



## **Groundwater quality prediction maps for risk-based decision making**

Michael Berg, Anja Bretzler, and Joel Podgorski

Eawag, Swiss Federal Institute of Aquatic Science and Technology, Dübendorf, Switzerland

Sampling and analysing groundwater for the presence of inorganic contaminants relevant to human health (e.g. arsenic, fluoride, nitrate) is time-consuming and costly, especially in regions with weak infrastructure. Predictive mapping and modelling techniques exist that have the capability to identify regions where aquifers are particularly vulnerable to certain contaminants, both from geological or anthropogenic sources. Over more than 10 years of applied research, we have refined an approach by which measurements of contaminants in groundwater are correlated with widely available spatial data of environmental parameters (e.g. geology, soil properties, climate, population density) using statistical and machine learning methods. The resulting hazard maps predict areas of contaminated groundwater by giving the probability of contaminant concentrations exceeding a given threshold value. Examples of produced maps include successful predictions of arsenic in groundwater for South-East Asia, China, Pakistan and Burkina Faso, and fluoride in India. Recently, we have also started assessing and mapping the vulnerability of aquifers to surface contaminants such as nitrate using these techniques.

To facilitate access to these maps and associated data, we have developed the Groundwater Assessment Platform (GAP), accessible at [www.gapmaps.org](http://www.gapmaps.org). GAP is a free, interactive online GIS platform for the mapping, sharing, analysis and statistical modelling of groundwater quality data. Maps of existing data and models can be viewed by all users, whereas a free login enables registered users to upload and process their own data in a secure environment and make use of the statistical modelling functionalities. Prediction maps of point data can be generated by means of gridding (inverse distance or nearest neighbour) or logistic regression, using either publicly available spatially continuous global datasets of various environmental parameters or one's own raster or polygon data as predictor variables. Despite statistical programming packages offering more options for manipulating and modifying data and models, the prediction maps produced using GAP are remarkably similar and require substantially less time to generate, due to GAP's simple, quick-to-use modelling interface.

Such contaminant hazard maps provide valuable baseline information for practitioners and scientists to help prioritize locations for detailed water quality surveys, to target specific contaminants and to site new boreholes. They are also useful in estimating the number of people potentially exposed to contaminants in drinking water and thereby at risk to related diseases such as arsenicosis or fluorosis.