

A novel computational green infrastructure design framework for hydrologic and human benefits

Ankit Rai^{a,*}, Barbara Minsker^b, William Sullivan^c, Lawrence Band^d

^a Informatics, University of Illinois at Urbana Champaign, United States

^b Department of Civil Engineering, Southern Methodist University, United States

^c Department of Landscape & Architecture, University of Illinois at Urbana Champaign, United States

^d Department of Civil and Environmental Engineering, University of Virginia, United States

ARTICLE INFO

Keywords:

Green stormwater infrastructure
Hydrologic modeling
Computer vision
Human preference
Machine learning

ABSTRACT

Increased storm-water runoff and flooding and poor ecosystem health have brought increasing attention to catchment-wide implementation of green infrastructure (e.g., bioswales, rain gardens, permeable pavements, tree box filters, urban wetlands and forests, stream buffers, and green roofs) to replace or supplement conventional storm water management practices and create more sustainable urban water systems. Current green infrastructure (GI) practice aims at mitigating the negative effects of urbanization by restoring pre-development hydrology and ultimately addressing water quality issues at an urban catchment scale. However, the benefits of GI extend well beyond local storm water management, as urban green spaces are also major contributors to human health. Considerable research in the psychological sciences have shown significant human health benefits from appropriately designed green spaces, yet impacts on human wellbeing have not yet been formally considered in GI design frameworks. This work develops a novel computational green infrastructure (GI) design framework that integrates storm water management requirements with criteria for human wellbeing. A supervised machine-learning model is created to identify specific patterns in urban green spaces that promote human wellbeing; the model is linked to RHESSYS hydrological model to evaluate GI designs in terms of both water resource and human health benefits. An application of the framework to tree-based GI design in Dead Run Watershed, Baltimore, MD, shows that image-mining methods are able to capture key elements of human preferences that could improve GI design. The results also show that hydrologic benefits associated with tree-based features are substantial, indicating that increased urban tree coverage and a more integrated GI design approach can significantly increase both human and hydrologic benefits.

1. Introduction

The rapid growth of urbanization interferes with natural water and nutrient cycling by increasing impervious surfaces, which in turn increases flashiness of urban drainage systems and reduces water quality, causing human health and ecosystem problems downstream (NRC, 2008; Wendel et al., 2011). This problem has increased interest in using green infrastructure in urban areas (e.g., bioswales, rain gardens, permeable pavements, tree box filters, cisterns, urban wetlands, and green roofs). LeRoy Poff (1997) suggested that implementation of these practices at the watershed scale could restore the riverine ecosystem along with addressing water quality issues. Roy et al. (2008) considered watershed-wide implementation of these approaches as a prerequisite for sustainable urban water systems.

Currently, green infrastructure design guidelines provide site-

specific (patch) design criteria with only qualitative discussion of catchment-scale impacts of GI installations (e.g., CalTrans (2010); City of Portland (2008); Harper and Baker (2008); NCDWQ (2007). Catchment-scale lumped-parameter stormwater models (e.g., MARC (2008); Tsihrintzis and Hamid (1998); and tools such as HSPF, SWMM and HEC-HMS) do not represent site-specific hydrology or GI processes. In these catchment-scale models, both traditional (“grey”) and green infrastructure have typically been modeled as “edge-of-field” or “in-line” filters and sinks for storm-water runoff received from source catchment areas. Attenuation of storm-water volumes and pollutants are often included as fixed reduction percentages or first-order decay reactions based on limited input and output water quality measurements (e.g., Lee and Riverson (2012), in the SUSTAIN modeling framework, and Wong et al. (2002), in the MUSIC framework).

Research over the past decade as part of the Baltimore Ecosystem

* Corresponding author.

E-mail address: rai5@illinois.edu (A. Rai).

Study suggests that significant carbon sequestration and nitrogen retention can occur in a range of urban ecosystem features, including lawns, gardens, and stormwater detention structures, but that these processes are sensitive to specific characteristics of the integrated drainage system, including contributing areas, flow regimes, soils, and structure design (e.g. [Raciti et al. \(2011\)](#); [Bettez and Groffman \(2012\)](#)). Living components of green infrastructure will grow and adjust to prevailing water, climate, and nutrient conditions, and there may be a long, transient development of ecosystem cycling and retention capacity following development. Design of sustainable green infrastructure as either edge-of-field or at-source treatment should incorporate transient development as the ecosystem develops in response to local climate, soil, and drainage position (e.g. location within a flow field). It is critical that GI modeling extend to encompass the full catchment as a continuum beyond the discrete GI sites, including runoff source areas in addition to edge-of-field or in-line treatment systems.

Most of the current research on green infrastructure designs considers only hydrologic benefits ([Macro et al. \(2018\)](#); [Massoudieh et al. \(2017\)](#); [Glenis et al. \(2018\)](#)). However, the benefits of green infrastructure extend well beyond local storm water control, as urban green spaces (e.g., lakes, parks, and community gardens) are also major contributors to human health. Forty years of research has established the powerful and consistent effects of the presence of particular natural elements in increasing human preferences for urban landscapes ([Kaplan and Kaplan \(1989\)](#)). These high-preference elements, in turn, are associated with faster recovery from stressful experiences, reduced physiological symptoms of stress ([Thompson et al. \(2012\)](#); [Chang and Chen \(2005\)](#)), and increased life expectancy after controlling for a host of features associated with mortality ([Mitchell and Popham \(2008\)](#); [Takano et al. \(2002\)](#)).

Furthermore, human preferences are also important to address major barriers to GI implementation that were identified in our work with advisors from five U.S. cities as part of a national GI working group at the National Socio-Environmental Synthesis Center (SESYN). Community attitudes and perceptions about GI installations can lead to resistance and even active vandalism (e.g., throwing trash into a bioswale or mowing “weedy” plants) of GI. Maintenance of distributed GI is also a major challenge for cities. If GI can be designed to better meet human perceptions, then community acceptance and adoption of GI maintenance would likely improve.

Researchers in both the field of psychology and computer science have been working together for more than two decades to solve the problem of capturing human perception by means of human computer interactions. Significant efforts have been made to build machine-learning models that replicate and predict human perception to improve social acceptability of designs, strategies, policies, and marketing approaches, among others. Machine perception cannot perform effectively without a wealth of experimental data about human perception. In this work, we explore how physiological preferences for GI can be captured with machine learning.

Despite the extensive research showing the benefits of green spaces on human health, GI design has not yet considered criteria for human wellbeing in any formal design frameworks. In this research we propose a new GI design framework that considers both human and hydrologic benefits at patch and catchment scales. To aid in rapid initial evaluation of potential GI designs, we have developed a human preference model that predicts increasing human benefits using supervised machine learning algorithms and computer vision techniques.

The training data for this work has been collected using social surveys conducted at Galesburg, IL and the model is validated using data from Amazon Mechanical Turk. The resulting algorithm is then coupled with a hydrologic model called RHESSYS to estimate human benefits for a GI design case study in Baltimore, MD. To our knowledge, this is the first study that quantifies the significance of urban green infrastructure design for both human wellbeing and hydrologic benefits.

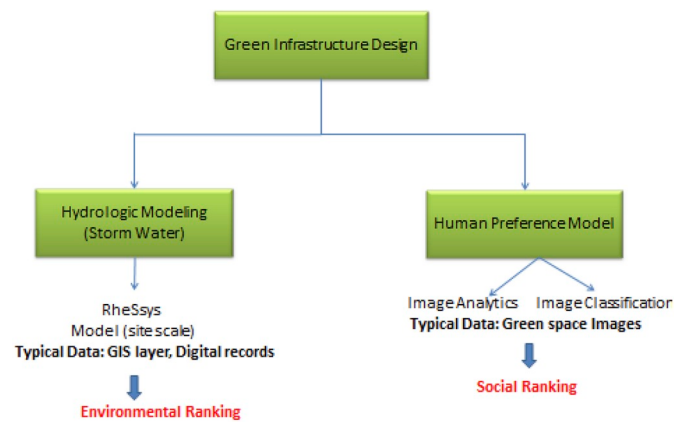


Fig. 1. Green infrastructure design framework. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

In the subsequent sections, we introduce the methodologies for quantifying human preference with an image-based machine learning approach and assessing hydrologic benefits, followed by a case study and results that explore the implications of this approach in Baltimore, MD. The final section provides conclusions and suggestions for future research on this topic.

2. Methodology

This section presents the fundamental concepts and technologies used to create a GI design framework that integrates human preferences for green spaces, which are correlated with improved human health ([Sullivan et al., 2004](#)), and hydrologic benefits.

An overview of the green infrastructure design framework is given in [Fig. 1](#). In order to predict GI human benefits, data acquisition (Section 2.1) and an image-based machine learning approach (Section 2.2) are used to visually identify landscape features in design images that humans prefer, and are therefore linked with improved human health, and provide a quantitative ranking of the design. To link these human preference criteria with storm water requirements, RHESSYS ([Tague and Band, 2004](#)) is used to predict hydrologic benefits of the design's landscape features, as described in Section 2.3 ([Table 2](#)).

2.1. Data acquisition

Previous research has shown that preferences of stakeholders can be predicted using digital images that visually represent potential GI designs ([Sullivan, 2004](#)). For example, [Fig. 1](#) shows GI design images from Galesburg, Illinois, and their human preference rankings.

Both images in [Fig. 2](#) are from the same setting, with the addition of trees to both sides of the street in (b) resulting in higher preference ratings.

Similar images of streetscapes with varying tree cover and density from Galesburg, IL, were shown to 30 participants in a survey conducted



(a) A low preference setting (b) A high preference setting

Fig. 2. Sample GI design settings in Galesburg, IL (Yuan, 2010).

Please take a quick look at the following photos. It is not necessary to study the picture in depth. How much do you approve of the streetscape design?



- Very much – Quite a bit – Somewhat – A little – Not at all

Fig. 3. Sample question from survey.

by co-author Sullivan at the University of Illinois (Yuan, 2010). The participants were asked how much they like/dislike each of 360 images, which were presented in random order. Below is a sample question from the survey.

Fig. 3 is illustrative of the inputs to the human preference model, which are the raw RGB images from Sullivan's study with labels spanning five categories.

2.2. Human preference modeling approach

An image-based supervised machine learning technique is trained to predict human preferences, or relative human health benefits, of tree-based GI as shown in Fig. 4. The approach automates the prediction of human preferences from GI design images by identifying (extracting)

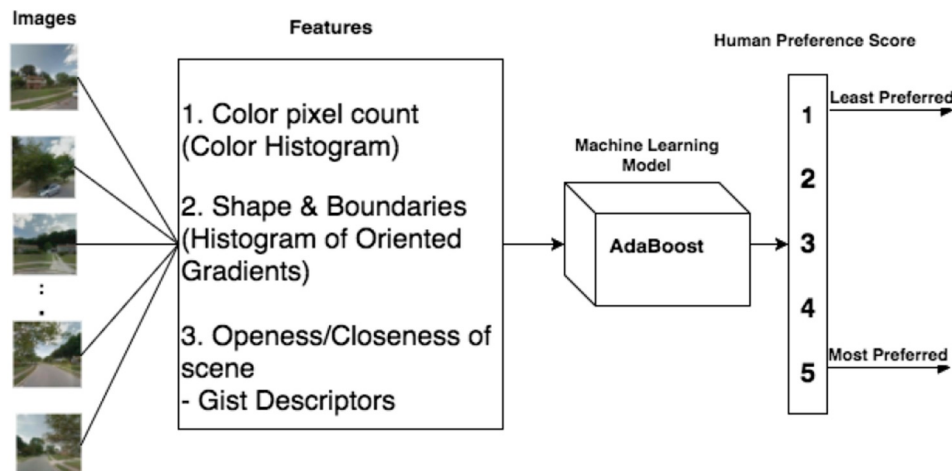


Fig. 4. Stages in the image-based human preference model for GI designs.

Table 1
Preference matrix.

	Understanding	Explanation
2D (Immediate/Present)	Coherence: How clear and orderly the green setting appears. Color histogram identifies green shapes and their layouts.	Complexity: richness & variety in visual components of the setting. GIST descriptor identifies openness.
3D (Future/Promised)	Legibility: distinctiveness of the setting EOH identifies distinctive shapes of trees and pathways.	Mystery: extent of features partially hidden from view (e.g., via curving pathways) GIST descriptor identifies their openness.

Table 2

Results for both the training and validation image sets from Galesburg.

Sample Set	Prediction Accuracy
Training	0.96
Validation	0.81

specific landscape features that correlate with high human preferences (Kaplan and Kaplan, 1989) using computer vision techniques (e.g., Fig. 2). A supervised machine-learning model is then trained to predict a human preference rating for the image based on the extracted features. Details on this approach are provided in the following sub-sections.

2.2.1. Feature extraction

To identify which image features to include in the model, Table 1 maps available image segmentation algorithms to a human preference matrix developed by Kaplan and Kaplan (1989) for urban green spaces, which gives GI characteristics that are most linked to human wellbeing. The column on the left gives landscape features that attract people and engage them longer by promoting understanding, while the column on the right are features that encourage exploration of the landscape. These features and algorithms are defined in more detail below.

2.2.2. Color histogram

Color histogram quantifies colors distribution of pixel intensities in an image (Anami et al., 2010). Since the color green plays an important role in GI design, we implemented RGB 256 bin color histogram (Chapelle et al, 1999). In this approach, color is represented by a three-dimensional vector for each color channel, i.e. RGB and it corresponding to a pixel point in color space. Cha Chapelle et. al, 1999 showed that for image classification the use of color space (i.e RGB (Red Green Blue) vs. HSV (Hue Saturation Value)) has very minimal effect on

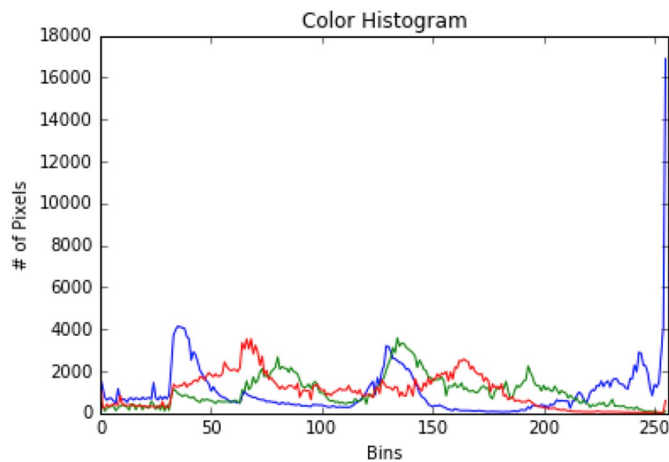


Fig. 5. Color histogram. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

classification accuracy. Therefore, we used RGB color space in this work. Fig. 5 shows the color histogram for the sample survey image shown in Fig. 3.

2.2.3. Histogram of oriented gradient

The histogram of oriented gradient (HOG) is used to extract the distribution of edges of objects in an image (Dalal and Triggs (2005)). HOG finds distinctive shapes of objects present in a design image (e.g., the shapes of trees shown in Fig. 3), which are needed to determine the legibility feature in Table 1 that affects human understanding of the scene. We implemented the HOG feature based on the approach of Dalal and Triggs (2005), which gives a 31-dimensional descriptor. HOG features are computed over rectangular regions represented by a grid of 3x3 blocks with an overlap of 8 pixels. Each block contains 16x16 pixels and HOGs are computed for each block with 9 unsigned orientations. Orientations are the gradient vectors at each pixel (for 16x16 pixel cells) that are categorized into a 9-bin histogram. The histogram ranges from 0 to 180° (because the orientations are unsigned), so there are 20° per bin.

2.2.4. Gist descriptors

The previous two image features identify specific objects from the images. Oliva and Torralba (2001) proposed a technique to estimate the structure or shape of a scene using spatial properties of the scene, called Gist descriptors (e.g., degree of naturalness, openness, roughness, or expansion). Fig. 6 shows the Gist descriptor for a sample input green space image. The results show the local energy spectrum, which measures the quantity of radiation passing through or emitted from the

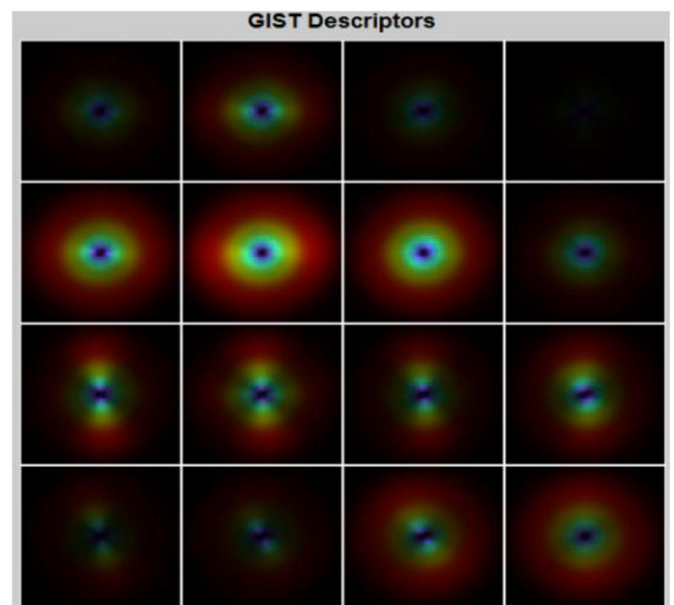


Fig. 7. Gist descriptor for sample image.

surface of an object, for all objects in the image. The image is divided into a 4x4 cell matrix for spatially grouping objects.

Each sub-image of the matrix corresponds to the amount of radiation (with multiple wavelengths) emitted or passed through objects placed in the corresponding spatial location. Dark regions in the gist descriptor matrix (e.g., the upper right and lower left sub-images in Fig. 10) thus identify openness in the image. Objects such as the tree in the middle of the input image on the left in Fig. 7 emit higher radiation and thus have more energy spectrum captured by the Gist descriptors in the corresponding sub-image. This spatial distribution of local energy spectrum provides a more holistic representation of a scene.

Once the GI design image features are extracted using the methods described above, a supervised machine learning approach is used as a regression model to predict human preferences for each design. Since we average ratings from all survey participants, the preference scores range continuously between 1 and 5 (i.e., real numbers). We concatenate the above three features into a combined {Color hist, HOG, Gist} feature set and use these features for training the model.

For the supervised machine learning, the adaptive boosting algorithm (AdaBoost) is used, which is an ensemble-learning algorithm that combines a weak classifier [e.g., decision trees or k-nearest neighbors (KNN)] to output a strong regressor (Freund and Schapire, 1999). AdaBoost gives high accuracy and prevents over-fitting of parameters to

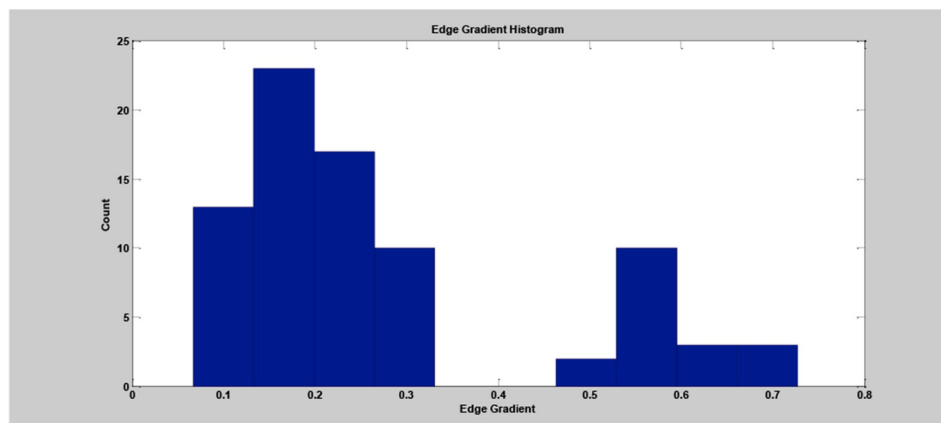


Fig. 6. Histogram of oriented gradient.



Fig. 8. Examples of validation results for Galesburg image data using the model.

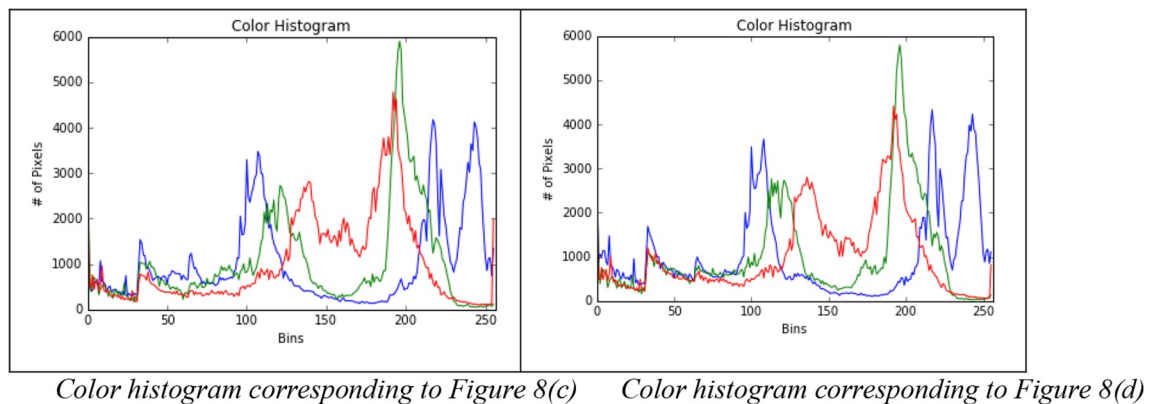


Fig. 9. Color histogram indicating differences in coherence. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

the training dataset by using an iterative learning model, which improves accuracy by learning from mistakes made in an earlier training step. We use decision trees as the weak classifier in this work.

2.2.5. Model validation & evaluation

We split the training data, which consisted of 360 images, into training (70%) and validation (30%) sets, tuning the model for optimal performance using the training set. The accuracy of the model is computed using the coefficient of r^2 .

Fig. 8 shows examples from the validation set. The predicted ratings for each of these images were identical to the actual human ratings.

The results indicate that the machine-learning model predicts human preferences reasonably well, particularly since human ratings can have considerable uncertainty (James et al., 2009).

The training set is predominantly tree-based green infrastructure, and generally human preferences increase with more trees in images, as shown in Fig. 8. However, merely adding more trees, bushes, plants, etc. are not sufficient to obtain high human preferences. The image features extracted from the training set (color histogram, EOH, spatial histogram, GIST descriptor), and their corresponding landscape features from Kaplan & Kaplan's preference matrix (1998), also play an important role.

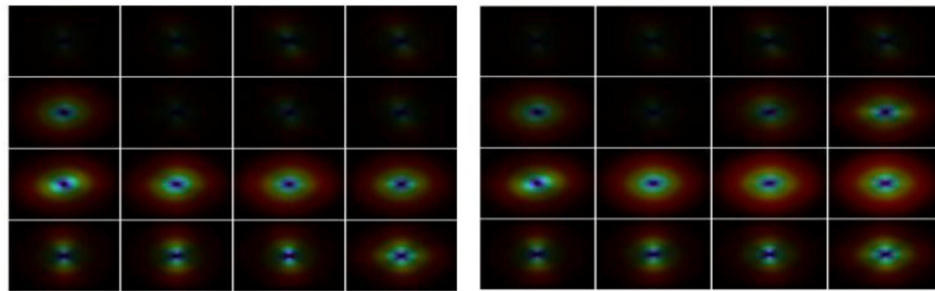
For example, recall that color histogram measures the green coloring and coherence in the design. Fig. 7 (c) and (d) show that the rating changes from 1 to 3 just by adding trees, which provides more green richness, and also by arranging the trees in an orderly manner along the sidewalks (e.g., Fig. 8(d) is more coherent than Fig. 8(c)). This added coherence can be seen in the color histogram results for these images in Fig. 9.

Complexity, legibility, and mystery parameters of the preference matrix also change human preferences. These parameters are captured by the Gist and HOG descriptors as shown in Fig. 10.

2.2.6. Evaluation using Amazon Mechanical Turk

Finally, to further evaluate the models with more diverse human preference scores (Buhrmester et al, 2011), we also conducted an online survey with additional participants using Amazon Mechanical Turk. The nature of the questionnaire was the same as shown in Fig. 2. We posted 30 different images of GI and asked 20 turkers to rate the images. After we obtained the turkers' ratings, we ran our human preference model to see how well it performs on the same 30 images using Pearson correlation coefficient (ρ). We obtained $\rho = 0.83$ for the turkers' ratings and model predictions, which shows that the model can identify key features of GI spaces affecting human preference from

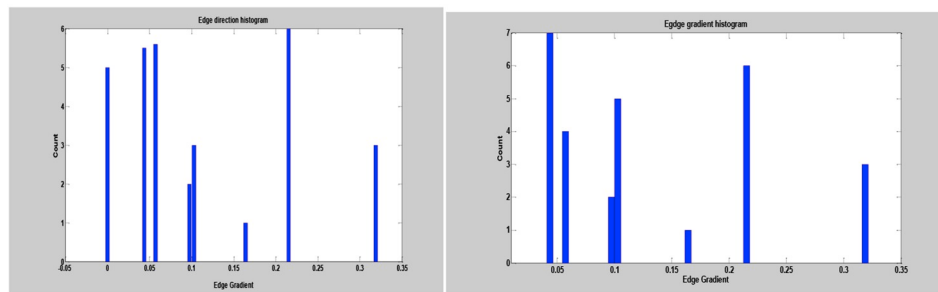
GIST Descriptors



(a)

(b)

HOG:



(c)

(d)

Fig. 10. GIST and HOG image descriptors for Fig. 8(a),(b) reveal differences in complexity and legibility.

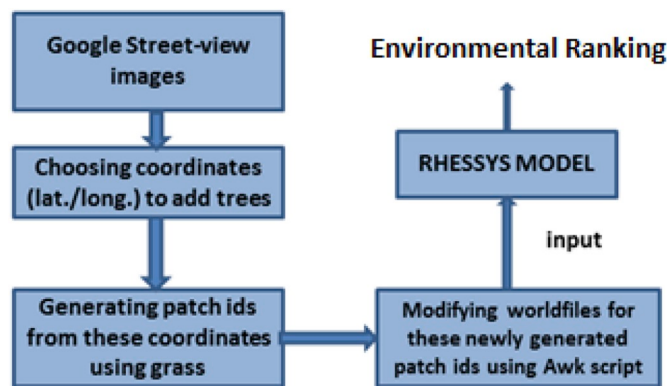


Fig. 11. Framework of the steps involved in estimating GI hydrologic benefits using RHESSYS.

visual scenes, despite the inherent variability in human preference ratings.

2.3. Evaluating the hydrologic benefits of GI design

In order to evaluate the hydrologic benefits of candidate GI designs, RHESSYS distributed parameter hydrologic model (Tague and Band, 2004) is used. RHESSYS is designed to simulate integrated water, carbon, and nutrient cycling and transport over spatially variable terrain at small to medium scales (i.e. from 1st to 3rd order watersheds). Its spatially distributed framework enables the modeling of spatio-temporal interactions between different eco-hydrological processes from patch to watershed scales (Hwang et al., 2012).

Fig. 11 gives an overview of the steps involved in preparing the required input for GI simulation using RHESSYS and integrating it with the human preference model for social ranking. Google street view

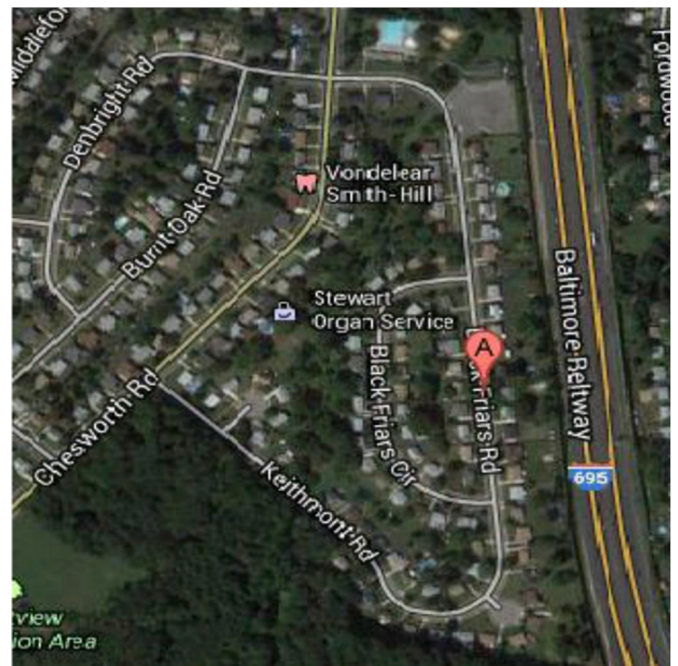


Fig. 12. Map of Baltimore neighborhood examined for GI design.

images are first extracted from the Google Maps application-programming interface (API) (<http://maps.google.com>) for the neighborhood where GI design is planned. The coordinates (latitude/longitude) are identified using Google Maps for the front yard of every house in the neighborhood where GI installation is under consideration. These coordinates are used to generate patch identifications (ids) that are used as inputs for GI parameters in RHESSYS. Patch ids are unique numbers

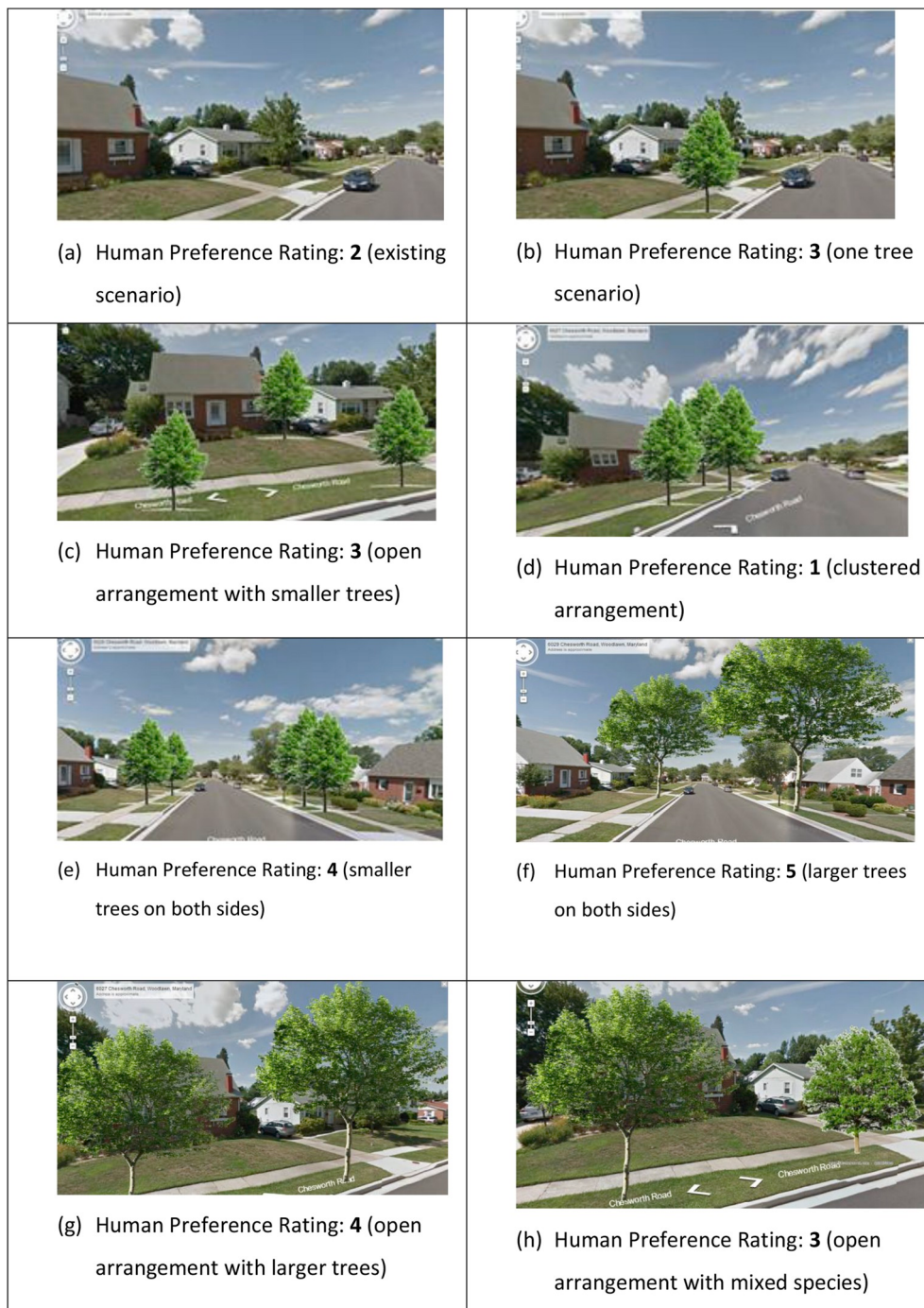


Fig. 13. Results of human preference modeling for different scenarios in Baltimore.

associated with particular ecosystem patches, which are the smallest-resolution spatial units that define areas of similar soil moisture and land-cover characteristics. In landscapes modified by humans, patches can also be defined to contain stream channels, road segments, storm sewers, etc. Human sources of water and nutrients are also defined at this level (Tague and Band, 2004).

Therefore, adding trees to the existing patches requires reflecting the presence of trees in the existing lawn by generating patch ids from their respective coordinate values. Patch id numbers are generated using GRASS GIS (<http://grass.osgeo.org>), which queries the existing raster map layers to output labels associated with input coordinate values using the ‘r.what’ command. These labels are patch id numbers that represent the current vegetation type present in the raster map.

The next step in executing RHESSYS is to prepare additional non-spatial files, called worldfiles, which represent landscapes patterns such as land use, tree canopy, etc. within RHESSYS. Worldfiles are generated with a GRASS interface program that references input raster maps and a text document defining initial state variables (e.g., saturation deficit within the patch level, which is a measure of the degree of saturation). Within the worldfile, each spatial layer is associated with the following identifiers: an assigned ID based on the map used and state variable values for stores and fluxes at each spatial level that are initialized at the start of the simulation. The *Grass2World* (g2w) program is used to generate the worldfiles automatically from spatial data layers within GRASS GIS. Next, in order to incorporate new trees into the existing worldfile based on the newly generated patch id numbers, an Awk

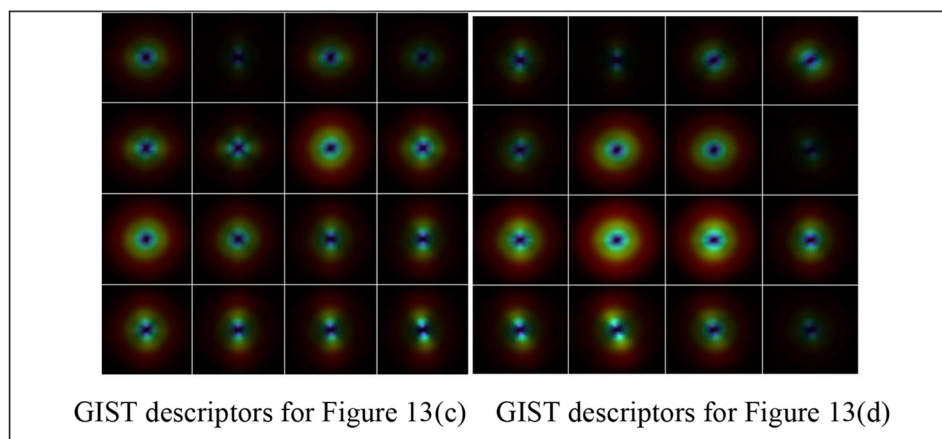


Fig. 14. GIST descriptors reflect the increased complexity of clustered trees in Fig. 13(d).

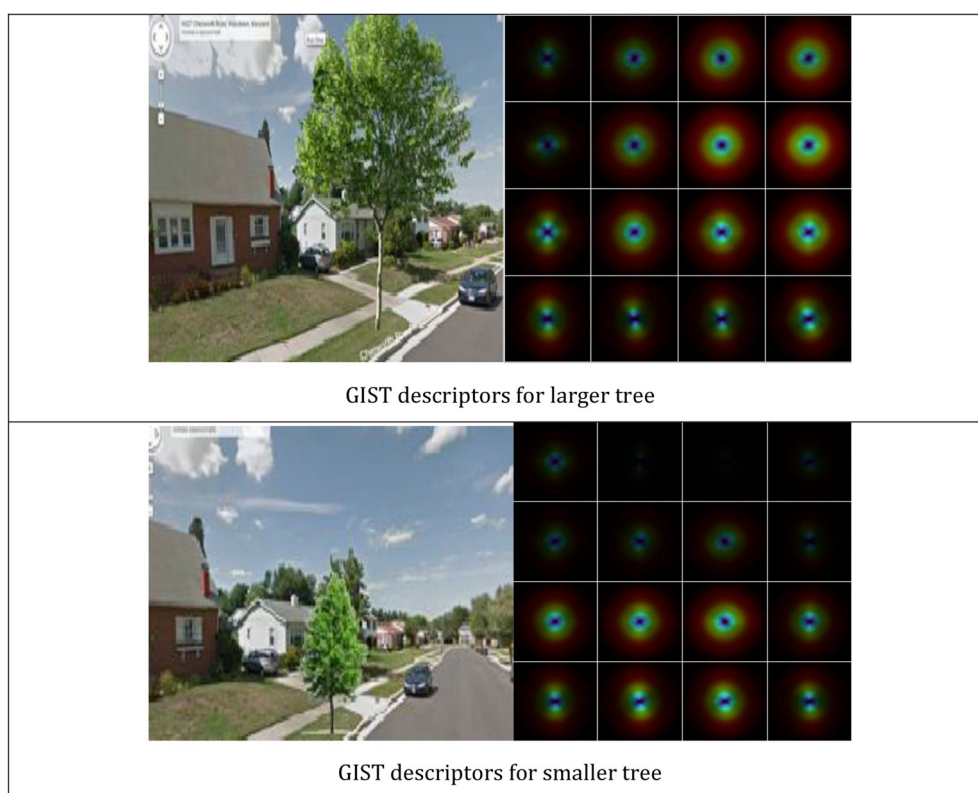


Fig. 15. GIST descriptors for GI setting with a larger and smaller tree.

script is used. Awk script is an interpreted programming language typically used for data extraction in UNIX operating systems (Aho et al., 1978).

Finally, the modified worldfile is used to execute the RHESSYS model and forecast the effects of the GI design on annual mean stream flow and base flow (Ritcher et al., 1996) at the outlet of the urban catchment where the neighborhood is located. These statistics have previously been identified as important indicators for measuring the alteration in hydrology of an urban catchment due to changes in land use and land cover. By simulating the impacts of GI designs on mean flows using RHESSYS and comparing flows with current and pre-development values, we can observe the efficiency of the GI designs in moving towards restoration of pre-development urban watershed hydrology. This provides an initial environmental rating for the GI designs, which can be supplemented with nutrient impacts in the future.

3. Case study: implication for GI design in Dead Run Watershed, Baltimore, MD

The GI design methodology described in Section 2 is applied in the Dead Run Watershed in Baltimore, MD, which is part of the Baltimore Ecosystem Study (BES). As the available human preference training set contains predominantly tree-based green infrastructure, the case study application focuses on the addition of trees to a neighborhood within Dead Run Watershed, shown in Fig. 12, and observing how the human and hydrologic benefits change with different designs. Six scenarios are examined: 1) Existing scenario, 2) Adding a single tree in the neighborhood, 3) Adding multiple trees in open vs. clustered arrangements, 4) Adding small vs. large trees, 5) Adding trees on one vs. both sides of the street, and 6) Adding single vs. mixed species of trees.

Google Street-view images of the entire neighborhood are extracted

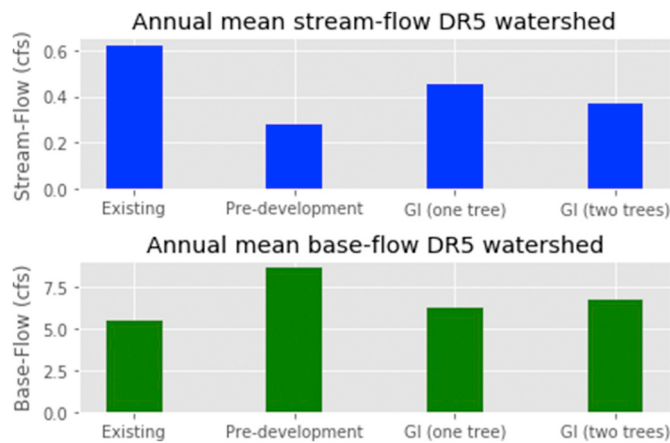


Fig. 16. Hydrologic impacts of GI (trees) on annual mean stream and base flow in Dead Run Watershed.

and used as a base case for human preferences. The coordinate values (latitude & longitude) of the front yard of every house in the region are chosen as potential locations to add trees. Google Street-view overlay API JavaScript is used to overlay trees in the Google image for estimating human preferences using the machine learning model trained and validated in Galesburg (Section 3). The results are given in the following sections.

Once the human preference ratings are obtained, RHESSYS is used to evaluate the hydrologic benefits of the added green infrastructure to the larger Dead Run Watershed using the methodology described in Section 2.2. The hydrologic data used in the simulations were collected by the Baltimore Ecosystem Study, as outlined by Band et al. (2012). The hydrologic impacts of the green infrastructure scenarios are compared with the existing scenario, using 2007 hydrologic and model data from Band et al. (2012), and a pre-development scenario. The earliest hydrologic data available for the pre-development scenario is 1960, which was before the design neighborhood was built. The hydrologic results are presented in Section 3.2, along with a discussion of the interactions between human preferences and hydrologic benefits.

3.1. Results of human preference modeling

The human preference results for the six scenarios in Baltimore are shown below for a sample street view image –

Settings that include trees, plants, bushes, etc. have been shown to reduce symptoms of both mental fatigue and stress (Coley et al., 1997; Sullivan, 2004). The existing scenario (Fig. 13(a)) lacks such a setting, and therefore the human preference model predicts a low rating. Simply by adding a single tree to the existing scenario (Fig. 13(b)), the model predicts a higher preference rating based on shifts in the trained image features for HOG and color histogram.

The spatial arrangement of trees in GI settings is also an important design factor that affects the level of complexity in the scene (Kaplan and Kaplan, 1989), which is captured in the model by the GIST descriptor feature. Highly complex settings are generally less preferable and thus the clustered arrangement of trees in Fig. 13 (d) gets a much

lower human preference rating than the open arrangement of trees in Fig. 13(c). Fig. 14 shows how shifting energy spectrums in the GIST descriptors captures this change.

The legibility and mystery factors of human preference require three-dimensional inference, allowing people to imagine themselves in the scene (Kaplan and Kaplan, 1989). The mystery parameter involves exploration, and therefore the view of angle plays a very important role in human perceptions. The arrangement of trees on both sides of the street in Fig. 13 (d) and (f) create a more three-dimensional view where people can imagine more exploration and thus prefer such settings. The spatial arrangement of trees on both sides of a street is well captured by the GIST feature and color histogram and HOG identify tree shape and color. The features have captured the visual change in GI setting, which encourages human beings to explore the setting.

Further, it is important to note that larger trees are more preferable than smaller trees and therefore have higher ratings (compare Fig. 13(e) vs. (f), as well as 13(c) vs. (g)). The reason for this result is the higher degree of richness and complexity in the setting, which are captured by GIST descriptors. The degree of richness and complexity for a large versus small tree is shown in Fig. 15. The setting with a large tree has fewer dark regions in the GIST matrix than the setting with a small tree.

3.2. Results of hydrologic modeling using RHESSYS

Fig. 16, which shows the hydrologic impact of the GI design scenarios, compares the modeled change in annual mean stream flow and base (groundwater) flow for the following scenarios: 1) Existing condition (for the year 2007), 2), Pre-development conditions in 1960, 3) Adding a single tree to every yard in the neighborhood, and 4) Adding two trees to every yard in the neighborhood. Because hydrologic impacts are measured at the watershed outlet, the physical arrangement of the trees, which are important for the human preference results in Section 3.1, do not change the estimated hydrologic impacts.

As shown in Fig. 16, the tree scenarios significantly reduce the impacts of development on the urban watershed, shifting both the mean stream and base flows much closer to the pre-development scenario. The extent of the shift is particularly impressive given that the neighborhood of the GI design is only 1.2 percent of the land area in the overall Dead Run Watershed. The Dead Run Watershed has 36% of its land as impervious (Gwynn Falls water quality management plan, 2004), and thus adding more trees in front of every house in the neighborhood has significant beneficial impacts. Table 3 shows a quantitative comparison of the two different GI scenarios, as well as a third scenario adding three trees to every yard.

From Table 3, it is clear that green infrastructure can play an important role in restoring the urban watershed to predevelopment hydrology. Adding two trees to every front yard in the neighborhood gives better results than one tree, but the incremental improvement declines as additional trees are added. In fact, adding three trees to every yard does not cause any additional changes to flow at the outlet of watershed. Moreover, the results in Section 1 indicate that the second and third trees in each yard do not cause significant differences in human preferences (compare Fig. 13(b) and (d)), although their arrangements across the neighborhood are important.

Table 3
Quantitative comparison of GI impacts for different scenarios.

GI Design Scenario	Stream Flow Comparison		Base Flow Comparison	
	% reduction from existing scenario	% difference from pre-development scenario	% reduction from existing scenario	% difference from pre-development scenario
One tree (every yard)	26	32	15	23
Two trees (every yard)	36	46	23	17
Three trees (every yard)	36	46	23	17

The hydrologic benefits of adding more green components to urban watersheds helps in restoration of pre-development flows (US EPA, 2009), but our results show that the human preference model can identify specific patterns for installing these green components that are more beneficial for human wellbeing.

4. Conclusions

This work develops a novel computational green infrastructure (GI) design framework that couples storm-water management requirements with criteria for human wellbeing. Current approaches to designing green storm water features tend to emphasize rapid removal of storm water runoff to reduce impacts of downstream flows and pollutant loads. The approach presented in this paper is a first step towards incorporating the benefits of human health associated with these urban green spaces into the design process. In order to map landscape features that are correlated with human wellbeing to features that can be used to train a supervised machine-learning model, a suite of computer vision algorithms and techniques have been used. The result is the first GI design model capable of predicting human preferences, with validation accuracy of 81%, which is quite high given the variability in human perceptions.

Acknowledgement

This material is based upon work supported by the National Science Foundation under Grant No. 1216817 and 1261582. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

- Aho, Alfred V., Kernighan, Brian W., Weinberger, Peter J., 1978. *Awk – A Pattern Scanning and Processing Language*, Second Edition. .
- Anami, B.S., Nandyal, S.S., Govardhan, A., 2010. A combined color, texture and edge features based approach for identification and classification of indian medicinal plants. *Int. J. Comput. Appl. IJCA* 6 (12), 45–51.
- Band, LE, Hwang, T, Hales, TC, et al., 2012. Ecosystem processes at the watershed scale: Mapping and modeling ecohydrological controls of landslides. *Geomorphology* 137, 159–167.
- Bettez, Neil, Groffman, Peter, 2012. Denitrification potential in stormwater control structures and natural riparian zones in an urban landscape. *Environ. Sci. Technol.* 46, 10909–10917. <https://doi.org/10.1021/es301409z>.
- Buhrmester, M., Kwang, T., Gosling, S.D., 2011. Amazon's mechanical Turk. A new source of inexpensive, yet high-quality, data? *J. Assoc. Psychol. Sci.* <https://doi.org/10.1177/1745691610393980>. 2011.
- Chang, C.Y., Chen, P.K., 2005. Human response to window views and indoors plants in the workplace. *Hortscience* 40 (5), 1354–1359.
- Dalal, N., Triggs, B., 2005. Histograms of oriented gradients for human detection. In: *Computer Society Conference on Computer Vision and Pattern Recognition. IEEE*.
- Freund, Y., Schapire, R.E., 1999. A short introduction to boosting. *J. Jpn. Soc. Artif. Intell.* 14 (5), 771–780.
- Glenis, A., Kutija, V., Kilsby, C.G., 2018. A fully hydrodynamic urban flood modelling system representing buildings, green space and interventions. *Environ. Model. Softw.* 109, 272–292.
- Harper, H.H., Baker, D.M., 2008. Evaluation of Current Stormwater Design Criteria within the State of Florida. prepared for Florida Department of Environmental protection Contract FDEP S0108.
- Kaplan, R., Kaplan, S., 1989. *The Experience of Nature: A Psychological Perspective*. Cambridge University Press.
- Lee, Joong Gwang, Riverson, John, 2012. “SUSTAIN: urban modeling systems integrating optimization and economics” in “Fifty Years Of Watershed Modeling - Past, Present And Future”. In: Donigan, A.S., Field, Richard, Baker Jr.Michael (Eds.), *ECI Symposium Series*, . <http://dc.engconfintl.org/watershed/23>.
- LeRoy Poff, N., June 1997. Landscape filters and species traits: towards mechanistic understanding and prediction in stream ecology. *J. North Am. Benthol. Soc.* 16 (2), 391–409.
- Mitchell, R., Popham, F., 2008. Effect of exposure to natural environment on health inequalities: an observational population study. *Lancet* 372, 1655–1660.
- Macro, K., Matott, S., Rabideau, A., Ghodsi, S.H., Zhu, Z., 2018. Ostrich-swmm: a new multi-objective optimization tool for green infrastructure planning with SWMM. *Environ. Model. Softw.* <https://doi.org/10.1016/j.envsoft.2018.12.004>.
- Massoudieh, A., Maghrebi, M., Kamrani, B., Nietch, C., Tryby, M., Aflaki, S., Panguluri, S., 2017. A flexible modeling framework for hydraulic and water quality performance assessment of stormwater green infrastructure. *Environ. Model. Softw.* 92, 57–73.
- NCDWQ, July 2007. Stream Restoration Monitoring Plan Unnamed Tributary to Crab Creek (UTCC). New River Basin Subbasin 05-07-03. Prepared by Division of Water Quality, Watershed Assessment Team.
- NRC, 2008. Committee on Reducing Stormwater Discharge Contributions to Water Pollution, Water Science Technology Board 2008. “Urban Stormwater Management in the United States”. National Academies Press, pp. 513p.
- Oliva, A., Torralba, A., 2001. Modeling the shape of the scene: a holistic representation of the spatial envelope. *Int. J. Comput. Vis.* 42 (3), 145–175.
- Raciti, S.M., Burgin, A.J., Groffman, P.M., Lewis, D.N., Fahey, T.J., 2011. Denitrification in suburban lawn soils. *J. Environ. Qual.* 40, 1932. <https://doi.org/10.2134/jeq2011.0107>.
- Ritcher, B.D., Baugartner, J.V., Powell, J., Braun, D.P., 1996. A method for assessing hydrologic alteration within ecosystems. *Conserv. Biol.* <https://doi.org/10.1046/j.1523-1739.1996.10041163.x>.
- Roy, A.H., Wenger, S.J., Fletcher, T.D., Walsh, C.J., Ladson, A.R., Shuster, W.D., Thurston, H.W., Brown, R.R., 2008. Impediments and solutions to sustainable, watershed-scale urban stormwater management: lessons from Australia and the United States. *Environ. Manag.* 42, 344–359. <https://doi.org/10.1007/s00267-008-9119-1>. [PubMed: 18446406].
- Sullivan, W.C., 2004. Forest, savanna, city: evolutionary landscapes and human functioning. In: In: Barlett, P. (Ed.), *Urban Place: Reconnections with the natural world Vol. 3. Academic Exchange, Atlanta*, pp. 55–63.
- Sullivan, W.C., Anderson, O.M., Lovell, S.T., 2004. Agricultural buffers at the rural-urban fringe: an examination of approval by farmers, residents, and academics in the Midwestern United States. *Landsc. Urban Plann.* 69 (2–3), 299–313.
- Tague, C.L., Band, L.E., 2004. RHESys: regional hydro-ecologic simulation system-an object-oriented approach to spatially distributed modeling of carbon, water, and nutrient cycling. *Earth Interact.* 8, 1–42. [https://doi.org/10.1175/1087-3562\(2004\)8<1:RRHSSO>2.0.CO;2](https://doi.org/10.1175/1087-3562(2004)8<1:RRHSSO>2.0.CO;2).
- Takano, T., Nakamura, K., Watanabe, M., 2002. Urban residential environments and senior citizens' longevity in megacity areas: the importance of walkable green spaces. *J. Epidemiol. Community Health* 56 (12), 913–918. <https://doi.org/10.1136/jech.56.12.913>.
- Thompson, C.W., Roe, J., Aspinall, P., Mitchell, R., Clow, A., Miller, D., 2012. More green space is linked to less stress in deprived communities: evidence from salivary cortisol patterns. *Landsc. Urban Plann.* 105, 221–229.
- Tsihrintzis, Vassilios A., Hamid, Rizwan, 1998. Runoff quality prediction from small urban catchments using SWMM. *Hydrol. Process.* 12 (2), 311.
- Wendel, H.E.W., Downs, J.A., Mihelcic, J.R., 2011. Assessing equitable Access to urban green space: the role of engineered water infrastructure. *Environ. Sci. Technol.* 45, 6728–6734 (Wiley-Interscience).
- Wong, Tony, Fletcher, Tim, Duncan, Hugh P., Coleman, John R., Jenkins, Graham, 2002. Model for Urban Stormwater Improvement Conceptualisation (MUSIC) Version 3. Global Solut. Urban Drain. [https://doi.org/10.1061/40644\(2002\)115](https://doi.org/10.1061/40644(2002)115).