Model-driven decision support for monitoring network design based on analysis of data and model uncertainties: methods and applications

Velimir V Vesselinov[1], Dylan Harp[1], Danny Katzman[2]

[1] Computational Earth Sciences, Earth and Environmental Sciences,
[2] Environmental Programs,
Los Alamos National Laboratory (LANL), Los Alamos, NM

LA-UR-12-26681

AGU Fall Meeting 2012
H32F. Uncertainty Quantification and Parameter Estimation:
Impacts on Risk and Decision Making
December 5, 2012
San Francisco, CA

Outline

- ✧ Model-driven (model-based) decision support
- ✧ Probabilistic vs Non-Probabilistic Decision Methods
- ✧ Information Gap (info-gap) Decision Theory
- ✧ Information Gap (info-gap) Applications:
  - Monitoring Network Design
  - Contaminant Remediation through Source Control
- ✧ Decision Support for Chromium contamination site @ LANL

✧ MADS: Model Analyses & Decision Support
Open source C/C++ computational framework
Publications, examples & tutorials @
http://mads.lanl.gov

✧ ASCEM: Advanced Subsurface Computing for Environmental Management; Multi-national lab code development project http://ascemdoe.org (U.S. DOE)

Model-driven (model-based) decision support

- provides decision makers (DM) with model analysis of decision scenarios taking into account site data and knowledge including existing uncertainties (uncertainties in conceptualization, model parameters, and model predictions)
- Model analysis: evaluation, ranking and optimization of alternative decision scenarios
- Decision metric(s): e.g. contaminant concentration at a monitoring well (environmental risk at a point of compliance)
- Decision goal(s): e.g. no exceedance of MCL at a compliance point and/or increase chance of detecting exceedance of MCL at a monitoring well
- Decision scenarios: combinations of predefined activities to achieve the decision goal(s)

Model-driven decision support (cont.)

- <u>Activities</u>:
  - data acquisition campaigns
  - field/lab experiments
  - monitoring
  - remediation
- Activities are analyzed in terms of their impact on decision making process (decision uncertainties)
- <u>Decision uncertainties</u>: uncertainties associated with selection of optimal decision scenarios, or performance of specific decision scenarios
- <u>The Game:</u> Decision maker (DM) vs Nature

<u>Important:</u>

- activities are selected only to reduce decision uncertainties
- activities are not selected to reduce model or parameter uncertainties per se (unconstrained problem).

Non-Probabilistic Decision Methods

- Lack of knowledge or information precludes decision analyses requiring unbiased probabilistic distributions or frequency of occurrence (e.g. Bayesian approaches)
- Severe uncertainties (black swans, dragon kings) can have important impact in the decision analyses
- Non-probabilistic decision methods can be applied to effectively incorporate lack of knowledge and severe uncertainties in decision making process
  - Minimax (Maximin) Theory (Wald, 1951)
  - Information Gap Decision Theory (Ben-Haim, 2006)
- Non-Probabilistic and Probabilistic methods can be coupled (e.g. unknown probability distribution parameters can be a subject of non-probabilistic analysis, e.g. info-gap)

Information Gap Decision Theory

- **Nominal ("best") model prediction intended for decision making (based on nominal / "best estimates" model parameter set)**
- **Decision metric(s) / performance goal(s)**
- **Decision scenarios: vector of alternative decisions *d* to compare**
- **Info-Gap Uncertainty Model (info-gap uncertainty metric = *α*)**
  - **energy bound (functional uncertainties: objective function, forcing functions, etc.)**
  - **envelope bound (domain uncertainties: model parameters, calibration targets, etc.)**
  - **nested sets of uncertain model entities ranked by the largest information gap *α* that can be included in the set**
  - **uncertain model entities: parameters, calibrations, functions, etc. with info-gap uncertainties**
  - **e.g. $U(\alpha,T) = \{ T: abs(T-T') < \alpha \}$ where T' are the nominal values for uncertain model entities**
- **Model predictions $C(d)$ constrained by $U(\alpha,T)$**

Ben-Haim (2006). *Info-gap decision theory: decisions under severe uncertainty*. Academic Press.

Information Gap Decision Theory

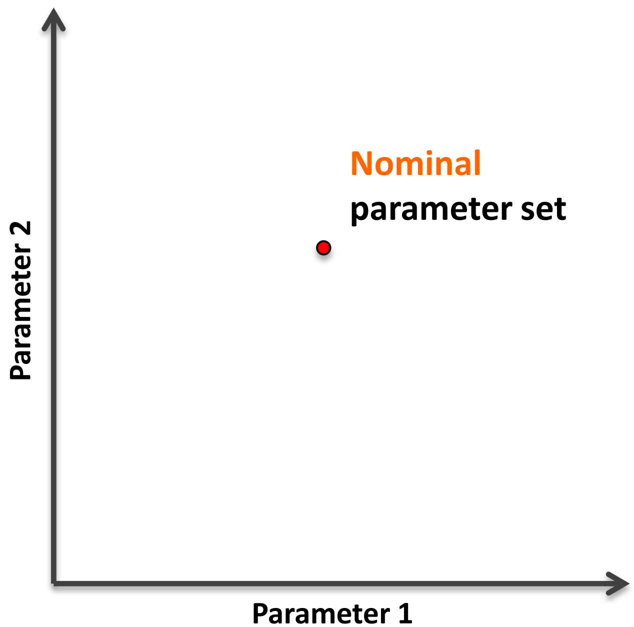
- **Decision uncertainty** is bounded by **robustness** and **opportuness** functions
- **Robustness** function (immunity to failure of alternate decisions d)
  - defines the maximum horizon of uncertainty
  - $R(d) = \max\{\ \alpha$: performance goal is satisfied $\}$

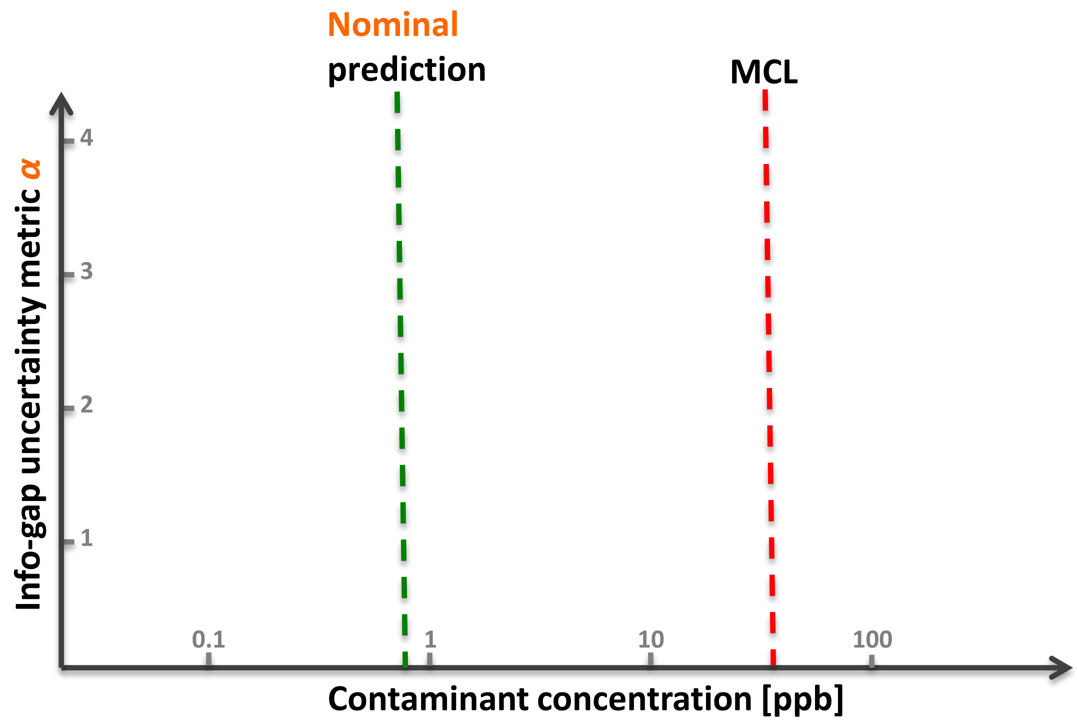
    e.g. $R(d) = \max\{\ \alpha: (\ \max C(d)\ ) < MCL\ \}$
- **Opportuness** function (immunity to windfall of alternate decisions d)
  - defines the minimum horizon of uncertainty
  - $O(d) = \min\{\ \alpha$: performance goal is satisfied $\}$

    e.g. $O(d) = \min\{\ \alpha: (\ \min C(d)\ ) < MCL\ \}$
- Analyses based on Decision Robustness and/or Decision Opportuness:
  - Model selection
  - Remedy selection
  - Performance assessment
  - …

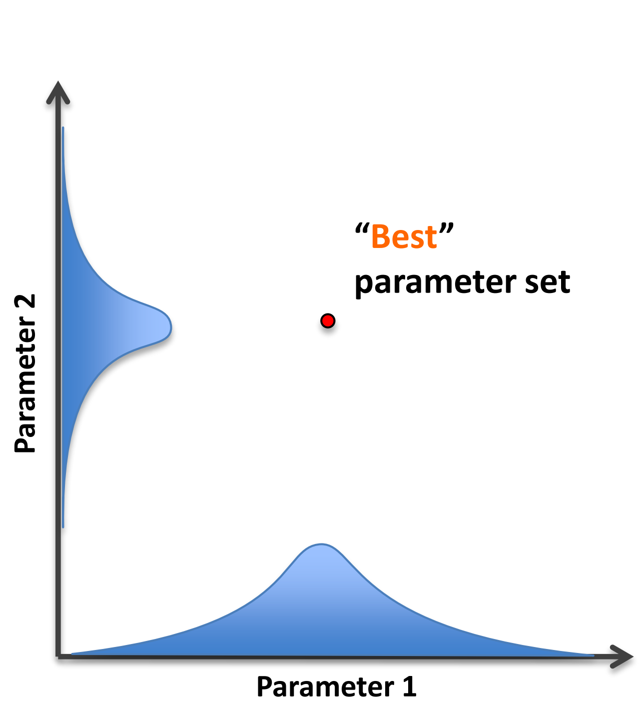
Ben-Haim (2006). *Info-gap decision theory: decisions under severe uncertainty*. Academic Press.

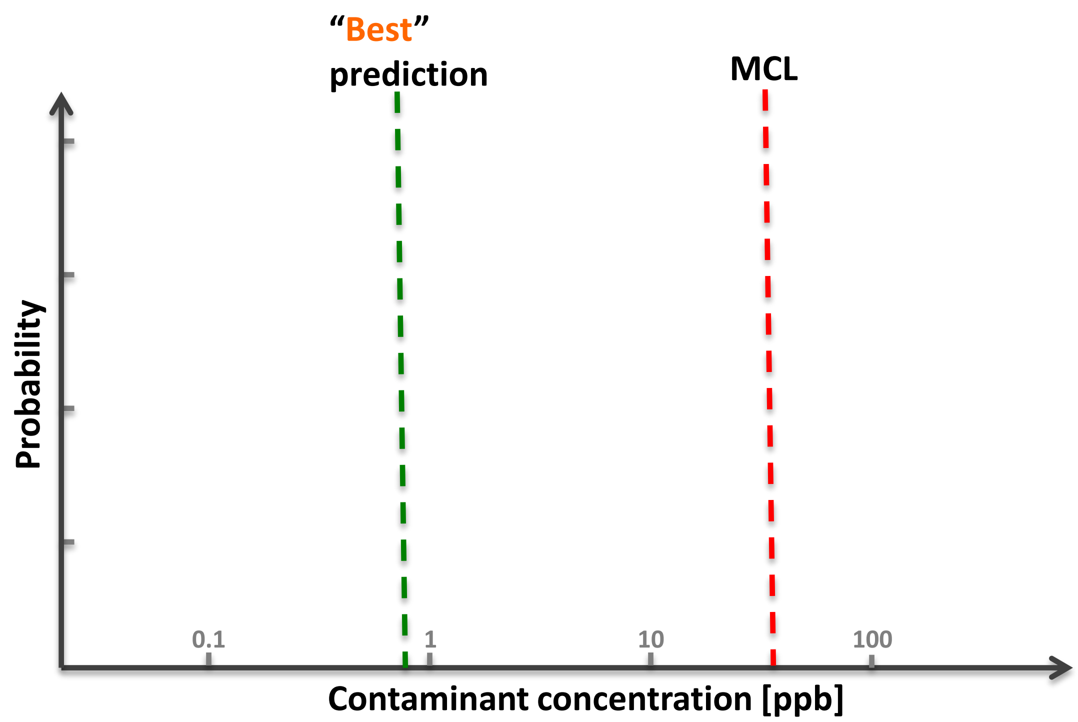
Info-Gap Analysis: Model parameters



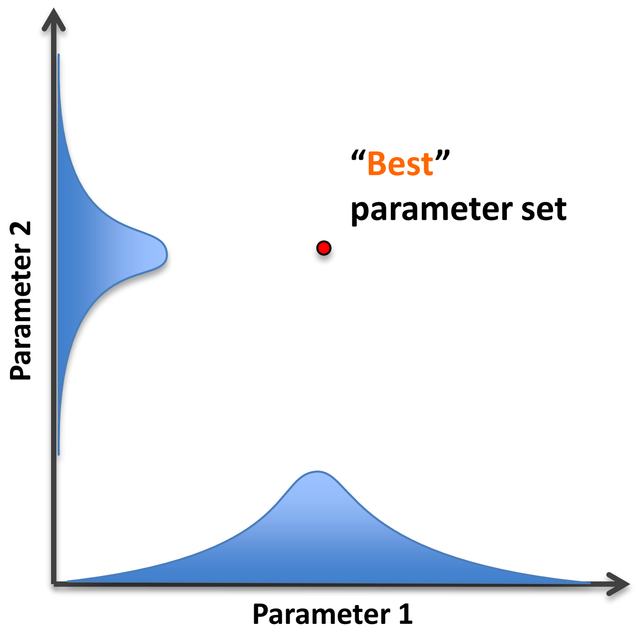


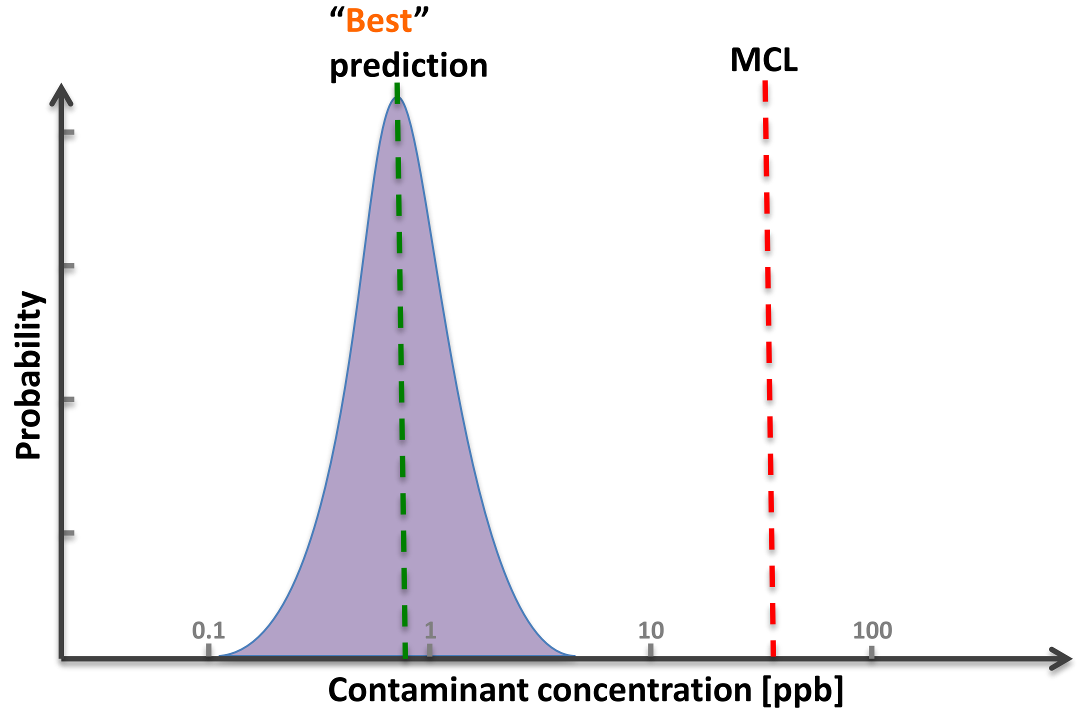
Bayesian Analysis: Model parameters



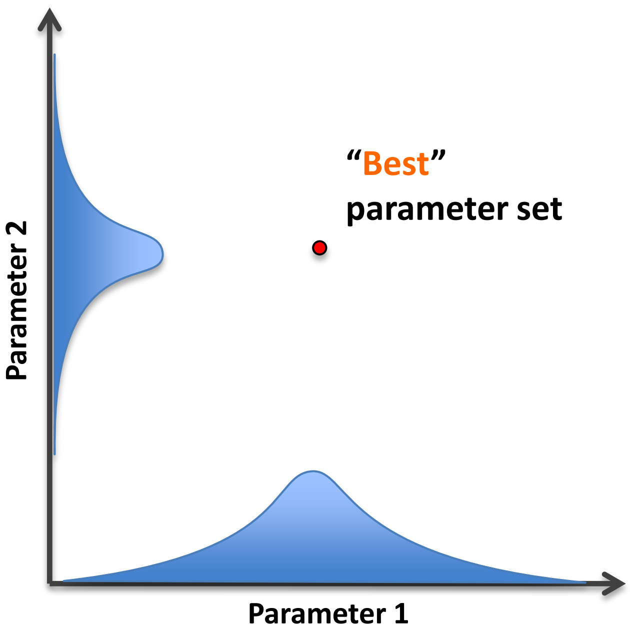


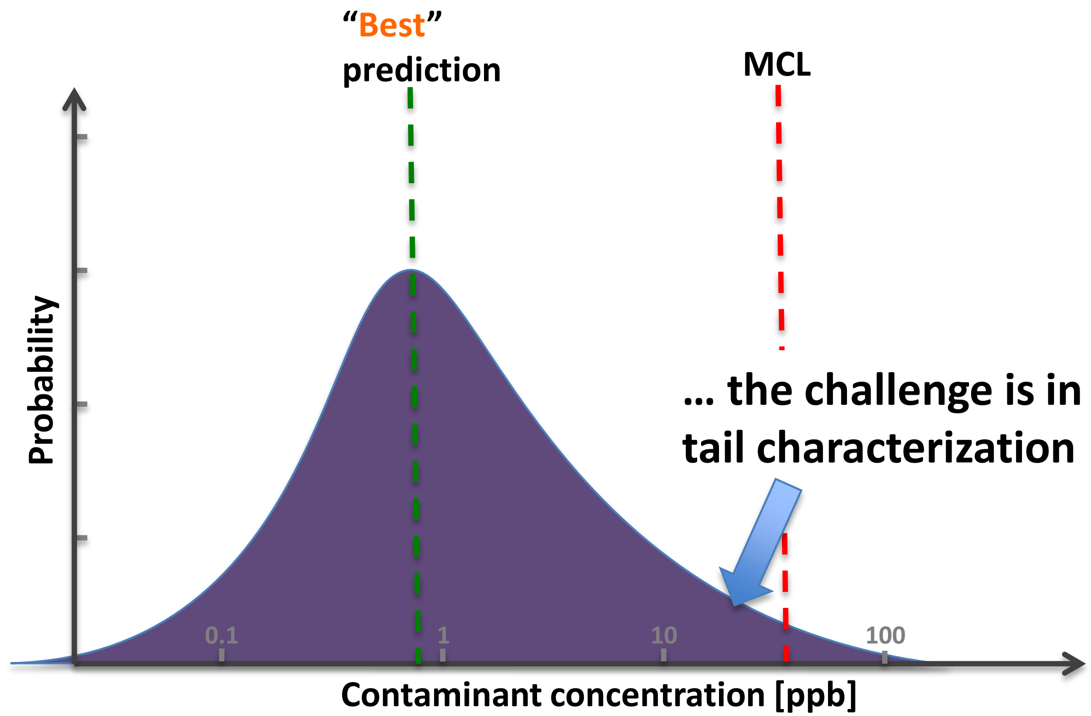
Bayesian Analysis: Model parameters



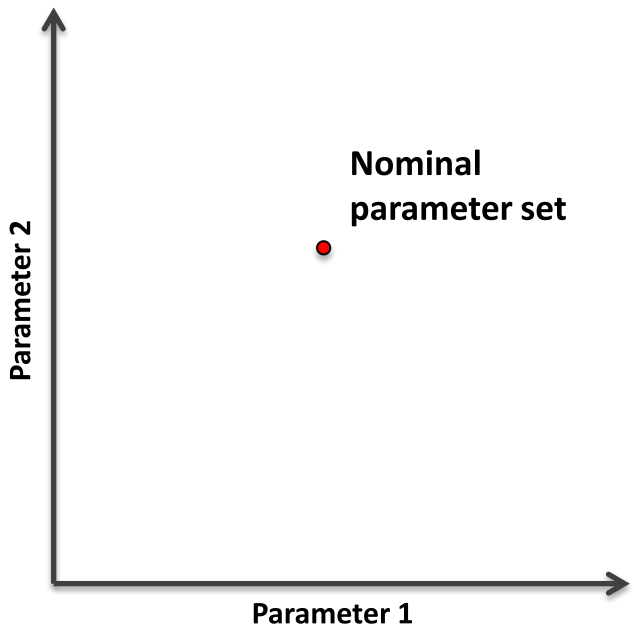


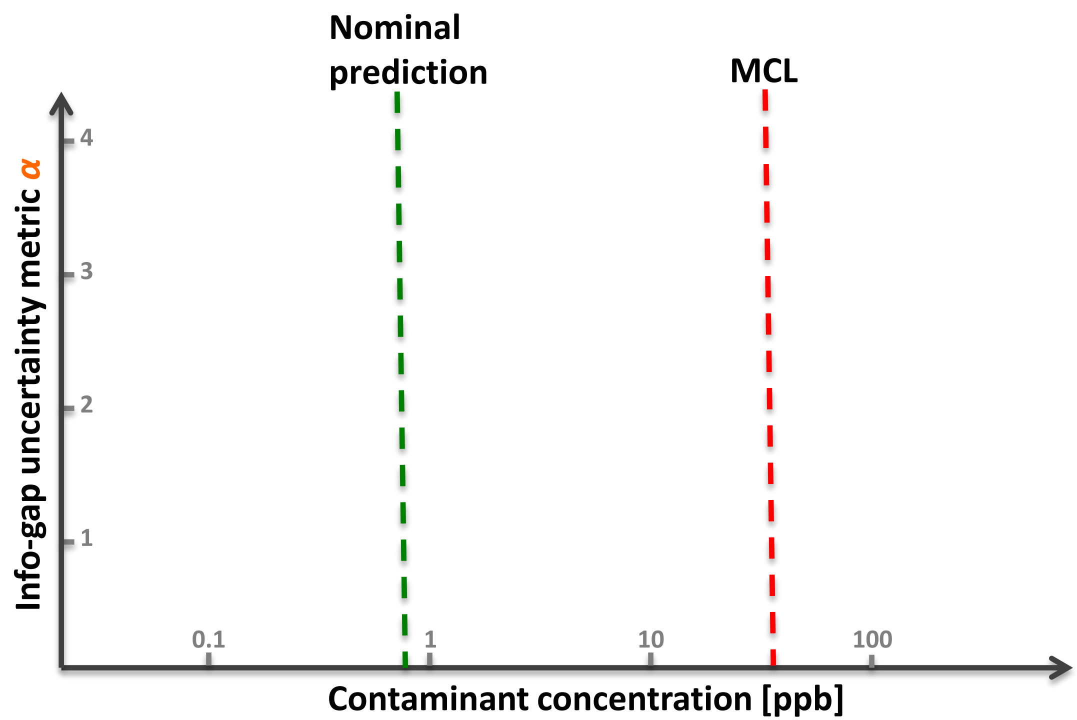
Bayesian Analysis: Model parameters



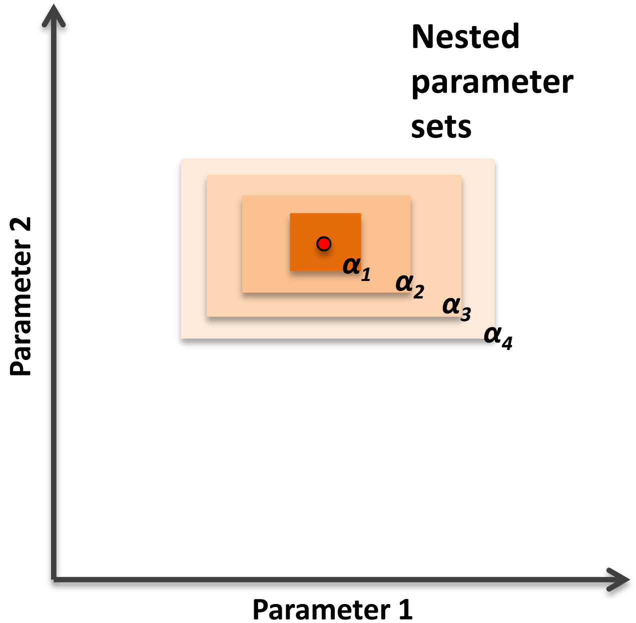


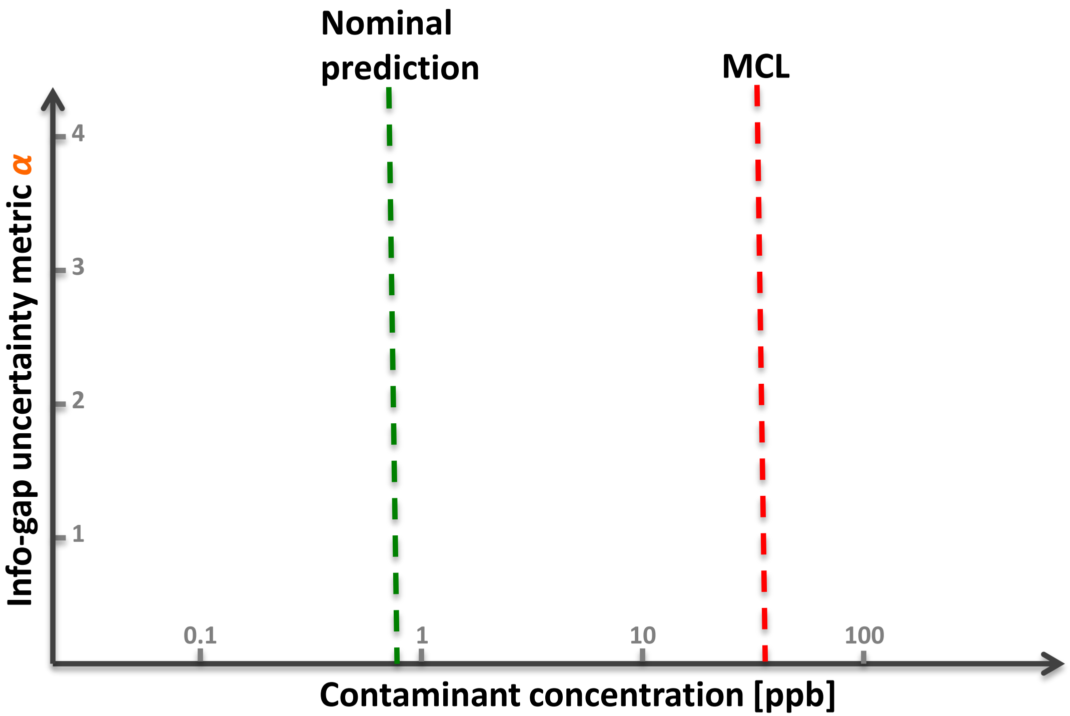
Info-Gap Analysis: Model parameters (envelope bounds)





Info-Gap Analysis: Model parameters (envelope bounds)

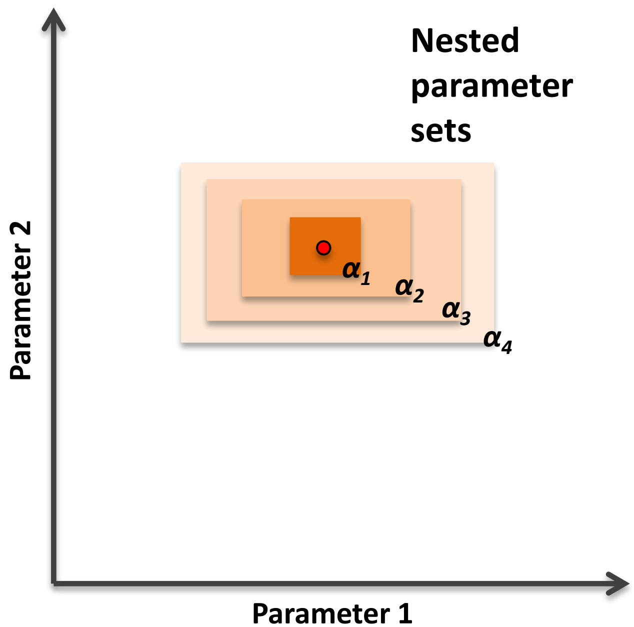


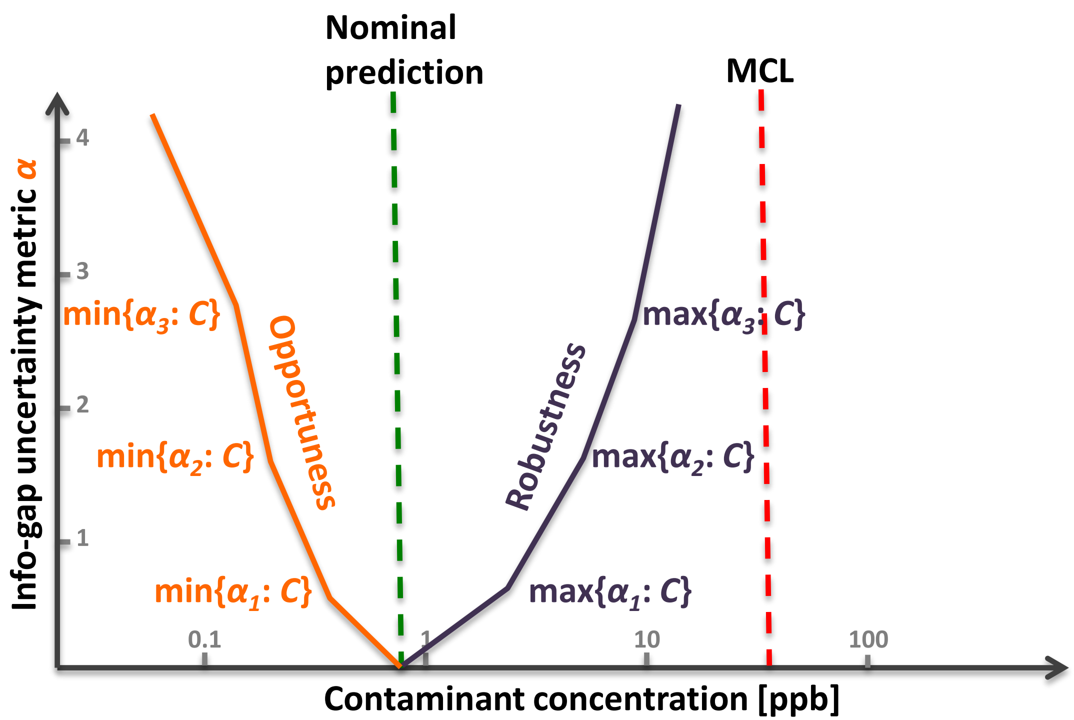


info-gap uncertainty metric = α

$\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$

Info-Gap Analysis: Model parameters (envelope bounds)

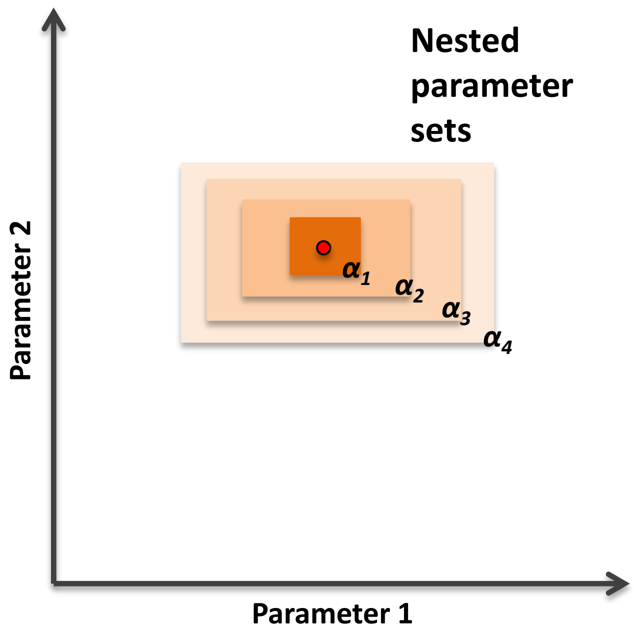


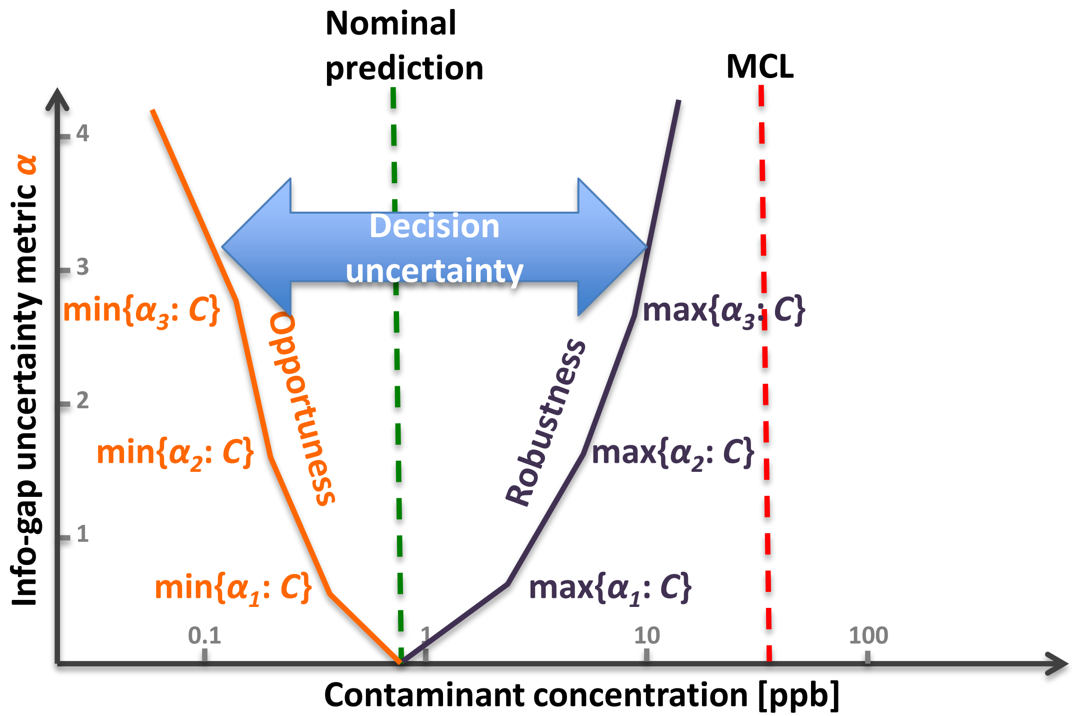


info-gap uncertainty metric = α

$\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$

Info-Gap Analysis: Model parameters (envelope bounds)

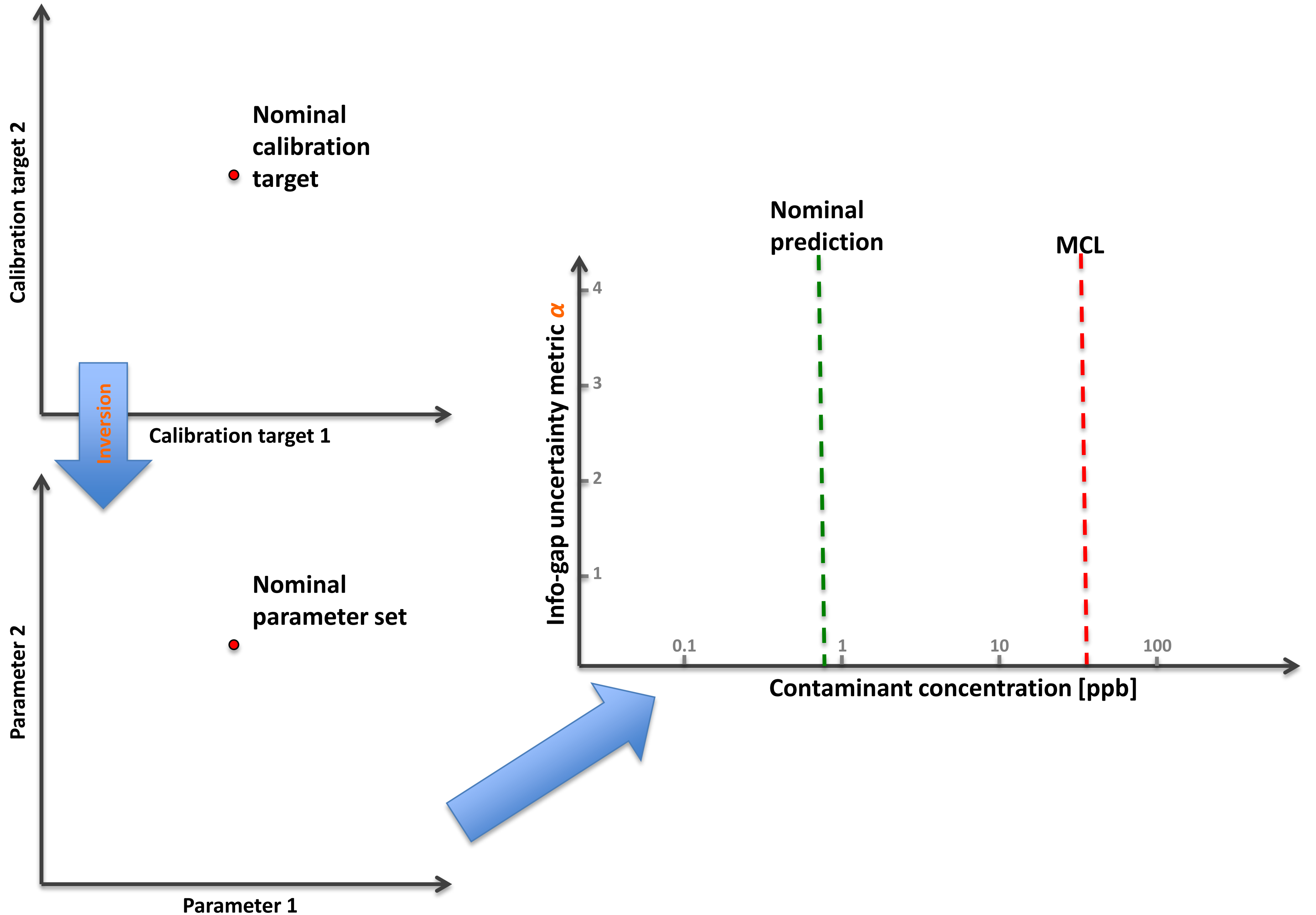




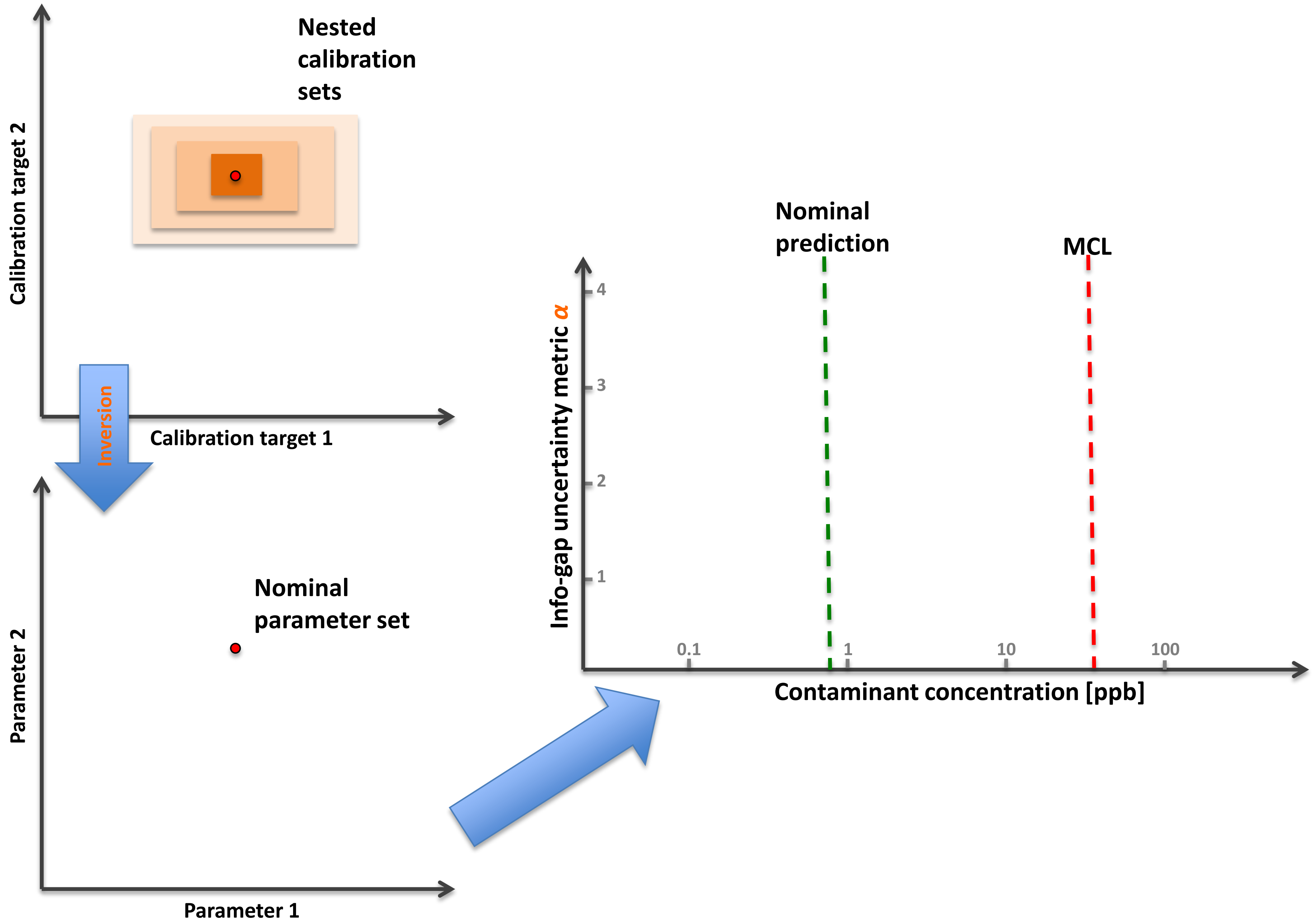
info-gap uncertainty metric = α

$\alpha_1 < \alpha_2 < \alpha_3 < \alpha_4$

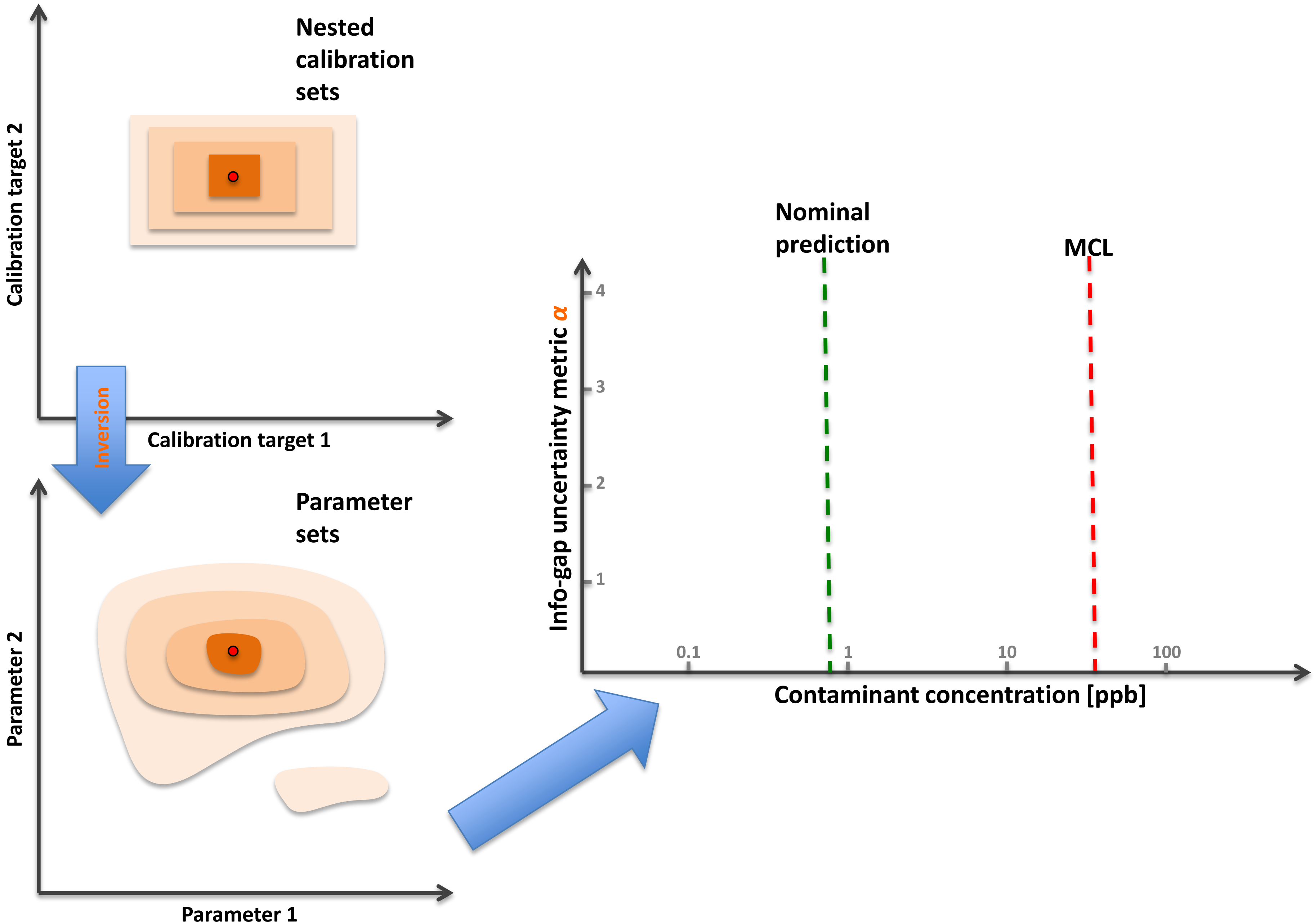
Info-Gap Analysis: Calibration Targets (envelope bounds)



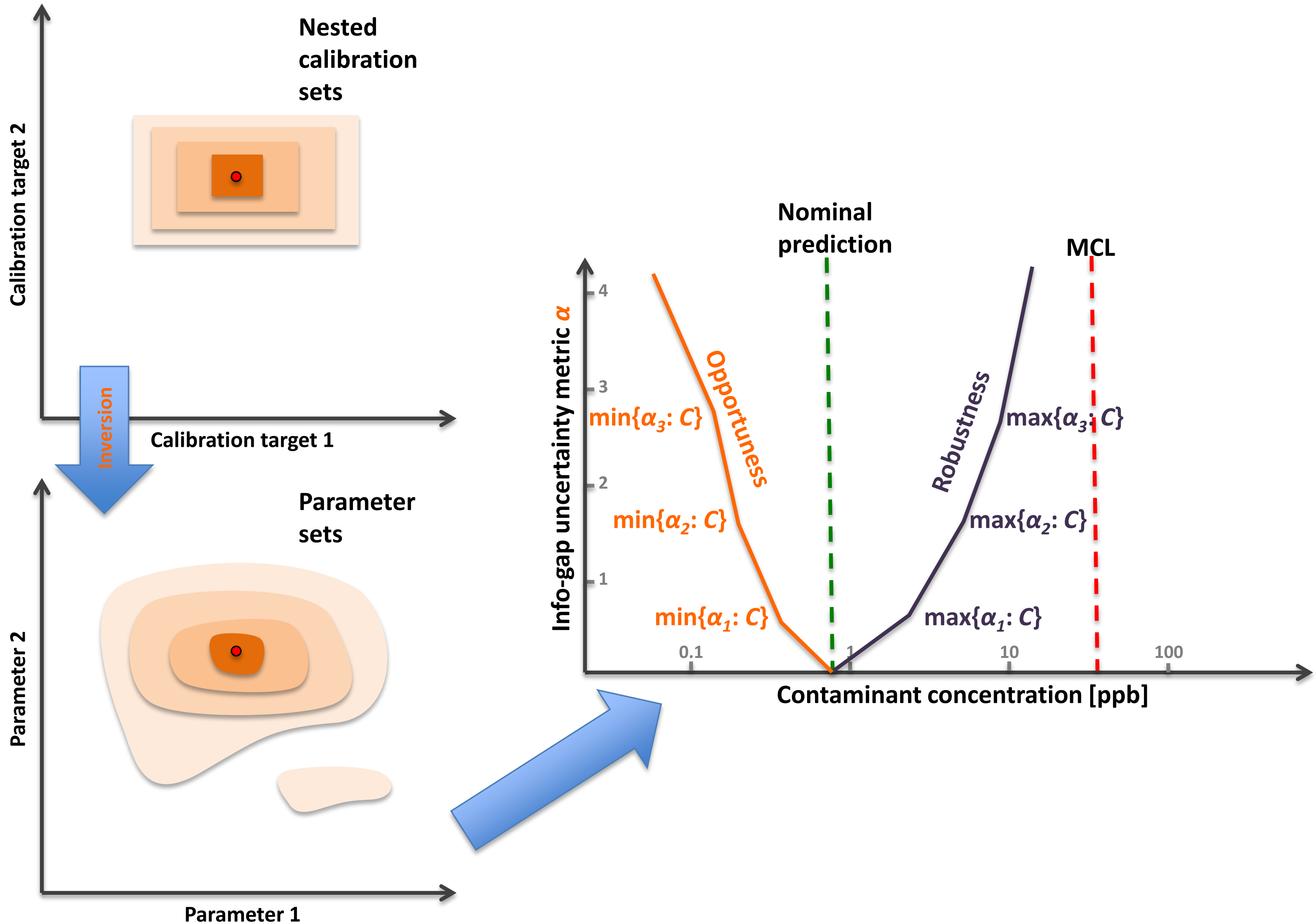
Info-Gap Analysis: Calibration Targets (envelope bounds)



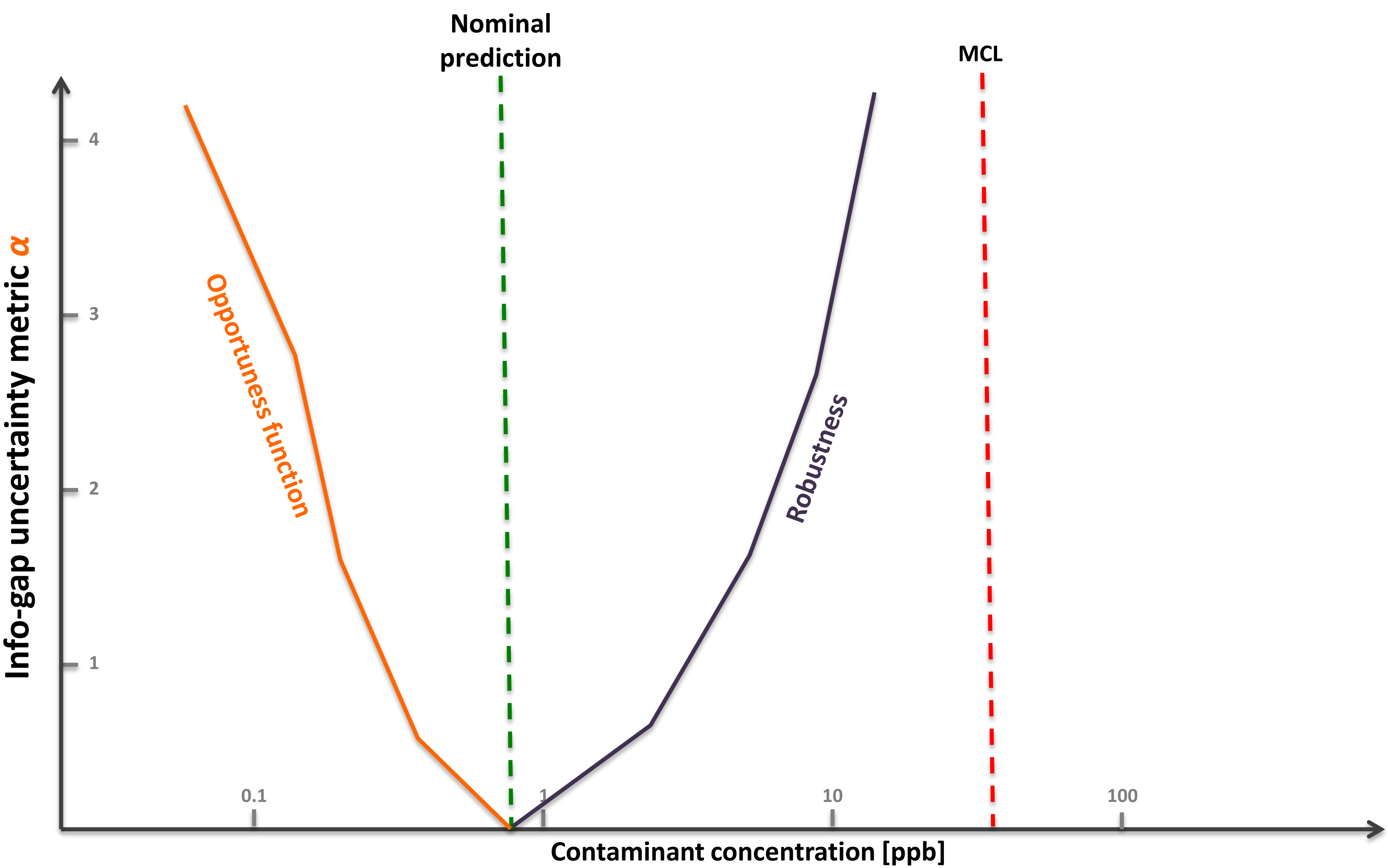
Info-Gap Analysis: Calibration Targets (envelope bounds)



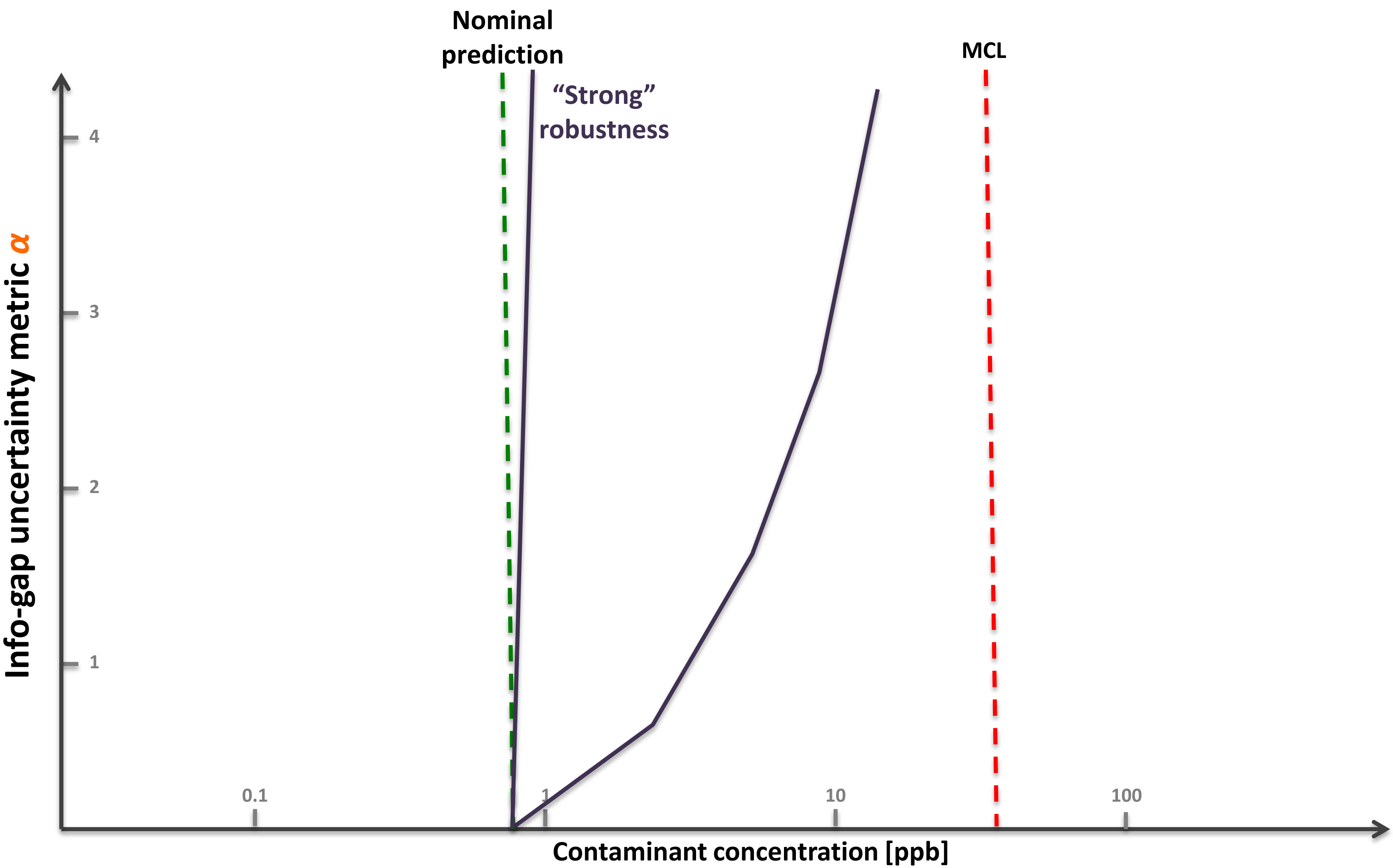
Info-Gap Analysis: Calibration Targets (envelope bounds)



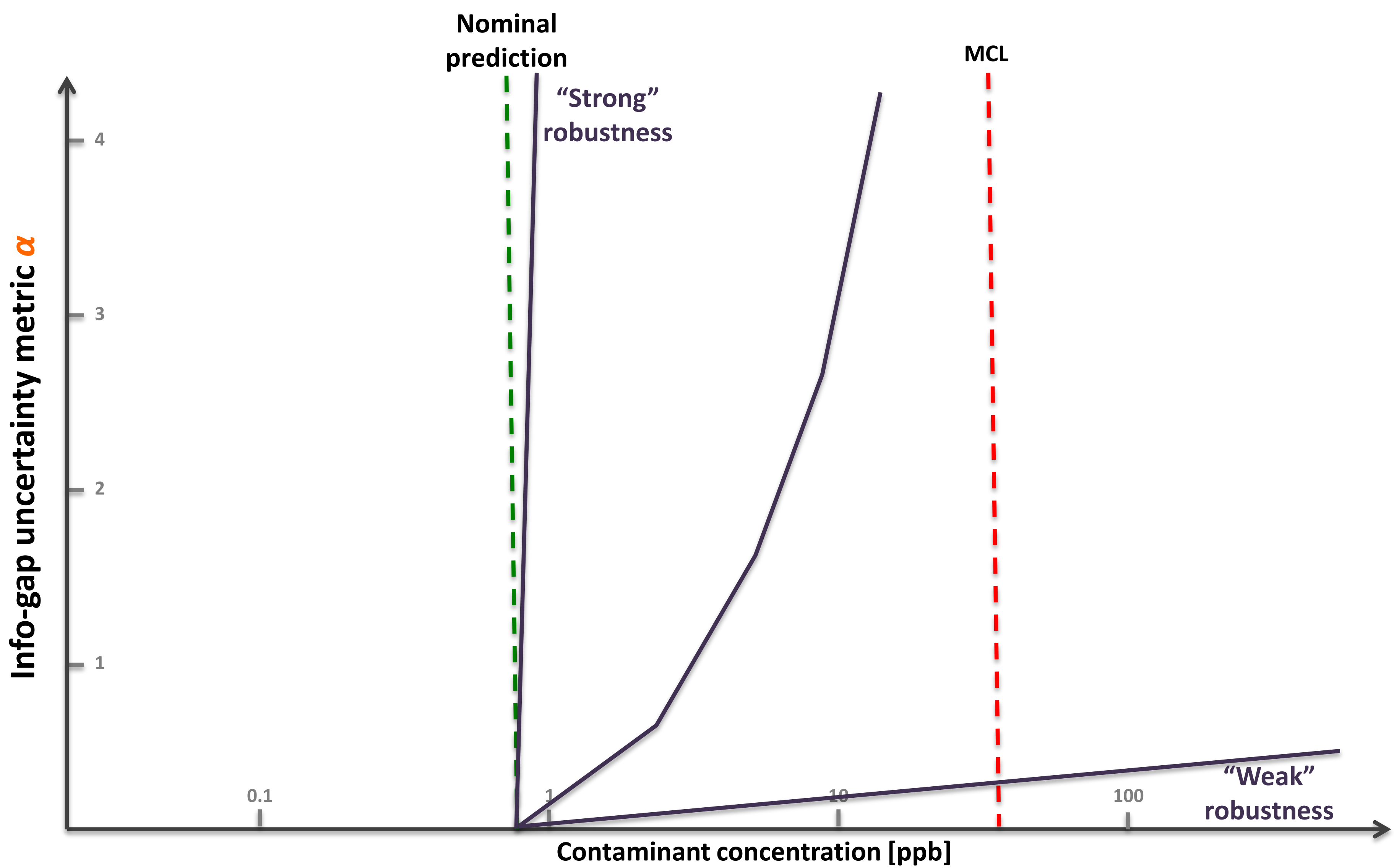
Info-Gap Analysis: Decision selection based on robustness



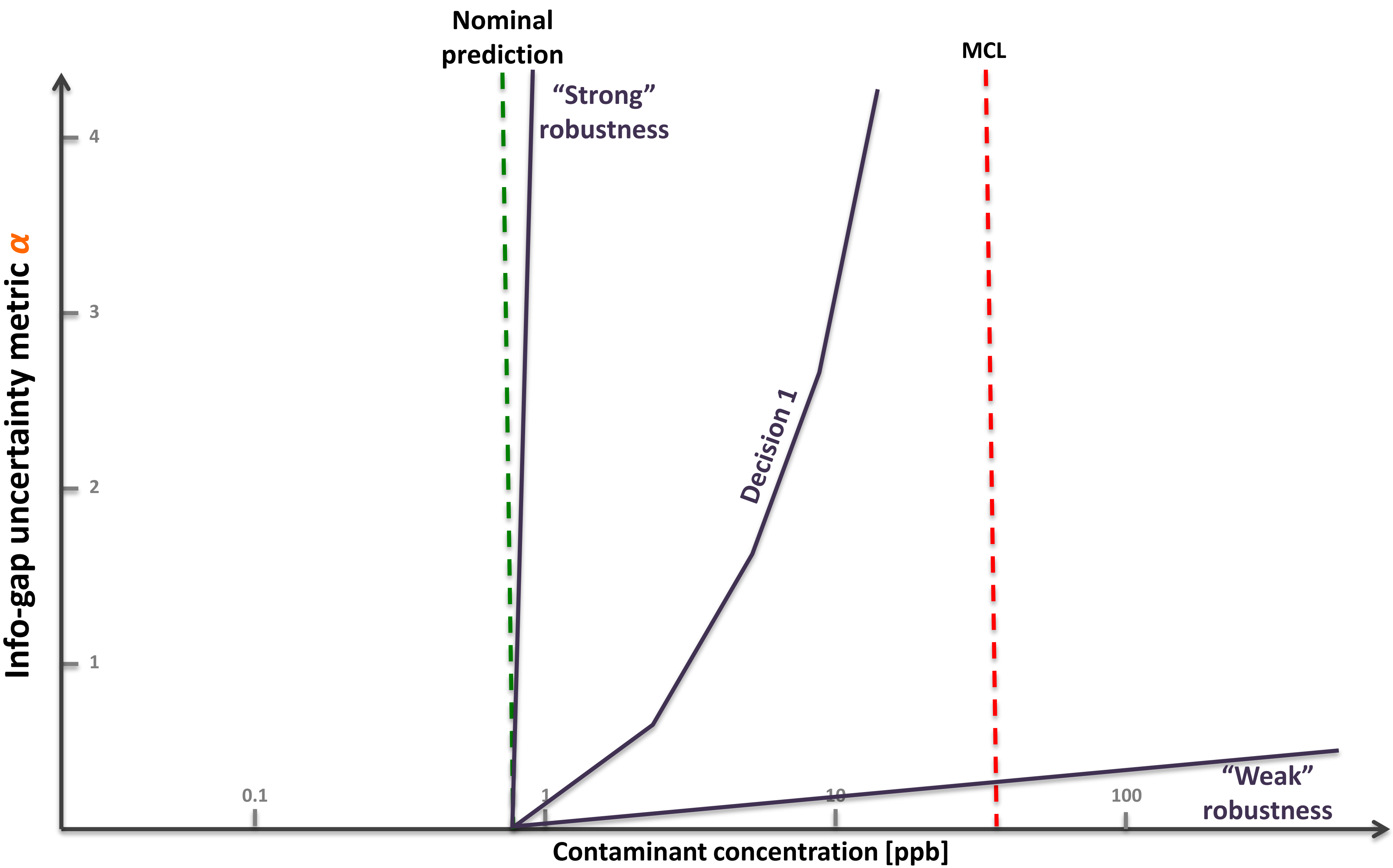
Info-Gap Analysis: Decision selection based on robustness



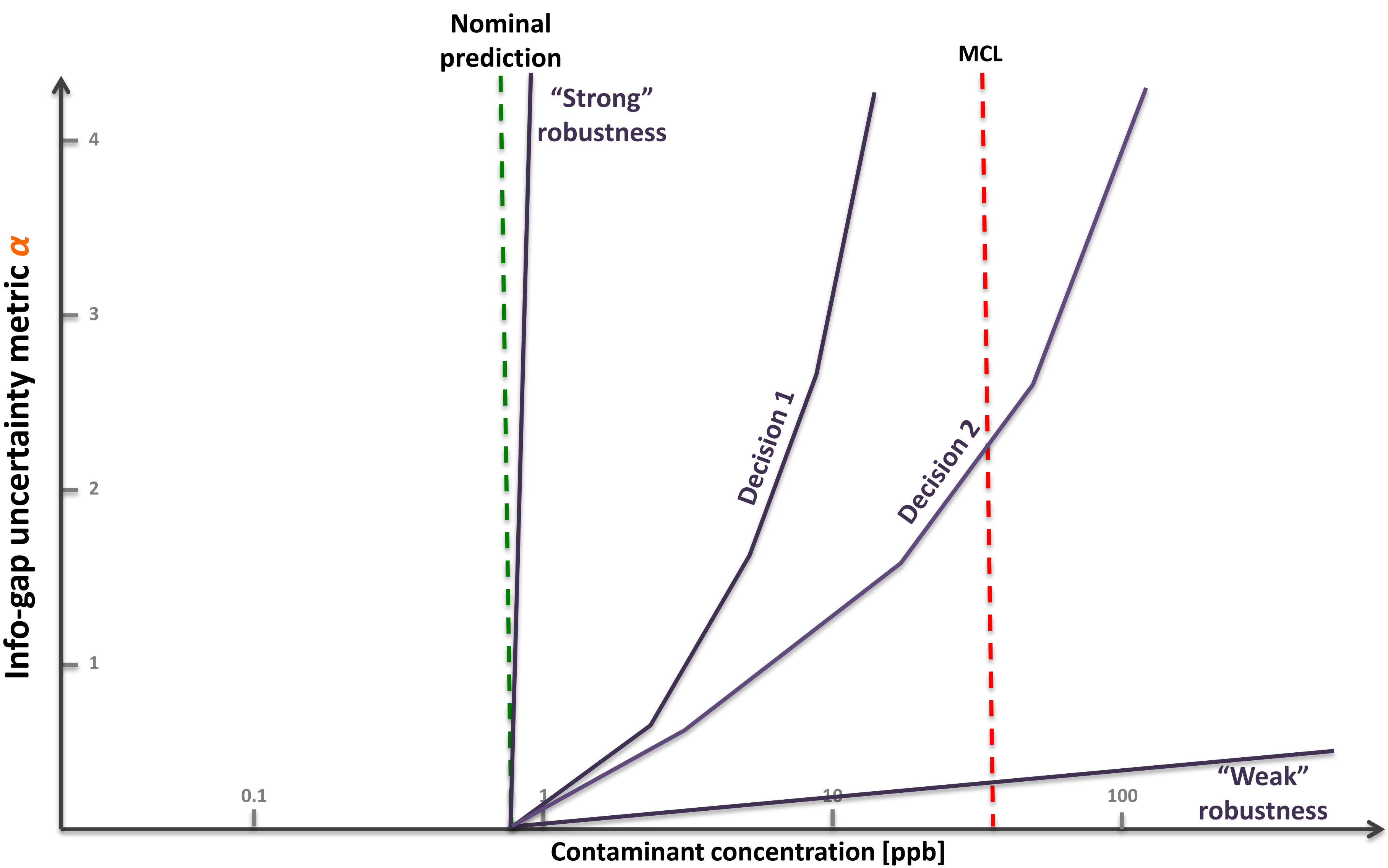
Info-Gap Analysis: Decision selection based on robustness



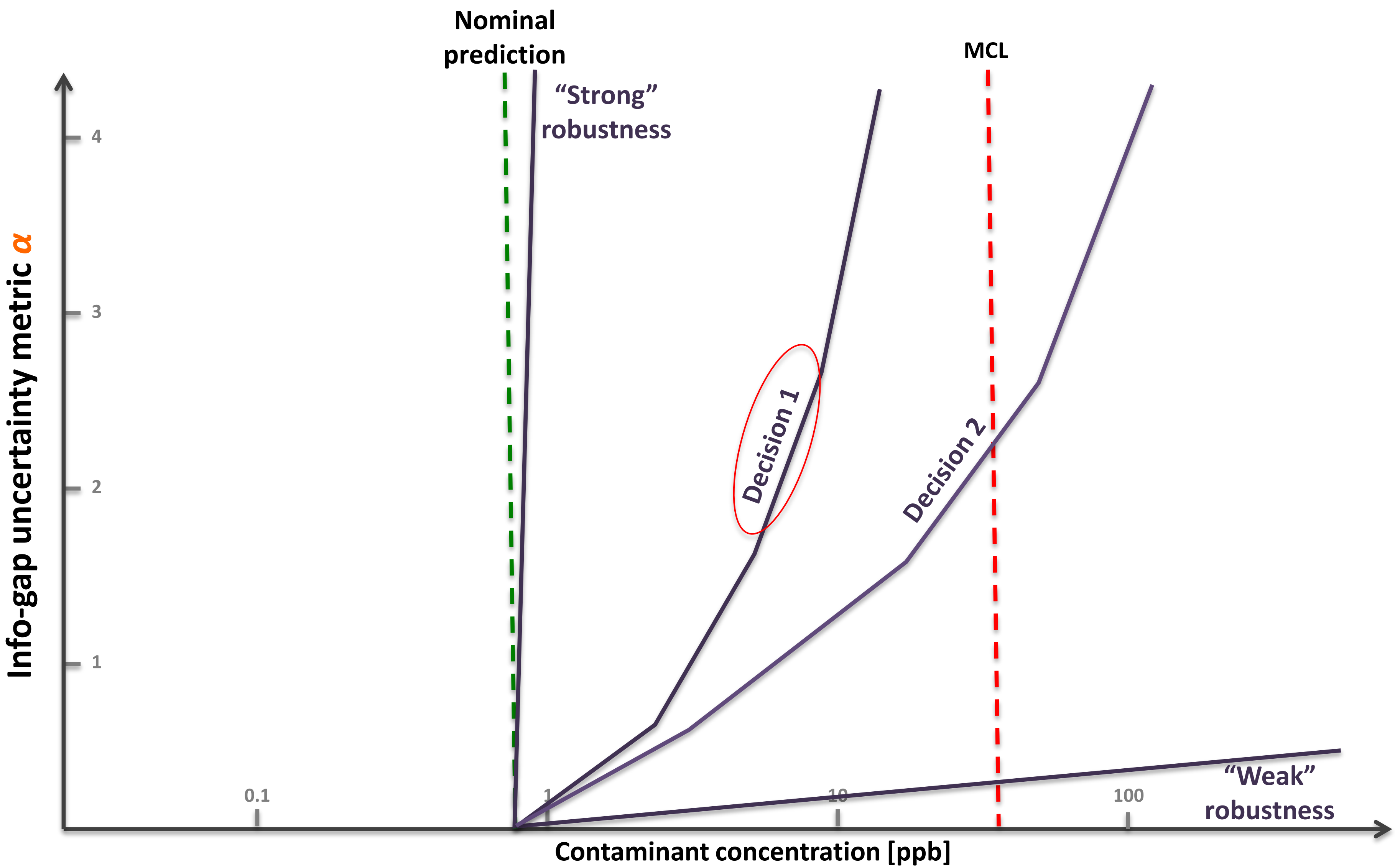
Info-Gap Analysis: Decision selection based on robustness



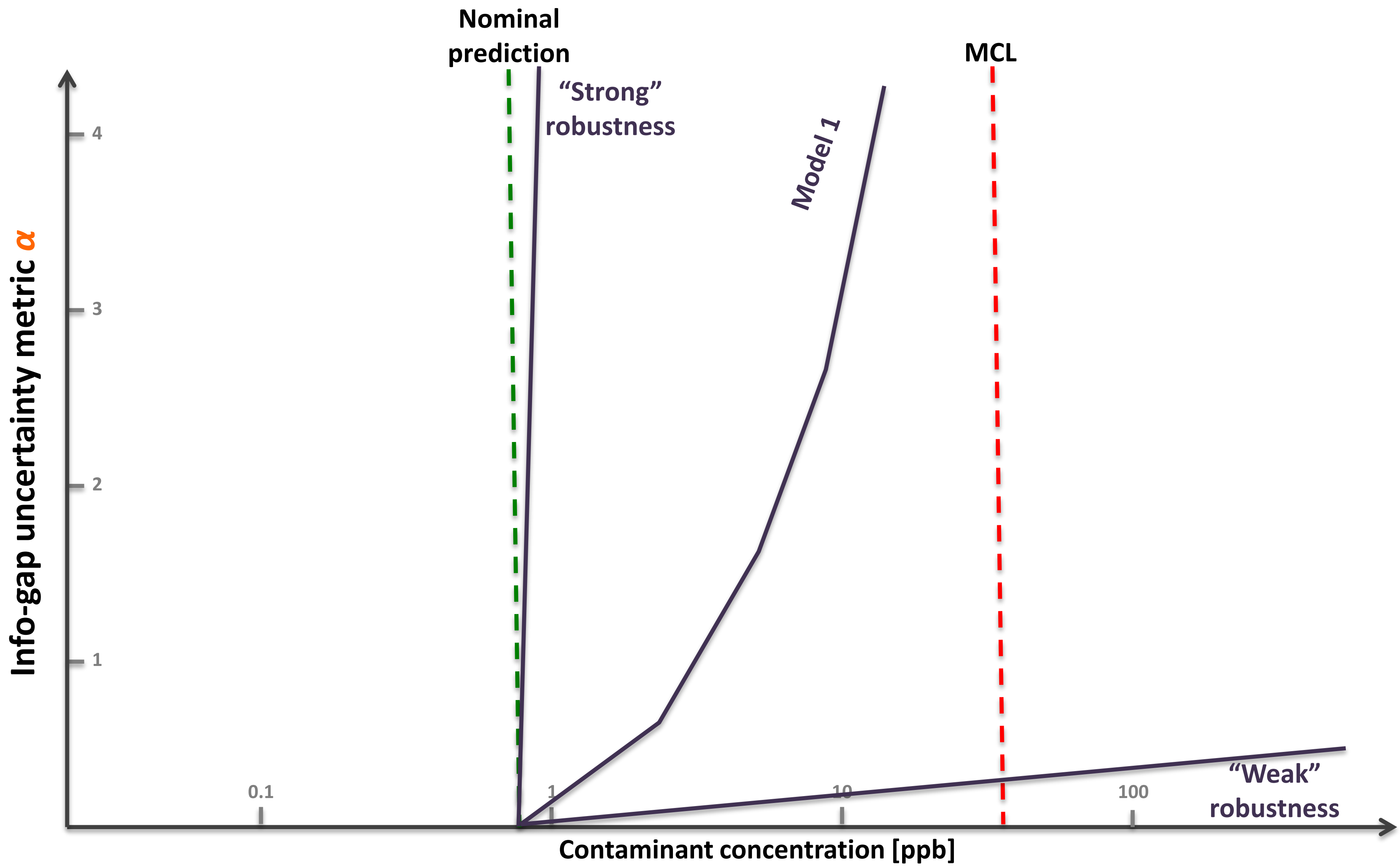
Info-Gap Analysis: Decision selection based on robustness



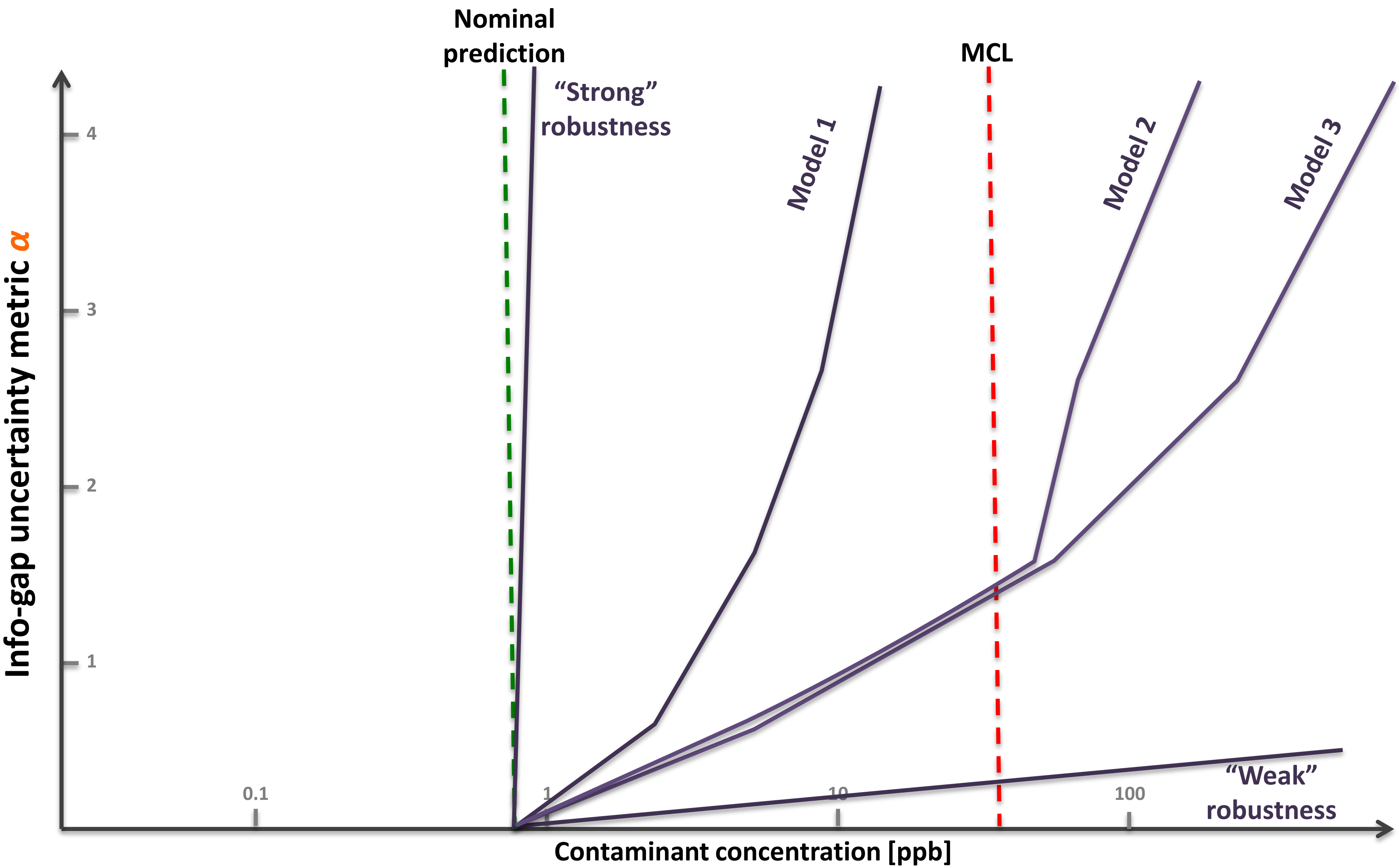
Info-Gap Analysis: Decision selection based on robustness



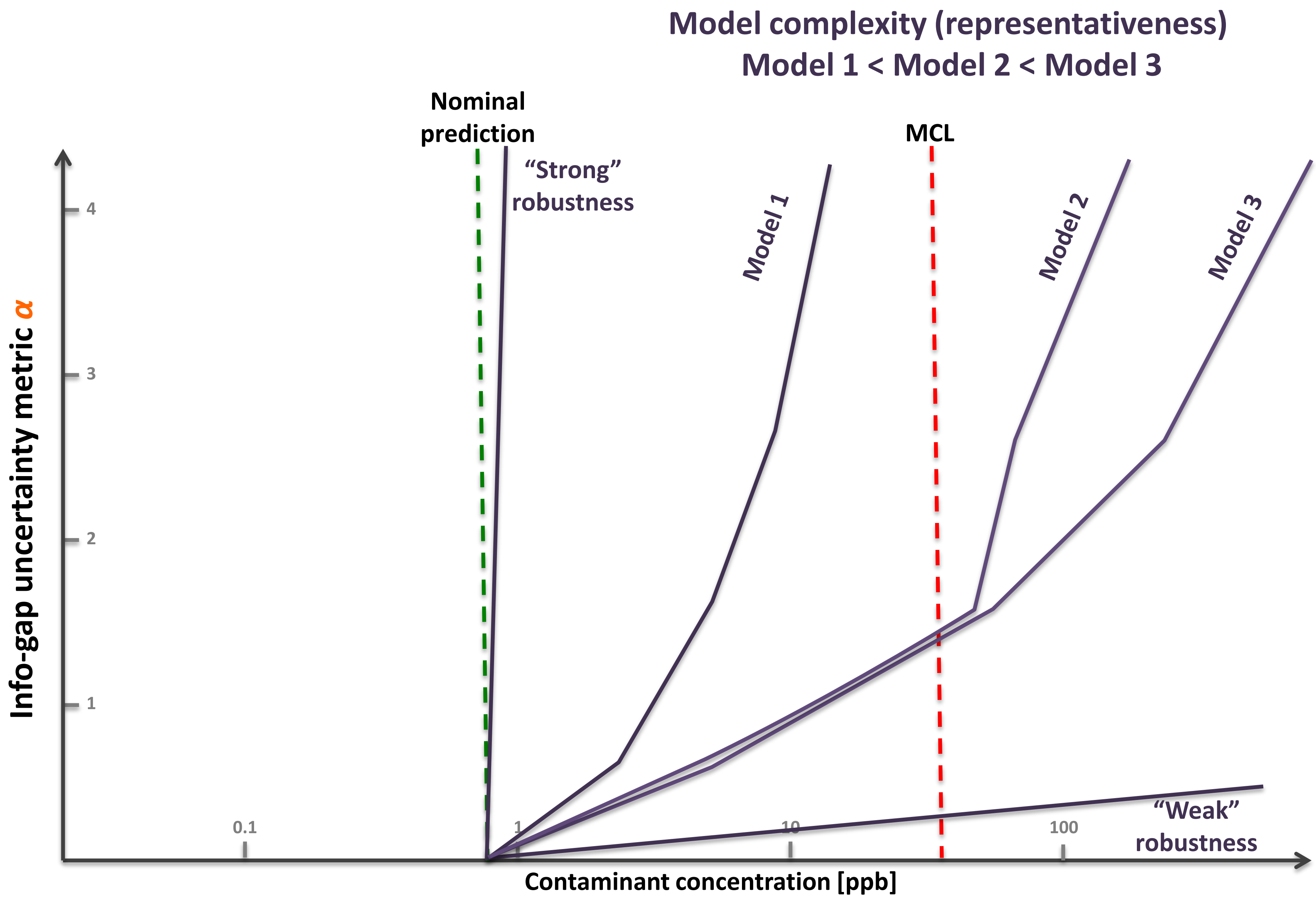
Info-Gap Analysis: Model selection based on robustness



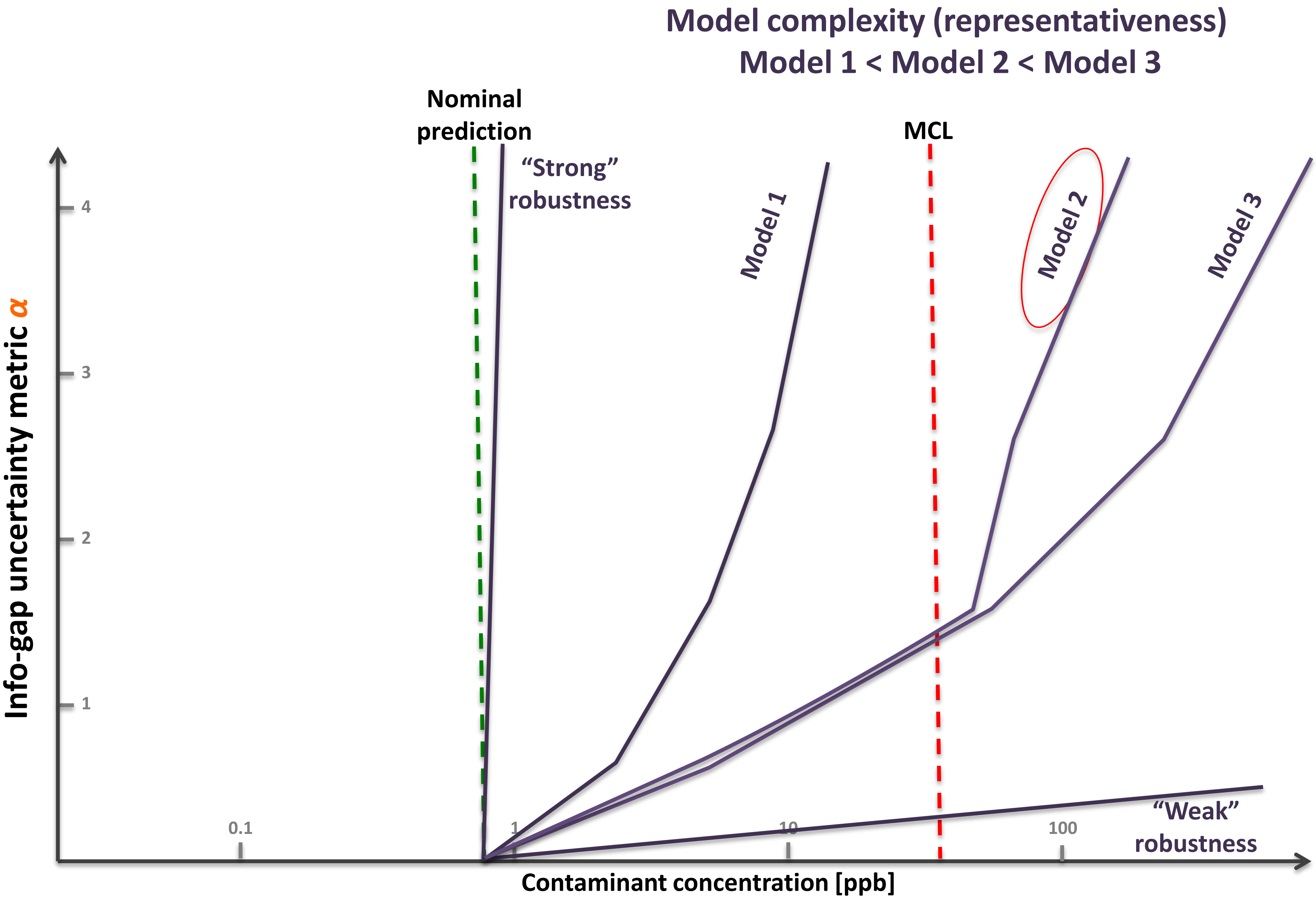
Info-Gap Analysis: Model selection based on robustness



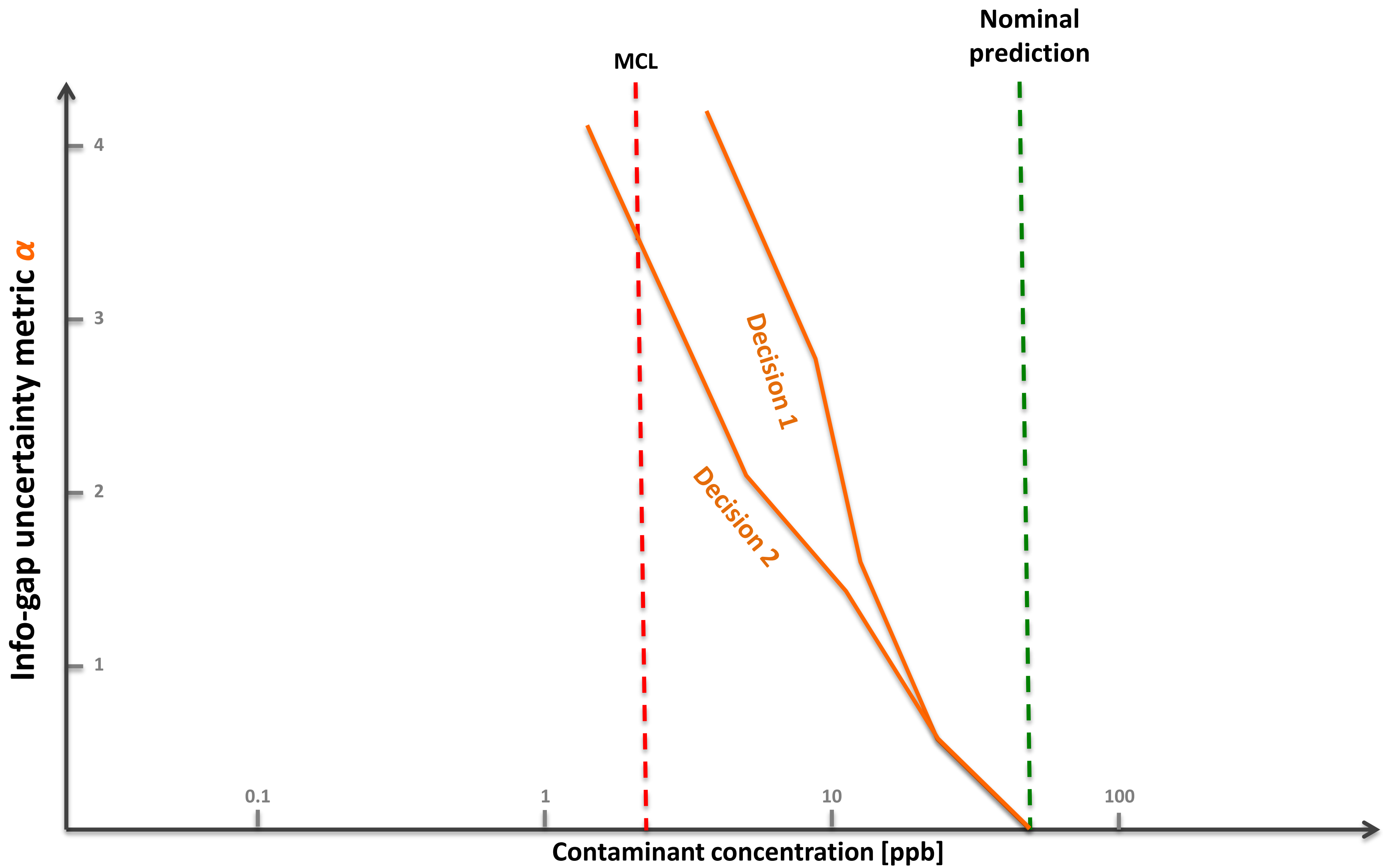
Info-Gap Analysis: Model selection based on robustness



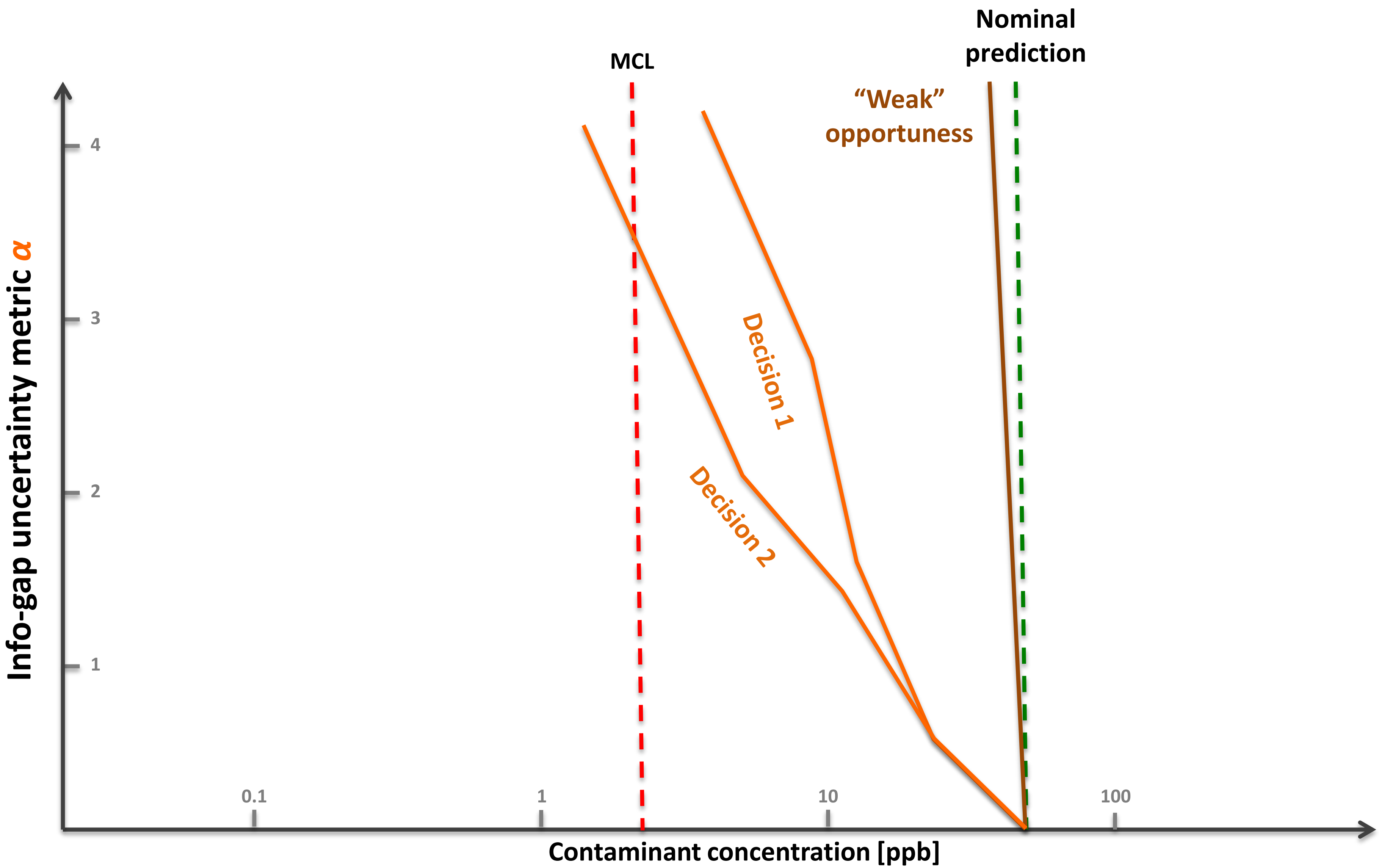
Info-Gap Analysis: Model selection based on robustness



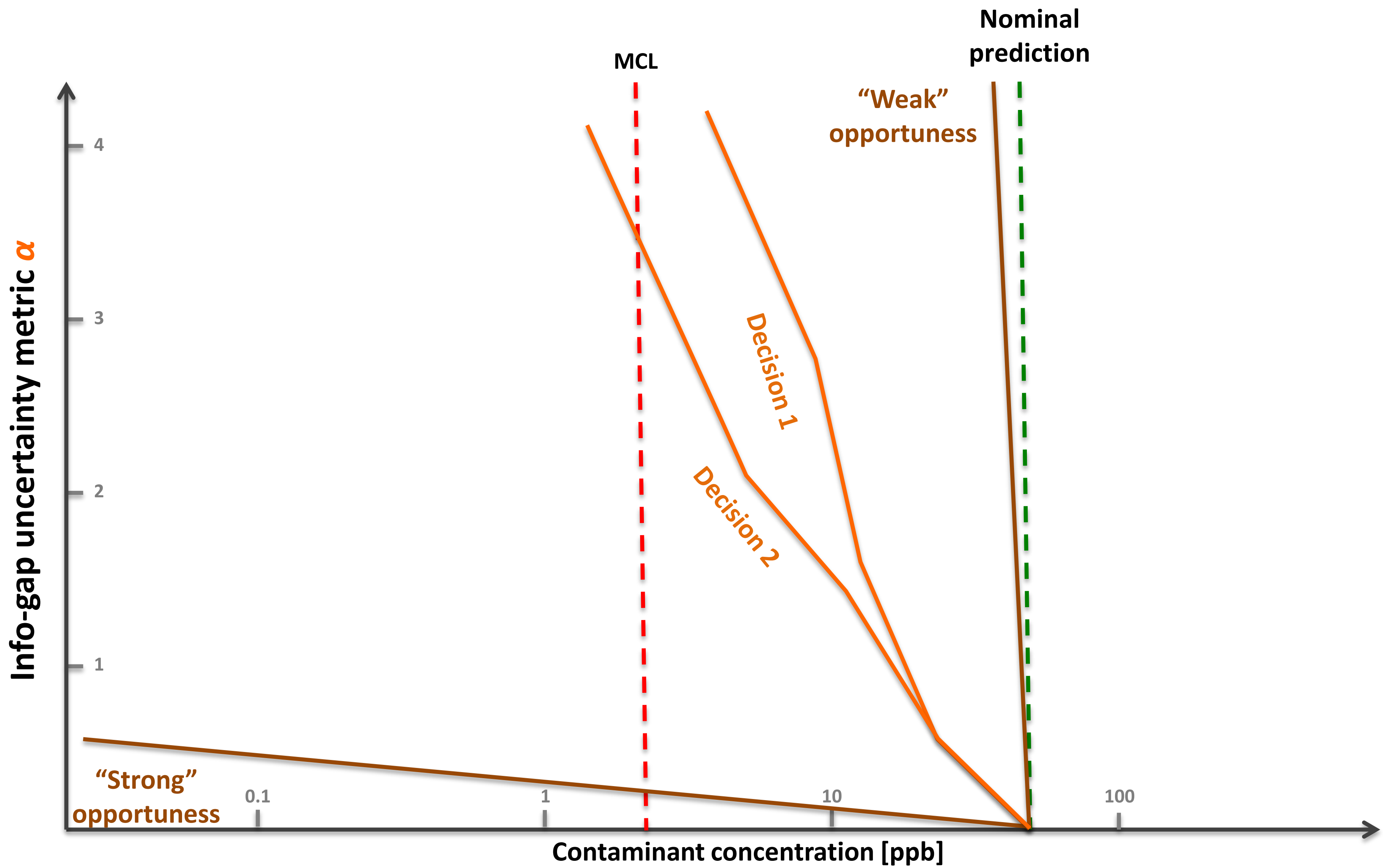
Info-Gap Analysis: Decision selection based on opportuness



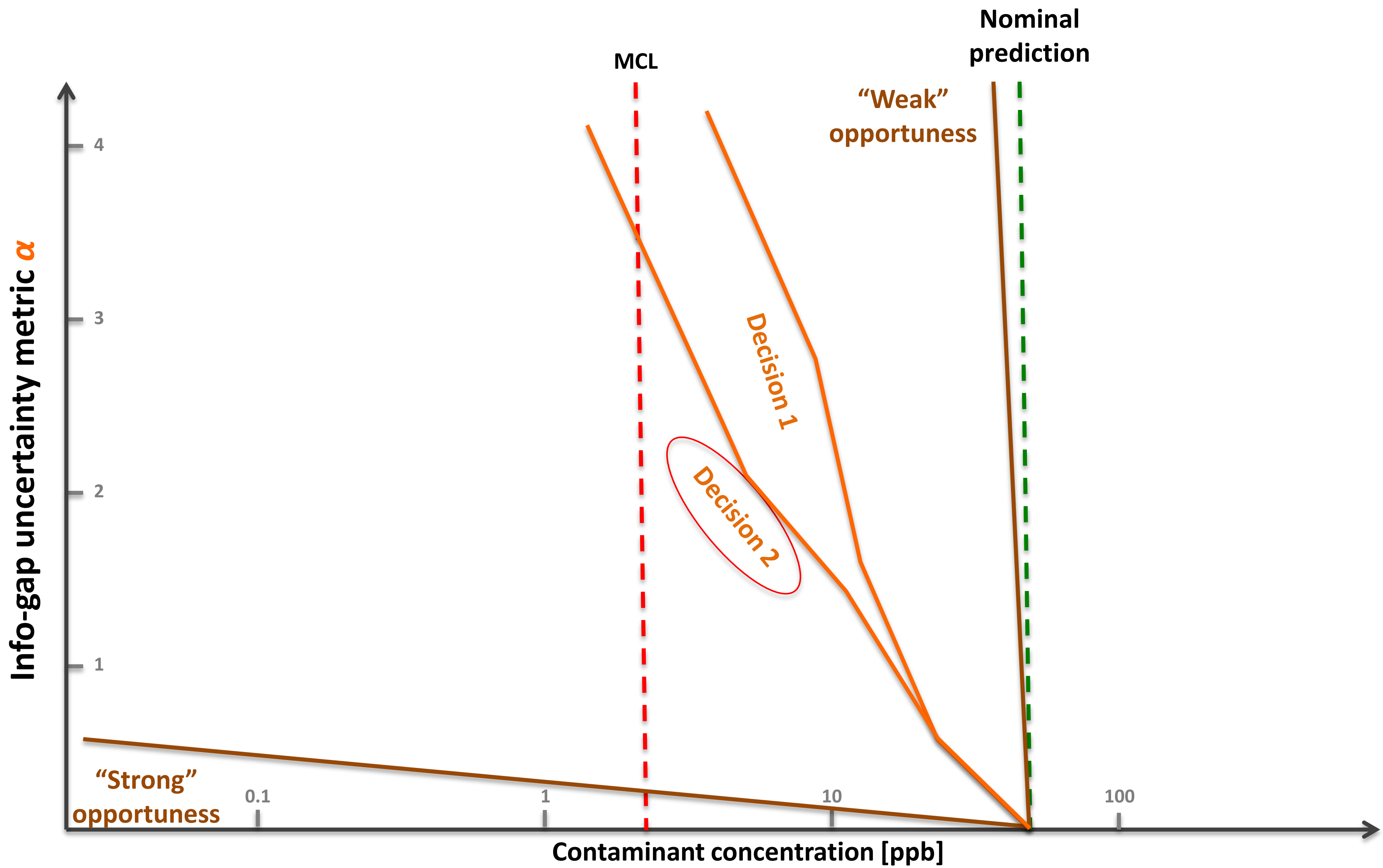
Info-Gap Analysis: Decision selection based on opportuness



Info-Gap Analysis: Decision selection based on opportuness

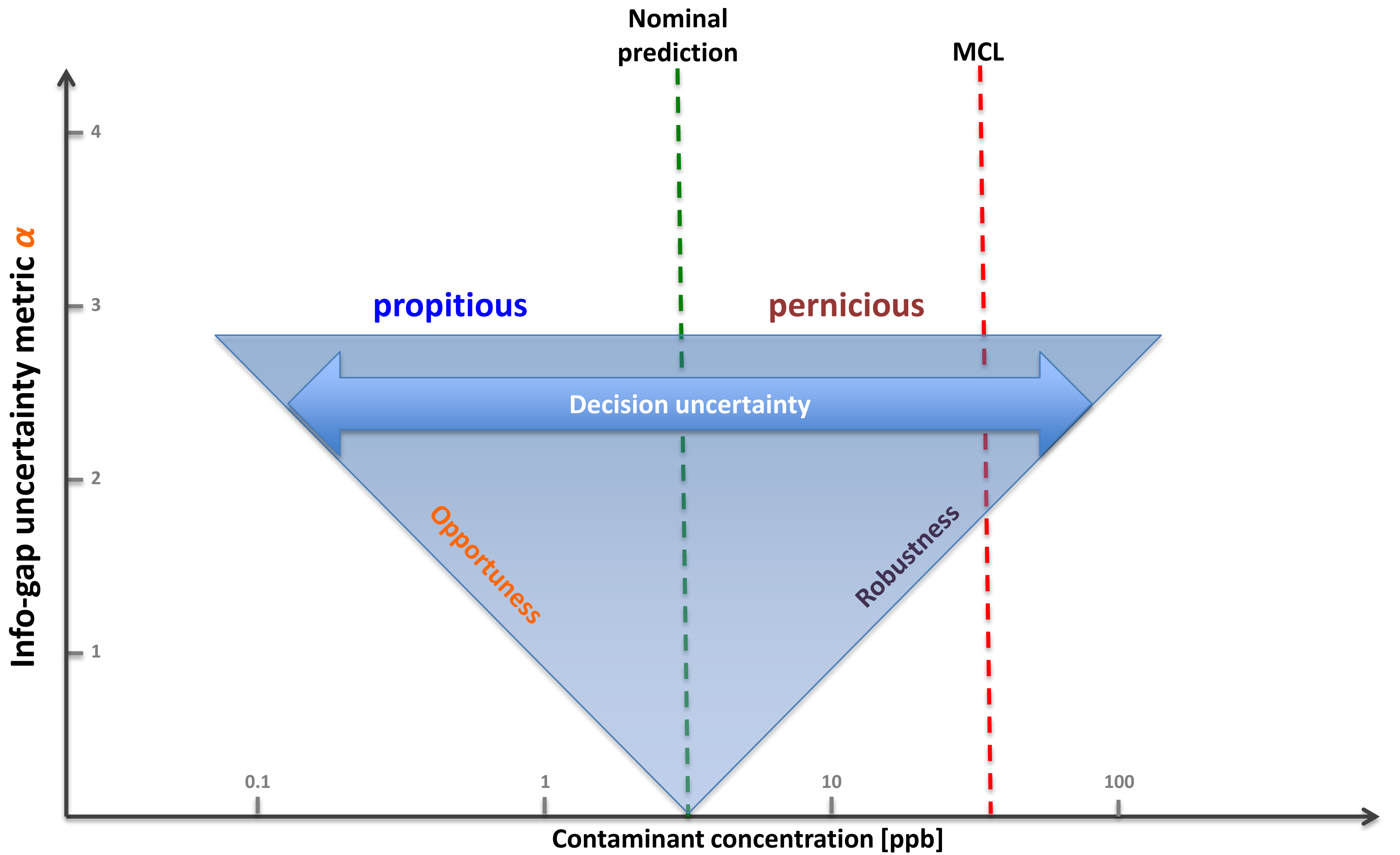


Info-Gap Analysis: Decision selection based on opportuness



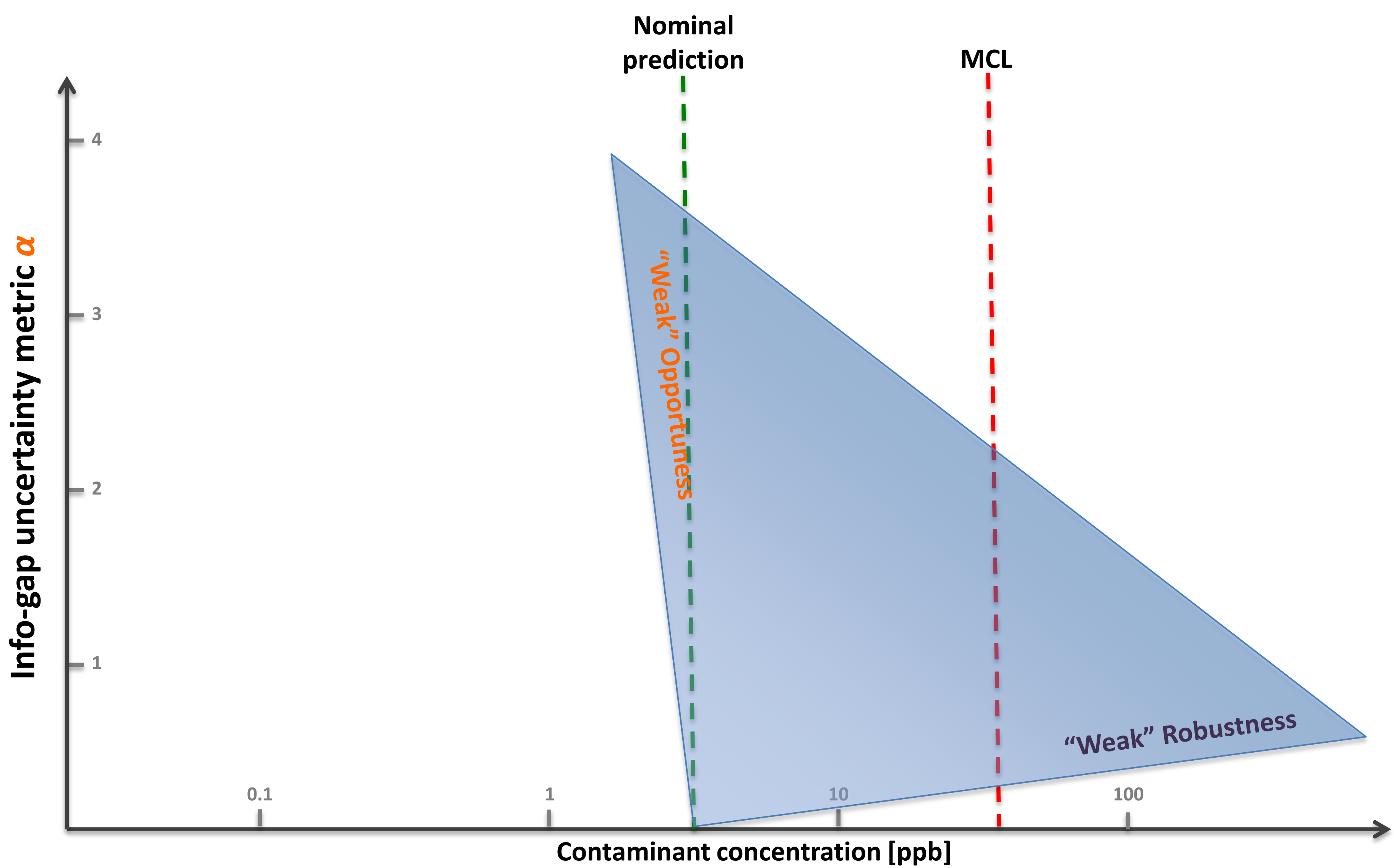
Info-Gap Analysis: Decision uncertainty

… duality of decision uncertainty



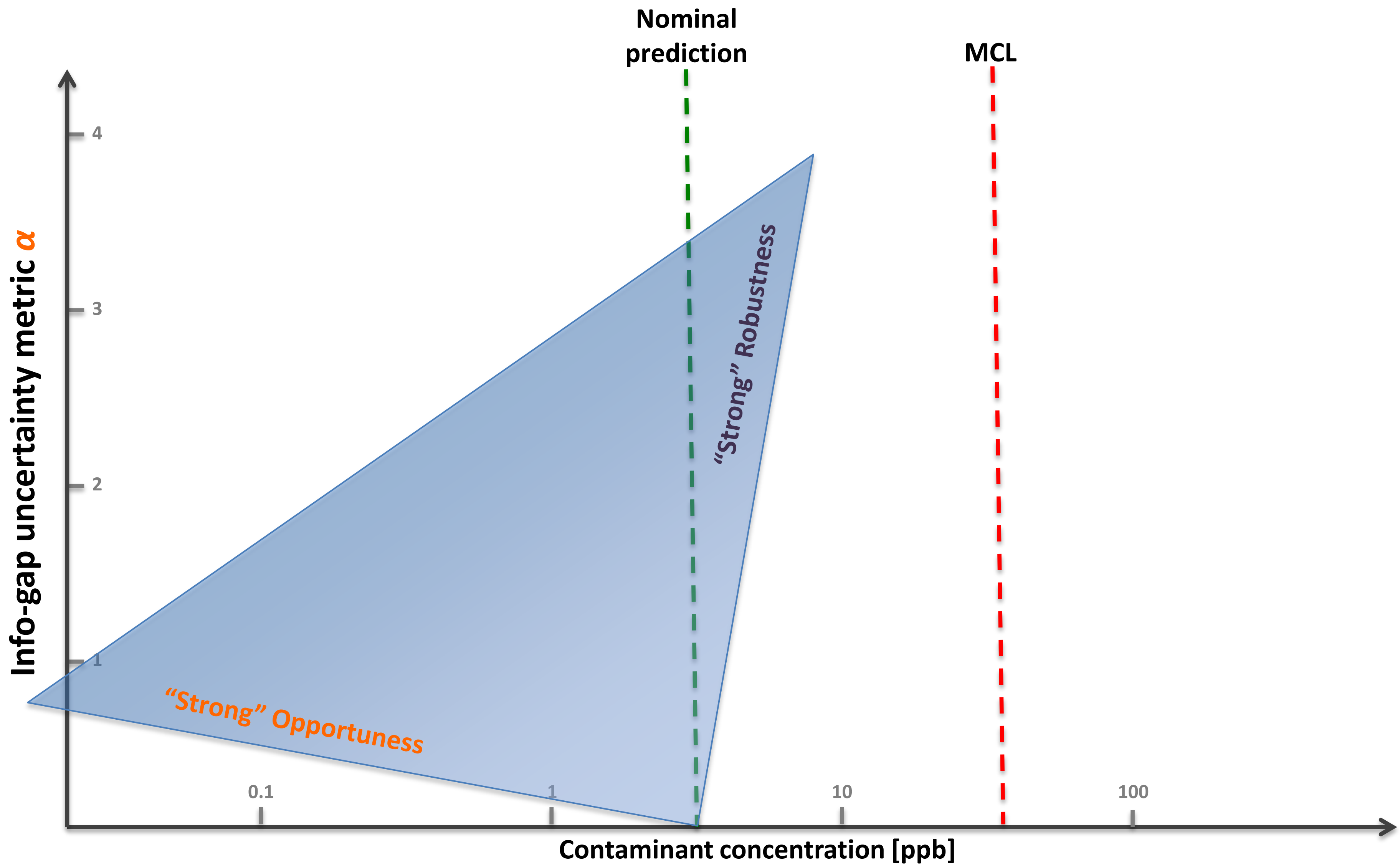
Info-Gap Analysis: Decision uncertainty

… not preferred decision bounds



Info-Gap Analysis: Decision uncertainty

… preferred decision bounds

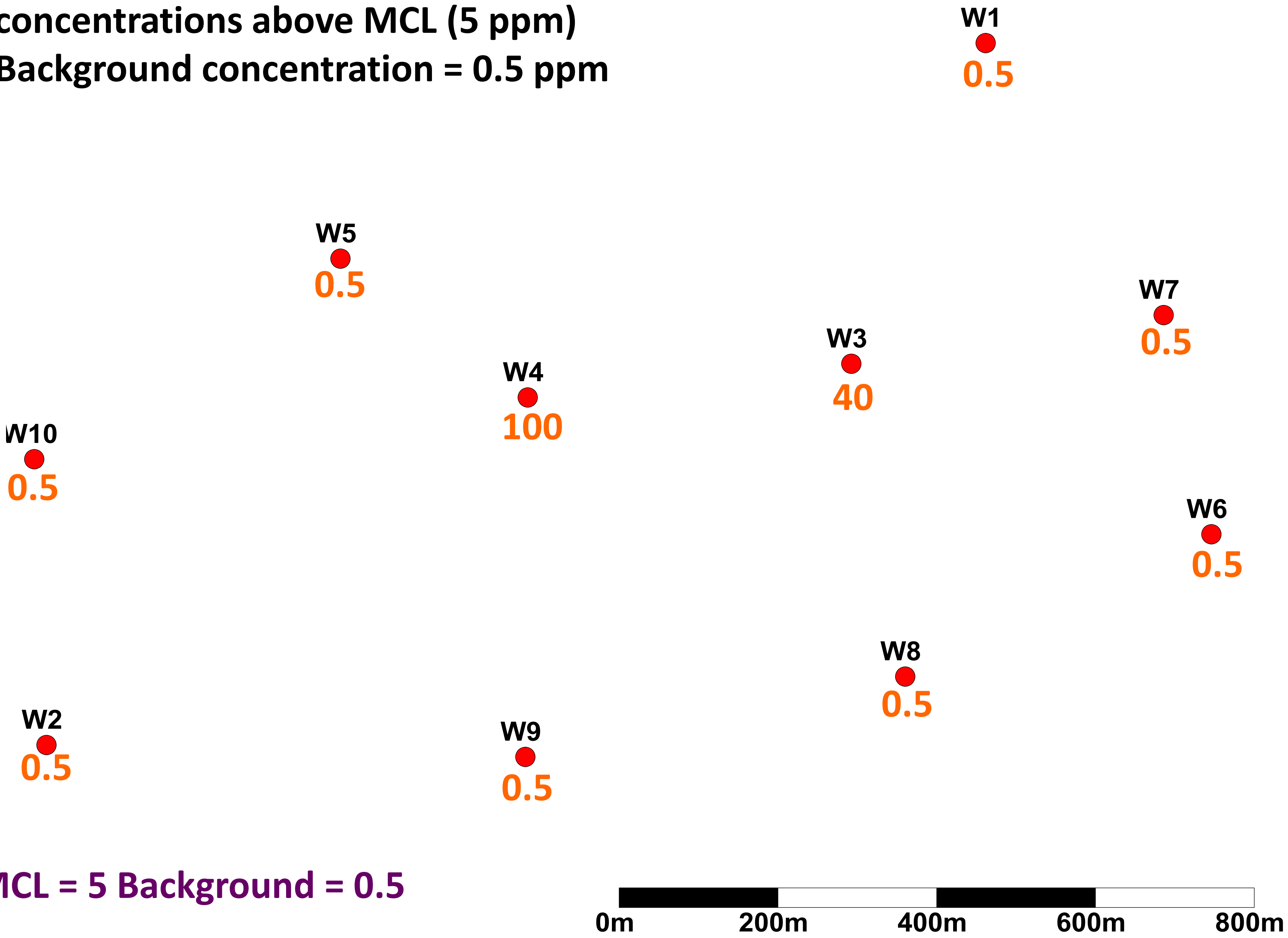


Info-Gap Application: Case 1

Optimization of monitoring network

Info-Gap Analysis: Network Design

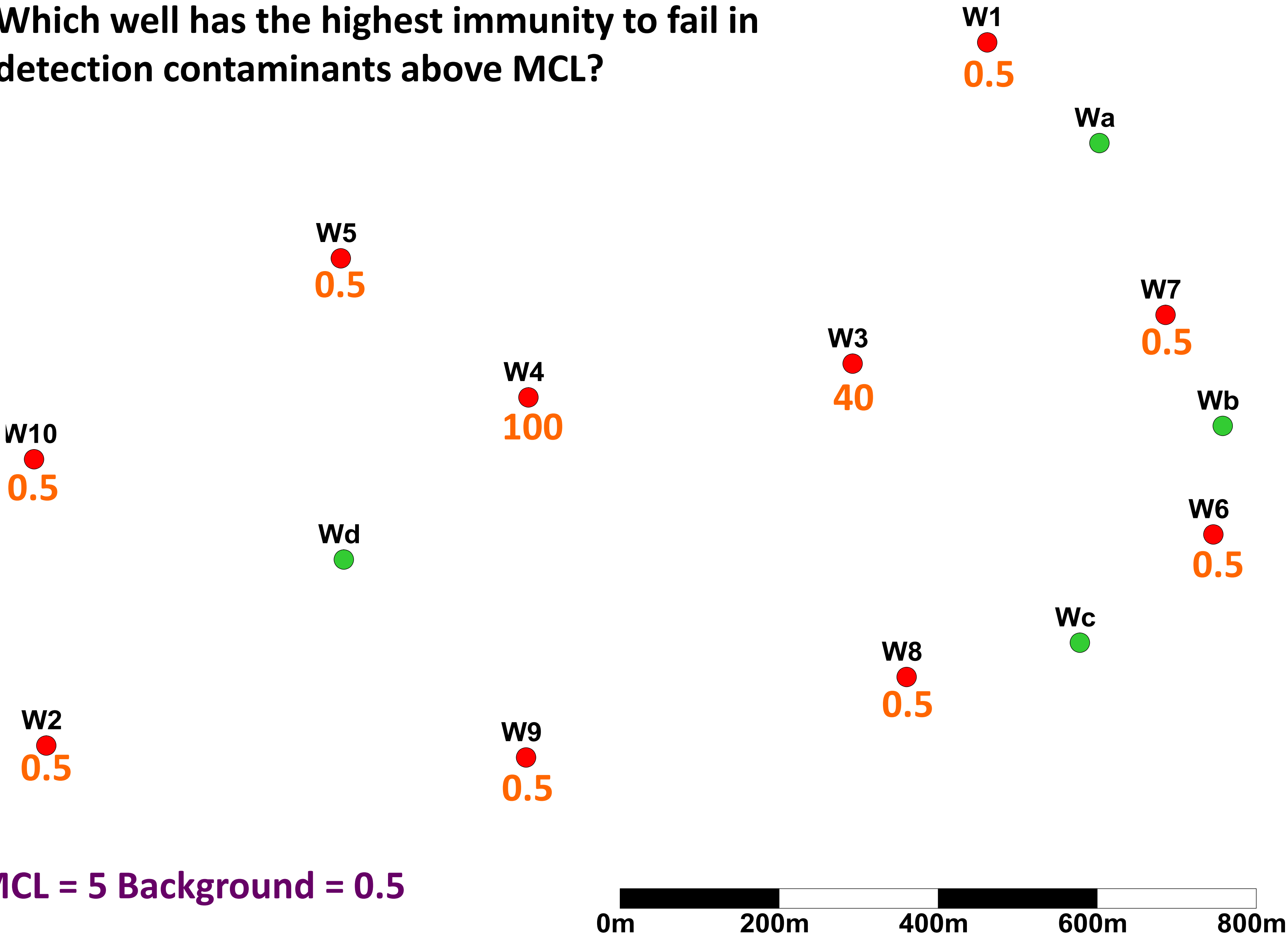
- Two monitoring wells in an aquifer with contaminant concentrations above MCL (5 ppm)
- Background concentration = 0.5 ppm



MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design

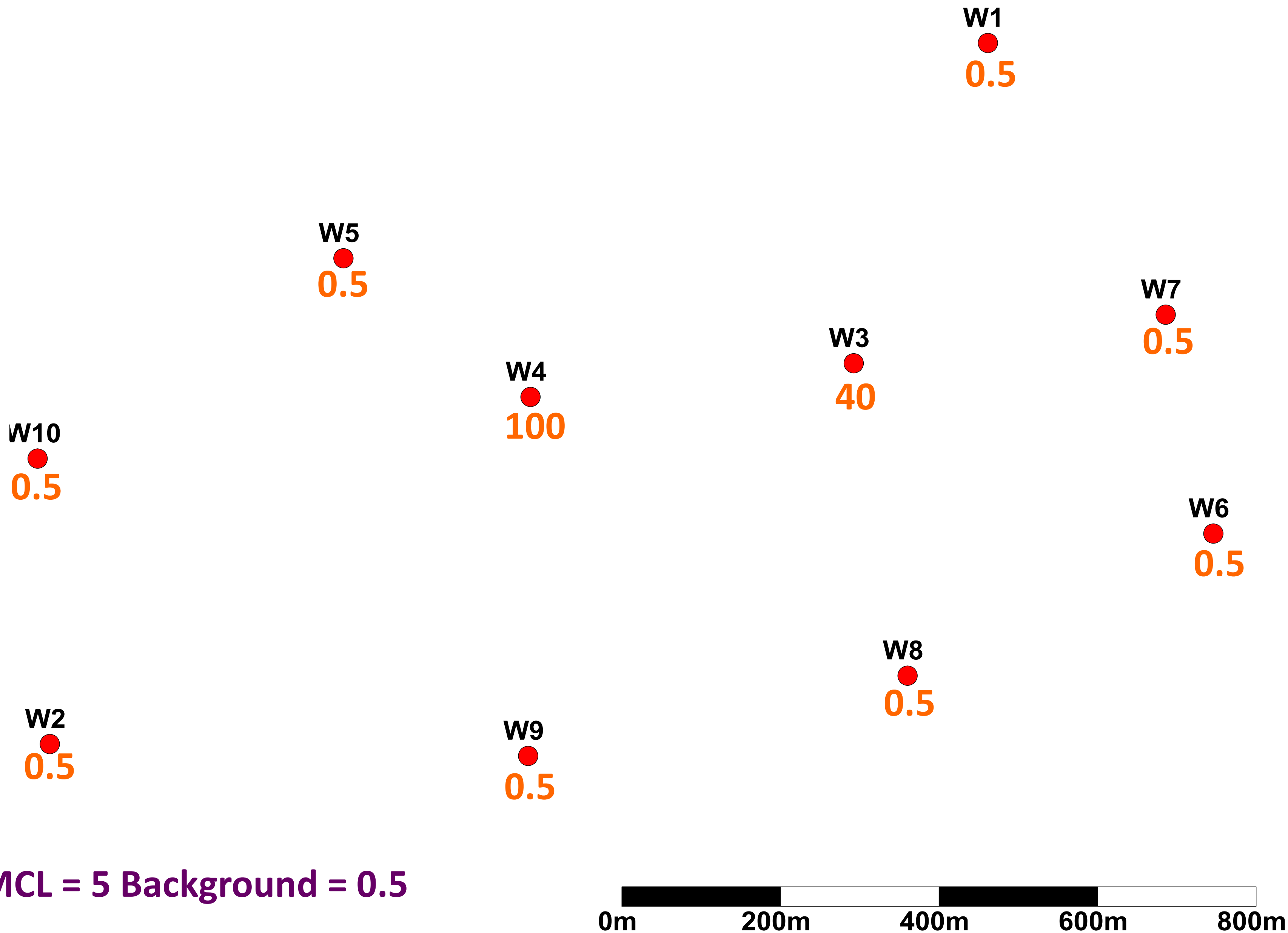
- 4 new proposed monitoring well locations
- Which well has the highest immunity to fail in detection contaminants above MCL?



MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design

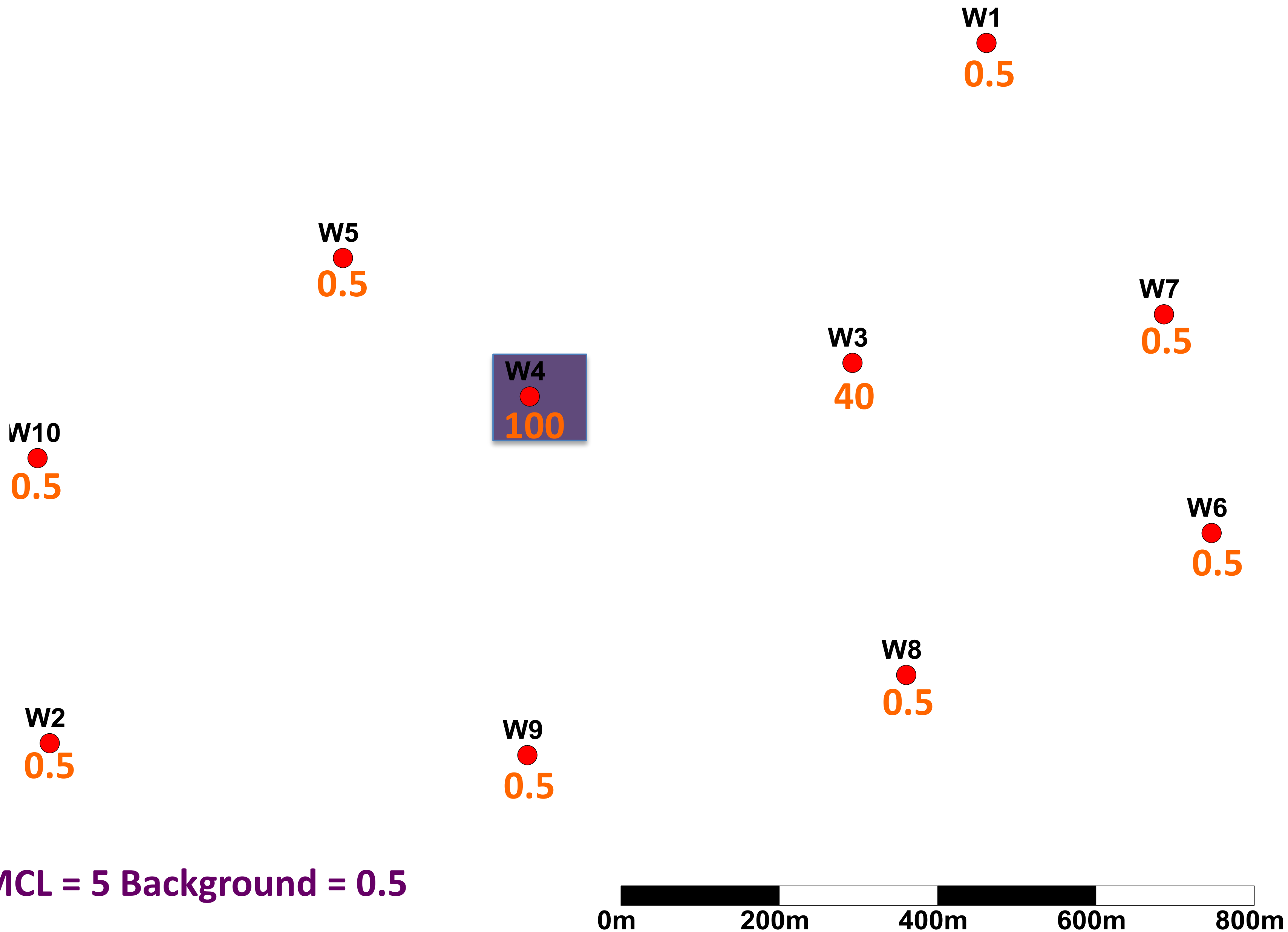
✧ Where is the contaminant source?



MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design

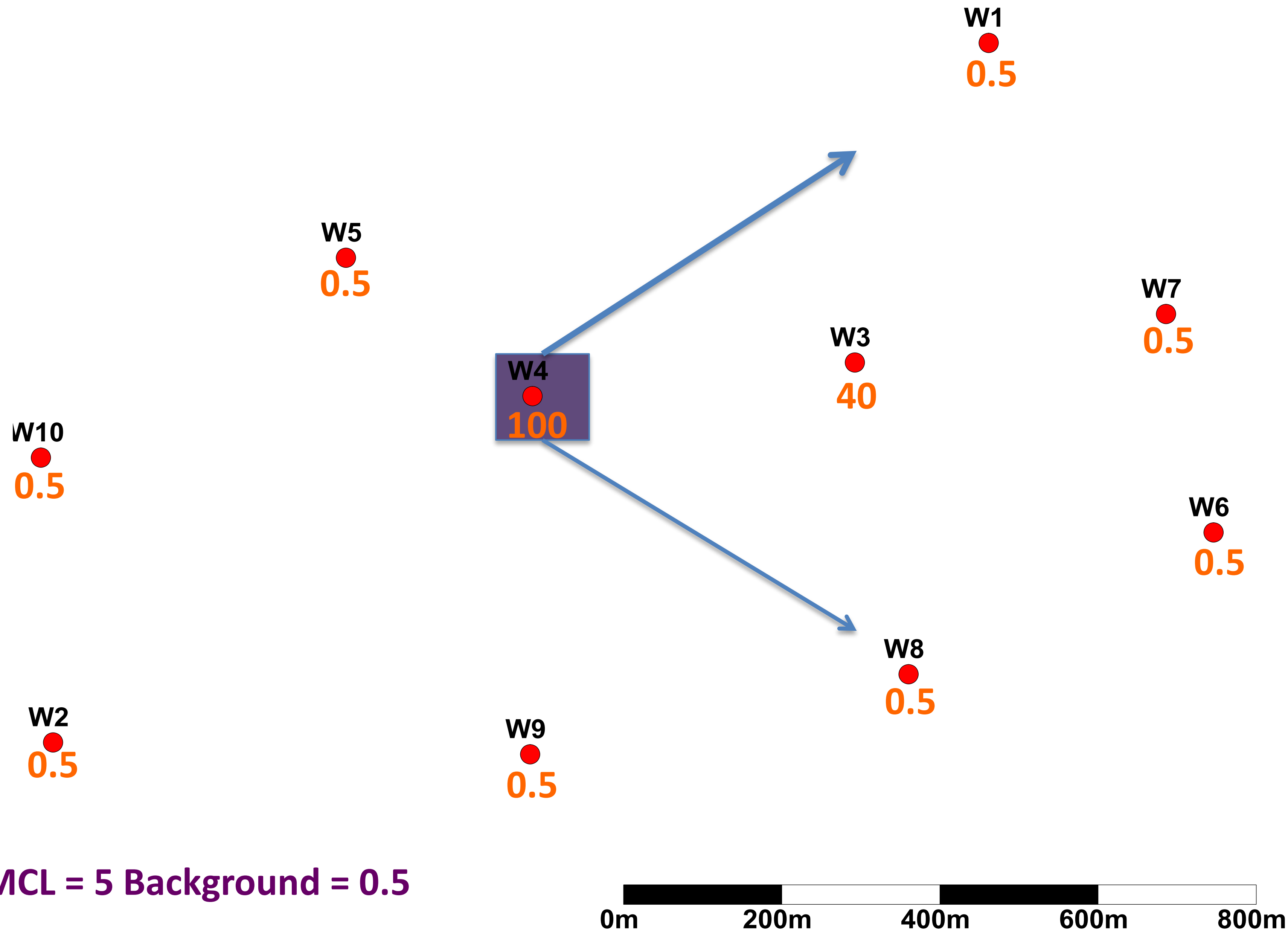
✧ Where is the contaminant source?



MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design

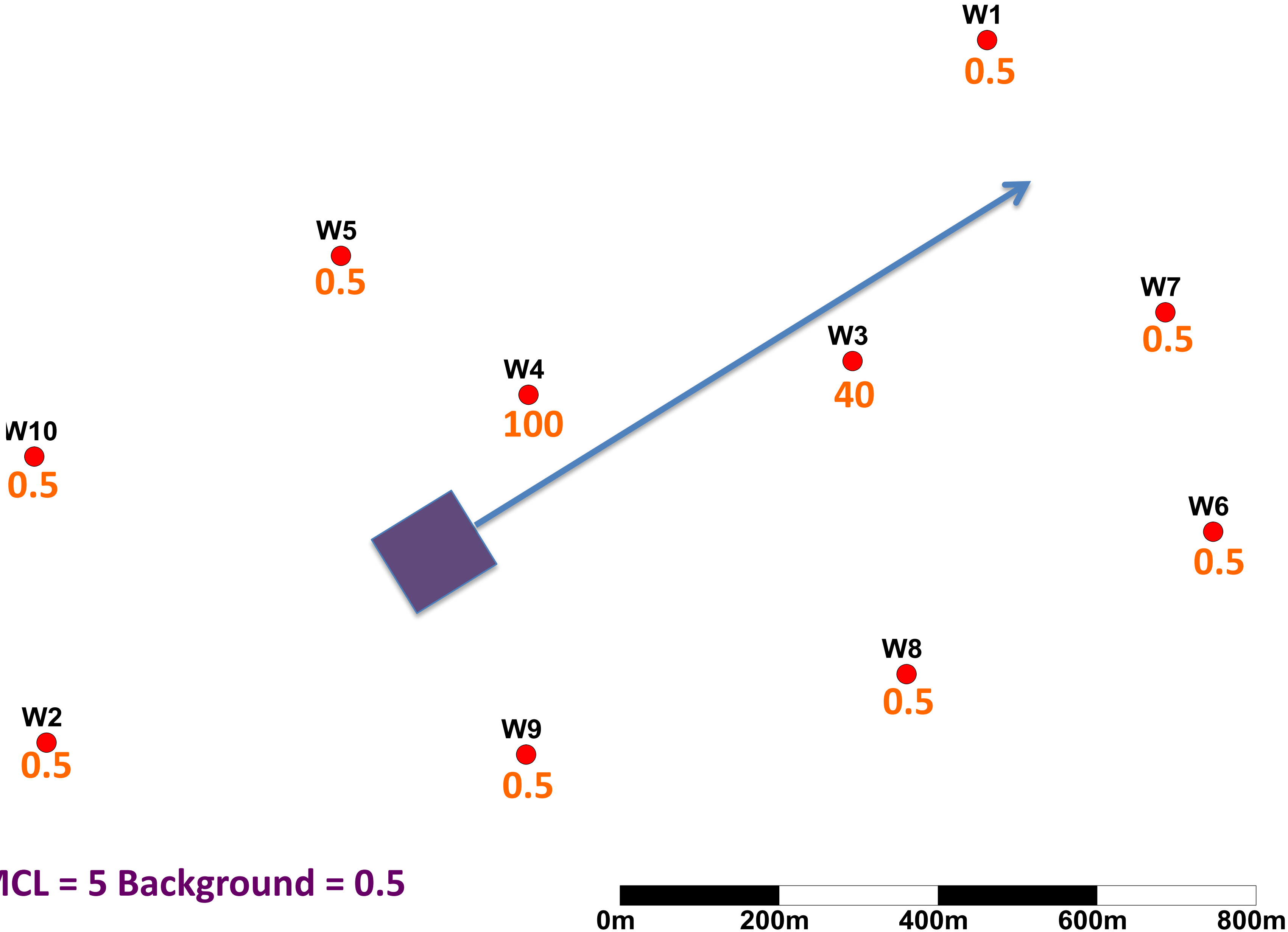
- Where is the contaminant source?



MCL = 5 Background = 0.5

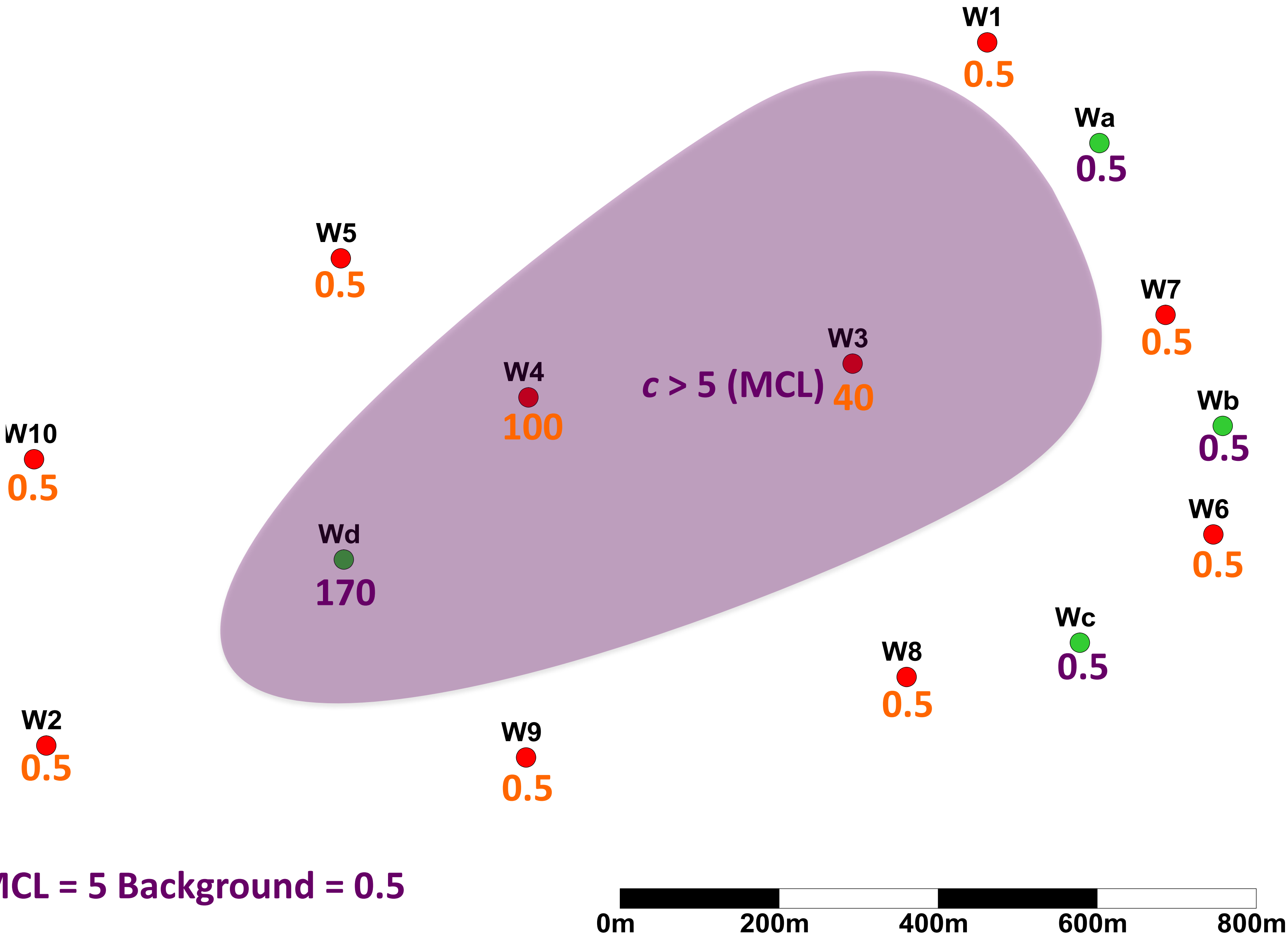
Info-Gap Analysis: Network Design

- Where is the contaminant source?



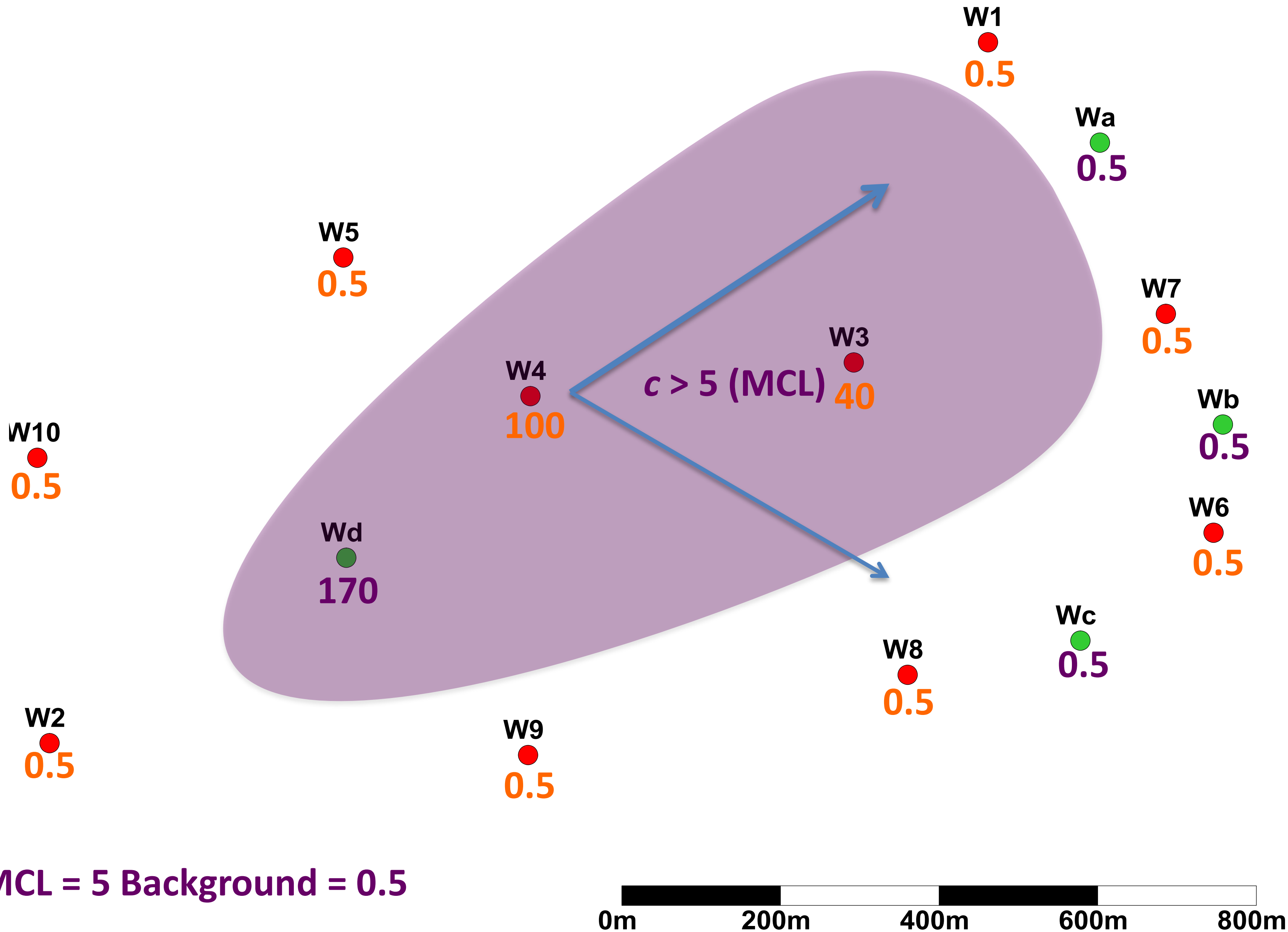
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



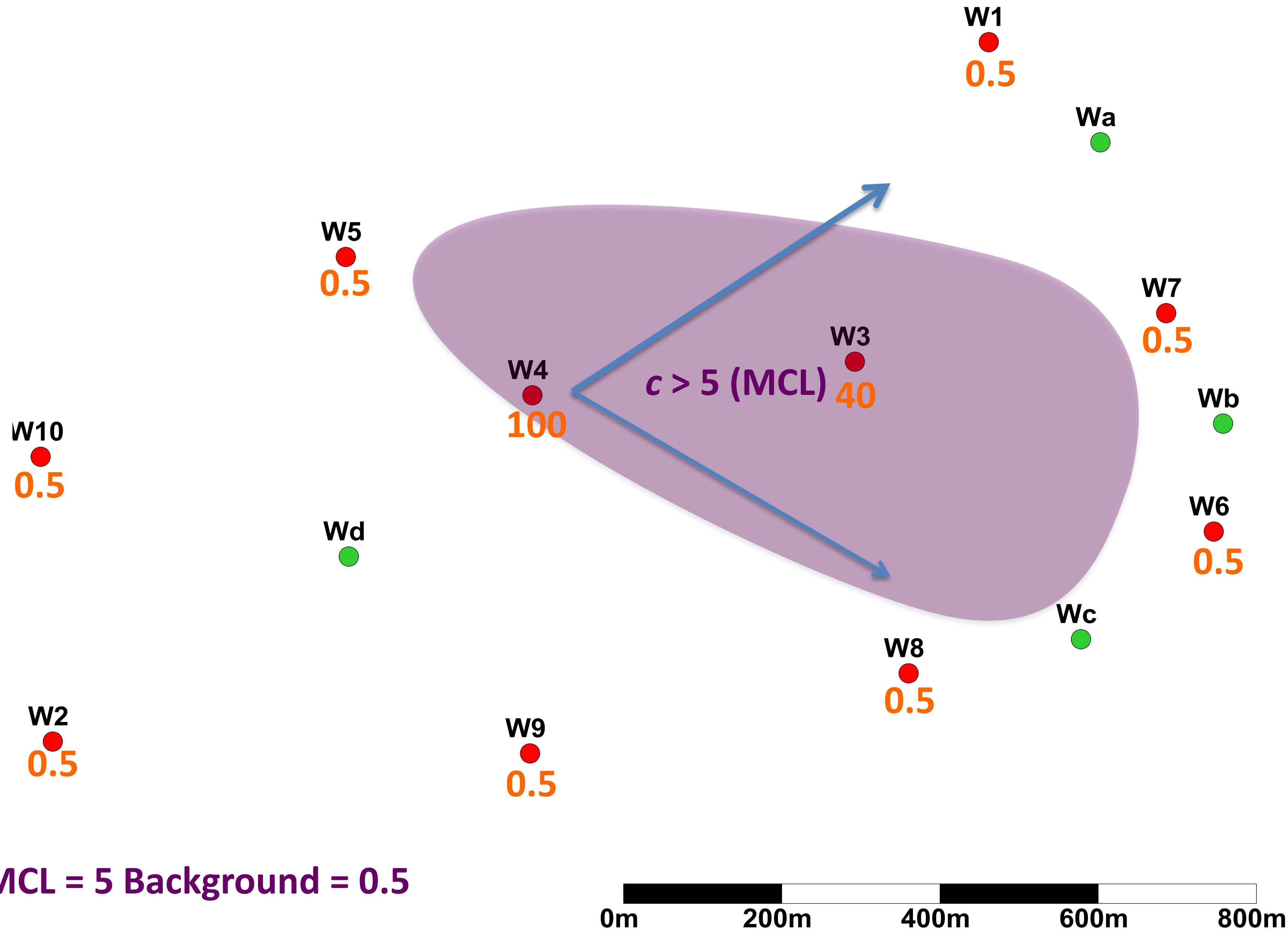
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



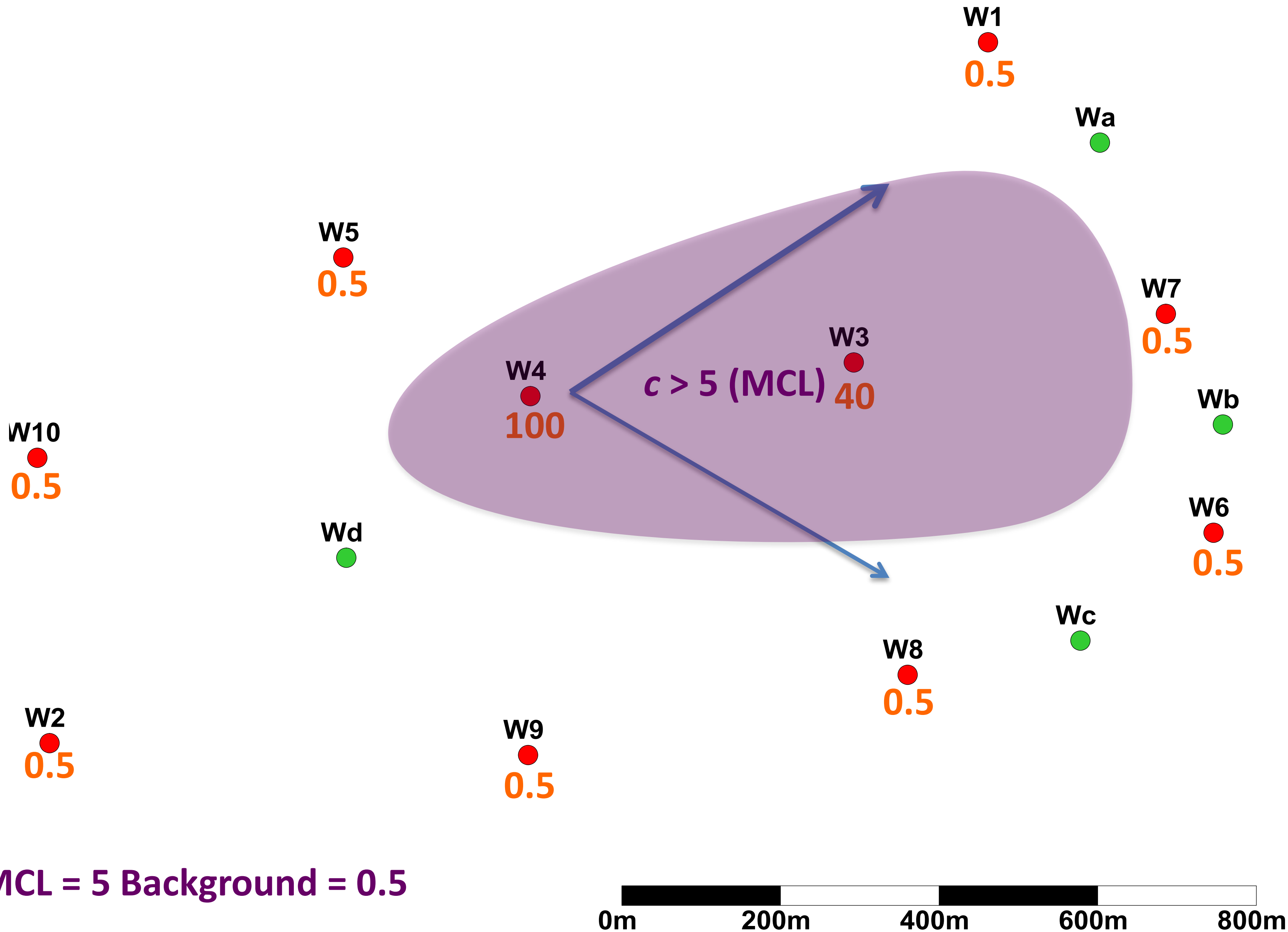
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



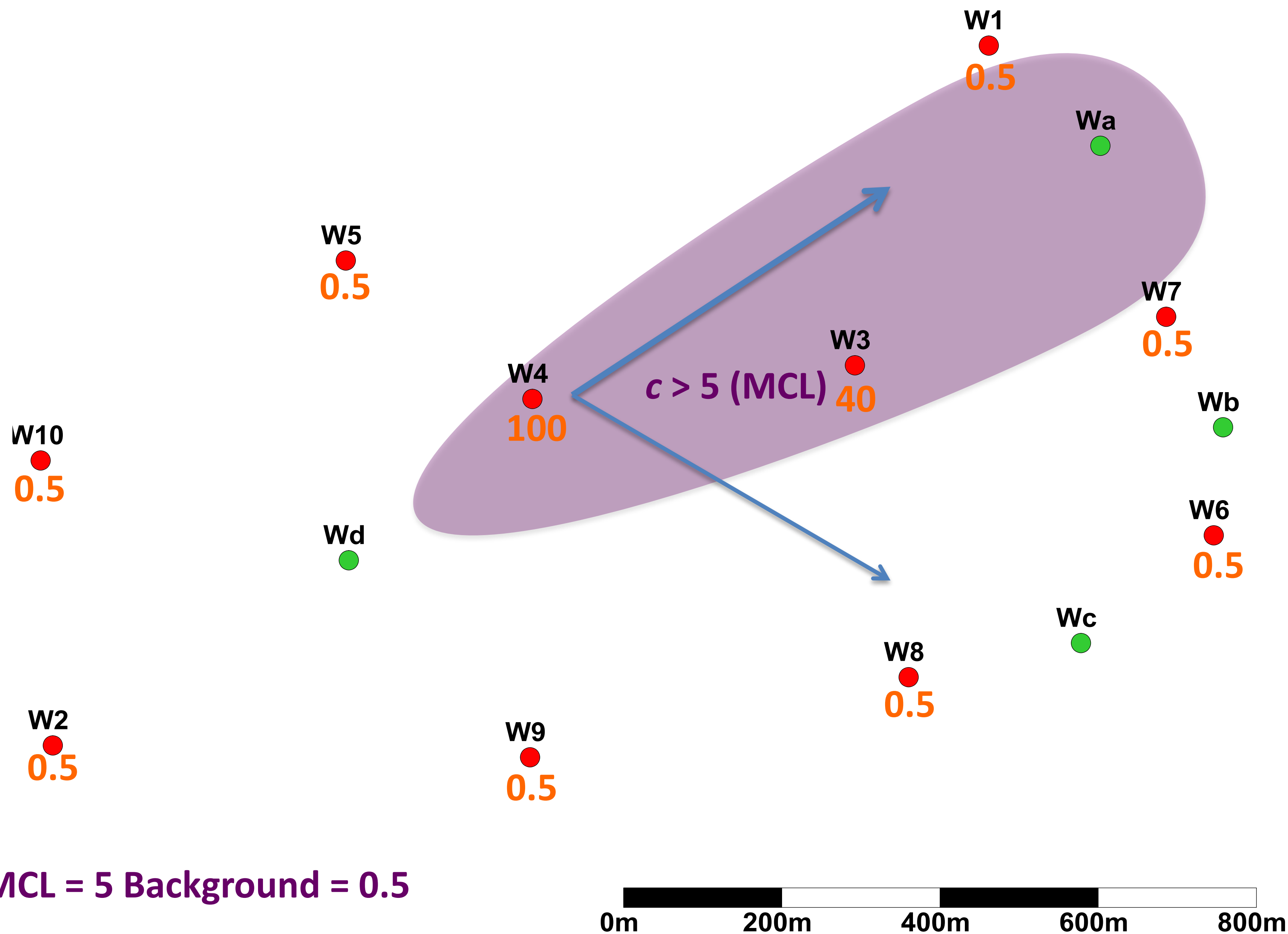
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



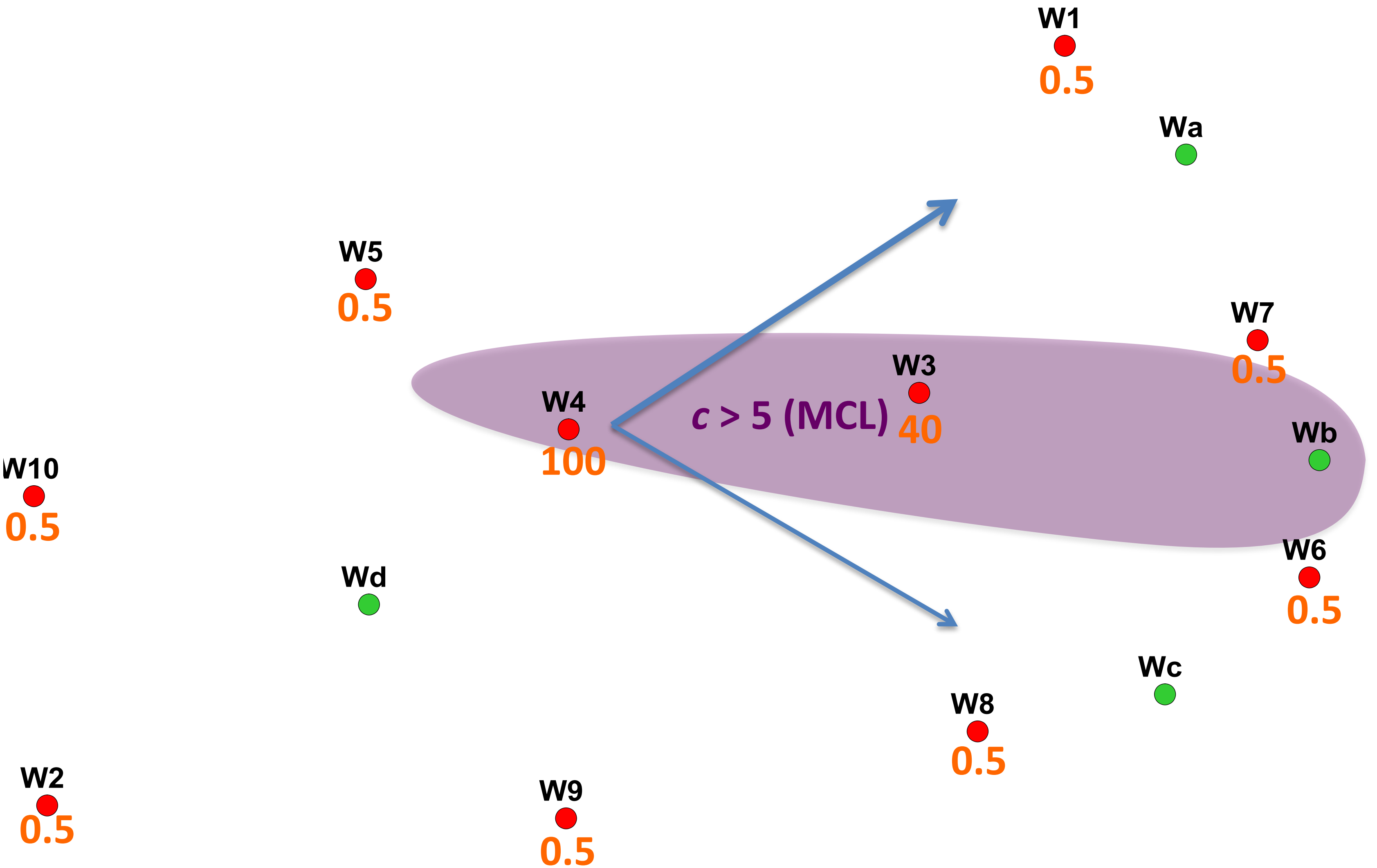
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



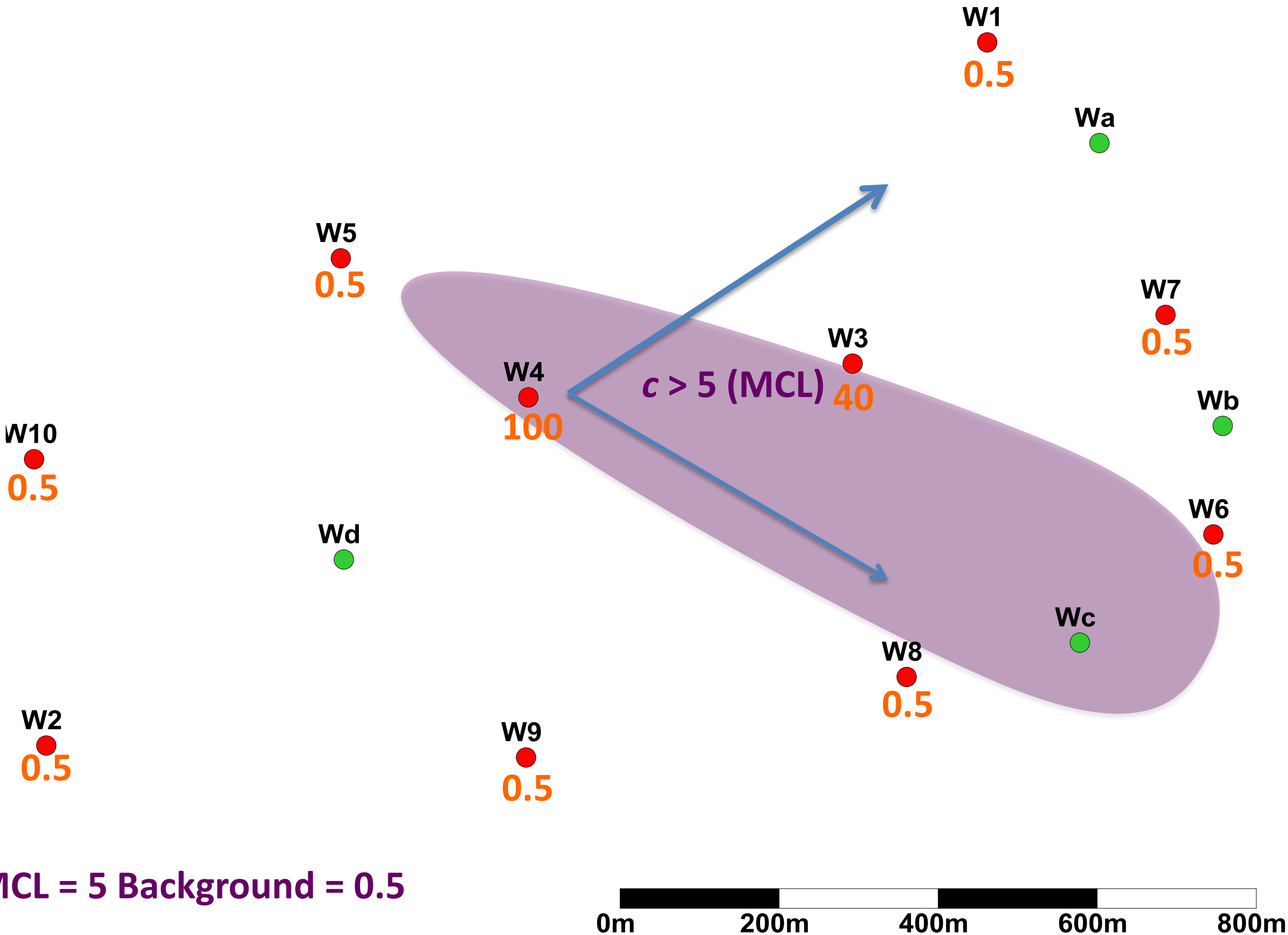
MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design



MCL = 5 Background = 0.5

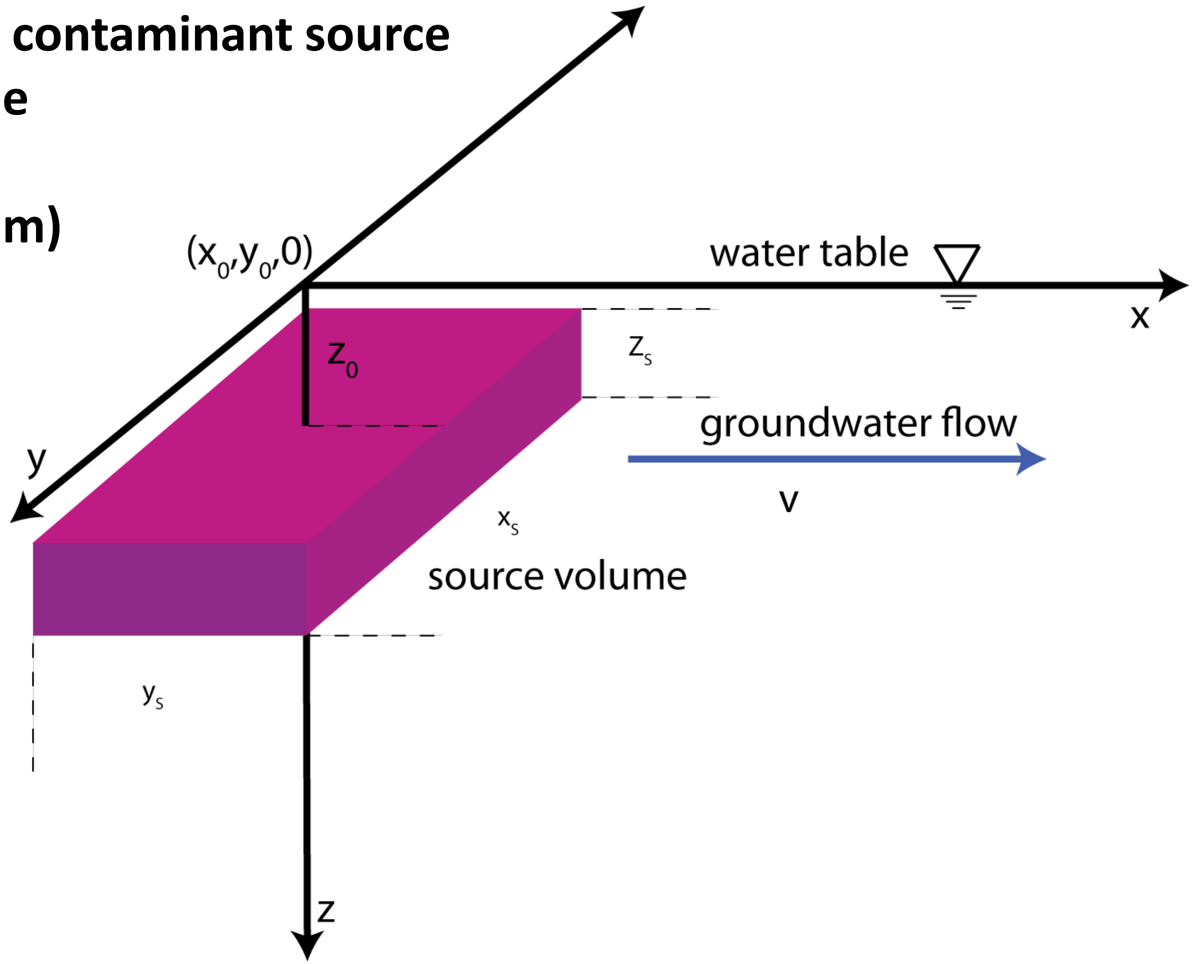
Info-Gap Analysis: Network Design



MCL = 5 Background = 0.5

Info-Gap Analysis: Network Design

- Analytical contaminant flow model:
  - 3D steady-state uniform groundwater flow in unbounded aquifer
  - 3D contaminant source at the top of the aquifer
  - 3D contaminant migration (advection, dispersion)
- Deterministic model parameters
  - contaminant flux at the contaminant source
  - contaminant arrival time
  - groundwater velocity
  - source thickness (z_S = 1 m)



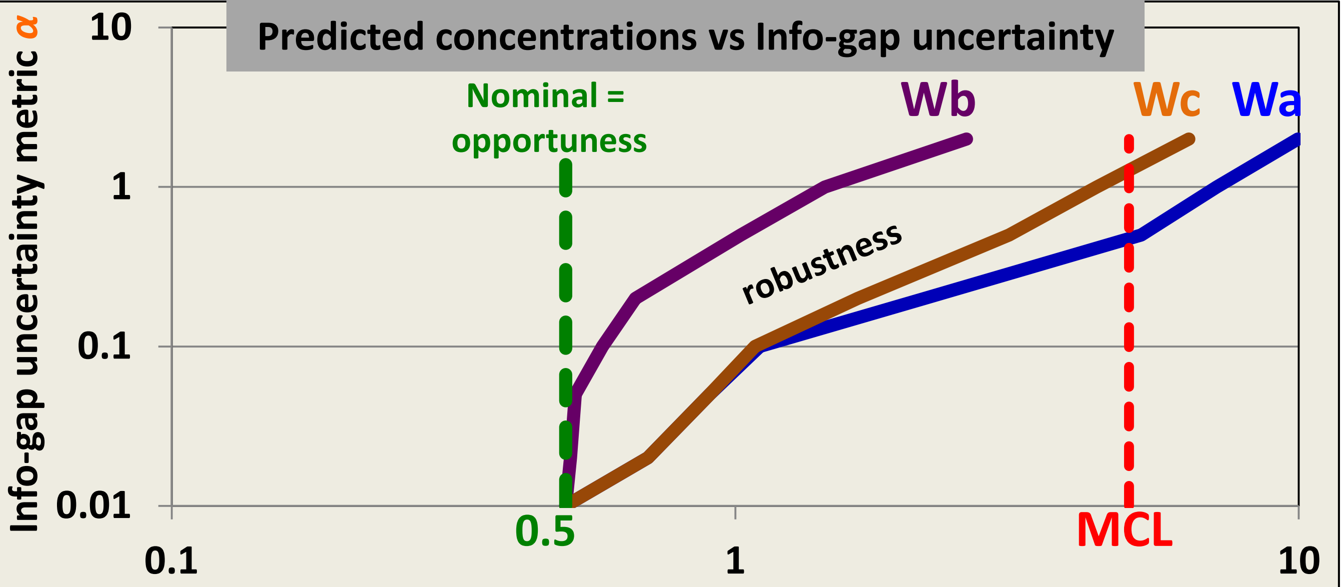
Info-Gap Analysis: Network Design

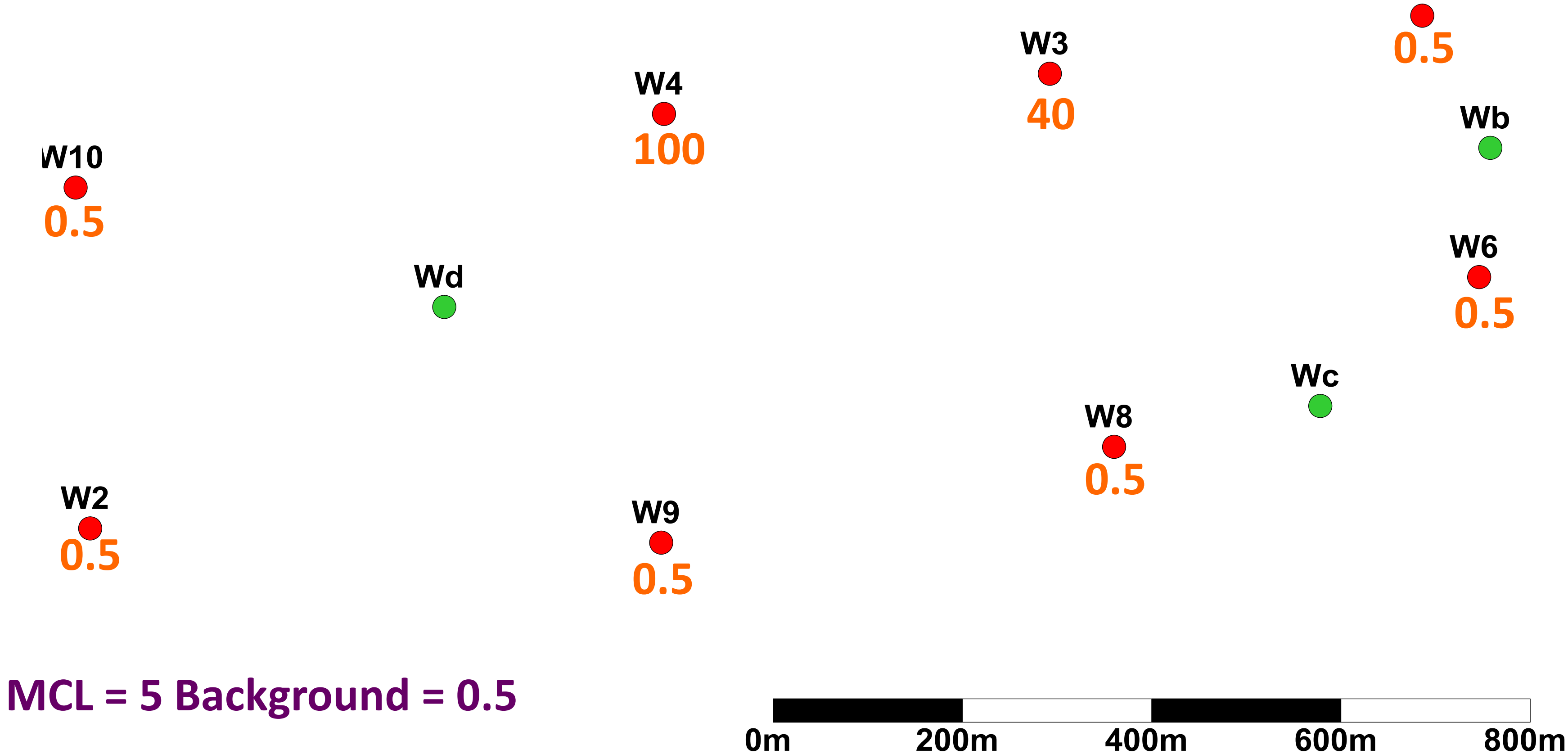
- Unknown model parameters (8)
  - source coordinates (x, y)
  - source size (x_S, y_S)
  - flow direction
  - aquifer dispersivities (longitudinal, horizontal/vertical transverse)
- Uncertain observations (calibration targets) (10):
  - concentrations at the monitoring wells
- Unknown model parameters estimated using inversion
- Impact of uncertainty in calibration targets on model parameters is estimated using info-gap analyses
- Robustness and opportuness functions associated with predicted contaminant concentrations at the proposed new well locations are applied for decision analyses
- Decision question: which of the new proposed well location has the highest immunity of failure/windfall to detect concentrations above MCL (c > 5 ppm)
  i.e. which well provides the most robust/opportune decision to improve the monitoring network

Info-Gap Analysis: Network Design

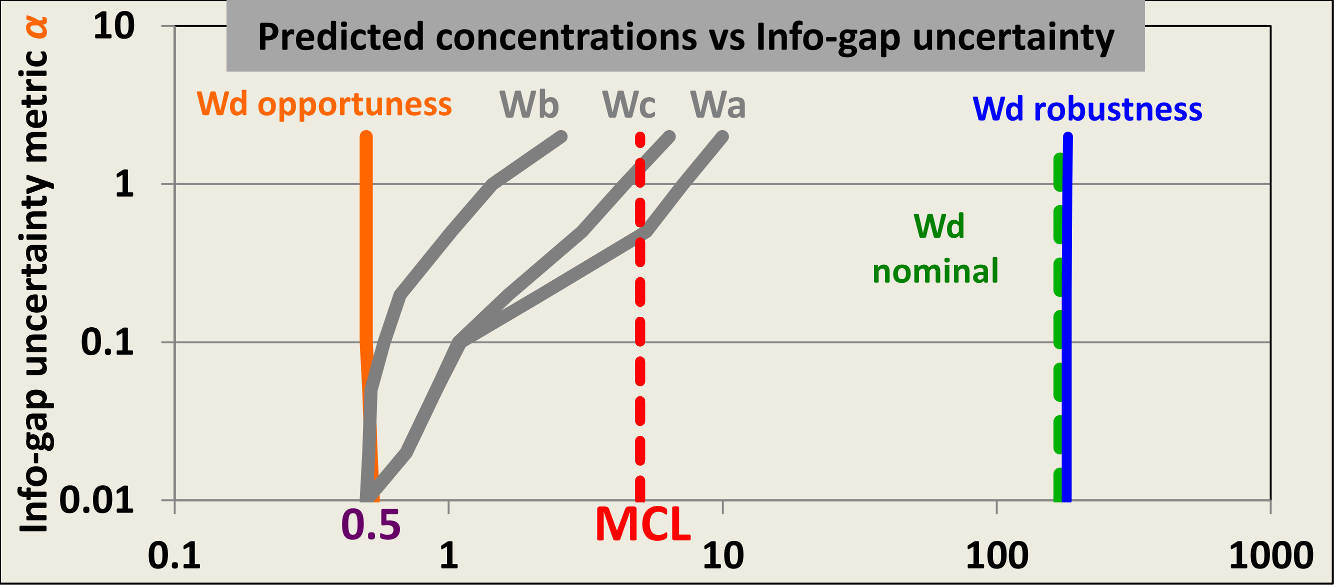
- Calibration targets are highly uncertain (PDF's cannot be defined) due to:
  - measurement errors
  - uncertain background concentrations
  - uncertain local hydrogeological and geochemical conditions

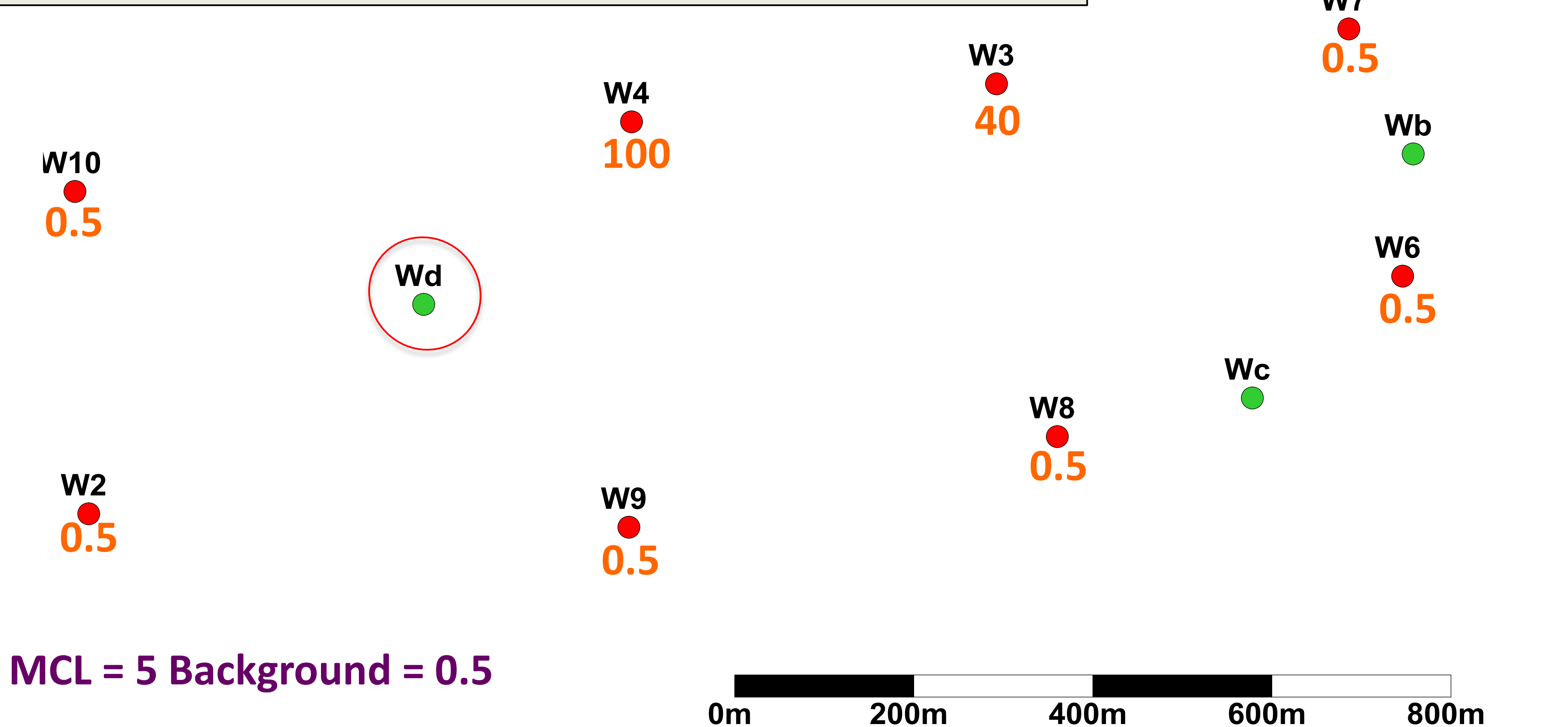
Info-Gap Analysis: Network Design





Info-Gap Analysis: Network Design





Info-Gap Application: Case 2

Remediation of contamination in a aquifer through contaminant source control

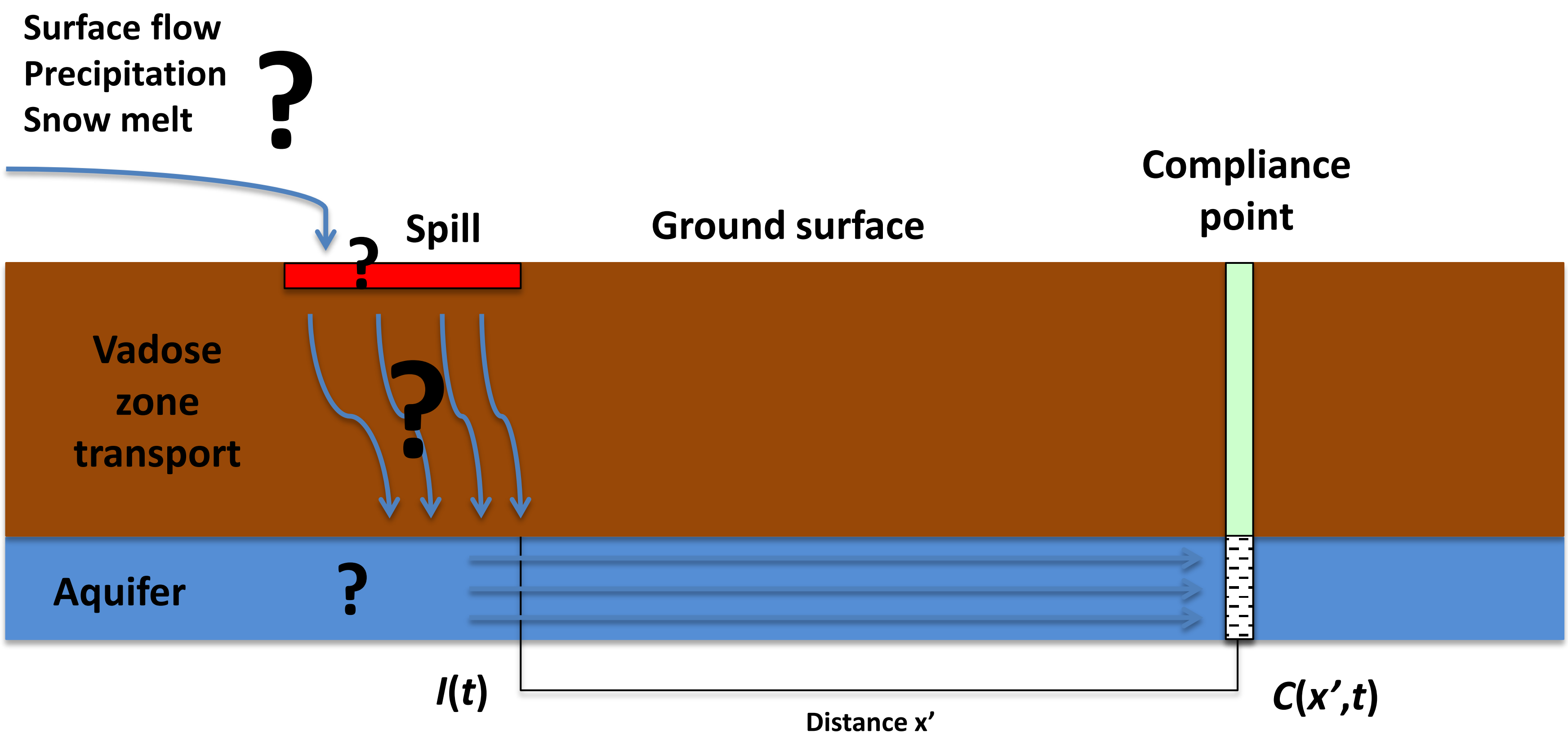
Harp & Vesselinov (2011). *Contaminant remediation decision analysis using information gap theory*. SERRA.

Info-Gap Analysis: Remediation of contaminant source

<u>Simple contaminant remediation problem</u>:

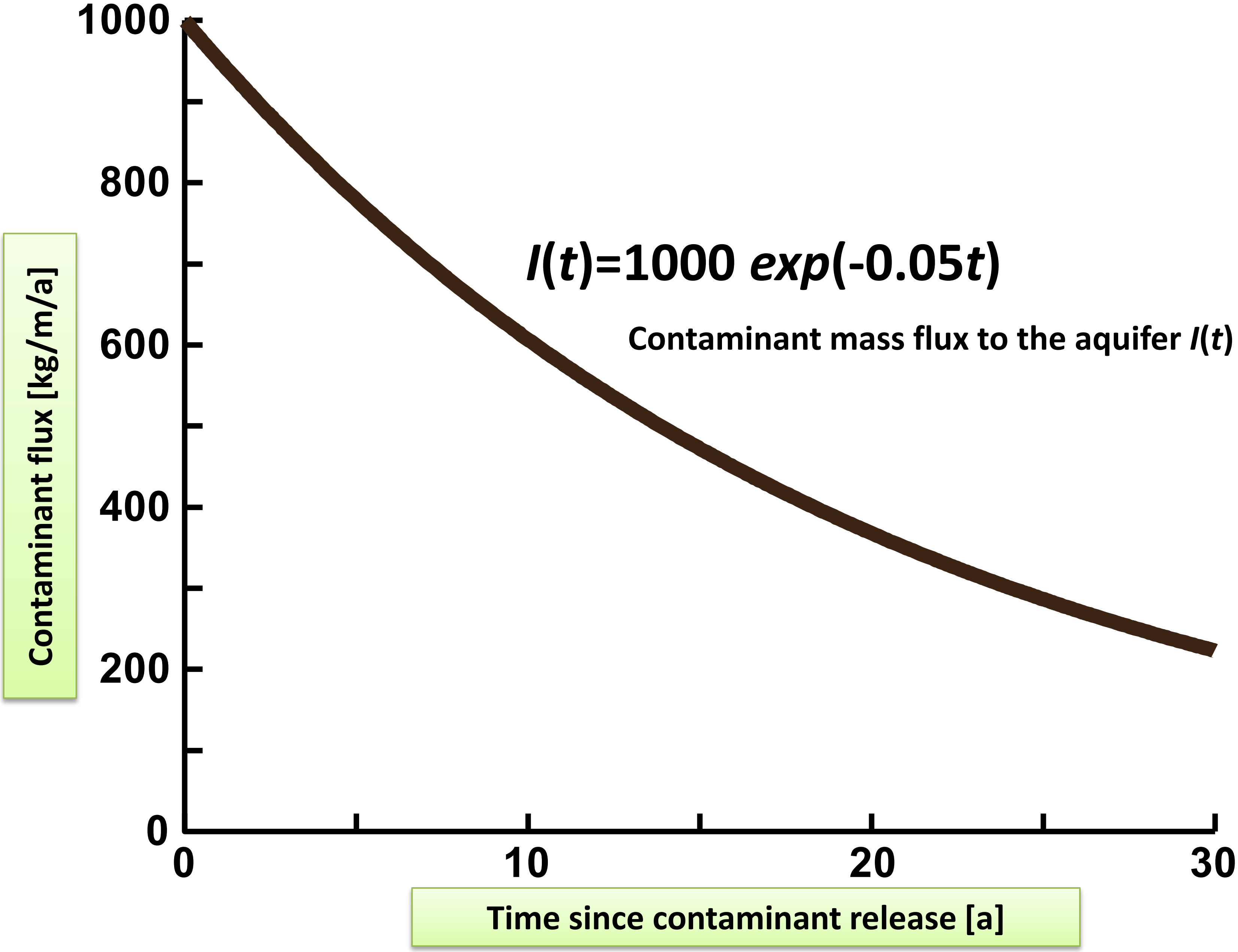
✧ how much contaminant mass needs to be removed to satisfy compliance requirement $C(x',t)$ < MCL

✧ lack of probabilistic (frequency of occurrence) information about the contaminant mass flux to aquifer $I(t)$

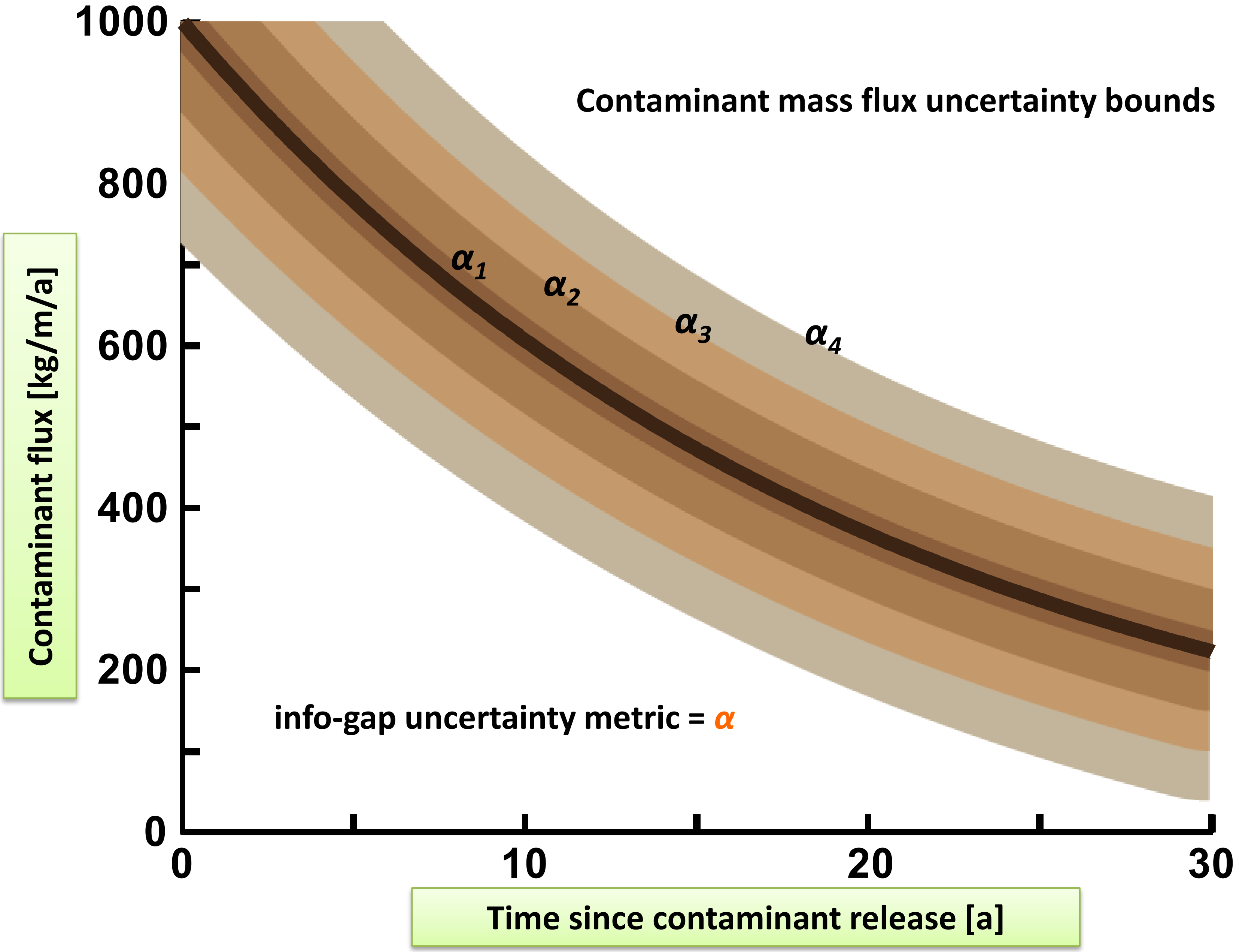


Harp & Vesselinov (2011). *Contaminant remediation decision analysis using information gap theory*. SERRA.

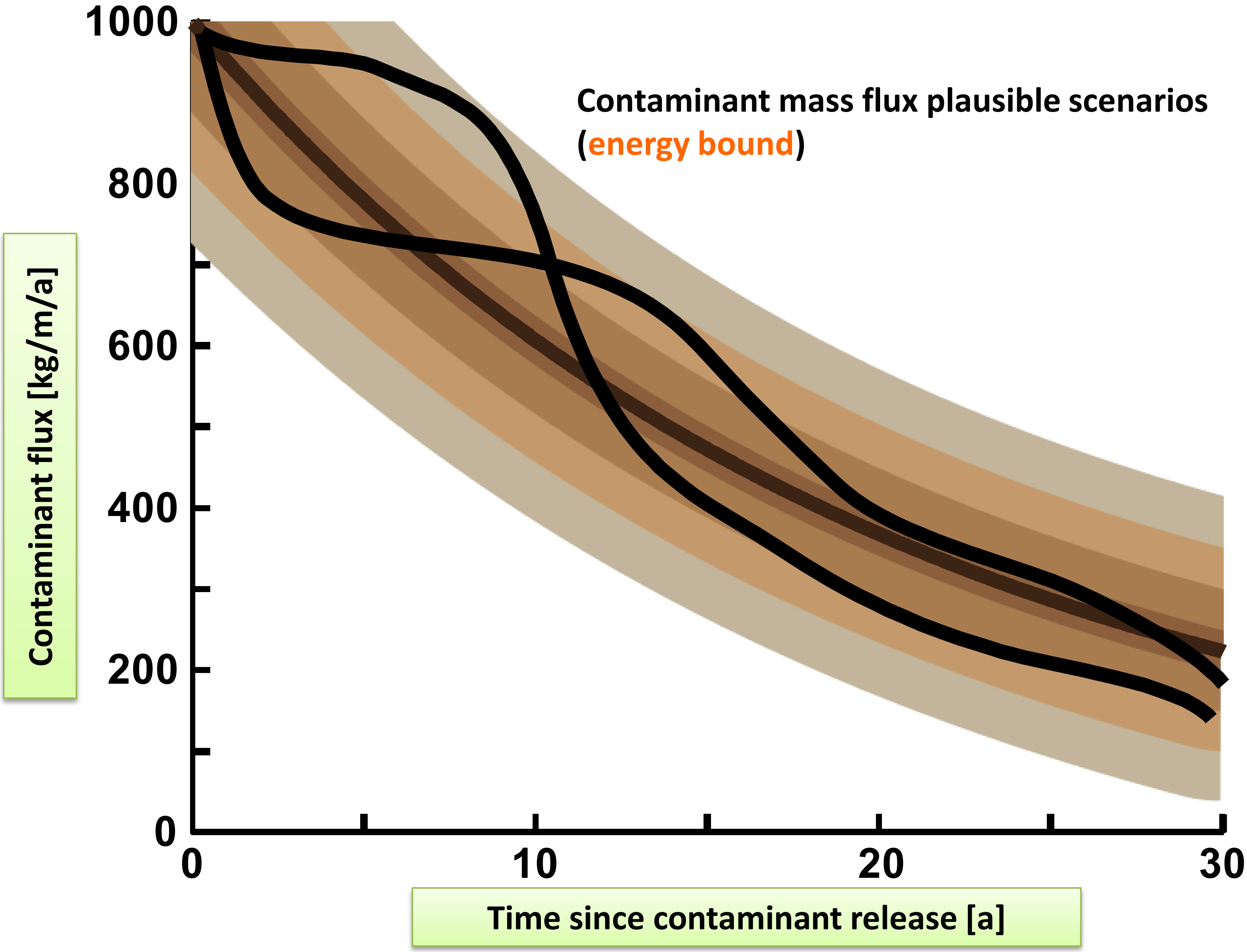
Info-Gap Analysis: Remediation of contaminant source



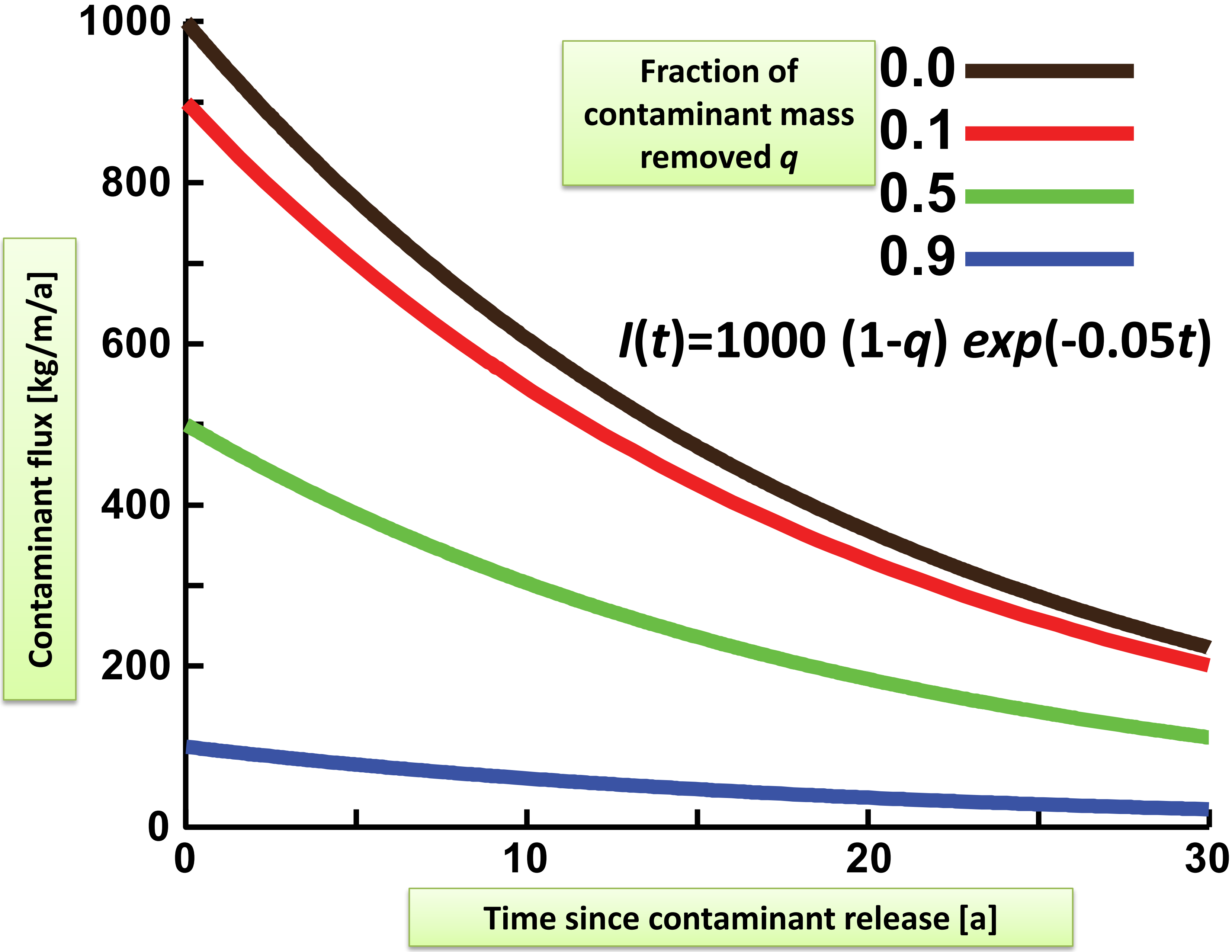
Info-Gap Analysis: Remediation of contaminant source



Info-Gap Analysis: Remediation of contaminant source

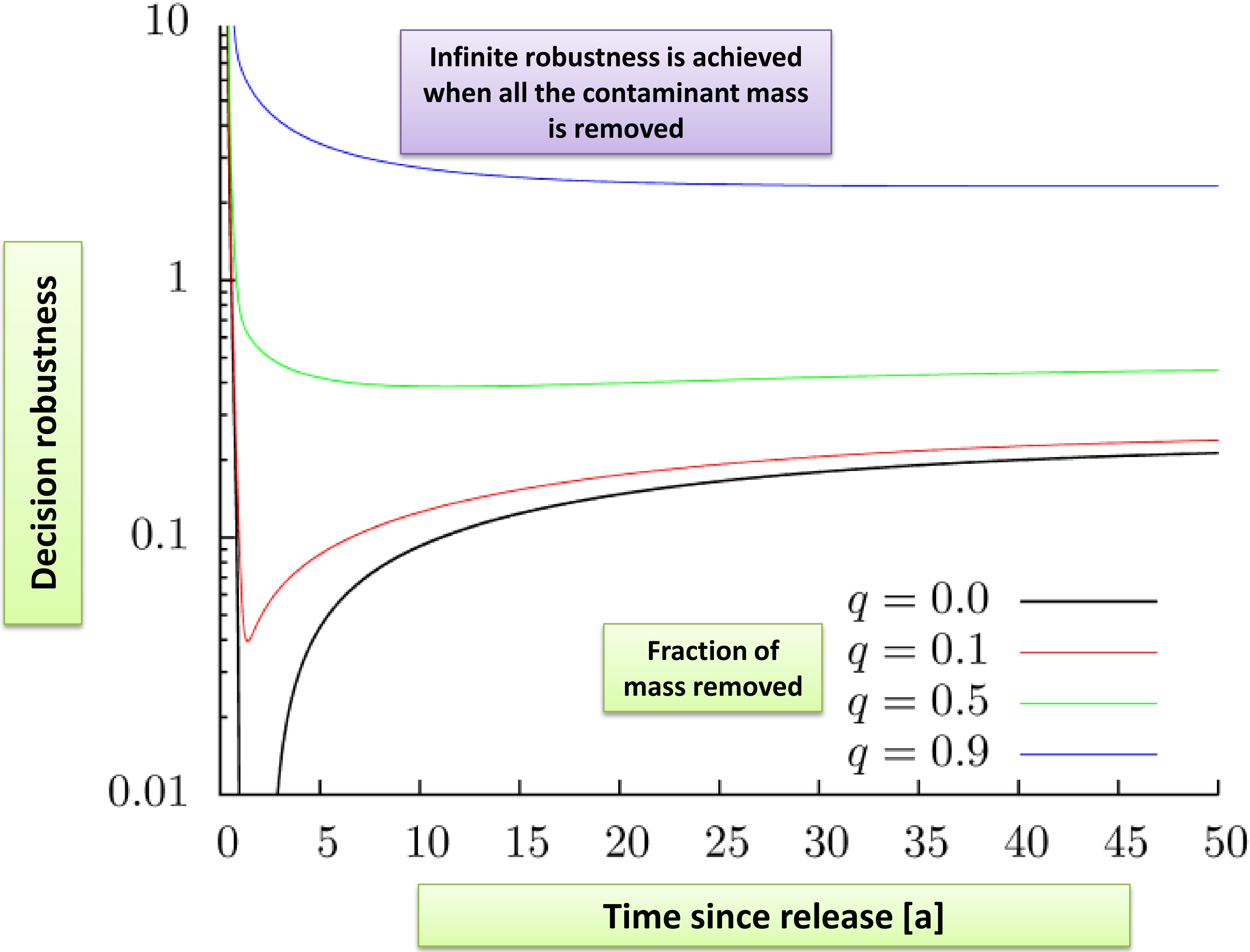


Info-Gap Analysis: Remediation of contaminant source

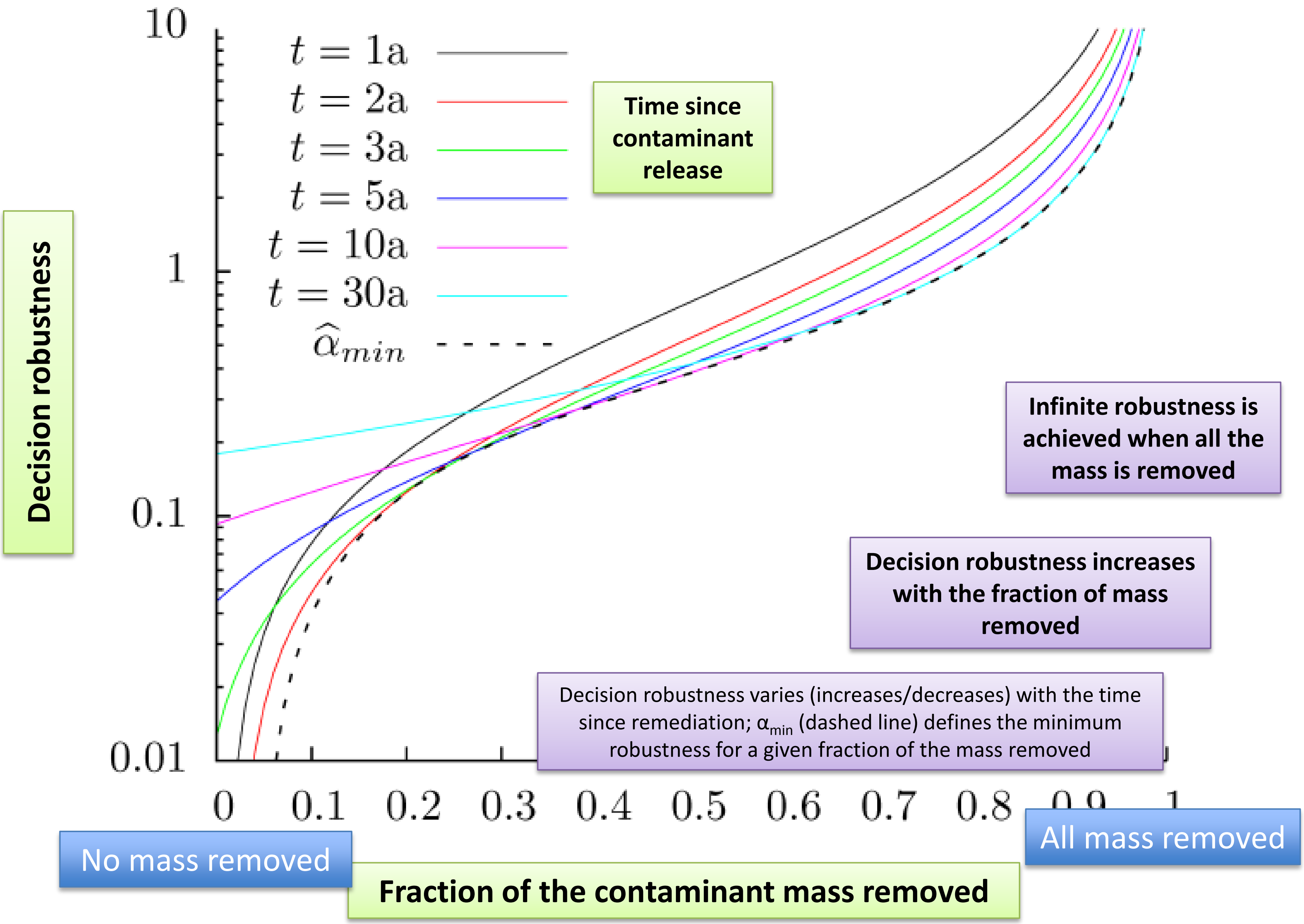


Info-Gap Analysis: Remediation of contaminant source

Decision robustness defines how much contaminant mass should be removed and still be immune to failure considering lack of information about the contaminant mass flux



Info-Gap Analysis: Remediation of contaminant source

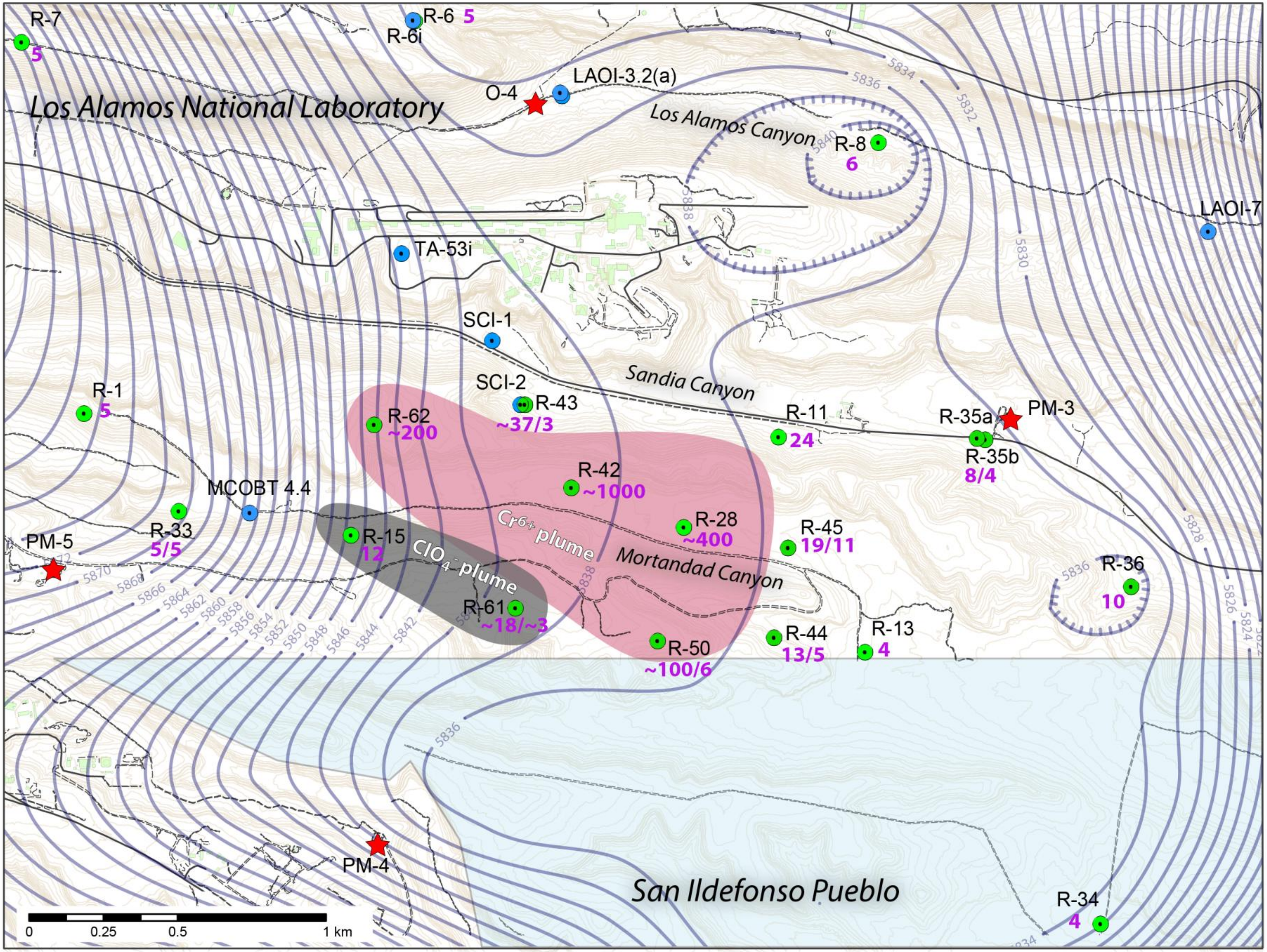


Chromium plume in the regional aquifer at LANL

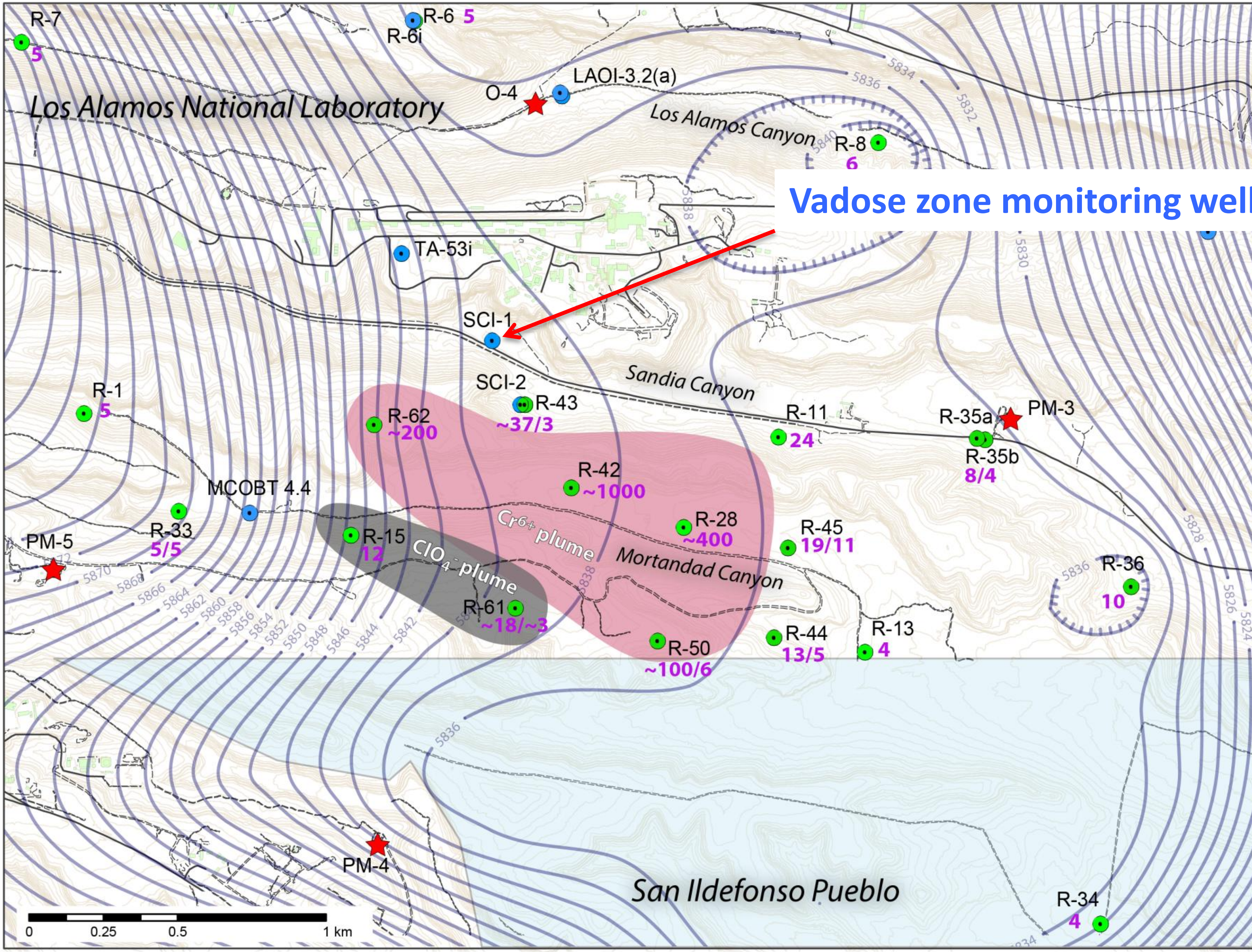
<u>GOALS</u>:

- **provide model-based decision support related to chromium transport in the vadose zone and regional aquifer at LANL**
- **apply advanced computationally efficient methods for:**
  - **parameter estimation (PE)**
  - **model calibration**
  - **model-based uncertainty quantification (UQ)**
  - **risk analysis (RA), and**
  - **decision support (DS)**
- **utilize high-performance computing due to high computational demands for model simulations and model analyses**

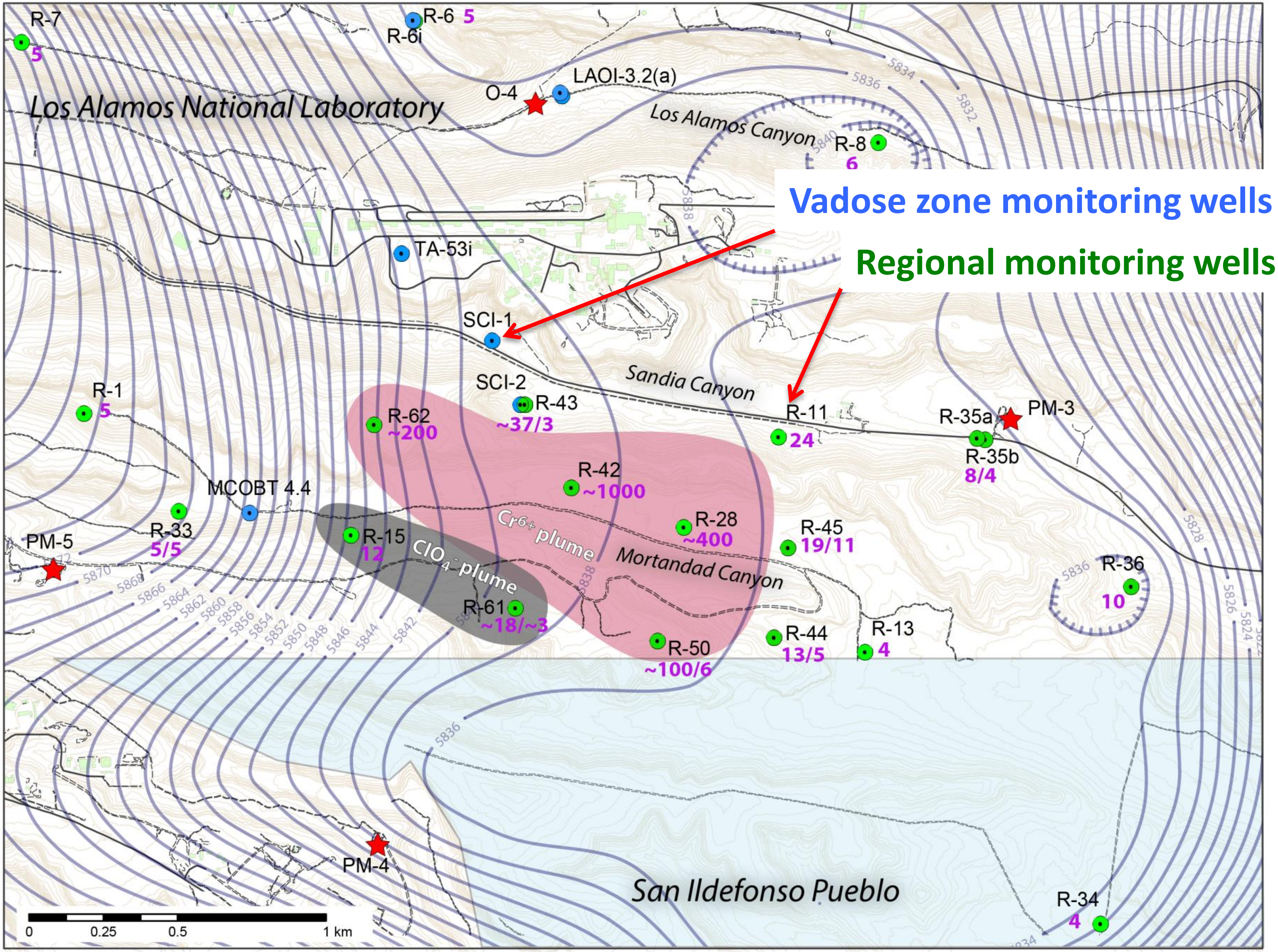
Chromium plume in the regional aquifer at LANL



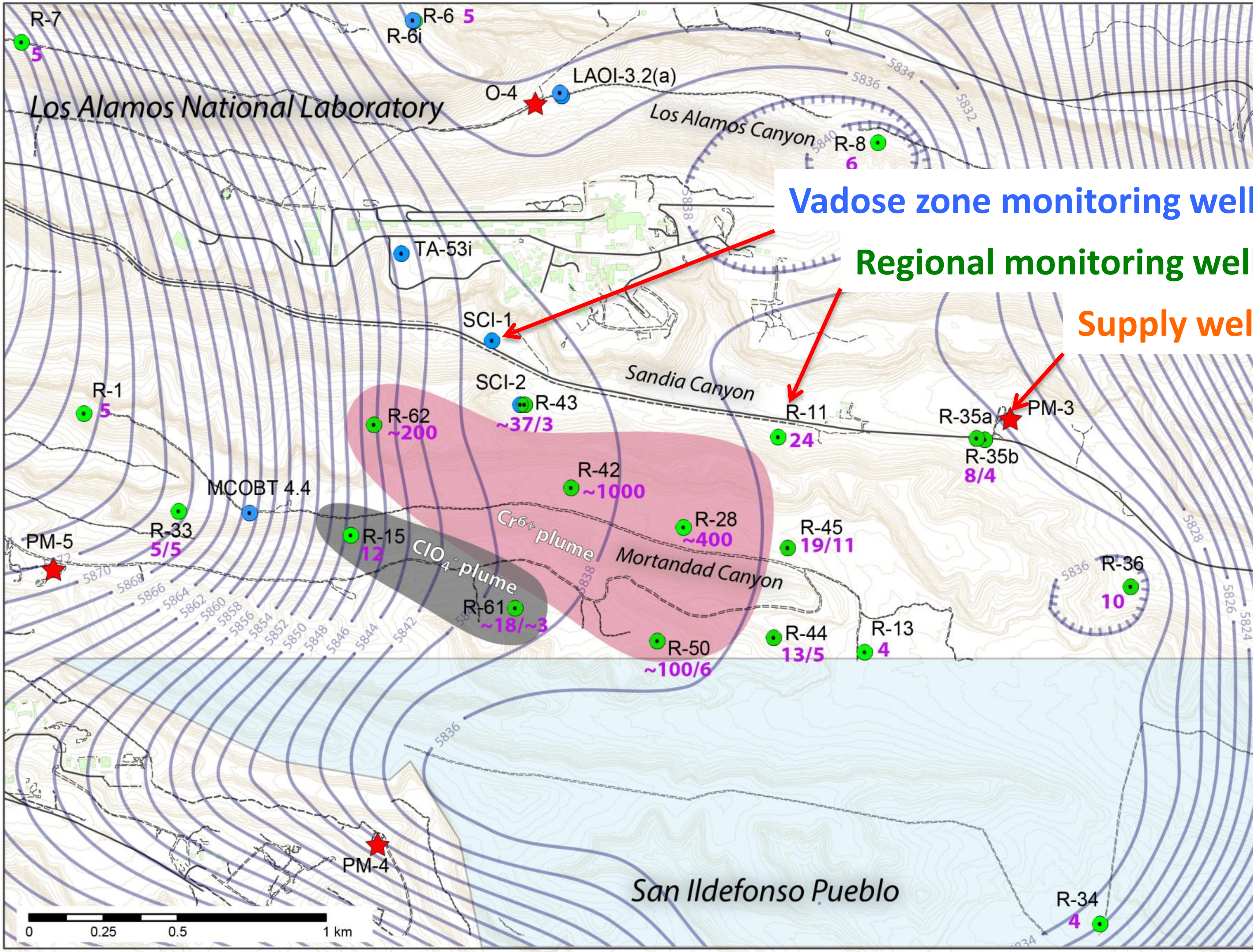
Chromium plume in the regional aquifer at LANL



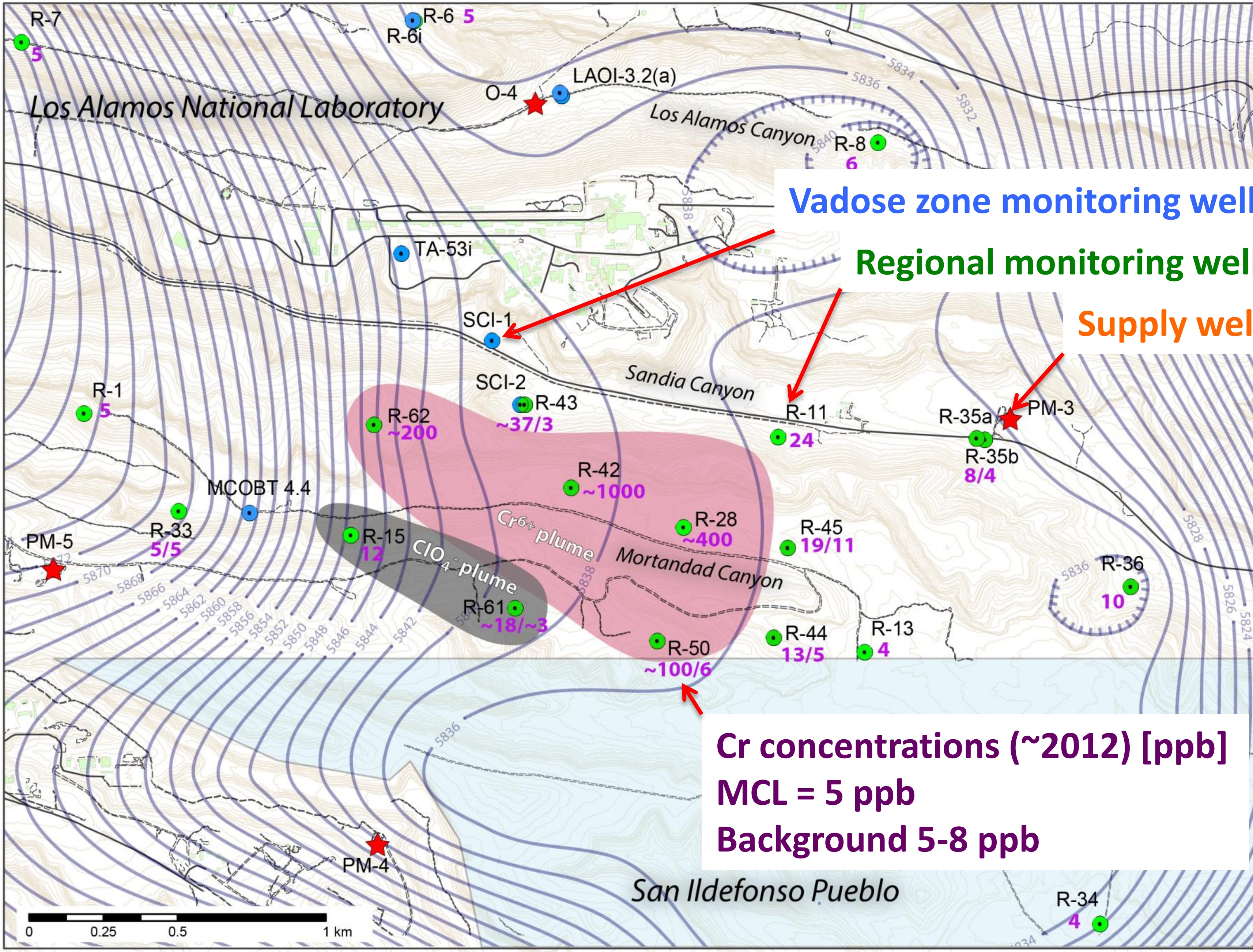
Chromium plume in the regional aquifer at LANL

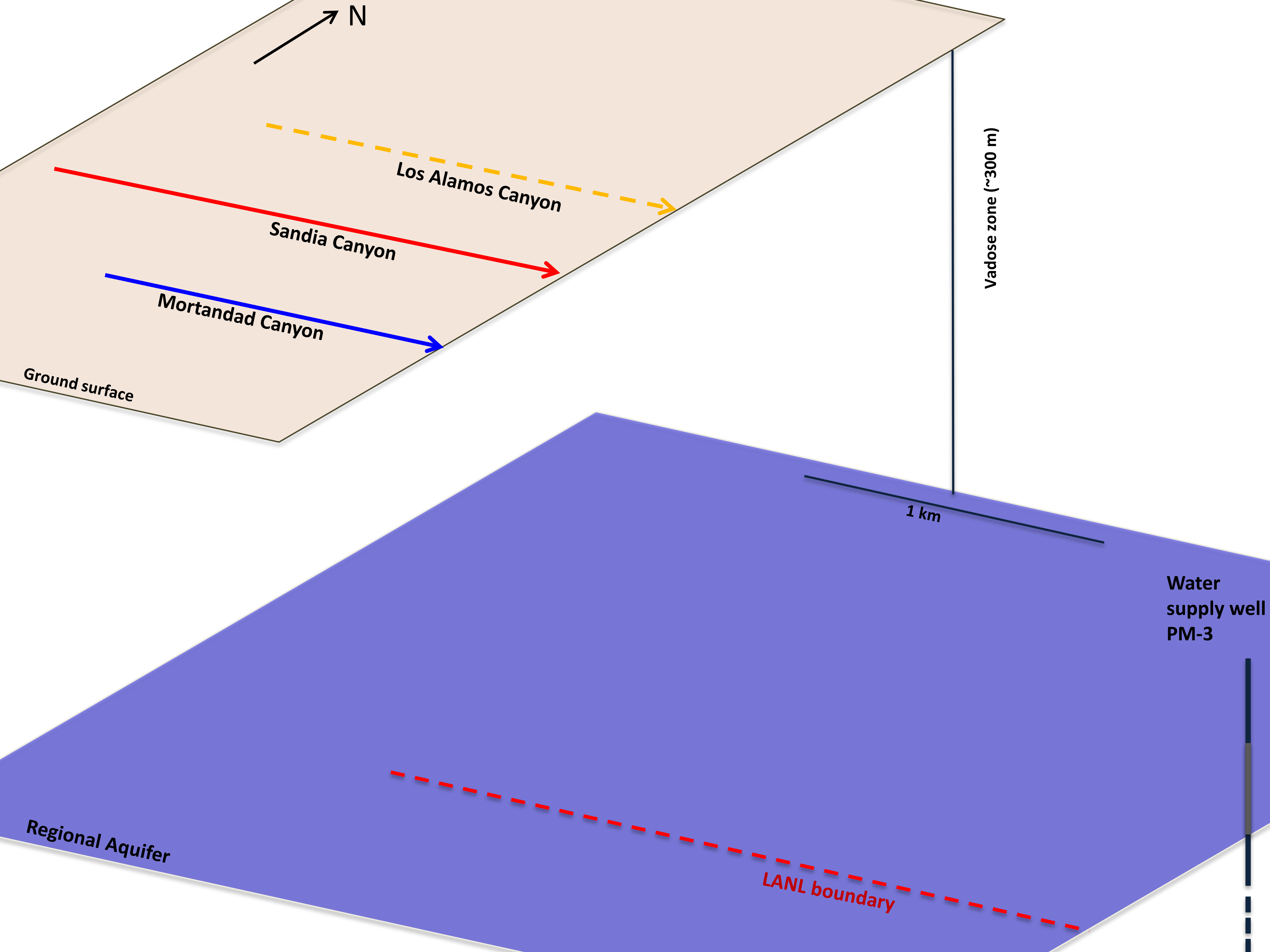


Chromium plume in the regional aquifer at LANL

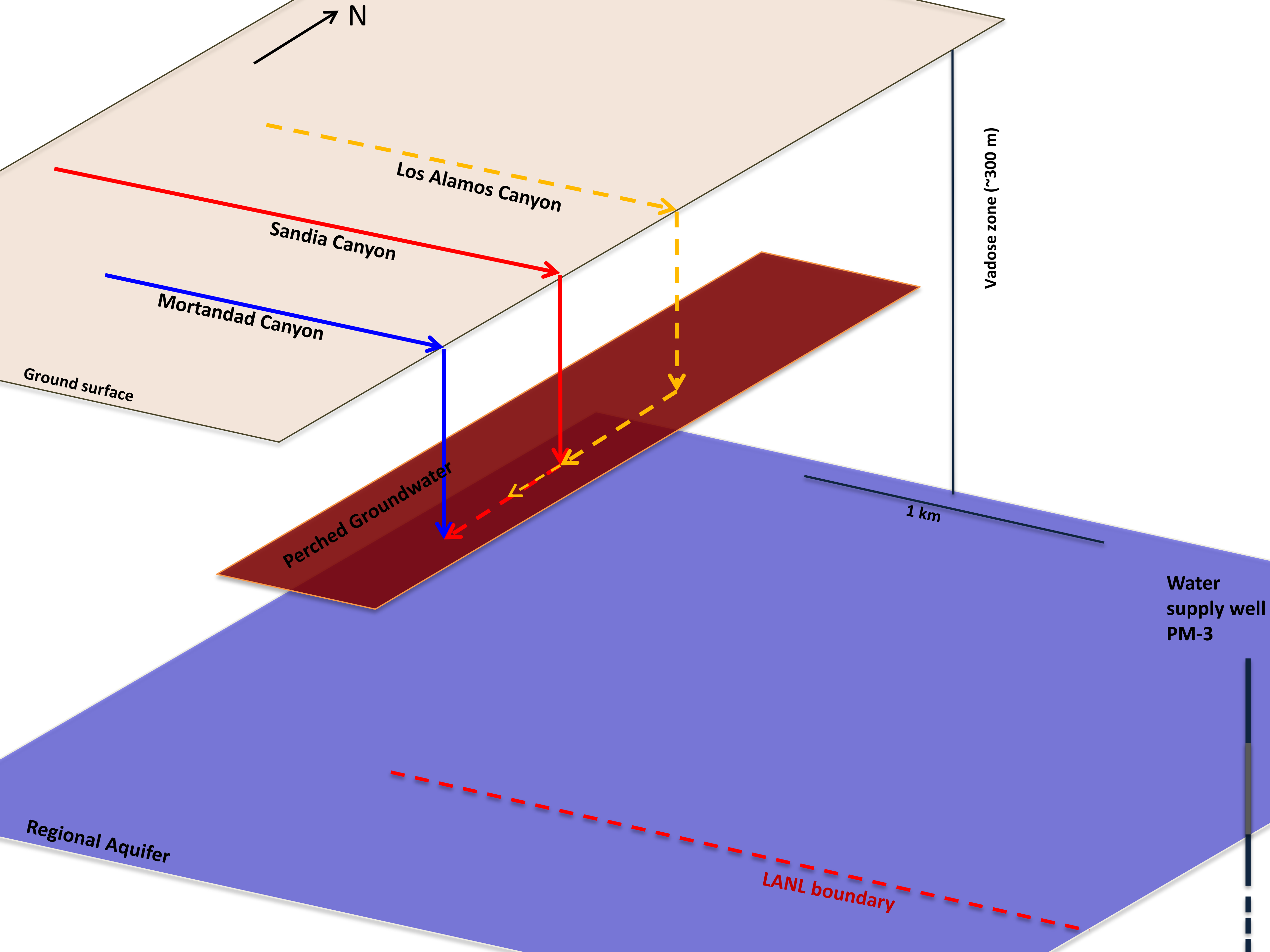


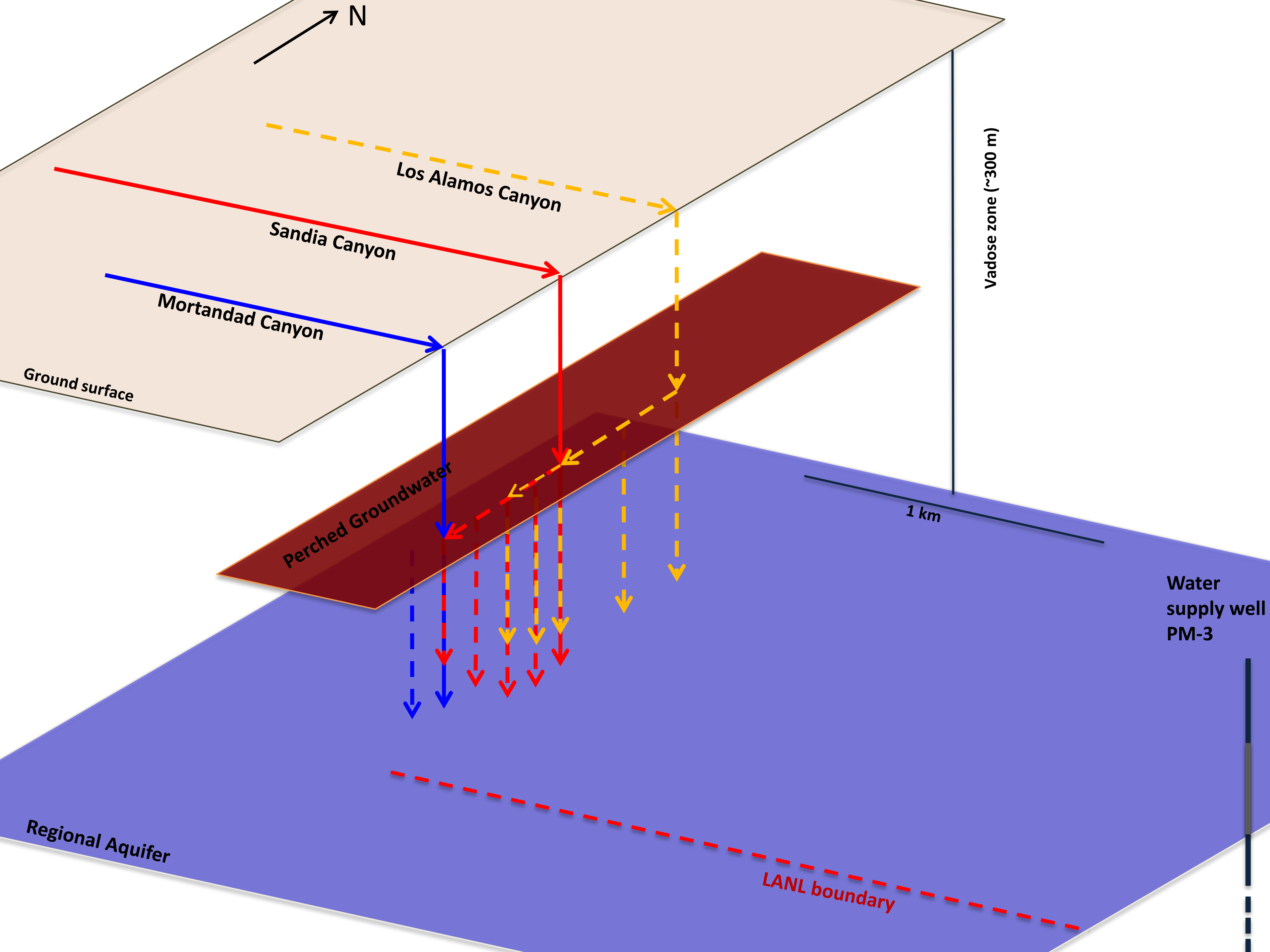
Chromium plume in the regional aquifer at LANL

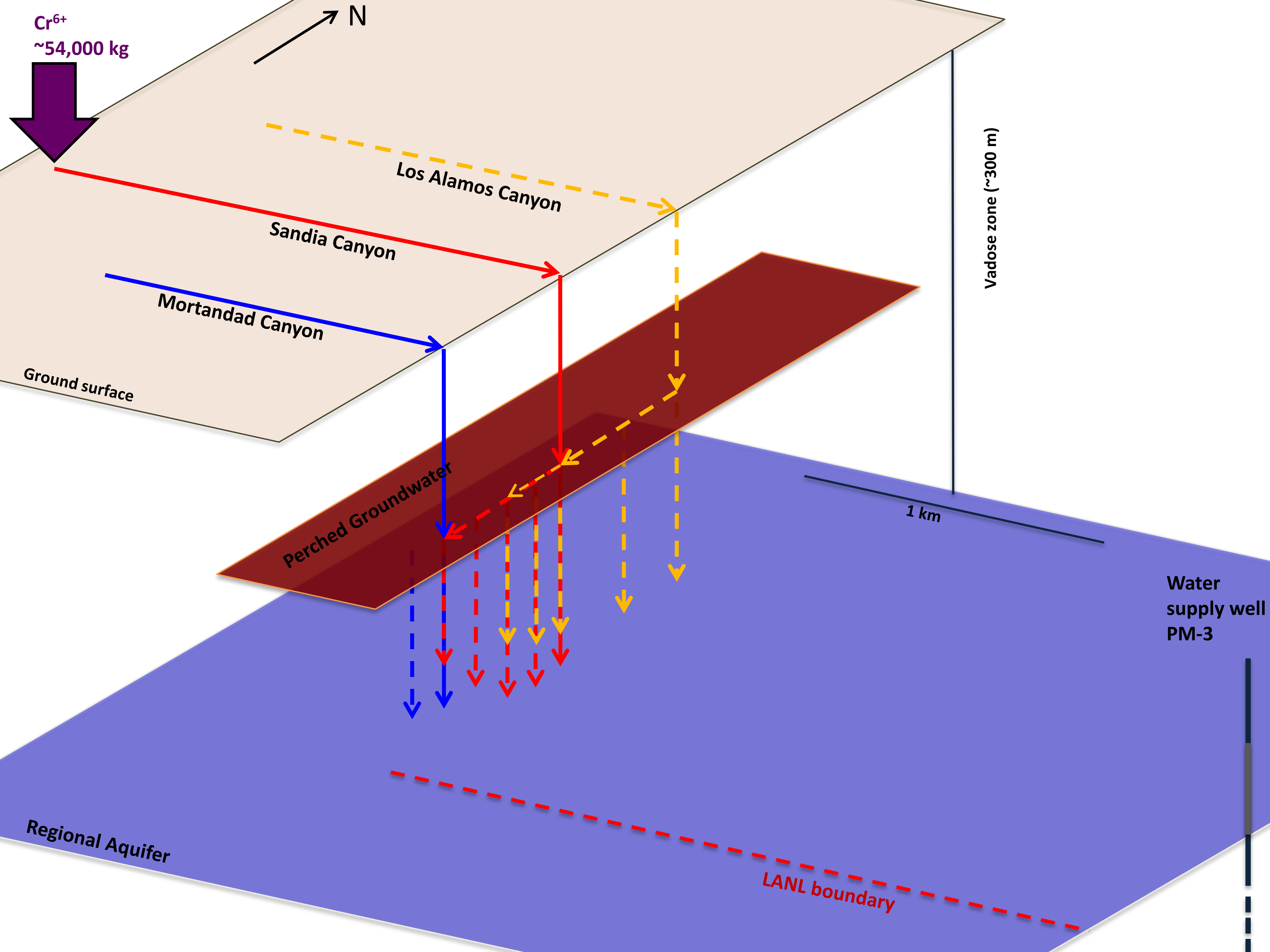




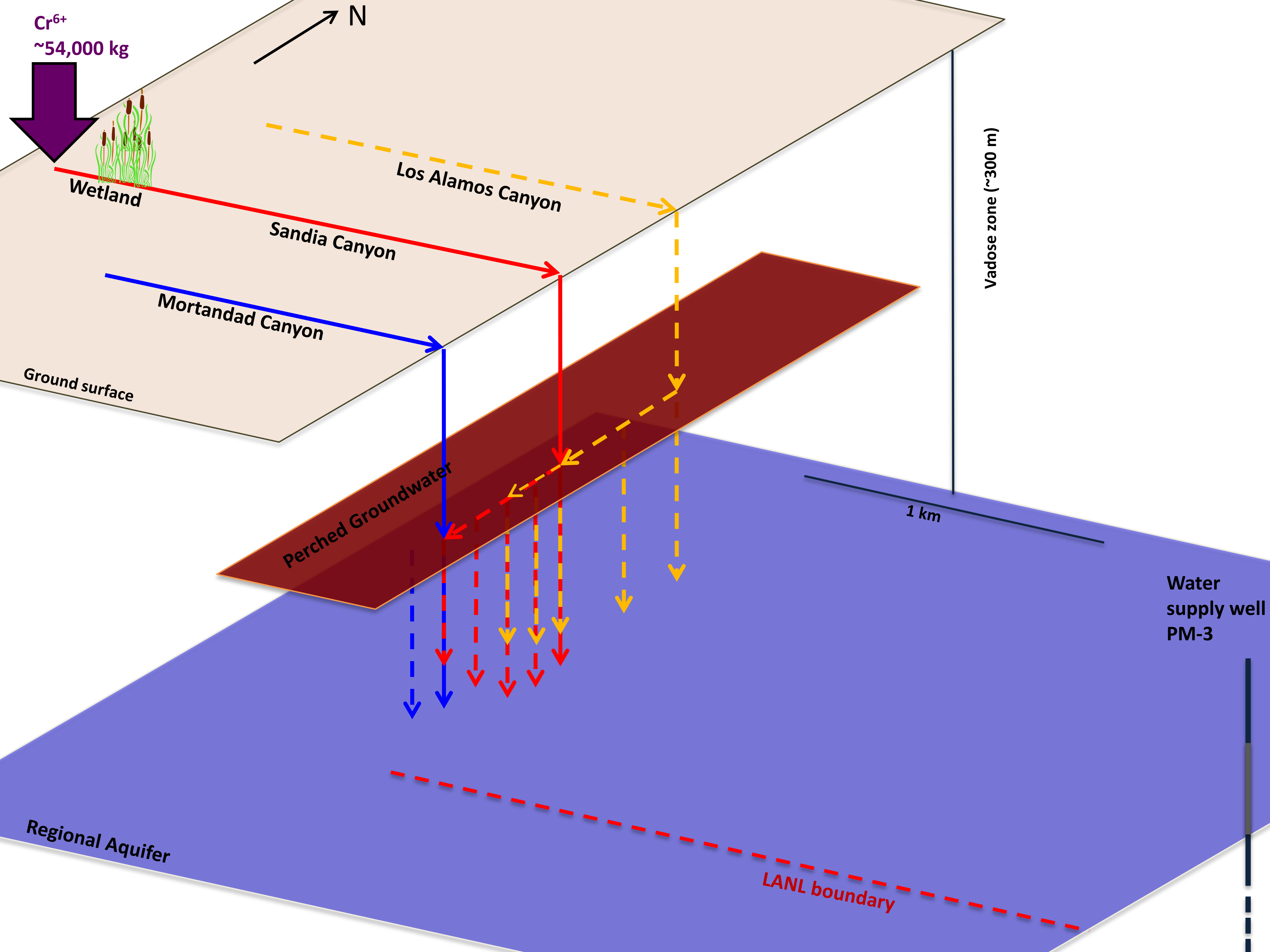
N
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
1 km
Water supply well PM-3
Regional Aquifer
LANL boundary


N
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water
supply well
PM-3
Regional Aquifer
LANL boundary

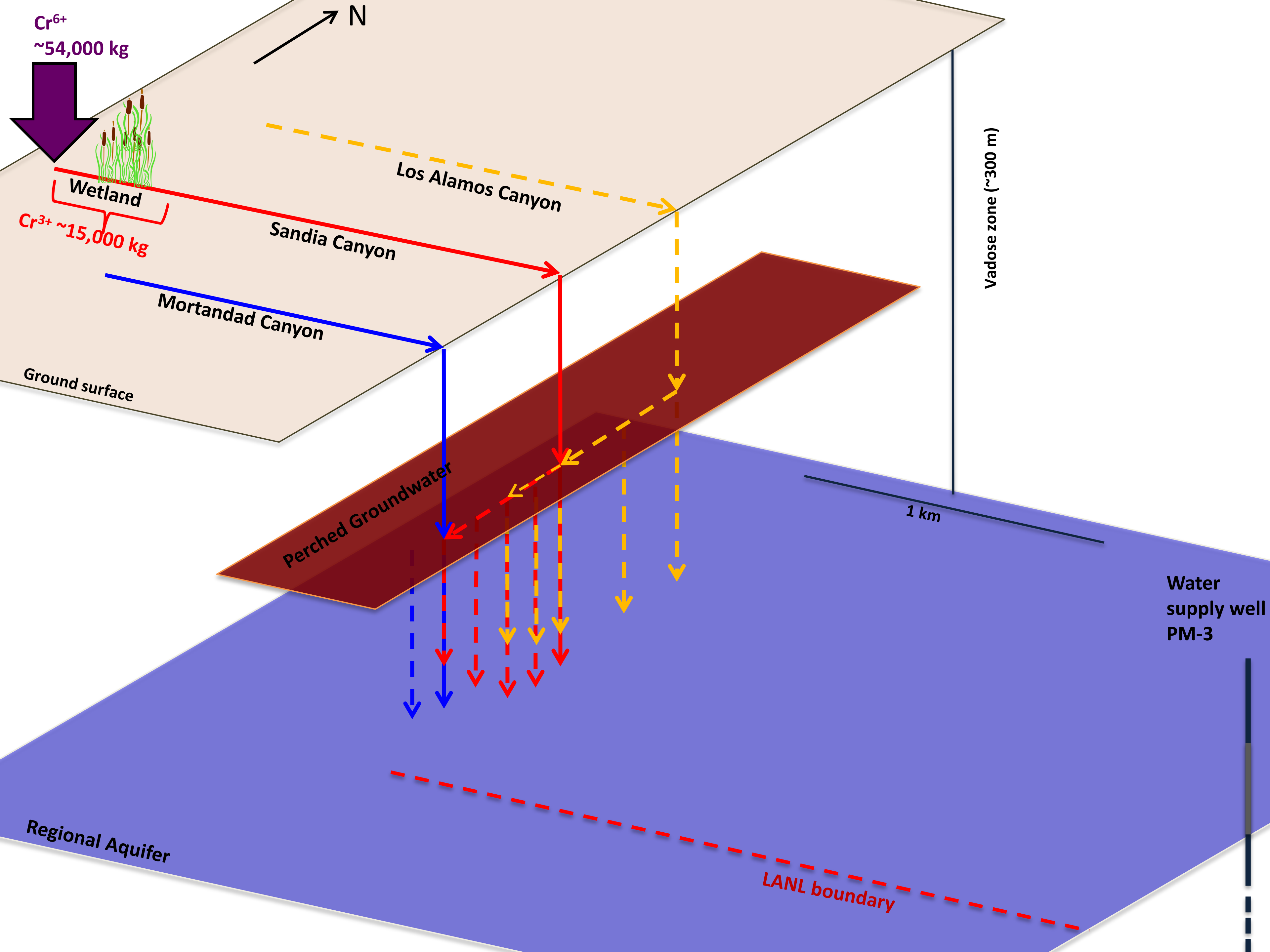

N
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water
supply well
PM-3
Regional Aquifer
LANL boundary

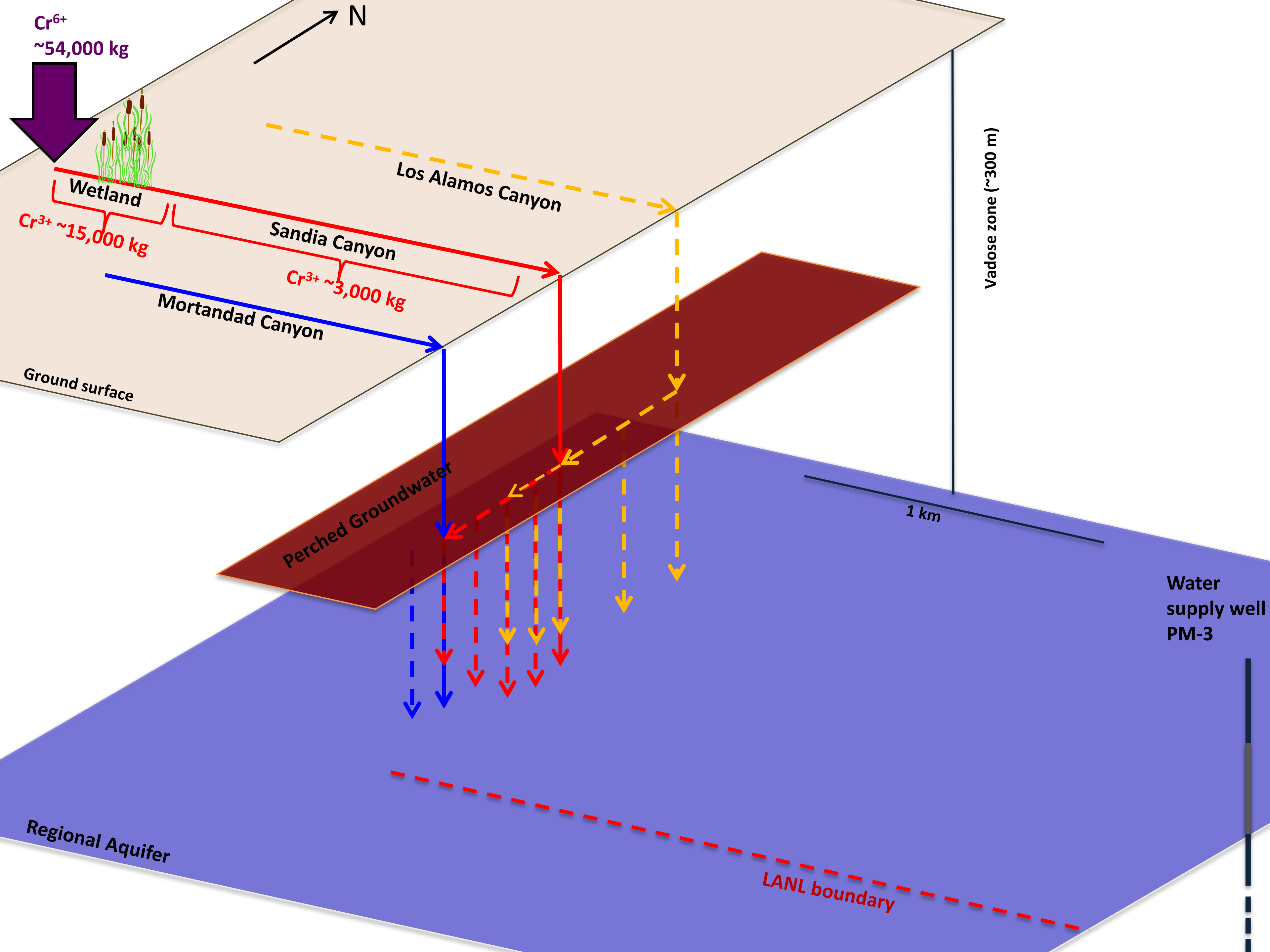


Cr6+
~54,000 kg
N
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water
supply well
PM-3
Regional Aquifer
LANL boundary

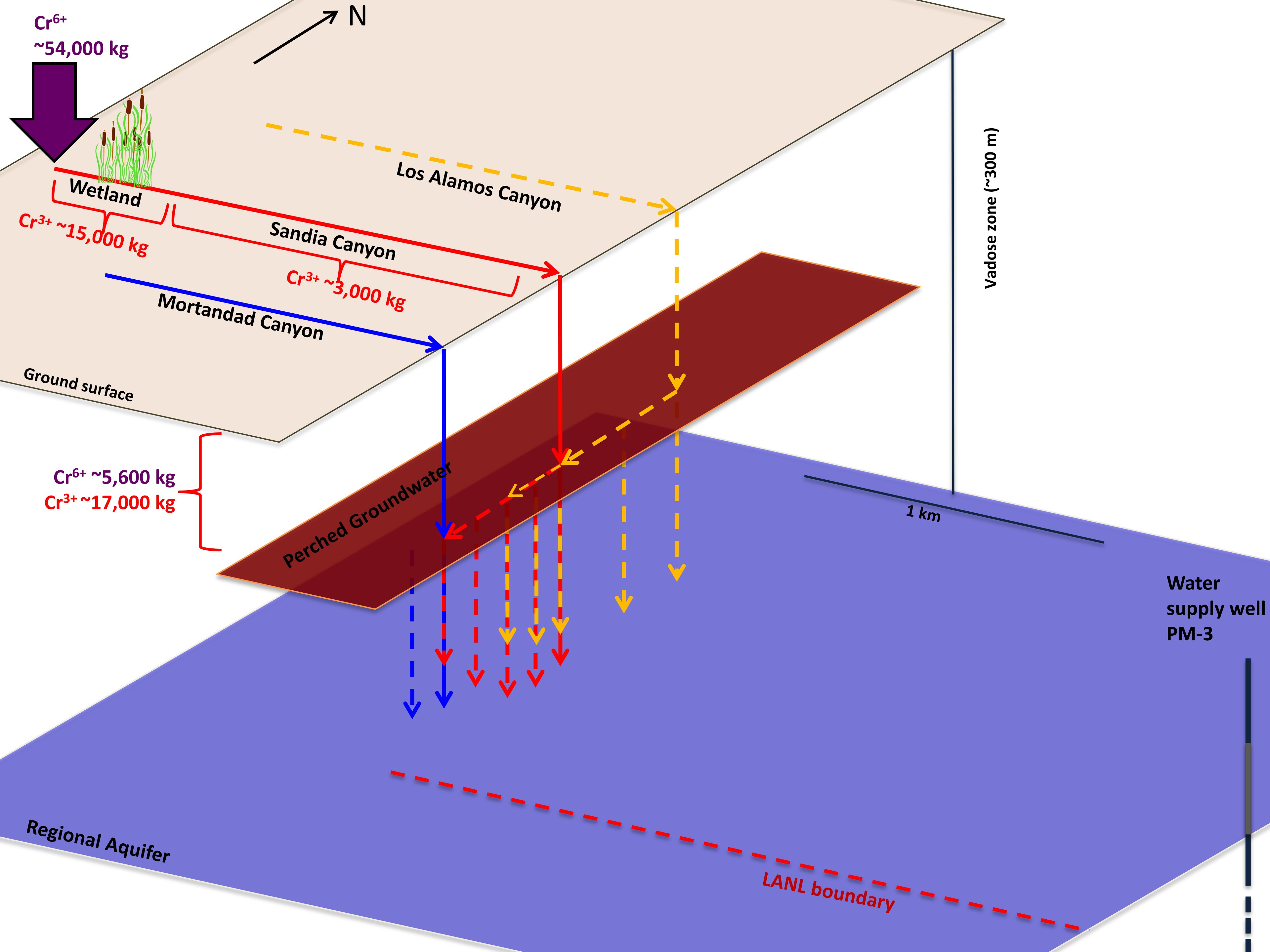


Cr^{6+}
~54,000 kg
N
Wetland
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water supply well PM-3
Regional Aquifer
LANL boundary

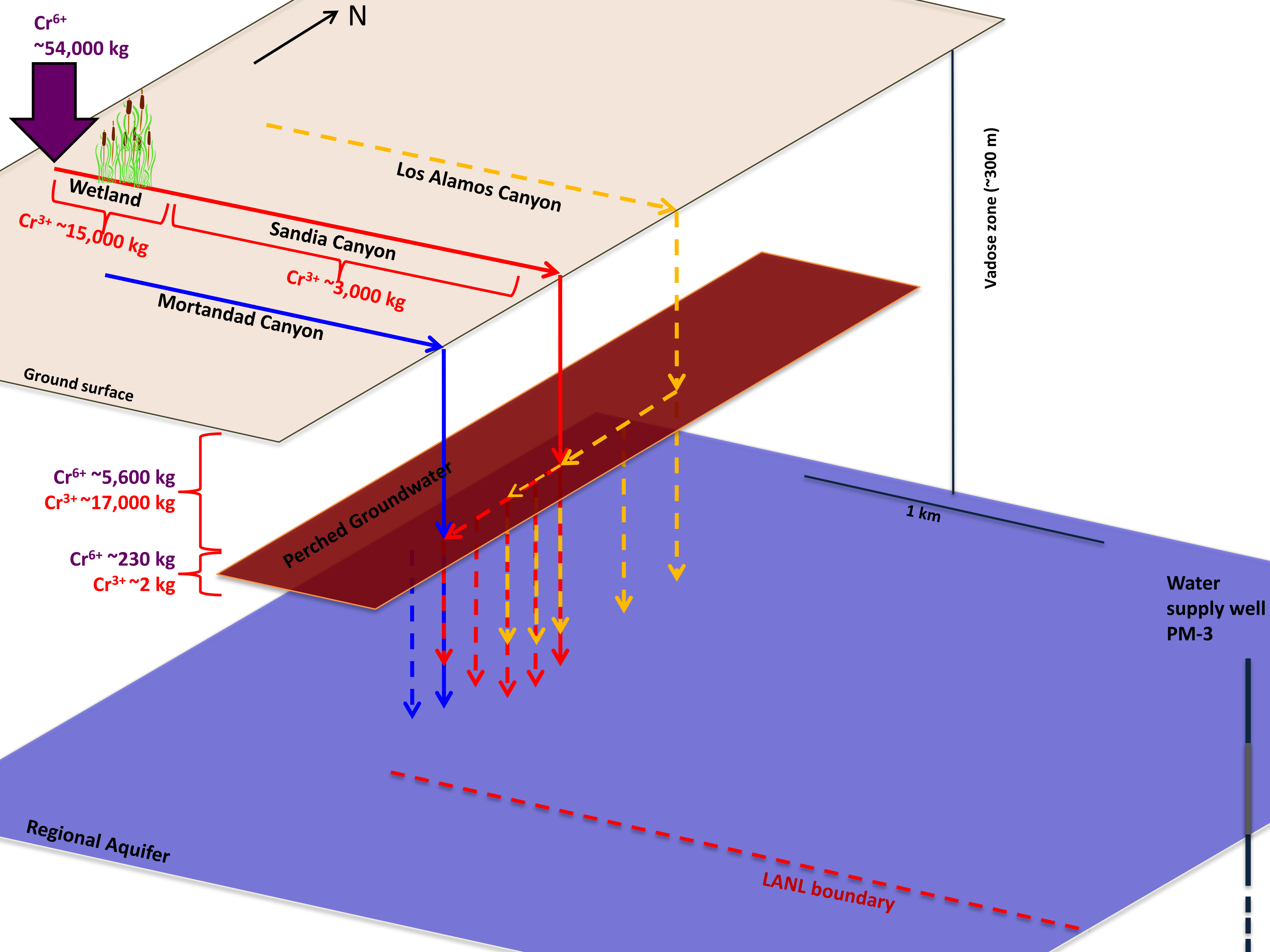

N
Cr^{6+}
~54,000 kg
Wetland
Cr^{3+} ~15,000 kg
Los Alamos Canyon
Sandia Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water supply well PM-3
Regional Aquifer
LANL boundary



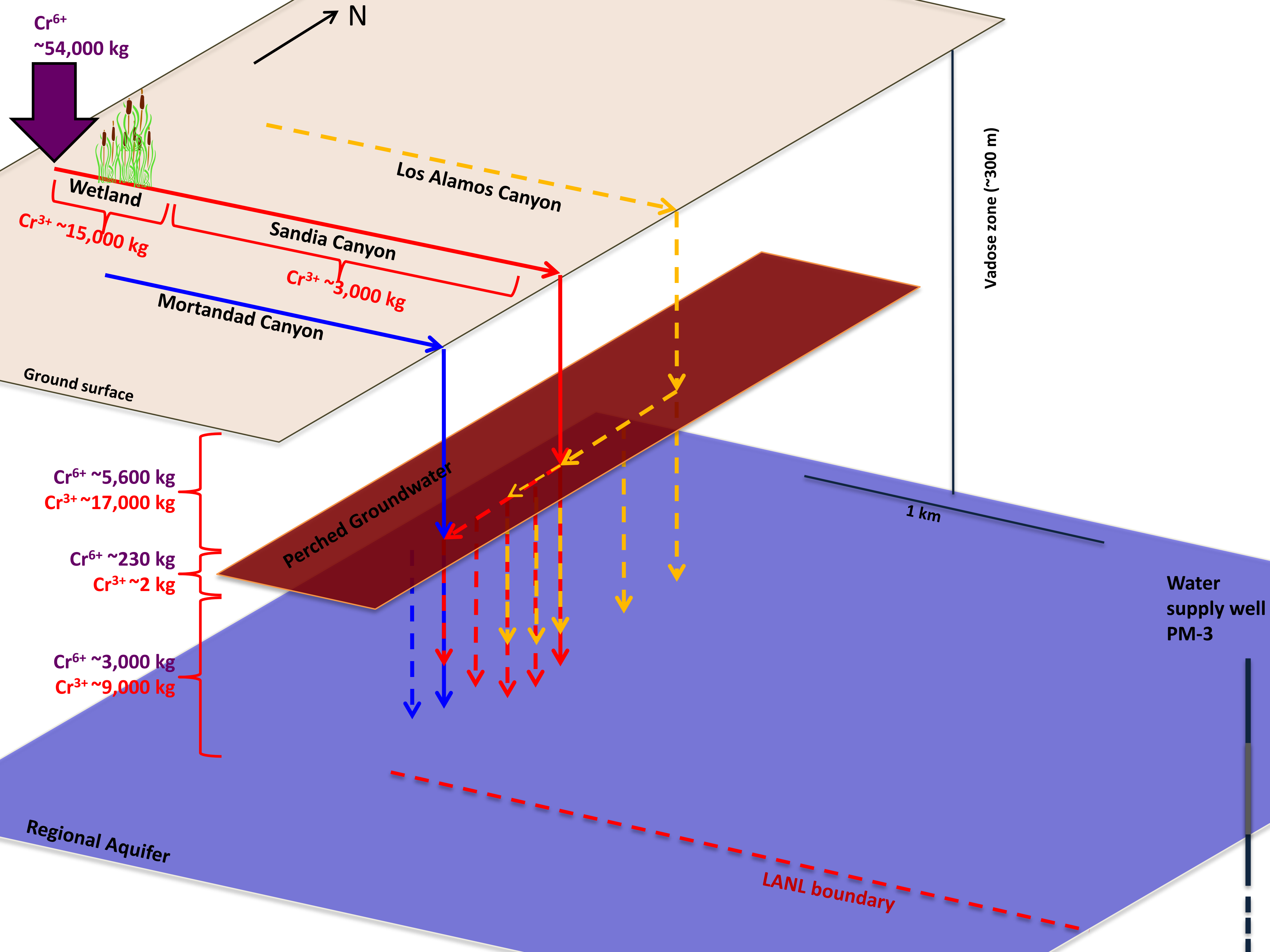
Cr6+
~54,000 kg
N
Wetland
Cr3+ ~15,000 kg
Sandia Canyon
Cr3+ ~3,000 kg
Los Alamos Canyon
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
1 km
Water
supply well
PM-3
Regional Aquifer
LANL boundary



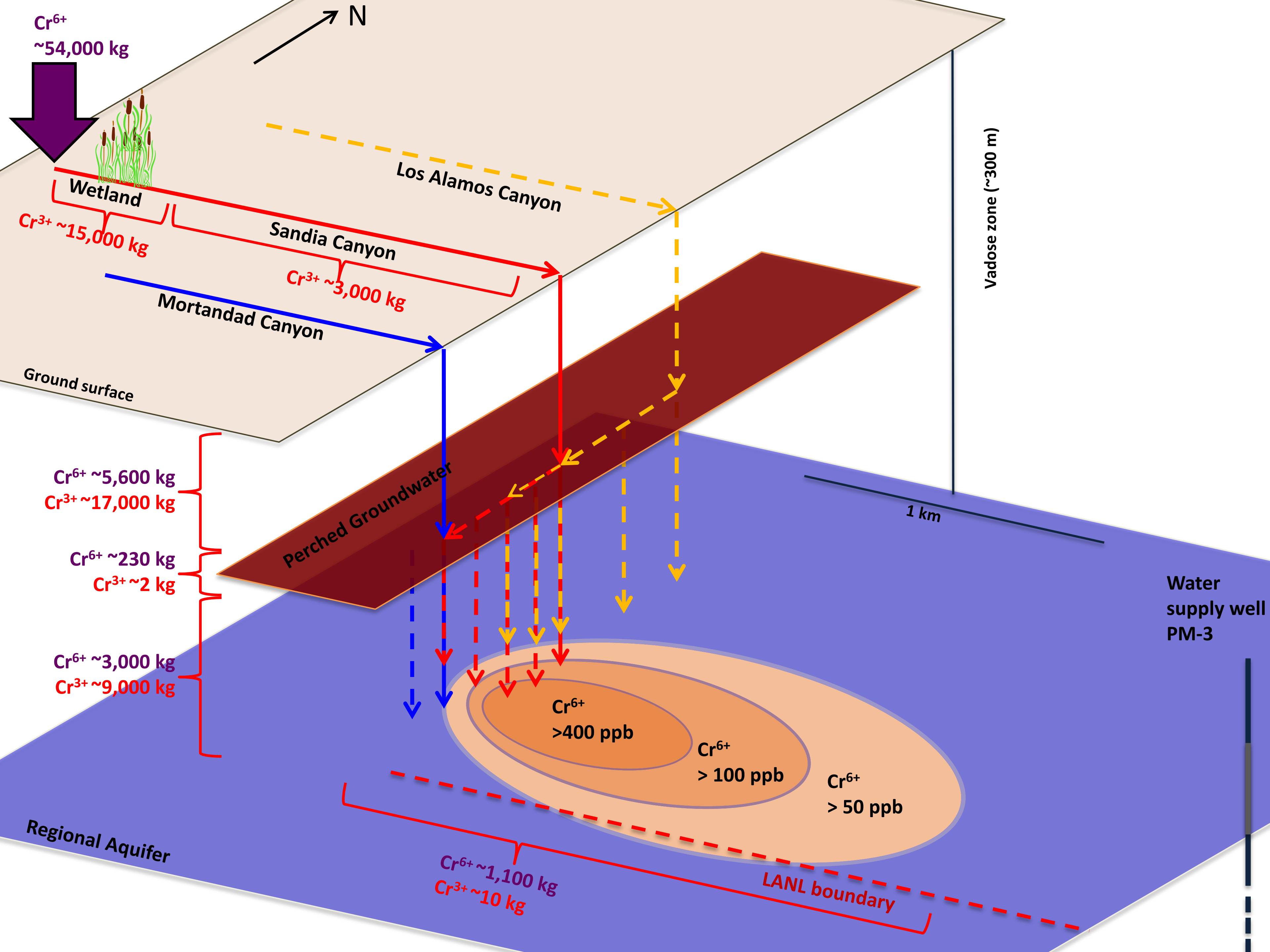
N
Cr6+
~54,000 kg
Wetland
Cr3+ ~15,000 kg
Los Alamos Canyon
Sandia Canyon
Cr3+ ~3,000 kg
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Cr6+ ~5,600 kg
Cr3+ ~17,000 kg
Perched Groundwater
1 km
Water supply well PM-3
Regional Aquifer
LANL boundary



N
Cr6+
~54,000 kg
Wetland
Cr3+ ~15,000 kg
Los Alamos Canyon
Sandia Canyon
Cr3+ ~3,000 kg
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Cr6+ ~5,600 kg
Cr3+ ~17,000 kg
Cr6+ ~230 kg
Cr3+ ~2 kg
Perched Groundwater
1 km
Water
supply well
PM-3
Regional Aquifer
LANL boundary



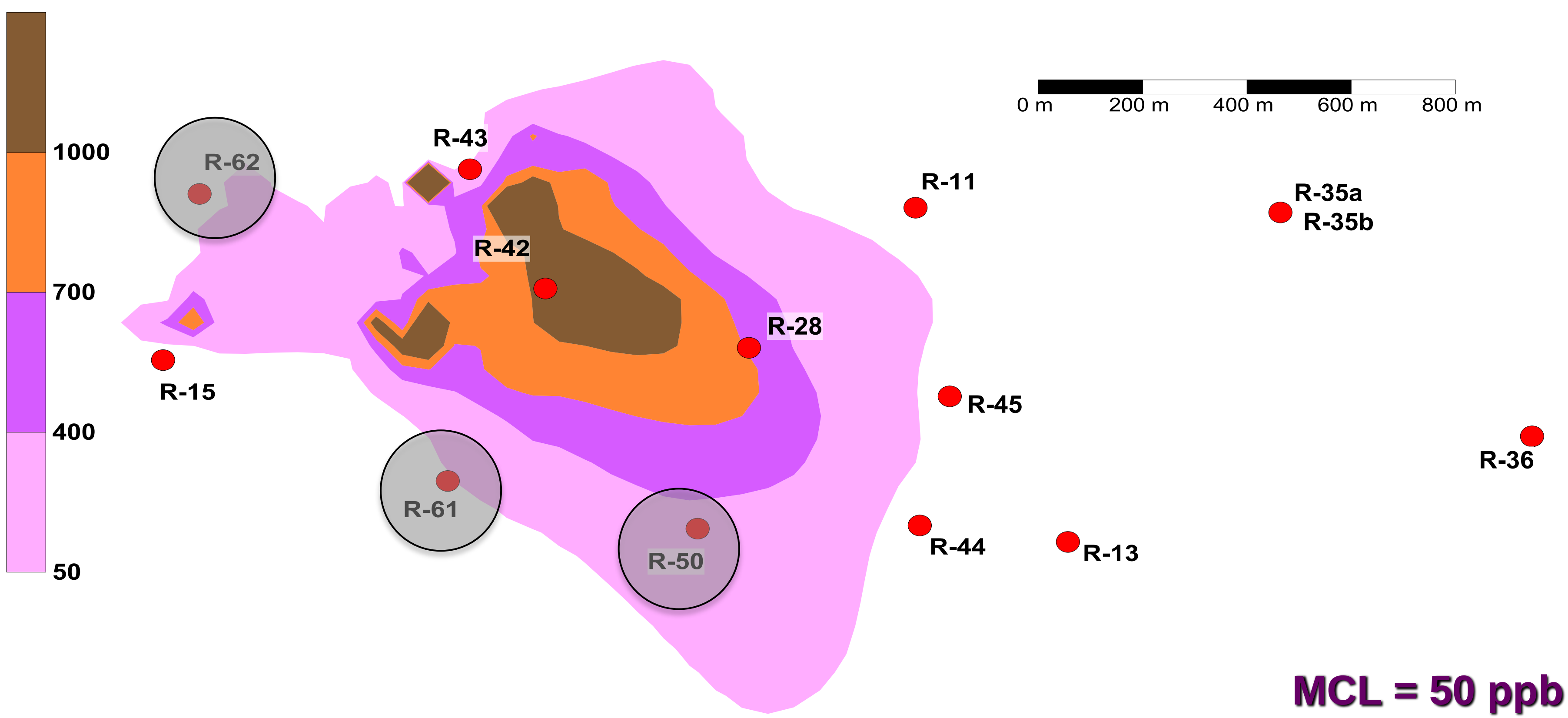
Cr^{6+}
~54,000 kg
N
Los Alamos Canyon
Wetland
Cr^{3+} ~15,000 kg
Sandia Canyon
Cr^{3+} ~3,000 kg
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Cr^{6+} ~5,600 kg
Cr^{3+} ~17,000 kg
Perched Groundwater
Cr^{6+} ~230 kg
Cr^{3+} ~2 kg
1 km
Water
supply well
PM-3
Cr^{6+} ~3,000 kg
Cr^{3+} ~9,000 kg
Regional Aquifer
LANL boundary



Cr6+
~54,000 kg
N
Wetland
Cr3+ ~15,000 kg
Los Alamos Canyon
Sandia Canyon
Cr3+ ~3,000 kg
Mortandad Canyon
Ground surface
Vadose zone (~300 m)
Perched Groundwater
Cr6+ ~5,600 kg
Cr3+ ~17,000 kg
Cr6+ ~230 kg
Cr3+ ~2 kg
Cr6+ ~3,000 kg
Cr3+ ~9,000 kg
1 km
Water
supply well
PM-3
Cr6+
>400 ppb
Cr6+
> 100 ppb
Cr6+
> 50 ppb
Regional Aquifer
Cr6+ ~1,100 kg
Cr3+ ~10 kg
LANL boundary

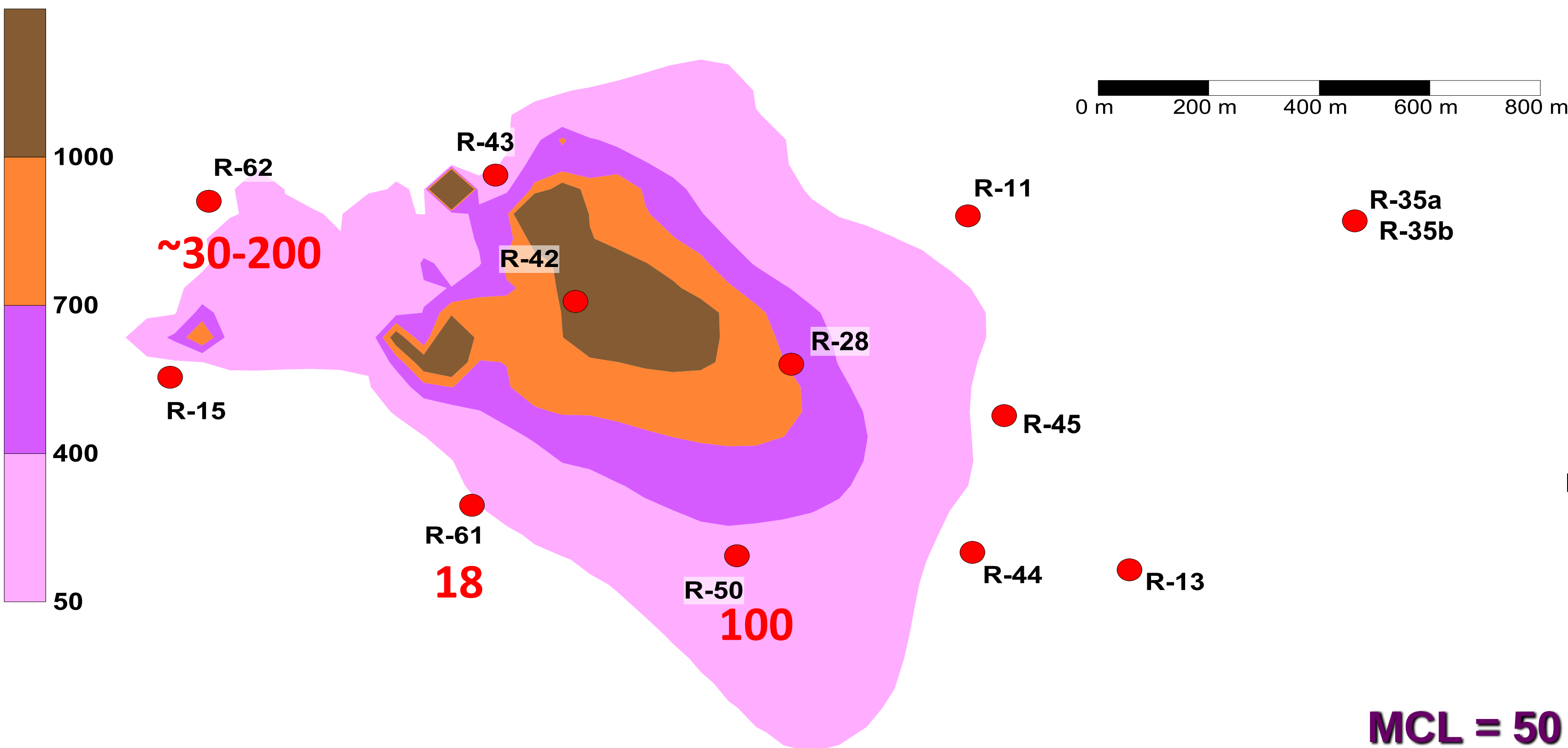
2009 model estimate of the plausible contaminant concentrations [ppb] along the regional aquifer water table

- Wells R-62, R-61 and R-50 were not drilled yet
- Locations of wells R-62, R-61 and R-50 were optimized based on model analyses
- Observed concentrations at R-62, R-61 and R-50 confirmed model predictions
- R-43 concentration were at background when the analyses were performed
- Since 2010, R-43 concentrations are increasing and approaching the model predicted concentration



2009 model estimate of the plausible contaminant concentrations [ppb] along the regional aquifer water table

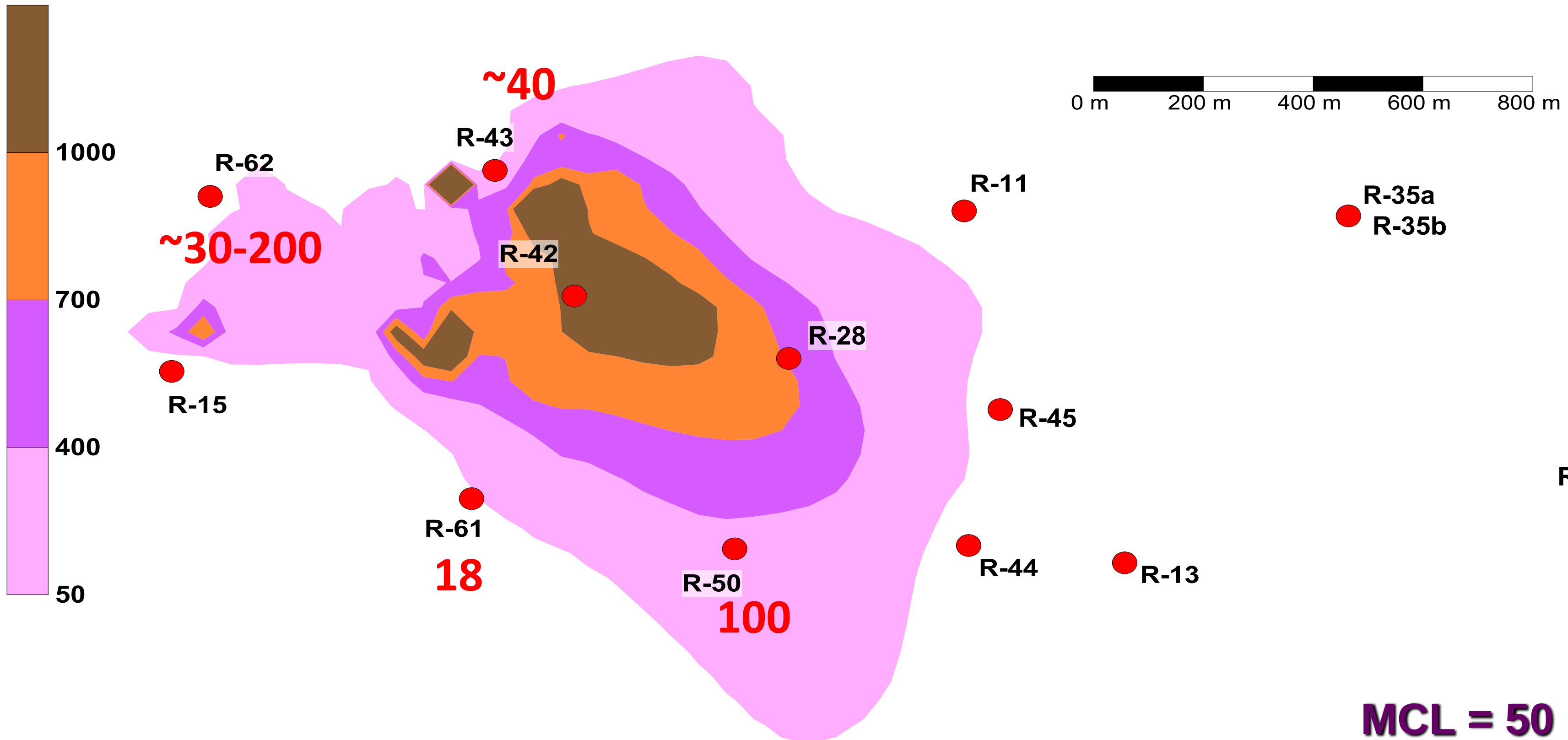
- Wells R-62, R-61 and R-50 were not drilled yet
- Locations of wells R-62, R-61 and R-50 were optimized based on model analyses
- Observed concentrations at R-62, R-61 and R-50 confirmed model predictions
- R-43 concentration were at background when the analyses were performed
- Since 2010, R-43 concentrations are increasing and approaching the model predicted concentration



MCL = 50 ppb

2009 model estimate of the plausible contaminant concentrations [ppb] along the regional aquifer water table

- Wells R-62, R-61 and R-50 were not drilled yet
- Locations of wells R-62, R-61 and R-50 were optimized based on model analyses
- Observed concentrations at R-62, R-61 and R-50 confirmed model predictions
- R-43 concentration were at background when the analyses were performed
- Since 2010, R-43 concentrations are increasing and approaching the model predicted concentration



MCL = 50 ppb

MADS

Model Analysis and Decision Support

Los Alamos NATIONAL LABORATORY

MADS is applied to perform all the presented info-gap decision analyses …

MADS
Model Analysis and Decision Support

Los Alamos NATIONAL LABORATORY

SerenDIPity

- an open-source high-performance computational framework for analyses and decision support based on complex process models
- advanced adaptive computational techniques:
  - sensitivity analysis (local / global);
  - uncertainty quantification (local / global);
  - optimization / calibration / parameter estimation (local / global);
  - model ranking & selection
  - decision support (GLUE, info-gap)
- novel algorithms
  - Agent-Based Adaptive Global Uncertainty and Sensitivity (ABAGUS)
    Harp & Vesselinov (2012) An agent-based approach to global uncertainty and sensitivity analysis. *Computers & Geosciences*.
  - Adaptive hybrid (local/global) optimization strategy (Squads)
    Vesselinov & Harp (2012) Adaptive hybrid optimization strategy for calibration and parameter estimation of physical process models. *Computers & Geosciences*.
- internal coupling with analytical contaminant transport solvers and test problems
- external coupling with existing process simulators (ModFlow, TOUGH, FEHM, eSTOMP, Amanzi, …)
- Source code, examples, performance comparisons, and tutorials @ http://mads.lanl.gov

ASCEM **Advanced Subsurface Computing for Environmental Management**

Regulatory

Public Interface

Reviews

Decision Making

Programmatic

Project Management

Oversight

Decision Making

Scientific

Model Setup and Execution

Model Analyses

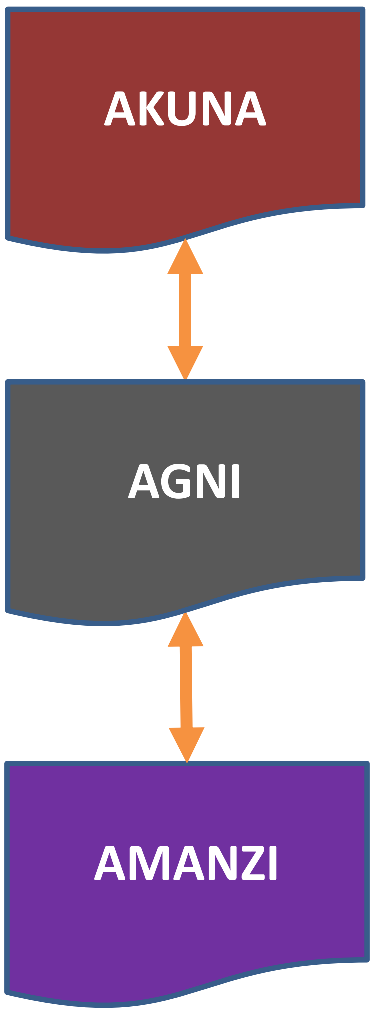
Decision Support

- an open-source interactive decision support system (Akuna/Agni) coupled a process simulator (Amanzi)
- high-performance computing (HPC)
- data- and model-driven decision support to provide standardized, consistent, site-specific and scientifically defensible decision analyses across DOE-EM complex
- **Challenge:**
  - develop tools to make better use of complex information and capabilities to explore problems in greater detail
  - address the most challenging performance assessment and waste-disposal problems
- **Impact:**
  - provide technical underpinnings for current U.S. DOE-EM risk and performance assessments
  - inform strategic data collection for model improvement and decision support
  - support scientifically defensible and standardized assessments and remedy selections

http://ascemdoe.org

BERKELEY LAB · Los Alamos NATIONAL LABORATORY EST.1943 · Pacific Northwest NATIONAL LABORATORY · OAK RIDGE National Laboratory · SRNL SAVANNAH RIVER NATIONAL LABORATORY

ASCEM Modules



Akuna ("no worries"): Graphic User Interface
(**Karen Schuchardt**, *PNNL*)

- Open Source Eclipse/Java based
- Incorporates data management, visualization, and model development tools

Agni ("fire"): Simulation controller and Toolset driver
(**George Pau**, *LBNL*, **Velimir Vesselinov**, *LANL*)

- Open Source C++ object oriented
- Provides coupling between **Akuna** and **Amanzi**
- Performs various model-based analyses (SA, UQ, PE, DS, …)

Amanzi ("water"): HPC Flow and Transport Simulator
(**David Moulton**, *LANL*)

- Open Source C++ object oriented
- Saturated / unsaturated groundwater flow, …
- Structured / unstructured / adaptive gridding
- …

http://ascemdoe.org

BERKELEY LAB · Los Alamos NATIONAL LABORATORY EST. 1943 · Pacific Northwest NATIONAL LABORATORY · OAK RIDGE National Laboratory · SRNL SAVANNAH RIVER NATIONAL LABORATORY

ASCEM Model-Analysis Toolsets in Agni

- Sensitivity Analysis (SA) (*Stefan Finsterle*, *Elizabeth Keating*)
- Parameter Estimation (PE) (*Stefan Finsterle*, *LBNL*)
- Uncertainty Quantification (UQ) (*Elizabeth Keating*, *LANL*)
- Risk Assessment (RA) (*Wilson McGinn*, *ORNL*)
- Decision Support (DS) (*Velimir Vesselinov*, *LANL*)

Conclusions and recommendations:

- Both Non-Probabilistic and Probabilistic uncertainties often exist in a decision problem
- Non-Probabilistic and Probabilistic methods should be applied to their appropriate uncertainties in the decision analyses
- In the case of probabilistic methods, definition of prior probability distributions for model parameters or calibration targets with unknown/uncertain distribution can produce biased predictions and decision analyses
- In the case of non-probabilistic methods, lack of knowledge and severe uncertainties can be captured
- Non-probabilistic methodologies have been successfully applied for a series of synthetic and real-world problems, though less often in hydrology
  - Remediation of unknown contaminant source
    Harp & Vesselinov (2011). *Contaminant remediation decision analysis using information gap theory*. SERRA
- MADS provides a computationally efficient framework for decision analyses using non-probabilistic and probabilistic methods (http://mads.lanl.gov)
- ASCEM tools are currently actively developed and will become available for testing and benchmarking in 2013 (http://ascemdoe.org)


SEREN
DIP
ITY
MADS
Monty Vesselinov
(C) 2011