j.cma.2019.112732)
- TensorFlow, 2015, URL <https://www.tensorflow.org>
- Waheed, U. B., Haghghat, E., and Alkhalifah, T., 2020, Anisotropic eikonal solution using physics-informed neural networks, *90th Ann. Internat. Mtg., Soc. of Expl. Geophys., Expanded Abstracts*, 1566-1570, doi: [10.1190/segam2020-3423159.1](https://doi.org/10.1190/segam2020-3423159.1)
- Waheed, U. B., Haghghat, E., Alkhalifah, T., Song, C., and Hao, Q., 2021, PINNeik: Eikonal solution using physics-informed neural networks. *Computers & Geosciences*, 155, 104833, doi: [10.1016/j.cageo.2021.104833](https://doi.org/10.1016/j.cageo.2021.104833)
- Wang, S., Yu, X., and Perdikaris, P., 2022, When and why PINNs fail to train: A neural tangent kernel perspective, *J. Comput. Phys.*, **449**, 110768, doi: [10.1016/j.jcp.2021.110768](https://doi.org/10.1016/j.jcp.2021.110768)
- Xu, Y., Li, J., and Chen, X., 2019, Physics informed neural networks for velocity inversion, *89th Ann. Internat. Mtg., Soc. Expl. Geophys., Expanded Abstracts*, 2584-2588, doi: [10.1190/segam2019-3216823.1](https://doi.org/10.1190/segam2019-3216823.1)
- Yilmaz, Ö., 2001, *Seismic data analysis: Processing, inversion, and interpretation of seismic data*, Society of exploration geophysicists, doi: [10.1190/1.9781560801580](https://doi.org/10.1190/1.9781560801580)
- Zhu, Y., Zabarar, N., Koutsourelakis, P. S., and Perdikaris, P., 2019, Physics-constrained deep learning for high-dimensional surrogate modeling and uncertainty quantification without labeled data, *J. Comput. Phys.*, **394**, 56-81, doi: [10.1016/j.jcp.2019.05.024](https://doi.org/10.1016/j.jcp.2019.05.024)