2023 SWAT Conference - Aarhus, Denmark

Friday, 30 June / Session H1: Hydrology

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water Conditions

Using Deep Learning and Physics-based Modeling

Soobin Kim, Eunhee Lee, Hyoun-Tae Hwang, Jongcheol Pyo, Daeun Yun, Sang-Soo Baek*, and Kyung Hwa Cho*

CONTACT

Soobin Kim
Ulsan National Institute of Science Technology, Republic of Korea
School of Urban and Environmental Engineering
Water-Environmental Informatics Laboratory (WEIL), Bldg. 113-405
Email: skim.env@gmail.com / Phone: +82-10-2584-6940

Soil & Water Assessment Tool SWAT

WEIL @UNIST

UNIST ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY 2009

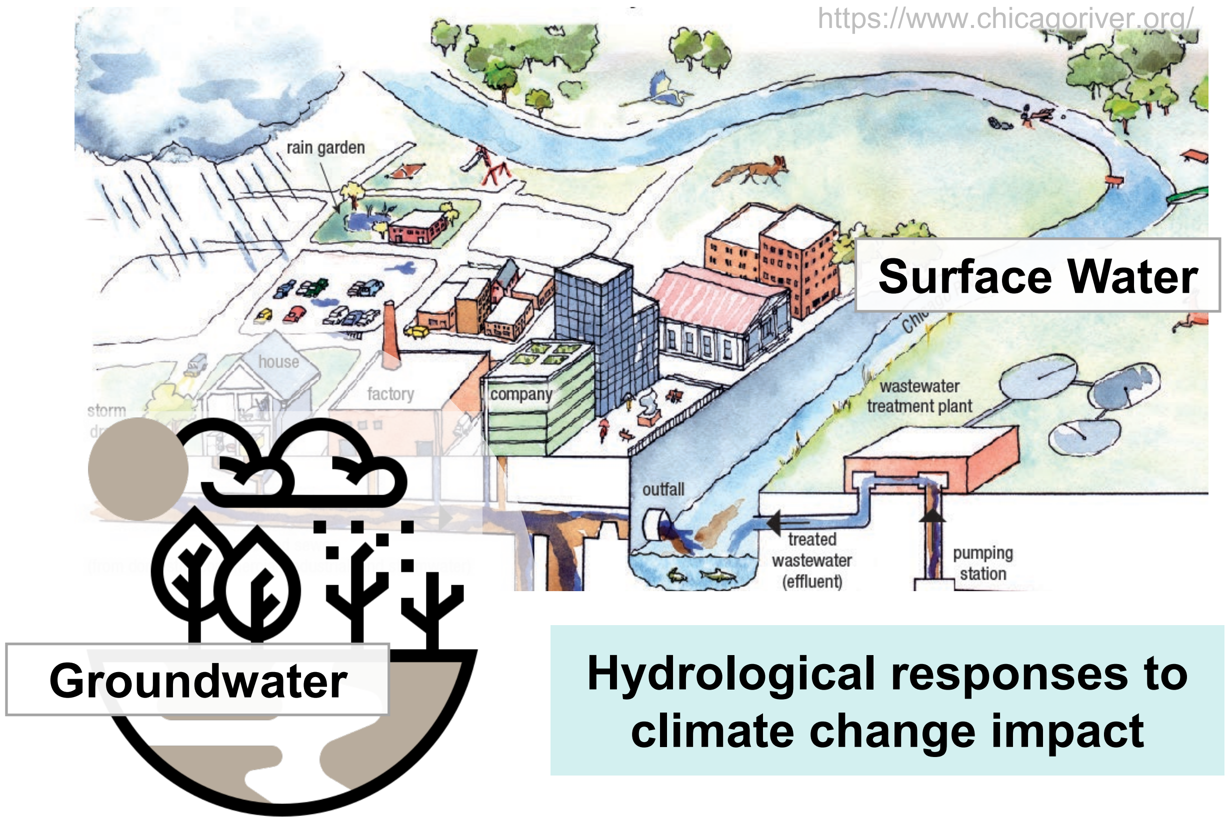
Contents

2023 SWAT Conference - Aarhus, Denmark / Soobin Kim

INTRODUCTION

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water Using Deep Learning and Physics-based Modeling

Water quantity concerns



Urbanization and **climate change** pose a risk to drinking water resources, aquatic ecosystems, public health, and the economy

To minimize the impacts, government agencies and research groups have developed advanced systems for emergency response, early warning, and water quality/quantity management

- **South Korea** operates the Water Pollution Control Information System, utilizing a hydrodynamic/water quality model (Kim, J. et al., 2022; Mun et al., 2012)
- **China** has established a rapid emergency response framework for detecting and removing water pollutants (Zhang, X.-j. et al. (2011))
- **Rui et al. (2015)** developed an emergency response system by integrating hydraulic and water quality models, and GIS

Modeling approaches

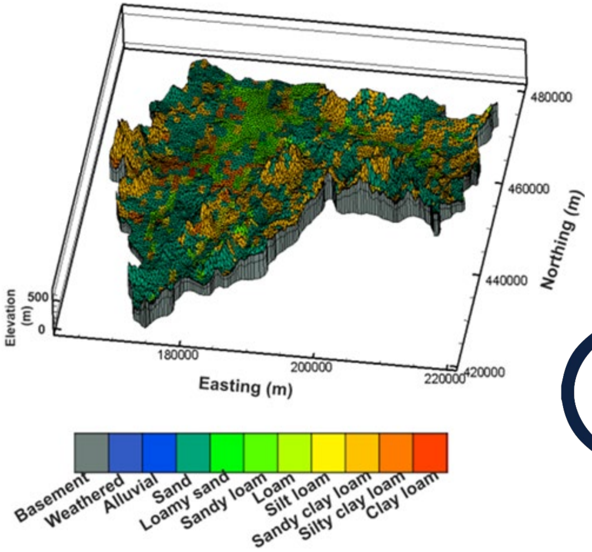
Provide valuable **spatiotemporal information on hydrological conditions and water quality**, which **supports decision-making** regarding flushing and dilution activities, vulnerability mapping, as well as risk assessment

(Guzman et al., 2015; Zhou et al., 2013; Choi et al., 2014; Martin et al., 2004)

Main modeling approaches

Process-based (Physics-based)

- A **mathematical representation** of the **environmental processes**
- Simulating spatial-temporal variations of water quantity/quality variables **by solving numerical solutions**
  - Watershed model (e.g., SWAT)
  - Integrated surface-subsurface models (e.g., HGS)
  - Urban catchment models (e.g., SWMM)
  - Hydrodynamic/water quality models (e.g., EFDC-NIER)



[E. Lee, Lee, Park, Hwang, and Park (2023)]

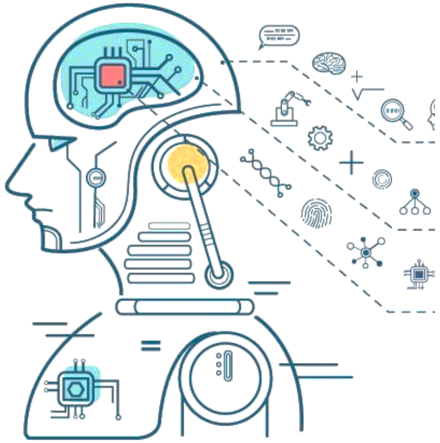
Complex Hydrologic/Hydrodynamic/ Water quality processes

Computationally expensive & Data-intensive

Data-driven (Deep learning)

- A subset of machine learning inspired by the structure and function of the human brain
- Deep learning (DL) algorithm **adjusts and fits itself using given data,** through **multiple processing layers,** and it allows the model to make predictions
- Outperforms in **processing complex data** (e.g., video and image) and **performing extensive computations**
  - Convolutional neural network (CNN)
  - Long short-term memory (LSTM)
  - Graph neural network (GNN)

To address the limitations of complex process-based modeling, …



Objective and approaches

Main objective	To evaluate **the applicability of deep learning (DL)** to simulate **spatiotemporal changes in water quantity**
Problem statement	**1. Most previous studies** have focused on **temporal hydrologic simulations using DL** 2. In previous studies, the **spatial resolution of DL models needs improvement** **3. Acquiring high-resolution data is challenging** due to high operational and labor costs
Approach	1. Employing **convolutional neural networks (CNNs)** to simulate hydrologic conditions in a high spatial resolution 2. Synthesizing **high-resolution spatial data** based on the simulation results generated by **a fully distributed hydrologic model**

2023 SWAT Conference - Aarhus, Denmark / Soobin Kim

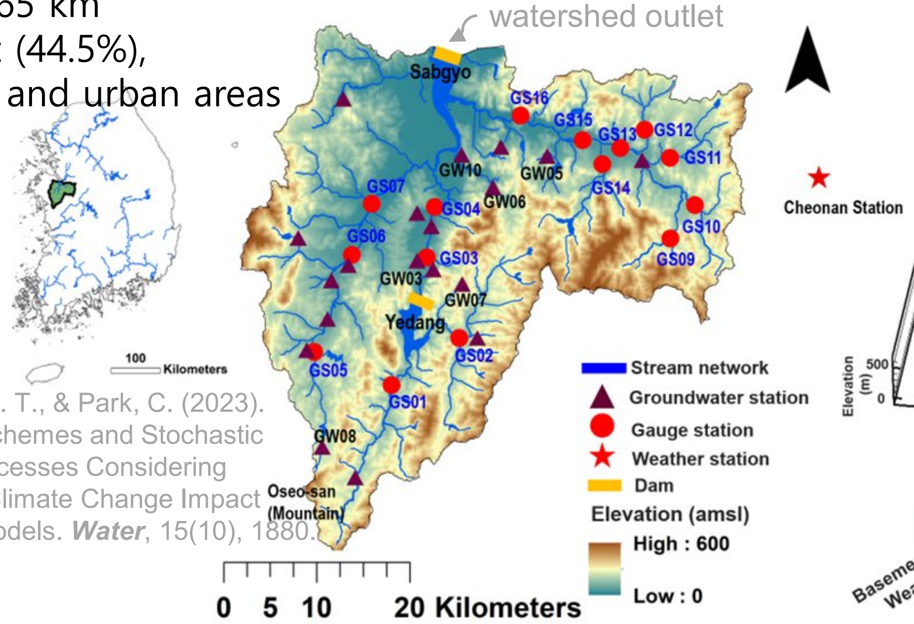
Materials and methods

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water Using Deep Learning and Physics-based Modeling

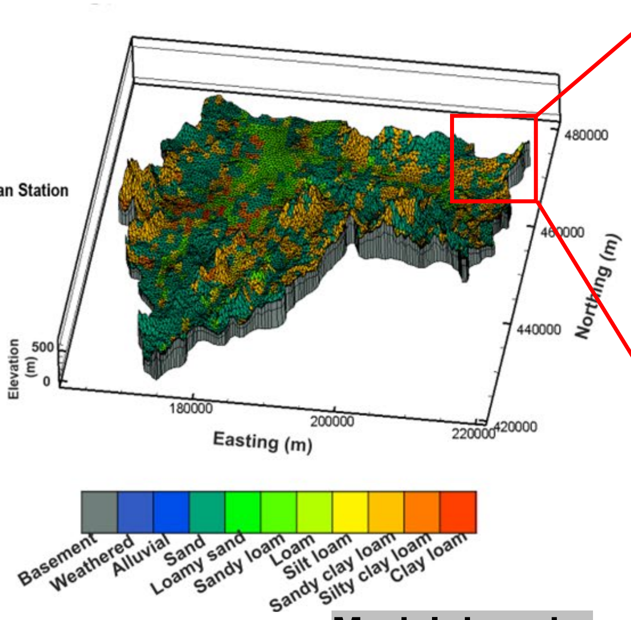
Study site & Physic-based model description

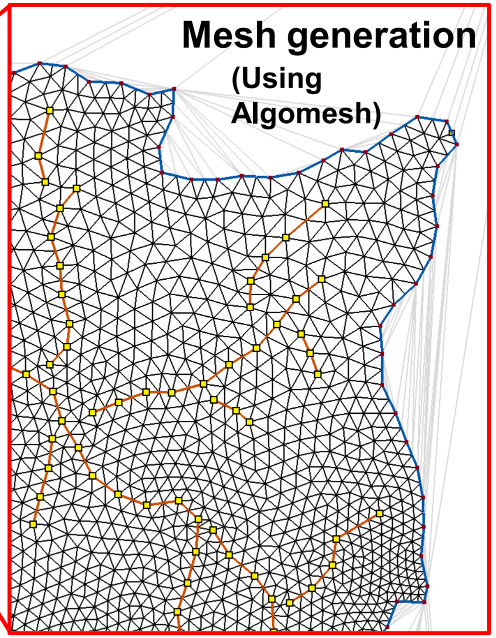
▶ Sabgyo Stream Watershed (SSW)

- Midwestern part of South Korea (36.395796–36.911621° N, 126.596445–127.213928° E)
- A drainage area: 1650 km^2
- A stream length: 65 km
- Covered by forest (44.5%), cropland (42.9%), and urban areas



Lee, E., Lee, H., Park, D., Hwang, H. T., & Park, C. (2023). Application of Different Weighting Schemes and Stochastic Simulations to Parameterization Processes Considering Observation Error: Implications for Climate Change Impact Analysis of Integrated Watershed Models. *Water*, 15(10), 1880.





[E. Lee, Lee, Park, Hwang, and Park (2023)]

Model domain:
3D domain: 89530 nodes&156420 elements
2D domain: 17380 nodes (a mean length of 420 m)

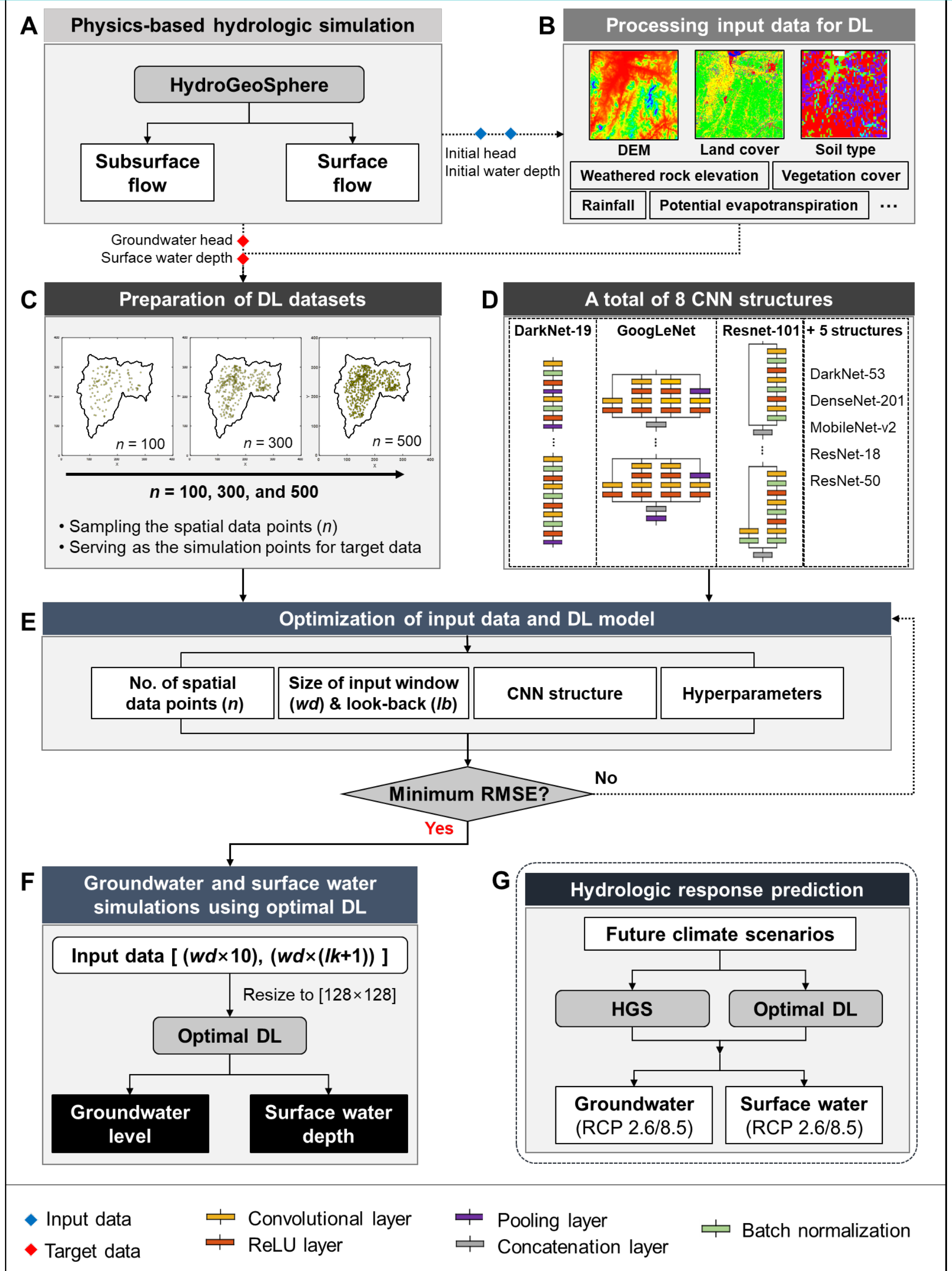
Model period: 2012-2018 (monthly)

Model calibration:
Groundwater level & surface water discharge by varying hydraulic conductivities

▶ HydroGeoSphere (HGS)

- A fully distributed, surface-subsurface integrated hydrologic model for watershed simulations
- Based on topological, geographical, geological, and meteorological datasets such as elevation, land cover, soil type, geology, and rainfall

Deep learning model for hydrologic simulations



▸ **Model implementation** involves **six steps:**

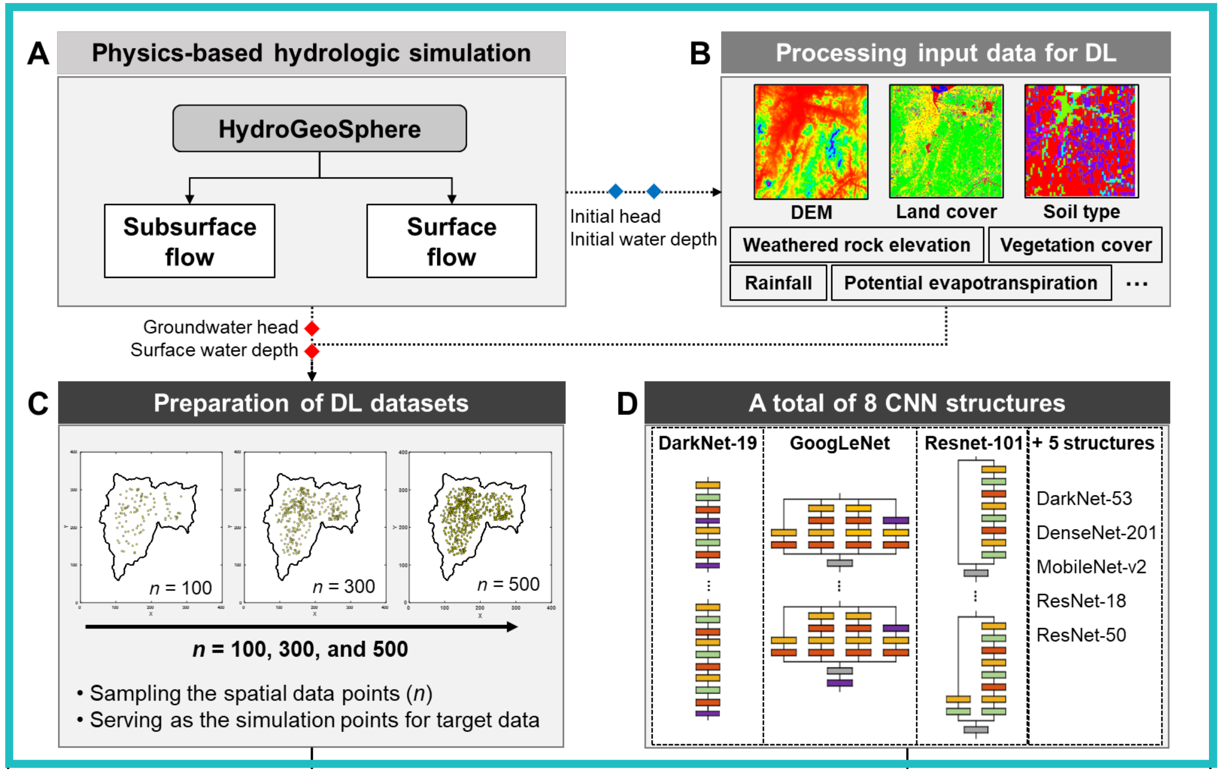
(A) Hydrologic simulation using the HGS model (obtained from Lee et al., 2023)
(B-C) Preparation of DL datasets
(D) DL model setup
(E) Optimization of the input data and DL model
(F) Hydrologic simulation using the optimal DL model
(G) Future hydrologic prediction

(using MATLAB software)

▸ List of input data for DL

Data type	Component	Source
Topographic data	Digital elevation model	National Geographic Information Institute
	Weathered rock elevation	1) Geotechnical Information DB System; 2) National Groundwater Information Center
Morphologic data	Stream network	Water Resources Management Information System
	Watershed boundary	
Geographic data	Land cover	Ministry of Environment
	Vegetation cover	National Institute of Agricultural Sciences
	Soil type	National Institute of Agricultural Sciences
Hydrometeorological data	**Rainfall**	Korea Meteorological Administration
	Potential evapotranspiration	Calculated using the simplified FAO Penman-Monteith equation (Valiantzas, 2006)
Hydrologic data	Initial hydraulic head	SSW model simulation (Lee et al., 2023)
	Initial water depth	SSW model simulation (Lee et al., 2023)
Temporal data	**Data time (month)**	-

Data processing for DL model



▶ **Data processing for the DL model** involved **three steps:**

(1) Converting the **unstructured SSW model outputs** (triangular) into a **gridded format (200 m resolution)** suitable for CNN algorithm (using a 'linear' interpolation method (MATLAB, R2022b))

(2) Extracting spatiotemporal information from the topological, morphological, geographical, hydrometeorological, and hydrologic data within the study site

(3) Transforming the input data into a 2D representation

For the DL modeling,

- **monthly results from 2012 to 2018** were utilized
- data from **2013 to 2018** were used for **training & validation**
- data from **2012 served as the look-back (*lb*)** period

Simulation target:

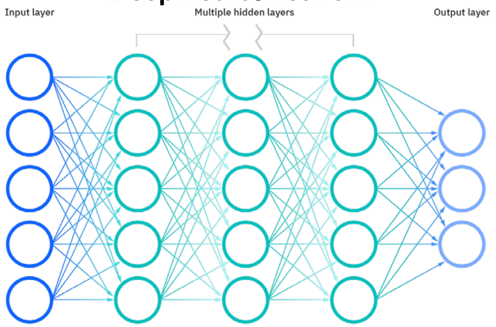
- Groundwater head & Surface water depth derived from the SSW model
- Considered the simulation results as a ground truth of DL

Convolutional neural networks (CNNs)

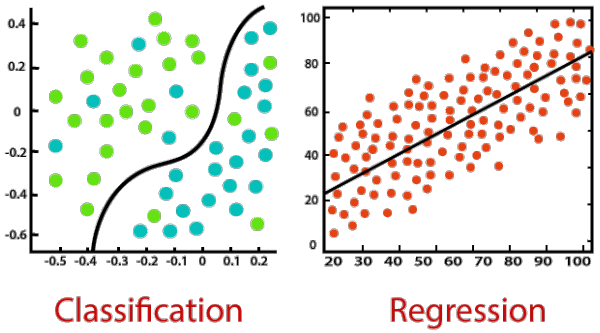
Data preprocessing	Model building & training	Model test & prediction

- **Meteorological data**
  - Temperature
  - Precipitation
- **Hydrologic data**
  - Water level
  - Water velocity
- **Geographic data**
  - Land cover
  - Soil type

Deep neural network

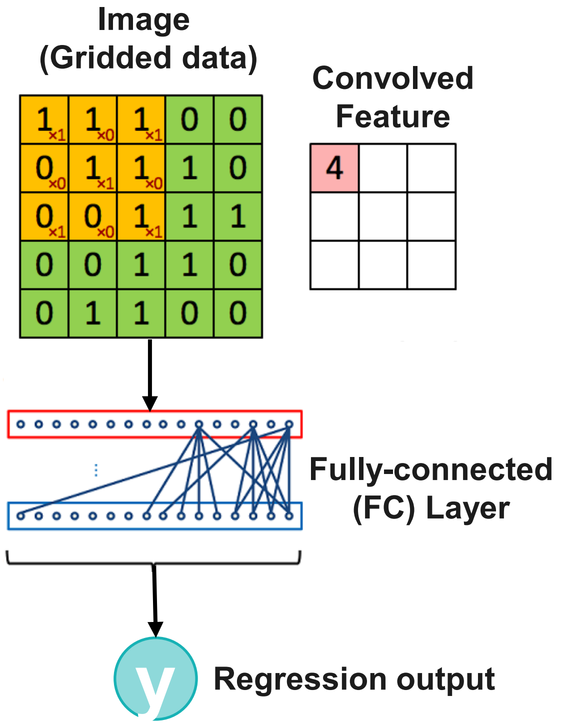


Source: IBM



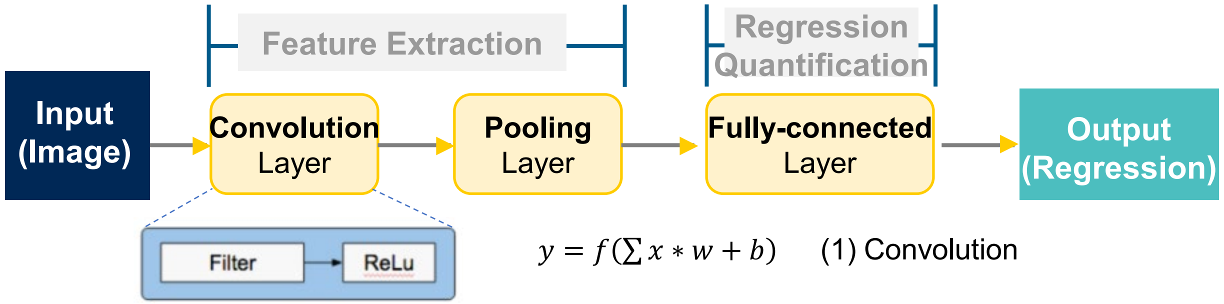
Source: Javapoint

CNN algorithm



▸ Convolutional neural networks (CNNs)

Promising DL technique for **multi-dimensional data processing** to extract **spatial features** using **convolutional filters** [LeCun, Bengio, & Hinton, 2015; Deng et al., 2009]



$y = f(\sum x * w + b)$ (1) Convolution

$z = \max(0, x)$ (2) Pooling (Max pooling)

- **Hussain et al. (2020)** used a 1D-CNN to predict streamflow for daily, weekly, and monthly forecasting
- **Pyo et al. (2020)** identified a potential of CNN model for short-term prediction of harmful algae in river
- **Xia et al. (2023)** proposed a residual dense CNN for groundwater contamination source identification

2023 SWAT Conference - Aarhus, Denmark / Soobin Kim

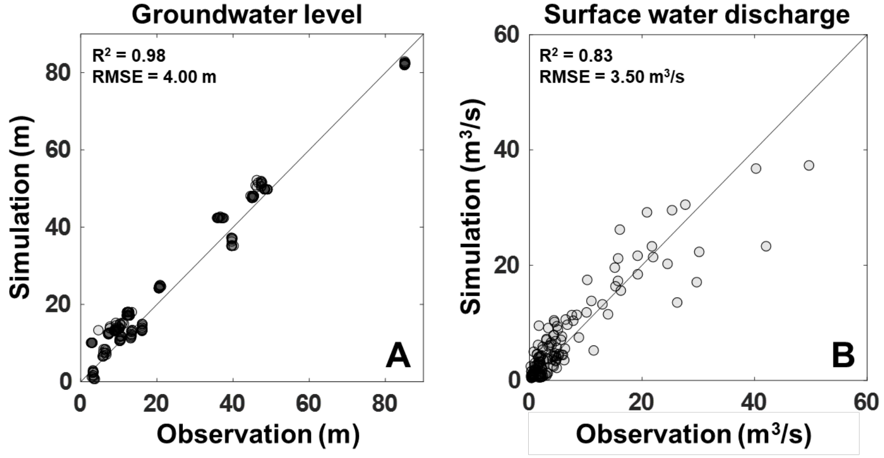
Results and discussions

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water Using Deep Learning and Physics-based Modeling

Groundwater and surface water simulations

Results of HGS model calibration

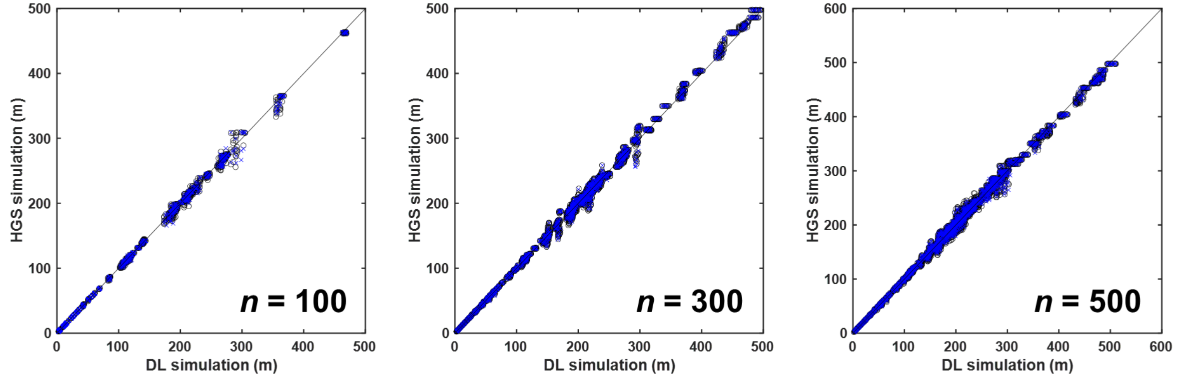
based on observed data prepared under monthly normal conditions



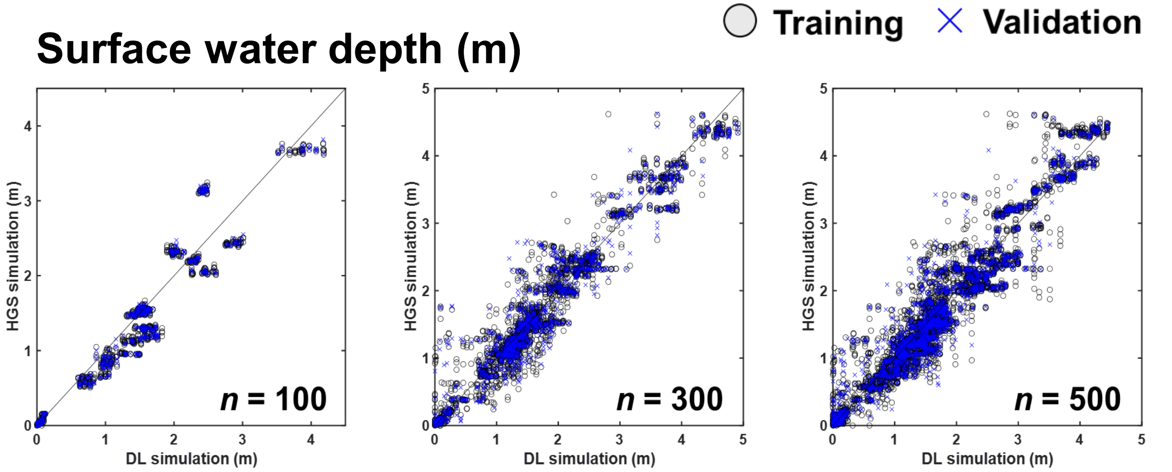
Results of CNN model training and validation

CNN performance with respect to the number of spatial data points (n)

Groundwater head (m)



Surface water depth (m)



n = 100, 300, and 500 were selected as simulation point among 41703 grid cells (within the watershed)

HGS model parameters

Parameter	Value	Source
Manning's roughness coefficients¶	0.0016 (Urban) – 0.03 (Forest)	Chow (1959); H.-T. Hwang et al. (2021)
Rill storage height¶	2.0×10^{-5} (urban) – 5.0×10^{-3} (Wetland)	H.-T. Hwang et al. (2021); Panday and Huyakorn (2004)
Obstruction storage height¶	1.0×10^{-5} (Urban) – 5.0×10^{-3} (Wetland)	H.-T. Hwang et al. (2021); Panday and Huyakorn (2004)
Hydraulic conductivity¶¶* (m/s)	1.0×10^{-10} (Basement rock) – 2.60×10^{-2} (Sand)	H.-T. Hwang et al. (2021)
Specific storage¶¶	1×10^{-4}–5.0×10^{-4}	Freeze and Cherry (1979)
Porosity¶¶ (-)	0.05 (Basement rock) – 0.35 (Soil)	Freeze and Cherry (1979)
Leaf area index (LAI)	0.26–2.9 (Vegetation) – 0.45–4.1 (Deciduous)	Myneni et al. (2015)

¶ indicates the surface flow parameters; ¶¶ indicates the subsurface flow parameters;
* indicates the calibration parameter.

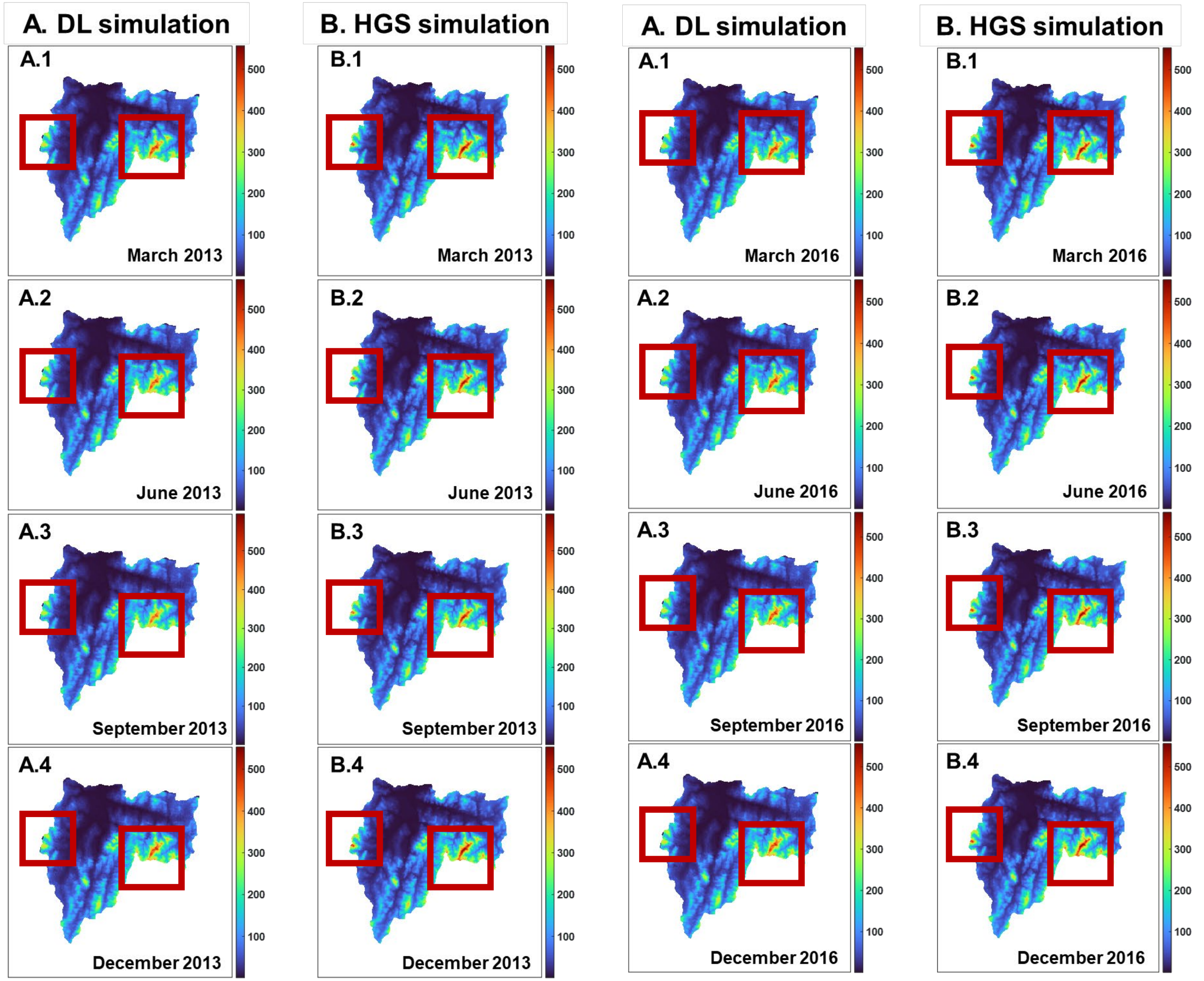
CNN model performance and optimal sets

▶ **Results of model optimization : A) groundwater level (m) and B) surface water depth (m)**

	Spatial	Input design		CNN type	Hyperparameters			Model performance						
								RMSE (m)			R^2			
	n	*wd*	*lb*	Structure	Mini-batch size	Optimizer	Learning rate	Tr¶	Val¶¶	Map¶¶¶	Tr¶	Val¶¶	Map¶¶¶	
A	100	15	12	MobileNet-v2	127	Adam	0.001	2.18	2.43	38.82	1.00	1.00	0.74	
	300	15	12	ResNet-101	130	RMSProp	0.001	2.98	3.12	20.89	1.00	1.00	0.91	
	500	**15**	**12**	**ResNet-101**	**130**	**RMSProp**	0.001	3.01	3.03	**19.63**	1.00	1.00	**0.93**	Optimal DL
B	100	15	12	ResNet-50	256	Adam	0.003	0.14	0.14	0.44	0.97	0.97	0.06	
	300	15	12	ResNet-101	132	Adam	0.001	0.13	0.14	0.41	0.98	0.98	0.17	
	500	**3**	**12**	**ResNet-50**	**256**	**RMSProp**	0.005	0.15	0.15	**0.32**	0.97	0.97	**0.47**	Optimal DL

- **Optimized** the number of spatial data points (*n*), input window (*wd*) size, look-back (*lk*) size of input data, CNN structure, and hyperparameters
- Investigated **DL model performance** based on **the amount of spatial information**
- **ResNet models** showed the highest model performance among eight CNN structures

Spatiotemporal mapping results (groundwater)



▶ Groundwater heads

Simulations results simulated by the A) optimal CNN and B) HGS models

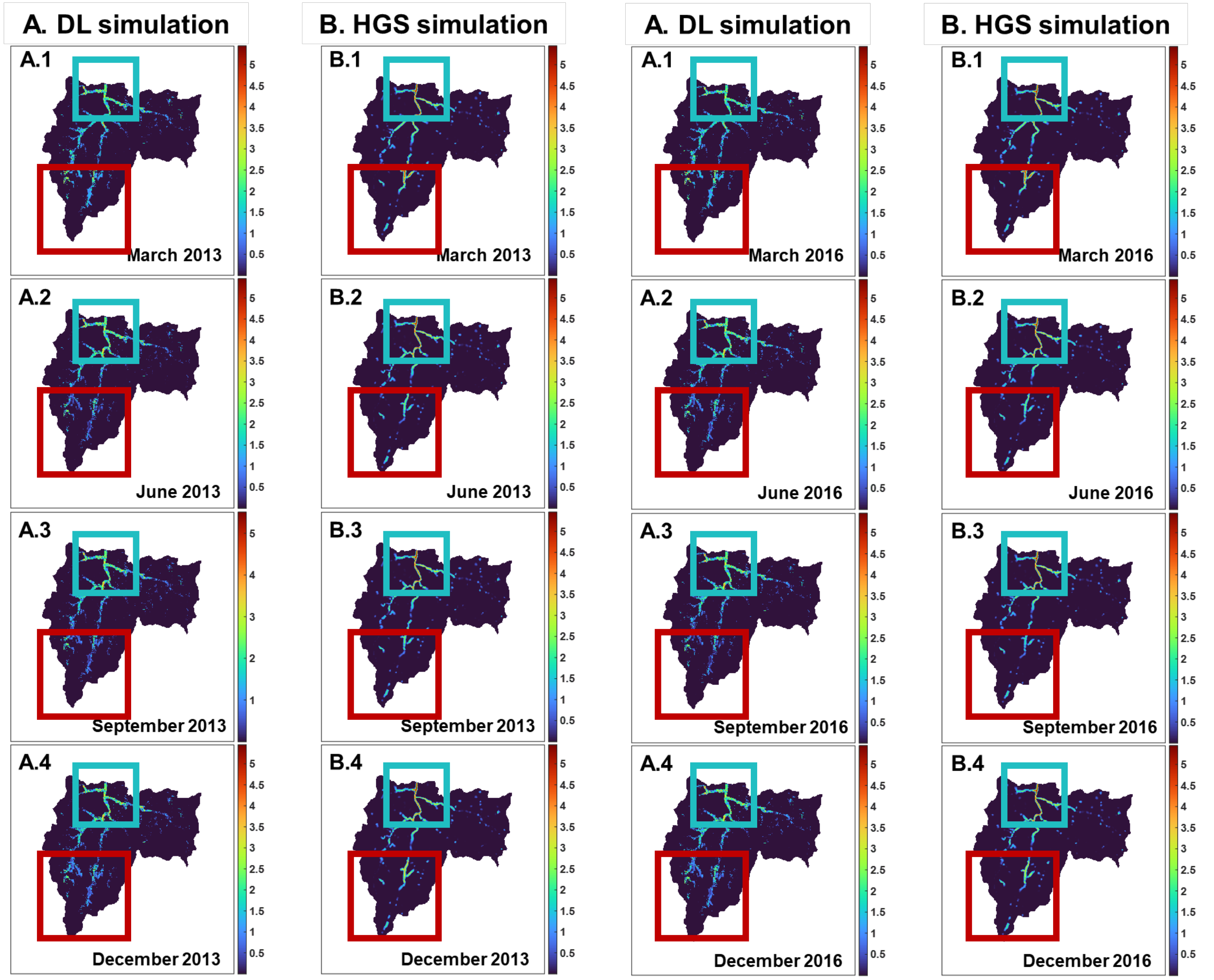
The **average RMSE and R^2 values** of **19.63 m** and **0.93**, respectively (during 2013-2018)

- Underestimated the groundwater head in **the mountainous areas of the watershed**

It might be because of the **uncertainties in data processing**

e.g., the conversion of **the HGS data from triangular mesh** into **gridded data for the CNN input**

Spatiotemporal mapping results (surface water)



▶ Surface water depth

Simulations results simulated by the A) optimal CNN and B) HGS models

The **average RMSE and R^2 values** of **0.32 m** and **0.37**, respectively (during 2013-2018)

- **Underestimated** the surface water depth **near the watershed outlet**
- **Estimated more waterbodies in the upstream region** than the HGS model

This was because **data close to the watershed boundary** were **not included in the input dataset**, resulting in a **low prediction performance at the boundary.**

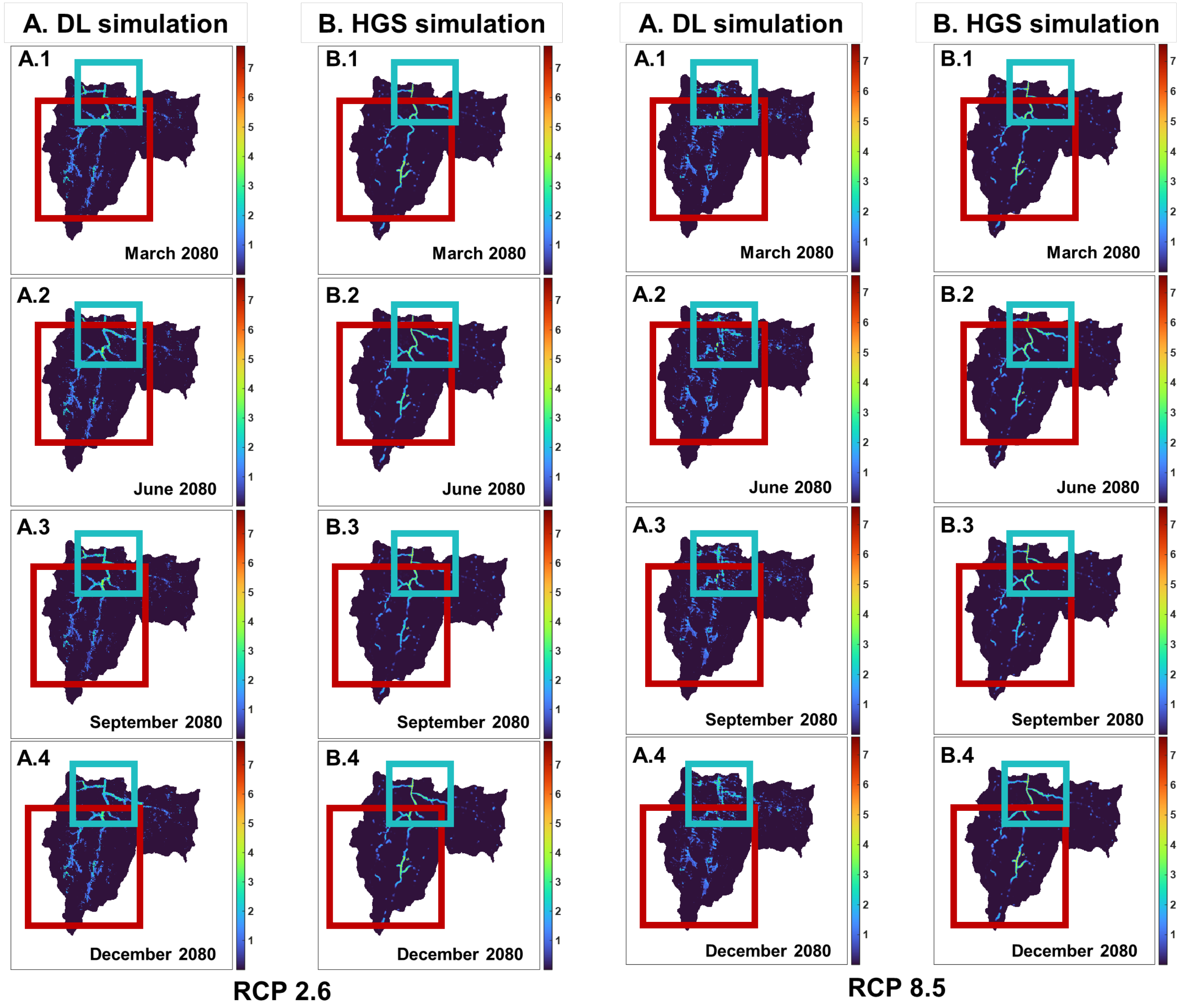
Simulation time:

HGS: 6.164 h (4 CPUs)

CNN: 0.138 h (1 GPU)

→ **44.54 times reduced!**

Future hydrologic response predictions using optimal DL



▸ **Projection of climate change in the study site**

Climate component		RCP 2.6	RCP 8.5
Monthly rainfall (mm)	2020s	99.28±117.07	92.33±109.70
	2080s	96.35±111.74	103.27±127.92
Mean temperature (°C)	2020s	12.77±9.51	12.51±9.55
	2080s	13.30±9.46	15.95±9.80
Relative humidity (%)	2020s	71.98±6.87	71.47±6.67
	2080s	71.70±6.38	72.84±6.35
Mean wind speed (m/s)	2020s	2.72±0.44	2.72±0.44
	2080s	2.73±0.43	2.66±0.44

2020s: 2011–2040 & **2080s**: 2071–2100

Using an optimal DL, spatiotemporal maps of predicted surface water depth under RCP 2.6, 8.5

The **DL model underestimated** the surface water depth **near the outlet of the watershed**

This was because **future climate conditions** were **not considered in training CNN model**

The **DL-based surface water predictions** particularly **deteriorated under RCP 8.5**, compared to the predictions under RCP 2.6

Conclusion

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water Using Deep Learning and Physics-based Modeling

Soil & Water Assessment Tool SWAT
WEIL @UNIST
UNIST ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY 2009

Conclusion

Conclusion	1. **Deep learning** models significantly **reduced simulation time**, compared to fully distributed physics-based model 2. **ResNet models** showed the highest model performance 3. **Combining the fully distributed model results, deep learning can** simulate both groundwater and surface water, providing a **high spatial-resolution results** 4. Our approach could be **a computational efficient method for simulating spatiotemporal changes in complex water systems**
Novelty	1. **Few studies** have applied deep learning models combined with a fully distributed hydrologic model to provide high spatial resolution results 2. Investigating deep learning model performance based on the amount of spatial information

Acknowledgment

This work was supported by Korea Environment Industry & Technology Institute (KEITI) through Advanced Technology Development Project for Predicting and Preventing Chemical Accidents Project, funded by Korea Ministry of Environment (MOE) (2022003620001).

Summary

Main findings

1) Deep learning models cost-effectively simulated the spatiotemporal groundwater/surface water conditions
2) Deep learning models significantly reduced simulation time, compared to fully distributed physics-based model

Research implications

Model Development
- Model integration
- Sophisticated deep learning model

Model Application
- HGS & CNN models for simulating hydrologic conditions

Water quality/quantity Management
- Investigating the optimal sets in hydrologic simulations
- Reducing computational costs

Publication information

Spatiotemporal Simulation of Groundwater and Surface Water Integrating Deep Learning and Physics-Based Watershed Models (Will be submitted to Water Research) Soobin Kim[a], Eunhee Lee[b], Hyoun-Tae Hwang[c,d], Jongcheol Pyo[e], Daeun Yun[a], Sang-Soo Baek[f,*], and Kyung Hwa Cho[a,g,*]

[a] School of Urban and Environmental Engineering, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea; [b] Korea Institute of Geoscience and Mineral Resources; 124 Gwahak-ro, Yuseong-gu, Daejeon 34132, Republic of Korea; [c] Aquanty, Inc., 564 Weber Street North, Unit 12, Waterloo, Ontario N2 L 5C6, Canada; [d] Department of Earth and Environmental Sciences, University of Waterloo, Waterloo, Ontario N2L 3G1, Canada; [e] Department of Environmental Engineering, Pusan National University, Busandaehak-ro 63 beon-gil 2, Geumjeong-gu, Busan 46241, South Korea; [f] Department of Environmental Engineering, Yeungnam University, 280 Daehak-Ro, Gyeongsan-Si, Gyeongbuk 38541, South Korea; [g] Graduate School of Carbon Neutrality, Ulsan National Institute of Science and Technology, 50 UNIST-gil, Eonyang-eup, Ulju-gun, Ulsan 44919, Republic of Korea

Thanks for your attention!

Enhancing Spatiotemporal Simulation of Groundwater and Surface Water
Using Deep Learning and Physics-based Modeling

CONTACT

Soobin Kim
Ulsan National Institute of Science Technology, South Korea
Department of Urban and Environmental Engineering
Water-Environmental Informatics Laboratory (WEIL), Bldg. 113-405
Email: skim.env@gmail.com / Phone: +82-10-2584-6940