Extension Philosophy

Extension is one of the important pillars of land-grant universities, bridging the gap between research and practice by providing non-formal education to farmers, agricultural producers, and other stakeholders. I believe that an impactful extension program successfully creates a critical link between scientific research and actionable practices, empowering agricultural stakeholders to address pressing challenges. My extension philosophy centers on four pillars:

1. Stakeholder Engagement and Need Assessment: Extension begins with actively engaging with farmers, county and regional extension agents, and stakeholders to understand their needs and challenges. At UF, I participated in various Potato Field Days at the Hastings Agricultural Extension Center, FL and in-service trainings (IST) conducted by UF/IFAS. While interacting with farmers at field days, I observed that while Florida farmers are aware of the risks associated with extreme weather events, yet many lack in implementing appropriate preventive strategies to counter these impacts. Similarly, at UIUC, I worked with stakeholders to develop the CoverCrop Analyzer, a web-based decision-support tool for cover crop management. Meetings with various Illinois farmers, I identified barriers to adopting cover crops, such as concerns about costs, labor requirements, and limited knowledge of cover crop management. These insights guided my research to ensure its alignment with farmers' needs and expectations. As a faculty member, I plan to expand stakeholder engagement by conducting regular needs assessments through surveys, focus groups, and field visits. By creating trust-based relationships, I aim to bridge gaps in understanding climate change and facilitate the adoption of climate-resilient practices.

2. Demonstration and Knowledge Transfer: Field demonstrations are essential for showcasing the practicality of sustainable practices and tools. At UF, I assessed the impact of different combinations of nitrogen fertilizers and irrigation practices on potato agroecosystem by incorporating field trials with crop modeling. These results were shared with farmers during various field days. Similarly, my involvement in cover crop project at UIUC helped promote the environmental and agronomical benefits of adopting cover crops in Midwestern cropping systems. I envision expanding these efforts by designing collaborative field trials and tools on climate-resilient irrigation and nutrient management practices. Partnering with growers to test these solutions will provide real-world evidence of their efficacy, encouraging broader adoption.

3. Educational Outreach: Effectively disseminating knowledge is at the heart of extension work. At UF, I co-authored an extension publication on "Assessing Nitrate Leaching in Crop Production through the Application of Crop Simulation Models with Experimental Data from Florida" (Sharma et al., 2024) to provide farmers and extension agents with actionable insights for mitigating nutrient loss. Additionally, I contributed to multiple extension articles at UIUC to introduce farmers with the cover crop management tool (Coppess et al. 2020, 2021a, and 2021b). These extension publications allowed me to address specific stakeholder concerns and provide accessible, practical solutions. My future outreach efforts will prioritize developing educational content and extension articles along with disseminating information through bulletins, social media, seminars, and workshops in state/local growers' meetings and conferences.

4. **Program Evaluation:** Continuous improvement in extension programs relies on robust assessment and feedback mechanisms. While developing the CoverCrop Analyzer, I actively sought feedback from stakeholders to refine the tool's usability and relevance. This iterative approach ensured the tool met farmer expectations and increased its adoption. Moreover, interacting with farmers and answering their questions during the field days or during break time in workshops provides us an opportunity to evaluate the efficacy of the extension program. I plan to institutionalize a feedback loop in all my extension programs, employing

surveys, focus groups, and digital analytics to measure effectiveness. By analyzing adoption rates and environmental outcomes, I aim to quantify the impact of my initiatives. This data will guide program refinement, ensuring relevance and efficacy over time.

Future Vision

Looking forward, my extension philosophy will evolve to address emerging challenges in agriculture, particularly climate change and sustainability. I aim to establish a comprehensive Climate-Smart Extension Program focused on the following objectives: 1) **Enhancing Climate Literacy:** Developing training programs to improve farmers' understanding of climate change and its implications, 2) **Promoting Regenerative Practices:** Conducting research and extension activities on soil, nutrient, and water conservation, 3) **Farmer Oriented Decision Support Tools:** Expanding the use of AI, remote sensing, and crop modeling in decision support systems to provide stakeholders with precise, actionable recommendations by developing web-based decision-supported irrigation and nutrient management tools, and 4) **Fostering Collaboration:** Building networks between researchers, farmers, extension agents, and policymakers to collectively solve agricultural challenges. My extension philosophy is rooted in the belief that effective stakeholder engagement, practical demonstrations, accessible educational outreach, and rigorous program evaluation are essential for driving sustainable agricultural practices. By leveraging these pillars, I aim to bridge the gap between research and practice, fostering climate-resilient farming practices and empowering communities to adapt to evolving challenges. Through innovative tools, collaborative field trials, and strategic outreach, I am committed to creating a lasting, positive impact in agriculture.

Research Statement

Agriculture accounts for approximately 72% of freshwater consumption and is a major contributor to non-point source pollution. Without improved water use efficiency, rising food demand and climate change could push freshwater withdrawals beyond current levels, causing immense pressure on water resources. Additionally, intensifying agricultural practices to boost productivity and increasing extreme weather events may worsen freshwater contamination and exacerbate water scarcity. Addressing these challenges requires innovative strategies to enhance food security while minimizing environmental impact—an urgency that drives my commitment to advancing agricultural sustainability.

My research focuses on developing climate-resilient, sustainable farming systems by integrating cutting-edge technologies in digital agriculture, including agroecological and hydrological models, remote sensing, GIS, and AI/ML, to improve its water and nutrient use efficiency. Over the past decade, I have contributed to interdisciplinary projects that deepened my expertise in watershed hydrology, crop modeling, soil nutrient cycling, agricultural water management, and climate change adaptation.

Currently, I am leading a multidisciplinary state legislature-backed project entitled '*Using AI for improved crop nutrient recommendations'* [funded by: FDACS] with Dr. Zotarelli, collaborating with Drs. Zare and Harley from Department of Electrical and Computer Engineering at UF. I am leading a transdisciplinary team to develop a uniform data collection framework, crop model-guided AI algorithms, and machine learning model to provide crop nutrient recommendations, while balancing productivity and environmental impact. Earlier, I developed a crop model-guided long short-term memory (LSTM)-based machine learning approach to estimate soil mineral nitrogen in seepage irrigation using data-sparse field observations (Gupta et al. 2024). Currently, I am leading my team to integrate unmanned arial vehicles-(UAVs) and satellite-derived multi-/hyperspectral imageries with crop model-integrated ML methods to provide site-specific fertilizer and irrigation application recommendations. Additionally, I participate in the *DSSAT Development Sprint* led by Prof. Gerrit Hoogenboom to contribute to the development and advancement of DSSAT by creating a tool (*DSSAT-SoilPro*) to optimize the soil hydraulic properties using soil moisture and water table observations. Moreover, I assessed the impact of irrigation (overhead sprinkler) and nitrogen management on potato performance under varying Florida climate conditions (da Silva et al., 2024). I am developing AI-based methodologies for site-specific irrigation scheduling maintaining optimum soil moisture based on crop water demand and soil mechanical impedance. While working on these projects, I developed key skills required to successfully lead interdisciplinary projects such as team management-supervision-mentorship, networking, and collaboration skills. Additionally, I gained extensive grant writing experience, drafted four grant proposals with Dr. Zotarelli (PI), one accepted (FDACS) and others pending (USDA-NIFA, SCBGP-FDACS, IFAS-UF).

As a graduate research assistant at UIUC, I worked on a project entitled- '*CoverCrop Analyzer: a web-based decision support tool for cover crop management*' [funded by: Illinois Nutrient Research and Education Council] with Dr. Bhattarai. To develop CoverCrop Analyzer, I developed the CERES-Wheat model of DSSAT as a proxy model for simulating the growth cereal rye as a winter cover crop (Gupta et al., 2022) and investigated its long-term impact and sustainability on tile drainage, water quality, and corn-soybean growth using Regional Climate Models at state scale (Gupta et al., 2023a; 2023b). My dissertation concluded that the cover crops might not be able to completely mitigate climate change impact by 2060 in a few climate divisions of Illinois. During my doctoral research, I also collaborated on a side project to assess the global climate change impact on stream low flows of the Great Miami River watershed, Ohio (Shrestha et al., 2019) using Soil and Water Assessment Tool (SWAT) hydrological model. During my master's at IIT-KGP, I assessed the climate change-induced impact and uncertainty of the rice yield over the agroecological zones of India using various Global Climate Models future projections and warming scenarios (Gupta and Mishra, 2019). Additionally, I developed various agroclimatic tools such as '*Climate Data Bias Corrector*' (Gupta et al., 2019) and '*Weather Data Interpolator*' to remove statistical bias and downscale climate model projections. While working on these projects, I gained foundational knowledge on subject matter such as soil water hydrology, irrigation and drainage, soil nitrogen cycling, water quality, cover crops, crop modeling, watershed hydrology, geospatial modeling, and climate change, which boosted my confidence and solidified my aspirations of becoming a prominent researcher in the field.

With profound expertise in agricultural water management, physical modeling, and integrating ML/AI techniques to provide data-driven computational solutions, I am well-equipped to develop competitive research and extension programs aimed at improving the climate resilience of farming systems regionally and globally by optimizing fertilizer and water use efficiency. In the next few years as a faculty member, my research program will focus on four primary areas: 1) **develop and evaluate site- and crop-specific irrigation and drainage systems** for sandy soils with shallow water-table condition to improve agricultural production and soil-nutrient-water conservation, 2) **advance AI-integrated scalable crop, hydrological, and nutrient transport modeling approaches** to enhance crop water-nutrient use efficiency while reducing nutrient losses and soil erosion amid water table management, 3) **develop web-based decision support system** to improve the state climate resilience by data sciences, remote sensing, and geospatial modeling approaches, and 4) **implement and deliver these strategies** by collaborating with growers/extension agents/stakeholders to improve drainage water quality preventing nutrient/sediment loads to water bodies while improving crop yield. Moreover, I plan to deliver my latest findings by actively writing extension articles and creating educational programs for extension agents, producers, and clienteles.

As I seek to define my own research and extension program to advance process-based modeling, data science driven computational approach and their application to improve water-nutrient use efficiency and climate resilience of the system, I plan to collaborate with faculties in various departments/schools/colleges, such as Agricultural and Biological Engineering, Agronomy, Horticultural Sciences, Soil, Water, and Ecosystem Sciences Departments, College of Agricultural and Life Sciences, School of Natural Resources and Environment along with various research/extension centers and advisory groups. Moreover, I plan to collaborate with various departments in the Herbert Wertheim College of Engineering such as Electrical and Computer Engineering, Environmental Engineering Sciences, and Civil & Coastal Engineering Departments to develop technologies and methods required for my research vision. To support my research program, I am well-equipped to secure competitive external funding from government entities such as USDA, NIFA, United States Geological Survey (USGS), National Science Foundation (NSF), Sustainable Agriculture Research and Education (SARE), Environmental Protection Agency (EPA), and other federal funding agencies. Additionally, I plan to secure state fundings from various programs and grant opportunity provided by FDACS and USDA-backed state Specialty Crop Block Grant Program. Moreover, I am fully prepared to serve as a principal investigator on USDA-NIFA Multistate Hatch projects. Through my research program, I aim to generate peer-reviewed journal articles to showcase my research activities and engage in regional/national/international recognition conferences and professional societies. I am committed to advising and mentoring undergraduate/graduate students, postdoctoral researchers, extension agents, biological scientists, and research technicians. With my research program, I aim to provide practical and actionable solutions to farmers, policymakers, and other stakeholders, optimizing agroecosystems while promoting climate change-resilient water and nutrient use practices.

Extension Philosophy

Extension is one of the important pillars of land-grant universities, bridging the gap between research and practice by providing non-formal education to farmers, agricultural producers, and other stakeholders. I believe that an impactful extension program successfully creates a critical link between scientific research and actionable practices, empowering agricultural stakeholders to address pressing challenges. My extension philosophy centers on four pillars:

1. Stakeholder Engagement and Need Assessment: Extension begins with actively engaging with farmers, county and regional extension agents, and stakeholders to understand their needs and challenges. At UF, I participated in various Potato Field Days at the Hastings Agricultural Extension Center, FL and in-service trainings (IST) conducted by UF/IFAS. While interacting with farmers at field days, I observed that while Florida farmers are aware of the risks associated with extreme weather events, yet many lack in implementing appropriate preventive strategies to counter these impacts. Similarly, at UIUC, I worked with stakeholders to develop the CoverCrop Analyzer, a web-based decision-support tool for cover crop management. Meetings with various Illinois farmers, I identified barriers to adopting cover crops, such as concerns about costs, labor requirements, and limited knowledge of cover crop management. These insights guided my research to ensure its alignment with farmers' needs and expectations. As a faculty member, I plan to expand stakeholder engagement by conducting regular needs assessments through surveys, focus groups, and field visits. By creating trust-based relationships, I aim to bridge gaps in understanding climate change and facilitate the adoption of climate-resilient practices.

2. Demonstration and Knowledge Transfer: Field demonstrations are essential for showcasing the practicality of sustainable practices and tools. At UF, I assessed the impact of different combinations of nitrogen fertilizers and irrigation practices on potato agroecosystem by incorporating field trials with crop modeling. These results were shared with farmers during various field days. Similarly, my involvement in cover crop project at UIUC helped promote the environmental and agronomical benefits of adopting cover crops in Midwestern cropping systems. I envision expanding these efforts by designing collaborative field trials and tools on climate-resilient irrigation and nutrient management practices. Partnering with growers to test these solutions will provide real-world evidence of their efficacy, encouraging broader adoption.

3. Educational Outreach: Effectively disseminating knowledge is at the heart of extension work. At UF, I co-authored an extension publication on "Assessing Nitrate Leaching in Crop Production through the Application of Crop Simulation Models with Experimental Data from Florida" (Sharma et al., 2024) to provide farmers and extension agents with actionable insights for mitigating nutrient loss. Additionally, I contributed to multiple extension articles at UIUC to introduce farmers with the cover crop management tool (Coppess et al. 2020, 2021a, and 2021b). These extension publications allowed me to address specific stakeholder concerns and provide accessible, practical solutions. My future outreach efforts will prioritize developing educational content and extension articles along with disseminating information through bulletins, social media, seminars, and workshops in state/local growers' meetings and conferences.

4. **Program Evaluation:** Continuous improvement in extension programs relies on robust assessment and feedback mechanisms. While developing the CoverCrop Analyzer, I actively sought feedback from stakeholders to refine the tool's usability and relevance. This iterative approach ensured the tool met farmer expectations and increased its adoption. Moreover, interacting with farmers and answering their questions during the field days or during break time in workshops provides us an opportunity to evaluate the efficacy of the extension program. I plan to institutionalize a feedback loop in all my extension programs, employing

surveys, focus groups, and digital analytics to measure effectiveness. By analyzing adoption rates and environmental outcomes, I aim to quantify the impact of my initiatives. This data will guide program refinement, ensuring relevance and efficacy over time.

Future Vision

Looking forward, my extension philosophy will evolve to address emerging challenges in agriculture, particularly climate change and sustainability. I aim to establish a comprehensive Climate-Smart Extension Program focused on the following objectives: 1) **Enhancing Climate Literacy:** Developing training programs to improve farmers' understanding of climate change and its implications, 2) **Promoting Regenerative Practices:** Conducting research and extension activities on soil, nutrient, and water conservation, 3) **Farmer Oriented Decision Support Tools:** Expanding the use of AI, remote sensing, and crop modeling in decision support systems to provide stakeholders with precise, actionable recommendations by developing web-based decision-supported irrigation and nutrient management tools, and 4) **Fostering Collaboration:** Building networks between researchers, farmers, extension agents, and policymakers to collectively solve agricultural challenges. My extension philosophy is rooted in the belief that effective stakeholder engagement, practical demonstrations, accessible educational outreach, and rigorous program evaluation are essential for driving sustainable agricultural practices. By leveraging these pillars, I aim to bridge the gap between research and practice, fostering climate-resilient farming practices and empowering communities to adapt to evolving challenges. Through innovative tools, collaborative field trials, and strategic outreach, I am committed to creating a lasting, positive impact in agriculture.

Research Statement

Agriculture accounts for approximately 72% of freshwater consumption and is a major contributor to non-point source pollution. Without improved water use efficiency, rising food demand and climate change could push freshwater withdrawals beyond current levels, causing immense pressure on water resources. Additionally, intensifying agricultural practices to boost productivity and increasing extreme weather events may worsen freshwater contamination and exacerbate water scarcity. Addressing these challenges requires innovative strategies to enhance food security while minimizing environmental impact—an urgency that drives my commitment to advancing agricultural sustainability.

My research focuses on developing climate-resilient, sustainable farming systems by integrating cutting-edge technologies in digital agriculture, including agroecological and hydrological models, remote sensing, GIS, and AI/ML, to improve its water and nutrient use efficiency. Over the past decade, I have contributed to interdisciplinary projects that deepened my expertise in watershed hydrology, crop modeling, soil nutrient cycling, agricultural water management, and climate change adaptation.

Currently, I am leading a multidisciplinary state legislature-backed project entitled '*Using AI for improved crop nutrient recommendations'* [funded by: FDACS] with Dr. Zotarelli, collaborating with Drs. Zare and Harley from Department of Electrical and Computer Engineering at UF. I am leading a transdisciplinary team to develop a uniform data collection framework, crop model-guided AI algorithms, and machine learning model to provide crop nutrient recommendations, while balancing productivity and environmental impact. Earlier, I developed a crop model-guided long short-term memory (LSTM)-based machine learning approach to estimate soil mineral nitrogen in seepage irrigation using data-sparse field observations (Gupta et al. 2024). Currently, I am leading my team to integrate unmanned arial vehicles-(UAVs) and satellite-derived multi-/hyperspectral imageries with crop model-integrated ML methods to provide site-specific fertilizer and irrigation application recommendations. Additionally, I participate in the *DSSAT Development Sprint* led by Prof. Gerrit Hoogenboom to contribute to the development and advancement of DSSAT by creating a tool (*DSSAT-SoilPro*) to optimize the soil hydraulic properties using soil moisture and water table observations. Moreover, I assessed the impact of irrigation (overhead sprinkler) and nitrogen management on potato performance under varying Florida climate conditions (da Silva et al., 2024). I am developing AI-based methodologies for site-specific irrigation scheduling maintaining optimum soil moisture based on crop water demand and soil mechanical impedance. While working on these projects, I developed key skills required to successfully lead interdisciplinary projects such as team management-supervision-mentorship, networking, and collaboration skills. Additionally, I gained extensive grant writing experience, drafted four grant proposals with Dr. Zotarelli (PI), one accepted (FDACS) and others pending (USDA-NIFA, SCBGP-FDACS, IFAS-UF).

As a graduate research assistant at UIUC, I worked on a project entitled- '*CoverCrop Analyzer: a web-based decision support tool for cover crop management*' [funded by: Illinois Nutrient Research and Education Council] with Dr. Bhattarai. To develop CoverCrop Analyzer, I developed the CERES-Wheat model of DSSAT as a proxy model for simulating the growth cereal rye as a winter cover crop (Gupta et al., 2022) and investigated its long-term impact and sustainability on tile drainage, water quality, and corn-soybean growth using Regional Climate Models at state scale (Gupta et al., 2023a; 2023b). My dissertation concluded that the cover crops might not be able to completely mitigate climate change impact by 2060 in a few climate divisions of Illinois. During my doctoral research, I also collaborated on a side project to assess the global climate change impact on stream low flows of the Great Miami River watershed, Ohio (Shrestha et al., 2019) using Soil and Water Assessment Tool (SWAT) hydrological model. During my master's at IIT-KGP, I assessed the climate change-induced impact and uncertainty of the rice yield over the agroecological zones of India using various Global Climate Models future projections and warming scenarios (Gupta and Mishra, 2019). Additionally, I developed various agroclimatic tools such as '*Climate Data Bias Corrector*' (Gupta et al., 2019) and '*Weather Data Interpolator*' to remove statistical bias and downscale climate model projections. While working on these projects, I gained foundational knowledge on subject matter such as soil water hydrology, irrigation and drainage, soil nitrogen cycling, water quality, cover crops, crop modeling, watershed hydrology, geospatial modeling, and climate change, which boosted my confidence and solidified my aspirations of becoming a prominent researcher in the field.

With profound expertise in agricultural water management, physical modeling, and integrating ML/AI techniques to provide data-driven computational solutions, I am well-equipped to develop competitive research and extension programs aimed at improving the climate resilience of farming systems regionally and globally by optimizing fertilizer and water use efficiency. In the next few years as a faculty member, my research program will focus on four primary areas: 1) **develop and evaluate site- and crop-specific irrigation and drainage systems** for sandy soils with shallow water-table condition to improve agricultural production and soil-nutrient-water conservation, 2) **advance AI-integrated scalable crop, hydrological, and nutrient transport modeling approaches** to enhance crop water-nutrient use efficiency while reducing nutrient losses and soil erosion amid water table management, 3) **develop web-based decision support system** to improve the state climate resilience by data sciences, remote sensing, and geospatial modeling approaches, and 4) **implement and deliver these strategies** by collaborating with growers/extension agents/stakeholders to improve drainage water quality preventing nutrient/sediment loads to water bodies while improving crop yield. Moreover, I plan to deliver my latest findings by actively writing extension articles and creating educational programs for extension agents, producers, and clienteles.

As I seek to define my own research and extension program to advance process-based modeling, data science driven computational approach and their application to improve water-nutrient use efficiency and climate resilience of the system, I plan to collaborate with faculties in various departments/schools/colleges, such as Agricultural and Biological Engineering, Agronomy, Horticultural Sciences, Soil, Water, and Ecosystem Sciences Departments, College of Agricultural and Life Sciences, School of Natural Resources and Environment along with various research/extension centers and advisory groups. Moreover, I plan to collaborate with various departments in the Herbert Wertheim College of Engineering such as Electrical and Computer Engineering, Environmental Engineering Sciences, and Civil & Coastal Engineering Departments to develop technologies and methods required for my research vision. To support my research program, I am well-equipped to secure competitive external funding from government entities such as USDA, NIFA, United States Geological Survey (USGS), National Science Foundation (NSF), Sustainable Agriculture Research and Education (SARE), Environmental Protection Agency (EPA), and other federal funding agencies. Additionally, I plan to secure state fundings from various programs and grant opportunity provided by FDACS and USDA-backed state Specialty Crop Block Grant Program. Moreover, I am fully prepared to serve as a principal investigator on USDA-NIFA Multistate Hatch projects. Through my research program, I aim to generate peer-reviewed journal articles to showcase my research activities and engage in regional/national/international recognition conferences and professional societies. I am committed to advising and mentoring undergraduate/graduate students, postdoctoral researchers, extension agents, biological scientists, and research technicians. With my research program, I aim to provide practical and actionable solutions to farmers, policymakers, and other stakeholders, optimizing agroecosystems while promoting climate change-resilient water and nutrient use practices.

Computers and Electronics in Agriculture 225 (2024) 109355

Contents lists available at ScienceDirect

Computers and Electronics in Agriculture

journal homepage: www.elsevier.com/locate/compag

Original papers

Estimating soil mineral nitrogen from data-sparse field experiments using crop model-guided deep learning approach

Rishabh Gupta [a,*], Satya K. Pothapragada [b,*], Weihuang Xu [b], Prateek Kumar Goel [a], Miguel A. Barrera [b], Mira S. Saldanha [b], Joel B. Harley [b], Kelly T. Morgan [c], Alina Zare [b], Lincoln Zotarelli [a]

[a] Horticultural Sciences Department, Institute of Food & Agricultural Sciences - University of Florida, Gainesville, FL 32611, USA
[b] Department of Electrical & Computer Engineering, University of Florida, Gainesville, FL 32603, USA
[c] Soil, Water, and Ecosystem Sciences Department, Institute of Food & Agricultural Sciences - University of Florida, Gainesville, FL 32611, USA

ARTICLE INFO

Keywords:
Crop System Model
DSSAT
Leaching
LSTM
Machine Learning
SUBSTOR-Potato

ABSTRACT

Sandy soils are susceptible to excessive nitrogen (N) leaching under intensive crop production which is linked with the soil's low nutrient holding capacity and high-water infiltration rate. Estimating soil mineral nitrogen (SMN) at the daily time-step is crucial in providing fertilizer recommendations balancing plant nitrogen use efficiency (NUE) and N losses to the environment. Crop models [e.g., Decision Support System for Agro-technology Transfer (DSSAT)] can simulate SMN trend under varied fertilizer application rates and timings but struggle with accuracy in high-water table conditioned sub-irrigated soil. Alternatively, time-series deep learning (DL) models based on a long short-term memory (LSTM) are promising in understanding nonlinearity among complex variables. Yet, purely data-driven DL models for crops are challenging to develop due to limited field data availability and the excessive costs to collect more field data. Hence, a hybrid model (hybrid-LSTM) was developed by leveraging both the DSSAT and LSTM models to estimate daily SMN primarily using daily weather, applied fertilizer rates-timings, and the SMN sparse observations. This study used the observations from field trials conducted between 2011–2014 in Hastings, FL. The first step was to calibrate the DSSAT-SUBSTOR-Potato model to produce reliable SMN of the topsoil for treatments with varied N applied fertilizer rates split among the pre-planting, emergence, and tuber-initiation stages of the potato. Thereafter, the hybrid-LSTM model was trained on the calibrated DSSAT simulated SMN time-series and fine-tuned its predictions using the observed SMN to improve DSSAT simulated SMN. The hybrid-LSTM model was then tested on both calibrated and un-calibrated DSSAT SMN simulations where it outperformed the DSSAT model (improvement ranged ~18–30 % on comparing the normalized root mean squared error) in providing reliable estimates of SMN across most of the farms and years. This novel hybrid modeling approach could guide stakeholders and farmers to build sustainable N management with improved crop NUE and yield and help in minimizing environmental losses.

1. Introduction

Florida is a significant off-season contributor of fresh potatoes (*Solanum tuberosum L.*) consistently producing approximately 30 % of the total spring production of the United States (NASS - Quick Stats, 2021). Most of the potatoes are grown in the northeastern region of the state, which is dominated by uncoated fine sand soil-type (USDA-NASS, 2019). These soils are highly vulnerable to excessive N leaching due to rapid water infiltration (Zotarelli et al., 2007). Moreover, there is an impermeable soil layer below the surface, which allows subirrigation by maintaining a high-water table throughout the growing season (da Silva et al., 2018; Dukes et al., 2010). At the same time, it also poses challenges in effectively draining excess water during heavy rainfall. The field relies on irrigation furrows to distribute fresh ground water into the fields, connected to surrounding drainage ditches. Under heavy rainfall conditions, the water table level in the field is receded and excess water is moved off-site by gravity carrying soluble nutrients, especially N, from the top layer of the soil. Hence, multiple N fertilizer applications are required to compensate for the excess N leaching (Errebhi et al., 1998; Scholberg et al., 2013; Simonne et al., 2010).

Several studies have tried to estimate ideal fertilizer application rates and timings for potatoes in Florida (da Silva et al., 2023; Rens et al.,

* Corresponding authors.
E-mail addresses: r.gupta@ufl.edu (R. Gupta), spothapragada@ufl.edu (S.K. Pothapragada).

https://doi.org/10.1016/j.compag.2024.109355
Received 9 February 2024; Received in revised form 28 June 2024; Accepted 12 August 2024
Available online 22 August 2024

2018, 2016b, 2016a, 2015b, 2015a; Zotarelli et al., 2015, 2014). Typically, the N fertilizer applications should synchronize with the potato growth and N uptake. The N uptake rates are higher between mid-vegetative growth to mid-tuber initiation; hence, N applications between emergence and tuber-initiation are crucial to maximize tuber yield (Rens et al., 2018, 2016a; Zotarelli et al., 2015). Zotarelli et al. (2015) and Rens et al. (2018) reported that N application at pre-planting could be highly prone to leaching particularly in cases of heavy rainfall projections before plant emergence. Hence, they recommended applying N fertilizer close to planting rather than at pre-planting. Another finding by Zotarelli et al. (2014) suggests keeping N application rates 224–280 kg-N ha^{-1} to maximize tuber yield under heavy rainfall conditions. Rens et al. (2018) found that tuber yield peaked when N application specifically at emergence varied from 128-168 kg-N ha^{-1} under the same circumstances. Similarly, N application at tuber-initiation was observed to be crucial with ~62 % NUE and found to have no significant increase in tuber yield when applied above 56 kg-N ha^{-1} (Rens et al., 2016a; Zotarelli et al., 2014). Therefore, it becomes crucial to adaptively manage fertilizer application rates and timings in highly uncertain rainfall conditions to minimize N leaching and maximize the potato yield and NUE.

Estimating SMN at a higher frequency is crucial for making informed decisions for fertilizer applications to aid growers and stakeholders in improving tuber yield/NUE without harming the environment. However, in-situ measurement of SMN is time-consuming, labor-intensive, and highly expensive. Consequently, there is a need to develop a reliable framework and methodology to estimate SMN at a higher frequency using the limited experimental data. Previous research employed process-based cropping system models, such as DSSAT (Hoogenboom et al., 2019; Jones et al., 2003), Agricultural Production Systems sIMulator (APSIM) (Holzworth et al., 2014; Keating et al., 2003), and other agroecological models to replicate the interaction between soil, water, and crop under different environmental and geographical setting (Gupta et al., 2022; Liu et al., 2011; Moriasi et al., 2013; Salmerón et al., 2014; Saseendran et al., 2007). However, these models may not capture the variability of SMN in the sandy soils. This variability is particularly difficult to capture when caused by uncertain rainfall events and water table fluctuations as these models could not simulate three-dimensional movement of N dynamics in the soil layers, necessary for precise estimation of SMN (Archontoulis et al., 2014; Raymundo et al., 2017; Wallach and Thorburn, 2014). Hence, these models may not provide accurate estimates of SMN required for narrowing weather-specific crop nutrient recommendations. Additionally, calibrating these models could be tedious and time-consuming as it would require fine-tuning the extensive list of parameters (Akhavizadegan et al., 2021; Seidel et al., 2018). Furthermore, one would need to re-calibrate these models for different climate and soil conditions. Therefore, current circumstances call for devising new approaches and modeling frameworks that possess transferable learning, reducing the need for repeated calibration.

DL, a subfield of machine learning (ML) based on artificial neural network, have been extensively embraced by researchers for interpreting intricate agroecological systems due to their ability to effectively learn complex relationships between plants, soil, and climate (Kamilaris and Prenafeta-Boldú, 2018). A significant advantage of DL is their transfer learning ability to adapt to completely new scenarios (Weiss et al., 2016), such as, for different crop, soil, or climate conditions. Most research in the field of agriculture used remote-sensing data (multi-spectral/hyperspectral dataset or satellite imageries) for building DL models to estimate the crop growth stages (Wang et al., 2022; Yue et al., 2020), detect pests/diseases (Hadipour-Rokni et al., 2023; Mohanty et al., 2016), predict leaf area indices (Ilniyaz et al., 2023), soil properties (Padarian et al., 2019; Zhang et al., 2022), and evapotranspiration (Sharma et al., 2022). A deep recurrent neural network-based LSTM (Hochreiter and Schmidhuber, 1997) has significant popularity and proven its effectiveness in capturing the temporal variations in the high-frequency time series data in the agricultural domain (Datta and Faroughi, 2023; Gauch et al., 2021; Zhang et al., 2022). For instance, Saha et al. (2023) built an LSTM model to estimate daily stream nitrate concentration for 42 sites in data-sparse watersheds in Iowa, U.S.A., with a Nash-Sutcliffe efficiency of ~75 %. Similarly, Datta and Faroughi (2023) successfully predicted the soil moisture content for the next hour, a day, a week, and a month in advance, achieving a R-squared value of ~0.95, developing a multiheaded-LSTM model, training the model using 15-minute interval field measurements.

Nevertheless, time-series DL models like LSTM necessitate daily, or at least, discrete long-term observations of SMN at a higher frequency, to effectively learn the temporal variation caused by weather, irrigation, and N applications (Hua et al., 2019). Building an effective time-series DL model like LSTM with limited observations (5–6 sampling per season) could not be possible. Hence, the DSSAT model was employed to accompany the LSTM model to accurately predict the response of weather and N application (rates/timings) on daily SMN values during the crop cycle (Fig. 1). Numerous studies have affirmed that ML could serve as a companion to crop models rather than a competitor (Everingham et al., 2016; van Klompenburg et al., 2020; Zhang et al., 2023). Feng et al. (2020) asserted that comprising ML-based models with biophysical agroecosystem models has the potential to provide an improved evaluation of the changing climate on wheat yield in Australia. Similarly, Shahhosseini et al. (2021) found promising results in adding APSIM model simulated features in the ML models for corn yield prediction in the U.S. Corn Belt. In the present study, a novel method was developed to integrate the DSSAT and LSTM models (hybrid-LSTM) to precisely estimate the SMN between the soil sampling dates. Integrating DSSAT simulations with LSTM could allow LSTM to learn the SMN dynamics in limited field data conditions and later, its prediction could be fine-tuned using field observations. Past studies mentioned earlier developed various such hybrid models incorporating different crop growth indicators with weather data for yield forecasting. However, to the authors' knowledge, there is no such hybrid model developed to improve the daily estimates of cropping system dynamics like SMN in sparse field observations. The hypothesis of this research was that the hybrid-LSTM model could reduce the gap between the DSSAT simulated and observed SMN to provide accurate estimates of SMN. The overall goal of this integration was to leverage the strengths of the DSSAT model (ability to simulate dynamics of the cropping system) and mitigate its weaknesses (inability to capture soil nutrients and water dynamics in a sub-irrigated high water-table conditioned cropping systems) to produce the daily SMN fluctuations with rainfall variability. The objective of this study was to develop the hybrid-LSTM model that accurately estimate daily SMN throughout a potato crop growing season using limited SMN field observations. The outcomes of this study can be further used to support informed decisions on fertilizer application rates and timings recommendations to improve the tuber yield and reduce the environmental impacts.

2. Materials and methods

2.1. Field experiments overview

A series of field experiments were conducted on several commercial potato fields in Hastings, FL, U.S. from 2011 to 2014. Experiments carried out in the spring of 2011 and 2012 used the 'FL 1867' potato cultivar in Farm 1 (F1), Farm 2 (F2), and Farm 3 (F3), and the 'Atlantic' potato cultivar in Farm 4 (F4). Field trials conducted in 2013 and 2014, used the 'FL 1867' potato cultivar in F1 and the 'Atlantic' potato cultivar in F4. Each experiment was comprised of eight treatments combining three application timings and N fertilizer rates (Table 1). All the treatments of 2011–2012 experiments received a fixed application (56 kg-N ha^{-1}) of granular ammonium nitrate during pre-plant followed by the combination of four different rates of liquid urea ammonium nitrate (0, 56, 112, and 168 kg-N ha^{-1}) at the plant emergence and two different rates of liquid urea ammonium nitrate (56, 112 kg-N ha^{-1}) as side-dress

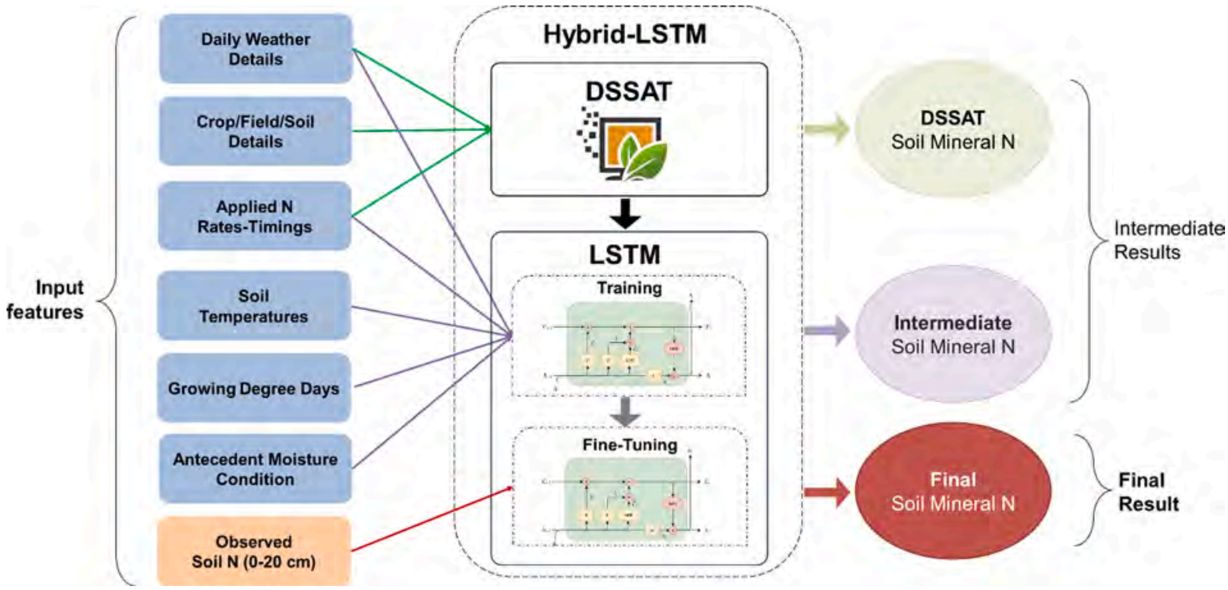

Fig. 1. Block diagram to demonstrate the integration of two models (DSSAT and LSTM), input features, and the workflow of the hybrid-LSTM model. LSTM picture source: LSTM architecture, 2021

Table 1
Nitrogen fertilizer application rates and timings (at pre-planting, plant emergence, and tuber initiation) along with total N applied during potato season for the two set of field experiments. One set of experiments was conducted in 2011 and 2012 (on F1, F2, F3, and F4), while other set of experiments was conducted in 2013 and 2014 (on F1 and F4).

Experiment Year	Treatments	N fertilizer rate applied (kg-N ha^{-1}) at			Total N applied (kg-N ha^{-1})
		Pre-planting	Plant Emergence	Tuber Initiation	
2011 and 2012	1	56	0	56	112
	2	56	0	112	168
	3	56	56	56	168
	4	56	56	112	224
	5	56	112	56	224
	6	56	112	112	280
	7	56	168	56	280
	8	56	168	112	280
2013 and 2014	1	0	0	56	56
	2	56	0	56	112
	3	0	56	56	112
	4	56	56	56	168
	5	0	112	56	168
	6	56	112	56	224
	7	0	168	56	224
	8	56	168	56	280

at the tuber initiation. The treatments of 2013–2014 experiments received N fertilizer rates of 0 and 56 kg-N ha^{-1} at the pre-plant followed by N rates of 0, 56, 112, and 168 kg-N ha^{-1} at the plant emergence and a fixed rate of 56 kg-N ha^{-1} at the tuber initiation. All the treatments in the last two years of experiments received granular ammonium nitrate. The predominant soil types in F1, F2, F3, and F4 are Ultic Alaquod, Arenic Glossaqualf, Arenic Endoaqualf, and Aeric Alaquod, respectively. All the field trails from 2011 to 2014 were conducted under seepage irrigated maintaining the water table at around 55 cm below the top of the raised soil bed for potato planting. More details of the experimental design and results can be found in (Rens et al., 2018, 2015a; Zotarelli et al., 2015, 2014).

2.2. Data description and preparation

This study uses weather-soil-crop management information to use the DSSAT and DL models for estimating daily SMN. Daily weather data of Hastings, FL (29.6933°N, 81.4449°W), comprised of rainfall, average, minimum, maximum air temperatures (at 60 cm height), solar radiation, dew point, relative humidity, and wind speed, was collected from the Florida Automated Weather Network data access portal (FAWN, 2022) from 2010 to 2014. Daily soil temperatures (average, minimum, maximum) were also collected from the FAWN data access portal for the same period. Field management information such as potato planting/emergence/tuber initiation/harvesting dates, details of fertilizer application (rates and timings), initial soil condition, previous cropping history, and other relevant information was collected from the earlier studies (mentioned in section 2.1). Soil surface information, such as soil albedo, slope, drainage class, runoff potential, and soil layered information, for instance, soil wilting point and field capacity, saturated water content, bulk density, organic matter content, and soil root growth factor were collected from previous studies and Soil Survey Geographic Database (da Silva et al., 2018, 2024; Rens et al., 2022; Reyes-Cabrera et al., 2016; SSURGO Database, 2018). The DSSAT model was set up using the weather-soil-crop information collected from various sources by creating its weather, soil, and experiment files. The observed datasets from the field trials such as SMN, potato aboveground biomass/N accumulation, and tuber weight (dry and fresh)/N accumulation was used to evaluate the performance of the DSSAT model for replicating the potato cropping system.

Likewise, daily weather information (rainfall, average air temperature), soil average temperature, applied N fertilizer rate-timing, and the DSSAT simulated daily SMN were considered in developing the hybrid-LSTM model. Later, all the dataset was arranged into different samples, where every sample corresponds to a specific farm and a specific year with all the N applied rates-timings treatments. Every sample contained training input of a specific farm-year (farm-year represents the data of a specific farm and a specific year), which includes data for all the input features and a target vector for the same. Since the hybrid-LSTM was first trained with the DSSAT simulated data and then with the observed SMN, we had two sets of training data, the difference being the number of data points per sample. The first training set with DSSAT simulated data had daily data of all the features for the whole crop cycle, while the second training set contained only those points when the SMN was sampled at the experiment site for fine-tuning the model (more detail provided in section 2.4.2).

2.3. Model description

2.3.1. DSSAT model

The process-based cropping system model DSSAT v4.8 can simulate soil water-nutrient dynamics and the growth and development of more than 42 crops (Hoogenboom et al., 2023, 2019; Jones et al., 2003). It employs the SUBSTOR-Potato model (Griffin et al., 1993; Ritchie et al., 1995) of DSSAT to simulate daily phenological development, biomass accumulation, and tuber yield in a climate-soil varied setting. SUBSTOR-Potato model of DSSAT has five cultivar-type parameters listed and explained in Table 2. It also has two radiation use efficiency (RUE) parameters- RUE1 (3.5 g plant dry matter per MJ photosynthetically active radiation) and RUE2 (4.0 g plant dry matter per MJ photosynthetically active radiation) to represent before and after tuber initiation, respectively. We used Ritchie water balance, FAO-56, soil conservation service, and Suleiman-Ritchie methods to estimate soil hydrology, evapotranspiration, soil infiltration, and soil evaporation, respectively (Ritchie et al., 1998). Active, intermediate, and passive soil organic matter pools were determined using the CENTURY model in DSSAT (Gijsman et al., 2002).

The DSSAT model was calibrated for plant/tuber biomass/N accumulation and SMN. The soil parameters listed in Table 3, initial soil conditions were adjusted carefully to closely match with the real field conditions based on previous literatures (da Silva et al., 2018; Rens et al., 2022; Reyes-Cabrera et al., 2016) and Soil Survey Geographical Database (SSURGO Database, 2018) to simulate SMN for the topsoil layer. Genotype coefficients of the potato cultivar were also calibrated to replicate tuber/plant biomass and N accumulation (Table 2). Other parameters, for example, sprout length and irrigation depth were also found sensitive to plant/tuber growth, hence these parameters were also adjusted accordingly for modeling purposes. The DSSAT model was calibrated using field observations of four farms (F1-F4) from 2011 to 2012; while the model was evaluated using field observations of two farms (F1, F2) from 2013 to 2014.

2.3.2. LSTM model

A long short-term memory (LSTM) is a type of recurrent neural network designed to process sequential (or time-series) data by effectively capturing long-term dependencies (Hochreiter and Schmidhuber, 1997). The structure of each LSTM cell is the fundamental block of the network. It consists of a memory state along with a gating mechanism to selectively store, forget, and process the output information (Staudemeyer and Morris, 2019). Whenever a new input sequence is passed through an LSTM cell, the input and forget gates within its cell determine what information is needed to be stored in and forgotten from the cell state, respectively, at a particular timestep. The output gate processes information from the cell state, which is the output of LSTM at the consecutive time step. This process is repeated for each time step to produce a time series output. Sequence length is one of the crucial parameters of LSTM architecture. The sequence length refers to the number of time steps or observations considered as input to predict the consecutive time step (Hua et al., 2019). It affects the computational complexity, memory requirements, and the model's ability to capture long-term dependencies. While longer sequences may lead to vanishing gradients, overfitting, and increased computational demands, shorter sequences offer computational efficiency at the cost of reduced context (Hua et al., 2019). Every input that goes into the LSTM network is a sequence of data which is represented as an input feature matrix and a target vector, fed into the model one after the other while training.

Table 2
Calibrated genotype coefficients to simulate the growth and development of 'Atlantic' and 'FL1867' potato cultivar.

Coefficients	Definitions	Values
G2	Leaf area expansion rate after tuber initiation (cm^2 m^{-2} d^{-1})	1000
G3	Potential tuber growth rate (g m^{-2} d^{-1})	25.0
PD	Index that suppresses tuber growth during the period that immediately follows tuber induction	0.9
P2	Tuber initiation sensitivity to long photoperiods	0.6
TC	Upper critical temperature for tuber initiation (°C)	17.0

2.3.3. The hybrid-LSTM model – An integration of the DSSAT and LSTM model

The main goal of the hybrid-LSTM model development was to estimate the daily time-series of the SMN using weather, soil and air temperatures, and fertilizer applications (rates, timings). However, training a DL model exclusively on experimental SMN data was infeasible due to the lack of observed data available (only 5–6 soil samplings per crop cycle per experiment) to effectively capture the complex relationships and estimate daily SMN. Hence, the DSSAT model was used, which could simulate daily SMN throughout the crop cycle along with other input features to train a time-series (LSTM) model. However, the current version of DSSAT (v4.8) is unable to replicate a sub-irrigated cropping system in a high-water table condition. Hence, the study was aimed to improve over DSSAT simulated SMN in such data-sparse, sub-irrigated, and high-water table conditioned potato cropping system in northeastern Florida using hybrid-LSTM model. The key concept was to train a time-series DL model (LSTM) on the DSSAT simulated SMN to learn the interaction between the features and the simulated SMN. The DL model (LSTM) then trained again with these features but with experimental SMN on those days experimental data is available, which is called as fine-tuning process. This integration of DSSAT and LSTM was called the hybrid-LSTM model for ease of explanation.

2.4. Setup and training of the hybrid-LSTM model

2.4.1. Feature engineering and parameterization

Identifying the right features for the ML/DL model, known as feature engineering, is crucial for efficient model training. As a part of the study, a comprehensive list of potential features was prepared to subsequently conduct a feature importance test using wrapper method, identifying the features that exhibited a strong correlation with the experimental SMN values. In this method, a linear regression model was trained considering all possible subsets of the potential features and the features subset with the best performance was chosen to be the optimal feature set. The comprehensive list of potential features included daily rainfall, minimum–maximum-average air and soil temperatures, evapotranspiration, and applied N fertilizer rate-timing. The list also included the DSSAT simulated data- daily SMN and plant N, in the feature selection process. Furthermore, a feature to capture the interaction of rainfall and applied N fertilizer rates (multiplying their normalized values only when rainfall occurred on the N application day) was also included in the potential features list to supplement the model understanding of potential N leaching. We also considered features such as antecedent soil moisture condition (it was assumed dry when last 5 days of rainfall accumulation was less than 35 mm, and wet in vice versa) (Gray et al., 1982) and growing degree days of potato crop (Hristine et al., 2005). After performing feature importance testing, the selected features were rainfall, average air temperatures, soil average temperatures, applied N fertilizer rate-timing, rainfall and applied N fertilizer rate interaction feature, and the DSSAT simulated SMN.

The sequence length of the hybrid-LSTM model was optimized to ten days by conducting a grid search through experimental runs to effectively capture meaningful patterns between input features and the target variable. The architecture of the LSTM model developed had one hidden layer with five hidden units and the hyperbolic tangent function (Tanh) was used as an activation function. The hidden units and the number of layers were chosen such that the total number of the hybrid-LSTM

Table 3
Soil layered parameters were used for all the farms to calibrate the DSSAT model..

Soil Depth (cm)	Wilting Point ($cm^3\ cm^{-3}$)	Field Capacity ($cm^3\ cm^{-3}$)	Saturated water content ($cm^3\ cm^{-3}$)	Root growth factor (fraction)	Bulk Density ($g\ cm^{-3}$)	Organic Carbon (%)
5	0.055	0.128	0.456	1.00	1.53	0.65
15	0.055	0.125	0.456	1.00	1.53	0.65
30	0.047	0.123	0.429	0.44	1.54	0.63
45	0.036	0.122	0.419	0.02	1.57	0.44
60	0.036	0.119	0.419	0.01	1.57	0.44
90	0.036	0.110	0.407	0.00	1.61	0.18
120	0.036	0.110	0.407	0.00	1.61	0.18
150	0.033	0.108	0.401	0.00	1.63	0.08

Source: da Silva et al., 2018, 2024; Rens et al., 2022; Reyes-Cabrera et al., 2016; SSURGO Database, 2018

parameters did not exceed the total number of input data points, minimizing the model complexity. A specific optimizer and learning rates were used for training (Adam optimizer) and fine-tuning (stochastic gradient descent optimizer) the hybrid-LSTM model. The L_2 regularization (weight decay of 0.01) was applied to the weights of the hybrid-LSTM model. Moreover, air and soil temperatures were detrended (to remove seasonality from winter to summer) to avoid any temporal dependencies in the hybrid-LSTM model.

2.4.2. Training procedure

The training procedure was divided into two parts. The first part included the training of the LSTM model on the calibrated DSSAT simulations of SMN. This part of the training allowed LSTM model to understand the relationship between input features and DSSAT simulated SMN. The second part of the training procedure involved refining the SMN estimated by the LSTM model further using sparsely available observed SMN (the training procedure are elaborated in Table S2 in the supplementary material). This integration of LSTM being trained on the DSSAT simulated SMN and later, fine-tuned using observed SMN to refine SMN estimates was referred to as the hybrid-LSTM model (Fig. 1). The developed hybrid-LSTM model's performance was tested against SMN observations on the sampling days while feeding the unseen-calibrated and -uncalibrated DSSAT simulations of different farm-year combinations in the hybrid-LSTM model.

The hybrid-LSTM model, trained using continuous DSSAT simulated data and subsequently, fine-tuned using sparsely observed ground truth data, could be prone to significant noise in its predictions during the fine-tuning process. This noise in SMN predictions could be attributed to the mismatch in trends and an offset between the DSSAT simulated SMN curve and field observed SMN in certain farm-years. Since the soil parameters of the DSSAT model were not adjusted for specific farm during calibration, assuming similar soil properties among all the farms, the performance and trend of the DSSAT simulated SMN was not consistent for all the farm-years compared to the observed SMN. These discrepancies could force the hybrid-LSTM model to learn new patterns during fine-tuning process with ground truth data which differed from those learned during initial training with the DSSAT simulations. To address this issue, the farm-years were qualitatively classified into "good", "bad", and "uncertain" categories based on the agreement between the DSSAT-simulated and observed SMN trends before the model training. This classification was performed based on the coefficient of determination (R^2, Section 2.5) and visual interpretation as it was difficult to only rely on R^2 to compare trends between two quantities which are not distributed similarly over time (DSSAT simulated SMN is continuous, daily data is available, but observed SMN is sparsely distributed over time). The good farm-years exhibited consistent trends and had lower offsets (e.g. Fig. S1 in the supplementary), while the bad farm-years showed contrasting patterns with higher offsets between the DSSAT simulated and field observed SMN (e.g. Fig. S2 in the supplementary). The uncertain farm-years had incomplete agreement across all sampling stages (e.g. Fig. S3 in the supplementary).

From the initial training stages of the hybrid-LSTM, it was understood that these bad and uncertain farm-years could influence the overall model performance and induce unexplainable uncertainty in the predictions. Moreover, given the limited data to train the hybrid-LSTM model, excluding these problematic farm-years would have further reduced the already scarce dataset. Hence, to obtain an upper bound of performance, a slightly different approach was adopted to develop an optimal hybrid model to improve SMN estimates at and between sampling days over the DSSAT simulations. This strategy involved training the hybrid-LSTM model on multiple sets of farm-year combinations, identifying the model with best farm-year combinations, and finally, confirming its performance on unseen calibrated as well as uncalibrated DSSAT data. Table S1 demonstrates how the hybrid-LSTM performance could be affected by training it on a randomly selected combination of farm-years. When the hybrid-LSTM model was trained with random combination of farm-years, the outcomes usually surpassed the DSSAT simulations, but fall shorted compared to those obtained with the optimal farm-years combination (Table S1). These results emphasize the significance of the (modified) training procedure explained in the subsequent paragraph.

For training hybrid-LSTM model, multiple combinations of farm-years were generated at random, emphasizing more on good farm-years followed by bad and uncertain farm-years. This process allowed us to assess the model's ability to handle variations introduced by challenging farm-years (bad and uncertain farm-years) simulations for effective learning. By incorporating this approach, we ensured that the hybrid-LSTM model effectively captured the complexities of different categories (good/bad/uncertain), leading to accurate predictions of SMN. For each farm-years combination, the hybrid-LSTM model was trained and fine-tuned using DSSAT simulations and field observations, respectively, applying a k-fold cross-validation approach, and evaluated on the field observations of remaining farm-years. This process was repeated several times (please refer to Table S2 for more detail on training process). The need of multiple iterations for training the hybrid-LSTM model stemmed from the SMN observations scarcity, posing a formidable challenge in estimating a precise SMN line curve for each crop cycle. Later, SMN predictions from multiple iterations were averaged to get a smooth line curve with reduced noise. The SMN area curve was derived using the standard deviation across all the iterations.

Finally, the combination of farm-years yielded the best model performance was selected and tested on unseen uncalibrated DSSAT simulations using the performance metrics mentioned in section 2.5. This approach produced the least noise in the prediction during model testing to determine whether the hybrid-LSTM model could improve the calibrated and uncalibrated DSSAT daily SMN simulations using field sampled SMN. The final set of farm-years used for training were F1-2011, F2-2011, F3-2011, F2-2012, F3-2012, F4-2012, F4-2013, and F4-2014 (50.0 %, 25.0 %, and 25.0 % of "good", "bad", and "uncertain" farm-years, respectively). The testing set with unseen calibrated DSSAT farm-years (Testing-1) included F4-2011, F1-2012, F1-2013, and F1-2014 (75 % "good" and 25 % "bad" farm-years) for all treatments (F3-2011 and F4-2013 were uncertain farm-years, and F1-2012, F3-2012, and F4-2012 were bad farm-years, rest are good farm-years) Later, the developed hybrid-LSTM model was tested on unseen uncalibrated DSSAT simulations (Testing-2) on F1-2011, F2-2011, F3-2011, F1-2012, F2-2012, F3-2012, F1-2013, F1-2014, and F4-2014..

2.5. Statistical approaches for the model evaluation

The hybrid-LSTM and DSSAT model performances were assessed by comparing the estimated or simulated SMN line and area curves from these models against the scarce observations. Since there were four replicates of the SMN observed and we estimated the SMN line/area curve using the hybrid-LSTM and DSSAT models, it was challenging to evaluate the model performance by comparing the line/area curve with the observed replicates. Hence, a metric was formulated to evaluate the model estimated SMN line/area curve considering the observed replicant bounds (standard deviation) of SMN. The metric calculates the absolute error using the lower (LB) and upper bounds (UB) of observed replicates given if the SMN predicted values at the sampling stages were closer to the lower and upper bounds of replicates, respectively (Eq. (1)). This error was called the passing error (*pE*). Later, the root mean square of *pE* was calculated and normalized (*nRMSpE*) using average range of LB and UB of observed replicates for respective farm-years (Eq. (2)). Lower *nRMSpE* values are better and $nRMSpE = 0$ would be an ideal case when the estimated SMN line curve passed through the replicant bounds of the observed SMN on all the sampling dates. This metric was used to compare the performance of the hybrid-LSTM and the DSSAT models.

$$pE = \begin{cases} 0; if SMN_{LB} \leq SMN_{pred} \leq SMN_{UB} \\ min\left(|SMN_{pred} - SMN_{LB}|, |SMN_{pred} - SMN_{UB}|\right); else \end{cases} \tag{1}$$

$$nRMSpE = \frac{\sqrt{\sum_{i=1}^{n} (pE_i)^2}}{\overline{(SMN_{LB} - SMN_{UB})}} \tag{2}$$

Other statistical performance metrics- normalized- root mean square error (*nRMSE,* Eq. (3)), and R^2 (Eq. (4)) were used to assess the simulated results from the DSSAT and the hybrid-LSTM models (Laperre et al., 2020; Moriasi et al., 2007). R^2 values above 0.5 and *nRMSE/nRMSpE* values below 30 % were considered indicative of satisfactory model performance. Moreover, a percentage of improvement (*PI*) was calculated by comparing *nRMSE* (PI_{nRMSE}) and *nRMSpE* (PI_{nRMSpE}) of both models to find out which model performed better (Eq. (5)). Due to the inadequate observed data (such as tuber and plant dry weights and N accumulation, tuber fresh weight, and SMN), the simulated outputs were also evaluated based on visual interpretations (randomly picking graphs and manually checking if the simulated/estimated values aligned with observed data) to ensure the coherence of the model.

$$nRMSE = \frac{\sqrt{\sum_{i=1}^{n} \left(y_i^{obs} - y_i^{pred}\right)^2}}{\overline{y^{obs}}} \tag{3}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(y_i^{pred} - \overline{y^{obs}}\right)^2}{\sum_{i=1}^{n} \left(y_i^{obs} - \overline{y^{obs}}\right)^2} \tag{4}$$

$$PI_{nRMSE/nRMSpE} = \frac{nRMSE/nRMSpE_{DSSAT} - nRMSE/nRMSpE_{hybrid-LSTM}}{nRMSE/nRMSpE_{DSSAT}} \times 100 \tag{5}$$

3. Results and discussion

The hybrid-LSTM model, an integration of a process-based crop simulation model- DSSAT and a time-series DL model- LSTM, was able to learn the complexity between the input features (e.g., rainfall, average air temperature, average soil temperature, and applied N rates) and the temporal variability in the SMN. The generalized hybrid-LSTM model trained on various experimental sites in Hastings, FL was able to predict the SMN daily values with the improvement of 22.2 % (PI_{nRMSE}) and 27.2 % (PI_{nRMSpE}) compared to the DSSAT simulated SMN estimates across all the farms and years. These improvements indicate the effectiveness of the hybrid modeling approach used in this study. In the subsequent sections, the detailed performance analysis of the DSSAT model (calibration and evaluation) and the hybrid-LSTM model (training and testing) is provided.

3.1. DSSAT model performance

3.1.1. Potato growth simulation by DSSAT

To evaluate the performance of the SUBSTOR-potato model of DSSAT for potato growth and development, the simulated results on tuber dry weight (DW), plant DW, tuber N, plant N, and tuber yield fresh weight (FW) were compared against observed values (Fig. 2). The R^2 and *nRMSE* values for plant/tuber growth/N ranged from 0.75 to 0.92 and 26.3–32.0 % during the DSSAT model calibration. The DSSAT model performance during the evaluation was also similar to calibration with the R^2 and *nRMSE* values for plant/tuber growth/N ranging from 0.51 to 0.92 and 29.1–51.3 %, respectively. However, the DSSAT model underperformed in simulating tuber yield during calibration (R^2 = −0.08, *nRMSE* = 14.6 %), caused specifically due to consistent poor performance in all four farms in 2011. The DSSAT model underestimated the tuber yield in 2011 (average observed yield = 46.2 Mg ha^{-1}, average simulated yield = 23.9 Mg ha^{-1}). In contrast, the DSSAT model performance was satisfactory (R^2 = 0.56, *nRMSE* = 8.1 %) during evaluation in 2013 and 2014 for two farms (F1, F2). Raymundo et al. (2017) also simulated similar results using the SUBSTOR-potato model of DSSAT with *nRMSE* of 21.0 %, 37.2 %, 85.3 %, 40.4 %, and 86.3 %, respectively, for tuber FW, tuber DW, aboveground DW, tuber N update, and above ground N. Similarly, Wang et al. (2023) observed *nRMSE* ranging between 12.3 and 69.7 % when simulating tuber DW using the same model from a two-year experiment in China with different irrigation and fertigation levels.

A possible reason behind the DSSAT model's poor performance in 2011 could be attributed to the atypical rainfall distribution in that year (Fig. 3). The 2011 season received less rainfall (80 mm, average of all farms) compared to other years (130, 401, and 145 mm in 2012, 2013, and 2014, respectively) after tuber-initiation when there is a rapid increase in the plant N uptake due to tuber bulking. The reduced rainfall after tuber initiation in 2011 and the upward soil water flux from the subirrigation could have induced less N leaching from the topsoil compared to other years. Moreover, poor drainage conditions due to an impermeable layer below the surface require minimal irrigation to maintain a high-water table in dry conditions. In contrast, the current version of the DSSAT model could not replicate such a complex high-water table system in the sub-irrigated potato fields. Moreover, the model treats the subsurface irrigation as surface irrigation, and hence, the model might have overestimated the N leaching by applying more than needed irrigation water amid dry conditions (not considering the impermeable layer to hold the water table level) which could have underestimated the simulated tuber yield. Overall, the DSSAT performance in simulating the potato growth was satisfactory well.

3.1.2. SMN simulation by DSSAT

Figs. 3 and 4 show the comparison of the DSSAT simulated and observed SMN values to illustrate the model performance during calibration and evaluation on F3-2011 and F1-2013, respectively, for all the N fertilizer rate and timing treatments. The DSSAT model performance was satisfactory during calibration and evaluation for simulating the daily SMN values. The *nRMSE* and *nRMSpE* for the observed vs. DSSAT simulated SMN values were 16.0 % and 8.0 %, respectively in calibration (Table 4). Whereas the *nRMSE* and *nRMSpE* for the observed vs. DSSAT simulated SMN values were 12.8 % and 9.3 %, respectively in the

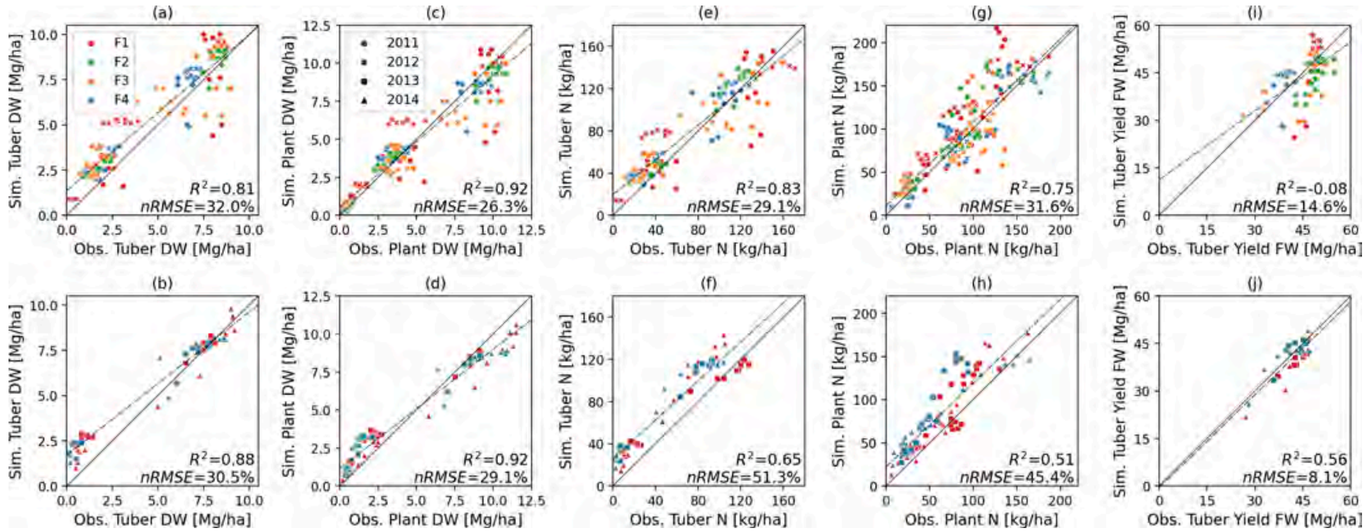

Fig. 2. Scatter plots comparing the observed and DSSAT – SUBSTOR Potato model simulated (a, b) tuber dry weight (DW), (c, d) plant DW, (e, f) tuber N, (g, h) plant N, and (i, j) tuber yield fresh weight (FW) while calibrating (first row, model calibrated using observations from four farms for 2011 and 2012) and evaluating (second row, model evaluated using observations from two farms for 2013 and 2014) the DSSAT model. [Note: Different colors of markers represent different farms whereas different shapes of markers represent different years, R^2 = Coefficient of determination, *nRMSE*=Normalized root mean squared error].

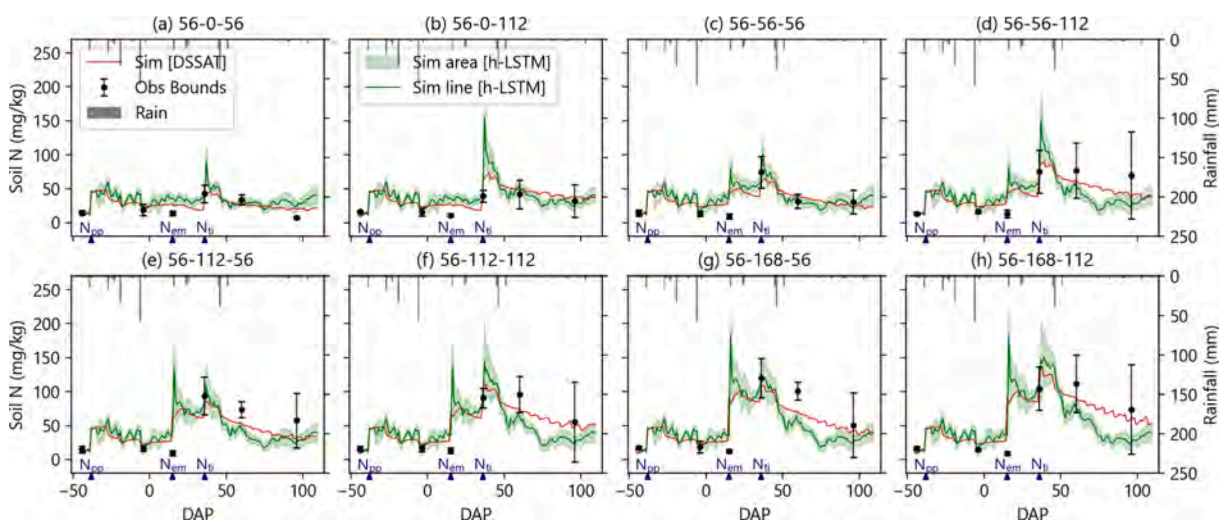

Fig. 3. Comparison of the observed (black error bars, error bar represents the standard deviation within the replicates), DSSAT simulated (red line) and hybrid-LSTM estimated (green line, green area curve) soil N concentration (0–15 cm) in F3-2011 for different fertilizer N rates and timing of application treatments (subplot title of (a) is 56-0-56 which means N_{pp} = 56 kg-N ha^{-1}, N_{em} = 0 kg-N ha^{-1}, and N_{ti} = 56 kg-N ha^{-1}, the solid blue triangles below x-axis are the days after planting (DAP) when N fertilizer was applied for N_{pp}, N_{em}, and N_{ti}) while training the hybrid-LSTM model [Sim = Simulated, Obs = Observed; N_{pp}, N_{em}, and N_{ti} = N fertilizer applied at planting, emergence, and tuber initiation, respectively].

model evaluation (Table 4). These results revealed an improvement in simulating SMN (nitrate-N and ammonium-N) over previous study by Raymundo et al. (2017) who observed *nRMSE* of 95.5 % and 140.1 % using the earlier version of the DSSAT model. Overall, the current version of DSSAT model (calibrated in this study) was able to simulate the daily variation of SMN, however, not the order of magnitude of variation of SMN. For example, in Fig. 3, the DSSAT simulated SMN curve followed the pattern closely in most of the treatments (Fig. 3). However, DSSAT simulated SMN values were consistently underestimated in all the treatments. The possible reason for such performance of the DSSAT model could be attributed to its inability to replicate SMN in a high-water table condition maintained by the subirrigation of the fields. The daily high-water table level oscillation might have influenced the dynamics of water and N movement in the sandy soil in the observed conditions (da Silva et al., 2018). In contrast, the DSSAT model performance in simulating the SMN trend was reasonably well in heavy and more uniformly distributed rainfall conditions (Figs. 3 and 4). For instance, the rainfall was more frequent and heavier after the tuber initiation (compared to before the tuber initiation) in F1-2013, and correspondingly, the DSSAT simulated SMN values also followed the same trend with observed SMN values after the tuber initiation N application (Fig. 4). This could be attributed due to limited need of irrigation in such conditions and the irrigation furrow would have been kept open to drain the excessive rainfall water. Overall, the DSSAT model was able to simulate the daily SMN values satisfactorily well given the continuous oscillation of the water table level and the presence of the sub-surface irrigation which is not supported as of the current version of the DSSAT model.

3.2. Performance of the hybrid-LSTM model in estimating SMN

Once the calibrated DSSAT model was evaluated, the hybrid-LSTM

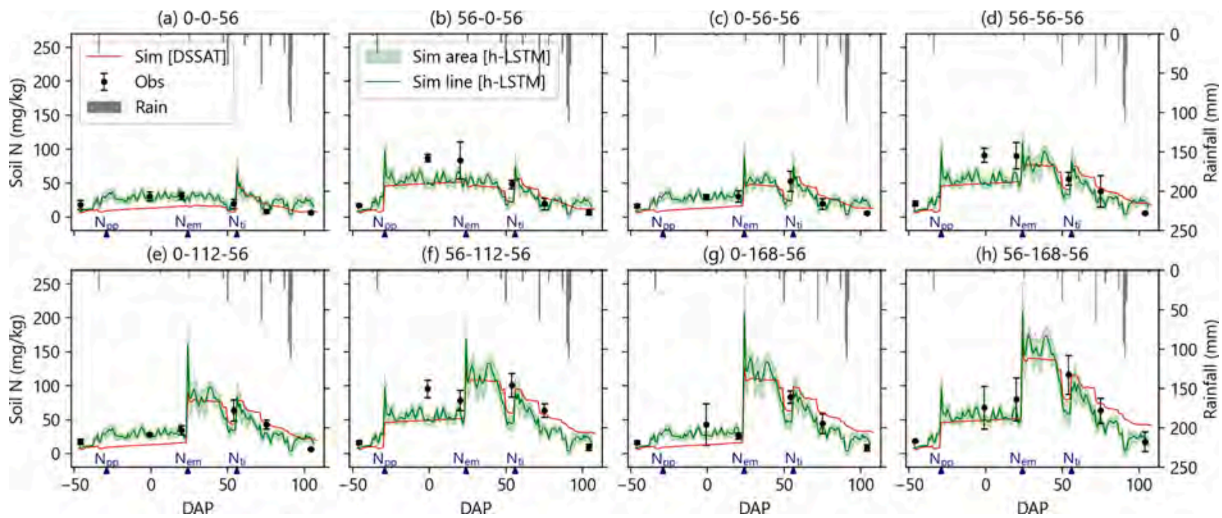

Fig. 4. Comparison of the observed (black error bars, error bar represents the standard deviation within the replicates), DSSAT simulated (red line) and hybrid-LSTM estimated (green line, green area curve) soil N concentration (0–15 cm) in F1-2013 for different fertilizer N rates and timing of application treatments (subplot title of (a) is 0–0-56 which means $N_{pp} = 0$ kg-N ha^{-1}, $N_{em} = 0$ kg-N ha^{-1}, and $N_{ti} = 56$ kg-N ha^{-1}, the solid blue triangles below x-axis are the days after planting (DAP) when N fertilizer was applied for N_{pp}, N_{em}, and N_{ti}) while testing the hybrid-LSTM model [Sim = Simulated, Obs = Observed; N_{pp}, N_{em}, and N_{ti} = N fertilizer applied at planting, emergence, and tuber initiation, respectively].

Table 4
Comparing the performance of the hybrid-LSTM (h-LSTM) and calibrated DSSAT model using the statistical metrics- *nRMSE* and *nRMSpE* for individual farms and years while hybrid-LSTM training and testing for estimating SMN daily time series.

	Farm-years	*nRMSE* (%)		*nRMSpE* (%)	
		h-LSTM*	DSSAT	h-LSTM	DSSAT
Training	F1-2011	13.0	17.7	8.8	12.0
	F2-2011	10.1	12.2	3.1	4.8
	F3-2011	12.8	21.4	6.2	12.5
	F2-2012	9.8	19.8	1.4	6.7
	F3-2012	15.0	23.7	5.4	12.2
	F4-2012	10.0	6.1	6.6	3.8
	F4-2013	11.9	16.5	7.2	11.7
	F4-2014	7.8	4.7	5.7	2.4
	Average	**11.3**	**15.3**	**5.6**	**8.3**
	PI	**26.1**		**32.5**	
Testing-1**	F4-2011	7.6	9.4	2.7	3.7
	F1-2012	12.4	17.6	4.3	8.0
	F1-2013	14.5	14.4	9.8	8.6
	F1-2014	12.1	15.6	10.5	14.6
	Average	**11.6**	**14.2**	**6.8**	**8.7**
	PI	**18.3**		**21.8**	
Testing-2***	F1-2011	9.7	19.1	4.6	11.5
	F2-2011	15.0	23.5	5.4	13.2
	F3-2011	16.5	26.5	8.2	15.2
	F1-2012	26.9	37.0	17.3	26.3
	F2-2012	26.5	38.7	13.9	23.5
	F3-2012	28.4	40.1	15.2	25.9
	F1-2013	20.6	25.2	15.5	19.1
	F1-2014	9.1	13.9	7.1	11.9
	F4-2014	11.4	11.9	8.5	8.4
	Average	**18.2**	**26.2**	**10.6**	**17.2**
	PI	**30.5**		**38.3**	

* h-LSTM means the hybrid-LSTM model.
** Testing-1 corresponds to testing results when the hybrid-LSTM was fed with unseen calibrated DSSAT simulated SMN.
*** Testing-2 corresponds to testing results when the hybrid-LSTM was fed with uncalibrated DSSAT simulated SMN.

model was developed using the calibrated DSSAT model simulated SMN and observed SMN samplings for the combination of various fertilizer N rate and timing of application treatments, and farm-years of datasets to improve the SMN estimates. The hybrid-LSTM model performance was then analyzed by feeding both unseen calibrated (section 3.2.1) and uncalibrated DSSAT simulated SMN (section 3.2.2) to determine whether it could improve the calibrated and uncalibrated DSSAT daily SMN simulations using field sampled SMN. Based on the analysis, as the developed hybrid-LSTM model had the knowledge about both the estimates given from the DSSAT model (daily estimates) and the ground truths (sampled estimates), it was able to provide improved estimate of the SMN over DSSAT simulated SMN on daily (Figs. 3 and 4) as well as on the sampling days (Fig. 5, Table 4).

3.2.1. The hybrid-LSTM model performance with unseen calibrated DSSAT simulated SMN

Figs. 3 and 4 demonstrate the comparison of the SMN values simulated from the hybrid-LSTM and the DSSAT model with observed SMN values for F3-2011 and F1-2013 farm-year combinations, respectively, while training and testing the hybrid-LSTM model (Please see supplementary Figs. S4 and S5 for the results of 2012 and 2014, respectively). Table 4 represents the performance comparison over two metrics of DSSAT simulated SMN and hybrid-LSTM predicted SMN for various farms and years (see section 2.5). As per the results, the hybrid-LSTM model was able to improve the daily SMN estimates over the DSSAT simulated SMN daily values in most of farm-years. The *nRMSE* and *nRMSpE* for SMN values estimated using the hybrid-LSTM model [*nRMSE*/*nRMSpE* = 11.3/5.5 % (training), *nRMSE*/*nRMSpE* = 11.6/6.8 % (testing)] were lower compared to the DSSAT simulated SMN values [*nRMSE*/*nRMSpE* = 12.3/11.3 % (training), *nRMSE*/*nRMSpE* = 14.2/11.6 % (testing)], both in training and testing the hybrid-LSTM model. Based on these metrics, the hybrid-LSTM model estimated the SMN with 26.1/32.5 % (PI_{nRMSE}/PI_{nRMSpE}) and 18.3/21.8 % improvement in training and testing farm-years over DSSAT simulations, respectively. In some years farm-years, the *nRMSE* and *nRMSpE* of the hybrid-LSTM model were higher compared to the DSSAT model. For instance, the *nRMSE*/*nRMSpE* increased from 6.1/3.8 % (DSSAT) to 10.0/6.6 % (hybrid-LSTM) and 4.7/2.4 % (DSSAT) to 7.8/5.7 % (hybrid-LSTM), for F4 in 2012 and 2014 used for training the DL model, respectively. Similarly, the *nRMSE* remained almost the same for DSSAT (14.4 %) and hybrid-LSTM models (14.5 %) while the *nRMSpE* increased slightly from

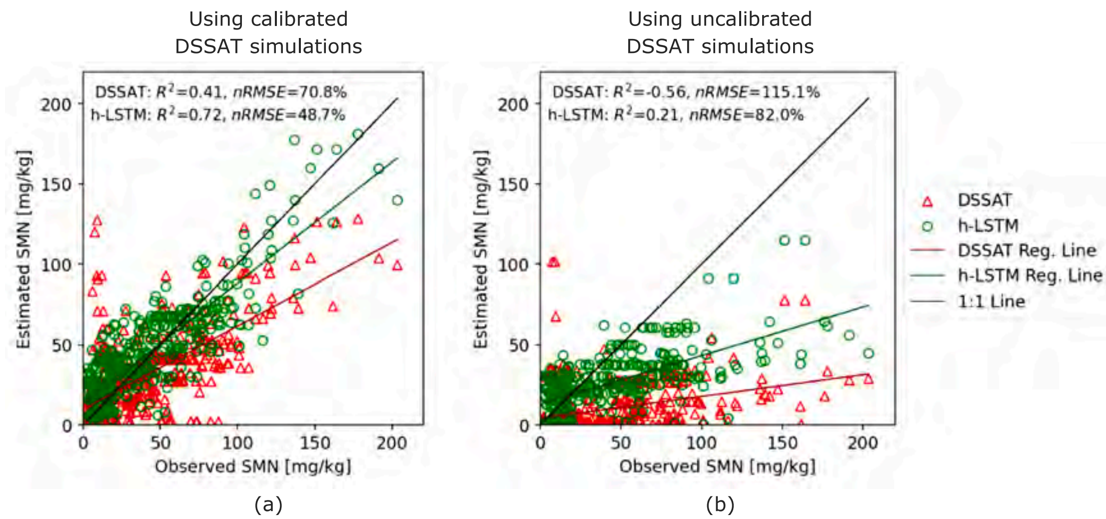

Fig. 5. Comparison of the hybrid-LSTM (h-LSTM) estimated and DSSAT simulated SMN against observed SMN at different sampling stages, treatments, and farm-years feeding (a) calibrated and (b) uncalibrated DSSAT simulated SMN in the hybrid-LSTM model. [Reg. = Regression, R^2 = Coefficient of determination, *nRMSE* = Normalized root mean squared error]

8.6 % (DSSAT) to 9.8 % (hybrid-LSTM) for F1 in 2013 used for testing the DL model. In all these farm-years (F4-2012, F1-2013, and F4-2014), the hybrid-LSTM model was unable to improve over the DSSAT simulated SMN daily estimates at the sampling stages. Later, the hybrid-LSTM estimated and DSSAT simulated SMN values were compared against those observed on only sampling day for all the farm-years (Fig. 5a). The comparison in Fig. 5a clearly illustrates that the hybrid-LSTM could outperform the DSSAT model in estimating SMN.

3.2.2. The hybrid-LSTM model performance with uncalibrated DSSAT simulated SMN

The hybrid-LSTM model was also evaluated for if the model could enhance the uncalibrated DSSAT simulated SMN for different farm-years. These farm-years were similar to what the model was already informed with while training/testing, the only difference being instead of inputting the calibrated DSSAT SMN simulations, the uncalibrated DSSAT SMN simulations were fed into the model. The hybrid-LSTM was able to improve the SMN estimates by 38.3 % (PI_{nRMSpE}) and 30.5 % (PI_{nRMSE}) over uncalibrated DSSAT simulated SMN, as seen in Table 4. The hybrid-LSTM outperformed the DSSAT model in estimating SMN in most farm-years except F4-2014 (please refer to the supplementary Fig. S6 comparing daily SMN estimated using the DSSAT and the hybrid-LSTM models against observed SMN). Fig. 5 illustrates a comparison between the hybrid-LSTM estimated and DSSAT simulated SMN against the observed SMN when feeding calibrated (Fig. 5a) and uncalibrated (Fig. 5b) DSSAT simulated SMN in the hybrid-LSTM model for all sampling stages, treatments, and farm-years. With these results, it can be interpreted that the hybrid-LSTM model could improve the DSSAT simulated SMN time-series even when uncalibrated DSSAT model simulated SMN values were fed in the hybrid-LSTM.

3.3. Discussion on the overall performance of the hybrid-LSTM model in estimating SMN

Overall, the hybrid-LSTM model outperformed across most farm-years compared to DSSAT. The hybrid-LSTM model improved on the DSSAT simulated SMN daily values regardless of whether the calibrated or uncalibrated DSSAT simulated SMN were fed into it. However, there were a very few exceptions where DSSAT simulations appeared slightly better.

These exceptions could be attributed to the limited SMN observations available for fine-tuning hybrid-LSTM model hyperparameters. Another reason for these exceptions could be due to generalized (same soil properties for all the farms) DSSAT calibration for all the farms which could have led to a distribution shift between both the datasets (DSSAT simulated and the observed SMN data). The generalized DSSAT model did not produce a similar trend for every farm compared to the observed SMN, especially in the case of F1-2011, F3-2012, and F4-2012. In these farms, the DSSAT simulated daily SMN did not align with the patterns observed SMN, despite the DSSAT model sometimes accurately predicted the daily SMN on a few sampling dates. It could be challenging to get a consistent performance with a generalized model in each farm with different soil properties (such as soil textural properties, soil water retention curve, soil organic matter content) (da Silva et al., 2018) under fluctuating water table level and irrigation water distribution in the root zone across the field. In this study, sandy soils were the predominant soil textural class with low soil organic matter. Even a small difference in their texture and organic matter content can change the soil water retention curve affecting the N movement in the soil. da Silva et al. (2018) reported different soil water retention curves due to sand content ranging from 81.9-91.6 % in nearby areas. In addition, a great variation in soil moisture is also expected due to the irrigation and drainage cycles during the crop season (Rens et al., 2022). These generalizations in using the DSSAT model generated a trend gap in the datasets (DSSAT simulated vs observed SMN) which might have hampered the overall hybrid-LSTM training. The hybrid-LSTM model trained on such an aberrated dataset would not always give consistent results. Consequently, refining these predictions with fine-tuning might bring the predictions closer to the experimental SMN values for that specific farm-year, but it could affect the overall learning for other farm-years.

Similarly, with cover crop inclusion every year, the soil organic matter from the cover crop residue could also vary among the farm-years and could cause differences in the SMN values with the changes in the net mineralization. Since there was no feature used to track these soil properties variations and cover crop growth dynamics among the farm-years, the hybrid-LSTM model could not incorporate the effect of these changes. Hence, these could be a few of the reasons that the hybrid-LSTM model could not improve the SMN daily estimates over DSSAT simulated SMN while being trained on the generalized DSSAT model simulations. The hybrid-LSTM model outperforms the DSSAT model for F1-2011 and F3-2012, while falling slightly short in the case of F4-2012.

Additionally, the influence of interannual factors on the performance of the hybrid-SLTM model was also investigated. The hybrid-LSTM model found to be performing better in 2011 and 2012 [PI_{nRMSE}/PI_{nRMSpE} = 30.0/39.4 % (training), PI_{nRMSE}/PI_{nRMSpE} = 25.9/40.2 % (testing)] compared to 2013 and 2014 [PI_{nRMSE}/PI_{nRMSpE} = 7.1/8.5 %

(training), $PI_{nRMSE}/PI_{nRMSpE} = 11.3/12.5$ % (testing)], in both training and testing. This could be attributed to the data imbalance between these two sets of years as 2011 and 2012 data originated from 4 farms (F1 to F4) while 2013 and 2014 included only 2 farms (F1 and F4). Alternatively, the DSSAT simulations for F2 and F3 in 2013 and 2014 using N application rates-timings treatments of F1 and F4 could be used to minimize the data imbalance that was leading to the interannual factors interfering in the simulation, however, the respective ground truth data was not available to fine-tune for these farms-years. To build a versatile DL model, it is crucial to have the model trained on various combinations of distributions of all the features. A limited number of farm-years might have restricted the hybrid-LSTM model to perceive such distribution variety and combinations among the features. SMN daily fluctuation depends on various features, primarily, N fertilizer application rates and timings, rainfall, N leaching, plant N uptake, irrigation, and water table depth fluctuation. While the features like fertilizer N application rates and timings, rainfall, and an interaction feature of rainfall and fertilizer application rates in the hybrid-LSTM model could have captured their direct influence on SMN. Average air and soil temperatures were included to indirectly capture the plant N uptake as plant growth is proportional of these temperatures. However, apart from the rainfall as a feature, there was no other feature incorporated to capture the anthropogenically managed water table in the hybrid-LSTM model. The current version of the DSSAT model did not have a water table management routine to simulate water-table conditioned sub-irrigated cropping system. While a manual water table depth management could have caused differences between DSSAT simulated and field observed SMN which challenged the fine-tuning process during the hybrid-LSTM training and hence, could have affected the hybrid-LSTM performance interannually. Rainfall distribution could also be one of crucial factors, as only four years of weather distribution and its interaction with multiple N application and plant growth would not be enough to capture its effect on SMN variability. To capture these dynamics, more years of field trials dataset would be required for training such hybrid models.

In summary, the hybrid modeling approach incorporating a time-series DL model and a crop simulation model DSSAT proved to be promising in estimating the SMN, especially between the sampling dates. While there were a few exceptions and inconsistencies in some farms and years generated due to several reasons mentioned above, these issues could be rectified with more data and improved data parsing approaches. The principal idea of the article was to propose a novel method of combining two models to develop the hybrid-LSTM model which could provide more accurate estimates of SMN in data-sparse agroecosystems. Although the hybrid-modeling approach relied on the simulated data from the DSSAT model (which requires field-specific configuration) to deliver precise estimates of daily SMN, making it inapplicable to different locations, soils, and crops, the approach served as a proof-of-concept. It demonstrated the potential to unravel the cropping systems with highly complex management practices where traditional crop models fail due to the under-representation of such practices, such as the high-water table conditioned subirrigation system in our case. Therefore, developing field-specific hybrid models could aid in understanding the dynamics of other soil nutrients, such as phosphorus, potassium, and sulfur in different cropping systems worldwide. In the future, the hybrid model developed could be improved by continuously training on newly available experiment observations from various farms and years. Moreover, encompassing a broader array of scenarios and facilitating a more comprehensive understanding of data patterns and features' relationships could further strengthen the modeling approach presented in this article.

4. Conclusions

The study was planned to present the novel approach of integrating the DSSAT model with the LSTM (hybrid-LSTM) to precisely estimate the SMN time series in the data-sparse high-water table conditioned sub-irrigated potato agroecosystem. The hybrid-LSTM model was able to estimate the daily SMN in such complex agroecosystem with a higher accuracy (reducing *nRMSE* between 18–30 % and *nRMSpE* between 21–38 %) than the DSSAT simulated SMN for most of the farms and years. However, there were a few farms and years where the hybrid-LSTM model could not outperform the DSSAT model. The possible reasons could be the limited experimental observations, incorporating a generalized modeling approach for the farms with slightly different soils, and inconsistencies in the DSSAT simulated SMN with observations on which the hybrid-LSTM model was initially trained. The limitation of this approach is its inapplicability for different locations as the approach uses simulated data from the DSSAT model which requires field-specific configurations. Overall, the hybrid modeling approach produced a more reliable SMN time series which could aid in developing sustainable N management and provide optimized N fertilizer recommendations minimizing environmental damages for various cropping systems around the world. The proposed modeling strategy could be adapted to estimate the daily time series of other soil nutrients, such as phosphorus, potassium, and sulfur.

CRediT authorship contribution statement

Rishabh Gupta: Writing – original draft, Visualization, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Satya K. Pothapragada:** Writing – original draft, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Weihuang Xu:** Methodology, Investigation, Conceptualization. **Prateek Kumar Goel:** Data curation. **Miguel A. Barrera:** Data curation. **Mira S. Saldanha:** Data curation. **Joel B. Harley:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization. **Kelly T. Morgan:** Project administration, Funding acquisition. **Alina Zare:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization. **Lincoln Zotarelli:** Writing – review & editing, Project administration, Methodology, Funding acquisition, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgements

This project was funded by the Florida Department of Agriculture and Consumer Services (Award AWD10728, AWD12838, and AWD15066). The authors acknowledge the support of Judyson de Matos Oliveira, Fernando Rodrigo Bortolozo, and Ayesha Naikodi.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compag.2024.109355.

References

Akhavizadegan, F., Ansarifar, J., Wang, L., Huber, I., Archontoulis, S.V., 2021. A time-dependent parameter estimation framework for crop modeling. Sci. Rep. 11 https://doi.org/10.1038/s41598-021-90835-x.

Archontoulis, S.V., Miguez, F.E., Moore, K.J., 2014. Evaluating APSIM maize, soil water, soil nitrogen, manure, and soil temperature modules in the Midwestern United States. Agron. J. 106, 1025–1040. https://doi.org/10.2134/agronj2013.0421.

da Silva, A.L.B.R., Dias, H.B., Gupta, R., Zotarelli, L., Asseng, S., Dukes, M.D., Porter, C., Hoogenboom, G., 2024. Assessing the impact of irrigation and nitrogen management on potato performance under varying climate in the state of Florida, USA. Agric. Water Manag. 295, 108769. https://doi.org/10.1016/j.agwat.2024.108769.

da Silva, A.L.B.R., Hashiguti, H.T., Zotarelli, L., Migliaccio, K.W., Dukes, M.D., 2018. Soil water dynamics of shallow water table soils cultivated with potato crop. Vadose Zone J. 17, 1–15. https://doi.org/10.2136/vzj2018.04.0077.

da Silva, A.L.B.R., Zotarelli, L., Dukes, M.D., van Santen, E., Asseng, S., 2023. Nitrogen fertilizer rate and timing of application for potato under different irrigation methods. Agric. Water Manag. 283, 108312 https://doi.org/10.1016/j.agwat.2023.108312.

Datta, P., Faroughi, S.A., 2023. A multihead LSTM technique for prognostic prediction of soil moisture. Geoderma 433, 116452. https://doi.org/10.1016/j.geoderma.2023.116452.

Dukes, M.D., Zotarelli, L., Morgan, K.T., 2010. Use of irrigation technologies for vegetable crops in Florida. HortTechnology Hortte 20, 133–142. https://doi.org/10.21273/HORTTECH.20.1.133.

Errebhi, M., Rosen, C.J., Gupta, S.C., Birong, D.E., 1998. Potato yield response and nitrate leaching as influenced by nitrogen management. Agron. J. 90, 10–15. https://doi.org/10.2134/agronj1998.00021962009000010003x.

Everingham, Y., Sexton, J., Skocaj, D., Inman-Bamber, G., 2016. Accurate prediction of sugarcane yield using a random forest algorithm. Agron. Sustain. Dev. 36, 27. https://doi.org/10.1007/s13593-016-0364-z.

FAWN, 2022. Florida Automated Weather Network [WWW Document]. University of Florida Institute of Food and Agricultural Sciences (UF/IFAS). URL https://fawn.ifas.ufl.edu/ (accessed 12.19.22).

Feng, P., Wang, B., Liu, D.L., Waters, C., Xiao, D., Shi, L., Yu, Q., 2020. Dynamic wheat yield forecasts are improved by a hybrid approach using a biophysical model and machine learning technique. Agric. For. Meteorol. 285–286, 107922 https://doi.org/10.1016/j.agrformet.2020.107922.

Gauch, M., Kratzert, F., Klotz, D., Nearing, G., Lin, J., Hochreiter, S., 2021. Rainfall-runoff prediction at multiple timescales with a single Long Short-Term Memory network. Hydrol. Earth Syst. Sci. 25, 2045–2062. https://doi.org/10.5194/hess-25-2045-2021.

Gijsman, A.J., Hoogenboom, G., Parton, W.J., Kerridge, P.C., 2002. Modifying DSSAT crop models for low-input agricultural systems using a soil organic matter–residue module from CENTURY. Agron. J. 94, 462. https://doi.org/10.2134/agronj2002.0462.

Gray, D.D., Katz, P.G., deMonsabert, S.M., Cogo, N.P., 1982. Antecedent moisture condition probabilities. J. Irrig. Drain. Eng. 108, 107–114.

Griffin, T.S., Johnson, B.S., Ritchie, J.T., 1993. A simulation model for potato growth and development: SUBSTOR-Potato Version 2.0. Honolulu, HI.

Gupta, R., Bhattarai, R., Coppess, J.W., Jeong, H., Ruffatti, M., Armstrong, S.D., 2022. Modeling the impact of winter cover crop on tile drainage and nitrate loss using DSSAT model. Agric. Water Manag. 272, 107862 https://doi.org/10.1016/j.agwat.2022.107862.

Hadipour-Rokni, R., Askari Asli-Ardeh, E., Jahanbakhshi, A., Esmaili paeen-Afrakoti, I., Sabzi, S., 2023. Intelligent detection of citrus fruit pests using machine vision system and convolutional neural network through transfer learning technique. Comput. Biol. Med. 155 https://doi.org/10.1016/j.compbiomed.2023.106611.

Hochreiter, S., Schmidhuber, J., 1997. Long short-term memory. Neural Comput. 9, 1735–1780. https://doi.org/10.1162/neco.1997.9.8.1735.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62, 327–350. https://doi.org/10.1016/j.envsoft.2014.07.009.

Hoogenboom, G., Porter, C.H., Boote, Kenneth J., Shelia, V., Wilkens, P.W., Singh, U., White, J.W., Asseng, S., Lizaso, J.I., Moreno, L.P., Pavan, W., Ogoshi, R., Hunt, L.A., Tsuji, G.Y., Jones, J.W., 2019. The DSSAT crop modeling ecosystem. In: Boote, K.J. (Ed.), Advances in Crop Modeling for a Sustainable Agriculture. Burleigh Dodds Science Publishing, Cambridge, United Kingdom, pp. 173–216. doi: 10.19103/as.2019.0061.10.

Hoogenboom, G., Porter, C.H., Shelia, V., Boote, K.J., Singh, U., Pavan, W., Oliveira, F.A.A., Moreno-Cadena, L.P., Ferreira, T.B., White, J.W., Lizaso, J.I., Pequeno, D.N.L., Kimball, B.A., Alderman, P.A., Thorp, K.R., Cuadra, S.V., Vianna, M.S., Villalobos, F.J., Batchelor, W.D., Asseng, S., Jones, M.R., Hopf, A., Dias, H.B., Hunt, L.A., Jones, J.W., 2023. Decision Support System for Agrotechnology Transfer (DSSAT) Version 4.8.0.

Hristine, C., Orthington, M.W., Had, C., Utchinson, M.H., 2005. Accumulated growing degree days as a model to determine key developmental stages and evaluate yield and quality of potato in northeast Florida, Proc. Fla. State Hort. Soc.

Hua, Y., Zhao, Z., Li, R., Chen, X., Liu, Z., Zhang, H., 2019. Deep learning with long short-term memory for time series prediction. IEEE Commun. Mag. 57, 114–119. https://doi.org/10.1109/MCOM.2019.1800155.

Ilniyaz, O., Du, Q., Shen, H., He, W., Feng, L., Azadi, H., Kurban, A., Chen, X., 2023. Leaf area index estimation of pergola-trained vineyards in arid regions using classical and deep learning methods based on UAV-based RGB images. Comput. Electron. Agric. 207 https://doi.org/10.1016/j.compag.2023.107723.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Wilkens, P.W., Singh, U., Gijsman, A.J., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron. 18, 235–265. https://doi.org/10.1016/S1161-0301(02)00107-7.

Kamilaris, A., Prenafeta-Boldú, F.X., 2018. Deep learning in agriculture: a survey. Comput. Electron. Agric. 147, 70–90. https://doi.org/10.1016/j.compag.2018.02.016.

Keating, B.A., Carberry, P.S., Hammer, G.L., Probert, M.E., Robertson, M.J., Holzworth, D., Huth, N.I., Hargreaves, J.N.G., Meinke, H., Hochman, Z., McLean, G., Verburg, K., Snow, V., Dimes, J.P., Silburn, M., Wang, E., Brown, S., Bristow, K.L., Asseng, S., Chapman, S., McCown, R.L., Freebairn, D.M., Smith, C.J., 2003. An overview of APSIM, a model designed for farming systems simulation. Eur. J. Agron. 18, 267–288. https://doi.org/10.1016/S1161-0301(02)00108-9.

Laperre, B., Amaya, J., Lapenta, G., 2020. Dynamic time warping as a new evaluation for dst forecast with machine learning. Front.Astron. Space Sci. 7 https://doi.org/10.3389/fspas.2020.00039.

Liu, H.L., Yang, J.Y., Tan, C.S., Drury, C.F., Reynolds, W.D., Zhang, T.Q., Bai, Y.L., Jin, J., He, P., Hoogenboom, G., 2011. Simulating water content, crop yield and nitrate-N loss under free and controlled tile drainage with subsurface irrigation using the DSSAT model. Agric. Water Manag. 98, 1105–1111. https://doi.org/10.1016/j.agwat.2011.01.017.

LSTM architecture, 2021. Insights into LSTM architecture [WWW Document]. URL https://thorirmar.com/post/insight_into_lstm/.

Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Front. Plant Sci. 7 https://doi.org/10.3389/fpls.2016.01419.

Moriasi, D.N., Arnold, J.G., Liew, M.W.V., Bingner, R.L., Harmel, R.D., Veith, T.L., 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Trans. ASABE 50. https://doi.org/10.13031/2013.23153.

Moriasi, D.N., Gowda, P.H., Arnold, J.G., Mulla, D.J., Ale, S., Steiner, J.L., 2013. Agricultural water management modeling the impact of nitrogen fertilizer application and tile drain configuration on nitrate leaching using SWAT. Agric. Water Manag. 130, 36–43. https://doi.org/10.1016/j.agwat.2013.08.003.

NASS - Quick Stats, 2021. USDA National Agricultural Statistics Service - Quick Stats [WWW Document]. United States Department of Agriculture, National Agricultural Statistics Service. URL https://data.nal.usda.gov/dataset/nass-quick-stats (accessed 3.19.23).

Padarian, J., Minasny, B., McBratney, A.B., 2019. Using deep learning to predict soil properties from regional spectral data. Geoderma Reg. 16 https://doi.org/10.1016/j.geodrs.2018.e00198.

Raymundo, R., Asseng, S., Prassad, R., Kleinwechter, U., Concha, J., Condori, B., Bowen, W., Wolf, J., Olesen, J.E., Dong, Q., Zotarelli, L., Gastelo, M., Alva, A., Travasso, M., Quiroz, R., Arora, V., Graham, W., Porter, C., 2017. Performance of the SUBSTOR-potato model across contrasting growing conditions. Field Crops Res. 202, 57–76. https://doi.org/10.1016/j.fcr.2016.04.012.

Rens, L.R., Zotarelli, L., Cantliffe, D.J., Stoffella, P.J., Gergela, D., Burhans, D., 2015a. Rate and timing of nitrogen fertilizer application on potato 'FL1867' part II: Marketable yield and tuber quality. Field Crops Res 183, 267–275. https://doi.org/10.1016/j.fcr.2015.08.008.

Rens, L.R., Zotarelli, L., Cantliffe, D.J., Stoffella, P.J., Gergela, D., Fourman, D., 2015b. Biomass accumulation, marketable yield, and quality of Atlantic potato in response to nitrogen. Agron. J. 107, 931–942. https://doi.org/10.2134/agronj14.0408.

Rens, L.R., Zotarelli, L., Alva, A., Rowland, D., Liu, G., Morgan, K., 2016a. Fertilizer nitrogen uptake efficiencies for potato as influenced by application timing. Nutr. Cycl. Agroecosyst. 104, 175–185. https://doi.org/10.1007/s10705-016-9765-2.

Rens, L.R., Zotarelli, L., Cantliffe, D.J., Stoffella, P.J., Gergela, D., Burhans, D., 2016b. Commercial evaluation of seasonal distribution of nitrogen fertilizer for potato. Potato Res. 59, 1–20. https://doi.org/10.1007/s11540-015-9304-6.

Rens, L.R., Zotarelli, L., Rowland, D.L., Morgan, K.T., 2018. Optimizing nitrogen fertilizer rates and time of application for potatoes under seepage irrigation. Field Crops Res. 215, 49–58. https://doi.org/10.1016/j.fcr.2017.10.004.

Rens, L.R., Zotarelli, L., Ribeiro da Silva, A.L.B., Ferreira, C.J.B., Tormena, C.A., Rowland, D.L., Morgan, K.T., 2022. Managing water table depth thresholds for potato subirrigation. Agric. Water Manag. 259 https://doi.org/10.1016/j.agwat.2021.107236.

Reyes-Cabrera, J., Zotarelli, L., Dukes, M.D., Rowland, D.L., Sargent, S.A., 2016. Soil moisture distribution under drip irrigation and seepage for potato production. Agric. Water Manag. 169, 183–192. https://doi.org/10.1016/j.agwat.2016.03.001.

Ritchie, J.T., Singh, U., Godwin, D.C., Bowen, W.T., 1998. Cereal growth, development and yield. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Springer Netherlands, Dordrecht, pp. 79–98. doi: 10.1007/978-94-017-3624-4_5.

Ritchie, J.T., Griffin, T.S., Johnson, B.S., 1995. SUBSTOR: functional model of potato growth, development and yield. In: Kabat, P., Marshall, B., van den Broek, B.J., Vos, J., van Keulen, H. (Eds.), Modelling and Parameterization of the Soil-Plant-Atmosphere System: A Comparison of Potato Growth Models. Wageningen Pers, Wageningen, Netherlands, pp. 401–435.

Saha, G., Rahmani, F., Shen, C., Li, L., Raj, C., 2023. A deep learning-based novel approach to generate continuous daily stream nitrate concentration for nitrate data-sparse watersheds. Sci. Total Environ. 162930 https://doi.org/10.1016/j.scitotenv.2023.162930.

Salmerón, M., Cavero, J., Isla, R., Porter, C.H., Jones, J.W., Boote, K.J., 2014. DSSAT nitrogen cycle simulation of cover crop-maize rotations under irrigated mediterranean conditions. Agron. J. 106, 1283–1296. https://doi.org/10.2134/agronj13.0560.

Saseendran, S.A., Ma, L., Malone, R., Heilman, P., Ahuja, L.R., Kanwar, R.S., Karlen, D.L., Hoogenboom, G., 2007. Simulating management effects on crop production, tile

drainage, and water quality using RZWQM-DSSAT. Geoderma 140, 297–309. https://doi.org/10.1016/j.geoderma.2007.04.013.
Scholberg, J.M.S., Zotarelli, L., Dukes, M.D., Ozores-Hampton, M., Liu, G., Tittonell, P., 2013. Enhancing fertilizer efficiency in high input cropping systems in Florida. In: Lichtfouse, E. (Ed.), Sustainable Agriculture Reviews: Volume 12. Springer Netherlands, Dordrecht, pp. 143–174. doi: 10.1007/978-94-007-5961-9_5.
Seidel, S.J., Palosuo, T., Thorburn, P., Wallach, D., 2018. Towards improved calibration of crop models – Where are we now and where should we go? Eur. J. Agron. 94, 25–35. https://doi.org/10.1016/j.eja.2018.01.006.
Shahhosseini, M., Hu, G., Huber, I., Archontoulis, S.V., 2021. Coupling machine learning and crop modeling improves crop yield prediction in the US Corn Belt. Sci. Rep. 11, 1606. https://doi.org/10.1038/s41598-020-80820-1.
Sharma, G., Singh, A., Jain, S., 2022. Hybrid deep learning techniques for estimation of daily crop evapotranspiration using limited climate data. Comput. Electron. Agric. 202, 107338 https://doi.org/10.1016/j.compag.2022.107338.
Simonne, E., Hutchinson, C., Devalerio, J., Hochmuth, R., Treadwell, D., Wright, A., Santos, B., Whidden, A., Mcavoy, G., Zhao, X., Olczyk, T., Gazula, A., Ozores-Hampton, M., 2010. Current knowledge, gaps, and future needs for keeping water and nutrients in the root zone of vegetables grown in Florida. HortTechnology 20, 143–152. https://doi.org/10.21273/HORTTECH.20.1.143.
SSURGO Database, 2018. Soil Survey Staff, Soil Survey Geographic (SSURGO) Database [WWW Document]. United States Department of Agriculture, Natural Resources Conservation Service. URL https://sdmdataaccess.sc.egov.usda.gov.
Staudemeyer, R.C., Morris, E.R., 2019. Understanding LSTM - a tutorial into Long Short-Term Memory Recurrent Neural Networks. CoRR abs/1909.09586.
USDA-NASS, 2019. 2017 Census of Agriculture: Florida State and County Data - Chapter 2.
van Klompenburg, T., Kassahun, A., Catal, C., 2020. Crop yield prediction using machine learning: a systematic literature review. Comput. Electron. Agric. 177, 105709 https://doi.org/10.1016/j.compag.2020.105709.
Wallach, D., Thorburn, P.J., 2014. The error in agricultural systems model prediction depends on the variable being predicted. Environ. Model. Softw. 62, 487–494. https://doi.org/10.1016/j.envsoft.2014.08.001.
Wang, H., Cheng, M., Liao, Z., Guo, J., Zhang, F., Fan, J., Feng, H., Yang, Q., Wu, L., Wang, X., 2023. Performance evaluation of AquaCrop and DSSAT-SUBSTOR-Potato models in simulating potato growth, yield and water productivity under various drip fertigation regimes. Agric. Water Manag. 276 https://doi.org/10.1016/j.agwat.2022.108076.
Wang, C., Pan, W., Song, X., Yu, H., Zhu, J., Liu, P., Li, X., 2022. Predicting plant growth and development using time-series images. Agronomy 12. https://doi.org/10.3390/agronomy12092213.
Weiss, K., Khoshgoftaar, T.M., Wang, D., 2016. A survey of transfer learning. J Big Data 3, 9. https://doi.org/10.1186/s40537-016-0043-6.
Yue, Y., Li, J.-H., Fan, L.-F., Zhang, L.-L., Zhao, P.-F., Zhou, Q., Wang, N., Wang, Z.-Y., Huang, L., Dong, X.-H., 2020. Prediction of maize growth stages based on deep learning. Comput. Electron. Agric. 172, 105351 https://doi.org/10.1016/j.compag.2020.105351.
Zhang, L., Cai, Y., Huang, H., Li, A., Yang, L., Zhou, C., 2022. A CNN-LSTM model for soil organic carbon content prediction with long time series of MODIS-based phenological variables. Remote Sens. (Basel) 14. https://doi.org/10.3390/rs14184441.
Zhang, N., Zhou, X., Kang, M., Hu, B.-G., Heuvelink, E., Marcelis, L.F.M., 2023. Machine learning versus crop growth models: an ally, not a rival. AoB Plants 15, plac061. https://doi.org/10.1093/aobpla/plac061.
Zotarelli, L., Scholberg, J.M., Dukes, M.D., Muñoz-Carpena, R., 2007. Monitoring of nitrate leaching in sandy soils. J. Environ. Qual. 36, 953–962. https://doi.org/10.2134/jeq2006.0292.
Zotarelli, L., Rens, L.R., Cantliffe, D.J., Stoffella, P.J., Gergela, D., Fourman, D., 2014. Nitrogen fertilizer rate and application timing for chipping potato cultivar atlantic. Agron. J. 106, 2215–2226. https://doi.org/10.2134/agronj14.0193.
Zotarelli, L., Rens, L.R., Cantliffe, D.J., Stoffella, P.J., Gergela, D., Burhans, D., 2015. Rate and timing of nitrogen fertilizer application on potato 'FL1867'. Part I: Plant nitrogen uptake and soil nitrogen availability. Field Crops Res 183, 246–256. https://doi.org/10.1016/j.fcr.2015.08.007.

RISHABH GUPTA, Ph.D.
Postdoctoral Research Associate
Horticultural Sciences Department (HOS), Institute of Food and Agricultural Sciences (IFAS)
University of Florida (UF), Gainesville, Florida (FL)
Email: gupta.rishabh.61@outlook.com, Ph: (217) 904-6041 (c)

Search Committee Chair
Department of Agricultural and Biological Engineering (ABE)
Everglades Research and Education Center (EREC)
IFAS, UF – Belle Glade, FL

Feb 10, 2024

Dear Search Committee Members,

I am writing to apply for the position of **Assistant Professor – Agricultural Water Management**, a 9-month tenure-track position, with research (60%) and extension (40%) responsibilities at EREC, IFAS-UF. Currently, I am working as a Postdoctoral Research Associate at the HOS Department, IFAS-UF with Dr. Zotarelli. I graduated with a Ph.D. in Agricultural and Biological Engineering at the University of Illinois Urbana-Champaign (UIUC) under the supervision of Dr. Bhattarai. I believe that my expertise in integrating various process-based models and artificial intelligence (AI) for effective water table management to improve and sustain agricultural production, uniquely qualifies me for this tenure-accruing role.

The position description emphasizes the key elements of developing internationally recognized, extramurally funded research and extension programs to address major challenges towards water resources and quality management for long-term agricultural sustainability in the Everglade area by collaborating with dynamics researchers and faculties of IFAS-UF. Developing such strong interdisciplinary extension-focused research programs requires deep foundational knowledge on plant-soil-climate nexus and soil nutrients-water management in sandy soils, irrigation and drainage cycles in shallow water table areas to effectively address stakeholders' needs. Additionally, expertise in synergizing state-of-the-art modeling methods, such as watershed modeling, crop modeling, climate modeling, decision support systems, data science, AI, applied machine learning (ML), remote sensing and GIS techniques–is crucial to excel in this role. These areas are integral to my professional, educational, and research background.

As a Postdoctoral Research Associate at UF, I am working on a multidisciplinary collaborative project funded by state legislature entitled *'Using AI for Improved Crop Nutrient Recommendation'* collaborating with Drs. Zare and Harley from the Department of Electrical & Computer Engineering. I am leading an interdisciplinary team of undergraduate/graduate students with multiple objectives 1) to establish a uniform data collection framework for agricultural field trials for future AI efforts, 2) to advance AI algorithms merging crop models and drone-/satellite-based multi-/hyperspectral imageries, and 3) to utilize AI to provide site-specific nutrient recommendations for potato grown under shallow water table condition, balancing productivity and nutrient leaching in uncertain Florida weather. While working as a postdoc, so far, I submitted 4 grants (1 accepted in FDACS, 1 pending in USDA-NIFA, 1 pending in SCBGP, FDACS, 1 pending in IFAS-UF (LIFT-AI), published 2 journal articles, 1 extension article, and 3 conference presentations (oral). Recently, I am collaborating to combine advanced AI methods and soil hydraulic models to develop site-specific irrigation and drainage scheduling for optimum water table recommendations to conserve water resources, reduce nutrient loss, and improve crop production. Moreover, I am engaged in developing remote sensing-based AI methods to reduce our dependance on destructive soil and plant sampling in a long-term for developing site-specific best management practices

(BMPs). Additionally, I actively participate in the DSSAT Development Sprint led by Prof. Hoogenboom and developed various tools and libraries such as *dssat-pylib* and *DSSAT-SoilPro*. In my role as a postdoc, I gained extensive experience in integrating advanced ML and data science methods with process-based agroecosystem models, leveraging sparse in-field observations for sustainable development. I developed crucial skills required to lead a diverse team that would be invaluable in advancing data-driven water resources management approaches for improved site-specific agricultural systems climate resiliency as a faculty member to lead BMP programs at EREC.

For my doctoral research at UIUC, I worked on a research project entitled-*'CoverCrop Analyzer: a web-based decision support tool for cover crop management'* for farmers, where I collaborated with various software/web developers, agronomists, climatologists, farmers, and stakeholders to deploy the DSSAT model in the web application providing field-specific cover crop management for reducing nutrient loss and improving water quality in Illinois. My Ph.D. dissertation aimed to investigate the long-term sustainability of cereal rye as a winter cover crop in improving drainage water quality with changing climate conditions in the Illinois cropping system using a large-scale geospatial crop modeling approach accompanied by high-performance computational (HPC) resources. Moreover, I collaborated to investigate the impacts of global climate change on stream low flows in the Great Miami River watershed, Ohio using a hydrological model- SWAT. During my master's research, I assessed the climate change-induced impact and uncertainty of rice yield in agro-ecological zones of India. From my doctoral and mater's research, I published 6 journal articles, 3 extension articles, 6 conference presentations (1 poster, 5 oral), and 1 conference proceeding. Having actively engaged in climate change-focused research projects, I am well-equipped to collaborate seamlessly with colleagues at UF to lead a constructive research program in developing and evaluating various carbon-neutral water resources conservation practices by leveraging my experience in HPC resources and large-scale geospatial climate modeling.

I believe that an effective extension program bridges the gap between research findings and practical application in the agricultural community. I actively participated in various extension activities such as Potato field days conducted at Hastings Agricultural Extension Center (HAEC), FL and in-service trainings (IST). I observed a tendency of farmers to over-drain or over-irrigate their fields which cause either excessive nutrient loss or anaerobic conditions in crop root zone, resulting in yield losses. With my extension program, I aim to provide practical and actionable site-specific solutions to farmers, policymakers, and other stakeholders, to improve agricultural productivity by maintaining optimum soil moisture while conserving soil, nutrients, and water (quantity and quality). I aim to drive my research-based extension program to educate among farmers, county, and extension agents in four steps approach: 1) **Stakeholder engagement and need assessment**, 2) **Demonstration and knowledge transfer**, 3) **Educational outreach** and 4) **Program Evaluation** to improve soil, nutrients, and water conservation literacy and agricultural sustainability. My approach would emphasize the transfer of cutting-edge research knowledge and technologies to end-users, empowering them to make informed decisions for BMPs.

In general, I understand rapidly evolving challenges in food production systems including but not limited to rising global temperatures, shifting rainfall patterns, increasing greenhouse gas emissions, and El Niño-Southern Oscillations phenomenon-driven climate anomalies. Moreover, Florida's unique topography with low-lying landscapes, sandy soils, and surficial aquifer system challenges an effective soil, nutrients, and water resources management to improve vegetables production. To address these challenges, I envision my research and extension program in four primary areas: 1) **develop and evaluate site- and crop-specific irrigation and drainage systems** for sandy soils with shallow water-table condition to improve agricultural production and soil-nutrient-water conservation, 2) **advance AI-integrated scalable crop, hydrological, and nutrient transport modeling approaches** to enhance crop water-nutrient use efficiency while reducing nutrient losses and soil erosion amid water table management, 3) **develop web-based decision support system** to improve the state climate resilience by data sciences, remote sensing, and geospatial modeling approaches, and 4) **implement and deliver these strategies** by collaborating with growers, extension agents, and other stakeholders to improve drainage water quality by reducing nutrient and sediment loadings to water bodies while improving crop yield. Addressing these issues requires not only cutting-edge research but also equipping future scholars, engineers, and researchers with the technological expertise necessary to meet these demands. I am committed to building an internationally recognized research program and generating peer-reviewed journal/extension articles and scholarships to showcase my research activities and integrate my research case studies into the extension programs. I am well-equipped to secure competitive external funding to support my research and educational programs.

Apart from my technical expertise, I have a strong track record of leadership in academic and professional settings. I have presented at numerous national/international conferences and published in internationally renowned journals which significantly enhanced my communication, presentation, and writing abilities. Moreover, I am committed to chairing and serving on graduate committees, advising, and mentoring undergraduate/graduate students, postdoctoral scholars, and research technicians. Additionally, I am prepared for any teaching responsibilities to actively contribute to all three components of the land-grant mission- Research, Teaching, and Extension. In my pursuit of continuous improvement, I have actively engaged in professional development. I successfully completed two workshops on '*Teaching for inclusivity and accessibility*' and the '*Preparing Future Faculty*' program at UF. The insights acquired from these workshops extend beyond the realm of teaching and can be effectively integrated into my research and extension programs, creating inclusive and accessible learning environment.

In summary, I believe that I am highly qualified for the position- Assistant Professor – Agricultural Water Management. Working seamlessly within multidisciplinary teams has prepared me to excel as a researcher and an educator. I planned to incorporate multidisciplinary cutting-edge methodologies by collaborating with faculties, students from various departments, and other stakeholders. Given my expertise in developing effective nutrient and water management strategies using scalable modeling, data sciences, knowledge-guided ML, remote sensing, and GIS techniques, I am confident in my ability to advance high-impact research, and extension programs for sustainable agricultural developments. I am fully devoted to creating meaningful and impactful contributions to the university's academic and research endeavors promoting the land-grant missions.

I have enclosed my curriculum vitae, research statement, extension philosophy, contact information for professional references, and unofficial transcripts. I am eager to discuss my qualifications further and answer any questions you may have. I look forward to hearing from you. Thank you for your consideration.

Sincerely,

Rishabh Gupta

Rishabh Gupta

RISHABH GUPTA, Ph.D.
Postdoctoral Research Associate
Horticultural Sciences Department, Institute of Food and Agricultural Sciences (IFAS)
University of Florida (UF), Gainesville, Florida (FL)
Email: gupta.rishabh.61@outlook.com, Ph: (217) 904-6041 (c)
Google Scholar | ResearchGate | GitHub | LinkedIn

EDUCATION

University of Illinois at Urbana-Champaign
***Ph.D.* - Agricultural Engineering** (Soil and Water Resources Engineering) **2022**

Indian Institute of Technology - Kharagpur
***M.Tech.* - Agricultural Engineering** (Land and Water Resources Engineering) **2016**

Jawaharlal Nehru Agricultural University
***B.Tech.* - Agricultural Engineering** **2014**

RESEARCH EXPERIENCE

Postdoctoral Research Associate, University of Florida **2022 - Present**
Horticultural Sciences Department
Advisor: Prof. Lincoln Zotarelli

- Spearheaded Florida state legislative-backed and Florida Department of Agriculture and Consumer Services (FDACS)-funded best management practices (BMPs) project
- Managed multi-disciplinary teams from various departments to lead the effort in improving crop nutrient recommendations by utilizing crop models-fused artificial intelligence (AI) and remote sensing, leveraging high-performance computing (HPC) system.
- Plan, coordinate, and assist with several ongoing nutrient and irrigation management and potato breeding program
- Wrote several grant proposals targeting various state and federal agencies such as- USDA-NIFA/Data Science for Food and Agricultural Systems (DSFAS), FDACS/Specialty Crop Block Grant Program (SCBGP).
- Collaborated with multiple research teams spearheaded by different PIs for conducting inter-disciplinary research to enhance crop productivity with minimal environment impact.
- Advanced the development of the cropping system model- DSSAT while being part of the DSSAT Development Team led by Prof. Gerrit Hoogenboom at UF.
- Advised and mentored various graduate and undergraduate students in the research group
- Participated in various extension activities such as growers field days and in-service trainings to interact with growers and extension agents

This work has resulted in 4 grants (3 pending, 1 accepted), 2 journal publications, and 1 extension article.

Graduate Research Assistant, University of Illinois at Urbana-Champaign **2017 - 2022**
Department of Agricultural and Biological Engineering
Advisor: Dr. Rabin Bhattarai

- Conceptualized a methodology to understand the behavior of climate-soil-water driven impact on the crops (corn-soybean-cereal rye) by developing a calibrated/validated process-based decision support system to address the soil nutrient loss and drainage water quality issue.
- Developed CERES-Wheat model of DSSAT as a proxy model to simulate the growth and development of cereal rye using field experimental observations.

- Translated a field scale crop modeling work to large scale (state scale) spatial model using high-resolution soil, weather, and cropland data.
- Investigated the long-term impact of winter cover crop on corn-soybean growth and soil nitrogen-water dynamics using spatial modeling approach by leveraging HPC system.
- Examined the sustainability of cover cropping practices in Illinois climate divisions using future projections from Regional Climate Models (RCM).
- Collaborated with the researchers and scientists to aid them in developing the CoverCropAnalyzer – web application for cover crop management (https://covercrop.ncsa.illinois.edu).

This work resulted in 4 journal publications and 3 extension articles.

Graduate Research Assistant, Indian Institute of Technology - Kharagpur **2014 - 2017**
Department of Agricultural and Food Engineering
Advisor: Prof. Ashok Mishra

- Analyzed the climate change-induced impact and uncertainty on rice yield of India using spatial crop modeling approach at the country-scale using Global Climate Models' (GCMs) projections.
- Designed various tools to downscale and bias correct (*WDI, CDBC*) the climate model historical and future projections using a quantile mapping approach.

This work resulted in 2 journal publications and 2 GitHub repositories.

Remote Sensing, GIS, Hydrological Modeling - Internship, Indian Institute of Remote Sensing, ISRO **2014**
Water Resources Department
Advisor: Dr. Vaibhav Garg, Co-Advisor: Dr. Bhaskar Nigam

- Analyzed the hydrological features of a watershed in India, by deploying Remote Sensing and GIS and the Soil and Water Assessment Tool (SWAT) modeling technique.

Natural Resources Management- Internship, Central Soil-Water Conservation Research & Training Institute, ICAR **2013**

- Conceptualized the soil erosion controlling practices with several field visits in a highly eroded- ravenous land beside Yamuna River, Agra to gain insights on natural resources management.

TEACHING EXPERIENCE

Post Doctoral Mentor **2022-Present**
Herbert Wertheim College of Engineering, University of Florida

- *EGN 4912 Engineering Undergraduate Research* - [1 Undergraduate Student]
  -The primary purpose of this course is to provide the undergraduate student an opportunity for firsthand, supervised research experience.
  -Mentored an undergraduate student to develop machine learning models for potato cropping system using DSSAT crop model simulations.

- *EEL 9635 Physics-Informed Machine Learning* - [20 Graduate Students]
  -Gave a talk on '*Evapotranspiration and its Modeling*' as a group project idea for graduate students class project.
  -Mentored a group of students (3 graduate students) interested in pursuing their class project to use ML for predicting evapotranspiration

Graduate Teaching Assistant **2018 - 2022**
Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign

- *ABE 224 Principles: Soil & Water* - [30-40 Students]
  -Instructed all the laboratory classes of the course every Fall semester from 2018-2020.
  -Topics instructed are *'Field survey and contour mapping', 'GIS, Watershed delineation', 'Runoff volume calculation –SCS methods', 'Rainfall simulator and soil erosion',* and *'Calibration of v-notch in the flume.'*
  -Designed ArcMap training guidelines for *'GIS, Watershed delineation'* and *'Contour Mapping'* lab classes.

- *ABE 199 (Campus Honors Program) Water in the Global Environment* - [20-30 Students]
  -Teaching assistant every Fall semester from 2019-2022.
  -Graded assignments and solved students' questions and problems.
  -Led office hours for one-to-one interaction with students.

SUPERVISING / MENTORING EXPERIENCE

Post-Doctoral Associate **2022 - Present**
Horticultural Sciences Department, University of Florida, Gainesville, FL

a. Supervised - Research/Student Assistants (6)
In my role as a Postdoctoral Research Associate, I lead the AI project, focusing on developing several crop model-guided machine learning models to offer nutrient recommendations for Florida's cropping systems. I am directly overseeing the work of five research and student assistants dedicated to this project.

Name	Designation	Project Activity	Year
Prateek K. Goel*	Research Assistant, HOS & Department of Electrical & Computer Engineering (ECE)	Supervised and trained to improve the understanding of the DSSAT model and `*dssat-pylib*` python scripts to simulate, extract, and plot the model output. Guided on developing transformer based neural network models for soil mineral nitrogen	2022-Present
Bhanu P. R. Lanka*	Student Assistant, M.S. Student, Department of Computer & Information Science & Engineering (CISE)	Mentored and trained to develop XGBoost and other machine learning models for estimating potential nitrogen leaching while reducing dependence on process-based crop model simulations	2024-Present
Miguel A. Barrera	Student Assistant, B.S. student, ECE	Advised and familiarized with the raw data collected for the potato trails in Hastings, Florida and explained the best way to convert the raw data into DSSAT model input files, explained the working on the DSSAT model and how to extract its output in required format.	2022-2024
Lahari Kethinedi	Student Assistant, M.S. Student, ECE	Supervised to develop machine learning models for potato agroecosystem using DSSAT simulated data for the hypothetical scenarios.	2023-2024
Geoffrey 'Austin' Simon	Graduate Research Assistant, M.S. student, ECE	Guided on to use Unmanned Aerial Vehicle (UAV)-captured multispectral data for monitoring potato growth and developing a XGBoost-based machine learning model.	2024
John Jack Upchurch	Graduate Research Assistant,	Guiding him with the potato farms so that he could use his UAV flying expertise to capture all the	2024

	M.S. student, ECE	farms capturing plot level and plant level data by mapping the plots and sampling plants in the plot, respectively.	

* These students are currently working with me on AI project.

b. Mentored - Ph.D. and Masters' Students (4)

During my postdoctoral training, I am providing mentorship to numerous Ph.D. and Masters students. While I was not their official advisor, the students frequently sought my counsel and guidance in formulating their thesis objectives. They approached me during moments of uncertainty related to their project, seeking advice on navigating challenges and formulating pertinent questions to discuss with their respective advisors.

Name	Designation	Project Activity	Year
Satya K. Pothapragada*	Ph.D. student, ECE	Guiding with different aspect and providing foundational knowledge of agricultural system to develop machine learning models.	2022-Present
Rakshya Dhakal*	Ph.D. student, HOS	Provided guidance on how to employ the DSSAT model and remote sensing techniques for modeling the potato breeding/variety trials.	2023-Present
Weihuang Xu	Ph.D. student [Graduated], ECE	Explained the soil nitrogen volatility in the sandy soils to efficiently predict the daily response of rainfall and subirrigation on the soil mineral nitrogen in the potato field trials using machine learning.	2022-2023
Ayesha Naikodi	M.S. Student, [Graduated], ECE	Enhanced the understanding on the fundamental agricultural processes related to potato agroecosystem modeling.	2022

* These students are currently working with me on AI project.

c. Mentored – Bachelor's Student

Name	Designation	Research Activity	Semester
Miguel A. Barrera	B.S. student, ECE	EGN-4921 Engineering Undergraduate Research (Credit - 3) Course objectives were to familiarize undergraduate students on how to search literature, approach a research problem, write a research report, and work in a team environment	Fall-2023

GRANT WRITING EXPERIENCE

1. Developing site-specific irrigation and nutrient prescription maps using AI and remote sensing associating soil spatial heterogeneity response with potato production. — $74,826 — 2025-2027 (Pending)
   [**Co-PI** with Drs. Zotarelli, Resende, Zare, and Sharma]
   UF/IFAS Dean for Research & Extension Office
2. Advancing Precision irrigation for specialty crops: Development of a statewide tool for site-specific subirrigation recommendations. — $249,930 — 2026-2028 (Pending)
   [**Co-PI** with Drs. Zotarelli, Oliveira, Nunes, Guzman, Tsouvaltzis, and Agehara]
   Specialty Crop Block Grant Program, USDA

3. DSFAS: AI-driven genomic and crop simulation models for enhancing plant breeding using high-throughput phenotyping. $649,956 2025-2028 (Pending)
   [**Co-PI** with Drs Zotarelli, Zare, Resende, Messina, Harley, and Sharma]
   National Institute of Food Agriculture, USDA
4. Using artificial intelligence for improved crop nutrient management. $248,035 2024-2025
   [**Co-PI** with Drs Zotarelli, Zare, Resende, Messina, and Harley]
   Florida Department of Agriculture & Consumer Services, FL

FUNDED PROJECTS

- Partnership: Integrating crop growth models and genomic prediction to advance the development of heat tolerant potatoes. 2024-Present
  National Institute of Food Agriculture, USDA.
- Using artificial intelligence for improved crop nutrient management. 2022-Present
  Florida Department of Agriculture & Consumer Services, FL.
- Web-based decision support tool for cover crop management. 2017-2022
  Illinois Nutrient Research & Education Council, IL.
- Climate change impact and adaptation options for sustaining rice-wheat crop production in India. 2015-2017
  Department of Science and Technology, Government of India.

PUBLICATIONS

a. Refereed Journal papers (7)

1. **Gupta, R.**, Pothapragada, S.K., Xu, W., Goel, P.K., Barrera, M.A., Harley, J., Morgan, K., Zare, A., Zotarelli, L. (2024). Estimating soil mineral nitrogen from data-sparse field experiments using crop model-guided deep learning approach. *Computers and Electronic in Agriculture*, 225, 109355. https://doi.org/10.1016/j.compag.2024.109355
2. da Silva, A.L.B.R., Dias, H.B., **Gupta, R.,** Zotarelli, L., Asseng, S., Dukes, M.D., Porter, C.H., Hoogenboom, G. (2024). Assessing the impact of irrigation and nitrogen management on potato performance under varying climate in the state of Florida, USA. *Agricultural Water Management*, 295, 108769. https://doi.org/10.1016/j.agwat.2024.108769
3. **Gupta, R.**, Bhattarai, R., Dokoohaki, H., Armstrong, S.D., Coppess, J.W. & Kalita, P.K. (2023b). Sustainability of cover cropping practice with changing climate. *Journal of Environmental Management*, 339, 117946. https://doi.org/10.1016/j.jenvman.2023.117946
4. **Gupta, R.**, Bhattarai, R., Kalita, P.K., Dokoohaki, H., Armstrong, S.D. & Coppess, J.W. (2023a). Evaluation of long-term impact of cereal rye as a winter cover crop in Illinois. *Science of the Total Environment*, 887, 162956. https://doi.org/10.1016/j.scitotenv.2023.162956
5. **Gupta, R.**, Bhattarai, R., Coppess, J.W., Jeong, H., Ruffatti, M., & Armstrong, S.D. (2022). Modeling the impact of winter cover crop on tile drainage and nitrate loss using DSSAT model. *Agricultural Water Management*, 272, 107862. https://doi.org/10.1016/j.agwat.2022.107862
6. **Gupta, R.**, & Mishra, A. (2019). Climate change induced impact and uncertainty of rice yield of agro-ecological zones of India. *Agricultural Systems*, 173, 1-11. https://doi.org/10.1016/j.agsy.2019.01.009

7. **Gupta, R.**, Bhattarai, R., & Mishra, A. (2019). Development of climate data bias corrector (CDBC) tool and its application over the agro-ecological zones of India. *Water (Switzerland)*, 11(5), [1102]. https://doi.org/10.3390/w11051102
8. Shrestha, S., Sharma, S., **Gupta, R.**, & Bhattarai, R. (2019). Impact of global climate change on stream low flows: A case study of the great Miami river watershed, Ohio, USA. *International Journal of Agricultural and Biological Engineering*, 12(1), 84-95. https://doi.org/10.25165/j.ijabe.20191201.4486

b. Peer-reviewed Extension Publications (4)

1. Sharma, L., **Gupta, R.**, Zotarelli, L., & Hoogenboom, G. (2024). Assessing Nitrate Leaching in Crop Production through the Application of Crop Simulation Models with Experimental Data from Florida: SL514/SS727, 9/2024. EDIS, 2024(5). https://doi.org/10.32473/edis-ss727-2024
2. Coppess, J., Navarro, C., Naraharisetty, V.V.G., Satheesan, S.P., Gatzke, L., Bhattarai, R., **Gupta, R.**, Armstrong, S.D., & Ford, T. (2021) Introducing Further Updates to the Cover Crop Decision Support Tool. *farmdoc daily* (11):148 https://farmdocdaily.illinois.edu/2021/10/introducing-further-updates-to-the-cover-crop-decision-support-tool.html
3. Coppess, J., Navarro, C., Satheesan, S.P., Naraharisetty, V.V.G., Bhattarai, R., Armstrong, S.D., & **Gupta, R.** (2021) Introducing an Update to the Cover Crop Decision Support Tool. *farmdoc daily* (11):18. https://farmdocdaily.illinois.edu/2021/02/introducing-an-update-to-the-cover-crop-decision-support-tool.html
4. Coppess, J., Navarro, C., Satheesan, S.P., Naraharisetty, V.V.G., Bhattarai, R., Armstrong, S.D., & **Gupta, R.** (2020). Introducing the Cover Crop Decision Support Tool. *farmdoc daily*, 10(176). https://farmdocdaily.illinois.edu/2020/10/introducing-the-cover-crop-decision-support-tool.html

c. Non-refereed Proceedings (1)

1. Satheesan, S.P., Bhattarai, R., Bradley, S., Coppess, J., Gatzke, L., **Gupta, R.**, ... & Navarro, C.M. (2019). Extensible framework for analysis of farm practices and programs. *In Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (learning)* (pp. 1-8). https://doi.org/10.1145/3332186.3337063

d. Newsletters (1)

1. **Gupta, R.**, Bhattarai, R., Cooke, R.A. (2018) Characterizing the response of four tile-drained watershed - Application of SWAT model. *Land Improvement Contractors of America (Illinois Chapter) newsletter*. Page (14-15): Jan-Feb 2018. https://www.illica.net/_files/ugd/f15a8b_e5800588eb414ffc93582efc8382bfd8.pdf

e. Published Abstracts (5)

1. **Gupta, R.**, Dias, H.B., Zotarelli, L., Porter, C.H., Hoogenboom, G. (2023) Evaluating the DSSAT-CSM-SUBSTOR model in sub-irrigated potato-agroecosystem under varied nitrogen fertilizer rates and application timings. ASA, CSSA, SSSA International Annual Meeting – 2023. https://scisoc.confex.com/scisoc/2023am/meetingapp.cgi/Paper/151749
2. Dias, H.B., **Gupta, R.**, da Silva, A.L.B.R., Zotarelli, L., Asseng, S., Porter, C.H., Hoogenboom, G. (2023) Evaluating and applying the DSSAT-CSM-SUBSTOR model to simulate water and nitrogen responses in spring potato in northeast Florida. ASA, CSSA, SSSA International Annual Meeting – 2023. https://scisoc.confex.com/scisoc/2023am/meetingapp.cgi/Paper/153900
3. Oliveira, J.M., Tormena, C.A., Zotarelli, L., Bortolozo, F.R., **Gupta, R.** (2023) Pedotransfer function to estimate soil penetration resistance of sandy soils. ASA, CSSA, SSSA International Annual Meeting – 2023. https://scisoc.confex.com/scisoc/2023am/meetingapp.cgi/Paper/149572

4. Satheesan, S.P., Navarro, C. Lee, J., Naraharisetty, G., Gatzke, L., **Gupta, R.**, Bhattarai, R., Ford, T., Armstrong, S.D., Coppess, J. (2022) Development and advancement of CoverCrop Analyzer: A decision support web application for cover crop management. American Geophysical Union - Fall Meeting Abstracts – 2022. https://agu.confex.com/agu/fm22/meetingapp.cgi/Paper/1090970
5. **Gupta, R.**, Bhattarai, R. Dokoohaki, H., Coppess, J., Armstrong, S.D. (2022) Analyzing sustainability of cover cropping practice with changing climate over Illinois state. Ohio River Basin Consortium for Research and Education - Symposium – 2022. https://event.fourwaves.com/orbcre/abstracts/5949b516-09aa-4bed-8cc7-4e14766462c5

f. Research Mentions in Media (2)

1. Top corn producing state to see future drop in yield, cover crop efficiency. *ScienceDaily* (2023). www.sciencedaily.com/releases/2023/07/230706124639.htm
2. Quinn, L. (2023) Winter cover crops could reduce nitrogen in Illinois drainage water by 30%. *NewsWise* (2023). https://www.newswise.com/articles/winter-cover-crops-could-reduce-nitrogen-in-illinois-drainage-water-by-30

SOFTWARES / TOOLS DEVELOPED

• DSSAT-SoilPro	*DSSAT-SoilPro* – A tool to optimize soil water-related parameters in the DSSAT model to calibrate soil moisture content of different soil layers.
• dssat-pylib	*DSSAT-Python Library* – A Python library to read/extract DSSAT outputs, create/modify its soil file, and run the model.
• TabulateQR	*TabulateQR*- A tool to create customized tables for organizing soil, water, or other sampled data linked with QR codes.
• CDBC	*Climate Data Bias Corrector* – A tool developed in Python to remove the bias from daily gridded climate projections of climate models using the probability distributions (gamma, gaussian, beta).
• WDI	*Weather Data Interpolator* – A tool developed in Python to downscale daily weather data from lower to higher resolution using linear and inverse distance interpolation techniques.

INVITED PRESENTATIONS / SEMINARS

- Presentation invitation by Dr. Joel Harley to give a talk on '*Evapotranspiration and its Modeling*' as a project idea for his class- Physics-Informed Machine Learning at the Department of Computer and Electrical Engineering, University of Florida, FL – 2023.
- Presentation invitation by Dr. Marcio Resende to give a talk on '*Soil mineral nitrogen modeling in the potato cropping system*' to his research group and a visiting professor from Brazil at the Horticultural Sciences Department, University of Florida, FL – 2023.
- Seminar invitation by Prof. Prasanta Kalita to give a talk on my Ph.D. research '*Modeling the impact of winter cover crop on soil water and nitrogen dynamics using DSSAT model*' at the Department of Agricultural and Biological Engineering, University of Illinois at Urbana-Champaign, IL – 2021.

PRESENTATIONS AT PROFESSIONAL CONFERENCES / MEETINGS

a. Oral Presentations (8)

1. **Gupta, R.**, Pothapragada, S.K., Goel, P.K., Lanka, B.P.R., Kethinedi, L., Barrera, M.A., Harley, J.B., Morgan, K.D., Zare, A., Zotarelli, L. (2024) Exploring crop model-informed machine learning approach for interpreting cropping system. ASABE Annual International Meeting – 2024.
2. **Gupta, R.**, Dias, H.B., Zotarelli, L., Porter, C.H., Hoogenboom, G. (2023) Evaluating the DSSAT-CSM-SUBSTOR model in sub-irrigated potato-agroecosystem under varied nitrogen fertilizer rates and application timings. ASA, CSSA, SSSA International Annual Meeting – 2023. https://scisoc.confex.com/scisoc/2023am/meetingapp.cgi/Paper/151749
3. **Gupta, R.**, Pothapragada, S.K., Xu, W., Goel, P.K., Barrera, M.A., Harley, J.B., Morgan K.D., Zare, A. Zotarelli, L. (2023) An effort to couple DSSAT model with machine learning model to estimate soil mineral nitrogen in potato fields. ASABE Annual International Meeting – 2023.
4. **Gupta, R.**, Bhattarai, R., Satheesan, S.P., Navarro, C., Ford, T., Armstrong, S.D., Coppess, J. (2023) CoverCrop Analyzer: a decision support web application for cover crop management. ASABE Annual International Meeting – 2023.
5. **Gupta, R.**, Bhattarai, R. Dokoohaki, H., Coppess, J., Armstrong, S.D. (2022) Analyzing sustainability of cover cropping practice with changing climate over Illinois state. Ohio River Basin Consortium for Research and Education - Symposium – 2022. https://event.fourwaves.com/orbcre/abstracts/5949b516-09aa-4bed-8cc7-4e14766462c5
6. **Gupta, R.**, Bhattarai, R., Dokoohaki, H., Coppess, J., Armstrong, S.D. (2022) A nutrient loss prevention outlook using extensive cover cropping in the croplands of Illinois. ASABE Annual International Meeting – 2022.
7. **Gupta, R.**, Bhattarai, R., Coppess, J., Roth, R.T., Armstrong, S.D. (2020) Modeling the impact of winter cover crop on corn-soybean and soil nitrogen dynamics using DSSAT model. ASABE Annual International Meeting – 2020.
8. **Gupta, R.**, Bhattarai, R., Roth, R.T., Armstrong, S.D. (2019) Analyzing winter cover crop impacts on soil nitrogen dynamics using DSSAT model. ASABE Annual International Meeting – 2019.

b. Poster Presentations (1)

1. **Gupta, R**., Bhattarai, R., Mishra, A. (2018) Climate Data Bias Corrector: A tool for bias correction of GCM/RCMs outputs. ASABE Annual International Meeting – 2018.

ACADEMIC SERVICES

- Served as a *manuscript reviewer* for the research articles submitted to various journals-
  - *Scientific Data (Nature Publishing Group), Journal of Environmental Management, Agricultural and Forest Meteorology, Journal of the American Water Resources Association, Agronomy Journal, American Journal of Potato Research, Journal of Natural Resources and Agricultural Ecosystems (ASABE), Environmental Processes, Agronomy, Climate, Applied Sciences, and Sustainability.*
- Served as a *moderator*:
  - Natural Resources & Environmental Systems (NRES) Oral Technical Session on '*Nutrient Transport and Cycling: Measurement and Data Synthesis*' at the ASABE AIM – 2024, Anaheim, CA.
  - NRES Oral Technical Session on '*Nutrient Transport and Cycling: Modeling*' at the ASABE AIM – 2024, Anaheim, CA.

  - NRES Oral Technical Session on ‘*Nutrient Transport and Cycling: Measurement and Modeling*’ at the ASABE AIM – 2023, Omaha, NE.

- *Judged* the students’ presentations
  - NRES Oral Technical Session on ‘*Nutrient Transport and Cycling: Modeling*’ at the ASABE AIM – 2024, Anaheim, CA.
  - NRES Oral Technical Session on ‘*Advances in Agro-EcoSystems Modeling and Data Analytics: Climate Impacts*’ at the ASABE AIM – 2023, Omaha, NE
  - NRES Oral Technical Session on ‘*Advances in Agro-EcoSystems Modeling and Data Analytics: Ecohydrology Applications*’ at the ASABE AIM – 2023, Omaha, NE
  - NRES Poster Technical Session at the ASABE AIM – 2023, Omaha, NE

PROFESSIONAL DEVELOPMENT TRAININGS

- Completed the ‘*Preparing Future Faculty*’ program offered by the Graduate School, the Office of Postdoctoral Affairs, and the Center for Teaching Excellence at the University of Florida, FL – 2024.
- Completed the workshop- ‘*Teaching for Inclusivity and Accessibility*’ offered by the Center of Instructional Technology and Training (CITT) at the University of Florida, FL – 2023.
- Completed the workshop- ‘*Grantsmanship 101: Keys to Writing Effective Proposals*’ offered by the Office of Graduate Professional Development at the University of Florida, FL – 2023.

PROFESSIONAL MEMBERSHIPS

- American Society of Agricultural and Biological Engineers (ASABE) (2018-Present)
  - Chair of NRES-224 Water Quality (2024-2026)
  - Vice Chair of NRES-224 Water Quality (2023-2024)
  - Member-at-Large of the Young Professional Community (YPC) (2023-2024)
  - Member-at-Large of the Association of Agricultural, Biological, and Food Engineers of Indian Origin (2023-2024)
  - Member of NRES -21 Hydrology Group Committee (2022-Present)
  - Member of NRES-23 Drainage Group Committee (2022-Present)
  - Member of NRES-22 Soil Erosion and Water Quality Committee (2022-Present)
- American Society of Agronomy (ASA) (2023-Present)

AWARDS / ACHIEVEMENTS / SCHOLARSHIPS

- Postdoctoral Travel Award from the Graduate School, UF, 2024
- Conference Presentation Award from the Graduate College, UIUC, 2022
- Graduate College Block Grant from the Graduate College, UIUC, 2018
- All India Council for Technical Education Postgraduate Scholarship from Govt. of India, 2014-17
- Editor-in-chief in a wall magazine- ‘*AgriNEST*’ at JNKVV, India, from 2012-14
- National Talent Scholarship from the Indian Council of Agricultural Council (ICAR), 2010-14

PROFILE LINKS

- Google Scholar: https://scholar.google.com/citations?user=txTgLqsAAAAJ&hl=en
- ResearchGate: https://www.researchgate.net/profile/Rishabh-Gupta-5
- LinkedIn: https://www.linkedin.com/in/rishabhgupta61/
- GitHub: https://github.com/AgroClimaticTools
- ORCID: https://orcid.org/0000-0002-8937-1910