2015

Hydrophobic organic contaminant transport property heterogeneity in the Borden Aquifer

Richelle M. Allen-King
*University at Buffalo, State University of New York*

Indra Kalinovich
*University at Buffalo, State University of New York*

David F. Dominic
*Wright State University*

Guohui Wang
*University at Buffalo, State University of New York*

Reid Polmanteer
*University at Buffalo, State University of New York*

*See next page for additional authors*

Follow this and additional works at: [http://digitalcommons.unl.edu/conservationsurvey](http://digitalcommons.unl.edu/conservationsurvey)

Part of the [Geology Commons](http://digitalcommons.unl.edu/conservationsurvey), [Geomorphology Commons](http://digitalcommons.unl.edu/conservationsurvey), [Hydrology Commons](http://digitalcommons.unl.edu/conservationsurvey), [Paleontology Commons](http://digitalcommons.unl.edu/conservationsurvey), [Sedimentology Commons](http://digitalcommons.unl.edu/conservationsurvey), [Soil Science Commons](http://digitalcommons.unl.edu/conservationsurvey), and the [Stratigraphy Commons](http://digitalcommons.unl.edu/conservationsurvey)


[http://digitalcommons.unl.edu/conservationsurvey/10](http://digitalcommons.unl.edu/conservationsurvey/10)

This Article is brought to you for free and open access by the Natural Resources, School of at DigitalCommons@University of Nebraska - Lincoln. It has been accepted for inclusion in Conservation and Survey Division by an authorized administrator of DigitalCommons@University of Nebraska - Lincoln.
Hydrophobic organic contaminant transport property heterogeneity in the Borden Aquifer

Richelle M. Allen-King1, Indra Kalinovich1,2,3, David F. Dominic4, Guohui Wang1,5, Reid Polmanteer1,6, and Dana Divine7

1Department of Geology, University at Buffalo, State University of New York, Buffalo, New York, USA, 2Now at Dillon Consulting Limited, Winnipeg, Manitoba, Canada, 3Also at Department of Civil Engineering, University of Manitoba, Winnipeg, Manitoba, Canada, 4Department of Earth and Environmental Sciences, Wright State University, Dayton, Ohio, USA, 5Now at Pacific Northwest National Laboratory, Geosciences Group, Richland, Washington, USA, 6Now at Water for People-Malawi, Penn Yan, New York, USA, 7Conservation and Survey Division, University of Nebraska-Lincoln, Lincoln, Nebraska-Lincoln

Abstract We determined that the spatial heterogeneity in aquifer properties governing the reactive transport of volatile organic contaminants is defined by the arrangement of lithofacies. We measured permeability ($k$) and perchloroethene sorption distribution coefficient ($K_d$) for lithofacies that we delineated for samples from the Canadian Forces Base Borden Aquifer. We compiled existing data and collected 57 new cores to characterize a 30 m section of the aquifer near the test location of Mackay et al. (1986). The $k$ and $K_d$ were measured for samples taken at six elevations from all cores to create a data set consisting of nearly 400 colocated measurements. Through analysis of variance (corrected for multiple comparisons), we determined that the 12 originally mapped lithofacies could be grouped into five relatively distinct chemohydrofacies that capture the variability of both transport properties. The mean of ln $k$ by lithofacies was related to the grain size and the variance was relatively consistent. In contrast, both the mean and variance of ln $K_d$ were greater for more poorly sorted lithofacies, which were also typically more coarse-grained. Half of the aquifer sorption capacity occurred in the three highest-sorbing lithofacies but comprised only 20% of its volume. The model of the aquifer that emerged is that of discontinuous scour-fill deposits of medium sand, generally characterized by greater $K_d$ and $k$, within laterally extensive fine-grained to very fine-grained sands of lower $K_d$ and $k$. Our findings demonstrate the importance of considering source rock composition, transport, and deposition processes when constructing conceptual models of chemohydrofacies.

1. Introduction

Groundwater contaminant transport simulations, used for risk assessment and remediation planning, require information on the spatial distributions of both hydraulic and reactive aquifer properties. A recent National Research Council (NRC) committee estimated that the United States has at least 126,000 sites with residual concentrations of legacy contaminants and an associated estimated clean-up cost of at least $110 billion [National Research Council, 2012]. Among the different programs that they surveyed, chlorinated volatile organic compounds (VOC) were consistently among the contaminants identified most frequently. For heterogeneous sedimentary aquifers, we can better predict contaminant transport and improve remediation if we can quantify the distributions of permeability and VOC reactivity.

In their recent review article, Anderson and Siegel [2013] describe how results from the legacy field research sites have demonstrated the importance of physical, chemical, and biological heterogeneity on contaminant transport in sedimentary aquifers. In the section "Hydrofacies and Preferential Flow Paths," they review the developments that correlated physical hydrologic properties to identifiable sedimentary units. These approaches have been particularly successful in capturing the effects of permeability ($k$) heterogeneity on transport.

Comparatively few studies have evaluated the heterogeneous distribution of reactivity to organic contaminants in sedimentary aquifers, and even fewer have also related reactivity to sedimentary units and their permeability. For hydrophobic and moderately hydrophobic organic contaminants (HOCs), of which the chlorinated VOCs are a subset, sorption is one important aspect of reactivity. Several studies have characterized the spatial variability of HOC sorption in the Canadian Forces Base Borden (Borden) Aquifer and most
of these used perchloroethene (PCE) as a test compound. As examples, both Mackay et al. [1986] and Allen-King et al. [2006] demonstrated that the vertical variability of the PCE sorption distribution coefficient ($K_d$) was significant and that the depth-averaged $K_d$ was relatively more consistent. The latter work also presented permeability measurements for the samples, but neither of these studies related results to sedimentary units. In one prior study, Allen-King et al. [1998] observed that $K_d$ and $k$ differed among sedimentary units, which suggested that defining such units could illuminate the spatial distribution of geochemical reactivity ($K_d$ in this case) and its relation to permeability. However, the sample set for this study was extremely limited, comprising data primarily from a single core.

The HOC $K_d$ in some Quaternary aquifers has been linked to the petrologic composition of the sediments. Barber et al. [1992] and Barber [1994] identified cooccurring trends of greater fraction organic carbon ($f_{oc}$) content and HOC $K_d$ in magnetically separable fractions of samples from the Cape Cod Aquifer, Massachusetts. The distribution of sedimentary rock fragments containing fossil-organic-carbon has been identified as a key to understanding HOC $K_d$ variability in a number of prior studies [Karapanagioti et al., 2000; Kleineidam et al., 1999a; Zhao et al., 2005; Kalinovich et al., 2012]. Comparing several glaciofluvial aquifers, Kleineidam et al. [1999a] found the highest $K_d$, greatest proportion of reactive sedimentary rock fragments, and greatest carbonate contents among gravels. With longer transport distances, carbonate grains are abraded more readily than those dominated by silicate minerals. Thus, Kleineidam et al. [1999a] found that carbonate was depleted and silicate was enriched with decreasing grain size and travel distance, from near-source gravels dominated by carbonate rock fragments, to distal sands dominated by silicate mineral grains. These workers also noted that carbonate minerals were greater in the silt and clay grain sizes, compared to the sands, because individual calcite crystals are generally smaller than quartz crystals. In a related study, Kleineidam et al. [1999b] also showed that the $K_d$ was dependent upon the carbonaceous matter form and degree of thermal alteration (or maturation), as well as the $f_{oc}$.

These previous studies illustrate that for sedimentary aquifers containing sedimentary rock fragments as the primary source of organic matter, the unit-specific HOC $K_d$ depends on the source rock properties (their $K_d$) and the sediment transport processes that sort and arrange these materials into sedimentary units. In this context, we use the term lithofacies to mean sedimentary units that are differentiable by both their reactivity and their permeability.

Prior work on the Borden Aquifer has identified findings parallel to some of those described above. For example, Ball et al. [1990] showed that grains sieved from a depth-integrated Borden Aquifer sample have a minimum carbonate content in the fine sand fraction (11% as CaCO$_3$) with greater carbonate concentration...
in the coarse sand to fine pebble (73%) and the silt (30%) fractions. Ball and Roberts [1991a] showed that the HOC $K_d$ and $f_{oc}$ are also related to grain size. Ran et al. [2004] showed that the primary HOC sorbent is kerogen, which Kalinovich et al. [2012] showed occurs within calcareous lithocomponents derived from the source sedimentary rocks, including both high-$f_{oc}$ shales and low-$f_{oc}$ limestones.

In order for contaminant transport modeling to be accurate, the complex spatial distributions of $K_d$ and $k$ within an aquifer must be described, understood, and quantified. Two steps are essential to the goal of quantifying spatial distributions using this paradigm: (1) define lithofacies based on sediment characteristics and demonstrate that the resulting $K_d$ distributions for each unit are different from each other (i.e., distinct) for at least some units and that the same is true for the $k$. (2) determine if the property distributions within the sampled volume are representative of the aquifer by analyzing the uniformity of $K_d$ and $k$ distributions for each lithofacies within the sampled region. A third step will increase the utility of such characterization: (3) determine how the spatial distribution of $k$ and $K_d$ can be simplified by grouping lithofacies by their effect on transport (i.e., as chemohydrofacies). This last step seeks to identify the most parsimonious model of $K_d$ and $k$ that preserves the unitwise differences that, in turn, capture transport property variability.

The conceptual model of chemohydrofacies (Figure 1) developed here extends the hydrofacies concept to include differentiating sediment units according to their reactivity, as well as to their permeability. This approach builds on prior work summarized above. This approach is anchored to the concept that in sorting sediment grains into lithofacies, sedimentologic processes also sort grains with respect to their chemical reactivity. Chemohydrofacies are thus units distinguishable using standard sedimentology approaches. The approach is related to, but different from, the concept of multimodal reactive mineral facies developed by Deng et al. [2010]. Both of these conceptual models, and others, assume that contaminant reactivity at the smallest scale is related to physicochemical properties that are quantifiable at the scale of one or a few grains, and that the properties of larger-scale units can be estimated through a knowledge of these smaller-scale properties. The chemohydrofacies model presented here embodies explicit consideration of the geo-logic processes that created the aquifer. Important facets include consideration of the geology of the source rocks and geologic processes that affected them, such as diagenesis, as well as processes such as weathering, transport, and deposition that result in the creation of units differentiable by standard sedimentology approaches.

Following this approach, we determined the relationships of aquifer transport properties ($k$ and $K_d$) to sedimentary architecture, as indicated by lithofacies, for the Borden Aquifer, a Quaternary glacial aquifer that includes Paleozoic marine carbonates and shales among the source rocks. Through a Scheffe analysis, we determined that some units, or groups of units, as defined by lithofacies, are demonstrably different, highlighting the utility of the approach for aquifer characterization. We evaluated the uniformity of the unitwise $k$ and $K_d$ means and standard deviations and concluded that the defined properties for each lithofacies are representative of the aquifer and hence, applicable across it. The Scheffe analysis also allowed us to consolidate lithofacies to develop a simplified chemohydrofacies classification for the aquifer that captures variability in both $K_d$ and $k$ for purposes of reactive transport modeling. In this manuscript, we relate $K_d$ variability among samples within units to specific lithocomponent types and relate this finding to overall aquifer variability. We illustrate how the relationships between $k$, $K_d$, and lithofacies can enable delineation of the spatial distributions of $k$ and $K_d$ through the link to sedimentary architecture and discuss the potential impact of the observed patterns on VOC transport.

The core samples used to develop the data for these analyses were taken from the Borden Aquifer at locations near well-known field experiments that studied the transport of chlorinated VOCs (Figure 2) [e.g., Mackay et al., 1986; Roberts et al., 1986; Thorbjarnarson and Mackay, 1994b]. We used perchloroethene (PCE) as the probe compound for the $K_d$ measurements because PCE was used in these initial tracer studies and it has been included in many subsequent transport studies at Borden [e.g., Mackay et al., 1994; Ptacek and Gillham, 1992; Rivett and Allen-King, 2003; Rivett et al., 1994; Thorbjarnarson and Mackay, 1994a].

The Borden field site has served as a system intermediate in scale between that of very large laboratory experiments and actual (uncontrolled) releases. This scale reveals details that promote an understanding of transport processes [Sudicky and Illman, 2011]. Prior studies that have sought to evaluate the impact of heterogeneity on reactive transport through analysis of the Borden plumes have been limited by lack of information on the spatial distribution of sorption.
The Borden field site has geologic characteristics that are common to many surficial sedimentary aquifers. Aspects of the Borden Aquifer geology relevant to the transferability of our findings are summarized here. The sediments have long been interpreted to be primarily beach and near-shore deposits from glacial Lake Algonquin [Burwasser and Cairns, 1974; Fitzgerald, 1982; Bohla, 1986]. The aquifer is characterized by a low $f_{OC}$ overall, of which a high proportion is kerogen [Ran et al., 2004]. The high deposition rates evident in many of the sandy units and the cold glacial lake and near-shore depositional environment would have resulted in little accumulation of organic matter derived from primary production in the contemporaneous aquatic environment. The presumed aquifer source rocks include Paleozoic formations from the Simcoe Group [Liberty, 1969] that were subjected to extensive diagenesis resulting in kerogen condensation. For example, the Collingwood Member of the Lindsay Formation has been characterized as an oil shale [Barker et al., 1983]. These kerogen-rich rocks are immature, to marginally mature with considerable potential to yield oil [Barker et al., 1983]. Geochemical parameters have pointed to a type II kerogen with minor type I kerogen. Such condensed organic matter has greater resistance to biochemical weathering than recent, labile organic matter [e.g., Keller and Bacon, 1998; Vandenbroucke and Largeau, 2007]. Further, the preservation of kerogen is facilitated by the lower oxygen fluxes below the water table (compared to vadose zone), as well as by its location within porous carbonate grains within which nutrient and microbial transport are limited. Preservation of the kerogen is important to the VOC sorption affinity of the aquifer [Allen-King et al., 2002; Huang et al., 2003; Xiao et al., 2004; Cornelissen et al., 2005].

It is useful to be explicit in delineating how the work presented herein relates to recent publications that explore transport property heterogeneity of the Borden Aquifer, the links to sedimentary architecture, and their impacts on transport. Allen-King et al. [2006] and Ritzi and Allen-King [2007] used portions of a data set containing the same types of measurements as those reported here from cores covering approximately 1.5 m vertical and 10 m horizontal extent. These data comprised detailed measurements in the vertical, but limited horizontal measurements. Allen-King et al. [2006] compared descriptive statistical and empirical semivariograms for $\ln k$ and $\ln K_d$. They determined that the variances were similar for the two properties, and that the vertical ranges for each of the two properties were much lower than their horizontal ranges. Ramanathan et al. [2010] and Ritzi et al. [2013] used portions of the data set presented herein. These workers and Ritzi and Allen-King [2007], who worked with the earlier data set, found that both the spatial bivariate correlation and the cross correlation of $\ln k$ and $\ln K_d$ are explained by the sedimentary architecture. Ramanathan et al. [2010] further showed how the architecture controlled the nonreactive tracer macrodispersion observed in the original tracer test. Here we give the first presentation of the extensive, complete data sets from which data in the studies by Ramanathan et al. [2010] and Ritzi et al. [2013] were derived. These data are provided with the supporting information. Furthermore, we give the first complete description of the field sampling and laboratory methods used to acquire these data, and provide an analysis of measurement error. The lithologic data are presented using a fundamental classification system which contains a larger number of samples and lithofacies than used in prior work. Here we examine how and, importantly, why the $k$ and particularly the $K_d$ vary as a function of lithology.

## 2. Methods

### 2.1. Field Methods

We have selected a relatively fine characterization scale (cm in the vertical and decimeters in the horizontal) because this is the scale that describes heterogeneity relevant to the relatively small contaminant plumes.
and limited transport distances of the original experiments. Core samples collected for use in this project infill and substantially extend the core transect collected previously by Allen-King et al. [2006] (Figure 2) near the field site of the Stanford-Waterloo natural gradient tracer test, as described by Mackay et al. [1986]. The purposes of collecting additional core samples were to provide both greater lateral extent and more refined spacing needed to identify the horizontal spatial patterns in the aquifer properties. We also use data from the earlier coring campaign in this work and refer to the Allen-King et al. [2006] work as “original” data or cores throughout.

Variable core spacing was used to facilitate development and testing of statistical models to describe horizontal heterogeneity while managing the time required for detailed data collection. The core spacing was different in each of the three approximately 10 m horizontal sections: 0.5 m in the south section (cores 5.0–15.0°N), 0.25 m in the central section (cores 15.5–25.25°N), and 1 m in the north section (cores 25.5–35.5°N) (Figure 3). The cores were numbered by northing. The core collection method was the same as in our prior work at Borden [Allen-King et al., 1998; Allen-King et al., 2006] in which cores were collected without a drill rig in 5 cm diameter aluminum tubing using the wireline piston coring method of Starr and Ingleton [1992]. In this method, the sediment was held under suction during extraction, and therefore, no catcher was required to obtain high recoveries of undisturbed sediment cores. In order to improve the consistency with which we viewed the sediment both between core locations and within cores, a new measure in this study was to record core rotation during extraction (supporting information Figure S1) and to note and account for alignment at all junctures of core cutting.

Once retrieved, each core was allowed to drain in a vertical position to preserve sediment structure during transport. The 1.5 m cores were cut into three approximately equal length pieces, end gaps (where present)
were wax filled to prevent the sediment from shifting in transit, and the capped core pieces were stored cool (with ice) until transport to the laboratory at University at Buffalo, where they were stored at 4°C (to inhibit biogeochemical reactions) until use. Prior to mapping and sampling, an approximately 90° arc of the aluminum tube was removed and an image of the sediment was recorded by digitally scanning.

### 2.2. Lithofacies Mapping and Indicator Data Set Generation

The lithofacies code used to map the sedimentary units for this work expanded on the code developed in prior work [Allen-King et al., 1998; Divine, 2002]. Table 1 includes lithofacies properties. We adopted a systematic nomenclature based on the style of the Unified Soil Classification System [ASTM, 2011]. The nomenclature captures the following information (Table 1): grain size is indicated by the capital letter and subscript; structure is indicated by the next lower case letter; and the final subscript, if present, identifies multiple units of the same grain size and structure type but distinguished by other characteristics. Because this nomenclature differs from that used in prior work, supporting information Table S1 clarifies the relationships among lithofacies designations. In the present work, we found that it was necessary to define a unit type, DS, to represent portions of cores thoroughly disturbed (mixed over cm) by the coring process. These typically occurred at the core ends. In our analyses, this unit type is treated separately from the 12, naturally occurring lithofacies.

A database representing the lithofacies occurrences in the 72 cores (21, 40, and 11 in each of the south, central, and north sections) was created by sampling the cores with a 0.01 m vertical spacing and coding an integer indicator for the lithofacies type present at each location. This gave 10,568 sample locations, and the indicator data from those locations are referred to as the exhaustive indicator data set hereafter (provided in supporting information Table S2).

### 2.3. k and PCE $K_d$ Measurements

Six elevation horizons were identified for $k$ and $K_d$ sample collection (Figure 3a). The samples for $k$ measurements were taken in 1.3 cm diameter stainless steel tubes perpendicular to the core axis (parallel to ground...
surface). Surrounding material from the same depth interval was removed and dried at low temperature (40°C) for the $K_d$ measurement. Five of the cores were reserved for continuous vertical determinations of $k$ and $K_d$ for a companion project. Therefore, samples from 67 cores were used for $k$ and $K_d$ measurements in this study.

The $k$ measurements were completed on samples dried at low temperature (40°C under N$_2$ flow) using an air minipermeameter following the methods established by Davis et al. [1994] and used in our prior work [Allen-King et al., 2006]. The reasons for missing samples include problems such as low core recovery in the field and lost measurements (e.g., $k$ sample fell out of tube before measurement, broken $K_d$ batch vial, etc.). In a few cases, a lost $k$ sample was replaced by the measurement of an adjacent sample from the same lithofacies. Comparisons of such samples showed that the $k$ were very similar (data not shown). The various occurrences of lost samples resulted in a lower number of colocated measurements compared to the total sample results for either parameter individually. The data set includes measurements for 396 $k$ samples and 395 $K_d$ samples, and 392 colocated measurements of a possible 402 ($6 \times 67$).

The PCE $K_d$ measurements followed the general batch system approach described in our prior work [Allen-King et al., 2006] with minor modifications to improve our ability to quantify measurement uncertainty and to reduce uncertainty for low $K_d$ measurements. In brief, each dried sample was divided into representative halves using a riffle splitter. In order to more rapidly determine the equilibrium $K_d$, the samples were pulverized to a fine powder prior to conducting the measurements. Ball and Roberts [1991a] have shown that pulverization facilitates more rapid equilibration without affecting the observed $K_d$ magnitude for PCE in Borden sediments. The sorbed PCE concentration was determined by difference between the PCE mass added to the system and the mass remaining in the water and headspace following a 3 day equilibration period and the reported $K_d$ is the sorbed-to-solution concentration ratio. Because a consistent PCE mass was added to all batch systems and the sorbed mass varied, the equilibrium aqueous concentration also varied over a relatively limited concentration range. The average and range of the equilibrium aqueous concentrations were 503 µg L$^{-1}$ and 56–734 µg L$^{-1}$, respectively. The bulk Borden control sample was a large, well-mixed sample with an average aqueous concentration of 514 µg L$^{-1}$ and range of 459–582 µg L$^{-1}$.

In this study, we reduced the uncertainty of low $K_d$ measurements by using a greater sediment-to-water ratio compared to our prior work [Allen-King et al., 2006]; 10 mL (nominal) ampoules received 17.7 g sediment and 6.4 mL water for the majority of measurements. Samples were analyzed in 19 experiment sets of approximately 20–25 samples each. The exceptions to the standard method described were experiment sets 1 and 19.5, which were completed in 5 mL ampoules. Experiment set 1 was completed using the batch system sediment and water amounts per Allen-King et al. [2006]. Experiment set 19.5 included reanalyses of high-sorbing samples or samples for which very limited sample mass was available (because of previously broken vials and the like). In either case, a lower sediment-to-water ratio was used. Therefore, set 19 included both 10 and 5 mL ampoule systems (designated as experiment sets 19.1 and 19.5, respectively).

While reducing the uncertainty on the low $K_d$ measurements, the method modification introduced a potential source of bias between samples measured in 5 mL ampoules in the original data set (integer numbered cores 5–15, Allen-King et al. [2006], and experiment set 1) and the rest of the samples reported in this study, measured in 10 mL ampoules (with exceptions noted above). Method-related bias was evaluated by comparing In $K_d$ by lithofacies for sample results of the original data set to sample results from the new cores, all taken from the south section. The result of this analysis is that measurements for two lithofacies (Ssp and Sfp, definitions in Table 1) from the original data set are not used in the subsequent statistical data analyses, while samples of all other lithofacies in the original data set are used. Additional description of these statistical comparisons and their results are provided in supporting information Table S5 and Figure S4. Removal of these $K_d$ results is another cause of the difference between the total number and location of $k$ and $K_d$ measurements used in our analyses.

Prior work on the Borden Aquifer has indicated that coarser grains have a higher $K_d$ and $f_{eq}$ than the aquifer as a whole [Ball and Roberts, 1991a]. Ball and Roberts [1991a] also observed trends of greater $K_d$ and $f_{eq}$ with greater grain size. Specifically, the $K_d$ of coarse to very-coarse sand sized grains was over 20 times greater than the $K_d$ of fine to very-fine sand grains. This led to the concern that the sample splitting procedure could yield inconsistent results for $K_d$ subsamples containing large grains. In order to maintain a relatively efficient sample preparation method (pulverization is time consuming) while also controlling the effect of
coarse grains on the observed $K_d$, we chose a pragmatic threshold of 2 mm (very coarse sand) and treated samples containing grains $\geq 2$ mm differently. A low proportion of the total number of samples contained grains $\geq 2$ mm diameter. The following procedure was developed to insure representation of these grains on the observed $K_d$. The sample was sieved into fractions containing $\geq 2$ mm and $< 2$ mm grains and the proportions of each of the fractions in the total sample were determined. The $< 2$ mm fraction was riffle split and a portion was pulverized for the sorption measurement: in other words, using the same procedure applied to samples with no $\geq 2$ mm grains. The $\geq 2$ mm fraction was pulverized in its entirety: pulverized material of the two fractions was recombined in the proportions determined for the sample.

Because the study is focused on determining the variability of system properties, we instituted several different analytical controls to allow us to quantify variability introduced by our measurement procedures. These controls allow us to distinguish sources of variability. Each experiment set included two to five duplicate measurements of samples from a previous experiment set, and three bulk Borden samples. Duplicate samples were prepared (pulverized and filled into the ampoule) from the reserved material completely independently from the initial sample preparation. Samples for duplicate analyses were originally selected at random. As the analyses progressed, samples for duplication were chosen with the purpose of representing a range of $K_d$ values. Per the standard procedures employed in our prior work, we also prepared and analyzed soil-free control vials and independently measured the initial PCE mass added to the samples for each set of experiment vials.

The analytical uncertainty on the $K_d$ measurements was low, as determined from the bulk Borden control sample results and error propagation analysis of individual samples [Allen-King et al., 2006]. The bulk Borden control sample demonstrated the high reproducibility of the method between experiment sets; the average $K_d \pm 1 \sigma$ was $0.29 \pm 0.03$ L kg$^{-1}$ over all experiment sets. The standard deviation of triplicate measurements on a single day was $\leq 0.03$ L kg$^{-1}$ for all except one experiment set, in which $1 \sigma = 0.04$ L kg$^{-1}$ (supporting information Figure S3). The standard error estimated for individual sample $K_d$ by propagation of measurement errors was $\leq 10%$ for most samples, and less than 15% overall. These calculations are illustrated using data for experiment set 12 that presented a particularly wide range of $K_d$ values (supporting information Figure S2).

### 3. Results

#### 3.1. Lithofacies and Their Spatial Patterns

The conceptual model of the aquifer is of scour-fill deposits that are coarser grained within an otherwise more monotonous field of finer sands. The scours likely represent channelized flows within the nearshore environment of a glacial lake. This character is illustrated in the cross section shown in Figure 3a. The bases of medium sand lithofacies define concave-upward surfaces that reflect erosion into the underlying, finer units before deposition of the overlying, coarser lithofacies. Above these scours (marked by yellow lines in Figure 3a), the deposits generally fine upward, as would be expected of depositional events that begin with erosion into the previously deposited sediment but then decrease in strength, allowing progressively finer grained sediments to deposit. The most clearly defined of these scours are approximately 4–7 m wide and likely indicate that paleoflow direction was largely perpendicular to the cross section. Less-well-defined examples likely reflect scours of similar shape and size, but with paleoflow direction oriented nearly parallel with the cross section. Scours extend over the southern half of the cross section at lower elevations (218–219 m above sea level; MASL) but over the entire section at higher elevations (219–220 MASL). The deposits above these scour surfaces are referred to herein as scour-fill deposits.

One lithofacies, $S_{np}$, extends discontinuously from approximately 16.75° N to the north end of the transect. Vertically, it extends from 219.3 MASL to near the tops of cores and is commonly 20 or more cm thick. Its vertical position and association with massive facies ($S_{mm}$ and $S_{mp}$) suggest that it was created from a primarily medium sand lithofacies in cores that were insufficiently drained in the field so that the uppermost portions were "mixed" during transport to the laboratory. This resulted in a mottled appearance caused by fine-grained swirls highlighted by dark minerals (potential former laminae) and patches within an overall medium-grained unit. Like DS, it may not be a naturally occurring lithofacies. However, it is extensive and its exact origin is unclear. Therefore, we treat it as one of the 12 lithofacies for our analyses.
The lithofacies proportions for the complete (exhaustively sampled) indicator data set and for each section are provided in Table 2. Because of low sample numbers and similar description, the results for the Smx2 and Smx3 lithofacies were combined with the Smx lithofacies and the Sfx2 was combined with Sfx. (These samples are identified in supporting information Table S3.) Thus, our analyses are carried out for 10 lithofacies. The DS data are retained in the overall data set but are not used in the analyses.

The samples immediately above concave-upward scour surfaces showed differences in $K_d$ when compared to the rest of the lithofacies in the aquifer. The $K_d$ results show that samples with greater magnitude ($C_21/[C_21]$ approximately 0.6 L kg$^{-1}$) are often clustered within scour-fill deposits of predominantly medium sand grain sizes (Figure 4). For example, elevations 218.4 MASL and 218.6 MASL both have high $K_d$ areas near northing 15–18m that correspond to the scour with a thalweg apparent at this northing (Figure 3). A more laterally extensive and thin scour is sampled at elevation 218.8 MASL, approximately 15–20 m north. Juxtaposed to these, the monotonous behavior at 219.0 MASL (with one prominent exception) illustrates the relative uniformity and lateral continuity of the fine-grained to very fine-grained tabular sands.

3.2. Ln $K_d$ and Ln $k$ Cumulative Distribution Functions by Unit

Consistent with our prior work, we use natural-logarithm transformed data to evaluate the relationships between $K_d$ or $k$ and lithofacies. The ln-transformed data are nearer to normal than the data without transformation.

The cumulative distribution functions (CDF) of the ln $k$ for the horizontal data set occur in relatively discrete textural groups; the medium-grained samples (warm colors) are relatively distinct from the fine-grained samples (cool colors) (Figure 5b). The lithofacies exhibit similar variance (with the exception of Silt), are well modeled by the normal distributions shown, and are distinguished primarily by differences in the means.

In contrast, the ln $K_d$ CDFs show differences in both the mean and variance between some

---

**Table 2. The Unit Proportions for the Complete (Exhaustively Sampled Indicator) Data Set Overall and Within Each Section**

<table>
<thead>
<tr>
<th></th>
<th>$S_{mx}$</th>
<th>$S_{mx2}$</th>
<th>$S_{mx3}$</th>
<th>$S_{mp}$</th>
<th>$S_{m}$</th>
<th>$S_{pk}$</th>
<th>$S_{p}$</th>
<th>$S_{pm}$</th>
<th>$S_{pm}$</th>
<th>$Z$</th>
<th>DS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>All Units</strong></td>
<td>0.149</td>
<td>0.044</td>
<td>0.019</td>
<td>0.092</td>
<td>0.073</td>
<td>0.004</td>
<td>0.125</td>
<td>0.013</td>
<td>0.147</td>
<td>0.195</td>
<td>0.088</td>
</tr>
<tr>
<td>South</td>
<td>0.220</td>
<td>-</td>
<td>0.096</td>
<td>0.075</td>
<td>0.005</td>
<td>0.130</td>
<td>-</td>
<td>0.152</td>
<td>0.216</td>
<td>0.091</td>
<td>0.014</td>
</tr>
<tr>
<td>Central</td>
<td>0.084</td>
<td>-</td>
<td>0.097</td>
<td>0.111</td>
<td>0.136</td>
<td>0.158</td>
<td>-</td>
<td>0.109</td>
<td>0.234</td>
<td>0.061</td>
<td>0.011</td>
</tr>
<tr>
<td>North</td>
<td>0.085</td>
<td>-</td>
<td>0.046</td>
<td>0.097</td>
<td>0.150</td>
<td>0.265</td>
<td>-</td>
<td>0.063</td>
<td>0.205</td>
<td>0.081</td>
<td>0.007</td>
</tr>
<tr>
<td>All data</td>
<td>0.124</td>
<td>-</td>
<td>0.090</td>
<td>0.099</td>
<td>0.099</td>
<td>0.164</td>
<td>-</td>
<td>0.115</td>
<td>0.225</td>
<td>0.072</td>
<td>0.011</td>
</tr>
<tr>
<td>Weighted average</td>
<td>0.132</td>
<td>-</td>
<td>0.081</td>
<td>0.094</td>
<td>0.094</td>
<td>0.181</td>
<td>-</td>
<td>0.110</td>
<td>0.219</td>
<td>0.078</td>
<td>0.011</td>
</tr>
</tbody>
</table>

---

*Sand units comprising a small proportion of the total were combined with more prominent units. The $S_{mx}$ includes the $S_{mx2}$ and $S_{mx3}$. The $S_{fp2}$ includes the $S_{fp2}$.*

---

**Figure 4.** The PCE sorption distribution coefficient plotted for each of the six sampling elevations noted on Figure 3.
lithofacies (Figure 5a). Two of the fine-grained lithofacies (Sp, Sp2) have the lowest mean values and are relatively distinct from all other lithofacies. For the other sand lithofacies, significant overlap between samples occurs among the low values regardless of textural classification. The Silt exhibits neither the lowest nor the highest mean in \( K_d \). The descriptive statistics for \( \ln k \) and \( \ln K_d \) by lithofacies are listed in Table 3. Complete listings of \( \ln k \), \( \ln K_d \) and associated indicator data are provided in supporting information Table S3.

The \( \ln \)-transformed \( K_d \) and \( k \) distributions of the sand lithofacies were tested for normality using the entire data set. The \( \ln K_d \) distributions of the Sp, Sm, and Smir lithofacies for the complete data set are significantly non-normal (\( p < 0.05 \), Shapiro-Wilk’s test), while the others are consistent with a normal distribution. The \( \ln k \) distribution of the Sp2 lithofacies is also nonnormal (\( p < 0.05 \), Shapiro-Wilk’s test). Neither the full \( \ln K_d \) nor \( \ln k \) data sets passed tests for homogeneity of variance, albeit most of the normally distributed \( \ln k \) or \( \ln K_d \) data for individual lithofacies do have homogeneous variance. Statistical tests allowing for nonnormality are used to analyze the data, where appropriate.

### 3.2.1. Comparisons of \( \ln K_d \) and \( \ln k \) by Lithofacies Between Sections

It is possible for the \( K_d \) and \( k \) CDFs by lithofacies to be uniform whether or not unit proportions are consistent. In order to determine whether the unitwise \( \ln k \) and \( \ln K_d \) CDFs are representative of the entire region sampled, we evaluated uniformity of the means for each lithofacies property with the null hypothesis that the means of the \( \ln \)-transformed data are not different between the south, central, and north sections. We also compared the standard deviations of each lithofacies property CDF between sections. Statistical tests were completed using PASW Statistics 18. We note that the aquifer area represented by each section (approximately 10 \( \times \) 1.5 m) is consistent among the sections and that sample spacing is approximately constant within each section, although different between sections. The univariate statistics by lithofacies for each section are listed in supporting information Table S4.

The individual sand lithofacies do not exhibit significantly different \( \ln K_d \) means between sections, with \( p = 0.05 \) as the criterion below which differences are considered significant (Table 4). Additionally, although the \( \ln K_d \) standard deviations of the sand lithofacies range from 0.3 to 0.9, the range for the sections of is \( \leq 0.2 \) for most lithofacies. These findings are consistent with and support the hypothesis that the properties of the lithofacies are uniform. While three of the lithofacies (Sp, Sm, Spm) exhibit differences in the means that cannot be considered significant at any reasonable confidence level, five of the lithofacies (Sp, Sp2, Sm, Smir, Smx) would be considered different if \( p = 0.11 \) was selected as the criterion for significance, leading to some ambiguity. However, upon further examination, we determined that for four of these lithofacies (Sp, Sm, Smir, Smx) and Spx), the “potentially significant” \( \ln K_d \) differences are subtle and attributable to a combination of low sample numbers, restricted sampling locations, and high within-unit variability. This examination leads us to judge that the differences in the means for these lithofacies are unimportant (not significant) with respect to the hypothesis test. The differences in the Sm lithofacies are discussed with the \( \ln k \) results in the following paragraphs. Additional information and the CDFs by section for these lithofacies are provided in supporting information Figures S5 and S7.

![Figure 5.](image-url)
Significant differences between the mean ln k by section were identified for two of the fine sand lithofacies (Sfx and Sm; Table 4b; supporting information Figure S6). Both of these lithofacies also had potentially significant differences in the means for ln Kd, and are discussed in the following paragraphs. The standard deviations of the ln k range from 0.3 to 0.6, less than observed for the ln Kd, and the range between the sections is <0.2 for most of the fine sands (Sfx excepted) and <0.3 for the medium sands.

The Sm lithofacies is represented by only 17 samples that occurred in very different contexts in the south compared to the central section. Unlike the CDFs by section for ln Kd, the CDFs by section for Sm ln k are completely distinct from one another (supporting information Figure S6b). The difference in the mean value is 0.5 ln k units and the greatest value in the central section is lower than the lowest value in the south section. Although not as distinct, the majority of the ln Kd CDF for this lithofacies occurs at greater values for the south compared to the central section (supporting information Figure S5e). Examination of the sample contexts is informative for this lithofacies. In the south section, the samples are at 219.5 m elevation and are associated with a scour fill. In the central section, most Sm samples were collected from the two lowest elevations from representatives that were interleaved between Sfx and Sfp2 lithofacies. In hindsight, we interpret that what was classified as Sm based on similar characteristics at the core scale should have been identified as two different lithofacies. Hence, the Sm lithofacies provides an illustrative counterpoint to the units for which the differences between sections are not significant. The other lithofacies appear to have relatively consistent means independent of section based on the original definitions while the Sm differs between sections because it represents a different lithofacies in the two sections in which it occurred. These observations suggest that dividing the Sm into two distinct lithofacies according to their associations would produce more consistent univariate statistics. However, because the Sm lithofacies represents a small proportion of the total sampled field, we retained this unit as originally defined for subsequent analysis.

The ln k CDF of the Sfx lithofacies of the south section differed from those of the north and central sections primarily because of a few high samples in the south section. This finding for ln k does not support uniformity of the unitwise properties. No patterns were identified by investigating potential commonalities among these highest samples (e.g., elevation, core number, etc). Additionally, there were no trends in ln k identified by examining samples by elevation in the same section or by section at the same elevation (supporting information Figure S7). Therefore, the reason(s) for the subtle, but significant, differences in the ln k mean and standard deviation for this lithofacies are not known.
ences identified for ln $K_d$ of facies (Sfx) with a significant difference identified for one property (ln $k$ in the data set overall. We created a minimally biased sample set (termed the "unbiased" data set hereafter) to use in the Scheffe analysis. Samples were selected for the unbiased data set with the intention of approximating the sample proportions for each lithofacies by section in the exhaustively sampled indicator data set (Table 2), which is our best estimator of the true lithofacies proportions. The following general procedure was followed to select samples for each lithofacies. For each unit, all samples were retained from the section in which the sample availability was most limiting. The other sections were randomly subsampled. The Sfx lithofacies provides an illustration. This lithofacies comprises 0.125, 0.147, and 0.221 of the section in which the sample availability was most limiting. The other sections were randomly sub-

Table 4. Results of Statistical Tests Comparing ln $K_d$ (a) or ln $k$ (b) by Lithofacies and Section

<table>
<thead>
<tr>
<th>Unit</th>
<th>South</th>
<th>Central</th>
<th>North</th>
<th>Average*</th>
<th>Standard Deviation*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(a) ln $K_d$</td>
<td></td>
</tr>
<tr>
<td>Sfx</td>
<td>22</td>
<td>16</td>
<td>3</td>
<td>−1.1</td>
<td>0.6</td>
</tr>
<tr>
<td>Sfp</td>
<td>9</td>
<td>23</td>
<td>4</td>
<td>−0.8</td>
<td>0.7</td>
</tr>
<tr>
<td>Sffm</td>
<td>10</td>
<td>28</td>
<td>4</td>
<td>−0.2</td>
<td>0.8</td>
</tr>
<tr>
<td>Smr</td>
<td>18</td>
<td>42</td>
<td>21</td>
<td>−1.2</td>
<td>0.3</td>
</tr>
<tr>
<td>Sfp</td>
<td>11</td>
<td>14</td>
<td></td>
<td>−1.9</td>
<td>0.3</td>
</tr>
<tr>
<td>Smr</td>
<td>15</td>
<td>46</td>
<td></td>
<td>−1.7</td>
<td>0.4</td>
</tr>
<tr>
<td>Smx</td>
<td>7</td>
<td>10</td>
<td>2</td>
<td>−0.7</td>
<td>0.9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(b) ln $k$</td>
<td></td>
</tr>
<tr>
<td>Sfx</td>
<td>22</td>
<td>16</td>
<td>3</td>
<td>−14.6</td>
<td>0.7</td>
</tr>
<tr>
<td>Sfp</td>
<td>9</td>
<td>24</td>
<td>7</td>
<td>−14.7</td>
<td>0.4</td>
</tr>
<tr>
<td>Sffm</td>
<td>10</td>
<td>28</td>
<td></td>
<td>−14.4</td>
<td>0.7</td>
</tr>
<tr>
<td>Smr</td>
<td>18</td>
<td>42</td>
<td>22</td>
<td>−15.2</td>
<td>0.5</td>
</tr>
<tr>
<td>Smr</td>
<td>25</td>
<td>13</td>
<td></td>
<td>−15.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Smx</td>
<td>7</td>
<td>10</td>
<td></td>
<td>−14.8</td>
<td>0.3</td>
</tr>
</tbody>
</table>

*Means reported only for sections used for comparisons. If no mean is reported, either there were too few samples to include in the contrast or the facies was not observed in the section. The latter is the case only for Sffm.

Overall, the finding that most lithofacies do not show significant differences in the means or differences in the standard deviations between sections for either property supports the hypothesis that the lithofacies properties identified are uniform (consistent in space) and representative of the sampled aquifer region. This condition is necessary, albeit not sufficient, to support the hypothesis. Uniform properties in lithofacies across the transect suggest that the properties are also representative of the observed aquifer, even though the lithofacies proportions are not stationary. Only two lithofacies for one property (ln $K_d$) differ significantly by section, and of these, our subsequent analysis indicates a misclassification for one lithofacies (Sfm). In other words, reclassification (after the fact) would remove this apparent difference. This leaves only one lithofacies (Sfx) with a significant difference identified for one property (ln $k$). The potentially significant differences identified for ln $K_d$ for four of the lithofacies were determined to be a consequence of limitations of the sampling scheme and are considered not significant. While it is frustrating that this relatively large data set is not less ambiguous on this point, these results underscore the need to employ sampling approaches that account for both random and nonrandom within unit variability (e.g., grain size, grain composition, deposition sequence, depth).

3.2.2. Minimally Biased Data Set Versus Full Data Set

The different sampling densities within each section create the potential for bias in the CDFs for ln $K_d$ and ln $k$ in the data set overall. We created a minimally biased sample set (termed the "unbiased" data set hereafter) to use in the Scheffe analysis. Samples were selected for the unbiased data set with the intention of approximating the sample proportions for each lithofacies by section in the exhaustively sampled indicator data set (Table 2), which is our best estimator of the true lithofacies proportions. The following general procedure was followed to select samples for each lithofacies. For each unit, all samples were retained from the section in which the sample availability was most limiting. The other sections were randomly subsampled. The Sfx lithofacies provides an illustration. This lithofacies comprises 0.125, 0.147, and 0.221 of the indicator samples within each of the south, central, and north sections, respectively (Table 2). Because each of the sections is of approximately equal size, these proportions were translated to the following desired sample percentages for each of the south, central, and north sections: 25%, 30%, and 45%, respectively. Samples were drawn randomly from those available for each of the south and central section and all 21 available samples from the north section were included (supporting information Table S4a). The unbiased Sfx sample set includes 23%, 30%, and 48% samples from the south, central, and north sections. The representation of each section in the unbiased Sfx sample set is similar to, thus representative of, its occurrence within each section in the indicator data. This approach was modified to include: samples from the south...
and center sections only for lithofacies in which the sample occurrence in the northern section was very low (Smx, Smm, Sp, Sm), samples from the central and north sections only for the Smi lithofacies because there were no Smi samples in the south section, and all the Z samples because of its low numbers overall. The same samples are used for the \( k \) and \( K_d \) unbiased data sets (Tables 3b and supporting information S4). Because many Sp and Sp2 samples were removed from the south section (as described in the Methods), the unbiased sample set overall has approximately 6% overrepresentation of the medium-grained and underrepresentation of the fine-grained samples compared to their weighted average proportions in the indicator data set. In the following sections, analyses of the \( \ln k \) and \( \ln K_d \) data by lithofacies use the unbiased data set. We also use the unbiased data set to represent sorption in the total sampled region of the aquifer. The \( \ln K_d \) and \( \ln k \) CDF and univariate descriptive statistics for most lithofacies are similar between the unbiased and overall data sets (compare Figures 5 and 6; supporting information Table S4). The exception among the \( \ln K_d \) is the shift of the Sfp and Sfp2 to lower values in the unbiased data set with removal of the samples from the original data set (see Methods). The \( \ln k \) data of the Sfm lithofacies shifts to greater magnitude (\(~0.2\) \( \ln \) units) with reduced skew in the unbiased compared to the full data set.

The general consistency between the \( \ln k \) and \( \ln K_d \) univariate statistics between the total and the unbiased data sets (supporting information Table S4) also supports the hypothesis that lithofacies properties are uniform. The aquifer area of each section is approximately consistent, and the sampling frequency is constant within each. Plotting the full data set is approximately equivalent to combining the CDFs with weighting proportional to sampling frequency. Therefore, uniform property distributions are expected to produce consistent results regardless of weighting scheme, provided the sample set from each section is representative (e.g., no data clustering effects, sufficient sample numbers, etc.). The important finding from comparison of the unbiased and overall data sets is that (with the exception of the Sfm), the lithofacies property statistics are qualitatively consistent independent of the sampling frequency or lithofacies proportions.

3.3. Characteristics That Differentiate \( \ln k \) and \( \ln K_d \) Within and Between Units

3.3.1. Development of Chemohydrofacies

Groups of lithofacies that exhibit relatively distinct \( \ln K_d \) and \( \ln k \) were identified with two intended purposes. The first determined sediment characteristics that correlate with and/or control the \( \ln k \) and \( \ln K_d \) patterns. The second defined a parsimonious classification system that captured differences in both \( k \) and \( K_d \). A Scheffe analysis was used to accommodate the nonnormality of some lithofacies distributions and because the variance of the data set as a whole is inhomogeneous for each property. The unbiased data set was used for this analysis with the Silt excluded because of its low sample numbers and obviously distinct character compared to the more subtle differences among the sand lithofacies. Thus, this analysis focuses on the more common and well-studied sand lithofacies.

The patterns exhibited by \( \ln k \) are relatively simple (Table 5). The \( \ln k \) can be divided into three relatively distinct groups; these are the fine sands excluding Sm, medium sands, and Sm, which is different from...
The duplicate $k$ measurements provided unexpected insights on the distribution and type of lithocomponents causing $K_d$ variability within lithofacies. A brief method review highlighting preparation differences...
between duplicates generally corresponded to sample pairs with greater \( K_d \). These differences are much greater than the analytical uncertainties, as indicated either by standard error propagation techniques (usually \( 0.05 \times K_d \approx 0.1 \times K_d \), supporting information Figure S2) or by replicated Borden reference sample observations (\( 1 \sigma \approx 0.1 \times K_d \), supporting information Figure S3). The contrast between the duplicate results and the bulk Borden results underscores the finding that high duplicate variability is not a product of analytical error, but a finding about the source of variability within lithofacies.

The differences between duplicate observations must arise primarily from the random distribution of grains within the \(< 2 \) mm fraction of each sample. It is highly likely that the duplicate differences are a consequence of one, or a few, very high \( K_d \) grain(s) falling into opposite duplicate splits. These findings lead to the obvious question, “how high must the \( K_d \) of a single grain be in order to cause the observed difference between the duplicates?” This can be estimated for our systems using equation (1),

\[
(K_d)_d = \frac{(\Delta K_d)_{sd} m_s}{m_d}
\]

where \( (K_d)_d \) is the sorption distribution coefficient of the single grain, and \( m_s \) and \( m_d \) are the masses of the sediment and grain in the batch vial, respectively. A standard value of 17.7 g is used for \( m_s \). The grain is assumed to be spherical and to have the density reported for Borden sand (2.71 g cm\(^{-3}\)) [Mackay et al., 1986]. A \( \Delta K_d \) of 0.3–1.6 L kg\(^{-1}\) requires a 2 mm grain with \( (K_d)_d = 500–2000 \) L kg\(^{-1}\). These values are approximately 1000–10,000 times greater than the aquifer average. Although extremely high for PCE, such values are not unprecedented. For example, Binger et al. [1999] reported PCE \( K_d \) of 900 L kg\(^{-1}\) and 2400 L kg\(^{-1}\) for two Paleozoic shale samples from southern Ontario and Grathwohl [1990] reported values of 300 L kg\(^{-1}\) for Tertiary and Jurassic shales from Germany. Moreover, this finding is consistent with carbonate granules sieved from Borden Aquifer samples. Very dark calcareous grains (2–4 mm) were observed to have \( K_d \) values ranging from 60 to 600 L kg\(^{-1}\), whereas the bulk aquifer sample \( K_d \) was observed to be approximately 0.3 L kg\(^{-1}\) [Kalinnovich et al., 2012].

The average duplicate \( \ln(K_d) \) variances of the well-sorted fine sand lithofacies (\( S_p \), \( S_{p2} \), and \( S_x \)) were similar to the units’ sample variances (Figure 8). We deduce that the variance of \( \ln(K_d) \) of these lithofacies is controlled by the essentially random distribution of shale, or shale-like, grains within and among samples. These fine-grained lithofacies rarely contained grains \( \geq 2 \) mm, therefore the duplicates fully captured the subsample variability. Further, the duplicate \( \ln(K_d) \) variances of the other sand lithofacies (medium grained and poorly sorted) are also similar to the range reported for the fine-grained lithofacies (Figure 8). Taken together, the above findings implicate the random distribution of small high sorbing shale or shale-like
grains to account for this level of \( \ln K_d \) variance within the aquifer overall. Because variability among grains greater than 2 mm diameter was controlled, the duplicates only capture variability attributable to sand-sized and smaller-sized grains, regardless of lithofacies.

In contrast to the fine-grained lithofacies, the average duplicate \( \ln K_d \) variances of the poorly sorted lithofacies (Smm, Smp, Smx, and Sfm) were a small proportion of the units’ sample variances (Figure 8). The standard deviations of \( \ln K_d \) for these poorly sorted lithofacies are 0.6–1.4 L kg\(^{-1}\) (compared to <0.2 L kg\(^{-1}\) for the well-sorted units). Differences of this magnitude among samples within a lithofacies could arise either as a consequence of small shale-like grains or because of a larger mass of large diameter grains with \( K_d \) less than the shale-like grains but greater than the average aquifer \( K_d \). Such grains are typified by the light carbonate grains separated from the Borden Aquifer sediments and characterized by Kalinovich et al. [2012].

In summary, small (<2 mm) high sorbing grains having \( K_d \) consistent with shale, appear to control both the magnitude of PCE sorption and its variability in the well-sorted fine-grained lithofacies. Additionally, we conclude that grains with shale-like sorption also contribute to sorption and its variability within the <2 mm fraction of the medium-grained lithofacies. In other words, the findings suggest that the sorption capacity of a significant volume fraction of the aquifer is derived from a minute volume of randomly distributed small grains with “super sorbent” (shale-like) character.

4. Discussion

4.1. Sorting Effects on Physical and Geochemical Heterogeneity

This study illustrates how sedimentary processes affect \( k \) and PCE \( K_d \) heterogeneity. Within the Borden Aquifer, lithologic sorting differences between and within lithofacies control the spatial distribution of \( K_d \). Importantly, sorting exhibits a strong influence on both the magnitude and variance of \( \ln K_d \), but has little impact on the \( \ln k \). These effects are strongest within scour-fill deposits that are zones of high PCE \( K_d \) with the capacity to store and retard PCE to a much greater extent than the aquifer overall (Figure 4). The conceptual model of the aquifer is of scour-fill deposits that are relatively higher and more variably sorbing, and also more permeable, arranged within an otherwise more monotonous field of mostly fine-to-very-fine sands of much lower and relatively uniform PCE \( K_d \) and \( k \).

As noted above, Kleineidam et al. [1999a] found the highest carbonate contents within the lithofacies with the coarsest grain sizes, the lowest carbonate content within lithofacies of sand size, and intermediate carbonate content within lithofacies of silt and clay. Similar patterns exist within the Borden Aquifer [Kalinovich et al., 2012]. The PCE sorption follows a similar pattern because the primary PCE sorbent is kerogen that occurs within porous limestone and shale grains derived from the bedrock source of the sediments. These trends occur because carbonate lithic fragments are softer and more readily degraded by physical erosion than are igneous and metamorphic rock fragments. Thus, carbonate grains are prevalent in those deposits exhibiting the least amount of sorting. In the Borden Aquifer, these are the lithofacies Smm and Smp.

It is important to note that although the mean grain size of these lithofacies is medium sand (0.25–0.50 mm), they contain particles of gravel (>2 mm) [Kalinovich et al., 2012]. The presence of carbonate grains in these lithofacies yields greater mean \( \ln K_d \) values and their distributions within these units leads to greater \( \ln K_d \) variance (i.e., greater geochemical heterogeneity) compared to the fine-grained lithofacies.
The fine sand lithofacies, Sp, Sp2, Sox, are depleted in the softer carbonate grains and contain a higher proportion of relatively inert (towards PCE sorption) monomineralic silicate grains, such as quartz and feldspars. As a consequence, these lithofacies also have lower $f_{\text{oc}}$ [Kalinovich et al., 2012] and lower $K_d$ compared to the other units [Kalinovich et al., 2012].

In the Borden Aquifer, the Silt lithofacies also follow the trend identified by Kleineidam et al. [1999a]. The carbonate content, as observed for a representative Silt sample in prior work by Kalinovich et al. [2012], is greater than that of any of the sand lithofacies and is consistent with accumulation of carbonate particles that have been abraded to sizes finer than sand size (that is, less than 62.5 μm). The mean $K_d$ is comparable to that of the highest sand unit and the $f_{\text{oc}}$ is also within the range observed for the sands [Kalinovich et al., 2012]. Therefore, it is likely that the sorbent in the Silt comprises kerogen-containing particles derived from the same source rocks as the sands. This explanation is consistent with work by Binger et al. [1999], who determined that the sorbent within a thick glacial silt/clay aquitard was consistent with the kerogen in the presumed shale source rocks. Hence, the physical processes that control grain sizes and sorting during transport and deposition also control the proportions and distributions of limestone and shale lithic fragments with high-PCE-affinity among the sand and silt lithofacies.

We do not mean to promote an overly simplistic representation of the heterogeneous aquifer properties. There are counterpoint examples to include. For example, samples of medium grained lithofacies at 219.5 MASL and 219.3 MASL and north of 15 m include adjacent lower $K_d$ samples within the medium-grained lithofacies. There are also regions of high $K_d$ that do not also have high $k$. These results hint at the potential for spatial trends within medium-grained lithofacies. Such trends could arise as a natural consequence of fining upward sequences, such as that shown in Allen-King et al. [1998]. However, such patterns would not be well delineated by this data set because of its emphasis on a horizontal sampling pattern. More rigorous analysis may be achievable with a complementary vertical data set that we have completed for the same field site.

The variability caused by <2 mm high-sorbing grains distributed throughout the aquifer volume provides a physical basis for the nugget apparent in the vertical and horizontal semivariograms observed by Allen-King et al. [2006]; the nugget values are 0.08 and 0.17, respectively. The observed nugget values are encompassed by sampling-related variance as quantified by the duplicate analyses; the average variance of the duplicate $K_d$ pairs is 0.04–0.17. The uncaptured variability occurred at a spatial scale smaller than the sample dimensions (1.5 cm vertical slice of core, 5.1 cm outside diameter). The within-lithofacies spatial distribution of these high-sorbing grains was likely determined by processes that created internal structures, such as laminations. The effects of subsample reactive grain distributions in other aquifers could similarly lead to observation of a nugget for an empirically determined semivariogram. In such a circumstance, the nugget may provide information about the origin of variability at the smallest sampled scale.

4.2. ln $K_d$ Skew and Potential Impacts on Mass Storage and Transport

The primary cause of skew in the overall ln $K_d$ data set is the correlated trend in ln $K_d$ mean and variance by lithofacies. The ln $K_d$ distributions for most of the lithofacies have similar mean and median and do not exhibit pronounced skewness. As shown in Figure 9 (ln transformed), normal distributions exhibiting the correlated means and variances of the lithofacies and their proportions observed in this study lead to an overall ln $K_d$ distribution with an obvious high-valued tail. The ln $k$ distribution for the aquifer, which has less skew, provides a useful contrast to the ln $K_d$ behavior. The ln $k$ distributions of individual lithofacies are also relatively normal with low skewness. The standard deviations of the lithofacies-specific ln $k$ distributions extend over a limited range and lack significant correlation to the means.

The positive skew of the overall ln $K_d$ distribution indicates that a small volume of the aquifer contains a large fraction of the sorbent. Within a region of the aquifer equilibrated with a uniform aqueous PCE concentration ($C$), the mass stored per volume of aquifer, $q_{\text{ph}}$, is directly proportional to the $K_d$ ($= \rho_{b} K_{d} C$), where $q_{i}$ and $K_{d}$ are the sorbed concentration and $K_{d}$ at location $i$, and $\rho_{b}$ is the bulk density. The fraction of total PCE mass in the sorbed phase (termed fractional mass sorbed) can be computed using equation (2),
where \( \phi_i \) is the cumulative fractional mass sorbed for the \( i \)th sample of the ranked list (low to high) and \( n \) is the total number of samples. Assuming that the samples are representative of the aquifer and that bulk density is uniform, the aquifer volume fraction, \( x_i \), is \( i/n \). This approach assumes a data set with samples representative of the proportions and properties of each lithofacies within the portion of the aquifer that is of interest. The unbiased data set is a reasonable choice for this purpose.

Approximately 20% of the aquifer with the greatest \( K_d \) accounts for >50% of the total aquifer sorption or mass storage capacity (Figure 10a). This high PCE affinity portion comprises predominantly \( S_{mm} \), \( S_{mx} \), \( S_{mp} \), with Silt, and some \( S_{mr} \) samples. In order to connect this result more directly to the potential to impact transport, we examine the retardation factor, \( R = (1 + K_d \rho_b / \eta) \), defined for saturated conditions, linear sorption, and local equilibrium. For the Borden Aquifer, the bulk density \( \rho_b \), and porosity, \( \eta \), are 1.81 g cm\(^{-3}\) and 0.33 \( [\ell] \), respectively [Mackay et al., 1986]. Approximately 80% of the aquifer has a relatively low \( K_d \) (\( \leq 0.5 \) L kg\(^{-1}\)), and, if examined separately, would produce a relatively symmetrical histogram with a median \( R \) of approximately 2.4 (range \( \sim 1.5-3.7 \), Figure 10b). In contrast, the upper 20% of the distribution extends to \( R > 20 \) with a median \( R > 6 \) (corresponding to \( \phi_9 = 0.9 \)). These calculations indicate that 20% of the aquifer has the potential to exert a dominant retardation effect on VOC transport.

### 4.3. Implications of \( K_d \) Spatial Distribution on Transport

The observed sorption spatial patterns are consistent with the apparent “macrokinetic” behavior previously observed in the natural gradient tracer test conducted in the 1980s at Borden. Roberts et al. [1986] observed that the PCE \( R \) increased rapidly from 2.7 to 3.9 during the initial 85 days after plume injection, and continued to increase more slowly to 5.9 during the nearly 2 years of subsequent monitoring, with no indication that a maximum had been reached by the end of the test. Similar temporal patterns were observed for the other HOCs that persisted throughout the test. Our results are consistent with sorption heterogeneity contributing greatly to these observed trends. The early monitoring period in the tracer test (\( \sim 85 \) days) produced \( R \) consistent with the \( K_d \) range determined to comprise 80% of the aquifer in our study (Figure 10). The arrangement of medium-grained lithofacies observed could cause an apparent increase in the retardation of a PCE plume with time or distance of travel as the plume encountered successive high-sorbing scour-fill deposits during transport. Our findings are qualitatively consistent with the conclusion that sorption heterogeneity, as dictated by the spatial distributions of the sedimentary units identified and their distinct properties, is causal to the observed field tracer behavior of increasing \( R \) with time and/or distance of transport.
Quantitative evaluation of the heterogeneous $K_d$ patterns observed in this study will require additional study. The results here provide both the impetus and supporting data for such analysis.

Our findings do not exclude the potential for other mechanisms to contribute importantly to the observed “nonideal” field behavior of PCE and the other HOCs, as originally indicated in Roberts et al. [1986]. These ideas have been expanded upon in later works and include nonlinear and competitive sorption, as described by Rivett and Allen-King [2003], and rate limited intragranular diffusion, as described by Ball and Roberts [1991b], Harmon and Roberts [1994], and Ball et al. [1997].

5. Concluding Remarks

This work demonstrates that chemohydrofacies can be distinguished using standard sedimentary characteristics that create distinct populations with different VOC sorption statistics and permeability statistics. The $\ln K_d$ and $\ln k$ means and standard deviations of each lithofacies are relatively uniform (constant) across the sampled transect, although the lithofacies proportions vary. This finding suggests that the defined properties for each lithofacies are representative of and applicable to other occurrences in the aquifer. Because we used more laterally extensive data than presented in prior studies, this work provides a more rigorous test of the approach than previously available.

Within the Borden Aquifer, the sorting differences among chemohydrofacies have a large impact on $\ln K_d$ magnitude and variability and little impact on $\ln k$. This is because the depositional processes that create the finer-grained and better-sorted sand lithofacies also remove the high $K_d$ limestone and shale lithic fragments containing kerogen. The sorting differences skew the CDF for $\ln K_d$ for the aquifer more than for $\ln k$. This result has implications for the correlation between the two and for contaminant transport. We include the entire data set in supporting information to this manuscript to facilitate the exploration of these issues.

We believe our results apply more generally beyond the Borden Aquifer because the geologic characteristics that underlie its heterogeneity are common to many surficial aquifers. Sedimentary rocks containing kerogen are widely distributed at Earth’s surface and many glacial deposits are derived from them. In north-central North America and central Europe, sequences of Paleozoic and Mesozoic marine shales and limestones occur at or near the surface and have provided sediment to many surficial glacial aquifers.

An important conclusion of this study is that VOC sorption is greater in poorly sorted and coarse-grained lithofacies than in well-sorted and fine-grained lithofacies. These findings have important implications for aquifer characterization in support of effective and efficient remediation wherever the primary organic matter sorbent occurs within porous sediment grains. In particular, conceptual models for glacial aquifers should include consideration of source rocks as well as transport and deposition processes.

References


