

Construction Safety Risk Modeling and Simulation

Antoine J.-P. Tixier,^{1,*} Matthew R. Hallowell,² and Balaji Rajagopalan²

By building on a genetic-inspired attribute-based conceptual framework for safety risk analysis, we propose a novel approach to define, model, and simulate univariate and bivariate construction safety risk at the situational level. Our fully data-driven techniques provide construction practitioners and academicians with an easy and automated way of getting valuable empirical insights from attribute-based data extracted from unstructured textual injury reports. By applying our methodology on a data set of 814 injury reports, we first show the frequency-magnitude distribution of construction safety risk to be very similar to that of many natural phenomena such as precipitation or earthquakes. Motivated by this observation, and drawing on state-of-the-art techniques in hydroclimatology and insurance, we then introduce univariate and bivariate nonparametric stochastic safety risk generators based on kernel density estimators and copulas. These generators enable the user to produce large numbers of synthetic safety risk values faithful to the original data, allowing safety-related decision making under uncertainty to be grounded on extensive empirical evidence. One of the implications of our study is that like natural phenomena, construction safety may benefit from being studied quantitatively by leveraging empirical data rather than strictly being approached through a managerial perspective using subjective data, which is the current industry standard. Finally, a side but interesting finding is that in our data set, attributes related to high energy levels (e.g., machinery, hazardous substance) and to human error (e.g., improper security of tools) emerge as strong risk shapers.

KEY WORDS: Construction safety; risk modeling; stochastic simulation

1. INTRODUCTION AND MOTIVATION

Despite the significant improvements that have followed the inception of the Occupational Safety and Health Act of 1970, safety performance has reached a plateau in recent years and the construction industry still suffers from a disproportionate accident rate. Fatalities in construction amounted to 885 in 2014, the highest count since 2008.⁽¹⁾ In addition to dreadful human costs, construction injuries

are also associated with huge direct and indirect economic impacts.

A very large portion of construction work, upstream or downstream of groundbreaking, involves making safety-related decisions under uncertainty. Partly due to their limited personal history with accidents, even the most experienced workers and safety managers may miss hazards and underestimate the risk of a given construction situation.^(2,3) Designers face an even greater risk of failing to recognize hazards and misestimating risk.⁽²⁾ In addition, when uncertainty is involved, humans often recourse to personal opinion and intuition to apprehend their environment. This process is fraught with numerous biases and misconceptions inherent to human cognition⁽⁴⁾ and compounds the likelihood of misdiagnosing the riskiness of a situation.

¹Computer Science Laboratory, École Polytechnique, Palaiseau, France.

²Department of Civil, Environmental, and Architectural Engineering, CU Boulder, USA.

*Address correspondence to Antoine J.-P. Tixier, Postdoctoral Researcher, Computer Science Laboratory, École Polytechnique, Palaiseau, France; antoine.tixier-1@colorado.edu.

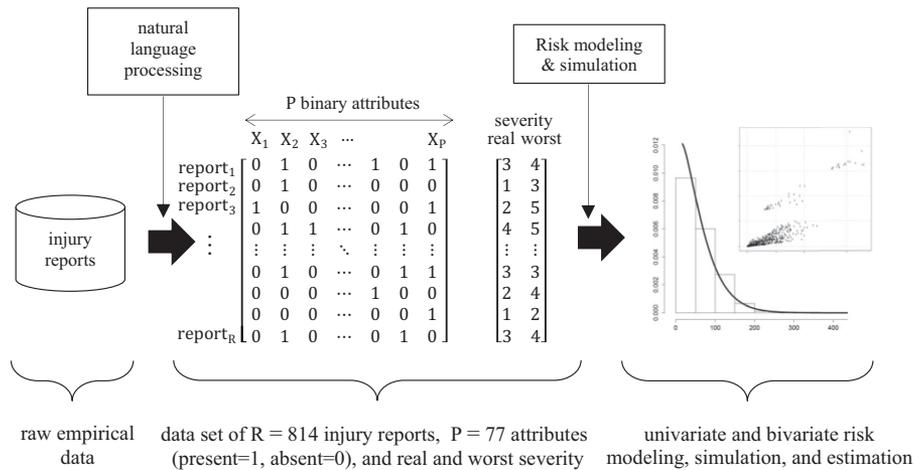


Fig. 1. Overarching research process: from raw injury reports to safety risk analysis.

Therefore, it is of paramount importance to provide construction practitioners with tools to mitigate the adverse consequences of uncertainty on their safety-related decisions. In this study, we focus on leveraging situational data extracted from raw textual injury reports to guide and improve construction situation risk assessment. Our methodology facilitates the augmentation of construction personnel’s experience and grounds risk assessment on potentially unlimited amounts of empirical and objective data. In other words, our approach combats construction risk misdiagnosis on two fronts, by jointly addressing both the limited personal history and the judgment bias problems previously evoked.

We used fundamental construction attribute data extracted by a highly accurate natural language processing (NLP) system⁽⁵⁾ from a database of 921 injury reports provided by a partner company engaged in industrial construction projects worldwide. Attributes are context-free universal descriptors of the work environment that are observable prior to injury occurrence. They relate to environmental conditions, construction means and methods, and human factors, and provide a unified, standardized way of describing any construction situation. To illustrate, one can extract four attributes from the following text: “worker is unloading a ladder from pickup truck with bad posture”: ladder, manual handling, light vehicle, and improper body positioning. Because attributes can be used as leading indicators of construction safety performance,⁽⁶⁾ they are also called injury precursors. In what follows, we will use the two terms interchangeably. Drawing from

national databases, Esmaeili and Hallowell^(7,8) initially identified 14 and 34 fundamental attributes from 105 fall and 300 struck-by high-severity injury cases, respectively. In this study, we used a refined and broadened list of 80 attributes carefully engineered and validated by Prades⁽⁹⁾ and Desvignes⁽¹⁰⁾ from analyzing a large database of 2,201 reports featuring all injury types and severity levels.

A total of 107 of 921 reports were discarded because they either were not associated with any attribute or because the real outcome was unknown. Additionally, 3 attributes out of 80 (pontoon, soffit, and poor housekeeping) were removed because they did not appear in any report. This gave us a final matrix of $R = 814$ reports by $P = 77$ attributes. Although other related studies concerned themselves with predictive modeling,⁽⁶⁾ here we focus on defining, modeling, and simulating attribute-based construction safety risk. The overall study pipeline is summarized in Fig. 1.

The contributions of this study are fourfold: (1) we formulate an empirically-grounded definition of construction safety risk at the attribute level, and extend it to the situational level, both in the univariate and the bivariate case; (2) we show how to model risk using kernel density estimators (KDE); (3) we observe that the frequency-magnitude distribution of risk is heavy-tailed, and resembles that of many natural phenomena; and finally (4) we introduce univariate and bivariate nonparametric stochastic generators based on kernels and copulas to draw conclusions from much larger samples and better estimate construction safety risk.

Table I. Counts of Injury Severity Levels Accounted for by Each Precursor

Precursors	Severity Levels				
	$s_1 = \text{Pain}$	$s_2 = \text{1st Aid}$	$s_3 = \text{Medical Case/Lost Work Time}$	$s_4 = \text{Permanent Disablement}$	$s_5 = \text{Fatality}$
X_1	n_{11}	n_{12}	n_{13}	n_{14}	n_{15}
X_2	n_{21}	n_{22}	n_{23}	n_{24}	n_{25}
...
X_{P-1}	$n_{(P-1)1}$	$n_{(P-1)2}$	$n_{(P-1)3}$	$n_{(P-1)4}$	$n_{(P-1)5}$
X_P	n_{P1}	n_{P2}	n_{P3}	n_{P4}	n_{P5}

Table II. Severity Level Impact Scores Adapted from Hallowell and Gambatese⁽¹⁶⁾

Severity Level (s)	Severity Scores (S_s)
Pain	$S_1 = 12$
1st aid	$S_2 = 48$
Medical case/lost work time	$S_3 = 192$
Permanent disablement	$S_4 = 1,024$
Fatality	$S_5 = 26,214$

2. BACKGROUND AND POINT OF DEPARTURE

The vast majority of construction safety risk analysis studies use opinion-based data,⁽⁹⁾ and thus rely on the ability of experts to rate the relative magnitude of risk based on their professional experience. This approach suffers two main limitations. First, prior ranges are very often provided by researchers to bound risk values. Second, and more importantly, even the most experienced experts have limited personal history with hazardous situations, and their judgment under uncertainty suffers the same cognitive limitations as that of any other human being,⁽¹¹⁾ such as overconfidence, anchoring, availability, representativeness, unrecognized limits, motivation, and conservatism.⁽¹¹⁻¹³⁾ It was also suggested that gender⁽¹⁴⁾ and even emotional state⁽¹⁵⁾ impact risk perception. Even if it is possible to somewhat alleviate the negative impact of adverse psychological factors,⁽¹⁶⁾ the reliability of data obtained from expert opinion is questionable. Conversely, truly objective empirical data, like the injury reports used in this study, seem superior.

Due to the technological and organizational complexity of construction work, most safety risk studies assume for simplicity that construction processes can be decomposed into smaller parts.⁽¹⁷⁾

Such decomposition allows researchers to model risk for a variety of units of analysis, like specific tasks and activities.⁽¹⁸⁻²⁰⁾ Most commonly, trade-level risk analysis has been adopted.⁽²¹⁻²³⁾ The major limitation of these segmented approaches is that because each one considers a trade, task, or activity in isolation, it is impossible for the end user to comprehensively characterize onsite risk in a standard, robust, and consistent way.

Some studies attempted to overcome these limitations. For instance, Shapira and Lyachin⁽²⁴⁾ quantified risks for generic factors related to tower cranes such as type of load or visibility, thereby allowing safety risk modeling for any crane situation. Esmaeili and Hallowell^(7,8) went a step further by introducing a novel conceptual framework allowing any construction situation to be fully and objectively described by a unique combination of fundamental context-free attributes of the work environment. This attribute-based approach is powerful in that it shows possible the extraction of structured standard information from naturally occurring, unstructured textual injury reports. Additionally, the universality of attributes allows to capture the multifactorial nature of safety risk in the same unified way for any task, trade, or activity, which is a significant improvement over traditional segmented studies. However, manual content analysis of injury reports is expensive and fraught with data consistency issues. For this reason, Tixier *et al.*⁽⁵⁾ introduced an NLP system capable of automatically detecting the attributes presented in Table III and various safety outcomes in injury reports with more than 95% accuracy (comparable to human performance), enabling the large-scale use of Esmaeili and Hallowell's attribute-based framework. The data we used in this study were extracted by the aforementioned NLP tool.

Table III. Relative Risks and Counts of the $P = 77$ Injury Precursors

Precursor	n	e (%)	Risk Based on		Precursor	n	e (%)	Risk Based on	
			Real	Worst Possible				Real	Worst Possible
			Outcomes				Outcomes		
Concrete	29	41	7	96	Unstable support/surface	3	32	1	2
Confined workspace	21	2	115	336	Wind	29	37	6	16
Crane	16	12	22	76	Improper body position	7	25	3	6
Door	17	21	11	174	Imp. procedure/inattention	13	16	10	44
Sharp edge	8	38	2	5	Imp. security of materials	78	12	77	1007
Formwork	22	5	63	135	Insect	19	18	8	21
Grinding	16	16	11	34	No/improper PPE	3	67	0*	1
Heat source	11	20	4	13	Object on the floor	41	43	9	22
Heavy material/tool	29	30	11	247	Lifting/pulling/handling	141	31	49	439
Heavy vehicle	12	12	12	307	Cable tray	9	27	4	11
Ladder	23	14	15	52	Cable	8	33	1	3
Light vehicle	31	59	7	123	Chipping	4	16	1	4
Lumber	69	14	53	158	Concrete liquid	8	41	2	4
Machinery	40	8	67	3159	Conduit	11	31	4	14
Manlift	8	8	16	50	Congested workspace	2	32	0*	1
Object at height	14	50	4	136	Dunnage	2	16	1	3
Piping	74	38	19	141	Grout	3	41	1	1
Scaffold	91	33	28	74	Guardrail handrail	16	40	4	8
Stairs	28	41	8	25	Job trailer	2	59	0*	1
Steel/steel sections	112	35	33	281	Stud	4	41	1	5
Rebar	33	4	76	251	Spool	9	33	2	9
Unpowered transporter	13	9	23	401	Stripping	12	22	7	18
Valve	24	27	9	22	Tank	16	31	5	115
Welding	25	22	10	34	Drill	16	43	5	88
Wire	30	43	5	19	Bolt	36	41	7	27
Working at height	73	40	18	46	Cleaning	22	56	5	12
Wkg below elev. wksp/mat.	7	17	3	21	Hammer	33	50	5	18
Forklift	11	9	9	380	Hose	11	41	3	8
Hand size pieces	38	47	7	95	Nail	15	50	4	10
Hazardous substance	33	1	590	6648	Screw	7	50	1	2
Adverse low temperatures	33	3	101	292	Slag	10	10	8	32
Mud	6	6	9	20	Spark	1	12	2	11
Poor visibility	3	23	2	3	Wrench	23	39	5	23
Powered tool	32	27	12	54	Exiting/transitioning	25	49	6	17
Slippery surface	32	25	13	40	Splinter/sliver	9	44	1	2
Small particle	96	31	28	105	Working overhead	5	40	1	3
Unpowered tool	102	44	24	352	Repetitive motion	2	51	0*	1
Electricity	1	33	0*	1	Imp. security of tools	24	22	12	314
Uneven surface	33	32	11	129					

*Values are rounded up to the nearest integer.

3. UNIVARIATE ANALYSIS

3.1. Attribute-Level Safety Risk

Following Baradan and Usmen⁽²¹⁾ we defined construction safety risk as the product of frequency and severity as shown in Equation (1). More precisely, in our approach, the safety risk R_p accounted for by precursor_p (or X_p in Tables I) was computed

as the product of the number n_{ps} of injuries attributed to precursor_p for the severity level s (given by Table II) and the impact rating S_s of this severity level (given by Table II, and based on Hallowell and Gambatese⁽¹⁶⁾). We considered five severity levels, s_1 = Pain, s_2 = First Aid, s_3 = Medical Case/Lost Work Time, s_4 = Permanent Disablement, and s_5 = Fatality. Medical Case and Lost Work Time were merged

because differentiating between these two severity levels was not possible based only on the information available in the narratives and associated databases. Equation (1) shows construction safety risk.

$$\text{risk} = \text{frequency} \times \text{severity}. \quad (1)$$

The total amount of risk that can be attributed to precursor_p was then obtained by summing the risk values attributed to this precursor across all severity levels, as shown in Equation (2):

$$R_p = \sum_{s=1}^5 (n_{ps} S_s), \quad (2)$$

where n_{ps} is the number of injuries of severity level s attributed to precursor_p, and S_s is the impact score of severity level s .

Finally, as noted by Sacks *et al.*,⁽²⁵⁾ risk analysis is inadequate if the likelihood of worker exposure to specific hazards is not considered. Hence, the risk R_p of precursor_p was weighted by its probability of occurrence e_p (see Equation (3)), which gave the relative risk RR_p of precursor_p. The probabilities e_p , or exposure values, were provided by the same company that donated the injury reports. These data are constantly being recorded by means of observation as part of the firm's project control and work characterization policy and therefore were already available.

$$RR_p = \frac{1}{e_p} \times R_p = \frac{1}{e_p} \sum_{s=1}^5 (n_{ps} S_s), \quad (3)$$

where R_p is the total amount of risk associated with precursor_p, and e_p is the probability of occurrence of precursor_p onsite.

To illustrate the notion of relative risk, assume that the precursor lumber has caused 15 first aid injuries, 10 medical cases and lost work time injuries, and has once caused a permanent disablement. By following the steps outlined above, the total amount of risk R_{lumber} accounted for by the attribute lumber can be computed as $15 \times 48 + 10 \times 192 + 1 \times 1,024 = 3,664$. Moreover, if lumber is encountered frequently onsite, e.g., with an exposure value $e_{\text{lumber}} = 0.65$, the relative risk of lumber will be $RR_{\text{lumber}} = 3,664/0.65 = 5,637$. However, if workers are very seldom exposed to lumber (e.g., $e_{\text{lumber}} = 0.07$), RR_{lumber} will be equal to $3,664/0.07 = 52,343$. It is clear from this example that if two attributes have the same total risk value, the attribute having the lowest exposure value will be associated with the greatest relative

risk. The assumption is that if a rare attribute causes as much damage as a more common one, the rare attribute should be considered riskier by proportion. Note that relative risk values allow comparison but do not have an absolute physical meaning. As presented later, what matters more than the precise risk value itself is the range in which a value falls.

Also, note that since Tixier *et al.*'s⁽⁵⁾ NLP tool's functionality did not include injury severity extraction at the time of writing, we used the real and worst possible outcomes manually assessed for each report by Prades.⁽⁹⁾ Specifically, in Prades,⁽⁹⁾ a team of seven researchers analyzed a large database of injury reports over the course of several months. High output quality was ensured by using a harsh 95% intercoder agreement threshold, peer reviews, calibration meetings, and random verifications by an external reviewer. Regarding worst possible injury severity, human coders were asked to use their judgment of what would have happened in the worst-case scenario should a small translation in time and/or space had occurred. This method and the resulting judgments were later validated by Alexander *et al.*,⁽²⁶⁾ who showed that the human assessment of maximum possible severity was congruent with the quantity of energy in the situation, which, ultimately, is a reliable predictor of the worst possible outcome.

For instance, in the following excerpt of an injury report: "worker was welding below scaffold and a hammer fell from two levels above and scratched his arm," the real severity is a first aid. However, by making only a small translation in space, the hammer could have struck the worker in the head, which could have yielded a permanent disablement or even a fatality. Coders in Prades⁽⁹⁾ were asked to favor the most conservative choice. Thus, in this case, permanent disablement was selected. Whenever mental projection was impossible or required some degree of speculation, coders were required to leave the field blank and the reports were subsequently discarded. As indicated, these subjective assessments were empirically validated.⁽²⁶⁾

By considering severity counts for both real outcomes and worst possible outcomes, we could compute two relative risk values for each of the 77 precursors. These values are listed in Table III, and were stored in two vectors of length $P = 77$.

For each attribute, we computed the difference between the relative risk based on worst possible outcomes and the relative risk based on actual outcomes. The top 10% attributes for this metric are

hazardous substance ($\Delta = 6,059$), machinery (3,092), improper security of materials (930), lifting/pulling/manual handling (390), unpowered transporter (378), forklift (371), unpowered tool (328), improper security of tools (302), and heavy vehicle (295). These attributes can be considered as the ones giving a construction situation the greatest potential for severity escalation in the worst-case scenario. Except lifting/pulling/manual handling and unpowered tool, all these precursors are directly associated with human error or high energy levels, which corroborates recent findings.⁽²⁶⁾ Furthermore, one could argue that the attributes lifting/pulling/manual handling and unpowered tool are still related to human error and high energy levels, as the former is often associated with improper body positioning (human factor) whereas the latter usually designates small and hand-held objects (hammer, wrench, screwdriver, etc.) that are prone to falling from height (high energy). Many attributes in Table III, such as sharp edge, manlift, unstable support/surface, or improper body position, have low risk values because of their rarity in the rather small data set that we used to provide a proof of concept for our methodology, but this does not incur any loss of generality.

3.2. Report-Level Safety Risk

As shown in Equation (4), we define safety risk at the situational level as the sum of the risk values of all the attributes that were identified as present in the corresponding injury report.

$$R_{\text{report}_r} = \sum_{p=1}^P (RR_p \cdot \delta_{rp}), \quad (4)$$

where RR_p is the relative risk associated with precursor_p, and $\delta_{rp} = 1$ if precursor_p is present in report_r ($\delta_{rp} = 0$ else).

In practice, computing real (or worst) safety risk at the report level comes down to multiplying the (R, P) attribute binary matrix (attribute matrix of Fig. 1) by the (P, 1) relative real (or worst) risk vector as shown in Equation (5). In the end, two risk values (real and worst) were obtained for each of the $R = 814$ incident reports.

For instance, in the following description of a construction situation: “worker is unloading a ladder from pickup truck with bad posture,” four attributes are present: namely (1) ladder, (2) manual handling,

(3) light vehicle, and (4) improper body positioning. The risk based on real outcomes for this construction situation is the sum of the relative risk values of the four attributes present (given by Table III), that is, $15 + 49 + 7 + 3 = 74$, and similarly, the risk based on worst potential outcomes is $52 + 439 + 123 + 6 = 620$. As already stressed, these relative values are not meaningful in absolute terms, they only enable comparison between situations and their categorization into broad ranges of riskiness (e.g., low, medium, high). Estimating these ranges on a small, finite sample such as the one we used in this study would have resulted in biased estimates. To alleviate this, we used stochastic simulation techniques to generate hundreds of thousands of new scenarios honoring the historical data, enabling us to make inferences from a much richer, yet faithful sample.

$$\begin{bmatrix} 0 & 1 & 0 & \cdots & 1 & 0 & 1 \\ 0 & 1 & 0 & \cdots & 1 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 1 & 1 & \cdots & 0 & 1 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 1 & 0 & \cdots & 0 & 1 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 1 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} RR_1 \\ RR_2 \\ RR_3 \\ \vdots \\ RR_{(P-2)} \\ RR_{(P-1)} \\ RR_P \end{bmatrix} = \begin{bmatrix} R_{\text{report}_1} \\ R_{\text{report}_2} \\ R_{\text{report}_3} \\ R_{\text{report}_4} \\ \vdots \\ R_{\text{report}_{(R-3)}} \\ R_{\text{report}_{(R-2)}} \\ R_{\text{report}_{(R-1)}} \\ R_{\text{report}_R} \end{bmatrix} \quad (5)$$

Multiplying the (R, P) attribute matrix by the (P, 1) vector of relative risk values for each attribute gives the (R, 1) vector of risk values associated with each injury report.

3.3. The Probability Distribution of Construction Safety Risk Resembles That of Many Natural Phenomena

For a given injury report, the risk based on real outcomes and the risk based on worst potential outcomes can each take on a quasi-infinite number of values ($2^P - 1$) with some associated probabilities. Therefore, they can be considered quasi-continuous random variables, and have legitimate probability distribution functions (PDFs). Furthermore, since a risk value cannot be negative by definition, these PDFs have $[0, +\infty[$ support.

The empirical PDF of the risk based on real outcomes for the 814 injury reports is shown as a histogram in Fig. 2. The histogram divides the sample

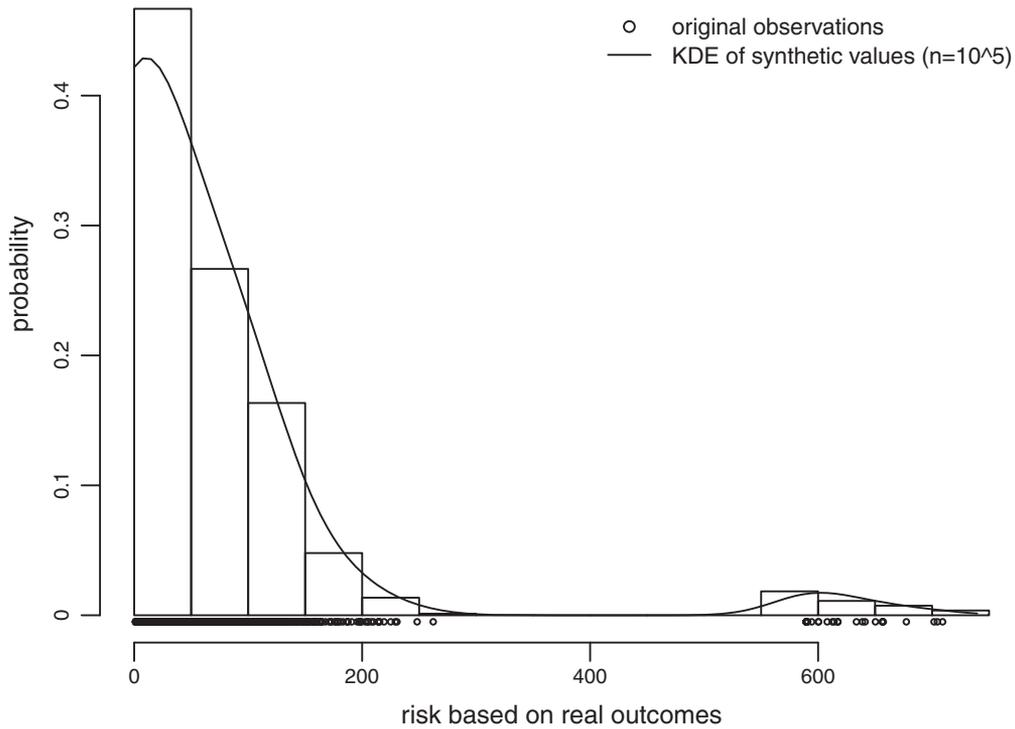


Fig. 2. Histogram of original observations ($n = 814$) with boundary corrected KDE of the simulated observations ($n = 10^5$).

space into a number of intervals and simply counts how many observations fall into each range. We can clearly see that the empirical safety risk is rightly skewed and exhibits a thick tail feature. In other words, the bulk of construction situations present risk values in the small-medium range, whereas only a few situations are associated with high and extreme risk. This makes intuitive sense and is in accordance with what is observed onsite, i.e., frequent benign injuries, and low-frequency high-impact accidents.

Such heavy-tailed distributions are referred to as “power laws” in the literature, after Pareto,⁽²⁷⁾ who proposed that the relative number of individuals with an annual income larger than a certain threshold was proportional to a power of this threshold. Power laws are ubiquitous in nature.^(28,29) Some examples of natural phenomena whose magnitude follow power laws include earthquakes, ocean waves, volcanic eruptions, asteroid impacts, tornadoes, forest fires, floods, solar flares, landslides, and rainfall.^(28,30–32) Other human-related examples include insurance losses and health-care expenditures,⁽³³⁾ hurricane damage,^(34,35) and the size of human settlements and of files transferred on the web.^(36,37)

To highlight the resemblance between construction safety risk and some of the aforementioned

natural phenomena, we selected four data sets that are standard in the field of extreme value analysis, and freely available from the “extRemes” R package.⁽³⁸⁾ We overlaid the corresponding PDFs with that of construction safety risk. For the sake of comparison, variables were first rescaled as shown in Equation (6). The output can be seen in Fig. 3. In what follows, each data set is briefly presented.

$$Z = \frac{X - \min(X)}{\max(X) - \min(X)} \quad (6)$$

where X is the variable in the original space and Z is the variable in the rescaled space.

The first data set reported summer maximum temperatures in Phoenix, AZ, from 1948 to 1990, measured at Sky Harbor Airport. The observations were multiplied by -1 (flipped horizontally) before rescaling. The distribution is named “max temperature” in Fig. 3. The second data set (“hurricane damage” in Fig. 3) consisted of total economic damage caused by every hurricane making landfall in the United States between 1925 and 1995, expressed in 1995 U.S. \$ billion. All individual storms costing less than \$0.01 billion were removed to minimize potential biases in the recording process. The final number

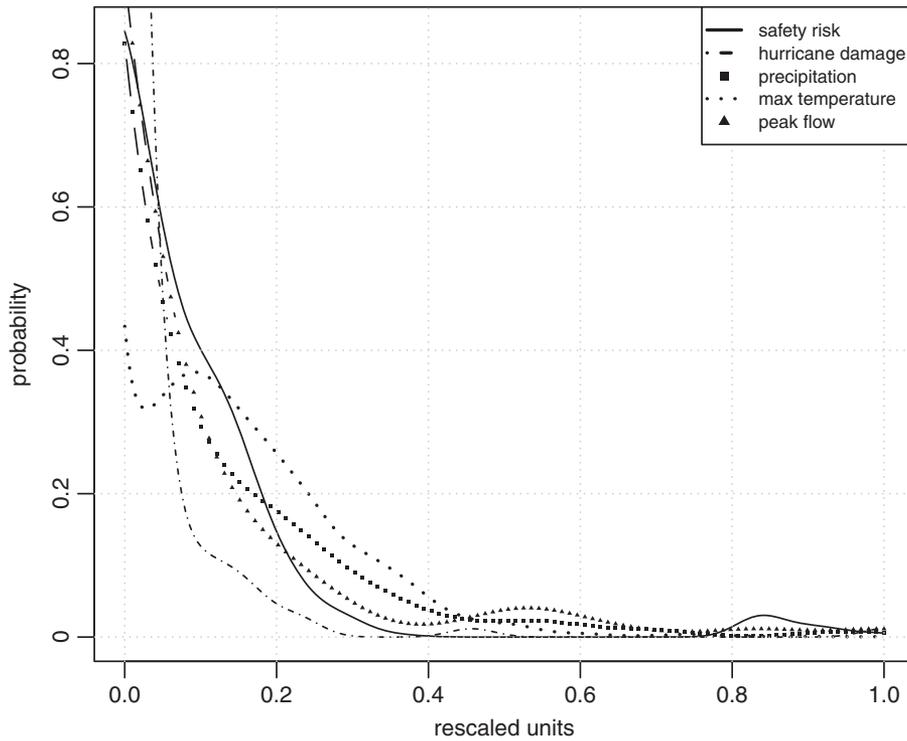


Fig. 3. Safety risk compared to natural phenomena.

of hurricanes taken into account was 86. The third data set included in our comparison was observations of Potomac River peak stream flow measured in cubic feet per second at Point Rocks, MD, from 1895 to 2000. The observations were divided by 10^5 before rescaling. The curve is labeled “peak flow” in Fig. 3. The fourth and last data set contained 36,524 daily precipitation amounts (in inches) from a single rain gauge in Fort Collins, CO. Only values greater than 1 inch were taken into account, giving a final number of 213 observations. The distribution is named “precipitation” in Fig. 3.

We estimated the PDFs shown in Fig. 3 via KDE because overlaying histograms would have resulted in an incomprehensible figure. The KDE is a non-parametric way to estimate a PDF. It can be viewed as a smoothed version of the histogram, where a continuous function, called the kernel, is used rather than a box as the fundamental constituent.⁽³⁹⁾ The kernel has zero mean, is symmetric, positive, and integrates to one. The last two properties ensure that the kernel, and as a result the KDE, is a probability distribution. More precisely, as shown in Equation (7), the KDE at each point x is the sum of the

weighted contributions from all the observations to the point x , the weights being assigned by the kernel function.^(39,40)

Put differently, the KDE at x is a local average of functions assigning weights to the neighboring observations x_i that decrease as $|x_i - x|$ increases.^(41,42) The “local” estimation is the key feature of this method in enabling to capture the features present in the data. KDEs converge faster to the underlying density than the histogram, and are robust to the choice of the origin of the intervals.⁽⁴²⁾

$$\widehat{f}_X(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}, h\right), \quad (7)$$

where $\{x_1, \dots, x_n\}$ are the observations, K is the kernel and h is a parameter called the bandwidth. Note that \widehat{f}_X is an estimator of the true unknown PDF.

h is a parameter called the bandwidth that controls the degree of smoothing and therefore affects the final shape of the estimate.⁽⁴⁰⁾ In this study, we used a standard and widespread way of estimating h called Silverman’s rule of thumb⁽³⁹⁾ and shown in Equation (8). We invite the reader to reference Rajagopalan *et al.*⁽⁴³⁾ for a good review of the

objective bandwidth selection methods.

$$h = \frac{0.9 \min(\hat{\sigma}_X, \frac{Q_3 - Q_1}{1.34})}{n^{1/5}}, \quad (8)$$

where Q_3 and Q_1 are the third and first quartiles (respectively), $\hat{\sigma}_X$ is the standard deviation of the sample, and n is the size of the sample. Here, $n = R = 814$.

Further, for our kernel K , we selected the standard normal distribution $N(0, 1)$, that is, the normal distribution centered on zero with unit variance. Because the PDF of $N(0,1)$ is $\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, the associated KDE can be written using Equation (7) as shown in Equation (9). Other popular kernels include the triangular, biweight, or Epanechnikov, but the consensus in the statistics literature is that the choice of the kernel is secondary to the estimation of the bandwidth.⁽⁴¹⁾

$$\widehat{f}_X(x) = \frac{1}{n(2h^2\pi)^{p/2}} \sum_{i=1}^n e^{-\frac{1}{2}(x_i - x/h)^2} \quad (9)$$

where $\{x_1, \dots, x_n\}$ are the observations, and h is the bandwidth. Here, $n = R = 814$.

It is well known that the KDE suffers a bias at the edges on bounded supports. Indeed, because the kernel functions are symmetric, weights are assigned to values outside the support, which causes the density near the edges to be significantly underestimated, and creates a faulty visual representation. In our case, safety risk takes on values in $[0, +\infty[$, so issues arise when approaching zero. We used the correction for the boundary bias via local linear regression⁽⁴⁴⁾ using the “evmix” package⁽⁴⁵⁾ of the R programming language.⁽⁴⁶⁾ Log transformation and boundary reflection are other popular approaches for controlling boundary bias.^(39,43)

3.4. Univariate Safety Risk Stochastic Generator

In this section, we present a computational method that can be used to generate synthetic safety risk values that honor the historical data. We focus on the risk based on real outcomes, but the same methodology can be used to simulate from any distribution. Note that although many techniques and concepts in risk modeling and management deal with extreme values only, here we seek to capture and simulate the entire risk spectrum (not only the extremes) in order to accurately and comprehensively assess the safety risk of any construction situation.

The quantile function (or simply quantile, for short) of a continuous random variable X is defined as the inverse of its cumulative distribution function (CDF) as shown in Equation (10). The CDF is obtained by integrating or summing the PDF, respectively, in the continuous and discrete case.

$$Q(p) = F_X^{-1}(p), \quad (10)$$

where F_X is the CDF of X defined as $F_X(x) = P[X \leq x] = p \in [0, 1]$

The quantile is closely linked to the concept of exceedances. In finance and insurance, for instance, the value-at-risk for a given horizon is the loss that cannot be exceeded with a certain probability of confidence within the time period considered, which is given by the quantile. For instance, the 99.95% value-at-risk $Q(99.95)$ at 10 days represents the amount of money that the loss can only exceed with 0.05% probability in the next 10 days. In other words, the corresponding fund reserve would cover 199 losses over 200 ($199/200 = 0.995$).

The quantile is also associated with the notion of return period T in hydroclimatology. For example, the magnitude of the 100-year flood ($T = 100$) corresponds to the streamflow value that is only exceeded on average by 1% of the observations, assuming one observation per year. This value is given by $Q(1 - 1/T) = Q(0.99)$, which is the 99th percentile, or the 99th 100-quantile. Similarly, the magnitude of the 500-year flood, $Q(0.998)$, is only exceeded on average by 0.2% of the observations. For construction safety, this quantity would correspond to the minimum risk value that is only observed on average in one construction situation over five hundred. The median value, given by $Q(0.5)$, would correspond to the safety risk observed on average in one construction situation over two.

3.4.1. Limitations of Traditional Parametric Techniques

Traditional approaches to quantile estimation are based on parametric models of PDF especially from the extreme value theory (EVT) framework.⁽⁴⁷⁾ These models possess fat tails unlike traditional PDFs, and thus are suitable for robust estimation of extremes. The candidate distributions from the EVT are Frechet, Weibull, Gumbel, GEV, generalized Pareto, or mixtures of these distributions.⁽⁴⁸⁾ These parametric models are powerful in that they allow complex phenomena to be entirely described by a single mathematical equation and a few

parameters. However, being parametric, these models tend to be suboptimal when little knowledge is available about the phenomenon studied^(48,49) and though they are heavy-tailed, they still are prone to underestimating the extreme quantiles.⁽⁵⁰⁾ A popular remediation strategy consists in fitting a parametric model to the tail only, such as the generalized Pareto, but selecting a threshold that defines the tail is a highly subjective task,⁽⁵¹⁾ and medium and small values, which represent the bulk of the data, are often overlooked.⁽⁵⁰⁾ What is clearly better, however, especially when the final goal is simulation, is to capture the entire distribution. As a solution, hydroclimatologists have proposed dynamic mixtures of distributions, based on weighting the contributions of two overlapping models, one targeting the bulk of the observations, and the other orientated toward capturing extremes.^(52,53) Unfortunately, threshold selection implicitly carries over through the estimation of the parameters of the mixing function, and with most mixing functions, conflicts arise between the two distributions around the boundary.⁽⁴⁵⁾ For all these reasons, we decided to adopt a fully data-driven, nonparametric approach that we describe below.

3.4.2. Proposed Approach

Our methodology consists in generating independent realizations from the nonparametric PDF estimated via the KDE described above. We base our generator on the smoothed bootstrap with variance correction proposed by Silverman.⁽³⁹⁾ Unlike the traditional nonparametric bootstrap⁽⁵⁴⁾ that simply consists in resampling with replacement, the smoothed bootstrap can generate values outside of the historical limited range, and does not reproduce spurious features of the original data such as noise.⁽⁵⁵⁾ The smoothed bootstrap approach has been successfully used in modeling daily precipitation,⁽⁵⁶⁾ streamflow,⁽⁵⁷⁾ and daily weather.⁽⁵⁵⁾ Kernel quantile function estimators⁽⁵⁸⁾ and local polynomial-based estimators⁽⁵⁹⁾ are other attractive options. Here, we propose simulation from *smoothed bootstrap*, which is easier to implement and competitive to other methods.

The algorithm we used to generate our synthetic values according to the smoothed bootstrap scheme can be broken down into the following steps:

For j in 1 to the desired number of simulated values:

- (1) choose i uniformly with replacement from $\{1, \dots, R\}$;

- (2) sample ϵ_X from the standard normal distribution with variance h_{X^2} ;
- (3) record $X_{simj} = \bar{X} + (X_i - \bar{X} + \epsilon_X) / \sqrt{1 + h_{X^2} / \hat{\sigma}_X^2}$.

where $R = 814$ is the sample size (the number of injury reports), \bar{X} and $\hat{\sigma}_X^2$ are the sample mean and variance, and h_{X^2} is the variance of the standard normal kernel (bandwidth of the KDE). Note that we deleted the negative simulated values to be consistent with the definition of risk.

Fig. 2 shows the KDE of 10^5 simulated values overlaid with the histogram of the original sample. It can be clearly seen that the synthetic values are faithful to the original sample, as the PDF from the simulated values fits the observations very well. Also, while honoring the historical data, the smoothed bootstrap generated values outside the original range, as desired. The maximum risk value in our sample was 709, whereas the maximum of the simulated values was 740 (rounded to the nearest integer). Table IV compares the quantile estimated via the `quantile()` R function of the original and simulated observations.

The quantile estimates of Table IV are roughly equivalent before reaching the tails. This is because the bulk of the original observations were in the low to medium range, enabling quite accurate quantile estimates for this range in the first place. The problem stemmed from the sparsity of the high to extreme values in the historical sample, which made estimation of the extreme quantiles biased. Our use of the smoothed bootstrap populated the tail space with new observations, yielding a slightly higher estimate of the extreme quantiles, as can be seen in Table IV. It makes sense that the extremes are higher than what could have been inferred based simply on the original sample, as the original sample can be seen as a finite window in time whereas our simulated values correspond to observations that would have been recorded over a much longer period. The chance of observing extreme events is of course greater over a longer period of time. Based on estimating the quantiles on the extended time frame represented by the synthetic values, we propose the risk ranges shown in Table V. As already explained, these ranges are more robust and unbiased than the ones that would have been built from our historical observations only. Thanks to this empirical way of assessing safety risk, construction practitioners will be able to adopt an optimal proactive approach by taking coherent preventive

Table IV. Quantile Estimates Based on Original and Simulated Values for the Risk Based on Real Outcomes

	Safety Risk Observed in One Situation Over						
	2	5	10	100	500	1,000	10,000
Original observations ($n = R = 814$)	57	110	152	649	703	706	709
Simulated observations ($n = 10^5$)	61	116	154	647	700	708	728

Table V. Proposed Ranges for the Risk Based on Real Outcomes

quantiles	0	0.25	0.50	0.75	0.99	1
risk value	0	29	61	105	647	740
range	<i>low</i>	<i>medium</i>	<i>high</i>	<i>very high</i>	<i>extreme</i>	

actions and provisioning the right amounts of resources.

4. BIVARIATE ANALYSIS

In what follows, we study the relationship between the risk based on real outcomes (X , for brevity) and the risk based on worst potential outcomes (Y). Rather than considering these two random variables in separation, we acknowledge their dependence and aim at capturing it, and faithfully reproducing it in our simulated observations. This serves the final goal of being able to accurately assess the potential of an observed construction situation for safety risk escalation should the worst-case scenario occur. Fig. 4 shows a plot of Y versus X , whereas a bivariate histogram can be seen in Fig. 5.

We can distinguish three distinct regimes in Fig. 4. The first regime, corresponding roughly to $0 < X < 70$, is that of benign situations that stay benign in the worst case. Under this regime, there is limited potential for risk escalation. The second regime ($70 < X < 300$) shows that beyond a certain threshold, moderately risky situations can give birth to hazardous situations in the worst case. The attribute responsible for the switch into this second regime is machinery (e.g., welding machine, generator, pump). The last regime ($X > 300$) is that of the extremes, and features clear and strong upper tail dependence. The situations belonging to this regime are hazardous in their essence and create severe outcomes in the worst-case scenarios. In other words, those situations are dangerous in the first place and unforgiving. The attribute responsible for this extreme

regime is hazardous substance (e.g., corrosives, adhesives, flammables, asphyxiants). Again, note that these examples are provided as a result of applying our methodology on a data set of 814 injury reports for illustration purposes but do not incur any loss of generality. Using other, larger data sets would allow risk regimes to be characterized by different and possibly more complex attribute patterns.

4.1. Copula Theory

Many natural and human-related phenomena are multifactorial by nature and as such their study requires the joint modeling of several random variables. Traditional approaches consist in modeling dependence with the classical family of multivariate distributions, which is clearly limiting, as it requires all variables to be separately characterized by the same univariate distributions (called the margins). Copula theory addresses this limitation by creating a joint probability distribution for two or more variables while preserving their original margins.⁽⁶⁰⁾ In addition to the extra flexibility they offer, the many existing parametric copula models are also attractive in that they can model the dependence among a potentially very large set of random variables in a parsimonious manner. For an overview of copulas, one may refer to Cherubini *et al.*⁽⁶¹⁾

Although the introduction of copulas can be tracked back as early as 1959 with the work of Sklar, they did not gain popularity until the end of the 1990s when they became widely used in finance. Copulas are now indispensable to modeling multivariate dependence,⁽⁶²⁾ and are used in various fields from cosmology to medicine. Because many hydroclimatological phenomena are multidimensional, copulas are also increasingly used in hydrology, weather, and climate research, for instance, for precipitation infilling and extreme storm tide modeling.^(63–65)

Formally, a d -dimensional copula is a joint CDF with $[0, 1]^d$ support and standard uniform margins.⁽⁶⁶⁾ Another equivalent definition is given by

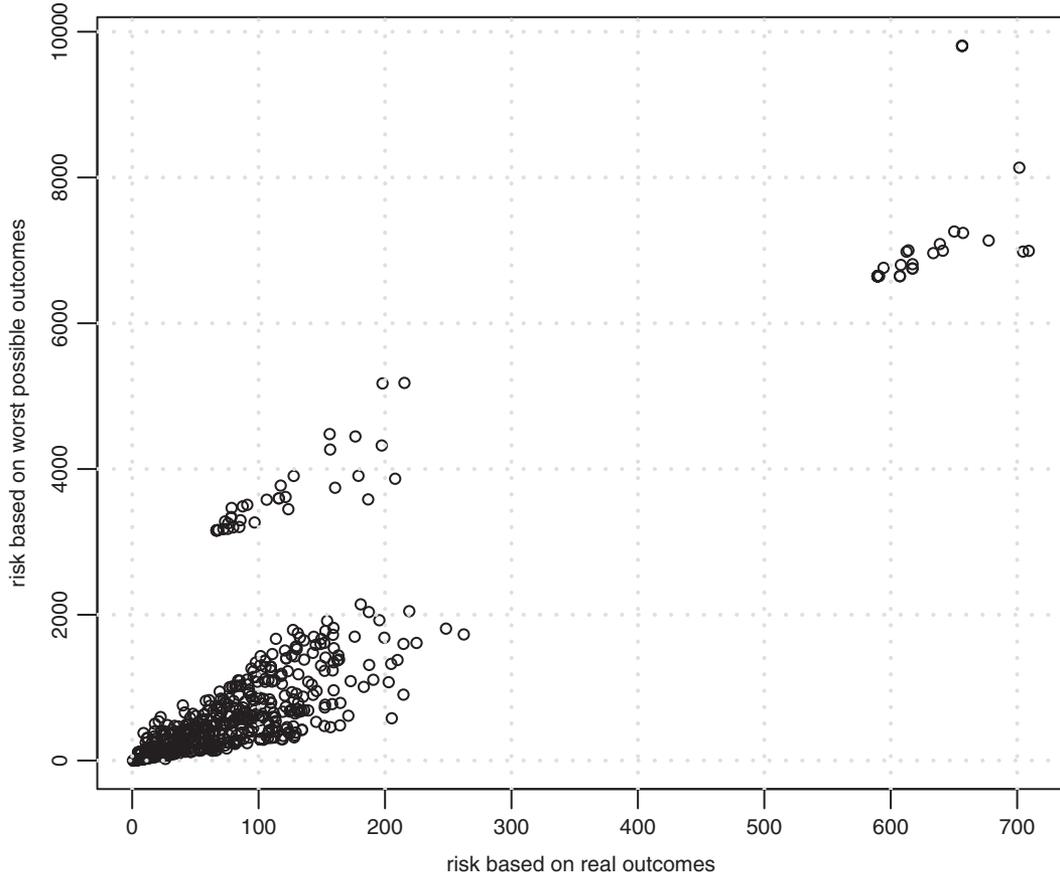


Fig. 4. Bivariate construction safety risk.

Sklar's theorem,⁽⁶⁷⁾ which states in the bivariate case that the joint CDF $F(x, y)$ of any pair (X, Y) of continuous random variables can be written in terms of a copula as shown in Equation (11).

$$F(x, y) = C\{F_X(x), F_Y(y)\}, (x, y) \in \mathbb{R}^2, \quad (11)$$

where F_X and F_Y are the respective margins of X and Y , and $C: [0, 1]^2 \rightarrow [0, 1]$ is a copula.

Note that Sklar's theorem is consistent with the first definition given because for any continuous random variable X of CDF F_X , $F_X(X)$ follows a uniform distribution (a result known as the probability integral transform).

Parametric copulas suffer from all the limitations inherent to univariate parametric models evoked previously. Therefore, like in the univariate case, we decided to adopt an empirical, nonparametric approach to copula density estimation. We used the bivariate KDE to estimate the empirical copula, which is defined as the joint CDF of the rank-transformed

(or pseudo) observations. The pseudo-observations are obtained as shown in Equation (12).

$$U_X(x) = \frac{\text{rank}(x)}{\text{length}(X) + 1}, \quad (12)$$

where U_X is the transformed sample of the pseudo-observations, and X is the original sample.

Because the copula support is the unit square $[0, 1]^2$, the KDE boundary issue arises twice this time, near zero and one, yielding multiplicative bias.⁽⁶⁸⁾ Therefore, the density is even more severely underestimated than in the univariate case, and it is even more crucial to ensure robustness of the KDE at the corners to ensure proper visualization. We used the transformation-based approach described by Charpentier *et al.*⁽⁶⁸⁾ as our boundary correction technique, using the inverse CDF of the normal distribution, $F_{N(0,1)}^{-1}$, as the transformation T . The resulting empirical copula density estimate of the original sample is shown in Fig. 6, and can be seen to capture the data very well.

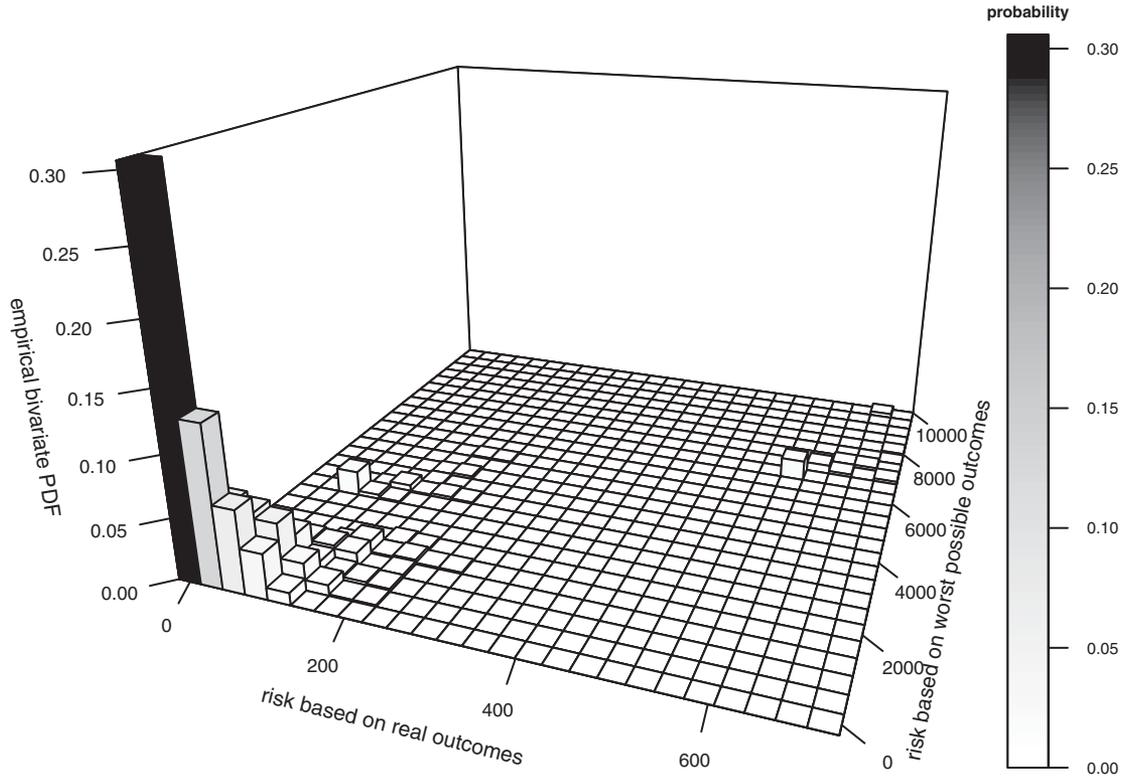


Fig. 5. Bivariate histogram with 25×25 grid.

4.2. Bivariate Safety Risk Stochastic Generator

Like in the univariate case, we used a nonparametric, fully data-driven approach, the smoothed bootstrap with variance correction, as our simulation scheme. Minor adaptations were needed due to the two-dimensional nature of the task. The steps of our algorithm are outlined below, and the resulting 10^5 simulated values are shown in Fig. 7. Note that the procedure is equivalent to simulating from the nonparametric copula density estimate introduced above. Like in the univariate case, we deleted the negative simulated values to ensure consistency with the definition of risk.

For j in 1 to the desired number of simulated values:

- (1) choose i uniformly with replacement from $\{1, \dots, R\}$;
- (2) sample ϵ_X from the standard normal distribution with variance h_X^2 , and ϵ_Y from the standard normal distribution with variance h_Y^2 ;
- (3) take:

$$X_{simj} = \bar{X} + (X_i - \bar{X} + \epsilon_X) / \sqrt{1 + h_X^2 / \hat{\sigma}_X^2}$$

$$Y_{simj} = \bar{Y} + (Y_i - \bar{Y} + \epsilon_Y) / \sqrt{1 + h_Y^2 / \hat{\sigma}_Y^2}$$

- (4) record: $U_{simj} = F_{N(0,1)}(X_{simj})$, $V_{simj} = F_{N(0,1)}(Y_{simj})$.

where $R = 814$ is the number of injury reports, \bar{X} and $\hat{\sigma}_X^2$ are the mean and variance of X ; \bar{Y} and $\hat{\sigma}_Y^2$ are the mean and variance of Y ; h_X^2 is the bandwidth of the KDE of X ; h_Y^2 is the bandwidth of the KDE of Y ; and $F_{N(0,1)}$ is the CDF of the standard normal distribution, that is, the inverse of our transformation T .

Note that step 1 selects a pair (x, y) of values from the original sample (X, Y) , not two values independently. This is crucial in ensuring that the dependence structure is preserved. Step 4 sends the simulated pair to the pseudo space to enable visual comparison with the empirical copula density estimate, which is defined in the rank space (i.e., the unit square). We can clearly observe in Fig. 7 that our sampling scheme was successful in generating values that reproduce the structure present in the original data, validating our nonparametric approach. For the sake of completeness, we also compared (see

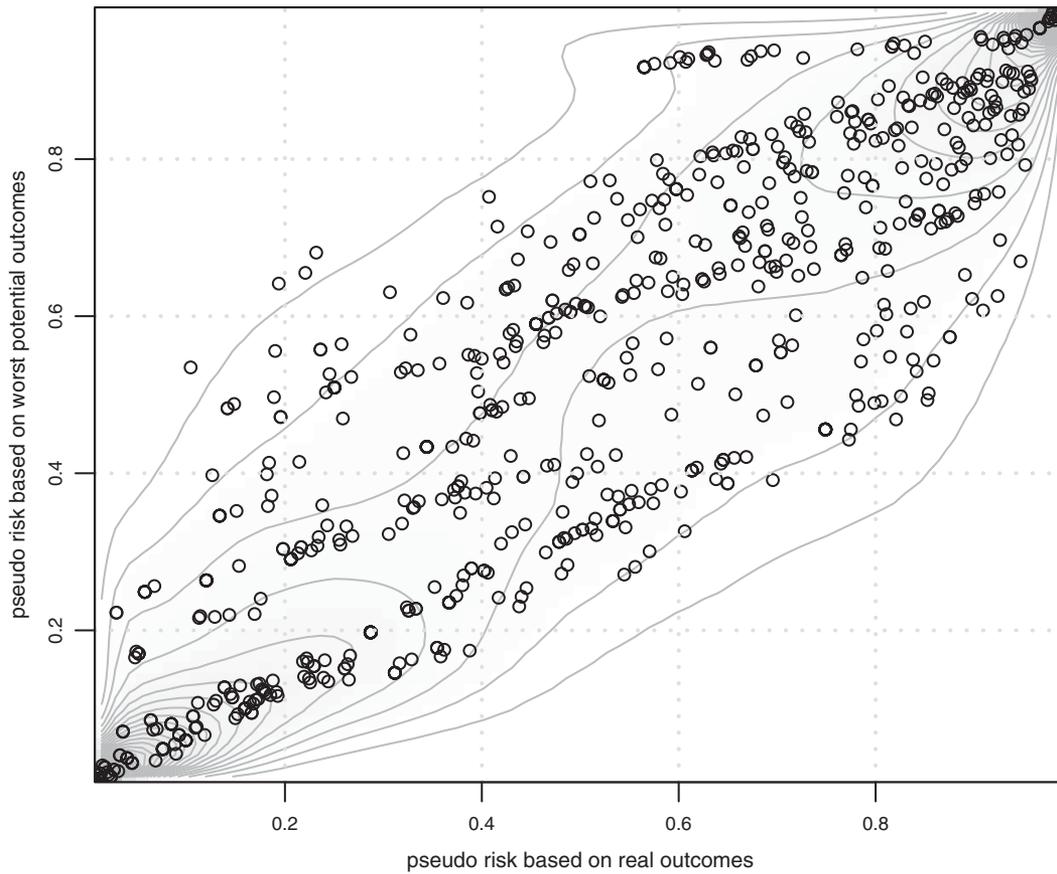


Fig. 6. Nonparametric copula density estimate with original pseudo-observations.

Table VI. Quantile Estimates Based on Original and Simulated Values for the Risk Based on Worst Potential Outcomes

	Safety Risk Observed in One Situation Over						
	2	5	10	100	500	1,000	10,000
Original observations ($n = R = 814$)	343	950	1,719	7,000	9,808	9,808	9,808
Simulated observations ($n = 10^5$)	395	1,061	1,953	7,092	9,765	9,586	10,045

Fig. 8 and 9) the simulated pairs in the original space with the original values. Once again, it is easy to see that the synthetic values honor the historical data. To enable comparison with the univariate case (see Table IV), Table VI summarizes the empirical quantiles for the historical and simulated observations of risk based on worst potential outcomes (i.e., Y). Like in the univariate case, we can observe that the differences between the estimates increase with the quantiles. Notably, simulation allows to obtain richer estimates of the extreme quantiles, $Q(1 - \frac{1}{1000}) = Q(0.999)$ and $Q(1 - \frac{1}{10000}) = Q(0.9999)$, whereas

Table VII. Proposed Ranges for the Risk Based on Worst Potential Outcomes

quantiles	0	0.25	0.50	0.75	0.99	1
risk value	0	183	395	837	7092	10126
range		<i>low</i>	<i>medium</i>	<i>high</i>	<i>very high</i>	<i>extreme</i>

with the initial limited sample, the values of the quantile function plateau after $Q(1 - \frac{1}{500}) = Q(0.998)$ due to data sparsity in the tail. Similarly to Table V, we also propose in Table VII ranges for the risk based on worst potential outcomes.

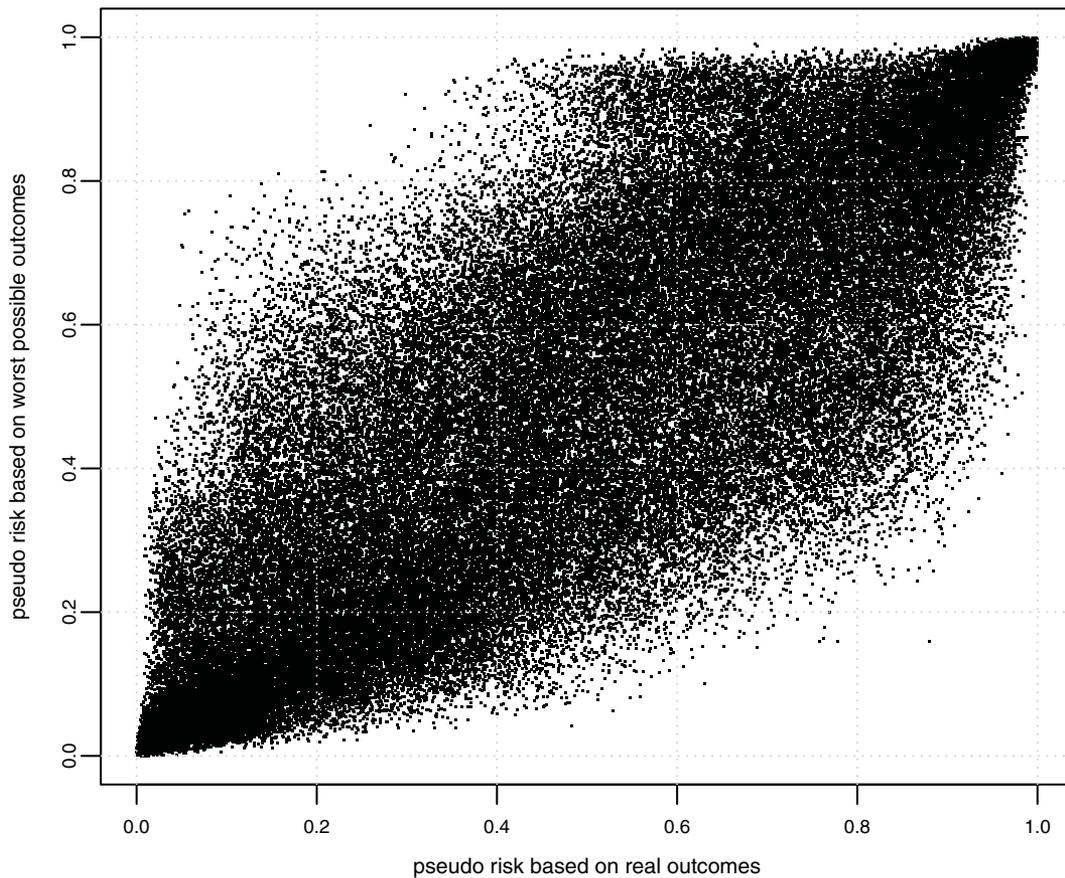


Fig. 7. Simulated values $n = 10^5$.

4.3. Computing Risk Escalation Potential Based on Simulated Values

Using the synthetic safety risk pairs obtained via our bivariate stochastic safety risk generator, and evidence provided by the user (i.e., an observation made onsite in terms of attributes), it is possible to compute and estimate the upper limit of risk, i.e., the safety risk presented by the observed construction situation based on worst-case scenarios. This estimate is based on large numbers of values simulated in a data-driven approach that features the same dependence structure as the original, empirical data. The technique we propose, based on conditional quantile estimation, consists in the steps detailed in what follows.

First, the attributes observed in a particular construction situation give the risk based on real outcomes for the construction situation, say x_0 . By fixing the value of X to x_0 , this first step extracts a slice from the empirical bivariate distribution of the simulated values. This slice corresponds to the

empirical probability distribution of Y conditional on the value of X , also noted $P[Y|X = x_0]$. Because only a few values of Y may exactly be associated with x_0 , we consider all the values of Y associated with the values of X in a small neighboring range around x_0 , that is, $P[Y|x_0 - x_- < X < x_0 + x_+]$. In our experiments, we used $x_- = x_+ = 5$; that is, a range of $[-5, +5]$ around x_0 , because it gave good results, but there is no absolute and definitive best range. The second step simply consists in evaluating the quantile function of $P[Y|x_0 - x_- < X < x_0 + x_+]$ at some threshold. The reader can refer to Equation (10) for the definition of the quantile function. In our experiments, we used a threshold of 80%, (i.e., we computed $Q(0.8)$ with the `quantile()` R function), but the choice of the threshold should be made at the discretion of the user, depending on the desired final interpretation. In plain English, the threshold we selected returns the risk based on worst possible outcomes that is only exceeded in 20% of cases for

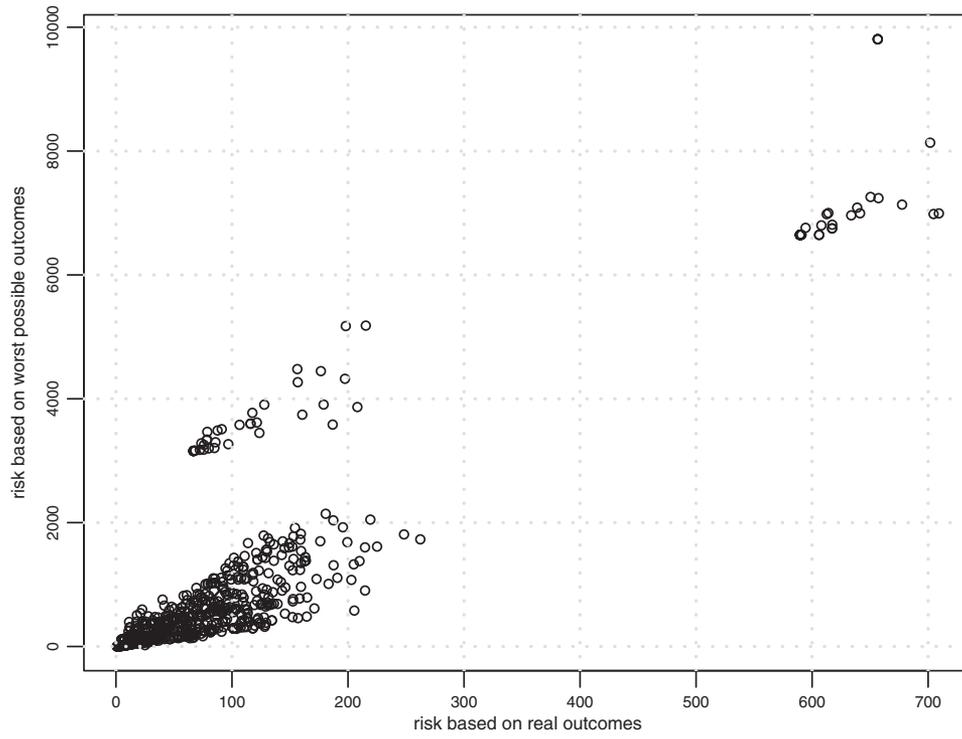


Fig. 8. Bivariate construction safety risk.

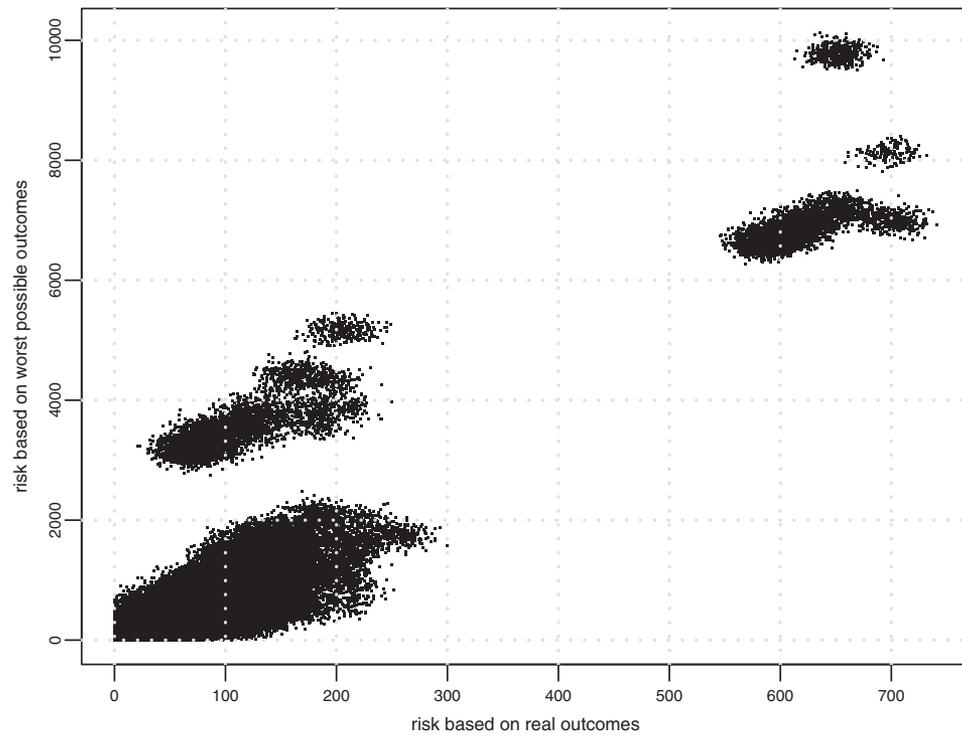


Fig. 9. Simulated values $n = 10^5$.

Table VIII. Illustration of the Proposed Risk Escalation Estimation Technique

Attributes	Step 1: Prior Evidence		Step 2: Conditional Quantile Estimate	
	Risk Based on Real Outcomes (x_0) and Associated Range ^a		Estimate Q(0.8) of Risk Based on Worst Potential Outcomes and Associated Range ^b	
Hazardous substance, confined workspace	590 + 115 = 705	Extreme	7,266	Extreme
Hammer, lumber	5 + 53 = 58	Medium	676	High
Hand size pieces	7	Low	145	Low

^aBased on the ranges proposed in Table V.

^bBased on the ranges proposed in Table VII.

the particular value of risk based on real outcomes computed at the first step. Finally, by categorizing this value into the ranges of risk based on worst possible outcomes provided in Table VII, we are able to provide understandable and actionable insight with respect to the most likely risk escalation scenario.

These steps are illustrated for two simple construction situations in Table VIII. For comparison purposes, we also show the range of risk based on real outcomes (provided in Table V) in which x_0 falls.

5. LIMITATIONS

Given the data-driven nature of our approach, attribute risk values are expected to change from one injury report database to another, and from one set of exposure values to another, even though the distributions of safety risk based on real and worst potential outcomes are expected to remain the same (i.e., heavy-tailed). Also, in this study, we used a rather small data set (final size of 814 injury reports) to provide a proof of concept for our methodology. With larger data sets, more attributes would play a role in characterizing the different regimes presented in Fig. 4, and their respective signature would therefore enjoy a higher resolution.

6. CONCLUSION

We defined construction safety risk at the attribute and situational level, and showed its empirical probability distribution to be strikingly similar to that of many natural phenomena. We then proposed univariate and bivariate safety risk stochastic generators based on nonparametric density estimation techniques. The combination of kernels and copulas and the introduction of these methods for

modeling construction safety risk make a unique and pioneering contribution. It provides a powerful methodology to model and visualize bivariate safety risks, which are ubiquitous in construction and whose understanding is of paramount importance to safety performance improvement. Our approach can be used as a way to ground risk-based safety-related decisions under uncertainty on objective empirical data far exceeding the personal history of even the most experienced safety or project manager. Additionally, the combined use of the attribute-based framework and raw injury reports as the foundation of our work allows the user to escape the limitations of traditional construction safety risk analysis techniques that are segmented and rely on subjective data. Finally, the attribute-based nature of our methodology enables easy integration with building information modeling (BIM) and work packaging. We believe this study gives promising evidence that transitioning from an opinion-based and qualitative discipline to an objective, empirically-grounded quantitative science would be highly beneficial to construction safety research.

ACKNOWLEDGMENTS

We would like to thank the National Science Foundation for supporting this research through an Early Career Award (CAREER) Program. This material is based upon work supported by the National Science Foundation under Grant No. 1253179. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation. Sincere thanks also go to Prof. Arthur Charpentier for his kind help on nonparametric copula density estimation and on the

bivariate smoothed bootstrap, and to Prof. Carl Scarrott for his insight on dynamic mixture modeling.

REFERENCES

- Bureau of Labor Statistics. Occupational injuries/illnesses and fatal injuries profiles, 2015. Available at <http://www.bls.gov/news.release/pdf/cfoi.pdf>.
- Albert A, Hallowell MR, Kleiner B, Chen A, Golparvar-Fard M. Enhancing construction hazard recognition with high-fidelity augmented virtuality. *Journal of Construction Engineering and Management*, 2014; 140(7):04014024.
- Carter G, Smith SD. Safety hazard identification on construction projects. *Journal of Construction Engineering and Management*, 2006; 132(2):197–205.
- Kahneman D, Tversky A. On the study of statistical intuitions. *Cognition*, 1982; 11(2):123–141.
- Tixier AJP, Hallowell MR, Rajagopalan B, Bowman D. Automated content analysis for construction safety: A natural language processing system to extract precursors and outcomes from unstructured injury reports. *Automation in Construction*, 2016; 62:45–56.
- Tixier AJP, Hallowell MR, Rajagopalan B, Bowman D. Application of machine learning to construction injury prediction. *Automation in Construction*, 2016; 69:102–114.
- Esmaili B, Hallowell M. Attribute-based risk model for measuring safety risk of struck-by accidents. *Construction Research Congress*, 2012:289–298.
- Esmaili B, Hallowell MR. Using network analysis to model fall hazards on construction projects. *Saf Health Construct CIB W*, 2011; 99:24–26.
- Prades Villanova M. Attribute-based risk model for assessing risk to industrial construction tasks (master's thesis). University of Colorado at Boulder, 2014.
- Desvignes M. Requisite empirical risk data for integration of safety with advanced technologies and intelligent systems (master's thesis). University of Colorado at Boulder, 2014.
- Capen EC. The difficulty of assessing uncertainty. *SPE-5579-PA. JPT*, 1976; 28(8):843–50.
- Rose PR. Dealing with risk and uncertainty in exploration: How can we improve? *AAPG Bulletin*, 1987; 71(1):1–16.
- Tversky A, Kahneman D. The framing of decisions and the psychology of choice. *Science*, 1981; 211(4481):453–458. doi:10.1126/science.7455683.
- Gustafsson Per E. Gender differences in risk perception: Theoretical and methodological perspectives. *Risk Analysis*, 1998; 18(6):805–811. doi:10.1111/j.1539-6924.1998.tb01123.x.
- Tixier AJP, Hallowell MR, Albert A, van Boven L, Kleiner BM. Psychological antecedents of risk-taking behavior in construction. *Journal of Construction Engineering and Management*, 2014:140(11).
- Hallowell MR, Gambatese JA. Qualitative research: Application of the Delphi method to CEM research. *Journal of Construction Engineering and Management*, 2009; 136(1):99–107.
- Lingard H. Occupational health and safety in the construction industry. *Construct Manage Econ*, 2013; 31(6):505–514.
- Hallowell MR, Gambatese JA. Activity-based safety risk quantification for concrete formwork construction. *Journal of Construction Engineering and Management*, 2009; 135(10):990–998.
- Navon R, Kolton O. Model for automated monitoring of fall hazards in building construction. *Journal of Construction Engineering and Management*, 2006; 132(7):733–740.
- Huang X, Hinze J. Analysis of construction worker fall accidents. *Journal of Construction Engineering and Management*, 2003; 129(3):262–271.
- Baradan S, Usmen M. Comparative injury and fatality risk analysis of building trades. *Journal of Construction Engineering and Management*, 2006; 132(5):533–539.
- Jannadi O, Almishari S. Risk assessment in construction. *Journal of Construction Engineering and Management*, 2003; 129(5):492–500.
- Everett JG. Overexertion injuries in construction. *Journal of Construction Engineering and Management*, 1999; 125(2):109–114.
- Shapira A, Lyachin B. Identification and analysis of factors affecting safety on construction sites with tower cranes. *Journal of Construction Engineering and Management*, 2009; 135(1):24–33.
- Sacks R, Rozenfeld O, Rosenfeld Y. Spatial and temporal exposure to safety hazards in construction. *Journal of Construction Engineering and Management*, 2009; 135(8):726–736.
- Alexander D, Hallowell M, Gambatese J. Energy-based safety risk management: Using hazard energy to predict injury severity. In ICSC15: The Canadian Society for Civil Engineering 5th International/11th Construction Specialty Conference, University of British Columbia, Vancouver, Canada. June 7–10, 2015. Available at: <https://open.library.ubc.ca/cIRcle/collections/52660/items/1.0076370>.
- Pareto V. *Cours d'économie politique*. Geneva, Switzerland: Droz, 1896.
- Pinto C, Mendes Lopes A, Machado JA. A review of power laws in real life phenomena. *Communications in Nonlinear Science and Numerical Simulation*, 2012; 17(9):3558–3578.
- Malamud Bruce D. Tails of natural hazards. *Physics World*, 2004; 17(8):31–35.
- Papalexiou SM, Koutsoyiannis D, Makropoulos C. How extreme is extreme? An assessment of daily rainfall distribution tails. *Hydrology and Earth System Sciences*, 2013; 17(2):851–862.
- Menendez M, Mendez FJ, Losada IJ, Graham NE. Variability of extreme wave heights in the northeast Pacific Ocean based on buoy measurements. *Geophysical Research Letters*, 2008; 35(22).
- Malamud BD, Turcotte DL. The applicability of power-law frequency statistics to floods. *J Hydrol*, 2006; 322(1):168–180.
- Ahn S, Kim JH, Ramaswami V. A new class of models for heavy tailed distributions in finance and insurance risk. *Insur Math Econ*, 2012; 51(1):43–52.
- Jagger TH, Elsner JB, Saunders MA. 2008: Forecasting U.S. insured hurricane losses. Pp. 189–208 in Diaz HF, Murnane RJ (eds). *Climate Extremes and Society*. Cambridge University Press, 2008.
- Katz RW. Stochastic modeling of hurricane damage. *Journal of Applied Meteorology*, 2002; 41(7):754–762.
- Reed WJ. The Pareto, Zipf and other power laws. *Economics Letters*, 2001; 74(1):15–19.
- Crovella ME, Bestavros A. Explaining World Wide Web traffic self-similarity. Boston University Computer Science Department, 1995. Available at: <http://dcommon.bu.edu/xmlui/handle/2144/1574>.
- Gilleland E, Katz RW. New software to analyze how extremes change over time. *Eos Transactions, American Geophysical Union*, 2011; 92(2):13–14.
- Silverman BW. *Density Estimation for Statistics and Data Analysis*, Vol. 26. CRC Press, 1986.
- Hastie T, Tibshirani R, Friedman J. *The Elements of Statistical Learning*, Vol. 2, No. 1. New York: Springer, 2009.
- Saporta G. *Probabilites, analyse des donnees et statistique*. Paris, France: Editions Technip, 2011.
- Moon YI, Rajagopalan B, Lall U. Estimation of mutual information using kernel density estimators. *Physical Review E*, 1995; 52(3):2318–2321.

43. Rajagopalan B, Lall U, Tarboton DG. Evaluation of kernel density estimation methods for daily precipitation resampling. *Stochastic Hydrology and Hydraulics*, 1997; 11(6):523–547.
44. Jones MC. Simple boundary correction for kernel density estimation. *Statistics and Computing*, 1993; 3(3):135–146.
45. Hu Y, Scarrott CJ. *evmix: An R package for extreme value mixture modelling, threshold estimation and boundary corrected kernel density estimation*, 2013.
46. R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing, 2015. Available at: <http://www.R-project.org/>.
47. Coles S, Bawa J, Trenner L, Dorazio P. *An Introduction to Statistical Modeling of Extreme Values*, Vol. 208. London: Springer, 2001.
48. Charpentier A, Oulidi A. Beta kernel quantile estimators of heavy-tailed loss distributions. *Statistics and Computing*, 2010; 20(1):35–55.
49. Breiman L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). *Statistical Science*, 2001; 16(3):199–231.
50. Vrac M, Naveau P. Stochastic downscaling of precipitation: From dry events to heavy rainfalls. *Water Resources Research*, 2007; 43(7).
51. Scarrott C, MacDonald A. A review of extreme value threshold estimation and uncertainty quantification. *REVSTAT Statistical Journal*, 2012; 10(1):33–60.
52. Furrer EM, Katz RW. Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resources Research*, 2008; 44(12).
53. Frigessi A, Haug O, Rue H. A dynamic mixture model for unsupervised tail estimation without threshold selection. *Extremes*, 2002; 5(3):219–35.
54. Efron B. *Bootstrap Methods: Another Look at the Jackknife*. New York: Springer, 1992.
55. Rajagopalan B, Lall U, Tarboton DG, Bowles DS. Multivariate nonparametric resampling scheme for generation of daily weather variables. *Stochastic Hydrology and Hydraulics*, 1997; 11(1):65–93.
56. Lall U, Rajagopalan B, Tarboton DG. A nonparametric wet/dry spell model for resampling daily precipitation. *Water Resources Research*, 1996; 32(9):2803–2823.
57. Sharma A, Tarboton DG, Lall U. Streamflow simulation: A nonparametric approach. *Water Resources Research*, 1997; 33(2):291–308.
58. Moon YI, Lall U. Kernel quantile function estimator for flood frequency analysis. *Water Resources Research*, 1994; 30(11):3095–3103.
59. Apipattanavis S, Rajagopalan B, Lall U. Local polynomial-based flood frequency estimator for mixed population. *Journal of Hydrologic Engineering*, 2010; 15(9):680–691.
60. Hull J. Defining copulas. *Risk*, 2006; 19(10):62–64.
61. Cherubini U, Luciano E, Vecchiato W. *Copula Methods in Finance*. John Wiley Sons, 2004.
62. Durante F, Sempi C. Copula theory: An introduction. Pp. 3–31 in *Copula Theory and its Applications*. Berlin, Heidelberg: Springer, 2010.
63. Bardossy A, Pegram G. Infilling missing precipitation records—A comparison of a new copula-based method with other techniques. *Journal of Hydrology*, 2014; 519:1162–1170.
64. Domino K, Blachowicz T, Ciupak M. The use of copula functions for predictive analysis of correlations between extreme storm tides. *Physica A: Statistical Mechanics and Its Applications*, 2014; 413:489–497.
65. Salvadori G, De Michele C, Kottegoda NT, Rosso R. *Extremes in Nature: An Approach Using Copulas*, Vol. 56. Springer Science Business Media, 2007.
66. Charpentier A. *Structures de dependance et resultats limites avec applications en finance assurance (doctoral dissertation)*. ENSAE ParisTech, 2006.
67. Sklar A. Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris*, 1959; 8:229–231.
68. Charpentier A, Fermanian JD, Scaillet O. The estimation of copulas: Theory and practice. *Copulas Theoretical Application Finance*, 2007; 35–62.