Overview of Task 4: Model Evaluation and Performance Quantification

Task 1
Data Collection and Pre-Processing

Task 2
Satellite Data Acquisition

Task 3
Model Development

Task 4
Model Evaluation and Performance Quantification

Task 5
Methodology Development and Reporting

In Task 4, Resolve Hydro will use the testing dataset to **evaluate the overall performance of the shortlisted models** developed in Task 3

Key Performance Metrics:

- Micro-average and macro-average precision, recall, and F1-score
- Cohen's Kappa
- Cross-entropy
- Matthew's correlation coefficient
- Accuracy
- False Positive Rate
- Critical Success Index
- False Alarm Rate
- Frequency Bias

Key Evaluation Contexts:

- Crop type
- County
- Hydrogeomorphic region (CBP)
- Major physiographic section (PA)
- Soil class and percent slope
- In regions for which historical data was used for model training
- In regions for which historical data was not used for model training

Metrics will be used to help explain model errors, provide recommendations regarding model application in new areas, and select a "best-performing model"

Overview of Task 5: Methodology Development and Reporting

Task 1
Data Collection and Pre-Processing

Task 2
Satellite Data Acquisition

Task 3
Model Development

Task 4
Model Evaluation and Performance Quantification

Task 5
Methodology Development and Reporting

Primary Approaches for Remote Sensing Model Application

Virtual Transect Survey (VTS)

- Use remote sensing to classify conservation tillage at existing observation points to reduce cost, time, and labor associated with the current in-situ approach

Virtual Field Survey (VFS)

- Use remote sensing to classify conservation tillage in a random sample of agricultural fields
- Similar to the VTS approach, but not limited to roadside fields

Total Area Classification (TAC)

- Use remote sensing to classify conservation tillage over all agricultural lands identified by the CBP LULC dataset

Note: In field verification will be required for all approaches

Recommended Guidance and Resources

- Appendix V – Protocols for Verification of Annual BMP Data Submissions (not included in Verification Framework document)
- Assessment of Potomac River Basin Remote Sensing Pilot Project
- Recommendation Report: Establishment of Uniform Evaluation Standards for Application of Remote Sensing to Identify and Inventory Ag Conservation Practices for the CBP Partnership's Watershed Model
- Example Methodologies:
  - VA Tillage/Residue Survey Pilot Project and Tetra Tech review document
  - LDG Non-Intrusive BMP Verification SOP

Review of VA DCR Virginia Tillage/Residue Survey (Tetra Tech 2023)

- Acknowledge how variances in calibration could contribute to the overall accuracy
- Geolocated locations should not be cited in publicly available report
- Commended for exceeding the 10% threshold identified in the CBP (2017) report by verifying 11.5% of the sites (1,561 out of 13,600) in their revised digital image verification procedure
- Reminder to revisit CBP 2017 Recommendations Report

Recommendation Report for the Establishment of Uniform Evaluation Standards for Application of Remote Sensing to Identify and Inventory Agricultural Conservation Practices for the Chesapeake Bay Program Partnership's Watershed Model (CBP 2017)

- Metrics assessed: False Alarm Ratio (FAR), Hit Rate (HR), and Frequency Bias (FB)
- Minimum sample size
- Options for crediting the results of data generated via suitable remote sensing

Methodology Elements

- Problem Statement and Scope
- Traditional Process Limitations
- Overview of Reporting Requirements
- Professional Qualifications Criteria
  - Training Activities
  - Verification Activities
- Data Source Description
- Description of the Technology
- Results, Comparison, and Selection of Final Approach
- Statistical Testing Results
- Multi-Jurisdictional Considerations and Recommendations
- Future Perspective and Lessons Learned

Specific Recommondations

- Include QA/QC on prevention of double-counting, as well as an exterior validation (i.e., ground-truthing) methodology
- Long-term maintenance considerations
- Primary constraints for application in other Bay jurisdictions:
  - Funding
  - Baseline data
  - Personnel/capacity
- Contributed example of how methodology could be applied in a different jurisdiction/ geography
- PA-specific project attributes:
  - Type of crop residue systems considered
  - Soils
  - Different attributes of agricultural production

Resources, Lessons, or Suggestions to Share?

Ashley Hullinger
Water Program Specialist, PA DEP
ahullinger@pa.gov

Scott Heidel
Environmental Group Manager, PA DEP
scheidel@pa.gov

Thomas Howard
Resolve Hydro
thoward@resolvehydro.com