California Water Quality Monitoring Council

Data Management and Integration Possibilities

December 7th, 2009

CALIFORNIA ENVIRONMENTAL PROTECTION AGENCY

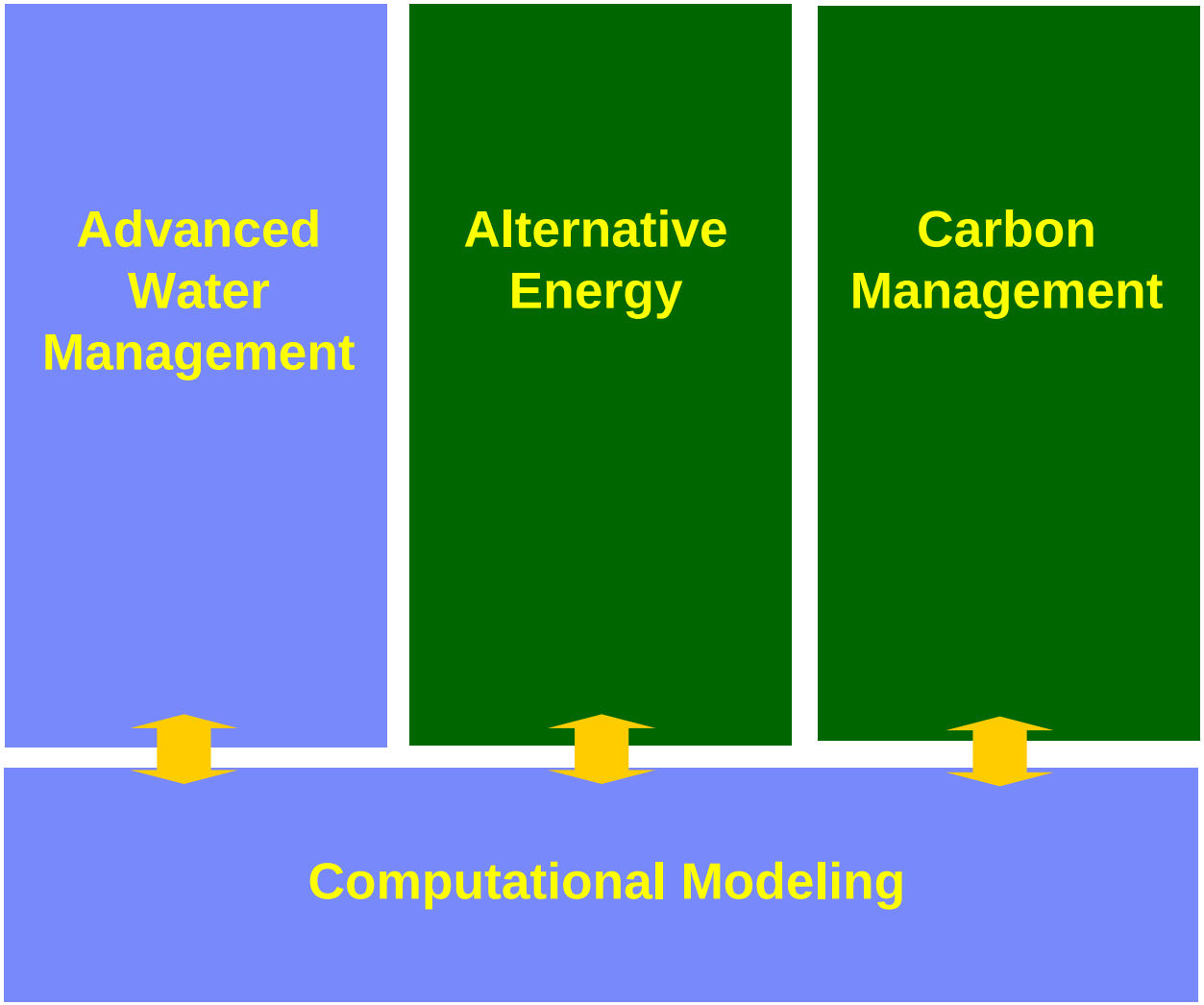
CALIFORNIA natural resources AGENCY

Introduction

- Intent:
  - To describe examples of IBM's portal creation, and water data management activities
  - To enable a discussion about data management issues relating to water (including habitat) quality management in CA
- Contents
  - IBM and Water Management
  - Example - Water For Tomorrow
  - Example - Smart Bay

IBM and Water Management

“Big Green Innovations”



Water management: improving information flow & use

"One barrier to better management of water resources is simply ***lack of data …. The water problem is largely an information problem***".[1]

"The effectiveness of [water management] decisions ***depends on the quality of information*** …In addition to improved water data the United States should ***develop and expand …forecasting and predictive models and systems…*** " [2]

- **Is our water safe to drink?**
- **Is it safe to swim in our waters?**
- **Is it safe to eat fish and shellfish from our waters?**
- **Are our aquatic ecosystems healthy?**
- **What stressors and processes affect our water quality?**

1. Source: Wired Magazine, "Peak Water: Aquifers and Rivers Are Running Dry. How Three Regions Are Coping", Matthew Power, April 21st, 2008

***2. Source*:** NSTC, "A Strategy For Federal Science And Technology To Support Water Availability And Quality In The United States, - Report Of The National Science And Technology Council Committee On Environment And Natural Resources Subcommittee on Water Availability and Quality", September 2007

Data "pathologies" - and a prediction

- **No data**
- Data is in the **wrong scale** (spatial or temporal) for the decision
- Data is **fragmented** between different stakeholders:
  - Different formats, scales, frequencies, standards
  - Re-capture of data many times
- **Too much data** to use
- **Incompatible or incomplete models** mean that data is not leveraged
- **Poor visualization** of information impedes effective decision-making
  - "So what's this telling us?" syndrome

- *Prediction:*
  - *"Unless we solve these data problems, some percentage of whatever we invest in water management in California (let's say, 30-50%) will be wasted.*
  - *But we won't know which 30-50% until something major breaks".*

The skills and assets IBM brings to water management

- World class skills in:
  - Data management and integration
  - Computational model creation and integration
- Portal creation and management
- “System S” - streaming data, live modeling
- Cognos and ILOG analysis & visualization
- Smart grid technologies and expertise
- Award-winning water and energy management in our semiconductor plants
- Maximo asset management
- Green Sigma™ process improvement
- “Deep Thunder” weather forecasting

- **In partnership with:**
  - Civil & environmental engineering companies
  - Sensor companies
  - Application vendors (eg ESRI, Derceto)
  - Universities
  - DOE and other research labs
  - NGOs
  - Others

Our Almaden, CA, lab undertakes much of our research into data management and integration technologies

- Unstructured text search/mining
- Multi-modal data mining - extraction and correlation across multiple records
- Adding higher level abstraction and query methods to Apache Hadoop (for distributed applications)
- Data cleansing technologies
- “System T” - information extraction from structured and unstructured data
- Visualization techniques
- Extreme database compression for faster retrievals
- … and more…

IBM’s Almaden lab near San Jose - original home of the hard disk drive, relational database, and many other innovations

We work at - and integrate between - three “scales”

Natural scale

- Water resource mapping and availability
- Water quality monitoring and management (surface and subsurface)
- Land use analysis
- Extraction monitoring (surface and subsurface)
- Flood control

Utility scale

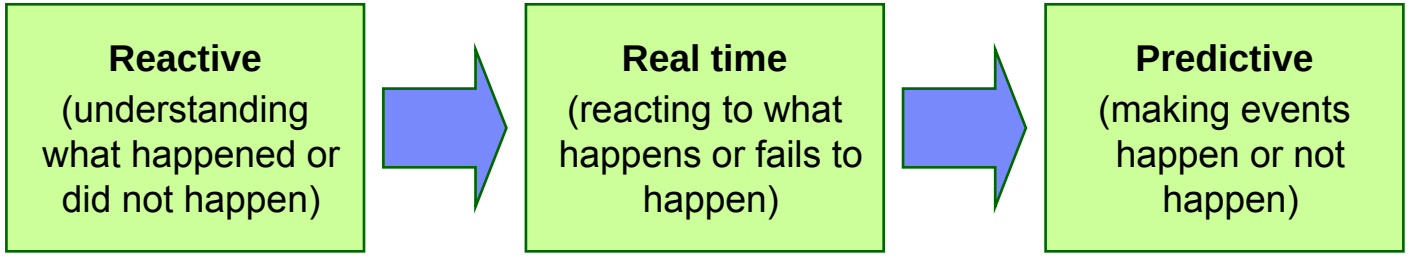
- Water quality and usage
- Discharge, combined sewer overflow
- Asset management
- “Smart levees” and levee monitoring systems
- Weather event assimilation
- Energy management

Enterprise Scale

- Water usage tracking
- Water quality control (into and within plants, discharges)
- Supply chain optimization
- Energy management
- Business process improvements
- Metrics and management

Water quality - transition to predictive management?

- We are detecting an interest among advanced water agencies world-wide to move to a predictive management decision paradigm, by using sophisticated models to anticipate the likelihood of quality issues and enable reactions in advance.



- We would be interested to discuss with CA Water Quality Monitoring Council its perspective on predictive management - for example, predicting at-risk times and locations for excess nutrient enrichment, water warming linked to low flows, and so on.
  - Also, intended approaches to issues such as pathogen sensing, “emergent” contaminants such as estrogen, pharmaceuticals, etc.

Example - Water For Tomorrow

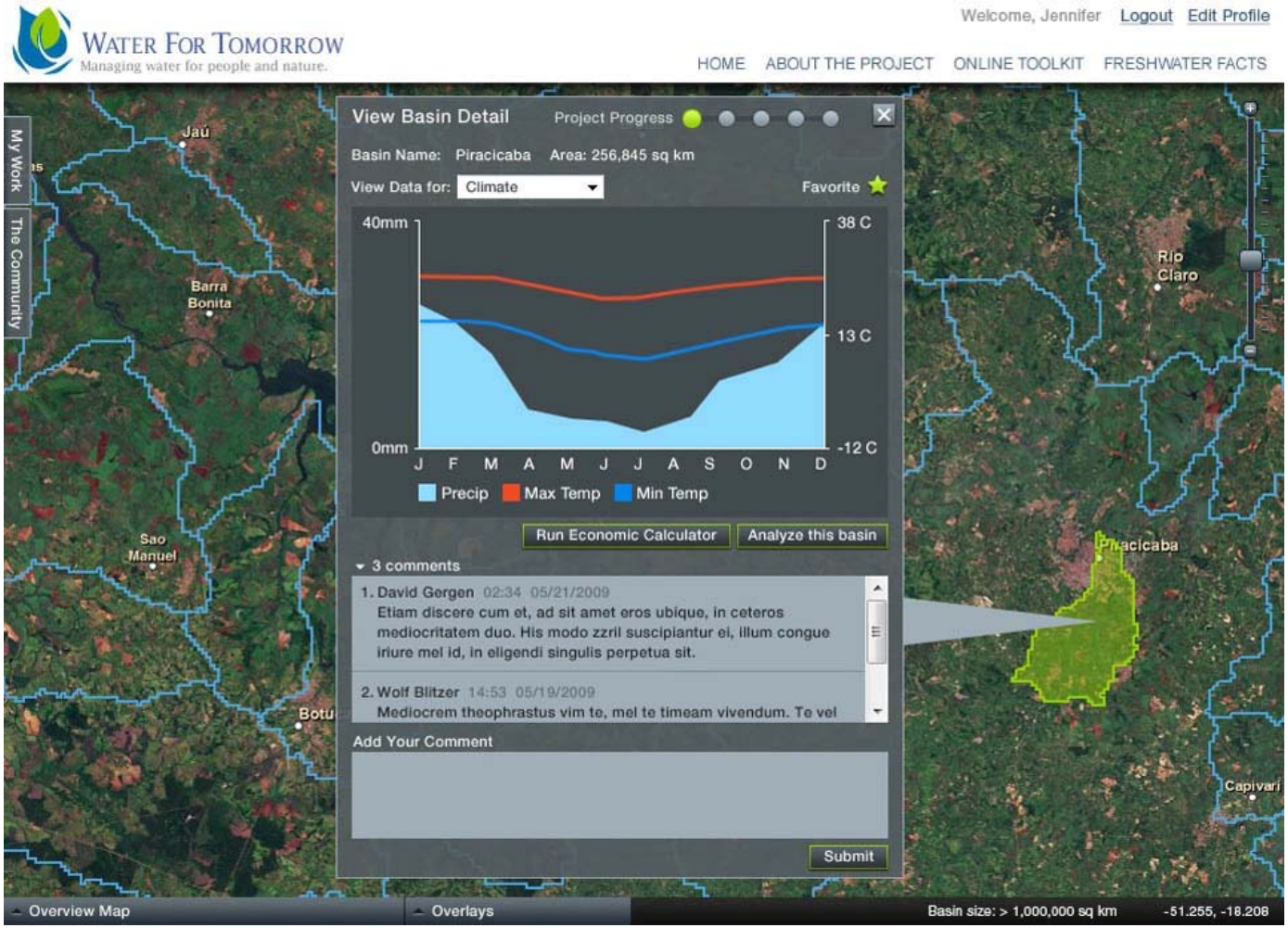
Water for Tomorrow - Overview

The Nature Conservancy

- Partnership between IBM, The Nature Conservancy, plus academic and public agency partners
- Joint effort to create state-of-the-art technology and collaboration tools to support conservation of the world's freshwater ecosystems
  - Combines rich graphics and dynamic mapping capabilities
- Helps planners and scientists analyze river basins and visualize the effects of different management scenarios on the overall basin health
  - In so doing, initiates collaboration to develop sustainable water management policies
- Grants free and open access to datasets and models to better support water planning science and management

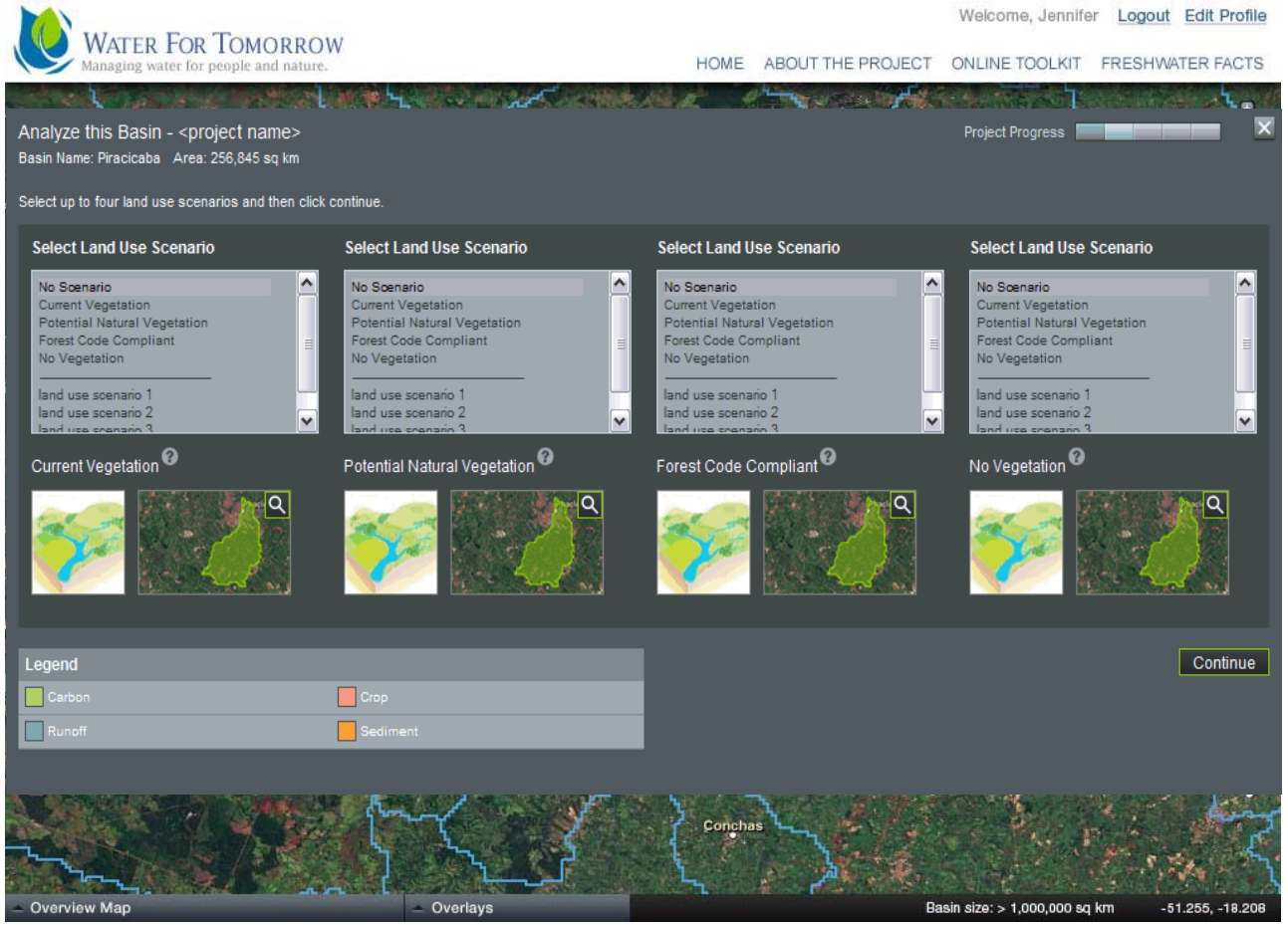
Basin Selection

- Basins selected by:
  - Use of zoom and pan controls on the interactive map
  - Creating a new basin by delineating areas on the map
  - Uploading a basin file
- This basin is part of the Piracicaba, Capivari and Jundiaí rivers (PCJ) pilot site in Brazil



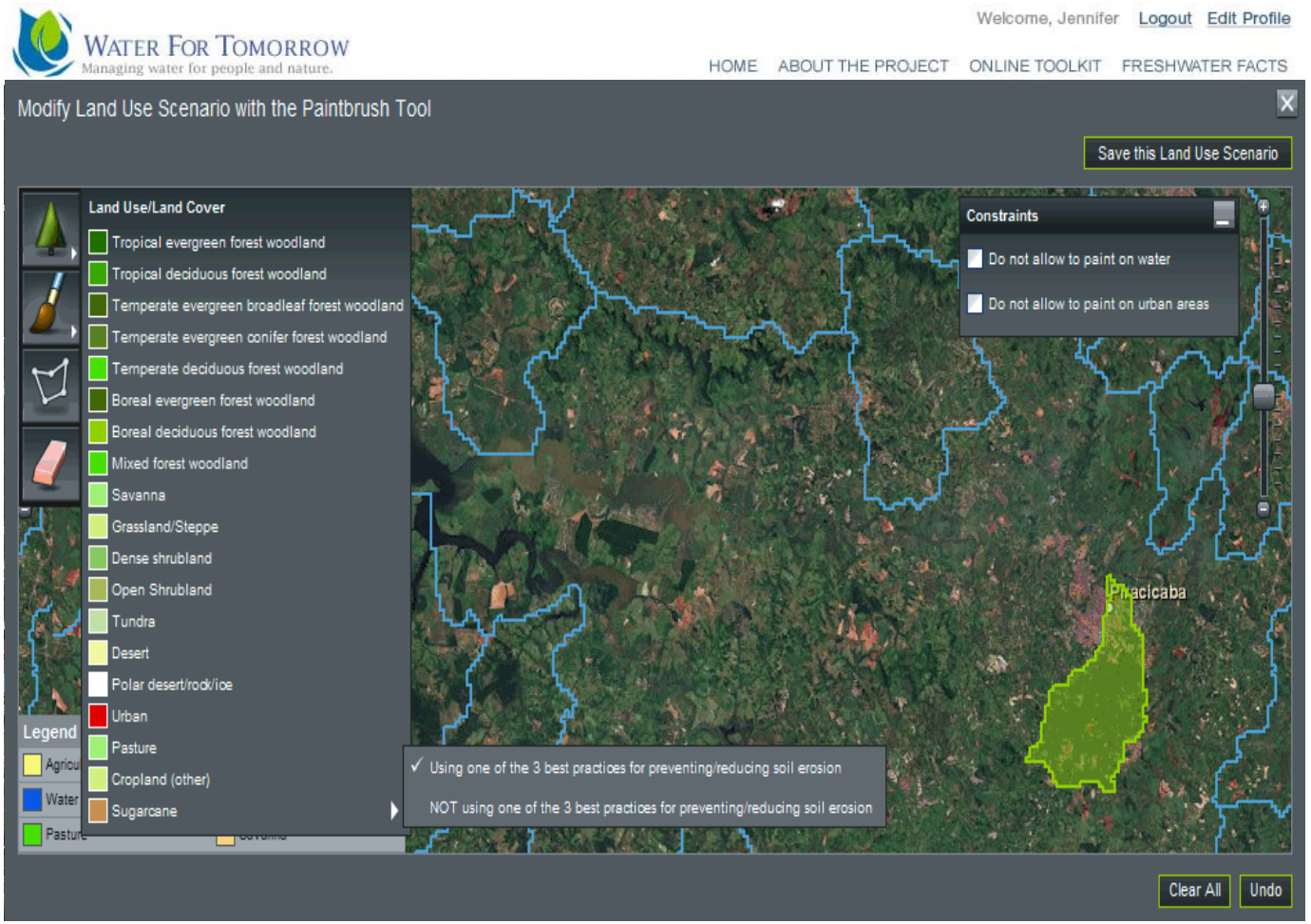
Scenario Selection

- A land use scenario is a set of land use practices. Scenarios are used by the WFT hydrology model to compute crop production, water quality, water balance and other variables
- Pre-canned scenarios are provided in the tool, or custom scenarios can be created or imported



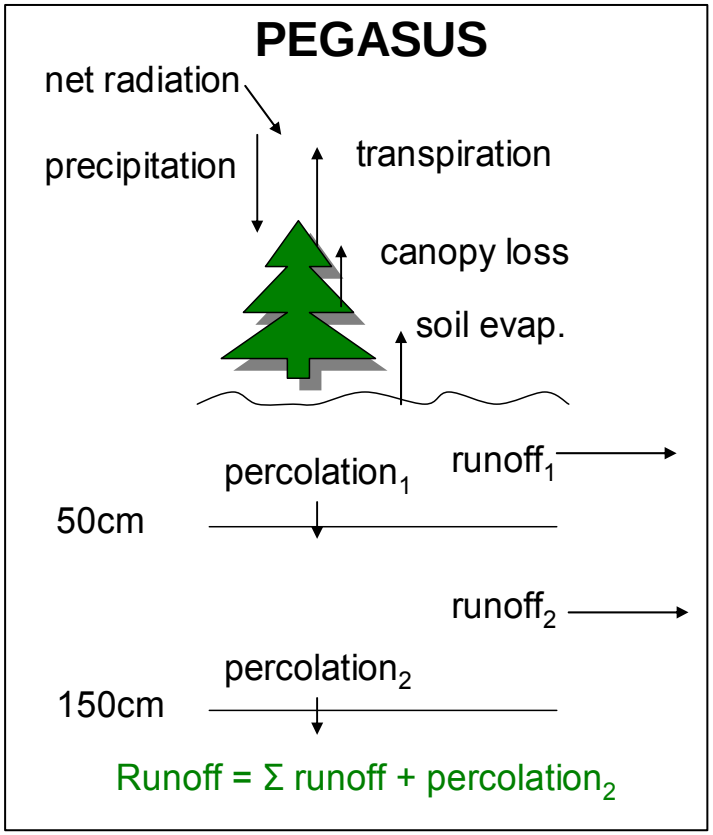
Scenarios Modification – Paintbrush Tools

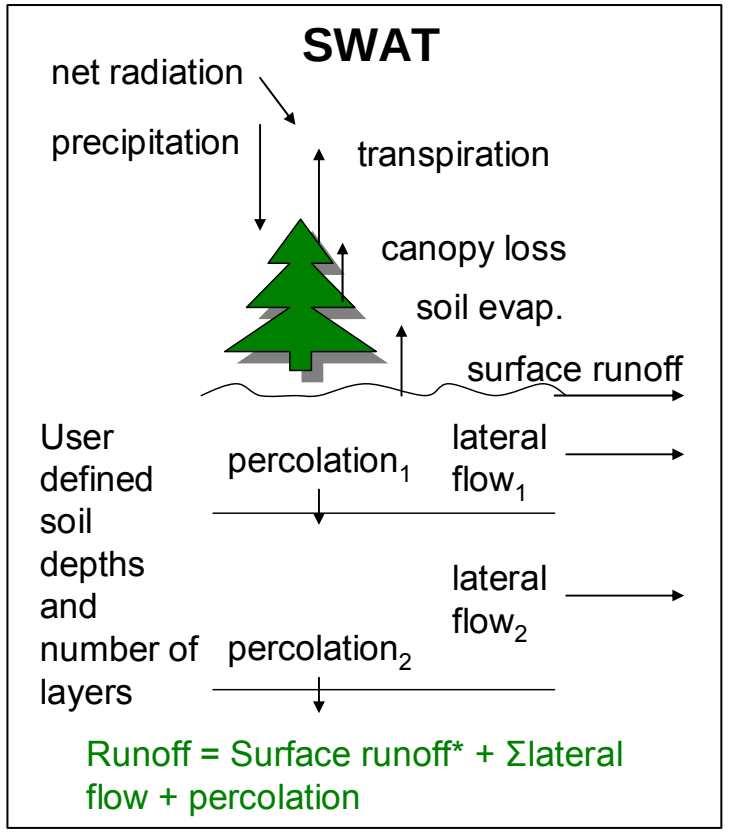
- Scenarios can be modified by using simple paintbrush tools



Model Integration

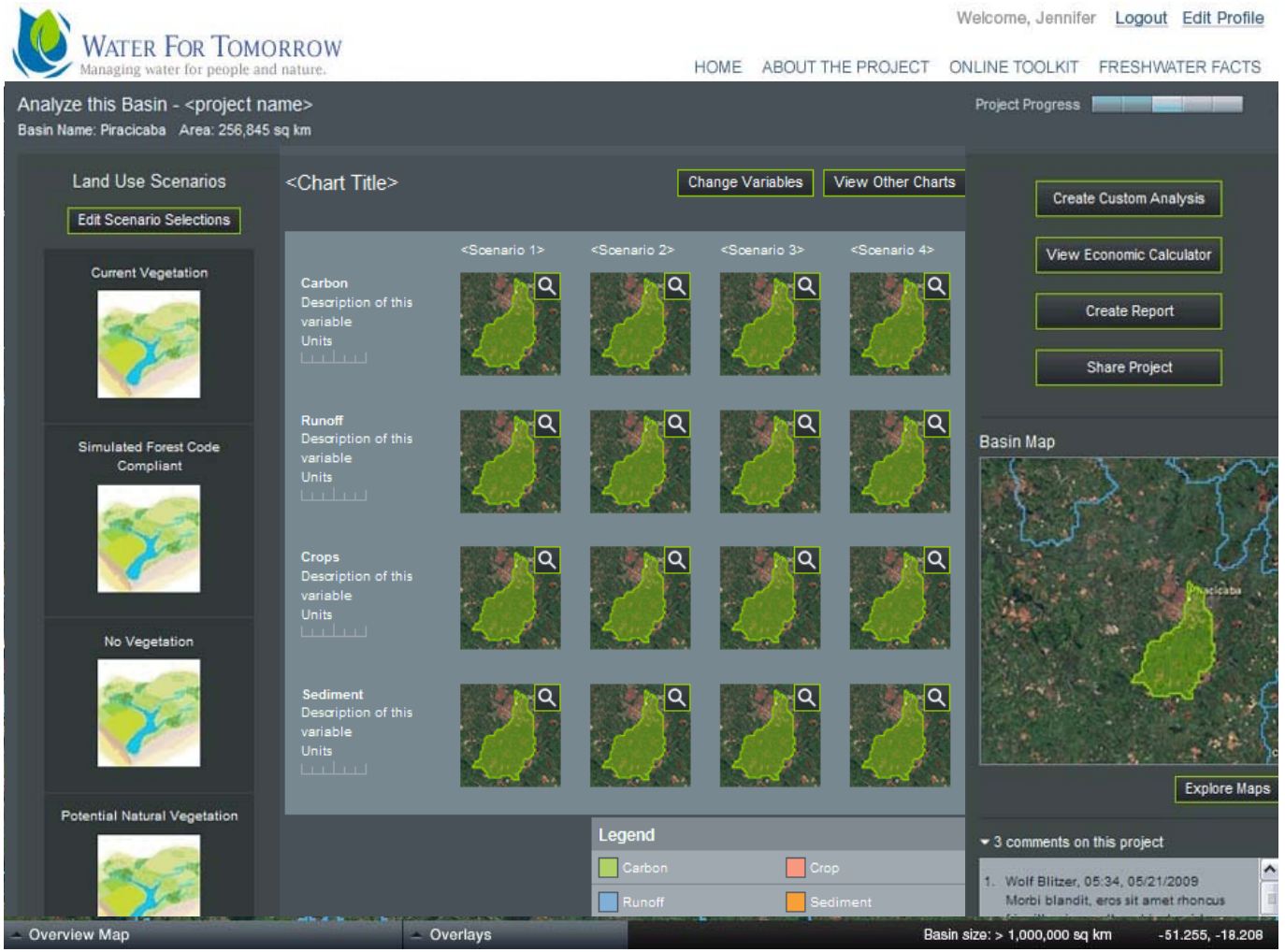
- WFT required extensive semantic harmonization between different hydrological models to ensure 100% valid results across all scenarios





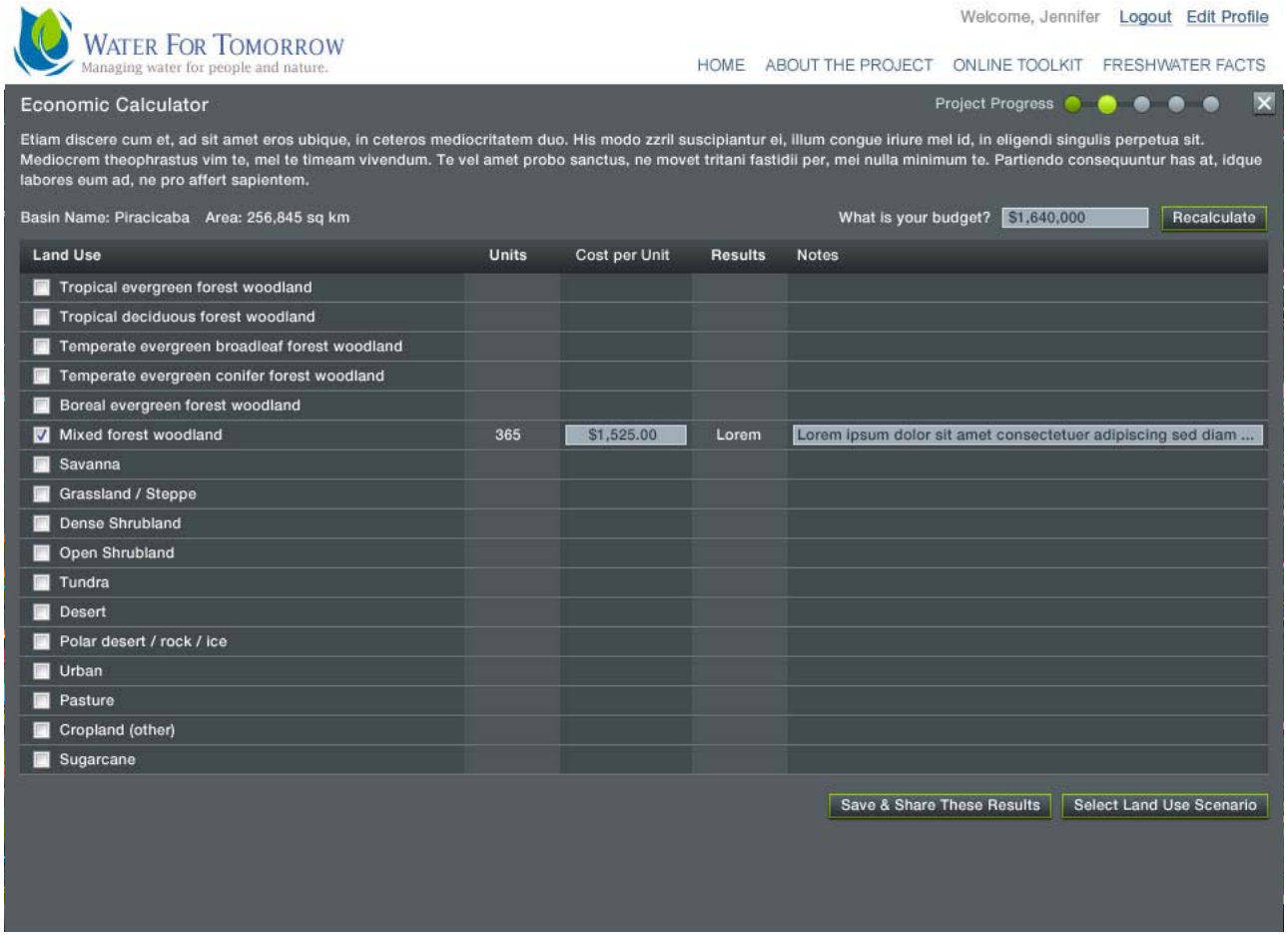
Interactive Visualizations

- Multiple visualization methods are offered – charts, graphs, maps etc
- The aim is to provide a "workbench" - analyses can be edited, saved, commented on and shared with the user community or pre-defined work groups

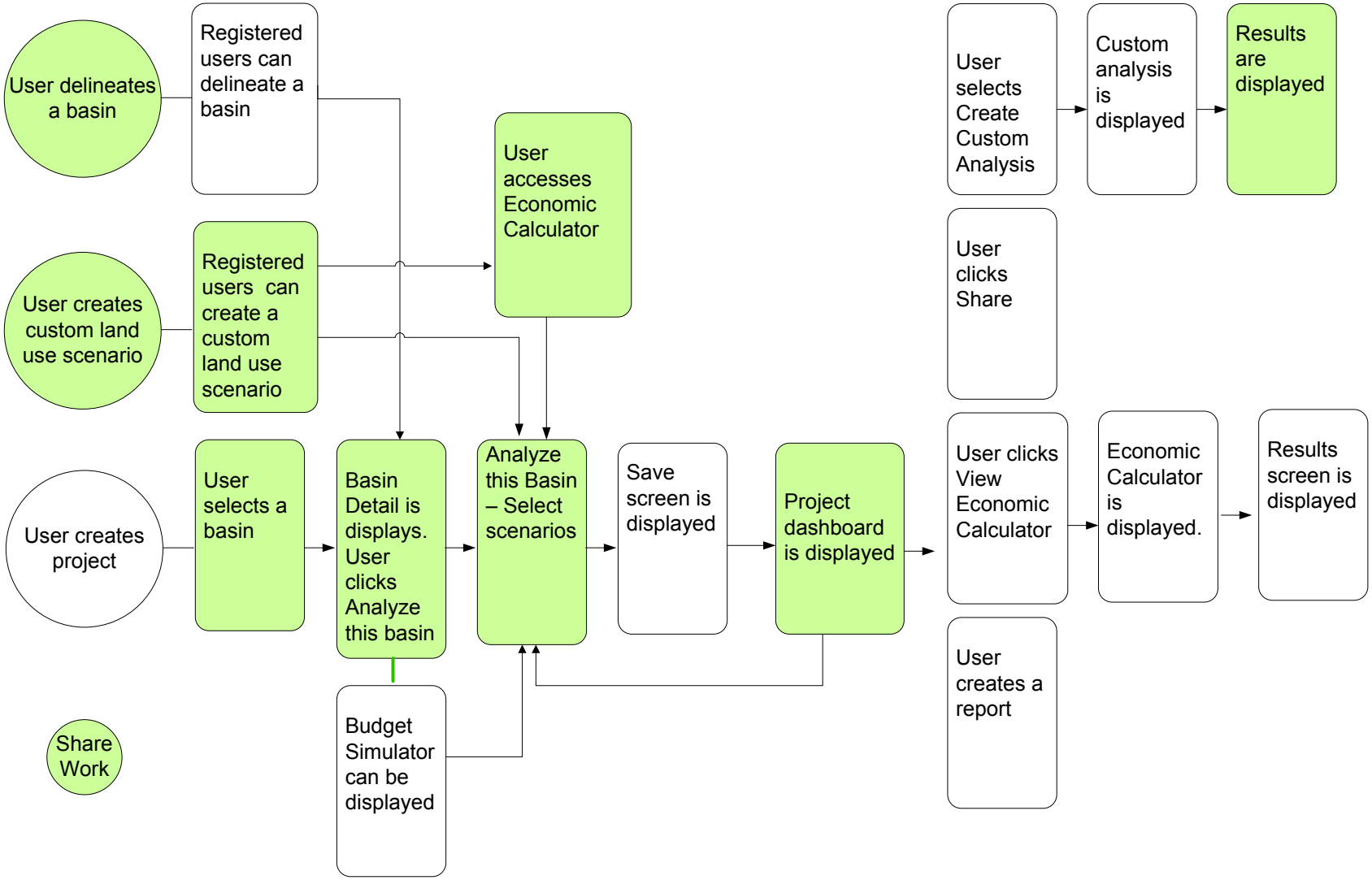


Budget Simulator

- Economic budget simulations can be created for the selected basin
- These enable economic cost-benefit analysis of current policies vs ecosystem remediation vs. restoration



Water for Tomorrow Toolkit: High-Level Flow



Example - Smart Bay

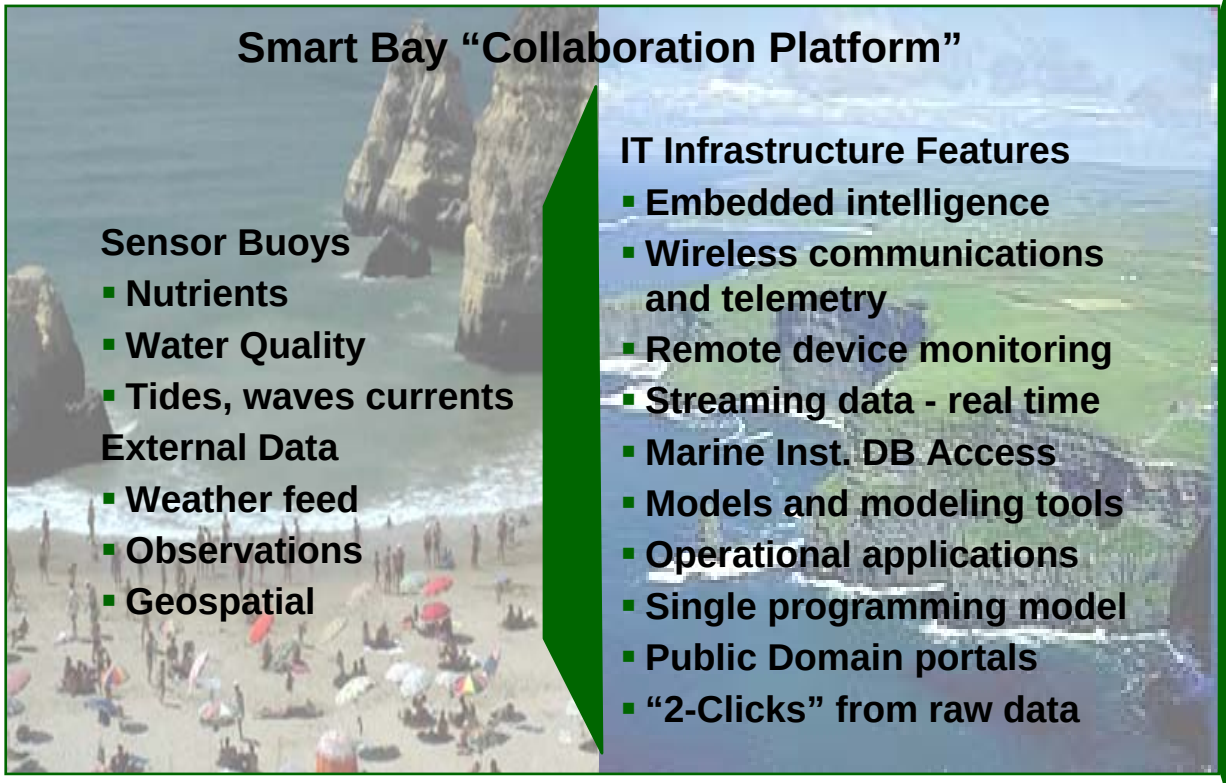
Smart Bay - Objectives

Marine Institute
Fiotas na Mara

- Design, test, and implement the next generation water management system (marine and freshwater) for the Marine Institute of Ireland via a collaborative approach
- "Fast-path" a number of new technologies for water management and advanced cyber-physical systems
- Build a "collaboration platform" that enables stakeholders to Galway Bay to work together
  - Integrate scientific and operational management concerns and perspectives

IDA IRELAND DCU Tyndall National Institute UCD DUBLIN sfi science foundation ireland fondúireacht eolaíochta éireann

One “collaboration platform” supports multiple stakeholders, uses and perspectives



Policy
- WFD Europe
- Fisheries protection

Weather
- Predictive local alerts

Tourism
- Volvo yacht race
- Beach management
- Pleasure craft
- LA infrastructure support

Fisheries
- Floating Hazards
- Biohazard monitoring
- Testing aquaculture

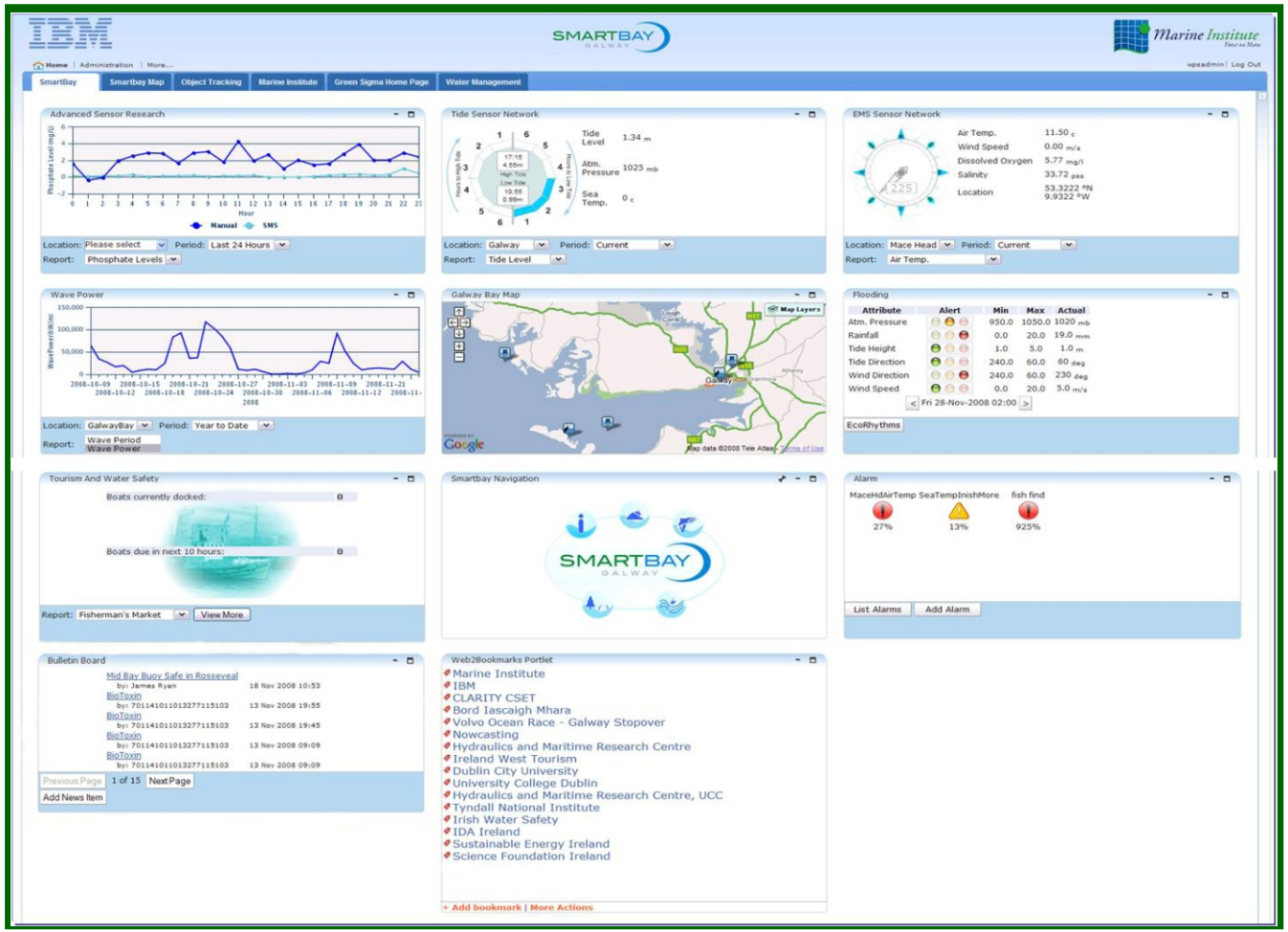
Energy
- Testing wave energy devices

Safety
- Flood alerts and management
- Water quality
- Water conditions
- Disaster management

Science and Education
- Primary, secondary, tertiary
- Research

Smart Bay User Portal and "Portlets"

- Consistent, unified user interface for diverse set of stakeholders
  - More than a dashboard
  - Analytics
  - Applications
- Each user has own "portlet" accessing single overall data set
- Drill down capabilities in all cases



Advanced capabilities for sensors/monitoring platforms – supporting new sensor technology development

- The next generation intelligent sensor platform
  - Real-time decision making at the measurement/monitoring source
  - Improved data quality and consistency
  - High-integrity communications (no critical data loss should communications fail)
  - Advanced remote device management
    - Ability to remotely provision sensor platform
    - Start/stop/load/unload software remotely

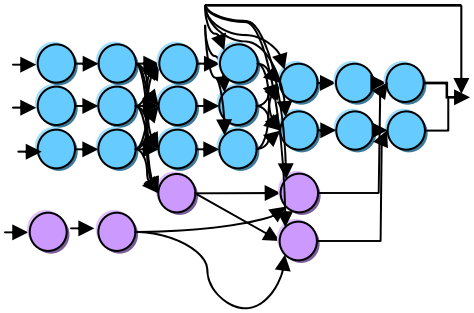
Sensor types included in the Smart Bay platform

- **Water Quality, including:**
  - PCO2 (measure of CO2 and H2CO3 - reflects alkalinity etc)
  - Colored dissolved organic matter
  - Nitrate, phosphate
  - Temperature, conductivity, salinity, dissolved oxygen, chlorophyll fl, and turbidity - by depth
- **Meteorology, including**:
  - Atmospheric pressure
  - Wind speed, max gust
  - Wind direction
  - Air temperature
  - Relative humidity
  - Rain intensity, hail intensity
- **Oceanographic, including:**
  - Spectral harmonic wave detail
  - Sig. wave height, wave period
  - Max wave height, wave period
  - Mean Direction

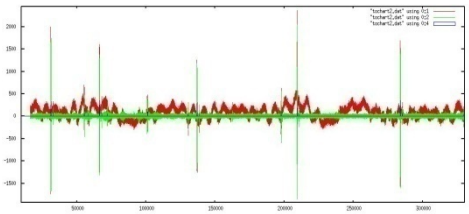
- Smart Bay offers the ability to manage and integrate these data streams to explore complex, multi-dimensional questions and issues

Current and planned activities

- **Stream analytics**
  - Real-time analytics for acoustic hydrophone data
  - Cetaceans - species i/d, protection, population measurement/projection, monitoring/ tracking
  - Also processing offline raw data
  - Live demonstrator for buoy mounted hydrophones with streaming telemetry in 4Q 09
  - Examining other applications for distributed analytics: undersea and research vessel-based processing of high bandwidth data including video streams



- **Data Federation**
  - Demonstrator to include new capabilities for data federation across Marine Institute - advanced tooling
  - Leveraging existing metadata efforts currently underway at MI
  - Incorporation of multilayered geospatial data
  - Modeling data



Benefits of Smart Bay

- SmartBay provides an open and extensible foundation for:
  - Improved exploitation of the Marine Institute's investment in data, information, and knowledge - supporting mixed research and operational uses
  - Enhanced support for the Marine Institute's industry sector enablement
    - Ability to add significant new functionality and expand support
  - Improved operational consistency with a fully integrated, robust and high-integrity infrastructure
  - Improved monitoring and management of natural assets with better quality, more timely data
  - More effective physical asset protection and management
  - Enhanced sharing of data across functional areas
    - Data standardization, data fusion, data federation
    - Ubiquitous accessibility and computing
  - New level of agility and responsiveness to changing environmental conditions

Thank you!

peter.r.williams@us.ibm.com