Zhongrun Xiang

PhD Candidate in Civil and Environmental Engineering, University of Iowa

Supervised by Associate Prof. Ibrahim Demir at Hydro-Informatics Lab https://HydroInformatics.uiowa.edu/

zhongrun-xiang@uiowa.edu | (240) 413‑1044 | Iowa City, IA 52246

Google Scholar | LinkedIn | GitHub

Education Background

- University of Iowa, IIHR–Hydroscience & Engineering — Aug 2017 – Dec 2021 (Expected)
  Ph.D. in Civil and Environmental Engineering — GPA 3.98/4.00
- University of Iowa, Department of Computer Science — Aug 2019 – Dec 2020
  Master of Computer Science — GPA 3.98/4.00
- University of Maryland, A. James Clark School of Engineering — Aug 2015 – May 2017
  M.S. in Environmental Engineering — GPA 3.50/4.00
- Troy University, Department of Biological and Environmental Sciences — Aug 2012 – May 2015
  B.S. in Environmental Science — GPA 3.50/4.00
- China Pharmaceutical University, College of Engineering — Sep 2010 – May 2015
  B.S. in Environmental Science — GPA 80/100

Publications

- Demir I., Xiang, Z., Demiray, B. Z., Sit, M. (2021). WaterBench: A Large-scale Benchmark Dataset for Data-Driven Streamflow Forecasting. (in preparation)
- Xiang, Z., Demir, I. (2021). Flood Markup Language – A Standards-based Exchange Language for Flood Risk Communication. *EarthArXiv*. DOI: 10.31223/X50S4V
- Xiang, Z., Demir, I., Mantilla R., & Krajewski, Witold. (2021). A Regional Semi-Distributed Streamflow Model Using Deep Learning. *EarthArXiv*. DOI: 10.31223/X5GW3V
- Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., & Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology, 82*(12), 2635-2670. DOI: 10.2166/wst.2020.369
- Xiang, Z., & Demir, I. (2020). Distributed long-term hourly streamflow predictions using deep learning–A case study for State of Iowa. *Environmental Modelling & Software*, 104761. DOI: 10.1016/j.envsoft.2020.104761
- Xiang, Z., Yan, J., & Demir, I. (2020). A Rainfall-Runoff Model with LSTM-based Sequence-to-sequence Learning. *Water Resources Research*. DOI: 10.1029/2019WR025326
- Sit, M., Demiray, B. Z., Xiang, Z., Ewing, G. J., Sermet, Y., & Demir, I. (2020). A comprehensive review of deep learning applications in hydrology and water resources. *Water Science and Technology*. DOI: 10.2166/wst.2020.369
- Xiang, Z. & Demir, I. (2020). Generalized Streamflow Forecast Model Using Deep Learning. In *AGU Fall Meeting 2020 Abstracts*.
- Demir, I., & Xiang, Z. (2019). Sequence-to-Sequence Learning with Deep Neural Networks in Rainfall-Runoff Modeling in Iowa. In *AGU Fall Meeting 2019 Abstracts*.
- Demir, I., & Xiang, Z. (2018). Runoff Prediction using Long-Short Term Memory Model. In *AGU Fall Meeting 2018 Abstracts*.
- Xiang, Z., Montas, H. J., Shirmohammadi, A., Leisnham, P. T., & Brubaker, K. (2018). Impact of Climate Change on Critical Source Areas in a Chesapeake Bay Watershed. In *2018 ASABE Annual International Meeting* (p. 1). American Society of Agricultural and Biological Engineers. DOI: 10.13031/aim.201801831
- Xiang, Z. (2017). *Hydrologic Response of a Suburban Watershed to Climate Models* (Master Thesis, University of Maryland, College Park). DOI: 10.13016/M22G50

Presentations

- Xiang, Z. & Demir, I. (2020). Generalized Streamflow Forecast Model Using Deep Learning. 2020 AGU Fall Meeting. (Poster).
- Xiang, Z., & Demir, I. (2020). An Application of Deep Learning-Based Streamflow Forecasting in the State of Iowa. AWRA 2020 Annual Water Resources Conference. (Oral and Poster).
- Demir, I., & Xiang, Z. (2019). Sequence-to-Sequence Learning with Deep Neural Networks in Rainfall-Runoff Modeling in Iowa. 2019 AGU Fall Meeting. (Poster).
- Demir, I., & Xiang, Z. (2018). Runoff Prediction using Long-Short Term Memory Model. 2019 AGU Fall Meeting. (Poster).
- Xiang, Z., Yan, J., Gu, Chen. 2018. Application of Long Short-Term Memory (LSTM) Model on Air Quality Forecast – Solution to KDD Cup of Fresh Air. 24th ACM SIGKDD Conference on Knowledge Discovery and Data Mining. London, UK. (Oral and Poster, the 5th place).
- Chanse, V., Xiang, Z., Rockler, A. 2016. Improving adoption of stormwater management practices: Integrating social and biophysical dimensions. Baltimore Urban Waters Partnership Meeting. Baltimore, MD. (Invited talk).

Research Experience

Streamflow and Runoff Modeling using Machine Learning and Deep Learning Approaches
Graduate Research Assistant in University of Iowa   Sep 2017 – Present
- Developed geospatial-related data-driven rainfall-runoff and streamflow forecast models for both short-term and long-term predictions based on the deep learning methods including Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), sequence learning, etc.
- Statistically analyzed the model accuracy and compared to existed statistical models (multivariate regression, ridge regression and Lasso regression) and machine learning models (random forest regression and support vector machine (SVM) regression).
- Managed the radar data on the PostgreSQL database with the extension of PostGIS, and visualized the results on QGIS for flood predictions on 125 USGS gages located in the State of Iowa.
- Exploring interpreting of geospatial information from radar data with deep learning methods, i.e. Convolutional Neural Networks (CNNs), ConvLSTM, and Graph Neural Networks (GNNs).
- Exploring the Predictions in Ungauged Basins (PUB) by developing generalized models using deep learning approaches.
- Establishing the real-time streamflow forecast system using deep learning approaches with the data integration of the real-time streamflow and weather forecast data from HRRR and USGS.
- Project GitHub: https://github.com/uihilab/HydroLSTM

Design of Flood Markup Language
Graduate Research Assistant in University of Iowa   Sep 2017 – Dec 2020
- Designed the flood forecast data sharing format Flood Markup Language (FloodML).
- Developed the FloodML schema based on the XML which supporting varieties of flood related data formats.
- Using HTML, JavaScript, and PHP, developed the real-time sample data converter from third-party organizations (i.e., NWS) and different data types (i.e., stream data and alerts) to FloodML.
- Developed the database schema for storing the data in a sample flood data-sharing platform.
- Project GitHub: https://github.com/uihilab/FloodML

Modeling and Visualization of Spatially Targeted Social Interventions to Different BMPs

Graduate Assistant Researcher in University of Maryland　　　　　　　　　　May 2017 – June 2017
- Assisted the social groups and analyzed the social willingness of the Best Management Practices (BMPs) including the native landscaping, fertilizer reduction, and infiltration trench.
- Visualized the SWAT model results, BMPs suggestions, and social willingness based on OpenLayers and Google Map API.
- Assisted preparing the EPA grant project report "Final Report: Sustainable Community Oriented Stormwater Management (S-COSM): A Sensible Strategy for the Chesapeake Bay". EPA Grant Number: R835284. https://cfpub.epa.gov/ncer_abstracts/index.cfm/fuseaction/display.abstractDetail/abstract/9911/report/F

Hydrologic Response of a Suburban Watershed to Climate Models

Master's Thesis in University of Maryland　　　　　　　　　　　　　　　　Sep 2016 – May 2017
- Analyzed the difference of input climate data from Climate Forecast System Reanalysis (CFSR) and National Climatic Data Center (NCDC) on the suburban watershed Wilde Lake in Maryland.
- Established a surface water quality model with SWAT on a suburban watershed Wilde Lake in Maryland.
- Calibrated and validated with SWAT-CUP SUFI2 program with the observations in the last ten years.
- Predicted the future water quality at outlet with the consideration of six different CMIP5 climate models.
- Analyzed the potential Critical Source Areas (CSAs) based on both historical and future climate with SWAT models.
- Statistically analyzed the results by RStudio, and geographically visualized the results by ArcGIS.
- Thesis published. Xiang, Z. (2017). *Hydrologic Response of a Suburban Watershed to Climate Models* (Master Thesis, University of Maryland, College Park). DOI:10.13016/M22G50.

Course Projects

Reddit User Preference on President Election using Natural Language Processing

CS:4440 Web Mining
- Extracted and labeled the political opinions of both parties from news as the training dataset.
- Selected users and posts using regular expression for election related comments as the model input.
- Lemmatized, vectorized, and mapped the words to vectors using the Natural Language Processing (NLP) tools including the pre-trained word embedding GoogleNews-vectors-negative300.
- Developed machine learning and deep learning models using scikit-learn and TensorFlow based on the word vectors from the Reddit posts to identify the Reddit user performance on the president election.

Prototype-based Interpretable Machine Learning on Credit Card Overdraft and Streamflow Forecast

MSCI:7000 Interpretable Machine Learning
- Applied linear machine learning explainer including LIME (Local Interpretable Model-agnostic Explanations) and SHAP (SHapley Additive exPlanations) on the machine learning models.
- Designed prototype-based deep learning models to provide reasonable explains for the black-box deep learning models using TensorFlow.
- Evaluated the machine learning model performance of LIME and SHAP The interpretable model provides the same model performance with reasonable explains.
- Proposed a prototype-based interpretable deep learning model, which has 8% reduced accuracy on the streamflow forecast dataset as the tradeoff for the explanations.

Modeling of COVID-19 using Dynamic SEIRD Model in a Data-Driven Approach

CS:4980 Computational Epidemiology
- Designed an epidemiology transportation model based on SEIRD (Susceptible-Exposed-Infectious-Removed-Death) considering the COVID-19 features and policy influence dynamically.
- Applied the simulated annealing algorithm on the dynamic SEIRD model for the parameter optimization.
- Calculated and analyzed the reproduction values and R_0 from our fitted model. This model is developed in a data-driven approach, which does not require numerical assumptions.

Database Design and Implementation

CS:4400 Introduction to Database Systems

- Designed a database for an online taxi booking system like Uber. The database includes three main entities: Passenger, Driver, and Ride, and 10 entities with 60 attributes are included in the entity-relationship model.
- Implemented the database in MongoDB through a .json file and designed the authorization actions from each end users including managers, employees, drivers, and passengers.
- Executed SQLs from our database for required tasks.

Competitions

Tianchi 2021 "AI Earth" Innovation Challenge   Mar 2021
- Participation in a prediction project of Oceanic Niño Index (ONI) for the next 36 months using the past 12 years of CMIP5 Global Climate Model (GCM) data.
- As a team leader and coder, developed a CNN-LSTM model for the ONI prediction.
- Rank #33 in the final round leaderboard out of 2849 teams.

Midwest Hackathon 2018 in the University of Iowa   Oct 2018
- Developed a global Flood Index Map to evaluate the flood risk during the 2-day non-stop hackathon in a three-man team.
- Visualized the global geological data with JavaScript with D3.js.
- Won the Water Data Analytics Award out of 500 participants.

ACM Data Mining and Knowledge Discovery (KDD) competition 2018   May 2018
- Participated in a real-life air quality forecast of PM2.5, PM10, and O_3 for Beijing and London with deep learning method in a 3-person team.
- Provided insights for the developing of Long Short-Term Memory (LSTM) model for air quality modeling.
- Data crawling of the future hourly weather data from the weather forecast websites.
- Outcome 5th place in main challenge and 3rd place in long-term prediction challenge out of over 4,000 teams.

Certifications

Coursera.com Course Certificates

Practical Machine Learning	Aug 2016, License GCT42HPKR369
Regression Models	Aug 2016, License QUYGH6SDMUNG
Exploratory Data Analysis	Apr 2016, License ZDC2BZN9B8AS
Fundamentals of GIS	Apr 2016, License BAUVZNDBCLLB
Getting and Cleaning Data	Mar 2016, License BVYBCYUV7HCT
R Programming	Feb 2016, License VJ2C77SHHKL4
The Data Scientist's Toolbox	Feb 2016, License 4J3Q9JZ75HXZ
An Introduction to Interactive Programming in Python (Part 1)	Oct 2015, License 8CGMZCE8DR

Teaching Experience

Teaching Assistant of *Large Data Summer School* in University of Iowa at Summer 2018.
- Worked as the Teaching Assistant for Big Data Summer School in the sections of Introduction to R, Probability & Statistics, Learning from Data, Parallelism for Large Data and Clustering for Data Analysis.
- Mentored over 20 students in freshman and sophomore and provided individual assistance in the classroom when conducting curriculum-based group lessons.

Teaching Assistant of *ENCE 200 Fundamental of GIS* in University of Maryland at Fall 2016.
- Graded assignments, papers and final exams for the ENCE200 part II - Fundamental of ArcGIS.
- Assisted professor and answer questions from 80 students at two sections.