Supporting Information Appendix

for article published in PNAS
P. Natl. Acad. Sci. USA, doi:10.1073/pnas.1011915108

# Arsenic pollution of groundwater in Vietnam exacerbated by deep aquifer exploitation for more than a century

Lenny H.E. Winkel[1,4], Pham Thi Kim Trang[2], Vi Mai Lan[2], Caroline Stengel[1], Manouchehr Amini[1], Nguyen Thi Ha[3], Pham Hung Viet[2], Michael Berg[1]*

[1] *Eawag, Swiss Federal Institute of Aquatic Science and Technology, Ueberlandstrasse 133, 8600 Dübendorf, Switzerland.*

[2] *Centre for Environmental Technology and Sustainable Development (CETASD), Hanoi University of Science, 334 Nguyen Trai Street, Hanoi, Vietnam.*

[3] *Vietnam Geological Survey, Northern Hydrogeological and Engineering Geological Division (NHEGD), Nghia Tan ward, Cau Giay district, Hanoi, Vietnam.*

[4] *present address: University of Grenoble, Institut des Sciences de la Terre (ISTerre), Rue de la Piscine, 38041 Grenoble, France.*

* To whom correspondence should be addressed: E-mail: Michael.Berg@eawag.ch

Data deposition: Data, hydrochemical maps, modeled risk maps, and movies reported in this paper were deposited on the web site of Eawag and can be downloaded from http://www.eawag.ch/arsenic-vietnam

## Table of Contents

## 1. Quality Assurance and Quality Control (QA/QC)

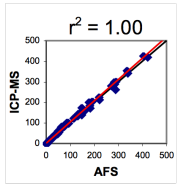
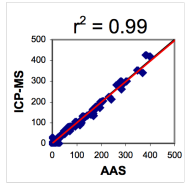
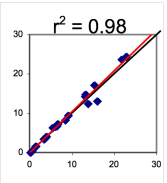
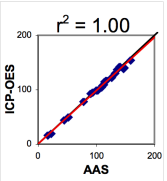
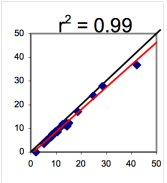
The robustness of the chemical analyses was assured by intermittent analysis of certified reference samples (SLRS-4 River Water Canada, TM-28.2 Lake Ontario, SPS-SW2 Surface Water Level 2 and reference samples from the international interlaboratory quality evaluations ARS13-16, ARS17-20, and ARS21-24 (*1*). In addition, cross-evaluation between different analytical techniques applied in our laboratories in Vietnam and Switzerland were carried out, e.g. AAS versus ICP-MS (see results in Table S1 below).

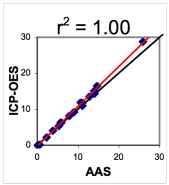
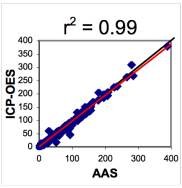
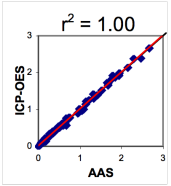
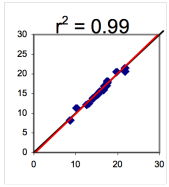
The results of certified samples and cross-checking agreed within ±5%. Calibration curves had $r^2$ >0.999 with the exception of Na and K where $r^2$ were 0.990 (ICP-OES). Standard deviations of triplicates were always <5%. The limits of quantification (LOQ, 10 x standard deviation of noise) were:

| | |
|---|---|
| 0.1 µg/L | for As, Cd, Co, Cr, Cu, Hg, Li, Ni, Pb, Sb, Se, U, and Zn |
| 0.5 µg/L | for Al |
| 1 µg/L | for B |
| 5 µg/L | for Ba |
| 0.01 mg/L | for Fe, Mn and Ammonium ($NH_4^+$-N) |
| 0.1 mg/L | for Ca, K, Mg, Na, Phosphate ($PO_4^{3-}$-P), and Sulphate ($SO_4^{2-}$) |
| 0.25 mg/L | for Nitrate ($NO_3^-$-N) |
| 0.5 mg/L | for Chloride ($Cl^-$) and Dissolved Organic Carbon (DOC) |
| 1 mg/L | for Br and I |
| 2 mg/L | for Si |
| 12 mg/L | for $HCO_3^-$ (0.2 mmol/L) |

**Table S1. Cross-correlations of selected parameters determined by various methods in our laboratories in Vietnam and Switzerland.**

(1) Analyses conducted by the Swiss Federal Institute of Aquatic Science and Technology (Eawag), Dübendorf, Switzerland.
(2) Analyses conducted by the Research Centre for Environmental Technology and Sustainable Development (CETASD), Hanoi University of Science, Vietnam National University.

| Parameter | Analytical methods | Samples | Cross-correlation |
|---|---|---|---|
| Arsenic (A) | ICP-MS(1) vs. AFS(1) | n = 216 |   |
| Arsenic (B) | ICP-MS(1) vs. AAS(2) | n = 461 |   |
| Ammonium | Photometry(1) vs. Photometry(2) | n = 21 |   |
| Calcium | ICP-OES(1) vs. AAS(2) | n = 21 |   |
| Chloride | IC(1) vs. IC(2) | n = 21 |   |

| | | | |
|---|---|---|---|
| Iron | ICP-OES[(1)] vs. AAS[(2)] | n = 21 |  |
| Magnesium | ICP-OES[(1)] vs. AAS[(2)] | n = 460 |  |
| Manganese | ICP-OES[(1)] vs. AAS[(2)] | n = 74 |  |
| Silicon | Photometry[(1)] vs. Photometry[(2)] | n = 21 |  |

## 2. Geology

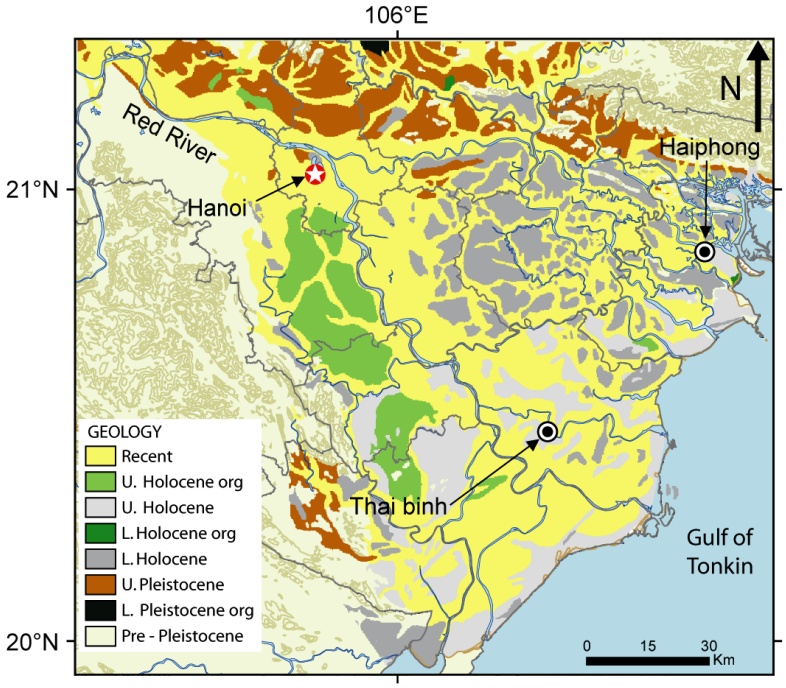



**Figure S1. Geological map of the Red River delta** indicating eight depositional environments. (Adapted from the geological map 1998, Northern Hydrogeological and Engineering Geological Division (NHEGD), Vietnam Geological Survey).

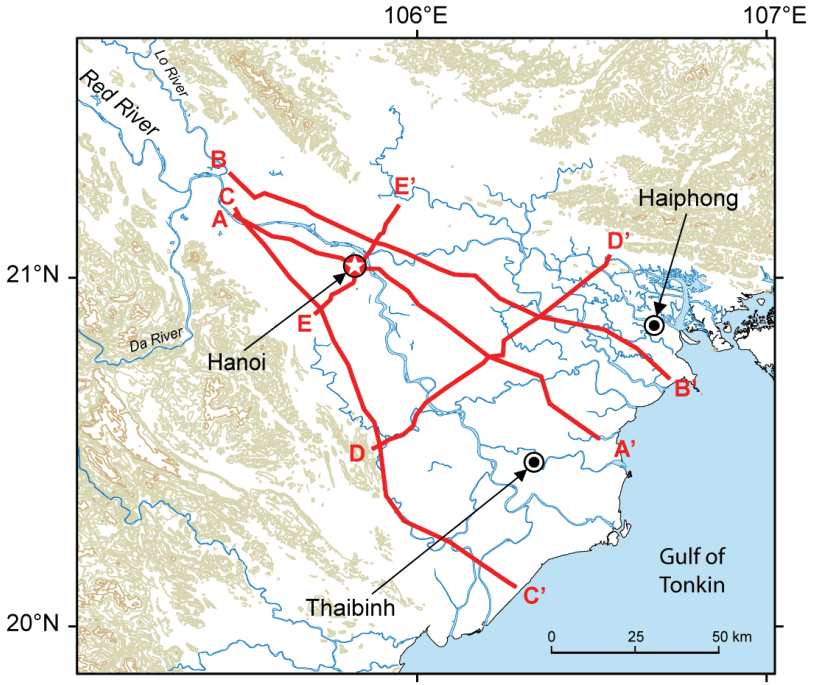



**Figure S2. Location of the geological cross-sections** A-A' to E-E'. The cross-sections are displayed in Figure S3.

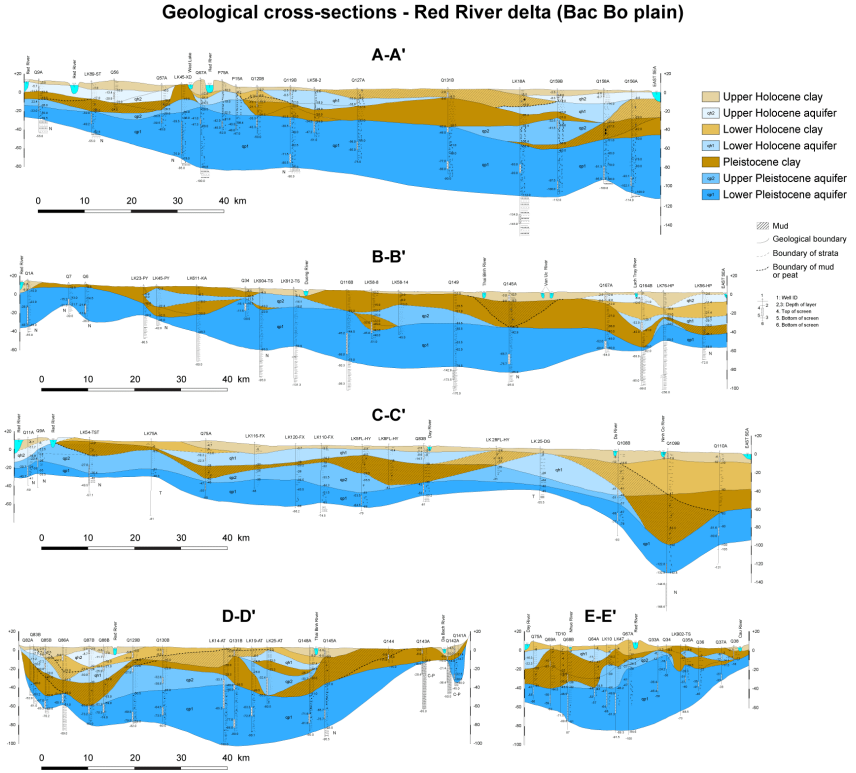



**Figure S3. Simplified geological cross-section along the transects A-A' to E-E' D–D'** indicated in Figure S2. At a regional scale, four different aquifers were formed in the Quaternary period: Lower Pleistocene (LP) aquifer (lower boundary 700,000 years BP), Upper Pleistocene (UP) aquifer (125,000 years BP), Lower Holocene (LH) aquifer (3000 years BP) and Upper Holocene (UH) aquifer (1000 BP). The three Quaternary aquitards are characterized by clay layers and occasionally intercalated peat lenses.

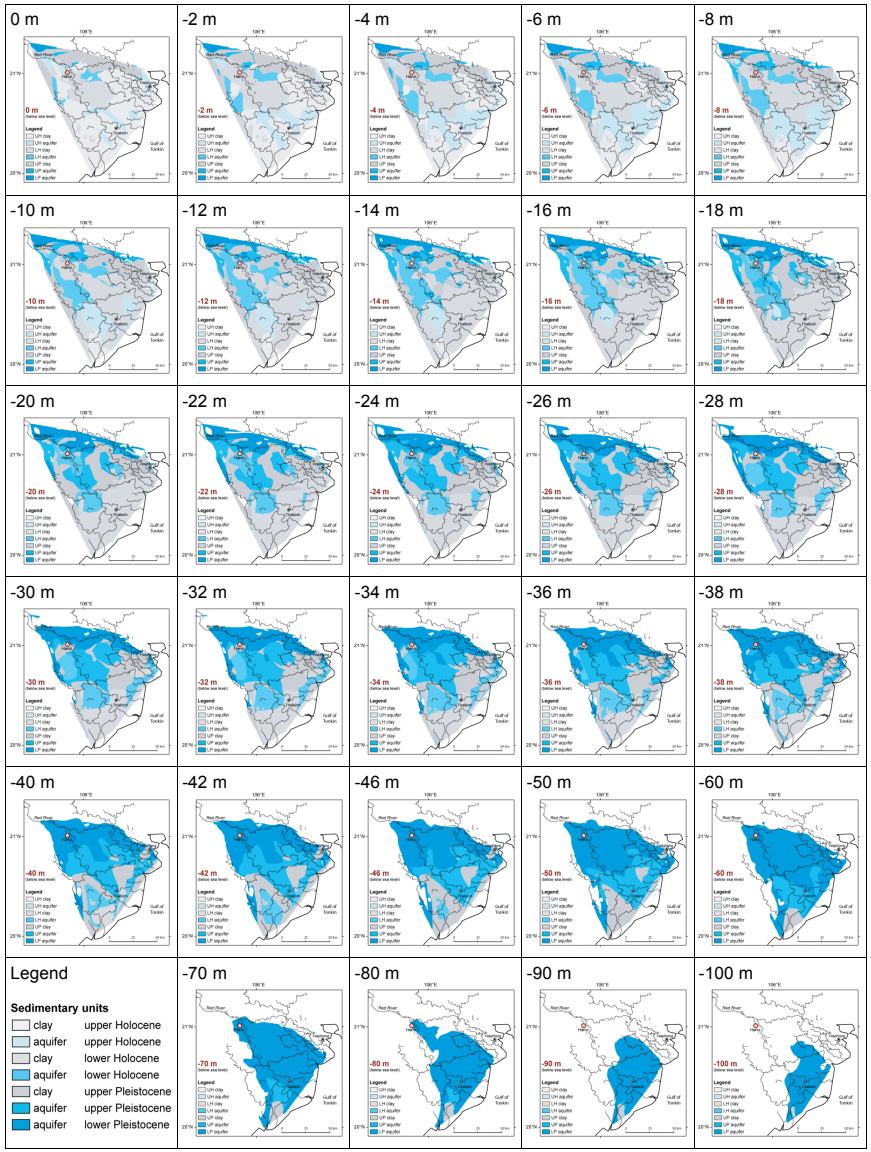



**Figure S4. Major aquifers of the** Red River delta **depicted at depth intervals of 2, 4 or 10 m.** Depths are given in meters below sea level. These maps are available online in high resolution as Supporting Information (SI) Movie S1. See http://dx.doi.org/10.1073/pnas.1011915108

# 3. Modeling Parameters and Model Description

## 3.1. Model based on three-dimensional (3D) geology

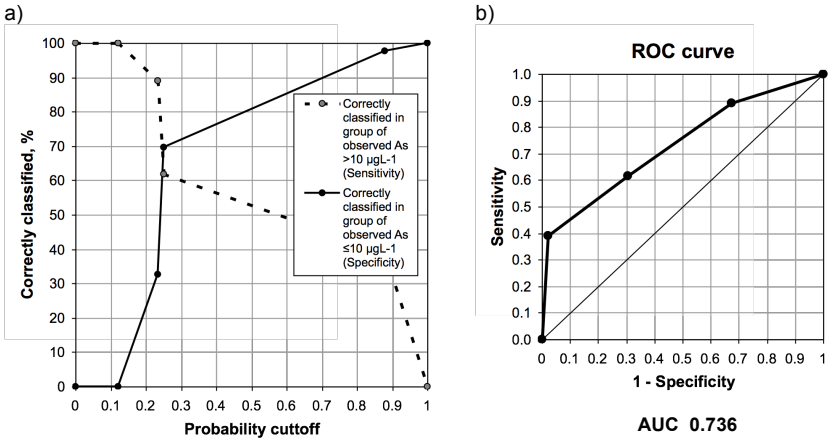



**Figure S5. a) Model classification results based on 3D geology of the Red River delta.** The graph shows the sensitivity (true positives) and specificity (true negatives) of the model for different probability cutoff values. The full classification table is provided in Table S2. **b) Receiver operation characteristics curve (ROC)** derived from the model classification table (Table S3). The area under the curve (AUC, also referred to as 'c' statistics) indicates the discriminative power of the logistic equation. It typically varies between 0.5 (random model) and 1.0 (entirely perfect model). The AUC value of this model is 0.736, which corresponds to 73.6% correctly classified cases of measured As concentrations.

**Table S2.** Hosmer-Lemeshow (2) goodness-of-fit table for the arsenic prediction model based on 3D geology.

| | | As ≤10 µgL$^{-1}$ | | As >10 µgL$^{-1}$ | | |
|---|---|---|---|---|---|---|
| group | probability | observed[a] | expected[a] | observed[b] | expected[b] | total |
| 1 | 0.18 | 84 | 83.7 | 11 | 11.26 | 95 |
| 2 | 0.23 | 3 | 3.3 | 1 | 0.74 | 4 |
| 3 | 0.25 | 92 | 89.3 | 22 | 24.72 | 114 |
| 4 | 0.26 | 63 | 66.1 | 23 | 19.89 | 86 |
| 5 | 0.28 | 19 | 18.6 | 10 | 10.40 | 29 |
| 6 | 0.34 | 6 | 6.0 | 43 | 43.00 | 49 |

[a] Average deviation of observed and expected As ≤10 µgL$^{-1}$ : 3.14% (absolute)
[b] Average deviation of observed and expected As >10 µgL$^{-1}$ : 9.66% (absolute)

**Table S3.** Classification table of model performance for different cutoff values of predicted probabilities (*2*) for the arsenic prediction model based on 3D geology.

| | As ≤10 ugL$^{-1}$ | | | As >10 ugL$^{-1}$ | | | correctly classified in group | |
|---|---|---|---|---|---|---|---|---|
| probability cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | As obs ≤10 "specificity" true negatives | As obs >10 "sensitivity" true positives |
| 0 | 267 | 0 | 267 | 110 | 0 | 110 | 0 | 100 |
| 0.121 | 267 | 0 | 267 | 110 | 0 | 110 | 0 | 100 |
| 0.233 | 267 | 87 | 180 | 110 | 12 | 98 | 32.6 | 89.1 |
| 0.250 | 267 | 186 | 81 | 110 | 42 | 68 | 69.7 | 61.8 |
| 0.878 | 267 | 261 | 6 | 110 | 67 | 43 | 97.8 | 39.1 |
| 1 | 267 | 267 | 0 | 110 | 110 | 0 | 100 | 0 |

## 3.2. Model based on surface parameters

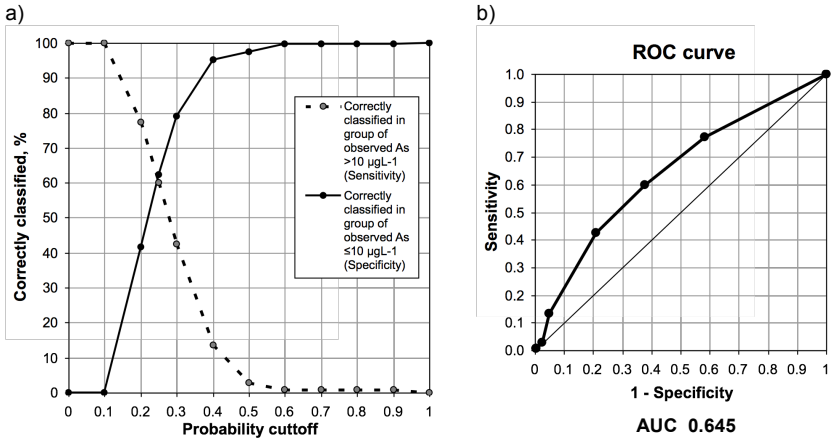



**Figure S6. a) Model classification results based on surface parameters of the Red River delta.** The graph shows the sensitivity (true positives) and specificity (true negatives) of the model for different probability cutoff values. The full classification table is provided in Table S4. **b) Receiver operation characteristics curve (ROC)** derived from the model classification table (Table S5). The area under the curve (AUC, also referred to as 'c' statistics) indicates the discriminative power of the logistic equation. It typically varies between 0.5 and 1.0. The AUC value of this model is 0.645, which corresponds to 64.5% correctly classified cases of measured As concentrations.

**Table S4.** Hosmer-Lemeshow (*2*) goodness-of-fit table for the arsenic prediction model based on surface parameters.

| | | As ≤10 µgL$^{-1}$ | | As >10 µgL$^{-1}$ | | |
|---|---|---|---|---|---|---|
| group | probability | observed[a] | expected[a] | observed[b] | expected[b] | total |
| 1 | 0.17 | 155 | 154.3 | 32 | 32.7 | 187 |
| 2 | 0.17 | 18 | 18.7 | 7 | 6.3 | 25 |
| 3 | 0.17 | 45 | 44.2 | 16 | 16.8 | 61 |
| 4 | 0.25 | 75 | 73.0 | 26 | 28.0 | 101 |
| 5 | 0.28 | 21 | 21.8 | 14 | 13.2 | 35 |
| 6 | 0.28 | 39 | 41.0 | 27 | 25.0 | 66 |
| 7 | 0.28 | 17 | 17.0 | 18 | 18.0 | 35 |

[a] Average deviation of observed and expected As ≤10 µgL$^{-1}$ : 2.55% (absolute)
[b] Average deviation of observed and expected As >10 µgL$^{-1}$ : 5.47% (absolute)

**Table S5.** Classification table of model performance for different cutoff values of predicted probabilities (*2*) for the arsenic prediction model based on surface parameters.

| | As ≤10 ugL$^{-1}$ | | | As >10 ugL$^{-1}$ | | | correctly classified in group | |
|---|---|---|---|---|---|---|---|---|
| probability cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | As obs ≤10 "specificity" true negatives | As obs >10 "sensitivity" true positives |
| 0 | 371 | 0 | 371 | 141 | 0 | 141 | 0 | 100 |
| 0.1 | 371 | 0 | 371 | 141 | 0 | 141 | 0 | 100 |
| 0.2 | 371 | 155 | 216 | 141 | 32 | 109 | 41.8 | 77.3 |
| 0.25 | 371 | 223 | 148 | 141 | 53 | 88 | 62.4 | 60.1 |
| 0.3 | 371 | 293 | 78 | 141 | 81 | 60 | 79.0 | 42.6 |
| 0.4 | 371 | 353 | 18 | 141 | 122 | 19 | 95.1 | 13.5 |
| 0.5 | 371 | 362 | 9 | 141 | 137 | 4 | 97.6 | 2.8 |
| 0.6 | 371 | 370 | 1 | 141 | 140 | 1 | 99.7 | 0.7 |
| 0.7 | 371 | 370 | 1 | 141 | 140 | 1 | 99.7 | 0.7 |
| 0.8 | 371 | 370 | 1 | 141 | 140 | 1 | 99.7 | 0.7 |
| 0.9 | 371 | 370 | 1 | 141 | 140 | 1 | 99.7 | 0.7 |
| 1 | 371 | 371 | 0 | 141 | 141 | 0 | 100 | 0 |

### 3.3. Model for Hanoi region based on 3D geology

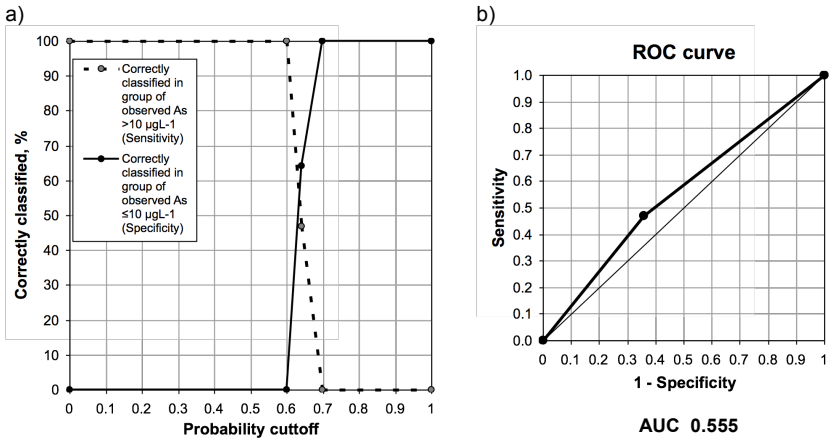



**Figure S7. a) Model classification results for the Hanoi region based on 3D geology at depth.** The graph shows the sensitivity (true positives) and specificity (true negatives) of the model for different probability cutoff values. The full classification table is provided in Table S6. **b) Receiver operation characteristics curve (ROC)** derived from the model classification table (Table S7). The area under the curve (AUC, also referred to as 'c' statistics) indicates the discriminative power of the logistic equation. It typically varies between 0.5 and 1.0. The AUC value of our model is 0.555, which corresponds to 55.5% correctly classified cases of measured As concentrations.

**Table S6.** Hosmer-Lemeshow (2) goodness-of-fit table for the arsenic prediction model of the Hanoi region based on 3D geology.

| | | As ≤10 µgL$^{-1}$ | | As >10 µgL$^{-1}$ | | |
|---|---|---|---|---|---|---|
| group | probability | observed | expected | observed | expected | total |
| 1 | | 57 | 57.0 | 43 | 43.0 | 100 |
| 2 | | 92 | 92.0 | 115 | 115.0 | 207 |
| 3 | | 35 | 35.0 | 111 | 111.0 | 146 |

**Table S7.** Classification table of model performance for different cutoff values of predicted probabilities (2) for the arsenic prediction model based on 3D geology.

| | As ≤10 ugL$^{-1}$ | | | As >10 ugL$^{-1}$ | | | correctly classified in group | |
|---|---|---|---|---|---|---|---|---|
| probability cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | observed cases | predicted cases ≤cutoff | predicted cases >cutoff | As obs ≤10 "specificity" true negatives | As obs >10 "sensitivity" true positives |
| 0 | 184 | 0 | 184 | 269 | 0 | 269 | 100 | 100 |
| 0.6 | 184 | 0 | 184 | 269 | 0 | 269 | 100 | 100 |
| 0.641 | 184 | 118 | 66 | 269 | 143 | 126 | 46.8 | 46.8 |
| 0.7 | 184 | 184 | 0 | 269 | 269 | 0 | 0 | 0 |
| 1 | 184 | 184 | 0 | 269 | 269 | 0 | 0 | 0 |

## 4. Model Output at Depth

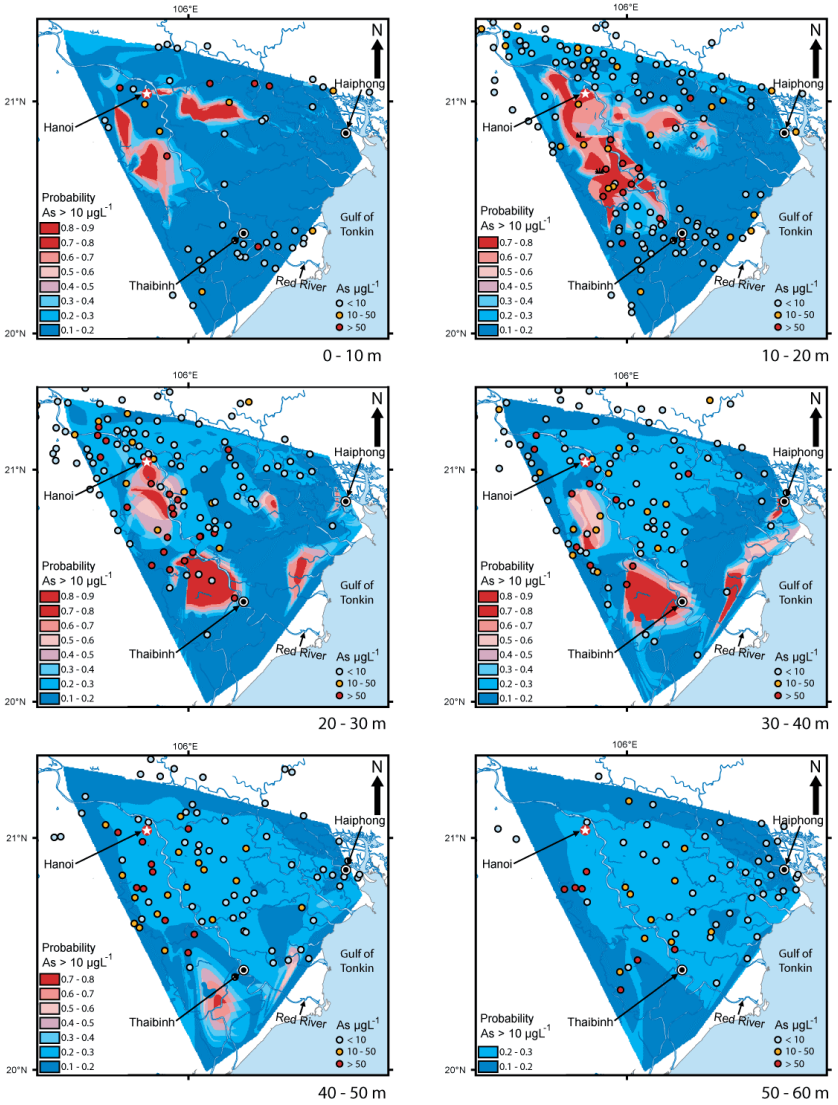



**Figure S8.** Average probability of As concentrations exceeding 10 µg L$^{-1}$ at 10 m depth intervals and measured arsenic concentrations at the same depths.

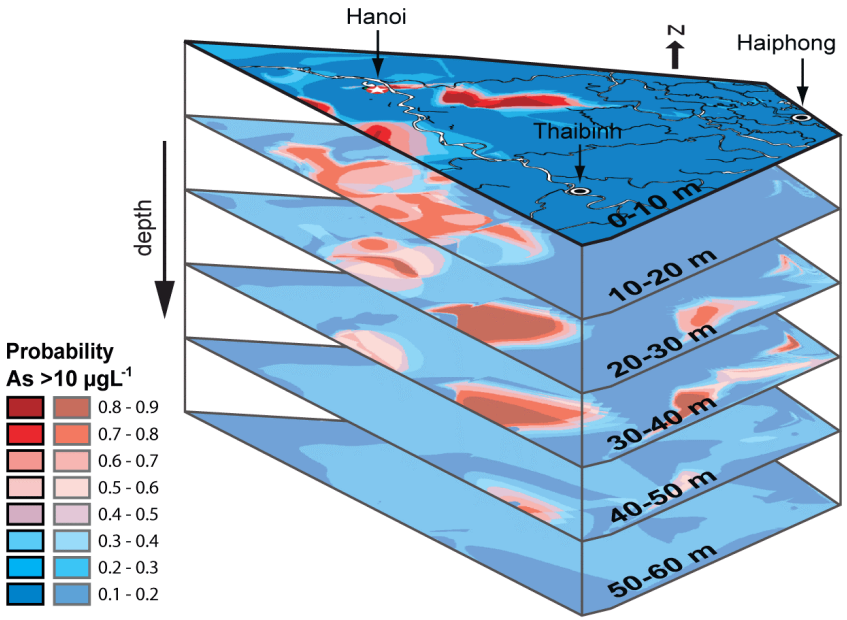



**Figure S9.** 3D distribution of As exceeding 10 µgL$^{-1}$ in the Red River delta, stacked in 10 m depth intervals.

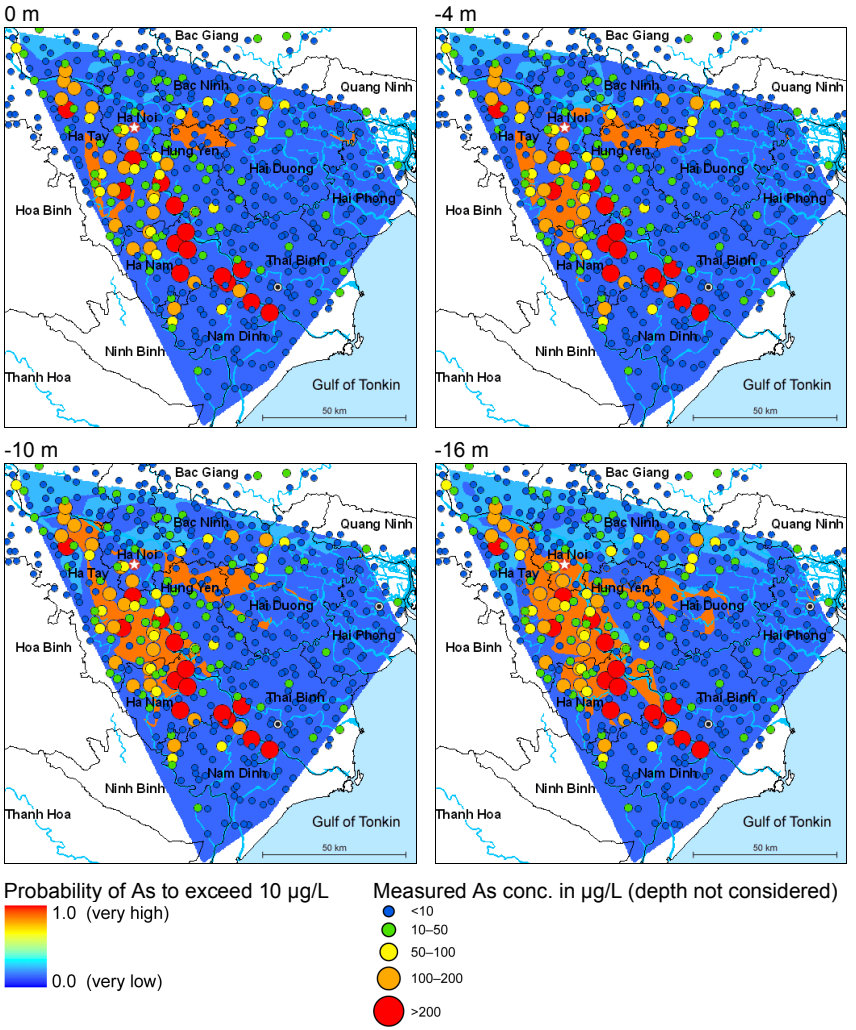



**Figure S10. Probability of As contamination at given depths.** Depths are given in meters below sea level. High resolution maps in 2-meter depth intervals are as Supporting Information (SI) Movie S2. See http://dx.doi.org/10.1073/pnas.1011915108

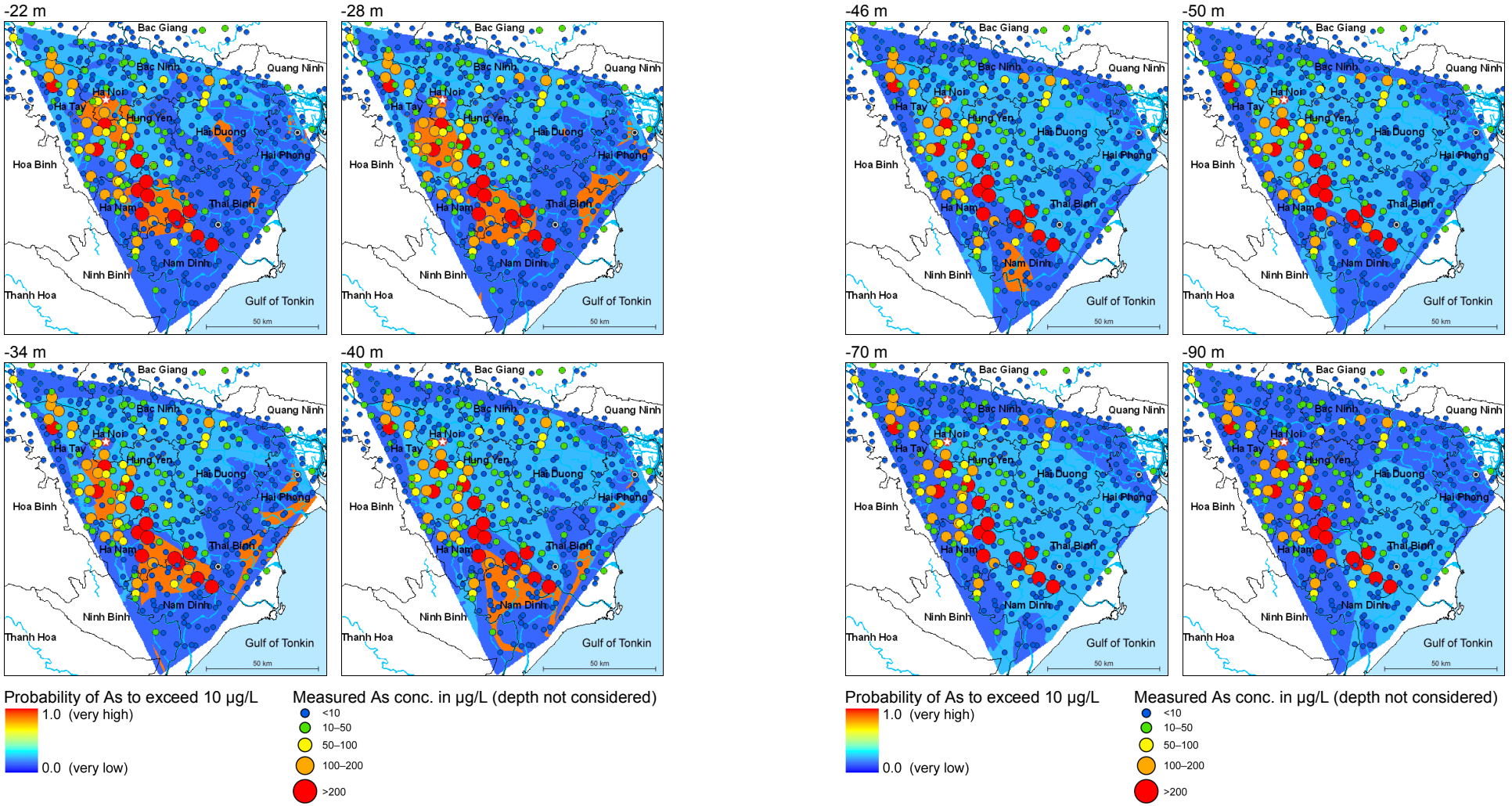



**Figure S10 (cont.). Probability of As contamination at various depths.** Depths are given in meters below sea level. High resolution maps in 2-meter depth intervals are as Supporting Information (SI) Movie S2. See http://dx.doi.org/10.1073/pnas.1011915108

**Figure S10 (cont.). Probability of As contamination at various depths.** Depths are given in meters below sea level. High resolution maps in 2-meter depth intervals are as Supporting Information (SI) Movie S2. See http://dx.doi.org/10.1073/pnas.1011915108

## 5. Hydrochemical Atlas

The complete geo-referenced database is provided online as Supporting Information (SI) Dataset S1. See http://dx.doi.org/10.1073/pnas.1011915108

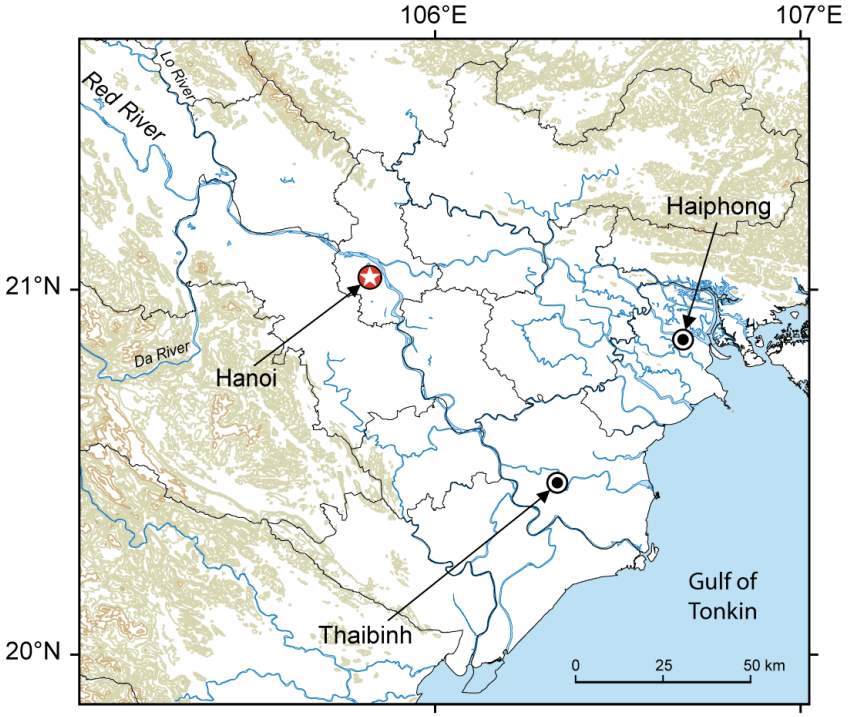
### Contents



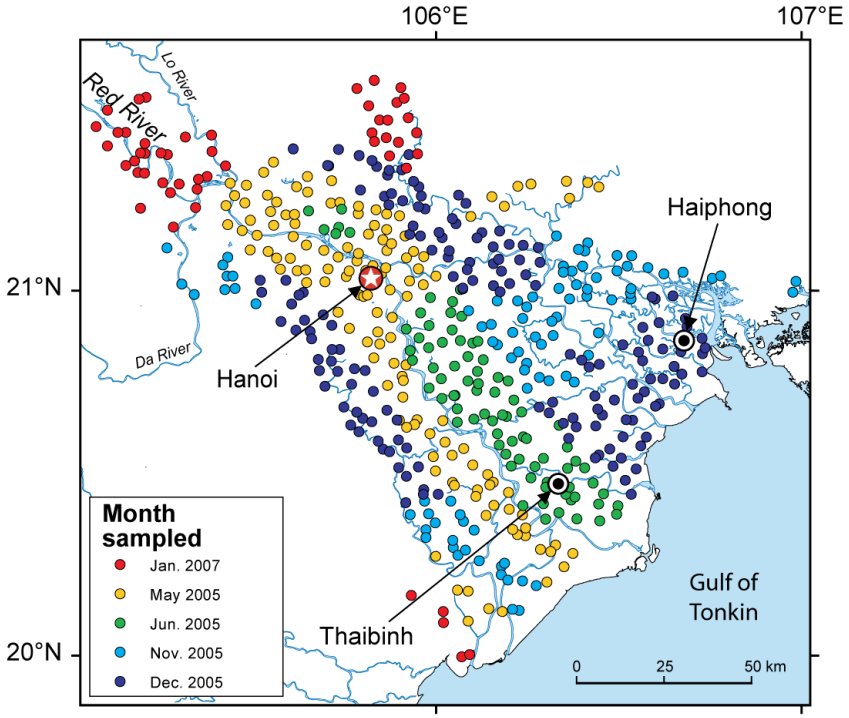

Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam


correspondence: Michael.Berg@eawag.ch

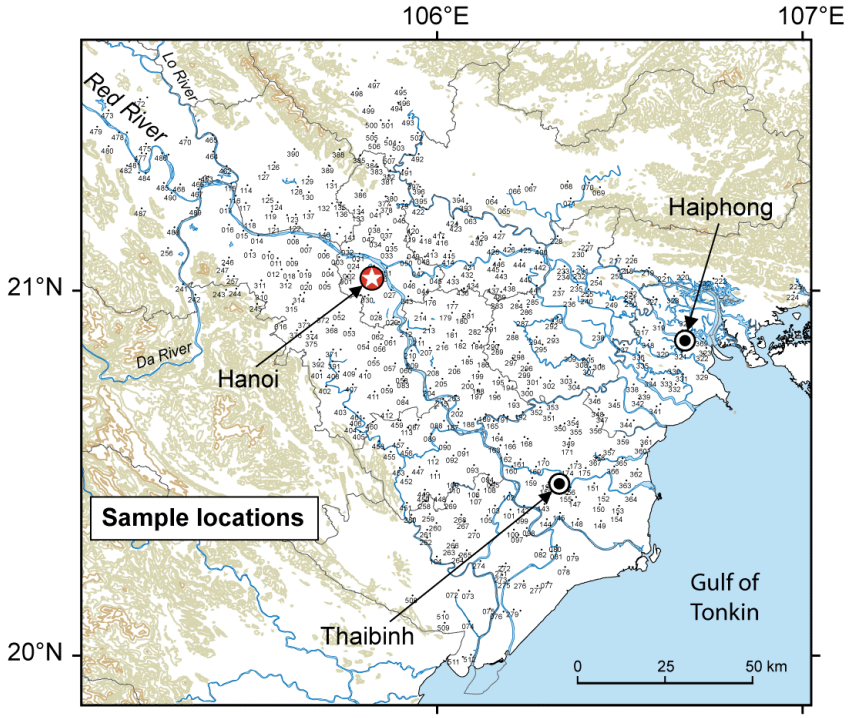

## 1 Topography, rivers, and province boundaries



## 2 Sampling date

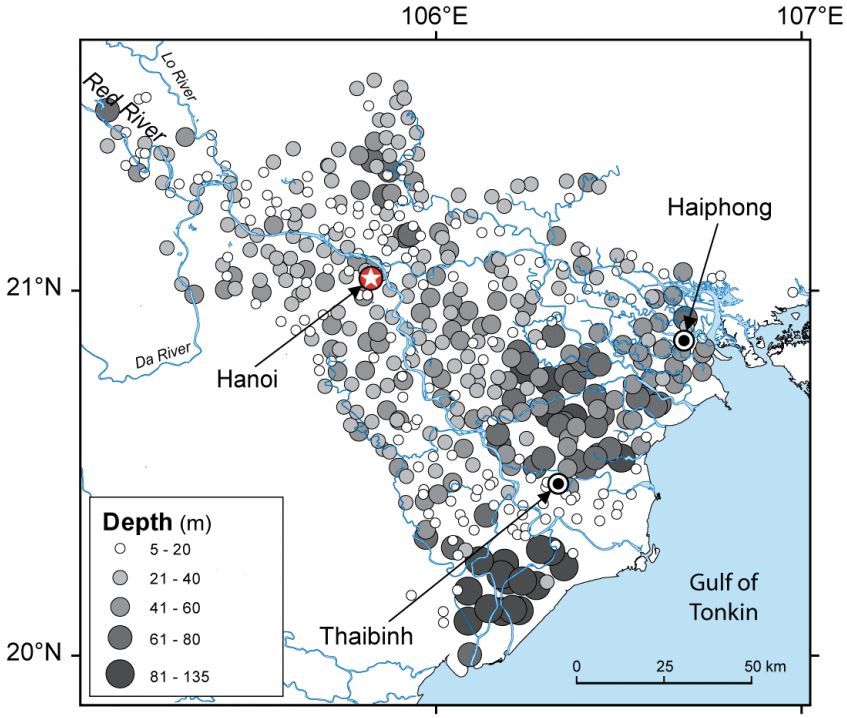



## 3 Sample locations and ID

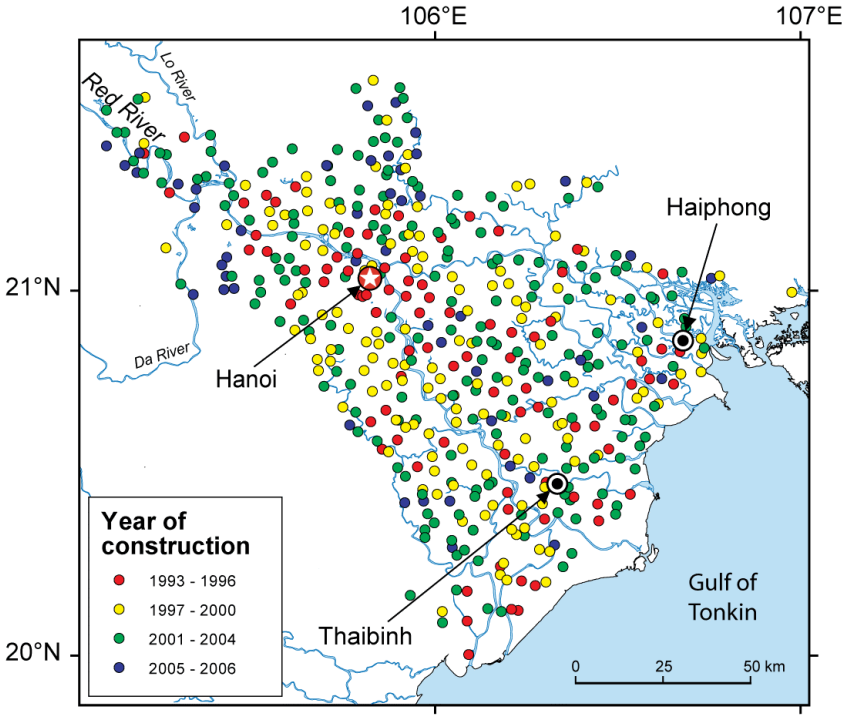



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 4 Well depth

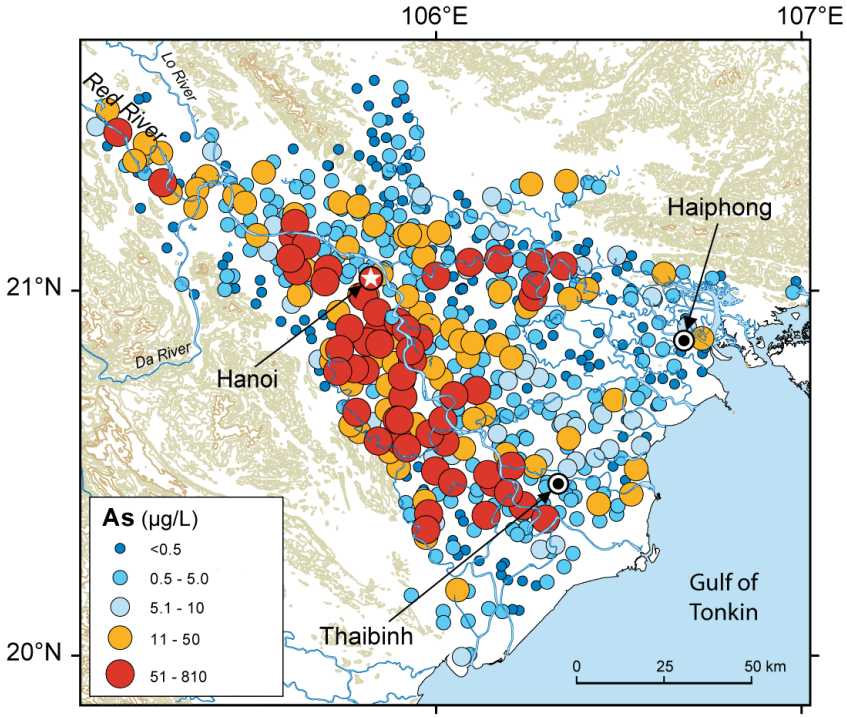



## 5 Age of wells

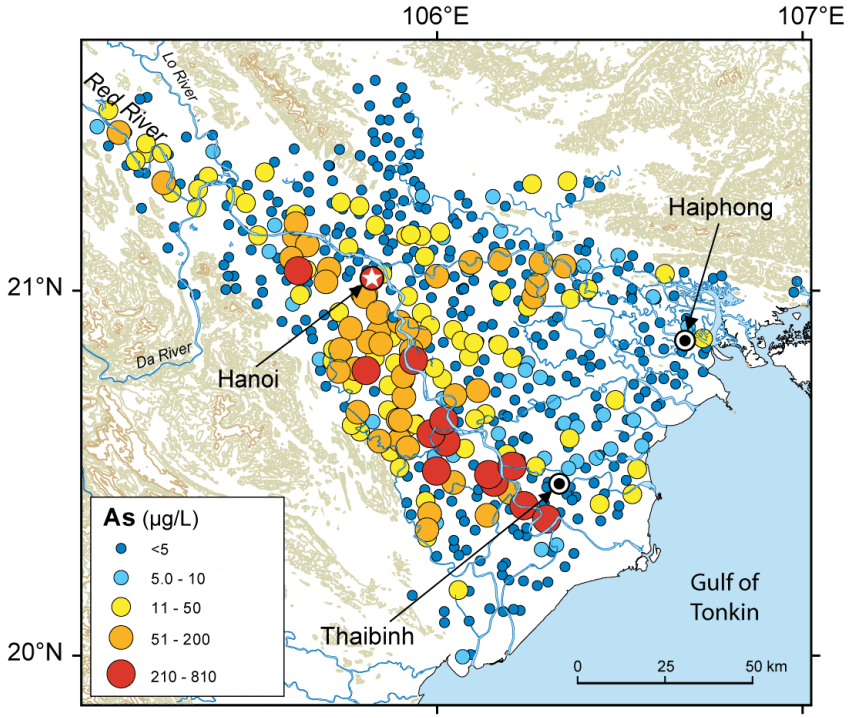



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

**6a Arsenic** (<0.5 to >50)

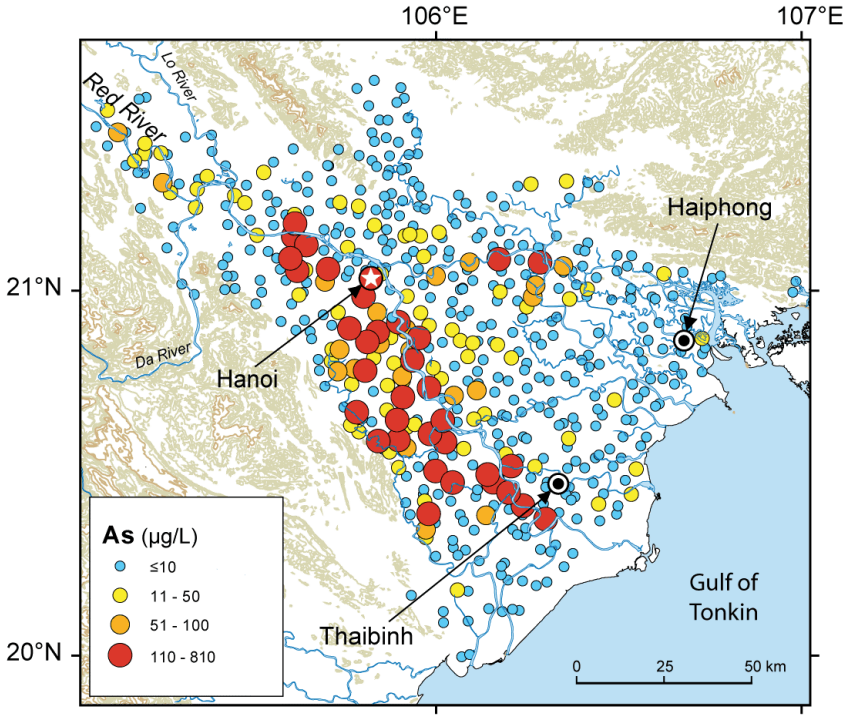



**6b Arsenic** (<5 to >200)

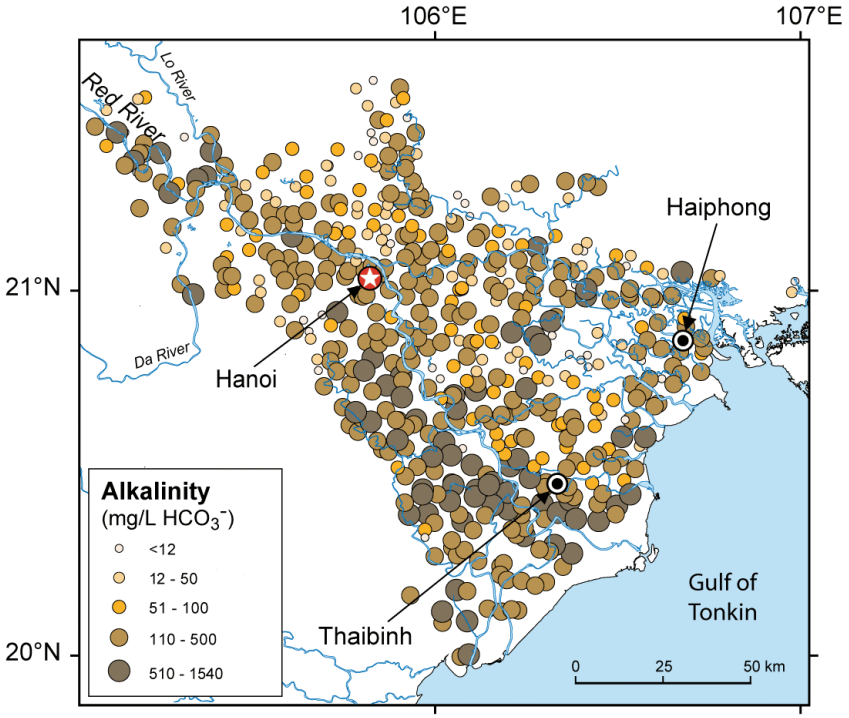



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 6c **Arsenic** (<10 to >100)

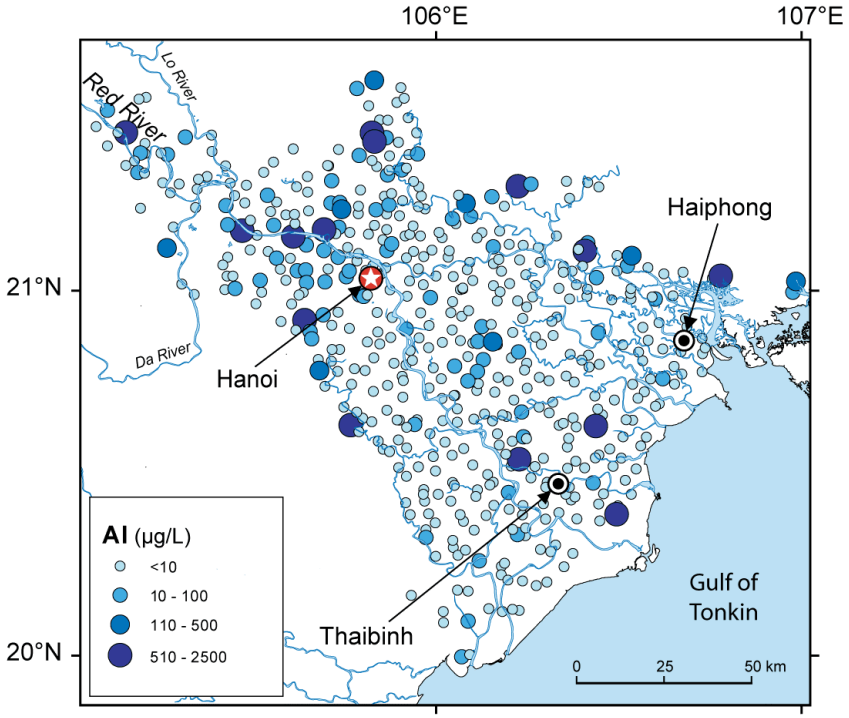



## 7 **Alkalinity**

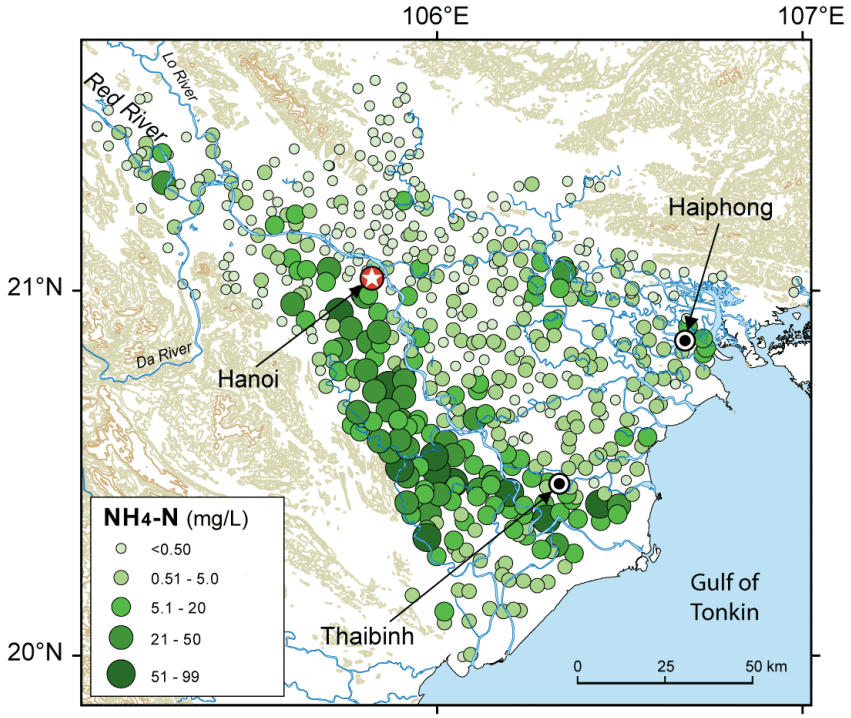



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 8 Aluminium

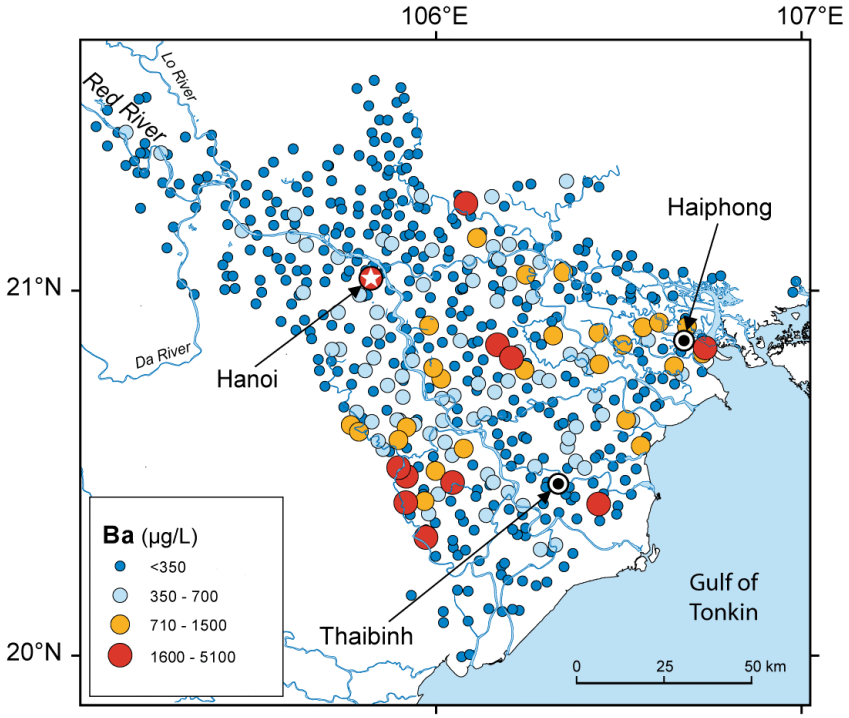



## 9 Ammonium

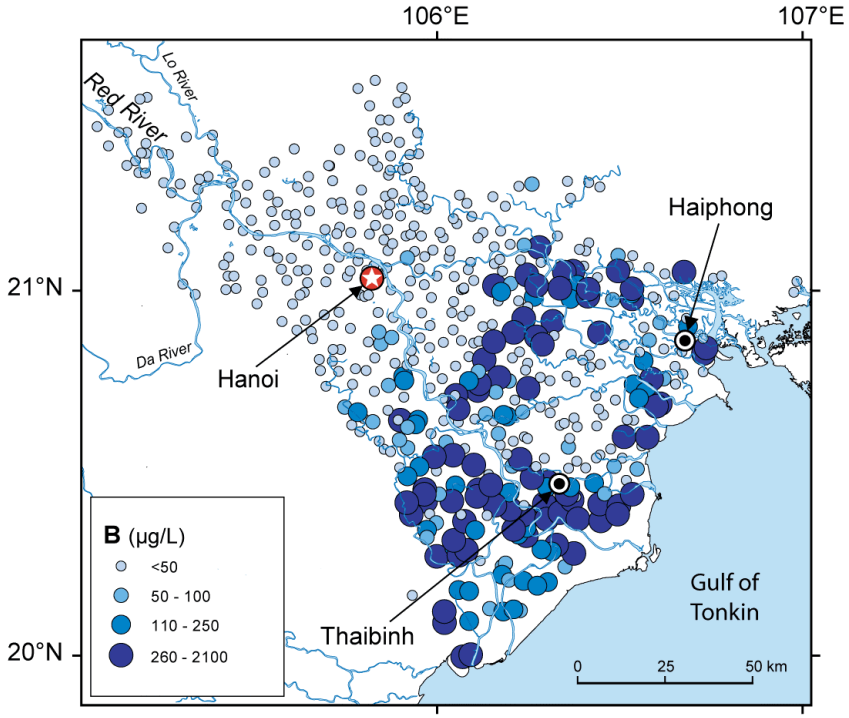



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 10 Barium

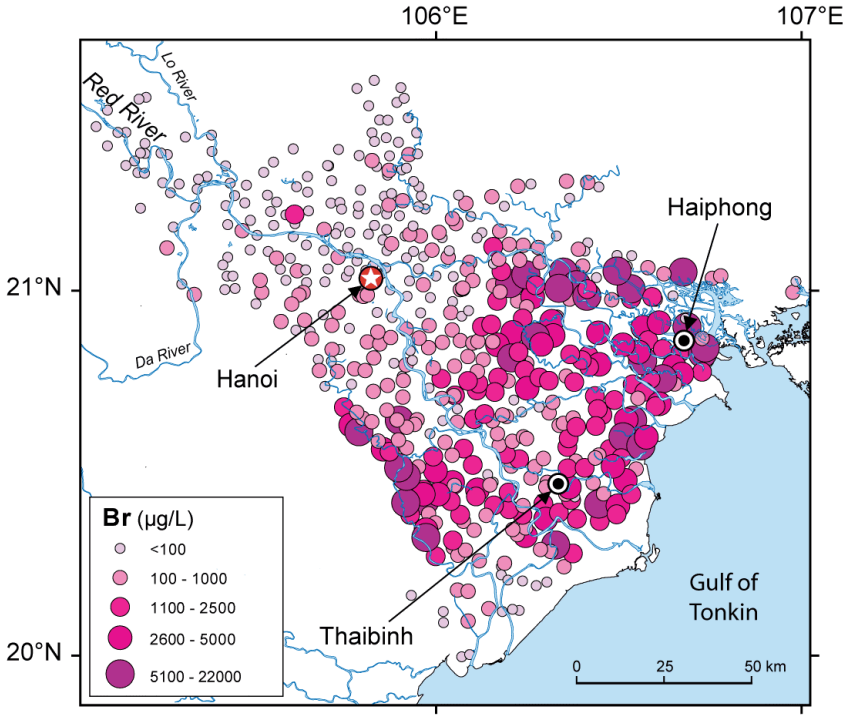



## 11 Boron

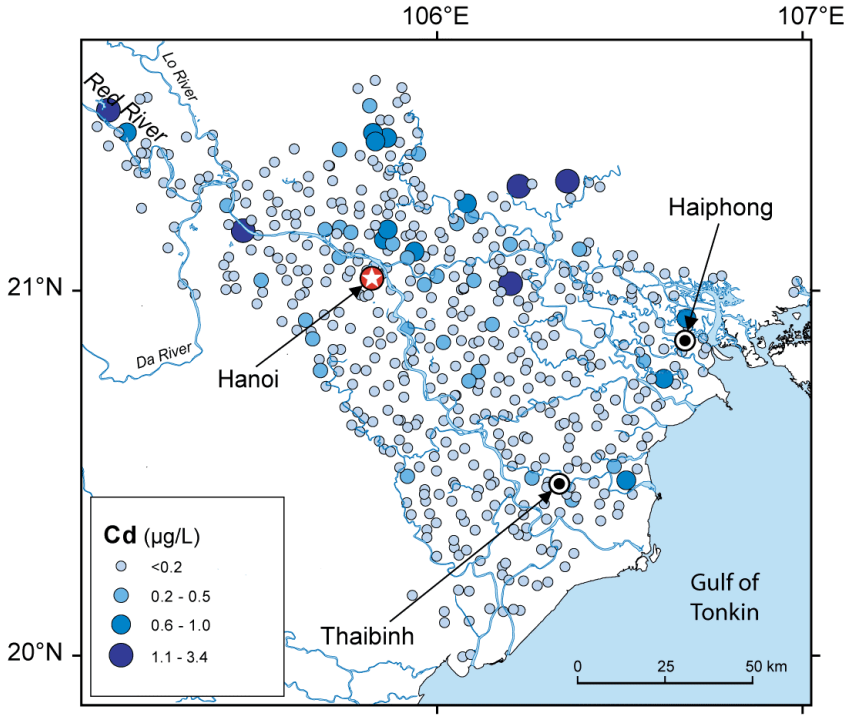



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 12 Bromine

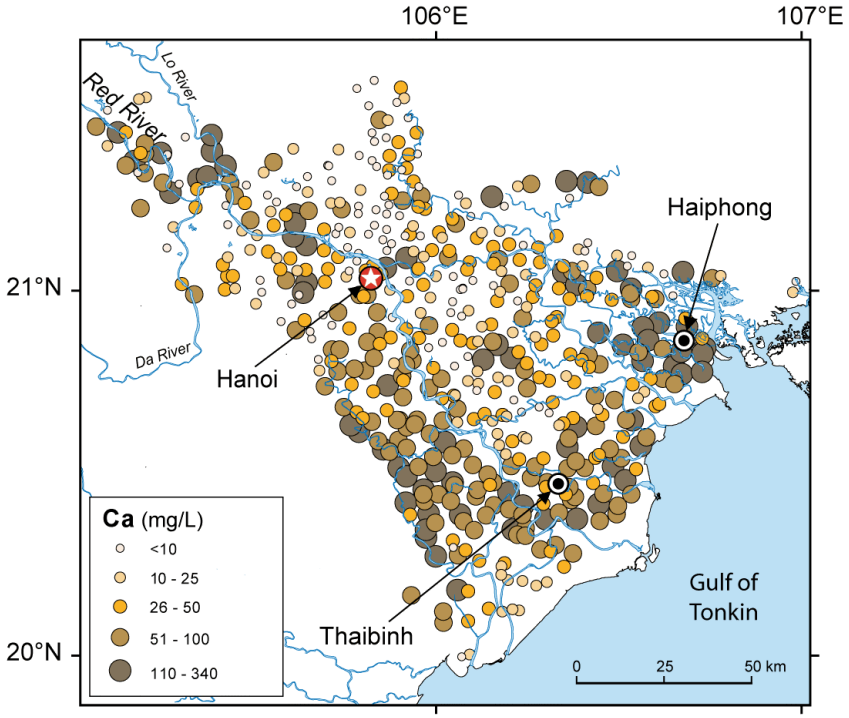



## 13 Cadmium

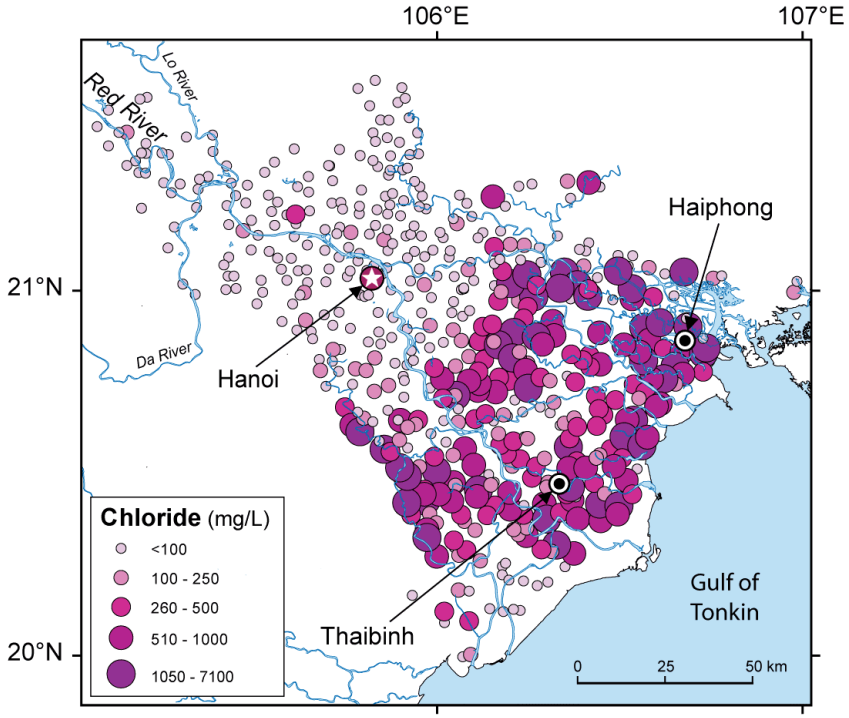



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 14 Calcium

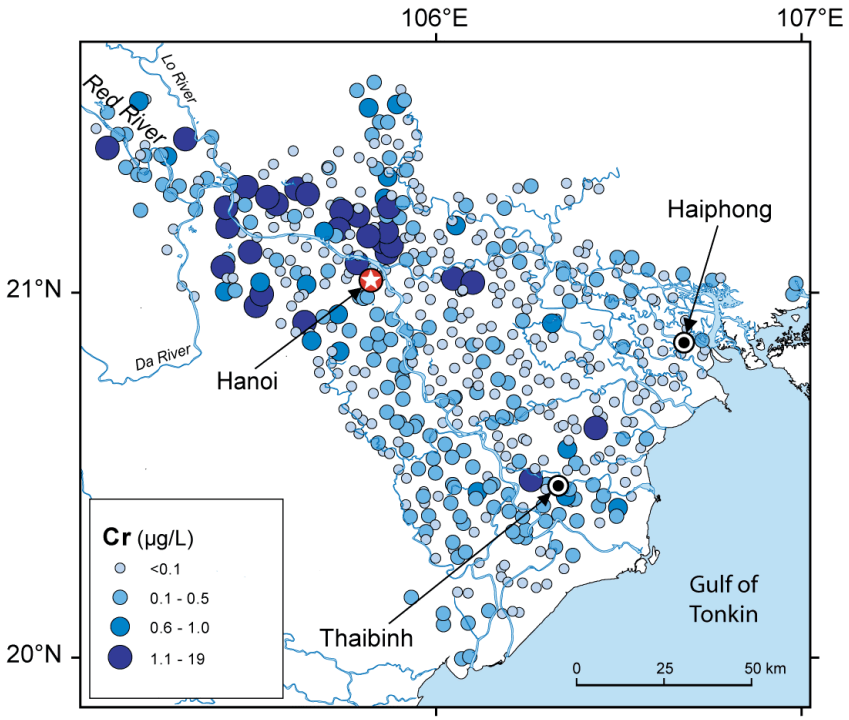



## 15 Chloride

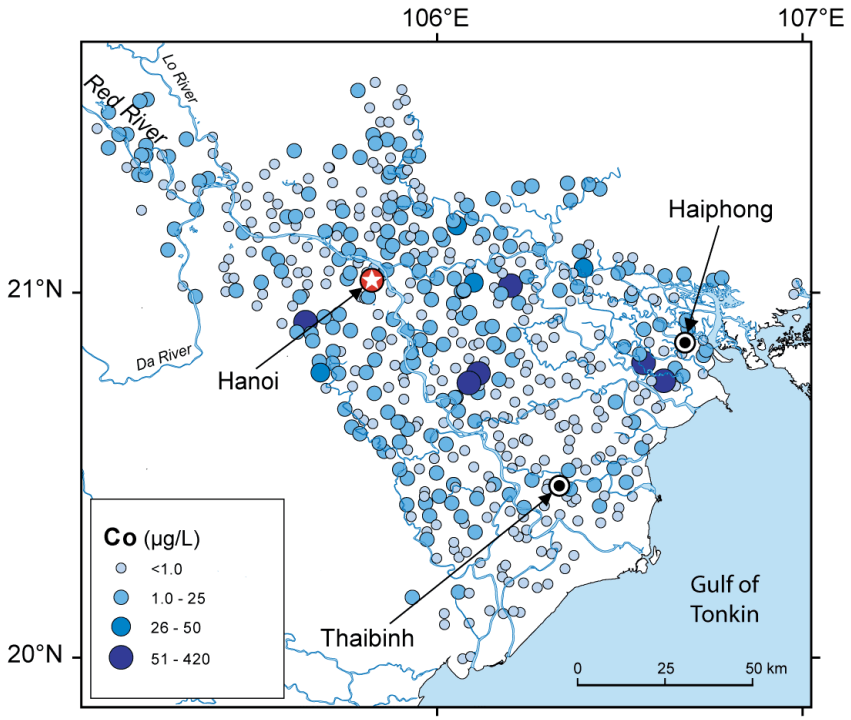



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 16 Chromium

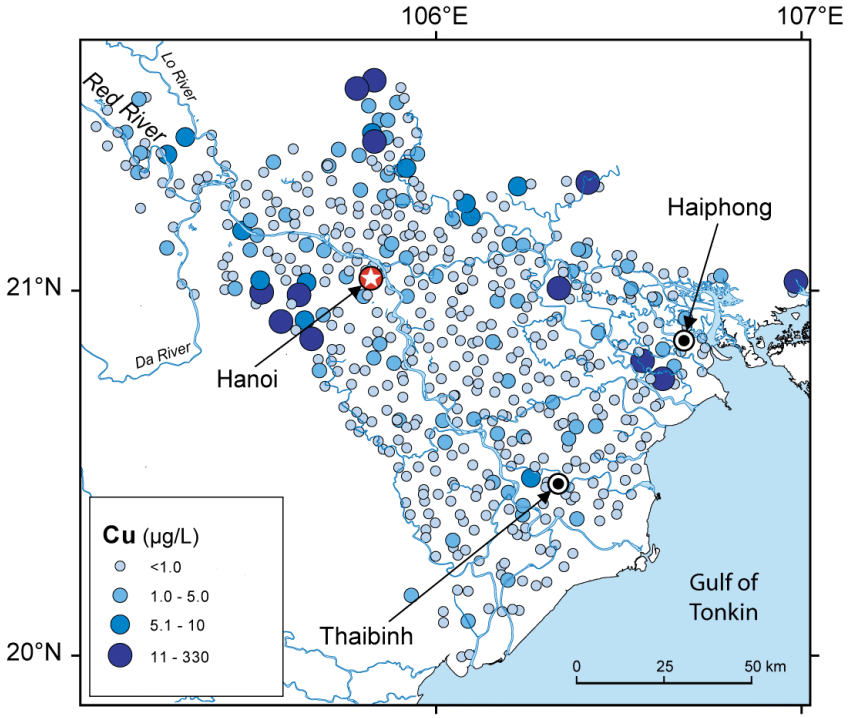



## 17 Cobalt

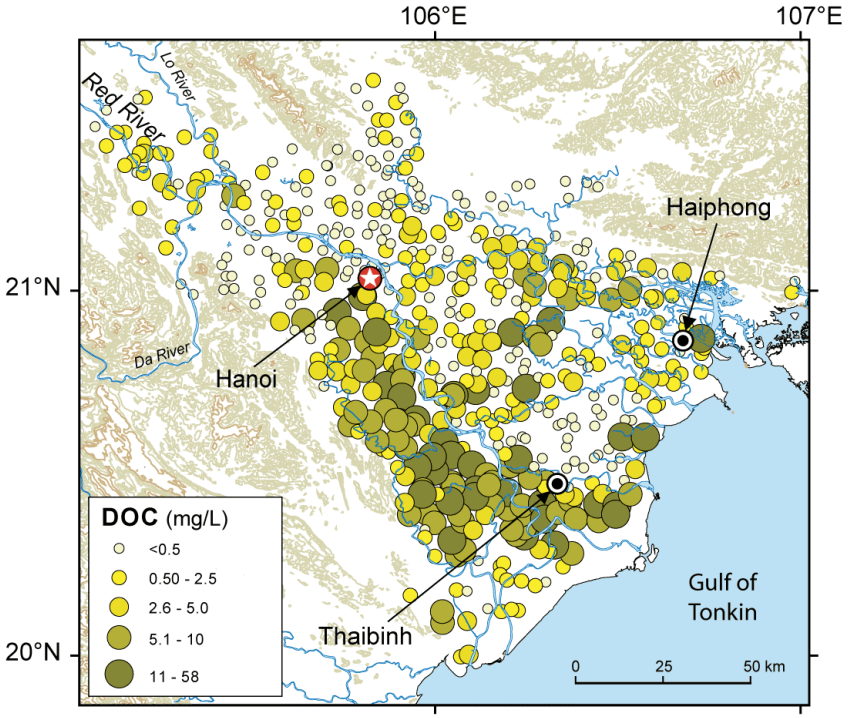



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 18 Copper

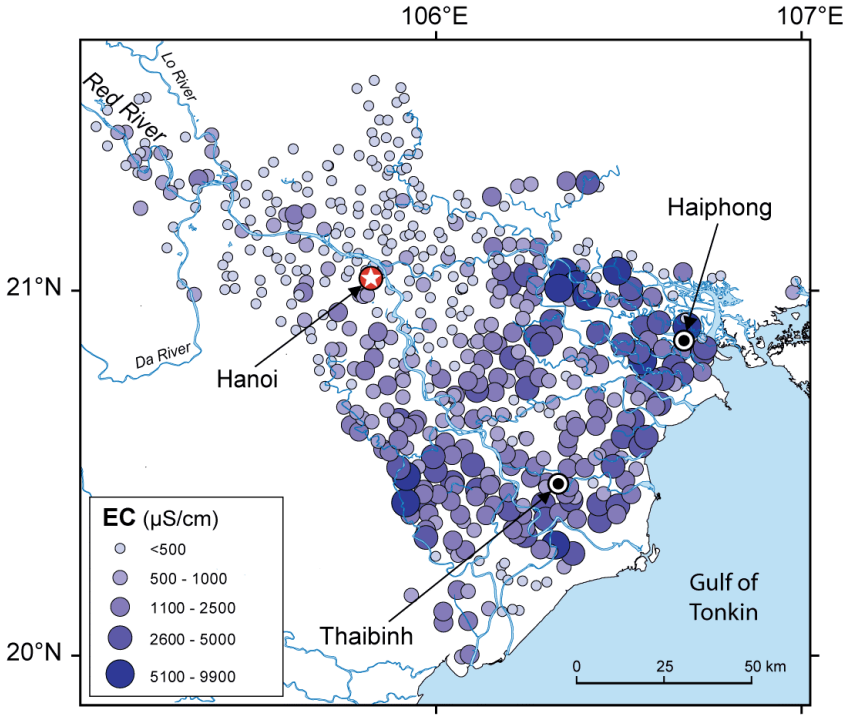



## 19 Dissolved organic carbon (DOC)

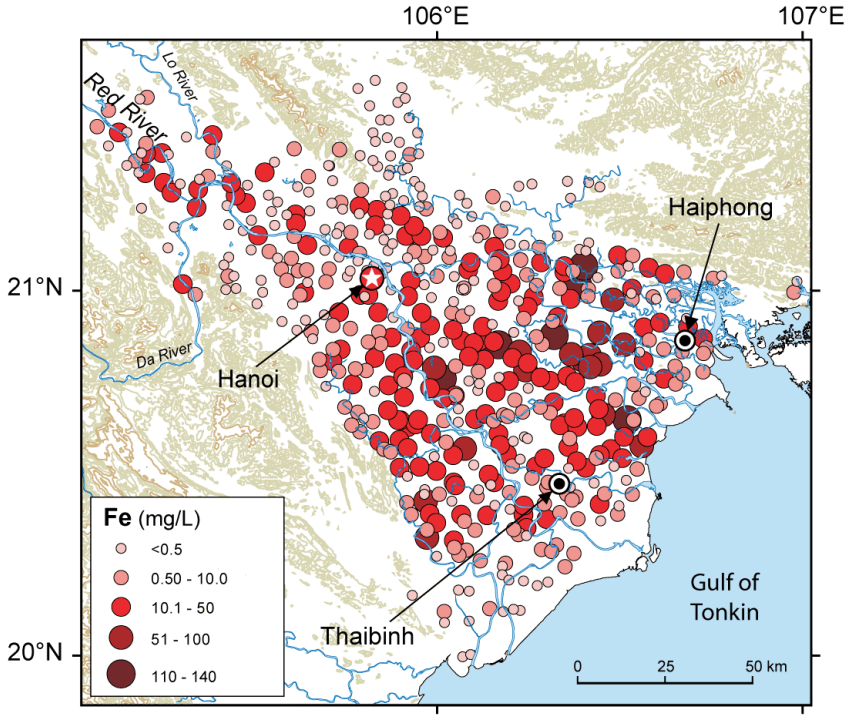



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 20 Electrical conductivity

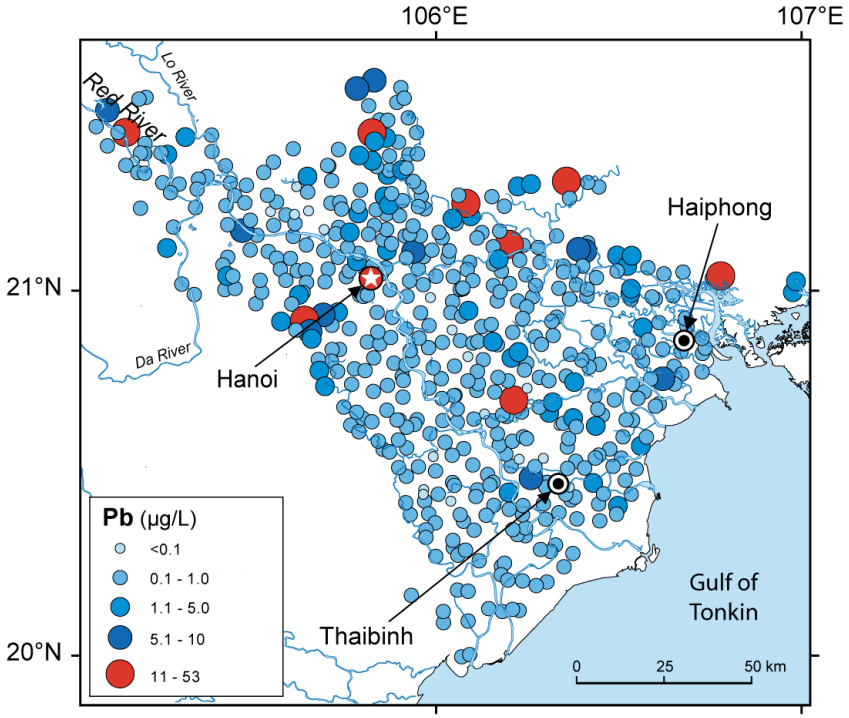



## 21 Iron

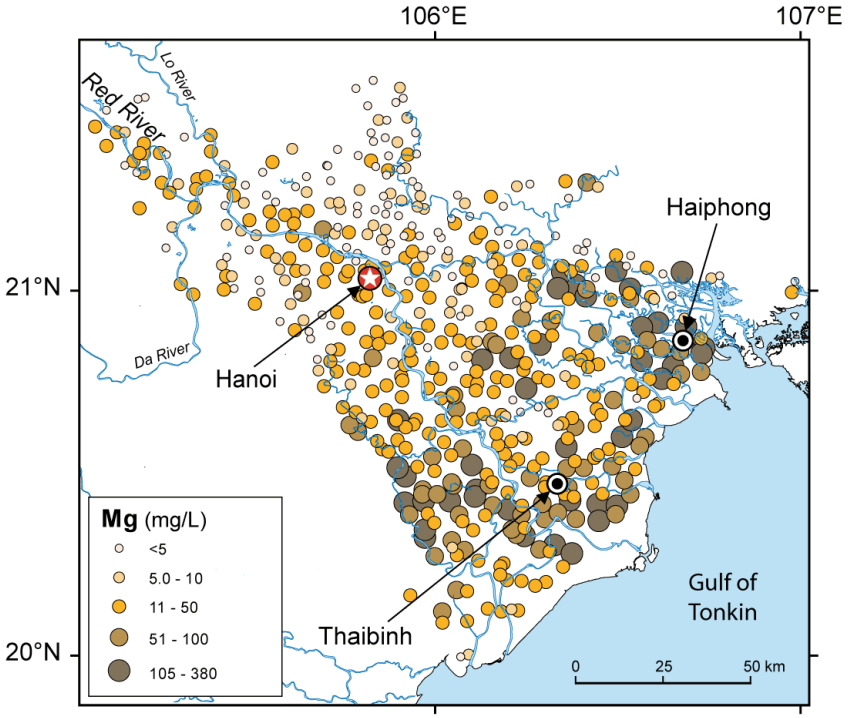



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 22 Lead

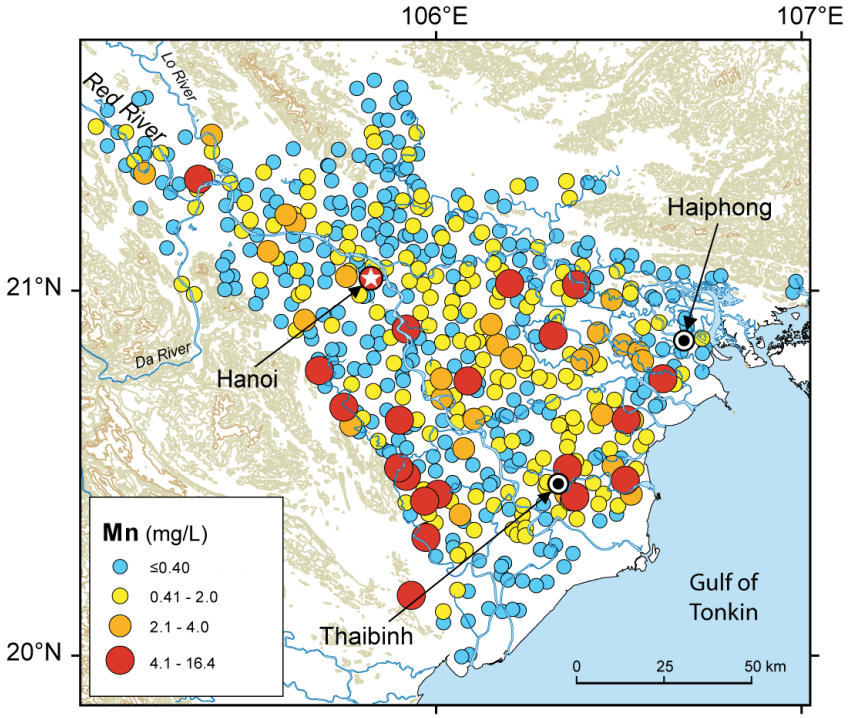



## 23 Magnesium

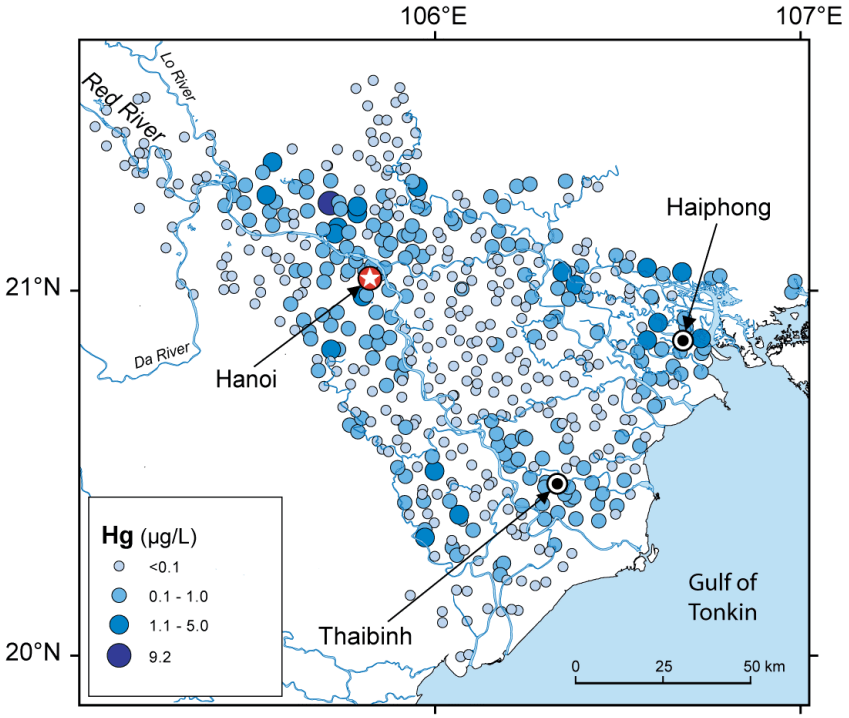



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 24 Manganese

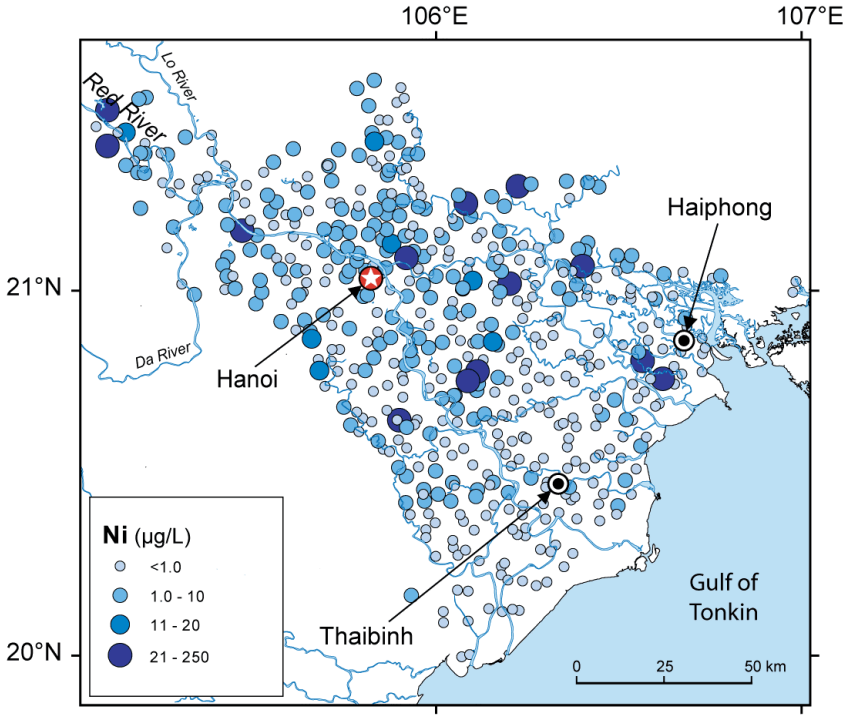



## 25 Mercury

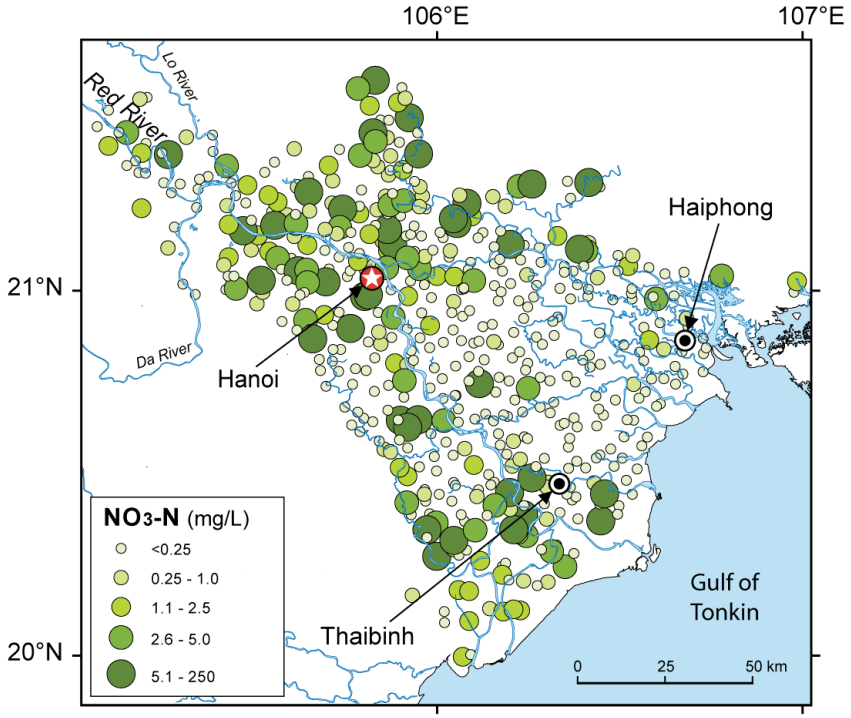



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 26 Nickel

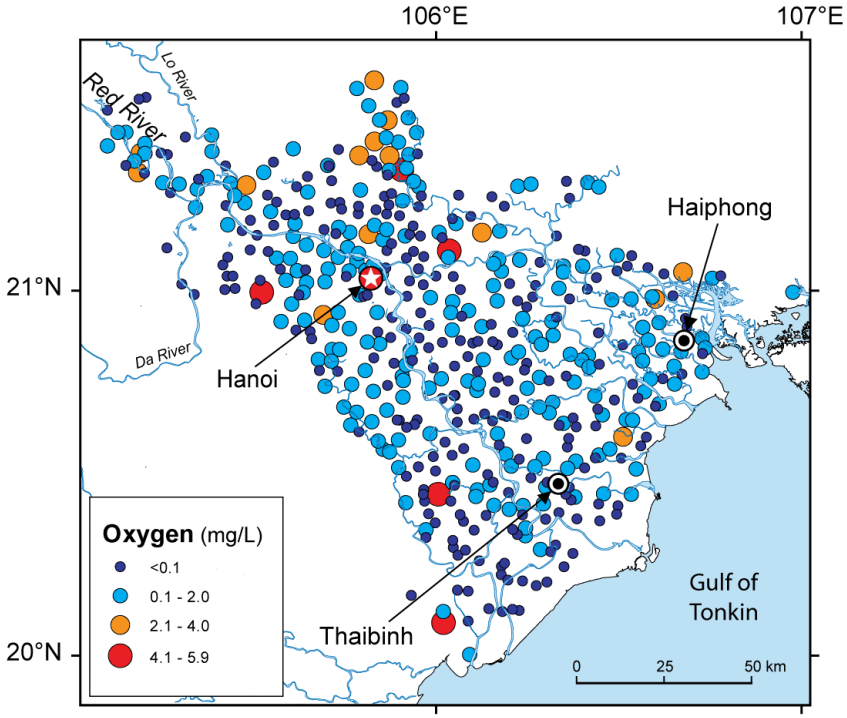



## 27 Nitrate

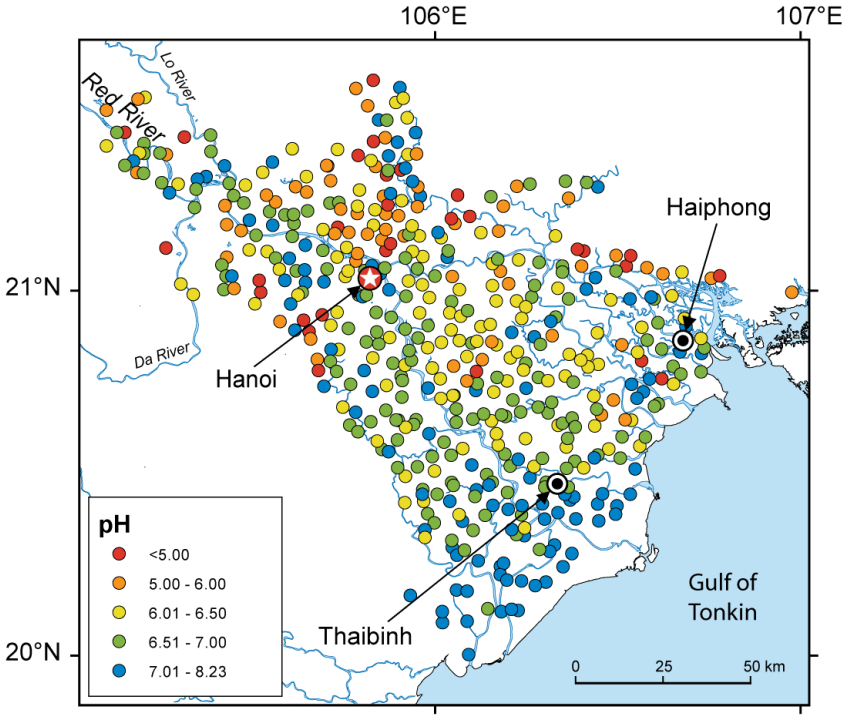



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 28 Dissolved Oxygen

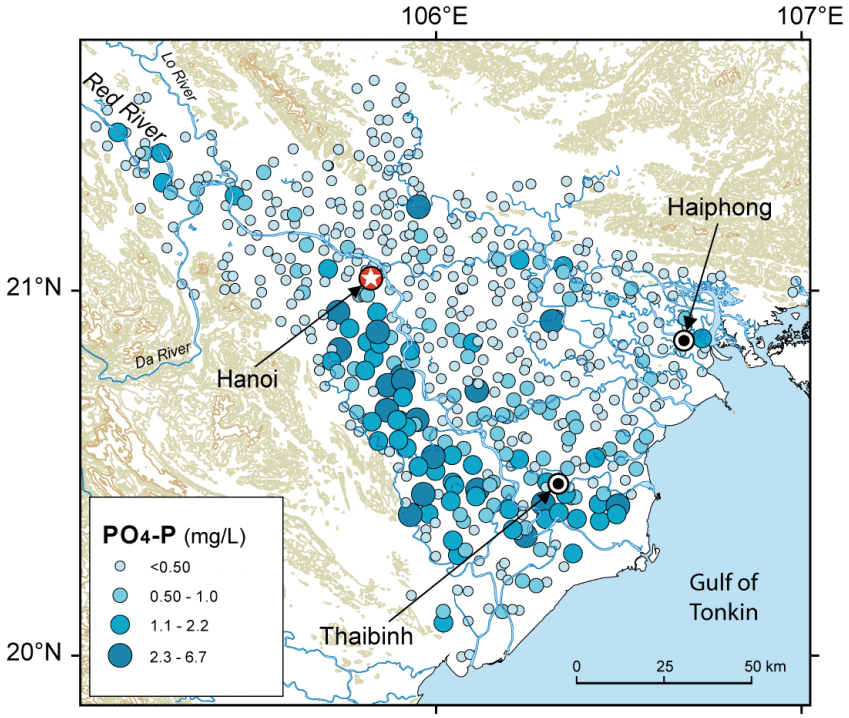



## 29 pH

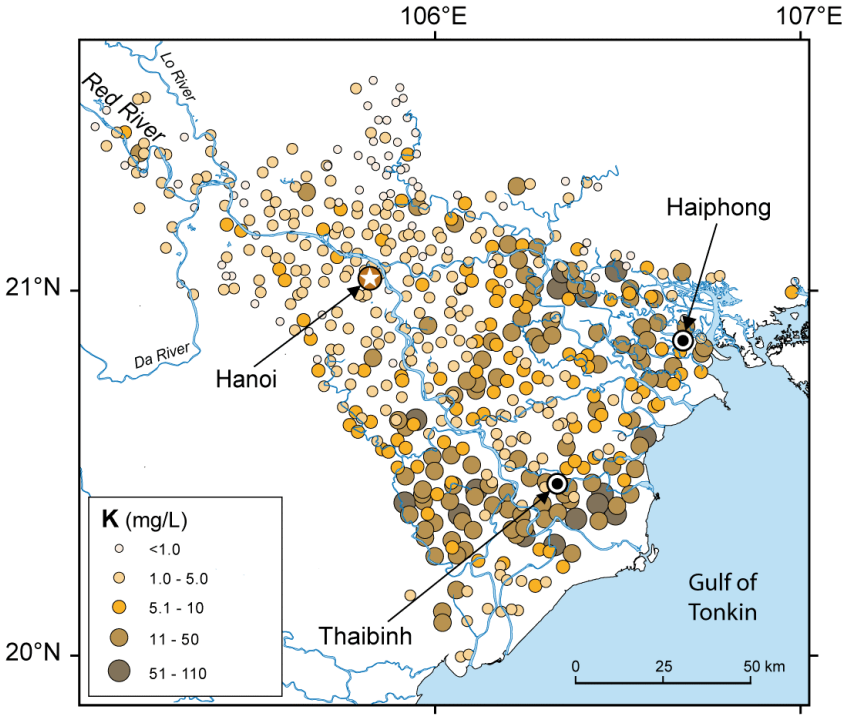



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 30 Phosphate

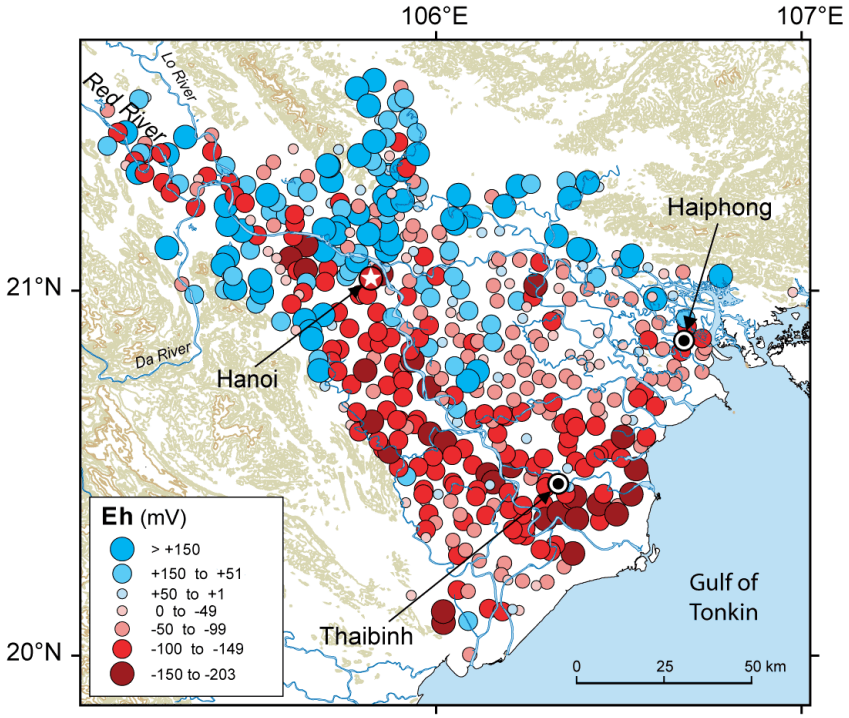



## 31 Potassium

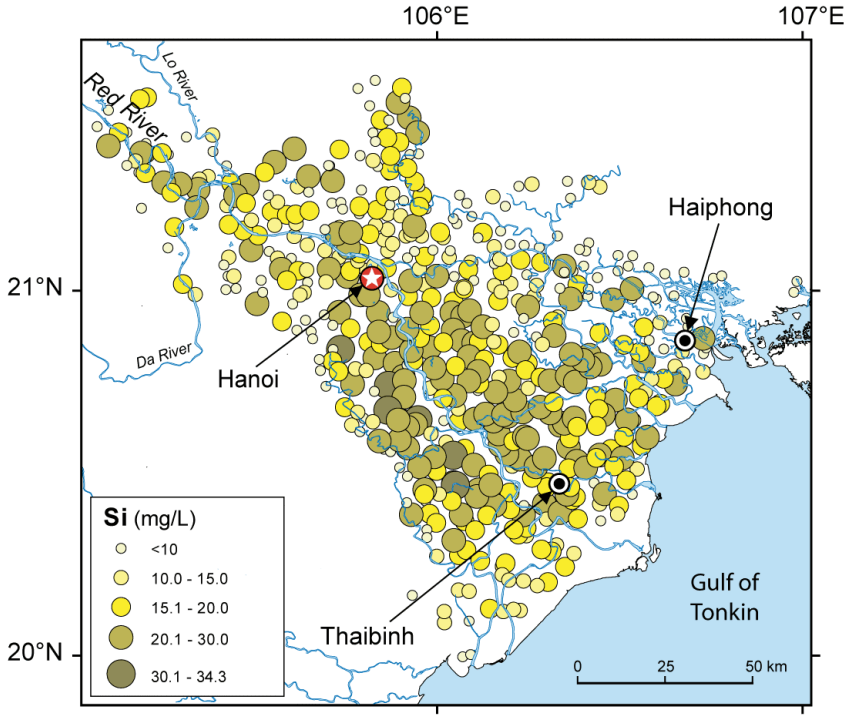



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 32 Redox potential (Eh)

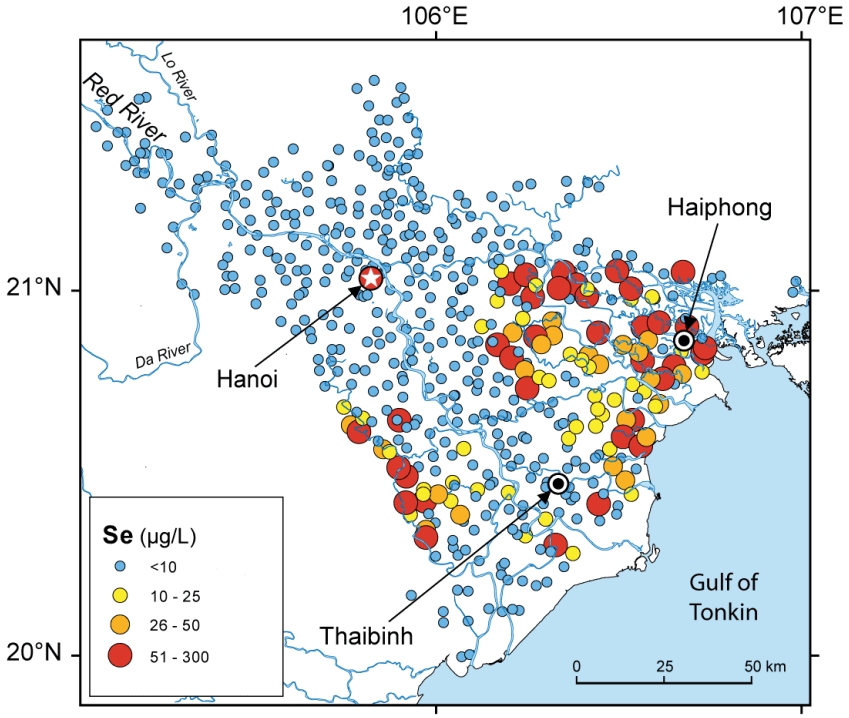



## 33 Silicon

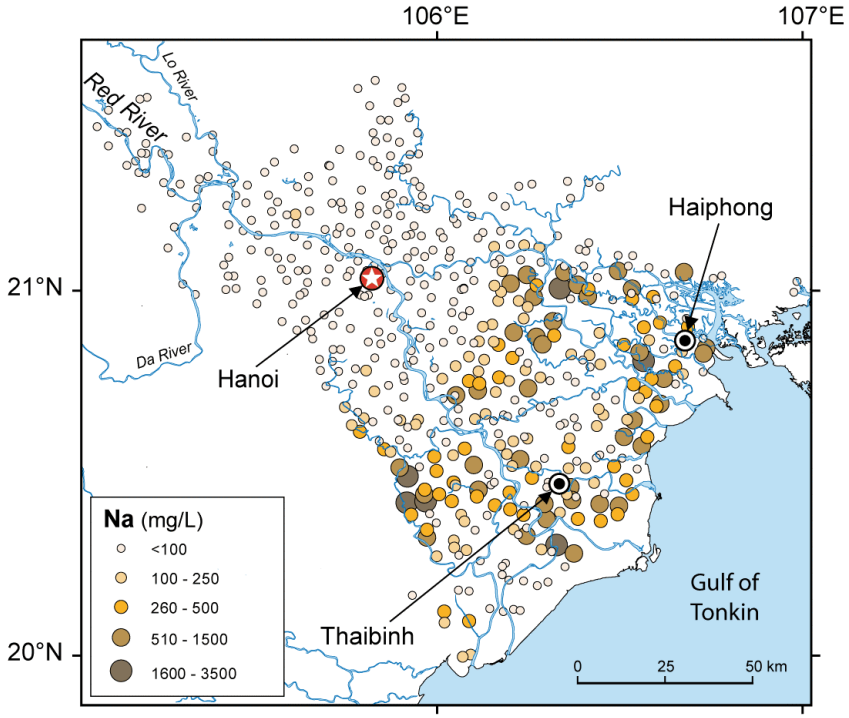



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 34 Selenium

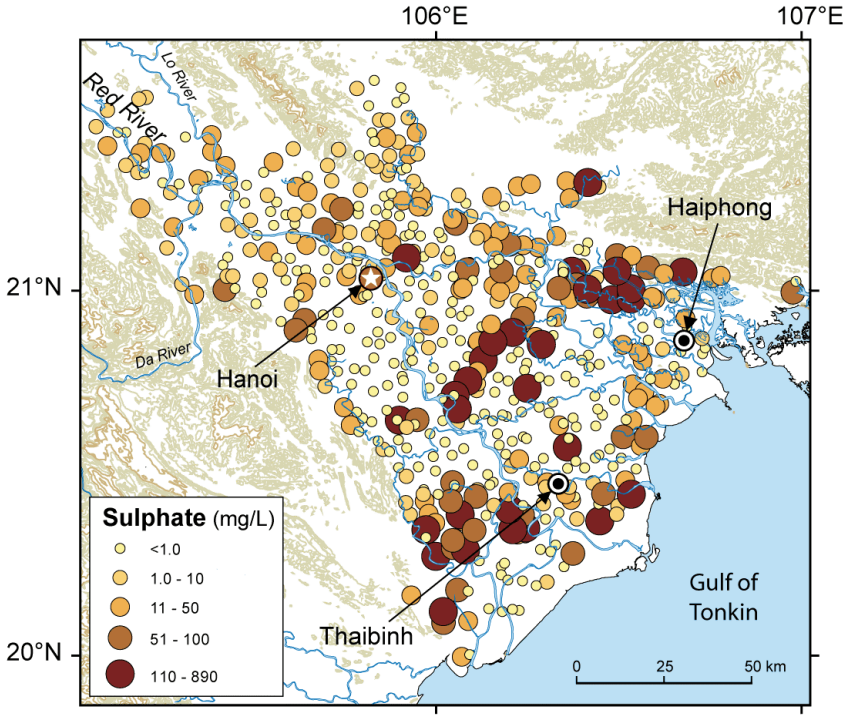



## 35 Sodium

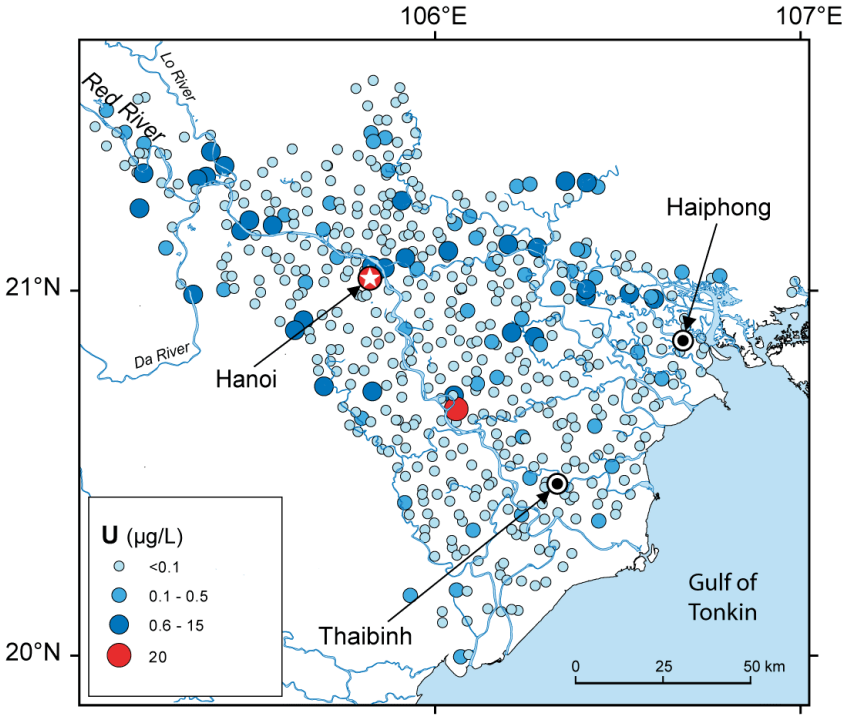



Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 36 Sulphate

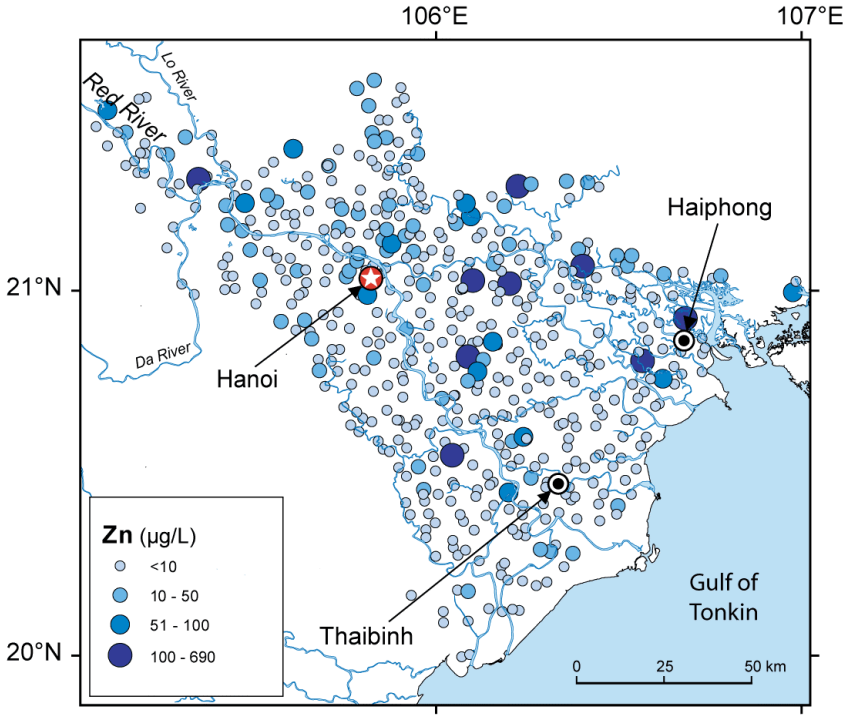



## 37 Uranium




Hydrochemical Atlas of the Red River delta, Vietnam
2010, Eawag, Swiss Federal Institute of Aquatic Science and Technology
www.eawag.ch/arsenic-vietnam
correspondence: Michael.Berg@eawag.ch

## 38 Zinc




## 6. References


1. Berg M, Stengel C (2006) ARS21-24 arsenic reference samples Interlaboratory Quality Evaluation (IQE). Report to Participants, Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dubendorf, Switzerland.

2. Hosmer DW, Lemeshow S (2000) Applied Logistic Regression, 2nd edition, John Wiley and Sons, New York.