COVER SHEET FOR PROPOSAL TO THE NATIONAL SCIENCE FOUNDATION

PROGRAM ANNOUNCEMENT/SOLICITATION NO./CLOSING DATE/if not in response to a program announcement/solicitation enter NSF 15-1 NSF 14-565 06/05/15	FOR NSF USE ONLY NSF PROPOSAL NUMBER
FOR CONSIDERATION BY NSF ORGANIZATION UNIT(S) (Indicate the most specific unit known, i.e. program, division, etc.) HRD - CENTERS FOR RSCH EXCELL IN S&T	

DATE RECEIVED	NUMBER OF COPIES	DIVISION ASSIGNED	FUND CODE	DUNS# (Data Universal Numbering System)	FILE LOCATION
				071298814	

EMPLOYER IDENTIFICATION NUMBER (EIN) OR TAXPAYER IDENTIFICATION NUMBER (TIN)	SHOW PREVIOUS AWARD NO. IF THIS IS	IS THIS PROPOSAL BEING SUBMITTED TO ANOTHER FEDERAL AGENCY?
650177616	☐ A RENEWAL ☐ AN ACCOMPLISHMENT-BASED RENEWAL	YES ☐ NO ☒ IF YES, LIST ACRONYM(S)

NAME OF ORGANIZATION TO WHICH AWARD SHOULD BE MADE Florida International University	ADDRESS OF AWARDEE ORGANIZATION, INCLUDING 9 DIGIT ZIP CODE Florida International University 11200 SW 8TH ST Miami, FL. 331990001
AWARDEE ORGANIZATION CODE (IF KNOWN) 0096354000	
NAME OF PRIMARY PLACE OF PERF Florida International University	ADDRESS OF PRIMARY PLACE OF PERF, INCLUDING 9 DIGIT ZIP CODE Florida International University FL ,331990001 ,US.

IS AWARDEE ORGANIZATION (Check All That Apply) (See GPG II.C For Definitions)
☐ SMALL BUSINESS ☐ MINORITY BUSINESS
☐ FOR-PROFIT ORGANIZATION ☐ WOMAN-OWNED BUSINESS
☐ IF THIS IS A PRELIMINARY PROPOSAL THEN CHECK HERE

TITLE OF PROPOSED PROJECT **CREST: Center for Aquatic Chemistry and the Environment**

REQUESTED AMOUNT	PROPOSED DURATION (1-60 MONTHS)	REQUESTED STARTING DATE	SHOW RELATED PRELIMINARY PROPOSAL NO. IF APPLICABLE
$ 5,000,000	60 months	01/01/16	1511553

THIS PROPOSAL INCLUDES ANY OF THE ITEMS LISTED BELOW
- ☐ BEGINNING INVESTIGATOR (GPG I.G.2)
- ☐ DISCLOSURE OF LOBBYING ACTIVITIES (GPG II.C.1.e)
- ☐ PROPRIETARY & PRIVILEGED INFORMATION (GPG I.D, II.C.1.d)
- ☐ HISTORIC PLACES (GPG II.C.2.j)
- ☒ VERTEBRATE ANIMALS (GPG II.D.6) IACUC App. Date Planned
  PHS Animal Welfare Assurance Number A3096-01
- ☒ FUNDING MECHANISM Research - other than RAPID or EAGER
- ☒ HUMAN SUBJECTS (GPG II.D.7) Human Subjects Assurance Number 00000060
  Exemption Subsection ________ or IRB App. Date Pending
- ☐ INTERNATIONAL ACTIVITIES: COUNTRY/COUNTRIES INVOLVED (GPG II.C.2.j)
- ☒ COLLABORATIVE STATUS
  Not a collaborative proposal

PI/PD DEPARTMENT Southeast Environmental Research Center	PI/PD POSTAL ADDRESS 11200 SW 8th St. Modesto A Maidique OE-148 Miami, FL 33199 United States
PI/PD FAX NUMBER 305-348-4094	

NAMES (TYPED)	High Degree	Yr of Degree	Telephone Number	Email Address
PI/PD NAME Todd A Crowl	PhD	1989	305-348-1666	tcrowl@fiu.edu
CO-PI/PD Shu-Ching Chen	PhD	1998	305-348-3480	chens@cs.fiu.edu
CO-PI/PD Rudolf Jaffe	PhD	1985	305-348-2456	jaffer@fiu.edu
CO-PI/PD Laird H Kramer	PhD	1992	305-348-6073	Laird.Kramer@fiu.edu
CO-PI/PD Rene M Price	PhD	2001	305-348-3119	pricer@fiu.edu

PROJECT SUMMARY

Overview:

During the past three decades, incidents involving pesticides, industrial chemicals, oil, pharmaceuticals, nutrients and metals have attracted worldwide attention and greatly affected environmental conditions (e.g., the Gulf of Mexico Deep-water Horizon Oil spill). These events demonstrate a regional, national and international need for enhanced research on the effects of toxic substances in the environment. The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) at Florida International University (FIU) will transform the institution by integrating discrete campus-wide programs across 10 departments and 4 colleges in fields from environmental chemistry through computer intensive data analysis and visualization, in order to tackle one of the regions most complex challenges: environmental contamination. CACE will create innovative opportunities for students, especially encouraging those from underrepresented minorities (URM), to participate in authentic research and foster their development as future STEM professionals. FIU CACE will unify this talented pool of researchers into a cohesive Center that will enhance collaborations, partnerships and synergies. The Center will bridge academic programs that exist across campuses by integrating graduate and undergraduate students into all research subprojects, emphasizing evidence-based educational approaches, technology advances, and analytical chemistry infrastructure, while providing authentic research experiences and solutions. CACE will transform cutting-edge research into technological and science-based solutions for various forms of water contamination using a framework that includes detection/identification, transport and fate in complex ecosystems, and data analytics and visualization. CACE will develop a modeling platform that will enable policy makers and managers to make informed decisions. FIU's CACE will work in collaboration with governmental and private sector partners in S. Florida to develop practical solutions to problems related to water quality in a natural-agricultural-urban setting. This partnership includes the South Florida Management District, the National Park Service, The Miccosukee Tribe of Indians, the Environmental Protection Agency, Everglades National Park, Department of Interior, and others.

Intellectual Merit :

FIU CREST CACE will increase opportunities for graduate and undergraduate students, especially encouraging those from URMs, to conduct authentic research while advancing aquatic and environmental chemistry research and data analytics, methodologies, ecological risk assessments. CACE will generate significant new knowledge regarding contaminants and pollutants in aquatic environments, as well as produce innovative new methodologies for detecting and assessing contaminant quantities and impacts, including the use of molecular detection techniques. Using new data analytic approaches for visualization and synthesis of complex data, CACE will provide managers and policy makers, including governmental and private sector partners in S. Florida, real-time, accessible decision tools. The proposed program will advance current efforts on the biological effects, transport, transformation and distribution of contaminants in the environment into new collaborative research areas that investigate the sources and transport of contaminants and pollutants in aquatic systems. The research conducted by the Center will inform the economic, environmental, societal, policy, regulatory, and legal implications of water quality issues.

Broader Impacts :

CACE will build on the success of FIU?s evidence-based transformation of STEM instructional practices to provide enhanced support for students to pursue and complete STEM graduate degrees, both at FIU or elsewhere. Through an innovative program that spans the graduate school to high school spectrum, CACE will increase the success of students in graduate programs, especially supporting participation of underrepresented students in aquatic chemistry and environment (ACE) fields and future professions. CACE will develop technologies for improving water quality analysis and contaminant detection, as well as translate research findings into actionable information for decision-makers and stakeholders. By providing potential scenarios for understanding the risks, sources, transport and impacts of chemical contaminants that threaten aquatic ecosystems and human wellbeing, CACE can impact global water quality.

1. Introduction - CREST: Center for Aquatic Chemistry & the Environment (CACE)

The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) at Florida International University (FIU) will transform our institution by integrating independent campus-wide programs across 10 departments and 4 colleges in fields from chemistry through computer intensive data analysis and visualization, in order to tackle one of the region's most complex challenges: ***environmental contamination***. CACE will create innovative opportunities for FIU students, of whom 75% are historically underrepresented minorities (URM), to experience authentic and socially relevant environmental research and foster their development as future STEM professionals. The identity of human-derived environmental contaminants has changed from traditional pollutants such as nutrients, trace metals, DDT and PCBs to other biologically active compounds such as antibiotics and pharmaceuticals (e.g. endocrine disrupters), mercury, black carbon, and fossil fuels (oil). These stressors are now recognized as having significant effects on ecosystems and biota as well as on human wellbeing (Kolpin et al. 2002). The ~7 million inhabitants of South Florida abut the Greater Everglades, the area's main source of freshwater, as well as a large expanse of agriculture. The hydrologic connectivity between the natural, agricultural, and urban landscapes (Allan 2004) results in a highly complex network of contaminant sources that are transported throughout the landscape (**Figure 1**). In the context of this closely-knit food-energy-water nexus (**Figure 2**), it is critical to understand the biogeochemical processes that govern the ultimate fate of these compounds and their impacts on the ecosystem. Once introduced into ecosystems, contaminants undergo various transformations, which may affect their toxicity. New methods of detection are required so that "exposure" can be characterized at very low concentrations for effective, early intervention. ***CACE research will address the sources, transport, transformation and ecosystem responses to contaminants, pollutants and other natural stressors, under changing land-use and environmental conditions.*** We will combine enhanced stressor detection with *in situ* measurements of contaminant transport and ecosystem and organism responses to provide tools for risk assessment and decision-making. We will develop new data analytic tools to fully quantify and model the agricultural and urban sources of stressors as well as their ultimate transformations and fates. FIU's CREST CACE will convene a multidisciplinary team involving chemistry, biology, hydrology, statistics, ecotoxicology, public health, and computer sciences to put FIU research at the forefront of this emerging field of science and technology. ***The guiding principle behind CACE is to understand nature, "one molecule at a time" and to train the next generation of scientists who will carry this initiative forward!***

Figure 1. South Florida transition boundaries between **Natural**, **Agricultural** and **Urban** landscapes. Arrows show hydrologic connectivity and direction of contaminant, pollutant and other natural stressor fluxes across landscape boundaries.

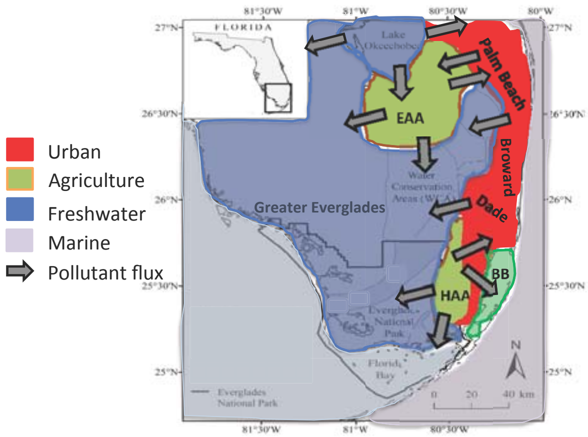
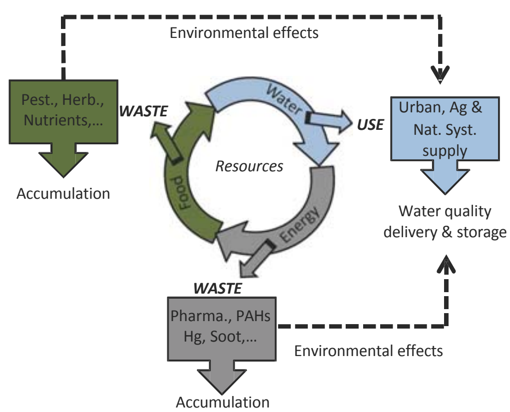


Figure 2. The production of contaminants, pollutants and other natural stressors in the context of the food-energy-water nexus.



2. Center Vision and Significance

The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) will not only transform our institution by creating this truly interdisciplinary cross-campus team, but also establish a sustainable infrastructure to enhance our research and education competitiveness in this evolving frontier area of STEM. CACE will unify FIU's talented, highly interdisciplinary pool of researchers into a cohesive Center that will enhance collaborations, partnerships and synergies. CACE will also transform cutting-edge research into technological and science-based solutions for various forms of water resource management in a framework that considers the social, policy, regulatory, and legal implications. The Center and its activities will integrate undergraduate and graduate students into all research subprojects, emphasizing technological advances, enhancing analytical chemical infrastructure, and providing authentic experiences and solutions. Through our partnerships with county, state, federal and Native American resource management agencies, students will be provided opportunities to tie university research with local management, as well as provide a pathway to permanent careers. Practical solutions to problems related to water quality in coastal and urban settings are critical components of the Center's mission.

CACE will embed innovative educational opportunities within the CACE science context for graduate, undergraduate, and high school students in order to increase access and participation for our highly diverse South Florida urban region. CACE's investments in faculty and student professional development will create a sustainable environment for continuous development of future STEM professionals. CACE also capitalizes on the experience and expertise of faculty carrying out educational transformation at FIU and in our local colleges and public schools, most notably through FIU's STEM Transformation Institute, as well as the unique demographics at our urban, majority-minority institution.

Specifically, the objectives of CACE are to:

1. Generate new knowledge for identifying and predicting future water quality issues, particularly eutrophication, industrial chemicals, pesticides, petroleum-based pollutants and pharmaceuticals in fragile coastal and urban areas in South Florida. We will focus on (a) advanced analytical methodologies for detecting low levels of chemical stressors; (b) stressor transport, storage, transformations and ecosystem responses; (c) ecotoxicological tools and non-traditional molecular tools for effects assessment; and (d) new prediction and decision-making tools to address water quality issues in South Florida wetland and coastal ecosystems.
2. Increase opportunities for education and careers in environmental and aquatic chemistry; ecosystem sciences; data mining, analysis and visualization; and related water sciences. Our program will deploy innovative, evidence-based approaches that integrate research with education activities to inform, encourage and facilitate student and faculty participation. Our discovery-based education environment will recruit, mentor, and graduate students from traditionally underrepresented groups. The Center will design professional development plans to help students identify career paths and develop skills to achieve them.
3. Build institutional capacity for interdisciplinary research and achieve sustainability beyond the NSF CREST. The Center has strong institutional support and will bring together independently operated environmental chemistry labs and facilities into a unified center with a core set of faculty and technicians to coordinate the multidisciplinary research and education team. ***CACE will thus provide a truly transformative, interdisciplinary approach to understanding risk and human wellbeing that will prepare students to solve real-world problems in a changing and uncertain future.***

Our objectives are perfectly aligned with the goals of the CREST Program. A key CACE objective is to attract, retain, and graduate students, especially underrepresented minorities (URMs), prepared for STEM careers. As a minority serving institution, FIU has a culture of cultivating and supporting student success, including advising, mentoring, tutoring, and multicultural programs. The Center will leverage these programs and build new institutional capacity for recruiting students by creating hands-on research projects for recruiting and engaging students, initiating long-term transects for data collection and modeling, and creating articulated and dual enrollment courses. The Center is committed to incorporating diversity at every level of organization, including project personnel, students, and collaborators.

3. Research Plan

3.1. Unifying Theme

Educational-experiential programs are designed to increase URM representation in STEM careers and are coupled to our research program organized around environmental chemistry, biogeochemistry, ecology and data synthesis and modeling as they pertain to regional water resources. Our proposed approach of *incorporating technologies* to measure and identify contaminants and pollutants with *field methodologies* to determine their ultimate transport and fate coupled with an enhanced ability to *mine, analyze, visualize and model* very large, complicated data sets (data analytics) is applicable to global environmental issues. Each of these research thrusts is individually challenging and warrants significant

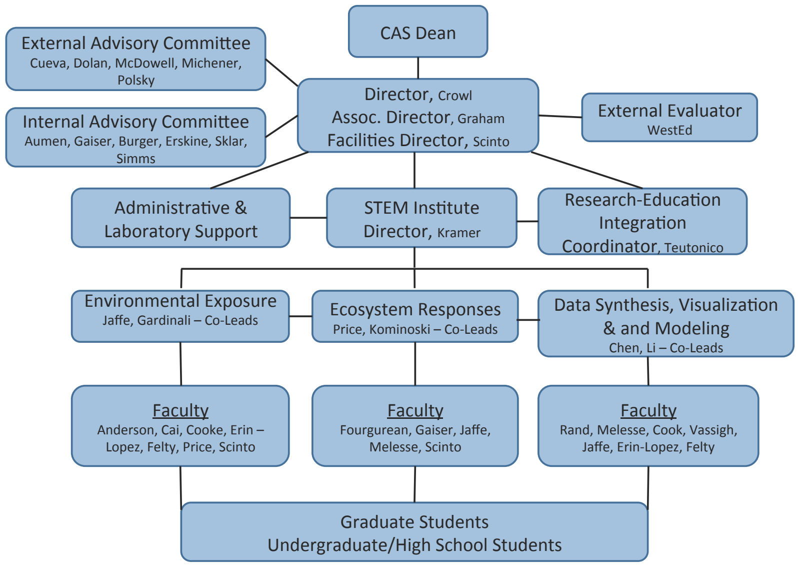


Figure 3. Center organization showing advisory committees, leadership and faculty involvement. Note that over 20 faculty from 4 colleges and 10 departments are fully integrated and invested in the CREST effort.

research. An organizational chart is provided in **Figure 3**.

3.2. Research Thrusts: Three Research Subprojects

The first subproject extends our current technologies and approaches for measuring anthropogenic contaminants, pollutants and other chemical stressors. The second uses these new sensing techniques to determine biogeochemical cycles including contaminant sources, storage, transport and transformations. The third develops data analytic methods to enable synthesis across large, complex data sets to allow holistic effects assessment for understanding South Florida's fragile aquatic ecosystem.

3.2.1. Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges

The goal of this subproject is to provide detailed characterization and measurement of the myriad stressors associated with our natural, urban and agricultural landscapes. The team will develop methodologies using advanced high resolution mass spectrometry coupled with online solid phase extraction (SPE) and liquid chromatography to extend our interpretation of environmental exposure from trace analysis of critical compounds such as **pharmaceuticals** (Wang & Gardinali 2013, Ramirez et al. 2014) to the detailed analysis of metabolites and degradates so that a more accurate account of the releases of emergent contaminants can be constructed (Wang & Gardinali 2014). The automation of analysis to produce high-throughput of specific wastewater intrusions using recalcitrant tracers has now connected the traditional water quality approach to a human derived signature enabling large-scale applications. We will perform unknown-unknown analysis using full scan, high resolution, tandem mass spectrometry so that not only suspect compounds such as drugs of abuse, but also their transformation products can be detected, opening the window for molecular epidemiological studies targeting consumption at population levels without intrusive monitoring (Heuett et al. 2014).

Metals, especially Hg, pose severe risks to wildlife in Florida Everglades (Wang and Gardinali 2014). Our goal is to identify the importance, mechanism, and diversity of Hg photoreactions (Yin et al. 2014) in aquatic environments and their relationship with other drivers of Hg transformations, such as dissolved organic matter (DOM), and to elucidate the pathways through which DOM (Maie et al. 2012) influences photochemical process in Hg biogeochemical cycling.

Carbon-based stressors are a direct result of the energy sector; Black carbon (BC) results from the partial combustion of organic matter during wildfires and fossil fuel burning and may affect pollutant transport and microbial loop dynamics. While ultrahigh resolution mass spectrometry (FT-ICR/MS) characterizations have resulted in a series of potential dissolved BC (DBC) structures, their identity has not been confirmed as isolation from the bulk DOM has not been achieved (Yamashita et al. 2010, 2013). Resolving the transformation steps of **energy derived emission products** to their molecular composition is a great challenge because of the complex nature of the samples (Marshall & Rodgers 2004). FT-ICR/MS has the potential to resolve components in fossil fuels and identify the elemental composition and double bond equivalents based solely on accurate mass measurements. Gas-phase post-ionization techniques provide additional separation dimensions (Fernandez-Lima et al. 2009). With the advent of new variants of IMS analyzers with higher resolving power and their coupling to FTICR-MS (Schenk et al. 2014) we will develop IMS-FT-ICR/MS methods to better characterize the transformation of fossil fuels emission products exposed to a variety of environmental conditions.

Organismal and molecular sensing: the roadmap from contamination to exposure to management. Ultimately, the biotic responses to stressors manifest as changes in survival, growth, development, and reproduction (SGDR) – which are also used in "classical" ecotoxicology testing. Given the complex nature of aquatic systems, a variety of molecular and biological endpoints are needed to understand the receptors, pathways and mechanism(s) of adverse outcomes and their environmental significance and predictive value. Molecular alterations are usually good indicators of exposure and effect for single chemical contaminants and mixtures, and are also typically more sensitive than those at higher levels of organization, and can represent surrogates for effects at the organ and whole organism level. Molecular biology tools, such as 'omics technology, can be applied to monitor the impact of contaminants in ecosystems (Bollati & Baccarelli 2010, Suarez-Ulloa et al. 2014). We propose to examine gene expression (Anderson et al. 2012) and epigenetic changes, together with whole genome (Karbaschi and Cooke 2014) and sequence-specific DNA damage (Hu et al. 2015) to sensitively biomonitor ecosystem and determine their relationship with effects in traditional ecotoxicological endpoints (i.e. SGDR).

Specific activities include: 1) Develop methods for analysis using full scan, high resolution, tandem mass spectrometry; 2) Develop IMS-FT-ICR/MS methods to better characterize at the molecular level the transformation of Carbon-based molecules including fossil fuels; 3) Examine gene expression and epigenetic changes, together with whole genome and sequence-specific damage to DNA.

3.2.2. Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems

Anthropogenic activities (land use, water management, population growth and consumptive activities) and natural forces (climate, hurricanes, storm surges, floods, droughts) influence the sources, mobilization, transport and transformations of pollutants across land-use boundaries. Biogeochemical processes govern the ultimate fate of these pollutants and their impacts on the environment (Dittmar et al. 2012). We propose through this CREST research to determine the hydrologic transport (flux), fate (biogeochemical processes), and environmental impact of pollutants across three major land-use boundaries in South Florida under current and potentially changing environmental conditions**.** We will establish three permanent transects that encompass the main transitions between agriculture, urban and natural landscapes. These transects will provide a common research platform for students and faculty to apply the sensor methodologies from Subproject 1, quantitatively sample water quality with a common experimental design, apply state of the art modeling (e.g., M3ENP) techniques to couple flows with contaminants and provide data that will be used to inform our data analytics research in Subproject 3. Finally, in using these data as our initial conditions, we will collaborate with our stakeholders and Subproject 3 to provide future scenarios models to predict future conditions.

Specific activities include: 1) Quantify the flux of water, nutrients, contaminants and pollutants along transects that cross major land-use boundaries (agriculture, urban, and natural; Figure 1); 2) Hydrodynamically model the transport of nutrients, contaminants and pollutants and their associated biogeochemical processes; 3) Predict (scenarios modeling) the potential transport of nutrients, contaminants and pollutants and their influences in any adverse biological outcomes with changing land-use and climate.

3.2.3. Subproject 3: Data Analytics for Effects Assessment and Decision-Making

The Ecotoxicology and Risk Assessment involve studies with non-chemical and organic and inorganic chemical stressors including **nutrients, contaminants and pollutants** with a multitude of exposure types (e.g., single-slug, intermittent and continuous) with native, exotic and standard test species. This entails collection of large volumes of data from various heterogeneous sources such as the data from analytical chemistry techniques and approaches and the data on biogeochemical cycles on how natural processes affect ecosystems. As the scale and complexity of these data types increase exponentially, it becomes challenging to effectively model the increasing volumes of data, discover useful information knowledge from the data, and provide data analytics capability to support effective and accurate assessment and decision-making capability for the scientists and their partners.

To address these challenges, the CREST center will provide a suite of data analytics algorithms, including computation modeling, data mining, and visualization. The computation-modeling component provides the computational and system support for diverse data analytics and decision-making tasks based on novel multi-tiered data analysis architecture. Data analysis tasks are computationally intensive and thus we will address system and architecture issues related to computational requirements for data gathering, data analysis, and decision-making (Liu et al. 2014, Ren & van der Schaar 2013, Xu et al. 2012, Xu and Shatz, 2003, Yu & Liu 2003).

Specific activities include:

1. Create novel multi-tiered data analysis architecture, consisting of sensors and cloud/HPC computing systems.
2. Utilize associations and correlations among the data to understand the characteristics and to extract semantics and patterns from the data.
3. Develop the dimension reduction and information fusion algorithms to address the scalability and multi-source issues.
4. Develop visual analytics and visualization algorithms to assist in assessment and strategic decision-making.
5. Develop computational methods for exploratory scenario development, evaluation and synthesis.

3.3. Integration across Research and Education

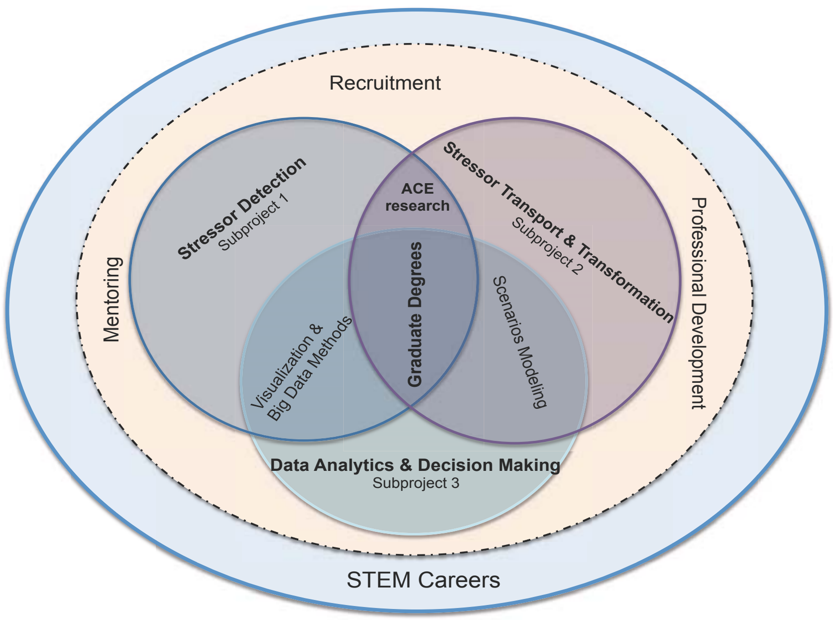


Figure 4. Integration across research areas and with education activities leads to the goal of graduate degrees and future STEM careers.

CACE will provide graduate students the opportunity to carry out their cutting edge research, as well as be immersed in the unique CACE environment to develop their multidisciplinary, collaborative skills, thus preparing them to solve future real-world problems for our ever-changing Nation. As shown in **Figure 4**, the CACE research areas and associated facilities are integrated with the main theme and central to the education plan and all of the planned activities.

Subprojects 1 and 2 provide the basic lab and field stressor methodologies and fine resolution response data needed to determine ecosystem risk and inform decision-making. Subproject 3 provides the data analytics necessary to interpret and fully utilize the suite of very large, complex data sets associated with environmental impacts. All aspects of our research, education and applications are designed to fully integrate across this research topic. The innovative educational opportunities are embedded within the CACE science context for graduate, undergraduate, and high school students.

4. Education Plan

4.1. Mission and Goals of the CACE Education Plan

The mission of CACE is to increase the number of students pursuing graduate STEM degrees, especially those from underrepresented groups (URMs), and, ultimately, the number of well-qualified professionals pursuing related careers. Specifically, CACE will promote graduate opportunities and training related to **aquatic chemistry and environment (ACE)** research and data analytics. To achieve this mission, the CACE Education plan will implement innovative experiential-centric and community-building mechanisms derived from evidence-based research on recruitment and retention of underrepresented students in

STEM. CACE will enhance the number of students completing graduate degrees by recruiting, mentoring and preparing graduate, undergraduate, and even high school students to enter STEM fields and succeed. Specifically, CACE students will engage in authentic research experiences in the laboratory and field to evaluate toxicity and risk of exposures to chemical and non-chemical stressors, ecosystem responses and the synthesis and visualization of this information for its application in decision-making using innovative data analytic tools. These experiences are investigated with regards to the South Florida environmental, as lack of relevance is often seen as a factor in URMs leaving the STEM disciplines.

The CACE education plan incorporates the following values and promising practices:

- Experience with authentic scientific practices
- Contextualization of the societal relevance of research projects
- Establishment of diverse learning communities
- Consistent and trained mentoring
- Formative evaluation that informs program development
- Professional development of students, postdocs, and faculty
- Integration with existing STEM programs on campus
- Development of models/modules that can be replicated STEM-wide across campus

4.2. Overview of Innovative Educational Objectives

The CACE education activities center on recruitment and development of the next and future generations of **aquatic chemistry and environment (ACE)** researchers and data analytics professionals, hereafter referred to as ACE careers. CACE will utilize a coherent program design that integrates across the entire student to graduate school to researcher spectrum. In parallel with developing future STEM professionals, CACE education activities will develop faculty to be better student mentors, especially for our highly diverse student population. Thus, experiencing authentic scientific practices, contextualizing CACE's research impact on society, creating diverse learning communities, and mentoring have emerged as core themes that run throughout all activities. The design leverages the wealth of education research in effective practices, as well as over a decade of successful education interventions and research at FIU. The evaluation plan provides continuous feedback on the overall project impact to guide implementation of our innovative model for STEM professional recruitment and development.

The core education activities provide high quality graduate research experiences in aquatic chemistry and the environment along with professional development activities that will prepare students to be future faculty and/or researchers. At least 30 graduate students will be supported by the Center over the first 5 years. We expect another 15 to 30 additional students that will be partially funded and/or participate in the center's research. These graduate students will develop research projects in one or more of the research thrusts while developing collaborative research skills across all three thrusts. To complement their research development, we will provide intentional professional development for them to become effective professionals, including interacting with multiple researchers, including academic, government, and industry partners. We expect at least 30 additional graduate students, not funded by CACE, to be engaged in professional development workshops and CACE colloquia each year. CACE student participants will engage in evidence-based educational practices, develop their mentoring skills, and develop skills that will lead to successful fellowship applications. The professional development is woven into the Center's activities, as is described in the next sections.

At its core, FIU's CREST CACE is about developing future ACE researchers, especially those from historically underrepresented groups. However, the need for well-qualified ACE professionals cannot be met through the graduate training component alone. Thus we propose several innovative research experiences to amplify the pathways into STEM research for students at FIU, in local community colleges, and in local high schools. ***In effect, we aim to 'prime the pump' and enhance the pool of potential students well qualified to enter STEM graduate degree programs*****.**

4.3. FIU: A Community that Supports Diversity

FIU is an institution dedicated to improving STEM education across the K20 spectrum and is rapidly becoming a living laboratory for developing future STEM professionals, especially from historically underrepresented groups. For over a decade, FIU has been using discipline-based education research (DBER) practices to transform undergraduate courses and programs, as well as partnering with local teachers and K12 districts to support teacher intensive professional development. These deep commitments to STEM education are driven by FIU's uniqueness as a large urban public research university that enrolled 54,099 students in Fall 2014, of which 63.1% were Hispanic, 14.0% were African American, and 56.2% were women. This diverse undergraduate population includes over 11,000 STEM majors and is majority minority. ***CACE will leverage FIU's powerful STEM foundation and diversity experience to become a top producer of future URM ACE researchers and professionals.***

4.4. Graduate Student Recruitment, Retention, and Mentoring
CACE includes explicit recruitment, mentoring, and professional development for our graduate students. CACE will institute a broad graduate student recruitment plan, implementing promising practices from successful minority bridge programs, such as the Fisk-Vanderbilt Masters-to-PhD Bridge program (Stassun et al. 2011). Our most effective strategies will leverage our broad environmental science networks to disseminate CACE opportunities. Our network of faculty partners has collaborators across the nation, which we will use to recruit potential applicants. We will use our discipline-specific networks to recruit students into the CACE program, such as the Ecological Society of America's SEEDS program that supports URM students in ecological fields. We also include an innovative undergraduate recruitment model to target FIU students, many of whom have strong ties to South Florida. Our plans for graduate student recruitment include: a) implementing the Discovery 1 and 2 courses (see 4.6 below) to provide students with ACE research experiences and career development; b) arranging informational meetings for prospective undergraduate students, especially in large introductory classes; c) collaborating with existing student science clubs; and d) coordinating with existing FIU diversity programs, such as the McNair Network and MBRS RISE, both of which are active on campus.

Our overall retention strategy involves consistent mentoring and professional development provided through various aspects of the Center. The CACE-funded applicants will be selected as Fellows by the CACE research faculty to ensure a good match with the overall objectives of the program. The transition to FIU and cohort community building activities will begin the week before the Fall semester starts. We will host a one-day orientation symposium, where the broad range of CACE research activities will be highlighted. It will also feature cohort development activities, introductions to the CACE team, and summary of the year's CACE activities. In addition, CACE faculty, in coordination with the Research/Education Coordinator, will implement a proactive retention plan to enhance the success of the project. The plan includes 1) an annual trainee survey to understand issues and expectations; and 2) an annual trainee progress review to facilitate the understanding of potential problems, continuous improvement of mentoring, and course corrections as needed.

CACE graduate students will conduct collaborative research projects with partners, which will provide real-world problems and a potential pipeline to hiring opportunities with State, Federal and Native American Water Management Agencies (see Collaborators section). This will help the graduate students engage with role models who offer their experience with graduate school, dissertation research, and successful careers after graduation as a source of inspiration and mentoring. We also plan a comprehensive and individualized mentoring plan for each trainee. Mentoring will start during the recruitment stage. We will also develop a CACE Colloquium that will convene the entire team of project participants quarterly to have graduate students present their research results for feedback and discussion. The Colloquium will also strengthen the student cohort and team cohesion. In addition, project collaborators will have opportunities to discuss their programs with the CACE team to enhance the development of future research projects.

4.5. Professional Development
Career development mentorship for both CACE-funded and other students will focus on helping them establish short and long term career objectives and skills needed to reach those objectives.

4.5.1. GRF workshop (Graduate, Undergraduate): We will develop a graduate student workshop focused on preparing an application to the NSF Graduate Research Fellowship program, based on a model of best practices developed at Carnegie Mellon University. The objective is for our students to develop competitive fellowship applications; thus the seminar will foster student understanding of the competition and then allow students to develop their application while receiving formative feedback throughout the summer. The Research-Education Coordinator will organize the seminar and convene faculty review boards to provide feedback to students. CACE Graduate Participants will enroll in the seminar in the summer after their first academic year to prepare applications for the fall deadlines. Successful applicants will be incorporated into the feedback mechanism for graduate students in later cohorts. The seminar will be open to at least 15 students each summer, including the CACE cohort of students plus additional students not funded through the project. The goal of the seminar will be to build a culture of successful fellowship applicants.

4.5.2. Career Skills Workshop (Graduate)
CACE will develop a graduate student workshop series that includes sessions on career options, resume/vita writing, interview skills, and strategies and job hunting. The core skills the workshop will convey include 1) *Discipline-specific knowledge:* concepts and principles, theories and analytic approaches in a particular domain; 2) *Research ability:* data management, literature review, critical thinking, presentation, proposal writing, technology disclosures, and patent applications; 3) *Communication skills:* writing, oral presentation, teaching, and inter-personal skills in different culture contexts; 4*) Project Management skills:* planning, status review, and collaboration with other trainee students, 5*) Leadership and Entrepreneurship skills:* developing vision and strategies for research and professional careers, as well as teamwork and decision-making skills; and 5) *Professional Ethics:* managing conflict of interest, responsible authorship, mitigating research misconduct and research with human subjects. In addition, the Center will compile all relevant post-doc opportunities as advertised (ECOLOG Society Bulletins, Science, Nature).

4.5.3. Experiential Teaching Development (Graduate, Faculty): CACE incorporates development of graduate student and faculty experiential teaching expertise through its courses and activities. The Discovery 1: CACE Research Course (below) will be integrated with CACE's research thrusts; thus it also provides professional development for the graduate students and faculty engaged in the course. CACE graduate students will be trained for, and serve as, TAs for the course. CACE faculty will identify appropriate research projects, kick off the course, and attend the final presentations. As the course operates in a discovery-based, guided-inquiry mode, the primary role of the graduate student is to provide feedback on student ideas and guidance for collecting, analyzing and interpreting data. This allows them to experience the benefits of discovery learning while developing their mentoring and teaching skills. For faculty, this offers a similar lens on student science discovery and mentoring.

4.5.4. Mentor Professional Development (Graduate, Faculty)
We will provide mentor professional development for faculty and graduate students using the Entering Mentoring (Handelsman et al. 2005) framework. The mentoring program will prepare faculty and graduates to better provide critical feedback for effective student engagement. The eight training sessions will be held over four meetings during the summer. An experienced, external trainer will run the initial training in Year 1, with Years 2-5 training led by Research-Education Coordinator.

4.6. Undergraduate Pathways
In order to augment the next generation of ACE graduate students, we will also focus on dramatically increasing the number of undergraduate minority students interested in pursuing ACE careers. Through CACE, we will develop an innovative experiential model where students carry out authentic research using either available data streams or their own field data, connect the research to its social relevance, and gain professional skills needed to pursue STEM careers. Our inspiration originates in successful approaches for engaging underrepresented students including *critical science* agency (Barton & Tan 2010, McNeill & Vaughn 2012) *wise schooling* (Steele 1997, Cohen et al. 1999, Taylor & Antony 2000) and *Third Space* (Gutierrez 2008). These ideas are instantiated in two programs currently providing profound teaching experiences for STEM students in order to recruit and prepare future teachers, FIU's Learning Assistants (LA) program and FIUteach, both of which have significantly increased the number of students interested in earning science and mathematics teaching degrees. We shall harness the knowledge and experience developed out of these programs to attract and prepare future environmental research professionals. This CACE model has the potential to be applied to other disciplines and thus transform the participation, recruitment and preparation of historically underrepresented students in many high-need STEM subfields.

4.6.1. Discovery 1: CACE Research Course. We will develop and offer a free one-credit semester-long course to recruit undergraduates (Juniors and Sophomores) into graduate research programs. Students will carry out authentic research using existing datasets and/or collecting CACE data to experience the thrills (and challenges) they would encounter in a research career. Students will be required to interpret how their research results impact their South Florida community and then present these results to the community. The goal is to allow students to develop their identity as scientists while experiencing the social relevance of their work first hand, thus motivating pursuit of environmental science research careers. Students will work in teams to use existing and field-collected data to develop community-relevant solutions. The CACE Research/Education Coordinator will lead the course. CACE faculty mentors will provide the research projects for students and team-teach the course with their research teams. Through this new course, students will not only learn inquiry-based approaches to environmental problems, but will also build a learning community to foster their long-term undergraduate success.

This innovative research experience course is modeled after the UTeach STEP 1 course (http://uteach-institute.org) which is a free course that provides science and mathematics experiences for STEM majors and can lead to secondary science and mathematics certification. In its first two semesters of operation, FIUteach has consistently filled its STEP 1 sections, with a high percentage continuing in the program. This CACE Research Course will be the first translation of the STEP 1 course to a research-specific professional recruitment model. We expect it to be utilized in other high-needs areas to recruit students, especially those from historically underrepresented groups. We are confident that we will be able to fill the course sections and are cautiously optimistic that we will measure a significant increase in the number of students entering graduate research programs at FIU and elsewhere.

CACE will provide in-state tuition-equivalent support for 50 students per semester (in 2 sections of 25 students each) to incentive participation and increase the number of students exposed to the authentic ACE research experience. Recruitment will be through direct email marketing, as well as science classroom visits. The CACE director, faculty, graduate students, and eventually prior CACE undergraduate participants, will visit introductory STEM courses for brief recruitment sessions and hand out CACE fliers. FIUteach, for example, hands out over 1,000 fliers to interested students every semester. Students will apply to be part of the program, allowing us to evaluate potential for graduate school, ensure they are not overloaded in courses and/or jobs, and monitor diversity. We expect over 100 applications each semester, based on previous LA/FIUteach recruiting experiences. The goal is to select sophomores and juniors with academic records showing potential for graduate school admission.

4.6.2. Discovery 2: CACE Professions (Undergrad): For those undergraduate students who wish to continue exploring research careers opportunities, we will develop and offer a second free course. In this one-credit, semester-long course, undergraduate students will develop their career and graduate school preparedness through hands-on professional development training. Firstly, students will prepare applications to relevant national REU site programs to gain additional research experience to make them more competitive for graduate programs. Discovery 2 students will also write graduate school applications and prepare an NSF GRF application, as outlined in Section 4.4.1. All writing projects include iterative formative feedback by CACE faculty and graduate students. In addition to graduate school preparedness, students will learn about ACE career pathways, meet alumni and early career professionals working in ACE fields, and develop their own professional career resumes. We will offer the Discovery 2 course to 50 undergraduates per year.

4.6.3. Community college and future transfer students: The Discovery 1 course has the potential to transform recruitment into graduate programs, especially for URM students; thus CACE includes a strategic investment to bring the Discovery 1 course to local colleges and high schools. This allows us to test the recruitment model as well as build collaboration with local students, faculty, and teachers.

Both Miami Dade College (MDC), a primarily two-year college with a few 4-year programs, and Florida Keys Community College (FKCC) have committed to partnering with us to prepare faculty to implement the course (see letters). By training the faculty, we achieve a multiplier effect that will allow us to engage students from a broader socioeconomic range and provide them with the opportunity to engage in authentic research and consider science careers. This will provide their students with a recruitment pathway both into the undergraduate ACE degree programs as well as graduate school.

We will prepare MDC and FKCC faculty to teach the Discovery 1: Research course in Years 2 and 4 of the project. The training will take place over one week in the summer, incorporating both the curriculum and explorations of the research projects. Faculty will be recruited by partner institute leaders (see letters) and earn stipends. The Florida College System has established articulation agreements with the state universities, simplifying course adoptions and transfer credit.

We will also prepare Miami-Dade County Public School teachers with science master's degrees to teach the Discovery 1: Research Course as a dual enrollment high school course in Years 2 and 4. In additional to earning college credit, the course will allow high schools to engage in CACE-based authentic research to increase science literacy and recruit future scientists. The training will be integrated into our partner college faculty preparation, but with separate breakouts to adapt to high school environment. The course will run over the entire school year, thus allowing sufficient time to complete their investigations. Teachers will be recruited by Mr. Cris Carranza (see letters) and earn stipends.

4.6.4. Amplifying Undergraduate Research Experiences

CACE will facilitate access to undergraduate research experiences across the nation for our students, by circulating opportunities as well as providing feedback on student applications. The aim is to increase students' interest as well as their competitiveness in the programs. Faculty, graduate students and the (director) will share the responsibility for providing student feedback. This facilitation is integrated into the Discovery 1 and 2 courses and also provided to other undergraduates in the project, including students at MDC and FKCC. Through these engagements, more of FIU, MDC, and FKCC's students will apply and participate in research experience across the nation, significantly increasing the diversity of those programs. Ultimately, this will support increased participation in graduate programs by our students.

5. Evaluation Plan:

Independent evaluation will be conducted by WestEd, the project's external evaluator. WestEd will conduct a formative and summative evaluation of CACE, based on the project's logic model (Figure 5). WestEd will use the formative aspect of the evaluation to provide feedback to CACE project staff after each data collection. The summative evaluation will be used to measure project impact. WestEd will report progress on the summative evaluation in annual reports. The evaluation will address the following questions and sub questions:

Question	Data Source
EQ1: To what extent does CACE advance interdisciplinary research in the ACE fields? • *EQ1a:* How many faculty members participate in CACE research (by discipline, gender, race/ethnicity)? • *EQ1b*: To what extent does the Center increase faculty productivity? • *EQ1c:* To what extent does the Center facilitate learning networks? How are these networks utilized for research and dissemination?	Interviews and publication data; Social Network and Bibliometric Analysis
EQ2: To what extent does the project successfully recruit URM students? • *EQ2a:* How many URM students participate in CACE research? What proportion of URM students complete their research experience? • E*Q2b*: What factors support completion?	Administrative data; document review; surveys;
EQ3: To what extent does the CACE research experience prepare undergraduate students for ACE careers and/or continuing education? • *EQ3a:* How do undergraduate students learn about the CACE research experience? • *EQ3b*: How many students participate in each of the CACE undergraduate activities? How does this vary by gender, race/ethnicity, major, and transfer status? • *EQ3c:* How do students perceive the value of their research experiences? • *EQ3d:* To what extent do students gain relevant knowledge, skills, and scientific identity through their research experience? • *EQ3e:* To what extent do students gain learning networks? How are these networks utilized during their research experiences and beyond?	Quasi-experimental design/ Online Surveys.

EQ4: To what extent does the CACE Fellows Program prepare graduate students for ACE professions and post-doctoral positions? • *EQ4a:* How are graduate students recruited? How many participate? How does this vary by gender, race/ethnicity, and major? • *EQ4b*: To what extent do graduate students value the career counseling? • *EQ4c*: What professional development training is offered? How do graduate students perceive the usefulness of the professional development for teaching inquiry-based courses? How useful do graduate students find teaching Discovery courses for their future goals? • *EQ4d:* To what extent do graduate students gain learning networks? How are these networks utilized during their research experiences and beyond?	Interviews/Focus Groups and Longitudinal Tracking
EQ5: To what extent does professional development successfully prepare faculty to teach inquiry-based courses? • *EQ5a*: What professional development training is offered and by whom? • *EQ5b:* How many faculty members participate in the professional development? How does this vary by institution and department, research field, gender, and race/ethnicity? How many MDCPS teachers participate in dual enrollment professional development? • *EQ5c:* How do faculty perceive the usefulness of each professional development activity offered by the Center?	Interviews and Focus Groups

As described below, WestEd will develop a mixed-methods approach to answering the evaluation questions and sub questions. The sub questions are generally provided to allow for the necessary data collection and analysis before answering the more impactful evaluation questions.

EQ1: To what extent does CACE advance interdisciplinary research in the ACE fields?
Method: Social Network and Bibliometric Analysis
Through the sub questions, WestEd will identify the network of faculty, students, and practitioners participating in each sub-project. At the kickoff meeting, we will collect from project leadership the contact information and role for each group of researchers on the subprojects. Every six months thereafter, we will interview network members to collect information about activities and interactions pertaining to the challenge, and identify new members of the network. Analysis will focus on the interdisciplinary nature of the networks, stability, and durability. WestEd will also track dissemination of published research by faculty members and students, as well as other dissemination to practitioners.

EQ2: To what extent does the project successfully recruit URM students?
Method: Longitudinal database tracking students who show interest in the program.
FIU and WestEd will develop a mechanism to collect information on students who show interest in CACE research experiences and track progress through FIU data and online surveys. The project will collect contact information (names, e-mail addresses) of those who express interest in CACE research and share those with WestEd. WestEd will contact those who express interest and ask them to complete a short survey about their career interests and educational experiences. WestEd will also collect the same data from program participants. WestEd will then create a sample of nonparticipants matched to the participants on academic measures on data supplied by FIU (SAT/ACT/GRE scores, grade point average, major). After selecting the comparison group, WestEd will invite them to participate in the evaluation and provide them with appropriate incentives to encourage their continued participation through the course of the project.

EQ3: To what extent does the CACE research experience prepare undergraduate students for ACE careers and/or continuing education?
Method: Quasi-experimental design/ Online Surveys.
WestEd will track student participation in each research experience through the longitudinal database and online surveys. At the beginning of each semester, a list of students enrolled in the D-1 courses will be matched to an appropriate comparison group of other science majors using a matching algorithm. If sample sizes are sufficient, students in the D-2 courses will also be matched to a comparison group of students who took the D-1 course, but not the D-2 course. Using this design, we will analyze outcomes such as GPA, retention, completion, employment, and/or graduate school enrollment for undergraduate

students participating in the research experiences. Surveys will then provide feedback on student experiences and perceived gains in knowledge, technical and non/technical skills, self-efficacy in science, and science identity.

EQ4: To what extent does the CACE Fellows Program prepare graduate students for ACE professions and post-doctoral positions?
Method: Interviews/Focus Groups and Longitudinal Tracking.
The sample size of graduate students participating in the Fellows Program will be insufficient to conduct a quasi-experimental design. Instead, WestEd will utilize the longitudinal database to track and analyze graduate student outcomes such as completion, employment, fellowships, and/or post-doctoral placements. Additionally, WestEd will conduct annual semi-structured interviews and focus groups with graduate students participating in the Fellows Program. Specific questions will ask about participants' opportunities to engage in professional development, mentoring, teaching, and research; their frequency of participation; their perceptions of the contributions of these components to participants' future goals; and their suggestions for improvement.

EQ5: To what extent does professional development successfully prepare faculty to teach inquiry-based courses?
Method: Interviews and Focus Groups.
WestEd will conduct annual semi-structured interviews and focus groups with faculty participating in professional development for the D-1, D-2, and dual enrollment courses. Specific questions will ask about components of the professional development offered, opportunities to engage in the professional development, their perceptions of the contributions of these components to teaching and research goals, and their suggestions for improvement.

6. Outcomes
The expected outcomes for the CACE Center are a combination of the outputs and outcomes columns described in detail on the project's logic model below (**Figure 5**).

Inputs	Activities	Outputs	Outcomes
Faculty and graduate student researchers Cyber-infrastructure, datasets, and models Long-term datasets, models, Multiple funding sources for faculty research in multiple disciplines Undergraduate student population with over 70% URMs Existing FL networks for recruitment Partners at local ACE scientists at government, industry, & academic institutions Partners at Community Colleges and Miami-Dade public schools Analytical facilities and lab equipment	Interdisciplinary research by faculty and student teams Recruitment of graduate students for CACE fellowships Develop and teach CACE undergraduate Discovery Courses Train graduate student as TAs for Discovery courses Mentoring training for faculty and graduate students Near-peer mentoring of undergraduates by grad students GRF workshop for graduate students and undergrads Professional development training for faculty and graduate students Faculty training for Discovery classes at community college and high school	Spatially explicit model(s) that incorporate environmental, chemical, bio-molecular data along landscape gradients Interdisciplinary publications in high-impact journals and presentations at frontier conferences Multidisciplinary proposals submitted by CACE research teams Increased number of ACE graduate students and PhD degrees Increased number of FIU students applying for NSF GRF fellowships Increased number of undergraduates pursuing graduate STEM degrees, especially URM students Increased number of community college and transfer students in ACE fields	World-class, cutting-edge technologies for contaminant detection, transport, fate, and impacts New knowledge generated from interdisciplinary research on environmental contaminants New models for effective STEM education, with specific support for URM students Institutional transformation through cross-college collaborations across ACE fields Well-qualified FIU students entering STEM graduate programs Greater pool of students prepared to enter into ACE and other STEM careers Engagement with community stakeholders about environmental solutions and decision-making tools

Figure 5. CACE Project Logic Model

7. Project Management
The structure of the project management consists of: 1) External Advisory Committee (Table 7.1); 2) Internal Advisory Committee (Table 7.2); 3) Internal Steering Committee (Table 7.3); and 4) Collaborations (Table 7.4).

8. Sustainability/Leveraging
The proposed Center will enhance institutional capacity by fostering collaboration on environmental

Table 7.1. External Advisory Committee (EAC): Our EAC consists of a diverse, highly qualified individuals that will guide us in establishing and sustaining NSF-CREST Program and will meet annually at FIU.
Dr. Elvira Cueva is the PI of the CREST - University of Puerto Rico and will provide first-hand experience associated with operating a CREST and will provide a direct recruitment pipe-line to Puerto Rican students.
Dr. Erin Dolan is the Executive Director of the Texas Institute for Discovery Education in Science at the University of Texas and has extensive experience with all levels of education.
Dr. William McDowell is Presidential Chair and Director of the New Hampshire Water Resources Center at the University of New Hampshire and will provide expertise in environmental chemistry and biogeochemistry.
Dr. William Michener is the Director of the e-Science Program, the EPSCoR Program and PI and Director of the NSF DataONE Project and will provide expertise in our Data Analytics research as well as extensive experience in large-project management.
Dr. Colin Polsky is the Director of the Center for Environmental Studies, Florida Atlantic University and has extensive experience with urban ecology and socioeconomic systems.

Table 7.2. Internal Advisory Committee (IAC): Our IAC includes University Administrators, other university Program Directors and Scientists from our collaborating partner Institutions. CV's are attached.
Dr. Nick Aumen is Regional Science Advisor for the US Geological Survey will Chair and has extensive research experience in modeling water resources and climate change.
Dr. Evelyn Gaiser is Executive Director of the School of the Environment, Arts and Society.
Dr. Kevin Burger is Deputy Director for the Department of Interior's Office of Everglades Restoration Initiatives and for the South Florida Ecosystem Restoration Task force.
James Erskine is Chief Environmental Officer for the Miccosukee Tribe of Indians and a member of many Everglades Restoration committees and working groups.
Dr. Fred Sklar is Director of the Everglades Ecosystem Assessment Section, South Florida Water Management District and a member of many Everglades restoration task forces and committees and has extensive experience in wetland restoration and landscape modeling.
Dr. George Simms is Assistant Vice President for Student Access and Success and Director of the FIU McNair Program.

chemistry and ecotoxicology, increasing the impact of individual research. We will build institutional capacity for recruiting and retaining URM students by constructing a demonstration project and hands-on activities focused on water issues, creating a water-focused competition, strengthening URM student organizations, and supporting mentoring, tutoring, and other programs for URM students. In concert with the initiation of CACE, this research group will be formalized as a new Center of Environmental Chemistry and the Environment including new common space for analytical chemistry and ecotoxicology. We are currently working with both federal and private sectors to provide additional funding. The center mission will be partially supported by existing centers such as our Environmental Research Center (SERC) that will provides staff and operations support.

9. Broader Impacts of the Proposed Work: CACE will build on the success of FIU's evidence-based transformation of STEM instructional practices to provide enhanced support for underrepresented students to successfully pursue STEM careers. Through an innovative program that spans high school through graduate school, CACE will increase the participation of underrepresented students in aquatic chemistry and environmental research and data analytics professions. Knowledge gained by the graduate and undergraduate students who conduct research in this Center will be key to enabling responsible development of water resources, as well as protecting human health and the environment.

Sustainability of water systems and the environment are wide reaching global and societal problems; thus, the very core of the proposed program addresses broad societal needs. CACE will develop technologies for improving water quality detection, as well as providing alternative management scenarios. The Center will develop tools for decision-making to be made based on an improved understanding of the costs/benefits of various possible actions. Results of this research will be broadly disseminated in venues and formats typically accessed by policy makers and water managers, such as regional magazines, newsletters, and executive summaries that can be distributed to lawmakers. With the decision-making processes developed in this research, policy-makers will be able to explore the implications and consequences of various water management goals and develop more sustainable water management strategies. Results of this research will also be reported through peer-reviewed publications and professional presentations by both faculty and students. A website will be created for the Center and will contain summaries of results, items of interest, and images. Modern society has unsustainable rates of resource use, particularly with respect water to resources. By investigating water quality, water use and contamination and coupling these water issues to decision-making processes, the project has the potential to improve water use management for future generations.

Table 7.3. **Internal Steering Committee (ISC):**
Todd Crowl will serve as the **Lead PI** and oversee the CREST Center. He founded and directed the NSF $20,000,000 iUTAH project on water sustainability prior to coming to FIU and has trained over 70 undergraduate and graduate students.
Dr. Scott Graham will serve as interim Associate Director and lead the search for a permanent director. Graham co-directs the NSF I/UCRC Center for Advanced Knowledge Enablement and was Program Coordinator for FIU's CREST Center for Innovative Information Systems Engineering.
Dr. Laird Kramer is Director of the STEM Transformation Institute will be a **Co-PI** and will direct and develop the education and outreach materials in concert with the STEM Transformation Unit and oversee the high school-undergraduate student and teacher component of the program.
Dr. Rita Teutonico is Associate Executive Director, School of Environment, Arts and Society will be the **Research-Education Integration Coordinator**.
Dr. Rudolf Jaffe will be a **Co-PI** and will lead Subproject 1 - Advanced Analytical Methodologies and supervise undergraduate and graduate student research.
Dr. Rene Price will be a **Co-PI** and will lead Subproject 2, supervise graduate students and participate in the environmental gradient synthesis activities.
Dr. Shu-Ching Chen will be a **Co-PI** and will lead Subproject 3, supervise graduate students and participate in the data analytics activities.

Table 7.4. Collaborations: In addition to providing research and resource management expertise, our collaborating partners will provide **access to data, potential students** and **internships** at state (*South Florida Water Management District*), Federal (*US Geological Survey, US Environmental Protection Agency, the National Park Service* and *the Department of the Interior*), Native American (*Miccosukee Tribe of Indians*) organizations, and Educational Institutions (Miami Dade College, Florida Keys Community College, Miami-Dade County Public Schools). We are in a unique (and enviable) position in that these collaborators cannot accept NSF funding and so will provide in-kind support and opportunities. Individuals representing the collaborating institutions will rotate onto our Internal Advisory Committee. CV's for Institutional representatives are attached.
Dr. Nick Aumen is Regional Science Advisor for the **US Geological Survey** will Chair and has extensive research experience in modeling water resources and climate change.
Dr. Kevin Burger is Deputy Director for the Department of Interior's Office of Everglades Restoration Initiatives and for the South Florida Ecosystem Restoration Task force.
Cristian Carranza is STEM Director for the Miami-Dade County Public Schools and will collaborate with Dr. Kramer developing and delivering research and teaching tools developed through CREST.
Dr. Joffre Castro is Water Quality Manger for the **Everglades National Park** and will provide environmental chemistry and environmental engineering expertise.
James Erskine is Chief Environmental Officer for the **Miccosukee Tribe of Indians** and will provide a recruiting avenue for Native American Students.
Dr. Peter Kalla is a Senior Scientist in **US EPA Region 4** Laboratory and has extensive experience in mercury research.
Dr. Michael Lewis is a Senior Scientist and Branch Chief at the US EPA and has extensive experience in ecosystems ecology and management.
Dr. David Rudnick is Science Coordinator for the Everglades National Park and has extensive experience in wetland ecology and management.
Dr. Fred Sklar is Director of the Everglades Ecosystem Assessment Section, South Florida Water Management District and a member of many Everglades restoration task forces and committees. He has extensive experience in wetland restoration and landscape modeling.
Dr. Heather Belmont is Dean of Sciences, Miami Dade College She will serve as the point of contact for the college, working with Dr. Kramer to recruit faculty to participate in the Discovery 1 course and facilitating its implementation.
Dr. Michael McPherson is Dean of Arts & Sciences, Florida Keys Community College He will serve as the point of contact for the college, working with Dr. Kramer to recruit faculty to participate in the Discovery 1 course and facilitating its implementation.

CREST: Center for Aquatic Chemistry & the Environment (CACE)

Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges

Project Summary

During the past three decades, incidents involving pesticides, industrial chemicals, oil, pharmaceuticals, nutrients and metals have attracted worldwide attention and greatly affected environmental conditions (e.g., the Gulf of Mexico Deep-water Horizon Oil spill). These events demonstrate a regional, national and international need for enhanced research on the effects of toxic substances in the environment. The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) at Florida International University (FIU) will transform the institution by integrating discrete campus-wide programs across 10 departments and 4 colleges in fields from environmental chemistry through computer intensive data analysis and visualization, in order to tackle one of the regions most complex challenges: **environmental contamination**. CACE will create innovative opportunities for students, especially encouraging those from underrepresented minorities (URM), to participate in authentic research and foster their development as future STEM professionals. FIU CACE will unify this talented pool of researchers into a cohesive Center that will enhance collaborations, partnerships and synergies. The Center will bridge academic programs that exist across campuses by integrating graduate and undergraduate students into all research subprojects, emphasizing evidence-based educational approaches, technology advances, and analytical chemistry infrastructure, while providing authentic research experiences and solutions. CACE will transform cutting-edge research into technological and science-based solutions for various forms of water contamination using a framework that includes detection/identification, transport and fate in complex ecosystems, and data analytics and visualization. CACE will develop a modeling platform that will enable policy makers and managers to make informed decisions. FIU's CACE will work in collaboration with governmental and private sector partners in S. Florida to develop practical solutions to problems related to water quality in a natural-agricultural-urban setting. This partnership includes the South Florida Management District, the National Park Service, The Miccosukee Tribe of Indians, the Environmental Protection Agency, Everglades National Park, Department of Interior, and others.

Intellectual Merit

FIU CREST CACE will increase opportunities for graduate and undergraduate students, especially encouraging those from URMs, to conduct authentic research while advancing aquatic and environmental chemistry research and data analytics, methodologies, ecological risk assessments. CACE will generate significant new knowledge regarding contaminants and pollutants in aquatic environments, as well as produce innovative new methodologies for detecting and assessing contaminant quantities and impacts, including the use of molecular detection techniques. Using new data analytic approaches for visualization and synthesis of complex data, CACE will provide managers and policy makers, including governmental and private sector partners in S. Florida, real-time, accessible decision tools. The proposed program will advance current efforts on the biological effects, transport, transformation and distribution of contaminants in the environment into new collaborative research areas that investigate the sources and transport of contaminants and pollutants in aquatic systems. The research conducted by the Center will inform the economic, environmental, societal, policy, regulatory, and legal implications of water quality issues.

Broader Impacts

CACE will build on the success of FIU's evidence-based transformation of STEM instructional practices to provide enhanced support for students to pursue and complete STEM graduate degrees, both at FIU or elsewhere. Through an innovative program that spans the graduate school to high school spectrum, CACE will increase the success of students in graduate programs, especially supporting participation of underrepresented students in aquatic chemistry and environment (ACE) fields and future professions. CACE will develop technologies for improving water quality analysis and contaminant detection, as well as translate research findings into actionable information for decision-makers and stakeholders. By providing potential scenarios for understanding the risks, sources, transport and impacts of chemical contaminants that threaten aquatic ecosystems and human wellbeing, CACE can impact global water quality.

Subproject Relevancy Statement

Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges

Technological advances in instrumentation and analytical approaches have significantly augmented the capacity of chemists and environmental scientists to quantify ultra-trace amounts of organic and inorganic substances with enormous precision, and characterize these substances on the molecular level. In addition, advanced fingerprinting using molecular biology methodologies such as gene expression, have significantly enhanced the conditions needed to sense environmental stress on aquatic organisms with high sensitivity and specificity. The combination of both approaches now allows scientists to conduct research to address complex challenges of environmental contamination and risk assessment. The CREST Team will not only enhance existing pollutant sensing methodologies, but also apply them to field studies across land-use boundaries (Subproject 2), and interface analytical data-intensive methodologies with advanced computational modeling and visualization, designed for the development of transformative and scalable methods for data mining and management (Subproject 3). As such, the research conducted by this Subproject will facilitate collaboration with the other two Subprojects and support researchers at the proposed CREST Center to better detect and understand the sources, transport, transformation and ecosystem responses to contaminants, pollutants and other natural stressors in the aquatic systems of south Florida. Using advanced analytical and molecular biology methodologies, researchers will be able to: 1) Develop sensing technologies to determine known traditional or emergent pollutants at environmentally relevant concentrations; 2) explore the relations between chemical stressors and biological responses through advanced molecular biology approaches, 3) generate products through collaboration efforts with the other Subprojects to communicate more effectively with decision and policymakers. **The ultimate goal of Subproject 1 is to advance and enhance the effectiveness of existing analytical approaches for the analysis of traditional pollutants, develop novel analytical methodologies and approaches for the identification, characterization and quantification of novel, previously unknown contaminants of concern, enhance and extend the applicability of molecular biology methodologies to assess environmental stressors to aquatic organisms across land-use boundaries**.

Links between CACE subprojects	**Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems**	**Subproject 3: Data Analytics for Effects Assessment and Decision Making**
Subproject 1: **Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges**	• Generate robust, environmentally relevant analytical chemistry and molecular biology data on contaminants, pollutants and other stressors in support of fate and transport process assessment through land-use boundaries • Provide spatially distributed water, sediment, soil and organism samples collected along land-use boundaries for pollutant and molecular biology analysis • Provide feedback on analytical and molecular biology methods' needs for adequate assessment of environmental risk and pollutant flux determinations	• Provide feedback on techniques for managing complex analytical data • Provide feedback on techniques for managing complex molecular biology data • Design and development trans-formative and scalable computational methods for sensor data mining and management • Generate robust, environmentally relevant analytical chemistry and molecular biology datasets on contaminants, pollutants and other stressors in support of the data mining, prediction and modeling efforts.

Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges

I.-Introduction:

Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors: Developing Cutting Edge Technologies for Tomorrow's Environmental Quality Challenges.

Despite a constant growth in numbers, the identity and nature of human-derived environmental contaminants have changed from traditional pollutants such as nutrients, trace metals, DDT and PCBs to other biologically active compounds such as endocrine disruptors, antibiotics and, more recently, a wide range of chemicals broadly classified as emergent chemicals of interest. Chemicals such as mercury, natural and synthetic estrogens, antibiotics, high use pharmaceuticals and even natural toxins produced as the result of algal blooms are now recognized as having significant effects both directly, in receiving ecosystems, and indirectly, to humans. Processes such as eutrophication, produce a direct and striking visual response for water quality degradation often evidenced by microorganism blooms, drastic changes of water color and in some cases animal deaths. The causes, however are often hidden in more subtle changes of the chemical ecology of an ecosystem triggered by chemicals (present in treated domestic effluents), contaminants (e.g. the combination of mercury and sulfate leading to mercury methylation) or natural stressors (wildfires, climatic disturbances and others) whose introduction affects the delicate balance of natural systems, particularly those in the direct path of urban development. The first signs of impact are often presented in water streams where small concentrations of these chemicals are constantly introduced and transported through multiple boundaries (air, water, soil, and organisms). Water quality is not only the most critical driver in ecosystem sustainability but also a major limiting factor for human development due to its effects on water scarcity. As a result, water tends to be progressively laced with anthropogenic signatures from components that elude treatment as it continues its journey from release to recharge to reuse. With few exceptions, in surface water bodies used as sources for drinking water, contamination is largely affected by dilution, so "exposure" needs to be characterized at very low concentrations for early intervention to be effective. The remaining challenges beyond the characterization of the environmental stressor, are establishing the links between the chemical indicator (CI) and its receptors, and determining the biological conduits and physiological events that will result in an adverse outcome pathway (AOP). Addressing these two key issues is essential in producing meaningful, population-based science in support of the decision making and regulatory processes.

The CREST team has taken on these challenges using multiple approaches creating new technologies, developing analytical methods, implementing extensive environmental assessments and characterizing the processes and mechanisms that control the "exposome" using engineering, chemistry, biology, statistics and molecular and traditional ecotoxicology. The principle behind the work of the sensing group is *to understand nature, one molecule at a time.*

II.-Research Plan:

Goals and Objectives:

The driving hypothesis for the sensing group is that *detection of environmentally relevant levels of pollutants, contaminants or stressors using high throughput technologies with high degree of specificity and at ultra-low concentrations will lead to the recognition of important pathways, interactions, changes and transport along environmental gradients and across land-use boundaries, that could negatively affect ecosystem functioning through anthropogenic activities, natural forces and altered biogeochemical cycles.*

Our research will:

1) Develop sensing technologies to determine known traditional or emergent pollutants at environmentally relevant concentrations (parts per billion to parts per-trillion) in multimedia samples (biotic and abiotic).
2) Take advantage of recently developed cutting-edge analytical chemistry tools to assess changes of the overall molecular composition of target, suspect and unknown components present in environmental samples through a relevant ecosystem boundary.
3) Apply molecular biology know-how to simultaneously assess the genetic and functional responses of relevant organisms or receptors to identify the role of these pollutants or stressors in the creation of adverse outcome pathways that may influence ecosystem functioning.

Methodologies and strategies developed in **Subproject 1** will be applied in field studies along relevant land-use boundaries defined by **Subproject 2** to generate information to feed the activities of **Subproject 3**. The availability of high resolution "big data" sets from the sensing group will greatly influence the capabilities of the ecosystem links and prediction and forecasting groups to create and fine-tune the management tools needed to assure South Florida's ecosystems future sustainability.

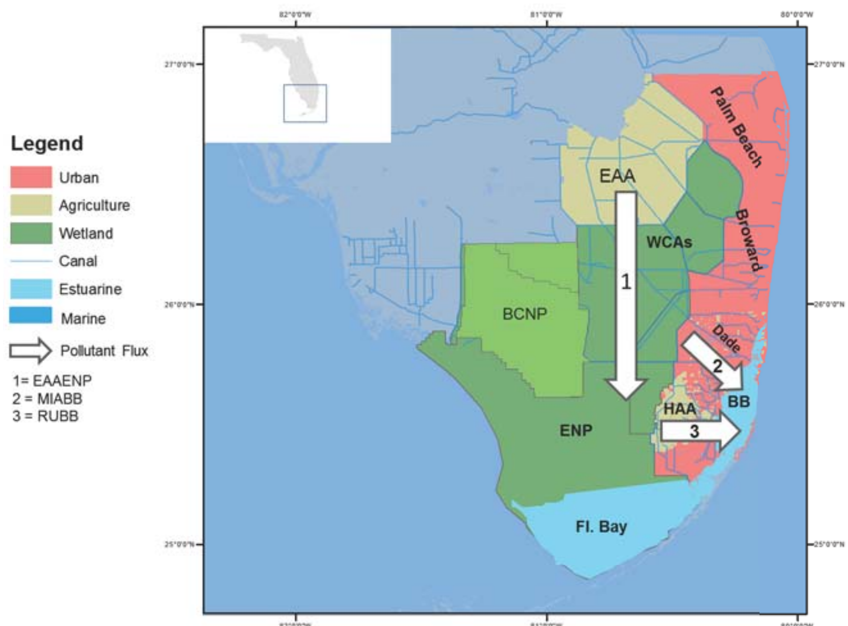


Figure 1. Major land-use and coastal areas of south Florida. Pollutant flux arrows represent three example transects investigated in this research: 1) EAAENP: Everglades Agricultural Area (EAA) to Everglades National Park (ENP) including the managed Water Conservation Areas (WCA); 2) MIABB: urban Miami (MIA)to Biscayne Bay (BB) via the Miami River; and 3) RUBB: the Redland Agricultural District across the suburban landscape of Miami and eventually to BB.

Experimental Approach:

Within this context, the research team has contributed to the advance of scientific inquiry in the following concentration areas and such expertise will be used to address processes described for the land-use transects described in Figure 1 as the basis for studies in **Subproject 2** (EAAENP, MIABB and RUBB).

Analytical Environmental Chemistry approaches:

Analytical Chemistry of Emergent Pollutants at Trace Levels. The use of advanced high resolution MS coupled with online solid phase extraction (SPE) and UHPLC has extended our interpretation of environmental exposure from trace analysis of critical compounds (Wang and Gardinali, 2013; Ramirez et al., 2014) to the detailed analysis of metabolites and degradation products so that a more accurate mass balance of the releases of emergent contaminants could be constructed (Wang and Gardinali, 2014). In addition, assay automation has produced high-throughput analysis of specific wastewater intrusions using recalcitrant tracers such as sucralose (Batchu et al., 2013), linking the traditional water quality approach, using non-specific nutrient and pathogen analysis, to a human-derived signature at costs which allow us to use the novel water tracer in large scale projects. Moving a step forward, the application of full scan mass spectrometry at resolution powers above 100K followed by all-ion fragmentation or specific (MS/MS) mass spectrometry experiments has greatly expanded the frontier of chemical characterization of the exposome by allowing retrospective identification of known-unknown or even unknown-unknown chemicals. Using these tools we can now identify suspect compounds, such as drugs of abuse and their transformation products, with the potential for molecular epidemiological studies targeting consumption at population levels without intrusive monitoring (Heuett et al., 2014). The same principle could be easily expanded to characterize the unknown recalcitrant portions of wastewaters in an effort to improve and redesign treatment technologies or even track previously undetected biological stressors in archived samples.

Analysis of emissions and byproducts from energy consumption in urban areas and intermediates during environmental transformation.

Resolving the transformation steps of energy consumption sources and byproducts to their molecular composition is a great challenge to analytical and environmental chemists because of the complex nature of the samples (Marshall and Rodgers, 2004; Qian et al., 2001a, 2001b). Moreover, commonly used energy sources (e.g., fossil fuel samples) and their transformation byproducts are composed of thousands of individual molecules derived from highly diverse populations of compound types (Labadibi et al., 2013; Gaspar et al., 2012; Panda et al., 2011; Altgelt and Boduszynski, 1992; Boduszynski, 1987,1988). Numerous approaches have been utilized to fully characterize energy sources and byproducts, but no single technique is fully descriptive, and some analyses generate conflicting data for the same sample (Aske et al., 2001; Zadro et al., 1985; Wang et al., 2005). As a result, complete description of the transformation of energy sources is often difficult, if not impossible. Identification and quantification of a limited number of a priori selected compounds have been used for their identification and fingerprint (Daling et al., 2002). Ultrahigh resolution MS (e.g., Fourier transform ion cyclotron resonance, FTICR- MS) has shown potential to resolve components from energy sources (e.g., fossil fuels) and identify the elemental composition, double bond equivalents (DBE = rings plus double bonds to carbon), and carbon number, based on accurate mass measurements at the required resolving power (Marshall and Rodgers, 2004; McKenna et al., 2014; Hughey et al., 2002; Qian et al., 2001a, 2001b). More recently, it has been shown that coupling fast, gas-phase post-ionization techniques (e.g., ion mobility spectrometry mass spectrometry, IMS-MS) provides an extra, orthogonal separation dimension for the analysis of fossil fuels (Fernandez-Lima et al., 2009; Ahmed et al., 2010,2014). In particular, ion mobility spectrometry (IMS) when combined with theoretical modeling, has proven to be the most versatile technique for conformational analysis of intermediate and equilibrium structures of molecular ions (Dugourd et al., 1997; Scott et al., 2007; Becker et al., 2008; Fernandez-Lima et al., 2008). That is, high resolution IMS-MS provides rapid separation of isomers (Kanu and Hill, 2007; Schenk et al., 2014a; Pierson et al., 2013; Merenbloom et al., 2009), conformers (Sawyer et al., 2005; Schenk et al., 2014b; Molano-Arevalo, 2014), and species of differing chemical class (Ruotolo et al., 2002; May et al., 2014) (based on differences in functional groups, polarities, and atomic compositions), which is advantageous for the rapid characterization and screening of intermediates and end products of the transformation of fossil fuel samples. Ion mobility measurements have been used to explore molecular dynamics and follow structural changes occurring on the millisecond time scale by comparison to CCS of candidate structures under controlled conditions (e.g., reactive/inert, polar/nonpolar bath gas at different temperatures, Kanu and Hill, 2007; Zilch et al., 2007, Fasciotti et al., 2013). With the advent of new variants of IMS analyzers with higher resolving power and the recently coupling to FTICR-MS in our laboratory, (Schenk et al., 2014a, 2014b, 2015 Hernandez et al., 2014; Castellanos et al., 2014; Fernandez-Lima et al., 2011a, 2011b) our efforts in CREST will be on the development of SA-TIMS-FTICR-MS (Benigni et al., 2015) methods to better understand and characterize at the molecular level the transformation intermediates and end products from energy consumption in urban areas during environmental transformation.

Black Carbon (BC) an overlooked source affecting global carbon dynamics. BC results from the partial combustion of organic matter during events such as wildfires and fossil fuel burning (Goldberg, 1985). A significant portion of BC is incorporated into soils, potentially affecting long-term carbon cycling in the environment. It has been estimated that BC comprises between 2 – 45% of soil organic carbon (Skjemstad et al., 1999; and others), and was initially thought to be refractory, persisting in soils for thousands of years. Recent studies however, have shown that turnover rates can occur on much shorter timescales (Singh et al., 2012 and references therein) and a significant portion of BC (char) is solubilized and exported as dissolved BC (DBC) (Jaffé et al., 2013; Ding et al., 2013). The environmental effects and implications of these chemicals remain undetermined to a large extent due to the lack of structural information. However, evidence for the effects of land-use on the distribution of DBC compounds has recently been observed (Wagner and Jaffe, unpublished). Thus, to gain a better understanding of the driving forces behind BC stability and mobility, a deeper knowledge of its chemical characteristics is necessary. While FT-ICR/MS characterizations have resulted in a series of potential DBC structures, their identity has not been confirmed as isolation from the bulk dissolved organic matter (DOM) has not been achieved. Heteroatomic compounds within the DBC pool may also be significant, such as dissolved black nitrogen (DBN) (Maie et al., 2006; Jaffé et al., 2012; Ding et al., 2014) and dissolved black sulfur (DBS) (Jaffé et al., unpublished) have been discovered, but very little is known about the molecular structure of these compounds (Wagner et al., 2015). While significant progress has been made in assessing the environmental dynamics of dissolved pyrogenic products, much work is still needed to constrain the molecular structures and reactivity of these molecules. This is needed to better assess their sources, transport and fate in the environment and predict their roles in global biogeochemical cycles and climate change. Through CREST we will focus on the physical isolation of DBC, DBN and DBS to enhance our capacity to fully characterize the molecular structures of these substrates.

Mercury in the Everglades. Mercury (Hg), a long recognized notorious pollutant, poses severe health risks to millions of people worldwide. The most recent global response, the signing of the Minamata Convention on Mercury by over 140 United Nation (UN) member states in Oct 2013, prompts the immediate needs for an advanced understanding of Hg biogeochemical cycling. Elevated levels of Hg have been found in fish, wading birds, and large predators in the Florida Everglades (Ware et al., 1990). Efforts have been made to investigate the scope and magnitude of Hg contamination and to understand Hg biogeochemical cycling in this subtropical wetland ecosystem (Gilmour et al., 1998; Krabbenhoft et al., 1998; Cai et al., 1999; Liu et al., 2008, 2009, 2011; Li et al., 2010, 2012). Several important aspects of Hg biogeochemical cycling, including source, transport, transformation (especially methylation/demethylation), and bioaccumulation have been investigated to various degrees. However, Hg cycling and the corresponding biogeochemical controls in this complicated ecosystem are not fully understood. In the CREST program, we will address a key issue that currently hampers our understanding of the mass budget and fate of mercury, i.e. the role of photochemical processes played in overall mercury cycling in the Everglades. One of our long-term research goals is to identify the importance, mechanism, and system-diversity of Hg photoreactions (oxidation, reduction, methylation, and demethylation) in aquatic environments (Li et al., 2012; Yin et al., 2014). The work proposed here is not only an expansion of our previous studies on methylmercury (MeHg) photodemethylation in the Florida Everglades, but also an essential step toward our long-term goal of producing a complete conceptual model of the Hg biogeochemical cycle. Photodemethylation was found to be a critical process for MeHg cycling in the Florida Everglades (Li et al., 2010; Tai et al., 2014), and could be driven by DOM. Our previous work further suggests that mechanisms of MeHg photodemethylation in aquatic systems may be system-specific (Li et al. 2010). Thus, a fundamental understanding of environmental chemical processes involved in MeHg photodemethylation is required to reveal the mechanisms of this process and its overall role in the cycling of MeHg (and Hg in general). This research aims to validate the relative importance of DOM, among other factors, in MeHg photodemethylation, to identify the major fractions and functional groups of DOM that are involved in the photodemethylation of MeHg, and to elucidate the pathways through which DOM is involved in MeHg photodemethylation in aquatic systems. Successful accomplishment of this proposed research will help us achieve our goals, providing essential information on understanding the fundamental chemistry of MeHg photo-demethylation and the role of this photochemical process in the overall Hg biogeochemical cycling.

Organismal and molecular biology approaches:

Organismal and molecular sensing: the roadmap from exposure to management. Aquatic ecosystems, especially coastal areas, can be subject to a variety of anthropogenic and natural stressors (e.g. nutrients, turbidity, pollutants, hypoxia, altered habitat/hydrologic cycles). Ultimately, the biotic responses to stressors manifest as changes in survival, growth, development, and reproduction (SGDR), affecting individuals, communities and populations. Given the complex nature of aquatic systems, a variety of molecular and biological endpoints are needed to understand the receptors, pathways and mechanism(s) of adverse outcomes and, more importantly, their environmental significance and prediction value. Among the levels at which single chemical contaminants and mixtures operate (Figure 2), molecular alterations stand out as being significant indicators of exposure and effect, and are also typically more sensitive than those at higher levels of organization. Furthermore, changes at the molecular level underlie the effects at higher levels of organization and, depending on the system affected and the type of response, can represent surrogates for effects at the organ and whole organism level.

Current advances in 'omics' research, and our basic toxicological understanding of molecular events, have increased the information available to make informed decisions. Accordingly, molecular biology tools can be applied to monitor the impact of contaminants in ecosystems (Fig. 2).

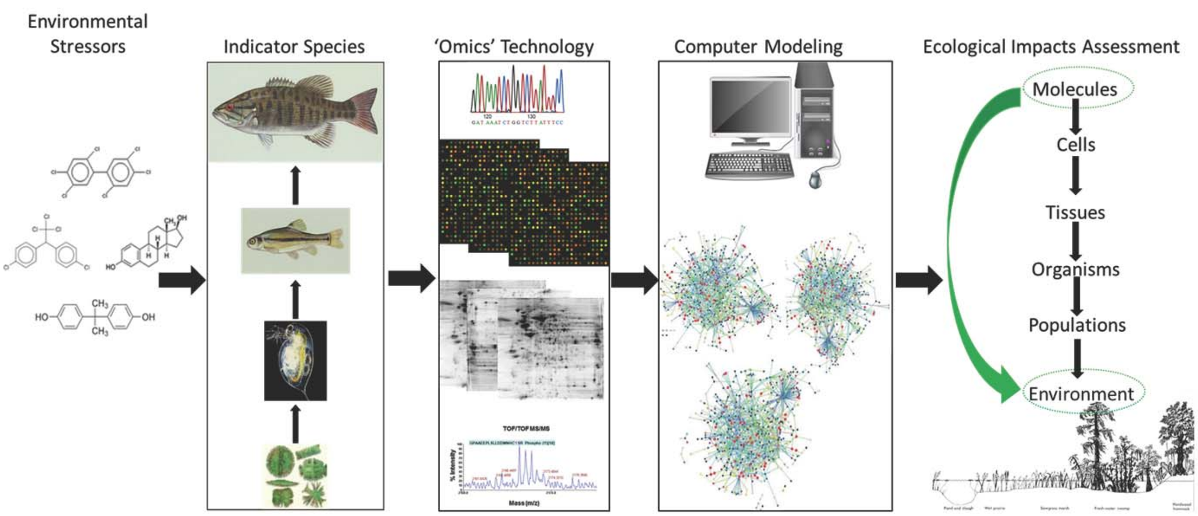


Fig. 2 High-Throughput Molecular Biological Analysis of Environmental Stressors: Since the function of an ecosystem can be viewed in the context of gene function rather than organism function, molecular biology tools such as high-throughput 'omics' technology that include genetic, epigenomic, and protein expression profiling can be applied to monitor the impact of environmental stressors on populations of diverse species and interacting communities at different strata of the ecosystem. Using high-throughput technologies to measure the molecular effects of stressors on biological indicator species, we will generate complex datasets to be analyzed computationally to identify key molecular pathways altered in response to stressors.

We propose to use a three-tier approach. In the first tier, we will evaluate DNA damage using a high throughput variant of the single cell gel electrophoresis assay (the comet assay; Karbaschi and Cooke, 2014) to identify the dose, mixtures of, and target organs for, environmental stressors causing DNA damage to biological indicator species. Single cell gel electrophoresis is widely used to evaluate DNA damage in regulatory toxicity testing, and therefore a perfect intermediate assessment between traditional primary ecotoxicology endpoints (SGDR), and the emerging 'omics approaches (see below). The comet assay has been applied to the detection of damage in cells of tissues such as gills, liver, and digestive glands, as a biomarker of deleterious effects of pollutants on fish and other aquatic organisms (e.g. Mitchelmore and Hyatt, 2004; Sullivan et al., 2007; Kang et al., 2014). Additionally, assessments of DNA damage at the single nucleotide resolution (using Damaged DNA Immuno-precipitation-seq, DDIP-seq, Hu et al. 2015; Yoshihara et al., 2015), have the potential to identify new genes important in the cellular response to environmental insult, and classify genotoxins based upon the distribution of damage.

Second, we will assess the impact of stressors on gene networks and damage at single nucleotide resolution. For that purpose we will develop the microarray characterization of gene expression responses to genotoxic compounds. Since these changes constitute the earliest response to adverse outcomes, the identification of specific group of genes involved in such process constitutes a highly sensitive biomarker of insult exposure (Venier et al. 2006; Steinberg et al. 2008). Additionally, the study of the epigenetic mechanisms mediating exposure-response relationships can elucidate how environmental factors influence phenotypic variation (Baccarelli and Bollati 2009; Bolatti and Baccarelli, 2010). The dynamic and potentially reversible nature of epigenetic changes has outstanding potential for the development of rapid and sensitive environmental biomonitoring programs in diverse ecosystems (Dolinoy and Jirtle, 2008; Huang et al., 2012; Suarez-Ulloa et al., 2015). DNA damage, determined by single cell gel electrophoresis, is widely used in regulatory toxicity testing, and therefore a perfect intermediate assessment between traditional primary ecotoxicology endpoints (SGDR), and the emerging 'omics' approaches. In the third-tier, we will take these large and complex gene datasets to construct ecologically relevant gene regulatory networks by computational model building and predict ecological impacts by model analysis methods. Accordingly the above experimental information will feed into our systems biology approach to develop mathematical models, visualization tools and prediction algorithms that could be expressed through the environmental gradients explored in Subproject 2 using the methods, techniques and technologies developed in Subproject 3. In addition, the large amounts of data resulting from molecular analyses will be organized into data warehouses providing full access to reads, consensus sequences and unigenes linked to gene ontology information (function, process and subcellular compartment).

In summary, the research focus from the organismal and biomolecular sensing group will (1) apply innovative, interdisciplinary approaches of genotoxicity profiling in biological indicator species, (2) determine the relationship between molecular effects/responses and effects in traditional ecotoxicological endpoints (i.e. SGDR); (3) develop a sophisticated and informative molecular ecological network modeling system to predict ecological impacts from complex mixtures; and (4) support a creative and integrated education program to attract and educate underrepresented students from our feeder magnet high-schools and throughout FIU's undergraduate and graduate education programs.

Broader Impacts

Subproject 1 pursues developing and applying a set of innovative molecular sensing technologies following a 3-tier approach. In doing so, this subproject will provide unique hands-on research and discovery experiences for minority students (undergraduate and graduate) and postdoctoral researchers, exposing them to state of the art research in analytical chemistry, genetics, toxicology and molecular biology. Additionally, as one of the major goals of this subproject is to trace relationships between stressors and biological responses using these tools, this work will set a framework for applying these results in environmental biomonitoring, improving quickness and sensibility of pollution detection. Therefore, an obvious societal impact of this subproject will be the transfer of these technologies, and the vast information they generate, to end-users and stakeholders, fostering the collaboration with industry partners and helping the decision making for managers and policy makers.

Subproject 1 References:

Ahmed, A., Y. Cho, K. Giles, E. Riches, J.W. Lee, H.I. Kim, C.H. Choi, S. Kim, 2014. Elucidating Molecular Structures of Nonalkylated and Short-Chain Alkyl (n < 5, (CH2)n) Aromatic Compounds in Crude Oils by a Combination of Ion Mobility and Ultrahigh-Resolution Mass Spectrometries and Theoretical Collisional Cross-Section Calculations, Analytical Chemistry 86 3300-3307.

Ahmed, A., Y.J. Cho, M.-h. No, J. Koh, N. Tomczyk, K. Giles, J.S. Yoo, S. Kim, 2010. Application of the Mason−Schamp Equation and Ion Mobility Mass Spectrometry To Identify Structurally Related Compounds in Crude Oil, Anal. Chem., 83, 77-83.

Altgelt, K.H., M.M. Boduszynski, 1992. Composition of heavy petroleums. 3. An improved boiling point-molecular weight relation, Energy & Fuels, 6, 68-72.

Aske, N., H. Kallevic, J. Sjoblom, 2001. Determination of Saturate, Aromatic, Resin, and Asphaltenic (SARA) Components in Crude Oils by Means of Infrared and Near-Infrared Spectroscopy, Energy Fuels, 15, 1304-1312.

Baccarelli A., V. Bollati, 2009. Epigenetics and environmental chemicals. Current Opinions in Pediatrics, 21, 243-51

Batchu, S.R., N. Quinete, V.R. Panditi, P.R Gardinali, 2013. Online solid phase extraction liquid chromatography tandem mass spectrometry (SPE-LC-MS/MS) method for the determination of Sucralose in reclaimed and drinking waters and its photo degradation in natural waters from South Florida" Chemistry Central Journal. 7, 141

Becker, C., K. Qian, D.H. Russell, 2008. Molecular weight distributions of asphaltenes and deasphaltened oils studied by laser desorption ionization and ion mobility mass spectrometry. Analytical chemistry, 80(22), 8592-8597.

Benigni, P., C.J. Thompson, M.E. Ridgeway, M.A. Park, F. Fernandez-Lima, 2015. Targeted high-resolution ion mobility separation coupled to ultrahigh-resolution mass spectrometry of endocrine disruptors in complex mixtures, Analytical Chemistry 87, 4321-4325.

Boduszynski, M.M., 1987. Composition of heavy petroleums. 1. Molecular weight, hydrogen deficiency, and heteroatom concentration as a function of atmospheric equivalent boiling point up to 1400.degree.F (760.degree.C), Energy & Fuels 1, 2-11.

Boduszynski, M.M., 1988. Composition of heavy petroleums. 2. Molecular characterization, Energy & Fuels 2, 597-613.

Bollati V, A. Baccarelli, 2010. Environmental Epigenetics. Heredity (Edinb) 105, 105-112.

Cai, Y., R. Jaffé, R.D. Jones, 1999. Interactions between dissolved organic carbon and mercury species in surface waters of the Florida Everglades. Appl. Geochem. 14, 395-407.

Castellanos, A., P. Benigni, D.R. Hernandez, J.D. DeBord, M.E. Ridgeway, M.A. Park, F.A. Fernandez-Lima, 2014. Fast Screening of Polycyclic Aromatic Hydrocarbons using Trapped Ion Mobility Spectrometry – Mass Spectrometry, Analytical Methods 6, 9328-9332.

Daling, P. S., L.G. Faksness, A.B. Hansen, S.A. Stout, 2002. Improved and standardized methodology for oil spill fingerprinting. Environmental Forensics, 3(3-4), 263-278.

Ding, Y., A. Watanabe, R. Jaffé, 2014. Dissolved black nitrogen (DBN) in freshwater environments. Organic Geochemistry 68, 1-4.

Ding, Y., Y. Yamashita, W.K. Dodds, R. Jaffé, 2013. Dissolved black carbon in grassland streams: Is there an effect of recent fire history? Chemosphere 90, 2557-2562.

Dolinoy D.C., R.L. Jirtle, 2008. Environmental epigenomics in human health and disease. Environ Mol Mutagen 49, 4-8.

Dugourd, P., R.R. Hudgins, M.F. Jarrold, 1997. High-resolution ion mobility studies of sodium chloride nanocrystals, Chemical Physics Letters 267, 186-192.

Fasciotti, M., P.M. Lalli, G. Heerdt, R.A. Steffen, Y.E. Corilo, G.F.d. Sá, R.J. Daroda, F.d.A.M. Reis, N.H. Morgon, R.C.L. Pereira, M.N. Eberlin, C.F. Klitzke, 2013. Int. J. Ion Mobil. Spec., 16, 117.

Fernandez-Lima, F., C. Becker, A.M. McKenna, R.P. Rodgers, A.G. Marshall, D.H. Russell, 2009. Petroleum Crude Oil Characterization by IMS-MS and FTICR MS, Analytical Chemistry 81, 9941-9947.

Fernandez-Lima, F., D.A. Kaplan, J. Suetering, M. Park, 2011. Gas-phase separation using a trapped ion mobility spectrometer, Int. J. Ion Mobility Spectrometry 14, 93-98.

Fernandez-Lima, F.A., D.A. Kaplan, M.A. Park, 2011. Note: Integration of trapped ion mobility spectrometry with mass spectrometry, Rev. Sci. Instr. 82, 126106.

Fernandez-Lima, F.A. , C. Becker, K. Gillig, W.K. Russell, M.A.C. Nascimento, D.H. Russell, 2008.Experimental and Theoretical Studies of (CsI)nCs+ Cluster Ions Produced by 355 nm Laser Desorption Ionization, The Journal of Physical Chemistry A 112, 11061-11066.

Gaspar, A., E. Zellermann, S. Lababidi, J. Reece, W. Schrader, 2012. Impact of Different Ionization Methods on the Molecular Assignments of Asphaltenes by FT-ICR Mass Spectrometry, Analytical Chemistry, 84, 5257-5267.

Gilmour, C.C., G.S. Riedel, M.C. Edrington, J.T. Bell, J.M. Benoit, G.A. Gill, M.C. Stordal, 1998. Methylmercury concentrations and production rates across a trophic gradient in the northern Everglades. Biogeochemistry 40, 327-345.

Goldberg, E., 1985. Black carbon in the environment. John Wiley and Sons, New York.

Hernandez, D.R., J.D. DeBord, M.E. Ridgeway, D.A. Kaplan, M.A. Park, F.A. Fernandez-Lima, 2014. Ion dynamics in a trapped ion mobility spectrometer, Analyst 139, 1913-1921.

Heuett, N., S. Rani, P. R. Gardinali, 2014. Understanding the magnitude of emergent contaminant releases through target screening and metabolite identification using high resolution mass spectrometry: illicit drugs in raw sewage influents. Journal of hazardous Materials. http://dx.doi.org/10.1016/j.jhazmat.2014.08.009.

Hu, J., S. Adar, C.P. Selby, J.D. Lieb, A. Sancar, 2015. Genome-wide analysis of human global and transcription-coupled excision repair of UV damage at single-nucleotide resolution. Genes & development, 29, 948-960.

Huang H.S., J.A. Allen, A.M. Mabb, I.F. King, J. Miriyala, B.Taylor-Blake, N. Sciaky, J.W. Dutton JW, Jr., H.M. Lee, X. Chen, J. Jin, A.S. Bridges, M.J. Zylka, B.L. Roth BL, B.D. Philpot, 2012. Topoisomerase inhibitors unsilence the dormant allele of Ube3a in neurons. Nature 481, 185-189.

Hughey, C.A., R.P. Rodgers, A.G. Marshall, 2002. Resolution of 11 000 Compositionally Distinct Components in a Single Electrospray Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrum of Crude Oil, Analytical Chemistry 74, 4145-4149.

Jaffé, R., Y. Ding, J Niggeman, A.V. Vähätalo, A. Stubbins, R.M. Spencer, J. Campbell, T. Dittmar, 2013. Global charcoal mobilization from soils via dissolution and riverine transport to oceans. Science 340, 345-347.

Jaffé, R., Y. Yamashita, N. Maie, W.T. Cooper, T. Dittmar, W.K. Dodds, J.B. Jones, T. Myoshi, J.R. Ortiz-Zayas, D.C. Podgorski, A. Watanabe, 2012. Dissolved organic matter in headwater streams: Compositional variability across climatic regions of North America. Geochimica et Cosmochimica Acta 94, 95-108.

Karbaschi, M., M.S. Cooke, 2014. Novel method for the high-throughput processing of slides for the comet assay. Scientific reports, 4.

Kang N., H.I. Kang, K.G. An, 2014. Analysis of fish DNA biomarkers as a molecular-level approach for ecological health assessments in an urban stream. Bull Environ Contam Toxicol 93, 555-560.

Kanu, A. B., H.H. Hill, 2007. Identity confirmation of drugs and explosives in ion mobility spectrometry using a secondary drift gas. Talanta, 73, 692-699.

Krabbenhoft, D., J. Hurley, M. Olson, L. Cleckner, 1998. Diel variability of mercury phase and species distributions in the Florida Everglades. Biogeochemistry 40, 311-325.

Lababidi, S., S.K. Panda, J.T. Andersson, W. Schrader, 2013. Direct Coupling of Normal-Phase High-Performance Liquid Chromatography to Atmospheric Pressure Laser Ionization Fourier Transform Ion Cyclotron Resonance Mass Spectrometry for the Characterization of Crude Oil, Analytical Chemistry, 85, 9478-9485.

Li, Y. Y. Mao, G. Liu, G. Tachiev, D. Roelant, X. Feng, Y. Cai, 2010. Degradation of Methylmercury and Its Effects on Mercury Distribution and Cycling in the Florida Everglades. Environ. Sci. Technol. 44, 6661−6666.

Li, Y, Y. Yin, G. Liu, G. Tachiev, D. Roelant, G. Jiang, Y. Cai, Y., 2012. Estimation of the Major Source and Sink of Methylmercury in the Florida Everglades. Environ. Sci. Technol. 46, 5885−5893.

Liu, G., Y. Cai, P. Kalla, D. Scheidt, J. Richards, L.J. Scinto, E. Gaiser, C. Appleby, C., 2008. Mercury Mass Budget Estimates and Cycling Seasonality in the Florida Everglades. Environ. Sci. Technol. 42 1954−1960.

Liu, G., Y. Cai, Y.Mao, D. Scheidt, P. Kalla, J. Richards, L.J. Scinto, G. Tachiev, D. Roelant, C. Appleby, 2009. Spatial Variability in Mercury Cycling and Relevant Biogeochemical Controls in the Florida Everglades. Environ. Sci. Technol. 43, 4361−4366.

Liu, G., M.G. Naja, P. Kalla, P., D. Scheidt, E. Gaiser, Y. Cai, 2011. Legacy and Fate of Mercury and Methylmercury in the Florida Everglades. Environ. Sci. Technol. 45, 496−501.

Maie, N., K.J. Parish, A. Watanabe, H. Knicker, R. Benner, T. Abe, K. Kaiser, R. Jaffé, 2006. Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem. Geochimica et Cosmochimica Acta 70, 4491-4506.

Marshall, A.G., R.P. Rodgers, 2004. Petroleomics: The Next Grand Challenge for Chemical Analysis, Acc. Chem. Res., 37, 53-59.

May, J.C., C.R. Goodwin, N.M. Lareau, K.L. Leaptrot, C.B. Morris, R.T. Kurulugama, A. Mordehai, C. Klein, W. Barry, E. Darland, G. Overney, K. Imatani, G.C. Stafford, J.C. Fjeldsted, J.A. McLean, 2014. Conformational Ordering of Biomolecules in the Gas Phase: Nitrogen Collision Cross Sections Measured on a Prototype High Resolution Drift Tube Ion Mobility-Mass Spectrometer, Analytical Chemistry, 86, 2107-2116.

McKenna, A.M., J.T. Williams, J.C. Putman, C. Aeppli, C.M. Reddy, D.L. Valentine, K.L. Lemkau, M.Y. Kellermann, J.J. Savory, N.K. Kaiser, A.G. Marshall, R.P. Rodgers, 2014. Unprecedented Ultrahigh Resolution FT-ICR Mass Spectrometry and Parts-Per-Billion Mass Accuracy Enable Direct Characterization of Nickel and Vanadyl Porphyrins in Petroleum from Natural Seeps, Energy & Fuels, 28, 2454-2464.

Merenbloom,S.I., R.S. Glaskin, Z.B. Henson, D.E. Clemmer, 2009. High-Resolution Ion Cyclotron Mobility Spectrometry, Analytical Chemistry, 81, 1482-1487.

Mitchelmore, C.L., S. Hyatt, 2004. Assessing DNA damage in cnidarians using the Comet assay. Marine Environmental Research, 58, 707-171.

Molano-Arevalo, J.C., D.R. Hernandez, W.G. Gonzalez, J. Miksovska, M.E. Ridgeway, M.A. Park, F.A. Fernandez-Lima, 2014. Flavin Adenine Dinucleotide Structural Motifs: From Solution to Gas Phase. Analytical chemistry, 86, 10223-10230.

Panda, S. K., K.J. Brockmann, T. Benter, W. Schrader, 2011. Atmospheric pressure laser ionization (APLI) coupled with Fourier transform ion cyclotron resonance mass spectrometry applied to petroleum samples analysis: comparison with electrospray ionization and atmospheric pressure photoionization methods. Rapid Communications in Mass Spectrometry, 25, 2317-2326.

Pierson, N.A., L. Chen, D.H. Russell, D.E. Clemmer, 2013. Cis–Trans Isomerizations of Proline Residues Are Key to Bradykinin Conformations, Journal of The American Chemical Society, 135, 3186-3192.

Qian, K., W.K. Robbins, C.A. Hughey, H.J. Cooper, R.P. Rodgers, A.G. Marshall, 2001a. Resolution and identification of elemental compositions for more than 3000 crude acids in heavy petroleum by negative-ion microelectrospray high-field Fourier transform ion cyclotron resonance mass spectrometry. Energy & Fuels, 15, 1505-1511.

Qian, K., R.P. Rodgers, C.L. Hendrickson, M.R. Emmett, A.G. Marshall, 2001b. Reading chemical fine print: Resolution and identification of 3000 nitrogen-containing aromatic compounds from a single electrospray ionization Fourier transform ion cyclotron resonance mass spectrum of heavy petroleum crude oil. Energy & Fuels 15, 492-498.

Ramirez, C., C. Wang, P. Gardinali, 2014. Fully automated trace level determination of parent and alkylated PAHs in environmental waters by online SPE-LC-APPI-MS/MS”. Analytical and Bioanalytical Chemistry, 406, 329-344.

Ruotolo, B.T., G.F. Verbeck IV, L.M. Thomson, A.S. Woods, K.J. Gillig, D.H. Russell, 2002. Distinguishing between phosphorylated and nonphosphorylated peptides with ion mobility-mass spectrometry. Journal of Proteome research, 1, 303-306.

Sawyer, H.A., J.T. Marini, E.G. Stone, B.T. Ruotolo, K.J. Gillig, D.H. Russell, 2005. The Structure of Gas-Phase Bradykinin Fragment 1-5 (RPPGF) Ions: An Ion Mobility Spectrometry and H/D Exchange Ion-Molecule Reaction Chemistry Study, J. American Society of Mass Spectrometry, 16, 893-905.

Schenk, E.R., M.E. Ridgeway, M.A. Park, F. Leng, F.A. Fernandez-Lima, 2014a. Isomerization kinetics of AT hook decapeptide solution structures. Analytical chemistry, 86, 1210-1214.

Schenk, E.R., V. Mendez, J.T. Landrum, M.E. Ridgeway, M.A. Park, F.A. Fernandez-Lima, 2014b. Direct observation of differences of carotenoid polyene chain cis/trans isomers resulting from structural topology. Analytical Chemistry 86(4), 2019-2024.

Schenk, E.R., R. Almeida, J. Miksovska, M.E. Ridgeway, M.A. Park, F.A. Fernandez-Lima, 2015. Kinetic intermediates of holo- and apo- myoglobin studied using HDX-TIMS-MS and molecular dynamic simulations, Journal of the American Society for Mass Spectrometry 26, 555-563.

Scott, C.D., M. Ugarov, R.H. Hauge, E.D. Sosa, S. Arepalli, J.A. Schultz, L. Yowell, 2007. Characterization of Large Fullerenes in Single-Wall Carbon Nanotube Production by Ion Mobility Mass Spectrometry, J. Phys. Chem. C. 111, 36-44.

Singh, N., S. Abiven, M.S. Torn, M.W.I. Schmidt, 2012. Fire-derived organic carbon in soil turns over on a centennial timescale. Biogeosciences 9, 2847-2857.

Skjemstad, J.O., J.A. Taylor, L.J. Janik, S.P. Marvanek, 1999. Soil organic carbon dynamics under long-term sugarcane monoculture. Australian Journal of Soil Research 37, 151-164.

Steinberg C.E., S.R. Sturzenbaum, R. Menzel, 2008. Genes and environment - striking the fine balance between sophisticated biomonitoring and true functional environmental genomics. Sci Total Environ 400,142-161.

Suarez-Ulloa, V., R. Gonzalez-Romero, J.M. Eirin-Lopez, 2015. Environmental epigenetics: a venue for developing next-generation biomarkers of marine pollution. Marine Pollution Bulletin, *in press.*

Sullivan C., C.L. Mitchelmore, R.C. Hale, P.A. Van Veld, 2007. Induction of CYP1A and DNA damage *in press*the fathead minnow (Pimephales promelas) following exposure to biosolids. Science of the Total Environment 384, 221-228.

Tai, C., Y. Li, Y. Yin, L.J. Scinto, G. Jiang, Y. Cai, 2014. Methylmercury photodegradation in surface water of the Florida Everglades: Importance of dissolved organic matter-methylmercury complexation. Environmental Science & Technology, 48, 7333-7340.

Venier P., C. De Pitta, A. Pallavicini, F. Marsano, L. Varotto, C. Romualdi, F. Dondero, A. Viarengo, G. Lanfranchi, 2006. Development of mussel mRNA profiling: can gene expression trends reveal coastal water pollution?. Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis, 602, 121-134.

Wagner, S., T. Dittmar, R. Jaffé, 2015. Molecular characterization of dissolved black nitrogen via electrospray ionization Fourier transform ion cyclotron resonance mass spectrometry. Organic Geochemistry, 79, 21-30.

Wang, C., P. R. Gardinali, 2013. Detection and occurrence of microconstituents in reclaimed water used for irrigation–a potentially overlooked source. Analytical and Bioanalytical Chemistry. 405, 5925-5935.

Wang, F., K.N. Wan, L.A. Green, 2005. GCxMS of diesel: A two dimensional separation approach, Anal. Chem. 77, 2777-2785.

Wang, J., P.R. Gardinali, 2014. Identification of phase II pharmaceutical metabolites in reclaimed water using high resolution benchtop Orbitrap mass spectrometry. Chemosphere, 107, 65-73.

Ware, F., H. Royals, F. Lange, 1990. Mercury contamination in Florida largemouth bass. Proceedings of Annual Conference of Southeastern Association of Fish & Wildlife Agencies 44, 5–12.

Yin, Y., Y. Li, C. Tai, Y. Cai, G. Jiang, 2014. Fumigant methyl iodide can methylate inorganic mercury species in natural waters. Nature Communications, 5.

Yoshihara, M., L. Jiang, S. Akatsuka, M. Suyama, S. Toyokuni, 2014. Genome-wide Profiling of 8-Oxoguanine Reveals Its Association with Spatial Positioning in Nucleus. DNA Research, 21, 603-612.

Zadro, S., J.K. Haken, W.V. Pinczewski, 1985. Analisis of Australian crude oils by high-resolution capillary gas-chromatography mass-spectrometry J. Chromatog. 323, 305-322.

Zilch, L.W., D.T. Kaleta, M. Kohtani, R. Krishnan, M.F. Jarrold, 2007. Folding and Unfolding of Helix-Turn-Helix Motifs in the Gas Phase, J. Am. Soc. Mass Spectrom. 18, 1239-1248.

CREST: Center for Aquatic Chemistry & the Environment (CACE)

Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems

Project Summary

During the past three decades, incidents involving pesticides, industrial chemicals, oil, pharmaceuticals, nutrients and metals have attracted worldwide attention and greatly affected environmental conditions (e.g., the Gulf of Mexico Deep-water Horizon Oil spill). These events demonstrate a regional, national and international need for enhanced research on the effects of toxic substances in the environment. The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) at Florida International University (FIU) will transform the institution by integrating discrete campus-wide programs across 10 departments and 4 colleges in fields from environmental chemistry through computer intensive data analysis and visualization, in order to tackle one of the regions most complex challenges: **environmental contamination**. CACE will create innovative opportunities for students, especially encouraging those from underrepresented minorities (URM), to participate in authentic research and foster their development as future STEM professionals. FIU CACE will unify this talented pool of researchers into a cohesive Center that will enhance collaborations, partnerships and synergies. The Center will bridge academic programs that exist across campuses by integrating graduate and undergraduate students into all research subprojects, emphasizing evidence-based educational approaches, technology advances, and analytical chemistry infrastructure, while providing authentic research experiences and solutions. CACE will transform cutting-edge research into technological and science-based solutions for various forms of water contamination using a framework that includes detection/identification, transport and fate in complex ecosystems, and data analytics and visualization. CACE will develop a modeling platform that will enable policy makers and managers to make informed decisions. FIU's CACE will work in collaboration with governmental and private sector partners in S. Florida to develop practical solutions to problems related to water quality in a natural-agricultural-urban setting. This partnership includes the South Florida Management District, the National Park Service, The Miccosukee Tribe of Indians, the Environmental Protection Agency, Everglades National Park, Department of Interior, and others.

Intellectual Merit

FIU CREST CACE will increase opportunities for graduate and undergraduate students, especially encouraging those from URMs, to conduct authentic research while advancing aquatic and environmental chemistry research and data analytics, methodologies, ecological risk assessments. CACE will generate significant new knowledge regarding contaminants and pollutants in aquatic environments, as well as produce innovative new methodologies for detecting and assessing contaminant quantities and impacts, including the use of molecular detection techniques. Using new data analytic approaches for visualization and synthesis of complex data, CACE will provide managers and policy makers, including governmental and private sector partners in S. Florida, real-time, accessible decision tools. The proposed program will advance current efforts on the biological effects, transport, transformation and distribution of contaminants in the environment into new collaborative research areas that investigate the sources and transport of contaminants and pollutants in aquatic systems. The research conducted by the Center will inform the economic, environmental, societal, policy, regulatory, and legal implications of water quality issues.

Broader Impacts

CACE will build on the success of FIU's evidence-based transformation of STEM instructional practices to provide enhanced support for students to pursue and complete STEM graduate degrees, both at FIU or elsewhere. Through an innovative program that spans the graduate school to high school spectrum, CACE will increase the success of students in graduate programs, especially supporting participation of underrepresented students in aquatic chemistry and environment (ACE) fields and future professions. CACE will develop technologies for improving water quality analysis and contaminant detection, as well as translate research findings into actionable information for decision-makers and stakeholders. By providing potential scenarios for understanding the risks, sources, transport and impacts of chemical contaminants that threaten aquatic ecosystems and human wellbeing, CACE can impact global water quality.

Subproject Relevancy Statement

Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems

Land use, water management, population growth and consumptive activities along with climate and disturbance regimes influence the sources, mobilization, transport and transformations of pollutants across land-use boundaries. The fate and environmental impact of the pollutants is governed by biogeochemical processes. We propose through this CREST research to determine the hydrologic transport, fate and environmental impact of pollutants/contaminants across three major land-use boundaries (agricultural, urban and natural) in South Florida under current and potentially changing environmental conditions**.**

We will establish three transects that encompass the main transitions between agriculture, urban and natural landscapes. These transects will provide a common research platform for students and faculty to apply the detection methodologies developed from Subproject 1 (Table 1), quantitatively sample water quality with a common experimental design, apply state of the art hydro-dynamic modeling (e.g., MIKE SHE/MIKE 11) techniques to couple flows with contaminants. We will combine pollutant source data from the analytics research in Subproject 3 along with pollutant concentrations determined in Subproject 1 into the hydro-dynamic model, and work with Subproject 3 to visualize the transport and fate of the pollutants across the landscape (Table 1). Finally, in using these data as our initial conditions, we will collaborate with our stakeholders and Subproject 3 to provide future scenarios models to predict future conditions of a changing landscape and climate (Table 1).

Table 1. Research links between CACE subgroups in this proposal.

Links between CACE subprojects	**Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors**	**Subproject 3: Data Analytics for Effects Assessment and Decision-Making**
Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems	• Collect surface water, groundwater, soil/sediment, and aquatic organisms for identification and quantification of nutrients, carbon and the suite of pollutants/contaminants • Collected organisms for gene expression in assessed for stress and nutrient retention indicators	• Provide Techniques for Managing Complex Data • Apply data analysis and trend identification techniques for studying current and historical sources and loading of pollutants, other agricultural chemicals and industrial emissions. • Provide visualization tools for hydro-dynamic models to determine cross-boundary fluxes • Develop computational methods to examine specific plausible and realistic scenarios for future changes in water and land resource management

In summary, CREST graduate students will gain experience in field sampling, laboratory detection, GIS analysis, experimental design, hydro-dynamic modeling and data analytics. The graduate students will become experts on the environmental detection, quantification, transport and fate of their chosen constituents. Their interdisciplinary research experience (analytical chemistry, hydrogeology, ecology, computer sciences) will provide them with a unique educational experience that can be translated into STEM careers in the private, government or educational sectors.

Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems

Introduction:

I. Interlinked Systems of South Florida

South Florida represents a unique triangular interface between three dominant systems - natural, agricultural and urban (Fig. 1), each of which has changed in area across the landscape through time. The natural landscape component consists of a diverse ecosystem gradient from uplands (Florida scrub, pine rocklands, tree islands, hammocks) to interior freshwater ridge and slough wetlands to coastal mangroves and estuarine seagrass beds (Lodge, 2010). In the last 100 years, more than 50% of the mainland portion of the natural system has been altered to urban and agricultural systems (McVoy, 2011). The products of the highly productive Florida agricultural sector are mainly destined to export and consumption elsewhere in the U.S. or abroad, resulting in an economic benefit locally, but leading to an accumulation of related pollution over time (Fig. 2). The greater risk of unintended potential effects of pesticides and fertilizers is the contamination of the soil-surface-groundwater system. This system supports aquatic life and related food chains and is used for recreation, drinking water, irrigation, and many other purposes. Regional agriculture is often cited for excessive loading of nutrients, pesticides, and herbicides into the Everglades (Daroub et al., 2011) and South Florida freshwater and marine ecosystems (Caccia and Boyer, 2007). At the same time, the urban sprawl of South Florida has experienced a tremendous expansion over the past few decades, leading to an increase in energy consumption, resulting in the local accumulation of energy-derived byproducts [e.g., gasoline, Polycyclic Aromatic Hydrocarbon (PAHs)] and waste products from human households (e.g., pharmaceuticals and personal care products (PPCPs) in treated domestic wastewaters). The Florida Everglades provides essential ecological and economic services to the larger and rapidly growing U.S. and global populations, making the world's largest, funded landscape restoration ($8 billion slated for the next 30 years) a critical endeavor. There is great uncertainty about how restored hydrologic and ecologic connectivity will enhance the movement of nutrients, pollutants, contaminants and water across the landscape.

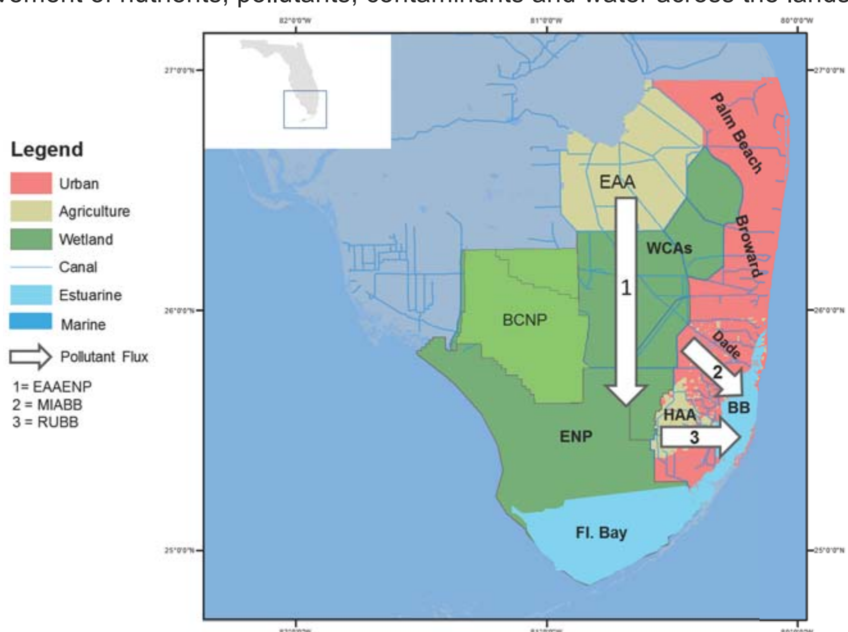


Figure 1. Major land-use and coastal areas of south Florida. Pollutant flux arrows represent three example transects investigated in this research: 1) EAAENP: Everglades Agricultural Area (EAA) to Everglades National Park (ENP) including the managed Water Conservation Areas (WCA); 2) MIABB: urban Miami (MIA)to Biscayne Bay (BB) via the Miami River; and 3) RUBB: the Redland Agricultural District across the suburban landscape of Miami and eventually to BB.

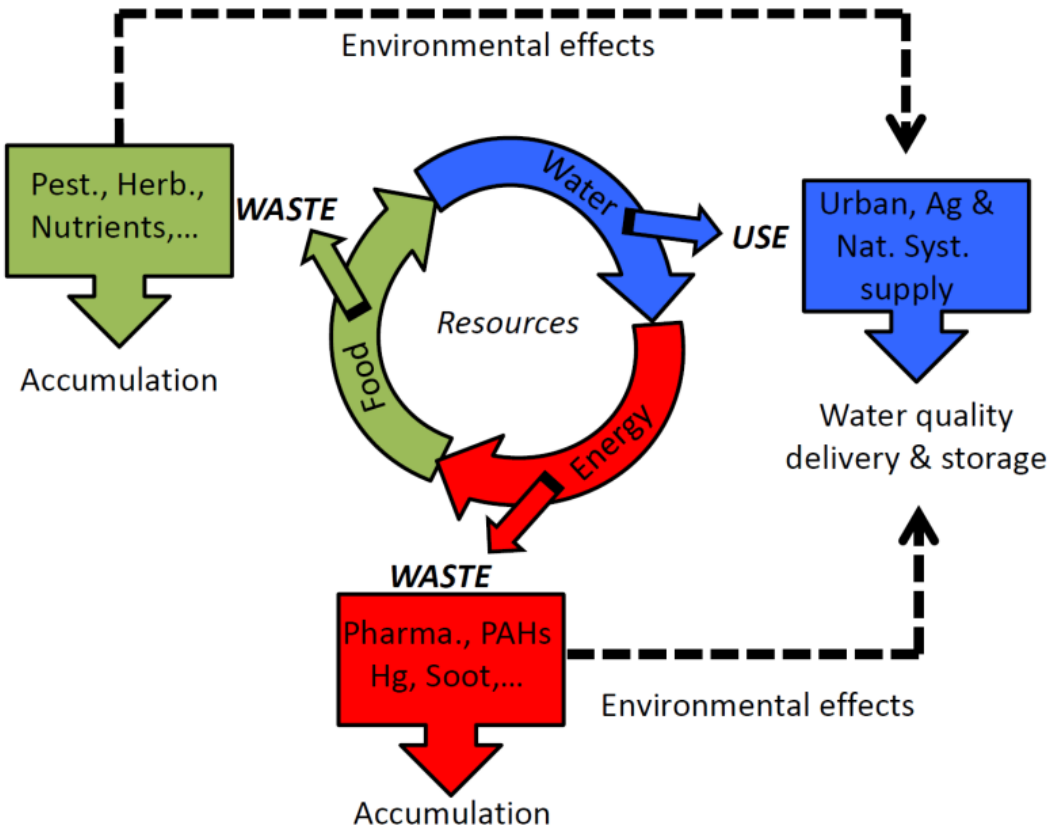


Figure 2. Conceptual diagram of the production of contaminants and pollutants in context of food, energy, and water supply

At the same time that Everglades and South Florida's natural ecosystems have been impacted by urban and agriculture nutrient and pollutant loads, agricultural acreage of South Florida has been decreasing and urban land cover has been increasing. Connectivity among these ecosystems is determined by hydrology, which is dynamically seasonal (distinct wet and dry seasons) and heavily managed through a complex system of canals, levees, and pump stations to protect the region from drought and flooding. Much of South Florida is underlain by the Biscayne Aquifer, an unconfined karst aquifer with numerous vertical and horizontal solution holes (Truss et al., 2007; Cunningham et al., 2006) that created one of the highest recorded hydraulic conductivities (DiFrenna et al., 2008) capable of transporting dissolved contaminants quickly (Renken et al., 2008). Development in South Florida was facilitated by the construction of approximately 2,500 km of drainage canals and approximately 17,000 retention ponds and borrows pits that now dot the landscape as "urban lakes". These excavations into the underlying karst limestone aquifer increase the occurrence of groundwater-surface water interactions (Price and Swart, 2006) and additionally result in land-applied contaminants having a high capacity for infiltration into the underlying aquifer.

In recent years, the composition of native and exotic flora and fauna across the south Florida landscape have shifted dramatically in response to fluxes of nutrients (Noe et al., 2001) other pollutants (e.g., mercury (Cai et al., 1999) as well as changes in climate and sea level rise (Saha et al., 2011). The overall impact of the release of pollutants/contaminants on a landscape under a seasonally defined and controlled water flow system experiencing a changing climate has yet to be addressed in a quantitative manner that would allow predictive scenarios for planning, decision-making, and management purposes. Ecologic and hydrologic connectivity alone will not ensure the resilience of a restored Everglades and other aquatic ecosystems in South Florida. Shifts in the direction, magnitude, and characteristics of ecological state changes are difficult to predict, requiring effective monitoring and adaptive management both during and after restoration. Increased understanding of effective restoration through adaptive management along with understanding the full consequences (and potential trade-offs) of restored connectivity are needed to maintain ecological integrity and ecosystem resilience of a restored Everglades. South Florida and the Greater Everglades landscape may well exemplify the positive decisions made by, and lessons learned by, human-intensified coastal communities worldwide, many with larger human populations and greater consequences for adaptive capacity and ecosystem resilience.

II. Research Plan:

Goals and Objectives:

Anthropogenic activities (land use, water management, population growth and consumptive activities) and natural forces (climate, hurricanes, storm surges, floods, droughts) influence the sources, mobilization, transport and transformations of pollutants across land-use boundaries. Biogeochemical processes govern the ultimate fate of these pollutants and their impacts on the environment. We propose through this CREST research to determine the hydrologic transport (flux), fate (biogeochemical processes), and environmental impact of pollutants/contaminants across three major land-use boundaries in South Florida under current and potentially changing environmental conditions**.**

Our research will:

1) Quantify the flux of water, nutrients, carbon, and pollutants/contaminants along transects that cross major land-use boundaries (agriculture, urban and natural; Fig. 1).
2) Hydro-dynamically model the flux of water and the transport of pollutants/contaminants and their associated biogeochemical processes along land-use boundary transects.
3) Predict the potential transport of pollutants/contaminants and any adverse biological outcomes with changing land use or natural forcings.

Experimental Approach:

The boundaries between the urban, rural/agricultural, and natural areas can be diffuse (e.g., low-density development grading to agricultural land) or, as is often the case in South Florida, sharply defined by water control structures such as canals and levees. The deposition of pollutants/contaminants on the land surface can be transported across the landscape via surface water sheet flow, canal flow, and groundwater flow, under a seasonal and managed hydrologic system. During transport, the pollutants/contaminants are expected to interact with soils, bedrock, vegetation, microbial communities, organic matter, other chemical constituents (e.g., sulfate), and the atmosphere resulting in transformations that may reduce or enhance their concentrations and toxicity. Ultimately these pollutants/contaminants applied on land may be discharged to coastal zones via surface water (rivers and canals) and/or groundwater (Stalker et al., 2009). Our proposed CREST research will quantify the flux of water and pollutants/contaminants across major land-use areas (agriculture, urban, natural) of South Florida. We highlight three example transects (not limited) to illustrate how differences in the sequence and intensity of different land uses can differentially affect the pollutant loading and the capacity of ecosystems to buffer or remove contaminants along the flow path (Fig.1).

Transect 1 **(EAAENP; Fig. 1)** represents the transfer from the Everglades Agricultural Area (EAA) to the adjacent natural areas of the Everglades including the managed Water Conservation Areas (WCA) and Everglades National Park (ENP). Much of the surface water flow through this system is hydrologically linked through water management and manipulated through pump stations, control gates and interior and exterior transfer canals. Agricultural land-use legacies in the EAA are at the heart of the challenge of restoring the Central Everglades and the explanation for an increase in droughts and fire in the southern Everglades. High concentrations of fertilized-derived nutrients (primarily phosphorus, P), especially in proximity to canals that bifurcate the surface water sheetflow across this landscape, have known, lasting effects on ecosystem state changes (ridge and slough transition to invasive *Typha*-dominated; Childers et al., 2003; Hagerthey et al., 2008). Other pollutants such as Mercury (Hg) and pesticides/herbicides (i.e. Atrazine) as well as natural dissolved organic matter (DOM) have been identified at concentrations higher than background along Transect 1 (Childers et al., 2003; Osborne et al., 2011; Pfeuffer, 2009; Yamashita et al., 2010) and are also known to be transferred from the EAA to down-stream aquatic systems including the WCAs and ENP (Yamashita et al., 2010). Additionally, the interactions of some of these chemicals can influence their transport and fate in the system. For example, transport, bioavailabilty, and photodegradation of various forms of Hg are influenced by DOM (Liu et al., 2009; Tai et al., 2014). **We hypothesize that transformations of pollutants/contaminants in Transect I will be influenced by seasonal hydrological cycles that influence water levels and residence time, and through interactions with dissolved and particulate organic matter.** Degradation and sequestration of pollutants/contaminants is expected along the southward flowpath from agricultural lands to stormwater treatment wetlands and water conservation areas of the Central Everglades. Current water management, however, compartmentalizes hydrology along this transect, artificially increasing and decreasing water residence times in a way that reduces the overall natural capacity of the landscape to remove pollutants/contaminants. A restored Everglades will enhance the removal capacity of these nutrients and agriculture-based contaminants along this transect, as well as reduce the loading of both to coastal South Florida during floods.

Transect 2 **(MIABB; Fig. 1)** represents the transfer from the urban area of Miami to the Miami River, transferring surface waters to the northern section of Biscayne Bay (surface water discharge), and ultimately to Biscayne National Park further towards the south through internal water currents and tidal activity. Concentrations of PPCPs, caffeine, and other compounds of purely anthropogenic sources are greatest in the most urban areas (Gardinali and Zhao, 2002; Rosi-Marshall and Royer, 2012, Rosi-Marshall et al., 2013). **We hypothesize that the loading of urban derived compounds to terrestrial and aquatic ecosystems along Transect II will vary with seasonal rainfall patterns, amount of impervious surface vs exposed soils, retention time, and biologic activity along the flow path.** Additionally, the design (e.g., lack of littoral zones) and operation of transfer canals will reduce water and DOM residence time, decreasing the ability of ecosystems to remove nutrients and other anthropogenic contaminants (Rosi-Marshall et al., 2013). This increasingly urban transect is expected to have the shortest water and OM residence times of the three transects and therefore have the highest loading of PPCPs, caffeine, and other human-derived organic compounds to downstream ecosystems with the most reduced removal capacity due to low organic matter inputs and retention that suppress biolfim growth and function. Everglades restoration is unlikely to enhance the removal capacity of these nutrients and urban-based contaminants.

Transect 3 **(RUBB; Fig. 1)** represents the transfer from the Redland Agricultural District across the suburban landscape of Miami and eventually to Biscayne Bay. The nursery industry in Miami-Dade County is the largest in Florida and second largest in the U.S. with 1,500 state certified nurseries (http://www.dade-agriculture.org/p/dade-agriculture.html) contributing nutrients and pesticides to nearby water canals and shallow groundwater systems. The greater Miami areas will also contribute contaminates from waste water as well as storm water drainage to canals and rivers and eventually to Biscayne Bay. Most importantly, a large waste water treatment plant (WWTP) in South Date is planned to discharge its treated waters into coastal wetlands adjacent to Biscayne Bay as part of a rehydration project (Miami-Dade County Water and Sewer Department, 2011; http://pbadupws.nrc.gov/docs/ML1226/ML12269A237.pdf). Finally, a landfill and nuclear power plant are located adjacent to Biscayne Bay along this transect. **We hypothesize that groundwater-surface water interactions will be most important in delivering nutrients and pollutants/contaminants along Transect III to Biscayne Bay.** This semi-restored agricultural transect is hypothesized to have the highest nutrient and other agriculture-based contaminant loads. The highly porous nature of the karst limestone aquifer can allow fertilizers, pesticides, herbicides to easily infiltrate from agricultural fields to the groundwater table. Canals located along this transect assist in keeping groundwater levels low to avoid flooding in the region, but also receive surface water runoff from the urban areas. The pollutants/contaminants released along transect 3 are expected to be transported with the groundwater flow, interact with the canals and urban lakes/ponds that cut through the aquifer and eventually discharge into Biscayne Bay. Everglades restoration projects aimed along the coastline (e.g., the C-111 Spreader

Canal and Biscayne Bay Coastal Wetlands rehydration projects) may enhance removal capacity of pollutants/contaminants along this flow path by reducing the amount of contaminant loading to groundwater and increasing water residence time in coastal wetlands.

Quantify the flux of water, nutrients carbon, and pollutants/contaminants along transects that cross major land-use boundaries

We will collect surface water, groundwater, soil/sediment, and aquatic organisms at key locations along each transect for identification and quantification of nutrients, carbon and the suite of pollutants/contaminants identified by **CREST Subproject 1**. Gene expression in collected organisms will be assessed for stress and nutrient retention indicators (see details in sub-project **CREST Subproject 1**). Current and historical sources and loadings of fertilizers, other agricultural chemicals as well as industrial emissions will be obtained from data-mining operations conducted by **CREST Subproject 3** and compared with the field observations. GIS mapping techniques will be employed to compare presence/absence and concentrations of nutrients, carbon and pollutants/contaminants with land use.

Trans-ecosystem boundary fluxes of water and associated contaminants will initially be modeled using advective-dispersion reactions considering the contaminants as conservative (non-reactive) in the environment. Conservative model results will be validated and revised based on observed contaminant concentrations from field samples. We propose to develop and use a hydrological model based on the MIKE SHE/MIKE 11 modeling system from DHI Water & Environment. This model couples surface and subsurface water flow using MIKE SHE, a 3-dimensional saturated and unsaturated groundwater flow, 2-dimensional overland/sheet flow model (Abbott et al., 1985; Refshaard et al., 1995;), and MIKE 11 (Havnø et al., 1995) a 1-dimensional river flow model which also includes management operations and schedules. A version of this model, Mike Marsh Model of Everglades National Park (M3ENP) was developed at FIU's Applied Research Center to simulate water flow in the Everglades (Cook, 2012) and used to model P transport in ENP with the addition of advection-dispersion transport equations (Long et al., 2015).

Hydrologic connectivity must be balanced by ecosystem fragmentation in order to use existing fragmentation to control the spread of pollutants and toxins (Jackson and Pringle, 2010; Rahel, 2013). Across all three proposed study transects, a network of canals both connects and fragments surface water, and changes in seasonal precipitation (i.e. wet and dry seasons) drive hydrologic connectivity. What is not understood is how well canals transport, transform and/or remove excessive nutrients as well as other pollutants such as pesticides, herbicides, pharmaceuticals, and trace metals. For example, mercury contamination has been one of the major concerns in south Florida, since elevated mercury levels have been observed in fishes in the Everglades and Florida Bay (Liu et al., 2008; Scheidt and Kalla, 2007). Mercury contamination has adversely affected recreational fishing activities, caused system-wide issuing of fish consumption advisories, and impaired wildlife and ecosystem health (Scheidt and Kalla, 2007). In addition to mercury, a limited number of studies have suggested that pollution by other metals, e.g., arsenic and lead, could be a problem in this area (Li et al., unpublished data). Mass budget estimation of mercury at the watershed level suggests that the mobility of toxic metals through hydrological connectivity and cross boundary transport needs to be considered for a better understanding of the cycling and fate of metals in this area (Liu et al., 2009). Similarly, through the application of pesticides and herbicides in local agriculture (including nurseries; Carriger et al., 2006; Quinete et al., 2013), and the discharge of reclaimed water (Wang and Gardinali, 2013; Rani-Batchu et al., 2013) and its use in coastal wetlands re-hydration projects, trace organic contaminants have entered local waterways (Singh et al., 2010). But still little is known about their loadings, seasonal and spatial distribution through hydrological connectivity.

In summary, along the proposed transects (Fig. 1), we will determine concentrations of the above-mentioned contaminants and biogeochemical components in different environmental compartments as applicable (water, soil, sediments, organisms), using methodologies described/determined by **CREST Subproject 1**. We will incorporate known sources and loadings of pollutants/contaminants as determined by **CREST Subproject 3** into a hydrodynamic model to determine cross-boundary fluxes assuming the constituents are conservative, and compare them to the observed field concentrations. Based on the hydrological model results, we will relate pollutant loadings, transport and concentrations to gene expression in organisms collected along the trans-boundary transects.

Hydro-dynamically model the flux of water and the transport of pollutants/contaminants and their associated transformations along the transects.

We will also study the effects of each land use type on biogeochemical cycling of nutrients, OM, as well as pollutants/contaminants in associated terrestrial and aquatic ecosystems using both field and laboratory experiments. CREST graduate students will be allowed to choose which pollutants/contaminants they are interested in and with their advisors design appropriate field and/or laboratory experiments to better quantify biogeochemical reactions associated with those constituents in the presence of the nutrient and OM concentrations observed in each of the land use areas. With regards to nutrients, we will investigate their removal rate as a function of the loading of nutrients and chemical toxins (e.g., pesticides), water column ecosystem metabolism, soil/sediment/bedrock interactions, water residence time, and atmospheric interactions. Specifically, we will measure how nutrient removal capacity and pollutant fluxes change as a function of the sequence and intensity of land use along the proposed transects. Lastly, we will incorporate into the database quantitative and molecular composition aspects of dissolved organic matter (DOM). DOM in south Florida has been found to mediate the transport and fate of mercury (Cai et al., 1999; Yamashita and Jaffé, 2008), to represent a large fraction of the total dissolved nitrogen (Maie et al., 2006), and to present optical properties that can be traces to different sources, including land-use and effects of fire (Yamashita et al., 2010; Chen et al., 2013; Ding et al., 2014). As such, it represents an important biogeochemical component intimately involved in the transport and fate of many other chemical parameters to be determined in this study.

To further model contaminant transport, fate and biogeochemistry process, we will use the ECO LAB (Butts et al., 2012) ecological modeling tool available with the MIKE SHE modeling framework. ECO Lab is a generic and open tool that can be customized to fit particular and unique ecosystems and is useful for developing problem-specific and need-based ecological models. We will customize ECO Lab to model ecological responses due to the selected scenarios including environmental consequences of developing new urban and industrial areas, impact of nutrient and pesticide loading on aquatic environment, risks in eutrophication, efficacy of nutrient and pesticide reduction plans, modeling of bacterial fate in wastewater contamination, impact of nutrient delivery on primary production and growth of benthic vegetation, and others.

Predicting the potential transport of contaminants with changing land-use or natural forcings:

Generation and evaluation of future scenarios of the dynamic variables affecting the hydrology, fate and contaminant transport in ecosystems of south Florida is an important component of this subproject. Scenarios are useful tools for impact modeling and help policy makers in making informed decisions. We will analyze past trends and future predictions of the selected variables to construct plausible scenarios. This component of this subproject will strive to determine the sensitivity of the selected hydrodynamic and waters quality models in south Florida, identify relevant dynamic variables for scenarios development and also evaluate the potential impacts of the selected scenarios on the hydrology, contaminant fate and transport and ecosystem health.

In this section, the hydrodynamic model MIKE SHE/MIKE 11 will be run to predict the transport of water and contaminants under various land-use change scenarios (population growth, water demand) as well as potential environmental forcings (rainfall and temperature changes due to climate change as well as sea level rise). In this study, we will examine specific and realistic scenarios of future change in driving variables (land use and climate change) evaluating the risks to the South Florida ecosystems and vulnerabilities of humans to projected changes.

In collaboration with decision makers and other local partners, our **CREST Subproject 3** will develop a reasonable number (6-9) of future scenarios that consider proposed plans and past trends in the nature and distribution of dynamic variables (e.g. land cover, management techniques, population, infrastructure, urban systems). Here we subscribe to the IPCC definition of a scenario as "a coherent, internally consistent, and plausible description of a possible future state of the world." Scenarios developed as part of this project are intended for the purposes of strategic planning and for guiding future policy development in South Florida. We expect to develop *exploratory scenarios*, which describe how the future of the South Florida might unfold given our understanding of change processes and an extrapolation of past trends. We will include a Business as Usual (BAU) scenario as a baseline extrapolation of all, relevant past trends. Due to global climate change, temperatures are predicted to rise across the region, sea-level will continue to rise and precipitation is expected to decrease by about 10% (Aumen et al. 2015). Based on plausible scenarios taking into account land use change (urbanization, intensive urban agriculture) and climate change (increase in temperature and evapotranspiration as well as a decrease in precipitation), we will conduct impact modeling on the fluxes of water (surface water, groundwater, evapotranspiration) and contaminants along the transects and across the land use boundaries.

Broader Impacts

The subgroup will promote science and increase our understanding of the *environmental benefits* of a *balanced* and *well-functioning natural-human ecosystem,* now and under future climate/land-use change scenarios. We will train graduate students in state-of-art methods of contaminant transport, biogeochemical processes, and ecosystem response. Through this subgroup and in collaboration with the other sub-groups, the students will gain interdisciplinary research in the fields of analytical chemistry, hydrogeology, hydro-dynamic modeling, ecology, and computer sciences providing them a one-of-a-kind educational experience. Their interactions with our collaborators in the private and governmental sectors will provide them with the understanding of how science can be used in local and regional decisions related to land use, water supply, and sustainability.

Subproject 2 References

Abbott, M.B., J.C. Bathurst, J.A. Cunge, P.E. O'Connell, J. Rasmussen, 1986. An introduction to the European Hydrological System—Systeme Hydrologique Europeen, SHE. 2 Structure of a physically-based distributed modelling system. Journal of Hydrology, 87, 61-77.

Aumen, N.G., K.E. Havens, G.R. Best, L. Berry, 2015. Predicting Ecological Responses of the Florida Everglades to Possible Future Climate Scenarios: Introduction. Environmental management, 55(4), 741-748.

Butts, M., M.C. Loinaz, P. Bauer-Gottwein, R. Unnasch, D. Gross, 2012. *Mike She-Ecolab: An Integrated Catchment-Scale Eco-Hydrological Modelling Tool*. Abstract from 19th International Conference on Computational Methods in Water Resources, Illinois, United States.

Caccia, V.G., J.N. Boyer, 2007. A nutrient loading budget for Biscayne Bay, Florida. Marine Pollution Bulletin, 54(7), 994-1008.

Cai, Y., R. Jaffé, R.D. Jones, 1999. Interaction between dissolved organic carbon and mercury species in surface waters of the Florida Everglades. Applied Geochemistry, 14, 395-407.

Carriger, J.F., G.M. Rand, P.R. Gardinali, W.B. Perry, M.S. Tompkins, A.M. Fernandez, 2006. Pesticides of potential ecological concern in sediment from south Florida canals: an ecological risk prioritization for aquatic arthropods. Soil & Sediment Contamination, 15, 21-45.

Chen, M., N. Maie, K. Parish, and R. Jaffé. 2013. Spatial and temporal variability of dissolved organic matter quantity and composition in an oilgotrophic subtropical coastal wetland. Biogeochemistry, 115, 167-183.

Childers, D.L., R.F. Doren, R. Jones, G.B. Noe, M. Rugge, L.J. Scinto, 2003. Decadal change in vegetation and soil phosphorus patterns across the Everglades landscape. Journal of Environmental Quality, 32, 344–362.

Cook, A., 2012. Development of an integrated surface and subsurface model of Everglades National Park. FIU Electronic Theses and Dissertations. Paper 634 (URL:http://digitalcommons.fiu.edu/etd/634).

Cunningham, K.J., R.A. Renken, M.A. Wacker, M.R. Zygnerski, E. Robinson, A.M. Shapiro, G.L. Wingard, 2006. Application of carbonate cyclostratigraphy and borehole geophysics to delineate porosity and preferential flow in the karst limestone of the Biscayne aquifer, SE Florida. Geological Society of America Special Papers, 404, 191-208.

Daroub, S.H., S. Van Horn, T.A. Lang, O.A. Diaz, 2011. Best Management Practices and Long-Term Water Quality Trends in the Everglades Agricultural Area. Critical Reviews in Environmental Science and Technology, 41(S1), 608-632.

DiFrenna, V. J., R.M. Price, M.R. Savabi, 2008. Identification of a hydrodynamic threshold in karst rocks from the Biscayne Aquifer, South Florida, USA, Hydrogeology Journal, 16, 31-42. DOI 10.1007/s10040-007-0219-4.

Ding, Y., K. Cawley, C. Nunes, R. Jaffé, 2014. Environmental dynamics of dissolved black carbon in wetlands. Biogeochemistry, 119, 259-273.

Gardinali, P.R., X. Zhao, 2002. Trace determination of caffeine in surface water samples by liquid chromatography–atmospheric pressure chemical ionization–mass spectrometry (LC–APCI–MS). Environment International, 28, 521-528.

Hagerthey, S.E., S. Newman, K. Rutchey, E.P. Smith, J. Godin, 2008. Multiple regime shifts in a subtropical peatland: community-specific thresholds to eutrophication. Ecological Monographs, 78, 547-565.

Havnø, K., M.N. Madsen, J. Dørge, V.P. Singh, 1995. MIKE 11-a generalized river modelling package. Computer models of watershed hydrology, 733-782.

Jackson, C.R., C.M. Pringle, 2010. Ecological benefits of reduced hydrologic connectivity in intensively developed landscapes. BioScience, 60, 37-46.

Liu, G.L., Y. Cai, P. Kalla, D. Scheidt, J. Richards, L.J. Scinto, E. Gaiser, C. Appleby, 2008. Mercury Mass Budget Estimates and Cycling Seasonality in the Florida Everglades. Environmental Science & Technology, 42, 1954-1960.

Liu, G. L., Y. Cai, Y.X. Mao, D. Scheidt, P. Kalla, J. Richards, L.J. Scinto, G. Tachiev, D. Roelant, C. Appleby, 2009. Spatial Variability in Mercury Cycling and Relevant Biogeochemical Controls in the Florida Everglades. Environmental Science & Technology, 43, 4361-4366.

Lodge, T.E. 2010. The Everglades Handbook: Understanding the Ecosystem. (3rd ed.): CRC Press, Taylor and Francis Group, Florida, 392 pp**.**

Long, S.A., G.I. Tachiev, R. Fennema, A.M. Cook, M.C. Sukop**,** F. Miralles-Wilhelm, 2015. Modeling the impact of restoration efforts on phosphorus loading and transport through Everglades National Park, FL, USA, Science of The Total Environment, 520, 1 July 2015. http://www.sciencedirect.com/science/article/pii/S0048969715001163

Maie N., K. Parish, A. Watanabe, H. Knicker, R. Benner, T. Abe, K. Kaiser, R. Jaffé, 2006. Chemical characteristics of dissolved organic nitrogen in an oligotrophic subtropical coastal ecosystem. Geochimica et Cosmochimica Acta, 70, 4491-4506.

McVoy, C.W., W.P. Said, J. Obeysekera, J.A. VanArman, T.W. Dreschel, 2011. *Landscapes and hydrology of the predrainage Everglades* Gainesville: University Press of Florida. 576p.

Noe, G.B., D.L. Childers, R.D. Jones, 2001. Phosphorus biogeochemistry and the impact of phosphorus enrichment: why is the Everglades so unique? Ecosystems, *4*(7), 603-624.

Osborne, T.Z., S. Newman, D.J. Scheidt, P.I. Kalla, G.L. Bruland, M.J. Cohen, L.J. Scinto, L.R. Ellis, 2011. Landscape patterns of significant soil nutrients and contaminants in the greater Everglades ecosystem: Past, present, and future. Critical Reviews in Environmental Science and Technology, 41, 121-148.

Pfeuffer, R.J, 2009. Ambient pesticide monitoring network: 1992 to 2007. Technical Publication SFWMD 105, South Florida Water Management District, West Palm Beach, FL, p. 57

Price, R.M., P. Swart, 2006. Geochemical indicators of groundwater recharge in the Surficial Aquifer System, Everglades National Park, Florida, USA, in Harmon, R.S., and Wicks, C., eds., Perspectives on karst geomorphology, hydrology, and geochemistry—A tribute volume to Derek C. Ford and William B. White: Geological Society of America Special Paper 404, 251-266. doi: 10.1130/2006.2404(21).

Quinete, N., J. Castro, A. Fernandez, I.M. Zamora-Ley, G.M. Rand, P.R. Gardinali, 2013. Occurrence and distribution of endosulfan in water, sediment, and fish tissue: An ecological assessment of protected lands in South Florida. Journal of Agricultural and Food Chemistry, 61, 11881-11892.

Rahel, F.J., 2013. Intentional fragmentation as a management strategy in aquatic systems. BioScience, 63, 362-372.

Rani-Batchu S., N. Quinete, V.R. Panditi, P.R Gardinali, 2013. Online solid phase extraction liquid chromatography tandem mass spectrometry (SPE-LC-MS/MS) method for the determination of Sucralose in reclaimed and drinking waters and its photo degradation in natural waters from South Florida. Chemistry Central Journal, 7, 141.

Refshaard, J.C., B. Storm, V.P. Singh, 1995. MIKE SHE. Computer models of watershed hydrology., 809-846.

Renken, R.A., K.J. Cunningham, A.M. Shapiro, R.W. Harvey, M.R. Zygnerski, D.W. Metge, M.A. Wacker, 2008. Pathogen and chemical transport in the karst limestone of the Biscayne aquifer: Revised conceptualization of groundwater flow. Water Resources Research 44(8).

Rosi-Marshall, E.J., D.W. Kincaid, H.A. Bechtold, T.V. Royer, M. Rojas, J.J. Kelly, 2013. Pharmaceuticals suppress algal growth and microbial respiration and alter bacterial communities in stream biofilms. Ecological Applications, 23, 583-593.

Rosi-Marshall, E.J., T.V. Royer, 2012. Pharmaceutical compounds and ecosystem function: an emerging research challenge for aquatic ecologists. Ecosystems, 15. 867-880.

Saha, A.K., S. Saha, J. Sadle, J. Jiang, M.S. Ross, R.M. Price, L.S.L.O. Sternberg, K.S. Wendelberger, 2011. Sea level rise and South Florida coastal forests. Climate Change, 107(1-2), 81-108.

Scheidt, D., P. Kalla, 2007. Everglades ecosystem assessment: water management and quality, eutrophication, mercury contamination, soils and habitat: monitoring for adaptive management: A R-EMAP status report, USEPA Region 4, Athens, GA.

Singh, S.P., A. Azua, A. Chaudhary, S. Khan, K.L. Willett, P.R. Gardinali, 2010. Occurrence and distribution of steroids, hormones and selected pharmaceuticals in South Florida coastal environments. Ecotoxicology, 19, 338-350.

Stalker, J. C., R.M. Price, P.K. Swart, 2009. Determining Spatial and Temporal Inputs of Freshwater, Including Groundwater Discharge, to a Subtropical Estuary Using Geochemical Tracers, Biscayne Bay, South Florida. Estuaries and Coasts, 32, 694-708.

Tai, C., Y. Li, Y. Yin, L.J. Scinto, G. Jiang, Y. Cai, 2014. Methylmercury photodegradation in surface water of the Florida Everglades: Importance of dissolved organic matter-methylmercury complexation. Environ Sci Technol, 48, 7333-7340.

Truss, S., M. Grasmueck, S. Vega, D.A. Viggiano, 2007. Imaging rainfall drainage within the Miami oolitic limestone using high-resolution time-lapse ground-penetrating radar. Water Resources Research 43(3).

Wang, C., P.R. Gardinali, 2013. Detection and occurrence of microconstituents in reclaimed water used for irrigation–a potentially overlooked source. Analytical and Bioanalytical Chemistry, 405, 5925-5935.

Yamashita, Y., L.J. Scinto, N. Maie, R. Jaffé, 2010. Dissolved organic matter characteristics across a subtropical wetland's landscape: Application of optical properties in the assessment of environmental dynamics. Ecosystems, 13, 1006-1019.

Yamashita, Y. R. Jaffé, 2008. Characterizing the Interactions between Trace Metals and Dissolved Organic Matter Using Excitation-Emission Matrix and Parallel Factor Analysis. Environmental Science & Technology, 42, 7374-7379.

CREST: Center for Aquatic Chemistry & the Environment (CACE)

Subproject 3: Data Analytics for Effects Assessment and Decision Making

Project Summary

During the past three decades, incidents involving pesticides, industrial chemicals, oil, pharmaceuticals, nutrients and metals have attracted worldwide attention and greatly affected environmental conditions (e.g., the Gulf of Mexico Deep-water Horizon Oil spill). These events demonstrate a regional, national and international need for enhanced research on the effects of toxic substances in the environment. The proposed CREST Center for Aquatic Chemistry & the Environment (CACE) at Florida International University (FIU) will transform the institution by integrating discrete campus-wide programs across 10 departments and 4 colleges in fields from environmental chemistry through computer intensive data analysis and visualization, in order to tackle one of the regions most complex challenges: **environmental contamination**. CACE will create innovative opportunities for students, especially encouraging those from underrepresented minorities (URM), to participate in authentic research and foster their development as future STEM professionals. FIU CACE will unify this talented pool of researchers into a cohesive Center that will enhance collaborations, partnerships and synergies. The Center will bridge academic programs that exist across campuses by integrating graduate and undergraduate students into all research subprojects, emphasizing evidence-based educational approaches, technology advances, and analytical chemistry infrastructure, while providing authentic research experiences and solutions. CACE will transform cutting-edge research into technological and science-based solutions for various forms of water contamination using a framework that includes detection/identification, transport and fate in complex ecosystems, and data analytics and visualization. CACE will develop a modeling platform that will enable policy makers and managers to make informed decisions. FIU's CACE will work in collaboration with governmental and private sector partners in S. Florida to develop practical solutions to problems related to water quality in a natural-agricultural-urban setting. This partnership includes the South Florida Management District, the National Park Service, The Miccosukee Tribe of Indians, the Environmental Protection Agency, Everglades National Park, Department of Interior, and others.

Intellectual Merit

FIU CREST CACE will increase opportunities for graduate and undergraduate students, especially encouraging those from URMs, to conduct authentic research while advancing aquatic and environmental chemistry research and data analytics, methodologies, ecological risk assessments. CACE will generate significant new knowledge regarding contaminants and pollutants in aquatic environments, as well as produce innovative new methodologies for detecting and assessing contaminant quantities and impacts, including the use of molecular detection techniques. Using new data analytic approaches for visualization and synthesis of complex data, CACE will provide managers and policy makers, including governmental and private sector partners in S. Florida, real-time, accessible decision tools. The proposed program will advance current efforts on the biological effects, transport, transformation and distribution of contaminants in the environment into new collaborative research areas that investigate the sources and transport of contaminants and pollutants in aquatic systems. The research conducted by the Center will inform the economic, environmental, societal, policy, regulatory, and legal implications of water quality issues.

Broader Impacts

CACE will build on the success of FIU's evidence-based transformation of STEM instructional practices to provide enhanced support for students to pursue and complete STEM graduate degrees, both at FIU or elsewhere. Through an innovative program that spans the graduate school to high school spectrum, CACE will increase the success of students in graduate programs, especially supporting participation of underrepresented students in aquatic chemistry and environment (ACE) fields and future professions. CACE will develop technologies for improving water quality analysis and contaminant detection, as well as translate research findings into actionable information for decision-makers and stakeholders. By providing potential scenarios for understanding the risks, sources, transport and impacts of chemical contaminants that threaten aquatic ecosystems and human wellbeing, CACE can impact global water quality.

Subproject Relevancy Statement
Subproject 3: Data Analytics for Effects Assessment and Decision Making
Technological advances in hardware, storage, and software have significantly increased scientists' ability to find new ways of generating and using data, thus creating new possibilities for conducting research to address the complex challenges of environmental contamination and ecological risk assessment. Conducting scientific research through high-resolution data acquisition, data mining, and visualization enables scientists to better understand the transient nature of aquatic data and pollutant movement across various boundaries.

However, data intensive science requires significant collaboration between environmental and computer scientists. This collaboration is becoming increasingly critical in finding better and more effective ways to research, discover and solve problems. The research conducted by this Subproject will facilitate such collaboration and support CACE researchers at the proposed CREST Center to better detect and understand the sources, transport, transformation and ecosystem responses to contaminants, pollutants and other natural stressors in the aquatic systems of south Florida.

Using a data-intensive approach, CACE researchers will be able to: 1) provide detailed characterization and measurement of the environmental pollutants, 2) improve predictive abilities on effects of pollutants and address future water quality issues, 3) explore, manipulate and visualize data thus collaborate more effectively for risk assessment, 4) conduct literature mining on the nature of contaminants and access relevant environmental information rapidly, and 5) communicate more effectively with decision makers and other stakeholders. **The ultimate goal of this Subproject is to support data-intensive research on aquatic chemistry and the environment by developing transformative and scalable methods for data mining and management, advanced computational modeling, and visualization.** The table below is a summary of how this Subproject relates to the two other Subprojects of the proposed CREST Center.

Links between CACE subprojects	**Subproject 1: Advanced Sensing of Environmental Exposure to Anthropogenic Contaminants, Pollutants and Other Natural Stressors**	**Subproject 2: Quantifying the Fate and Transport of Contaminants across Natural, Agricultural and Human Systems**
Subproject 3: Data Analytics for Effects Assessment and Decision Making	• Create novel multi-tiered data analysis architecture, consisting of sensors and cloud/HPC computing systems • Provide mining capability investigation by utilizing associations and correlations among the data to understand the characteristics and to extract semantics and patterns from the data • Provide techniques for managing complex analogue environmental and digital molecular biology digital data • Provide synthesis and analysis of gene function networks • Create data visualization and decision making support tools	• Provide mining capability investigation by utilizing associations and correlations among the data to understand the characteristics and to extract semantics and patterns from the data • Provide data analysis support for quantifying and trend identification of current and historical sources and loading of pollutants, • Develop appropriate visualization tools to examine specific plausible and realistic scenarios for future changes in water and land resource management • Provide capacity for literature mining of biological and visual analytics and visualization algorithms to assist in assessment and strategic decision-making

Subproject 3: Data Analytics for Effects Assessment and Decision Making

1. Introduction

Computational research and tools developed under this Subproject are designed to support the CREST Center's team efforts in identification of source, transport, transformation and ecosystem responses to contaminants, pollutants and other natural stressors. Researching environmental contamination and ecological risk assessment requires investigation of non-chemical, as well as organic and inorganic chemical stressors including **nutrients, contaminants and pollutants** with a multitude of exposure types (e.g., single-slug, intermittent and continuous) with native, exotic and standard test species.

Conducting this research entails collection of large volumes of data from various heterogeneous sources such as data from analytical chemistry techniques and data from biogeochemical cycles used to determine how natural processes affect ecosystems. As the scale and complexity of these data types increase exponentially, it becomes challenging to effectively model the increasing volumes of data, discover useful information, and provide data analytics capability to support effective and accurate assessment and decision-making capability for the scientists and their partners.

To address these challenges, the CREST center will provide a suite of data analytics algorithms, including computation modeling, data mining, and visualization tools. The computation-modeling component provides the computational and system support for diverse data analytics and decision-making tasks based on novel multi-tiered data analysis architecture. As data analysis tasks are computationally intensive, we will address system and architecture issues related to computational requirements for data gathering, data analysis, and decision-making (Liu et al., 2014; Ren and van der Schaar, 2013; Xu et al., 2012; Xu and Shatz, 2003).

Two areas of computing research will be developed. First area will focus on creating novel multi-tiered data analysis architecture, consisting of sensors and cloud/HPC computing systems. Second area will be to support data mining and visualization research components. The novelty of this architecture is in allowing data analysis to be conducted at different granularities and satisfy different timing requirements (from near-real-time, global static data analysis to real-time, local dynamic data analysis).

The data mining component proposes a comprehensive investigation on utilizing associations and correlations among the data to understand the characteristics and to extract semantics and patterns from the data (Chen et al., 2007; Ha et al., 2013; Ha et al., 2015; Lin et al., 2012; T. Meng and M.-L. Shyu, 2013; Shyu et al., 2005; Thompson 2005; Yang et al., 2014). This component will also develop the dimension reduction and information fusion algorithms to address the scalability and multi-source issues (Gehler and Nowozin, 2009; Ha et al., 2013; Liu et al., 2009; Yu and Liu, 2003).

Once large amounts of heterogeneous data are processed by the computation modeling component, it will be ready for easy access to useful information and pattern extraction that can assist quick evaluation and assessment of contaminants transport and fate. This will be support a collaborative effort among various stakeholders including scientists, local government and industry partners to create plausible and realistic scenarios for evaluating the risk and possible course of action for the South Florida Region.

The visualization component focuses on developing visual analytics and visualization algorithms to assist in assessment and strategic decision-making. This research will explore novel ways of displaying information visually, aggregating existing techniques into visual ensembles that are tailored for solving specific problems and providing the interactive means for users to work with these systems in their specific context (Li et al., 2009; Saleem et al., 2007; Zhang et al., 2006). In particular, we would like to create novel visualization and visual analytics methods for displaying complex data types, data aggregates, and analytic concepts at the border between humans and computing.

2. Research Plan

Our research plan is focused on developing new methodologies to support detection and evaluation of trends, analyzing contaminant transport, and creating visualization tools for querying data that allows for early intervention and restoration of the water and the ecosystems of South Florida. This plan will be conducted in four research thrusts as described below.

2.1 Multidimensional Data Analytics of Environmental and Molecular Biology Information

The research conducted by scientists in Subproject 1 requires characterization and measurement of a myriad of stressors associated with urban and agricultural landscapes. Using advanced methodologies they will collect vast amounts of data to determine the environmental exposure from trace analysis of critical pollutants such as nutrients, trace metals, DDT and PCBs to other biologically active compounds such as antibiotics and pharmaceuticals (e.g. endocrine disrupters), mercury, black carbon and fossil fuels (oil). These data will be of two types: 1) environmental data that are analog parameters; and 2) molecular biology information that have a digital signal.

The environmental data have identity (parameter), intensity above a threshold (signal or concentration), and potentially an environmental limitation benchmark as well as toxicological indicators. The parameter list could be as large as 100 to 200 items. The molecular biology data will be much larger and will include a "gene identity" (related to a function i.e., gene responsible for metal detoxification); and whether exposure to environmental stressors caused the gene to be under-expressed (-1), over-expressed (+1), or showed no change (0) when compared to the untreated group. These data could rank in the thousands depending on the generating methods. All these data can be represented and expanded on in a matrix such as the one shown below:

Sample ID	Temperature	Concentration	Toxicity effect	Location	Gene Signal	Epigenomic Markers	Other Parameters …
A1	Analog	Analog	Analog	Analog	Digital	Digital	Digital
A2	Analog	Analog	Analog	Analog	Digital	Digital	Digital
…	Analog	Analog	Analog	Analog	Digital	Digital	Digital
B1	Analog	Analog	Analog	Analog	Digital	Digital	Digital
…	Analog	Analog	Analog	Analog	Digital	Digital	Digital
C1	Analog	Analog	Analog	Analog	Digital	Digital	Digital

In Subproject 2, CACE scientists will examine the biogeochemical processes that govern the ultimate fate of these pollutants and their impacts on the environment**.** They will establish three transects that encompass the main transitions between agriculture, urban and natural landscapes. These transects will provide a common platform for detection and measurement for Subproject 1, sample water quality with a common experimental design, and application of advanced modeling techniques to couple flows with contaminants. The generated data from this process will provide a third dimension to the matrix provided above to inform our data analytics research.

***Low-rank matrix factorization*:**

The matrices generated from Subprojects 1 and 2 could involve low-dimensional structures (e.g., sparse or low-rank). In particular, the numbers of samples in these cases are typically far less than the total number of degrees of freedom (i.e. determined by the analog parameters and molecular biology information). The goal of our analysis is to understand the relationships between the samples and the parameters and the relationships among various parameters. Factor analysis (Child, 2006; Mulaik, 1972) or topic modeling (Blei et al., 2003) can be used to establish the relationships. In this project, we will jointly perform factor analysis and topic modeling while exploiting the low-dimensional matrix structure.

Compute the densely connected components:

In genomics, given a protein interaction network, it is often useful to compute the densely connected components as protein interaction modules. In these cases, the input data is the adjacency matrix *A* of an undirected graph with weights in {0;1}. We can formulate the problem as computing maximal cliques, although the rigorous definition of a clique is often unnecessary. We propose to solve the following optimization problem:

$$\max_X X^T AX, \;\; s.t. \; \sum_{i=1}^{n} x_i^{\alpha} = 1; \; x_i > 0$$

where $\alpha \in [1,2]$ is a parameter (Ding et al., 2008). The nonzero entries in the solution vector correspond to the vertices of the densely connected component we are seeking. It can be shown that: (1) a maximal clique is obtained when $\alpha = 1 + \varepsilon,\ 0 < \epsilon < 1$, while setting *Aii* = 1 (this enables us to generalize this approach to bipartite graphs). (2) When $\alpha = 1$, this formulation reduces to the Motzkin and Strauss formulation (Motzkin & Straus, 1965), where *Aii* = 0 is required. (3) As α goes close to 1, the sparsity of the solution increases steadily, reflecting the close relation between L_1 constraints and sparsity. At α = 2, the solution is given by the principal eigenvector of *A*.

2.2 Synthesis and Analysis of Gene Function Networks

Computational methods for gene functional prediction fall into two categories: direct annotation schemes, which infer the function based on the functional annotations of genes in its neighborhood in the network, and module-assisted schemes, which first identify modules of related genes and then annotate each module based on the known functions of its members. In this project, we are interested in computation methods of the second category (i.e., module-based methods). The key step of methods falling in this category is to identify biologically meaningful functional modules. Cluster analysis is a popular methodology for the extraction of function modules from genes and protein interaction networks since it has been observed by biologists that groups of highly interacting proteins could be involved in common biological processes (Spirin and Mirny, 2003).

However, the special characteristics of the data obtained from high-throughput experiments make the clustering task for identifying the function modules very challenging. These challenges include: (a) Poor data quality: The data obtained from the high-throughput experiments are quite noisy and contain many false positives. (b) Specific topological and network properties: The network structure from the data has been observed to have high clustering coefficients and modularity (Yook et al., 2004; Jeong et al., 2001; Ravasz et al., 2002).

A few genes/proteins in the network may have very large degrees while most others only have very few interactions. Clustering algorithms in this context need to pay attention to these topological and network properties. (c) Huge Volume: The datasets obtained from the experiments are of a large volume including tens of thousands of interactions among thousands of proteins even for a unicellular eukaryotic organism. Hence, clustering algorithms need to be fast and scalable. (d) Multi-functional: A gene/protein is often multi-functional and often involved with multiple modules. Hence clustering algorithms need to support “soft assignments”, i.e., assigning a protein into multiple groups.

Clustering algorithms for extracting function modules should address these challenges. Despite the significant progress that has been made in the area, existing clustering methods for extracting function modules are far from satisfactory due to the presence of noisy false positives, specific topological challenges, and the huge amount of data (Asur et al., 2007; Jaimovich et al., 2006). In addition, most of the existing methods do not support the “soft assignment” as they assign each gene/protein into a specific group.

In this project, we propose to develop ensemble-clustering methods for combining multiple, diverse and independent clustering results to improve the quality and robustness of identification. Different base clustering algorithms (e.g., spectral clustering algorithms and graph clustering algorithms) might have their own strengths and limitations. Ensemble clustering offers an appealing framework for taking advantage of the strengths of individual clustering algorithms and for improving the quality of identification.

<u>Base Clustering:</u>

In addition to the conventional clustering algorithms, e.g., network motifs, local cluster growing, graph-theoretic, and hierarchical clustering (King et al., 2004; Brun et al., 2004; Arnau et al., 2005; Dunn et al, 2005; Enright et al., 2002), we will also explore the use of spectral clustering algorithms with diverse yet informative topological and graph properties (e.g. edge-betweenness and clustering coefficients) as base clustering algorithms. Spectral clustering algorithms have well-motivated objective functions that can easily incorporate the graph properties and can be computed efficiently using mature scientific computing software tools.

A spectral clustering algorithm is obtained by recursively applying a spectral method for graph partitioning (Shi and Malik, 2000; Dhillon et al., 2007). Let *Q* denote the Laplacian matrix of a graph with weights w_{ij} on its edges <i; j >: the diagonal elements q_{ii} of Q is the sum of the weights of the edges incident on the vertex i, and for other elements, $q_{ij} = w_{ij}$. The partition problem is first modeled as the minimization of a quadratic program $p^T Qp$ over all partition vectors p, whose elements are either 1 or -1. The integer constraints can be relaxed and we can then solve the continuous version of the optimization problem over real vectors with components bounded in the interval [-1;+1]. The solution to the continuous optimization problem is obtained by computing the eigenvalues or singular values of $p^T Qp$. The eigenvector components provide a natural soft assignment since the values in the components reflect the degree of association between the vertices and the clusters.

<u>Ensemble Clustering:</u>

Ensemble clustering, also called aggregation of clustering, refers to the situation in which a number of different clustering results have been obtained for a particular dataset and it is desired to find a single (combined or consensus) clustering which is a better fit in some sense than the existing clustering results (Hu et al., 2006; Gionis et al., 2005). Empirical evidence has suggested that ensemble clustering can improve clustering robustness and discover useful cluster structures even if the data is quite noisy (Topchy et al., 2004).

However, there is a significant drawback in current ensemble clustering approaches (Strehl and Ghosh, 2002; Asur et al., 2007), i.e., all input clustering solutions are treated equally, despite the facts that: (1) different input clustering results could differ significantly, and (2) subsets of input clustering results could be highly correlated. As a result, when collecting a large number of input clustering results, quite often many clustering results could be close (similar) to each other. These would easily skew the final consensus clustering. Hence, simply applying current ensemble clustering for extracting protein function modules is inadequate.

In this Subproject, we propose the weighted ensemble clustering for extracting protein function modules. In (Li et al., 2007), we show that the ensemble-clustering problem can be efficiently solved within the nonnegative matrix factorization framework. Building on our previous work, the weighted ensemble clustering can also be formulated as an optimization problem. In weighed ensemble clustering, different input clustering results weigh differently, i.e., a weight for each input clustering is introduced, but the weights are automatically determined by an optimization process similar to a kernel matrix learning (Lanckriet et al., 2006). It should also be noted that the weights obtained in the weighted ensemble clustering could be useful for selecting input clustering. Clearly, an input clustering with larger weight contributes more to the final consensus clustering.

<u>Heterogeneous Data Integration:</u>

The data from heterogeneous data sources (e.g., gene expression and protein interactions) are useful for inferring gene functions (Bhardwaj and Lu, 2005; Jansen et al., 2002; Tu et al., 2006, Wang et al., 2012). Despite previous efforts in the integration of heterogeneous data, there is still a lot of room for improvements since the information enriched in each biological source has not been fully utilized.

The integration of different types of experimental data into an overall model is a critical and challenging task because of the vast difference in data type, dimension and quality (Shannon et al., 2003; Cline et al. 2007; Camargo and Azuaje, 2007). Two major problems must be addressed in order to integrate the heterogeneous data sources/types and extract the optimal conclusions from the combined data: (a) Data types must be unified or scaled in order to allow comparison and combination. For example, gene expression data is continuous and relative in nature while protein interaction data is pairwise and binary; (b) the data must be weighted or verified in a quantitative and consistent manner.

In this project, we will use the following three approaches for data integration.

1. Feature Integration: This approach enlarges the feature representation to incorporate all data and produces a unified feature space. In particular, continuous data types will be converted into discrete levels and categorical data type will be mapped into similar discrete levels. The data are then transformed into the same feature space and standard computational methods, such as prediction and clustering, can be performed. The advantage of feature integration is that the unified feature representation is often more informative and also allows many different data

mining methods to be applied and systematically compared. One disadvantage is the increased learning complexity and difficulty as the data dimension becomes large.

2. Semantic Integration: This approach keeps data intact in their separate original form. Computational methods are applied to each dataset separately. Results on different datasets are then combined by either voting (Carter et al., 2001), Bayesian averaging (Bishop, 2006), or the hierarchical expert system approach (Jordan and Jacobs, 2004). This approach seems to work reasonably well. One advantage of semantic integration is that it can implicitly learn the correlation structure between different sets of features (Li and Ogihara, 2005).

3. Intermediate Integration: This approach can be viewed as a compromise between the feature-level integration and the semantic integration. The data is kept in their original form and they are integrated at the similarity computation or the Kernel level (Lanckriet et al., 2006). For example, for protein p_i and p_j, their total pairwise similarity or affinity is $S_{ij} = A_{ij} + B_{ij}$, where A_{ij} is computed from gene expression profiles and B_{ij} is obtained from protein interaction. Standard computational methods can then be applied once the total similarity is computed.

We will carefully compare these integration methods in this project and explore their trade-offs through the design of suitable experiments.

2.3 Literature Mining and Curating Biological and Environmental Information

The biological and environmental literature databases provide knowledge warehouses to cross-reference experimental and analytical results with previously known biological facts, theories, and results. They can also be used to identify function commonalities of genes. In this Subproject, we propose to incorporate expert genetic knowledge for function discovery, instead of relying on purely empirical methods. Each domain expert only knows a few objects well. Hence relating the measurements in observation data with existing knowledge is a key part for data analysis. Mining on observation data alone may not be able to reveal the biological information and the pollutants impact. On the other hand, references on the literature will provide additional information. We will facilitate using text literature as a guide for detection, identification, and effect of chemical stressors in the ecosystem.

There are many literature databases publicly available. One good source is KEGG (Kyoto Encyclopedia of Genes and Genomes). This is a particularly high-quality data source, as it is curated by a knowledgeable team based on reported information in the scientific literature and is continuously updated. Text mining techniques can be applied to provide descriptive information from the literature.

There are two steps when performing text mining on a set of literature documents: (1) Document Pre-processing; (2) identification of text summaries with observation data. Document pre-process includes stripping unwanted characters/markup (e.g, HTML tags, punctuation, numbers, etc.), removing common stop words (e.g. a, the, it, etc.) and stemming keywords into “root” words etc. developing a synonyms list. Note that there are many words and phrases that refer to the same entity, hence a synonyms list will also be developed in pre-processing. After document preprocessing, step (2) is to thus provide meaningful textual summaries with the information that domain experts may be interested in. Therefore, techniques to perform text summarization will be studied (Mittal et al., 2000; Lin & Hovy, 2002).

In this project, we will investigate keyword search based algorithm and sentence extraction for literature summarization (Kankar et al., 2002; Masys et al., 2001; Jurafsky and Martin, 2008).

After we get the literature description, the remaining question is how we should combine the literature information with observation information. There are some existing approaches such as MedMiner (Tanabe et al., 1999) and PubGene (Jenssen et al., 2001). Medminer first performs clustering on observation data and then interprets textually while PubGene first performs clustering on textual data and then interprets numerically). These two techniques will be studied initially. However, both types of approaches ignore the correlation structure between different sources. We discuss the integration of different data sources in a separate section.

2.4 Data Visualization and Decision Making Support

Visualization of the dynamic variables affecting the hydrology, fate and contaminant transport in ecosystems of south Florida can be critical in understanding, identifying and acting upon valuable information produced by Subprojects 1 and 2. First, visualization can help the scientists in both groups to understand and interpret patterns in the data (Goldstein et al., 1994; Ward, 1994). For example, a scatter plot can help to identify patterns of significance from a large amount of monitoring data. To display categorical data in the matrices, the categorical values need to be mapped into numerical values. One challenge is to choose an effective mapping, as a random order may not be effective, as it tends to spread the data across the visual space. We will combine clustering and dimensionality reduction techniques for better visualization.

Second, by developing event relationship networks, a graphical representation of event correlation (Burns et al., 2001), we will enable the domain experts to easily review and understand information. Formally, an event relationship network is a directed graph where the vertices are events and the edges indicate the dependence relationships among the events. In addition, it also serves as a concise representation of the domain knowledge. We propose to develop techniques to construct, validate and complete event relationship networks using the discovered temporal patterns.

Finally visual data analysis, facilitated by interactive interfaces, enables the detection and validation of expected results while enabling unexpected discoveries in science (Hansen et al., 2008) The scientists in Subproject 1 will develop scenarios for evaluating the impact of water and land resources management decisions on the hydrology to determine the transport of contaminates and eventual ecological vulnerabilities. Our research and development of visualization tools, will support these scientists and decision makers to explore “what if” scenarios, define hypotheses, and examine data using multiple perspectives and assumptions on fate and transport of the contaminants (Hansen et al., 2008). Utilizing our research and tools they can identify coherent patterns and assess the reliability of their assumptions.

We will develop tools to visualize information including condition indexes, ecosystem maps, and genomic responses maps. We will design and develop techniques to bridge the gap between the application and intelligent techniques. Specifically, we will develop an interactive tool that can present patterns visually, in a way that is intuitive and easily understandable for the users. In addition, we will use the discovered patterns to evaluate and validate the relationships among samples and observations.

Software framework for synthesizing decision recommendations:

We propose to develop a framework for synthesizing decision recommendations to aid decision makers. This framework will leverage FIU-Miner: A Fast, Integrated, and User-Friendly System for Data Mining in Distributed Environment to develop system components (Zeng et al., 2013). FIU-Miner allows users to rapidly configure a complex data

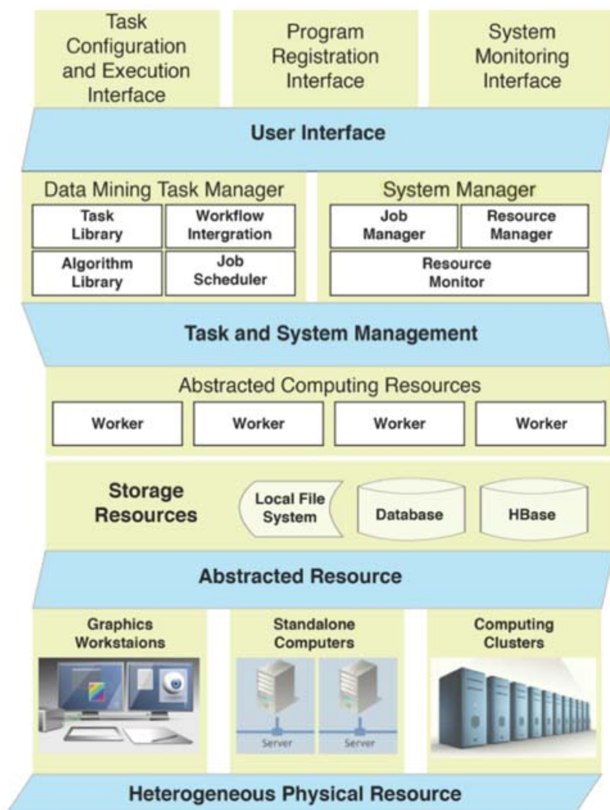


Figure 1: FIU-Miner System Architecture

analysis task without writing a single line of code. It also helps users conveniently import and integrate different analysis programs. Figure 1 shows the system architecture of FIU-Miner.

3. Broader Impacts

This Subproject addresses one of the significant scientific and engineering challenges by enabling a diverse group of environmental scientists to understand and make sense of "big data" ecological information. By developing new computational methodologies to support detection of pollutants, evaluation of trends, analyzing contaminant transport, and visualization of transient aquatic data, scientists and stakeholders can engage in early intervention and restoration of water in order to help the public live safely in their environment. The Subproject's literature mining research not only will help scientists in the other Subprojects more readily link their findings to those made by others, but also will facilitate the work of other environmental scientists in using text literature as tool for their research.

In addition, the computing research findings of FIU's CREST Center will be applicable to other scientific fields, as we will develop a novel multi-tiered data analysis architecture through data mining and visualization research. Developing this research is significantly important, as science is becoming increasingly data intensive with heavy reliance on using large datasets and visualization for discovery and problem solving. The developed software tools by this Subproject will be released to the open source community for further development and dissemination.

Subproject 3 References

Arnau, V., S. Mars, I. Marin. Iterative cluster analysis of protein interaction data. *Bioinformatics*, 21(3):364.378, 2005.

Asur, S., D. Ucar, S. Parthasarathy. An ensemble framework for clustering protein-protein interaction networks. In *Proceedings of the 15th Annual International Conference on Intelligent Systems (ISMB)*, 2007.

Bhardwaj, N., H. Lu. Correlation between gene expression profiles and protein-protein interactions within and across genomes. *Bioinformatics*, 21(11):2730.2738, 2005.

Bishop, C.M. *Pattern Recognition and Machine Learning*. Springer, 2006.

Blei, D.M., A.Y. Ng, M.I. Jordan. Latent Dirichlet Allocation. Journal of Machine Learning Research, Vo. 3, PP. 993-1022, Jan. 2003.

Brohee, S., J. van Helden. Evaluation of clustering algorithms for protein-protein interaction networks. *BMC Bioinformatics*, 7:488, 2006.

Brun, C., C. Herrmann, A. Gunoche. Clustering proteins from interaction networks for the prediction of cellular functions. *BMC Bioinformatics*, 5:95, 2004.

Burns, L., J. L. Hellerstein, S. Ma, C.-S Perng, D.A. Rabenhorst, D. Tayler. A systematic approach to discovering correlation rules for event management. In International Symposium on Integrated Network Management, 2001.

Camargo, A. F. Azuaje. Linking gene expression and functional network data in human heart failure. *PLoS ONE*, 2(12):e1347, 2007.

Carter, R., I. Dubchak, S. Holbrook. A computational approach to identify genes for functional rnas in genomic sequences. *Nucl. Acids Res*, 29:3928.3938, 2001.

Chen, M., S.-C. Chen, M.-L. Shyu. Hierarchical temporal association mining for video event detection in video databases. In Proceedings of the Second IEEE International Workshop on Multimedia Databases and Data Management, in conjunction with IEEE International Conference on Data Engineering, pp. 137-145, Istanbul, Turkey, April 15, 2007.

Child, D., The essentials of factor analysis, 3rd ed. New York, NY：Continuum Intl Pub Group, 2006.

Cline, M.S., M. Smoot, E. Cerami, A. Kuchinsky, N. Landys, et al. Integration of biological networks and gene expression data using cytoscape. *Nat Protoc.*, 2(10):2366.2382, 2007.

Dhillon, I.S., Y. Guan, S. Kulis. Weighted graph cuts without eigenvectors a multilevel approach. *IEEE Trans. Pattern Anal. Mach. Intell.*, 29(11):1944.1957, 2007.

Ding, C., T. Li, M.I. Jordan. Nonnegative Matrix Factorization for Combinatorial Optimization: Spectral Clustering, Graph Matching, and Clique Finding. In Proceedings of 2008 IEEE International Conference on Data Mining (ICDM 2008), Pages 183-192, 2008.

Dunn, R., F. Dudbridge, C.M. Sanderson. The use of edge-betweenness clustering to investigate biological function in protein interaction networks. *BMC Bioinformatics*, 6:39, 2005.

Enright, A.J., S.V. Dongen, C.A. Ouzounis. An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res*, 30(7):1575.1584, 2002.

GehlerP., S. Nowozin, “On Feature Combination for Multiclass Object Classification,” IEEE 12th International Conference on Computer Vision, 2009.

Gionis, A., H. Mannila, P. Tsaparas. Clustering aggregation. In *ICDE*, pages 341.352, 2005.

Goldstein, J., S.F. Roth, J. Kolojejchick, J. Mattis. A framework for knowledge-based, interactive data exploration. Journal of visual languages and computing, 5:339–363, 1994.

Ha, H.-Y., F. C. Fleites, S.-C. Chen, “Content-Based Multimedia Retrieval Using Feature Correlation Clustering and Fusion,” International Journal of Multimedia Data Engineering and Management (IJMDEM), Volume 4, No. 2, pp. 46-64, 2013.

Ha, H.-Y., S.-C. Chen, M. Chen, “FC-MST: Feature Correlation Maximum Spanning Tree for Multimedia Concept Classification,” Ninth IEEE International Conference on Semantic Computing, Anaheim, California, USA, pp. 276-283, February 7-9, 2015.

Hansen, C., C. Johnson, V. Pascuuchi, S. Claudio (2009). Visualization for Data-Intensive Science in The Fourth Paradigm: Data-Intensive Scientific Discovery, Microsoft Research, Redmond, Washington

Hu, X., I. Yoo, X. Zhang, P. Nanavati, D. Das. Wavelet transformation and cluster ensemble for gene expression analysis. *International Journal of Bioinformatics Research and Application*, 1(4):447.460, 2006

Jaimovich, A., G. Elidan, H. Margalit, N. Friedman. (2006). Towards an integrated protein-protein interaction network: a relational Markov network approach. Journal of Computational Biology. 13(2):145-64.

Jansen, R., D. Greenbaum, M. Gerstein. Relating whole-genome expression data with protein-protein interactions. *Genome Res.*, 12:37.46, 2002.

Jenssen, T.-K., A. Lagreid, J. Komorowski, E. Hovig (2001), A Literature network of human genes for high-throughput analysis of gene expression. Nature Genetics, 2001 28: 21-28.

Jeong, H., S.P. Mason, A.L. A.L. Barabasi, Z.N. Oltvai. Lethality and centrality in protein networks. *Nature*, 411(6833):41.42, 2001.

Jordan, M.I., R.A. Jacobs. Hierarchical mixtures of experts and the em algorithm. *Neural Computation*, 6:181.214, 1994.

Jurafsky, D., J. H. Martin. Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition. Pearson & Prentice Hall, 2008.

Kankar, P., S. Adak, A. Sarkar, K. Murari G. Sharma (2002), Medmesh Summarizer: Text Mining For Gene Clusters. In Proceedings of the Second SIAM International Conference on Data Mining, 2002.

King, A.D., N. Przulj, I. Jurisica. Protein complex prediction via cost-based clustering. *Bioinformatics*, 20(17):3013.3020, 2004.

Lanckriet, G.R., N. Cristianini, P.L. Bartlett, L.E. Ghaoui, M.I. Jordan. Learning the kernel matrix with semi-de_nite programming. In *Proceedings of International Conference on Machine Learning (ICML)*, pages 323.330, 2006.

Li, T., C. Ding, M. I. Jordan. Solving consensus and semi-supervised clustering problems using nonnegative matrix factorization. In *Proceedings of 2007 IEEE International Conference on Data Mining (ICDM 2007)*, 2007.

Li, T., M. Ogihara (2005). Semi-supervised learning from different information sources. Knowl. Inf. Syst. 7(3): 289-309.

Li, Y., K. Chatterjee, S.-C. Chen, K. Zhang, “A 3-D Traffic Animation System with Storm Surge Response,” IEEE International Symposium on Multimedia, San Diego, California, USA, pp. 257-262, December 14-16, 2009.

Lin, C.-Y., E. Hovy. From single to multi-document summarization: a prototype system and its evaluation. In ACL 02: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, 2002.

Lin, L., M.-L. Shyu, S.-C. Chen, “Association Rule Mining with a Correlation-based Interestingness Measure for Video Semantic Concept Detection," International Journal of Information and Decision Sciences*,* Vol. 4, Nos. 2/3, pp. 199-216, 2012.

Liu, D., S. Hua, Z. Ou, “IR and Visible-light Face Recognition using Canonical Correlation Analysis,” Journal of Computational Information Systems, 5(1), pp. 291-297, 2009.

Liu, J., Y. Liu, Z. Du, T. Li, “GPU-Assisted Hybrid Network Traffic Model,” 2014 ACM SIGSIM Conference on Principles of Advanced Discrete Simulation, Denver, Colorado, May 18-21, 2014.

Masys, D.R., J.B. Welsh, J.L. Fink, M. Gribskov, I. Klacansky J. Corbell (2001), Use Of Keywords Hierarchies To Interpret Gene Expression Patterns. Bioinformatics, 17(4): 319-326.

Meng, T., M.-L. Shyu, “Concept-Concept Association Information Integration and Multi-Model Collaboration for Multimedia Semantic Concept Detection,” International Journal of Information Systems Frontiers*,* pp. 1-13, April 2013.

Mittal, V., J. Carbonell Goldstein, J. M. Kantrowitz. Multi-document summa- rization by sentenceextraction. In NAACL-ANLP 2000 Workshop on Automatic summarization, 2000.

Motzkin, T.S., E.G. Straus. Maxima for graphs and a new proof of a theorem of turan. Canad. J. Math., 17:533-540, 1965.

S. Mulaik, S., The foundations of factor analysis. McGraw-Hill New York, 1972.

Ravasz, E.E., A.L. Somera, D.A. Mongru, Z.N. Oltvai, A.L. Barabasi. Hierarchical organization of modularity in metabolic networks. *Science*, 297(5586):1551.1555, 2002.

Ren, S., M. van der Schaar, “Efficient Resource Provisioning and Rate Allocation for Stream Mining in a Community Cloud,” IEEE Transactions on Multimedia, vol. 15, no. 4, pp. 723-734, Jun. 2013.

Saleem, K., S.-C. Chen, K. Zhang, “Animating Tree Branch Breaking and Flying Effects for a 3D Interactive Visualization System for Hurricanes and Storm Surge Flooding,” the Third IEEE International Workshop on Multimedia Information Processing and Retrieval, in conjunction with IEEE International Symposium on Multimedia, pp. 335-340, Taichung, Taiwan, R.O.C. , December 10-12, 2007.

Shannon, P., A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*, 13(11):2498.2504, 2003.

Shi, J., J. Malik. Normalized cuts and image segmentation. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 22(8):888.905, 2000.

Shyu, M.-L., I.P. Kuruppu-Appuhamilage, S.-C. Chen, L. Chang. Handling missing values via decomposition of the conditioned set. In Proceedings of the 2005 IEEE International Conference on Information Reuse and Integration, pp. 199-204, August 15-17, 2005, Las Vegas, Nevada, USA.

Spirin, V., L.A. Mirny. Protein complexes and functional modules in molecular networks. *Proceedings of the National Academy of Sciences*, 100:12123.12128, 2003.

Strehl, A., J. Ghosh. Cluster ensembles - a knowledge reuse framework for combining multiple partitions. *Journal on Machine Learning Research (JMLR)*, 3:583.617, December 2002.

Tanabe, L., L.H. Smith, J.K. Lee, U. Scherf, L. Hunter, J.N. Weinstein (1999). MedMiner: An internet tool for filtering and organizing biomedical information, with application to gene expression profiling. BioTechniques. 1999; 27: 1210-1217.

Thompson, B., “Canonical Correlation Analysis,” Encyclopedia of statistics in behavioral science, 2005.

Topchy, A.P., M. Law, A.K. Jain, A.L. Fred. Analysis of consensus partition in cluster ensemble. In *Proceedings of International Conference on Data Mining*, pages 225.232, 2004.

Tu, K., H. Yu, Y.X. Li. Combining gene expression profiles and protein-protein interaction data to infer gene functions. *Journal of Biotechnology*, 124:475.485, 2006.

Wang, D., M. Ogihara, E. Zeng, T. Li. Combining Gene Expression Profiles and Protein-Protein Interactions for Identifying Functional Modules, In Proceedings of 11th International Conference on Machine Learning and Applications (ICMLA 2012), pages 114-119, 2012.

Ward, M.O, Xmdvtool: Integrating multiple methods for visualizing multivariate data. In Proceedings of Visualization, 1994.

Xu, H., S. M. Shatz, "A Framework for Model-based Design of Agent-oriented Software," IEEE Transactions on Software Engineering, pp. 15–30, 2003.

Xu, Y., D. Arteaga, M. Zhao, Y. Liu, R. Figueiredo, S. Seelam, "vPFS: Bandwidth Virtualization of Parallel Storage Systems," the 28th IEEE Conference on Massive Data Storage, April 2012.

Yang, Y., H.-Y. Ha, F. C. Fleites, S.-C. Chen, "A Multimedia Semantic Retrieval Mobile System Based on Hidden Coherent Feature Groups," IEEE Multimedia, Volume 21, Number 1, pp. 36-46, January-March, 2014.

Yook, S.H., Z.N. Oltvai, A.L. Barabasi. Functional and topological characterization of protein interaction networks. *Proteomics*, 2004.

Yu, L., H. Liu, "Feature Selection for High-dimensional Data: A Fast Correlation-based Filter Solution," Twentieth International Conference on Machine Learning, 2003.

Zeng, C., Y. Jiang, L. Zheng, J. Li, L. Li, H. Li, C. Shen, W. Zhou, T. Li, B. Duan, M. Lei, P. Wang. "FIU-Miner: A fast, integrated, and user-friendly system for data mining in distributed environment". In Proceedings of the 19th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD'13), pages 1506-1509, 2013.

Zhang, K., S.-C. Chen, P. Singh, K. Saleem, N. Zhao, "A 3D Visualization System for Hurricane Storm Surge Flooding," IEEE Computer Graphics and Applications, Vol. 26, Issue 1, pp. 18-25, Jan.-Feb. 2006.

Zhang, Y., T. Li, "DClusterE: A Framework for Evaluating and Understanding Document Clustering Using Visualization**,"** ACM Transactions on Intelligent Systems and Technology**, 3**(2): 24, 2012.

Zheng, L., T. Li, C. Ding, "Hierarchical Ensemble Clustering," IEEE International Conference on Data Mining, pp. 1199-1204, 2010.