Integrated Catastrophic Risk Management: Robust Balance between Ex-ante and Ex-post Measures

By Yuri M. Ermoliev, Stephen M. Robinson, Elena A. Rovenskaya, and Tatiana Y. Ermolieva

Humans continually face catastrophes involving natural disasters, such as floods, droughts, hurricanes, and large-scale fires. In today's highly interconnected world, losses from such incidents have increased greatly due to growing population densities, asset concentration in disaster-prone areas, and environmental change from anthropogenic impacts.

Catastrophic natural disasters are random events that are rare but very impactful. Traditionally, most catastrophic losses are paid ex-post (adaptively) by individuals (property owners), government agencies, insurers and reinsurers, charity institutions, and international organizations, rather than through explicit ex-ante (forecast-based) arrangement via long-term strategic decisions [7].

Moreover, there is typically little or no prior agreement as to who should bear what portions of the monetary cost. In anticipation of the need to cover potentially large losses in an ad-hoc way, responsible agencies retain certain budget resources for this purpose. However, such retention reduces the options for profitable investment; in the case of large funds, it can potentially stifle economic growth.

We propose that intensification of ex-ante measures—combined with a more intelligent method for setting aside resources to build adaptive capacities for ex-post compensations, contingent credits, catastrophic bonds, monitoring, and regulation—can significantly reduce the overall burden on national economies and strike a healthy balance between economic growth and security. Integrated long-term approaches to risk management and economic development, with an explicit emphasis on the possibility of rare high-consequence catastrophes, enable effective decisions in this context. This tactic requires one to account for the dependence between decisions and risk distributions.

Existing observations demonstrate the increasing magnitude and variability of risks, indicating that one cannot assume catastrophic risk distribution to be Gaussian; in fact, they are skewed and have fat tails. Their focus on tails makes quantile-based risk measures—e.g., value at risk (VaR) and conditional value at risk (CVaR)—more appropriate than variance-based measures applicable only to Gaussian distributions. We have developed and applied a new approach to stochastic optimization in a number of case studies. Our strategy allows us to include quantile-based performance functions in decision support models for integrated catastrophic risk management. These models are characterized by complex nested distributions shaped by the decisions of policymakers. Here we briefly outline this approach, its

See **Risk Management** *on page 6*

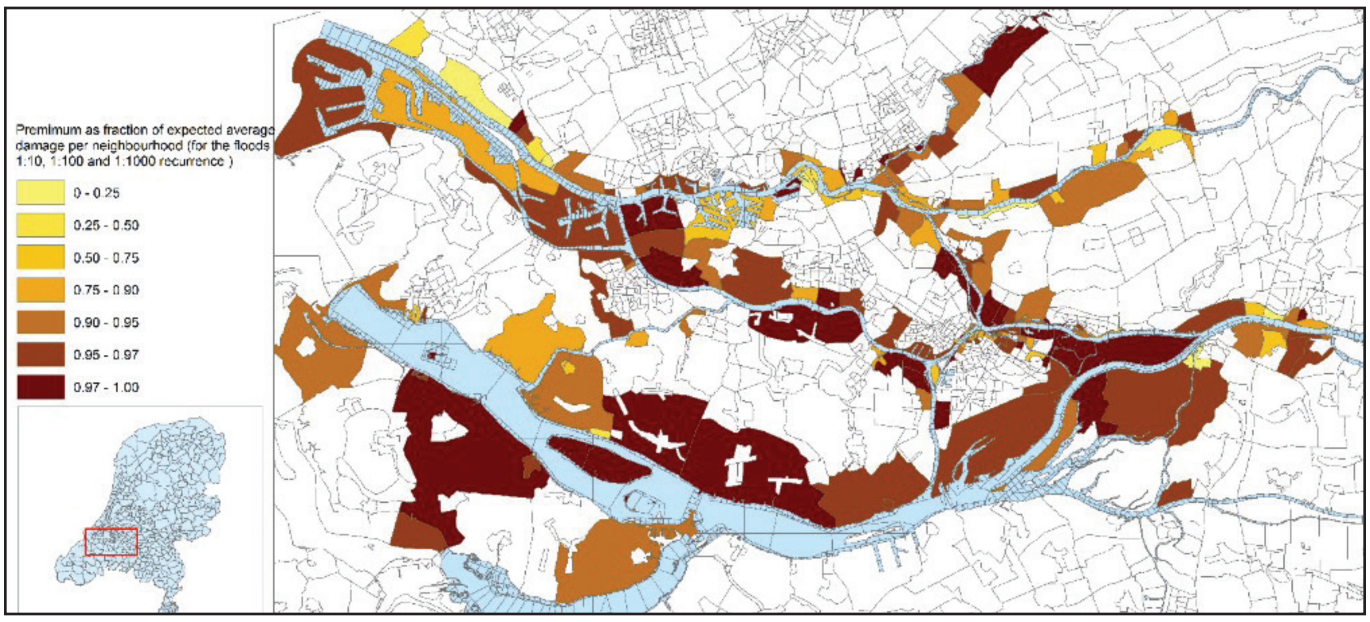

Figure 1. *Geographical distribution of robust premiums as percentage of the 100-year flood damages. Figure courtesy of [6].*

Deep Learning

Continued from page 1

well-publicized success in image classification has encouraged continued work and produced other amazing technologies, such as real-time text translation.

Unfortunately, DNN adoption powered by these successes—combined with the open-source nature of the machine learning community—has outpaced our theoretical understanding. We cannot reliably identify when and why DNNs will make mistakes. Though this does admittedly provide comic relief and fun fodder in research talks about applications like text translation, a single error can be very costly in tasks such as medical imaging. Additionally, DNNs have shown susceptibility to so-called adversarial examples, or data specifically designed to fool a DNN. We can generate such examples with imperceptible deviations from an image, causing the system to misclassify an image that is nearly identical to one that is correctly classified. Adversarial examples in audio applications can also exert control over popular systems like Amazon's Alexa or Apple's Siri, allowing malicious access to devices containing personal information. As we utilize DNNs in increasingly sensitive applications, a better understanding of their properties thus becomes imperative.

Early DNN theory employed learning and function approximation theory to analyze quantities like the Vapnik-Chervonenkis dimension. Although such quantities characterize DNN complexity with respect to training data, many important questions pertaining to generalization, expressibility, learning rule efficiency, intuition, and adversarial example susceptibility remain. More recent interpretations begin to address these questions and fall into three main analysis styles. First are methods to understand the explicit mathematical functions of DNNs by demonstrating the ways in which specific combinations of nonlinearities and weights recover well-known functions on the data. The second approach analyzes theoretical capabilities and limitations of the sequence of functions present in all DNNs — again, given assumptions on the nonlinearities and weights. These analyses include quantifications of the data-dependent cost-function landscape. Finally, a third class of techniques focuses on learning algorithms that solve the high-dimensional, nonlinear optimization programs required to fit DNNs, and attempts to characterize the way in which these algorithms interact with specific DNN architectures.

Advances in DNN theory include many different sources of intuition, such as learning theory, sparse signal analysis, physics, chemistry, and psychology. For example, researchers have related the iterative affine-plus-threshold structure to algorithms that find sparse representations of data [3]. A generalization of this result temporally unrolls the algorithmic iterations that solve regularized least-squares optimization programs

$$\arg\min_{\boldsymbol{x}} \left[\left\| \boldsymbol{y} - \boldsymbol{A}\boldsymbol{x} \right\|_2^2 + \lambda R(\boldsymbol{x}) \right], \quad (1)$$

via a proximal projection method that iteratively calculates

$$\widehat{\boldsymbol{x}}_{t+1} = P_{\lambda}\left(\widehat{\boldsymbol{x}}_t + \boldsymbol{A}^T\left((\boldsymbol{y} - \boldsymbol{A}\widehat{\boldsymbol{x}}_t) \right) \right), \quad (2)$$

where $P_{\lambda}(\boldsymbol{z})$ is the nonlinear proximal projection

$$\min_{\beta} \left\| \boldsymbol{z} - \boldsymbol{x} \right\|_2^2 + \lambda R(\boldsymbol{x}).$$

When the regularization function $R(\cdot)$ is separable, $R(\boldsymbol{z}) = \sum_k R(z_k)$, the proximal projection is a pointwise nonlinearity that mimics DNN architectures. Treating $\widehat{\beta}_t$ as different vectors at each algorithmic iteration, these variables can map to the node values at subsequent DNN layers, with weights $\boldsymbol{w} = \boldsymbol{A}^T\boldsymbol{A} + \boldsymbol{I}$ between layers, a bias $\boldsymbol{b} = \boldsymbol{A}^T\boldsymbol{y}$, and nonlinearity defined by the proximal projection. This example offers a sense of the intuitions gleaned by mapping the network operations onto well-known algorithms. And this single interpretation is just the tip of the iceberg; a larger, non-exhaustive list of additional explanations is available in [1].

The sheer quantity of recent publications on DNN theory demonstrates just how relentless the search for meaning has become. An interesting pattern begins to emerge in the breadth of possible interpretations. The seemingly limitless approaches are mostly constrained by the lens with which we view the mathematical operations. Physics-based interpretations stem from researchers with a physics background. Connections to sparsity and wavelets come from well-known scientists in those fields. Ultimately, the interpretation of DNNs appears to mimic a type of Rorschach test — a psychological test wherein subjects interpret a series of seemingly ambiguous ink-blots (see Figure 1b, on page 1). Rorschach tests depend not only on *what* (the result) a subject sees in the ink-blots but also on the *reasoning* (methods used) behind the subject's perception, thus making the analogy particularly apropos.

On the one hand, these diverse perspectives are unsurprising, given DNNs' status as arbitrary function approximators. Specific network weights and nonlinearities allow DNNs to easily adapt to various narratives. On the other hand, they are not unique in permitting multiple interpretations. We can likewise view standard, simpler algorithms through many lenses. For example, we can derive the Kalman filter—a time-tested algorithm that tracks a vector over time—from at least three interpretations: the orthogonality principle, Bayesian maximum *a-priori* estimation, and low-rank updates for least-squares optimization. These three derivations allow people with different mathematical mindsets (i.e., linear algebra versus probability theory) to understand the algorithm. Yet compared to DNNs, the Kalman filter is simple; it consists of only a handful of linear-algebraic operations. Its function is completely understood, allowing for validation of each viewpoint despite the different underlying philosophies.

Similar validation for DNN theory requires a convergence of the literature. We must distinguish between universal results that are invariant to the analysis perspective and those that are specific to a particular network configuration. A healthy debate is already underway, with respect to the information bottleneck interpretation of DNNs [4, 5]. We should also work to better understand the interactions between functions that DNNs perform, their mathematical properties, and the impact of optimization methods. Unfortunately, DNN complexity introduces numerous challenges. Many standard tools, such as those that attempt to comprehend a model's generalization from training data [6] or empirically assess important network features [2], are difficult to apply to DNNs. Luckily, there is no shortage of excitement, and we continue to enhance our understanding of DNNs with time. The community is also beginning to coalesce, and dedicated meetings—like workshops at the Conference on Neural Information Processing Systems and the recent Mathematical Theory of Deep Neural Network symposium at Princeton University—will further accelerate our pace.

References

[1] Charles, A.S. (2018). Interpreting deep learning: The machine learning rorschach test? Preprint, *arXiv:1806.00148.*

[2] Ghorbani, A., Abid, A., & Zou, J. (2017). Interpretation of neural networks is fragile. Preprint, *arXiv:1710.10547*.

[3] Papyan, V., Romano, Y., & Elad, M. (2016). Convolutional neural networks analyzed via convolutional sparse coding. *J. Mach. Learn. Res., 18*, 1-52.

[4] Saxe, A.M., Bansal, Y., Dapello, J., Advani, M., Kolchinsky, A., Tracey, B.D., & Cox. D.D. (2018). On the information bottleneck theory of deep learning. In *Sixth International Conference on Learning Representations*. Vancouver, Canada.

[5] Tishby, N. & Zaslavsky, N. (2015). Deep learning and the information bottleneck principle. In *2015 IEEE Information Theory Workshop* (pp. 1-5). Jeju, Korea.

[6] Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2017). Understanding deep learning requires rethinking generalization. In *Fifth International Conference on Learning Representations*. Toulon, France.

Adam S. Charles received his Ph.D. in electrical and computer engineering with a focus on signal processing from the Georgia Institute of Technology, and is currently a postdoctoral associate at the Princeton Neuroscience Institute. He works in the areas of signal processing, machine learning, and computational neuroscience, with a focus on novel neural recording technologies and analysis methods for biological neural networks and the artificial networks used to model them.