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: LGIT, University of Grenoble, 1381 rue de la Piscine, 38400 Saint Martin d'Heres, France.*

*To whom correspondence should be addressed. E-mail: michael.berg@eawag.ch (M.B.).

Revised Research Report submitted to *PNAS*, 2. December, 2010

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

Major Category: PHYSICAL SCIENCES

Minor Category: Environmental Sciences

Winkel, L. H. E., Trang, P. T. K., Lan, V. M., Stengel, C., Amini, M., Ha, N. T., … Berg, M. (2011). Arsenic pollution of groundwater in Vietnam exacerbated by deep aquifer exploitation for more than a century. *Proceedings of the National Academy of Sciences of the United States of America PNAS*, *108*(4), 1246-1251. http://doi.org/10.1073/pnas.1011915108

Arsenic contamination of shallow groundwater is among the biggest health threats in the developing world. Targeting uncontaminated deep aquifers is a popular mitigation option although its long-term impact remains unknown. Here we present the alarming results of a large-scale groundwater survey covering the entire Red River delta and a novel probability model based on 3D Quaternary geology. Our unprecedented dataset reveals that ~7 million delta inhabitants use groundwater contaminated with toxic elements, including manganese, selenium and barium. Depth-resolved probabilities and arsenic concentrations indicate drawdown of arsenic-enriched waters from Holocene aquifers to naturally uncontaminated Pleistocene aquifers as a result of >100 years of groundwater abstraction. Vertical arsenic migration induced by large-scale pumping from deep aquifers has been discussed to occur elsewhere, but has never been shown to occur at the scale seen here. The present situation in the Red River delta is a warning for other As-affected regions where groundwater is extensively pumped from uncontaminated aquifers underlying high arsenic aquifers or zones.

\body

Geogenic arsenic (As) contamination of groundwater is a major health problem that has been recognized in several regions of the world, especially in South and Southeast Asia (Bengal delta (*1, 2*), Vietnam (*3-5*), Cambodia (*6, 7*), Myanmar (*8*) and Sumatra (*9*)). In 2001 it was reported for the first time that groundwater used as drinking water in the densely populated Red River delta in Vietnam contains high As levels (*3*). Since then, regional groundwater studies have been carried out in the vicinity of Hanoi city (10-30 km distance), on the banks of the Red River and its adjacent floodplains (*5, 10-14*) and along a 45 km transect across the southern and central part of the delta (*15*). High As levels were found in both the Holocene and Pleistocene aquifers (*3, 5, 10, 13*). Private wells predominantly extract water from the Holocene aquifers, whereas wells of urban treatment facilities tap Pleistocene aquifers (*3*). As is the case in other areas in SE Asia, the mechanism responsible for high groundwater As levels is the microbial and/or chemical reductive dissolution of As-bearing iron minerals in the aquifer sediments (*3-5, 10*).

The Red River delta is one of the most densely populated regions in the world, with a population density of about 1,160 people/km^2 covering an area of some 14,000 km^2 (*16*). Of the 16.6 Mio people that live in the Red River delta, 11 Mio have no access to public water supply and are therefore depending on other drinking water resources such as private tubewells. Given that groundwater is the main source of drinking water (*4*), it is of crucial importance that contaminated wells be identified. Here we present and discuss the results of an unprecedented groundwater study covering the entire Red River delta. We report delta-wide concentrations of As and 32 other chemical parameters and provide the complete geo-referenced database as Supporting Information (SI) Dataset 1. We show that 65% of studied wells exceed the WHO guidelines for safe drinking water for one or more chemical elements.

Arsenic risk maps for Southeast Asia were recently generated using surface information such as surface geology and soil properties (*8*). In the present study we improved these sub-continental scale predictions by developing a regional probability model for the Red River delta based on a new set of 3D-geological data (see Methods Section). Our data indicate that As enrichment in aquifers has been exacerbated by human activities, i.e. by the abstraction of large volumes of groundwater from Pleistocene aquifers. This finding has important implications for other As-tainted regions in the world with comparable groundwater flow systems and where water is pumped from deep aquifers at high rates.

Results and Discussion

Arsenic Distribution in the Delta. The distribution of groundwater As concentrations is illustrated in Figure 1a. Maps depicting the spatial distribution of an additional 32 chemical parameters are provided in the hydrochemical atlas (Supporting Appendix, section 5). Arsenic concentrations were found to vary greatly throughout the delta (<0.1–810 μgL^{-1}) and 27% of the wells exceeded the WHO guideline value of 10 μgL^{-1}. Our results imply that some 3 million people are currently using groundwater burdened with As concentrations >10 μgL^{-1} and one million people consume groundwaters containing >50 μgL^{-1}, with both rural and urban populations being affected by toxic levels of As. The highest concentrations are present in a 20 km wide band along the NW-SE boundary of the delta plain, to the SW of the modern Red River course and coinciding with the location of the palaeo-Red River channel (9,000 year BP) (*15*). The spatial distribution of As in this region matches a pattern of elevated PO_4^{3-}, NH_4^+ and DOC concentrations, along with negative Eh values and low sulfate (SO_4) concentrations indicating anoxic groundwaters (Figure 1c,d and section 5 in SI Appendix). These conditions are the trigger for reductive dissolution of iron phases and subsequent release of surface-bound As (*1, 17-21*).

However, as is evident from Figure 1e, As concentrations only become particularly elevated (>50 μgL^{-1}) where dissolved sulfate levels are low, i.e., where sulfate reduction accompanied by As sequestration in sulfide minerals is limited (*20*). Despite the typically reducing conditions, at the scale of the delta, the concentrations of As and Fe do not show a correlation. This has previously been attributed to differential sequestration of As and Fe into sulphide minerals (*17, 20, 22*), or the formation of other phases (e.g. siderite $FeCO_3$) (*10, 23*).

Arsenic is the element of greatest toxicological concern in the well waters. Second comes manganese (Mn) which can cause malfunction in children's development. Selenium (Se) and barium (Ba) are of lesser concern. With an average concentration of 0.83 mgL^{-1} (max. 16.4 mgL^{-1}), 44% of the wells exceed the Mn WHO guideline of 0.4 mgL^{-1}. We estimate that this percentage corresponds to nearly 5 million people who thus consume water with health-threatening Mn levels. Exposure to elevated Mn in drinking water is associated with neurotoxic effects in children, for example, a diminished intellectual function (*24*). The spatial distribution of Mn (<0.01–16.4 mgL^{-1}) (Figure 2a) and Fe (<0.05-140 mgL^{-1}) is heterogeneous throughout the delta (Fe map provided in SI Appendix), with Mn and As showing an anti-correlation (R^2=0.00). The highest concentrations of Mn and Fe are mainly found at negative Eh values (see Figure 1d,e and SI Appendix), indicative of the reductive dissolution of Fe and Mn-oxides according to the redox sequences of Fe and Mn reduction. However, some overlap between Fe and Mn reduction zones might occur (see Figure 1e), as has also been observed on a local scale (*12*). Further elements that notably exceed the WHO guidelines are Se (19% >10 μgL^{-1}, max. 300 μgL^{-1}) and Ba (7% >700 μgL^{-1}, max. 5100 μgL^{-1}). The distribution of elevated Ba and Se (Figure 2b) closely resembles the distribution of Cl, SO_4 and Na in the coastal stretch, indicating a marine source. Nevertheless, Se concentrations are considerably higher than can be expected from the Se/B ratio for seawater, which has an average concentration of 0.45 μgL^{-1} Se compared to 4.5 mgL^{-1} B (*25*).

In summary, 65% of all studied wells exceed the WHO guideline values for As, Mn, Ba, Se or a combination of these elements. Correspondingly, geogenic groundwater pollution in the Red River delta poses a serious long-term health threat to about 7 million people. This is particularly worrying since groundwater is the main source of drinking water (*4*).

Risk Modelling. Logistic regressions were applied to compute weighting coefficients of independent variables for the two regional As risk models: one based on surface information and the other based on 3D geological data (see SI Figures S1 and S3, SI Appendix; and SI Video 1). Table 1 lists the importance of, and weighting factors (λ) from the independent variables that showed significance for the models. In agreement with the recently published sub-continental As prediction model for Southeast Asia (*8*), sedimentary depositional environments make a larger contribution to the model than soil variables. Young organic-rich sediments (λ=1.46) play a larger role than recent deltaic deposits (λ=0.60), which supports the importance of organic matter in the mobilization of As (*5, 26-28*).

In the logistic regression model based on 3D geology data, the Lower Holocene (LH) aquifers (λ=3.95) clearly show the highest probability (P) of being contaminated with As. The sediments of this aquifer (lower boundary 3000 years BP; part of the Vinphuc and Haihung formations) are predominantly present in the incised valley of the Palaeo-Red River, where they unconformably lie over the Pleistocene sediments (Figure 1f and geological cross-sections in SI Figure 1). The LH aquifer has a very irregular thickness and partly exists only as large sandy lenses imbedded in a more silty matrix. The lithology is characterized by gray, very fine-to-medium sands laminated with greenish-gray silty-clays and organic-rich peat layers (*5, 29, 30*). There are two Pleistocene aquifers. The Lower Pleistocene (LP) aquifer, part of the Hanoi formation (lower boundary: 700,000 years BP), mainly consists of coarse yellow and brown sediments (*15, 29*) and is the only aquifer in the delta with an almost homogeneous presence. The Upper Pleistocene (UP) aquifer (lower boundary 125,000 years BP; part of the Vinphuc formation) has a more irregular appearance and generally shows a fining-upward structure, starting off with pebbly sands and ending with fine sands. Both Pleistocene aquifers play a minor role in the model

(λ=0.88 (LP) and 0.79 (UP)). The youngest aquifer (Upper Holocene (UH), lower boundary 1,000 years BP) mostly lies on top of a massive clay layer and is part of the Thai Binh formation. It consists of sandy silt and clay deposited in a delta plain environment (*29, 31*). The UH aquifer did not show significance during logistic regressions (p-value>0.05). The shallow depth and near-coastal location of the UH aquifer indicate saline groundwaters, which are generally not suitable for consumption. Furthermore, the unconfined character of this aquifer in combination with high SO_4 levels and low organic matter minimizes the probability of high As levels in the UH aquifer (*20*).

Arsenic Probability Maps. Figures 3a and 3b illustrate the probability of groundwater As exceeding 10 μgL^{-1}, computed with the model based on 3D geology and surface information, respectively. The probability map derived from 3D geology (Figure 3a) presents the average probability for all depths between 0 and 50 m. The individual probability maps (at given depths) indicate local probabilities up to 0.9 (see Figure 4). The classification results of both models are given in sections 3.1 and 3.2 in the SI Appendix. The model based on geology at depth is statistically better than the model based on surface parameters (74% vs. 65% correct classifications). Apart from the soil imprint in the surface model (P=0.4, orange colour, Figure 3b) which coincides with the modern Red River course (medium soil), the distribution of high and low probability levels is quite similar. The highest probabilities are found where organic-rich sediments are present, either at the surface (Figure 3b) (organic-rich deposits) or at depth (LH aquifer) (Figure 3a), and both models correctly delineate the 20 km wide strip with elevated As levels to the SW of the modern Red River course. This result underlines the strength of predictions solely based on surface parameters. 3D As risk modeling is a very valuable tool that can be applied in other As-affected regions of the world, but it must be kept in mind that aquifers are complex and heterogeneous and that misclassifications at a local scale are inevitable. Monitoring of groundwater quality will therefore remain an important task in the future. Furthermore, actual groundwater flow paths can't be modeled with a static approach and therefore 3D risk modeling would ideally be complemented with dynamic hydrological models that could indicate flow directions and changes of flow.

Arsenic Risk Areas at Depth and Indication of Downward Arsenic Migration.

Probability maps derived from the 3D model can potentially be an important resource for mitigation of As since they indicate where and at which depths tubewells can be expected to produce safe (low-As) groundwater. In the last part of this section, we will interpret the probability maps and we will also show that depth-resolved probabilities in combination with measured As concentrations can indicate a vertical transport of As from shallower Holocene aquifers into naturally uncontaminated Pleistocene aquifers.

Figure 4 shows the 3D distribution of As exceeding 10 μgL^{-1}, stacked at 10 m depth intervals. Selected probability maps thereof are overlain by As concentrations at different well-depth ranges (Figures 4b-d). Individual probability maps at depths of 0-60 m and 0-100m with As concentrations at corresponding depths are provided in SI Figure S8 and S10, respectively, SI Appendix; and SI Video 2). The high-risk area ($P>0.4$) at 10-20 m depth (Figure 4a) has a NW-SE trend and largely coincides with the position of the former Palaeo-Red River where sediments of the LH aquifer unconformably overlie the Pleistocene sediments (see SI Video 1; and SI Figure S4, SI Appendix). The 84% correctly classified As concentrations in the 10-20 m depth interval are an excellent result (see Figure 4b), particularly in light of the frequently observed heterogeneity of As concentrations, even over short distances (*5, 21, 32*). With increasing depth (Figure 4b and 4c), the high-risk area in the west splits up into two main patches. The spatial agreement between predicted and measured As concentrations is somewhat lower at 20-30 m than at 10-20 m (72% correctly classified, see Figure 4c) and especially the percentage of false-negative classifications is higher (25% vs. 13%), indicating that As-tainted wells (>10 µgL^{-1}) are present in low-risk areas. Moreover, the As concentrations at a depth of 20-30 m show a better match with the probability map for 10-20 m, which is supported by a better classification result (Figure 4d). Furthermore, a McNemar's chi-squared test and a Kappa test showed that the agreement between measured and predicted data is statistically significant different ($p<0.05$) between data shown in Figure 4d and data in Figures 4b and 4c. Particularly, the number of false negative cases was lowered from 25 to 17%, indicating that the number of As-tainted wells lying in a low-risk area is markedly lower. This can be illustrated by the five high-As wells (>50 µgL^{-1}) located in the low-risk area between the two high-risk patches (Figure 4c). These five wells actually tap the UP aquifer below the As-contaminated LH aquifer (Figure 4b).

The high As concentrations in the generally low-As UP aquifer could be explained by the reduction and mobilization of As adsorbed to sediments triggered by the leaching of organic matter from peat deposits above (*5, 21, 26, 27, 33*). However, considering the high As concentrations (>50 µgL^{-1}) in those five wells, a more plausible explanation would be vertical leaching of As-enriched groundwater from the LH aquifer or clay-dominated layer into the UP aquifer. This explanation is supported by the results of in-depth groundwater studies conducted at Hoang Liet village and in the area of Nam Du, where LH aquitards were found to be leaky, causing vertical percolation of As-rich groundwater from the LH to the Pleistocene aquifers (*5, 13*).

Impact of Long-term Pumping. Below 50 meters depth, no Holocene aquifers are present in the delta, and therefore the calculated probabilities of finding As are low (see probability map 50-60 m, Figure 4d and SI Figure S8, SI Appendix). However, also in the Pleistocene aquifers at these depths, As concentrations in groundwater exceed 10 or even 50 μgL^{-1} (max 330 μgL^{-1}). It is noteworthy that the highest As concentrations (>100 μgL^{-1}) are present in the same stretch in which the Holocene aquifers are contaminated by high As levels. Upon closer inspection, wells with the highest As concentrations in the Pleistocene aquifers (LP and UP) are mainly localized south of Hanoi, i.e. in the densely populated former province of Ha Tay (2,386,000 inhabitants in 1999; www.hanoi.gov.vn) which merged with Hanoi in 2008, and in the vicinity of the cities Ninbinh, Namdinh and Thaibinh (see Figure 5a). Berg et al. (*5*) have shown that the area south of Hanoi contains elevated As concentrations (130 μgL^{-1}) in the Pleistocene aquifer due to groundwater abstraction by the Hanoi water works, resulting in the vertical downward migration of reducing conditions and/or downward transport of As-tainted waters to the Pleistocene aquifers (see Figure 5b).

To get a better understanding of the presence of As in Pleistocene aquifers of Hanoi, we established a local prediction model based on 3D geology (see SI Figure S7 and SI Tables S6 and S7, SI Appendix). This Hanoi model performs poorly, with an AUC value of 0.555, which indicates that in this area natural variables fail to explain the As concentrations in the groundwater. This suggests the strong impact of human activities, i.e. large-scale groundwater pumping, on the As concentrations in the Pleistocene aquifers below Hanoi.

Groundwater exploitation from the deep aquifers in Hanoi began more than 110 years ago (1894) (*3*) to meet the water needs of the growing city under the French administration.

The demand for water for domestic and industrial purposes has gradually increased since then, and the enormous amount of 750,000 m^3/day of groundwater is pumped today from the deep aquifers in the Hanoi area alone, with an additional 500,000 m^3/day withdrawn in the southern part of the Red River delta (*34*). Our data indicate that large-scale groundwater abstraction from deep aquifers has actually impacted a much larger area of Pleistocene groundwater resources in the Red River delta than has been previously known. Consequently, elevated As concentrations in the Pleistocene aquifers in Hanoi and in the vicinity of Ninh Binh, Nam Dinh and Thai Binh seriously threat the quality of urban drinking water derived from these aquifers.

Implications and Future Prospects. It has been discussed in literature that excessive groundwater withdrawal could induce downward migration of As-enriched groundwater or organic matter and eventually lead to the contamination of currently As-free Pleistocene aquifers, for example in the most severely As-affected Bengal Basin, and elsewhere (*21, 33, 35-38*). Both Vietnam and Bangladesh exploit deep aquifers for urban water supply but whereas groundwater in Bangladesh is heavily used for irrigation, in Vietnam, agricultural fields are irrigated with river water. Previously, it has been suggested that oxidized sediments in Pleistocene aquifers have a significant capacity to attenuate As over hundreds of years because of adsorption (39). However, our present results indicate that this assumption might be proven wrong in situations where groundwater drawdown is pronounced. The lithologic composition and chemical conditions of Pleistocene sediments (i.e. oxidized pebbly coarse sand to fine sand) as well as of Holocene sediments in the Red River delta are comparable to those in the Bengal Basin (*14, 21*), but groundwater exploitation from Pleistocene aquifers in Vietnam began some 50–70 years earlier than in Bangladesh. Therefore, the present situation in Vietnam should be considered a warning of what can happen as a result of decades of groundwater abstraction from deep aquifers located below As-rich zones: the significant propagation of As to previously safe aquifers.

Use of groundwater that contains elevated concentrations of As and other geogenic contaminants, as well as groundwater pumped from deep aquifers in the vicinity of shallow high-As aquifers, should, in the long term, be avoided by the utilization of other sources of drinking water. Alternatively, appropriate water treatment technologies must be evaluated and installed to produce sustainable drinking water that meets safe water-quality standards for both rural and urban populations.

Methods

Groundwater Data. Groundwater samples were collected from 512 private tubewells in the Red River delta floodplains during three field campaigns (May-June 2005, November-December 2005, and January 2007), according to a random sampling strategy. The delta area was divided into grid cells of 25 km^2 (5x5 km) and in each cell one tubewell was chosen (sampling locations are shown in the hydrochemical atlas of the SI Appendix section 5). The study area is positioned at a latitude of 20.00°N to 21.57°N and a longitude of 105.07°E to 106.99°E.

Procedures of sampling and analysis were carried out as described in Berg et al. (2008)(*5*). Briefly, samples were collected after 15-30 minutes of pre-pumping to obtain stable levels of dissolved O_2 and Eh. Two samples were collected from each groundwater well. One of these two samples was filtered in the field (0.45 μm) and acidified (1% HNO_3). All samples were immediately shipped to the laboratory and stored at 4 °C in the dark until analysis. The chemical constituents were quantified from triplicate analyses. As concentrations were measured with high-resolution, inductively-coupled-plasma mass spectrometry (HR ICP-MS, Element 2, Thermo Fisher, Germany) and cross-checked by atomic fluorescence spectroscopy (AFS, PS Analytical, UK) or AAS (see SI Table S1, SI Appendix. Fe, Mn, Na, K, Ca, Mg, and Ba concentrations were measured by inductively-coupled-plasma optical emission spectroscopy (ICP-OES, Spectro Ciros CCD, Kleve, Germany); Co, Ni, Cu, Zn, Pb, Cr, Cd and Ba by ICP-MS; ammonium and phosphate by photometry; nitrate, sulfate and chloride by ion chromatography (Dionex, Switzerland); alkalinity by titration; and dissolved organic carbon (DOC) with a TOC 5000 A analyzer (Shimadzu, Switzerland). Details on the robustness of the measurements and limits of quantification are provided in section 1 of the SI Appendix and in Berg et al. (2008) (*5*).

Model Variables: Geological Data. The 3D geological data between 0 and -100m were obtained by the interpretation and interpolation (ordinary kriging) of 94 sediment cores in the Red River delta (drilled by NHEGD, Hanoi, Vietnam). Quaternary sedimentary units recognized in these sediment cores were correlated and subsequently classified into aquifers and aquitards of the Holocene or Pleistocene periods based on predominant lithology (grainsize) and age (^{14}C dating(*40, 41*)). On a regional scale, four different aquifers of the Quaternary period are present: 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). Three Quaternary aquitards were identified based on a lithology dominated by clay layers and occasionally intercalated peat lenses (NHEGD, Northern Hydrogeological and Engineering Geology division, Hanoi, Vietnam).

From the classified 3D geology data, five litho-stratigraphical cross-sections were derived (Figure 1f and SI Figure S3, SI Appendix) and 36 geological maps were constructed for specific depths: 2 m depth intervals for depths of 0-50 m b.s.l. and 10 m depth intervals for depths of 50-100 m b.s.l. (see SI Video 1). These maps were used as independent variables in our As prediction model for the Red River delta. A second model was made of the same area, but using surface data as independent variables. For this second model the same independent variables were used as in the SE-Asia model (*8*). These variables are deltaic deposits, alluvial deposits, organic-rich deposits, tidal deposits, other and pre-Holocene deposits, as well as percentages of silt, clay and sand in both the topsoil (0-30 cm) and subsoil (30-100 cm) and coarse, medium and fine soil textures. For information on data sources, see Winkel et al. (2008)(*8*).

As Prediction Model Development. As prediction models were obtained by: i) binary coding of As groundwater concentration data (dependent variable), using the WHO guideline value for As in drinking water (10 µgL^{-1}) as a threshold; ii) conducting logistic regression; and iii) calculating of the probability of As contamination based on the threshold value. Our data set of groundwater As concentrations in the Red River delta (see SI Dataset 1) was used as a dependent variable. Well-depths were corrected using the local elevation to depths, expressed relative to the mean sea level (m.s.l.).

Logistic regression was applied to determine the weighting of the independent variables(*8*). Briefly, log(odds) was modelled, which is defined as the ratio of the probability (P) that an event occurs to the probability that it fails to occur log(P/(1−P)):

$$\ln(odds) = C + \sum_{i=1}^{n} \lambda_i X_i \qquad (1)$$

where C is the intercept of regression, X_i are independent variables, and λ_i are the weighting coefficients that were obtained using the maximum likelihood procedure(*42*). Exponential values of coefficients, Wald statistics, and p-values (Table 1) indicate the importance of each variable. Independent variables that were statistically proven insignificant were excluded from the model during one of the subsequent regression steps. The threshold for maintaining a variable in the model was determined by the 95% significance level ($p<0.05$). According to the calculated odds, the probability (P) of having an As concentration above 10 μgL^{-1} was calculated as follows:

$$P = \frac{\exp(C + \sum_{i=1}^{n} \lambda_i X_i)}{1 + \exp(C + \sum_{i=1}^{n} \lambda_i X_i)} \qquad (2)$$

Acknowledgements

This project was substantially funded by the Swiss Agency for Development and Cooperation in the framework of the Swiss–Vietnamese cooperation project, ESTNV. We gratefully acknowledge Dao Manh Phu and Bui Hong Nhat for excellent support with groundwater sampling, M. Langmeier and R. Illi for anion analyses, A. Ammann and D. Kistler for assistance in ICP analyses, Luis Rodriguez-Lado for statistical tests, and Nguyen Van Dan for access to geological data.

References

1. Nickson R*, et al.* (1998) Arsenic poisoning of Bangladesh groundwater. *Nature* 395:338–338.
2. Chowdhury UK*, et al.* (2000) Groundwater arsenic contamination in Bangladesh and West Bengal, India. *Environ. Health Persp.* 108:393–397.
3. Berg M*, et al.* (2001) Arsenic contamination of groundwater and drinking water in Vietnam: A human health threat. *Environ. Sci. Technol.* 35:2621–2626.
4. Berg M*, et al.* (2007) Magnitude of arsenic pollution in the Mekong and Red River Deltas - Cambodia and Vietnam. *Sci. Total Environ.* 372:413–425.

5. Berg M*, et al.* (2008) Hydrological and sedimentary controls leading to arsenic contamination of groundwater in the Hanoi area, Vietnam: The impact of iron-arsenic ratios, peat, river bank deposits, and excessive groundwater abstraction. *Chem. Geol.* 249:91–112.
6. Polya DA*, et al.* (2005) Arsenic hazard in shallow Cambodian groundwaters. *Mineral. Mag.* 69:807–823.
7. Buschmann J*, et al.* (2008) Contamination of drinking water resources in the Mekong delta floodplains: Arsenic and other trace metals pose serious health risks to population. *Environ. Int.* 34:756–764.
8. Winkel L, Berg M, Amini M, Hug SJ, Johnson CA (2008) Predicting groundwater arsenic contamination in Southeast Asia from surface parameters. *Nature Geosci.* 1:536–542.
9. Winkel L, Berg M, Stengel C, Rosenberg T (2008) Hydrogeological survey assessing arsenic and other groundwater contaminants in the lowlands of Sumatra, Indonesia. *Appl. Geochem.* 23:3019–3028.
10. Postma D*, et al.* (2007) Arsenic in groundwater of the Red River floodplain, Vietnam: Controlling geochemical processes and reactive transport modeling. *Geochim. Cosmochim. Ac.* 71:5054–5071.
11. Larsen F*, et al.* (2008) Controlling geological and hydrogeological processes in an arsenic contaminated aquifer on the Red River flood plain, Vietnam. *Appl. Geochem.* 23:3099–3115.
12. Eiche E*, et al.* (2008) Geochemical processes underlying a sharp contrast in groundwater arsenic concentrations in a village on the Red River delta, Vietnam. *Appl. Geochem.* 23:3143–3154.
13. Norrman J*, et al.* (2008) Arsenic mobilisation in a new well field for drinking water production along the Red River, Nam Du, Hanoi. *Appl. Geochem.* 23:3127–3142.
14. van Geen A*, et al.* (2008) Comparison of arsenic concentrations in simultaneously-collected groundwater and aquifer particles from Bangladesh, India, Vietnam, and Nepal. *Appl. Geochem.* 23:3244–3251.
15. Jessen S*, et al.* (2008) Palaeo-hydrogeological control on groundwater As levels in Red River delta, Vietnam. *Appl. Geochem.* 23:3116–3126
16. Luu TNM*, et al.* (2010) Hydrological regime and water budget of the Red River Delta (Northern Vietnam). *J. Asian Earth Sci.* 37:219–228.
17. McArthur JM, Ravenscroft P, Safiullah S, Thirlwall MF (2001) Arsenic in groundwater: testing pollution mechanisms for sedimentary aquifers in Bangladesh. *Water Resour. Res.* 37:109–117.
18. Islam FS*, et al.* (2004) Role of metal-reducing bacteria in arsenic release from Bengal delta sediments. *Nature* 430:68–71.
19. Ford RG, Fendorf S, Wilkin RT (2006) Introduction: Controls on arsenic transport in near-surface aquatic systems. *Chem. Geol.* 228:1–5.
20. Buschmann J, Berg M (2009) Impact of sulfate reduction on the scale of arsenic contamination in groundwater of the Mekong, Bengal and Red River deltas. *Appl. Geochem.* 24:1278–1286.
21. Fendorf S, Michael HA, van Geen A (2010) Spatial and Temporal Variations of Groundwater Arsenic in South and Southeast Asia. *Science* 328:1123–1127.
22. Moore JN, Ficklin WH, Johns C (1988) Partitioning of arsenic and metals in reducing sulfidic sediments. *Environ. Sci. Technol.* 22:432–437.
23. Herbel M, Fendorf S (2006) Biogeochemical processes controlling the speciation and transport of arsenic within iron coated sands. *Chem. Geol.* 228:16–32.
24. Wasserman GA*, et al.* (2006) Water manganese exposure and children's intellectual function in Araihazar, Bangladesh. *Environ. Health Perspect.* 114:124-129.
25. Hem JD (1985) Study and interpretation of the chemical characteristics of natural water.263.
26. McArthur JM*, et al.* (2004) Natural organic matter in sedimentary basins and its relation to arsenic in anoxic ground water: the example of West Bengal and its worldwide implications. *Appl. Geochem.* 19:1255–1293.
27. Meharg AA*, et al.* (2006) Codeposition of organic carbon and arsenic in Bengal Delta aquifers. *Environ. Sci. Technol.* 40:4928–4935.
28. Rowland HAL*, et al.* (2007) The control of organic matter on microbially mediated iron reduction and arsenic release in shallow alluvial aquifers, Cambodia. *Geobiology* 5:281–292.
29. Nghi T, Ngo QT, Do TVT, Nguyen DM, Nguyen VV (1991) Quaternary sedimentation of the principal deltas of Vietnam. *J. Southeast Asian Earth Sci.* 6:103–110.

30. Tanabe S, *et al.* (2006) Holocene evolution of the Song Hong (Red River) delta system, northern Vietnam. *Sediment. Geol.* 187:29–61.
31. Hanebuth TJJ, Saito Y, Tanabe S, Vu QL, Ngo QT (2006) Sea levels during late marine isotope stage 3 (or older?) reported from the Red River delta (northern Vietnam) and adjacent regions. *Quatern. Int.* 145-146:119–134.
32. Papacostas NC, Bostick BC, Quicksall AN, Landis JD, Sampson M (2008) Geomorphic controls on groundwater arsenic distribution in the Mekong River Delta, Cambodia. *Geology* 36:891–894.
33. Ravenscroft P, McArthur JM, Hoque BA (2001) *Geochemical and palaeohydrologicical controls on pollution of groundwater by arsenic. In Arsenic exposure & Health Effects IV* (Elsevier, Oxford).
34. Cao TH, *et al.* (2005) Improving the supply water quality of Hanoi. Part 1: Current situation of supply water and challenges for treatment technology. *Magazine of the Vietnam Water Supply and Sewerage Association* 7:31–35 (in Vietnamese).
35. Harvey CF, *et al.* (2002) Arsenic mobility and groundwater extraction in Bangladesh. *Science* 298:1602-1606.
36. van Geen A, *et al.* (2003) Spatial variability of arsenic in 6000 tube wells in a 25 km(2) area of Bangladesh. *Water Resour. Res.* 39:1140, doi:10.1029/2002WR001617.
37. Michael HA, Voss CI (2008) Evaluation of the sustainability of deep groundwater as an arsenic-safe resource in the Bengal Basin. *Proc. Natl. Acad. Sci. U. S. A.* 105:8531–8536.
38. Polizzotto ML, Kocar BD, Benner SG, Sampson M, Fendorf S (2008) Near-surface wetland sediments as a source of arsenic release to ground water in Asia. *Nature* 454:505-U5.
39. Stollenwerk KG, *et al.* (2007) Arsenic attenuation by oxidized aquifer sediments in Bangladesh. *Sci. Total Environ.* 379:133–150.
40. Tanabe S, *et al.* (2003) Sedimentary facies and radiocarbon dates of the Nam Dinh-1 core from the Song Hong (Red River) delta, Vietnam. *J. Asian Earth Sci.* 21:503–513.
41. Tanabe S, *et al.* (2003) Song Hong (Red River) delta evolution related to millennium-scale Holocene sea-level changes. *Quaternary Sci. Rev.* 22:2345–2361.
42. Kleinbaum DG, Klein M (2002) *Logistic Regression: A self-learning text* (Springer, New York).

Figures

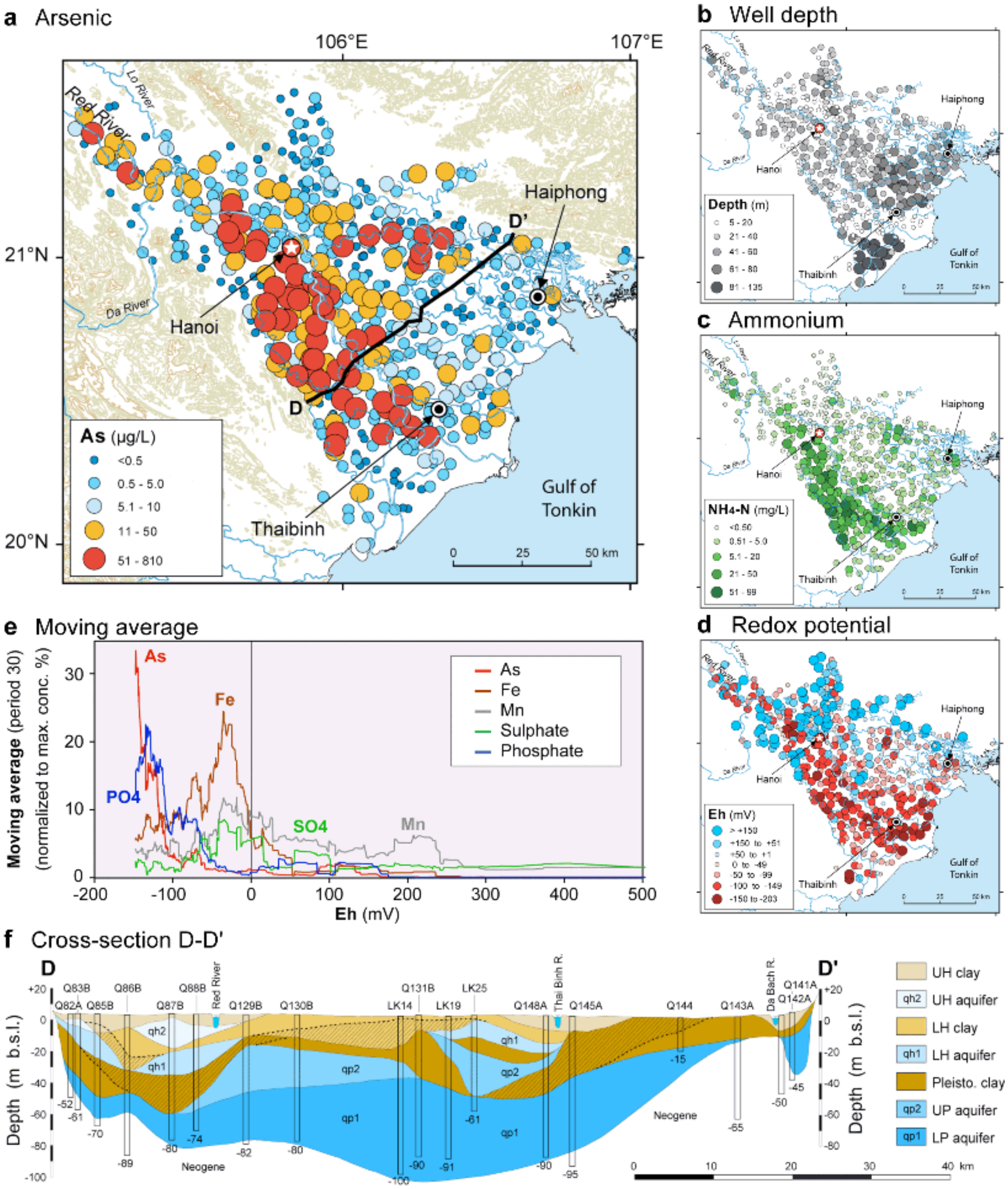


Figure 1. Concentrations of As and selected parameters observed in groundwater of the Red River delta. High-resolution maps of each parameter are in the SI Appendix. **a,** Arsenic concentrations in groundwater collected in the period from 2005 to 2007. **b,** Depth of sampled tubewells. **c,** Ammonium (NH_4^+) concentration. **d,** Redox potential (Eh). **e,** Concentration trends of As, Fe, Mn, phosphate (PO_4^{3-}) and sulfate (SO_4^{2-}) plotted against measured redox potential (Eh). Concentrations were normalized w.r.t. maximum concentrations and smoothed, using a moving average filter with a period of 30. **f,** Simplified geological cross-section along the transect D–D' indicated in Figure 1a. Further geological transects are presented in SI Figure S3, SI Appendix.

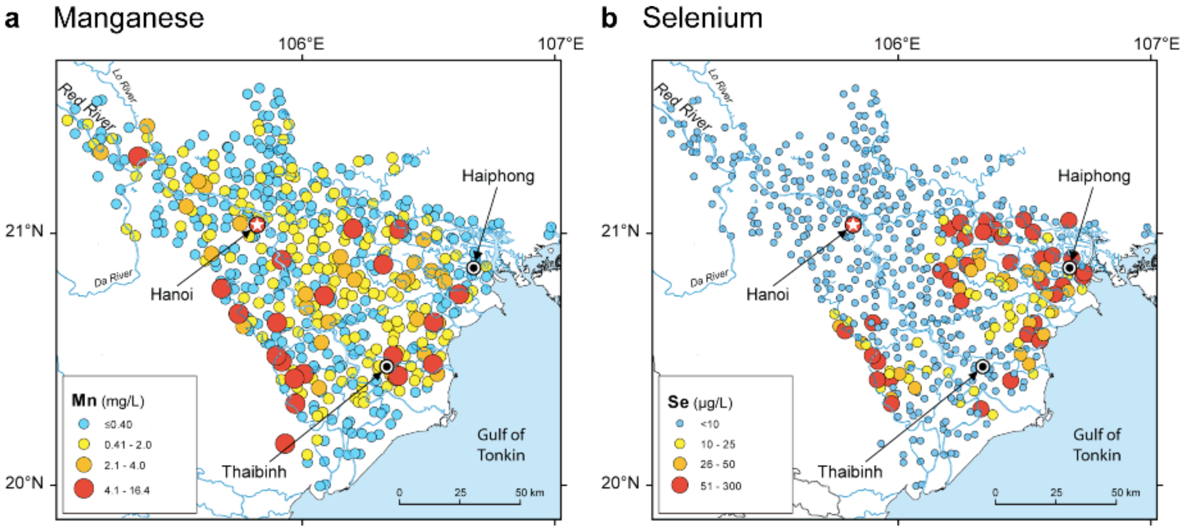
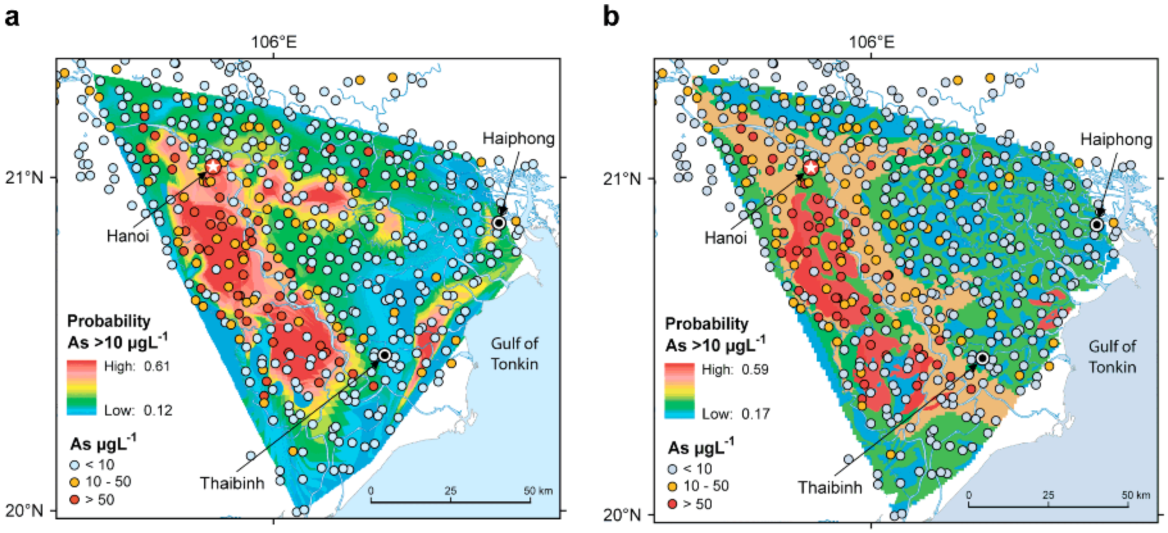


Figure 2. Concentrations of Mn and Se observed in groundwater of the Red River delta. a, Mn concentrations show a heterogeneous distribution throughout the delta. **b,** Elevated Se concentrations are found mainly along the coast and in aquifers affected by seawater intrusions.



Probability map for the Red River delta based on geology at depth (average over depth)

Probability map for the Red River delta based on surface parameters (soil and geology)

Figure 3. Modeled probability of As concentrations exceeding 10 µgL^{-1}. a, Average probabilities based on 3D geology integrated over the depth range of 0-50 m (74% correctly classified). **b,** Probabilities obtained from the prediction model based on land-surface geology and soil data (65% correctly classified).

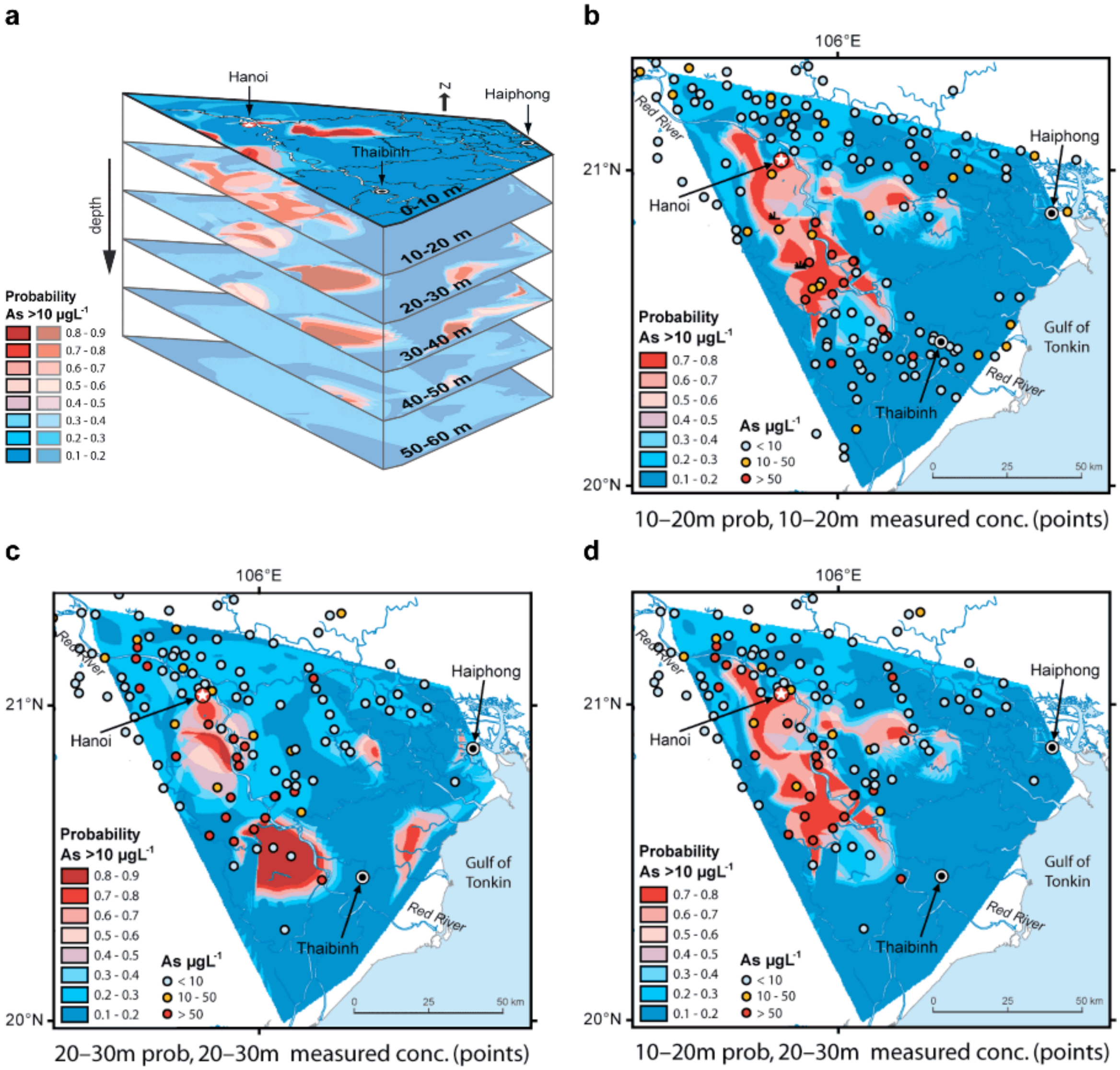


Figure 4. Risk of As pollution plotted in 3D and at 10 m depth intervals. a, 3D distribution of As exceeding 10 μgL^{-1}, stacked at 10 m depth intervals (see also SI Figure S8, SI Appendix). **b,** Average probability and measured As concentrations at a depth of 10-20 m (m.s.l.). Model classification results based on a probability cut-off value of 0.4 are: 84% correctly classified, 3% false-positive (As <10 μgL^{-1} in high-risk areas), and 13% false-negative (As >10 μgL^{-1} in low-risk areas). **c,** Average probability and measured As concentrations at a depth of 20-30 m (m.s.l.). Classification results are: 72% correct, 3% false-positive and 25% false-negative. **d,** Average probability and measured As concentrations at a depth of 10-20 m (same probability data as in Figure 4b) overlain by As concentrations from 20-30 m. Classification results are better than those for Figure 4c: 74% correct, 9% false-positive and 17% false-negative.

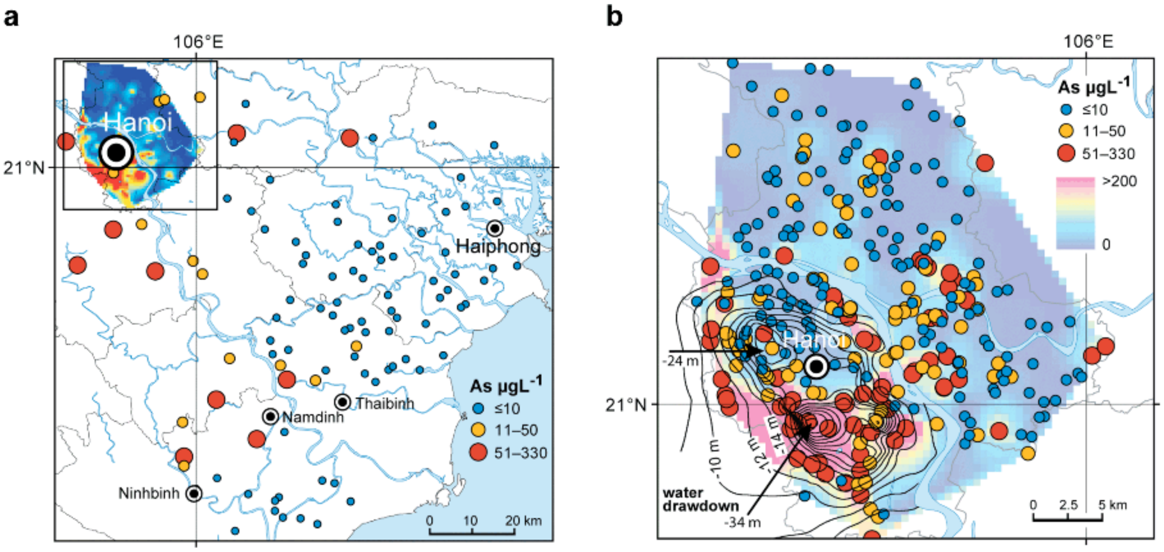


Figure 5. As concentrations in Pleistocene aquifers of the Red River delta at depths >50 m. a, Highest As concentrations (up to 330 µgL^{-1}) in the Pleistocene aquifer (at depths >50 m) are found in the same area where high As concentrations are present in shallower, Holocene aquifers (see also Figure 1a). **b,** The Hanoi area outlined by the box in Figure 5a. As concentrations in the Hanoi area were provided by the Vietnam Geological Survey. The interpolated As concentration map was obtained by ordinary kriging of this detailed data set (n=307). Contour lines of piezometric heads (recorded in Dec. 2006) depict the pronounced drawdown of Pleistocene groundwater levels (down to -34 m), caused by extensive groundwater pumping by the Hanoi Water Works (*5*).

Table

Table 1. Results of logistic regression analysis.

Prediction model	Output Variable	λ	Wald	p-value
Surface variables	Organic-rich deposits	1.46	14.44	0.000
	Deltaic deposits	0.60	5.53	0.019
	Alluvial deposits	0.59	4.08	0.043
	Medium-textured soils	0.46	4.19	0.041
–	Regression constant	–1.55	73.65	0.000
3D geology	Lower Holocene aquifer	3.95	54.81	0.000
	Lower Pleistocene aquifer	0.88	5.26	0.022
	Upper Pleistocene aquifer	0.79	4.48	0.034
–	Regression constant	–1.98	41.38	0.000

Statistically evaluated weighting coefficients of the independent variables in this study that were used to compute probabilities of As contamination are denoted by λ. Wald and p-values indicate the significance of the variables. Wald values give the relative importance in percentages and p-values the absolute significance, where a value <0.05 indicates a significance of at least 95%. Variables that were not statistically significant (p >0.05) were not considered in the modelling, i.e., other Holocene deposits, pre-Holocene sediments, coarse and fine soil textures, sand, silt and clay soil contents in the surface-based model, and the Upper Holocene aquifer in the model based on 3D geology.