



|                  |  |
|------------------|--|
| Title            | Heat as a Tracer for Examining Depth-Decaying Permeability in Gravel Deposits  |
| Author(s)        | Sakata, Yoshitaka  |
| Citation         | Groundwater, 53(S1), 21-32<br><a href="https://doi.org/10.1111/gwat.12236">https://doi.org/10.1111/gwat.12236</a>  |
| Issue Date       | 2015-04  |
| Doc URL          | <a href="http://hdl.handle.net/2115/61038">http://hdl.handle.net/2115/61038</a>  |
| Rights           | This is the peer reviewed version of the following article: Groundwater, Volume 53, Issue S1, pages 21–32, April 2015, which has been published in final form at <a href="http://onlinelibrary.wiley.com/doi/10.1111/gwat.12236/abstract">http://onlinelibrary.wiley.com/doi/10.1111/gwat.12236/abstract</a> . |
| Type             | article (author version)   |
| File Information | Final manuscript for HUSCAP.pdf  |



[Instructions for use](#)

# Heat as a tracer for examining depth-decaying permeability in gravel deposits

Yoshitaka Sakata

First affiliation: Division of Earth and Planetary Dynamics, Faculty of Science, Hokkaido University

North-10, West-8, Kita-ku, Sapporo, Hokkaido 060-0810, Japan

Second affiliation: Geotechnical Department, Division of Environmental Engineering, Docon Co., Ltd

1-5 Atsubetsu-chuo, Atsubetsu-ku, Sapporo, Hokkaido 004-8585, Japan

E-mail: yoshitaka\_sakata@mail.sci.hokudai.ac.jp

Telephone/Fax: +81-11-801-1570

The author has no conflicts of interest to disclose.

## Abstract

Depth dependence of permeability can appear in any geologic setting; however, vertical trends in alluvial gravel deposits are poorly understood because of the high variability of hydraulic conductivity  $K$  in monotonic sequences. This paper examines the sensitivity of depth-decaying permeability through heat transport simulation around a river's losing reach in the Toyohira River alluvial fan, Japan. Observed variations in groundwater temperature indicate that heat fluxes are dominant in the shallow zone, despite a vertical hydraulic gradient. In eight cases with different conditions (presence or absence of exponential decay trend, large or small variogram range, and cell isotropy or anisotropy) 1,000  $K$  realizations are stochastically generated throughout a cross-sectional

21 model. The groundwater flow and heat transport are transiently calculated, and the averaged root  
22 mean square error  $\overline{RMSE}$  is used for sensitivity comparison. The variance of  $\overline{RMSE}$  shows that small  
23  $\overline{RMSE}$  realizations are effectively reproduced with vertical trend assumed. Plausible realizations of  
24  $\overline{RMSE}$  below a given threshold were obtained only when a vertical trend was assumed. The most  
25 plausible realization almost completely matched the observations. However, the number of plausible  
26 realizations per case was  $\leq 10$  and the median  $\overline{RMSE}$  were insensitive to all the conditions. Statistical  
27 testing suggested that these plausible realizations may be statistically significant, aiding in generating  
28 a connected  $K$  zone for high heat flows. The cell anisotropy condition had the smallest effect on the  
29 simulation. Thus, effective modeling of the vertical trend contributes to heat transport; however, the  
30 model's efficiency is low without detailed information about the sedimentary structure.

31

## 32 **Introduction**

33 Depth-dependence of the permeability resulting from depositional and post-depositional  
34 processes may be found in any type of geologic formation. Geologic processes such as deltas,  
35 alluvial fans, and fluvial flows create unique patterns of hydrogeologic structure, not only in the

36 lateral direction but also vertically (Freeze and Cherry 1979). In alluvial fans, for example,  
37 increasing coarsening and fining with depth are typically seen as a result of progradation,  
38 retrogradation, and basin subsidence (Neton et al. 1994). On a smaller scale, streambeds may also  
39 exhibit vertical trends in grain size, probably related to hydraulic conductivity  $K$  (Marion et al. 2008;  
40 Chen 2011). Post-depositional processes also affect geological formation: compaction with buried  
41 depth, diagenesis, metamorphism, chemical cementation, and other processes (Domenico and  
42 Schwartz 1990; Ingebritsen et al. 2006). Such permeability decay with depth significantly affects  
43 groundwater flow systems not only on a regional scale (Saar and Manga 2004; Jiang et al. 2009), but  
44 also at multiple scales including hyporheic zones (Zlotnik et al. 2011). Permeability decay with  
45 depth was also discussed in relation to groundwater transport and residence time, as well as flow  
46 (Cardenas and Jiang 2010; Jiang et al. 2010a).

47 In a coarse sediment mixture,  $K$  is sensitive to changes in porosity, as shown in experimental  
48 studies (e.g., Koltermann and Gorelick 1995), laboratory tests (Major et al. 1997), and in-situ  
49 measurements (Yamaguchi and Nakamura 1994). However, the vertical trend of  $K$  has rarely been  
50 determined in alluvial gravel deposits. Multilevel in-situ measurements of  $K$  by pumping tests, slug

51 tests, flow meters and other means are expensive and susceptible to many sources of error such as  
52 skin effects (Butler et al. 1994). Even with many in-situ measurements, the trend would not be  
53 necessarily detected because of the strong heterogeneity of  $K$  on a smaller scale (e.g., Eggleston and  
54 Rojstaczer 1998). The sedimentary structure in alluvial gravel deposits consists of various  
55 components such as clast-supported gravel, matrix-supported gravel, and sand and finer materials  
56 (Rust 1979) and detailed information has only been obtained from a few outcrops and trenches  
57 (Jussel et al. 1994; Zappa et al. 2006). Even when detailed information is obtained, it is not valid to  
58 assume that the sedimentary texture of the upper layers extends to deeper layers. Generally,  
59 groundwater modeling of an unconsolidated gravel aquifer frequently assumes a constant value or  
60 homogeneity of  $K$  in the vertical direction.

61 The purpose of this study is to assess the effect of a decreasing trend of  $K$  on groundwater  
62 flow and transport in a numerical model of a typical monotonic gravelly aquifer. For this purpose,  
63 the author performed a sensitivity analysis of heat transport simulation around a losing river in the  
64 alluvial fan of the Toyohira River, Sapporo, Japan. Heat as a groundwater tracer is applicable in  
65 terms of robustness, convenience, and the low cost of field measurements (Anderson 2005).

66 Groundwater temperature can also be used as an alternative to measuring solute components, owing  
67 to the analogy of the governing equations in groundwater transport (Langevin et al. 2008). In  
68 particular, the thermal properties of rocks and fluids can be specified a priori or may be less sensitive  
69 to heat transport simulation than  $K$  (Anderson 2005), and thus temperature data can be used for  
70 estimating aquifer properties, especially as an inverse problem of  $K$  (e.g., Woodbury and Smith  
71 1988; Wang et al. 1989; Bravo et al. 2002).

72         However, using groundwater temperature as a tracer is not necessarily an advantage relative to  
73 other tracers, especially for inferring a  $K$  pattern. Heat is transferred through the solid as well as the  
74 fluid, such that the thermal diffusivity of heat is several orders of magnitude larger than the  
75 molecular diffusion coefficient of other solute components (Domenico and Schwartz 1998). As a  
76 result, heat may be transferred mainly by conduction, including thermal diffusivity, and will be  
77 insensitive to convection, including groundwater velocity (i.e.,  $K$ ). In other words, measuring  
78 groundwater temperature would be effective in a situation of strong groundwater flow through  
79 relatively thick and permeable layers, for example, around a losing river in an alluvial fan. Several  
80 unique features such as the delay of the peak temperature arrival and the seasonal envelope are seen

81 in temperature variations at observation wells near a losing reach (Constantz 2008).

82 This study focuses on a shallow heat transport system in a recharge zone along the Toyohira  
83 River. Temperature observations show heat transport is dominated near the water table, indicating a  
84 vertical decrease of  $K$ . Here, a Monte Carlo approach is applied to simulate the anisotropic heat  
85 transport system. One thousand realizations of  $K$  were stochastically generated for eight study cases  
86 with a combination of the following three conditions (i.e.,  $2 \times 2 \times 2$ ): exponentially decreasing or  
87 constant vertical trend, large or small variogram of the residuals after detrending, and vertical  
88 isotropy or anisotropy at each cell. A coupled groundwater flow and heat transport model was  
89 constructed as a cross-sectional model of SEAWAT v.4. The average values of the root-mean-square  
90 residuals at the observation points were compared for the eight modeled cases. The sensitivity of the  
91 vertical trend and other heat transport simulation comparisons are discussed in terms of the  
92 probability of occurrence of the small error realizations.

93

## 94 **Materials and methods**

### 95 **Study site**

96 The field site is the Toyohira River alluvial fan in Sapporo, near 43°N, 141°E in Hokkaido,  
97 the northernmost island of Japan (Figure 1). The city of Sapporo is the largest in northern Japan, and  
98 with population currently over 1.9 million and increasing. The alluvial fan extends across ~35 km<sup>2</sup>,  
99 and is surrounded by Tertiary volcanic mountains to the west, hilly lands covered by Pleistocene  
100 volcanic ash to the east, and lowland to the north. The fan comprises a western Holocene fan (gray  
101 tone in Figure 1) 70–10 m above mean sea level (amsl) and an eastern Pleistocene fan (dark gray  
102 tone) 90–20 m asl. The Toyohira River is about 72.5 km long with a watershed area of about 900  
103 km<sup>2</sup>. It flows south to north, almost through the center of the fan with an average inclination of about  
104 7:1000, to the lowland around the Ishikari River (one of the main rivers in Japan). The median  
105 values of the daily mean and drought discharges of this river in the fan were 12.6 and 4.0 m<sup>3</sup>/s,  
106 respectively, during the period 1974–2011. Synoptic discharge surveys (Sakata and Ikeda 2012)  
107 revealed that the losing rate  $Q$  was about 1 m<sup>3</sup>/s between kilometer point (KP) 15.5 and KP 17.0,  
108 where KP is the distance units in kilometers along the river, as shown in Figure 1.  $Q$  was about 80%  
109 of the total pumping rate of the city. The seepage rate per unit meter over distance  $q$  was about 2  
110 m<sup>3</sup>/(m hr). Similar seepage values have been estimated for other losing reaches in Japan. This large

111 infiltration has been problematic for management of surface water use, especially during drought  
112 periods.

113 Figure 1c shows the water table elevation contours (bold lines) and the vertical hydraulic  
114 gradient contours (dashed lines) from a previous study (Sakata and Ikeda 2013a). The water table  
115 gently declines northward in the downstream direction, with horizontal hydraulic gradients in the  
116 order of magnitude of  $10^{-2}$ . The water table rises along the river between KP 15 and 17; this section  
117 includes the distinct losing section. On the other hand, the vertical hydraulic gradients are negative in  
118 the fan area, that is, downward fluxes of groundwater are dominant there because of groundwater  
119 withdrawals. The negative peak also extends to the river between KP 15.5 and 19.0, and the values  
120 in the losing section (K.P. 15.5–17.0) are ranged between  $-0.1$  and  $-0.2$ . The values are about one  
121 order of magnitude larger than the horizontal gradients of the water table. The anisotropy of  
122 hydraulic gradient also appears in the observation wells for this analysis, as described below.

123 Various Quaternary sediments such as gravel, sand, silt, clay, and humus underlie the middle  
124 to distal parts of the Toyohira River alluvial fan. The hydrogeologic basement comprises Tertiary  
125 volcanic formations in the upper fan from KP 18, and Pleistocene marine sediments in the middle

126 and distal fan. A subsurface gravel aquifer above the basement is divided into four sub-aquifers (Hu  
127 et al. 2010): Pleistocene Nos. IV and III, and Holocene Nos. II and I. The lowermost Pleistocene No.  
128 IV consists of alternating gravel, sand, and silt layers – the result of the transition of fluvial, delta,  
129 and shallow marine depositional systems. Both Pleistocene No. III and Holocene No. II aquifers  
130 consist mainly of alluvial fan gravel deposits, i.e., of poorly sorted andesite gravel with less than  
131 several percent fine content. The thickness of the gravel aquifer is over 100 m at the upper middle  
132 part of the Holocene fan, and less than 50 m at the distal part. Holocene No. I also consists of finer  
133 materials and appears only in the northern part of the fan (not in the study area).

134

### 135 **Field observation**

136 For the analysis data, the author conducted field observations of groundwater level and  
137 temperature in one deep well (DW-1) and three shallow wells (SW-1 to 3) along an off-stream  
138 transect at KP 16.6 (Figure 2). Distances from the center of the river channel are 76 m to DW-1, 80  
139 m to SW-1, 140 m to SW-2, and 50 m to SW-3. The stratigraphic column at DW-1 shows a  
140 monotonic gravel sequence (i.e., Aquifers No. III and II) to a depth of about 63 m. A finer layer

141 appears at that depth, continuously lain with a gentle inclination to the east. The finer layer  
 142 corresponds to the top of Pleistocene aquifer IV, and serves as a hydrologic basement of gravel  
 143 deposits in the local area. Direct measurements of  $K$  by slug tests (solid squares at DW-1) might  
 144 show unobvious decrease of  $K$  with depth because of the large variation of  $K$  in the Holocene gravel  
 145 aquifer. Sakata and Ikeda (2013b) estimated the effective hydraulic conductivity  $\bar{K}$  (open squares)  
 146 from the core properties. A moving average analysis (solid line) revealed a vertical trend from the  
 147 surface down to 30 m depth, and was established as an exponential function of depth  $z$  (m):

$$148 \quad \bar{K} = \bar{K}_0 \exp(-Az) = 10^{-2.3} \exp(-0.11z) \quad \text{at } z < z_0 \quad (1)$$

149 where  $\bar{K}_0$  is the average  $\bar{K}$  value on the surface, from DW-1 and another borehole;  $A$  ( $\text{m}^{-1}$ ) is a  
 150 decay exponent; and  $z_0$  is the depth limit of the exponential decay ( $z_0 = 30$  m in the fan). The depth  
 151 was also equal to the boundary between the Holocene No. II and Pleistocene III aquifers (Hu et al.  
 152 2010). This exponential model was used in previous studies (Saar and Manga 2004; Jiang et al.  
 153 2009; Zlotnik et al. 2011; Cardenas and Jiang 2010; Jiang et al. 2010a) and verified in a theoretical  
 154 study (Jiang et al. 2010b). However, the decay exponent in the fan was 10 to 1000 times that in  
 155 consolidated rocks,  $1 \times 10^{-2}$  to  $1 \times 10^{-4} \text{ m}^{-1}$  (Jiang et al. 2009). The exponent decay reduced the

156 value of  $K$  at  $z = 30$  m to only  $1/25$  of its value on the surface. Thus, the previous study suggests that  
157 the large exponent decay may significantly affect the shallow systems of groundwater flow and  
158 transport.

159 The groundwater level and temperature were automatically measured in each well using a  
160 pressure transducer (HOBO U20 Water Level Data Logger U20-001-01; accuracy  $\pm 0.05\%$ ). At deep  
161 well DW-1, the hydraulic head and temperature were observed at 15 m (DW-1-1), and the  
162 groundwater temperature was also observed at two additional depths, 30 m (DW-1-2) and 60 m  
163 (DW-1-3) using a thermal sensor (HOBO Water Temperature Pro v2 Data Logger U22-001;  
164 accuracy  $\pm 0.2^\circ\text{C}$ ). The screen interval of DW-1 was 1 m between 63 and 64 m depth, and those of  
165 the shallow wells SW-1, SW-2, and SW-3 were a few meters below the water table. The set of deep  
166 well DW-1 and shallow well SW-1 (2 m apart) formed a piezometer nest. Observations were  
167 conducted from 10 June 2010 to 28 February 2011 (264 days). The river stage was observed at the  
168 upper gauging station of KP 17.8, and the data were converted to those measured at KP 16.6 using a  
169 linear relationship ( $R^2 = 0.96$ ). All the original data were acquired hourly, but daily average data were  
170 used in the analysis to smooth uncertain fluctuations, including measurement errors.

171 Figure 3 shows the observed water level and temperatures of the groundwater and surface  
172 water. The water table levels in the shallow wells were usually lower than the river stage, but among  
173 them they varied almost in parallel. Sudden rises caused by precipitation were rarely seen during the  
174 observation period, indicating that the groundwater recharge depends mainly on surface water  
175 infiltration within the local area. The water table levels at SW-1 and SW-2 were nearly identical, and  
176 the water table at SW-3 was lower than at other shallow wells, although SW-3 was closer to the  
177 river channel. Thus, the average horizontal hydraulic gradients (dimensionless) from the river were  
178 substantially different, 0.007 at SW-1, 0.004 at SW-2, and 0.018 at SW-3. The hydraulic head at  
179 DW-1 was about 7–8 m lower than the water table at SW-1; the average vertical hydraulic gradient  
180 was  $-0.11$ , indicating a downward flow. The magnitude was about one order larger than that of the  
181 horizontal hydraulic gradient, as shown in Figure 1c. The water-level variation at DW-1 was  
182 relatively smooth, and the seasonal fluctuation probably reflected the water cycle (e.g., reduced  
183 recharge in winter) and the intake of the city (e.g., pumping for cooling in summer).

184 The surface water temperature varied daily and seasonally, reaching a low of  $1.0^{\circ}\text{C}$  on 18  
185 December and a high of  $20.6^{\circ}\text{C}$  on 6 September (Figure 3b). In the well nearest the river, SW-3, the

186 groundwater temperature almost matched the smoothed temperature of the surface water, probably  
187 indicating preferential flows east of the river. The groundwater temperature at SW-1 and SW-2  
188 showed smoother fluctuations and delayed peaks (13.4°C on 28 October at SW-1 and 12.0°C on 2  
189 December at SW-2). At the deep well DW-1, the groundwater temperatures fluctuated in a much  
190 narrower range than at the shallow wells, and the temperature at DW-1-3 at 60 m depth was constant  
191 at 10.3°C, indicating an isothermal layer near the bottom, despite the large hydraulic gradient in the  
192 downward direction. In other words, the shallow heat transfer might be affected by permeability, in  
193 terms of Darcy's law.

194

### 195 **Heat transport model**

196 A simulation of the coupled groundwater flow and heat transport was performed on a cross  
197 section similar to that of the observation transect. The SEAWAT v.4 simulation program (Thorne et  
198 al. 2006; Langevin et al. 2008) in Visual MODFLOW 2009.1 (Schlumberger) was used in this study.  
199 Figure 4 illustrates the grid design and boundary conditions. The total length between SW-2 (left;  
200 west) and SW-3 (right; east) was 190 m (lateral coordinate  $x = -140$  m to 50 m), and the total depths

201 varied between 54 m on the eastern side and 64 on the western side. The vertical grid spacing was 1  
202 m, except for the lowermost layer. The horizontal grid spacing was also 1 m between the river  
203 channel and SW-1 ( $x = 80$  m), and 5 m on the east and west sides. Anderson and Woessner (1992)  
204 recommended that the increase in cell widths in such finite-difference models should be no more  
205 than 1.5 times the original width. However, in this study the grid was designed so that the stochastic  
206 realizations of  $K$  per 1-m regular grid were smoothly input into the numerical model. The inflow  
207 condition was set for the upper cells of the river channel ( $x = -15$  m to 5 m) using the “River  
208 Boundary” function of MODFLOW. The river stage and temperature were variable in time as  
209 shown in Figure 3a (black line), and  $K$  of the streambed was assumed to be  $1 \times 10^{-2}$  m/s, which was  
210 sufficient to allow for water infiltration. The outflow boundaries on both sides were represented  
211 using the “Drainage Boundary” condition of MODFLOW rather than the “Constant Head  
212 Boundary.” This was because the constant head boundary compels unrealistic downward flows from  
213 the groundwater table to the bottom. The drainage elevations varied daily, and were linearly  
214 interpolated from the water table to the west of SW-2 or the east of SW-3 vertically downward  
215 according to the vertical hydraulic gradient between SW-1 and DW-1. No vertical flows and heat

216 fluxes were assumed for the other upper cells based on the assumption that the seepage loss was the  
217 dominant recharge source within the local area. No flow and fluxes were also assumed for the  
218 bottom cells based on the assumption that the permeability of the bottom fine layer was negligible.  
219 The seasonal temperature profiles at DW-1 commonly showed that the temperature gradient was  
220 only about 0.3 °C /100m at the bottom depth (Sakata 2013). The initial conditions of the hydraulic  
221 head and temperature were estimated from maps drawn manually from observation data at the start  
222 date, 10 June 2010.

223

#### 224 **Stochastic realization of hydraulic conductivity**

225 This study generated stochastic realizations of  $K$  following the modeling concept of a vertical  
226 trend on a scale of tens of meters, and small-scale heterogeneity on a 1-m scale. One simple way to  
227 apply this is to detect a trend, simulate the residuals, and then return to the trend (Nowak and Verly  
228 2004). The logarithmic conductivity  $Y$  [ $\log$  (m/s)] was used instead of  $K$ , because  $K$  was assumed to  
229 be log-normal (Domenico and Schwartz 1998; ASCE 2008).  $Y(s)$  at any location  $s$  ( $x, z$ ) in the cross  
230 section is formulated as a sum of the trend component  $M(z)$  and residual component  $R(s)$ :

231  $Y(s) = \log K(s) = M(z) + R(s) \quad (2)$

232  $M(z)$  was assumed a function only of the depth  $z$  (m), and independent of distance  $x$  (m).  $M(z)$  was  
233 assumed a common logarithm of Equation (1), in which  $\overline{K}_0$  was replaced by  $10^{-3.07}$  as the value at  
234 the location of DW-1.  $M(z)$  was also assumed to be constant below the limited depth  $z_0$  as:

235  $M(z) = -3.07 - 0.047z \quad \text{at } z < 30 \text{ m} \quad (3a)$

236  $M(z) = -4.50 \quad \text{at } z \geq 30 \text{ m} \quad (3b)$

237 The equation also assumes heterogeneity with a vertical trend in Holocene aquifer no. II and  
238 stationary heterogeneity in Pleistocene aquifer no. III. This study assumed  $R(s)$  as a regional  
239 variable with the same variogram throughout the analysis domain, as detailed geologic information  
240 for dividing the monotonic sequence was not obtained. Sequential Gaussian simulation (SGS) was  
241 used for its stochastic realization as in other sandy gravel aquifer modeling (Eggleston et al. 1996;  
242 Jang and Liu 2004; Lee et al. 2007). The SGS process has six steps (Deutsch and Journel 1998): 1)  
243 input the honor data (solid squares in Figure 2), which are equal to the estimates at the nodes. 2)  
244 Generate a random path though the grid nodes. 3) Select the first node in that path and estimate a  
245 mean and standard deviation at that node by normal or ordinary kriging based on the surrounding

246 honor data. 4) Obtain a random value from the mean and standard deviation and set the value at the  
247 node. 5) Include the newly simulated value as part of the honor data. 6) Repeat steps 2 to 5 until all  
248 the grid nodes have simulated values. The SGS in this study was performed using a free Fortran  
249 code *sgsim* in GSLIB (Deutsch and Journel 1998).

250 The realizations by SGS are significantly dependent on the variogram model. The variogram  
251 of  $R(s)$  in the fan was previously formulated by Sakata (2013) who used 134 slug-test results and  
252 performed variogram analysis using original data of  $R(s)$  and obtained a histogram that was  
253 approximately Gaussian at 95% significance level. Consequently, a spherical anisotropic model of  
254 distance  $h$  was obtained as:

$$255 \quad \gamma(h) = c + b \left[ \frac{3}{2} \frac{h}{a} - \frac{1}{2} \left( \frac{h}{a} \right)^3 \right] \quad (4)$$

256 In Sakata (2013), the sill  $b$  and nugget  $c$  were determined to be nearly identical in the horizontal and  
257 vertical directions ( $b = 1.0$  and  $c = 0.5$ ). The range  $a$  differed in the  $x$  and  $z$  directions,  $a_x = 1000$  m in  
258 the lateral direction and  $a_z = 20$  m in the vertical. Such large range anisotropy was also observed in  
259 nature (Deutsch 2007). However, the estimated ranges may be too large on the local scale used in  
260 this analysis compared with other field studies (e.g., Hess et al. 1992; Rubin 2003). Gelhar (1993)

261 summarized the correlation scales of horizontal and vertical  $K$  on an overall scale less than 1 km.  
262 The relationship indicated that a correlation scale of 100 m was probably 10 to 20 m in the  
263 horizontal direction, and a few meters in the vertical direction (Figure 6.6 in the text of Gelhar).  
264 Thus, the author examined the residual variogram on two different scales, large ( $a_x = 1000$  m and  $a_z$   
265 = 20 m) and small ( $a_x = 10$  m and  $a_z = 2$  m), in addition to the presence or absence of a vertical  
266 decreasing trend. The sill and nugget parameters were probably also different among the local and  
267 basin scales as well as the ranges. However, the values in the previous analysis were consistently  
268 used in this study because only eight  $Y$  honor data were obtained in the cross section. The sill and  
269 nugget values generated a relatively random field of  $K$  and the value of  $K$  changed by several orders  
270 of magnitude between adjacent points. The variance of the  $Y$  honor data was 1.2, which was nearly  
271 equal to the sum of the nugget and sill.

272 One thousand realizations of the residuals were generated on the 1-m regular grid  
273 throughout the cross-section for the two variogram ranges (large and small). Next, the logarithmic  
274 conductivity  $Y(s)$  was obtained by adding the simulated residuals to the trend component (Equation  
275 (3)). For this process, another constant trend  $M_0$  was also assumed instead of the vertical trend.  $M_0$

276 was simply calculated at  $-4.17$ , as an integral average of Equation (3) from the surface to the bottom  
277 of the aquifer at DW-1. Figure 5 shows extracted examples of the residuals (panels a and b) and the  
278 resulting logarithm of  $K$  (panels c to f). Panel c shows the most plausible realization from a total of  
279 8000, as described below. The residuals were generated using the large (panel a) and small (panel b)  
280 variogram ranges, respectively. Both realizations demonstrate variability in about four orders of  
281 magnitude among the 1-m regular grids. These images also include slightly decreasing trends with  
282 depth as a result of the conditional simulation with the honor data at DW-1. The simulation that used  
283 the large-range condition (panel a) produced relatively continuous zones of high  $K$  value (red and  
284 orange) and low  $K$  value (light and dark blue). On the other hand, the small range condition (panel  
285 b) shows almost randomly patched zones, probably typical alluvial gravel deposits. The images of  $Y$   
286 (panels c and e) were obtained by adding the vertical trend to panels a and b, resulting in relatively  
287 high and low  $Y$  values in the shallow and deep layers. The lower images (panels d and f) were  
288 obtained by adding a constant mean to the same residuals, producing mostly vertical trends that are  
289 masked in the large heterogeneity.

290 The stochastic realizations of  $Y(s)$  were inversely log-transformed to  $K(s)$ , which was assigned

291 to each cell in the numerical model.  $Y(s)$  in the lowermost layer and on both sides of the 5-m-wide  
292 cells was obtained by averaging all the  $Y(s)$  values of the 1-m grid cells. In the final modeling step,  
293 the ratio of vertical/horizontal anisotropy was also considered for assessing further small-scale  
294 heterogeneity within a depth of 1 m (Freeze and Cherry 1979; Burger and Belitz 1997). The  
295 vertical/horizontal anisotropy ratio  $R$  varies in space; therefore, individual values should be given at  
296 each cell. However, the author prepared two simplified cases in which all the cells had a constant  
297 isotropic ratio ( $R = 1$ ) or were assigned an anisotropy value ( $R = 0.1$ ). Other aquifer and fluid  
298 properties were assumed to be constant in this simulation. The total porosity was 0.2 and the specific  
299 storage was set to  $10^{-5} \text{ m}^{-1}$ . The heat capacities and densities of the andesitic gravel deposits were  
300  $850 \text{ J/kg/}^\circ\text{C}$  and  $2600 \text{ kg/m}^3$ , and those of the fluid were 4183 and 1000, respectively. The thermal  
301 conductivities of the bulk and fluid were 2.5 and  $0.6 \text{ W/m/}^\circ\text{C}$ , respectively. The longitudinal and  
302 transverse thermal diffusivity values were 10 and 0.1. These thermal parameters were less sensitive  
303 to the heat transport simulation than  $K$  (Sakata 2013). The dependencies of the fluid density and  
304 viscosity on temperature were assigned from the literature (Voss 1984; Langevin et al. 2008).  
305

306 **Error estimation**

307 The coupled groundwater flow and heat transport simulation was run in transient mode from  
308 10 June 2010 through 28 February 2011 (264 days) with 24 time steps (hourly). The calculated  
309 temperature was extracted each day at six observation points, i.e., at three depths of DW-1 (index  $j =$   
310 1, 2, 3) and at the three shallow wells, SW-1 ( $j = 4$ ), SW-2 ( $j = 5$ ), and SW-3 ( $j = 6$ ). The model  
311 fitting in each realization was assessed by least squares; the fit between simulated values and field  
312 measurements was quantified using squared differences. The individual root mean square error  
313  $\overline{RMSE}_j$  values were calculated at each observation point ( $j = 1$  to 6), and the average root mean  
314 square error  $\overline{RMSE}$  was then obtained from the six values of the individual  $\overline{RMSE}_j$ :

315 
$$\overline{RMSE} = \frac{1}{N_w} \sum_1^{N_w} RMSE_j \quad (5a)$$

316 
$$RMSE_j = \sqrt{\frac{1}{N_d} \sum_{i=1}^{N_d} (T_j(t_i) - T_j^*(t_i))^2} \quad (5b)$$

317 where  $N_w$  is the number of observation points for assessing the fit (6 here).  $N_d$  is the number of  
318 transient calculation outputs (264).  $T^*(t_i)$  and  $T(t_i)$  are the calculated and observed temperatures at  
319 each calculation time  $t_i$  ( $i = 1$  to 264 day), respectively. The statistical values of the average  $\overline{RMSE}$   
320 were summarized, and the cumulative percentage curves of  $\overline{RMSE}$  plotted for comparison. In this

321 study, the criterion of  $\overline{RMSE}$  for a good match between the model and observations was defined as  
322 1.0 °C. The average range of the groundwater fluctuations is 6.52 °C; therefore, the normalized error  
323 of the criterion is about 15 % (1.0/6.52). The author believes that obtaining an error range within  
324 about 10–20 % is probably difficult in groundwater transport simulation. Indeed, the error estimator  
325  $\overline{RMSE}$  was affected not only by the comparison conditions but also by various factors such as  
326 measurement errors, uncertainty in initial and boundary conditions, and model restrictions in the  
327 cross section. Here, the criterion is considered to be suitable because it revealed a statistical  
328 discrepancy of  $\overline{RMSE}$  among the eight study cases, and the calculated temperature time-series  
329 visually match those of the observations, as shown below. Consequently, the realizations of  $\overline{RMSE}$   
330 that were less than the criterion value were extracted as “plausible.”

331

## 332 **Results**

### 333 **Error Statistics**

334 For each of the eight cases, Table 1 summarizes the statistical values of the  $\overline{RMSE}$ , median  
335 ( $Me$ ), variance ( $Var$ ), minimum value ( $Min$ ), and the number of  $\overline{RMSE}$  realizations less than 1°C ( $N$ )

336  $\overline{RMSE}$ ) for 1000 realizations. Figure 6 shows the cumulative percentage curves of  $\overline{RMSE}$  for each case.

337  $Me$  was in a narrow range between 2.5°C and 2.9°C, and was insensitive to any of the model

338 conditions. This means that most of the realizations failed to calculate the high heat flow around the

339 losing river. This was because  $\overline{RMSE}$  was calculated at 2.92°C in the special case where the

340 temperatures in all the cells did not vary from the initial condition. The actual average seepage rate

341 from the river boundary was about 0.2 m<sup>3</sup>/(m hr), and the value was only one tenth of the observed

342 seepage rate  $q$  (2 m<sup>3</sup>/(m hr)). The root mean square errors of the groundwater head in the realizations

343 were over about 1 m in the deep and shallow wells. The failures in most realizations were probably

344 due to the limitation of Gaussian simulation in this study. Gaussian simulation cannot reproduce

345 extremely high values of simulated variables (Koltermann and Gorelick 1996; Lee et al. 2007).

346 However, the exchange of heat between the surface water and groundwater is dependent on such

347 preferential flow paths such as buried channels (e.g., Woessner 2000). Moreover, in this study, the  $Y$

348 honor data were limited to only one borehole and detailed geologic information around the losing

349 river. This Monte Carlo approach using a total of eight thousands realizations indicated that the heat

350 transport simulation may be inefficient unless the typical conduits of high heat flows are identified in

351 advance, at least to an approximate extent.

352 Despite the low efficiency of the method, small  $\overline{RMSE}$  values were frequently obtained as *Var*  
353 increased since *Me* was nearly equal in all the study cases. In Figure 6, when *Var* increased, the  
354 cumulative percentage curves of  $\overline{RMSE}$  had relatively gentle slopes. The cases that had a vertical  
355 trend (red lines) had a more gentle slope than those without the trend (blue lines). The large  
356 variogram range (bold lines) also tended to produce a leftward shift in the area compared with the  
357 cases with a small range (thin lines), although the change appears smaller than that associated with  
358 the presence or absence of a vertical trend. The condition of cell anisotropy appears to have a smaller  
359 effect than the other two conditions. The relation of *Min* among the eight cases was opposite to that  
360 of *Var*; *Min* in the vertical trend cases was consistently smaller than in the constant trend cases,  
361 independent of the other conditions. *Min* in the large range cases was also smaller than in the small  
362 range cases, as well as in the vertical trend cases. However, the relation of *Min* with regard to cell-  
363 isotropy and anisotropy was not obvious. Especially in the large range cases, *Min* of the cell-isotropy  
364 cases was nearly equal to that of the anisotropy case.

365  $N_{\overline{RMSE}}$  is another probability index that indicates how often plausible realizations were

366 generated. As shown in Table 1, the largest  $N_{\overline{RMSE}}$  was 10 in Case 2. The second largest was 4 in  
367 Case 1. The smallest number was 2 in Case 3. These cases were common in terms of the  
368 exponentially decaying pattern of  $K$ . No plausible realizations were achieved in the other study  
369 cases. This also indicates that modeling the exponentially decreasing trend may be essential for  
370 obtaining plausible realizations. However, no plausible realization was achieved in Case 4, which  
371 also contained a vertical trend as did Cases 1 through 3. Moreover, where plausible cases were  
372 found, their occurrence corresponds to a probability of  $\leq 1\%$  of the 1000 realizations. Thus, one  
373 might argue that  $N_{\overline{RMSE}}$  in the latter three cases may not be as statistically significant as the other  
374 cases ( $N_{\overline{RMSE}} = 0$ ). To verify this, a statistical test was performed on the hypothesis that  $N_{\overline{RMSE}}$  was  
375 obtained rarely (less than or equal to 1%) but randomly, that is, independent of the vertical trend,  
376 variogram range, and cell anisotropy (i.e., the Poisson distribution). The probability  $P_0$  of no  
377 occurrence ( $N_{\overline{RMSE}} = 0$ ) among  $n$  trials was simply calculated with the mean probability  $p$ :  $P_0 = (1 -$   
378  $p)^n$ . For the statistical hypothesis, the mean probability of the occurrence of  $\overline{RMSE} < 1^\circ\text{C}$  in 1000  
379 samples was assumed to be  $p_1 = 0.002$  ( $N_{\overline{RMSE}} = 2$  in Case 3),  $p_2 = 0.004$  ( $N_{\overline{RMSE}} = 4$  in Case 1), and  $p_3$   
380  $= 0.01$  ( $N_{\overline{RMSE}} = 10$  in Case 2). Thus, the probabilities of  $N_{\overline{RMSE}} = 0$  were 0.14, 0.02, and  $< 0.01$ ,

381 respectively. For  $N_{\overline{RMSE}} = 2$  in Case 3, the hypothesis was accepted with a significance level of 95%.  
382 This indicates that Case 3 with a vertical trend, small variogram range, and cell anisotropy  
383 conditions may not be different in terms of reproducing plausible realizations from other cases of  $N$   
384  $\overline{RMSE} = 0$ . On the other hand, the hypothesis was rejected for  $N_{\overline{RMSE}} = 4$  in Case 1 and  $N_{\overline{RMSE}} = 10$  in  
385 Case 2. Thus, it can be concluded that the plausible realizations of  $\overline{RMSE} < 1^\circ\text{C}$  were statistically  
386 significant in Cases 1 and 2, i.e., those with a combination of the vertical trend and large variogram  
387 range, relative to the other cases.

388

### 389 **Temperature time series**

390 Figure 7 shows the calculated temperature time-series for the smallest  $\overline{RMSE}$ , i.e., *Min*,  
391 realizations in each study case. At the deepest point DW-1-3, all the results were nearly identical  
392 with the observation of a constant temperature of  $10.1^\circ\text{C}$ . This indicates that the isothermal layer  
393 was independent of the spatial variability of  $K$ , and was realized based on the boundary conditions.  
394 For the shallow SW-3 well, all the calculations also achieved almost satisfactory matching with  
395 observations. The average  $\overline{RMSE}$  values were obtained for individual  $\overline{RMSE}_j$  of six points, but the

396 averaging was most affected by  $\overline{RMSE}_6$  at SW-3 because  $\overline{RMSE}_6$  varied the most among the  
397 different realizations, with relation to large fluctuations of groundwater temperature as a result of  
398 high heat flows.

399         The temperatures observed in the shallow wells SW-1 and SW-2 showed typically smooth  
400 fluctuations with different peak delays. All the results of Cases 1 to 5 calculated the typical delay at a  
401 relatively close shallow well, SW-1. In Cases 6 and 7 the peak delay was longer while in Case 8 the  
402 calculated temperature at SW-1 was almost constant, indicating no heat fluxes in the west direction.  
403 At SW-2, however, the peak delay was in good agreement with the observations in Cases 1, 2, and  
404 3. The results were equal to those in which the plausible realizations were obtained ( $N_{\overline{RMSE}} > 0$ ). On  
405 the other hand, in Cases 4 to 8, almost no fluctuation indicated that most of the heat fluxes infiltrated  
406 into the deeper layer, due to downward flows.

407         Finally, Cases 1, 2, and 3 were compared at the deep well DW-1-1 and DW-1-2. Case 1  
408 showed little fluctuations compared with the observation at DW-1-2 while in Case 3, very little  
409 fluctuations were calculated at DW-1-1. Only the realization in Case 2 provided some fluctuations at  
410 both DW-1-2 and DW-1-1 as a result of heat flow toward these points. Therefore, the realization in

411 Case 2, as previously shown in Figure 5c, may be the most plausible among the 8000 realizations.  
412 The average seepage rate in the realization was  $2 \text{ m}^3/(\text{m hr})$ , which was equal to the observed rate  $q$ .  
413 The RMSE of the hydraulic head was only 0.12 m in SW-1 and 0.6 m in DW-1. These values are  
414 within the accepted values of anisotropic heat transfer around the losing river. Note that the  
415 calculated temperatures at DW-1-2 were not well matched because the heat fluxes from the previous  
416 winter occurred within the calculation period.

417

## 418 **Discussion**

419 This study used heat as a groundwater tracer for examining depth-decaying permeability in a  
420 monotonic gravel aquifer, in the Toyohira River alluvial fan. Around the losing river reach,  
421 groundwater temperature was applied as a tracer for the sensitivity analysis of  $K$ . Heat transport  
422 simulation was performed using stochastic realizations of  $K$ . Another unique aspect of this study was  
423 the simulation of the spatial variability of  $K$  using a combination of multiple-scale heterogeneity: the  
424 exponential decreasing trend with depth on a scale of tens of meters, the stochastic residuals between  
425 1-m regular grids, and cell anisotropy within a depth of 1 m. One thousand realizations of the

426 residuals were generated by SGS and converted to  $Y$  by adding the trend component. Back log-  
427 transformed  $K$  were input to the SEAWAT model in eight comparison cases with different  
428 conditions: the presence or absence of vertical trend, large or small variogram range, and cell-  
429 isotropy or anisotropy. The average  $\overline{RMSE}$  at six observation points was used for comparison with  
430 the observations, and the statistical parameters were compared.

431         The variance of the  $\overline{RMSE}$  indicates that the cases with the vertical trend condition had smaller  
432  $\overline{RMSE}$  values. Plausible realizations of  $N_{\overline{RMSE}}$  were obtained only in the cases that simulated a vertical  
433 trend. When the vertical trend was assumed, the plausible realizations were obtained in cases of  
434 either large or small ranges, and cell isotropy or anisotropy. The relative sensitivity is noteworthy,  
435 because the change of  $K$  by vertical trend (one to two orders of magnitude) was smaller than the  
436 random variation by SGS (four orders of magnitude). In this study, several realizations with the  
437 smallest  $\overline{RMSE}$  values provide a good match with observations when the vertical trend is included.  
438 The most plausible model showed an almost perfect match to the observations in terms of  
439 groundwater temperature, hydraulic head, and seepage rate. Therefore, modeling the exponential  
440 decrease may be an important factor in the description of the shallow heat transport, and of other

441 solute components such as artificial contaminants infiltrated from the surface.

442 In this study, however, the effect of modeling the vertical trend was not pronounced because  
443 the number of plausible realizations  $N_{\overline{RMSE}}$  was only ten or less, i.e., 1% of 1000 realizations. The  
444 mean  $\overline{RMSE}$  was insensitive to the introduced conditions and nearly equal to the value for the case  
445 when the groundwater temperature is assumed to remain steady at its initial level. These results also  
446 provided much lower seepage rates from the river boundary, compared with the observations. In  
447 particular, a statistical test suggested that  $N_{\overline{RMSE}}$  may be significant only when the vertical trend and  
448 large range conditions are present. The large range condition was particularly helpful for SGS in  
449 reproducing a high- $K$  connected area as a conduit. Thus, the vertical trend was probably essential to  
450 the anisotropic heat transport system; however, a sedimentary structure that included preferential  
451 flow paths is also required for a high heat flow system around the losing river. In this study, the  
452 preferential flow paths were specified among thousands of realizations with a (maybe unrealistic)  
453 large range condition. This approach was too inefficient, and the simulations were validated using a  
454 limited number of observation data (only six points in this study). Additional information about  
455 zones of increased groundwater flow may be combined with the stochastic modeling to improve the

456 efficiency of obtaining plausible realizations. For example, a delineation of the hydrofacies,  
457 including preferential flow paths (e.g., open-framework gravel) can be obtained near the surface.  
458 The shallow information is then interpolated to the deeper layers, based on data from a number of  
459 high-quality cores, a high-resolution geophysical survey (e.g., Regli et al. 2002; Bennett et al. 2006)  
460 and other advanced methods such as hydraulic tomography (Berg and Illman 2011). In future work,  
461 indicator simulation or other advanced methods could provide good realizations of complex  
462 heterogeneity, as shown in several innovative studies (Weissmann et al. 1999; Felletti et al. 2006;  
463 Ramanathan et al. 2010; Guin et al. 2010). These field and numerical methods are probably  
464 applicable but practically limited, as suggested in this paper. Another approach based on calibration  
465 such as the parameter estimation method PEST may be effective when field data are limited.

466         The vertical/horizontal ratio of the anisotropy of  $K$  at each 1-m cell helps prevent downward  
467 groundwater flow, as well as the vertical decreasing trend on a scale of tens of meters. This study  
468 also showed that the  $N_{\overline{RMSE}}$  in Case 2 with anisotropy was more than double the  $N_{\overline{RMSE}}$  in Case 1 with  
469 isotropy. The most plausible realization included the anisotropic condition; however, the statistical  
470 test suggested that the effect of the cell anisotropy condition was not larger than that of the other two

471 comparison conditions. The heterogeneity within 1 m was probably insensitive to the shallow heat  
472 transport on this scale. In this study, the vertical/horizontal anisotropy was assumed to be constant  
473 throughout the area even though the value actually varies in space. Modeling the cell anisotropy will  
474 also be required in future work.

475 Finally, several other uncertainties and limitations in the present analysis are discussed. First,  
476 the exponential decay trend (Equation (1)) is probably unique to this alluvial fan. Other gravelly  
477 aquifers may show different magnitudes of decay exponent, other functions (e.g., a simple linear  
478 trend), or unobvious trends with a small exponent. In this heat transport simulation, similar  
479 conclusions may be derived using a linear trend fitted to the exponential model in the shallow depth.  
480 The vertical change of  $K$  must be established individually at each site using multi-level in-situ tests,  
481 core-analysis, and other methods. Additionally, the number of observation points (six points in this  
482 study) was probably not sufficient for inferring the complex heterogeneity in the gravel deposits. To  
483 improve the reliability of the results, several shallow wells should be installed between SW-1 and  
484 SW-2, and at least one deep well in addition to DW-1. The error criterion should be changed to a  
485 more reasonable one, considering several other factors of error. Finally, restricting the model within

486 the cross section may affect the small number of plausible realizations, because the generated  $K$   
487 patterns act as if the domain was extended infinitely in the  $y$  plane. 3D modeling would be required  
488 for more rigorous treatment, although its computational cost may be a limiting factor.

489

## 490 **Conclusions**

491 In this study heat was used as a groundwater tracer to analyze the modeling of the vertical  
492 decrease of  $K$  in monotonic gravel sequences of the Toyohira River alluvial fan. The observations  
493 indicated a dominant heat flux in the shallow zone, despite a vertical hydraulic gradient, implying a  
494 hydrogeologic structure, including a vertical decrease. Heat transport simulation was performed  
495 using stochastic realizations of  $K$ , based on the Monte Carlo approach. One thousand realizations of  
496  $K$  were reproduced using a combination of the exponentially decreasing trend and small-scale  
497 heterogeneity generated by SGS. Eight different conditions were modeled: the presence or absence  
498 of a vertical trend, large or small variogram range, and cell isotropy or anisotropy. The averaged root  
499 mean square error  $\overline{RMSE}$  at six observation points was calculated for each realization, and the  
500 statistical values were compared.

501           The mean of  $\overline{RMSE}$ ,  $Me$ , was insensitive to the eight modeled conditions, and in many  
502 realizations the seepage rate from the river boundary was much smaller than in the observations.  
503 This means that the SGS did not generate reliable preferential flow paths for high heat flow with the  
504 limited honor data available. Despite its inefficiency, plausible realizations with  $\overline{RMSE}$  below the  
505 threshold (the criterion temperature of 1°C) were obtained when the vertical trend condition was  
506 present, although the systematic change of  $K$  with depth was two orders of magnitude smaller than  
507 the small-scale heterogeneity. Several realizations with vertical trend agreed with the observations,  
508 and the most plausible one was almost a complete match in terms of groundwater temperature,  
509 hydraulic head, and seepage rate. This indicates that the vertical trend effectively contributed to the  
510 anisotropic heat transport system.

511           However, there were fewer than ten plausible realizations among the 1000 realizations,  
512 including those with the vertical decrease. A statistical test suggests that the plausible realizations  
513 with the vertical trend and large variogram range conditions may be statistically significant. The  
514 variogram range condition was probably unrealistic, but aided in generating a high-connected zone.  
515 The cell-anisotropy condition also increased the plausibility of the realizations, but the effect was

516 slight compared with that of the other conditions. In conclusion, the modeling in this study showed  
517 that the vertical trend was essential in the anisotropic heat transport; however, for describing high  
518 heat transport, detailed information of the smaller-scale heterogeneity, including preferential flow  
519 paths, is required to produce statistically significant results.

520

## 521 **Acknowledgements**

522 The author sincerely thanks Prof. Ryuji Ikeda of Hokkaido University for the meaningful  
523 discussions, suggestions, and comments, and would like to acknowledge the Hokkaido Regional  
524 Development Bureau for the use of the observation wells. This study greatly benefited from the  
525 assistance of Sayuri Watanabe, Yoko Yamamoto, Hideki Otake, and Akinari Honda, who provided  
526 technical support and prepared the numerous datasets. The comments of two anonymous reviewers,  
527 the associate Editor, and Dr. F Schwartz, Chief Editor of this journal, led to substantial  
528 improvements of this manuscript.

529

530 **References**

531 Anderson, M.P., and W.W., Woessner. 1992. *Applied Groundwater Modeling Simulation of Flow*  
532 *and Advective Transport*. London: Academic Press.

533 Anderson, M.P. 2005. Heat as a ground water tracer. *Ground Water* 43, no. 6: 951–968.

534 DOI:10.1111/j.1745-6584.2005.00052.x

535 ASCE (American Society of Civil Engineers). 2008. *Standard guideline for fitting saturated*  
536 *hydraulic conductivity using probability density function*. ASCE/EWRI 50–08. Reston, Virginia:

537 ASCE.

538 Bennett, V.G.L., G.S., Weissmann, G.S, Baker, and D.W., Hyndman. 2006. Regional-scale  
539 assessment of a sequence-bounding paleosol on fluvial fans using ground-penetrating radar,

540 eastern San Joaquin Valley, California. *Bulletin of the Geological Society of America* 118: 724–

541 732.

542 Berg, S.J., and W.A., Illman. 2011. Three-dimensional transient hydraulic tomography in a highly  
543 heterogeneous glaciofluvial aquifer-aquitard system. *Water Resources Research* 47, no. 10,

544 W10507. DOI:10.1029/2011WR010616

545 Bravo, H.R., F., Jiang, and R.J., Hunt. 2002. Using groundwater temperature data to constrain  
546 parameter estimation in a groundwater flow model of a wetland system. *Water Resources*  
547 *Research* 38, no. 8: 28–1–28–14. DOI:10.1029/2000WR000172

548 Burger, R.L., and K., Belitz. 1997. Measurement of anisotropic hydraulic conductivity in  
549 unconsolidated sands: A case study from a shoreface deposit, Oyster, Virginia. *Water Resources*  
550 *Research* 33, no. 6: 1515–1522. DOI:10.1029/97WR00570

551 Butler, Jr., J.J., G.C., Bohling, Z., Hyder, and C.D., McElwee. 1994. The use of slug tests to describe  
552 vertical variations in hydraulic conductivity. *Journal of Hydrology* 156: 137–162. DOI:  
553 10.1016/0022-1694(94)90075-2

554 Cardenas, M.B., and X.-W., Jiang. 2010. Groundwater flow, transport, and residence times through  
555 topography-driven basins with exponentially decreasing permeability and porosity. *Water*  
556 *Resources Research* 46, no. 10. DOI:10.1029/2010WR009370

557 Chen, X. 2011. Depth-dependent hydraulic conductivity distribution patterns of a streambed.  
558 *Hydrological Processes* 25, no.2: 278–287. DOI:10.1002/hyp.7844

559 Constantz, J. 2008. Heat as a tracer to determine streambed water exchanges. *Water Resources*

560        *Research* 44, no. 4, W00D10. DOI:10.1029/2008WR006996

561        Deutsch, C.V., and A.G., Journel. 1998. *GSLIB Geostatistical Software Library and User's Guide*

562        2nd ed. New York: Oxford University Press.

563        Deutsch, C.V. 2007. A review of geostatistical approaches to data fusion. In *Subsurface Hydrology:*

564        *Data Integration for Properties and Processes*, ed. D.W., Hyndman, F.D., Day–Lewis, and K.,

565        Singha, 7–18. Geophysical Monograph 171. Washington, D.C.: AGU. DOI:10.1029/171GM03

566        Domenico, P.A., and F.W., Schwartz. 1998. *Physical and Chemical Hydrogeology* 2nd ed. New

567        York: Wiley.

568        Eggleston, J.R., S.A., Rojstaczer, and J.J., Peirce. 1996. Identification of hydraulic conductivity

569        structure in sand and gravel aquifers: Cape Cod data set. *Water Resources Research* 32, no. 5:

570        1209–1222. DOI:10.1029/96WR00272

571        Eggleston, J.R., and S., Rojstaczer. 1998. Identification of large-scale hydraulic conductivity trends

572        and influence of trends on contaminant transport. *Water Resources Research* 34, no. 9: 2155–

573        2168. DOI:10.1029/98WR01475

574        Felletti, F., R., Bersezio, and M., Giudici. 2006. Geostatistical simulation and numerical upscaling, to

575 model ground-water flow in a sandy-gravel, braided river, aquifer analogue. *Journal of*  
576 *Sedimentary Research* 76, no. 11: 1215–1229. DOI:10.2110/jsr.2006.091

577 Freeze, R.A., and J.A., Cherry. 1979. *Groundwater*. Englewood Cliffs, New Jersey: Prentice-Hall.

578 Gelhar, L.W. 1993. *Stochastic Subsurface hydrology*. Englewood Cliffs, New Jersey: Prentice Hall.

579 Guin, A., R., Ramanathan, R.W., Ritzi, Jr., D.F., Dominic, I.A., Lunt, T.D., Scheibe, and V.L.,  
580 Freedman. 2010. Simulating the heterogeneity in braided channel belt deposits: 2. Examples of  
581 results and comparison to natural deposits. *Water Resources Research* 46, no. 4, W04516.  
582 DOI:10.1029/2009WR008112

583 Hess, K.M., S.H., Wolf, and M.A., Celia. 1992. Large-scale natural gradient tracer test in sand and  
584 gravel, Cape Cod, Massachusetts: 3. Hydraulic conductivity variability and calculated  
585 macrodispersivities. *Water Resources Research* 28, no. 8: 2011–2027.  
586 DOI:10.1029/92WR00668

587 Hu, S.G., S., Miyajima, D., Nagaoka, K., Koizumi, and K., Mukai. 2010. Study on the relation  
588 between groundwater and surface water in Toyohira-gawa alluvial fan, Hokkaido, Japan. In  
589 *Groundwater Response to Changing Climate*, ed. M., Taniguchi, and I.P., Holman, 141–157. AK

590 Leiden, Netherlands: CRC Press. DOI:10.1201/b10530-13

591 Ingebritsen, S.E., W.E., Sanford, and C.E., Neuzil. 2006. *Groundwater in Geologic Processes*, 2nd  
592 ed. Cambridge: Cambridge University Press.

593 Jang, C.S., and C.W., Liu. 2004. Geostatistical analysis and conditional simulation for estimating the  
594 spatial variability of hydraulic conductivity in the Choushui River alluvial fan, Taiwan.  
595 *Hydrological Processes* 18, no. 7: 1333–1350. DOI:10.1002/hyp.1397

596 Jiang, X.W., L., Wan, X.S., Wang, S., Ge, and J., Liu. 2009. Effect of exponential decay in hydraulic  
597 conductivity with depth on regional groundwater flow. *Geophysical Research Letters* 36, no. 24.  
598 DOI:10.1029/2009GL041251

599 Jiang, X.W., X.S., Wang, and L., Wan. 2010a. Semi-empirical equations for the systematic decrease  
600 in permeability with depth in porous and fractured media. *Hydrogeology Journal* 18, no. 4: 839–  
601 850. DOI:10.1007/s10040-010-0575-3

602 Jiang, X.W., L., Wan., M.B., Cardenas, S. Ge, and X.S., Wang. 2010b. Simultaneous rejuvenation  
603 and aging of groundwater in basins due to depth decaying hydraulic conductivity and porosity.  
604 *Geophysical Research Letters* 37, no. 5, L05403. DOI:10.1029/2010GL042387

605 Jussel, P., F., Stauffer, and T., Dracos. 1994. Transport modeling in heterogeneous aquifers: 1.  
606 Statistical description and numerical generation of gravel deposits. *Water Resources Research*  
607 30, no. 6: 1803–1817. DOI:10.1029/94WR00162

608 Koltermann, C.E., and S.M., Gorelick. 1995. Fractional packing model for hydraulic conductivity  
609 derived from sediment mixtures. *Water Resources Research* 31, no. 12: 3283–3298.  
610 DOI:10.1029/95WR02020

611 Koltermann, C.E., and S.M., Gorelick. 1996. Heterogeneity in sedimentary deposits: a review of  
612 structure-imitating, process-imitating, and descriptive approaches. *Water Resources Research* 32,  
613 no. 9: 2617–2658. DOI:10.1029/96WR00025

614 Langevin, C.D., D.T., Thorne, Jr., A.M., Dausman, M.C., Sukop, and W., Guo. 2008. *SEAWAT*  
615 *Version 4: A computer program for simulation of multi species solute and heat transport.*  
616 Techniques and Methods 6, A22. Reston, Virginia: USGS.

617 Lee, S.Y., S.F., Carle, and G.E., Fogg. 2007. Geologic heterogeneity and a comparison of two  
618 geostatistical models: Sequential Gaussian and transition probability-based geostatistical  
619 simulation. *Advances in Water Resources* 30, no. 9: 1914–1932.

620 DOI:10.1016/j.advwatres.2007.03.005

621 Major, J.J., R.M., Iverson, D.F., McTigue, S., Macias, and B.K., Fiedorowicz. 1997. Geotechnical  
622 properties of debris-flow sediments and slurries, In *Debris flow Hazards Mitigation: Mechanics,*  
623 *Prediction, and Assessment*, ASCE Proceedings of First International Conference, ed. C.L.  
624 Chen, 249–259. San Francisco: ASCE.

625 Marion, A., A.I., Packman, M., Zaramella, and A. Bottacin-Busolin. 2008. Hyporheic flows in  
626 stratified beds. *Water Resources Research* 44, no. 9. DOI:10.1029/2007WR006079

627 Neton, M.J., J., Dorsch, C.D., Olson, and S.C., Young. 1994. Architecture and directional scales of  
628 heterogeneity in alluvial-fan aquifers. *Journal of Sedimentary Research* 64, no. 2b : 245–257.  
629 DOI:10.1306/D4267FA0-2B26-11D7-8648000102C1865D

630 Nowak, M. and G., Verly. 2004. The practice of sequential Gaussian simulation. In *Geostatistics*  
631 *Banff*, Ed., O., Leuangthong, and C.V., Deutsch, 387–398. Netherlands, Dordrecht: Springer.  
632 DOI:10.1007/978-1-4020-3610-1\_39

633 Ramanathan, R., A. Guin, R.W., Ritzi, Jr., D.F., Dominic, V.L., Freedman, T.D., Scheibe, and I.A.,  
634 Lunt. 2010. Simulating the heterogeneity in braided channel belt deposits: 1. A geometric-based

635 methodology and code. *Water Resources Research* 46, no. 4. DOI:10.1029/2009WR008111

636 Regli, C., P., Huggenbergera, P., and M., Rauberb. 2002. Interpretation of drill core and georadar  
637 data of coarse gravel deposits. *Journal of Hydrology* 255 no. 1–4: 234–252. DOI:  
638 10.1016/S0022-1694(01)00531-5

639 Rubin, Y. 2003. *Applied stochastic hydrogeology*. New York: Oxford university press.

640 Rust, B.R. 1979. Facies Models 2. Coarse Alluvial Deposits. In *Facies Models*, ed. R.G., Walker,  
641 11–21. Toronto: Geological Association of Canada.

642 Saar, M.O., and M., Manga. 2004. Depth dependence of permeability in the Oregon Cascades  
643 inferred from hydrogeologic, thermal, seismic, and magmatic modeling constraints. *Journal of*  
644 *Geophysical Research* 109: B04204. DOI:10.1029/2003JB002855

645 Sakata, Y., and R., Ikeda. 2012. Quantification of longitudinal river discharge and leakage in an  
646 alluvial Fan by synoptic survey using handheld ADV. *Journal of Japan Society of Hydrology*  
647 *and Water Resources* 25, no. 2: 89–102 (in Japanese).

648 Sakata, Y., and R., Ikeda. 2013a. Regional mapping of vertical hydraulic gradient using uncertain  
649 well data: a case study of the Toyohira River alluvial fan, Japan. *Journal of Water Resource and*

650        *Protection* 5, no. 8: 823–834. DOI:10.4236/jwarp.2013.58083

651        Sakata, Y., and R., Ikeda. 2013b. Depth dependence and exponential models of permeability in  
652        alluvial-fan gravel deposits. *Hydrogeology Journal* 21, no. 4: 773–786. DOI:10.1007/s10040-  
653        013-0961-8

654        Sakata, Y. 2013. 5 Stochastic simulation of groundwater-flow and heat-transport. In *Geostatistical*  
655        *reservoir modeling of trending heterogeneity specified in focused recharge zone: a case study of*  
656        *Toyohira River alluvial fan, Sapporo, Japan*, 112–149. Ph.D. dissertation, Hokkaido University,  
657        Sapporo, Japan.

658        Thorne, D., C.D., Langevin, and M.C., Sukop. 2006. Addition of simultaneous heat and solute  
659        transport and variable fluid viscosity to SEAWAT. *Computer and Geosciences* 32: 1758–1768.  
660        DOI:10.1016/j.cageo.2006.04.005

661        Voss, C.L. 1984. *A finite-element simulation model for saturated-unsaturated, fluid-density-*  
662        *dependent ground-water flow with energy transport or chemically-reactive single-species solute*  
663        *transport*, Water-Resources Investigations Report 84–4369. Reston, Virginia: USGS.

664        Wang, K., P.Y., Shen, and A.E., Beck. 1989. A solution to the inverse problem of coupled

665 hydrological and thermal regimes. In *Hydrogeological Regimes and Their Subsurface Thermal*  
666 *Effects*, ed. A.E. Beck, G. Garven, and L. Stegena, Geophysical Monograph Series 47. no. 2: 7–  
667 21. Washington, D.C.: American Geophysical Union.

668 Weissmann, G.S., S.F., Carle, and G.E., Fogg. 1999. Three-dimensional hydrofacies modeling based  
669 on soil survey analysis and transition probability geostatistics. *Water Resources Research* 35, no.  
670 6: 1761–1770. DOI:10.1029/1999WR900048

671 Woessner, W.W. 2000. Stream and fluvial plain ground water interactions: rescaling hydrogeologic  
672 thought. *Ground Water*, 38 No. 3: 423–429, DOI:10.1111/j.1745-6584.2000.tb00228.x

673 Woodbury, A.D., and L., Smith. 1988. Simultaneous inversion of hydrogeologic and thermal data: 2.  
674 Incorporation of thermal data. *Water Resources Research* 24, no. 3: 356–372.  
675 DOI:10.1029/WR024i003p00356

676 Yamaguchi, Y., and A., Nakamura. 1994. Change of permeability of shallow foundation for fill dams  
677 due to excavation and embankment. *Journal of Groundwater Hydrology* 36, no. 4: 423–438 (in  
678 Japanese).

679 Zappa, G., R., Bersezio, F., Felletti, and M., Giudici. 2006. Modeling heterogeneity of gravel-sand,

680 braided stream, alluvial aquifers at the facies scale. *Journal of Hydrology* 325, no. 1–4: 134–153.

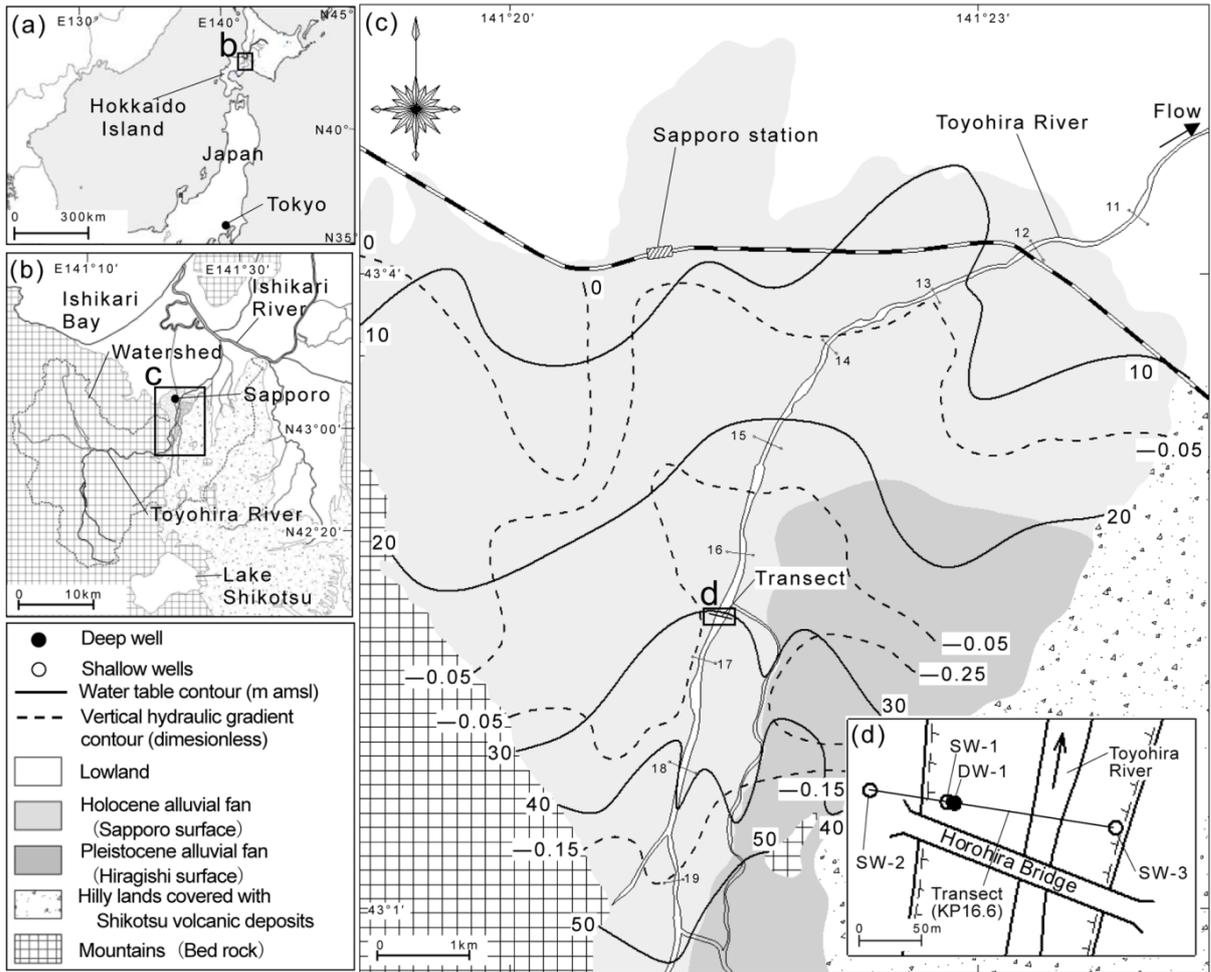
681 DOI:10.1016/j.jhydrol.2005.10.016

682 Zlotnik, V.A., M.B. Cardenas, and D. Toundykov. 2011. Effects of multiscale anisotropy on basin

683 and hyporheic groundwater flow. *Ground Water*, 49 No. 4: 576–583, DOI: 10.1111/j.1745-

684 6584.2010.00775.x

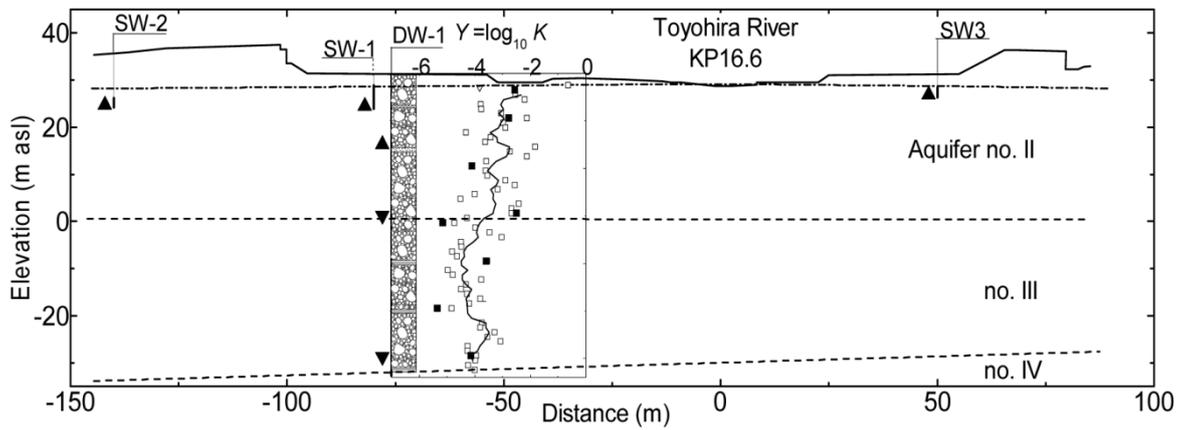
685



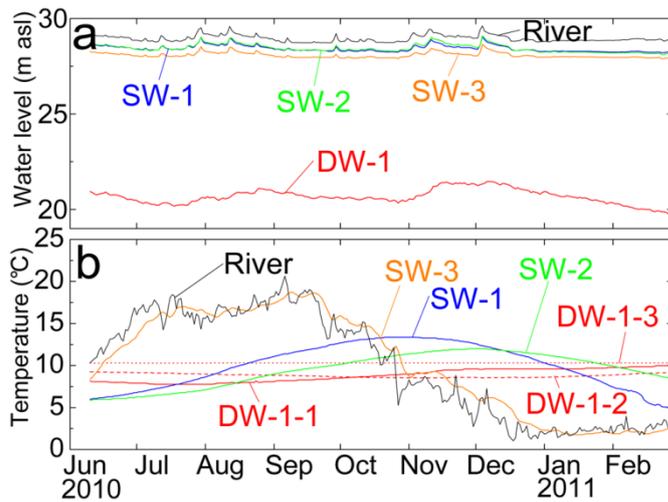
686

687 Figure 1 Location maps and topographic features of study area of Toyohira River alluvial fan.

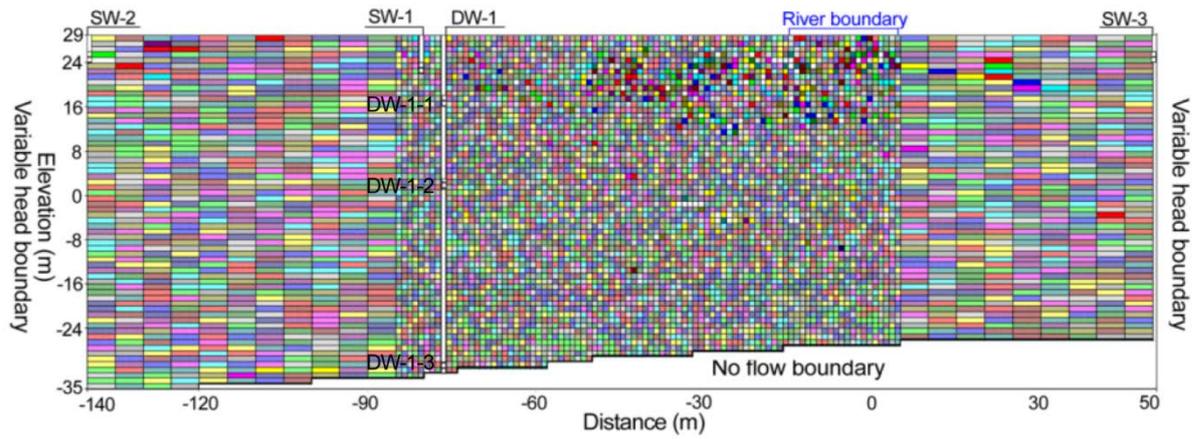
688



689  
 690 Figure 2 Cross section of the study area showing observation locations. Solid triangles represent  
 691 water pressure and temperature observations, and inverted triangles are those of temperature alone.  
 692 Aquifers II, III, and IV are described in the text. The vertical profile of  $Y$  was taken from Sakata and  
 693 Ikeda (2013b); solid squares, open squares, and solid line denote  $Y$  values measured by slug tests,  $Y$   
 694 values estimated from undisturbed cores, and moving averages of estimated  $Y$ , respectively.  
 695



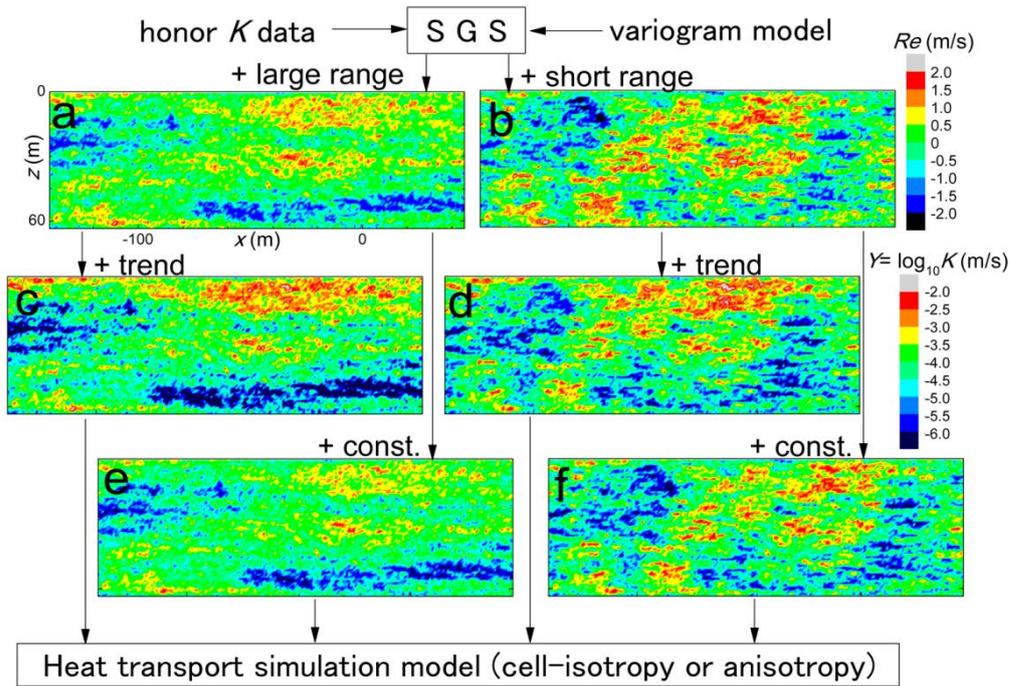
696  
 697 Figure 3 Daily mean variations of water level (a) and temperature (b) in the river and observation  
 698 wells.



699

700 Figure 4 Cross-sectional simulation model of SEAWAT. The random color of each cell was assigned  
 701 from the realization of  $K$ .

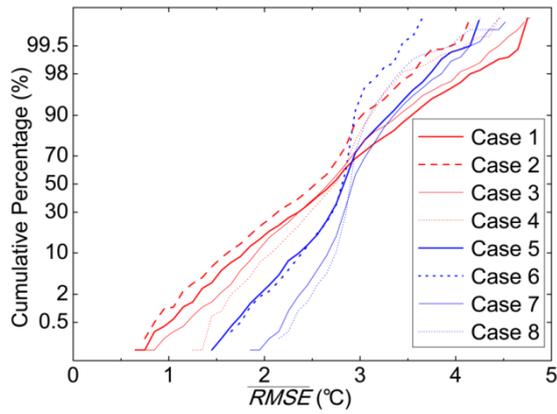
702



703

704 Figure 5 Conceptual flow chart showing the generating algorithm for the realization of the residuals  
 705  $R$  using large (a) and small (b) ranges, and of the logarithm of hydraulic conductivity  $Y$  (c to f).

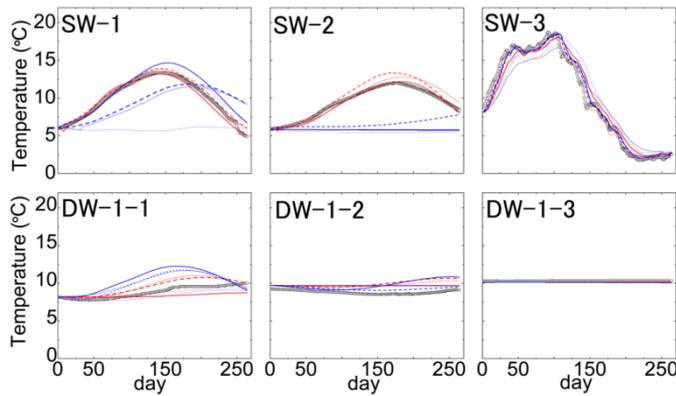
706



707

708 Figure 6 Cumulative percentage curves of  $\overline{RMSE}$  for the eight simulation cases. Red lines indicate  
 709 exponential decrease of  $K$  and blues lines have constant mean change; bold and thin lines represent  
 710 large and small variogram ranges, respectively, and solid and dotted lines indicate cell isotropy and  
 711 anisotropy, respectively.

712



713

714 Figure 7 Calculated time series of groundwater temperature by extracted realizations of  $Min$  for the  
 715 eight cases. Bold squares are observations, and color lines are the eight simulation cases denoted as  
 716 in Fig. 6.

717

718 Table 1 Summary of  $\overline{RMSE}$  in 1000 realizations of each study case.

| Case | Conditions               |              |                   | Results |       |       |             |
|------|--------------------------|--------------|-------------------|---------|-------|-------|-------------|
|      | Trend                    | Range        | Cell              | $Me$    | $Var$ | $Min$ | $N_{RMSE}$  |
| 1    | <i>exponential decay</i> | <i>large</i> | <i>isotropy</i>   | 2.76    | 0.47  | 0.66  | 4           |
| 2    |                          |              | <i>anisotropy</i> | 2.58    | 0.33  | 0.75  | 10          |
| 3    |                          | <i>small</i> | <i>isotropy</i>   | 2.73    | 0.35  | 0.62  | 2           |
| 4    |                          |              | <i>anisotropy</i> | 2.79    | 0.19  | 1.27  | <i>none</i> |
| 5    | <i>constant mean</i>     | <i>large</i> | <i>isotropy</i>   | 2.89    | 0.15  | 1.43  | <i>none</i> |
| 6    |                          |              | <i>anisotropy</i> | 2.87    | 0.07  | 1.59  | <i>none</i> |
| 7    |                          | <i>small</i> | <i>isotropy</i>   | 2.96    | 0.10  | 1.87  | <i>none</i> |
| 8    |                          |              | <i>anisotropy</i> | 2.93    | 0.04  | 2.17  | <i>none</i> |

719