Data Article

Hydro-stratigraphic datasets for the reconstruction of a large scale 3D FEM numerical model in the Milan metropolitan area (northern Italy)

Alberto Previati*, Mattia De Caro, Giovanni B. Crosta

DISAT – Department of Earth and Environmental Sciences, University of Milano Bicocca, Piazza della Scienza, 4, Milan 20126, Italy

A R T I C L E   I N F O

Article history:
Received 20 October 2020
Revised 10 November 2020
Accepted 12 November 2020
Available online 19 November 2020

Keywords:
Urban scale groundwater modeling
Stratigraphic database
Aquifer heterogeneities
3D hydrostratigraphy

A B S T R A C T

One of the objectives of groundwater numerical modeling is to accurately reproduce the flow velocity field and the flow and transport pathways. In this article the hydro-stratigraphic dataset, used in the co-submitted article “Modeling the interference of underground structures with groundwater flow and remedial solutions in Milan” (De Caro et al., 2020) [1], is presented. The work aims to reconstruct the spatial variability of the hydraulic parameters in the shallow aquifers of the Milan City area (northern Italy) and to integrate them in a groundwater flow 3D finite element method (FEM) numerical model. This objective is achieved by converting qualitative borehole logs stratigraphic information into hydrogeological parameters (e.g. hydraulic conductivity and porosity) and by interpolating these parameters over the finite element mesh nodes by means of 3D kriging techniques. The modeling domain and the mesh nodes, the boundary surfaces between the aquifers as well as some of the piezometric data used to calibrate the model are presented to make the numerical experiment reproducible.

DOI of original article: 10.1016/j.enggeo.2020.105652

* Corresponding author.
E-mail address: a.prevati1@campus.unimib.it (A. Previati).

https://doi.org/10.1016/j.dib.2020.106541
2352-3409/© 2020 Published by Elsevier Inc. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Specifications Table

| Subject | Environmental Science |
|---------|-----------------------|
| Specific subject area | Reconstruction of heterogeneous hydrogeological properties of shallow aquifers from borehole logs to support regional scale groundwater flow modeling |
| Type of data | Table |
| How data were acquired | Borehole logs lithological descriptions completed during the drilling and acquired from the Lombardy region geological database [2]. Piezometric map [3] elaborated from 2016 piezometric measurements, available at [4]. High resolution hydraulic head measurements recorded by means of pressure transducers at three selected boreholes from 2016 to 2019 [5]. |
| Data format | Raw/Analyzed |
| Parameters for data collection | Only the borehole logs crossing the phreatic and the semi-confined aquifers and with a complete and clear lithological description were considered. Daily hydraulic head values are representative for the shallow unconfined aquifer with a resolution of ±1 cmH2O |
| Description of data collection | The stratigraphic and piezometric data were collected from the regional geological database and organized in lithofacies classes. Hydraulic head values were recorded at three specific locations (piezometers) and corrected by the atmospheric pressure fluctuations. |
| Data source location | Milan metropolitan area (8°55'E – 9°25'E, 45°19'N – 45°37'N) Lombardy Po plain Northern Italy |
| Data accessibility | Direct URL to data: Dataset [3] http://dx.doi.org/10.17632/nc5v37t425.1 Regional geological borehole logs database (free-access link) [2] http://www.cartografia.regione.lombardia.it/viewer25/index.jsp?config=config-caspita.xml Dataset [5] http://dx.doi.org/10.17632/kg3yy9cz7y.1 |
| Related research article | De Caro et al., Modeling the interference of underground structures with groundwater flow and remedial solutions in Milan, Engineering Geology, vol. 272, 2020 https://doi.org/10.1016/j.enggeo.2020.105652 [1] |

Value of the Data

- The data are useful to understand how the groundwater flow model in the related research article [1] was developed and how the hydraulic parameters were assigned to the 3D FEM mesh
- The data and the proposed methodology can be useful to integrate the heterogeneities of regional aquifer complexes when flow pathways play a significant role in the objectives of the model
- The presented datasets can be used to reproduce the methodology presented in the research article [1] by any interested user

1. Data Description

In the related research article [1] a large-scale 3D FEM groundwater-flow numerical model was developed by means of the commercial code FeFlow 7.2 [6] to assess the impact of
underground subsurface infrastructures (i.e. tunnels) on the hydraulic head distribution in the Milan City area. In the online repository [3] the boundary of the model, the hydraulic head isolines as from 2016, the boundary surfaces between the aquifer complexes, and some of the hydraulic head measurements used for model calibration are stored. The piezometric map [3], elaborated from 2016 regional-scale piezometric measurements (available online at [4]), was essential to define the initial conditions and to calibrate the flow model.

Fig. 1.a shows a map of the location and the length of the available borehole logs. The stratigraphic database [2] collects the borehole logs available for the Lombardy-Po Plain area. The database contains information regarding the position, the elevation, the depth, and the lithological description of the layers crossed by each available borehole or well.

Each stratigraphic layer description was codified with a hierarchical approach as from the codification proposed by De Caro et al. (2020) [7] to obtain 29 lithofacies according to the relative abundance of grain size classes. In multiple term units, the first term represents the most abundant grain size class, the second and the third account for 25±50% and 5±25% of the total weight, respectively (Table 1).

Following the work by De Caro et al. (2020) [7], 133 grain-size distribution data, covering all the sediment-type spectra and depositional environments within the study area, were analyzed with different empirical equations to estimate the hydraulic conductivity and the porosity of each specific sediment type. Table 1 shows the hydrogeological parameters derived for each lithofacies class and Fig. 2 the relative frequencies in the modeling domain.

In the research article the interference of underground infrastructures with groundwater flow is examined, to this aim the assessment of natural fluid flow pathways (i.e. considering only the effect of hydrogeological parameters heterogeneity) is relevant to a deterministic comparison.

In this context, the regional geological borehole logs database and the parametrization presented in Table 1 were essential to spatialized into the model by means of 3D geostatistical techniques the variability of hydrogeological parameters.

Subsequently, the hydrogeological parameters were calibrated (details are provided in the next section), and the calculated hydraulic head distribution was validated by comparing the computed hydraulic head with the available piezometric maps and with hydraulic head fluctuations time series at specific locations. In Fig. 1.b daily averaged hydraulic head values recorded at three monitoring stations (see Fig. 1.a for their location) are presented, the raw dataset is stored in the online repository as a CSV table [5].

The continuous groundwater monitoring network in the Milan city area was installed in February 2016 by the Department of Earth and Environmental Sciences of the University of Milano-Bicocca. Open pipe piezometers were selected according to their location and depth, the degree of structural protection, and the integrity of the borehole structure. The submersed transducers are installed at a constant depth in the piezometers by means of a stainless steel cable of known length \(L_{\text{cable}}\) and autonomously record the pressure (±0.1 cmH2O) and the temperature (±0.01 °C) every 15 min. Pressure data recorded by each sensor are corrected by the atmospheric pressure and converted in meters above sea level (masl) knowing the ground surface elevation at each measuring location. The piezometric head (masl) is obtained through the following equation:

\[
H(t) = Z_{\text{sens}} + P_{\text{coal}}(t) - P_{\text{air}}(t)
\]

Where

- \(H(t)\) = piezometric head (masl) at time \(t\),
- \(Z_{\text{sens}}\) = elevation (masl) of the sensor obtained by subtracting \(L_{\text{cable}}\) to the ground surface elevation,
- \(P_{\text{coal}}(t)\) = pressure measured by the sensor at time \(t\),
- \(P_{\text{air}}(t)\) = atmospheric pressure at time \(t\).

In the dataset [5] we provided the average daily piezometric head at three locations in the Milan city area.
Fig. 1. (a) Map view of the study area showing the spatial distribution of the available borehole stratigraphic logs. The dots are colored by the depth of the borehole bottom. The extent of the modeling area is marked with the black line. Regional hydraulic head isolines as on 2016 and boundaries between different depositional fan systems as from Fontana et al., (2014) [8] are also represented with violet and blue lines, respectively. (b) High resolution hydraulic head time-series, values in meters above sea level (masl), recorded at locations 1, 2, 3 reported in Fig 1a (the raw dataset is stored in the online repository [5]).
Table 1
Hydraulic conductivity and porosity values assigned to each one of the 29 identified lithofacies. The following codification was adopted: G for gravel, S for sand, M for silt, C for clay, R for partially cemented deposits, Ar for arenaceous units, Cong for conglomeratic units, P for peat, and Landfill for shallow incoherent artificial deposits. In multiple term units the first term represents the most abundant grain size class, the second and the third account for 25-50% and 5-25% of the total weight, respectively.

| Facies Type | % | Hydrofacies | % | Lithofacies | Description | % | Hydraulic Conductivity (m/s) | Porosity |
|-------------|---|-------------|---|-------------|-------------|---|-----------------------------|----------|
| Gravel-Sand | 31.7 | G | 4.9 | G | gravel | 4.9 | 9.05E-02 | 0.26 |
| Sand-Gravel | 30.5 | GC | 1.4 | GC | gravel and clay | 1.4 | 6.68E-04 | 0.38 |
| Partially Cemented | 4.6 | R | 4.6 | Ar | arenaceous units | 4.6 | 3.63E-04 | 0.28 |
| Aquitard | 14.1 | M | 3.2 | M | silt | 3.2 | 9.00E-08 | 0.50 |
| Aquiclude | 19.1 | C | 19.1 | C | clay | 19.1 | 5.00E-06 | 0.50 |

Note: The codification used includes G for gravel, S for sand, M for silt, C for clay, R for partially cemented deposits, Ar for arenaceous units, Cong for conglomeratic units, P for peat, and Landfill for shallow incoherent artificial deposits.
2. Experimental Design, Materials and Methods

In this section the approach adopted in the related research article “Modeling the interference of underground structures with groundwater flow and remedial solutions in Milan” (De Caro et al., 2020) [1] is presented. This methodology aims to integrate the spatial variability of the hydrogeological parameters of the shallow aquifers into a 3D FEM numerical model as from borehole logs lithological descriptions. To this aim, 3677 borehole logs were collected and stored in a georeferenced database (Fig. 1).

The hydraulic parameters obtained from the analysis of grain-size distribution data (Table 1) were spatialized in the 3D FEM numerical modeling domain by interpolating the parameters over the nodes of the FEM mesh. The 3D FEM mesh (25.6 x 31.9 x 0.2 km) has a layered structure with 7181,880 triangular prismatic elements divided in 30 layers (239,396 elements per each layer). The element size along the XY plane ranges from 4 m (near the pumping wells or the axes of the underground tunnels) to at most 500 m and is about 4 m along the Z direction (i.e., the average distance between the FEM layers). The element size was chosen to account for strong hydraulic head changes near the boundary conditions or impervious elements and aims to satisfy the dimensions of the main stratigraphic structures at the scale of interest, both in the horizontal and in the vertical direction (i.e., the element size is always smaller than the expected size of the main stratigraphic bodies which are typically widespread in the horizontal direction and arranged in layers [7]).

The spatial interpolation of hydraulic parameters was carried out by means of the 3D ordinary kriging technique. The experimental variograms and the kriging model parameters used for the interpolation are presented in Fig. 3.

To assign the material properties to the FEM model, the values obtained from the interpolation at the mesh nodes were moved to the mesh elements by averaging the values of the surrounding nodes. Then, the modeling domain was divided into smaller sub-domains based on specific k-values intervals, on the distinction between the phreatic and the semi-confined aquifer.
A. Previati, M. De Caro and G.B. Crosta / Data in Brief 33 (2020) 106541

| Parameter          | Model  | Nugget  | Sill   | Range  | Anisotropy |
|--------------------|--------|---------|--------|--------|------------|
| Hydraulic Conductivity | Exponential | 0.05    | 1.78   | 130    | 0.05       |
| Porosity           | Exponential | 0.000024 | 0.0036 | 130    | 0.05       |

**Fig. 3.** 3D Kriging: experimental variograms and kriging model parameters used for the interpolation of hydraulic conductivity (a) and porosity (b). The ratio between the longest (X Y plane - Horizontal) and the shortest (Z direction - Vertical) axes of the semi-variogram ellipsoid is expressed with the anisotropy coefficient.

**Fig. 4.** (a) Spatial distribution of the hydraulic conductivity values in the FEM modeling domain. Cross sections showing (b) the raw hydraulic conductivity distribution obtained from the interpolation of k-values as from the hydro-stratigraphic dataset and (c) homogeneous zones obtained by reclassifying data from Fig. 4b into discrete classes of hydraulic conductivity. This reclassified distribution forms the input to the inverse calibration of the model with PEST.
(see the limit surface stored at [3]) and on different depositional fan systems as from the work by Fontana et al. (2014) [8]. Those sub-domains were used as homogeneous zones for the inverse calibration of the model with PEST V15 [9] taking the piezometric head, as on 2016 [3], as the steady-state reference level and the hydraulic head time series at specific locations, from 2016 to 2019 [5], as reference for the calibration in transient state.

Fig. 4 shows the spatial distribution of the hydraulic conductivity values in the modeling domain. Table 2 (the dataset is stored in the online repository [3]) shows the location of the mesh nodes and the associated parameters (hydraulic conductivity and porosity) as well as the grain size class, the aquifer system and the depositional system.

Declaration of Competing Interest

CASPITA [2] and PTUA [4] are free access datasets provided by Regione Lombardia. The authors declare that they have no known competing financial interests or personal relationships which have, or could be perceived to have, influenced the work reported in this article.

Acknowledgments

The authors are grateful to Gruppo-CAP and MM S.p.a for providing essential data in the studied area. This article is an outcome of MIUR Project URGENT - Urban Geology and Geohazards: Engineering geology for safer, resilieNt and smart ciTies - 2017HPJLPW, Project MIUR – Dipartimenti di Eccellenza 2018–2022 and PerFORM WATER 2030 – Regione Lombardia project.

References

[1] M. De Caro, G.B. Crosta, A. Previati, Modeling the interference of underground structures with groundwater flow and remedial solutions in Milan, Eng. Geol. 272 (2020) https://doi.org/10.1016/j.enggeo.2020.105652.
[2] Regione Lombardia, Banca dati geologica sottosuolo, Campioni Analisi Sondaggi Penetrometrie e Indagini Territoriali Ambientali, (CASPITA) (regional geological borehole logs database) http://www.cartografiaregione.lombardia.it/viewer25/index.jsp?config=config-caspita.xml, last access 10/01/2020.
[3] A. Previati, M. De Caro, Boundary surfaces and Hydraulic parameters of a 3D large scale FEM numerical model in the Milan City area (northern Italy), Mendeley Data v1 (2020) http://dx.doi.org/10.17632/mc5v37t425.1.
[4] Regione Lombardia, Programma di tutela e Uso delle Acque (PTUA) (regional database on groundwater levels) https://www.regione.lombardia.it/wps/portal/istituzionale/HP/DettaglioRedazionale/servizi-e-informazioni/Enti-e-Operatori/territorio/governodelleacque/piano-tutela-acque-pta-2016/piano-tutela-acque-pta-2016, last access 10/01/2020.
[5] A. Previati, M. De Caro, Hydraulic head fluctuations from 2016 to 2019 at three locations in the Milan city area (northern Italy), Mendeley Data v1 (2020) http://dx.doi.org/10.17632/kg3yy9cz7y.1.
[6] H.J. Diersch, in: FEFLOW: Finite Element Modeling of Flow, Mass and Heat Transport in Porous and Fractured Media, Springer-Verlag, Berlin, 2014, p. 996.
[7] M. De Caro, R. Perico, G.B. Crosta, P. Frattini, G. Volpi, Regional scale conceptual and numerical groundwater flow model in fluvi-glacial sediments for the Milan metropolitan area (northern Italy), J. Hydrol.: Reg. Stud. (2020).
[8] A. Fontana, P. Mozzì, M. Marchetti, Alluvial fans and megafans along the southern side of the Alps, Sediment. Geol. (2014) https://doi.org/10.1016/j.sedgeo.2013.09.003.
[9] J.E. Doherty, PEST: Model-Independent Parameter Estimation User Manual, Watermark Numerical Computing, Brisbane, Australia, 2016.

Table 2

Example of the structure of the file containing information about the nodes of the FEM mesh and the corresponding hydrogeological parameters and zones used for the calibration (the dataset is stored in the online repository [3]). See the research article for the alphabetical codification used for “Aquifer”. “K_LOG” is the base-10 logarithm of the hydraulic conductivity values expressed in m/s.

| X         | Y         | Z         | K_LOG   | Poro | Aquifer | Megafan | Grain Size |
|-----------|-----------|-----------|---------|------|---------|---------|------------|
| 1,521,490.55 | 5,032,648.47 | 105.09 | −3.22 | 0.34 | A       | Lambro  | Sand       |
| 1,521,467.78 | 5,032,653.75 | 105.1  | −3.21 | 0.34 | A       | Lambro  | Sand       |
| 1,521,448.64 | 5,032,633.6  | 105.03 | −3.24 | 0.34 | A       | Lambro  | Sand       |
| 1,521,428.63 | 5,032,625.16 | 104.95 | −3.25 | 0.34 | A       | Lambro  | Sand       |
| ...        |           |          |         |      |         |         |            |