Although the choice of probability distribution for parameters has a limited influence on the overall uncertainty of LCA results, an effort was made to extract accurate representative distributions based on the available data. For rainfall (R), the annual depth was sampled from a normal distribution fitted to the historical annual records of 54 years (Figure S1). Since the Hydrologic and Hydraulic (H&H) model works with hourly data, the hourly pattern of the representative year was applied to the sampled annual depth. The available capacity (C) data were not as extensive as the available rainfall data. Therefore, based on engineering judgment and consultation with RWH planners, a gamma distribution was considered to reflect the higher possibility of installing smaller RWH systems and the unlikelihood of installing cisterns larger than 50 m^3. This distribution considers demands of different buildings in the area (estimated based on size of buildings) as well as building owners' willingness to participate in the RWH program (estimated based on available RWH plan experiences). The uncertainty analysis approach in the present study considered RWH implementation for every building at the watershed with the same C, although C may get a value of zero that is equivalent to no RWH implementation. Lastly, LCA model parameters were sampled from lognormal distributions based on the Ecoinvent data quality guideline. Additionally, the pedigree matrix approach adapted by this guideline was used to calculate the variance of the distributions. The highest level of recommended uncertainties for the pedigree matrix were considered in order to comprehend the threshold of their effects compared to hydrologic data. Specifically, the variances of the underlying normal distributions for reliability, completeness, temporal correlation, geographical correlation, and technological correlation were considered as 0.04, 0.008, 0.04, 0.002, and 0.12.

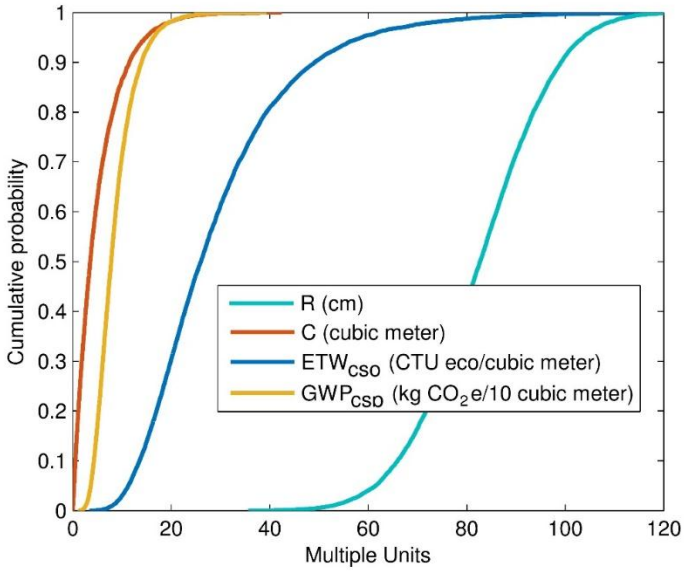

Figure S1. Cumulative probably density functions for parameters selected to perform the uncertainty analysis. A normal distribution for R (μ = 82.8, σ = 12.8), Gamma for C (α = 1, β = 0.2), lognormal for GWP_{CSD} (μ = -0.2, σ = 0.5) and lognormal for ETW_{CSO} (μ = 32.5, σ = 5.0) are considered.

The convergence criteria for the MC simulation was selected according to the Central Limit Theorem. For MC simulations, such theorem states that the term $\frac{\sqrt{n}}{\sigma}(E(Y) - \bar{Y})$ converges to a Gaussian random variable with a mean of 0 and a variance of 1, where n and σ respectively denote the number and standard deviation of samples, $E(Y)$ is the expected values of the selected output, and $\bar{Y}$ represents the average of the sampled outputs through Monte Carlo iterations. Since it is impossible to bound a random term, an inequality form of the above theorem is utilized in this research considering a confidence level of 95%. With a probability of 95% and for large values of n, convergence occurs when the absolute value of $E(Y) - \bar{Y}$ approaches values smaller than $1.96\frac{\sigma}{\sqrt{n}}$. The convergence criteria were checked after each run, and it was found that the algorithm was converged after 5,657 iterations. However, the calculations were continued until 10,000 iterations, and the results were plotted according to this number of iterations to be conservative.

After performing the MC simulation, the portion of uncertainty propagated by each parameter was calculated using the First-Order Sensitivity Analysis method for the MC results as:

$$Var(O) = \sum_{i=1}^{n} \left(\frac{\delta F}{\delta X_i}\right)^2 \times Var(X_i)$$

where F is the uWISE model, X denotes the parameters, O represents the model output and n is the number of considered parameters (4 in this study). This equation is applicable only when all the parameters are independent of each other. In this study, independency of parameters was verified using the Mutual Information (MI) index. Values of MI close to zero show the statistical independency of studied variables, while values higher than 1 represent a meaningful dependency. MI for LCA model parameters presented the highest value among all the permutations of the parameters. However, this value was 6.4×10^{-6}, which represent an insignificant dependency. In addition, $\frac{\delta F}{\delta X_i}$ was estimated using a numerical approximation, $\frac{\Delta F}{\Delta X_i}$, according to points adjacent to the output expected value. To control the accuracy of this approximation, the output variance calculated by the above equation was compared with the variance of the MC simulation for the uWISE outputs.

After performing iterative simulations for uncertainty/sensitivity analysis, engineers are usually interested in understanding different gradient behavior of system responses by means of scatter plot analysis. A traditional approach to perform such analysis is manually dividing the domains of the scatter plots achieved from uncertainty/sensitivity analyses. The Morse-Smale complex decomposition suggests a scientific solution for this need, since it can routinely divide the domain into regions of uniform gradient flow, as such each partition is associated with exactly one minimum and one maximum on the partition's boundary. In sum, the goal of Morse-Smale complex decomposition is to capture the geometry of the regression surface, instead of focusing on a quality of fit measure for splitting the domain that pays low attention to the geometry outputs of the studied system.

Therefore, the results of Morse-Smale complex decomposition consist of regions of uniform gradient flow that may be satisfactory represented by linear regression models (Morse-Smale regression), if model simplification is the goal of study. The other results are a set of local optimal points that can provide insight into interesting system responses. Compared to global optima, local optima are of a greater interest for system response analysis, because global optima are only two points (one global minimum and one global maximum) that are unable to provide information on different system responses.

In the case of this research, global optima happen with extrema in rainfall data because rainfall appeared to be the most sensitive dimension. However, we are unable to govern the rainfall to achieve minimized impacts. Thus, we are interested in understanding the effects of other system components, e.g. RWH system, on life cycle environmental impacts. The values of RWH system capacity that lead to optimal impacts are in fact local optima in our case.

However, due to the sparse nature of sampling, it is often the case that extraneous local minima and maxima may occur in the data when Morse-Smale decomposition is used. In order to filter out such insignificant features occurring in the data, a measure, called "persistence", is presented. Persistence is a measure of the amount of change in the function (uWISE in this case) required to remove a topological feature (i.e., local minimum or maximum), and thus merge two (or more) partitions. Low values of persistence are referred to as "noise" in order to emphasize the insignificance of such points (Figure S2). On the other hand, high persistence values are named "pattern" (Figure S2). In order to filter out insignificant local optima, those that have a persistence lower than a user-defined threshold are discarded.

Using this technique, the main drivers within local regions of the uWISE domain were discovered, as illustrated in Figure S2. The results of this analysis were two partitions with significant optima shown in gold and blue. In Figure S2a, x-axis is persistence and y-axis is Global Warming Potential (GWP). Red triangles show the maxima and blue triangles show the minima. For optima that have a persistence higher than the threshold, the triangles are shown with a bigger size to emphasize their significance. The two partitions had the same minimum but different maxima. Analysis showed that the gold partition result in loss of RWH potable water savings and CSO control benefits, while blue partition incurs excessive wastewater treatment burden. Number of points in each partition is also presented in this figure. Note that the sum of counts is 10,001 because the two clusters share one point (a minimum). Curvature of the gold partition is only used for aesthetics and does not have a value.

Additional information was also provided by the Morse-Smale regression tool. The partitions boundary appeared to be predominantly governed by C, which is RWH capacity (Figure S2b). The partitions exhibited the strongest linear regression parameters to R (annual rainfall depth), then to C, and lastly to GWP_{CSD} (Figure S2c), where CSD denotes the combined sewage volume delivered to WWTP. The partitions had two different responses to C, i.e. a direct correlation in the gold partition and an inverse correlation in the blue (Figure S2c). Figure S2d shows the R^2 coefficients for stepwise regression, i.e. adding the parameters one by to a linear model (from left to right as presented in Figure S2d). This figure shows after adding C, the improvement in R^2 is trivial, which affirms the results of Figure S2c.

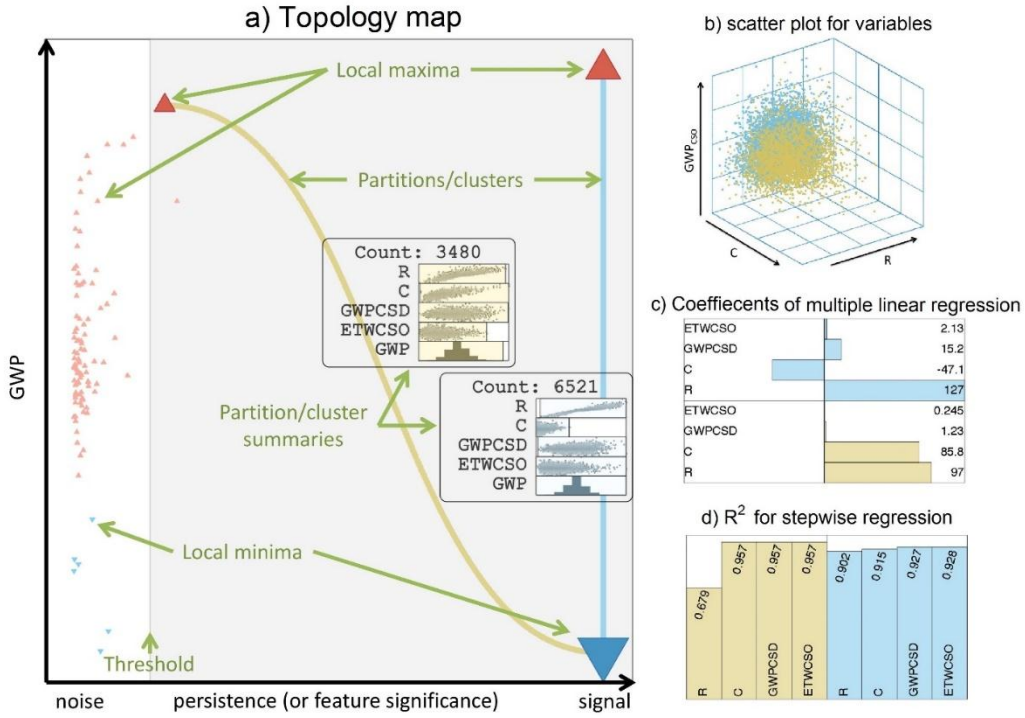

Figure S2. Results of topology-inspired regression model: (a) topology map and detected partitions, (b) scatter plot on variable space for the detected partitions, (c) coefficient of determination for linear regression in each partition, (d) R^2 for stepwise regression in each partition.

The studied watershed, namely Eastside, consists of 41 subwatersheds with a total area of 9.54 km^2. A location map of the tributary subwatersheds, major pipes and the interceptor of the study area is provided in Figure S3. An interceptor in the Eastside with the capacity of 0.25 MCM/day collects combined sewage and delivers it to the Bay View wastewater treatment plant (WWTP) located at the northeast of the study area.

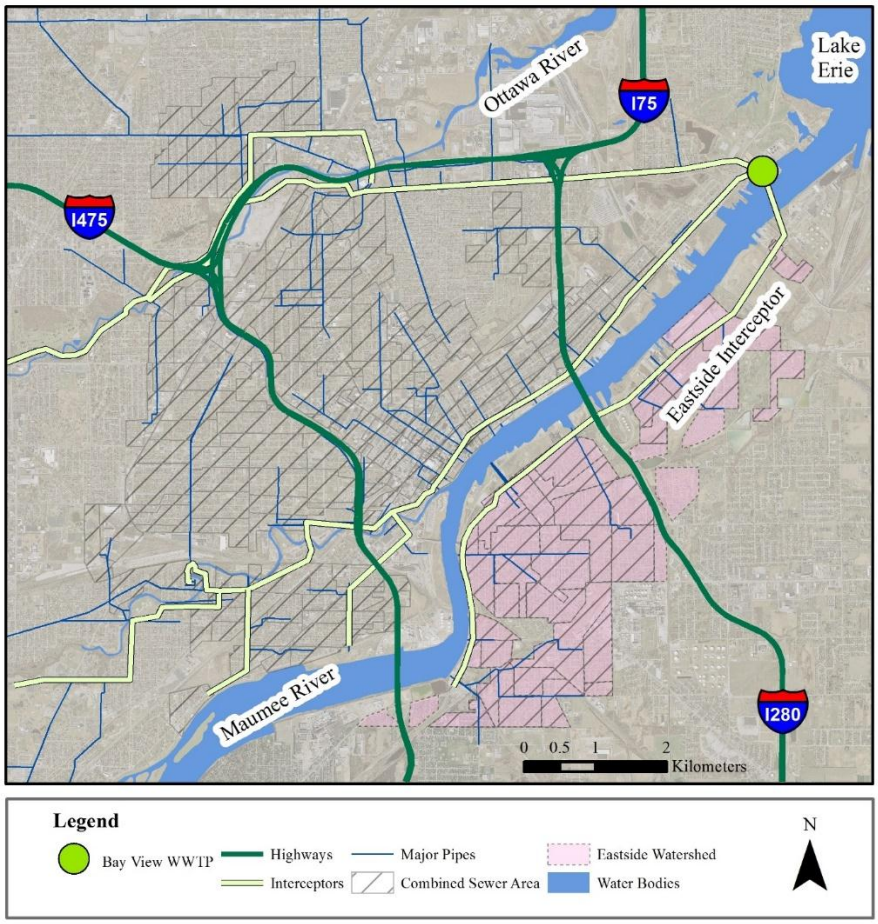

Figure S3. Tributary subwatersheds, major pipes and the interceptor at the case study. Subwatersheds have predominantly combined sewers, with a few cases of separate sewers in the southern and eastern areas. Data are adapted from City of Toledo and AREIS.