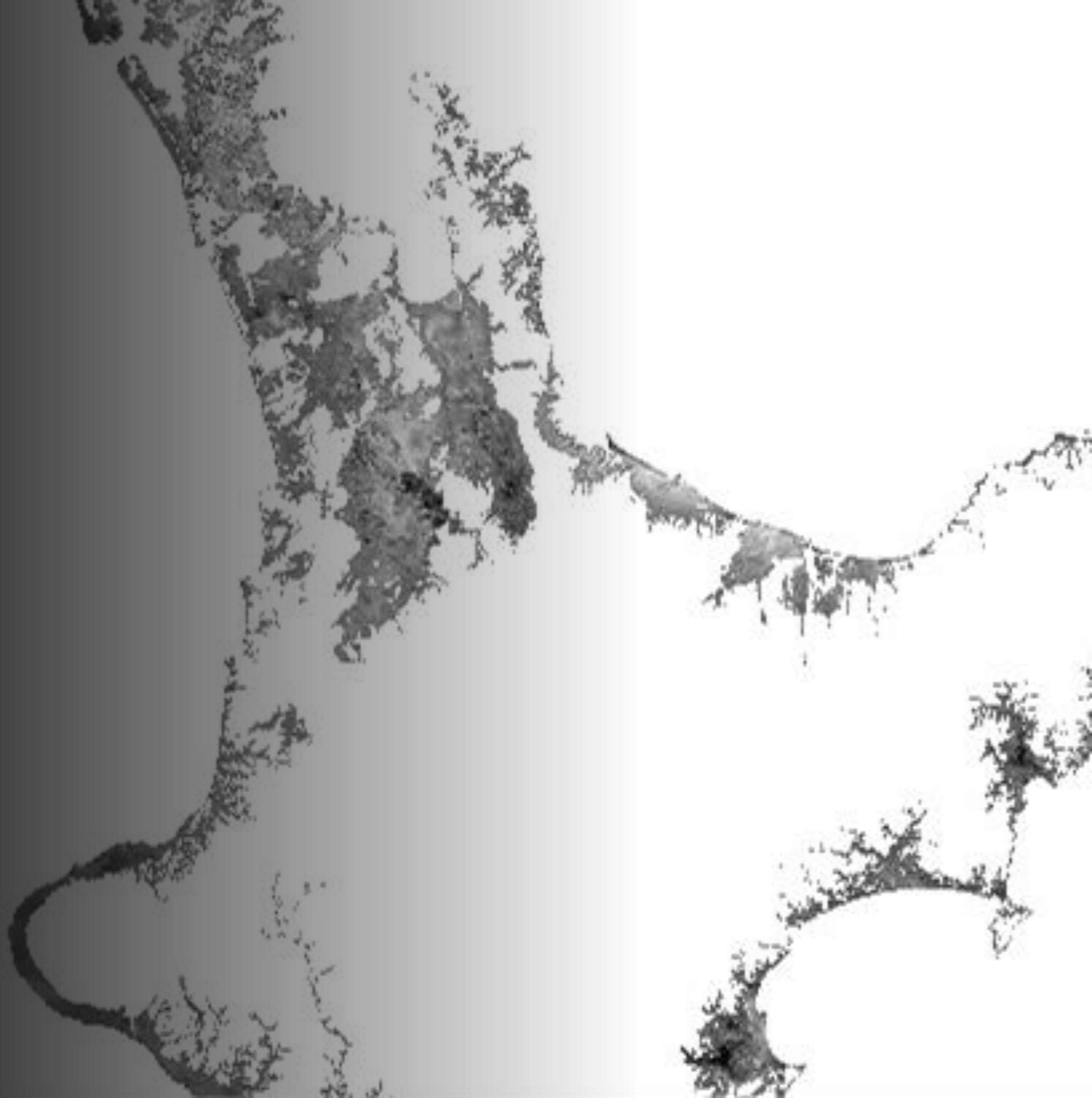


Modelling Shallow Groundwater Risk in New Zealand Using Categorical Machine Learning Models

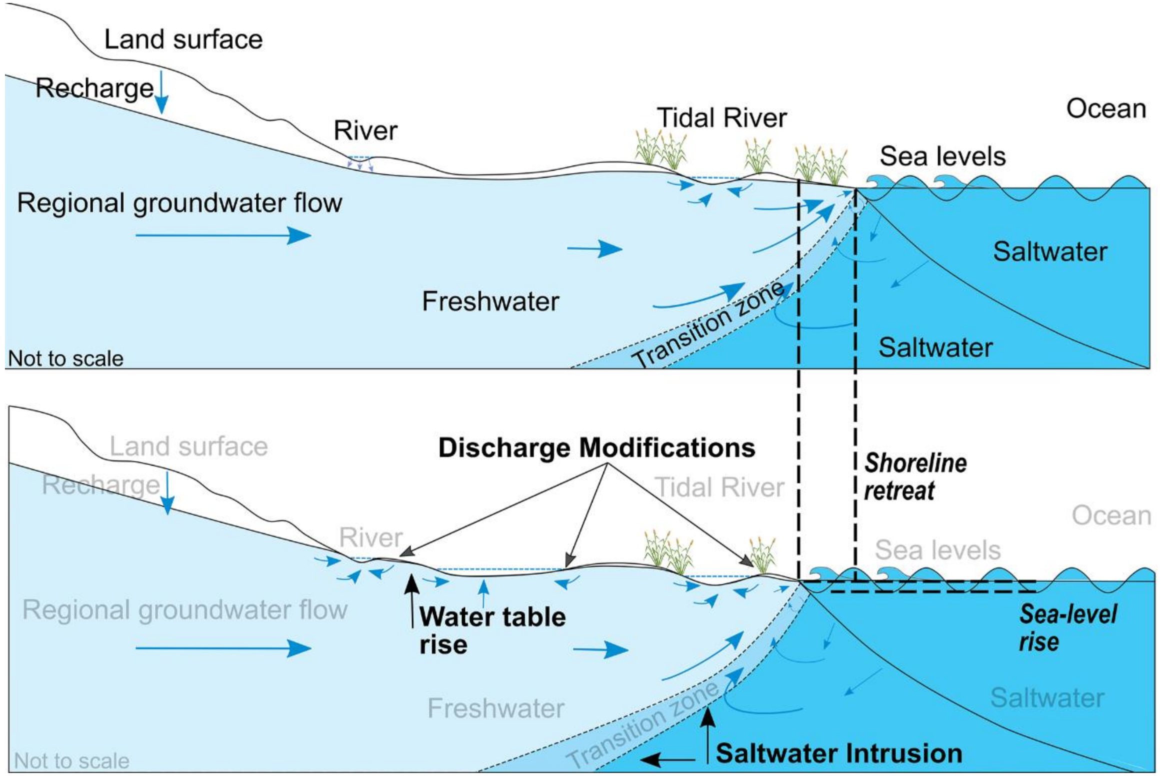
Patrick Durney, Matt Dumont, Zeb Etheridge, Christo Rautenbach,

Komanawa Solutions Ltd

NIWA

The Challenge

- Rising sea levels threaten coastal communities
- Groundwater shoaling - a hidden risk
- PREMISE 1: Need to know vulnerable areas
- Critical knowledge gap: Where is groundwater already shallow?
- PREMISE 2: Acceptable prediction error is dependent on acceptable risk



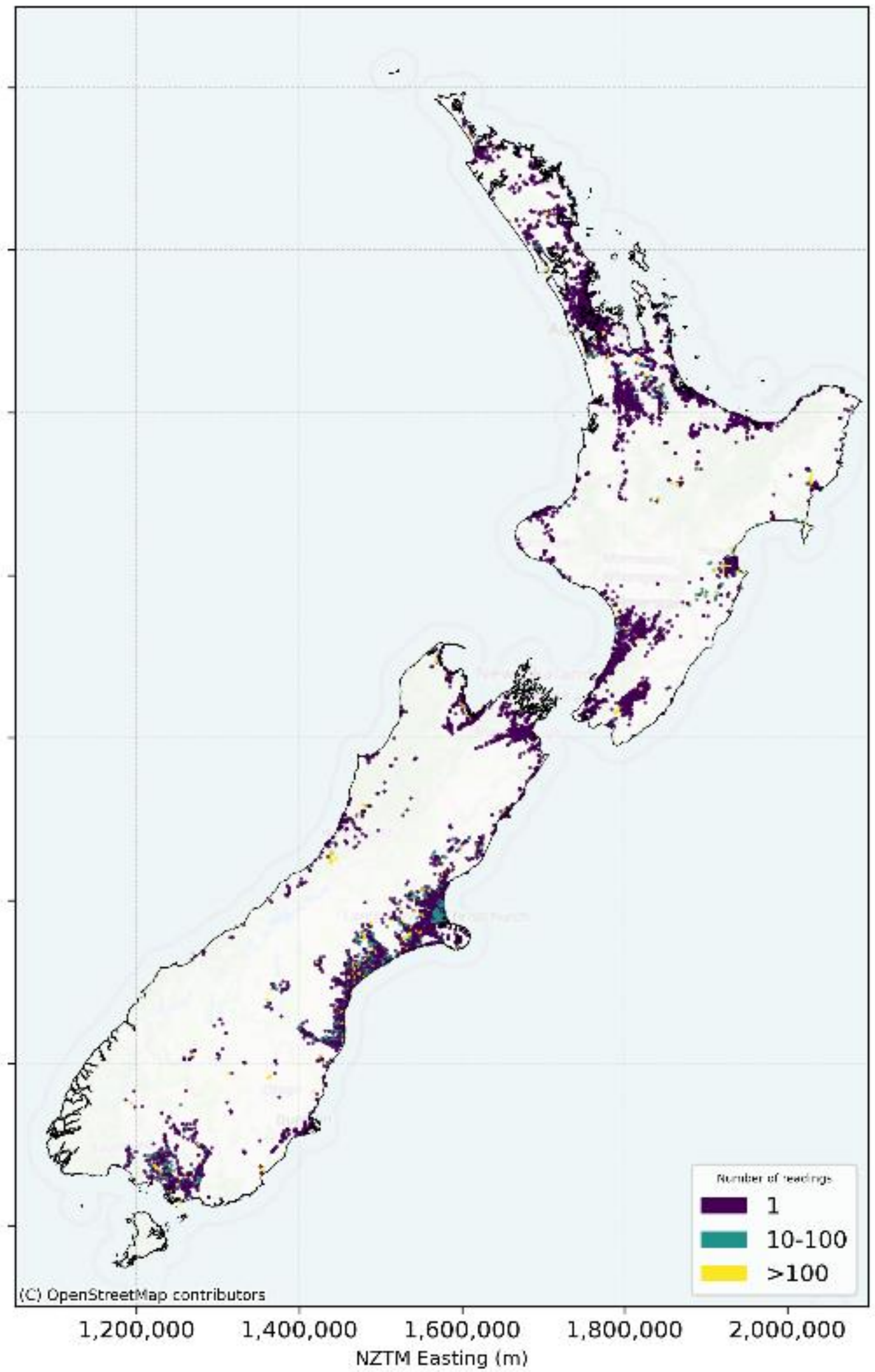
Bosserelle, A. L., Morgan, L. K., & Hughes, M. W. (2022). Groundwater rise and associated flooding in coastal settlements due to sea-level rise: A review of processes and methods. *Earth's Future*, 10, e2021EF002580. https://doi.org/10.1029/2021EF002580

What do we have to work with:

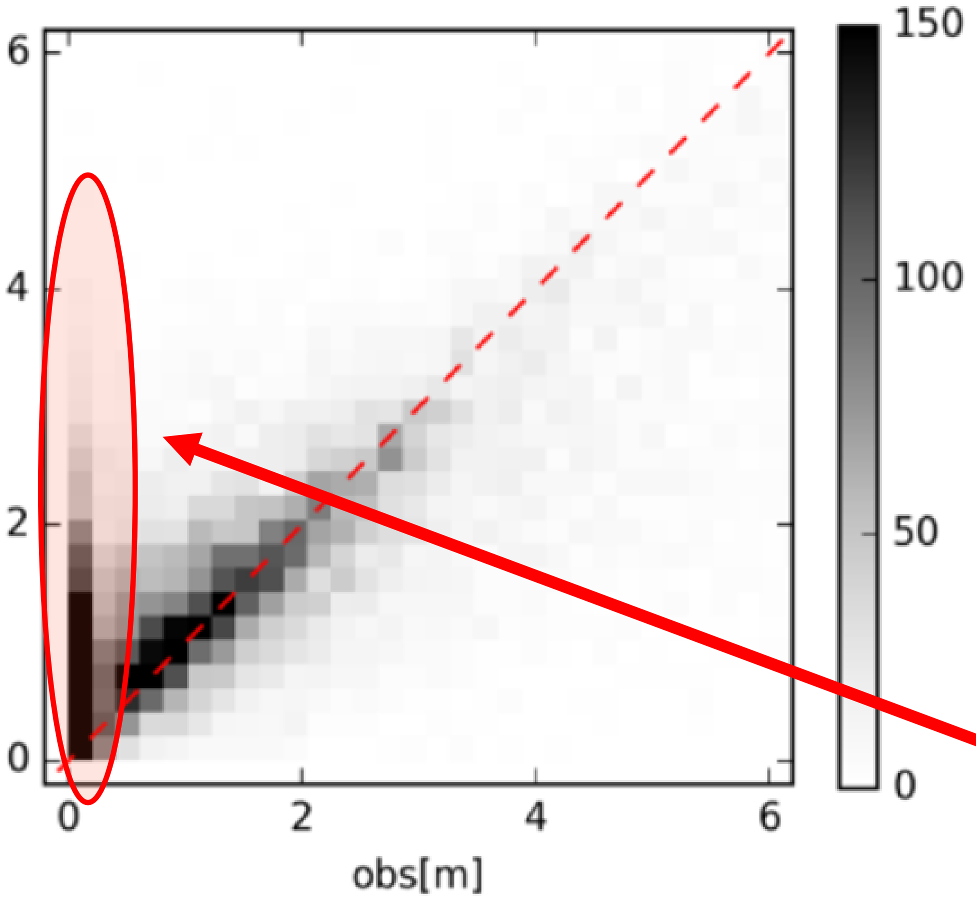
- Depth to groundwater dataset
  - 5.7M observations from ~113,000 locations available
  - 2.4M observations used from ~70,000 locations that met criteria (well depth, unconfined, non-artesian)
  - Cleaned and standardized
  - This dataset is in publication and will (hopefully) be freely available → In the interim contact us if you need it
- 199 predictor variables –precip, et, distance to coast…

How:

- Random Forest classification
- Multiple depth thresholds 0.5 m - 5.0 m
- Multiple probability thresholds



Rethinking the Problem – what's different about our approach?



(b) Depth to water estimates from Koch et al. (2019)

Figure 3

Traditional Approach:

- Predictions precise depths to groundwater (regression | modelling)
- Struggles with shallow groundwater (not a much monitoring here)
- Uncertainty is presented in +- depth
- Decision makers need to interpret uncertainty and decide what this means for their planning objectives

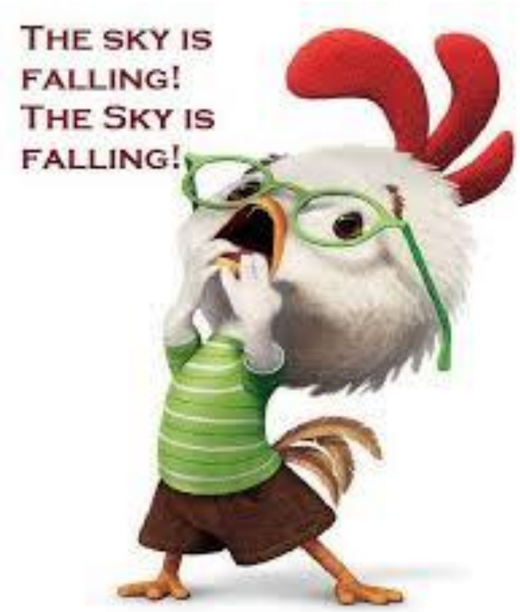
Huston, we have a problem…

Rethinking the Problem – what's different about our approach?

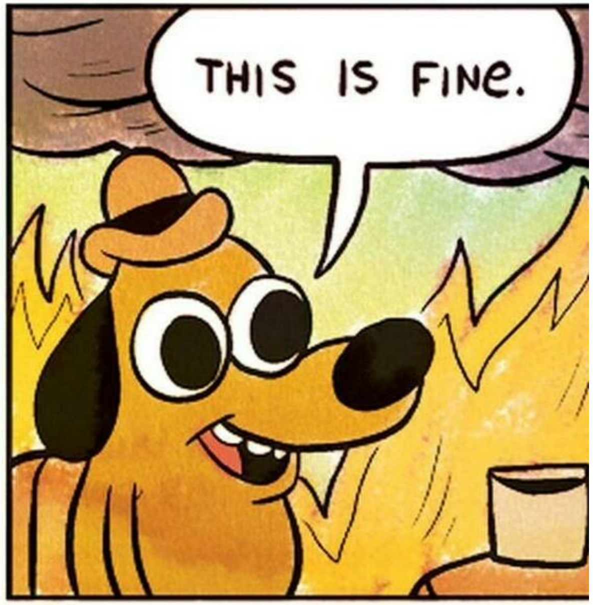
Our Approach

- Answer the question (**classify**): *is groundwater shallower than x m (e.g., 1 m)*
- Classification lets us handle uncertainty as **Type I and Type II errors**

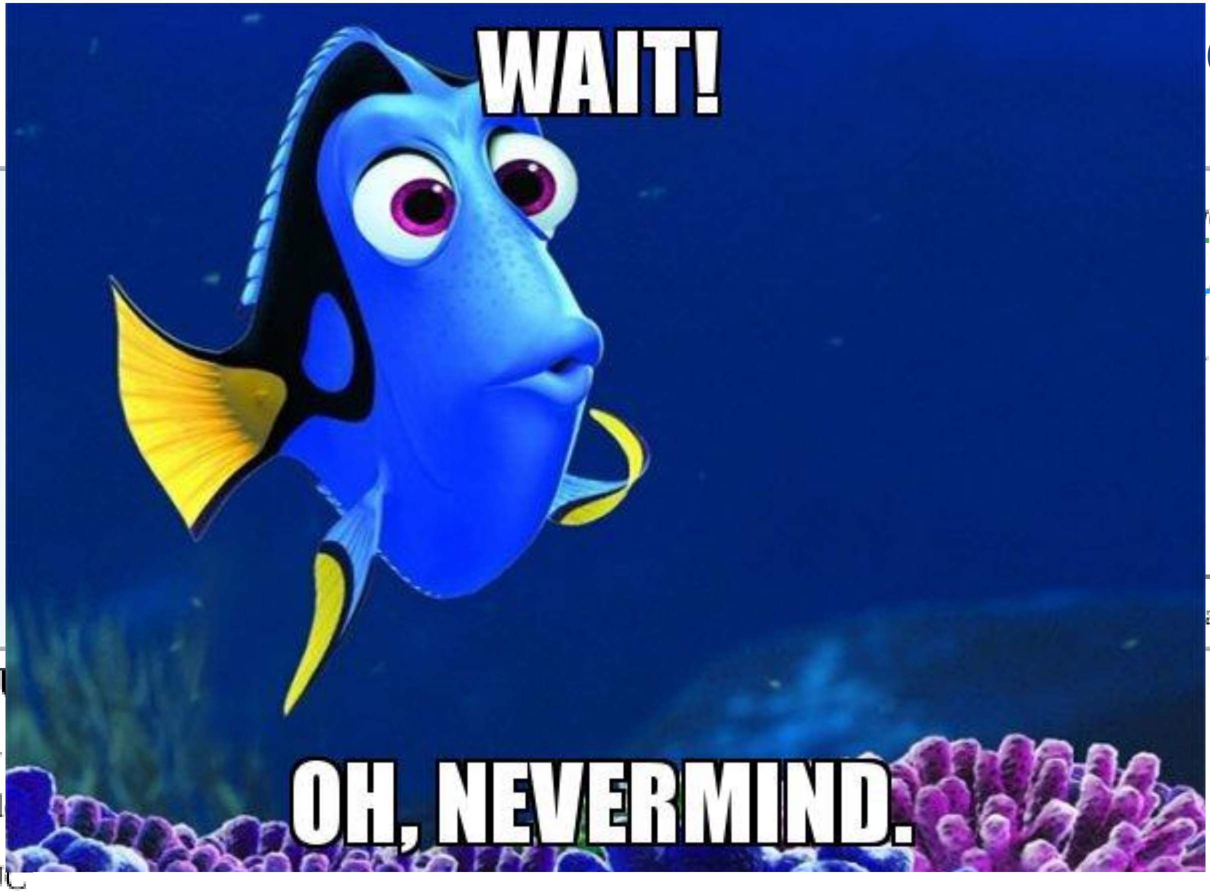
TYPE I Error

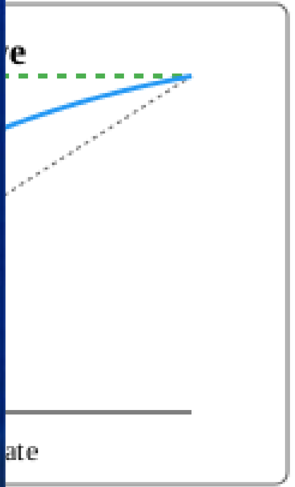


TYPE II Error



c Choice





ROC-A

- Accuracy
  - Simpl
- ROC-AUC
  - True Po
  - False Pc

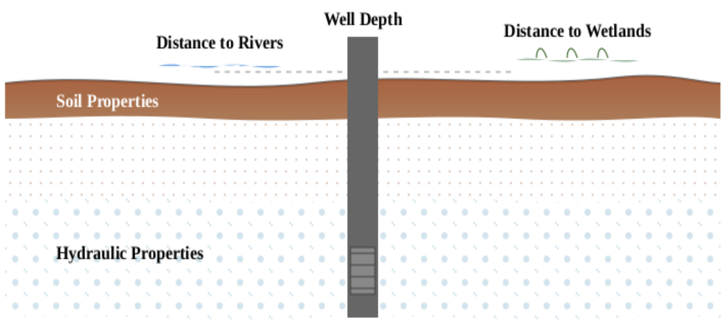
Why ROC

1. Handles in
2. Shows per
3. Less sensit

… Instead, let's talk about what the model means and how the results can be used

Our model's ROC-AUC (0.823 - 0.962) indicates strong predictive power

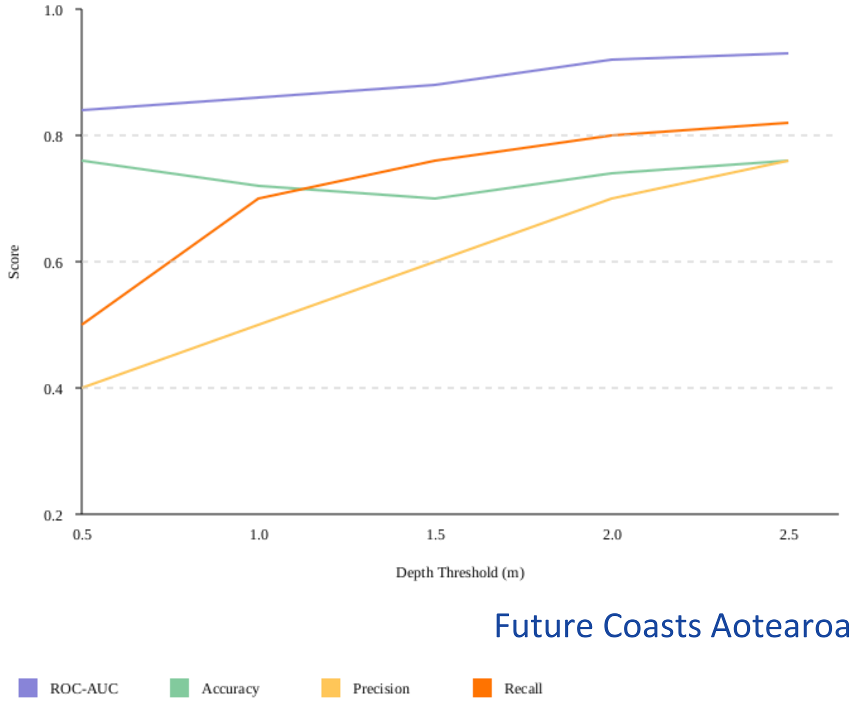
Variables of Importance



Model Variables:

- Soil properties (texture, drainage)
- Hydraulic properties (porosity, conductivity)
- Distance to wetlands
- Distance to rivers
- Well depth

Model Performance Metrics Across Depths



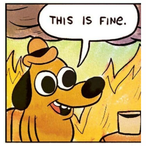
Error Classification

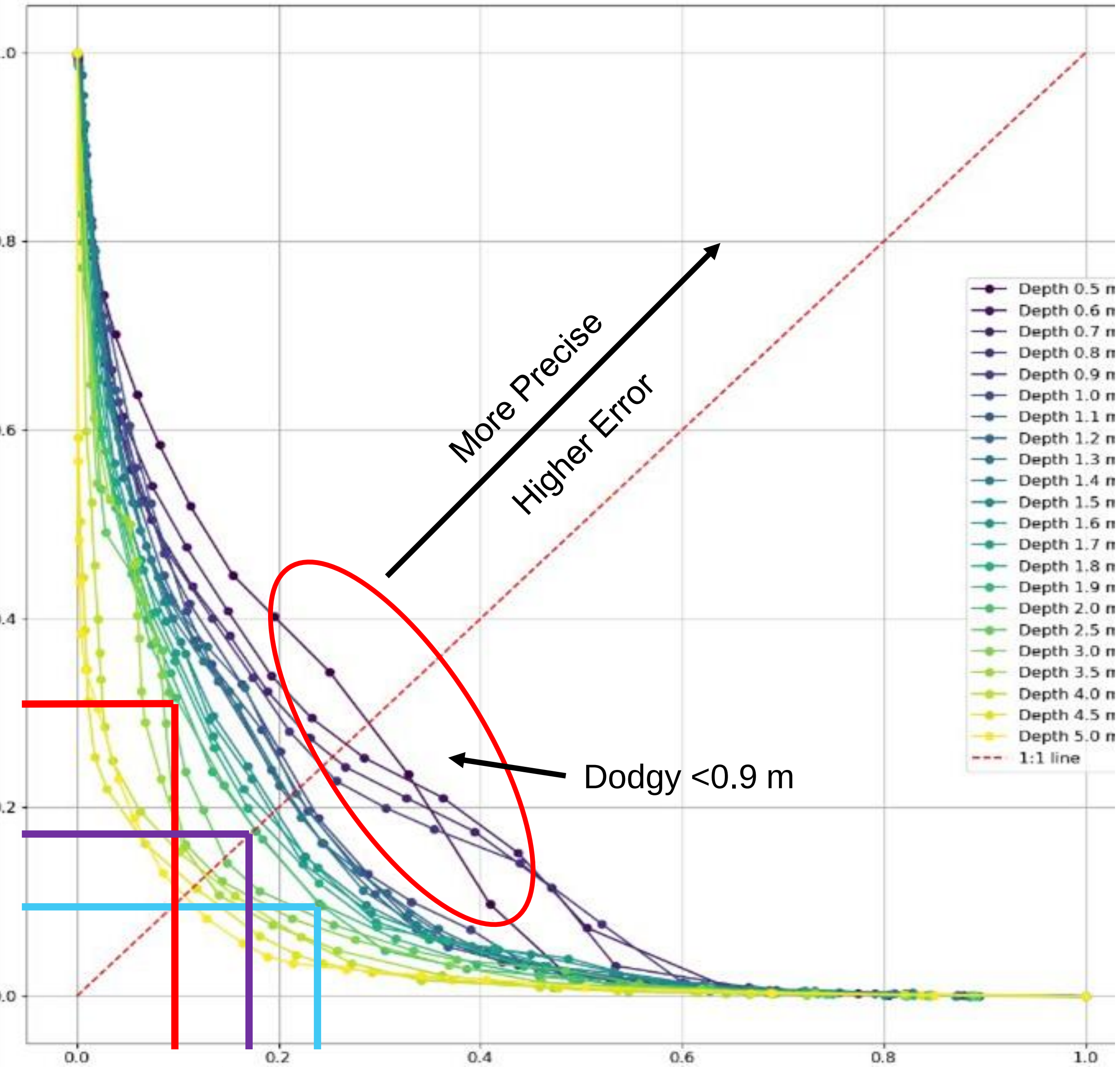
- Trade-off between:
  - **Precision (depth),**
  - **Type I error (sky is falling),**
  - **Type II error (this is fine)**

Red: 10% probability of Type 1,
30% probability of Type II

Purple: 18% probability of Type 1,
18% probability of Type II

Blue: 25% probability of Type 1,
10% probability of Type II

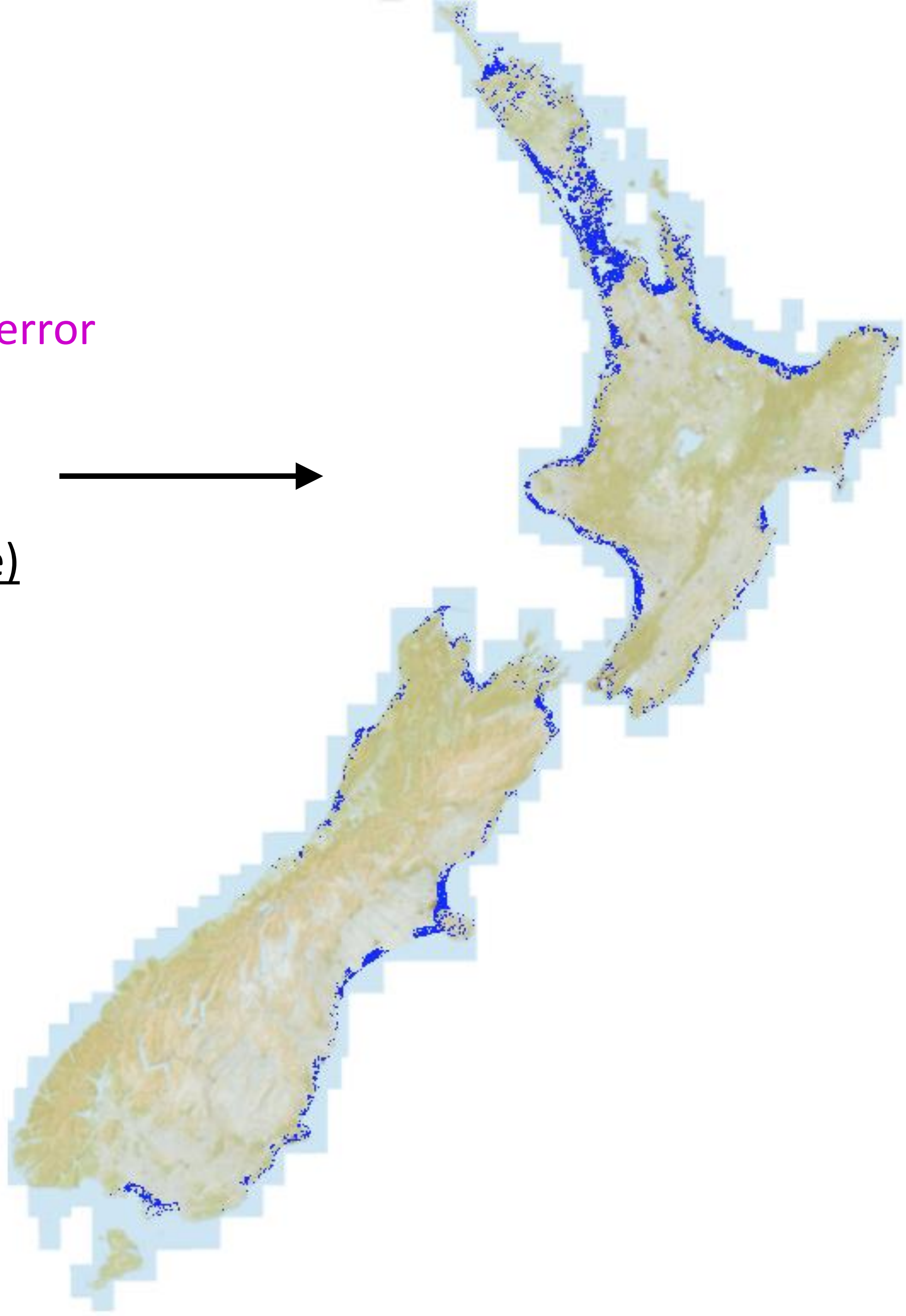




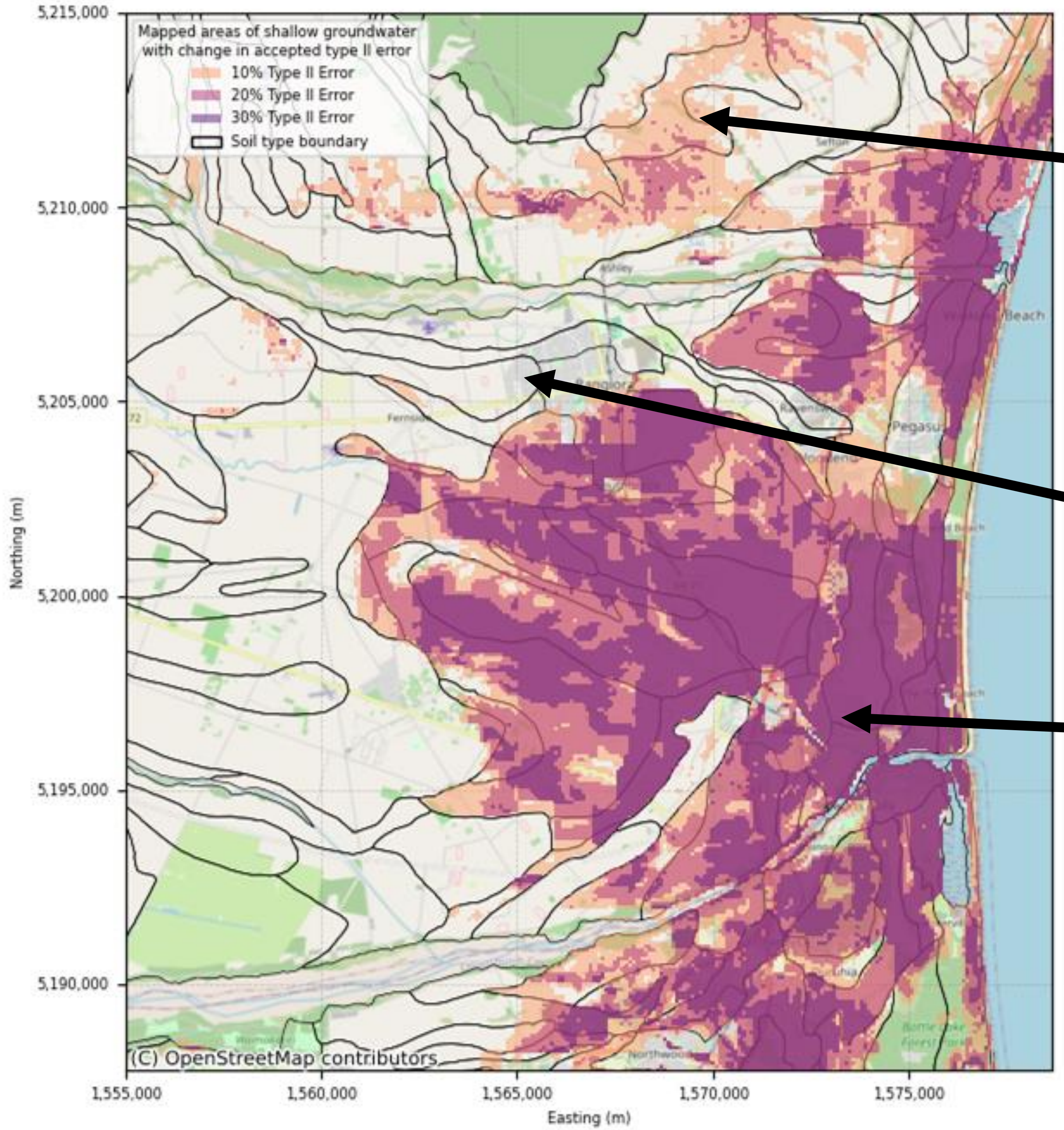
TYPE I Error

Results – within 10 km of coast

- Assuming "shallow gw" = ≤2 m depth and we want 10% Type II error
  - 10% of our deep groundwater is shallow
  - 25% of what's marked shallow is deep
  - Exposed area = **1060 km²** (**0.6%** of NZ ex conservation estate)
- ≤2 m depth and 30% Type II error
  - 30% of our deep groundwater is shallow
  - 10% of our shallow groundwater is actually deep
  - Exposed area = **85 km²** (**0.05%** of NZ ex conservation estate)



A Local Scale Example



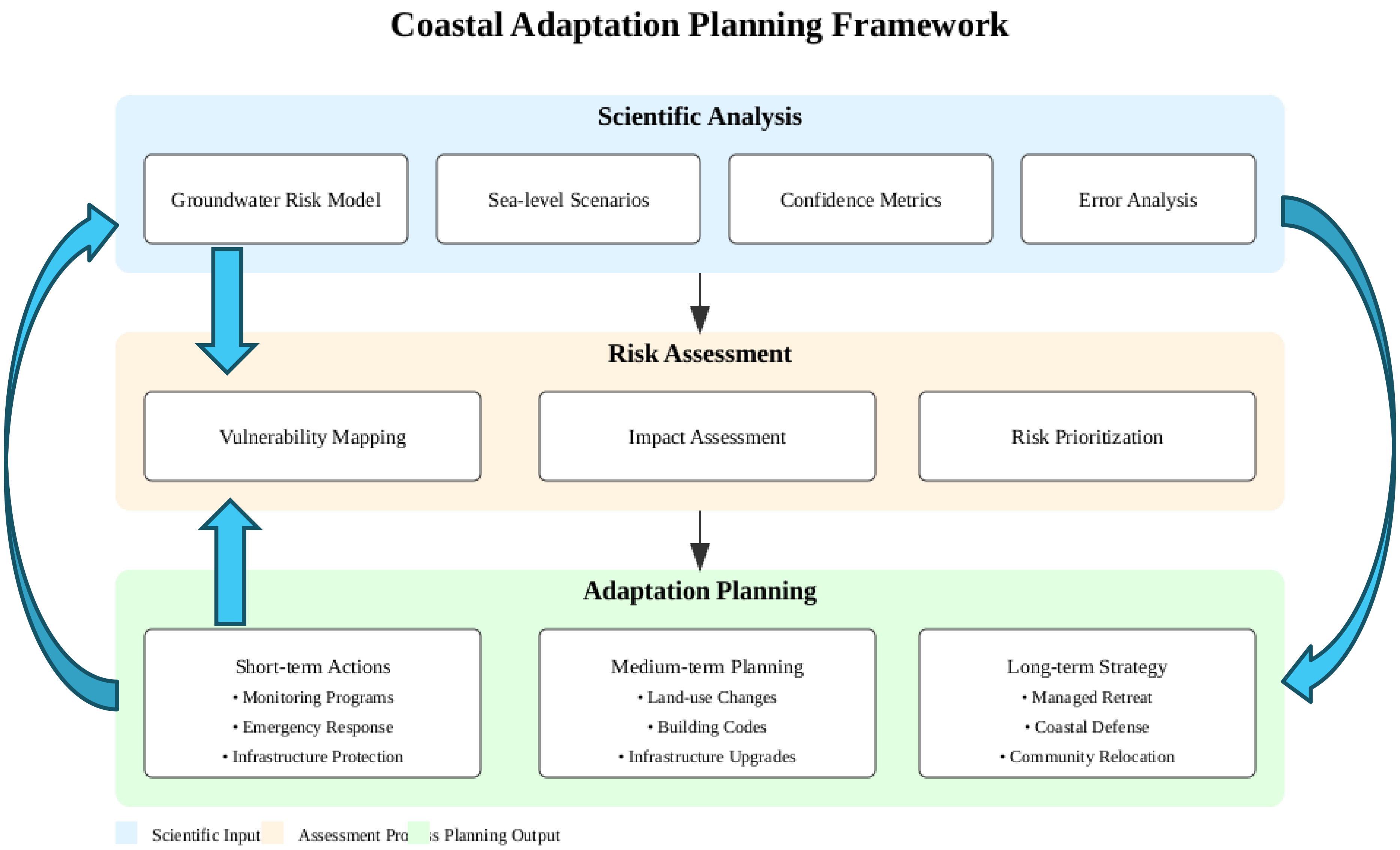
Type II error = 10%
Only for applications requiring a high certainty that we've have not mis-classified shallow groundwater

Shallow groundwater is unlikely

Type II error = 30%
Shallow groundwater even when we don't need to be that risk adverse

Flips the adaptation script

- Flexible risk thresholds for different needs
- Direct integration with planning decisions
- Quantified confidence in predictions



Coastal Adaptation Planning Framework

Scientific Analysis

Groundwater Risk Model

Sea-level Scenarios

Confidence Metrics

Error Analysis

Risk Assessment

Vulnerability Mapping

Impact Assessment

Risk Prioritization

Adaptation Planning

Short-term Actions
- Monitoring Programs
- Emergency Response
- Infrastructure Protection

Medium-term Planning
- Land-use Changes
- Building Codes
- Infrastructure Upgrades

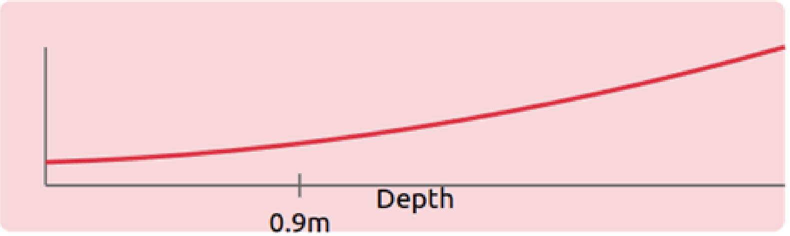
Long-term Strategy
- Managed Retreat
- Coastal Defense
- Community Relocation

🟦 Scientific Input 🟧 Assessment Pro 🟩s Planning Output

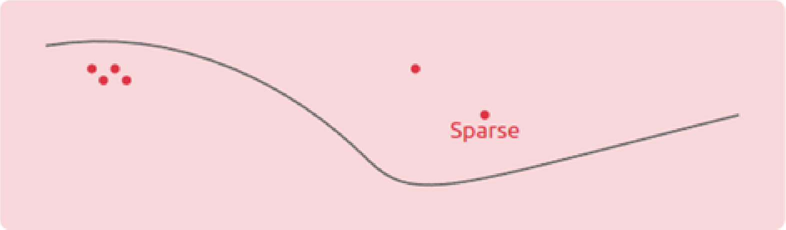
Current Limitations & Future Directions

Current Limitations

1. Reduced Performance < 0.9m Depth



2. Data Sparsity in Some Regions

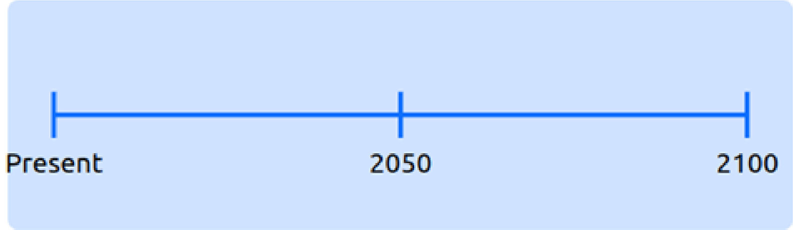


3. Can't Predict Sea Level Rise Effects

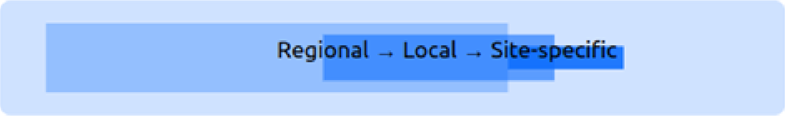
Future Directions

1. Sea Level Rise Integration

2. Enhanced Temporal Predictions



3. Local-scale Refinements



Conclusions

1. Novel approach to shallow groundwater modelling
2. Prioritises identification of at-risk areas
3. Balances accuracy with practical utility
4. Provides framework for coastal adaptation
5. Bridges gap between science and policy

What are your questions?

- Thank you for your attention
- Contact: patrick@komanawa.com

Get out your calculators! We’re going to talk machine learning, linear algebra and detailed model performance