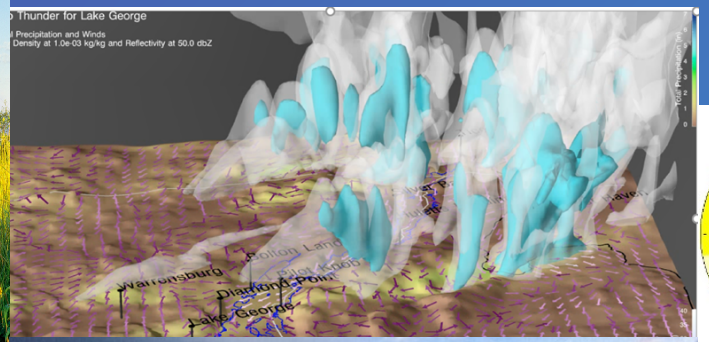
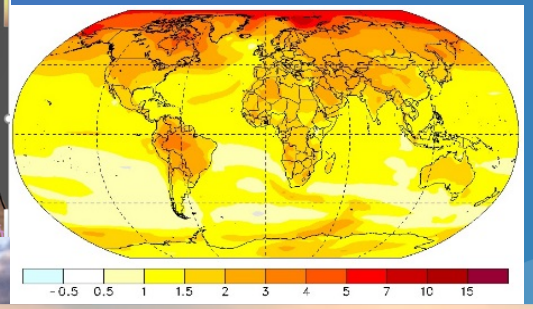
2019 Mekong Research Symposium

IBM Environmental Platform for Modeling, Impacts, and Mitigation Strategies

December 18, 2019





Lee Angelelli
Executive Client Architect
langelel@us.ibm.com
(c) +1-571-244-5970

IBM

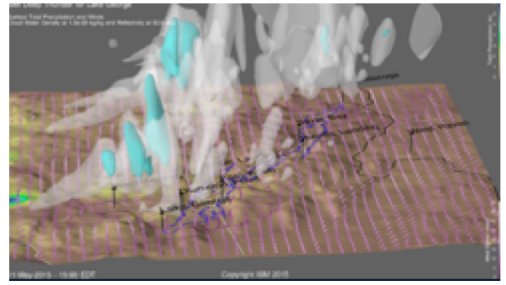
IBM Environmental Insights-as-a-Service Enterprise Platform

Client Challenges: – Ensure ecosystem resilience in the face of long-term pressures from climate change and intensifying human use

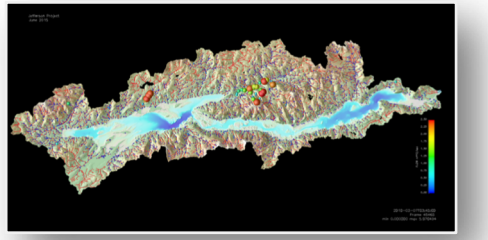
Platform Capabilities:

- **Weather model forecasts** atmospheric conditions for other models at 333m horizontal resolution
- **Runoff model predicts** water (precipitation, snowmelt) transport, hydrology, and salt / nutrient / pollutant conveyance in streams and into to the lake (or river or ocean)
- **Circulation model** (DeepCurrent) predicts transport, hydrodynamics, water quality and chemistry
- **IoT sensor data** is ingested into models for assimilation at run time
- Prior day model forecasts are compared to IoT sensor data for validation
- **Scenario Engine** provides for historic (hindcast) and future conditions (forecast, seasonal, climactic, …)

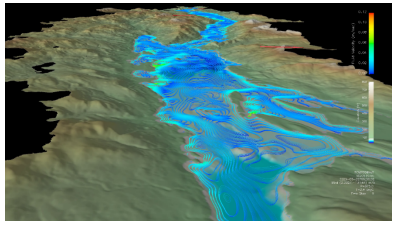
Weather – Land Surface



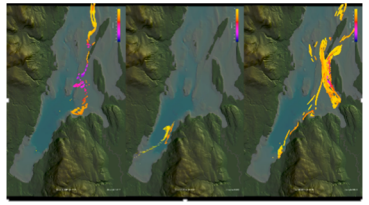
Runoff-Pollutants Particles



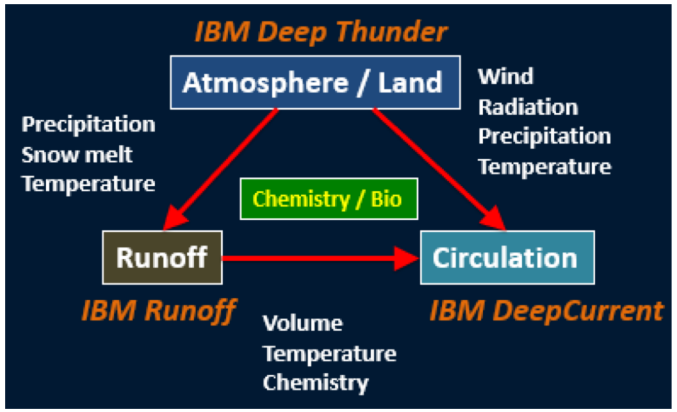
Circulation-Chemistry Pollutants Particles



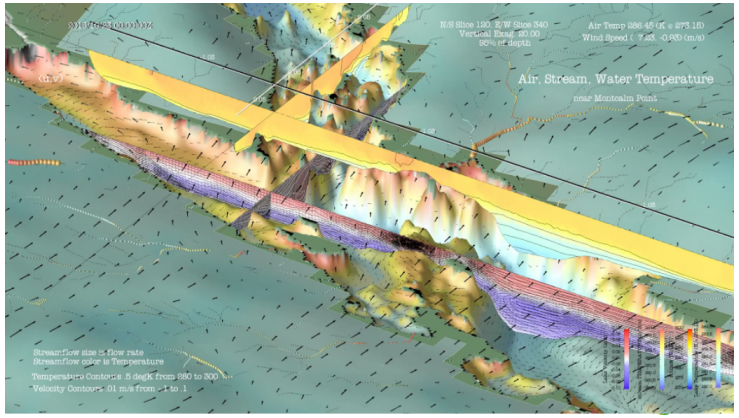
Scenario Engine



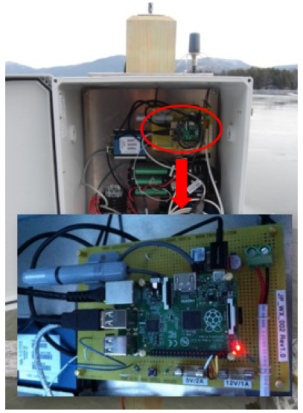
Operational Model Forecasting Platform



Fully Coupled Forecast

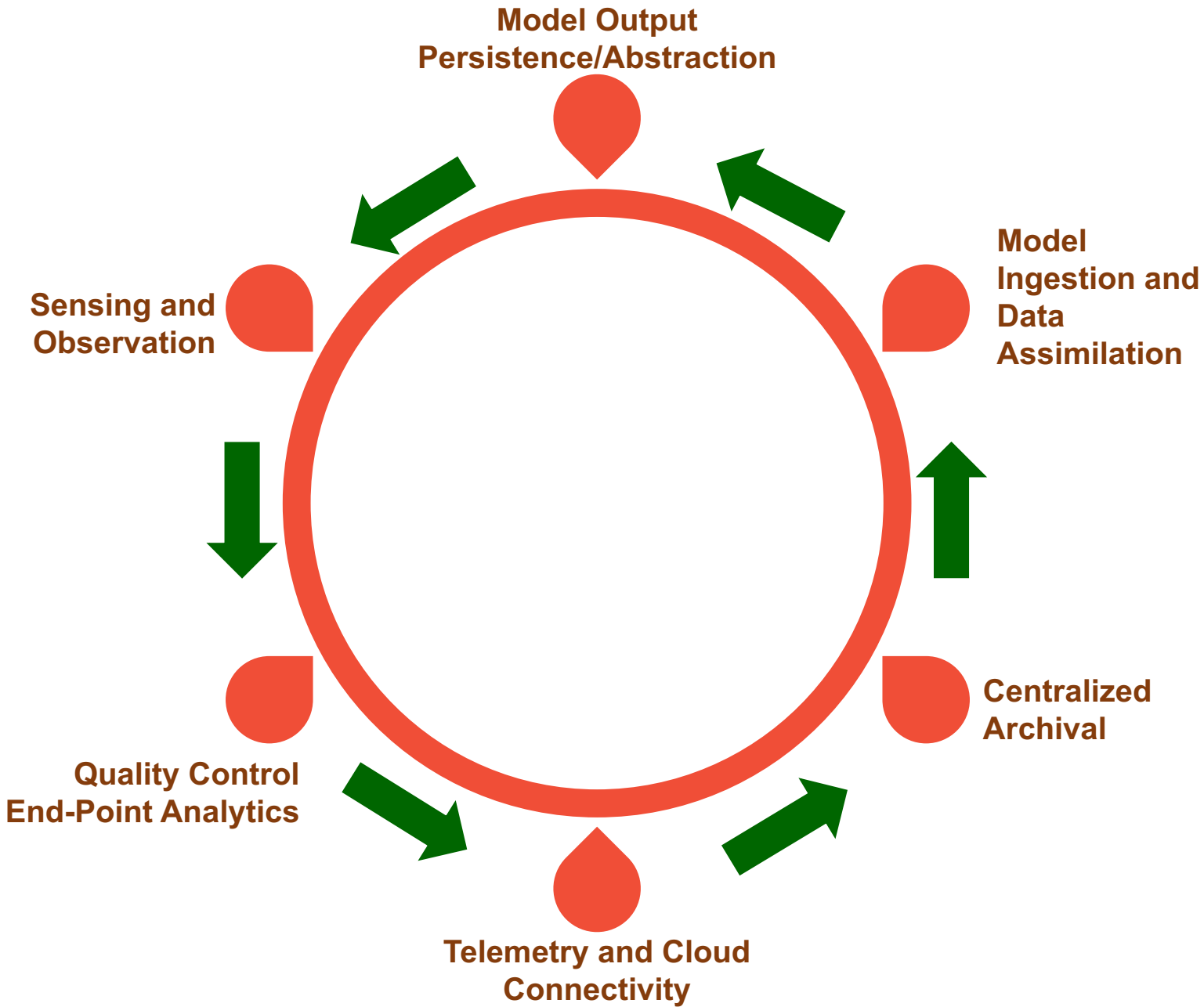


Edge Embedded Intelligence



IBM's Environmental Platform Infrastructure Design

- Provides a flexible, extendable, reliable, and reusable infrastructure to provide a common platform for environmental monitoring, collaboration and innovation
- Supports environmental sensors and other data sources as inputs to real time analytics, modeling, and operational decision making
- Provides an information hub to provide rapid access to model-derived and observational data
- Enables an operational coupling framework for independent models that scales in a generic way
- Distinction between model generated data and observational data



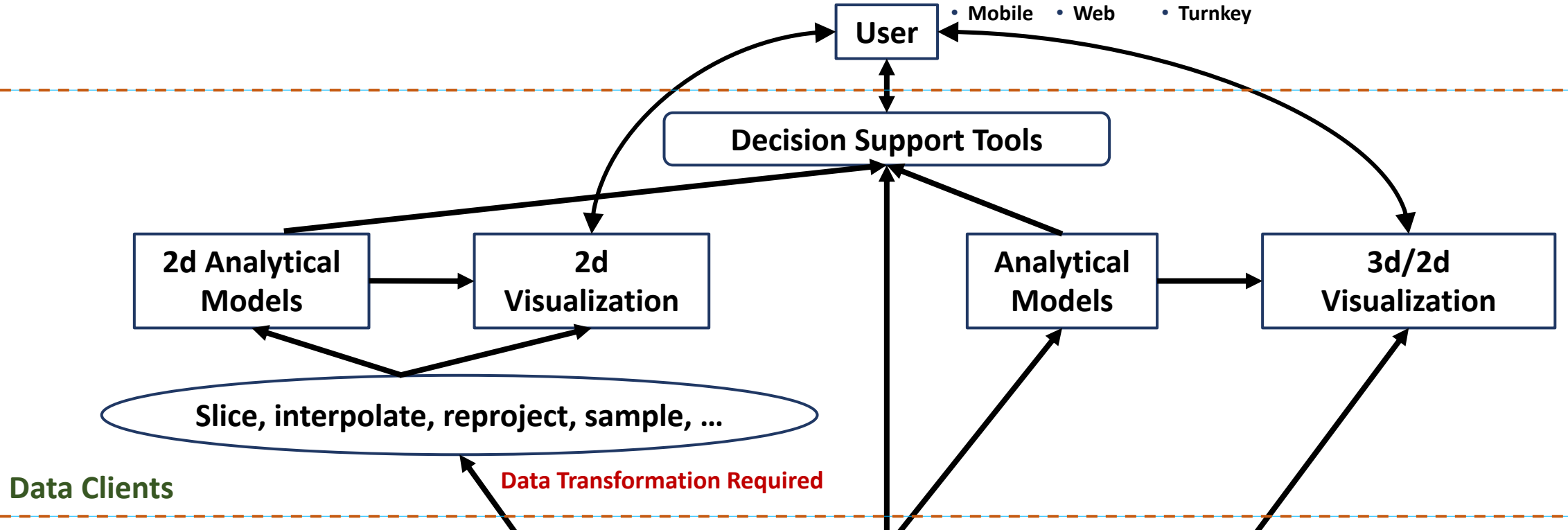
Enable Different Roles for Observation and Simulation

Sample (physically|virtually) the environment when and where needs are greatest

1. **Typical independent uses for observations (both remote and in situ sensors)**
   - Real-time monitoring
   - Analysis/data mining from collections
2. **Typical independent uses for simulations (models)**
   - Forecasting (e.g., prediction of what will be observed)
3. **Uses of observations for simulations (coupling)**
   - Data assimilation (e.g., forecast initialization)
   - Model verification and tuning
4. **Uses of simulations for observations (coupling)**
   - Observing system simulation experiments (e.g., sensor network design)
   - Sensor steering

- **Note: Don't force a data client on customers but enable support for their choice!!!**

Conceptual Architecture (Independent of Application)



Observations
- In situ and remote sensors
- 2d and 3d sampling
- Air, land and water

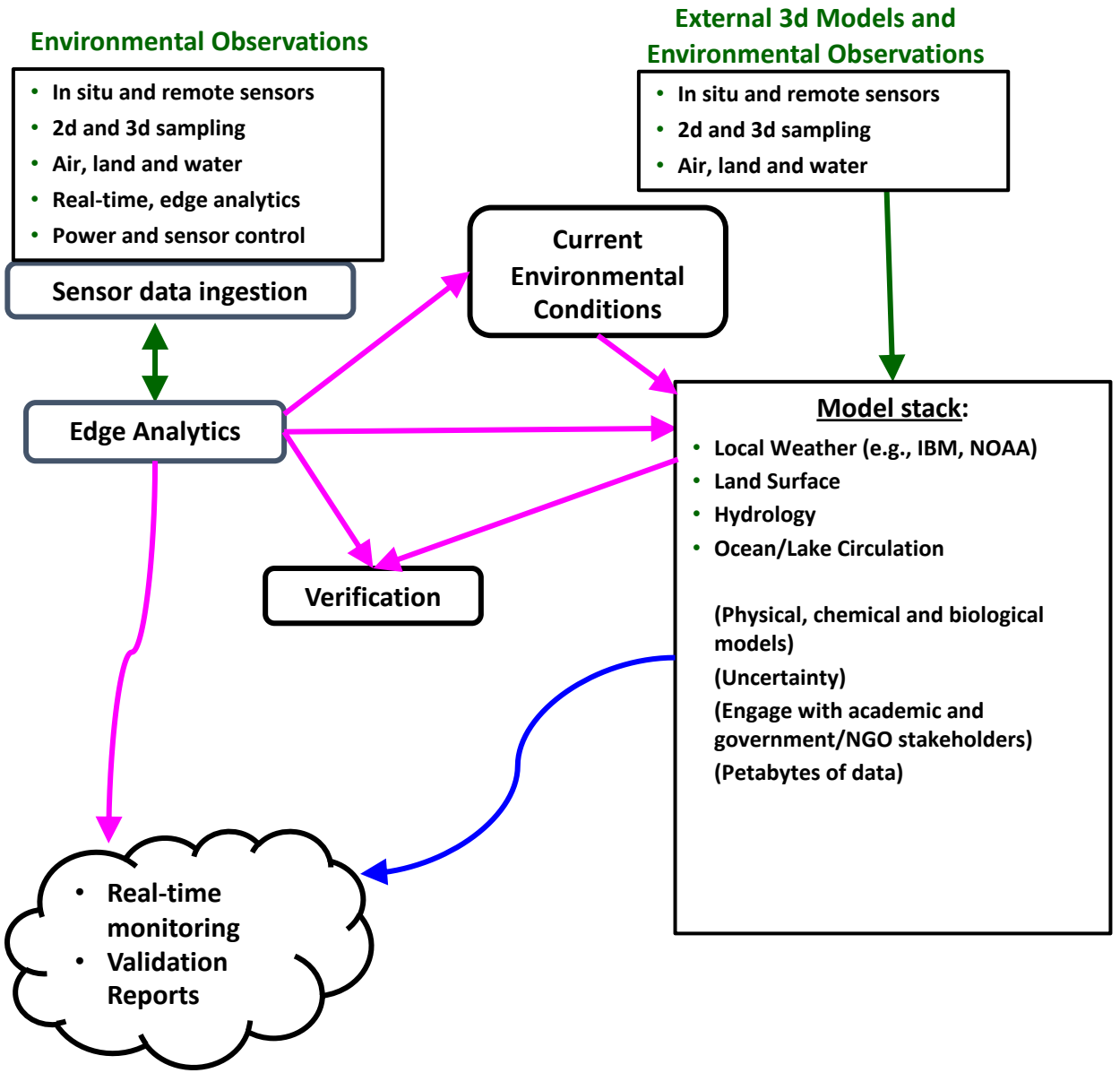
3d/2d Geo-Spatial Multi-variate timeseries, models and observations*

Public and private data sources
- Observations: Current (near real-time) conditions and history
- Models: Predictions and retrospective analyses

Physical, chemical and biological models
- Water/Hydrology/Hydrodynamics
- Air/Weather
- Land/Hydrology/Soil
- Crop/Vegetation
- Ecosystem
- Disease

*Requires comprehensive data model(s)

Data Flow and Components of the Environmental Platform (Data Collection and Generation)



Data Servers

Curation
Scale
Availability

- Data standards
- Preservation of full 3d data fidelity
- Support for basic queries, analysis and visualization
- Integration with the community
- Dissemination

Data Insight

AI
Uncertainty
Augmentation

- Data exploration
- Anomaly identification
- Feature extraction
- Ensemble blending
- Uncertainty analysis
- Deep learning backed regression

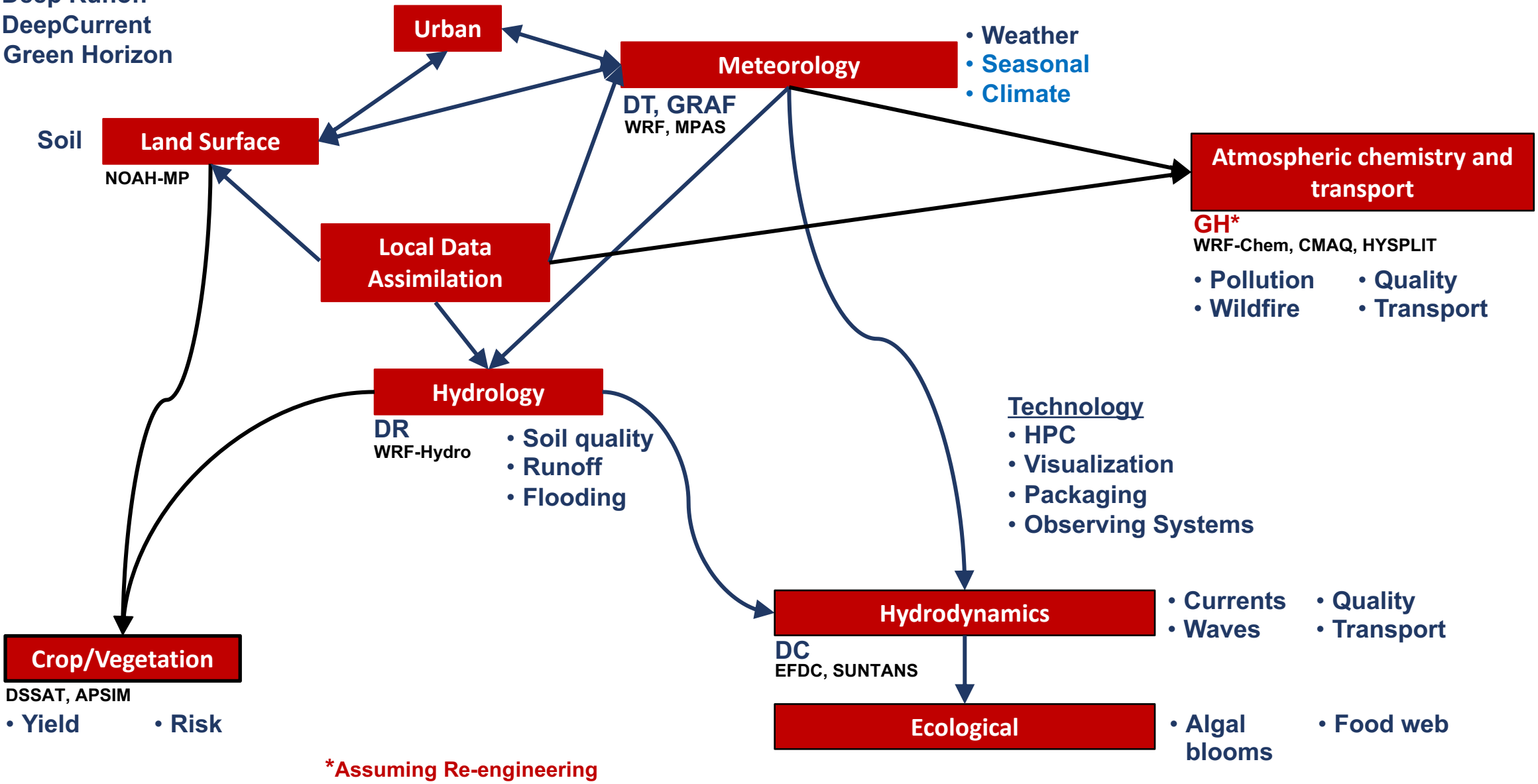
Visualization

Geospatial
Communicate
Innovate

- Spatio-temporal at scale
- Communicate information not data
- Augmentation (e.g. streamline, particle)
- Availability
- Community engagement

IBM's Environmental Science Research Areas: Model Coupling

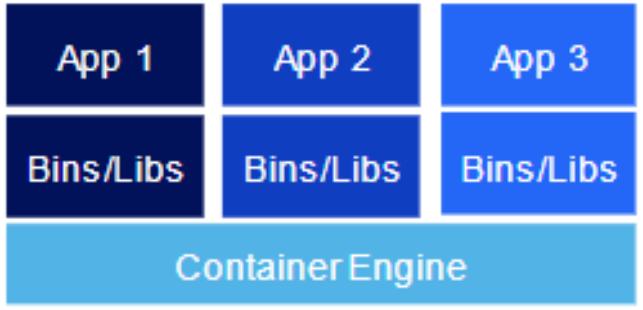
- DT = Deep Thunder
- DR = Deep Runoff
- DC = DeepCurrent
- GH = Green Horizon



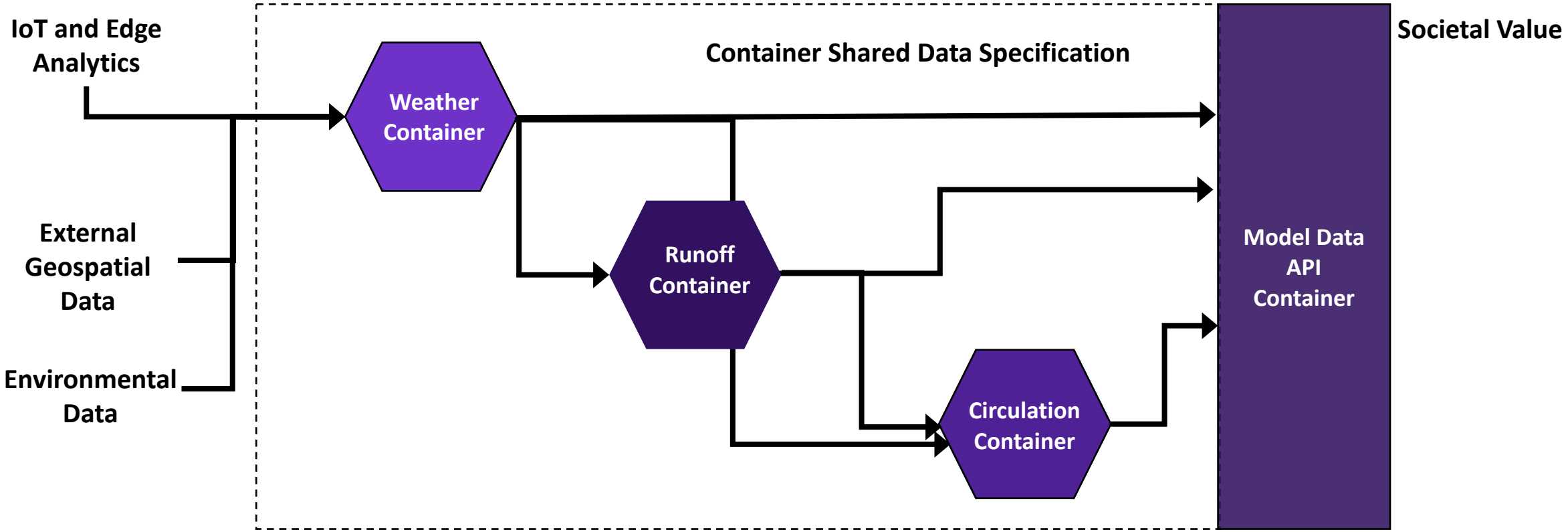
*Assuming Re-engineering

Packaging Approach / Philosophy

- Create a standalone package of the runtime environment for each model using Docker/OpenShift container technology
  - Each model can be directly deployed in the cloud or local IT environment
  - No separate code dependencies other than container itself
- Use Docker container technology to run a complete forecast (or hindcast or reanalysis) and provide an API for access to the data
- Enable minimal input from user to set up and run via a GUI
- Enable expert users that can run/schedule from command line/APIs



How to package environmental models to simplify reuse

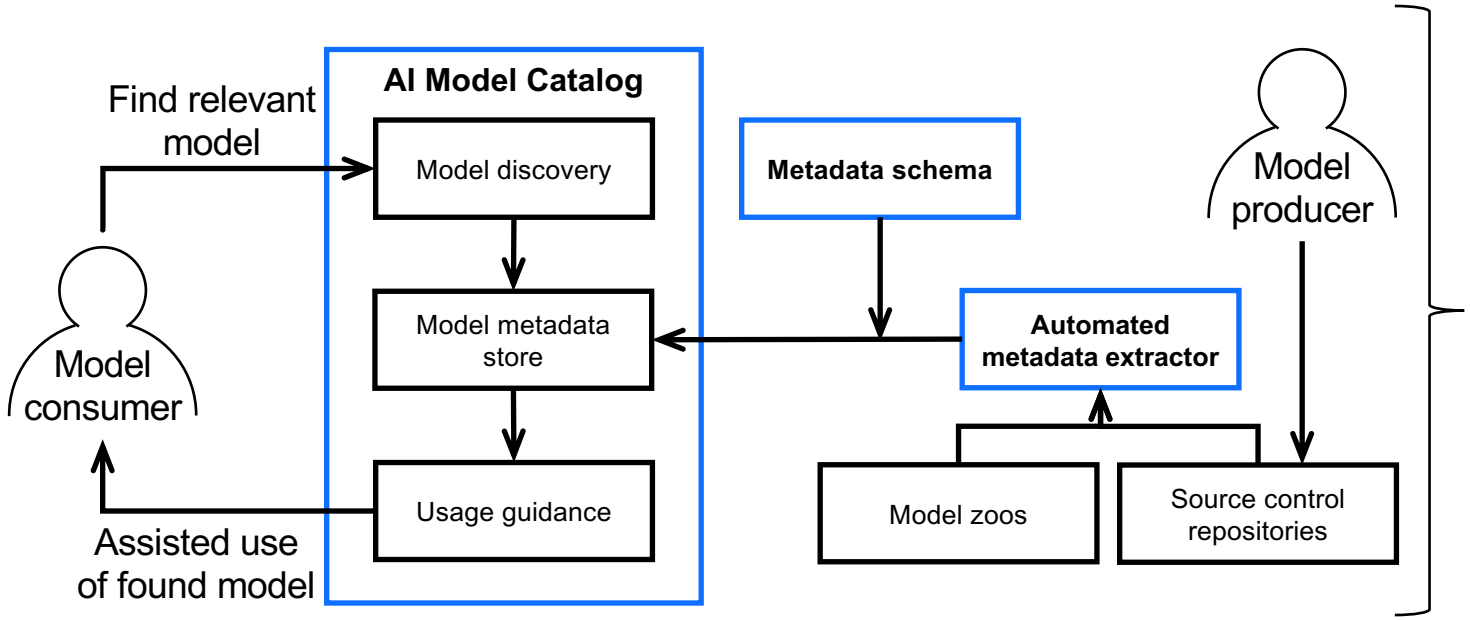


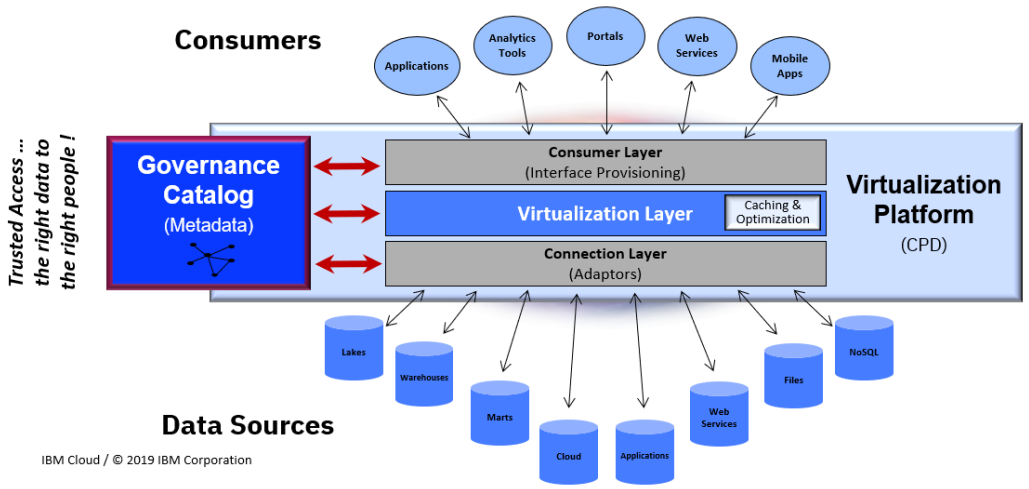
Cloud Native for integration with public/private cloud

Interconnected with the other assets and data

IBM Cloud – Watson Knowledge Catalog

- Quickly build a 360 view of all assets and provide them for AI and Analytics
- Discovery – Intelligent discovery of data and AI assets with advanced classification and profiling to provide context
- Catalog – A rich metadata index of all data and AI assets with social collaboration and enhanced findability
- Activate – Powerful governance tools to control and protect access to data with visibility to data use
  - Policy Engine to author, activate and enforce business policies and rules





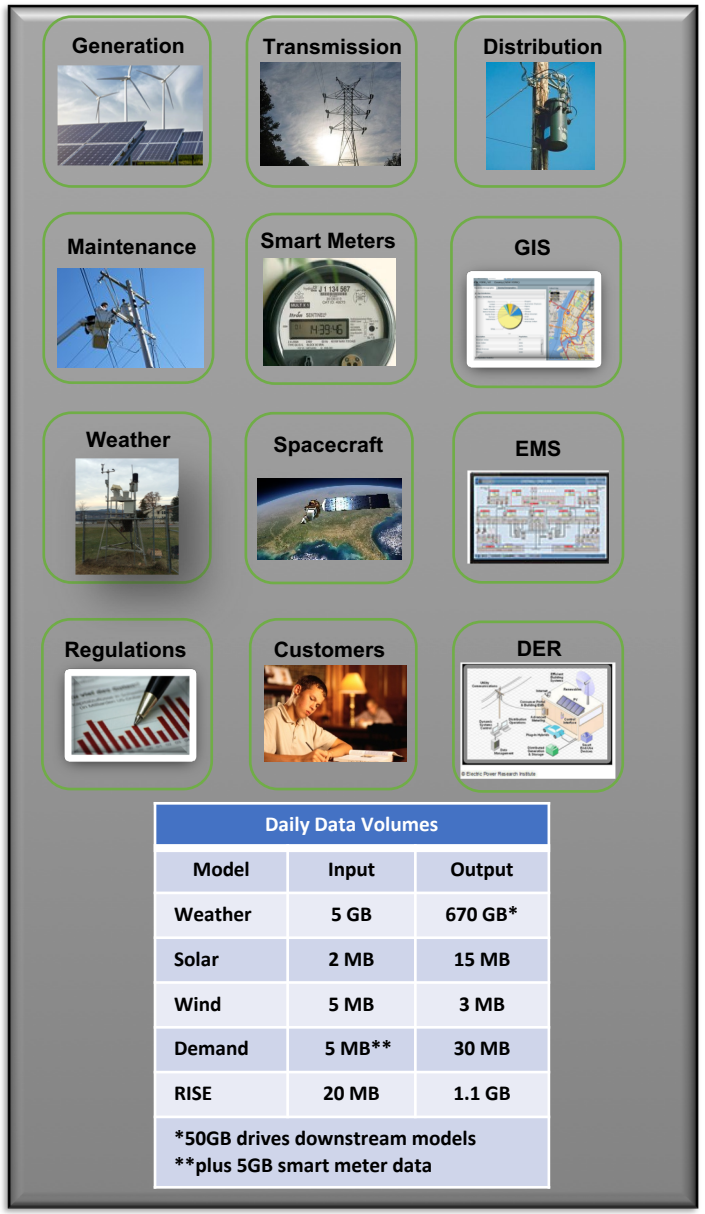
Use Cases

Modern Agriculture Technology/IoT Data Can Improve Harvests/Sustainability

- Use automatic data collection and AI-driven insights to increase crop yields and improve sustainability
  - Combine predictive analytics, artificial intelligence, weather data, and Internet of Things sensors
    - Give farmers insights and best practices on plowing, planting, spraying, and harvesting
  - Fuse geospatial data sets from satellites, drones, flights, weather models, and IoT sensors
  - Ingest data from satellite imagery, soil data, farm practice and workflow data, sensors on farm equipment that monitor, among other things, seed counts, nutrient levels and fertilizer flow
  - Leverage the Weather Company historical data, near-realtime observations, and seasonal forecasts
- Results
  - Improved sustainability with deeper insights into factors such as crop input optimization, energy consumption, land and water use, ….

Weather and Renewable Energy

Data Sources

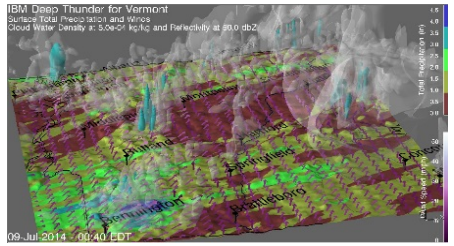


Daily Data Volumes		
Model	Input	Output
Weather	5 GB	670 GB*
Solar	2 MB	15 MB
Wind	5 MB	3 MB
Demand	5 MB**	30 MB
RISE	20 MB	1.1 GB

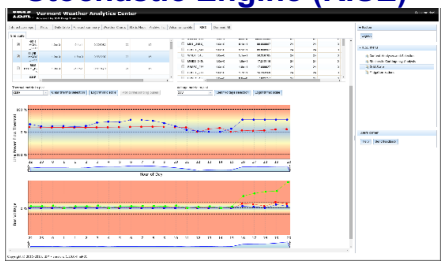
*50GB drives downstream models
**plus 5GB smart meter data

Models

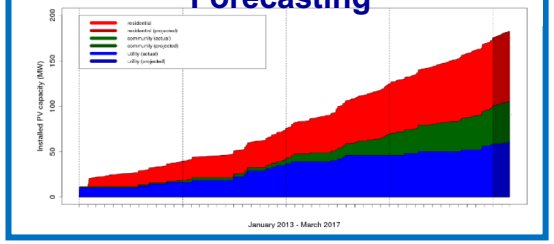
***Deep Thunder* Weather Forecasting**



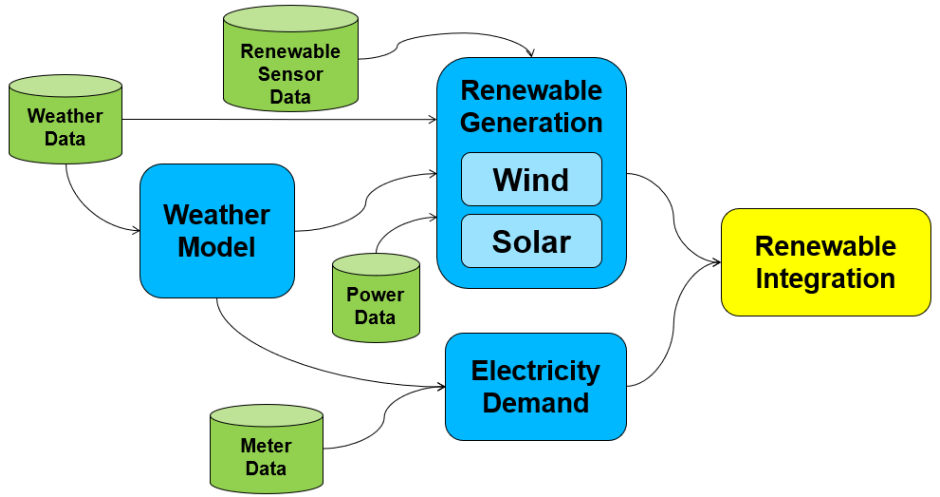
Renewable Integration Stochastic Engine (RISE)



Electric Demand + Net Metering Forecasting



R^2: Renewable Power Forecasting (Solar and Wind)



Decision Support

Manage Demand
Maximize Supply
Maintain the Grid

- Accurate weather and demand forecasts
- Schedule sufficient power generation between renewable energy sources and fossil fuels to match demand
- Integrated approach using a common platform, data model, visualization, etc.

Quick Demo of Watson Knowledge Studio

https://dataplatform.cloud.ibm.com/docs/content/wsj/getting-started/welcome-main.html

The Task Force for Climate-related Financial Disclosure (TCFD)

Physical Risk

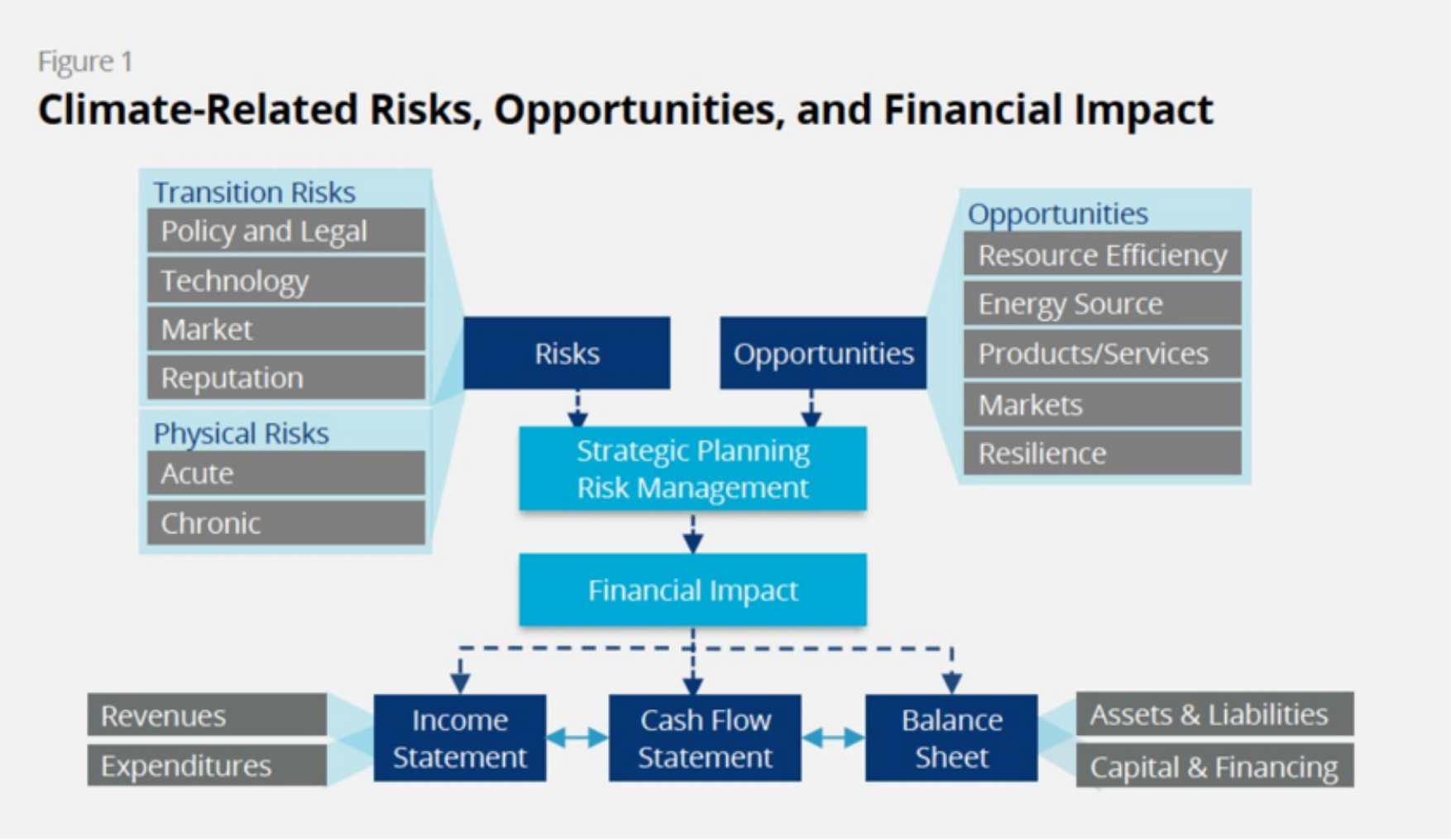
- Heatwaves / Floods / Wildfires / Storms
- Sea level rise / Storm surge / Extreme variabilities

Transitional Risk

- Moving to low-carbon …asset base / economy / operations / society and investments
- Changes in regulation, policy, technology, business models, legal frameworks

Liability Risk

- Having suffered loss or damage from physical or transition risk factors seeking to recover losses




IBM
Thanks

Appendixes

Environmental Insights-as-a-Service Platform Assets and Capabilities

High Resolution Modeling

Operational Model Forecasts (36 hour, automated, coupled, modular):

- Weather – conditions at 333m horizontal resolution
- Land Surface – surface and subsurface runoff
- Hydrology – surface runoff and streams, particle transport
- Hydrodynamics – water body circulation, particle transport
- Chemistry / Biology – reactions and model integration

Scenario Engines:

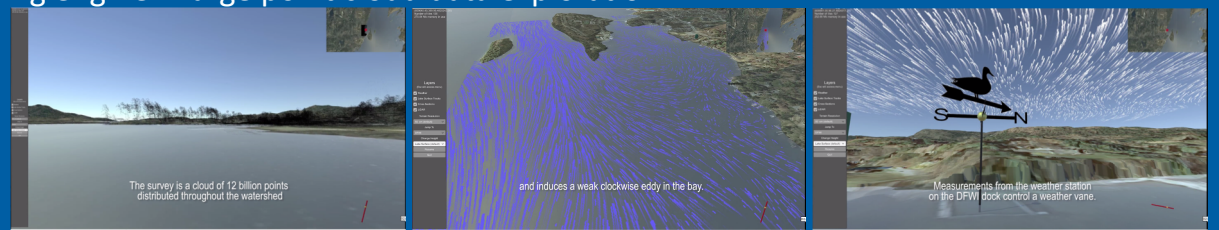
- Cognitive – model integration
- Salt, Chemicals, Biology
- Particles / Pollutants
  - Forecast Position – location and time in future
  - Source Attribution – likely location and track in the past
- Climate Impacts – for different warming scenarios

Model Coupling / Operational Framework:

- Automated Model Execution
- Live IoT Data Assimilation – model alignment to actual conditions
- Automated Model Validation – validates forecast to measurements
- ADCP Validation – big data processing, 3D current ingestion

Scientific Visualization:

- 3D
- Gaming engine – large point cloud data exploration



IoT / AI Platform

DiamondPoint (IoT edge node on each sensor platform):

- Sensor Integration – control of all sensors (sensor agnostic), hardware, operations and communications
- Real time data QA/QC and Analytics –
  - Streaming Statistics: on live sensor measurements
  - AI Prediction: of the next sensor reading before the measurement is made
- Live Data Uncertainty Determination – continuous determination of data quality (uncertainty) while streaming at the maximum measurement speed of the sensor
- Peer Communication – directly with neighbor or surrogate sensors for information query, data comparison, episodic event alerting with dynamic sampling strategy changes
- Adaptive / Autonomous Decisions – sensor control and sampling plan modification based on live measurements
- Model Forecast Analytics at Edge – adapt operations to forecasted conditions at the sensor platform location

Data Aggregator:

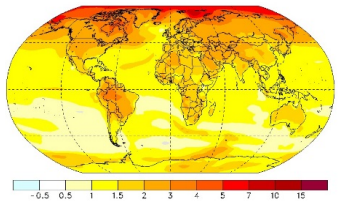
- Cloud, messaging hub, database, modeling platform , web / GIS integration for data, metadata and logistics

AssemblyPoint:

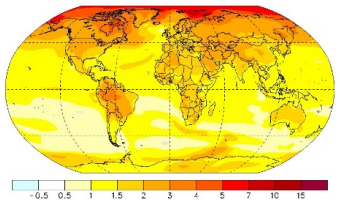
- Map / GIS user interface
- Operational model forecast framework data integration
- Automatic registration of new sensors and data fields
- Automated graphing and validation report generation
- Slack Integration
- TWC Mobile app. integration

Our Approach to Regional Climate Modelling

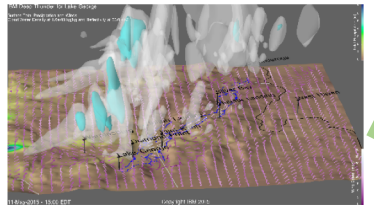
1



Global Climate Model (Future)

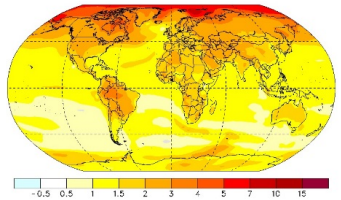


Global Re-Analysis (Past)

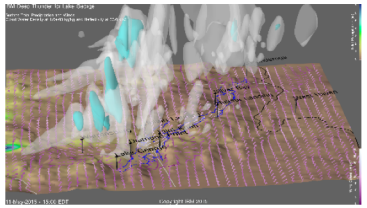


Regional Climate Model (Future)

2



Global Re-Analysis (Past)



Regional Re-Analysis (Past)

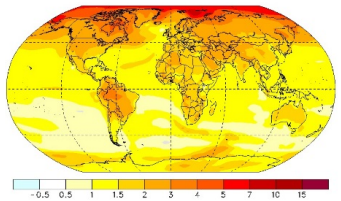
Observations (Past)

Validation

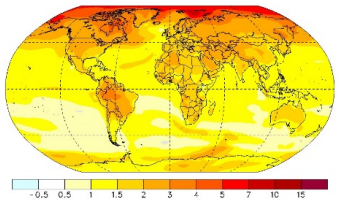
4

Effect of Mitigation Strategy

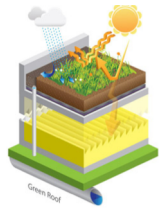
3



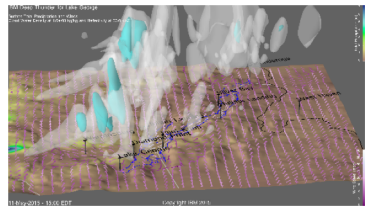
Global Climate Model (Future)



Global Re-Analysis (Past)



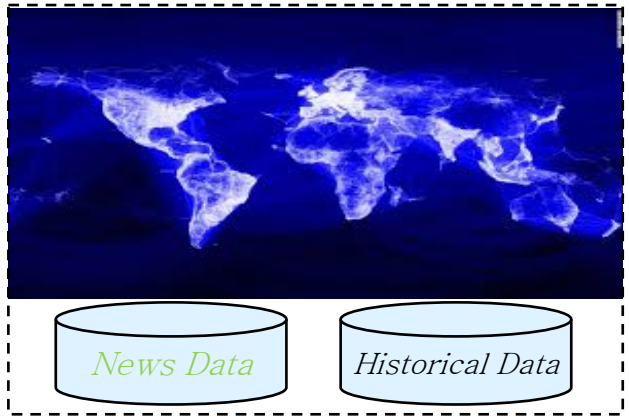
Mitigation (e.g., Green Roofs)



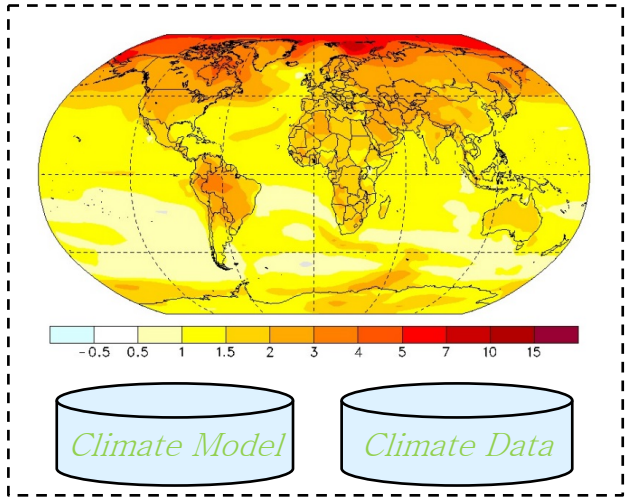
Perturbed Regional Climate Model

How AI Can Help Decision Making for Climate Change Mitigation

Couple processing of structured and unstructured data to provide support on climate action decision making



Global Coverage

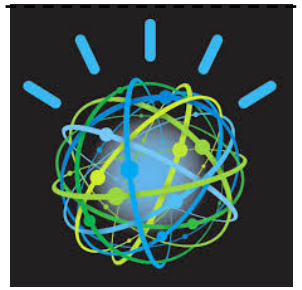


Unstructured Data

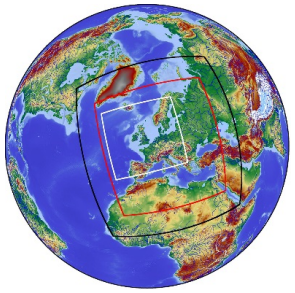
Structured Data

Watson/AI capabilities

Excluded image recognition



Monitoring & Prediction



Regional Climate Simulation

Examples of Actions for Heat Waves

- ✓ *Safety programs*
- ✓ *Cool roofs*
- ✓ *Green roofs*
- ✓ *Planting trees*
- ✓ *Cool pavements*
- ✓ *White roofs*
- ✓ *Early warning systems*
- ✓ *Community cooling centers*

Focus on a selected subset of actions/topics based data availability and analysis results

Decision Analysis

Evaluate the impact of different Climate Action Risks & Decisions

Focus on influence diagram(s) built and configured using experts data & input

Advise on policy decisions

IBM Cloud Private is a platform for Application Modernization

IBM software solutions including Open Source – e.g. Data, Analytics and Developer Services

Cloud-enabled middleware, application runtimes, messaging, databases & analytics to optimize current investments and rapidly innovate

Core Operational Services

To simplify Operations Management, Security, DevOps, and hybrid integration

Kubernetes-based Container Platform

Industry leading container orchestration platform

Cloud Foundry

For prescribed application development & deployment

Terraform (CAM)

Infrastructure as Code for multi-cloud provisioning to public and on-premise private clouds

YOUR DATA CENTER

IBM Environmental Platform Data Model Principles

- Support of models and/or observations, which may be coupled
- Leverage community, where appropriate => utilize open source standards
- Plethora of open source tools for development and end users
- Ease for data exchange and interoperability with scientific community
  - Separate data access from usage (enable application building, querying and data exchange independent of as well as for visualization)
  - Simplify use of tools (e.g., enables polymorphism)
- Preserve full (3d) data fidelity (avoid unnecessary interpolation or resampling) and semantics, including traceability
- Provide http-based access to data to compose higher-level services
- Don’t force a data client on customers but enable support for their choice

Environmental Platform Data Model Design Choices

- **Model data**
  - **CF-compliant netCDF (generalized data model as abstraction for self-describing multidimensional arrays) -- `https://www.unidata.ucar.edu/netcdf/`**
  - **Software interface and applications available on many platforms**
  - **Interfaces to many relevant data sets, data models and physical models**
  - **Metadata conventions for some complex topologies and coordinate systems**
  - **Familiar in relevant communities (e.g., oceanographic, atmospheric, hydrologic)**
  - **http access via ncWMS -- `ncwms.sourceforge.net`**
- **Sensor data**
  - **.csv (ingestion standard) for simplicity given diversity at sensor level**
  - **MQTT messaging protocol from sensor to platform**
    - **Integration with any IoT platform**

Precision modeling, coupled with high frequency data acquisition and IoT edge analytics with AI, provides insights leading to new discoveries and better understanding of existing phenomena such as…

Physical:

- Internal waves: sediment and nutrient re-distribution → understanding of conditions driving biological events (HABs)
- Flows under ice: basin exchange without wind → nutrient and pollutant distribution forecasting without wind
- Model forecasting scenario engine → eColi location predictions, source attribution and “what-if” scenario planning
- Enhanced weather and flow predictions for optimization of sea port operations

Chemical:

- Salt modeling and enhanced stream and near shore monitoring → salt truck sensors, application zone modification, alternative product investigation, real-time road condition monitoring, proactive treatments
- Pollutant tracking and source attribution

Biological:

- Enhanced stream salt loading measurements: salt / chemistry experiments → alteration of ratio of male / female frog development; salt tolerance changes in species; migration potential of specific species (including invasives)
- High frequency sonar integration: daily vertical cycle → diurnal feeding pattern migration of zooplankton identified; source of Harmful Algal Bloom (HAB) event initiation determined

Reduce Water Loss Through Intelligent Asset Management Practices

- Perform proactive wastewater management and situational awareness through intelligent asset management
- Leverage intelligent instrumented approaches that includes IoT sensors, advanced communication networks and advanced analytics.
- Use descriptive/predictive analytics and operational dashboards solution to help understand asset condition, usage and allocation in near real time to more effectively manage water resources and costs
- Create an interconnected value chain, linking utility participants at multiple levels
- Enable intelligent flows of information – transform/optimize asset performance, worker productivity
- Create a common operating picture that consolidates data from disparate sources to provide situational awareness
- **Results**
  - Better manage the life cycles of physical assets such as meters, pumps, valves and sensor devices to improve preventive maintenance
  - Reconcile deployed asset information with information collected by meter data management systems
  - Correlate the knowledge of physical assets with problems experienced with the operations infrastructure to better analyze a problem's root cause.
- **Client reference – Miami Dade County**
  - Miami-Dade County implemented IBM Intelligent Water to track water leaks in parks
  - Alerts park managers whenever a leak was detected, minimizing water waste and cutting costs
  - Saved Miami-Dade County over USD 1 million in the first year