Standardization and assessment of geological descriptions from water well records, Greater Toronto and Oak Ridges Moraine areas, southern Ontario[1]

H.A.J. Russell, T.A. Brennand[2], C. Logan, and D.R. Sharpe
Terrain Sciences Division, Ottawa

Russell, H.A.J., Brennand, T.A., Logan, C., and Sharpe, D.R., 1998: Standardization and assessment of geological descriptions from water well records, Greater Toronto and Oak Ridges Moraine areas, southern Ontario; in *Current Research 1998-E; Geological Survey of Canada, p. 89–102.*

Abstract:
Archival drilling records from water wells, geotechnical, mineral exploration, and hydrogeological studies provide subsurface information for regional geological and hydrogeological investigations. This paper evaluates methods by which water well material descriptions may be standardized. In Ontario, material descriptions are reported in three attribute fields using 82 terms, thus theoretically permitting over 500 000 permutations. Materials descriptions are rationalized to ten classes then reclassified according to two methods, 1) first-attribute method (FAM), and 2) rule-based method (RBM). The first-attribute method is presently applied by hydrogeologists in southern Ontario and uses only the first attribute field; it is a simple, effective method able to broadly delimit aquifers and nonaquifers. The rule-based method applies conditional rules developed from regional geological models. This method is more geologically accurate, and is recommended where water well data are to be integrated into geological and hydrogeological investigations. Successful applications are summarized and general recommendations made.

Résumé :
Des données provenant de forage de puits d'eau et d'études géotechniques, hydrogéologiques et de prospection minière, renseignent sur les conditions souterraines nécessaires aux études géologiques et hydrogéologiques régionales. Le présent article évalue les méthodes permettant de normaliser les descriptions des matériaux des puits d'eau. En Ontario, les descriptions de matériaux sont consignées dans trois champs d'attributs à l'aide de 82 termes, ce qui, théoriquement, permet de faire plus de 500 000 permutations. La rationalisation des descriptions de matériaux a permis de distinguer dix classes qui ont par la suite été reclassées selon deux méthodes : (1) la méthode des premiers attributs et (2) la méthode basée sur des règles. La méthode des premiers attributs, actuellement appliquée par les hydrogéologistes dans le sud de l'Ontario, utilise uniquement le premier champ d'attributs. Il s'agit d'une méthode simple et efficace pouvant délimiter grossièrement les formations aquifères et non aquifères. La méthode basée sur des règles applique des règles conditionnelles élaborées à partir de modèles géologiques régionaux. D'un point de vue géologique, cette méthode est plus précise. Il est recommandé de l'appliquer là où les données sur les puits d'eau doivent être intégrées aux études géologiques et hydrogéologiques. Les applications réussies sont présentées brièvement dans le présent article, lequel contient également des recommandations d'ordre général.

[1] Contribution to the Oak Ridges Moraine NATMAP project
[2] Department of Geography, Simon Fraser University, Burnaby, British Columbia V5A 1S6

Contact Information: Dr. David Sharpe, Terrain Sciences Division, Geological Survey of Canada, 601 Booth St. Ottawa, Ontario, K1A 0E8, ph: (613) 992-3059,
fax: (613) 992-0190, e-mail: dsharpe@nrcan.gc.ca

INTRODUCTION

Bedrock lithology and sediment texture (materials) are important to geological, geotechnical, and hydrogeological subsurface investigations as both allow inference of depositional environments and both are important controls on geotechnical and hydrogeological parameters. These data are commonly attained by expensive drilling programs, with large quantities of subsurface information archived in hard copy and digital formats (e.g. Belanger, 1975; McCleneghan and Dilabio, 1995). The availability and quantity of such data varies by region, related to the nature of natural resource exploitation (agriculture, mineral, forestry), the degree of infrastructure development (highways, railroads, bridges), the population density, and the timing of development. In glaciated terrain the most extensive archival data sets are geotechnical reports and water well records (e.g. Belanger and Harrison, 1980). Appropriate utilization of these data in regional geological (e.g. Russell et al., 1996b) and hydrogeological investigations (e.g. Anonymous, 1994; Holysh, 1995; LeGrand and Rosen, 1998) can significantly reduce project costs.

THE PROBLEM

Public concern regarding sustainable water resources and water quality have resulted in renewed interest in aquifer delineation (Kehew et al., 1998) and classification (Kreye et al., 1994; Fagan et al., 1997). These objectives are best achieved through basin analysis methodologies (e.g. Miall, 1984; Eyles et al., 1985; Sharpe et al., 1992). The Oak Ridges Moraine NATMAP–Hydrogeology Project (e.g. Sharpe et al., 1996) in the Greater Toronto and Oak Ridges Moraine areas is an example of such a project (Fig. 1). This project is focusing attention on the need for a more accurate model of the regional three-dimensional geology and hydrostratigraphy of this glaciated terrain (e.g. Sharpe et al., 1997). To delineate stratigraphic units, local site investigations often entail drilling numerous, continuously cored drillholes through the thick overburden (up to 160 m, Fenco-MacLaren, 1994). This methodology is economically prohibitive for regional investigations. Rather, interpretations from expensive, strategically placed and continuously-cored drillholes (high-quality data) must be extended regionally by integration with existing, spatially extensive, archival data. To this end, a multicomponent database composed of a relational database, a GIS database, and flat file data assemblage, has been developed (Russell et al., 1996b; Brennand et al., 1997b; Brennand, 1998). A MOE water well data set, with about 33 000 water well records, is the largest single data contributor to the relational database. This paper explores the task of extracting meaningful and standardized geological descriptions from this data set.

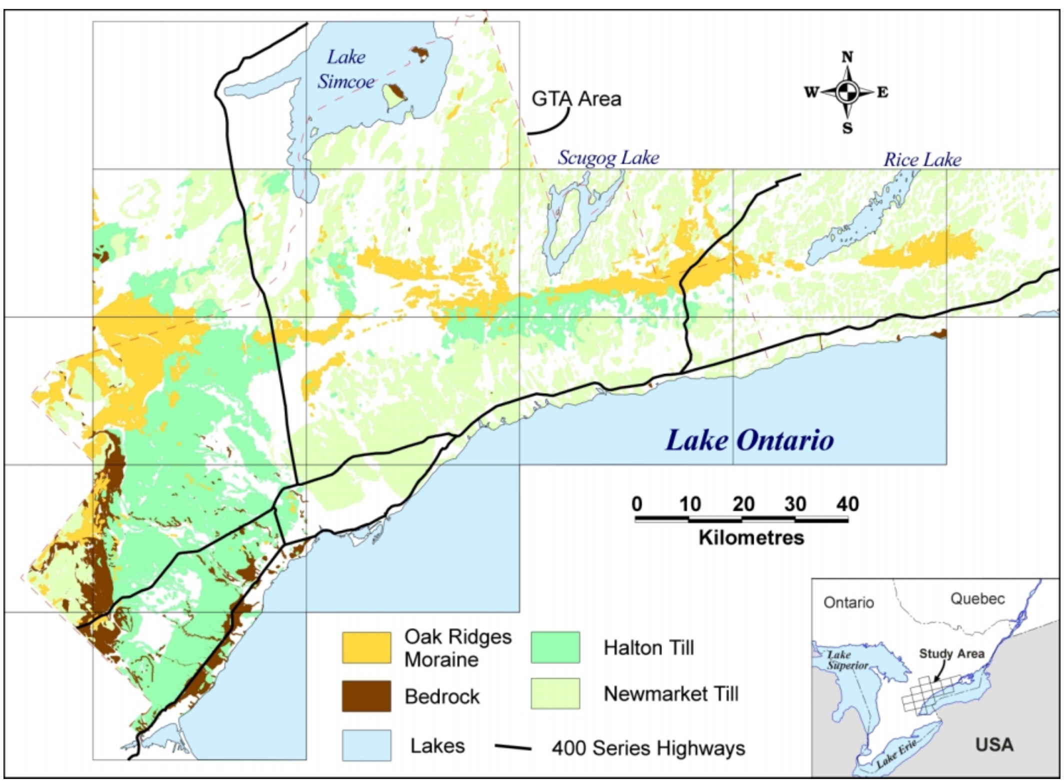

Figure 1. *Location map and simplified regional geology of Greater Toronto and Oak Ridges Moraine areas (modified from Sharpe et al., 1997).*

THE DATA: BACKGROUND TO THE MOE WATER WELL DATABASE

Legislation requiring well contractors to submit water well reports to the Ontario Department of Mines was passed in 1946 (Watt, 1952). In that year less than 500 reports were submitted from across the Province, but this number rapidly increased in succeeding decades. The management of these reports passed to the Ontario Water Resources Commission (OWRC) in 1956, and they have since published periodic summaries (e.g. Watt, 1961). A computerized database, the Water Well Information System (WWIS), was instituted by MOE in 1972 to aid input and retrieval of records (Mantha, 1988). This database (complete to 1992) contains information, on up to 212 parameters, for over 325 000 water wells across Ontario. In this paper only water wells records that fall within the study area and are accurately located (Kenny et al., 1997) (~33 000 wells with >142 000 geological units; Fig. 1) are analyzed; this data subset is henceforth called the MOE data set.

The MOE database contains 212 data fields per well. These fields can be subdivided into 1) header (location), 2) geology, 3) hydrogeology, and 4) construction categories (Table 1). Fields in the header category include location and elevation. The geology category contains material, depth, and description. The hydrogeology and construction categories contain fields relevant to the verification of sediment texture (e.g. flow rate and first screen depth). All fields are available for use with each well but, with the exception of location, many fields are incomplete (e.g. screen intervals are reported for only ~20 % of the water wells).

Table 1. Summary of data fields in the MOE water well database, grouped by four principal catagories. Numbers in brackets indiate multiple fields in MOE database.

Header	Geology	Hydrogeology	Construction
Well Number (2)	Unit - Depth to Top (24)	Piezometer Indicator (1)	Casings (18)
Municipality (2)	Unit - Colour (24)	Water - Depth Found (5)	Screens (8)
Concession – Range (4)	Unit - Materials (3x24)	Water - Kind (5)	Plugs (6)
Lot (2)		Test Method (1)	
Owner (1)		Pumping (2)	
Completion Date (3)		Levels (2)	
UTM Location (4)		Pumping/Recovery Indicator (1)	
Elevation (2)		Level During Pumping (4)	
Basin (5)		Flow Rate (1)	
Water Use (2)		Clear - Cloudy (1)	
Drill Method (1)		Recommended Setting (1)	
Data Source (1)		Recommended Rate (1)	
Contractor Code (1)		Specific Capacity (1)	
Date Received (3)		Final Status (1)	**total fields = 212**

The water well reports were primarily designed to protect the interests of the well owner, specifically with regard to well construction (Singer et al., 1997). Consequently, geological reporting (material description) has been given less attention. Material descriptions were captured in the database from these reports without alteration (Watt, 1952). A basic list of terms to be used for material description was not introduced until computerization in 1972 (S. Singer, pers. comm., 1998; Table 2). It should also be noted that geological descriptions in the MOE database are generally based on examination of drill chips and sediment flushed to the surface in the drilling process; rather than on continuously cored samples that are more commonly encountered by geologists and hydrogeologists.

The geological attributes consist of colour and material descriptors, the latter chosen from 82 terms (materials and descriptive terms) (Table 2). In the MOE database, material descriptors for each unit are applied to three fields (Table 1). Consequently, the number of possible descriptive permutations theoretically exceeds 500 000; the actual number used in the MOE data set exceeds 1800. Not all units in the MOE data set have descriptors entered in all three material fields: the first field is used for about 99% of the units, the second for 28%, the third for only 2.5%, leaving about 52% of all material fields blank (null). More importantly, the material for about 70% of the units is described in a single field only, whereas multiple fields are used in the geological description of about 30% of the units. Of the 49 materials specified (Table 2), 19 describe bedrock lithology, two are minerals, and those remaining either describe sediment texture (e.g. sand, gravel) or the character of the drilled hole (e.g. previously drilled). Of the 33 'descriptive terms' specified, 27 are of questionable geological value (asterisks, Table 2). A decision regarding geological validity was based on the drilling technique and the drillers ability to resolve salient details concerning units encountered, e.g. 'thin', and 'cemented'. Other terms such as 'dirty' or 'loose' simply fail to convey information that could be extracted in a meaningful way. The frequency of usage of each material descriptor (materials and descriptive terms, Table 2) varies; only two descriptors are used in more than 10% of the fields, and 13 descriptors in more than 1%. Combined, these 15 descriptors account for about 96% of all descriptor entries (Table 2).

There is a clear need for data standardization in this extensive data set. Material descriptions require standardization when either 1) there is a need for integration and/or comparison between disparate data sets; or 2) the number of geological descriptions within a single data set thwarts internal comparisons. The Oak Ridges Moraine NATMAP–Hydrogeology Project faced both of these challenges with the MOE water well data set.

Table 2. Geological descriptions used in the MOE water well database. Numbers in brackets indicate per cent usage in all three material fields in MOE data set. Asterisk indicates modifiers of questionable lithological value.

Materials					
00	unknown (49.92)	17	shale (0.87)	34	till (0.01)
1	fill (0.17)	18	sandstone (0.02)	35	wood fragments (0.00)
2	topsoil (3.54)	19	slate (0.00)	36	basalt (0.00)
3	muck (0.08)	20	quartzite (0.00)	37	chert (0.00)
4	peat (0.02)	21	granite (0.01)	38	conglomerate (0.00)
5	clay (16.12)	22	greenstone (0.00)	39	feldspar (0.00)
6	silt (1.39)	23	previously bored (0.65)	40	flint (0.00)
7	quicksand (0.29)	24	previously drilled (0.07)	41	gneiss (0.00)
8	fine sand (1.72)	25	overburden (0.01)	42	greywacke (0.00)
9	medium sand (4.74)	26	rock (0.05)	43	gypsum (0.00)
10	coarse sand (1.09)	27	- (0.00)	44	iron formation (0.00)
11	gravel (5.33)	28	sand (4.21)	45	marble (0.00)
12	stones (3.31)	29	fine gravel (0.05)	46	quartz (0.00)
13	boulders (0.79)	30	medium gravel (0.02)	47	schist (0.00)
14	hardpan (0.51)	31	coarse gravel (0.08)	48	soapstone (0.00)
15	limestone (0.75)	32	pea gravel (0.00)		
16	dolomite (0.01)	33	marl (0.00)		
Descriptive Terms					
60*	cemented (0.06)	71*	fractured (0.01)	82	shaly (0.03)
61	clayey (0.00)	72	gravelly (0.06)	83*	sharp (0.00)
62*	clean (0.08)	73*	hard (1.01)	84	silty (0.07)
63*	coarse-grained (0.01)	74*	layered (0.25)	85*	soft (0.58)
64*	crystalline (0.00)	75*	light-coloured (0.02)	86*	sticky (0.01)
65*	dark-coloured (0.03)	76*	limy (0.00)	87	stoney (0.08)
66*	dense (0.27)	77*	loose (0.48)	88*	thick (0.00)
67*	dirty (0.04)	78*	medium-grained (0.03)	89*	thin (0.00)
68*	dry (0.10)	79*	packed (0.30)	90*	very (0.03)
69*	fine-grained (0.02)	80*	porous (0.17)	91*	water-bearing (0.09)
70*	fossiliferous (0.01)	81	sandy (0.28)	92*	weathered (0.00)

STANDARDIZATION PROCEDURES FOR GEOLOGICAL DESCRIPTIONS

The existing literature contains little information regarding standardization procedures for geological descriptions such as those contained in water well data sets. Several sediment-coding schemes have been proposed (i.e. lithofacies codes, Miall, 1977; Eyles et al., 1983). Whereas these may be appropriate for classifying sediment in vertical exposures or in continuously-cored drillholes, such schemes generally require detailed textural and structural information and are thus inappropriate for standardizing the MOE water well data set. Consequently, the project defined and assessed its own standardization procedures. This paper addresses two phases of data standardization, 1) rationalization and 2) reclassification, and assesses two methods of attaining the latter.

Rationalization

Following an audit of material descriptors in the MOE data set, it became apparent that the first step in standardization should be to rationalize descriptors. Such rationalization is justified when the low frequency of usage of many individual descriptors is reviewed (Table 2); these usage statistics suggest that the data collection process, predominately wash-boring, was unable to attain the level of detailed description implied in Table 2. The rationalization process was achieved by 1) removing selected descriptive terms (asterisks, Table 2); 2) simplifying adjective-noun combinations; and 3) reassigning the remaining descriptors to appropriate groups (Table 3). Only three material terms were found to have ambiguous meaning: hardpan, marl, and muck. Their assignment to rationalized descriptors (Table 3) was based on conversations with drillers and consultants, and regional mapping experience. Rationalization resulted in a reduction of the 82 descriptors (Table 2) to ten major descriptors and two subcategories (Table 3), and reduced the number of theoretically possible descriptive permutations from more than 500 000 to more than 1400; the actual number used in the MOE data set was 464.

Table 3. Rationalized descriptors for MOE water well materials. Numbers in brackets indicate per cent usage in all three material fields in MOE data set.

Rationalized Descriptor	MOE Dataset Descriptor	
1 bedrock (0.1%)	18 sandstone	40 Flint
	20 quartzite	41 gneiss
	21 granite	42 greywacke
	22 greenstone	43 gypsum
	26 rock (bedrock)	44 iron fm.
	36 basalt	45 marble
	37 chert	46 quartz
	38 conglomerate	47 schist
	39 feldspar	48 soapstone
1.1 limestone (0.8%)	15 limestone	16 dolostone
1.2 shale (0.9%)	17 shale	82 shaly
	19 slate	
2 gravel (9.7%)	11 gravel	31 coarse gravel
	12 stones	32 pea gravel
	13 boulders	72 gravelly
	29 fine gravel	87 stoney
	30 medium gravel	
3 sand (12.3%)	7 quicksand	10 coarse sand
	8 fine sand	28 sand
	9 medium sand	81 sandy
4 silt (1.5%)	6 silt	84 silty
5 clay (16.1%)	5 clay	61 clayey
6 diamicton (0.5%)	14 hardpan	34 till
7 organic (0.1%)	3 muck	33 marl
	4 peat	35 wood frags.
8 fill (3.7%)	1 fill	25 overburden
	2 topsoil	
9 previously dug (0.7%)	23 previously bored	24 previously drilled
99 null (53.6%)	27 -	00 unknown

Reclassification

The second phase of standardization is reclassification. This process assigns a single, standardized material description to each water well unit. Two reclassification methods for unit materials are described, the first-attribute method and the rule-based method.

First-attribute method (FAM)

The first-attribute method reclassifies water well materials by considering only the rationalized material descriptions in the first material field (i.e. disregarding the second and third fields). This method reduces the number of possible material descriptions to those in the rationalized categories (Table 3). This is a method often implemented by hydrogeologists (e.g. Beckers and Frind, 1997; Holysh and Kassenaar, 1997; Martin et al., 1997).

Rule-based method (RBM)

The rule-based method assigns a single, material descriptor to each water well unit by applying conditional rules to descriptors and, where necessary, integrating descriptors across multiple fields (~30% of well units). Conditional rules were developed from an understanding of the regional geology. Following extensive regional mapping in the Oak Ridges Moraine area, a regional geological model was developed (e.g. Sharpe et al., 1996). This knowledge resulted in the formulation of an inclusive, yet simple, list of regionally relevant and geologically accurate material descriptions (Table 4). The task of the rule-based method was to apply conditional rules in order to reclassify unit materials to these geologically meaningful categories. For example, diamicton (till) is regionally extensive in southern Ontario, both at the surface and in the subsurface (e.g. Sharpe et al., 1996), yet it is only explicitly described in about 0.05% of units in the data set (diamicton, Table 3); if 'diamicton' is to be better extracted from this data set, then the designation must be applied from conditional rules controlling descriptor integration across multiple fields.

Conditional rules (programming statements) were defined and applied in four steps: I) simplify attribute strings, II) apply global rules, III) apply bedrock rules, and IV) apply sediment texture rules (Table 5). These routines were performed sequentially. Two subroutines were then applied to determine A) diamicton texture and B) bedrock lithology (Table 5). Routines I–III are discussed in more detail below; routines IV, and A and B are self-explanatory.

Routine I was designed to simplify descriptor strings (Table 5, routine I). This routine involved removal of duplicate descriptors (often an artifact of the rationalization process), deletion or modification of descriptors when in geologically unlikely combinations, and removal of leading null fields. The net effect of this routine was to reduce the occurrence of 'clay' which, based on regional mapping (e.g. Sharpe et al., 1997), is over-emphasized in the MOE data set.

Routine II applied six global rules (Table 5, routine II). Most of these rules are self-explanatory. In an attempt to balance the under-representation of regionally relevant, geologically accurate descriptions such as 'diamicton' and 'gravel', conditional rules were applied that emphasized these designations where warranted (Table 5, routines II.2 and II.3).

Table 4. Geological (material) description informaed by regional geological knowledge and applied in rule-based method. Percentage usage reported for all three material fields in MOE data set.

	Geologic Description	Usage %
99	no obvious material code	0.15
11	covered; missing; previously bored	0
10	fill (incl. topsoil, waste)	13.29
9	organic	0.15
8	clay, silty clay	21.21
7	silt, sandy silt, clayey silt	2.52
6	sand, silty sand	28.68
5	gravel, gravelly sand	14.89
4	clay-clayey silt diamicton	0.22
4-1	clay-clayey silt diamicton, stoney	
3	silt-sandy silt diamicton	
3-1	silt-sandy silt diamicton, stoney	13.90
3-3	diamicton, texture unknown	
2	silty sand-sand diamicton	0.41
2-1	silty sand-sand diamicton, stoney	
1	bedrock	
1-1	limestone	
1-2	shale	4.58
1-4	dolomite	
1-5	potential bedrock	
1-7	interbedded limestone/shale	

Routine III applied bedrock rules (Table 5, routine III). This routine was only applied to units that were located at the bottom of wells, or to units whereby all deeper units were classified as bedrock (i.e. it was an iterative routine starting at the lowest unit in every well). The 5 m depth-below-surface condition is based on the assumption that the driller can most probably correctly identify bedrock at shallow depth, and would be more likely to make mistakes at greater depths (Table 5, routine III.2 and III.3). This rule ensures a conservative estimate; it prevents under-estimating the depth to bedrock across the region (e.g. Brennand et al., 1997a).

Examples of the reclassification of geological units based on the application of conditional rules (Table 5) is presented in Table 6. A complete listing of each rationalized attribute string and its reclassification is presented in Russell et al. (1998a).

ASSESSMENT OF STANDARDIZATION (RECLASSIFICATION) PROCEDURES

Testing is required to assess confidence in the material reclassifications of units achieved by the first-attribute method and the rule-based method. Two tests are presented and discussed, 1) a comparison of the geological reclassification of well units at 1 m depth below the surface with the surficial geology map unit (defined as the material at 1 m depth below the surface, e.g. Sharpe et al., 1997) at each well location, and 2) an assessment of the reclassified material at the first screen depth in the well. The latter test is based on the assumption that the first screen is located in aquifer material (sand or gravel).

Comparison to surficial geology

The Oak Ridges Moraine NATMAP–Hydrogeology Project regional surficial geology compilation map was the benchmark for this test (Sharpe et al., 1997; simplified in Fig. 1). The results of the comparison for four of the major map units are presented in Figure 2 (data for all map units are presented in Russell et al., 1998a).

The Oak Ridges Moraine is a regional aquifer (Fig. 1). Its sediments generally range from silty sand to gravel; minor beds of clay (Fig. 3; Gilbert, 1997) and diamicton have been observed (Sharpe et al., 1997). In two boreholes through the Oak Ridges Moraine only about 1% of the total thickness was clay, whereas up to about 80% was silt and sand (Russell et al., 1998c). Using 'sand' and 'gravel' as acceptable materials for this map unit, the rule-based method performs marginally better than the first-attribute method, especially with regard to identifying gravel in the Oak Ridges Moraine. Both methods are able to appropriately describe this map unit and confirm its aquifer status in more than 50% of the wells drilled through it.

Table 5. Conditional rules applied by the rule-based method of data standardization to MOE data set.

i. Simplify Descriptor Strings:

- I.1. If organic has an accompanying descriptor, then organic is treated as null (see I.5).
- I.2. If clay occurs with sand or gravel, then clay is treated as silt.
- I.3. If bedrock descriptor is not for the last unit of the well or not with continuous bedrock beneath, then bedrock is treated as gravel.
- I.4. If clay or silt is with shale, then clay and/or silt is treated as null (see I.5).
- I.5. Remove duplicate attributes and spaces; eliminate all leading null fields.

II. Apply Global Rules

- II.1. If a single descriptor, then the description is based directly on that descriptor (see Table 4).
- II.2. If till in any field, then treat as diamicton (see Subroutine A below).
- II.3. If gravel is in any field with no bedrock and clay is not first descriptor, then treat as gravel.
- II.4. If fill and previously dug in any field are without bedrock, then treat as fill; else with bedrock, then treat as potential bedrock.
- II.5. If all fields are null, then treat as no obvious material code.
- II.6. If previously dug or fill are in any field except with bedrock, then treat as fill.

III. Apply Bedrock Rules (Apply if last unit in well or if continuous bedrock beneath)

- III.1. < 5 m depth, bedrock anywhere then code = bedrock (see Subroutine B below for categories)
- III.2. > 5 m depth, bedrock anywhere and gravel anywhere then code = gravel
- III.3. > 5 m depth, bedrock anywhere and sand/silt/clay and no gravel then code = diamicton (see Subroutine A below)

IV. Apply Sediment Texture Rules

- IV.1. Attribute 1 = sand with attribute 2/3 = silt or clay and no gravel, then treat as sand
- IV.2 Attribute 1 = silt with attribute 2/3 = sand or clay and no gravel, then treat as silt
- IV.3 Attribute 1 = clay with attribute 2/3 = sand or silt and no gravel, then treat as silt

Subroutine A: Determine Diamicton Texture (Use textural sand-silt-clay attribute in highest attribute position)

1. If sand, then treat as silty sand diamicton
2. If silt, then treat as silt diamciton
3. If clay, then treat as clay silt diamicton
4. If no texture indicated, then treat as silt diamicton

Subroutine B: Determine Bedrock Lithology

1. If bedrock in any field and not with limestone or shale, then treat as bedrock
2. If limestone in any field and not with shale, then reclassify as limestone
3. If shale in any field and not with limestone, then reclassify as shale
4. If limestone and shale in any field, then reclassify as interbedded limestone - shale

Halton Till is a complex fine-grained unit, often forming an aquitard (nonaquifer) in the western part of the region. It is best characterized as a stone-poor (<1%), silt to clay diamicton, interbedded with glaciolacustrine silt and clay; the diamicton component is thicker in the western part of the region and becomes thinner and more sandy in the east where it overlies the Oak Ridges Moraine. In comparison, Newmarket Till is generally an overconsolidated stone-rich (5-10%), silty sand diamicton; relatively thin sand and gravel beds may be observed within it. Newmarket Till is generally characterized as a dissected regional aquitard (e.g. Sharpe et al., 1996). The most accurate geological description for both of these units is ‘diamicton’ (till). For both map units the rule-based method clearly out-performs the first-attribute method, but in neither case exceeds 30% accuracy (Fig. 2). The better performance of the rule-based method is expected as the conditional rules were, in part, designed to integrate descriptors and generate more meaningful geological descriptions. If the descriptions ‘sand’ and ‘gravel’ are also accepted for Newmarket Till, the rule-based method continues to out-perform the first-attribute method, but if the descriptions ‘clay’ and ‘silt’ are accepted for Halton Till, then this performance reverses (first-attribute method, 70.5%; rule-based method, 58.4%; Fig. 2). Despite numerous misclassification (Fig. 2), both methods confirm the nonaquifer status of both the Halton Till and Newmarket Till, the rule-based method being more geologically accurate.

When compared with bedrock polygons both methods have a poor correlation (13%, Fig. 2). Despite the application of conservative rules, the rule-based method correlates marginally better than the first-attribute method. The poor correlation of both methods with bedrock polygons may reflect the character of weathered silt and shale of the Queenston and Georgian Bay formations (e.g. White, 1975), both of which may

Table 6. Examples of geological reclassification based on the application of conditional rules (RBM = rule-based method; Table 5)

Descriptor 1	Descriptor 2	Descriptor 3	RBM reclassification
Bedrock			Bedrock (1)
Clay	Gravel		Silt diamicton (3)
Silt	Shale		shale (1-2)
Gravel	Sand		Gravel (5)

have been understandably misclassified by drillers as 'clay' (Fig. 2). Alternatively, it may suggest that much of the area within bedrock polygons actually has a thin cover (~1 m thick) of surficial sediments. This issue of cartographic resolution was tested independently by comparing well unit reclassification (rule-based method) with known bedrock outcrops within bedrock map polygons. This comparison indicates that within 100 m of an outcrop, about 60% of the water wells intercept bedrock at less than 5 m, but this frequency declines to 39% at 500 m (Russell et al., 1998a). From this comparison it would appear that assessment of the performance of the bedrock reclassification is hindered by cartographic resolution.

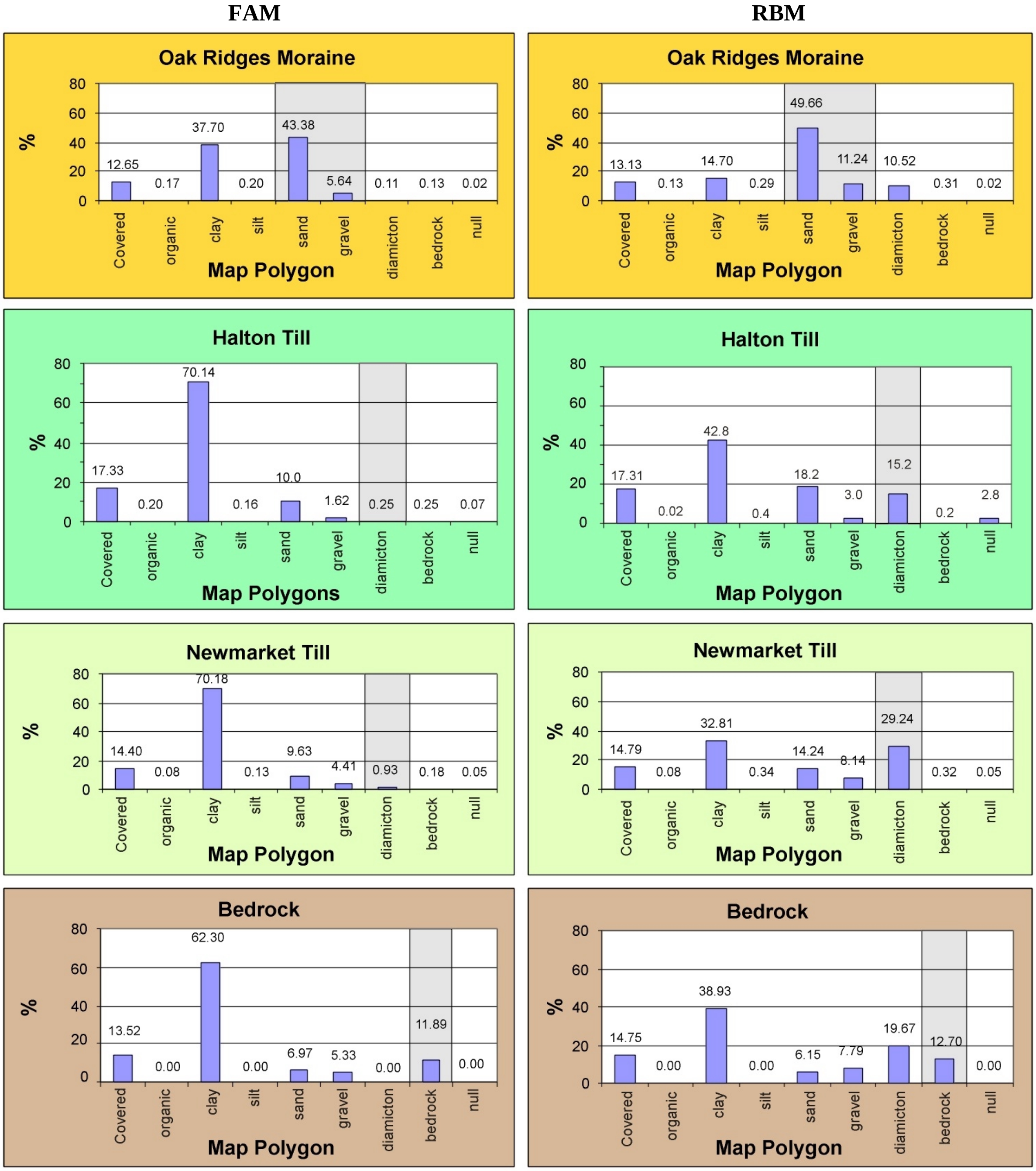

Figure 2. *Comparision ofsurficial geology map units (Sharpe et al., 1997) with MOE unit descriptions (from 1m depth below surface) reclassified by the first attribute method (left) and the rule based method (right). Shaded zones indicate the most appropriate reclassification. FAM – first-attribute method, RBM – rule-based method.*

The predominance of 'clay' as the description for all map units is overwhelming in the first-attribute method; clay exceeds all other descriptions by more than a 2:1 ratio, except within the Oak Ridges Moraine map unit were sand is dominant (Fig. 2). The first-attribute method truncates descriptions at the first field. For entries in the first field the term 'clay' accounts for 31% of single attribute entries and 70% of multiple attribute entries. Consequently, by ignoring entries in the second and third fields, the first-attribute method is expected to over-emphasize 'clay'. Furthermore, as material descriptions are based on drill chips and sediment flushed to the surface in the drilling process, inaccurate description (e.g. of wet silt or stone-poor diamicton) should be expected throughout the data set.

Assessment at first screen depth

Drillers place screens within water-bearing units to prevent water well resedimentation during pumping. The screen location can thus be used as a proxy for aquifer locations and thus for sand and/or gravel units. An assessment of all reclassified well units coinciding with first screen depth indicates that both the first-attribute method and the rule-based method reclassify more than 85% as sand and/or gravel. However, the rule-based method is more likely to reclassify a unit as gravel rather than sand (Fig. 4).

In summary, the first-attribute method permits a moderately reliable bipartite assignment of sediment to either aquifer or nonaquifer status. This method is generally adequate for about 70% of the units (described by a single field in the MOE data set), but may misclassify about 30% of the units (described by multiple fields in the MOE data set). The rule-based method allowed material standardization (reclassification) to be informed by knowledge of the regional geology. This method again allows moderately reliable aquifer and nonaquifer assignment, but is more geologically accurate.

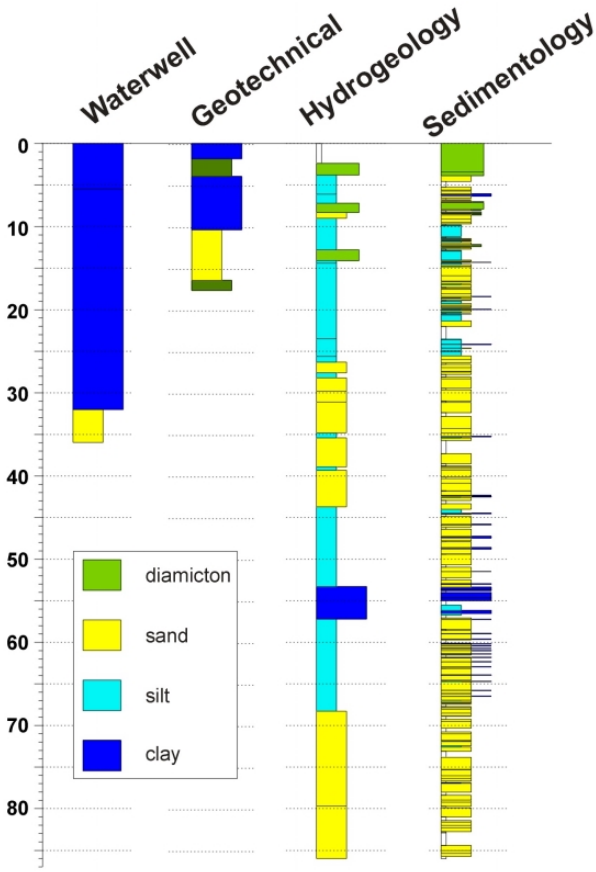

Figure 3. *Comparison of the geological resolution available from four types of drilling and logging procedures. Drillholes are from within several kilometres of one another. Note the contrast in unit thickness and variability of sediment textures. Depth scale to left is in metres.*

DISCUSSION

Limitations of the MOE database for geological and hydrogeological investigations

The MOE water well database is a valuable resource for regional geological and hydrogeological investigations. However, standardization procedures and assessments tested by the Oak Ridges Moraine NATMAP–Hydrogeology Project highlight the limitations of integrating this database into geological or hydrogeological investigations. These limitations include a) location accuracy, b) coverage, and c) geological accuracy and resolution.

Location accuracy

Water well locations are initially reported to the MOE using Lot and Concession designation and a sketch map. This information is then converted to a Universal Transverse Mercator (UTM) co-ordinate, and an elevation is assigned by MOE staff from Ontario Bureau of Mines maps. Spatial comparisons between UTM and Lot and Concession co-ordinates, and between assigned elevation and the Oak Ridges Moraine digital elevation model reported that 27% of well records in the MOE database had planimetric and/or elevation inaccuracies (Kenny et al., 1997).

Coverage

Water well records have a limited vertical coverage as they are drilled to exploit aquifers; in areas of shallow aquifers few wells continue to depth. The Oak Ridges Moraine is a regional aquifer and, as such, less than one third of the total sediment thickness to bedrock is recorded by the water wells intersecting it (Fig. 5; Brennand et al., 1997a; Russell et al., 1998b). Water wells are thus poor data sets to use for assessing potential deeper aquifers and thus for regional groundwater inventory . Whereas water well records can form the most extensive spatial data set in regional geological and hydrogeological investigations, such records may be sparse in metropolitan areas (e.g. Toronto).

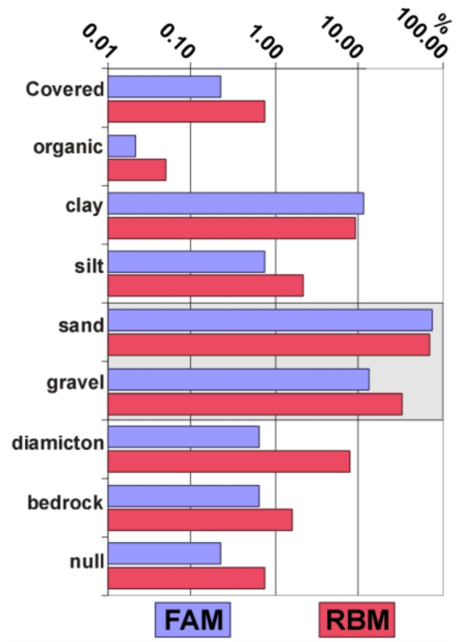

Figure 4. *Assessment of reclassified MOE material oresent at the first screen depth, first attribute method (FAM), rule-based method (RBM). Shaded zones indcate the most appropriate reclassficication. Note percent scale is logarithmic*

Geological accuracy and resolution

The glaciated terrain of southern Ontario is underlain by Paleozoic bedrock, with a terminal grain-size of silt (Dreimanis and Vagners, 1971), and thus silt may be expected to dominate glacial sediments above bedrock. This conclusion is supported by detailed sedimentological reports from continuously-cored drillholes in the Oak Ridges Moraine area. These reports clearly document a very low clay:silt ratio (e.g. Gilbert, 1997; Fig. 3). For two boreholes in the Humber River watershed, clay units formed less than 2% of the total sediment thickness (Russell et al., 1998c). In contrast, disregarding the null fields in the MOE data set, about 40% of all entries in the three material fields are 'clay'. This comparison thus highlights a geologically inaccurate overuse of 'clay' as a descriptor in the MOE data set. Possible reasons for this overuse include 1) likely underestimation of grain size due to the liquified nature of the samples from which drillers make descriptions (a product of the wash-boring drilling technique); and 2) no training requirement in regionally relevant geological descriptions for drillers.

Geological mapping (Sharpe et al., 1997) has shown that diamicton (till) outcrops across about 42% of the region and that these surficial units can attain thicknesses of 30 m, yet in the MOE data set 'till' accounts for only 0.05% of entries in material fields. The rule-based method was able to improve extraction of this regionally relevant, geologically accurate, heterogeneous material from the MOE data set by applying conditional rules and integrating descriptors across multiple fields.

Comparison of geological units from water well records with those from geological and hydrogeological reports highlights a clear disparity in unit resolution (Fig. 3). This disparity likely results from differences in drilling technology, wash-bored and continuously cored drillholes. In the Oak Ridges Moraine area sediment descriptions from continuously cored drillholes (e.g. Gilbert, 1997; Russell et al., 1998c) relate to units with thicknesses generally on the order of decimetres, and for clay units generally less than 1 cm. Material descriptions from water wells relate to units metres to tens of metres thick (Fig. 3). Consequently, the use of water well records for geological purposes should be purely supplemental.

Successful applications of the MOE database

The quantity and spatial distribution of water wells make them an attractive supplemental data source for both geological (e.g. Brennand et al., 1997b) and hydrogeological investigations (e.g. Kreye et al., 1994; Holysh, 1995; Fagan et al., 1997; Singer et al., 1997). However, appropriate integration and application of this data lies in recognizing its limitations and, perhaps, in devising innovative approaches to extracting meaningful information.

The recognition of data limitations and the enhancement of geological accuracy (utilization of the rule-based method) in the MOE database underpin the successful applications of this database in the Oak Ridges Moraine NATMAP–Hydrogeology Project. Successful applications have included 1) structural surface maps (e.g. bedrock topography; Brennand et al., 1997a), 2) isopach maps (e.g. sediment thickness; Russell et al., 1998b), and 3) construction of materials cross-sections to a) delineate aquifers and nonaquifers and thus facilitate the interpretation of stream gauging data (Hinton and Bowen, 1997) and b) interpolate subsurface basin stratigraphy (Russell et al., 1996a).

SUMMARY AND RECOMMENDATIONS

Two procedures for standardizing geological descriptions in the MOE database have been described and assessed. The first-attribute method (FAM) is a simple, effective method for characterizing units described by a single material field, but fails to capitalize on the additional information provided in units described by multiple fields. This procedure is able to broadly delimit aquifers and nonaquifers. The rule-based method (RBM) benefits from regional geological knowledge; the conditional rules applied here should be modified as knowledge is gained and should be regionally specific. The rule-based method is able to broadly delimit aquifers and nonaquifers, and is more geologically accurate. This procedure is recommended where water well data are to be integrated into geological and hydrogeological investigations. In the latter case, this may be important if regional hydrogeological parameters (e.g. hydraulic conductivity) are to be estimated from material descriptions. Limitations to the application of the MOE database in geological and hydrogeological investigations include: 1) location accuracy (Kenny et al., 1997); 2) coverage; and 3) geological accuracy and resolution. As water well descriptions have recognized regional variations, any rationalization and coding process should first carefully screen data sets.

Water well databases are valuable data sets for regional geological and hydrogeological investigations. Despite its limitations, the MOE database has become a cornerstone for the assessment of subsurface conditions by hydrogeologists, and for the creation of municipal groundwater management and extraction plans by Ontario hydrogeologists and planners. Most criticisms of such databases relate to the simple fact that they were never intended for such use. Consequently, in order to enhance water well databases for geological and hydrogeological purposes several recommendations are proposed with regard to standardization of geological descriptions: 1) reduce the attribute fields to two only, 2) reduce the number of material descriptors to about 10-12 (e.g. Tables 3 and 4), 3) implement a standardized description form, and 4) implement a training session on geological descriptions. The value of standardized reporting of water well data has been recognized at the national scale (Gilliland, 1990). Standardization of geological descriptions will reduce confusion, duplication of effort, and cost. All revisions to reporting protocol should be reviewed by user groups, and by provincial or state and national groundwater associations.

ACKNOWLEDGMENTS

The Oak Ridges Moraine Project has been funded by the GSC National Mapping and Hydrogeology programs. External funding has been received from a number of Ontario provincial ministries, notably the Ontario Ministry of Environment. The invaluable assistance of I. Pawlowski and F. Johnson is greatly appreciated. Critical reviews by R.J. Belanger and M. Hinton helped improve the paper. The assistance of P. Stacey and S. Lambert are appreciated.

REFERENCES

Anonymous
1994: Long-term water strategy, GW-1 additional groundwater task; prepared for Regional Municipality of Waterloo by CH2M Hill, 75 p.
Beckers, J. and Frind, E.
1997: The role of geology in the uncertain response of a groundwater system; *in* Geological Association of Canada, Program with Abstracts, Ottawa'97, May, 1997, Ottawa, Ontario, p. A-9.
Belanger, J.R.
1975: UGAIS data record instruction manual; Geological Survey of Canada, Open File 292, 11 p.
Belanger, J.R. and Harrison, J.E.
1980: Regional geoscience information: Ottawa-Hull; Geological Survey of Canada, Paper 77-11, 18 p.
Brennand, T.A.
1998: Urban geology note: Oshawa Ontario; *in* Urban Geology of Canadian Cities, (ed.) P.F. Karrow; Geological Association of Canada, Special Paper 42.
Brennand, T.A., Moore, A., Logan, C., Kenny, F.M., Russell, H.A.J., Sharpe, D.R., and Barnett, P.J.
1997a: Bedrock topography of the Greater Toronto and Oak Ridges Moraine areas, southern Ontario; Geological Survey of Canada, Open File 3419, scale 1: 200 000.
Brennand, T.A., Russell, H.A.J., and Logan, C.
1997b: Development of a regional geoscience database: application to hydrogeologic investigations, Oak Ridges Moraine area, southern Ontario; *in* Geological Association of Canada, Program with Abstracts, Ottawa'97, May, 1997, Ottawa, Ontario, p. A-17.
Dreimanis, A. and Vagners, U.
1971: The bimodal distribution of rock and mineral fragments in basal tills; *in* Till-a symposium, (ed.) R.P. Golthwait; Ohio State University Press, 237-250 p.

Eyles, N., Clark, B.M., Kaye, B.G., Howard, K.W.F., and Eyles, C.H.
1985: The application of basin analysis techniques to glaciated terrains; an example from the Lake Ontario Basin, Canada; Geoscience Canada, v. 12, p. 22-32.
Eyles, N., Eyles, C.H., and Miall, A.D.
1983: Lithofacies types and vertical profile models: an alternative approach to the description and environmental interpretation of glacial diamict and diamictite sequences; Sedimentology, v. 30, p. 393-410.
Fagan, F., Bourque, E., Lefebvre, F., Michaud, Y., Martel, R., Bolduc, A.M., Parent, M., and Boisvert, E.
1997: Hydrogeological mapping applied to land use planning: the Laurentian Piedmont Project; *in* Geological Association of Canada, Program with Abstracts, Ottawa '97, May 19-21, 1997, p. A-48.
Fenco-MacLaren, I.
1994: IWA landfill site search, Metro/York region: Step 6 hydrogeological report sites V3B, V4A, V4D; Interim Waste Authority Report, prepared by M. Gomer, Fenco MacLaren Inc., January 1994, Willowdale, Ontario.
Gilbert, R.
1997: Glaciolacustrine sedimentation in part of the Oak Ridges Moraine; Geographie physique et Quaternaire, v. 7, p. 55-66.
Gilliland, J.
1990: Report on federal/provincial workshop on groundwater data banks; National Hydrology Research Centre, Saskatoon, Saskatchewan, p. 75.
Hinton, M.J. and Bowen, G.S.
1997: Stream baseflow surveys to identify regional groundwater discharge and chemistry in the Oak Ridges Moraine; *in* Geological Association of Canada, Program with Abstracts, Ottawa '97, May 19-21, 1997, Ottawa, Ontario, p. A-68.
Holysh, S.
1995: Halton Aquifer Management Plan, Phase 1 report, background hydrogeology; Regional Municipality of Halton, p. 92.
Holysh, S. and Kassenaar, D.C.
1997: Regional ground water modelling of complex bedrock valley aquifer systems; *in* Geological Association of Canada, Program with Abstracts, Ottawa'97, May, 1997, Ottawa, Ontario, p. A-69.
Kehew, A., Straw, W., Thomas, N., Nicks, L., Kendzierski, S.K., Gardner, R.C., Flint, A.C., Hughes, L.D., and Baker, C.
1998: Glacial terrain mapping with hydrogeological applications in St. Josephs County, Michigan; *in* Abstracts with Program, Northeast Sectional Meeting of the Geological Society of America, Portland, Maine, March, 1998.
Kenny, F.M., Chan, P., and Hunter, G.
1997: Quality control of the positional accuracy of records in Ontario's water well database using automated GIS techniques; *in* Conference Proceedings Geomatics in the Era of RADARSAT GER 97, Ottawa, May 24-30, 1997.
Kreye, R., Ronneseth, K., and Wei, M.
1994: An aquifer classification system for groundwater management in British Columbia; *in* 6th National Drinking Water Conference, Victoria, British Columbia.
LeGrand, H.E. and Rosen, L.
1998: Putting hydrogeology site studies on track; Ground Water, v. 36, p.193.
Mantha, L.
1988: Ministry of the Environment Water Well Information System – selecting, reporting and plotting of wells and their characteristics; Ontario Ministry of Environment and Energy, 29 p.
Martin, P.J., Guiger, N.,and Van Vliet, D.J.
1997: Regional modelling of complex multi-aquifer systems; *in* Geological Association of Canada, Program with Abstracts, Ottawa'97, May, 1997, Ottawa, Ontario, p. A-98.
McCleneghan, M.B. and Dilabio, R.N.W.
1995: Overburden database compilation, Timmins, Ontario (42A/11, 12,13,14); Geological Survey of Canada, Open File 3086, 12 p.
Miall, A.D.
1977: A review of the braided-river depositional environment; Earth-Science Reviews, v. 13, p.1-62.
1984: Principles of Sedimentary Basin Analysis; Springer-Verlag, New York, 490 p.
Russell, H.A.J., Brennand, T.A., Kenny, F.M., Sharpe, D.R., and Barnett, P.J.
1996a: The application of GIS to a Quaternary basin analysis, Oak Ridges Moraine area, southern Ontario; *in* Geomatics 1996, 8th International Conference on Geomatics, May 28-30 (abstract), Ottawa, Ontario.
Russell, H.A.J., Brennand, T.A., Logan, C., and Sharpe, D.R.
1998a: Coding sediment descriptions fom drilling records: documentation of MOE water well coding Greater Toronto and Oak Ridges Moraine areas southern Ontario; Geological Survey of Canada, Open File 3616, 30 p.
Russell, H.A.J., Logan, C., Brennand, T.A., Hinton, M., and Sharpe, D.R.
1996b: A regional geoscience database: an example from the Oak Ridges Moraine NATMAP / Hydrogeology Project; *in* Current Research 1996-E; Geological Survey of Canada, p. 191-198.
Russell, H.A.J., Moore, A., Logan, C., Kenny, F., Brennand, T.A., Sharpe, D.R., and Barnett, P.J.
1998b: Sediment Thickness of the Greater Toronto and Oak Ridges Moraine areas, southern Ontario; Geological Survey of Canada, Open File 2892, scale 1:200 000.
Russell, H.A.J., Sharpe, D.R., and Arnott, R.W.C.
1998c: Sedimentology of the Oak Ridges Moraine, Humber River Watershed, southern Ontario: a preliminary report; *in* Current Research 1998-C; Geological Survey of Canada, p. 155-166.

Sharpe, D.R., Barnett, P.J., Brennand, T.A., Finley, D., Gorrell, G., Russell, H.A.J., and Stacey, P.
1997: Surficial geology of the Greater Toronto and Oak Ridges Moraine areas compilation map sheet; Geological Survey of Canada, Open File 3062, scale 1:200 000.

Sharpe, D.R., Dyke, L.D., Hinton, M.J., Pullan, S.E., Russell, H.A.J., Brennand, T.A., Barnett, P.J., and Pugin, A.
1996: Groundwater prospects in the Oak Ridges Moraine area, southern Ontario: application of regional geological models; *in* Current Research 1996-E; Geological Survey of Canada, 181-190 p.

Sharpe, D.R., Pullan, S., and Warman, T.
1992: A basin analysis of the Wabigoon area of Lake Agassiz, a Quaternary clay basin in northwestern Ontario; Geographie physique et Quaternaire, v. 46, p. 295-309.

Singer, S., Cheng, C.K., and Scafe, M.G.
1997: The hydrogeology of southern Ontario, volume 1; Ontario Ministry of Environment and Energy.

Watt, A.K.
1952: Ground water in Ontario, 1947; Ontario Department of Mines, 116 p.
1961: Ground water Ontario, 1953 and 1954; The Ontario Water Resources Commission, 302 p.

White, O.L.
1975: Quaternary geology of the Bolton area, Southern Ontario; Ontario Division of Mines, Ministry of Natural Resources, Geological Report 117, 119 p.