Groundwater Quality Sampling at Contaminated Sites: The Long And The Short Of It.

Current guidance and practice

Within the UK there are thousands of monitoring wells (or boreholes) used for groundwater quality sampling. A significant proportion are relatively long screened wells (>3m) especially where the aquifer thickness and/or the dynamic range in the water table is large or the monitoring facility is old. A regulatory objective of all groundwater sampling is to obtain a "representative" sample (e.g. BS ISO 5667-22, 2009). There is however no clear definition of what this means. The origin of a groundwater sample taken from a monitoring well is fundamentally controlled by a complex interplay of factors including the prevailing hydrogeological conditions, the screen length and the sampling methodology employed. Different sampling methods used in the same well will often yield different water quality samples - so which of these is the representative sample?

> *"Understanding groundwater quality is fundamentally underpinned by understanding groundwater flow. We have a long way yet to go in the development of new instrumentation before flows can be routinely measured alongside water quality sampling."*

Author/Contact Details:
Lindsay A. McMillan, Michael O. Rivett, John H. Tellam
Water Sciences
School of Geography, Earth and Environmental Sciences, University of Birmingham B15 2TT, UK

Peter Dumble
Independent Hydrogeologist
Tiverton, Devon, EX16 7TA, UK


Low-flow groundwater sampling

Field sampling is typically undertaken by technicians who, ideally, have been trained to follow standard sampling protocols specified by experienced professionals (preferably including a hydrogeologist) to match a set of sampling objectives. Such protocols are generally founded upon industry, national or international standards and guidance (e.g. BS ISO 5667-11, 2009; NIGLQ, 2015).

Without adequate training or regard to science or guidance, sampling practice may be poor. The economic and environmental consequences of poor sampling and/or the misinterpretation of groundwater quality data, particularly for remediation schemes, can be significant. For example, inadvertent loss of volatile organic contaminants during sampling may lead to missed plumes.

Representative groundwater samples

So, what methodologies should we use when sampling groundwater and what issues should be considered when interpreting groundwater quality data? Technically, what do water quality samples tell us about the distribution, origin and concentration of contaminants in the surrounding aquifer? Are these the "representative" samples required by our regulators and are they sufficient for decision making? When are short screens more appropriate than long screens? Are high resolution (costly) multi-level samplers necessary? Are in-vogue sampling methodologies such as low-flow or passive / zero-purge more appropriate than long-established 3 to 5 well volume purge protocols (Figure 1)? These questions have recently been brought into focus by researchers at the University of Birmingham, both in the interpretation of water quality data from long-screened wells (McMillan et al, 2014) and monitoring of heterogeneous source zones (Rivett et al, 2014).

Guidance increasingly requires screened intervals to be as short as possible, ideally 3m or less (BS ISO 5667-22, 2010). USEPA operating procedures for low-flow sampling now explicitly excludes the use of low-flow methodologies in well screens exceeding 3m in length (USEPA, 2010), yet many practitioners in the UK, encouraged perhaps by some ambiguous guidance in BS ISO 5667-11, 2009, increasingly use low-flow sampling in longer screened wells driven by the need to minimise sample disturbance and reduce contaminated purge volumes requiring disposal.

Guidance reflects concerns that increased screen lengths introduce greater potential for bias in samples with resulting uncertainty when interpreting sample origin and aquifer concentrations. So how real are these concerns?

Understanding the origin of a groundwater sample

Understanding groundwater sample origins is critical and is underpinned by knowledge of flows, both the natural (ambient) groundwater flow regime and the localised flow regime induced by the sampling process.

The three principal sampling methodologies were all originally conceived to deal with collecting samples from short -screen monitoring wells. Theoretically, a pumped sample is not a simple average of concentrations spanning the screened section, but rather a 'flow-weighted average' sample with proportions of contributed water weighted according to the permeability's and hence flows from the various geological units intersected – the rate of flow through highly permeable gravel being significantly greater than that through lower permeability sands or silts (Figure 1a, b). Hence a sample may be dominated by contributions from a thin, but very permeable gravel bed and receive limited contributions from thicker beds of sand or silt. The latter may in fact be much more contaminated but contribute little to the water sample recovered; consequently if the permeable gravel layer has low contamination, the sample may be relatively clean.

Where pumping rates are small relative to the screen volume we may have to pump for many hours to achieve a flow-weighted average sample which is representative of inflow across the entire well screen (Figure 2a, b, c). Research by Martin-Hayden et al (2012) and McMillan et al (2014) demonstrates the need to purge between 2 and 3 times the "screen volumes" (i.e. screen length x cross-sectional area) regardless of sampling protocol in order to achieve a fully-purged flow-weighted average sample. This implies the removal of significantly greater volumes and hence longer purge times than is normally the case when applying standard low-flow sampling protocols.

But why purge at all? If groundwater is already mixing freely into and through the screened section of a well, a flow-weighted average chemistry may already be present without the need to purge (Figure 1c). Passive or zero-purge samples will then produce similar results to low-flow sampling. This natural mixing process also explains why partially-purged low flow samples can achieve flow-weighted average conditions after short pumping times. Assessing the ambient natural flow through wells is difficult, but important to the validation of zero or low-flow options. Tracer-labelled passive samplers that measure flows by assessing mass loss of the tracer dissolved into the flowing groundwater offer a promising option. These samplers also measure contaminant flux via simultaneous contaminant sorption on to the sampler (Annable et al. 2005).

Heterogeneity (even at 3m monitoring scales)

Geological and contamination heterogeneities potentially sampled by even 3m well screens can be considerable as shown by adjacent monitoring well and multilevel sampler data (Rivett et al., 2014). On-going analysis of this dataset shows that low-flow sampling of the 3m well screens could effectively monitor aquifer contaminant flux. This is because concentrations were dominated by flow through a thin gravel bed and were moderately representative of a flow-weighted average concentration. In contrast, the multilevel sampler data were effective at identifying contamination in the lower permeability units, but these contributed negligible water to the low-flow sample.

Influence of vertical flows

Vertical flows, which occur to some extent in all aquifers, further complicate the picture. In long-screen wells, vertical flows can be considerable and the well provides an open conduit to flow. Low-flow sampling rates are unlikely to fully counter ambient flows and even at greater pumping rates and duration, the sample may remain biased towards the quality of water derived from the point of inflowing water to the well and may never be drawn from the entire screen interval (Figure 2d). Such vertical flow bias begins when the ambient up-flow of water through the well is much less (<20%) than the pumping rate; and it may be necessary to pump at many times higher than this rate to obtain a sample from across the entire screen interval (McMillan et al., 2014). In some cases, old long-screen wells may have served as a significant local conduit of contamination to deeper aquifers; consideration should be given to decommissioning and sealing of such wells or retrofit with shallow and deep short-screen wells.

Where moderate vertical flows exist, a pragmatic line would be to set a sampling objective which targets that flow stream. In these instances purging is unnecessary. The interpretation of the sample, however, demands an understanding of the flow regime, and recognition that it provides a sample drawn from a specific section of the well screen.

At the present time there are few reliable or economic methods for measuring often low, but nevertheless important, ambient flows in monitoring wells. In the absence of flow measurements, uncertainty is inevitable when interpreting sample origin and water quality concentration in the adjacent aquifer. Where screen lengths are long, this uncertainty is greater. Recent application of fibre optic distributed temperature sensor (DTS) methods to infer well flows from temperature data and passive flux meter methods of Annable et al. (2005) offer promising potential.

Improving the science and practice of groundwater sampling

The ongoing development of smaller and smarter sampling equipment and increased use of shorter well screens and multi-level samplers is leading to a better understanding of the inherent difficulties with interpretation of water quality results.

There is significant temptation to move (perhaps unthinkingly) from onerous volume purge to low-flow or passive methods. However, with long established monitoring locations (that can often include relatively long-screen wells), changes in sampling protocol needs to ensure that the data collected remain fit-for-purpose and the historical data record is transferable. An overlap period of old and new methodology is warranted.

The longer the well screen the greater the potential for increased complexity of contributing geological horizons with differing contamination. Whilst this provides a key driver towards using short-screen wells, even sample results from 3 m or shorter screens can be misinterpreted. Longer screen wells are more prone to significant flows up or down the well driven by natural hydraulic head variations that require recognition in sampling protocols and data interpretation.

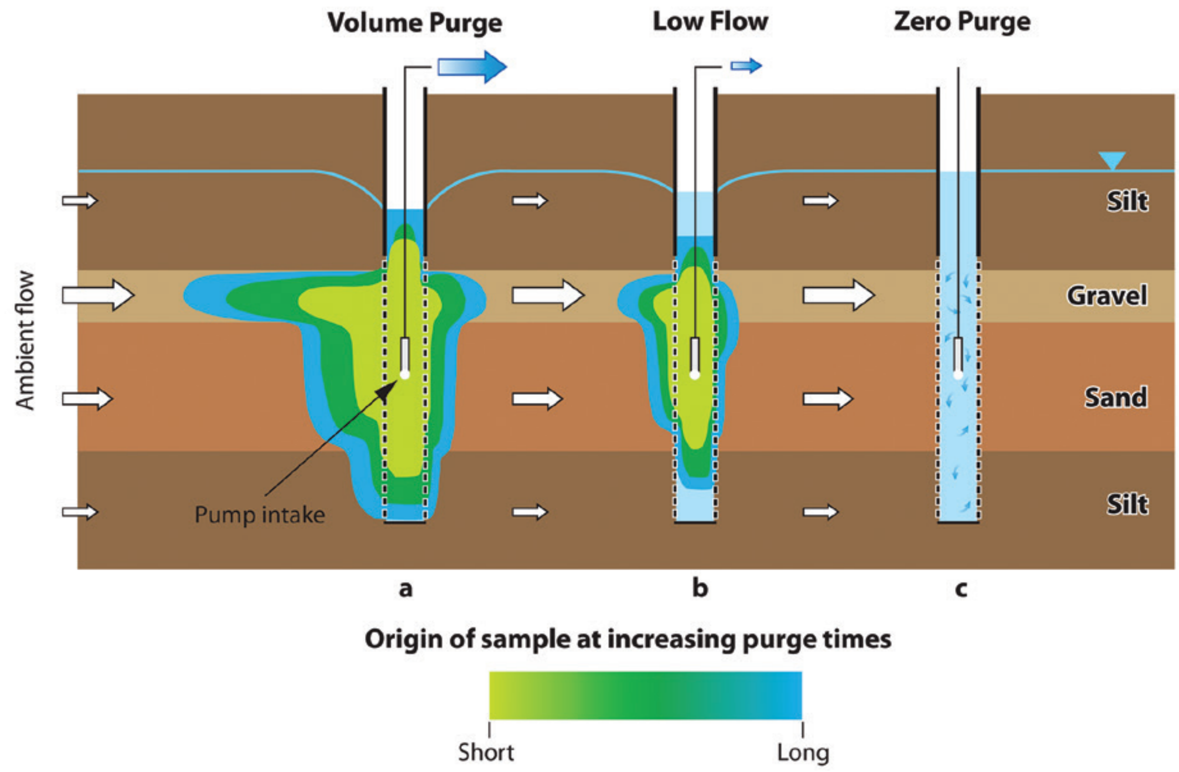

Figure 1: Groundwater sampling methodologies and flow-weighted averaging.

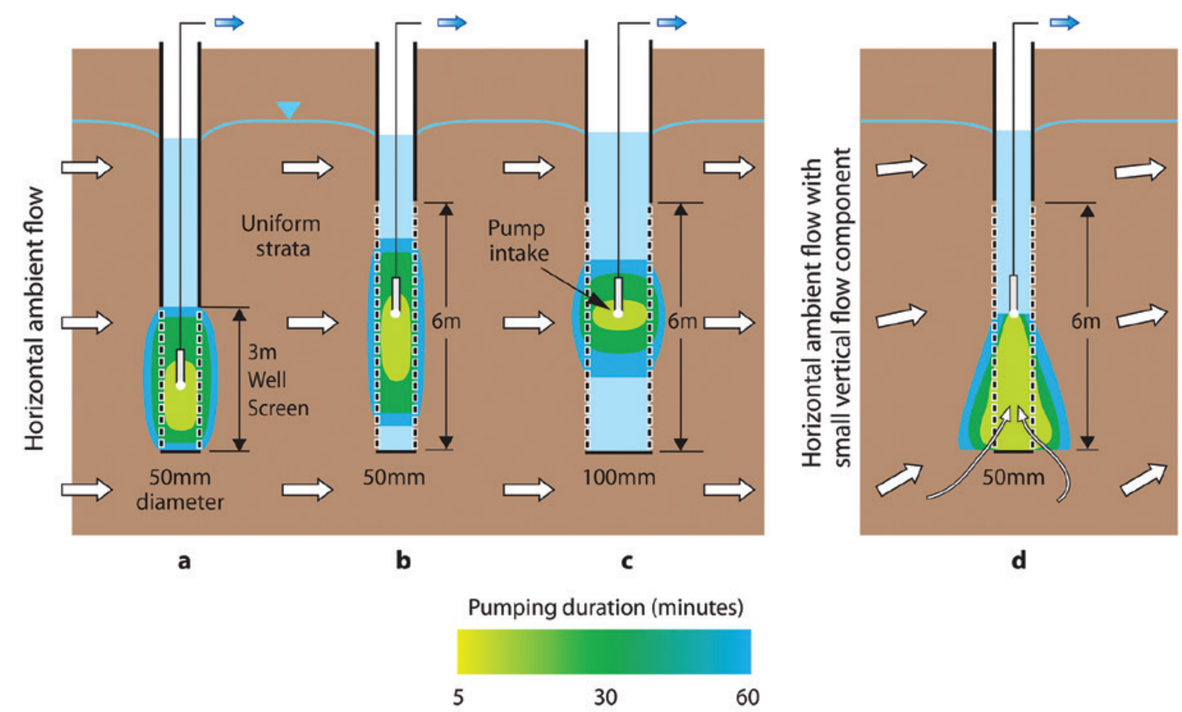

Figure 2: How increasing screen length, diameter and ambient vertical flows influence the origin of sample in a monitoring well – the outer blue line illustrates the zone of inflow to pump after 1 hour of low-flow pumping (adapted from McMillan et al (2014)).

The long and the short of it

We cannot simply abandon the legacy of long-screened wells in routine use for monitoring purposes. Rather, we need to be careful that sampling methodologies used and interpretations applied to the data generated, expresses an appropriate degree of uncertainty which reflects the hydrogeological environment and screen interval in relation to the contaminants of interest.

Groundwater quality sampling at contaminated sites requires particular care and attention to meet sampling objectives. It is vital that sampling is knowledge based, critically reflective and forward thinking in the choice and application of methodologies. Understanding groundwater quality is fundamentally underpinned by understanding groundwater flow. We have a long way yet to go in the development of new instrumentation before flows can be routinely measured alongside water quality sampling. In the meantime we need to be willing to understand and express uncertainty in our interpretations of sampling results – and wherever possible use shorter-screened wells.

Acknowledgements

Research work at Birmingham University referenced in this article includes NERC Open Case Studentship 741 NE/H019170/1 part funded by Waterra-In-Situ (now In-Situ Europe Ltd) and the Environment Agency (for England). This article solely represents the views of the authors.

References

Annable, M.D., Hatfield, K., Cho, J., Klammler, H., Parker, B.L., Cherry, J.A., Rao, P.S., 2005. Field-scale evaluation of the passive flux meter for simultaneous measurement of groundwater and contaminant fluxes. Environ. Sci. Technol. 39(18), 7194-201.

BS ISO 5667-11, 2009. Water quality – Sampling – Part 11: Guidance on sampling of groundwaters.

BS ISO 5667-22, 2010, Water quality – Sampling – Part 22: Guidance on the design and installation of groundwater monitoring points.

NIGLQ (Nuclear Industry Group for Land Quality), 2015. Nuclear industry code of practice for routine water quality monitoring. Version 1, February 2015. http://www.nda.gov.uk/publication/nuclear-industry-code-of-practice-for-routine-water-quality-monitoring

Martin-Hayden, J.M., Plummer, M., Britt, S.L., 2014. Controls of wellbore flow regimes on pump effluent composition. Ground Water Monit. Rem. 52 (1), 96–104. http://dx.doi.org/10.1111/gwat.12036.

McMillan, L.A., Rivett, M.O., Tellam, J.H., Dumble, P. and Sharp, H., 2014. Influence of vertical flows in wells on groundwater sampling, J. Contam. Hydrol., 169, 50-61. http://dx.doi.org/10.1016/j.jconhyd.2014.05.005.

Rivett, M.O., Dearden, R.A., Wealthall, G.P., 2014. Architecture, persistence and dissolution of a 20 to 45 year old trichloroethene DNAPL source zone. J. Contam. Hydrol., 170, 95-115. http://dx.doi.org/10.1016/j.jconhyd.2014.09.008

US EPA, 2010. Low Stress (Low Flow) Purging and Sampling Procedure for the Collection of Groundwater Samples from Monitoring Wells, EPASOP- GW 001. http://www.epa.gov/region1/lab/qa/pdfs/EQASOP-GW001.pdf.