GROW: Dataset Description

GROW (global integrated GROundWater package) is a global, analysis-ready, quality-controlled dataset that combines groundwater depth and level time series from around the world with associated Earth system variables. The dataset contains > 180,000 time series from 41 countries in a daily, monthly, or yearly temporal resolution, accompanied by 35 time series or attributes of meteorological, hydrological, geophysical. vegetation, and anthropogenic variables (e.g., precipitation, drainage density, aquifer type, NDVI, land use). 33 data flags regarding well features (e.g., location coordinates and country), as well as time series characteristics (e.g., gap fraction or length), facilitate quick data filtering.

The dataset is organized in two files: A table containing time series (grow_time_series.csv/.parquet) and one table with static attributes (grow_attributes.csv/.parquet). An example of use that demonstrates how the data is subset and prepared is given on GitHub (https://github.com/EarthSystemModelling/GROW).

The columns of the attributes and time series table are described in Table 1 & 2.

Citation

When using the data, please cite:

Bäthge, A., Ruz Vargas, C, Lischeid, G., Collenteur, R., Cuthbert, M., Fleckenstein, J., Flörke, M., De Graaf, I., Gnann, S., Hartmann, A., Huggins, X., Moosdorf, N., Wada, Y., Wagener, T., Reinecke, R. GROW: A Global Time Series Dataset for Groundwater Studies within the Earth System. Preprint (2025)

Table 1: Overview of columns in the attributes table.

Column name	Unit	Description	Data source	Licensed under
GROW_ID		A unique identifier that can be used to assign the time series to the attributes		
original_ID_groundwater		(Adopted from the original datasets) The groundwater well identifier that is used in the original groundwater datasets	IGRAC 2024a & IGRAC 2024b	
name		(Adopted from the original datasets) Name of the well	IGRAC 2024a & IGRAC 2024b	
feature_type		(Adopted from the original datasets) Well type (e.g., "Water well")	IGRAC 2024a & IGRAC 2024b	
purpose		(Adopted from the original datasets) Purpose of the well (e.g., "Observation / monitoring")	IGRAC 2024a & IGRAC 2024b	
status		(Adopted from the original datasets) Status if well is still running ("Active") or not ("Abandoned" or "Collapsed")	IGRAC 2024a & IGRAC 2024b	
description		(Adopted from the original datasets) Further information about the well; usually a web link	IGRAC 2024a & IGRAC 2024b	
latitude		(Adopted from the original datasets) Latitude coordinates in WGS 84	IGRAC 2024a & IGRAC 2024b	
longitude		(Adopted from the original datasets) Longitude coordinates in WGS 84	IGRAC 2024a & IGRAC 2024b	

surface_elevation_m_asl	m	(Adopted from the original datasets) Surface elevation at the location of the well in meter above mean sea level; given for 19 % of the wells	IGRAC 2024a & IGRAC 2024b	
top_of_well_elevation_m_asl	m	(Adopted from the original datasets) Elevation of the top of the well in meter above mean sea level; given for 2 % of the wells	IGRAC 2024a & IGRAC 2024b	
country		(Adopted from the original datasets) country in which the well is located	IGRAC 2024a & IGRAC 2024b	
address		(Adopted from the original datasets) Address of the well location	IGRAC 2024a & IGRAC 2024b	
license		(Adopted from the original datasets) Data license for that particular groundwater time series + attributes	IGRAC 2024a & IGRAC 2024b	
aquifer_name		(Adopted from the original datasets) Name of the aquifer	IGRAC 2024a & IGRAC 2024b	
confinement		(Adopted from the original datasets) Information if well is confined or unconfined	IGRAC 2024a & IGRAC 2024b	
organisation		(Adopted from the original datasets) Organisation that provided the groundwater data	IGRAC 2024a & IGRAC 2024b	
manager		(Adopted from the original datasets) Name of the well manager	IGRAC 2024a & IGRAC 2024b	
drilling_total_depth_m	m	(Adopted from the original datasets) Total drilling depth for well construction	IGRAC 2024a & IGRAC 2024b	
interval		Temporal resolution of the time series; either daily("d"), monthly ("MS") or yearly ("YS")		
starting_date		First date of the time series in "YYYY-mm-dd" format		

ending_date		Last date of the time series in "YYYY-mm-dd" format		
length_years		Number of years that occur in a time series		
autocorrelation		Indicates if the time series is autocorrelated (threshold: Spearman correlation >= 0.6); either "True" or "False"		
aggregated_from_n_values_median		Median number of records that were aggregated to the daily, monthly or yearly means; "NA" indicates that the time series was not aggregated		
gap_fraction		Fraction of gaps in the time series		
jumps		Indicates if 1-2 >=50 m groundwater table changes between adjacent time steps (jumps) occur (True) or not (False)		
plateaus		Indicates if sequences with the exact same value for: • >7 days (daily time series), • >3 months (monthly time series), • >1 year (yearly time series that are 1-4 years long), • >2 years (yearly time series that are 5-9 years long), • >3 years (yearly time series that are 10-14 years long), • >4 years (yearly time series that are >15 years long) or • >5 years (yearly time series that are >20 years long) occur (True) or not (False)		
trend_direction		Trend direction that is derived with the Mann Kendall test; either "no trend", "increasing" or "decreasing"		

		For autocorrelated time series, the Hamed and Rao Modified Mann Kendall test was used		
trend_slope_m_year-1	m/year	In case the trend was significant (p-value < 0.05), the slope is given		
reference_point		(Adopted from the original datasets) Reference point for the groundwater table measurements; either 'Water depth [from the top of the well]', 'Water level elevation a.m.s.l.' or 'Water depth [from the ground surface]'	IGRAC 2024a & IGRAC 2024b	
groundwater_mean_m	m	Mean groundwater table per time series; either mean depth to water table or level elevation (see reference_point)		
groundwater_median_m	m	Median groundwater table per time series; either median depth to water table or level elevation (see reference_point)		
koeppen_geiger_class		Original Koeppen-Geiger classification from Koeppen & Geiger (1939) Af: Equatorial fully humid Am: Equatorial monsoonal As: Equatorial summer dry Aw: Equatorial winter dry BWk: Cold desert BWh: Hot desert BSk: Cold steppe BSh: Hot steppe Cfa: Warm temperate fully humid hot summer Cfb: Warm temperate fully humid warm summer Cfc: Warm temperate fully humid cool summer Csa: Warm temperate summer dry hot summer LCsb: Warm temperate summer dry warm summer Csc: Warm temperate summer dry cool summer Cwa: Warm temperate winter dry hot summer Cwb: Warm temperate winter dry warm summer	CHELSA v2.1 - kg0 by Karger et al. (2021)	

		Cwc: Warm temperate winter dry cool summer Dfa: Snow fully humid hot summer Dfb: Snow fully humid warm summer Dfc: Snow fully humid cool summer Dfd: Snow fully humid extremely continental Dsa: Snow summer dry hot summer Dsb: Snow summer dry warm summer Dsc: Snow summer dry cool summer Dsd: Snow summer dry extremely continental Dwa: Snow winter dry hot summer Dwb: Snow winter dry warm summer Dwc: Snow winter dry cool summer Dwd: Snow winter dry extremely continental ET: polar tundra EF: polar frost		
hydrobelt_class		Hydrobelt classification after Meybeck et al. 2013. The classification was reduced to 5 classes (based on Table 1. in Meybeck et al. 2013) which are characterized by average annual temperature and runoff	Meybeck et al. 2013	CC-BY-3.0
ground_elevation_m_asl	m	Ground elevation in meter above mean sea level Data source: MERIT DEM by Yamazaki et al. (2017)	MERIT DEM by Yamazaki et al. (2017)	CC-BY-NC 4.0
topographic_slope_degrees	°	Topographic slope in degrees	Geomorpho90m by Amatulli et al. (2020)	CC-BY-4.0
rock_type_class		Rock type class	GLiM by Hartmann & Moosdorf (2012)	CC-BY-3.0

aquifer_type_class		Estimated aquifer type derived from rock_type_class and World Karst Aquifer Map; more information about method in Bäthge et al.; either "porous", "fractured", "karst" or "water_body"	WHYMAP WOKAM by Chen et al. (2017) + GLiM by Hartmann & Moosdorf (2012)	GLiM: CC-BY-3.0
rock_permeability_m2	m^2	Rock permeability	GLHYMPS by Gleeson et al. (2014)	CC-BY-4.0
rock_porosity_fraction		Rock porosity	GLHYMPS by Gleeson et al. (2014)	CC-BY-4.0
soil_texture_0-30_cm_class		Soil texture class for the topsoil (0-30 cm); either "Organic", "Very Fine", "Fine", "Medium Fine", "Medium" and "Coarse"	HiHydroSoil by Simons et al. (2020)	
soil_texture_30-200_cm_class		Soil texture class for the subsoil (30-200 cm); either "Organic", "Very Fine", "Fine", "Medium Fine", "Medium" and "Coarse"	HiHydroSoil by Simons et al. (2020)	
soil_saturated_conductivity_0-30_cm_cm_d-1	cm/day	Saturated Hydraulic Conductivity of topsoil (0-30 cm)	HiHydroSoil by Simons et al. (2020)	
soil_saturated_conductivity_30-200_cm_cm_d-1	cm/day	Saturated Hydraulic Conductivity of subsoil (30 - 200 cm)	HiHydroSoil by Simons et al. (2020)	

distance_perennial_streams_m	m	The distance between perennial streams with 0.1 cubic metres per second flow threshold	Cuthbert, Gleeson, et al. (2019)	CC-BY-4.0
drainage_density_m-1	1/m	Drainage density, calculated as sum of the river lengths per basin divided by the area of that basin	HydroRivers by Lehner & Grill (2013) + BasinATLAS by Linke et al. (2019)	Both: CC-BY-4.0
groundwater_dependent_ ecosystems_class		Groundwater-dependent ecosystems class, either 'Lotic', 'Terrestrial, lentic, and lotic (all)', 'Terrestrial and lentic', 'Lentic and lotic', 'Terrestrial', 'Terrestrial and lotic', 'Lentic' or 'No GDE'	Huggins et al. (2023)	CC BY-NC-SA 4.0
groundwaterscapes_ID_class		Groundwaterscapes - "Landscape units with specific and broadly occurring configurations of groundwater-connected system functions" (Huggings et al., 2024, p.2) • 1: Arid and desert regions with minimal functions - Large storage capacity • 2: Arid and desert regions with minimal functions - Small storage capacity and moderately effective national governance • 3: Arid and desert regions with minimal functions - Small storage capacity and ineffective national governance • 4: Underserved populations and ineffective national governance - Some terrestrial GDEs amid generally limited functions • 5: Underserved populations and ineffective national governance - Large storage capacity, moderate climate coupling, and some terrestrial GDEs • 6: Earth system functions in non-agricultural regions - Some terrestrial GDEs	Huggings et al. (2024)	CC-BY-4.0

		• 7: Earth system functions in non-agricultural regions - Effective national governance and some terrestrial GDEs • 8: Moderate GDEs and/or Earth system functions in nations with effective governance - Large storage capacity and moderate climate coupling • 9: Moderate GDEs and/or Earth system functions in nations with effective governance - Small storage capacity and very effective national governance • 10: Agricultural regions with high groundwater dependence - Smallholder farming and moderately effective national governance • 11: Agricultural regions with high groundwater dependence - Large farms with effective national governance • 12: Agricultural regions with lower dependence on groundwater - Moderate climate interactions, few aquatic GDEs, large farms • 13: Agricultural regions with lower dependence on groundwater - Large farms situated among GDEs • 14: Extensive GDEs in non-agricultural regions - Small storage capacity, underserved populations, and ineffective national governance • 15: Extensive GDEs in non-agricultural regions - Large storage capacity, large range in underserved populations, and ineffective national governance		
main_landuse		Main land use type per well; main land use is that land use whose mean fraction over time is the highest.	ISIMIP3a input data, Volkholz & Ostberg (2024)	CC0 1.0

Table 2: Overview of columns in the time series table.

Column name	Unit	Description	Data source	Licensed under
GROW_ID		A unique identifier that can be used to assign the time series to the attributes		
country		(Adopted from the original datasets) country in which the well is located	IGRAC 2024a & IGRAC 2024b	
interval		Temporal resolution of the time series; either daily("d"), monthly ("MS") or yearly ("YS")	IGRAC 2024a & IGRAC 2024b	
date		Timestamp of the record in "YYYY-mm-dd" format		
year		Year of the record in "YYYY" format		
month		Month of the record in "YYYY-mm" format		
aggregated_from_n_values		Number of values that were aggregated to this record; "NA" indicates that the time series was not aggregated		
plateaus		Marks sequences with the exact same value for: • >7 days (daily time series), • >3 months (monthly time series), • >1 year (yearly time series that are 1-4 years long), • >2 years (yearly time series that are 5-9 years long), • >3 years (yearly time series that are 10-14 years long), • >4 years (yearly time series that are >15 years long) or • >5 years (yearly time series that are >20 years long)		

		With the length of the sequence		
groundwater_depth_from_ ground_m	m	(Adopted from the original datasets) groundwater depth from ground surface	IGRAC 2024a & IGRAC 2024b	
groundwater_depth_from_ well_top_m	m	(Adopted from the original datasets) groundwater depth from the top of the well	IGRAC 2024a & IGRAC 2024b	
groundwater_level_m_asl	m	(Adopted from the original datasets) groundwater table level elevation	IGRAC 2024a & IGRAC 2024b	
groundwater_filled_depth_from _ground_m	m	(Adopted from the original datasets) groundwater depth from ground surface; gaps are linearly filled	IGRAC 2024a & IGRAC 2024b	
groundwater_filled_depth_from_ well_top_m	m	(Adopted from the original datasets) groundwater depth from the top of the well; gaps are linearly filled	IGRAC 2024a & IGRAC 2024b	
groundwater_filled_level _m_asl	m	(Adopted from the original datasets) groundwater table level elevation; gaps are linearly filled	IGRAC 2024a & IGRAC 2024b	
precipitation_gpcc_mm_year-1	mm/year	Precipitation sum in mm/year based on monthly data; for daily time series, each day within a month is matched with the corresponding value for that month	GPCC by Schneider et al. (2022)	
precipitation_mswep_mm_year-1	mm/year	Precipitation sum in mm/year based on 3-hourly data	MSWEP V2 by Beck et al. (2019)	CC-BY-NC 4.0
potential_evapotranspiration_era5 _mm_year-1	mm/year	Potential evapotranspiration sum in mm/year based on daily data	ERA5-Land by Muñoz-Sabater et al. (2021); Copernicus Climate Change Service (2024)	

potential_evapotranspiration_gleam_mm_year-1	mm/year	Potential evapotranspiration sum in mm/year based on daily data	GLEAM4 by Miralles et al. (2025)	
actual_evapotranspiration_ mm_year-1	mm/year	Actual evapotranspiration sum in mm/year based on daily data	GLEAM4 by Miralles et al. (2025)	
air_temperature_°C	°C	Mean air temperature based on daily data	ERA5-Land by Muñoz-Sabater et al. (2021); Copernicus Climate Change Service (2024)	
snow_depth_m	m	Mean snow depth based on daily data	ERA5-Land by Muñoz-Sabater et al. (2021); Copernicus Climate Change Service (2024)	
interception_mm_year-1	mm/year	Sum of rain interception in mm/year based on daily data	GLEAM4 by Miralles et al. (2025)	
ndvi_ratio		Mean NDVI (Normalized Difference Vegetation Index) based on daily data	AVHRR NDVI by Vermote & NOAA CDR Program (2018)	
lai_low_vegetation_ratio		Mean Leaf Area Index of low vegetation based on daily data	ERA5-Land by Muñoz-Sabater et al. (2021); Copernicus Climate Change Service (2024)	
lai_high_vegetation_ratio		Mean Leaf Area Index of high vegetation based on daily data	ERA5-Land by Muñoz-Sabater et al. (2021); Copernicus Climate Change Service (2024)	

withdrawal_industrial_m3_year-1	m^3/year	Modelled total water withdrawal for the industrial sector in 0.5° raster grid; multi-model ensemble mean of H08, PCR-GLOBWB, and WaterGAP; based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data, Wada et al. (2022)	CC0 1.0
withdrawal_domestic_m3_year-1	m^3/year	Modelled total water withdrawal for the domestic sector in 0.5° raster grid; multi-model ensemble mean of H08, PCR-GLOBWB, and WaterGAP; based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data, Wada et al. (2022)	CC0 1.0
urban_area_fraction		Fraction of urban area in 0.5° raster grid based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data; Volkholz & Ostberg (2024)	CC0 1.0
pastures_fraction		Fraction of area with pastures in 0.5° raster grid based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data; Volkholz & Ostberg (2024)	CC0 1.0
cropland_rainfed_fraction		Fraction of area with rainfed cropland in 0.5° raster grid based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data, Volkholz & Ostberg (2024)	CC0 1.0
cropland_irrigated_fraction		Fraction of area with irrigated cropland in 0.5° raster grid based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data, Volkholz & Ostberg (2024)	CC0 1.0

forests_natural_vegetation_fraction		Fraction area with forests and natural vegetation in 0.5° raster grid based on yearly data; for daily and monthly time series, each day[month] within a month[year] is matched with the corresponding value for that month[year]	ISIMIP3a input data, Volkholz & Ostberg (2024)	CC0 1.0

References:

Amatulli, G., McInerney, D., Sethi, T., Strobl, P., & Domisch, S. (2020). Geomorpho90m, empirical evaluation and accuracy assessment of global high-resolution geomorphometric layers. Scientific Data, 7(1), 162. https://doi.org/10.1038/s41597-020-0479-6

Beck, H. E., Wood, E. F., Pan, M., Fisher, C. K., Miralles, D. G., Van Dijk, A. I. J. M., McVicar, T. R., & Adler, R. F. (2019). MSWEP V2 Global 3-Hourly 0.1° Precipitation: Methodology and Quantitative Assessment. *Bulletin of the American Meteorological Society*, *100*(3), 473–500. https://doi.org/10.1175/BAMS-D-17-0138.1

Chen, Z., Goldscheider, N., Auler, A., Bakalowicz, M., Broda, S., Drew, D., Hartmann, J., Jiang, G., Moosdorf, N., Richts, A., Stevanovic, Z., Veni, G., Dumont, A., Aureli, A., Clos, P., & Krombholz, M. (2017). World Karst Aquifer Map (WHYMAP WOKAM) [Dataset]. BGR , IAH, KIT, UNESCO. https://doi.org/10.25928/B2.21_SFKQ-R406

Cuthbert, M. O., Gleeson, T., Moosdorf, N., Befus, K. M., Schneider, A., Hartmann, J., & Lehner, B. (2019). Global patterns and dynamics of climate–groundwater interactions. Nature Climate Change, 9(2), 137–141. https://doi.org/10.1038/s41558-018-0386-4

Copernicus Climate Change Service, Climate Data Store, (2024): ERA5-land post-processed daily-statistics from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.e9c9c792 (Accessed on 27-JAN-2025)

Gleeson, T., Moosdorf, N., Hartmann, J., & Van Beek, L. P. H. (2014). A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity. Geophysical Research Letters, 41(11), 3891–3898. https://doi.org/10.1002/2014GL059856

Hartmann, J., & Moosdorf, N. (2012). The new global lithological map database GLiM: A representation of rock properties at the Earth surface. Geochemistry, Geophysics, Geosystems, 13(12), 2012GC004370. https://doi.org/10.1029/2012GC004370

Huggins, X., Gleeson, T., Serrano, D., Zipper, S., Jehn, F., Rohde, M. M., Abell, R., Vigerstol, K., & Hartmann, A. (2023). Overlooked risks and opportunities in groundwatersheds of the world's protected areas. Nature Sustainability, 6(7), 855–864. https://doi.org/10.1038/s41893-023-01086-9

Huggins, X., Gleeson, T., Villholth, K. G., Rocha, J. C., & Famiglietti, J. S. (2024). Groundwaterscapes: A Global Classification and Mapping of Groundwater's Large-Scale Socioeconomic, Ecological, and Earth System Functions. Water Resources Research, 60(10), e2023WR036287. https://doi.org/10.1029/2023WR036287

IGRAC. (2024a). The Global Groundwater Monitoring Network (GGMN) [Dataset]. IGRAC. https://doi.org/10.58154/6Z0Y-DA34 (Accessed [26-06-2024]).

IGRAC. (2024b). Well and monitoring data [Dataset]. Retrieved from https://ggis.un-igrac.org/view/well-and-monitoring-data/ (Accessed [26-06-2024]).

Karger, D. N., Conrad, O., Böhner, J., Kawohl, T., Kreft, H., Soria-Auza, R. W., Zimmermann, N. E., Linder, H. P., & Kessler, M. (2021). Climatologies at high resolution for the earth's land surface areasCHELSA V2.1 (current) (Version 2.1, S. 2.1 KB) [Geotiff]. EnviDat. https://doi.org/10.16904/ENVIDAT.228.V2.1

Koeppen, W., & Geiger, R. (1936). Handbuch der Klimatologie. Gebrüder Borntraeger.

Lehner, B., & Grill, G. (2013). Global river hydrography and network routing: Baseline data and new approaches to study the world's large river systems. Hydrological Processes, 27(15), 2171–2186. https://doi.org/10.1002/hyp.9740

Linke, S., Lehner, B., Ouellet Dallaire, C., Ariwi, J., Grill, G., Anand, M., Beames, P., Burchard-Levine, V., Maxwell, S., Moidu, H., Tan, F., & Thieme, M. (2019). Global hydro-environmental sub-basin and river reach characteristics at high spatial resolution. Scientific Data, 6(1), 283. https://doi.org/10.1038/s41597-019-0300-6

Meybeck, M., Kummu, M., & Dürr, H. H. (2013). Global hydrobelts and hydroregions: Improved reporting scale for water-related issues? Hydrology and Earth System Sciences, 17(3), 1093–1111. https://doi.org/10.5194/hess-17-1093-2013

Miralles, D. G., Bonte, O., Koppa, A., Baez-Villanueva, O. M., Tronquo, E., Zhong, F., Beck, H. E., Hulsman, P., Dorigo, W., Verhoest, N. E. C., & Haghdoost, S. (2025). GLEAM4: Global land evaporation and soil moisture dataset at 0.1° resolution from 1980 to near present. Scientific Data, 12(1), 416. https://doi.org/10.1038/s41597-025-04610-y

Muñoz-Sabater, J., Dutra, E., Agustí-Panareda, A., Albergel, C., Arduini, G., Balsamo, G., Boussetta, S., Choulga, M., Harrigan, S., Hersbach, H., Martens, B., Miralles, D. G., Piles, M., Rodríguez-Fernández, N. J., Zsoter, E., Buontempo, C., & Thépaut, J.-N. (2021). ERA5-Land: A state-of-the-art global reanalysis dataset for land applications. Earth System Science Data, 13(9), 4349–4383. https://doi.org/10.5194/essd-13-4349-2021

Simons, G., Koster, R., & Droogers, P. (2020). HiHydroSoil v2.0—A high resolution soil map of global hydraulic properties (Wageningen, The Netherlands). FutureWater report 134.

Vermote, E. & NOAA CDR Program. (2018). NOAA Climate Data Record (CDR) of AVHRR Normalized Difference Vegetation Index (NDVI), Version 5 [Dataset]. NOAA National Centers for Environmental Information. https://doi.org/10.7289/V5ZG6QH9

Volkholz, J., & Ostberg, S. (2024). ISIMIP3a landuse input data (Version 1.3) [Dataset]. ISIMIP Repository. https://doi.org/10.48364/ISIMIP.571261.3

Yamazaki, D., Ikeshima, D., Tawatari, R., Yamaguchi, T., O'Loughlin, F., Neal, J. C., Sampson, C. C., Kanae, S., & Bates, P. D. (2017). A high-accuracy map of global terrain elevations. Geophysical Research Letters, 44(11), 5844–5853. https://doi.org/10.1002/2017GL072874