Conference on Computational Science

October 18, 2023 – 8:00 a.m. to 5:00 p.m. – NDSU Memorial Union's Oceti Sakowin Ballroom

Co-organized by NDSU's Information Technology Division and the Office of Research and Creative Activity

08:00–08:30	CHECK IN/BREAKFAST
08:30–08:40	Opening remarks. Vice President **Marc Wallman**. Conference overview. **Khang Hoang** (IT/CCAST)
08:40–09:00	*In silico examination of breast cancer and prostate cancer bone metastasis.* **Dinesh Katti** (Civil, Construction and Environmental Engineering)
09:00–09:20	*Text mining for U.S. pension de-risking analysis.* **Limin Zhang** (Accounting and Information Systems), **Ruilin Tian** (Transportation, Logistics, and Finance), **Jeffrey Chen** (Transportation, Logistics, and Finance)
09:20–09:40	*NDSU's comprehensive approach to agriculture, food systems, and natural resources.* **Frank Casey** (ND Agricultural Experiment Station)
09:40–10:00	*"Prevent brain cancer with just three easy steps!": Identifying medical misinformation on social media with a neural network using rhetorical feature sets.* **Zoltan Majdik** (Communication)
10:00–10:20	COFFEE BREAK
10:20–10:40	*Toward exascale computing for patient-specific design of cardiovascular prosthesis.* **Trung Le** (Civil, Construction and Environmental Engineering)
10:40–11:00	*Modernizing agriculture with data-driven solutions: The NDSU's Big Data Pipeline Unit Initiative.* **Ana Heilman-Morales** (Big Data Pipeline Unit)
11:00–11:20	*Machine learning in materials design.* **Bakhtiyor Rasulev** (Coatings and Polymeric Materials)
11:20–11:40	*From cells to societies: Unveiling patterns with multiplex network analysis.* **Harun Pirim** (Industrial and Manufacturing Engineering)
11:40–12:00	*Advanced cyberinfrastructure for computational research and education.* **Pete Lambertz** (IT/Enterprise Computing & Infrastructure) and **Khang Hoang** (IT/CCAST)
12:00–01:00	LUNCH. Special remarks/legislative updates from Provost **David Bertolini** and Vice President **Marc Wallman**
01:00–01:20	Research updates/special remarks. Vice President **Colleen Fitzgerald**
01:20–01:40	*Do sanctioning and monitoring affect elites' online toxicity? Evidence from a field experiment on U.S. general election candidates.* **Dan Pemstein** (Political Science)
01:40–02:00	*Emerging technology driven memory system designs.* **Sumitha George** (Electrical and Computer Engineering)
02:00–02:20	COFFEE BREAK
02:20–02:40	*Redesigning the next-generation pulse crops with genetic innovations.* **Nonoy Bandillo** (Plant Sciences)
02:40–03:00	*Leveraging large language models for genomic sequence analysis.* **Changhui Yan** (Computer Science)
03:00–03:20	*Impact of covalent functionalization of nanostructures on their optical properties: Computational insights.* **Svetlana Kilina** (Chemistry and Biochemistry)
03:20–05:00	POSTER SESSION (>40 poster presentations). Poster Award Ceremony (at about 4:30 p.m.)

Notes: Research talk (17min plus 3min for Q&As), coffee break (20min), lunch break (1h), poster session (~1h30min).

POSTER PRESENTATIONS

NS1 Improving photoluminescence properties of carbon nanotubes with combined machine learning, computational chemistry and chemical substructure encoding approaches.
Gerardo M. Casanola-Martin, Amirreza Daghighi, August Amb, Grace Tiffany, Dmitri Kilin, Svetlana Kilina, and Bakhtiyor Rasulev

NS2 Enhanced photoluminescent quantum yield in functionalized single walled carbon nanotube.
Dinesh Thapa, Dmitri Kilin, and Svetlana Kilina

NS3 Understanding Yellowstone cutthroat trout hybridization and connectivity within the Teton River system, a combined genomics and modeling approach.
Meggan Alston, Michael Youngwirth, Alexandra Fraik, Paul Hohenlohe, and Travis Seaborn

NS4 Relaxation of photoexcited electron-hole pairs at Si(111) surfaces with adsorbed Ag cluster monolayers.
Yulun Han, Tijo Vazhappilly, David A. Micha, and Dmitri S. Kilin

S01 Human colon cancer dual active drug–metal organic framework nanomaterial (DADMOM) systems: Dataset compilation and artificial intelligence modeling.
Shan He, Cristian R. Munteanu, Begoña de Bilbao, Harbil Bediaga, Mariana Chelu, Adina Magdalena Musuc, Sonia Arrasate, Alejandro Pazos, and Humberto González-Díaz

S02 PCR IF+PT+ML metagenomic study of lung microbiome in cystic fibrosis: Toward GDPR personalized medicine compliant decision making.
Estefania Asencio-Medina, Carmen Velásquez, Amaia González-Magaña, Maria D. Pastor-Vivero, Laura Acosta, Aliuska Duardo-Sánchez, Sonia Arrasate, Humberto González-Díaz, and David Albesa-Jové

S03 A hybrid staggered/non-staggered formulation for simulating incompressible flows with block-structured mesh refinement.
Thien-Tam Nguyen, Rajasekhar Dathi, Andy J. Nonaka, and Trung B. Le

S04 A heuristic agent-based model of the market as a distributed genetic algorithm.
Zach Johnson and James Caton

S05 Effect of the crystal structure on optical response of HgSe quantum dots.
Steven Westra and Svetlana Kilina

S06 Discovery of an effective α-amylase inhibitor via machine learning-based QSAR approach.
Mariam Zamani, Gerardo Casanola-Martin, and Bakhtiyor Rasulev

S07 Tungsten coordinated complexes: The effect of DFT functionals in comparison with experimental data.
Tori Oas, Grace Tiffany, Hang Wang, Wenfang Sun, and Svetlana Kilina

S08 A numerical study of particle motion in a turbulent pipe flow.
Rajasekhar Dathi, Imtiaj Nahin Ahmed, Hong Pan, Trung B. Le, Yan Zhang, and Zhibin Lin

S09 qsarify: High performance machine learning software package for QSAR model development.
Stephen Szwiec and Bakhtiyor Rasulev

S10 Predicting the toxicity of fullerene derivatives using classification machine learning models.
Madeleine Williams, Amirreza Daghighi, and Bakhityor Rasulev

S11 Machine learned force fields assisted molecular dynamics in VASP.
David Graupner and Dmitri Kilin

S12 Streamlining data querying with exlibris: A data pipeline approach.
Grant Hoff, Souradip Roy, Robert Fedor, Matheus Siqueira, Reed Fox, Sikiru Atanda, Didier Florez, Lane O'Brien, David Froslie, and Ana Heilman-Morales

S13 From raw images to vegetation indexes: Automating the UAV image data processing workflow for precision agriculture.
Robert Fedor, Matheus Siqueira, Paulo Flores, Reed Fox, Grant Hoff, Sikiru Atanda, Didier Murillo-Florez, Souradip Roy, Lane O'Brien, David Froslie, and Ana Heilman-Morales

S14 Modernizing NDSU pest management app: Enhancing pesticide knowledge and accessibility for North Dakota's agricultural stakeholders.
Reed Fox, Robert Fedor, Andrew Friskop, Matheus Siqueira, Grant Hoff, Sikiru Atanda, Didier Murillo-Florez, Souradip Roy, Lane O'Brien, David Froslie, and Ana Heilman-Morales

S15 Integrating experimental and computational approaches to assess the performance of organic corrosion inhibitors through QSAR modeling and extrapolation techniques.
Chameli Samarawickrema, Gerardo Casanola-Martin, Zachary Bergseth, Xiaoning Qi, and Bakhtiyor Rasulev

S16 β-cyclodextrin–phenolic acid complexation study to mask bitterness in wheat bran (A QSPR approach).
Kweeni Iduoku, Marvellous Ngongang, Jayani Maddakandaged, Amir Daghighi, Gerardo M. Casanola-Martin, Senay Simsek, and Bakhtiyor Rasulev

S17 Development of machine learning-based QSAR models to predict green algae growth inhibition by polymeric particles.
Seyedehelham Shirvani-Hosseini, Gerardo M. Casanola-Martin, and Bakhtiyor Rasulev

S18 Are we ready for hydrogen: A comprehensive evaluation of existing energy pipeline infrastructure for its transport.
Mohsin Ali Khan, Hong Pan, Kevin Wang, and Zhibin Lin

S19 Ab initio calculations of through-space and through-bond charge-transfer properties of interacting Janus-like PbSe and CdSe quantum dot heterostructures.
Hadassah B. Griffin, Andrei B. Kryjevski, and Dmitri Kilin

S20 Demystifying cyber insurance adoption for small businesses.
Zia Muhamad and Jeremy Straub

S21 Multi-endpoint acute toxicity assessment of organic compounds using large-scale machine learning modeling.
Amirreza Daghighi, Gerardo M. Casanola-Martin, Kweeni Idouku, and Bakhtiyor Rasulev

S22 Transfer kinetics of cargo items among mobile nanocarriers.
Faruk Hossain, Guilherme Volpe Bossa, and Sylvio May

S23 Optical properties of charged non-stoichiometric CdSe quantum dots: Computational insights.
Omolola Eniodunmo and Svetlana Kilina

S24 In search of improved, stabilized doped LLZO electrolytes for all solid-state Li-ion batteries through first-principles calculations.
Md Mozammal Raju, Syed Rizvi, and Qifeng Zhang

S25 High-throughput phenotyping of seed yield and breeder's score in canola.
Mohammad Jony, Md. Fahad Hasan, Paulo Flores, Mukhlesur Rahman

S26 Time-resolved properties of vibrational polariton model.
Patricia Adeoye and Dmitri Kilin

S27 Ab initio modeling of superconducting materials by exploring excited electronic configurations.
William Tupa and Dmitri Kilin

S28 FHGAON – A user friendly tool for fungal whole genome assembly using MinION sequencing.
Jatinder Singh, Rachel Konshok, Santosh Gudi, and Upinder Gill

S29 Development of numerical features/descriptors to describe complex materials for machine learning modeling.
Rahil Ashtari Mahini, Gerardo Casanola-Martin, Simone Ludwig, and Bakhtiyor Rasulev

S30 Development and application of KASP marker for screening sclerotinia stem rot resistance in canola (Brassica napus L.).
Md Zahangir Alam, Mohammad Jony, Luis E. del Rio Mendoza, and Mukhlesur Rahman

S31 Computational studies of droplet dynamics near a rough surface.
Julie Melbye and Yechun Wang

S32 Multilayer network analysis of energy networks.
Muhammad Kazim and Harun Pirim

S33 Nanomechanics of actin filament deformation and fracture: A steered molecular dynamics approach.
Sharad Jaswandkar, Kalpana Katti, and Dinesh Katti

S34 Unraveling integrins' conformational changes and mechanics at the molecular scale.
Hanmant Gaikwad, Sharad V. Jaswandkar, Kalpana S. Katti, Amanda Haage, and Dinesh Katti

S35 Coarse-grained modeling of polymer–clay nanocomposites.
Wenjian Nie, Sarah Ghazanfari, Kalpana Katti, Dinesh Katti,and Wenjie Xia

S36 Computational insights: Electronic structure and optical properties of CdSe quantum dots as a photocatalyst for lignin depolymerization reaction.
Shabin Gurung and Svetlana Kilina

S37 Relationships between electronic structures and structural properties of polymer systems.
Meade Erickson, Yulun Han, Gerado M. Casañola-Martin, Bakhtiyor Rasulev, and Dmitri Kilin

INVITED TALKS

#01 **In silico examination of breast cancer and prostate cancer bone metastasis**

Dinesh R. Katti (Department of Civil, Construction and Environmental Engineering)

One in eight women and one in seven men will be diagnosed with breast and prostate cancer, respectively, in their lifetime. Over one million people diagnosed with these cancers globally die annually. Most of these deaths are attributed to cancer metastasis, where cancer cells from the breast or prostate migrate to a distant organ. Breast cancer and prostate cancer have the propensity to metastasize to bone. There is no cure when cancer cells colonize at the bone site. Our research group has built novel in vivo models for breast and prostate cancer bone metastasis using tissue engineering approaches that include nanoclay-based scaffolds. We have developed multiscale models of tissue engineering scaffolds that allow for simulations-based design of in vivo models. Additionally, we have conducted steered molecular dynamics simulations of a critical cellular protein, actin, and elucidated the key mechanisms contributing to cancer's mechanobiology during cancer progression at the bone metastasis site. Simulations also provide a rich insight into the mechanics of adhesion molecule integrin and mechanisms contributing to human bone marrow mesenchymal stem cell differentiation leading to bone tissue regeneration on nanoclay-based scaffolds. Using principal component analysis of Raman data, we identified spectral biomarkers of cancer progression at the bone metastasis site. A novel cluster analysis approach allowed for imaging cellular proteins in tumors. Simulations provide an unprecedented view into cancer progression and the discovery of fundamental mechanisms that would open doors for cure.

#02 **Text mining for U.S. pension de-risking analysis**

Limin Zhang (Department of Accounting and Information Systems), **Ruilin Tian** (Department of Transportation, Logistics, and Finance), **Jeffrey Chen** (Department of Transportation, Logistics, and Finance)

The financial technology (or fintech) industry has witnessed tremendous growth in recent years thanks to the technological advancements in areas such as the Internet, artificial intelligence, machine learning, blockchain, etc. In this fintech-related research, we investigate the changes and activities in the defined benefit (DB) pension market with a focus on de-risking strategies. We build a comprehensive database that contains different types of pension de-risking strategies that US-based companies have used between 1993 and 2018 using various technologies including Web scraping, natural language processing, and machine learning. The resulting database provides valuable insights to stakeholders including companies with DB plans, pensioners, and practitioners in pension de-risking markets.

#03 **NDSU's comprehensive approach to agriculture, food systems, and natural resources**

Frank Casey (ND Agricultural Experiment Station)

Established in 1890, the legacy of the North Dakota Agricultural Experiment Station (NDAES) is intertwined with that of the North Dakota State University (NDSU). Our team comprises not only faculty scientists based at NDSU's main campus but also experts stationed at eight branch locations across North Dakota. These strategic placements allow us to directly address the unique needs of various local communities. NDAES's research spans a wide array of topics encompassing agriculture, food systems, and natural resources. This presentation offers an insightful look into NDSU's agricultural legacy and the transformative research conducted by NDAES.

#04 **"Prevent brain cancer with just three easy steps!": Identifying medical misinformation on social media with a neural network using rhetorical feature sets**

Zoltan Majdik (Department of Communication)

This presentation outlines a process for developing neural network-based classifiers for identifying image- or video-based cancer misinformation on (social) media platforms. Medical misinformation on social media is rampant. Often, medical misinformation is about various facets of cancer diagnosis, treatment possibilities,

and expected outcomes. And frequently, that information is presented through visual communication means: memes, infographics, images, video clips, or infomercial-style segments. Identifying such cancer-related misinformation computationally and at scale would help prevent harms by providing content providers a mechanism with which to identify and label misinformation. But it is challenging: visual classification models built on artificial neural network architectures have not made the same leap in accuracy as text-based classifiers, in large part because Convolutional Neural Networks used for classifying visuals are tuned for shape recognition where text-based model designs (word embeddings or attention-based transformers) are tuned to recognize semantic complexity. We developed a dual-step computational classification pipeline that leverages advancements in text-based classification/Large Language Model (LLM) architectures in ways that allow for application in visual communication domains, using rhetorical feature sets as text-based abstractions from image-based content.

#05 **Toward exascale computing for patient-specific design of cardiovascular prosthesis**

Trung Le (Department of Civil, Construction and Environmental Engineering)

Given the complexity of vasculature anatomy and valvular structures, the fluid–structure interaction (FSI) simulation of prosthetic devices and blood flow poses a significant challenge for numerical simulation. In this talk, recent advancements in exascale computing for simulating blood flow in patient-specific anatomy are systematically considered, emphasizing the numerical treatments of blood flow and prosthetic valves, which are the most critical aspects for accurate simulations. Numerical methods for hemodynamics are considered under both the continuum and discrete (particle) approaches. The numerical treatments for the structural dynamics of aortic/mitral valves and FSI coupling methods between the solid and fluid domain are also reviewed. Our current work toward multi-scale simulations in patient-specific anatomy is also discussed, including the blood damage modeling at the scale of red blood cells.

#06 **Modernizing agriculture with data-driven solutions: The NDSU's Big Data Pipeline Unit Initiative**

Ana Heilman-Morales (ND Agricultural Experiment Station, Big Data Pipeline Unit)

Advances in agricultural research currently rely on the utilization vast amounts of data, the adaptation of high throughput phenotyping and genotyping methods, high computational processing and fast speed networking technologies. The AES Big Data Pipeline Unit is pioneering efforts across diverse agronomic fronts, including the establishment of data pipelines and agronomic databases, the development of cutting-edge software applications, the inception of impactful extension projects, and the fusion of Unmanned Aerial Systems (UAS) technology software products, all of which collectively underline the unit's impact on our effort to modernize NDSU AES agricultural practices.

A pivotal dimension of the unit's work lies in the creation of databases tailored for the management of agronomic data. By integrating disparate datasets, employing data structuring methodologies, and harnessing the potential of scalable databases, the unit empowers researchers to effectively access, analyze, and manipulate agricultural information. This approach facilitates and enables data-driven decision-making and resource optimization.

The unit's software prowess is exemplified by two groundbreaking applications: FielDHub and MrBean. FielDHub, a bespoke software tool, offers researchers an intuitive click and point platform for designing experiments with great precision and simplifying the complex process of experimental planning and design. It caters to researchers across various domains by streamlining the creation of experiment layouts that adhere to specific requirements and objectives. In tandem, MrBean undertakes the intricate analysis of agronomic data by harnessing the potency and adaptability of Linear Mixed Models (LMM) to make data processing and interpretation both simple and accurate. A standout feature of Mr. Bean is its integration with robust analytical packages like ASReml and SpATS both integrated with strong visualization capabilities built in R to present the findings in an easily comprehensible format. This unique combination guarantees not just speed, but also analytical depth, ensuring our users get precise results swiftly.

Moreover, the unit is actively engaged in developing PredictPro, a tool driven by statistical and AI/ML techniques, dedicated to dissecting intricate genetic and OMICS data. Concurrently, we are also embarking on integrating Unmanned Aerial Systems (UAS) technology with the objective of speed up the final stages of imagery analysis and generation of indexes.

In conclusion, the Big Data Pipeline Unit's endeavors at NDSU represent an exemplary fusion of computational science and agricultural advancement. From innovative database creation and software development to the generation of impactful extension projects and pioneering UAS technology. The unit's holistic approach underscores the transformative potential of computational approaches in addressing contemporary agricultural necessities.

#07 **Machine learning in materials design**

Bakhtiyor Rasulev (Department of Coatings and Polymeric Materials)

The exponential growth and success of Artificial Intelligence (AI) and machine learning (ML) has resulted in their application in all scientific domains including material science. Advancement in experimental techniques has led to an increase in the volume of material science data, encouraging material scientists to investigate data-driven solutions to scientific problems. In this talk will discuss the potential and application of AI and ML in materials development and design. The focus on four aspects of the ML workflow will be discussed: (1) data collection, (2) encoding of material, i.e. feature representation, (3) model generation and selection, and (4) model validation and interpretation. Several cases will be shown, where predictive structure-property ML models development for nanomaterials and polymeric materials to predict various properties will be discussed.

#08 **From cells to societies: Unveiling patterns with multiplex network analysis**

Harun Pirim (Department of Industrial and Manufacturing Engineering)

This talk delves into a comprehensive exploration of multiplex networks, a specialized subtype of multilayer networks. These networks offer a unique approach to illustrating the intricate connections and interactions within complex systems, allowing for a detailed examination of the underlying relationships, persistent patterns, and latent communities. One use case applies a multiplex network framework innovatively to represent U.S. senators via their unique hashtag associations. This creative use of the multiplex network framework provides an efficient platform to scrutinize the layers within the political system using network metrics. The relationships between senators, usually challenging to discern due to the layered complexity of political connections, become more transparent and comprehensible when viewed through the lens of multiplex networks. Another significant use case explores the construction of a multiplex network by merging electric and gas networks. Five distinct scenarios, each representing a different network configuration, were considered in this process. These scenarios served as individual layers within the multiplex network. The integration of these diverse network conditions into a unified multiplex network provides valuable insights into the operation and interactions of energy systems. Moreover, another use case extends the application of multiplex networks to the field of biology, demonstrating their versatility in representing unique protein relationships. By harnessing the power of multiplex networks, researchers can gain a new perspective on protein interactions, offering valuable insights into cellular functions and disease processes. In conclusion, this talk underscores the significant importance and practical utility of multiplex networks. The ability of these networks to capture and depict the varying importance of connections within complex systems makes them an invaluable tool in various fields, ranging from cells to societies.

#09 **Advanced cyberinfrastructure for computational research and education**

Pete Lambertz (IT/Enterprise Computing & Infrastructure) and **Khang Hoang** (IT/CCAST)

This talk provides basic information about advanced cyberinfrastructure available for computational research and education at NDSU and beyond, including network, computing, and data storage infrastructure. We also

discuss possible challenges when transitioning to using advanced research computing, and CCAST's efforts and resources to help advance researcher capabilities and hence accelerate scholarly discovery.

#10 **Do sanctioning and monitoring affect elites' online toxicity? Evidence from a field experiment on U.S. general election candidates**

Dan Pemstein (Department of Political Science)

Toxic speech by electoral candidates can swamp out constructive political debate and exacerbate polarization. We conducted a field experiment among candidates for the 2022 U.S. general election to assess whether Twitter-based co-partisan sanctioning and email-based monitoring by researchers can reduce online toxicity. We use computational natural language processing techniques to systematically assess tweet toxicity. We find that candidates do not reduce their toxicity in response to in-group sanctioning. In fact, we find some evidence of a boomerang effect: in certain models, sanctioned candidates appear to become about 10% more likely to post a toxic tweet in the week following sanction. We also find a modest, but statistically significant, mitigating effect of monitoring on online toxicity among electoral candidates. At the same time, we find no evidence that monitoring causes candidates to tweet less often. Our findings suggest that making candidates aware of monitoring has the potential to reduce toxicity in political campaigns without reducing interaction between candidates and their voters.

#11 **Emerging technology driven memory system designs**

Sumitha George (Department of Electrical and Computer Engineering)

Technology changes can bring fundamental transformations on how a system works. As CMOS scaling reaching its physical limits and with the increasing demand for large volume data applications and low energy devices, emerging technologies seem to offer promising opportunities. In this talk, I will present the intriguing possibilities by the co-design of the device-circuit-architecture in the context of emerging devices focusing on the memory aspect. I will discuss the feasibility of using Ferroelectric FET (FEFET) as a potential choice for memory exploiting its unique features. In addition, I will introduce how emerging technologies enable new features such as Multi-Dimensional Access (MDA) in memories. Many applications, especially vision and matrix applications frequently generate sequences of data requests that belong to different rows in the memory and are therefore less efficient. This is solved by flexible access of MDA memories and would expedite applications from the fields of big data analytic, artificial navigation, and the Internet of Things (IoT) etc. This talk will also introduce future research directions including the utilization of emerging technologies for hardware security/peripheral designs.

#12 **Redesigning the next-generation pulse crops with genetic innovations**

Nonoy Bandillo (Departmnent of Plant Sciences)

The North Dakota State University Pulse Crops Breeding Program is critical to the development of new cultivars and germplasm of dry pea, chickpea, and lentil for the U.S. Northern Plains region. The program uses the conventional breeding method, and a network of field and controlled experiments to successfully develop elite cultivars. While the program has successfully used the conventional method to develop superior cultivars, making the conventional breeding process faster and more efficient remains to be a perennial challenge. I will talk about results and ongoing efforts for vetting and integrating genomic selection and high-throughput phenotyping into applied breeding efforts for redesigning next-generation varieties of pulse crops.

#13 **Leveraging large language models for genomic sequence analysis**

Changhui Yan (Department of Computer Science)

In various domains, such as bioinformatics and genomics, the transformative potential of Large Language Models (LLMs) has become increasingly evident. Notably, biological sequences exhibit remarkable parallels with natural languages, prompting a surge of interest in applying LLMs to unravel the intricacies of genomic

sequences. Our research group explores LLMs as a powerful tool for comprehending the evolution of SARS-CoV-2, the virus responsible for the COVID-19 pandemic.

#14 **Impact of covalent functionalization of nanostructures on their optical properties: Computational insights**

Svetlana Kilina (Department of Chemistry and Biochemistry)

In this talk, I will review current research of my group focused on computations of processes taking place at organic-inorganic interfaces in covalently functionalized nanomaterials such as, quantum dots (QD) and carbon nanotubes (CNTs), upon photoexcitation. Recent focus on assemblies of the QDs functionalized by various organic and metal-organic dyes is dictated by their promise to serve as a key element for both solar-to-electrical and solar-to-chemical energy conversion processes. Our simulations of QD/dye composites have led us to predictions of conditions that govern the direction and rates of the charge transfer from the QD to the dye and interpretation of transit-spectroscopy data. Our simulations of covalently functionalized CNTs have complimented and advanced recent experimental efforts in spectroscopy and chemistry of CNTs. A covalent binding of molecular adduct to the CNT, introducing sp3-hybridized defects into the sp2-lattice of the nanotube, results in redshifted optically active transitions providing defect-originated emission at the near infrared (NIR) range. Also, the sp3-defect in CNTs creates a required condition for single photon emission tunable from NIR to telecom wavelengths and achievable at room temperature. We have compared calculated and experimental results from pump-dependent low-temperature photoluminescence spectroscopy and identified the role of tube's chirality, tube's mode, adduct polarity, and defect-defect interactions in selective control of defect-associated emission of CNTs. Our results demonstrate that manipulation of the tube chirality together with the polarity and bond character of molecular adducts is a practical strategy for precise tuning of light emission in functionalized CNTs. Overall, our research outcomes have established foundation for the novel material design for solar energy conversion, sensing, and quantum technologies.

POSTER PRESENTATIONS

NS1 **Improving photoluminescence properties of carbon nanotubes with combined machine learning, computational chemistry and chemical substructure encoding approaches**

Gerardo M. Casanola-Martin,[1] Amirreza Daghighi,[1,2] August Amb,[3] Grace Tiffany,[3] Dmitri Kilin,[3] Svetlana Kilina,[3] and Bakhtiyor Rasulev[1,2]*

[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND, 58108, USA.
[2]Biomedical Engineering Program, North Dakota State University, Fargo, ND 58105, USA
[3]Department of Chemistry and Biochemistry, North Dakota State University, Fargo, ND, 58108, USA

Single-walled carbon nanotubes (SWCNTs) are nanomaterials widely used in industrial and biomedical applications. Computational studies have been actively applied for the last couple of decades to investigate physical, chemical, and mechanical properties of carbon nanotubes. Here we report for the first time a data-driven nano-QSAR study for a series of 85 SWCNTs with sp3-hybridized defects, to investigate various contributing factors and predict photoluminiscence properties of SWCNTs. Two machine learning (ML) models, namely Quantitative Structure-Property Relationships (QSPR) models were developed, to predict photoluminescence wavelength (nm) and oscillator strength of investigated SWCNTs. Molecular structures of SWCNTs were used as input information, on several levels, where encoded information was convalent functional groups, as well as the part of the SWCNT of a certain size. The obtained ML models showed good performances for the Emission Wavelength (nm) in the training set, R2=0.98 and validation set, R2=0.96. The model for Oscillator Strength values was also developed, where the determination coefficients were R2=0.92 and R2=0.87 for training and validation set, respectively. The QSPR models were internally validated by 10-fold cross-validation and y-scrambling experiments, confirming that the models have high levels of robustness and predictivity. The proposed approach provides an important tool to predict the photoluminescence

properties of SWCNTs, considering the part of the structure, related to the defect location, ligand and type of the nanotube, avoiding lengthy and complex calculations, assisting in design of SWCNTs with improved photoluminescence properties.

NS2 **Enhanced photoluminescent quantum yield in functionalized single walled carbon nanotube**

Dinesh Thapa, Dmitri Kilin, and Svetlana Kilina
Department of Chemistry and Biochemistry, North Dakota State University, Fargo, ND 58108

The local sp3-hybridized lattice defects in semiconducting single walled carbon nanotube (SWCNT), form new electronic states that efficiently trap excitons, leading to red-shifted emission features in the near infra-red (NIR) with increased photoluminescence quantum yields (PLQYs). In this context, we have used non-adiabatic molecular dynamics (NAMD) as implemented in VASP, to study the time dependent electronic and optical properties in (11,0) SWCNT functionalized with aryl molecule consisting of electron withdrawing group (NO2). We dynamically coupled electronic and nuclear degrees of freedom to go beyond Born Oppenheimer approximation by computing nonadiabatic couplings which then used in Redfield theory to explore nonradiative relaxation of the excited states to the band edges. The electron dynamics and photoluminescence of six different defect positions at the nanotube surface have been investigated and compared with that of pristine nanotube for different excitation energies. At the low excitation energy, 1.3 eV, hole relaxes faster than electron, whereas in larger excitation energies (>1.3 eV), hot electrons relaxes faster than hole, independent on the defect. The time evolution of charge density distribution indicates the electron and hole localization is separated in the tube with the dominant electron localization around the sp3 defect site, while holes have more delocalized character over the tube. Our calculations of PLQY based on Kasha's rule demonstrated the enhanced PLQY value for defect cases compared to the pristine SWCNT with significant enhancement by 72.0 % and 79.0 % for mostly localized and red-shifted ortho and para defects, respectively.

NS3 **Understanding Yellowstone cutthroat trout hybridization and connectivity within the Teton River system, a combined genomics and modeling approach**

Meggan Alston,[1*] Michael Youngwirth,[2] Alexandra Fraik,[3] Paul Hohenlohe,[2] and Travis Seaborn[1]
[1]School of Natural Resource Sciences, North Dakota State University, Fargo, ND
[2]Department of Biological Sciences, University of Idaho, Moscow, ID
[3]USDA Forest Service Rocky Mountain Research Station, National Genomics Center for Wildlife and Fish Conservation, Missoula, MT

Connectivity restoration efforts are an important strategy to assist native trout persistence across Western US river systems. However, increased connectivity may have unintended negative consequences, such as increasing hybridization rates between native and invasive species. This is particularly relevant for migratory species with multiple life history strategies. Here we use SNP genotyping data from single-digest RAD sequencing to assess population structure and estimate existing rates of hybridization and strength of assortative mating between Yellowstone cutthroat trout (Oncorhynchus clarkii boulveri) and non-native rainbow trout (Oncorhynchus mykiss) from the Teton River Basin. Samples were collected from approximately 470 individuals from several locations on the river mainstem and four tributaries. These include fish visually identified as cutthroat trout, rainbow trout and putative hybrids. Results from genetic analyses will be incorporated into demographic-genetic individual-based models to predict how different stream connectivity scenarios identified from stakeholder engagement may alter Yellowstone cutthroat trout life history variation and future hybridization outcomes. Understanding which restoration actions can facilitate native trout persistence while also limiting increased hybridization with heterospecifics will be key to informing future conservation and management decisions.

NS4 **Relaxation of photoexcited electron-hole pairs at Si(111) surfaces with adsorbed Ag cluster monolayers**

Yulun Han,[†] Tijo Vazhappilly,[‡] David A. Micha,[§] and Dmitri S. Kilin[†]
[†]Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58102, USA
[‡]Theoretical Chemistry Section, Chemistry Division, Bhabha Atomic Research Centre, Mumbai 400085, India

and Homi Bhabha National Institute, Mumbai 400094, India
[§]Quantum Theory Project, Departments of Chemistry and of Physics, University of Florida, Gainesville, Florida 32611, USA

Optoelectronic properties of semiconducting nanostructures have attracted much attention due to their applications in photovoltaics and photocatalysis. Transition metal nanoclusters show fascinating optical and chemical properties that depend on the size, shape, and composition of the clusters. Deposition of transition metal clusters onto semiconductor surfaces leads to stronger absorption of light and additional pathways for charge transfer compared to pure surfaces. In this work, we study relaxation dynamics of electron-hole pairs following photoexcitations for various Agn (n=0, 3, 7 and 10) cluster monolayers adsorbed on Si(111)/H surfaces. We first carry out density functional theory (DFT) calculations to obtain electronic structures. Results are presented in the form of density of states, band gaps, and light absorption, which allow for the investigation of the interaction of Ag clusters with Si. Different behavior can be expected depending on the size of the deposited Ag clusters. We then compute dynamics of electron-hole pairs using the combined electronic structure and reduced density matrix (RDM) treatment.[1-3] Specifically, we compute on-the-fly nonadiabatic couplings from DFT and process them using the RDM approach within the Redfield formalism. Nonradiative relaxation rates are noticeably different for various structures and transitions. One observes smaller relaxation rates for surfaces with adsorbates than pure Si surfaces due to charge transfer events with the contribution of Ag orbitals. Thus, introduction of adsorbates is advantageous for applications to photovoltaics and photocatalysis.

References: 1. D. S. Kilin, D. A. Micha. Surface Photovoltage at Nanostructures on Si Surfaces: Ab Initio Results. J. Phys. Chem. C 2009, 113, 9, 3530-3542.
2. D. S. Kilin, D. A. Micha. Relaxation of Photoexcited Electrons at a Nanostructured Si(111) Surface. J. Phys. Chem. Lett. 2010, 1, 7, 1073-1077.
3. T. Vazhappilly, Y. Han, D. S. Kilin, D. A. Micha. Electronic Relaxation of Photoexcited Open and Closed Shell Adsorbates on Semiconductors: Ag and Ag2 on TiO2. J. Chem. Phys. 2022, 156, 10, 104705.

S01 **Human colon cancer Dual Active Drug–Metal Organic Framework nanoMaterial (DADMOM) systems: Dataset compilation and artificial intelligence modeling**

Shan He,[1,2,3] Cristian R. Munteanu,[4,5,*] Begoña de Bilbao,[2,3] Harbil Bediaga,[3,6] Mariana Chelu,[7] Adina Magdalena Musuc,[7] Sonia Arrasate,[2] Alejandro Pazos,[4,5] and Humberto González-Díaz,[2,8,9,*]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58102, United States.
[2]Department of Organic and Inorganic Chemistry, University of Basque Country UPV/EHU, 48940 Leioa, Greater Bilbao, Basque Country, Spain.
[3]IKERDATA S.L., UPVEHU ZITEK, Rectorate Building, 48940 Leioa, Basque Country, Spain.
[4]Computer Science Faculty, University of A Coruna, CITIC, 15071 A Coruña, Spain.
[5]Biomedical Research Institute of A Coruña (INIBIC), University Hospital Complex of A Coruña (CHUAC), 15006 A Coruña, Spain.
[6]Department of Physical Chemistry, University of Basque Country UPV/EHU, 48940 Leioa, Greater Bilbao, Basque Country, Spain.
[7]"Ilie Murgulescu" Institute of Physical Chemistry, 202 Spl. Independentei, 060021 Bucharest, Romania.
[8]BIOFISIKA: Basque Center for Biophysics, CSIC-UPVEH, 48940 Leioa, Spain.
[9]IKERBASQUE, Basque Foundation for Science, 48011 Bilbao, Biscay, Spain.

Dual Active Drug – Metal Organic Framework (DADMOM) nano-materials are interesting for Human Colon Cancer (HCC) therapy. In these systems either the Drug (D) and/or the nanomaterial (NM) per se, by exert biological activity. However, the high number of combinations of HCC cell lines, anti-cancer drugs, and/or NM to be tested slows down the assay process of new DADMOM-HCC systems. Nevertheless, the data scarcity and high complexity of these systems make difficult AI/ML studies. Here, a new dataset of DADMOM-HCC

systems was compiled from literature. Herein 11 different Artificial Intelligence/Machine Learning (AI/ML) algorithms were used to seek the predictive models. The LDA and Random Forest (RF) models showed high values of sensitivity and specificity (> 0.9) in training/validation series and 3-fold cross validation respectively. The new AI/ML models are able to predict 14 output properties (CC50 (μM), EC50 (μM), Inhibition (%), etc.) for all combinations of 54 different NM cores classes vs. 15 different coats and vs. 41 different cell lines allowing to short list the more interesting results for experimental assays. It may reduce the cost of the traditional trial and error procedures.

Keywords: decorated NM; drug delivery; colon cancer; perturbation theory; machine learning.

S02 **PCR IF+PT+ML metagenomic study of lung microbiome in cystic fibrosis: Toward GDPR personalized medicine compliant decision making**

Estefania Asencio-Medina,[1,2,3,4] Carmen Velásquez,[2] Amaia González-Magaña,[2] Maria D. Pastor-Vivero,[2,5] Laura Acosta,[3] Aliuska Duardo-Sánchez,[6] Sonia Arrasate,[7] Humberto González-Díaz,[2,7,8] David Albesa-Jové,[2,7,8,*]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, North Dakota 58102, United States.
[2]BIOFISIKA: Basque Center for Biophysics CSIC, University of The Basque Country (UPVEHU), Barrio Sarriena s/n, 48940, Leioa, Basque Country, Spain.
[3]CITIC-Research Center of Information and Communication Technologies, Department of Computer Science and Information Technologies, University of A Coruña (UDC), 15071, A Coruña, Spain.
[4]IKERDATA S.L., ZITEK, University of The Basque Country (UPVEHU), Rectorate Building, 48940 Leioa, Spain.
[5]Cruces University Hospital, IIS Biocruces-Bizkaia, 48903, Barakaldo, Basque Country, Spain.
[6]Department of Public Law, University of the Basque Country UPV/EHU 48940, Leioa, Biscay, Spain.
[7]Department of Organic and Inorganic Chemistry, Faculty of Science and Technology, University of The Basque Country (UPV/EHU), P.O. Box 644, 48080, Bilbao, Spain.
[8]IKERBASQUE, Basque Foundation for Science, 48011, Bilbao, Spain.

Cystic Fibrosis (CF) is a genetic disease caused by mutations on chromosome 7, leading to thick lung mucus and increased susceptibility to microbial colonization. Although antibiotic therapy is the primary treatment, antibiotic resistance can develop over time. Traditional microbiological diagnosis via culture has limitations due to sensitivity issues and delayed results.

Fortunately, Next-generation sequencing (NGS) offers a promising solution by enabling comprehensive identification of known and novel pathogens. Targeted NGS, using conserved genetic sequences, allows for precise identification of bacteria at the species level.

However, to find optimal solutions for CF patients, it is crucial to consider various patient variables, including exacerbations, treatment history, age, CFTR protein mutations, and the presence of other health issues, along with metagenomic data. This is where Artificial Intelligence/Machine Learning (AI/ML), particularly the IFPTML model (Information Fusion + Perturbation Theory + Machine Learning) is needed.

Our study is based on a dataset collected from CF patients in Bilbao, Spain, which includes clinical, pharmacological, and microbiome data obtained through both clinical microbiology culture (CMC) and PCR metagenomics sequencing.

We have leveraged IFPTML predictive models to enhance patient follow-up and develop personalized treatment plans. These models encompass both Lineal Discriminant Analysis (LDA) and non-linear approaches, including Artificial Neural Networks (ANN). We have also conducted comparisons between IFPTML model outcomes and results obtained through CMC and PCR techniques. To ensure GDPR compliance and patient privacy, we've generated Synthetic Data (SD) using Monte Carlo methods while rigorously testing the model's robustness.

In summary, this study integrates AI/ML and NGS technologies to advance CF patient care. This integration enhances diagnostic accuracy, enables precise prognosis, and facilitates the creation of personalized treatment strategies.

S03 **A hybrid staggered/non-staggered formulation for simulating incompressible flows with block-structured mesh refinement**

Thien-Tam Nguyen,[1] Rajasekhar Dathi,[1] Andy J. Nonaka,[2] and Trung B. Le[1]
[1]Civil, Construction, and Env. Eng. Dept., North Dakota State University.
[2]Lawrence Berkeley National Laboratory.

We present a novel approach for simulating incompressible flows with local mesh refinement. In many biological problems such as simulations of cells, it is highly desirable to accommodate high-resolution regions near the surface of moving bodies in flows. In this work, we present a new approach for local mesh refinement with the dual use of staggered and non-staggered grid layouts. Our finite volume solver for incompressible flows is based on a fractional step method. The fluxes are stored at the surface centers whereas the pressure field and the Cartesian velocity components are at the volume centers. This hybrid staggered – non-staggered approach allows the flexibility of prescribing the boundary conditions on the moving bodies while satisfying the incompressibility constraint exactly. We use the Adaptive Mesh Refinement for Exascale (AMReX) framework for designing our grid infrastructure. The momentum equation is solved iteratively with an implicit Runge-Kutta method. The native Poisson solver of AMReX is utilized for the projection step. Our preliminary work includes two benchmark simulations: (a) lid-driven flow; and (b) Taylor-Green Vortex to demonstrate the feasibility of our approach. We will report on the efficiency and scalability of this approach for different grid sizes in different heterogenous computing infrastructures.

Acknowledgements: This work is supported by the Department of Energy (Sustainable Research Pathways project) and the National Science Foundation Grant #1946202.

S04 **A heuristic agent-based model of the market as a distributed genetic algorithm**

Zach Johnson and James Caton
North Dakota State University

Economic theory indicates that competition promotes minimization of production costs while allocating resources toward their most highly valued use. The static neoclassical model, however, does not capture economic evolution, wherein firms learn more efficient methods of producing and exchanging goods. In macroeconomics, such learning is approximated by the total factor productivity of capital. Our dynamic agent-based simulation on the Sugarscape (Epstein and Axtell, 1996) models the market process as a genetic algorithm wherein naïve, heuristic-based firms have inefficient strategies removed as a result of natural selection. It is found that like the neoclassical model, competition amongst the naïve agents promotes price equilibration and sustainably high market population, while allowing agents to make economic decisions in constant runtime. The dynamic model allows for learning that tends to positively influence total factor productivity, a consequence of the removal of inefficient strategies.

References: Epstein, Joshua M., and Robert Axtell. Growing artificial societies: social science from the bottom up. Brookings Institution Press, 1996.

S05 **Effect of the crystal structure on optical response of HgSe quantum dots**

Steven Westra and Svetlana Kilina
Chemistry and Biochemistry Department

Our research focuses on computations based on Density Functional Theory (DFT) of three distinct crystal structures of stoichiometric mercury selenide (HgSe) quantum dots (QDs): One characterized by the Wurtzite structure, and two representing isomers of the Zinc Blend structure, differentiated by symmetric and asymmetric configurations. Additionally, we explore the size-dependent behavior of HgSe QDs varying the

number of Hg/Se atoms from 33/33 and 34/34 to 48/48, and 54/54 atoms. All structures show a significantly perturbed crystal lattice symmetry. To gain deeper insights into the excited state properties of QDs, we employ Time-dependent Density Functional Theory (TD-DFT) to calculate absorption spectra and natural transition orbitals (NTOs). Calculated absorption spectra reveal a consistent energy shift in the brightest transition states as the quantum dot size increases, following the sequence 33/33 > 48/48 > 54/54 in accordance with the quantum confinement effects. Independent on the QD size, Zinc Blend structures exhibit a manifold of optically inactive transitions at the lowest-energy spectral contributed from surface-localized "trap" states. In contrast, the lowest energy transitions are highly optically active in Wurtzite QDs. The number of low-energy states with low optical activity in Zinc Blend QDs is expected to lead to competitions between radiative emission to the ground state and non-radiative relaxation and redistribution channels between the low-energy, surface-associated states, which significantly reduce their overall emission efficiency, compared to Wurtzite QDs. Our calculations provide important insights for controlling optical response in HgSe QDs via their structural engineering.

S06 **Discovery of an effective α-amylase inhibitor via machine learning-based QSAR approach**

Mariam Zamani,[1,2] Gerardo Casanola-Martin,[1] and Bakhtiyor Rasulev[1,2]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND
[2]Biomedical Engineering Program, North Dakota State University, Fargo, ND

α-Amylase (1,4-α-d-glucan-glucanohydrolase, EC 3.2. 1.1) is one of the key enzymes that help digest complex carbohydrates and convert them into simple sugar in the gastrointestinal tract. This process contributes to post-prandial (post-meal) hyperglycemia (Dona et al., 2010). Inhibitors of α-amylase slow down the carbohydrate digestion process by inhibiting the enzyme. As a result, the entry of glucose into the circulation is controlled. Our focus in this project is an investigation of the structure-activity relationship for a series of small molecule amylase inhibitors using a computational approach, for further search and discovery of effective α-amylase inhibitors. Such effective inhibitors can help delay glucose absorption, control blood glucose, maintain glycemic levels, and minimize hypoglycemic risks.

Methods: A series of small molecules with α-amylase inhibitory activity were analyzed via machine learning-based Quantitative Structure-Activity Relationship (QSAR), applying genetic algorithm and multiple linear regression (GA-MLR) techniques.

Results: The GA-MLR model was developed using a substantial dataset comprising compounds with α-amylase inhibitory activity for identifying optimum lead anti-diabetic compounds.

Conclusion: Through the ML-based QSAR techniques, our model showcases the strong structure-activity of compounds demonstrating α-amylase inhibitory activity. This approach holds great potential for identifying anti-diabetic therapeutic applications.

S07 **Tungsten coordinated complexes: The effect of DFT functionals in comparison with experimental data**

Tori Oas,[1] Grace Tiffany,[1] Hang Wang,[2] Wenfang Sun,[2] and Svetlana Kilina[1]
[1]Department of Chemistry and Biochemistry, North Dakota State University, Fargo ND 58108
[2]Department of Chemistry and Biochemistry, The University of Alabama, Tuscaloosa AL 35487

Obtained insights into optical properties of metal coordinated complexes are important for potential applications as red to near-infrared emitters. This leads to applications in isomerization, photocatalysis, and photodynamics. Our focus on tungsten is especially important, as working with earth abundant metals leads to a more economic research and development process. In our studies, we utilize Density Functional Theory (DFT) and Time-Dependent DFT (TD-DFT) to investigate the ground and excited state properties of tungsten carbonyl pyridyl-imidazole coordinated complexes. To confirm the validity of the computations, calculated optical spectra are compared with experimental data. It is found the B3LYP and PBE0 hybrid functionals best aligned with experimental data. Comparisons were also made between the different ligands and substituting groups. It is found that ligands with increasing conjugation have more redshifted absorption spectra out of all

ligands examined. It is also found that increasing the length of carbon chains on the substitution group has no effect on the lowest energy optical band, while adding an electron donating (EDG) or an electron withdrawing (EWG) group to the substitution group causes a slight blue-shift and red-shift, respectively.

S08 **A numerical study of particle motion in a turbulent pipe flow**

Rajasekhar Dathi,[1] Imtiaj Nahin Ahmed,[2] Hong Pan,[1] Trung B. Le,[1] Yan Zhang,[2] and Zhibin Lin[1]
[1]Civil, Construction, and Env. Eng. Dept., North Dakota State University.
[2]Mechanical Eng. Dept., North Dakota State University.

We investigate dynamics of sand particles in gas pipelines using experimental measurements and numerical simulations. Two configurations (45 and 90 degree) of pipeline bends are considered under working conditions of gas pipeline. A structured block mesh is created with 2.8 million grid points to simulate the turbulent gas flows using Large Eddy Simulation. An uniform flow profile is applied at the inlet while a no-slip and no flux boundary condition are applied on the pipe surface. The numerical method for particle-laden flows is based on one-way coupling approach, which only consider the impact of the turbulent flows on the particles. The particle is tracked continuously through the domain and its governing equation is based on the Newtons second law, which includes the drag force, the pressure gradient force, the added mass force, and the buoyancy force. The computational results on the particle deposition are compared with the erosion behavior in the corresponding experiments, which illustrate the erosion process mainly at the pipe bend due to the impacting of the sand particles. Our results show the critical role of the separation region in the pipeline bend in sweeping of the particles on the surface of the pipe, which suggest the dependence of erosion rate on the flow pattern in the bend.

Acknowledgements: This work is supported by US DOT (Pipeline and Hazardous Materials Safety Administration) Agreement number 693JK32110003POTA. This work used resources of the Center for Computationally Assisted Science and Technology (CCAST) at North Dakota State University, which were made possible in part by NSF MRI Award No. 2019077.

S09 **qsarify: High performance machine learning software package for QSAR model development**

Stephen Szwiec and Bakhtiyor Rasulev
Department of Coatings and Polymeric Materials, North Dakota State University

qsarify is a new software package for the development, validation, and visualization of machine learning models for Quantitative Structure-Activity Relationship (QSAR) and Quantitative Structure-Property Relationship (QSPR) studies. Written in Python and freely available under the GNU General Public License, this software package provides a focused workflow for the generation of predictive statistical models to better explain and predict the relationship between molecular structure and biological activities or chemical properties. qsarify implements an innovative algorithm to reduce input dimensionality during the feature selection process, utilizing cophenetic clustering followed by a genetic algorithm (GA) for variable selection. Implementing both serial and parallel processing, this algorithm allows for rapid predictive model development and validation of large chemical datasets on low performance computers, while also allowing for complex model development and validation on high performance computers by utilizing multi-processing. Finally, the software also provides for output the of statistical validation metrics and generates plots for model diagnostics, including Williams Plot and Y-scrambling tests.

S10 **Predicting the toxicity of fullerene derivatives using classification machine learning models**

Madeleine Williams, Amirreza Daghighi, and Bakhityor Rasulev

The toxicity of carbon-based nano molecules gives important insight into application of the molecules. Fullerene and fullerene derivatives (FDs) have previously been shown to be highly customizable to many applications. In this work a Quantitative Structure Activity Relationship (QSAR) model has been developed to predict the toxicity of the FDs. The ability of twelve machine learning models to consistently classify the FDs as toxic or nontoxic was examined. Information on 169 FDs was obtained from a previous work. The end point

of the model was binary, the FD is either toxic or nontoxic. 5,666 descriptors were calculated in alvaDesc, and a subset of important features was created using forward feature selection. External validation and AUROC were used to determine the most effective model. Internal validation was used to confirm the consistency of the model. The applicability domain was defined with the range, Euclidean distance, and probability calculated using Ambit Discovery software. Logistic regression was found to be the model with the highest external evaluation metrics with an accuracy of 98.6%. Internal validation metrics were also high with a mean average accuracy of 97.8%. Three influential descriptors were found, ATSC8i, DP08 and CATS2D_04_DA. A physical understanding of the descriptors and correlation to the output of the model was determined. A virtual library of 41803 FDs was generated and then categorized with the QSAR model. With the results of this work, query FDs' toxicity can be predicted. Understanding of this prediction can be established from the formula for the logistic regression algorithm. Future manipulation of the molecule will be assisted by the physical representation of the descriptors.

S11 **Machine learned force fields assisted molecular dynamics in VASP**

David Graupner and Dmitri Kilin
Department of Chemistry and Biochemistry, North Dakota State University, Fargo, ND 58108

Molecular dynamics (MD) is a vital method for understanding dynamics properties, but the quality of the MD simulation depends on the accuracy of the calculation method. However, more accurate calculations usually come at the cost of increased computational demand. One highly accurate method is ab initio MD, which treats the interactions of atoms and electrons fully quantum mechanically from ab initio calculations, however this method is usually limited to small computation times and simulation cells. Using force fields, which are parametrizations of the potential energy, can greatly speed up the calculations, but determining accurate force fields is time consuming and requires tremendous expertise. Machine learning is a way to significantly reduce computational cost by interpolating between known training systems, calculated ab initio. Selecting the appropriate training data remains a challenge. This is addressed by using on-the-fly learning, which continuously builds a force field using the ab initio MD data. At each step, it is determined whether to perform an ab initio calculation or to use the force field and skip learning for that step. Vienna Ab initio Simulation Package (VASP) uses a Bayesian-learning algorithm for on-the-fly machine learning. The total energy and forces are predicted based on the machine-learned force field at each time step of the MD simulation and if the Bayesian error estimate exceeds a threshold an ab initio calculation is performed. Here we use the machine learning force fields determined by VASP to calculate the adiabatic molecular dynamics trajectory for a Dion-Jacobson lead chloride perovskite and compare it to the ab initio adiabatic molecular dynamics trajectory for the same model. In a future work, it is hoped that we can use this method to calculate accurate molecular dynamics trajectories that can then be used to calculate the "on-the-fly" nonadiabatic couplings to be used for excited-states dynamics.

DG thanks NSF-2004197, DK thanks NSF-1944921. Both DG & DK acknowledge CCAST, NSF MRI-2019077, and NERSC, DE-AC02-05CH11231 using allocation award m1251, for resources.

S12 **Streamlining data querying with exlibris: A data pipeline approach**

Grant Hoff,[1] Souradip Roy,[1] Robert Fedor,[1] Matheus Siqueira,[1] Reed Fox,[1] Sikiru Atanda,[1] Didier Florez,[1] Lane O'Brien,[1] David Froslie,[1,2] and Ana Heilman-Morales[1]
[1]AES Big Data Pipeline Unit, NDSU
[2]Department of Computer Science, NDSU

Exlibris, developed by the AES Big Data Pipeline Unit at NDSU, is a powerful tool designed to streamline data querying from Genovix, a proprietary plant breeding database management software. In pursuit of improved efficiency in data retrieval through Exlibris, we have introduced a robust data pipeline system. This pipeline is dedicated to fetching, transforming, and storing data sourced from Genovix's SQL Server database into a PostgreSQL database, meticulously optimized to facilitate continuous access. Traditionally data retrieval from Genovix needed intricate real-time queries, involving the joining of multiple tables and on-the-fly result

filtering. Our innovative approach involves preprocessing of the data by performing anticipatory joins, saving the outcomes in a PostgreSQL database, and categorizing data into raw data, means data, and experiment categories, all organized by year. To maintain data currency, we have implemented an automated script that updates the tables daily. Previously, querying Genovix's database directly incurred a complexity of O(nlogn), a result of the intricate table joins (O(nlogn)) and subsequent filtering (O(n)). With the implementation of our data pipeline, users can now opt for pre-joined data tables from PostgreSQL (O(1)), combine them through union operations (O(1)), apply data access-based filtering (O(n)), and sort the results (O(n)) when requesting data for a specific range of years. This optimized approach reduces data retrieval complexity to O(n), significantly improving the efficiency of data access. Furthermore, additional filtering happens at the front-end level, further enhancing query efficiency (O(n)). In conclusion, our data pipeline represents a significant leap in improving Exlibris's performance by simplifying the complexity of operations executed when users submit queries. Looking ahead, we are committed to further enhancing efficiency by pre-sorting data and partitioning PostgreSQL tables to allocate dedicated tables for each dataset, category, and year. This future development has the potential to reduce data retrieval complexity to a remarkable O(1), ensuring a smoother and more efficient data querying experience.

Keywords: data pipeline, PostgreSQL database, optimized querying, ExLibris application.

S13 **From raw images to vegetation indexes: automating the UAV image data processing workflow for precision agriculture**

Robert Fedor,[1] Matheus Siqueira,[1] Paulo Flores,[2] Reed Fox,[1] Grant Hoff,[1] Sikiru Atanda,[1] Didier Murillo-Florez,[1] Souradip Roy,[1] Lane O'Brien,[1] David Froslie,[1,3] Ana Heilman-Morales[1]
[1]AES Big Data Pipeline Unit, NDSU
[2]Agricultural and Biosystems Engineering, NDSU
[3]Department of Computer Science

Advancements in Unmanned Aerial Systems (UAS) technology have started to transform research endeavors at North Dakota State University, with a particular focus on precision agriculture, agronomy, and plant breeding research. Within the Department of Plant Science, which has ten diverse breeding research programs, five have adopted UAS technology with aims to expedite their selection process. Simultaneously, research extension centers are utilizing UAVs with the hope to streamline the labor-intensive task of collecting phenotypic data. However, the integration of UAV data presents challenges in handling the voluminous terabytes of data generated during these missions. While UAVs significantly enhance data acquisition efficiency, data processing remains burdened with manual tasks. Currently, a multitude of software applications exist in the market and these help to convert raw image data into essential vegetation indexes for phenotypic analysis. Unfortunately, many of these applications are either specialized and require researchers' having an in-depth knowledge on programming languages or are cost prohibitive for some research programs.

At NDSU, initial attempts to automate UAV workflows have resulted in a collection of a dozen Python scripts. In response, we proposed to generate a comprehensive software solution called AgSkySight that seamlessly integrates these scripts into a unified application. The proposed software automates the entire workflow, encompassing image stitching, shapefile generation for plot identification, and vegetation index extraction. Importantly, this application centralizes computing resources, relieving users of the computational burden. Notably, we have condensed the Python scripts from a dozen to just two shorter scripts, thereby streamlining the process significantly. This integrated approach marks a substantial advancement in the workflow automation. Preliminary results indicate that ongoing development of the AgSkySight software will dramatically reduce the workload for numerous researchers, translating into substantial cost and time savings. This software not only enhances data processing efficiency, but also ensures that computing resources are maximized, ultimately fostering the advancement of precision agriculture and plant breeding research at NDSU.

Key words: UAV (Unmanned Aerial Vehicle), precision agriculture, vegetation index, image data processing, workflow automation, phenotypic data, data integration, python scripts, image stitching, shapefile generation.

S14 **Modernizing NDSU pest management app: Enhancing pesticide knowledge and accessibility for North Dakota's agricultural stakeholders**

Reed Fox,[1,2] Robert Fedor,[1,2] Andrew Friskop,[3] Matheus Siqueira,[1,2] Grant Hoff,[1,2] Sikiru Atanda,[1] Didier Murillo-Florez,[1] Souradip Roy,[1] Lane O'Brien,[1] David Froslie,[1,3] Ana Heilman-Morales[1]
[1]Big Data Pipeline Unit, Ag Experiment Station, NDSU
[2]Department of Plant Pathology, NDSU
[3]Department of Computer Science, NDSU

The increase in pesticide use to protect crops in agriculture has resulted in higher yields and profitability for growers. However, careless use of pesticides can lead to negative drawbacks including environmental contamination. To counteract and reduce the frequency of negative pesticide impacts, Extension programs have been implemented to promote pesticide stewardship. The Extension programs are extremely successful, but there is still an information gap from specialists to producers. This includes easy to obtain information on pesticides such: what to use, classes and nomenclature of pesticides, how much to use, and/or when to apply. To address these questions, the Pest Management Application emerges as a valuable tool, acting as a medium to disseminate unbiased scientific knowledge on pesticide use in the Upper Great Plains.

Currently the tool is undergoing a comprehensive modernization effort, transitioning away from its older mobile iOS version built on the Laravel framework version 5.0, which was susceptible to several security vulnerabilities. To ensure enhanced performance and security, we have undertaken a substantial revamp of the application, adopting a more contemporary architecture using the React framework and migrating the data to a PostGRE SQL database. Furthermore, we are in the process of developing an efficient admin panel using Django. The original application's user interface had significantly deteriorated over time, and the anticipated expenses for updating the framework were prohibitively high, rendering it challenging to maintain and utilize. Our current iteration has addressed these issues by not only redesigning the UI, but also optimizing the database structure, facilitating seamless data entry through a web-based application. The AES Big Data Pipeline unit eagerly anticipates the utilization of the Pest Management Application and its potential impact on both research and local farmers. The application's capacity to efficiently query pesticide information represents a significant advancement. We are confident that this tool will play a pivotal role in aiding decision-making processes related to pesticide use. We remain committed to further refining the application to enhance its efficiency and effectiveness, with a strong focus on promoting sustainable pesticide practices and supporting the vital agricultural landscape of the Upper Great Plains.

Keywords: Pest Management Application, Extension NDSU, PostGRE SQL, React, pesticide stewardship, software development.

S15 **Integrating experimental and computational approaches to assess the performance of organic corrosion inhibitors through QSAR modeling and extrapolation techniques**

Chameli Samarawickrema,[1] Gerardo Casanola-Martin,[2] Zachary Bergseth,[2] Xiaoning Qi,[2*] Bakhtiyor Rasulev[1*]
[1]Computational Polymer Science Lab, Department of Coatings & Polymeric Materials, North Dakota State University, Fargo, ND, USA
[2]Corrosion Science Lab, Department of Coatings & Polymeric Materials, North Dakota State University, Fargo, ND, USA

Corrosion inhibition is a critical aspect of materials science and engineering, particularly in the context of safeguarding infrastructure and industrial equipment. Organic compounds are significant as effective corrosion inhibitors due to their versatile chemical structures and inhibition mechanisms. Analytical identification of potential corrosion inhibitors is extremely time consuming, resource-consuming, and costly,

and primarily relies on trial-and-error experimentation. Consequently, there is an urgent need to explore the relationship between a compound's structure and its corrosion inhibition efficiency, which would provide crucial theoretical insights into designing effective corrosion inhibitors. In the first part of this study, a preliminary predictive QSAR model is constructed using in-house data, establishing correlations between the molecular structure and corrosion inhibition efficiency of diverse organic compounds. In the second part of this study, the formidable challenge of extrapolating corrosion inhibition efficiency data from specific experimental conditions to diverse conditions documented in broader scientific literature has been explored. The study explores the influence of various experimental parameters on corrosion inhibition efficiency. By systematically altering these conditions, a comprehensive dataset is compiled, representing a wide range of potential scenarios. Statistical and computational approaches, such as regression analysis and machine learning algorithms, are employed to develop predictive models that can account for variations in experimental conditions. These models allow for the extrapolation of corrosion inhibition efficiency data to conditions that may not have been directly studied. A comprehensive review of the existing literature on corrosion inhibition is conducted to identify trends and patterns in inhibitor performance across different experimental setups. The extrapolated data is compared with literature-reported values to assess the validity and accuracy of the developed models.

S16 **β-cyclodextrin–phenolic acid complexation study to mask bitterness in wheat bran (A QSPR approach)**

Kweeni Iduoku,[1,2] Marvellous Ngongang,[1] Jayani Maddakandaged,[3] Amir Daghighi,[1,2] Gerardo M. Casanola-Martin,[1] Senay Simsek,[3] and Bakhtiyor Rasulev[1,2*]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND 58105
[2]Biomedical Engineering Program, North Dakota State University, Fargo, ND 58105
[3]Department of Plant Sciences, North Dakota State University Fargo, ND 58105

The need for solvating and encapsulating hydro-sensitive molecules drives noticeable trends in the applications of cyclodextrins in pharmaceutical, food, polymer, materials, and agricultural science. Among them, β-cyclodextrin is one of the most used for the entrapment of phenolic acid compounds to mask the bitterness of wheat bran. This study uses a dataset of 20 phenolic acids docked to the β-cyclodextrin cavity to generate three different binding constants. Those values from docking methods were used with topological, topographical, semi-empirical, and quantum-chemical descriptors from the ligands in a QSAR study. Three different models for each binding constant were computed using a Genetic Algorithm (GA) and Multi-Linear Regression (MLR) combined approach. The predictive ability of the QSAR-MLR models was determined, and the binding score showed the best performance in training and test sets with correlation coefficients of 0.969 and 0.984, respectively. The relationship between the models and molecular descriptors was analyzed, and some factors were revealed, showing positive contributions toward the binding constants affinity values like the presence of rings in the molecules. Other factors influencing the complexes' binding constants, such as branching, electronegativity values, and polar surface area, were also discussed.

S17 **Development of machine learning-based QSAR models to predict green algae growth inhibition by polymeric particles**

Seyedehelham Shirvani-Hosseini, Gerardo M. Casanola-Martin, and Bakhtiyor Rasulev
Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND 58102, USA

Polymer products, including rubbers, synthetic plastics, coatings, fibers, and adhesives, have found increasingly broad applications and have become essential materials in modern life. However, the extensive use of polymeric materials has also brought a rising risk of ecosystem contamination, particularly in aquatic environments. As a consequence of the utilization and disposal of chemical products, diverse polymer particles enter aquatic ecosystems, creating potential toxicity hazards for various organisms, including microorganisms, such as green algae. The utilization of in silico methods for predicting material properties and bioactivities has proven to be a cost-effective and time-efficient approach in assisting scientists. We have collected a dataset with toxicity to green algae tested on various polymers and their particles. To address this

task, we have applied machine learning-based cheminformatics technique called a Quantitative Structure–Activity Relationship (QSAR) to find a structure-toxicity correlation and develop a predictive model for the toxicity of a set of polymeric materials which can estimate the green algae growth inhibition (EC50). This cheminformatics-based approach has been employed to identify the high-relevance features/descriptors responsible for growth inhibition using various machine learning methods by addressing the structure-toxicity relationship for investigated polymeric materials. The best models are selected based on their high-performance accuracy (R2training, Q2, RMSE) and validation (R2test, y-scrambling).

S18 **Are we ready for hydrogen: A comprehensive evaluation of existing energy pipeline infrastructure for its transport**

Mohsin Ali Khan,[1] Hong Pan,[1] Kevin Wang,[2] and Zhibin Lin[1]
[1]Department of Civil, Construction and Environmental Engineering, North Dakota State University, Fargo, ND 58018, USA
[2]Department of Aerospace and Ocean Engineering, Virginia Polytechnic Institute and State University, Blacksburg, VA 24061, USA

The synergy between natural gas pipelines and the growing hydrogen industry holds great potential. Yet, this agreement must be weighed against the potential ecological implications of establishing extensive hydrogen transportation networks. While the scientific literature of experimental investigations into pipeline steel, their collective contribution remains insufficient to guarantee for it as a robust industrial solution. The current study endeavors to bridge this gap through a scientometric-based systematic analysis, scrutinizing the available research on pipeline steel's fatigue crack growth rate, with a keen focus on hydrogen embrittlement. Scientometrics target to deliver findings that are not just impartial and rigorous but also all-encompassing, free from the influence of individual subjectivity. The complicated fatigue crack growth rate of pipeline steels under the influence of hydrogen, affected by a group of variables such as pressure, load dynamics, gas composition, temperature fluctuations, and the deployment of inhibitors, underlines the complexity of the challenge. Drawing upon the insights derived from our scientometric evaluation and traditional literature review, we confidently assert that existing pipeline infrastructure offers a compelling and cost-effective path for blending hydrogen into established natural gas pipelines. However, it is strongly advocated that upcoming research endeavors prioritize a deeper exploration of fatigue behavior concerning parental cracks, dent formation, gouge development, corrosion-induced impacts, and weldments distressed with diverse imperfections. Additionally, we emphasize the requirement of investigating fatigue behavior under the concurrent influence of several flaws, ensuring the secure and effective transportation of hydrogen within the existing pipeline network.

S19 ***Ab initio* calculations of through-space and through-bond charge-transfer properties of interacting Janus-like PbSe and CdSe quantum dot heterostructures**

Hadassah B. Griffin,[1] Andrei B. Kryjevski,[1] and Dmitri Kilin[2]
[1]Department of Physics, North Dakota State University, Fargo, ND 58108, United States
[2]Department of Chemistry and Biochemistry, North Dakota State University, Fargo, ND 58108, United States

Heterostructure quantum dots (QDs) are composed of two QD nanocrystals (NCs) conjoined at an interface. They are useful in applications such as photovoltaic solar cells. The properties of the interface between the NCs determine the efficiency of electron-hole recombination rates and charge transfer. Therefore, a fundamental understanding of how this interface works between the two materials is useful. To contribute to this understanding, we performed ab initio calculations to simulate two isolated heterostructure QD models with Janus-like geometry composed of Cd33Se33+Pb68Se68 NCs. The first Janus-like model has a bond connection between the two NCs and is approximately 16x17x29 Å3 in size. The second model has a through-space connection between the NCs and is approximately 16x17x31 Å3. Individual NC composites of PbSe and CdSe were also simulated for comparison. We use density functional theory to simulate the ground state properties of these models. Nonadiabatic on-the-fly couplings calculations were then used to construct the

Redfield Tensor, which described the excited state dynamics due to nonradiative relaxation. From our results, we identified a qualitative trend which shows that having a bond connecting the two NCs reduces hole relaxation time. We also identified, for a sample of electron-hole excitations pairs, that the through-bond model allows for a net positive or negative numerical net charge transfer, depending on the excitation pair.

Support of National Science foundation via NSF CHE-2004197 is gratefully acknowledged.

S20 **Demystifying cyber insurance adoption for small businesses**

Zia Muhamad and Jeremy Straub
Department of Computer Science, North Dakota State University (NDSU)

In the interconnected world of digitization, small businesses are vulnerable to cyber threats due to limited resources and a lack of protection. These businesses are required to evolve to defend services and business operations from cyber criminals. This requires a proactive risk mitigation approach and a risk transfer strategy. The poster will explain why cyber threats are inevitable and how every business is vulnerable. It will discuss some statistics and graphs on the trends, frequency, and severity of cyber-attacks on small businesses and the benefits of cyber insurance. The poster identifies risk mitigation strategies that companies can adopt to improve their cybersecurity posture, which will significantly reduce their cyber insurance costs. Finally, the poster proposes a risk transfer strategy in terms of a 'Cyber Insurance Benchmark', a three-tiered diagram that helps small businesses determine the optimal coverage level. The benchmark has three levels, ranging from low to high, and each level corresponds to different types of cyber risks and coverages. The goal is to explain how small businesses can use these strategies to build cyber resilience that best meets their situation.

S21 **Multi-endpoint acute toxicity assessment of organic compounds using large-scale machine learning modeling**

Amirreza Daghighi,[1,2] Gerardo M. Casanola-Martin,[1] Kweeni Idouku,[1,2] and Bakhtiyor Rasulev[1,2*]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND 58102, USA
[2]Biomedical Engineering Program, North Dakota State University, Fargo, ND, 58102, USA

In recent years, alternative animal testing methods such as computational and machine learning methods, have become increasingly crucial for toxicity testing as they allow for minimizing animal examination while reducing costs and time. Machine learning is a powerful tool for in silico discovery in drug development and environmental chemical screening. However, the complexity and scarcity of available biomedical data challenge the development of predictive models. A combined approach of state-of-the-art non-linear machine learning methods and multi-condition descriptors offers a solution for combining data from various assays to create a more accurate and robust model. This study uses multi-condition descriptors (MCDs) to develop a QSTR (Quantitative Structure-Toxicity Relationship) model based on a large toxicity dataset of over 80,000 compounds and 59 endpoints, leading to 122,572 data points. The study discusses the prediction capabilities of developed seven single-task multi-endpoint machine learning models and a novel method of data analysis using Convolutional Neural Networks (CNN) to develop QSTR models. The results show that using MCDs significantly improves the predictability power of the model and using them with CNN-1D yields the best results (R2=0.93, R2ext =0.70). Several structural features, such as van der Waals surface area, number of nitrogen-containing fragments, presence of S-P fragments, ionization potential, and presence of C-N fragments, showed the highest contribution to toxicity. These models can be useful tools for predicting the toxicity of various compounds under different conditions, enabling quick toxicity assessments of new compounds, and understanding their environmental impact.

S22 **Transfer kinetics of cargo items among mobile nanocarriers**

Faruk Hossain,[1] Guilherme Volpe Bossa,[2] and Sylvio May[1]
[1]North Dakota State University, Department of Physics, Fargo, ND 58108-6050, USA
[2]Universidad Austral de Chile, Institute of Mathematical and Physical Sciences, Valdivia, Los Rios, Chile

Micelles, liposomes, microgels, dendrimers, and nanoparticles represent nanocarriers used to deliver cargo items–often drug molecules–to a target. We calculate the kinetics of collision-mediated transfer of cargo items within ensembles of chemically distinct mobile nanocarriers in the Gaussian regime. To this end, the relevant rate equations for collision-mediated transfer of cargo items are expressed in the continuum limit as a set of Fokker-Planck equations and solved analytically. The solutions fully describe the time evolution of an arbitrary initial distribution of the cargo items among the nanocarriers toward equilibrium.

S23 **Optical properties of charged non-stoichiometric CdSe quantum dots: Computational insights**

Omolola Eniodunmo and Svetlana Kilina
Department of Chemistry and Biochemistry, North Dakota State University

Colloidal Semiconductor Quantum dots are promising structures and have attracted great attention for several applications because of their ease of reproducibility and highly size-tunable properties. Experimentally, non-stoichiometric QDs are easily achievable. In this work, we study charged non-stoichiometric CdSe QDs using density functional theory (DFT) and time-dependent DFT(TD-DFT). We explore two methods of introducing charge to Cd-rich and Se-rich non-stoichiometric QDs – via direct electron/hole injection and formation of a dangling bond by losing a surface ligand – and their effect on the ground and excited state properties. Our calculations show that ligand removal is similar to electron injection for Cd-rich and hole injection for Se-rich QDs. Both methods lead to the half-field mid-gap states strongly localized on a few surface cadmiums in Cd-rich systems and surface seleniums in Se-rich QDs. Coupling of a negative charge to the lowest-energy bright exciton in Cd-rich structures results on optically weak, red-shifted charge-related states. In contrast, coupling of a positive charge to the lowest-energy optically inactive exciton in Se-rich systems leads to strongly red-shifted optically inactive states. The result from this study provides insights that would be useful to facilitate easy tunability of their photophysical properties for applications ranging from photocatalysis, chemical sensing to bioimaging and quantum computing.

S24 **In search of improved, stabilized doped LLZO electrolytes for all solid-state Li-ion batteries through first-principles calculations**

Md Mozammal Raju,[1] Syed Rizvi,[1] and Qifeng Zhang,[1,2,*]
[1]Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND 58108, USA
[2]Materials and Nanotechnology, North Dakota State University, Fargo, ND 58105, USA

Different doping techniques such as single site doping, multiple site doping and co-doping (single site with multiple dopants) have been adopted to find suitable dopant to achieve solid-state electrolytes with higher ionic conductivity, lower activation energy and kinetically stable structure. With varying number or concentration of dopants, the ionic conductivity and phase stability of the solid-state electrolyte changes. Vienna Ab-initio Simulation Package (VASP) is a powerful tool that uses first principles method and density functional theory to run molecular dynamics (MD) of atoms in a crystal structure. Using MD simulation one can easily compute the ionic conductivity, activation energy and stability of a doped LLZO (Li7La3Zr2O96) crystal structure. In this work, we used four different cations Al, Ga, Rb and Ta to dope LLZO via single site, multiple sites and co doping, and calculated the ionic conductivity, activation energy of the structures. It was found, among Al, Ga, Rb, Ta, Al-Ga (co-doping), and Rb-Ta (multiple site doping) dopants - single site Ta doped LLZO, Li54La24Zr14Ta2O96 yields the highest ionic conductivity of ~2.14×10-2 S/cm with an activation energy of ~0.155 eV compared to undoped pure LLZO which shows ionic conductivity of ~2.38×10-3 S/cm with an activation energy of ~0.23 eV. Li44Al4La24Zr16O96 also yields a significantly higher ionic conductivity, ~1.44×10-2 S/cm. Li44Ga4La24Zr16O96, Li50Al2Ga2La24Zr16O96, and Li52La23Rb1Zr10Ta6O96 result in ~3.73×10-3 S/cm, ~4.2×10-3 S/cm, and ~3.2×10-3 S/cm respectively, which are just slightly higher than that of the undoped LLZO. However, Li58La23Rb1Zr16O96 gives a conductivity of ~2.36×10-3 S/cm which is lower than that of the undoped LLZO. The high conductivity of Li54La24Zr14Ta2O96 and Li44Al4La24Zr16O96 is attributed to the formation of a denser crystal structure, which leads to a high repulsive force between Ta-Li and Al-Li, keeping the cubic phase of the structure intact.

S25 **High-throughput phenotyping of seed yield and breeder's score in canola**

Mohammad Jony,[1] Md Fahad Hasan,[1] Paulo Flores,[2] and Mukhlesur Rahman[1]
[1]Department of Plant Sciences, NDSU
[2]Department of Agriculture and Biosystem Engineering, NDSU

Yield is a quantitative trait controlled by a small effect of many independent variables. Identifying new variables that correlate with yield holds significant importance in crop breeding programs. High-throughput phenotyping via image-based techniques can strengthen breeding program by reducing both the cost and time. In our ongoing efforts to integrate the high-throughput phenotyping, we are currently using DJI Matrice 200 equipped with a 10-band multispectral camera (MicaSense Dual Camera system) and small plot autonomous ROBOT for different agronomic traits. Multispectral images from the unmanned aerial system (UAS) were captured at 100 feet above ground level at 51, 66, 72, and 77 days after seeding (DAS) of the canola crop. Two experimental trials: a wide-area yield trial (WAT) featuring 108 research plots and an advanced yield trial (AYT) comprising 243 research plots were used in this study. The images were processed into orthomosaics using Pix4DMapper from Pix4D. Subsequently, an in-house Python script was employed to compute and extract single-plot statistics, specifically the mean values, for a total of 38 vegetation indices (VIs). We conducted Pearson's correlation analyses to explore the relationships between these VIs and four key agronomic traits: early vigor, breeder's score, seed yield, and seed oil content. Notably, breeder's score and seed yield exhibited significant correlations with several of these indices. Particularly, at 66 DAS in both experimental trials, indices such as NDVI, ENDVI, VEG, GRRI, NGRDI, MGRVI, VDVI, VARI, SAVI, and OSAVI demonstrated notably high correlations ranging from 0.60 to 0.74 with breeder's score and seed yield. This finding suggests that breeder's scores and multispectral images captured by UAVs could serve as reliable proxies for predicting seed yield. Additionally, we are in progress to employ a small-plot robot in both greenhouse and field settings to measure stem diameter and stand count.

S26 **Time-resolved properties of vibrational polariton model**

Patricia Adeoye and Dmitri Kilin
Department of Chemistry and Biochemistry, North Dakota State University, Fargo ND 58102

Vibrational polariton can occur due to the coupling of an IR-active molecular transition with a resonant mode of an optical cavity. The formation of polaritons creates opportunities for new chemical reactivities by trying to manipulate the basic characteristic of photons and provide a new pathway to enable chemical transformation. This system's 2D Hamiltonian is composed of the Hamiltonian of the molecule (phonon), the Hamiltonian of the field (photon) and the Hamiltonian of the field-to-molecule (photon-to-photon) coupling and it is explored by time-independent and dynamics approaches. In a time-independent approach, one explores the eigenstates of this Hamiltonian. In a dynamic approach, one explores time-resolved propagation of the initial coherent state via the evolution operator technique. The time-dependent wavefunction is used to monitor trajectories of expectation values of several important observables. The photon-to-phonon coupling results in three major computed effects as follows: (i) It facilitates the formation of polaritons and Rabi splitting between energies of the first eigenstate and second eigenstate. The Rabi splitting has been observed to be directly proportional to the coupling strength in the small coupling regime. (ii) It facilitates change in the equilibrium bond length in the molecule and equilibrium strength of the electric field. (iii) It is responsible for the time-dependent transition of energy from molecular oscillation to field oscillation and back. The observed effects have been rationalized in terms of the coupled quantized harmonic oscillators model,1 to quantify how the coupling strength affects the expectation value and eigenstate of the hybrid photon-to-phonon polaritonic state and also to explore the expectation values of the amplitude of the electric field and the interatomic bond displacement in the molecule. The finding of this exploration has a triple potential impact on basic science and technology: (a) Energy conversion from molecular bond oscillations to electromagnetic field amplitude serves as the foundation for the design of chemical lasers. (b) Energy conversion from cavity field to molecular motion contributes to the novel field of polaritonic chemistry for the

control of chemical reactions. (c) A focus on a phase of pure quantum states is tangential to quantum information processing.

"The parameters of the model are chosen to represent ground state potential energy surface of diatomic hydrogen iodine molecule in a cavity resonantly tuned to the vibrational frequency of this molecule"

Acknowledgement: This research is supported by NSF CHE-1944921

(1) Cohen-Tannoudji, C.; Diu, B.; Laloë, F. Quantum Mechanics; Wiley, 1977.

S27 **_Ab initio_ modeling of superconducting materials by exploring excited electronic configurations**

William Tupa[1,2] and Dmitri Kilin[3]
[1]Department of Physics, North Dakota State University, Fargo, North Dakota, 58108
[2]Department of Mathematics, North Dakota State University, Fargo, North Dakota, 58108
[3]Department of Chemistry and Biochemistry, North Dakota State University, Fargo, North Dakota 58108

The high-pressure hydrogen sulfide keeps its superconductive properties when at room temperature. This makes it a great candidate for industrial applications such as railway systems and power plants. We explore the electronic configuration features of this material relevant to superconductivity. Modification of electronic configurations includes changing the population of electronic states with certain values of spin projection, momentum, and energy. We specifically focus on configurations of electrons with opposite spins and opposite values of momentum interacting via lattice, referred to as Cooper pairs. The goal of this ab initio research is to show that these pairs are available to form in the sulfur trihydride material and can serve as the reason for superconductivity. The computed total energy of each electronic configuration is compared to the total energy of the reference neutral singlet configuration.

S28 **FHGAON–A user friendly tool for fungal whole genome assembly using MinION sequencing**

Jatinder Singh, Rachel Konshok, Santosh Gudi, and Upinder Gill
North Dakota State University, Fargo, ND, United States

Fungi play an important role in global agriculture due to their engagements in symbiotic and parasitic relationships with crop plants. Recent advancements in third-generation sequencing (TGS) technologies and easy accessibility of portable sequencers such as MinION, have revolutionized the fungal genomics. This genomics data holds immense potential to decipher the genomic landscapes and genetic diversity of fungal species. However, the analysis of this wealth of genomic information is bottlenecked due to the lack of user-friendly and streamlined computational tools and analysis pipelines. To address these challenges, we have developed an easy-to-use genome assembly pipeline, FHGAON (Fungal High-quality Genome Assembly using Oxford Nanopore), to assemble high-quality chromosome-scale fungal genomes from raw sequencing data. This pipeline supports multiple platforms including Windows and Linux and requires little familiarity with command line operations. It employs a docker container to install all the required softwares and their dependencies. We tested the reliability and simplicity of this tool for two fungal species, Trichoderma longibrachiatum and Fusarium oxysporum f. sp. spinaciae. FHGAON provides near-complete genome assembly of T. longibrachiatum (~31 Mb) in about 2 hours and for F. oxysporum f. sp. spinaciae (~62 Mb) in about 3 hours on Windows and Linux. This pipeline requires 10 CPU cores and 64 GB RAM. Overall, FHGAON helps in generating faster and more accurate chromosomal scale assemblies of fungal genomes with limited computational knowledge and resources.

S29 **Development of numerical features/descriptors to describe complex materials for machine learning modeling**

Rahil Ashtari Mahini,[1*] Gerardo Casanola Martin,[2] Simone Ludwig,[1] and Bakhtiyor Rasulev[2]
[1]Computer Science Department, North Dakota State University, Fargo, ND, United States
[2]Coatings and Polymeric Materials Department, North Dakota State University, Fargo, ND, United States

Multi-component materials/compounds and polymeric/composite systems pose structural complexity that challenges the conventional methods of molecular representation in cheminformatics, which have limited applicability in such cases. Therefore, we have introduced innovative structural representation techniques tailored for complex materials. We developed different formulations based on linear relationships and the possible interactions between different components in reaction that treat each multi-component material as a mixture system. We developed and improved mixture descriptors that are based on 13 different formulas and are divided to four categories including property-based descriptors, concentration-weighted descriptors, deviation-combination descriptors and combinatorial descriptors. The combinatorial descriptors rely on quantities of components ($1 \le i \le N$) and individual descriptors ($1 \le j \le M$) leading to an exponentially large value of M^N for cartesian product. To calculate that, Dask distributed computing framework has been used to distribute workloads on a PBS cluster within a high-performance computing environment. The objective of this study is to utilize these newly developed mixture descriptors for predicting a range of chemical and physical properties across various complex systems using mixture-based Quantitative Structure–Activity Relationship (QSAR) models. This would enable the prediction of the optimal combination of source components to create complex materials with desired properties or uncover the mechanisms behind the biological activity and toxicity of the existing one.

S30 **Development and application of KASP marker for screening sclerotinia stem rot resistance in canola (Brassica napus L.)**

Md Zahangir Alam,[1] Mohammad Jony,[1] Luis E. del Rio Mendoza,[2] and Mukhlesur Rahman[1]
[1]Department of Plant Sciences, North Dakota State University, Fargo 58102
[2]Department of Plant Pathology, North Dakota State University, Fargo 58102

Brassica napus with a trace amount of erucic acids and glucosinolates, as well as rich amounts of omega-3-fatty acids, is known as canola, the second most important oil crop in the world next to soybean. Canola oil has the potential to promote brain development in infants, prevent heart disease in adults and protect men and women from colon & breast cancer. Sclerotinia sclerotiorum is a notorious soil-borne and necrotrophic fungus that infect almost all growth stages and causes Sclerotinia Stem Rot (SSR) disease in canola, resulting in up to 80% global yield loss. Applying chemical fungicides is not always effective, economical, and eco-friendly. Therefore, it is important to identify the SSR resistant canola germplasms. However, morphological screening for SSR resistance is time-consuming, labor-intensive, and costly, which necessitates the use of molecular screening. KASP (Kompetitive Allele Specific PCR) marker has the potential to identify the SNP (Single Nucleotide Polymorphism) markers associated with disease resistance. Our objectives were to select stable and significant SNP markers associated with SSR resistance, then design KASP primers, and finally evaluate the efficacy of KASP markers for screening SSR resistance in canola. We selected 69 stable and significant SNPs, extracted 100 bases both upstream and downstream of the selected SNP position from the reference genome (ZS 11) of B. napus with the CCAST platform and designed two forward and one common reverse primer for each SNP marker. DNA was extracted from both resistant and susceptible B. napus accessions. PCR was carried out for the most important 21 sets of SNP markers, and data were recorded with the qPCR machine to read the fluorescence signal. Among them, sixteen sets of KASP primers amplified the target region of the genome, of which six sets of primers could screen out the resistant B. napus germplasms from the susceptible ones. Our primary data indicate the potential application of KASP markers for screening SSR disease resistance in canola.

S31 **Computational studies of droplet dynamics near a rough surface**

Julie Melbye and Yechun Wang
Mechanical Engineering, NDSU

In droplet-based microfluidics, particles, droplets, and microcapsules could be generated and employed in biomolecular synthesis, drug delivery, and disease diagnosis. One such real-world case is that of biological hemodynamics, here, the “droplet” represents a blood cell, and the “rough surface” represents the

endothelial cell monolayer coating the inner surface of a blood vessel or a microfluidic channel. By employing a 3D Spectral Boundary Element Method for interfacial dynamics, a high-order high-accuracy computational method, this work aims to numerically study viscous droplets interacting with a rough substrate by varying material properties and surface topographical characteristics. This study highlights the influence of fluid properties, such as viscosity and interfacial tension, substrate pattern, and operational parameters on the deformation, trajectory, and velocity of a viscous droplet as it passes over a rough surface. The results of this study can be used to further the understanding of the intricacies of the behavioral relationships between viscous droplets and conditions of blood vessels or microfluidic channels. The current computational work has the potential to provide fundamental knowledge for the development of microfluidics-based cell transportation device.

S32 **Multilayer network analysis of energy networks**

Muhammad Kazim and Harun Pirim
IME, North Dakota State University, Fargo ND, United States

In this pioneering research, we investigate the dynamics and interdependencies of energy networks that comprise multi layers. Network science offers unique tools to represent such real-world relationships. Our motivation stems from a desire to understand the same energy network from two distinct perspectives: a multi-layer perspective that assumes distinct nodes at each layer and a multiplex framework that holds same type of nodes at each layer. Within the multi-layer paradigm, we focus on the primary layers of Electricity and Gas. These layers are intricately connected through hubs, which act as inter-layer bridges, facilitating energy flow between them. Through this lens, our analysis is concentrated on understanding node prominence using centrality metrics and gauging the network's robustness via the Coverage Evolution metric.

Transitioning to the multiplex network analysis, we construct a network encompassing five unique layers, each representing a specific node shutdown scenario. This design choice is driven by a quest to ascertain the cascading effects of node failures across layers and to discern the interdependencies that bind these layers together. Layer similarities are compared, and community structures are revealed. Robustness due to random and targeted attacks is analyzed.

Our dual analysis approach is justified by the distinct insights each perspective offers. While the multi-layer analysis provides a foundational understanding of the network's architecture and key nodes, the multiplex analysis delves deeper into inter-layer relationships and the ramifications of disruptions across interconnected layers. This research stands as a seminal study in the domain, presenting an innovative approach to understanding energy networks. By dissecting the network both as a multi-layer and a multiplex structure, we illuminate the multifaceted dynamics of energy distribution, offering invaluable insights for future infrastructure planning and energy management endeavors.

S33 **Nanomechanics of actin filament deformation and fracture: a steered molecular dynamics approach**

Sharad Jaswandkar, Kalpana Katti, and Dinesh Katti
Civil, Construction, and Environmental Engineering, North Dakota State University

In eukaryotic cells, the cytoskeleton is a three-dimensional dynamic structure that helps cells retain their form, internal organization, and mechanical rigidity. The polymer biomolecules in eukaryotic cells that function as structural elements of the cytoskeleton are the actin microfilaments, microtubules, and intermediate filaments. The diseases like cancer transform the cell's cytoskeletal structure. Altered cytoskeletal structures modify cancer cells' ability to contract or stretch by affecting their deformation mechanics. The abnormal cellular motility is a signature characteristic of metastatic cancer cells, promoting tumor cells' spread to both local and distant locations in the body. The dynamic reorganization of the actin cytoskeleton is a fundamental requirement of this process. We have also shown that the cancer cells' altered mechanical properties during disease progression are connected with the actin reorganization. In actin assembly dynamics, actin filaments are severed constitutively by an essential regulatory protein, ADF/Cofilin.

Therefore, Several studies were performed to examine the mechanical properties of F-actins. However, the fundamental mechanism that controls F-actin's response to deformation is not understood.

The present study uses the steered molecular dynamics (SMD) simulation approach to understand actin filament mechanics. The work was performed in four computational sets of experiments, namely bending, compression, tension, and torsion simulations. Our findings demonstrate that F-actin's deformation response is regulated by the dissociation pattern of conformational locks at intra-strand and inter-strand G-actin interfaces. F-actin elongation enabled salt bridge formation at the inter-strand interfaces, improving the G-actin-G-actin bond strength. Furthermore, we noticed an inter-strand serrated locking pattern connecting G-actin subunits, restricting their relative movement, thus enabling F-actin's ability to resist deformation. Additionally, we found ADF/Cofilin to cause structural transmutations in f-actin, thus altering their physical properties. The F-actin mechanics described here are vital for constructing a mechanobiological eukaryotic cell model that mimics cell mechanics with disease progression.

S34 **Unraveling integrins' conformational changes and mechanics at the molecular scale**

Hanmant Gaikwad,[1] Sharad V. Jaswandkar,[1] Kalpana S. Katti,[1] Amanda Haage,[2] and Dinesh Katti[1]
[1]Department of Civil, Construction and Environmental Engineering, North Dakota State University,
[2]Department of Biomedical Sciences, University of North Dakota

Integrin as a mechanotransducer establishes the mechanical reciprocity between extracellular matrix (ECM) and cells at integrin-based adhesion sites. This study uses steered molecular dynamics (SMD) simulations to investigate the mechanical responses of integrin αvβ3 with and without ligand binding for tensile, bending, and torsional loading conditions. The ligand-binding integrin confirmed the integrin activation during equilibration by opening the hinge between βA and the hybrid domain. This activation of liganded αvβ3 integrin influenced the stiffness of the molecule observed during tensile loading. Furthermore, we observed that the interface interaction between β-tail, hybrid, and epidermal growth factor domains altered integrin dynamics. The bending deformation of extended integrin models reveals the different force responses corresponding to the application of force in bending and unbending directions of integrin. Along with available experimental data, the SMD simulation results were used to predict the mechanical properties of integrin underlying the mechanism of integrin-based adhesion. These evaluation of integrin mechanics provides new insights into understanding the mechanotransmission (force transmission) between cells and ECM.

S35 **Coarse-grained modeling of polymer–clay nanocomposites**

Wenjian Nie,[1] Sarah Ghazanfari,[1] Kalpana Katti,[1] Dinesh Katti,[1] and Wenjie Xia[1,2,*]
[1]Department of Civil, Construction and Environmental Engineering, North Dakota State University, Fargo, North Dakota 58108, USA
[2]Department of Aerospace Engineering, Iowa State University, Ames, Iowa 50011, USA

Polymer-clay nanocomposites (PCNs) are commonly applied as multi-functional structural materials with superior thermomechanical and dynamic properties while maintaining the characteristics of lightweight and optical clarity. These nanocomposites are obtained by adding a specific amount of clay nanofillers to a polymer matrix. In our work, building upon previously developed coarse-grained (CG) models of nanoclay and poly (methyl methacrylate) (PMMA), we employ molecular dynamics simulations to systematically investigate thermomechanical properties of PCNs with different microstructural features of nanoclays, i.e., exfoliated and stacked configurations, as well as varying their weight percentage and nanoclay size. Specifically, we perform tensile and shear simulations and detailed analyses of molecular properties, i.e., energy contribution and molecular stiffness to achieve an improved understanding of the influences of microstructural features in their physical properties at a fundamental level. We find that nanoclay weight percentage of nanocomposites could have significant influences on Shear Modulus and Young’s Modulus, and those influences have their own characteristics due to different configurations and different force directions. Global dynamics of PMMA slow down due to the interaction from nanoclay and the magnitude of the effect is larger for a bigger

nanoclay percentage. Moreover, the constraining effect on segmental dynamics of PMMA in the interfacial region decays from the region of nanoclay layers surfaces to a neat PMMA matrix. Our findings provide molecular insights into the arrangement of the constitutive components of PCNs under deformation and highlight the importance of nanoclay and interface in the mechanical and dynamic properties of these multi-functional materials.

S36 **Computational insights: Electronic structure and optical properties of cdse quantum dots as a photocatalyst for lignin depolymerization reaction**

Shabin Gurung and Svetlana Kilina
Department of Chemistry and Biochemistry, North Dakota State University, Fargo ND 58108

Lignin is a natural polymer found in vascular plants and consists of phenolic compounds with beta-O-4 bond. Phenol can be used as an alternative biofuel source, as it produces less greenhouse gases (GHGs). CdSe QDs can serve as a photocatalyst which uses light to catalyze the bond cleavage of lignin to produce the phenolic compounds. It is known that the rate of the reaction strongly depends on the surface chemistry and stoichiometry of the QDs. To get atomistic insights into the structural and surface effects of QDs on their photocatalytic abilities, we computationally study non-stoichiometric QDs (NS-QD), which do not have an equal number of cations and anions. The Cd28Se17 are used to model the Cd-rich NS-QDs. Halide ions are used as surface ligands neutralizing the surface charge. Using DFT calculations, we determine the alignment of the electronic levels between the Cd-rich NS-QD and a molecule modeling lignin to determine thermodynamic conditions for the hole transfer governing the first step of the cleavage of the beta-O-4 bond. Our calculations show that the alignment of the highest occupied orbitals of the molecule with respect to the valence band of the NS-QD is sensitive to the solvent polarity, with a non-polar or slightly polar solvents favoring thermodynamic conditions for the hole transfer from the excited NS-QD to the molecule.

S37 **Relationships between electronic structures and structural properties of polymer systems**

Meade Erickson,[1] Yulun Han,[2] Gerado M. Casañola-Martin,[1] Bakhtiyor Rasulev,[1] and Dmitri Kilin[2]
[1]Department of Coatings and Polymeric Materials, North Dakota State University, Fargo, ND 58108, USA
[2]Department of Chemistry and Biochemistry, North Dakota State University, Fargo, ND 58108, USA

Photodegradation of polymers is a complex phenomenon that is influenced by both topological and electronic structure of polymers.1 Regarding the process of exciting an electron from ground to excited state for photodegradation to occur, it is known that various relationships exist between ground and excited states. The goal of this work is to establish quantitative relations between rate of photodegradation, defined as inverse duration of irradiation by light of certain wavelength and intensity and several factors describing polymer geometry and electronic structure, computed by density functional theory. Specific factors include oscillator strength and transition energy of excitations matching resonance condition with incident radiation have been investigated heavily and have shown influence.2 Another factor that is believed to effect photodegradation is the spatial localization between the two orbitals relative to each other and along the structure. To investigate the influence of spatial localization on rate of photodegradation, we employ a cooperative approach between time dependent excited state molecular dynamics (TDESMD) and cheminformatics techniques where topological information is used to help describe the influential path of electron movements.3

References: 1. Wang, F.; Li, F.; Xu, M.; Yu, H.; et al.: J. Mater. Chem. A, 2015, 3, 5908- 5916.
2. Seki S.; Yoshida Y.; Tagawa, S; Macromolecules 1999, 32, 1080-1086.
3. Erickson, M.; Han, Y.; Rasulev, B.; Kilin, D. J. Phys. Chem. Lett. 2022, 13, 4374–4380.