



## How do flow peaks and durations change in suburbanizing semi-arid watersheds? A southern California case study

Robert J. Hawley<sup>a,1</sup>, Brian P. Bledsoe<sup>b,\*</sup>

<sup>a</sup> Colorado State University, Principal Scientist, Sustainable Streams, LLC, 2038 Eastern Parkway, Suite 1, Louisville, KY 40204, USA

<sup>b</sup> Colorado State University, Department of Civil and Environmental Engineering, 1320 Campus Delivery, Fort Collins, CO 80523, USA

### ARTICLE INFO

#### Article history:

Received 22 October 2010

Received in revised form 15 April 2011

Accepted 9 May 2011

Available online 19 May 2011

This manuscript was handled by K. Georgakakos, Editor-in-Chief, with the assistance of Enrique R. Vivoni, Associate Editor

#### Keywords:

Flow duration

Hydrology

Hydromodification

Peak flow

Semi-arid

Urbanization

### SUMMARY

Forty-three US Geological Survey gauges with records greater than ~15 yrs located in watersheds less than ~250 km<sup>2</sup> were used to model the effects of suburbanization on streams in semi-arid southern California. The watersheds spanned a gradient of urban development, ranging 0–23% total impervious area in 2001. With little flow control at the subdivision scale, most impervious area in the region is relatively well-connected to surface-drainage networks and hydrologically effective. Consequently, total impervious area was an effective hydrologic surrogate for urbanization, emerging from an expansive array of geospatially-derived hydrologic variables as a statistically-significant ( $p < 0.05$ ) predictor of instantaneous peak-flow rates at the 1.5- and 2-yr recurrence intervals and the durations of all geomorphically-important flows. To represent the effects of urbanization on flow durations, we developed duration density functions by using power functions (typical  $R^2 > 0.95$ ) to predict occurrence of logarithmically-binned mean daily discharges greater than some nominal value. This approach expands on previous scaling procedures to produce histogram-style cumulative flow durations for ungauged sites using urbanization extent and other watershed descriptors. For a particular watershed size and climatic setting, urbanization resulted in proportionally-longer durations of all geomorphically-effective flows, with a more pronounced effect on the durations of moderate flows. For example, a representative watershed with ~20% imperviousness could experience five times as many days of mean daily flows on the order of 100 cfs (3 m<sup>3</sup>/s) and approximately three times as many days on the order of 1000 cfs (30 m<sup>3</sup>/s) relative to the undeveloped setting. Increased duration of sediment-transporting flows is a primary driver of accelerated changes in channel form that are often concurrent with urbanization throughout southern California, particularly in unconfined, fine-grained geomorphic settings. We did not have comparable studies on flow durations from other regions; however, the peak factors presented herein (e.g., sixfold increase in  $Q_2$  at 20% imperviousness) are greater than studies from humid temperate regions suggesting that semi-arid regimes may be more susceptible to urbanization than other climatic settings.

© 2011 Elsevier B.V. All rights reserved.

### 1. Introduction

By decreasing infiltration and increasing direct runoff, impervious surfaces can create larger peaks, less groundwater recharge, and increased variability, especially if stormwater is routed directly to streams. These fundamental hydrologic interrelations, such as larger peaks and increased flashiness, have been demonstrated regionally (Beighley and Moglen, 2002; Konrad and Booth, 2002; Smith et al., 2002) and on a national scale (Poff et al., 2006; Sauer et al., 1983) using US Geological Survey (USGS) gauge data. In California, much research has attributed higher peak flows to

urbanization (Durbin, 1974; Rantz, 1971; Sheng and Wilson, 2009; Waananen, 1969; White and Greer, 2006) as quantified by surrogate measures such as development extent and population density.

Such changes in flow, broadly associated with urbanization, are documented as having profound effects on biologic and geomorphic processes, so much so that the US Environmental Protection Agency (EPA) has recently begun to mandate ‘hydromodification’ regulations (EPA, 2006). Channel instability and complex responses have been associated with urbanization across hydro-climatic regimes (Bledsoe and Watson, 2001; Booth, 1990; Chin, 2006; Chin and Gregory, 2001; Simon and Downs, 1995; Trimble, 1997), while altered flow and sediment regimes affect aquatic life cycles, habitats, food webs, and facilitate colonization by invasive species, among other types of degradation (Poff et al., 2006; Roesner and Bledsoe, 2002; Waters, 1995).

\* Corresponding author. Tel.: +1 970 491 8410; fax: +1 970 491 8671.

E-mail addresses: [bob.hawley@sustainablestreams.com](mailto:bob.hawley@sustainablestreams.com) (R.J. Hawley), [Brian.Bledsoe@colostate.edu](mailto:Brian.Bledsoe@colostate.edu) (B.P. Bledsoe).

<sup>1</sup> Tel.: +1 502 718 2912.

Semi-arid systems are reported to have an increased sensitivity to urbanization in terms of their geomorphic response potential (Trimble, 1997). The hydrogeomorphic setting in southern California (i.e., steep topography, flashy regimes, high-sediment loads, and largely nonresistant bed material) compounds risk factors for channel responses such as headcutting, mass-wasting, and planform shifts (Hawley, 2009). This paper is a critical first step in a broader study focused on understanding the physical effects of hydromodification in southern California, i.e., how urbanization has affected the flow regimes. With a focus on relatively-small unregulated streams, the ensuing investigation had the following objectives:

1. offer an updated alternative to the USGS (Waananen and Crippen, 1977) regional equations for instantaneous peak flows;
2. develop an empirical method for estimating long-term cumulative duration histograms for ungauged sites; and
3. determine how urbanization affects peak flows and cumulative durations for all geomorphically-important flows by including urban components (if statistically significant) in Objectives 1 and 2.

In filling these knowledge gaps, we offer the following hypotheses:

$H_0$ : urban influence on the magnitudes of peak flows and the durations of all flows will be higher at the more frequent/moderate flows and lower at rare/largest flows; and

$H_0$ : the lack of semi-arid gauges used to develop the USGS national urban equation (Sauer et al., 1983) should result in better performance by models calibrated directly to the semi-arid southern California region.

### 1.1. Research foundations and justification

This paper builds on the work and ongoing data collection of the USGS. As a first step in understanding the effects of urbanization on the entire flow regime, we revisited Waananen and Crippen's (1977) simple power functions of drainage area and mean annual precipitation, which to this day serve as a primary method of peak-flow estimation in southern California. Limited by an overall lack of data on "streams with drainage areas generally less than 25 mi<sup>2</sup>, and particularly less than 10 mi<sup>2</sup>," the models came with substantial standard errors and were deemed "generally applicable for streams with drainage areas greater than 10 mi<sup>2</sup>" (Waananen and Crippen, 1977). As a part of this study focused on the entire flow regime, we revisit the peak-flow equations to incorporate additional years of data, and especially more data on smaller streams. In this paper, we go beyond the Log-Pearson Type III distribution to seek a more regionally-appropriate statistical distribution. With several gauges in developed watersheds, urbanization was included in the models using direct measures of total impervious area (TIA). This approach is arguably less subjective and more parsimonious than the USGS national approach to urban flow augmentation (Sauer et al., 1983), which can be time intensive and is subject to user interpretation of "basin development factors" that are typically immeasurable with widely available Geographic Information System (GIS) data. Moreover, of the 199 gauges used to develop the USGS national equations, few gauges were from semi-arid settings, with only one from southern California (San Diego Creek, gauge no. 11048500).

### 1.2. Toward cumulative durations

Peak flows alone can be useful in understanding potential erosive energy at an individual recurrence interval; however, they

have less meaning when considered independent of durations. Whether an erosive flow lasts for minutes or days has substantial implications for sediment transport. Researchers have begun to favor cumulative sediment-transport models based on continuous or cumulative flows over extended periods (e.g., years/decades). One of the only published approaches to addressing geomorphic aspects of hydromodification in California to date uses flow-duration histograms produced from 50-yr rainfall-runoff simulations in Hydrologic Engineering Center – Hydrologic Modeling System (Santa Clara, 2004).

An alternative to solely using rainfall-runoff models to develop flow-frequency curves is to base them on local gauge data. Using the nearest upstream/downstream gauge (Hey, 1975) or a gauge from a similar watershed, frequency curves have typically been scaled using a nondimensional index such as  $Q/Q_{\text{bankfull}}$  (Emmett, 1975; Leopold, 1994) or  $Q/Q_2$  (Watson et al., 1997). We expand on the previous approaches by developing a statistical model to estimate synthetic flow-duration histograms as a function of watershed-scale physical descriptors such as drainage area and precipitation. The resulting *conditional* Probability Density Functions that predict cumulative durations of geomorphically-effective flows in a histogram format are henceforth referred to as Duration Density Functions (DDFs). The logarithmically-distributed histogram bins are represented by power functions (i.e., #days = coef \*  $Q^{\text{EXP}}$ ) and scaled by the maximum daily flow of record. Given a way to predict the shape (exponent), magnitude (coefficient), and scale ( $Q_{\text{max}}$ ) based on physical parameters, one could estimate long-term durations of sediment-transporting flows for any ungauged watershed. More importantly regarding hydromodification, DDFs could simulate the increases in durations of sediment-transporting flows associated with unmitigated urbanization by including a statistically-significant surrogate measure (e.g., TIA) in the model. In this light, DDFs can become a central tool in understanding, modeling, and mitigating the effects of hydromodification in southern California and offer a framework for other regions.

### 1.3. Study domain

Southern California is generally described in this study as the greater Los Angeles/San Diego area within about 100 km of the Pacific coast, including portions of Ventura, Los Angeles, San Bernardino, Orange, Riverside, and San Diego Counties and ca. 20–25 million residents. Mountain ranges to the north (Transverse Ranges) and east (Peninsular Ranges) offer fairly well-defined geologic bounds, with a total relief of up to 3500 m. The climate is broadly characterized as Mediterranean, but precipitation and vegetative influences tend to increase with elevation, although there are clear differences between the west and east slopes of the Peninsular Ranges due to the associated orographic lifting. Precipitation generally falls in the form of winter frontal storms, which can be intense (i.e., the 2-yr 24-h rainfall ranges ~2–6 in. (50–160 mm) across the domain). This leads to a flashy regime with short-lived instantaneous peak flows that are much larger than the corresponding mean daily flows. For example, a 10-yr instantaneous event would typically attenuate to a mean daily flow on the order of a 2- to 3-yr event, with the former likely ten to 20 times the latter.

During field investigations of recently-developed suburban neighborhoods, we saw little evidence of stormwater retention/detention. Developed watersheds often had lined channels with occasional energy dissipaters at stormwater outfalls. Large regional basins and dammed reservoirs do exist; however, flow controls in watersheds less than ~100 km<sup>2</sup> were largely lacking. With the understanding that unmitigated urbanization typically increases flow variability, and that streams in southern California are

inherently flashy, we hypothesize that the effects of urbanization may be especially pronounced in this context.

## 2. Methods

Gauge data are made publicly available by the USGS, which adheres to strict quality assurance procedures prior to publishing flows as approved. To ensure comparable quality in processing and analysis, we developed the following methods.

### 2.1. Unit disclaimer

Acknowledging the general preference of International System (SI) units among the scientific community, we developed these models in the US Customary System units for more direct comparisons to the USGS models. We offer SI units in introductory tables and figures for ease of comparison to studies from other regions.

### 2.2. Gauge-selection criteria

Our first step was the systematic selection of regional gauges for model development. We focused on watersheds less than ca. 250 km<sup>2</sup>, primarily because most of the region's larger streams have been affected by dams and diversions and our goal was to isolate the effects of urbanization relative to the undeveloped, free-flowing setting. We strove for a balance between a large representation of watersheds and gauges with sufficiently long records for our analyses. A relatively wide data gap segregated the gauges with longer records (i.e., greater than ca. 15 yrs) from those with shorter records (i.e., less than ca. 8 yrs). This made a minimum record length of ca. 15 yrs a logical selection criterion and resulted in 52 gauges with record lengths normally distributed about a mean of ca. 45 yrs and a broad spatial distribution (Fig. 1). A summary of selected gradients such as drainage area and mean annual precipitation is provided in Table 1. These gradients also serve as bounds to the applicable ranges of our models when considering application in southern California or other semi-arid regions.

The gauges had relatively-normal distributions of variables such as precipitation and surface slope, although drainage area and density showed a small positive skew. Imperviousness,

however, had a highly-positive skew of 2.2. As of 2001, only fifteen gauges had watersheds with more than 1% TIA, with only six greater than 10% imperviousness.

Another notable spatial trend was that eight gauges located in the eastern-most portion of the domain and one in the far south-east at the Mexican border ('dry' subset Fig. 1) lie in what is effectively a rain shadow. Stratified by USGS 8-digit Hydrologic Unit Codes (HUCs) of 18100200 or 18070305, the so-called 'dry' gauges were subject to less mean-annual precipitation as well as different types of events (i.e., local convective thunderstorms in addition to winter frontal storms). Hawley (2009) demonstrated significantly-different hydrologic behavior in the 'dry' subset, justifying their exclusion to develop more targeted models for the 43 watersheds examined in this paper.

### 2.3. Instantaneous peak flows

Procedures were developed to populate recurrence-interval flows for the 1-, 1.5-, 2-, 5-, 10-, 25-, 50-, and 100-yr events from annual peak-flow data as recorded by the USGS. Flows were proportionally ranked to determine recurrence probabilities via the Weibull plotting position (Chow, 1964; Yevjevich, 1972). Several probability distributions were tested using the method of moments to represent the flow-frequency relationship at each gauge, including the normal, lognormal (LN), exponential, gamma, and the Log-Pearson Type III (LP3), a log-transformed three-parameter gamma distribution that has been the standard USGS flow-frequency method since 1967 (US Water Resources Council, 1967).

Despite application in previous studies, the LP3 performed relatively poorly due to the flashy regimes and the corresponding effect on the skew factor. In contrast, the inverse gamma was superior in every case in terms of homoscedasticity of residuals and  $R^2$  (e.g., mean and median  $R^2$  of 0.95 and 0.97, respectively, with only three cases <0.90). Bounded by zero by definition, the gamma function is ideal for modeling skewed distributions without the need for a log transformation (Chow et al., 1988). Befitting for the flashy ephemeral regimes of southern California, gamma-distribution flows were used to develop models for recurrence intervals greater than 2 yrs, while the Weibull plotting position was used for the 1-, 1.5-, and 2-yr events.

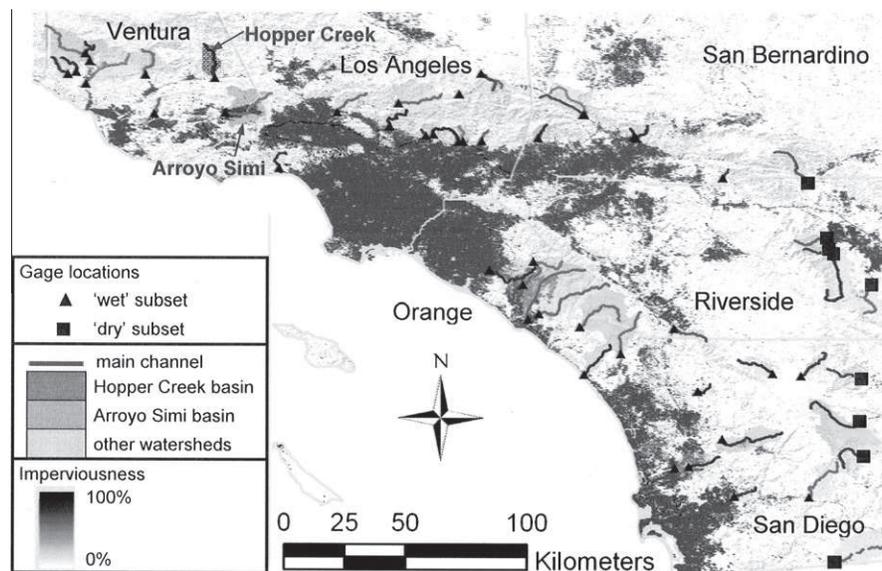


Fig. 1. Locations of gauges used in equation development ('wet' subset) with corresponding watershed and main channel, overlaid by a gradient of imperviousness and county boundaries, with rural (Hopper) and urban (Arroyo Simi) case study gauges (Section 4.2).



50	SANVELLPECNIBJULIAN	10255700	33.1186	-116.4344	18100200	1958	1983	25	1.08	0.76	0.4	0.3	0.3	445	65	230.8	1.46	864	1.8%	24%	4.8%
51	TAHQUTZCNRPALM SPRINGS	10258000	33.8050	-116.5583	18100200	1947	2008	59	0.00	0.00	0.0	0.0	0.0	562	92	44.1	1.43	1563	15.0%	41%	6.4%
52	VALLECTOCNIBJULIAN	10253850	32.9861	-116.4194	18100200	1963	1983	20	0.58	0.38	0.2	0.2	0.2	400	67	102.4	1.58	988	5.0%	30%	2.2%
Gradients of the 43 gauged watersheds used in model development (and model application bounds)					min	1905	1962	14	0.14	0.01	0.0	0.0	0.0	353	53	1.4	0.96	134	1.2%	8%	0.1%
					max	1940	1989	44	1.78	1.06	3.3	2.7	1.8	633	103	71.1	1.41	745	6.2%	34%	3.3%
						1987	2008	94	6.80	3.95	23.4	20.1	14.2	908	159	273.9	2.56	2228	20.2%	59%	21.1%

<sup>a</sup> Table includes all USGS gauges in the study domain with watersheds less than ~250 km<sup>2</sup>, flow records greater than ~15 yrs, and no upstream dams/diversions.  
<sup>b</sup> Gaps in US Department of Agriculture (USDA) geospatial soil coverages precluded the inclusion of soil characteristics in the analysis; however, a representative sample of regional watersheds ranged from 100% Natural Resources Conservation Service (NRCS) Type D to 100% NRCS Type B and up to 10% Type A soils with undeveloped NRCS Curve Numbers that ranged 77–88 with a mean of 83.4.  
<sup>c</sup> The nine gauges in Hydrologic Unit Code (HUC) 18100200 and 18070305 were excluded from model development due to their significantly ( $p < 0.05$ ) different hydrogeomorphic setting on the east slope of the Peninsular Range.  
<sup>d</sup> Total years of annual maximum instantaneous peak records as recorded and made available by the USGS (i.e., not necessarily equal to “End” minus “Begin” due to intermittent records at several gauges).  
<sup>e</sup> Average and maximum road density and impervious values based on integration of spatial extent over the gauge records using three to four measures of spatial extent in time, delineated from historical USGS quadrangle maps and contemporary geospatial coverages from USGS and CalAtlas in a GIS.  
<sup>f</sup> Mean annual precipitation integrated over the watershed using USGS shapefile developed using regional precipitation data from 1900 to 1960.  
<sup>g</sup> Total precipitation volume over 24-h duration with a probability of occurrence once every 2 yrs, spatially integrated over the watershed using regional precipitation data from 1961 to 1990.  
<sup>h</sup> Contributing watershed area delineated in a GIS using the USGS HUC boundaries and a 10-m National Elevation Dataset (NED).  
<sup>i</sup> Drainage density developed using total stream length in basin as delineated in a GIS using the USGS National Hydrography Dataset (NHD) developed at a 1:24,000 scale.  
<sup>j</sup> Average basin elevation and channel slope measured after USGS protocol using points at 10% and 85% of the main-channel distance from gauge to basin divide.  
<sup>k</sup> Average surface slope of the entire watershed using clipped NED model from USGS.  
<sup>l</sup> Representative valley slope over reach at gauge location measured across a geomorphically-continuous valley up to 10% of the main-channel length or 500 m.

2.4. Long-term cumulative durations

Next, procedures were developed to represent all mean daily flows with cumulative duration curves. Mean daily flows were binned via a histogram procedure analogous to the initial steps of an effective-discharge calculation after Biedenharn et al. (2000, 2001). Histogram bins were scaled by the maximum mean daily flow on record ( $Q_{max}$ ). The extreme flashiness of ephemeral streams in southern California made logarithmic bins the only practical way to represent flow frequency without discontinuities. The following equation was used to size logarithmically-equivalent bins after Raff et al. (2004):

$$H_{B-log} = \{\ln(Q_{max}) - \ln(Q_{min})\} / (N_B - 1) \tag{1}$$

where  $H_{B-log}$  is the bin size of logarithmically-spaced histogram bins;  $Q_{max}$  the maximum flow of record;  $Q_{min}$  the minimum flow of record; and  $N_B$  is the number of bins.

For consistency across all gauges toward development of a regional equation, we set  $Q_{min}$  equal to 0.01 cfs at all sites, the lowest non-zero mean daily flow reported by any gauge. Lower and upper bounds of each logarithmically-spaced bin were determined using the following equations after Raff et al. (2004):

$$B_{lwr-log} = e^{\ln(Q_{min}) + (B-2) * H_{B-log}} \tag{2}$$

$$B_{upr-log} = e^{\ln(Q_{min}) + (B-1) * H_{B-log}} \tag{3}$$

where  $B_{lwr-log}$  is the lower logarithmically-spaced bound of bin number ( $B$ );  $B_{upr-log}$  the upper logarithmically-spaced bound of bin number ( $B$ ); and  $B$  is the bin number (i.e., 1 to  $N_B$ , where  $N_B$  = total number of bins).

Setting  $N_B$  equal to 25 provided a reasonable balance of resolution (small bin sizes) and continuous frequency distributions, with all but four gauges having daily-flow records sufficient to populate 25 bins. This resulted in 39 gauges that could be included in the DDF models.

To represent the histograms in a concise, transferable format, the next step was to convert them into conditional Probability Density Functions by fitting power functions to the centroids of the bins representing the geomorphically-effective flows. We fit the DDFs to the arithmetic-bin centroids, as opposed to the geometric centroids, because sediment transport increases non-linearly with flow, and such a scheme would better approximate the composite transport of the individual flows within the bin.

The distributions were sufficiently continuous over bins 12–25, such that they could be well-represented with simple power functions (i.e.,  $R^2$  typically >0.95). Those bins also coincided with the range of flows that were important for sediment transport based on a Shields parameter mobility criterion ( $\tau_{*c}$ ) of 0.047 (Hawley, 2009). Fig. 2 offers an example of a DDF fit to bins 12–25 at Arroyo Trabuco (gauge no. 11047300).

The general form of the power function used in the DDF scheme discussed above is:

$$\text{days} = \text{day1} * Q^{\text{day2}} \tag{4}$$

where days is the number of days of occurrence at flow rate ( $Q$ );  $Q$  the arithmetic average of mean daily flows corresponding to the lower- and upper-bin boundaries defined by Eqs. (2) and (3), respectively (cfs); day1 the coefficient for power function fit to bins 12–25; and day2 the exponent for power function fit to bins 12–25.

2.5. Measures of urbanization

We had several goals regarding the quantification of urbanization for our models. Despite being an empirical approach, assurance of fidelity to hydrologic processes was desired. Additionally,

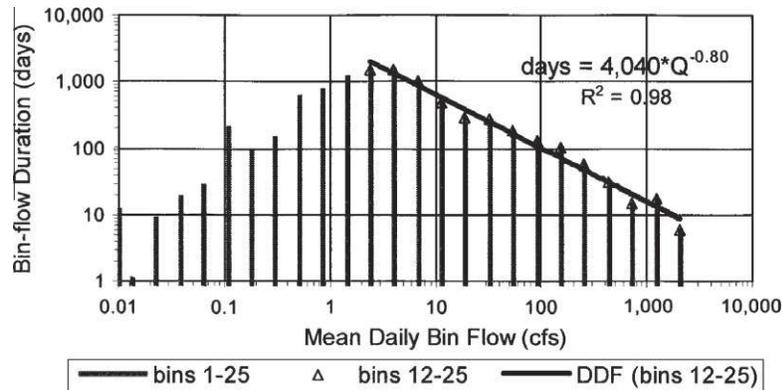


Fig. 2. DDF of Arroyo Trabuco (gauge no. 11047300) fitted to centroids of logarithmically-distributed histogram bins 12–25 over a composite record of 26 yrs of mean daily flows.

measures should be readily quantifiable via publically-available GIS data and be in the form of continuous metrics wherever possible. Finally, because urbanization is not constant through time, we needed to be able to measure the variability in spatial extent over the gauge records.

Arguably, the measure of urbanization that is most rooted in theory and most important hydrologically is imperviousness (Novotny, 2003), but it is whether an impervious surface is *connected* to the drainage network that determines if the potential effects are transferred downstream (i.e., Effective Impervious Area (EIA); *sensu* Booth and Jackson (1997)). Because stormwater at the subdivision scale in southern California has largely gone unmitigated to date, TIA is much more representative of EIA than in many other regions. Additionally, TIA is readily quantifiable in GIS via the USGS national impervious raster from 2001. Road density has also been used as a surrogate for EIA (Konrad and Booth, 2002) and is objectively quantifiable via State of California road vectors from 2000 and 2007.

To track the change of spatial extent through time, contemporary coverages of road density were clipped to match georeferenced historical USGS topographic quadrangle maps, providing two additional measurements in time (typically from the 1950s to 1980s). Going backwards through time, as the roads that accessed contemporary subdivisions were tracked out of existence, we removed the associated impervious areas that accompanied each suburban development. As a part of a broader study, we had access to historical aerial photography at several locations that confirmed the accuracy of the USGS quadrangle maps and served as a validation step to our approach. An example at one of the most urban gauges, Arroyo Trabuco, is presented in Fig. 3, along with 2001 impervious levels.

Acknowledging the uncertainty in modeling summaries of time-series hydrologic data (i.e., recurrence interval peak flows and cumulative flow durations) using time-integrated (average) values of non-stationary drivers such as urbanization, we quantified several time-integrated measures of both road density and TIA to test in the models (described in Section 2.7).

## 2.6. Other hydrogeomorphic metrics

We populated a matrix of over 50 geospatially-derived hydrogeomorphic metrics across varying temporal and spatial scales to test the influence of a multitude of factors, including measures tested by the USGS (e.g., average channel slope, average basin elevation, etc.) and other process-based metrics (e.g., average surface slope, drainage density, etc.). The geospatial data (“Internet references – geospatial data sources”) were acquired from public-domain sources such as the USGS and USDA. Empty fields in some

USDA polygons precluded a complete analysis of NRCS soil types; however, most source data were complete. General resolution of these data was such that their precision was typically on the order of 1% of the measurement (e.g., 10-m NED over 1 km of channel).

ArcMap software by Environmental Systems Research Institute (ESRI), including extensions such as ‘spatial analyst’, was used to optimize GIS measurements such as delineating watersheds and flow paths. Automated results from NED processing were cross-checked with existing shapefiles such as USGS HUC boundaries and National Hydrography Dataset (NHD) flowlines to verify estimates of drainage area, drainage density, etc.

We also accounted for the inter-annual, decadal, and multi-decadal trends in regional precipitation as recorded at the two long-term precipitation gauges in Los Angeles and San Diego. The data suggest that the more urban period of record (post ~1970) potentially had larger volumes of precipitation than the pre-urban period. To test the influence of this temporal variability, composite climate records were compiled for the active years of each gauge.

## 2.7. Analytical methods and model design

Beyond representing physical processes with appropriate, quantitative variables, it was also important to guide their combination in model design to obviate potential collinearity issues. A cross-validation step was performed prior to final model design using a 33/10 calibration/validation split. Multivariate power functions via regression analysis have been widely used by the USGS in developing regional equations for recurrence-interval flows (Jennings et al., 1994), such that our analyses continue in this tradition. We used Statistical Analysis Software (SAS) to perform ordinary least squares regression. Hundreds of iterations of models were run with diverse withholding schemes using forward, backward, and best subset selection to determine the most consistently-significant individual variables and combinations thereof to determine candidate models for final testing. Due to a non-normal distribution of some variables, many variables were tested in both log-transformed and arithmetic forms. To preclude collinear variables from attempting to represent the same process within the same model, we populated candidate models with up to one variable from distinct hydrogeomorphic categories. The process-based categories included:

- *watershed/network size*: drainage area ( $A$ ) or total stream length ( $Stm$ );
- *spatial efficiency*: shape ( $Shp$ ) or drainage density ( $DD$ );
- *precipitation*: mean annual ( $P$ ), 2-yr 24-h volume ( $P_{224}$ ), or 2-yr 24-h relative to mean annual ( $IP$ );

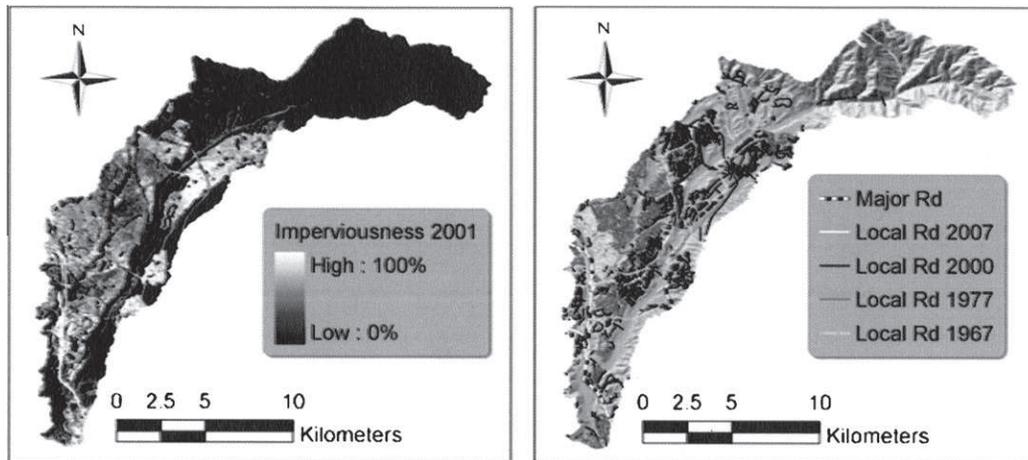


Fig. 3. 2001 imperviousness and road vectors tracked through time per USGS historic quadrangle maps at Arroyo Trabuco (Orange County, California, near intersection of Interstates 5 and 405).

- *topographic efficiency*: average slope of watershed surface ( $S_{rf}$ ), average channel slope ( $S_{chn}$ ), valley slope at site ( $Vly$ ), total relief along the main channel ( $Rlf$ ), and average basin elevation ( $Elv$ );
- *probabilistic category*: number of years of gauge record ( $Yr$ ), the relative difference from long-term mean annual precipitation recorded at Los Angeles during gauged years ( $LA_{hst}$ ), and the number of active gauge years that were exceptionally 'wet', that is, 50% greater than the mean ( $LA_{wt}$ );
- *urbanization extent*: time-integrated measures of TIA and road density, including average extent over record ( $Imp_{av}$  and  $Rd_{av}$ ), maximum extent of record ( $Imp_{max}$  and  $Rd_{max}$ ), fraction of record length greater than 5%, 7.5%, and 10% TIA ( $Imp_5$ ,  $Imp_7$ , and  $Imp_{10}$ ), and fraction of record length greater than 2, 4, and 6  $mi/mi^2$  road density ( $Rd_2$ ,  $Rd_4$ , and  $Rd_6$ , respectively).

Model forms that were congruent with hydrologic theory and had high performance in the cross-validation phase were selected for final model calibration. Model performance was measured using several indicators including (1) having individually-significant variables at the  $p < 0.05$  level, (2) high Adjusted  $R^2$  and/or minimum corrected Akaike Information Criterion ( $AIC_c$ ; Akaike, 1974; Sugiura, 1978), and (3) homoscedastic residuals across both calibration and validation samples. We assessed model performance, including standard diagnostics, in both geometric and arithmetic space. Outliers were identified using standard diagnostics such as Cook's D or RStudent residual; however, to be withheld from the model there needed to be supporting *a priori* evidence and/or compelling hydrologic justification (e.g., the climatically-distinct 'dry' subset of gauges east of the Peninsular Range discussed above). In general, we attempted to follow the guideline of ca. 10 observations per predictor variable.

### 3. Results

In both the cross-validation and final calibration of predictive models, urbanization was statistically significant ( $p < 0.05$ ) in predicting peak flows with return periods less than ca. 5 yrs and durations of all sediment-transporting flows. For example, after watershed size and precipitation, TIA was typically the third most powerful variable in predicting peak flows and accounted for up to 10% of the variance for DDF magnitude (i.e., the number of days). These results are summarized in the three subsections that follow: (1) cross-validation summary, (2) peak-flow equations, and (3) DDF models.

#### 3.1. Cross-validation summaries and individual variable performance

Cross-validation models of  $Q_i$  ( $n_{calibration} = 33$ ,  $n_{validation} = 10$ ) and DDF ( $n_{calibration} = 30$ ,  $n_{validation} = 9$ ) components are summarized in Tables 2a and b, respectively. Measures of watershed size ( $Stm$ ,  $A$ ) and precipitation ( $P$ ,  $P_{224}$ ) accounted for the most variance across all return intervals. Measures of imperviousness accounted for up to 10% of the variance of the 1.5-yr flow, with decreasing significance for higher flows (e.g., partial  $R^2$  of 0.06 and 0.02 for 2- and 5-yr flows, respectively). At higher return intervals (i.e.,  $\geq Q_{10}$ ), the size of the watershed accounted for so much of the variance that high performance resulted from relatively-simple models. For example, for return intervals 10, 25, 50, and 100 yrs,  $R^2$  in arithmetic space ranged from 0.7 to 0.9 for both calibration and validation sub-samples using two-variable functions of ( $A$ ,  $P$ ), ( $Stm$ ,  $P$ ), and ( $Stm$ ,  $P_{224}$ ) (Fig. 4).

Recall that DDFs have three components:  $Q_{max}$  (scale), day1 (magnitude), and day2 (shape). Watershed size and precipitation generally explained most of the variance of  $Q_{max}$  (partial  $R^2$  of 0.6 and 0.2, respectively). Record length ( $Yr$ ) was the next most significant variable in forward selection, explaining 3–4% of the variance, and the most significant measure of network spatial efficiency was DD (2–3% of the variance). Urbanization was insignificant in predicting the maximum mean daily flow on record, consistent with the models of the rarest and largest peak flows (i.e.,  $\geq Q_{25}$ ).

Forward selection for predicting DDF magnitude typically identified the following form, with corresponding partial  $R^2$  in parentheses:

$$\text{day1} = f(Yrs(0.46), A(0.07 - 0.09), P(0.18), \\ Imp_x(0.10 - 0.11), S_{chn}(0.02 - 0.03))$$

One of three similarly performing impervious descriptors (i.e.,  $Imp_x$  representing  $Imp_{av}$ ,  $Imp_5$ , or  $Imp_7$ ) was typically the third variable added during forward selection. Exponential forms of the impervious terms consistently explained more variance than the power form. Models of day1 with  $S_{chn}$  had improved calibration accuracy but reduced validation performance compared to the base model (i.e.,  $A$ ,  $P$ ,  $Yrs$ , and  $Imp_x$ ).

The shape of the DDFs (day2) was highly influenced by its magnitude ( $Q_{max}$ ) and scale (day1), explaining 51 and 26% of the variance in day2, respectively.  $Q_{10}$  could explain slightly more variance (31%) when used in place of  $Q_{max}$ . Models that intentionally withheld such measures were not only poorly fit but had severely patterned residuals. Variables that competed for the third and

**Table 2**  
Summary of cross-validated models.

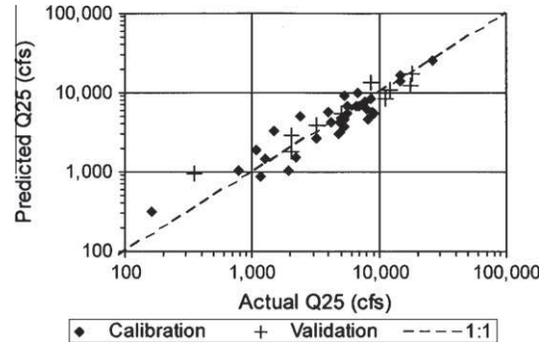
Dependent variable	Urbanization significant ( $p < 0.05$ ) in validated model?	Best predictor variables <sup>a, b</sup>	Average calibration standard error <sup>c</sup> (%)	Average validation standard error <sup>c</sup> (%)
<i>(a) For instantaneous peak flows (<math>n_{calibration} = 33</math>; <math>n_{validation} = 10</math>)</i>				
$Q_{1.5}$	↘	A, Strm (0.2); $P$ ; $P_{224}$ ; IP; (0.2); $Imp_{max}$ (0.1)	80	100
$Q_2$	↘	A, Strm (0.4); $P$ ; $P_{224}$ ; IP (0.1); $Imp_{max}$ (0.06)	80	80
$Q_5$	↘	A, Strm (0.7); $P$ ; $P_{224}$ ; IP (0.1); $Imp_{max}$ (0.02)	60	70
$Q_{10}$	$p = 0.12$	A, Strm (0.8); $P$ ; $P_{224}$ ; IP (0.05)	40	50
$Q_{2.5}$		A, Strm, (0.8); $P$ ; $P_{224}$ ; IP (0.07)	30	50
$Q_{5.0}$		A, Strm (0.8); $P$ ; $P_{224}$ ; (0.08)	30	50
$Q_{10.0}$		A, Strm (0.7); $P$ ; $P_{224}$ (0.1)	40	60
<i>(b) For DDFs</i>				
$Q_{max}$		Urbanization significant ( $p < 0.05$ ) in validated model?	Average calibration $R^{2d}$	Average validation $R^{2d}$
day1	↘		0.8	0.8
day2	↘		0.7	0.7
		Best predictor variables <sup>a, b</sup>	0.9	0.8
		A, Strm (0.6); $P$ ; $P_{224}$ (0.2)		
		A (0.1); $P$ (0.2); Yr (0.5); $Imp_x$ (0.1)		
		$Q_{10}$ (0.3); day1 (0.5)		

<sup>a</sup> Corresponding partial  $R^2$  in parentheses.

<sup>b</sup> Variables defined as: A = drainage area; Strm = total stream length;  $P$  = mean annual precipitation;  $P_{224}$  = 2-yr 24-h precipitation; IP =  $P_{224}/P$ ;  $Imp_{max}$  = maximum extent during gage record of TIA as fraction of total drainage area; Yr = number of years of gage record;  $Imp_x$  represents three similarly performing impervious descriptors: (1)  $Imp_{av}$  = average extent during gage record of TIA as fraction of total drainage area, (2)  $Imp_5$  = fraction of gage record length greater than 5% TIA, and (3)  $Imp_7$  = fraction of gage record length greater than 7.5% TIA;  $Q_{10}$  = 10-yr instantaneous peak flow; and day1 = coefficient of DDF model.

<sup>c</sup> Standard error of estimate reported from arithmetic space as a percentage of the sample mean.

<sup>d</sup>  $R^2$  reported from arithmetic space.



**Fig. 4.** Cross-validation performance of  $Q_i = f(\text{Stm}, P_{224})$  for 25-yr return interval (predicted  $Q_{25}$  versus actual) with superimposed 1:1 'perfect-fit' line.

fourth positions in the day2 models were Yr,  $Imp_x$ , and  $P_x$ . In summary, for each dependent variable, cross-validation produced five to six reasonably performing candidate models that were advanced to final calibration. In the sections below, we present the models that represent the median influence of urbanization for peak flows and DDFs, respectively.

3.2. Peak-flow equations

Of the five final ( $n = 43$ ) models that were calibrated to predict peak flows, the model that most consistently represented the median influence of urbanization across all return periods was a variation of the 1977 USGS model (function of A and P). The added exponential term of  $Imp_{max}$  is convenient because it models the effects of urbanization in a simple continuous form, becoming a rural equation when impervious cover is zero. Although it had slightly poorer  $R^2$  values during final calibration relative to other models (especially for the more frequent return periods), it was the best-overall performing model during the cross-validation step, suggesting that the standard errors may be more representative of what one could expect during application. The model is presented below with corresponding parameters for specific return intervals in Table 3a. In these equations, uppercase terms indicate variables and lowercase nomenclature indicates the corresponding  $\beta$  parameter from the regression.

$$Q_i = e^{(Incpt)} * A^a * P^p * e^{(impmax - Imp_{max})} \tag{5}$$

where  $Q_i$  is the instantaneous peak flow at return interval  $i$  yrs (cfs);  $Incpt$  the vertical axis intercept of the log-transformed linear regression model; A is the total contributing drainage area ( $mi^2$ ); P the average annual precipitation, USGS: 1900–1960 (in.); and  $Imp_{max}$  the maximum spatial extent of TIA during gage record expressed as a fraction of total drainage area ( $mi^2/mi^2$ ).

Model performance generally increased up to  $Q_{10}$ , with relatively consistent precision at higher return intervals. Performance of our peak-flow model relative to the USGS rural (1977) and urban (1983) equations is depicted in Figs. 5 and 6, respectively. The disparity generally decreases with increasing return period (Table 4). Given the longer record lengths, the focus on smaller watersheds, and the exclusion of the 'dry' subset, our models outperform the USGS models in every case in terms of Adjusted  $R^2$ , Sum of Squared Errors, Standard Error of Estimate, etc.

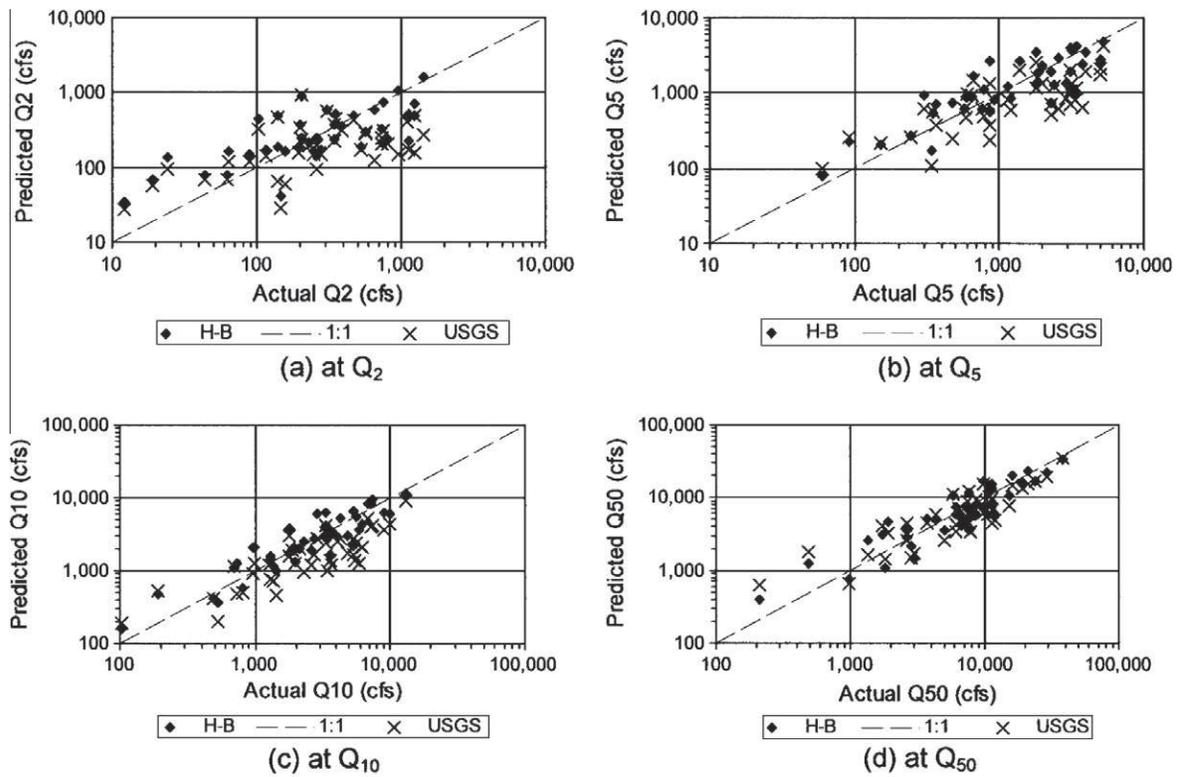
3.3. Duration density functions

DDFs use a power function (Eq. (4)) to predict durations of logarithmically binned mean daily flows as scaled by  $Q_{max}$  using Eq. (1). Models of  $Q_{max}$  that were advanced from cross-validation

**Table 3**  
Corresponding parameters, units, and performance measures for equations.

Return period (yrs)	Incpt (-)	a (mi <sup>2</sup> )	p (in.)	Imp <sub>max</sub> (-)	Adjusted R <sup>2a</sup>	Standard error <sup>b</sup> (%)	AIC <sub>c</sub> <sup>c</sup>	p-exceptions (p > 0.05)		
<i>(a) For Eq. (5)</i>										
1.5	-2.03	0.592	1.55	11.6	0.37	85	461			
2	-0.644	0.667	1.29	8.61	0.47	76	501			
5	2.137	0.838	0.773	3.23	0.70	59	603	P 0.08, Imp <sub>max</sub> 0.17		
10	2.90	0.868	0.767	0	0.81	45	637			
25	2.68	0.891	1.01	0	0.83	37	673			
50	2.63	0.902	1.11	0	0.82	37	700			
100	2.62	0.909	1.19	0	0.81	38	724			
Eq. number	Dependent variable	Incpt (-)	a (mi <sup>2</sup> )	yr (yrs)	p (in.)	Imp <sub>av</sub> (-)	Adjusted R <sup>2a</sup>	Standard error <sup>b</sup> (%)	AIC <sub>c</sub> <sup>c</sup>	p-values for Imp <sub>av</sub>
<i>(b) For DDF components Q<sub>max</sub> (scale) and day1 (magnitude) Eqs. (6) and (7)</i>										
(6)	Q <sub>max</sub>	-2.24	0.979	0.341	1.79	-	0.80	51	632	
(7)	day1	-12.9	0.676	1.85	3.71	13.8	0.75	92	709	0.002
Dependent variable	Incpt (-)	β <sub>Q10</sub> (cfs)	β <sub>day1</sub> (days, cfs)	β <sub>yr</sub> (yrs)	β <sub>impav</sub> (-)	Adjusted R <sup>2d</sup>	Standard error <sup>b</sup>	AIC <sub>c</sub> <sup>c</sup>	p-value for Imp <sub>av</sub>	
<i>(c) For DDF component day2 (shape) Eq. (8)</i>										
day2	-1.60	0.166	-0.138	0.129	0.720	0.85	9.1%	-188	0.060	

<sup>a</sup> Adjusted R<sup>2</sup> reported from geometric space.  
<sup>b</sup> Standard error of estimate expressed as percentage of sample mean in arithmetic space.  
<sup>c</sup> Corrected AIC reported from arithmetic space.  
<sup>d</sup> Adjusted R<sup>2</sup> reported from arithmetic space (for linear model).



**Hawley-Bledsoe (H-B):**  $Q = f(A, P, Imp)$   
**USGS (1977) Rural:**  $Q_r = f(A, P)$

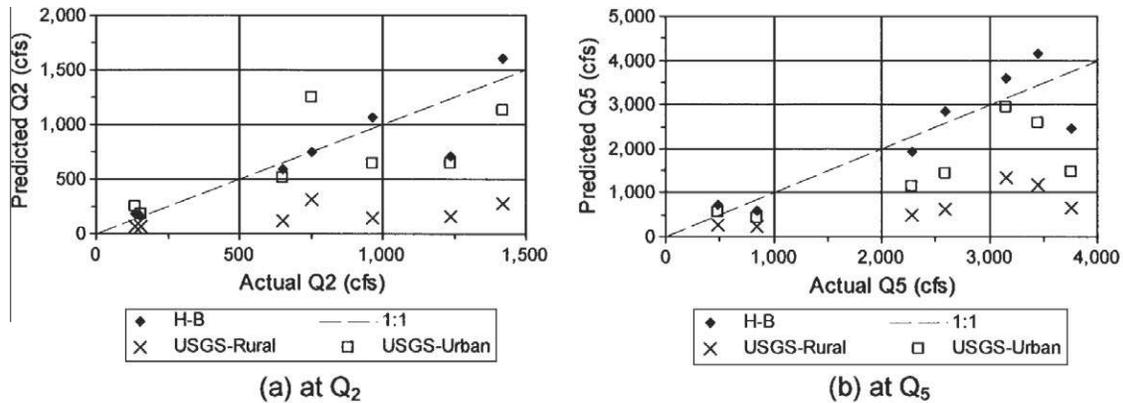
**Fig. 5.** Model performance of Hawley–Bledsoe and USGS (1977) rural at all 43 gauges: (a) at Q<sub>2</sub>, (b) at Q<sub>5</sub>, (c) at Q<sub>10</sub>, (d) at Q<sub>50</sub>.

performed comparatively well during final calibration (n = 43) in both geometric and arithmetic spaces, the only notable exception being that DD became less significant with p-values greater than 0.05. Consequently, we present the simplified model that had both high performance and shares a consistent format with the models of other DDF components. The model is presented below with corresponding performance measures in Table 3b:

$$Q_{max} = e^{(-2.24)} * A^{0.979} * P^{1.79} * Yr^{0.341} \tag{6}$$

where Q<sub>max</sub> is the maximum mean 24-h flow (cfs); and Yr is the length of mean daily flow record (yrs).

The other components of DDFs are the magnitude (day1) and shape (day2). Six models of day1 and day2 were advanced to final (n = 39) calibration from the cross-validation step. With similar



**Hawley-Bledsoe (H-B):**  $Q = f(A, P, Imp)$   
**USGS (1977) Rural:**  $Q_r = f(A, P)$   
**USGS (1983) Urban:**  $Q_u = f(A, S_{chn}, P_{22}, ST, BDF, Imp, Q_r)$

where:  $P_{22}$  = 2-yr 2-hr rainfall (in.);  
 $ST$  = basin storage (% of watershed); and  
 $BDF$  = Basin Development Factor, which is an index (0-12) of urban drainage improvements including channel improvements, channel linings, storm sewers, and curb and gutter.

**Fig. 6.** Model performance of Hawley-Bledsoe, USGS (1977) rural, and USGS (1983) urban in seven most urbanized watersheds (i.e., TIA > 5%): (a) at  $Q_2$ , (b) at  $Q_5$ .

performance from each model, we present the models that most regularly represent the median influence of urbanization on the durations of mean daily flows. The models are presented below with corresponding performance measures in Tables 3b and c:

$$\text{day1} = e^{(-12.9)} * A^{0.676} * P^{3.71} * Yr^{1.85} * e^{(13.8 * Imp_{av})} \quad (7)$$

$$\text{day2} = -1.60 + 0.166 * \ln(Q_{10}) - 0.138 * \ln(\text{day1}) + 0.129 * \ln(Yr) + 0.720 * Imp_{av} \quad (8)$$

where day1 is the coefficient of DDF calibrated in 'days' and 'cfs';  $Imp_{av}$  the average spatial extent of TIA during gauge record expressed as a fraction of the total drainage area ( $\text{mi}^2/\text{mi}^2$ ); day2 the exponent of DDF calibrated in 'days' and 'cfs';  $\ln()$  the natural logarithm (i.e., logarithm to the base e) of parenthetical variable; and  $Q_{10}$  is the 10-yr instantaneous peak flow (cfs).

#### 4. Implications and discussion

The models predict higher peak flows (especially for  $\leq Q_5$ ), and longer durations across all sediment-transporting flows in urban watersheds. In the following sections, we (1) provide an example application of the models to a hypothetical watershed with average conditions, controlling for all factors except imperviousness, and (2) present a case study contrasting two of the study gauges with differing land uses over identical periods to illustrate the effects of urbanization on flow peaks and durations in the study region.

##### 4.1. Effects of urbanization predicted by models

Large increases were found in instantaneous-peak flows of more frequent return periods, for example, ca. sixfold increase in  $Q_2$  in a watershed with 20% TIA relative to a rural setting (i.e.,  $\leq \sim 1\%$  TIA), with decreasing influence for less frequent storms (i.e.,  $Q_5$  peak factor ca. 2). Such attenuating influence of urbanization with return period is generally consistent with both theory and previous studies (Bledsoe and Watson, 2001; Durbin, 1974; Hollis, 1975; Rantz, 1971; Sauer et al., 1983). During very intense, infrequent events, precipitation rates can substantially overwhelm infiltration capacity and soil surfaces begin to behave more simi-

larly to impervious surfaces; however, it is noted that extreme events may still experience increased flows in urban basins due to greater hydraulic efficiency (Smith et al., 2002).

The effects of TIA as predicted by all calibrated models are summarized for  $Q_{1.5}$ ,  $Q_2$ ,  $Q_5$  and  $Q_{10}$  in Fig. 7, with the peak model presented above (Eq. (5)) indicated by the solid black line. The ephemeral nature of 15 of the 43 gauges ( $Q_1 = 0$  cfs) made models of  $Q_1$  too skewed for publication; however, the influence of urbanization was evident in that the four most urban gauges accounted for the four largest 1-yr flows.

The peak factors presented here are generally larger than those from previous studies. For example, Hammer (1972) and Hollis (1975) suggested that 1.5- to 2-yr flows could double or triple at 10–20% imperviousness, and Bledsoe and Watson (2001) found peak factors ranging 1.5–4 for  $Q_2$  at 20% imperviousness when controlling for other statistically-significant factors such as drainage area, precipitation (mean annual, 2-yr 24-h event, and 2-yr 24-h standardized by mean annual), average basin elevation, total stream length, drainage density, basin shape (length of main channel divided by drainage area), and reach-averaged valley slope at gauge location. The hydrogeomorphic differences of the southern California watersheds—both accounted and unaccounted for (e.g., soil type, hydrophobicity, and vegetative cover) in the model—could explain the differences from the more humid-temperate studies. The limited extent of flow-control practices in southern California could also play a role in the higher peak factors.

Finally, regarding peak-flow models, the finding that  $Imp_{max}$  accounted for more variance than other impervious measures such as  $Imp_{av}$  shows how the most developed portion of a gauge record can overwhelm peak flows from the undeveloped record years, especially for the more frequent return intervals. This suggests the potential for a statistically-significant influence at higher return intervals (e.g.,  $Q_{10}$ ) in the future as gauges have more time to capture large precipitation events at contemporary impervious levels.

Current gauge records did not show urbanization as statistically significant in explaining DDF scale ( $Q_{max}$ ), but urbanization had an exponential effect on the magnitude (day1), with a linear effect on day2 (shape). The combined effect tends to magnify durations of the moderate flows more than durations of the largest flows.

**Table 4**  
Model performance of Hawley–Bledsoe and USGS rural (Waananan and Crippen, 1977) at all 43 gauges, and USGS urban (Sauer et al., 1983) in seven most urbanized watersheds (i.e., TIA > 5%).

Return period (yrs)	All 43 (predominantly rural) gauges		Standard error of estimate as % of sample mean (arithmetic space)		Seven most urban gauges (TIA > 5%) <sup>a</sup>		R <sup>2</sup> (arithmetic space)	
	Adjusted R <sup>2</sup> (geometric space)		H–B		USGS (1983) urban <sup>b,c</sup>		USGS (1983) urban <sup>b,c</sup>	
	H–B <sup>b</sup>	USGS (1977) rural <sup>b</sup>	H–B	USGS (1977) rural	H–B	USGS (1983) urban <sup>b,c</sup>	H–B	USGS (1983) urban <sup>b,c</sup>
1.5	0.37	–	85%	–	0.87	–	0.66	–
2	0.47	0.18	76%	104%	0.92	0.76	0.77	0.45
5	0.70	0.38	59%	80%	0.84	0.39	0.72	0.09
10	0.81	0.47	45%	67%	0.81	0.47	0.70	0.24
25	0.83	0.67	37%	51%				
50	0.82	0.73	37%	43%				
100	0.81	0.75	38%	40%				

<sup>a</sup> Seven most urban watersheds (and respective USGS gauge numbers) and maximum TIA during records include: Aliso (11047500, 8.1%), Arroyo Simi (11105850, 8.6%), Los Coches (11022200, 9.1%), San Diego (11048500, 14.9%), Los Penasquitos, US (11023330, 15.2%), Arroyo Trabuco (11047300, 18.8%), and Los Penasquitos\_DS (11023340, 20.1%).

<sup>b</sup> Hawley–Bledsoe (H–B):  $Q = f(A, P, Imp_{max})$ ; USGS (1977) rural:  $Q_r = f(A, P)$ ; and USGS (1983) urban:  $Q_u = f(A, S_{ch}, P_{25}, ST, BDF, Imp, Q_2)$ ; where: A = drainage area;  $P$  = mean annual precipitation;  $Imp_{max}$  = maximum extent during gauge record of TIA as fraction of total drainage area (value also used for “Imp” in the USGS equation);  $S_{ch}$  = main channel slope as measured between 10% and 85% of the distance from gage to basin divide;  $P_{25}$  = 2-yr 2-h precipitation; ST = percentage of basin occupied by lakes, reservoirs, swamps, and wetlands; and BDF = basin development factor (0–12) index representing the extent of drainage improvements including (1) channel improvements, (2) channel linings, (3) storm sewers, and (4) curb and gutter streets (after Sauer et al. (1983)).

<sup>c</sup> Standardized performance measures such as standard error or Adjusted R<sup>2</sup> could not be determined for the USGS urban equation due to the number of samples (7) relative to the number of predictor variables (7).

Fig. 8a presents a 25-yr simulation of an average watershed across a gradient of TIA, depicting substantial increases in durations of the respective flow bins with increasing TIA. For example, at 20% TIA, mean daily flows on the order of 100 cfs are projected to have five-fold duration increases, with the highest flows (on the order of 1000 cfs) increasing in duration by ca. threefold (Fig. 8b).

The findings of this study indicating decreasing influence of urbanization on flow duration with increasing flow magnitude are consistent with the findings regarding peak flows: urbanization tends to show higher influence on more frequent events, with decreasing influence over the largest, rarest storms. In conclusion, the fact that  $Imp_{av}$  outperformed  $Imp_{max}$  in DDF models suggests that it may take longer for urbanization to show an effect on the cumulative durations of all flows than to appreciably affect instantaneous peaks at small return intervals.

#### 4.2. At-a-station effects of urbanization

Two watersheds, located in the southeast corner of Ventura County and separated by less than 6 km (Fig. 1), had gauges that were active during identical years (1934–1983). The Hopper Creek gauge (No. 11110500) occupied a watershed that remains entirely undeveloped. In contrast, the Arroyo Simi gauge (No. 11105850) spanned equal periods of relatively undeveloped (1934–1958) and developed (1959–1983) land use. By dividing each record into equal sub-samples, the paired data lend support to the findings discussed above. Regarding peak flows, Arroyo Simi (Fig. 9a) had more than a 10-fold increase in the 2-yr flow and a threefold increase in the 25-yr flow between the two time periods. In contrast, peak flows differed by an average of only 20% across the same periods in the rural watershed (Fig. 9b), and are likely attributable to the variability in the inter-period precipitation. The change in land use is also evident when comparing the durations of mean daily flows between the two periods at Arroyo Simi (Fig. 10a), with the undeveloped regime, for example, incurring only 4 days at 500 cfs and the post-developed regime having 21 days at 600 cfs. By comparison, DDFs of the two identical time periods at the rural gauge are nearly overlaid (Fig. 10b).

These differences in flow magnitudes and durations between undeveloped and developed periods at the same gauge and the relative similarity during identical periods at the nearby rural gauge add to the weight of evidence that such changes are largely attributable to urbanization. The disparities could also be attributable, in part, to differences in surface slope and watershed size: 42 vs. 23% and 38 vs. 112 km<sup>2</sup> in the Hopper Creek and Arroyo Simi basins, respectively. However, other potentially-important hydrogeomorphic characteristics such as elevation, annual precipitation, and drainage density were relatively comparable between the basins. Most importantly, their proximity, similar elevations, and identical years of operation mean that they should have experienced very similar precipitation events and climatic trends during their records, making their inter-period cross-comparison relevant to this study.

In fact, the inter-period changes in flow regimes that were observed concomitant with development at Arroyo Simi were larger than what is predicted by the regional models developed herein, particularly in terms of the more rare events such as  $Q_{25}$  and  $Q_{max}$ . For example,  $Q_{max}$  was over three times larger during the urban portion of the record at Arroyo Simi, but  $Q_{max}$  actually decreased between the two identical periods at the rural gauge. Indeed, the effects of urbanization captured in the regional models may have been dampened by the widespread variability across all sites, most of which were still relatively undeveloped. As more years of data are gathered at urban gauges, the models could be further refined to account for urbanization with a more equitable sampling of urban data.

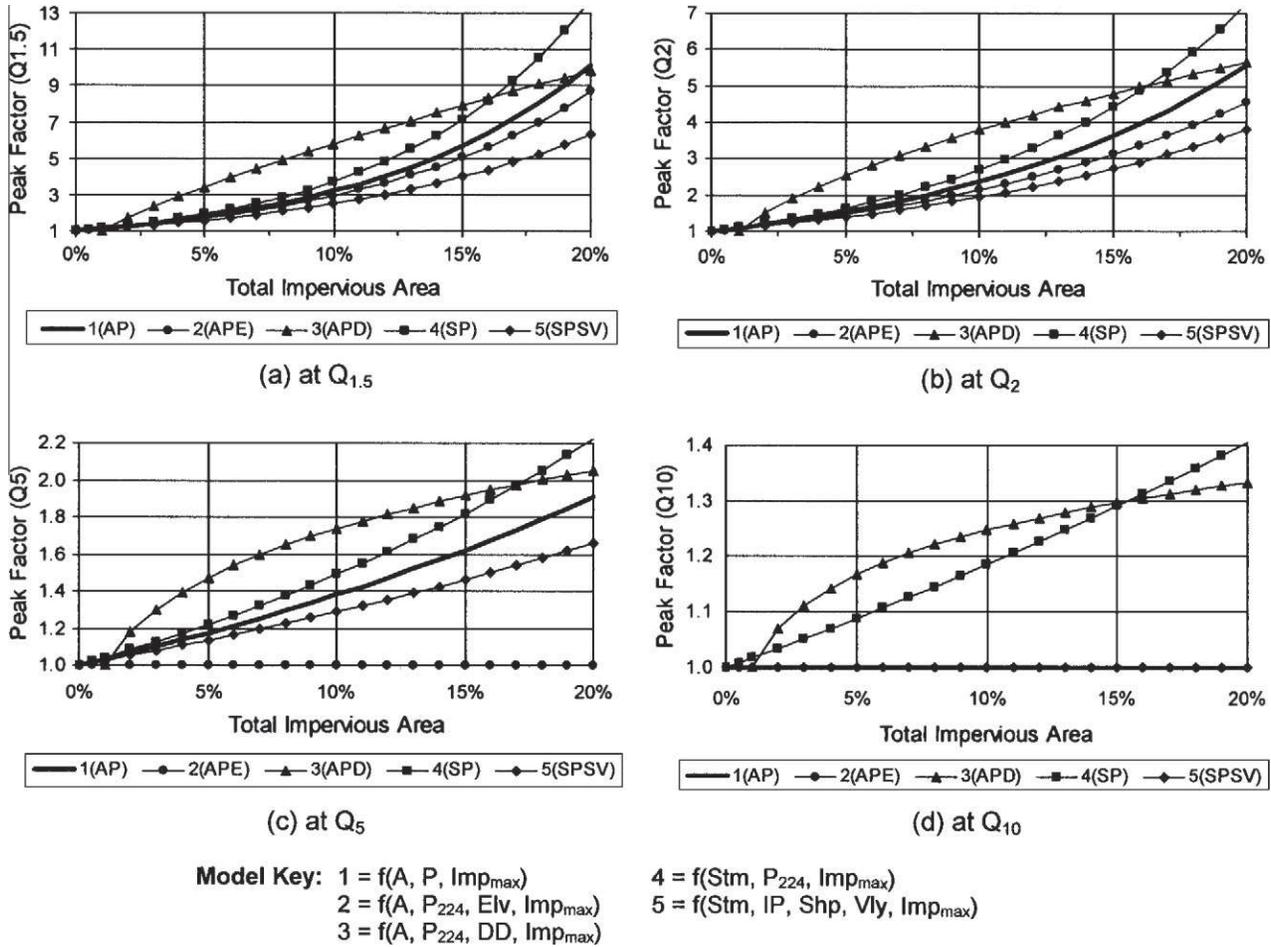
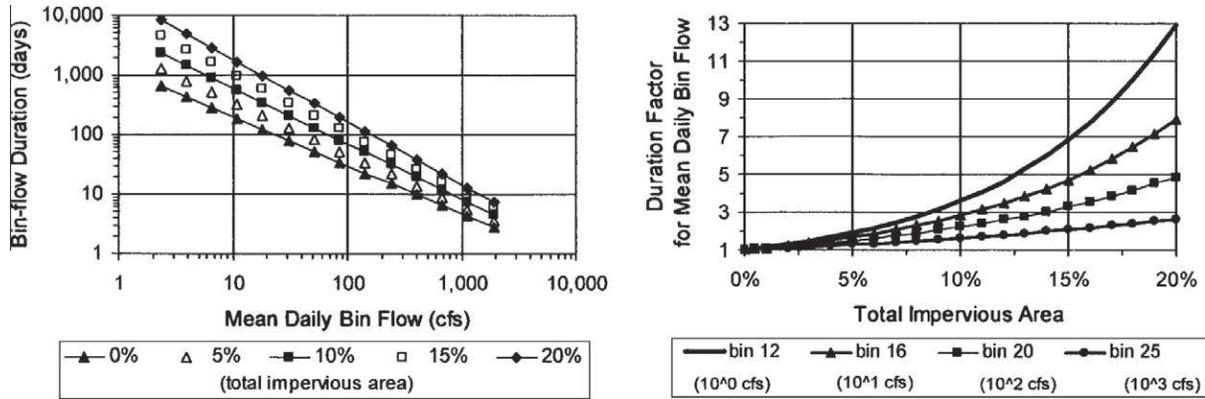


Fig. 7. Peak factors for instantaneous peak flows as a function of TIA for all five calibrated models: (a) at  $Q_{1.5}$ , (b) at  $Q_2$ , (c) at  $Q_5$ , (d) at  $Q_{10}$ .



(a) bin-flow duration versus bin-flow centroids (b) duration factors for respective bins versus TIA  
 Average Watershed:  $A = 25 \text{ mi}^2$  and  $P = 25 \text{ in.}$

Fig. 8. Twenty five year DDF simulations of an average watershed across a gradient of TIA: (a) bin-flow duration versus bin-flow centroids and (b) duration factors for respective bins versus TIA.

5. Summary and conclusions

The overarching objective of this paper was to understand the effects of suburbanization on the flow regimes (i.e., ‘hydromodification’) of semi-arid streams using a southern California case study. As a first step and in direct support of our work on understanding

the flow regime effects, we developed updated alternatives to the USGS regional equations for instantaneous peak flows, which outperformed earlier rural (Waananen and Crippen, 1977) and urban (Sauer et al., 1983) models, with particularly substantial improvements for return periods  $\leq 10$  yrs (e.g., Adjusted  $R^2$  0.81 versus 0.47 at  $Q_{10}$ ). In order to meet our primary objective regarding both

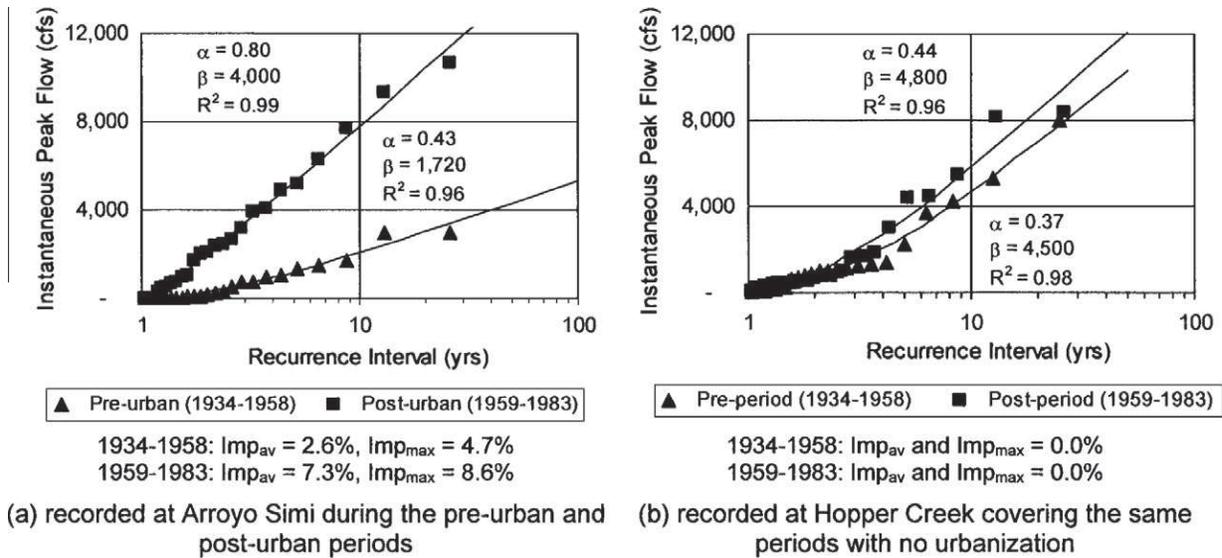


Fig. 9. Instantaneous-peak flow relative to recurrence interval, with fitted gamma distributions: (a) recorded at Arroyo Simi during the pre-urban and post-urban periods and (b) recorded at Hopper Creek covering the same periods with no urbanization.

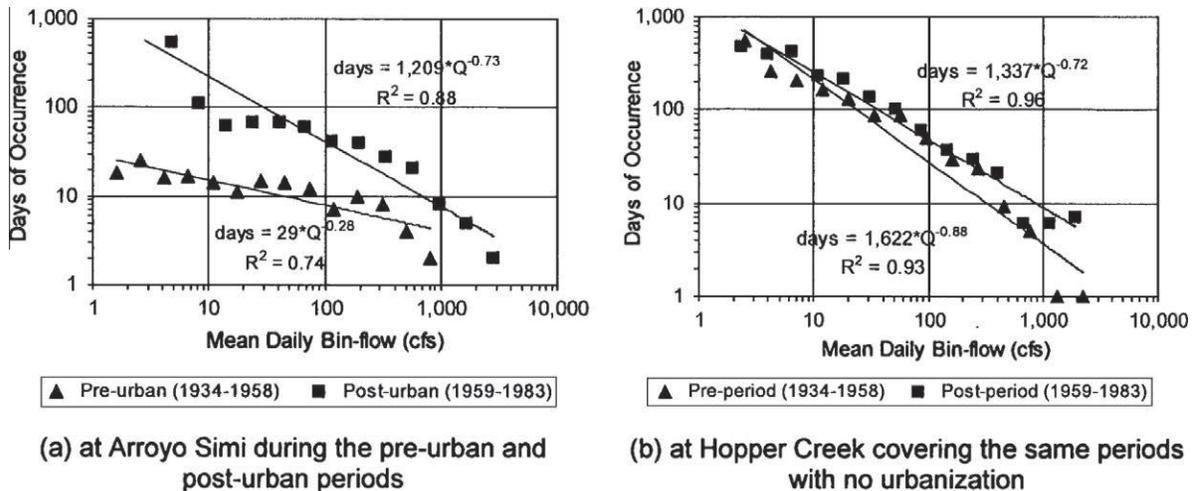


Fig. 10. Cumulative-duration histogram centroids, with fitted DDFs: (a) at Arroyo Simi during the pre-urban and post-urban periods and (b) at Hopper Creek covering the same periods with no urbanization.

flow magnitudes and durations, we developed a novel method for estimating long-term cumulative durations at ungauged sites that expands on previous approaches to histogram-style duration curves by predicting their magnitude, shape, and scale using watershed physical properties. Most importantly in the context of hydromodification, both the peak flow and DDF models account for urbanization using measures of total impervious area, which were statistically significant ( $p < 0.05$ ), particularly for peak flows  $\leq Q_2$  and the magnitude (coefficient) component of DDFs.

Multivariate regression that tested and controlled for other potentially-significant hydro-climatic variables (e.g., surface slope, drainage density, etc.) correlated urbanization to higher peaks and longer durations of all geomorphically-significant flows. This effect was also documented at an individual gauge whose records spanned both pre-urban and post-urban periods. Moreover, these effects were not linear. Although several metrics and forms were tested for modeling the effects of urbanization, the form that was most powerful was typically the exponential of total impervious area. That is, flow magnitudes and durations associated with identical watersheds differing only by measures of imperviousness (e.g., ~1% and

~20%) would be disproportionately larger, with the most substantial differences at the more frequent events. For example, instantaneous peak flows would increase by factors of ~10, 6, and 2 for  $Q_{1.5}$ ,  $Q_2$ , and  $Q_5$ , respectively. Durations would increase by factors of ~8, 5, and 3 for mean daily flows on the order of 10, 100, and 1000 cfs, respectively, in an average study basin (i.e.,  $A = 25 \text{ mi}^2$ ,  $P = 25 \text{ in.}$ ). Although we do not have duration factors from other regions for comparison, the peak factors presented herein are greater than studies from humid-temperate climates, suggesting that semi-arid flow regimes may be more sensitive to the effects of urbanization.

Such changes in the hydrologic regime can have far-reaching effects on receiving channels in terms of cumulative erosive energy and channel stability. The relatively dramatic responses in channel form that have been observed throughout the region are better explained in the context of such equally compelling changes in flow durations of sediment-transporting events. The empirically-calibrated DDF models presented here may become important tools in developing an improved understanding of hydromodification effects on fluvial systems in southern California, and offer a framework for studies in other regions.

## 6. Future work

A logical next step regarding hydromodification research would be to apply these hydrologic models to channels where geomorphic data have been collected to evaluate whether changes in flows correspond to sediment discontinuities that in turn correlate to the degree of channel degradation. For example, these hydrologic models could provide a starting point for developing risk-based models of channel stability. Future work could also focus on the refinement of the DDF models developed in this study. For example, we were limited to mean daily flow data for these analyses, but more precise models could be developed using finer resolution data (e.g., 15-min or hourly) where available.

## Acknowledgements

We would like to thank organizations and individuals who contributed to this paper in terms of funding, data collection, and analysis. This research was funded in part by the California State Water Resources Control Board (Agreement No. 06-295-559-0) and by the County of San Diego, for which we are very grateful. Eric Stein of the Southern California Coastal Water Research Project (SCCWRP) played a primary role in acquiring research funds and project guidance. Becky Schaffner of SCCWRP was extremely helpful in providing GIS assistance and training. The USGS continues to deliver spatial and hydrologic data to the public domain that are unparalleled in terms of scale and quality. We would like to extend our gratitude to all staff (past and present), and particularly to Scott Patterson of the South Poway Field Office whose time and care in answering questions provided clarity regarding individual gauges. We would also like to thank the anonymous reviewers who provided constructive suggestions on earlier drafts and subsequent reviewers who offered very helpful and specific suggestions. This paper is much improved as a result. Finally, our sincere thanks go to Gloria Garza for her dedicated assistance in preparing the manuscript for journal submission.

## References

- Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. Autom. Control* AC-19, 716–723.
- Beighley, R.E., Moglen, G.E., 2002. Trend assessment in rainfall–runoff behavior in urbanizing watersheds. *J. Hydrol. Eng.* 7 (1), 27–34.
- Biedenham, D.S., Copeland, R.R., Thorne, C.R., Soar, P.J., Hey, R.D., Watson, C.C., 2000. Effective Discharge Calculation: A Practical Guide. ERDC/CHLI TR-00-15, US Army Corps of Engineers, Engineer Research and Development Center, Coastal and Hydraulics Laboratory, Vicksburg, Miss.
- Biedenham, D.S., Thorne, C.R., Soar, P.J., Hey, R.D., Watson, C.C., 2001. Effective discharge calculation guide. *Int. J. Sediment Res.* 16 (4), 445–459.
- Bledsoe, B.P., Watson, C.C., 2001. Effects of urbanization on channel instability. *J. Am. Water Resour. Assoc.* 37 (2), 255–270.
- Booth, D.B., 1990. Stream-channel incision following drainage-basin urbanization. *Water Resour. Bull.* 26 (3), 407–417.
- Booth, D.B., Jackson, C.R., 1997. Urbanization of aquatic systems: degradation thresholds, stormwater detection, and the limits of mitigation. *J. Am. Water Resour. Assoc.* 33 (5), 1077–1090.
- Chin, A., 2006. Urban transformation of river landscapes in a global context. *Geomorphology* 79 (3–4), 460–487.
- Chin, A., Gregory, K.J., 2001. Urbanization and adjustment of ephemeral stream channels. *Ann. Assoc. Am. Geogr.* 91 (4), 595–608.
- Chow, V.T., 1964. *Handbook of Applied Hydrology*. McGraw-Hill, Inc., New York, 1467 pp.
- Chow, V.T., Maidment, D.R., Mays, L.W., 1988. *Applied Hydrology*. McGraw-Hill, Inc., New York, 572 pp.
- Durbin, T.J., 1974. Digital Simulation of the Effects of Urbanization on Runoff in the Upper Santa Ana River Valley, California. US Geological Survey, Menlo Park, Calif.
- Emmett, W.W., 1975. The Channels and Waters of the Upper Salmon River, Idaho. Professional Paper 870A, US Geological Survey, Washington, D.C.
- Environmental Protection Agency, 2006. Management measures for hydromodification. In: *Guidance Specifying Management Measures for Nonpoint Pollution in Coastal Waters*. US Environmental Protection Agency (Chapter 6).

- Hammer, T.R., 1972. Stream channel enlargement due to urbanization. *Water Resour. Res.* 8, 139–167.
- Hawley, R.J., 2009. Effects of Urbanization on the Hydrologic Regimes and Geomorphic Stability of Small Streams in Southern California. Ph.D. Dissertation, Colorado State University, Department of Civil and Environmental Engineering, Fort Collins, Colo., 393 pp.
- Hey, R.D., 1975. Design discharge for natural channels. In: Hey, R.D., Davies, T.D. (Eds.), *Science, Technology and Environmental Management*. Saxon House, pp. 73–88.
- Hollis, G.E., 1975. Effect of urbanization on floods of different recurrence interval. *Water Resour. Res.* 11 (3), 431–435.
- Jennings, M.E., Thomas, W.O.J., Riggs, H.C., 1994. Nationwide Summary of US Geological Survey Regionally Regression Equations for Estimating Magnitude and Frequency of Floods for Ungaged Sites, 1993. Water-Resources Investigation Report 94-4002, US Geological Survey, p. 196.
- Konrad, C.P., Booth, D.B., 2002. Hydrologic Trends Associated with Urban Development for Selected Streams in the Puget Sound Basin, Western Washington. Water-Resources Investigations Report 02-4040, US Geological Survey, Tacoma, Wash.
- Leopold, L.B., 1994. *A View of the River*. Harvard University Press, Cambridge Mass, 29 pp.
- Novotny, V., 2003. *Water Quality: Diffuse Pollution and Watershed Management*. John Wiley & Sons, New York.
- Poff, N.L., Bledsoe, B.P., Cuhacian, C.O., 2006. Hydrologic variation with land use across the contiguous United States: geomorphic and ecological consequences for stream ecosystems. *Geomorphology* 79 (3–4), 264–285.
- Raff, D.A., Bledsoe, B.P., Flores, A.N., 2004. *GeoTool User's Manual*. Colorado State University, Fort Collins, Colo.
- Rantz, S.E., 1971. Suggested Criteria for Hydrologic Design of Storm-drainage Facilities in the San Francisco Bay Region, California. US Geological Survey.
- Roesner, L.A., Bledsoe, B.P., 2002. Physical Effects of Wet Weather Flows on Aquatic Habitats – Present Knowledge and Research Needs. WERF Project Number 00-WSM-4, Water Environment Research Foundation.
- Santa Clara, 2004. Hydromodification Management Plan Report. Santa Clara Valley Urban Runoff Pollution Prevention Program, Sunnyvale, Calif.
- Sauer, V.B., Thomas, W.O.J., Stricker, V.A., Wilson, K.V., 1983. Flood Characteristics of Urban Watersheds in the United States. US Geological Survey.
- Sheng, J., Wilson, J.P., 2009. Watershed urbanization and changing flood behavior across the Los Angeles metropolitan region. *Nat. Hazards* 48, 41–57.
- Simon, A., Downs, P.W., 1995. An interdisciplinary approach to evaluation of potential instability in alluvial channels. *Geomorphology* 12, 215–232.
- Smith, J.A., Baeck, M.L., Morrison, J.E., Sturdevant-Rees, P.L., Turner-Gillespie, D.F., Bates, P.D., 2002. The regional hydrology of extreme floods in an urbanizing drainage basin. *Am. Meteorol. Soc.* 3, 267–282.
- Sugiura, N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Commun. Stat. A*(7), 13–26.
- Trimble, S.W., 1997. Contribution of stream channel erosion to sediment yield from an urbanizing watershed. *Science* 278, 1442–1444.
- US Water Resources Council, 1967. A Uniform Technique for Determining Flood Flow Frequencies. Bulletin 15 of the Hydrology Subcommittee, US Water Resources Council, Washington, D.C., p. 15.
- Waananen, A.O., 1969. Urban effects on water yield. In: Moore, W.L., Morgan, C.W. (Eds.), *Water Resources Symposium*. University of Texas Press, pp. 169–182.
- Waananen, A.O., Crippen, J.R., 1977. Magnitude and Frequency of Floods in California. US Geological Survey.
- Waters, T.F., 1995. Sediment in streams – sources, biological effects, and control. *Am. Fish. Soc. Monograph* 7, 251.
- Watson, C.C., Dubler, D., Abt, S.R., 1997. Demonstration Erosion Control Project Report. CSU Report submitted to US Army Engineer Waterways Experiment Station, Vicksburg, Miss.
- White, M.D., Greer, K.A., 2006. The effects of watershed urbanization on the stream hydrology and riparian vegetation of Los Penasquitos Creek, California. *Landsc. Urban Plan.* 74, 125–138.
- Yevjevich, V., 1972. *Probability and Statistics in Hydrology*. Water Resources Publications, Fort Collins, Colo, 302 pp.

## Internet references – geospatial data sources

- Cal-Atlas: 2000 and 2007 Roadway Shapefiles, State of California Geospatial Clearinghouse. <<http://www.atlas.ca.gov>>.
- Google Earth: Present-day Aerial Photography. <<http://earth.google.com>>.
- National Oceanic and Atmospheric Administration: Precipitation Intensities for 2-year, 24-hour Storm. <<http://www.nws.noaa.gov/oh/hdsc/noaaatlas2.htm>> (Atlas 2), National Weather Service, Hydrometeorological Design Studies Center; and [hdsc.nws.noaa.gov/hdsc/pfds/pfds\\_gis.html](http://hdsc.nws.noaa.gov/hdsc/pfds/pfds_gis.html) (Atlas 14), National Weather Service, Hydrometeorological Design Studies Center, Precipitation Frequency Data Server.
- US Department of Agriculture, Natural Resources Conservation Service: Soil Surveys, Average Annual Precipitation Shapefile (1961–1990), Geospatial Data Gateway. <<http://www.datagateway.nrcs.usda.gov/>>.
- US Geological Survey: Historical Aerial Photography and Quadrangle Topographic Maps, National Elevation Dataset, 2001. Impervious Raster, National Hydrography Dataset, Average Annual Precipitation Shapefile (1900–1960), The National Map Seamless Server. <<http://www.seamless.usgs.gov>>.