

수자원분야 인공지능 활용 연구 동향 및 소고



신주영
 국민대학교
 건설시스템공학부
 조교수
 jshin@kookmin.ac.kr

인공지능(Artificial Intelligence, AI)은 현재 우리 삶의 일부분이 되어 많은 변혁을 이끌고 있다. 특히, 생성형 AI를 이용하여 많은 인간의 창의성을 요구하는 문제들을 해결하고자 시도하고 있다. 예를 들어 ChatGPT는 주어진 프롬프트에 의해 대화하는 형식으로 주어진 질문에 대한 답을 생성한다. 이미지 생성 분야에서는 Midjourney가 콜로라도 주립박람회 미술대회 디지털 아트 부분에서 1등을 차지하였다. 현재 음악 분야에서는 생성형 AI를 이용하여 작곡을 하는 시도를 하고 있다.

어느 순간부터 AI는 특정 분야에 대해 인간보다 우수한 능력을 보이기 시작했다. 일반적으로 AI가 인간보다 뛰어난 능력을 보인 역사적인 사건으로 Imagenet Large Scale Visual Recognition Challenge(ILSVRC)의 2015년 결과를 많이 이야기 한다. ILSVRC는 기계학습모형을 이용하여 주어진 이미지의 사물이 무엇인지를 판독하는 정확도를 경쟁하는 대회이다. 그림 1은 매년 치루어 졌던 ILSVRC의 우승 모형의 깊이와 정확도를 나타낸 그림이다. 2012년부터 Deeplearning 모형이 우승을 차지 하였고, 2015년부터는 인간이 판독했을 시 발생하는 오차인 5%보다 낮은 오차를 보이며 교육받은 인간보다 높은 정확도를 보이는 판독모형이 개발되었다. ILSVRC는 2017년 이후로는 대회를 개최하는 의미를 상실하였기 때문에 더 이상 대회가 개최가 되지 않았다.

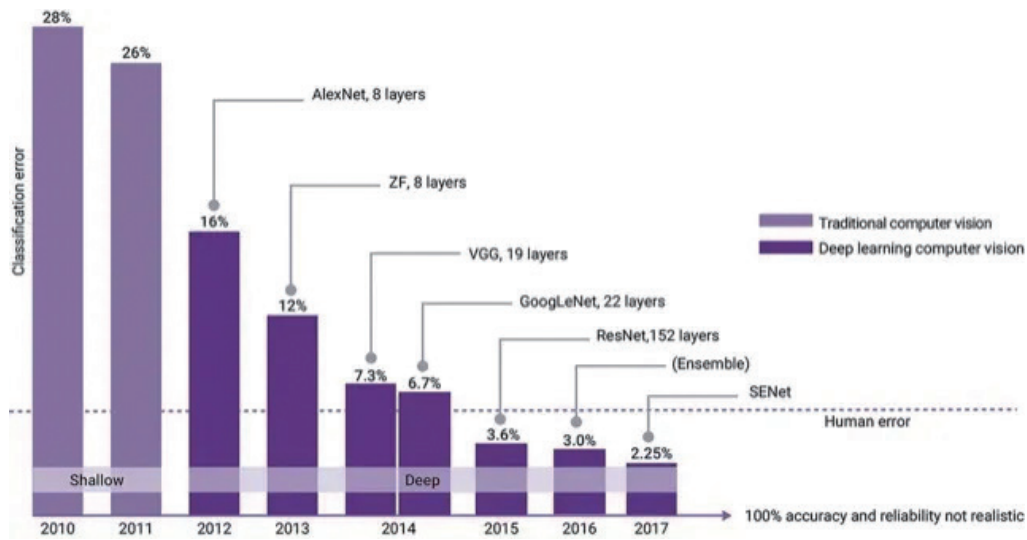


그림 1. Imagenet Large Scale Visual Recognition Challenge(ILSVRC)의 우승 모형의 정확도(<https://semiengineering.com/achieving-greater-accuracy-in-real-time-vision-processing-with-transformers/>)

AI의 획기적인 성능 개선으로 인하여 수자원분야에도 AI를 이용한 연구, 특히 Deeplearning에 대한 연구가 폭발적으로 늘어나게 되었다. 최근 자료는 아니나 Sit et al.(2020)의 조사에 따르면 최근에 합성곱신경망(Convolutional Neural Network, CNN)와 장단기 메모리 모형(Long Short Term Memory, LSTM)이 수자원분야에 활발히 적용되고 있다(그림 2).

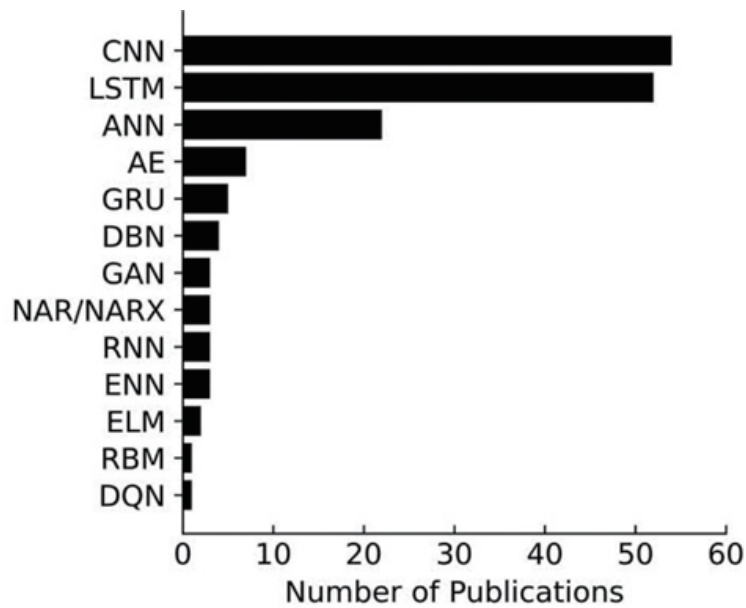


그림 2. 수자원분야 학술 논문에 사용된 AI 알고리즘의 숫자(Sit et al., 2020)

AI를 수자원분야에 활용할 경우 다양한 이익이 발생하는데, Richards et al.(2023)은 AI 활용으로 발생하는 전반적인 이점을 그림 3과 같이 정리하였다. 유역수준에서는 물공급, 유역내 수자원관리 및 긴급상황에 대한 대응에서 AI이 적용이 이점을 얻을 수 있고, 관망수준에서는 물분배와 수공구조물 설계, 운영과 유지보수에서 AI를 적용하는 것이 큰 이점을 얻을 수 있을 것으로 예상하였다. 마지막으로 사용자(end-user)수준에서는 수요 예측, 수요 관리 및 물 정의(water justice)에 대한 문제에서 AI적용이 큰 이점을 발생시킬 수 있을 것이라고 예상하였다.

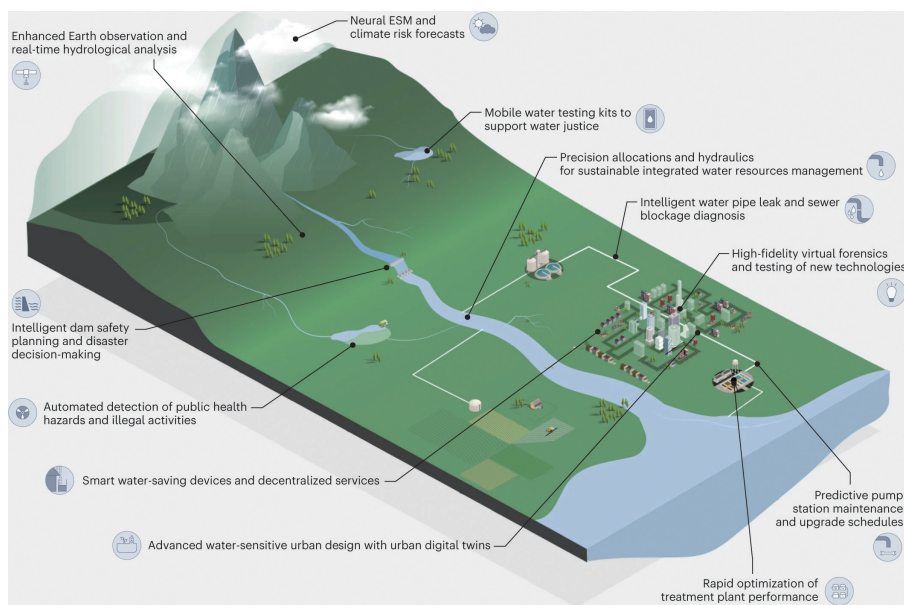


그림 3. 수자원 시스템 전반의 문제 해결을 위한 AI의 이점(Richards et al., 2023)

유역단위에서 AI적용은 빅데이터 처리, 결측자료 보완 등을 통하여 수자원의 시공간적 분포에 대하여 과거보다 정확한 정량적인 추정치를 제공할 수 있게 할 것이고, AI로 생성된 자료들은 물 스트레스를 정량화하는 데 도움을 줄 수 있다(Kadow et al., 2020; Pekel et al., 2016). 예를 들어, 원격영상(드론, 위성 등)을 이용하여 저수지 정보를 예측하는 모형에서 AI를 활용할 경우 실시간 관측, 이상 징후 감지, 수문 변수, 기상 변수의 단기 예측을 지원할 수 있다(Larson, 2022; Koppa et al., 2022; Espeholt et al., 2022). 수량적인 측면뿐만 아니라 수질적인 부분에 대하여서도 AI활용은 많은 이점을 제공할 수 있을 것이다. AI의 활용을 통하여 취수 일정을 최적화하여 저수지의 수위를 한도 내에서 유지하고 저수지 운영을 최적화하여, 상하류의 수문학적 변화에 의한 수생태의 피해 저감 가능성이

있다(Sharafati et al., 2020; Zaniolo et al., 2021). 비상사태에 대한 예방, 대비 및 대응 분야에서도 AI는 실시간 강수량 데이터를 조기 경보 시스템 및 제어 기술과 통합하여 저수지 유입을 모니터링하고 댐 원격 측정과 통신하여 안전한 여수로 방류를 관리할 수 있다(Zarei et al., 2021). 통합 물관리를 수행하려면 단기 기상 및 장기 기후 변동성과 가뭄, 폭풍 해일, 수인성 질병 유병률, 강우강도 등에 대한 영향 예측이 필요하고 이런 예측을 수행하기 위해서는 지구 시스템 모델(Earth System Model, ESM)내의 물 순환 역학에 대한 이해가 필요하다(Bauer et al., 2021). 초기 단계이지만 Neuro-ESM은 물리학에 대한 이해를 개선하고 숨겨진 매개변수를 발견하며 모의 성능을 확장할 수 있다(Irrgang et al., 2021). 이러한 개선된 예측 자료를 기반으로 최적화 알고리즘을 지속 가능한 장기 유역관리 및 인프라 계획에 활용할 수 있을 것이다(Hess et al., 2022). 유역의 물 흐름 경로, 유속, 홍수 추적, 조수위 등과 관련된 AI기반의 수리학적 모델을 통해 하천 엔지니어링, 댐, 제방, 폭풍 해일 방벽 구현을 개선할 수 있다(Hosseiny et al., 2020).

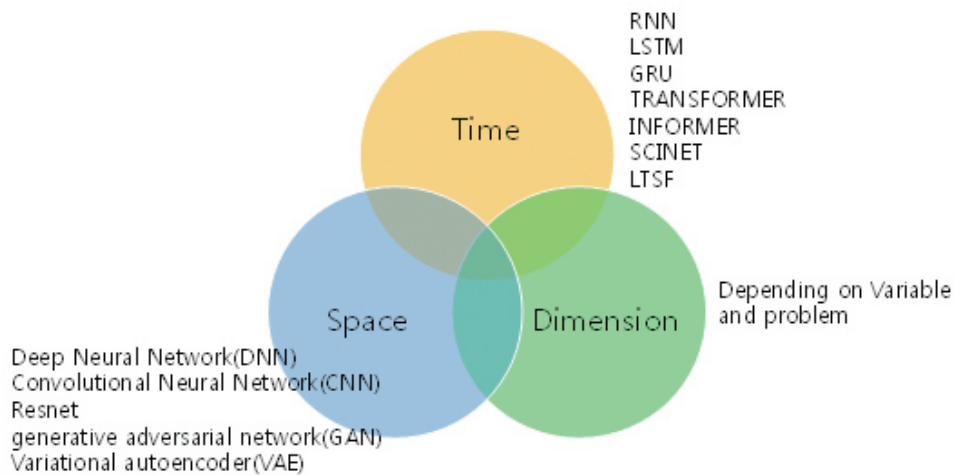


그림 4. AI활용을 통해 해결하고자 하는 수자원분야 문제의 분류

AI활용은 수자원 분야의 전반적인 부분에서 많은 이점을 가져올 것으로 기대되고 있다. 필자는 AI활용을 통해 해결하고자 하는 수자원분야 문제를 크게 세 분야로 분류하였다(그림 4). 첫 번째가 시간과 관련된 문제, 두 번째가 공간과 관련된 문제, 마지막으로 차원과 관련된 문제를 해결하기 위해 AI가 활용되고 있다고 판단하였다. 각각의 문제를 해결하기 위해 적용되는 AI알고리즘들이 다르다. 시간과 관련된 문제를 풀기 위해서는 순환 신경망(Recurrent Neural Network, RNN), LSTM, 게이트 순환 유닛(Gated Recurrent Unit, GRU), 트랜스포머(Transformer),

Informer, SCINet, Long term time series forecast(LTSF) 등이 적용되고 있다. 공간과 관련된 문제를 풀기 위해서는 합성곱 신경망(Convolution Neural Network, CNN), Resnet, Generative adversarial network(GAN), Variational autoencoder(VAE)가 폭넓게 사용되고 있다. 새로운 차원에 관련된 문제의 경우 예측 또는 분석하고자 하는 변수와 문제에 따라서 적용하는 AI알고리즘이 달라진다. 시간과 관련된 문제에서 위에서 언급된 AI알고리즘외에 다양한 AI알고리즘이 적용될 수 있으며, 공간과 관련된 문제에서도 마찬가지이다.

시간과 관련된 문제에서 자주 쓰이는 AI알고리즘들의 경우 sequential data 또는 자연어처리(natural language process, NLP)를 위해서 개발되었다. 현재는 Transformer를 이용한 large language model(LLM)의 발달로 인하여 다양한 분야에서 혁신을 이끌고 있다. 특히 Transformer는 attention과 인코딩, 디코딩 과정을 통해 단어들의 순서정보 보다는 단어간의 숨겨진 관계들을 찾아내어 문장이 갖고 있는 문맥을 잘 파악하는 장점이 있다. 그림 5은 Transformer의 구조로 복잡한 구조를 갖고 있는 것을 확인할 수 있다. 이런 이유 때문인지는 몰라도 Transformer는 기존의 sequential 자료 처리를 위한 AI알고리즘인 RNN이나 LSTM보다 수문 시계열자료 분석에서 좋은 성능을 보이는 것으로 알려졌다(Castangia et al., 2023; Xu et al., 2023).

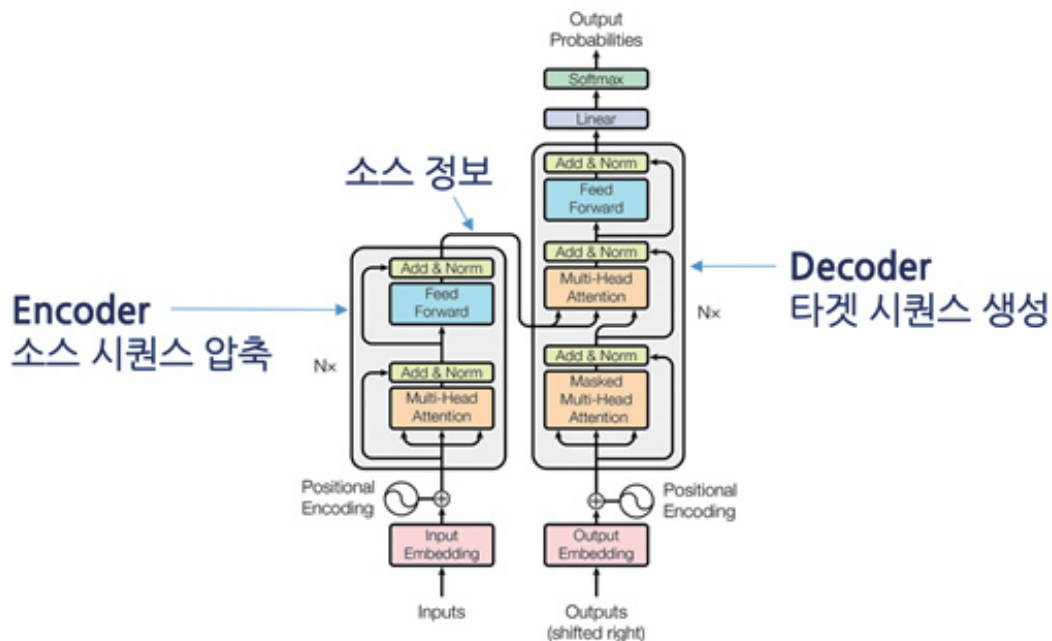


그림 5. The Transformer-model architecture (Vaswani et al., 2017).

최근에 2023년 Association for the Advancement of Artificial Intelligence(AAAI)학회에서 장기 시계열 예측을 위한 모형으로 LTSF(Zeng et al., 2023)이 제안 되었다. LTSF는 Transformer 계열의 AI알고리즘보다 시계열 예측에서 높은 정확도를 보이는 것을 확인 하였다. LTSF는 (1)Linear, (2)Dlinear, (3) Nlinear라는 세 개의 세부 모형을 갖고 있다. LTSF에서 Linear모형은 일반적인 선형 회귀와 같은 형태의 구조를 갖고 있다(그림 6). Linear모형은 기존에 수문 시계열 모델링에서 사용되어 온 Autoregressive모형과 비슷한 특성을 갖고 있다. Zeng et al.,(2023)의 연구에서도 Linear 모형은 Transformer와 같은 복잡한 구조를 갖고있는 알고리즘보다 낮은 예측 성능을 보였다. Dlinear의 구조는 그림 7에 표시되어 있다. Dlinear는 AI model안에서 자체적으로 시계열 자료를 주기성분과 경향성분으로 나누어서 예측 하고, 마지막에 다시 합하여 최종적인 예측 시계열을 생성한다. Dlinear모형에 적용된 개념은 전통적인 시계열 모델링에서 이미 적용되어 오고 있는 개념이다. 예를 들어서 Seasonal AutoRegressive Integrated Moving Average(SARIMA)의 경우 계절(주기)에 대한 차수와 경향성에 대한 차수가 포함되어 있어, 기존 시계열자료에서 주기성과 경향성을 제거한 후 모델링을 하고 추후에 주기성과 경향성을 복원하여 예측자료를 생산한다. 마지막 Nlinear의 경우 입력자료 중 최근값을 전체자료에서 빼다음 Linear 모형으로 학습 하고, 그 결과에 최근값을 다시 더해서 최종 예측값을 생산한다. Nlinear의 경우 AutoRegressive Integrated Moving Average(ARIMA)과 모델링 철학이 비슷한 것을 확인 할 수 있다. 이 연구결과는 기존의 시계열분석에서 사용되어 온 전통적인 방법론들이 최신 AI기법들과 비교하여도 그 성능이 부족하지 않다는 것을 증명한다. 현재 시계열 예측에 대한 State Of The Art(SOTA) 알고리즘으로 LTSF가 선정된 것을 확인할 수 있다(<https://paperswithcode.com/sota>).

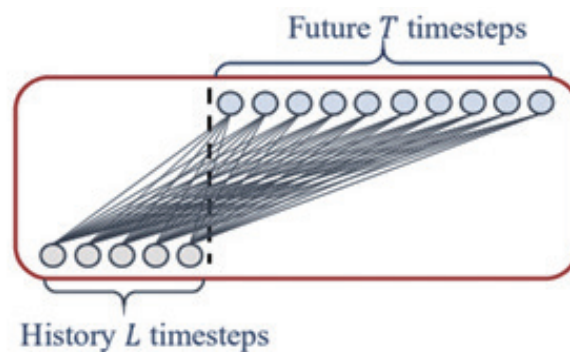


그림 6. LTSF에서 Linear 모형의 구조(Zeng et al., 2023)

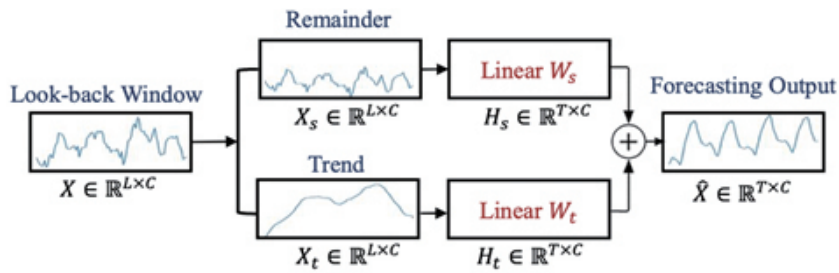


그림 7. LTSF에서 Dlinear모형의 구조

LTSF의 구조를 보면 Transformer와는 비교도 할 수 없을 정도로 단순한 형태를 갖고 있다. 모형 구조의 복잡도만 보았을 때, Transformer가 LTSF보다 높은 예측 정확도를 보이는 것이 당연하다. 이 연구 결과를 토대로 우리는 AI알고리즘을 활용할 때 알고리즘에 대한 이해가 얼마나 중요한지를 유추해 낼 수 있다. 위에서 언급했듯이 Transformer는 자연어처리를 위해 개발된 알고리즘으로 시계열 모델링을 위해서 개발된 알고리즘이 아니다. 모형 구조의 복잡성으로 인해 Transformer가 기존의 자연어처리모형(RNN, LSTM, GRU 등)보다는 시계열자료 모델링에서 더 좋은 성능을 보이는 것이 가능했으나, 시계열자료는 시간 순서에 따른 수치화된 자료 수열이지 자연어와 같은 언어로써의 자료가 아니기 때문에 Transformer가 갖고 있는 잠재적인 성능을 제대로 발휘하지 못한다고 판단된다. 즉, 최신 AI알고리즘을 바로 수자원분야에 활용한다고 했을 때 AI알고리즘이 우리가 해결하고자 하는 문제에 적절하지 않으면 우리가 기대한 만큼의 성능이 구현되지 않을 수 있다는 것이다. 즉, 수자원분야에서 AI를 활용하기 위해서는 수자원현상을 잘 이해하고 있는 AI알고리즘이 필요하다.

AI를 활용함에 있어 단순히 computer science에서 개발된 AI알고리즘을 활용하는 것도 좋지만 중요한 것은 수문 및 수리 모델링의 특성을 잘 반영할 수 있는 알고리즘을 적용되는 것이 필요하다. 전세계 많은 연구자들이 수문 및 수리 현상을 이해하는 AI알고리즘을 개발하고자 시도하고 있다. 현재는 크게 두 가지의 연구방향이 있다. 첫 번째는 수자원 분야 특화 AI알고리즘을 개발하는 것이다. Nearing et al.(2020)은 2020 NEURIPS에서 수자원분야에 적용하기 위한 LSTM기반의 Mass conservative-LSTM(MC-LSTM)을 발표하였다. Nearing et al.(2020)은 LSTM을 상태-공간 모형(state-space model)으로 생각하였다. 해당 논문은 LSTM을 수정 하여 그림 8과 같은 MC-LSTM을 제안하였다. 이 알고리즘의 구조는 수자원 분야에서 널리 사용되고 있는 저류함수와 비슷한 철학을 가지고 있다고 보면 된다. 예를들어 x_t 는 t 에서의 저류량 값을 의미하고, i_t 는 유입량, o_t 는

유출량을 의미한다. \hat{y}_t 가 실제 예측되는 유출량을 의미한다. x_t 에 대한 부분이 상태 방정식이고, \hat{y}_t 에 대한 부분이 관측방정식이라고 생각하면 된다. a_t 와 c_t 는 강우량, 토양습윤과 같은 유출 및 유입에 영향을 주는 변수들을 의미한다. 다시 LSTM으로 돌아 간다면 x_t 가 cell state, a_t 와 c_t 가 input, \hat{y}_t 가 hidden state(output)에 대응한다고 할 수 있다. 즉, MC-LSTM은 기존 LSTM의 구조적인 형태를 차용하여 질량, 에너지, 물리량 등을 보존하는 특성을 구현한 AI알고리즘이다.

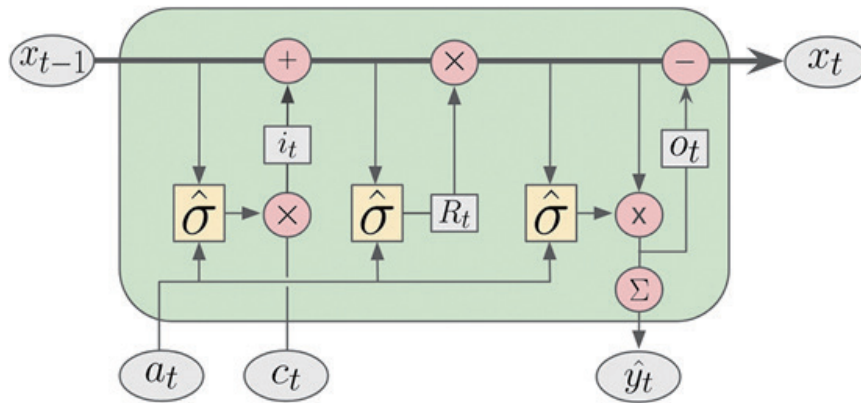


그림 8. MC-LSTM의 구조도 (Nearing et. al., 2020)

MC-LSTM이 LSTM보다 강우-유출과 같은 수문현상 모델링에 대하여 성능이 더 좋으나 하면, MC-LSTM이 LSTM보다 더 뛰어나다고 확답할 수는 없다. 현재까지 연구결과로는 MC-LSTM과 LSTM은 성능의 차이가 거의 없는 것으로 밝혀졌다. MC-LSTM이 LSTM보다 비슷한 성능을 보여서, 수자원 특화 AI알고리즘을 개발하는 것이 큰 의미가 없다고 판단할 수 있으나, 오히려 더 낮은 복잡도에서 비슷한 성능을 보이는 알고리즘을 개발되었다는 것은 큰 의미를 갖는다고 생각한다. 추후에 다양한 연구자들이 새로운 시도와 개선을 하게 된다면 수자원분야 특화 AI알고리즘의 성능이 computer science에서 제안한 기존 AI알고리즘보다 우수할 것으로 예상된다.

두 번째 연구방향으로는 Physics-informed neural network(PINN)이 있다. PINN은 AI알고리즘을 이용하여 특정 변수를 예측 할 때, 이 변수가 어떤 물리적인 법칙을 따르면 그 물리법칙의 특성을 AI모형에 주입시키는 AI알고리즘 학습전략이다. 그림 9는 AI알고리즘을 이용하여 유속과 수심을 예측하는 모형을 만들 때, Saint-Venant 방정식을 손실함수(loss function)와 결합하여, AI모형이 Saint-Venant 방정식의 특성을 반영하여, 실제값에 가까운 예측을 할 수 있도록

한다(Feng et al., 2023). 손실함수에 물리적 수리모형을 결합시켜서 AI알고리즘이 실제 관측값만을 추구하는 것이 아닌, 물리적인 특성도 같이 고려할 수 있게 하는 장점이 있다고 할 수 있다. PINN도 MC-LSTM과 같이 AI알고리즘을 바로 적용하는 것보다 우수한 성능을 보이지는 못하고 있으나, 평가 지표상으로 표현되기 어려운 모형의 불확실성을 완화한다는 측면에서 PINN이 기존 AI활용 방법론보다 우수한 성능을 보일 것으로 판단된다.

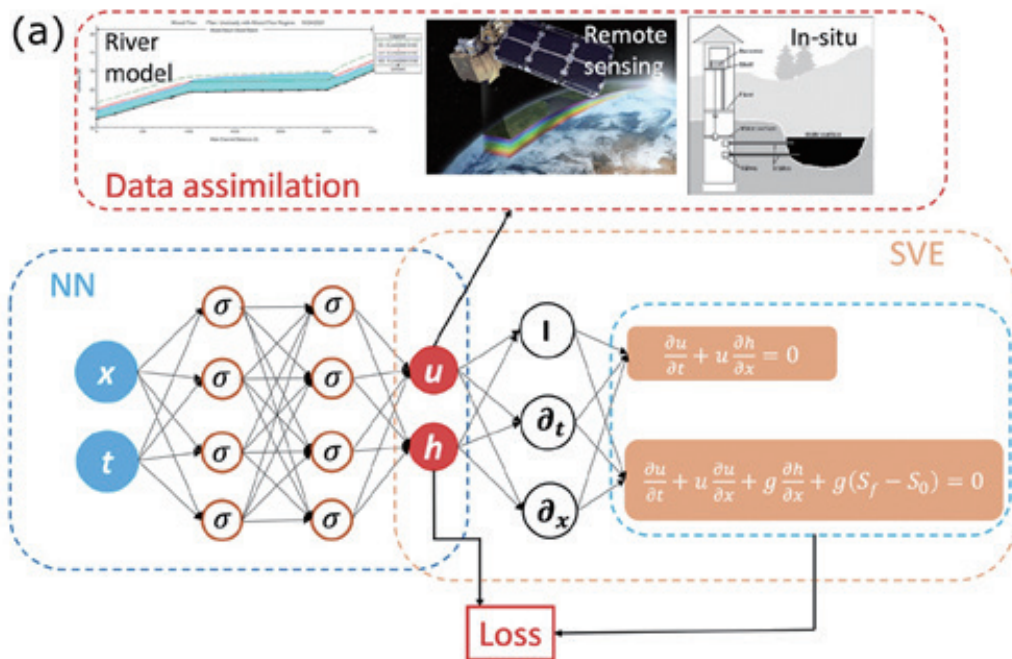


그림 9. The schematic of the physics-informed neural networks (PINN)-based data assimilation model(Feng et al., 2023).

현재도 새로운 AI알고리즘이 지속적으로 제안되고 있고, 이런 AI알고리즘을 수자원 현안에 어떻게 적용하는가에 대한 많은 고민들이 있다. 발전 속도가 매우 빠른 AI분야의 특성상 연구자가 본인의 연구 분야를 소홀히 하면서까지 AI알고리즘 공부에 매진한다면 주객이 전도되는 상황이 될 수 있을 것이다. AI의 최신기술을 활용하려 한다면, 지속적으로 새로운 AI 알고리즘에 대한 이해와 적용이 필요하다. 하지만 수자원 분야에 대한 연구를 소홀히 하면서 까지 할 필요는 없다고 판단된다. 오히려 전통적인 AI에 대한 이해도를 높여, 수문 및 수리 현상에 특화된 AI알고리즘을 개발하는 것이 수자원분야 AI활용에 대한 적절한 연구방향이 아닐까 필자 개인적으로 생각하고 추후에는 이런 AI알고리즘 개발이 AI활용 연구의 주류가 되지 않을까 예상된다.

참고문헌

- Bauer, P. et al., 2021, The digital revolution of Earth-system science. *Nat. Comput. Sci.* 1, 104-113.
- Castangia et al., 2023, Transformer neural networks for interpretable flood forecasting, *Environmental Modelling & Software*, 160, 105581.
- Espeholt, L. et al., 2022, Deep learning for twelve hour precipitation forecasts. *Nat. Commun.* 13, 5145.
- Feng, D., Tan, Z., and He, Q.Z., 2023, Physics-informed neural networks of the Saint-Venant equations for downscaling a large-scale river model. *Water Resources Research*, 59.
- Hess et al., 2022, Physically constrained generative adversarial networks for improving precipitation fields from Earth system models. *Nat. Mach. Intell.* 4, 828-839.
- Hosseiny, H., Nazari, F., Smith, V., and Nataraj, C., 2020, A framework for modeling flood depth using a hybrid of hydraulics and machine learning. *Sci. Rep.* 10, 8222.
- Irrgang, C. et al., 2021, Towards neural Earth system modelling by integrating artificial intelligence in Earth system science. *Nat. Mach. Intell.* 3, 667-674.
- Kadow, C., Hall, D. M. and Ulbrich, U., 2020, Artificial intelligence reconstructs missing climate information. *Nat. Geosci.* 13, 408-413.
- Koppa, A., Rains, D., Hulsman, P., Poyatos, R., and Miralles, D. G. 2022, A deep learning-based hybrid model of global terrestrial evaporation. *Nat. Commun.* 13, 1912.
- Larson, A., 2022, A clearer view of Earth's water cycle via neural networks and satellite data. *Nat. Rev. Earth Environ.* 3, 361.
- Nearing et al., 2020. A deep learning architecture for conservative dynamical systems: Application to rainfall-runoff modeling. In *AI for Earth Sciences Workshop at NeurIPS*.
- Pekel, J.-F, Cottam, A., Gorelick, N. & Belward, A. S. High-resolution mapping of global surface water and its long-term changes. *Nature* 540, 418-422 (2016).
- Richards, C.E., Tzachor, A., Avin, S., and Fenner, R., 2023. Rewards, risks and responsible deployment of artificial intelligence in water systems. *Nature Water*, 1, 422-432.
- Sharafati, A., Asadollah, S. B. H. S., and Neshat, A., 2020, A new artificial intelligence strategy for predicting the groundwater level over the Rafsanjan aquifer in Iran. *J. Hydrol.* 591, 125468.
- Sit et al., 2020, A comprehensive review of deep learning applications in hydrology and water resources, *Water Science and Technology*, 82 (12): 2635-2670.
- Vaswani et al., 2017, Attention is all you need. *Advances in neural information processing systems*.
- Xu et al., 2023, Transformer Based Water Level Prediction in Poyang Lake, China. *Water*, 15, 576.
- Zaniolo, M., Giuliani, M., Sinclair, S., Burlando, P., and Castelletti, A., 2021, When timing

matters-misdesigned dam filling impacts hydropower sustainability. Nat. Commun. 12, 3056.

Zarei, M. et al., 2021, Machine-learning algorithms for forecast-informed reservoir operation (FIRO) to reduce flood damages. Sci. Rep. 11, 24295.

Zeng, A., Chen, M., Zhang, L., and Xu, Q. 2023, Are Transformers Effective for Time Series Forecasting?, Proceedings of the AAAI Conference on Artificial Intelligence
