

Notes on groundwater data collection and pre-processing

Groundwater level (GWL) data was searched on the internet on governmental websites between 2019 and 2022. The national or state governmental agencies from which data were compiled are listed in Table A1 in the publication. Below, we provide notes on groundwater data collection and pre-processing for each country.

5 Australia

We downloaded data from the Australian Groundwater Explorer embedded on the website of the Australian Government – Bureau of Meteorology, on 11/4/2021 from the website: <http://www.bom.gov.au/water/groundwater/explorer>. The downloaded data version is "data VERSION 1.7.0." It contains GWL measurements for each state and secondary data from the National Groundwater Information System (NGIS). The secondary data were extracted from the geodatabase into csv files.

10 The GWLs were then linked to the NGIS well data using the "BoreID" to obtain the well location. The reference point for the GWL varies across the datasets. Ground surface, top of casing, and sea level are all included. GWLs were selected with a preference for the ground surface if there were duplicates and remaining GWLs referenced to sea level were converted to ground surface reference using SRTM elevation at the well site. Well coordinates, given in the Geocentric Datum of Australia 1994 (GDA94), were converted to WGS84.

15 Belgium

We downloaded data from Flanders in Belgium on 9/23/2020 from Databank Ondergrond Vlaanderen (DOV; <https://www.dov.vlaanderen.be>) using the package "pydov" for Python. The reference point for the GWLs in the original data set is sea level, but the ground surface elevation is given and enabled us to convert the GWLs to ground surface reference. The well coordinates, given in the Belgian Lambert 72 system, were converted to WGS84.

20 Brazil

We downloaded data from Brazil through the online data portal "RIMAS" (http://rimasweb.cprm.gov.br/layout/pesquisa_complexa.php?rimas=true), operated by Servicio Geologico de Brazil, on 9/15/2022 using Python.

Canada

25 GWL data were obtained from three sources: the Nova Scotia and British Columbia government agencies and the Groundwater Information Network (GIN; <http://gw-info.net>). First, we obtained GWL data from several Canadian provinces via a download file from GIN on 3/30/2020. In the dataset, we were unable to identify the reference point for the GWL observations from British Columbia, while for Nova Scotia we noted reversed months and days in the date. Therefore, we downloaded GWL data referenced to the ground surface from the British Columbia website

30 (http://www.env.gov.bc.ca/wsd/data_searches/obswell/map/data/ObservationWellDataAll_DailyMean.csv) on 3/1/2021, but used well location information from the GIN dataset, via the original identifiers. On 8/4/2021, we downloaded both GWL data and well information from the Nova Scotia website (<https://novascotia.ca/nse/groundwater/groundwaternetwerk.asp>). In the

dataset, GWLs are referenced to sea level but can mostly be converted to top of casing. GWL data from a few wells in Quebec remained from GIN. Here, GWLs are referenced to sea level. The remaining GWLs with sea level reference were converted to ground surface reference using SRTM elevation at the well site.

Denmark

We extracted GWL and well data from the Jupiter database of the Geological Survey of Denmark and Greenland (GEUS) that was downloaded on 5/16/2019 (<https://www.geus.dk/produkter-ydelser-og-faciliteter/data-og-kort/national-boringsdatabase-jupiter/>). We used the tables “watlevel”, “borehole” and “intake”, where GWLs provided with the “watlevel” table are connected to “borehole” locations via the “intake” table. The reference point for the GWLs is different throughout the datasets. Ground surface, top of casing, and sea level are all included, with the GWLs selected in this order when there were duplicates and remaining GWLs referenced to sea level were converted to ground surface reference using SRTM elevation at the well site.

France

GWL and secondary data were downloaded from the ADES database (<https://ades.eaufrance.fr/Recherche>) on 10/30/2020, for which professional access was granted for research purposes. GWLs are given with reference to ground surface, top of casing, and sea level. Where duplicates exist, the GWLs were selected in that order. Where well elevation is available, we converted GWLs from sea level to ground surface reference. Only one well referencing sea level remained that was converted to ground surface reference using SRTM elevation at the well site.

Germany

We received GWL data and secondary information from four government agencies of German coastal regions (sources 10 to 13 in Table A1 in the publication) between 11/24/2020 and 4/19/2021. GWLs are referenced to sea level everywhere, but a conversion to the ground surface reference with provided elevation at the well site is possible. Well coordinates, given in projected coordinate systems for Germany, were converted to WGS84.

Ireland

We downloaded GWL data from the Environmental Protection Agency (EPA) of Ireland (<https://epawebapp.epa.ie/hydronet/#Groundwater>) on 2/26/2020 using Python. Well locations were provided by the EPA via email. GWLs are referenced to sea level and we had no elevation at the well site available for conversion. GWLs were therefore converted to ground surface reference using SRTM elevation at the well site. Well coordinates, given in a projected coordinate system for Ireland, were converted to WGS84.

South Africa

We received GWL time series from the national Hydstra database via email request to HydstraS@dwa.gov.za between 3/27/2021 and 4/1/2021. We received feedback that the reference point for GWL observations in South Africa is the top of the casing.

Sweden

We downloaded GWL time series from the Geological Survey of Sweden using Python following the instructions given (<https://www.sgu.se/produkter/geologiska-data/oppna-data/grundvatten-oppna-data/grundvattennivaer-tidsserier/>) on 10/22/2021. GWLs are either reported for the ground surface or top of casing reference. Well coordinates, given in a projected coordinate system for Sweden, were converted to WGS84.

70 The Netherlands

We downloaded GWL time series and well locations from the DINOLOket database (<https://www.dinoloket.nl>) on 8/25/2020. GWLs are reported with reference to both the ground surface and sea level. We selected the GWL data referenced to the ground surface that was available for all wells. Well coordinates are reported in a projected coordinate system for the Netherlands. They were converted to WGS84.

75 United States

We downloaded GWL time series and well data from the United States Geological Survey (USGS) via two different websites (<https://waterdata.usgs.gov/nwis/gw> and <https://waterservices.usgs.gov/rest/GW-Levels-Test-Tool.html>) on 5/26/2020 and from California Department of Water Resources (<https://wdl.water.ca.gov>) on 6/16/2022. GWLs are available with either reference to ground surface or sea level from the USGS, where we preferred ground surface if available and otherwise converted GWLs referenced to sea level to ground surface reference using SRTM elevation at the well site. Well coordinates, reported in NAD83 geodetic coordinate system for North America in both datasets, were converted to WGS84.