Electronic Supplementary Information

A. Sensors and sites

Fig. A1. The sensor validation setup using a gage plate and camera.[1]

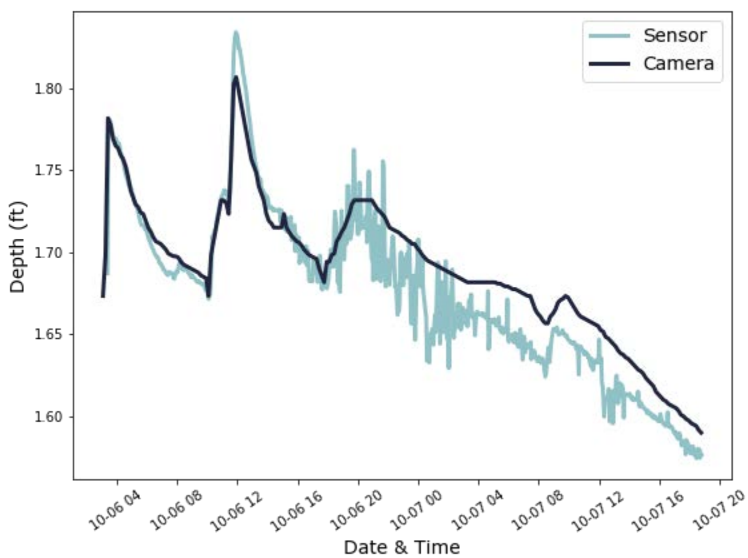

Fig. A2. An example of sensor and camera-measured depths in one GI asset from the validation study completed by an outside consultant (for more details, please refer to Dierks 2019[1]).

Table A1. List of the 14 monitored GI sites with their design features and physiographic features. The site type refers to if the site is a rain garden (RG) or a bioretention cell (BRC). The land use type refers to if the site is classified as a developed low intensity (DLI), developed medium intensity (DMI), or developed high intensity (DHI).

Site	Site Type	Surface Area (sq. m)	Drainage Area (sq. m)	DA/SA Ratio	Storage Volume (cu. m)	GI Depth (m)	Install Year	Percent Impervious	Land Use Type	Latitude	Longitude	Elevation (m)	Slope (%)	Hydrologic Soil Group	Depth to Groundwater (m)	Decay Constant (1/hr)	Avg. Drawdown Rate (cm/hr)	Avg. Captured Volume (cu. m)
S1	RG	15.8	60.4	3.8	3.5	0.3	2020	53%	DMI	42.38413	-83.07061	191.76	1.85	C	5.99	-0.040	0.614	1.974
S2	RG	18.6	241.5	13.0	2.2	0.3	2017	71%	DMI	42.38746	-83.07504	192.31	1.12	A	6.35	-0.011	0.261	2.759
S3	RG	16.4	23.2	1.4	1.0	0.3	2017	59%	DMI	42.36798	-83.09664	192.62	0.19	C	5.77	-0.044	0.626	2.839
S4	RG	27.9	173.7	6.2	6.6	0.3	2018	60%	DMI	42.42695	-83.10778	194.77	2.02	C	11.1	-0.305	7.317	3.681
S5	RG	11.1	23.2	2.1	2.5	0.3	2018	52%	DMI	42.38802	-83.07315	192.21	1.68	C	6.36	-0.146	0.751	1.489
S6	BRC	260.9	3106.7	11.9	14.1	1.0	2017	98%	DHI	42.36578	-83.06965	192.68	0.94	D	5.02	-0.024	0.255	29.423
S7	RG	14.9	40.5	2.7	3.3	0.3	2020	69%	DMI	42.37563	-83.10217	193.98	0.11	C	5.88	-0.069	0.771	2.126
S8	RG	20.4	49.1	2.4	4.1	0.3	2015	46%	DLI	42.39698	-83.26136	192.1	0.60	C	8.46	-0.397	2.972	2.102
S9	RG	39.0	163.5	4.2	8.9	1.0	2018	87%	DHI	42.40581	-82.94867	182.28	2.04	D	6.79	0.102	3.981	7.167
S10	RG	21.4	24.5	1.1	4.7	0.3	2019	61%	DMI	42.39074	-82.93565	178.21	0.42	D	6.37	0.119	4.521	4.218
S11	RG	9.3	53.0	5.7	0.9	0.3	2017	54%	DMI	42.37914	-83.09715	194.13	1.07	C	6.03	-0.200	4.713	1.738
S12	BRC	138.9	1740.1	12.5	17.3	1.0	2019	61%	DLI	42.39801	-82.96338	182.88	3.51	D	6.76	-0.047	0.797	14.197
S13	RG	33.4	371.6	11.1	8.0	1.00	2019	98%	DHI	42.41205	-83.06913	192.4	0.64	D	8.63	-0.072	0.494	5.216
S14	RG	13.9	40.1	2.9	1.6	0.3	2016	68%	DMI	42.36806	-82.95522	175.3	0.69	D	5.95	-0.021	0.410	1.680

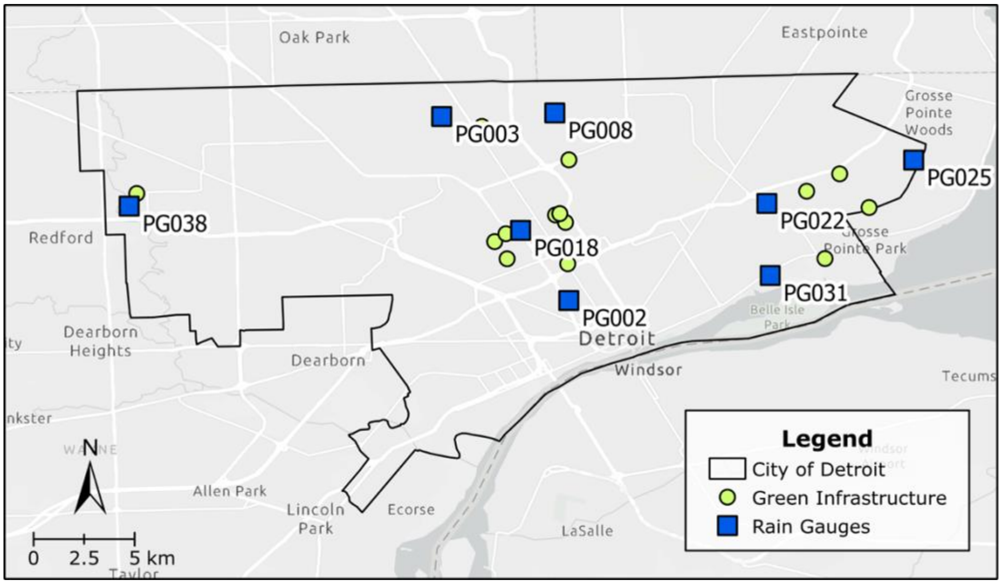

Fig. A3. The eight Detroit rain gauges used to analyze rainfall.

Table A2. Records of the field visits (deployment and maintenance) for the 14 GI devices.

Site ID	Deployment	Maintenance
S1	6/11/21	N/A
S2	6/18/21	N/A
S3	6/11/21	N/A
S4	6/11/21	N/A
S5	7/2/21	N/A
S6	6/11/21	N/A
S7	6/11/21	N/A
S8	7/2/21	N/A
S9	6/18/21	N/A
S10	6/18/21	N/A
S11	6/11/21	N/A
S12	6/18/21	8/11/21
S13	6/11/21	N/A
S14	6/18/21	N/A

B. GIS

GIS data pre-processing

The following steps were taken to download and pre-process the GIS datasets used in the correlation analysis using ArcGIS Pro.

1. Added a csv file with the spatial coordinates and decay constants of each GI location.
2. The GI data was displayed using the *Display XY Data* tool. Set the X field as longitude and the Y field as Latitude.
3. Reprojected the GI layer using the *Project* tool to GCS_WGS_1984.
4. Downloaded the City of Detroit Boundary JSON file from https://data.detroitmi.gov/datasets/detroitmi::city-of-detroit-boundary/about.
5. Converted Detroit boundary to shapefile and deleted center cutout of Hamtramck and Highland Park using the *Edit Vertices* tool.
6. Downloaded 3m (1/9th arc second) elevation data for Wayne County, Michigan, US from https://earthexplorer.usgs.gov/.[2]
7. Combined the eleven elevation images into one using the *Mosaic to New Raster* tool.
8. Added the USA SSURGO – Soil Hydrologic Group,[3] USA NLCD Land Cover, [4] and USA NLCD Impervious Surface Time Series[4] raster datasets from the ArcGIS Virtual Portal.
9. Reprojected the four raster datasets using the *Project Raster* tool to GCS_WGS_1984.
10. Clipped the four raster layers to Detroit boundary shapefile the *Extract by Mask* tool
11. Made all four raster files have the same resolution (30 m) using the *Resample* tool.
12. Used the elevation layer to create a new slope layer using the *Surface Parameters* tool.[5] Selected the quadratic option (which is default and recommended option for most data and applications), the default calculated neighborhood distance, the z unit was set to meter, and the output slope measurement was set for percent rise.
13. The land cover layer contains categorical data, so we needed to change these to numerical values for the correlation analysis using the *Reclassify* tool. The “High Developed Intensity” value was set to 3, “Medium Developed Intensity” to 2, “Low Developed Intensity” to 1, and all other categories to 1 since they are most closely related to “Low Developed Intensity”. “Open Water” was set to NODATA since we cannot install GI there.
14. Obtained the well data for Michigan from the State of Michigan’s Water Well Viewer,[6] Wellogic System,[7] and the US Geologic Survey’s Groundwater Watch.[8] Combined the three well datasets into one Excel file. Plotted histogram of static water level to check for outliers, kurtosis, and skewness to show it’s a normal distribution. The csv file was then added to ArcGIS Pro. The data is displayed using the *Display XY Data* tool. Set the X field as longitude and the Y field as Latitude.
15. Reprojected the wells layer using the *Project* tool to GCS_WGS_1984.
16. Interpolated groundwater levels using the *Empirical Bayesian Kriging* tool. The output cell size was set to the same size as the other raster datasets (30 m). Data transformation type was none and the semivariogram type was Power. Additional model parameters were 50 for the maximum # of points in each local model; 1 for the local model area overlap factor; and 1,000 for the number of simulated semivariograms. The search neighborhood parameters used a Smooth Circular search neighborhood with a smoothing factor of 0.85 and the default calculated radius (21,399 m).

17. Used *GA Layer to Rasters* to convert the geostatistical layer to a raster file for both the prediction and the prediction standard error. Masked it to the Detroit boundary shapefile and selected GCS_WGS_1984 as the projection.
18. Used the *Extract Multi Values to Points* tool to extract the values from each raster layer (elevation, slope, HSG, groundwater, imperviousness, land use type) at the GI locations.
19. Converted this data to an Excel file using the *Table to Excel* tool and loaded it into Python for the correlation analysis.

Table B1. Details on the GIS datasets including the year, source, type, and resolution.

Dataset	Year	Source	Type	Resolution
City of Detroit Boundary	2021	City of Detroit	Vector	N/A
National Elevation Dataset (NED) 1/9 Arc Second (3m)[2]	2017	USGS	Raster	3 m
USA SSURGO - Soil Hydrologic Group[3]	2021	Esri	Raster	30 m
USA NLCD Land Cover[4]	2019	Esri	Raster	30 m
USA NLCD Impervious Surface[4]	2019	Esri	Raster	30 m
Well Records[6]	2021	State of Michigan	CSV	N/A
Well Records[7]	2021	State of Michigan	CSV	N/A
Well Records[8]	2021	US Geologic Survey	CSV	N/A

Groundwater interpolation

Detroit has a shallow groundwater system, with the groundwater table being one to three meters below the surface in some regions.[9] For this reason, it is critical to include groundwater in the analysis. To the best of our knowledge, the only available groundwater data for the State of Michigan are a collection of water well records which provide the static water level (ft), or depth to the groundwater, for each well. Three sets of well records were found: (1) Water Well Viewer by the State of Michigan's Department of Environmental Quality;[6] (2) Wellogic System by the State of Michigan's Department of Environment, Great Lakes and Energy (previously the Department of Environmental Quality); [7] and (3) Groundwater Watch by the US Geologic Survey.[8] It is important to note that although the derived data in these files represents the best readily available data, they do not represent a complete database of all wells or well records in existence. The well records include three csv files with each well's ID, location (latitude, longitude), static water level, and other data that is irrelevant for this analysis. A limitation of these records is that there is a single static water level reading for each well. And since groundwater fluctuates, the variation in groundwater is missing, creating some uncertainty.

We want to use the groundwater data to see if there is a correlation between GI drawdown rates and the depth to groundwater. To do this, we need an estimate of groundwater depth for each GI, which requires the well records need to be manipulated into a usable form. The three water well records are combined into a single Excel file. A histogram of the static water level is plotted to check for outliers (Fig. B1). In addition, summary statistics are computed (Table B2). Since the values for kurtosis and skewness are between $\pm$ 2, this is considered acceptable to prove the data follows a normal univariate distribution.[10] Therefore, the full dataset is used (no outliers removed), and no data transformations are used. The dataset is added to ArcGIS Pro using the

Excel to Table tool, displayed using the *Display XY Data* tool, and then reprojected to "GCS WGS 1984". The next step is to interpolate the static water level across Detroit.

Table B2. Groundwater well record summary statistics.

Total Wells	1670
Mean (ft)	34.6
St. Dev. (ft)	22.7
Max (ft)	120.0
Min (ft)	1.0
Kurtosis	0.93
Skewness	0.96

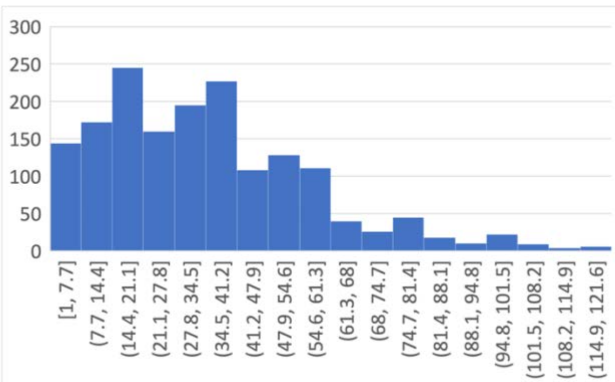

Fig. B1. Histogram of the static water level (ft) in the Detroit region water wells.

Kriging is an accepted method of estimating groundwater at sites where the water level data are available but where there may be insufficient additional data necessary for groundwater flow modeling.[11] Traditional kriging methods estimate a variogram that is considered the true variogram of the observed data without explicitly considering uncertainty. Recently, a new form of kriging, *Empirical Bayesian Kriging* (EBK), has been shown to perform better than other types of kriging methods.[12] It has been successfully used to evaluate inter-annual water-table evolution in Mexico[13] and to quantify uncertainty in groundwater modeling.[14] The fundamental advantage of EBK over classical kriging methods is that it creates a spectrum of variograms which account for the uncertainty introduced by estimating a variogram in the first place.[13] Therefore, EBK is used to interpolate groundwater following the methodology of Li et al. (2020).[13]

In ArcGIS Pro, the EBK tool is selected with the groundwater point data as the input feature. The Z value field was set to groundwater depth. The output cell size is set to the same size as the other raster datasets (30 m). The data transformation type is set to "None" because the data is normally distributed. There are three semivariogram options when the data transformation is "None": power, linear, and thin plate spline. Power is selected because it is relatively fast, flexible, and balances performance and accuracy. The search neighborhood parameters used a "Smooth Circular" search neighborhood with a smoothing factor of 0.85 and the default calculated radius (21,399 m).[13] ArcGIS Pro's leave-one-out cross-validation is used to find the remaining model parameters: maximum number of points in each local model, the local model area overlap factor, and the number of simulated semivariograms. Following the methodology of

Li et al. (2020),[13] we calculate the mean error (ME), root mean square error (RMSE), average standard error (ASE), mean standardized error (MSE), and root mean square standardized error (RMSSE) for each subset of parameters. The different errors from cross-validation ware analyzed with the rules in Table B3 to assess the variability of predictions and evaluate the performance (under or over-estimation) of the EBK model.

Table B3. Conditions to evaluate the performance of the EBK model.[13]

Check if the following hold:	**The prediction variability is:**
ASE ≈RMSE and RMSSE ≈ 1	Correctly assessed
ASE > RMSE and RMSSE < 1	Overestimated
ASE < RMSE and RMSSE > 1	Underestimated

Cross-validation results found the optimal parameters to be: 50 for the maximum number of points in each local model; 1 for the local model area overlap factor; and 1,000 for the number of simulated semivariograms. The interpolation errors suggest the EBK model performs correctly and does not under or overestimate groundwater (Table B4). As a secondary check, Li et al. reported an RMSE ≈ ASE ≈ 13.1614.43, and our values are close (16.4-16.7). [13] Fig. B2 shows the predicted versus true groundwater depths.

Table B4. ArcGIS Pro cross-validation report for the optimal parameter values: 50 for the maximum number of points in each local model; 1 for the local model area overlap factor; and 1,000 for the number of simulated semivariograms.

Inside 90% Interval	91.1
Inside 95% Interval	94.4
Mean	0.230
RMSE	16.7
MSE	0.00979
RMSSE	0.998
ASE	16.4

The EBK model with the optimal parameters is used to interpolate groundwater across Detroit. The *GA Layer to Rasters* tool is used to convert the geostatistical layer to a raster file for both the prediction and the prediction standard error. The output is set to be masked to the Detroit boundary shapefile and projected to “GCS WGS 1984”. The interpolated groundwater along with the groundwater wells are shown in Fig. B3a. The interpolated groundwater map aligns with the literature. Teimoori et al. found that the depth to groundwater is deepest in the northwest and gradually decreases as you move southeast.[9] The standard error of prediction plot shows the errors are higher in areas where well data does not exist (Fig. B3b).

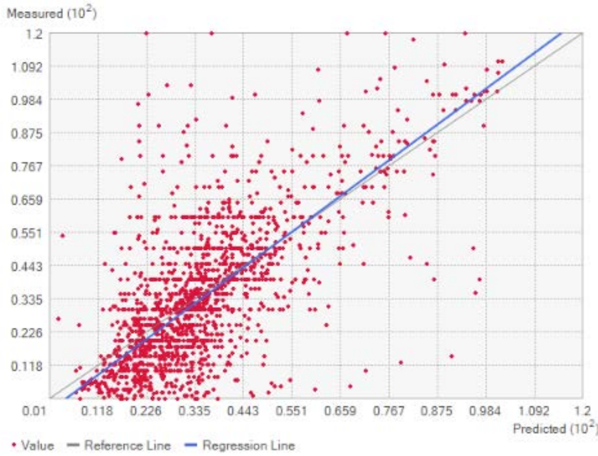

Fig. B2. The predicted versus true groundwater depth from the EBK model.

The final steps are to prepare the data for the correlation analysis. The *Extract Multi Values to Points* tool is used to extract the values from the groundwater raster layer at the GI locations and then it is converted to an Excel file using the *Table to Excel* tool. Then the groundwater data is also converted into an Excel file using the *Table to Excel* tool. Finally, these files are loaded into Python.

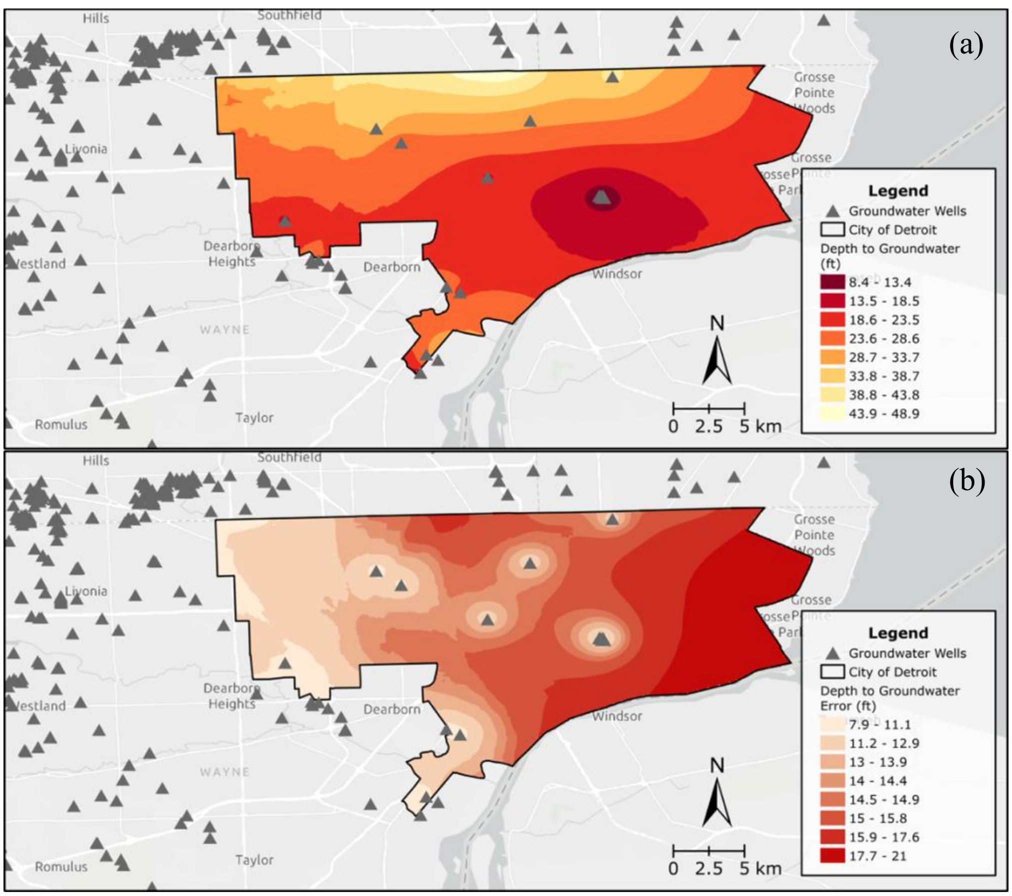

Fig. B3. (a) The groundwater wells and interpolated depth to groundwater (ft) for Detroit, Michigan, US. (b) The standard error of prediction of the interpolated depth to groundwater.

References

1 S. Dierks, in *Proceedings of the Water Environment Federation*, Water Environment Federation, 2019.
2 U.S. Geological Survey, *1/9th Arc-second Digital Elevation Models (DEMs) - USGS National Map 3DEP Downloadable Data Collection: U.S. Geological Survey.*, U.S. Geological Survey, 2017.
3 Soil Survey Staff, *Web Soil Survey*, Natural Resources Conservation Service, United States Department of Agriculture, 2021.
4 J. Dewitz and U.S. Geological Survey, 2021.
5 T. Hengl and H. I. Reuter, Eds., *Geomorphometry: concepts, software, applications*, Elsevier, 2008, vol. 33.
6 Department of Environment, Great Lakes, and Energy, *Water Well Viewer*, State of Michigan, 2021.
7 Department of Environmental Quality, *Wellogic System*, State of Michigan, 2021.
8 U.S. Geological Survey, *Michigan Real-Time Groundwater Level Network*, U.S. Department of the Interior, 2021.
9 S. Teimoori, B. F. O'Leary and C. J. Miller, Modeling Shallow Urban Groundwater at Regional and Local Scales: A Case Study in Detroit, MI, *Water*, 2021, **13**, 1515.
10 D. George and P. Mallery, *SPSS for Windows step by step: a simple guide and reference 18.0 update*, Allyn & Bacon/Pearson, Boston, MA, 11th ed., 2011.
11 R. M. Uetrecht, A. C. Elmore, J. D. Guggenberger and Z. D. Helwig, Practical considerations for kriging groundwater surfaces, *Remediat. J.*, 2019, **29**, 83–91.
12 K. Krivoruchko and A. Gribov, Evaluation of empirical Bayesian kriging, *Spat. Stat.*, 2019, **32**, 100368.
13 Y. Li, J. H. Hernandez, M. Aviles, P. S. K. Knappett, J. R. Giardino, R. Miranda, M. J. Puy, F. Padilla and J. Morales, Empirical Bayesian Kriging method to evaluate inter-annual water-table evolution in the Cuenca Alta del Río Laja aquifer, Guanajuato, México, *J. Hydrol.*, 2020, **582**, 124517.
14 H. M. Cooper, C. Zhang and D. Selch, Incorporating uncertainty of groundwater modeling in sea-level rise assessment: a case study in South Florida, *Clim. Change*, 2015, **129**, 281–294.