Title
HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community

Journal
Hydrology and Earth System Sciences, 22(11)

ISSN
1027-5606

Authors
Shen, Chaopeng
Laloy, Eric
Elshorbagy, Amin
et al.

Publication Date
2018

DOI
10.5194/hess-22-5639-2018

Peer reviewed

HESS Opinions: Incubating deep-learning-powered hydrologic science advances as a community

Chaopeng Shen[1], Jerad Bales, Amin Elshorbagy[5], Eric Laloy[2], Adrian Albert[3], Fi-John Chang[4], Sangram Ganguly[6], Kuo-lin Hsu[7], Daniel Kifer[8], Zheng Fang[9], Kuai Fang[1], Dongfeng Li[9], Xiaodong Li[10], and Wen-Ping Tsai[1]

1. Civil and Environmental Engineering, Pennsylvania State University, University Park, PA 16802
2. Institute for Environment, Health and Safety, Belgian Nuclear Research Centre, Mol, Belgium
3. Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA 94720
4. Department of Bioenvironmental Systems Engineering, National Taiwan University, Taipei, 10617, Taiwan
5. Dept. of Civil, Geological, and Environmental Engineering, University of Saskatchewan, Saskatoon, Canada
6. NASA Ames Research Center/ BAER Institute, Moffett Field, CA 94035
7. Civil and Environmental Engineering, University of California, Irvine, Irvine, CA 92697
8. Computer Science and Engineering, Pennsylvania State University, University Park, PA 16802
9. Civil Engineering, University of Texas at Arlington, Arlington, TX 76013
10. State Key Laboratory of Hydraulics and Mountain River Engineering, Sichuan University, Sichuan, China

Correspondence to: Chaopeng Shen (cshen@engr.psu.edu)

Abstract. Recently, deep learning (DL) has emerged as a revolutionary and versatile tool transforming industry applications and generating new and improved capabilities for scientific discovery and model building. The adoption of DL in hydrology has so far been gradual, but the related fields are now ripe for breakthroughs. This paper suggests that DL-based methods can open a viable, complementary avenue toward knowledge discovery in hydrologic sciences. In the new avenue, machine-learning algorithms present competing hypotheses that are consistent with data for scientists to further evaluate. Interrogative studies are then invoked to interpret DL models. However, hydrology presents many challenges to DL-powered scientific advances, such as data limitations, model diversity and variability, and the general inexperience of the hydrologic field with DL. The roadmap toward DL-powered scientific advances will need the coordinated effort from a large community involving scientists and citizens. Integrating process-based models with DL models will help alleviate data limitations. The sharing of data and baseline models will improve the efficiency of the community as a whole. Open competitions will greatly propel growth in hydrology and further enhance data science education in hydrology, which demands a grass-root collaboration. There is a great number of research opportunities in this new area which may stimulate advances in machine learning as well.

1. Overview

Deep learning (DL) is a suite of tools centered on artfully designed large-size artificial neural networks. The deep networks at the core of DL are said to have "depth" due to their multi-layered structures, which help represent abstract concepts about the data . Given input attributes that describe an instance, deep networks can be trained to make predictions of some dependent variables, either continuous or categorical, about this instance. For example, for standard computer vision problems, deep networks can recognize the theme or objects from a picture (Guo et al., 2016; He et al., 2016; Simonyan and Zisserman, 2014) or remotely sensed images (Zhu et al., 2017). For sequential data, DL can associate natural language sequence to commands (Baughman et al., 2014; Hirschberg and Manning, 2015) or predict the action of an actor in the next video frame (Vondrick et al., 2016). DL can also generate (or synthesize) images that carry certain artistic styles (Gatys et al., 2016) or a natural language response to questions (Leviathan and Matias, 2018; Zen and Sak, 2015). With the support of deep architecture, deep networks can automatically engineer relevant concepts and features from large datasets, instead of requiring human experts to define these features (Section 2.2.2). As a foundational component of modern artificial intelligence (AI), DL has made substantial strides in recent years and helped solve problems that have resisted AI for decades (LeCun et al., 2015).

While DL has stimulated exciting advances in many disciplines and has become the method of choice in some areas, hydrology so far have only had a very limited set of DL applications (Shen, 2018) (hereafter referred to as Shen18). Despite scattered reports of promising DL results (Fang et al., 2017; Laloy et al., 2017, 2018; Tao et al., 2016; Vandal et al., 2017; Zhang et al., 2018), water scientists seemed ~~to have reservations about~~have not widely adopted these new tools~~, perhaps with some good reasoning~~. This collective opinion paper argues that there are many opportunities in hydrological sciences where DL can help provide both stronger predictive capabilities and a complementary avenue toward scientific discovery. We then reflect on why it has been challenging to harness the power of DL and big data in hydrology and explore what we can do as a community to incubate progress. Readers who are less familiar with machine learning or deep learning are referred to a companion review paper (Shen18), which provides a more comprehensive and technical background than this opinion paper. Many details behind the arguments in Section 2 are provided in Shen18.

~~We first voice the opinions that elements of a complementary machine learning-based scientific discovery avenue are taking shape, and this avenue should at least be considered for problems with large data (section 2). Then, we propose several ways to accelerate this avenue (section 3). Finally, we argue that hydrology offers a unique set of challenges for DL research (section 4).~~

2. The emergence of a complementary avenue

We are witnessing the growth of three pillars needed for DL to support a research avenue that is complementary to traditional hypothesis-driven research: big hydrologic data, powerful machine learning algorithms, and interrogative methods to extract interpretable knowledge from the trained networks. This new avenue starts from data, uses DL methods to generate hypotheses, and applies interrogative methods to help us understand hydrologic system functioning. We discuss these aspects in the following sections.

2.1. With more data, opportunities arise

The fundamental supporting factor for emerging opportunities with DL is the growth of big hydrologic data, with all surface, sub-surface, urban, infrastructure, and ecosystem dimensions. Here, hydrology includes both the complete natural and engineered water cycle, and associated processes in the ecosystem and geologic media. There are ever increasing amounts of hydrologic data available through remote sensing (see a summary in Srinivasan, (2013)) and data compilations. For example, satellite-based datasets include precipitation, surface soil moisture (Entekhabi, 2010; Jackson et al., 2016; Mecklenburg et al., 2008), vegetation states and indices, e.g., (Knyazikhin et al., 1999), and derived evapotranspiration products (Mu et al., 2011), terrestrial water storage (Wahr et al., 2006), snowcover (Hall et al., 2006), and a planned mission for estimating streamflows (Pavelsky et al., 2014), etc. On the data compilation side, there are now compilations of geologic (Gleeson et al., 2014) and soil datasets; centralized management of streamflow and groundwater data in the United States, Europe, parts of South America and Asia, or globally for some large rivers (GRDC, 2017); water chemistry, groundwater samples and other biogeophysical datasets. The Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) operates two ~~data~~systems for the discovery and archival of water data: the Hydrologic Information System (CUAHSI, 2018c) for time series, and HydroShare for all water data types (Horsburgh et al., 2016). An Internet of Water (Aspen, 2017) has been proposed and is beginning to develop, thereby improving access to these emerging data sets.

Moreover, unconventional data sources are starting to emerge. High-resolution sensing of Earth will be provided by increasing amount of CubeSats, Unmanned Aerial Vehicles, balloons, inexpensive photogrammetric sensing and many other sources (McCabe et al., 2017). These new sources provide new forms of measurements not envisioned before. For example, cell phone signal strength and cell-phone pictures can contribute to high resolution monitoring of rainfall intensity (Allamano et al., 2015). Inexpensive infrared camera images can detect water levels in complex urban water flows (Hiroi and Kawaguchi, 2016). Internet-of-Things (IoT) sensors embedded in water infrastructure can transmit data about the states of water in our environment (Zhang et al., 2018). These new sources of information provided unprecedented volumes and multi-faceted coverages of the natural and built environment. However, since each new data source has its own characteristics and peculiarities, the identification of the appropriate approaches to fully exploit their value, especially synergistically, creates a significant challenge. In contrast, DL models can be built without significant human expertise and extensive manual labour, to rapidly derive useful information from these data.

2.2. DL: A big step forward

2.2.1. Rapid adoption

The field of hydrology has witnessed flows and ebbs of several generations of machine learning methods in the past few decades. From regularized linear regression (Tibshirani and Tibshirani, 1994) to Support Vector Regression (Drucker et al., 1996), from genetic programming (Koza, 1992) to artificial neural networks (Chang et al., 2014; Chen et al., 2018; Hsu et al., 1995, 1997, 2002), from classification and regression tree to random forest (Ho, 1995), from Gaussian Process (Snelson and Ghahramani, 2006) to Radial Basis Function Network (Moradkhani et al., 2004), each approach offered useful solutions to a set of problems, but each also faced its own limitations. As a result, over time, some may have grown dispassionate about progress in machine learning, while some others may have concerns about whether DL represents real progress or is just ~~a~~ “hype.”

The progress in AI brought forth by DL to various industries and scientific disciplines is revolutionary (Section 4 in Shen18) and can no longer be ignored by the hydrology community. Major technology firms have rapidly adopted and commercialized DL-powered AI (Evans et al., 2018). For example, Google has re-oriented its research priority from “mobile-first” to “AI-~~-~~first” (Dignan, 2018). The benefits of these investments can now be felt by ordinary users of their services such as machine translation and digital assistants who can engage in conversations sounding like a human (Leviathan and Matias, 2018). AI patents grew at a 34% compound annual growth rate between 2013 and 2017, apparently after DL’s breakthroughs in 2012 (Columbus, 2018). More than 65% of data professionals responded to a survey indicating AI as their company’s most significant data initiative for next year.

DL is gaining adoption in a wide range of scientific disciplines and, in some areas, has started to substantially transform those disciplines. The fast growth is clearly witnessed from literature searches. Since 2011, the number of entries with DL as a topic increased almost exponentially, showing around 100% compound annual growth rate before 2017 (Table 1). DL evolved from occupying less than 1% of machine learning entries in computer science in 2011 to 46% in 2017. This change showcases massive conversion from traditional machine learning to DL within computer science. Other disciplines lagged slightly behind, but also experienced exponential increase. They also saw the DL/ML ratio jumping from 0% in 2011 to 33% in 2017. As reviewed in Shen18, DL has enhanced the statistical power of data in high energy physics, and the use of DL can be considered to be equivalent to a 25% increase in the experimental dataset (Baldi et al., 2015). In biology, DL has been used to predict potential pathological implications from genetic sequences (Angermueller et al., 2016). DL models in computational fed with raw-level data have been shown to outperform those using expert-defined features when they predict high-level outcomes, e.g., toxicity, from molecular compositions (Goh et al., 2017). Just like other methods, DL may eventually be replaced by newer ones, but that is not a reason to hold out on possible progress.

Table 1. Number of papers returned from searches on ISI Web of Science.

year	DL-nonCS	DL-CS	ML-non-CS	ML-CS	DL/ML-CS	DL/ML-nonCS
2011	0	23	1068	1838	1%	0%
2012	15	25	1310	1899	1%	1%
2013	35	80	1677	2360	3%	2%
2014	84	238	2228	3050	8%	4%
2015	308	709	3074	4405	16%	10%
2016	841	1462	4414	5361	27%	19%
2017	2035	2723	6125	5860	46%	33%

DL-CS results were obtained by searching for topic (TS)="Deep Learning" AND "Research area" (SU)= "Computer Science"; ML-CS was obtained the same way as DL-CS, only that "Deep learning" was replaced by "machine learning"; DL-nonCS was obtained by TS="Deep Learning" NOT SU="Computer Science" NOT SU=education. Education was removed because entries in this category were not related to our definition of DL. There were also 19 articles in 2011 where deep learning was about education in disciplines other than SU=Education. Therefore, 19 was used as a blank value and also subtracted from the DL-nonCS column. DL/ML-CS is ratio of DL-CS to ML-CS expressed as a percentage. DL/ML-nonCS was obtained similarly.

Many of the abovementioned advances were driven by DL's domination in AI competitions:

- The ImageNet Challenges is an open competition to evaluate algorithms for object detection and image classification (Russakovsky et al., 2014). Topics change during each contest, and a dataset of ~14M tagged images and videos were cumulatively compiled, with convenient and uniform data access provided by the organizers. The 2010 Challenge was won by a large-scale Support Vector Machine (SVM). Convolutional Neural Network, a kind of deep network, first won this contest in 2012 (Krizhevsky et al., 2012a). This victory heralded the exponential growth of DL in popularity. Since then, and until 2017 (the last contest), the vast majority of entrants and all contest winners used CNNs, which edges out other methods by large margins (Schmidhuber, 2015).
- The IJCNN traffic sign recognition contest, which is composed of 50,000 images (48 pixels x 48 pixels), witnessed superhuman visual recognition performance (greater than human recognition) from CNN-based methods (Stallkamp et al., 2011). CNNs also performed better than humans on recognition of cancers from medical images (Yu et al., 2016).
- The TIMIT speech corpus is a dataset that holds the recordings from 630 English speakers. LSTM-based models showed a large edge over Hidden Markov Model (HMM) results (Graves et al., 2013) in recognizing the speeches. Similarly, LSTM-based methods significantly outperformed all statistical approaches in keyword spotting (Indermuhle et al., 2012), optical character recognition, language identification, text-to-speech synthesis, social signal classification, machine translation and Chinese handwriting recognition.
- An LSTM-based speech recognition system has achieved "human parity" in conversational speech recognition on the Switchboard corpus (Xiong et al., 2016). A parallel version achieved best-known pixel-wise brain image segmentation results on the MRBrainS13 dataset (Stollenga et al., 2015). The improvement in language translation software can be witnessed by ordinary web users.

- A time-series forecasting contests, Computational Intelligence in Forecasting Competition, was won by a combination of fuzzy and exponential models in 2015 when no LSTM was present, but LSTM won the contest in 2016 (CIF, 2016).

In contrast, only a handful of applications of big data DL could be found in hydrology. Vision DL has been employed to retrieve precipitation from satellite images, where exihibited a materially-superior performance than earlier-generation neural networks (Tao et al., 2017, 2018). GAN was used to imitate and generate scanning images of geologic media (Laloy et al., 2018). Time-series deep learning network was employed to temporally extend satellite-sensed soil moisture observations (Fang et al., 2017) and was found to be more reliable than simpler methods. Regionalized time series DL rainfall-runoff models have been created (Kratzert et al., 2018). There are also DL studies, based on smaller dataset, to help predict water flows in the urban environment (Assem et al., 2017) and water infrastructure (Zhang et al., 2018).- In addition to utilizing big data, DL was able to create valuable, big datasets that could not have been otherwise possible. For example, utilizing DL, researchers were able to generate new datasets for Tropical Cyclones, Atmospheric Rivers and Weather Fronts (Liu et al., 2016; Matsuoka et al., 2017) by tracking them. Machine learning has also been harnessed to tackled the convection parameterization issue in climate modelling (Gentine et al., 2018).

2.2.2. Technical advances

Underpinning the powerful performance of DL are its technical advances. The deep architectures have several distinctive advantages: (1) deep networks are designed with the capacity to represent extremely complex functions. (2) After training, the intermediate layers can perform modular functions which can be migrated to other tasks, in a process called transfer learning, and extend the value of the training data. (3) The hidden layer structures have been designed to automatically extract features, which helps dramatically reducing labor, expertise and the trial and error time needed for feature engineering. (4) Compared to earlier models like classification trees, deep networks are differentiable, meaning that we can calculate derivatives of outputs with respect to inputs or the parameters in the network. This feature enables highly efficient training algorithms that exploit these derivatives. Moreover, the differentiability of neural networks enables querying DL models for sensitivity analysis of outputs to input parameters, a task of key importance in many scientific applications, including hydrology. Metaphorically, the intermediate (or hidden) layers in DL algorithms can be understood as placeholders for tools that are to be built by deep networks themselves. These hidden layers are trained to calculate certain features from the data, which are then used to predict the dependent variables. For example, Yosinski (2015) showed that some intermediate layers of a deep vision recognition network are responsible for identifying the location of human or animal faces; Karpathy et al., (2015) showed that some hidden cells in a text prediction network act as length counters of a line while some others keep track of whether the text is in quotes or not. These functionality were not bestowed by the network designers, but emerged by themselves after network training. Earlier network architectures either did not have the needed depth, or were not designed in an artful way such that the intermediate layers could be effectively trained. For more technical details, refer to an introduction in Schmidhuber (2015) and Shen (2018).

Given that deep networks can identify features without guide, it follows that they may extract features that the algorithm designers were unaware of, or did not intentionally encode the network to do, leading to a potential pathway toward knowledge discovery. For example, deep networks recently showed that grid-like neuron response structures automatically emerge at intermediate network layers for a network trained to imitate how mammals perform navigation, providing strong support to a Nobel-winning neuroscience theory about the functioning of these structures (Banino et al., 2018).

Deep networks may be more robust than simpler models despite their large size, if they are regularized properly (regularization techniques apply penalty to model complexity to make the model more robust) and are chosen based on validation errors in a two-stage approach (Kawaguchi et al., 2017). Effective regularization techniques include (i) early stopping: monitor the training progress on a separate validation set and stop the training once validation metrics start to deteriorate; and/or (ii) novel regularization techniques such as dropout (Srivastava et al., 2014). DL models can be easier to train than previous networks, as their architectures and new stochastic gradient techniques (Kingma and Ba, 2014) address issues like vanishing gradient (Hochreiter, 1998). Training large networks used today was computationally implausible until scientists started to exploit the parallel processing power of graphical processing units (GPUs). Nowadays new application-specific integrated circuits have also been created to specifically tackle DL, although DL architectures are rapidly evolving.

Primary types of successful deep learning architectures include convolutional neural networks (CNN) for image recognition (Krizhevsky et al., 2012b; Ranzato et al., 2006), Long short-term memory (LSTM) (Greff et al., 2015; Hochreiter and Schmidhuber, 1997) for time series modeling, variational auto-encoders (VAE) (Kingma and Welling, 2013), and deep belief networks for pattern recognition and data (typically image but also text or sound, etc) generation (section 3.2 in ~~Shen17~~Shen18). Besides these new architectures, a novel generative model concept called generative adversarial networks (GANs) has become an active area of research. The key characteristic of GANs is that they are learned by creating a competition between the actual generative model or “generator” and a discriminator in a zero-sum game framework (Goodfellow et al., 2014), in which these components are learned jointly. Compared to other generative models, GANs potentially offer much greater flexibility in the patterns to be generated. The power of GANs has been recognized recently in the geoscientific community, especially in machine learning research inspired by physics, where ~~deep generative models~~GANs have been used ~~for~~ to generate certain complicated physical, environmental, and socio-economic systems ~~with deep generative models~~ (Albert et al., 2018; Laloy et al., 2018).

2.3. Network interrogative methods to enable knowledge gain from deep networks

Conventionally, neural networks were primarily used to approximate mappings between inputs and outputs. The focus was put on improving predictive accuracy. In terms of the use of neural networks in scientific research, then, there have been a major concern that DL and more generally machine learning (ML) are referred to as black boxes that cannot be understood by humans and, thus, cannot serve to advance scientific understanding. At the same time, data-driven research may lack clearly-stated hypotheses which is in contrast to traditional hypothesis-driven scientific methods. There has been significant pressure from inside and outside the DL community to make the network decisions more explainable. For example, new (as of January 2018) European data privacy laws dictate that automated individual decision making, which significantly influences the algorithm's users must provide a "right to explanation" where a user can ask for an explanation of an algorithmic decision (Goodman and Flaxman, 2016).

Some recent progress in DL research focused on addressing these concerns. Notably, a new sub-discipline, known as "AI neuroscience" has produced useful interrogative techniques to help scientists interpret the DL model (see literature in Section 5.2 in Shen18). The main classes of interpretive methods include (i) relevance back-propagation: attributing deep network decisions to input features or a subset of inputs; (ii) transferring knowledge from deep networks to interpretable, reduced-order models such as classification trees. (iii) visualization of network activations. Many scientists have also devised case-by-case adhoc methods, e.g., to investigate the correlation between inputs and cell activations (Shen, 2018; Voosen, 2017).

Interpretive DL methods have so far not been employed in hydrology or even geosciences. However, to give some examples from other domains, in medical image diagnosis, some researchers used relevance back-propagation methods to show which pixel on an image led the network to make its decision about anatomy classifications (Kumar and Menkovski, 2016). They found that the network traced its decisions to image landmarks mostly often used by human experts. In more recent research, AI researchers trained their network to not only classify an image, but also didactically explain why the decision was made and why an image is one class instead of another (Figure 1). Extending this idea to the precipitation retrieval problem in hydrology as in (Tao et al., 2017, 2018), we could let DL inform us what features on the satellite cloud image is helpful for reducing bias in precipitation retrieval.

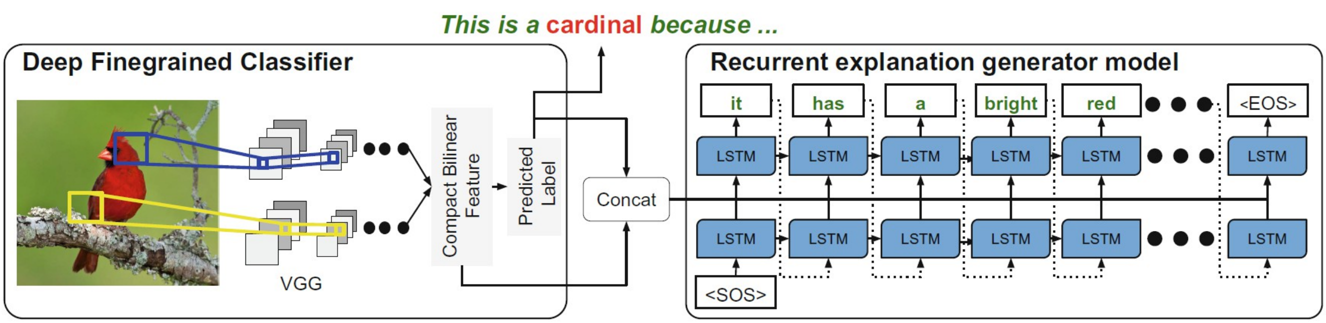

Figure 0. (Reprinted from Hendricks et al., 2016 with permission) Authors trained a joint classification and explanation network for image classification. The bolded text is a "class-relevant" attribute (a distinguishing attribute for the class) in the explanation. Their classification network extracts visual features (regions on the image) responsible for the decision. Then, the explanation network links these regions to distinguishing words in a dictionary to form an explanation that explains the reason for the classification, and why it is not other classes. This level of explanation may be difficult to achieve for hydrologic problems due to limited supervising data (annotated dictionary for classes), but it is possible to borrow the idea of isolating features in the input data and associating them with some descriptive words.

2.4. The complementary research avenue

As the interrogative methods further grow, there emerges a research avenue toward attaining knowledge that is complementary to the traditional hypothesis-driven one (Figure 2). The data-driven research avenue can be divided into four steps: (i) hypotheses are generated by machine learning algorithms from data; (ii) the validation step is where data withheld from training, and different from training, are employed to evaluate the machine-learning-generated hypotheses; (iii) interpretive methods are employed to extract data-consistent and human-understandable hypotheses (Mount et al., 2016) (described in Section 2.3); and (iv) the retained hypotheses are presented to scientists for analysis and further data collection, and the process iterates.

The classical avenue, especially when applied to modelling studies, faces non-uniqueness and subjectivity. To give a concrete example, consider a classical problem of rainfall-runoff modeling. Suppose a hydrologist found that hydrologic responses in several nearby basins are different. Some basins produce flashier peaks while others have smaller peaks in summer, large seasonal fluctuation and large peak streamflows only in winter. Taking a modeling approach, the hydrologist might invoke a conceptual hydrologic model, e.g., Topmodel (Beven, 1997) and find that the model results do not adequately describe the observed heterogeneity in the rainfall-runoff response. It might be hypothesized that the different behaviors are due to heterogeneity in soil texture, which is not well represented in the model. Subsequently, the hydrologist incorporates processes that represent soil spatial heterogeneity, such as modified soil pedo-transfer functions that can differentiate between the soil types in different regions. Perhaps with some parameter adjustment, this model can provide streamflow predictions that are qualitatively similar to the observations. This procedure then increases the hydrologist's confidence that the heterogeneity in soil hydraulic parameters is responsible for their different hydrologic responses. However, this improvement is not conclusive due to process equifinality: there can be alternative processes that can also result in similar outcomes, e.g., the influence of soil thickness, Karsted geology, terrain or drainage density. The identification of potential improvement might be dependent on the hydrologist's intuition or pre-conceptions, which are nonetheless important but potentially biased. Furthermore, incorporating all the physics into the model may prove technically challenging or too time-consuming.

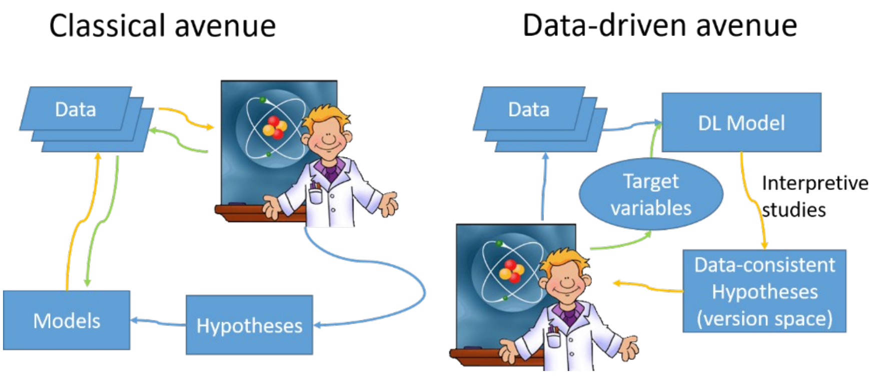

Figure 0. Comparing two alternative avenues toward gaining knowledge from data. In the classical avenue, scientists compile and interpret data, form hypotheses, (optionally) build models to describe data and hypotheses (the green path). Then the model results with data to affirm or reject the hypotheses and the feedbacks (the yellow path) allow the scientist to revise the model and iterate. In the data-driven avenue, scientists collect data and define the target variables of DL models (the green path). Then interpretive methods are invoked to extract data-consistent and human-understandable hypotheses (the yellow path). There must be a hypotheses validation step where data withheld from training is used to evaluate or reject the hypotheses.

Compared to the classical avenue, the data-driven approach allows us to more efficiently explore a larger set of hypotheses. Although it cannot be said that the machine learning algorithms present no human bias (because inputs are human-defined and some hyper-parameters are empirically adjusted), the larger set of hypotheses presented will at least greatly reduce that risk. First, let us examine a CART-based data-driven approach. We could start with physiographic data for many basins in this region, including terrain, soil type, soil thickness, etc. We can use CART to model the process-based model's errors, which allows us to separate out the conditions under which these errors occur more frequently. We let the pattern emerge out of data without enforcing a strong human pre-conceived hypothesis. Attention must be paid to the robustness of the data mining and utilize holdout dataset or cross-validation to verify the generality of the conclusion. Data may suggest that soil thickness is the main reason for the error. Or, if data do not prefer one hypothesis over the other, then all hypotheses are equally possible and cannot be ruled out. This advantage of DL can be summarized in a short phrase, "*an algorithm has no ego*." On a practical level, this approach can more efficiently and simultaneously examine multiple competing hypotheses.

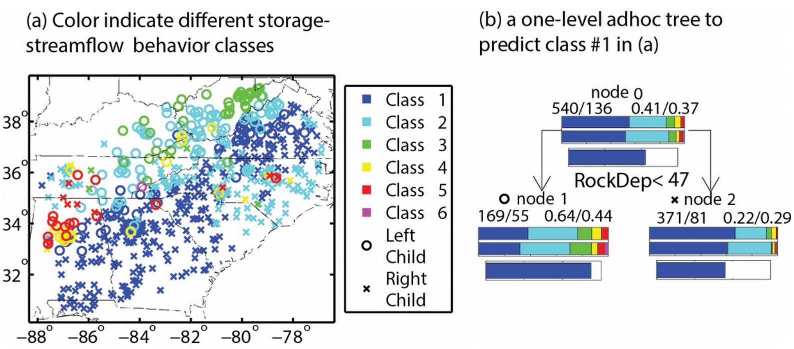

Figure 3. (adapted from Fang and Shen 2017. Reprint permission obtained). We calculated storage-streamflow correlation patterns over continental United States (CONUS) and divided small or mesoscale basins into multiple classes. We studied what physical factors most cleanly separate different correlation patterns. In this case, what separates the blue class (storage and streamflow are highly correlated across all flow regimes) and the green class turned out to be soil thickness. It suggests the blue basins in the south has good correlation because they have thick soils, which facilitates infiltration, water storage, and groundwater-dominated streamflow.

One example of such analyses was carried out in Fang and Shen, (2017) where differences in basin storage-streamflow correlations were explained by physical factors using CART, an earlier-generation data mining method (Figure 4). The data mining analysis allowed patterns to emerge, which inspired hypotheses about key factors that control the hydrologic functioning of different systems, such as soil thickness and soil bulk density are important controls of streamflow-storage relationships. For another example, data-mining analysis showed that drought recovery time is associated to temperature and precipitation, while biodiversity only has secondary importance (Schwalm et al., 2017). Scientists need to define the predictors and general model types, but they do not pose strongly constraining hypotheses about the controlling factors, and instead "let the data speak". The key to this approach is a large amount of data from which patterns emerge.

However, working with DL models, we need to further resort to interrogative methods to make the results understandable (Figure 2 right panel). For example, we can construct DL models to predict the errors of the process-based model, and then use visualization techniques to see which variable, under which condition, lead to the error. Because DL can absorb a large amount of data, it can find commonality among data as well as identify differences. Whereas CART models are limited by the amount of data and face stability problems in lower branches (data are exponentially less at lower branches), DL models may produce a more robust interpretation.

The machine learning paradigm lends us to finding "unrecognized linkages" (Wagener et al., 2010) or find complex patterns in the data that humans could not easily realize or capture. Owning to the strong capability of DL, it can better approximate the "best achievable model" (BAM) for the mapping relations between inputs and output. As such, it lends support to measuring the information content contained in the inputs about the output. Nearing et al., (2016) utilized Gaussian Process regression to approximate the BAM. DL can play similar roles and can also allow for modelling, perhaps in a more thorough way. The simplicity of building DL model and altering inputs makes it an ideal testbed for new ideas.

Outputs from the hidden layers of deep networks can now be visualized to gain insights about the transformations performed on the input data by the network (Samek et al., 2017). For image recognition tasks, one can invert the DL model to find out the parts of the inputs that led the network to make a certain decision (Mahendran and Vedaldi, 2015). There are also means to visualize outputs from recurrent networks, e.g., showing the conditions under which certain cells are activated (Karpathy et al., 2015). These visualizations can illustrate the relationships that the data-driven model has identified.

Considering the above potential benefits, the data-driven avenue should at least be considered or given an opportunity to play a role in hydrological sciences discovery. However, this avenue may be uncomfortable to some ~~researchers~~. In the classical avenue, the scientist must originate the hypotheses before constructing models; in the data-driven one~~avenue~~, the data mining/knowledge discovery process is a precursor step to the main hypotheses formation-- hypotheses cannot be generated before the data mining analysis (Mount et al., 2016). ~~This feature is a natural consequence of handing part of the work to an algorithm but may cause some disarray for those who follow what has been perceived as structured scientific methods.~~ Especially, hypotheses can no longer be unequivocally stated during the proposal stage of research.

Granted, the interrogative methods as a whole are new and time is required for them to grow. We need to note that the nascent "DL neuroscience" literature did not exist until 2015. However, if we outright reject the complementary avenue based on our habitual thinking that neural networks are black boxes, we may deny ourselves opportunities for breakthroughs.

3. Challenges and opportunities for DL in hydrology

The field of hydrology has a unique set of challenges that are also research opportunities for DL. Many of these science challenges have, to date, not been effectively addressed using traditional methods, and cannot be sufficiently tackled by individual research groups. Some challenges for which DL approaches might be exploited are presented below.

Observations in hydrology and water science generally are regionally and temporally imbalanced. For example, while streamflow observations are relatively dense in the United States, such data are sparse in most other parts of the world~~.~~, either because measurements have not been made or are not made accessible. There is often a dearth of observations that can be used as comprehensive training datasets for DL algorithms. Few hydrologic applications have as much data as what standard AI research applications such as imagine recognition or natural language processing need. Remote sensing of hydrologic variables also has limitations, including effects of canopy and clouds which can limit observations, temporal density of observations because of orbital paths, and observation footprints, which create challenges when trying to validate satellite observations with field point measurements. A body of literature studying this problem across different geographic regions can be loosely summarized under the topic of “prediction in ungauged basins” (PUB) (Hrachowitz et al., 2013). PUB problems pose a significant challenge to data-driven methods.

Global change is altering the hydrologic and related cycles, and hydrologists must now make predictions in anticipation of changes, beyond previously observed ranges (Wagener et al., 2010). Especially, more frequent extremes have been observed for many parts of the world and such extremes have been projected to occur more frequently in the future. Data-driven methods must demonstrate their capability to make reasonable predictions when applied out of the range of the training dataset.

Observations of the water cycle tend to focus on one aspect of the water cycle, and seldom offer a complete description. For example, we can estimate total terrestrial water storage (Wahr, 2004) or top 5-cm surface soil moisture via multiple satellite missions. It is difficult, however, to directly combine such observations of components of the water cycle into a complete picture of the water cycle. A challenge, then, is merging distinct observations, with all their space-time discontinuities to aid predictions, model validation, and to provide a more complete understanding of the global water cycle.

Hydrologic data are accompanied by a large amount of strongly heterogeneous (Blöschl, 2006) “contextual variables” such as land use, climate, geology, and soil properties. The proper scale at which to represent heterogeneity in natural systems is a vexing problem (Archfield et al., 2015), as micro-scale of soil heterogeneity, for example, is not computationally realistic in hydrologic models. The scale at which heterogeneity should be represented varies with setting and elements of the water cycle (Ajami et al., 2016). Moreover, while we recognize that heterogeneity exists in contextual features, many of these features, such as soil properties and hydrogeology, are poorly characterized across landscapes, but both features play important role in controlling water movement. Heterogeneity needs to be adequately represented without radically bloating the parameter space of the models. Moreover, the heterogeneous physiographic factors covary (Troch et al., 2013) and exert complicated causal and non-causal connections, but we have limited knowledge of their covariation. Consequently, training with insufficient data may result in many alternative DL models that cannot be rejected.

Hydrologic problems fit poorly into the template of problems for which standard network structures (Section 3.2 in Shen18) are designed, i.e., purely image recognition or time series prediction problems, or a mixture of both. For example, catchment hydrologic problems are characterized by both spatially heterogeneous but temporally static attributes (topography and hydrogeology) and temporal (atmospheric forcing) dimensions. Such input dimension are not efficiently represented by with typical input dimensions of LSTM or CNN.

Because large and diverse datasets are needed for DL application, access to properly pre-processed and formatted present practical challenges. These steps of data compilation, pre-processing, and formatting often occupy too much unnecessary time for researchers. Many of the processing tasks for images cannot be handled by individual research groups. Compared to the DL community in AI and chemistry, etc., DL learning community in hydrology is not sufficiently coordinated, resulting in significant waste of effort and “~~recreation~~ reinvention of wheels”.

Deep generative models such as GANs can be used for the stochastic generation of natural textures. This has recently led to methodological advances in subsurface hydrology (Laloy et al., 2017, 2018, Mosser et al., 2017) where the ability to efficiently and accurately simulate complicated geologic structures with given (non-Gaussian) geostatistical properties is of paramount importance for uncertainty quantification of subsurface flow and transport models. However, amongst other directions for future research, more work is needed (i) to generate the complete range of structural complexity observed in geologic layers, (ii) deal efficiently with large 3D domains and (iii), account for various types of direct (e.g., observed geologic facies at a given location, mean property value over a specific area, etc) and indirect (e.g., measured hydrologic state variables to be used within an inverse modeling procedure) conditioning data in the simulation.

4. A community roadmap toward DL-powered scientific advances in hydrology

Despite the challenges articulated above, here we offer a shared vision for a community roadmap for advancing hydrologic sciences using DL (Figure 3). A well-coordinated community is much more efficient and powerful in resolving problems, as we have seen in other scientific endeavours. Montanari et al. (2013) noted, "future science must be based on an interdisciplinary approach" and "the research challenges in hydrology for the next 10 years should be tackled through a collective effort". We see that several steps are crucial in this roadmap: devising ways to integrate physical knowledge, use DL to infer unknown quantities, process-based models (PBMs) and DL, community approaches in sharing and accessing data, open and transparent model competitions, baseline models and visualization packages and an education program that introduces data-driven methods at various levels.

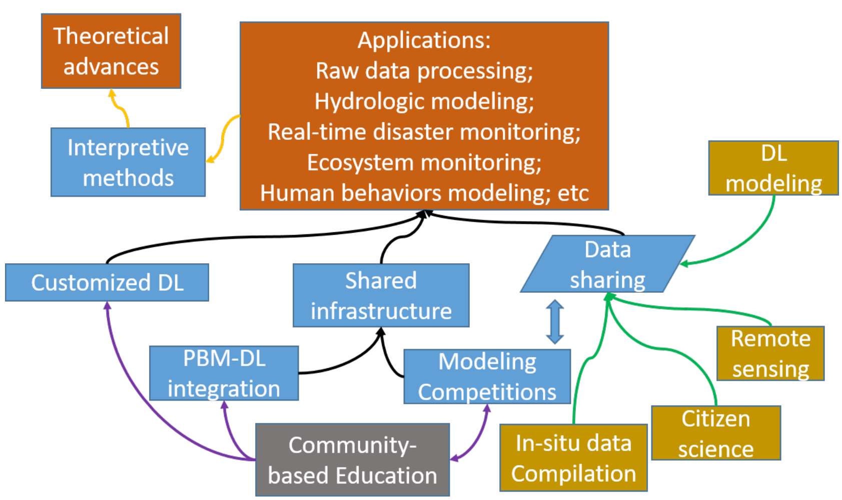

Figure 0. A roadmap toward DL-powered scientific discovery in hydrologic science. Data availability can be increased by (green arrows) collecting and compiling existing data, incorporate novel data sources such as those collected by citizen scientists, remote sensing and modelled dataset. The modelling competitions and the integration between PBM and DL will build important shared computing and analytic infrastructure, which, together with data sources, support a wide range of hydrologic applications. Interpretive methods should be attempted to extract knowledge from trained deep networks (orange arrows). Underpinning these

activities is the enhanced, community-based educational program for machine learning in hydrology (purple arrows). However, these activities, especially the modelling competitions, might in turn feedback to the educational activity.

4.1. Integrating physical knowledge, process-based models, and DL models

To address the challenge of data limitations, we envision that a critical and necessary step is to more organically integrate hydrologic knowledge, process-based models, and DL. Process-based models, as they are derived from underlying physics, require less data for calibration than data-driven models. They can provide estimates for spatial and temporal data gaps and unobserved hydrologic processes. Well-constructed PBMs should also be able to represent temporal changes and trends. However, because data-driven models directly target observations, these models may have better performance in locations and periods where data are available. Also, as discussed earlier, data-driven models are less prone to *a priori* model structural error than are PBMs. We should aim to maximally utilize the best features of each type of models.

There will be a diversity of approaches with which PBMs and data-driven models could be combined. Karpatne et al., (2017) compiled a list of approaches of what they collectively call “theory-guided data science,” which include () using knowledge to design data-driven model; () using knowledge to initialize network states; () using physical knowledge to construct priors states to constrain the data-driven models; () using knowledge-based constrained optimization (although this may be difficult to implement in practice); () using theory as regularization terms for the data-driven model, which will force the model to respect these constraints; and () learn hybrid models, where the data-driven method is used as surrogate for certain partpartpartparts of the physical model. One may also impose multiple learning objectives based on the knowledge of the problem.

There are a multitude of potential approaches and this list can be further expanded to accommodate various objectives. First, we can focus on PBM errors (difference between PBM simulation and observations). Non-deep machine learning has already shown promise in correcting PBM errors. *Abramowitz et al.* (2006) developed an ANN to predict the error in net ecosystem exchange from a land surface model, and achieved 95% reduction in annual error. More importantly, an ANN trained to correct the error at one biome corrects the PBM in another biome with a different temperature regime (Abramowitz et al., 2007). In the context of weather forecasts, machine learning methods were used to learn the patterns from past forecasting errors (Delle Monache et al., 2011, 2013), leading to a 20 percent improvement in performance for events of similar characteristics (Junk et al., 2015). Their results suggest PBMs make structural errors that are independent of the state-variable regimes. We envision that PBMs can better resolve the impacts of regime changes, while DL can better capture state-independent error patterns and do mild state-dependent extrapolations. A co-benefit of modelling PBM error is that insights are gained about the PBM. Using interrogative methods to reverse engineer what DL has learned about PBM error provides possible avenues for improving the underlying PBM processes.

Second, PBMs can augment input data for DL models. PBMs can be used to increase supervising data for DLs, for example, for climate or landuse scenarios that have not existed presently, to augment existing data. Furthermore, if the DL training is limited by available data, we may not be able to reject many alternative DL models that could generate unphysical or unrealistic outputs. Providing PBM simulations as either training data or regularization terms can help to nudge DL models to generate physically meaningful outputs. The extent to which errors in PBM model results affects DL outcomes remains to be explored. A theoretical framework is lacking for separately estimating aleatory uncertainty (resulting from data noise), and epistemic uncertainty (resulting from PBM error and training data paucity) and uncertainty due to regime-shift. There are significant research opportunities in this area. The advantages and disadvantages of various approaches could be systematically and efficiently evaluated in community-coordinated fashion.

4.2. Multi-faced, community-coordinated hydrologic modeling competitions

There are many possible approaches and many alternative model structures for using DL to make hydrologic predictions and to provide insight into hydrologic processes. In the light of these challenges, we argue that open, fast and standardized competitions are one effective way of accelerating the progress. The competitions can evaluate the models not only in terms of predictive performance but also the attainment of understanding.

The impacts of competitions are best evidenced in the community-coordinated challenges in computer science using a standardized set of problems. These competitions have strongly propelled the advances in AI. Some have argued that the contributions of the ImageNet dataset and the competition may be more significant than stimulating the winning algorithms (Gershgorn, 2017). New methods can be evaluated objectively and disseminated rapidly through competitions. Because the problems are standardized, they remove biases due to data sources and pre-processing. The community can quickly learn advantages and disadvantages of alternative model design through these competitions, which also encourage reproducibility.

We envision multi-faceted hydrologic modeling competitions where various models ranging from process-based ones to DL ones are evaluated and compared. The coordinators can, for example, provide a set of standard atmospheric forcings, landscape characteristics, and observed variables~~.~~ and provide targeted questions which participants must address. Importantly, the evaluation criteria should include not only performance-type criteria such as model efficiency coefficients and bias but also **qualitative/explanatory** ones such as explanations for control variables and model errors. Over-simplified or poorly-constructed models may provide more accessible explanations, but they might be misleading because the models may be overfitted to a given situation. Their simplicity may also constrain their ability to digest large datasets as a way of reducing uncertainty. Multi-faceted competitions allow us to also identify a "Pareto front" of explainability and performance and help rule out "false explanations". The objective of the competition is not only to seek the best simulation performance, but also those methods that offer deeper insight into hydrologic processes.

Another important value of competitions is that organizers will provide a standard input dataset and well-defined tasks. The entire community can leverage such effort. A substantial amount of effort is required to establish such a dataset, which may only be possible under a specifically designed project. Moreover, open competitions in the computer science field has produced well-known models such as AlexNet (Krizhevsky et al., 2012b), GoogLeNet (Szegedy et al., 2015), etc. These models serve as benchmarks and quick entry points for others and greatly improve~~d~~ reproducibility and the effectiveness of comparisons. Standard models, datasets and evaluation metrics will greatly improve DL adoption and hydrologic sciences.

4.3. Community-shared resources and broader involvement

A useful approach to address the major obstacle of data limitation is to increase our data repositories and to open access to existing data. Data value can be greatly enhanced by centralized data compilations, a task many institutions are already undertaking. For example, the Consortium of Universities for the Advancement of Hydrologic Science, Inc. (CUAHSI) provides access to large amounts of hydrologic data (CUAHSI, 2018a). As another example, in 2015, a project called Collaborative Research Actions (Endo et al., 2015) was proposed in Belmont Forum, which is a group of the world's major and emerging funders of global environmental-change research. Many scientists from different countries join the project and focus on the same issue, Food-Energy-Water Nexus. They shared their data (heterogeneous data) and research results from different regions.

Using data sharing standards will advance data sharing across domains (WaterML2, 2018). Providing access to data through web services, such as used by CUAHSI, negates the problem of storing data in a single location and enhances discoverability. Data brokers also provide more channels to share experiences, scholarly discussions, and debates along with the generation of data.

An important area where DL is expected to deliver significant value is the analysis of big and sub-research-quality data such as those collected by citizen scientists. Many aspects of the water cycle are directly accessible by everyone. Citizen scientists already gather data about precipitation (CoCoRaHS, 2018), temperature, humidity, soil moisture, river stage (CrowdHydrology, 2018), and potentially groundwater levels. These quantities can be measured using inexpensive instruments~~-~~such as cameras, pressure gauges and moisture sensors. Volunteer scientists can also be solicited for data in places where such data can best reduce the uncertainty of the DL model, as in a framework called active learning (Settles, 2012). Social data have been used to help monitoring flood inundation (Sadler et al., 2018; Wang et al., 2018). Crowd-sourced data have played roles in DL research, where a large but noisy dataset was argued to be more useful than a much smaller but well-curated dataset (Huang et al., 2016; Izadinia et al., 2015). Even though there are problems related to data quality which can be overcome using AI approaches. An important co-benefit of involving citizen scientists is the education and outreach to the public. The active engagement is much more effective when the public has a stake in the research outcomes.

4.4. Education

A major barrier to realizing the benefits of more data science lies in our undergraduate and DL lies in our undergraduate and graduate curriculum. Little in a standard hydrologist's curriculum prepares students for a future with substantially more data-driven science and engineering. Statistical courses often do not cover machine learning basics, yet data mining courses offered by computer science departments lack the connections to the water discipline. Given the interdisciplinary nature of hydrology, it has been long recognized that it takes a community to raise a hydrologist (Merwade and Ruddell, 2012; Wagener et al., 2012). We propose a concerted effort by current hydrologic machine learning researchers, with participation from computer scientists, to pool and share educational content. Such effort will form the basis of a hydrologic data mining curriculum and leverage the wit of the community. Collaborations may form through either grassroot collaborations or institutionally-supported education projects, e.g., (CUAHSI, 2018b). The open competitions would be a great source of education materials. A diversity of models that have been evaluated and contrasted help clarify Pros and Cons of different methods. Shared datasets, DL algorithms and data pre-processing software can be leveraged in classrooms.

As with the design of any education effort, it is important to consider inclusiveness and diversity. Research has found that the introductory computer science classes, especially those taken by non-majors, are instrumental in developing a desire to stay in the field (Lehman, 2017). In addition, the portrayal of gender stereotypes regarding computing and the increase in weed-out courses (Aspray, 2016) have both discouraged women students in computer science (Sax et al., 2017). To counter such negative impacts, the introductory courses in the curriculum need to assume little prior programming experiences. Moreover, the richness of natural science in hydrology may help bridge the gender gap.

5. Concluding remarks

In this opinion paper, we argue that hydrologic scientists ought to give thoughts to a research avenue that complements traditional approaches, wherein DL-powered data mining is used to generate hypotheses, predictions, and insights. Although in the past there may have been strong reservations toward black-box approaches, recent efforts have been put in the interpretation and understanding of deep learning networks, and hydrologists have the opportunity to push research forward in this regard. Progress in hydrology and other disciplines show that there is substantial promise in incorporating DL into toolbox. However, challenges such as data limitation and model variability demand a community-coordinated approach.

We have also argued for open hydrologic competitions that emphasize both performance and explainability. These competitions, along with shared data and DL models, will greatly improve the growth of the field as a whole. Hydrologists should make use of the potential of citizen science, and exploit DL as a valuable tool toward scientific discovery.

References

Abramowitz, G., Gupta, H., Pitman, A., Wang, Y., Leuning, R., Cleugh, H., Hsu, K., Abramowitz, G., Gupta, H., Pitman, A., Wang, Y., Leuning, R., Cleugh, H. and Hsu, K.: Neural Error Regression Diagnosis (NERD): A Tool for Model Bias Identification and Prognostic Data Assimilation, J. Hydrometeorol., doi:10.1175/JHM479.1, 2006.

Abramowitz, G., Pitman, A., Gupta, H., Kowalczyk, E., Wang, Y., Abramowitz, G., Pitman, A., Gupta, H., Kowalczyk, E. and Wang, Y.: Systematic Bias in Land Surface Models, J. Hydrometeorol., doi:10.1175/JHM628.1, 2007.

Ajami, H., Khan, U., Tuteja, N. K. and Sharma, A.: Development of a computationally efficient semi-distributed hydrologic modeling application for soil moisture, lateral flow and runoff simulation, Environ. Model. Softw., 85, 319–331, doi:10.1016/J.ENVSOFT.2016.09.002, 2016.

Albert, A., Strano, E., Kaur, J. and Gonzalez, M.: Modeling urbanization patterns with generative adversarial networks, arXiv:1801.02710 [online] Available from: http://arxiv.org/abs/1801.02710 (Accessed 24 March 2018), 2018.

Allamano, P., Croci, A. and Laio, F.: Toward the camera rain gauge, Water Resour. Res., 51(3), 1744–1757, doi:10.1002/2014WR016298, 2015.

Angermueller, C., Pärnamaa, T., Parts, L. and Stegle, O.: Deep learning for computational biology., Mol. Syst. Biol., 12(7), 878, doi:10.15252/MSB.20156651, 2016.

Archfield, S. A., Clark, M., Arheimer, B., Hay, L. E., McMillan, H., Kiang, J. E., Seibert, J., Hakala, K., Bock, A., Wagener, T., Farmer, W. H., Andréassian, V., Attinger, S., Viglione, A., Knight, R., Markstrom, S. and Over, T.: Accelerating advances in continental domain hydrologic modeling, Water Resour. Res., 51(12), 10078–10091, doi:10.1002/2015WR017498, 2015.

Aspen: Internet of Water: Sharing and Integrating Water Data for Sustainability, A Rep. from Aspen Inst. Dialogue Ser. Water Data [online] Available from: https://www.aspeninstitute.org/publications/internet-of-water/ (Accessed 27 August 2018), 2017.

Aspray, W.: Women and underrepresented minorities in computing : a historical and social study., 2016.

Assem, H., Ghariba, S., Makrai, G., Johnston, P., Gill, L. and Pilla, F.: Urban Water Flow and Water Level Prediction Based on Deep Learning, in ECML PKDD 2017: Machine Learning and Knowledge Discovery in Databases, pp. 317–329, Springer, Cham., 2017.

Baldi, P., Sadowski, P. and Whiteson, D.: Enhanced Higgs Boson to τ + τ − Search with Deep Learning, Phys. Rev. Lett., 114(11), 111801, doi:10.1103/PhysRevLett.114.111801, 2015.

Banino, A., Barry, C., Uria, B., Blundell, C., Lillicrap, T., Mirowski, P., Pritzel, A., Chadwick, M. J., Degris, T., Modayil, J., Wayne, G., Soyer, H., Viola, F., Zhang, B., Goroshin, R., Rabinowitz, N., Pascanu, R., Beattie, C., Petersen, S., Sadik, A., Gaffney, S., King, H., Kavukcuoglu, K., Hassabis, D., Hadsell, R. and Kumaran, D.: Vector-based navigation using grid-like representations in artificial agents, Nature, 1, doi:10.1038/s41586-018-0102-6, 2018.

Baughman, A. K., Chuang, W., Dixon, K. R., Benz, Z. and Basilico, J.: DeepQA Jeopardy! Gamification: A Machine-Learning Perspective, IEEE Trans. Comput. Intell. AI Games, 6(1), 55–66, doi:10.1109/TCIAIG.2013.2285651, 2014.

Beven, K.: Topmodel : A Critique, Hydrol. Process., 11(December 1996), 1069–1085, 1997.

Blöschl, G.: Hydrologic synthesis: Across processes, places, and scales, Water Resour. Res., 42(3), doi:10.1029/2005WR004319, 2006.

Chang, L.-C., Shen, H.-Y. and Chang, F.-J.: Regional flood inundation nowcast using hybrid SOM and dynamic neural networks, J. Hydrol., 519, 476–489, doi:10.1016/J.JHYDROL.2014.07.036, 2014.

Chen, I.-T., Chang, L.-C. and Chang, F.-J.: Exploring the spatio-temporal interrelation between groundwater and surface water by using the self-organizing maps, J. Hydrol., 556, 131–142, doi:10.1016/J.JHYDROL.2017.10.015, 2018.

CIF: Results, Int. Time Ser. Forecast. Compet. - Comput. Intell. Forecast. [online] Available from: http://irafm.osu.cz/cif/main.php?c=Static&page=results (Accessed 24 March 2018), 2016.

CoCoRaHS: Community Collaborative Rain, Hail and Snow Network (CoCoRaHS), “Volunteers Work. together to Meas. Precip. across nations.” [online] Available from: https://www.cocorahs.org/ (Accessed 23 August 2018), 2018.

Columbus, L.: Roundup Of Machine Learning Forecasts And Market Estimates, 2018, Forbes Contrib. [online] Available from: https://www.forbes.com/sites/louiscolumbus/2018/02/18/roundup-of-machine-learning-forecasts-and-market-estimates-2018/#7d3c5bd62225 (Accessed 30 July 2018), 2018.

CrowdHydrology: CrowdHydrology, [online] Available from: http://crowdhydrology.geology.buffalo.edu/ (Accessed 23 August 2018), 2018.

CUAHSI: Consortium of Universities Allied for Water Research, Inc (CUASI) Cyberseminars, [online] Available from: https://www.cuahsi.org/cyberseminars (Accessed 19 July 2016a), 2018.

CUAHSI: Data Driven Education, [online] Available from: https://www.cuahsi.org/education/data-driven-education/ (Accessed 26 September 2018b), 2018.

CUAHSI: HydroClient, [online] Available from: http://data.cuahsi.org/. (Accessed 19 August 2018c), 2018.

Delle Monache, L., Nipen, T., Liu, Y., Roux, G. and Stull, R.: Kalman Filter and Analog Schemes to Postprocess Numerical Weather Predictions, Mon. Weather Rev., 139(11), 3554–3570, doi:10.1175/2011MWR3653.1, 2011.

Delle Monache, L., Eckel, F. A., Rife, D. L., Nagarajan, B. and Searight, K.: Probabilistic Weather Prediction with an Analog Ensemble, Mon. Weather Rev., 141(10), 3498–3516, doi:10.1175/MWR-D-12-00281.1, 2013.

Dignan, L.: Google Research becomes Google AI to reflect AI-first ambitions, zdnet.com [online] Available from: https://www.zdnet.com/article/google-research-becomes-google-ai-to-reflect-ai-first-ambitions/ (Accessed 20 August 2018), 2018.

Drucker, H., Burges, C. J. C., Kaufman, L., Smola, A. and Vapnik, V.: Support vector regression machines, Proc. 9th Int. Conf. Neural Inf. Process. Syst., 155–161 [online] Available from: https://dl.acm.org/citation.cfm?id=2999003 (Accessed 5 January 2018), 1996.

Endo, A., Burnett, K., Orencio, P., Kumazawa, T., Wada, C., Ishii, A., Tsurita, I. and Taniguchi, M.: Methods of the Water-Energy-Food Nexus, Water, 7(10), 5806–5830, doi:10.3390/w7105806, 2015.

Entekhabi, D.: The Soil Moisture Active Passive (SMAP) mission, Proc. IEEE, 98(5), 704–716, doi:10.1109/JPROC.2010.2043918, 2010.

Evans, H., Gervet, E., Kuchembuck, R. and Hu, M.: Will You Embrace AI Fast Enough?, ATKearney Oper. Perform. Transform. Rep., 2018.

Fang, K. and Shen, C.: Full-flow-regime storage-streamflow correlation patterns provide insights into hydrologic functioning over the continental US, Water Resour. Res., doi:10.1002/2016WR020283, 2017.

Fang, K., Shen, C., Kifer, D. and Yang, X.: Prolongation of SMAP to Spatio-temporally Seamless Coverage of Continental US Using a Deep Learning Neural Network, Geophys. Res. Lett., doi:10.1002/2017GL075619, 2017.

Gatys, L. A., Ecker, A. S. and Bethge, M.: Image Style Transfer Using Convolutional Neural Networks, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414–2423, IEEE., 2016.

Gentine, P., Pritchard, M., Rasp, S., Reinaudi, G. and Yacalis, G.: Could Machine Learning Break the Convection Parameterization Deadlock?, Geophys. Res. Lett., 45(11), 5742–5751, doi:10.1029/2018GL078202, 2018.

Gershgorn, D.: The data that transformed AI research—and possibly the world, Quartz, 2017.

Gleeson, T., Moosdorf, N., Hartmann, J. and van Beek, L. P. H.: A glimpse beneath earth's surface: GLobal HYdrogeology MaPS (GLHYMPS) of permeability and porosity, Geophys. Res. Lett., 41(11), 3891–3898, doi:10.1002/2014GL059856, 2014.

Goh, G. B., Hodas, N. O. and Vishnu, A.: Deep learning for computational chemistry, J. Comput. Chem., 38(16), 1291–1307, doi:10.1002/jcc.24764, 2017.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y.: Generative Adversarial Networks, in Proceedings of the 27th International Conference on Neural Information Processing Systems (NIPS'14). [online] Available from: http://arxiv.org/abs/1406.2661 (Accessed 25 February 2017), 2014.

Goodman, B. and Flaxman, S.: European Union regulations on algorithmic decision-making and a "right to explanation", arXiv:1606.08813 [online] Available from: http://arxiv.org/abs/1606.08813 (Accessed 7 February 2018), 2016.

Graves, A., Mohamed, A. and Hinton, G.: Speech Recognition with Deep Recurrent Neural Networks, in ICASSP 2013., 2013.

GRDC: River Discharge Data, Glob. Runoff Data Cent. [online] Available from: http://www.bafg.de/GRDC/EN/02_srvcs/21_tmsrs/riverdischarge_node.html (Accessed 28 July 2017), 2017.

Greff, K., Srivastava, R. K., Koutník, J., Steunebrink, B. R. and Schmidhuber, J.: LSTM: A Search Space Odyssey, http://arxiv.org/abs/1503.04069 [online] Available from: http://arxiv.org/abs/1503.04069 (Accessed 18 July 2016), 2015.

Guo, Y., Liu, Y., Oerlemans, A., Lao, S., Wu, S. and Lew, M. S.: Deep learning for visual understanding: A review, Neurocomputing, 187, 27–48, doi:10.1016/J.NEUCOM.2015.09.116, 2016.

Hall, D. K., Riggs, G. A. and Salomonson., V. V.: MODIS/Terra Snow Cover 5-Min L2 Swath 500m. Version 5., Boulder, Colorado USA., 2006.

He, K., Zhang, X., Ren, S. and Sun, J.: Deep Residual Learning for Image Recognition, in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 770–778, IEEE., 2016.

Hiroi, K. and Kawaguchi, N.: FloodEye: Real-time flash flood prediction system for urban complex water flow, in 2016 IEEE SENSORS, pp. 1–3, IEEE., 2016.

Hirschberg, J. and Manning, C. D.: Advances in natural language processing., Science, 349(6245), 261–6, doi:10.1126/science.aaa8685, 2015.

Ho, T. K.: Random decision forests, in Proceedings of 3rd International Conference on Document Analysis and Recognition, vol. 1, pp. 278–282, IEEE Comput. Soc. Press., 1995.

Hochreiter, S.: The Vanishing Gradient Problem During Learning Recurrent Neural Nets and Problem Solutions, Int. J. Uncertainty, Fuzziness Knowledge-Based Syst., 6(2), 107–116, doi:10.1142/S0218488598000094, 1998.

Hochreiter, S. and Schmidhuber, J.: Long Short-Term Memory, Neural Comput., 9(8), 1735–1780, doi:10.1162/neco.1997.9.8.1735, 1997.

Horsburgh, J. S., Morsy, M. M., Castronova, A. M., Goodall, J. L., Gan, T., Yi, H., Stealey, M. J. and Tarboton, D. G.: HydroShare: Sharing Diverse Environmental Data Types and Models as Social Objects with Application to the Hydrology Domain, JAWRA J. Am. Water Resour. Assoc., 52(4), 873–889, doi:10.1111/1752-1688.12363, 2016.

Hrachowitz, M., Savenije, H. H. G., Blöschl, G., McDonnell, J. J., Sivapalan, M., Pomeroy, J. W., Arheimer, B., Blume, T., Clark, M. P., Ehret, U., Fenicia, F., Freer, J. E., Gelfan, A., Gupta, H. V., Hughes, D. A., Hut, R. W., Montanari, A., Pande, S., Tetzlaff, D., Troch, P. A., Uhlenbrook, S., Wagener, T., Winsemius, H. C., Woods, R. A., Zehe, E. and Cudennec, C.: A decade of Predictions in Ungauged Basins (PUB)—a review, Hydrol. Sci. J., 58(6), 1198–1255, doi:10.1080/02626667.2013.803183, 2013.

Hsu, K., Gupta, H. V. and Sorooshian, S.: Artificial Neural Network Modeling of the Rainfall-Runoff Process, Water Resour. Res., 31(10), 2517–2530, doi:10.1029/95WR01955, 1995.

Hsu, K., Gao, X., Sorooshian, S., Gupta, H. V., Hsu, K., Gao, X., Sorooshian, S. and Gupta, H. V.: Precipitation Estimation from Remotely Sensed Information Using Artificial Neural Networks, J. Appl. Meteorol., 36(9), 1176–1190, doi:10.1175/1520-0450(1997)036<1176:PEFRSI>2.0.CO;2, 1997.

Hsu, K., Gupta, H. V., Gao, X., Sorooshian, S. and Imam, B.: Self-organizing linear output map (SOLO): An artificial neural network suitable for hydrologic modeling and analysis, Water Resour. Res., 38(12), 38-1-38–17, doi:10.1029/2001WR000795, 2002.

Huang, W., He, D., Yang, X., Zhou, Z., Kifer, D. and Giles, C. L.: Detecting Arbitrary Oriented Text in the Wild with a Visual Attention Model, in Proceedings of the 2016 ACM on Multimedia Conference - MM ’16, pp. 551–555, ACM Press, New York, New York, USA., 2016.

Indermuhle, E., Frinken, V. and Bunke, H.: Mode Detection in Online Handwritten Documents Using BLSTM Neural Networks, in 2012 International Conference on Frontiers in Handwriting Recognition, pp. 302–307, IEEE., 2012.

Izadinia, H., Russell, B. C., Farhadi, A., Hoffman, M. D. and Hertzmann, A.: Deep Classifiers from Image Tags in the Wild, in Proceedings of the 2015 Workshop on Community-Organized Multimodal Mining: Opportunities for Novel Solutions, pp. 13–18, ACM., 2015.

Jackson, T., O'Neill, P., Njoku, E., Chan, S., Bindlish, R., Colliander, A., Chen, F., Burgin, M., Dunbar, S., Piepmeier, J., Cosh, M., Caldwell, T., Walker, J., Wu, X., Berg, A., Rowlandson, T., Pacheco, A., McNairn, H., Thibeault, M., Martínez-Fernández9, J., González-Zamora, Á., Seyfried10, M., Bosch, D., Starks, P., Goodrich, D., Prueger, J., Su, Z., van der Velde, R., Asanuma, J., Palecki, M., Small, E., Zreda, M., Calvet, J., Crow, W., Kerr, Y., Yueh, S. and Entekhabi, D.: Soil Moisture Active Passive (SMAP) Project Calibration and Validation for the L2/3_SM_P Version 3 Data Products, SMAP Proj. JPL D-93720 [online] Available from: http://nsidc.org/data/docs/daac/smap/sp_l2_smp/pdfs/L2SMP_validated_assess_rpt_rel2_v10a_final.pdf (Accessed 27 July 2017), 2016.

Junk, C., Delle Monache, L., Alessandrini, S., Cervone, G. and von Bremen, L.: Predictor-weighting strategies for probabilistic wind power forecasting with an analog ensemble, Meteorol. Zeitschrift, 24(4), 361–379, doi:10.1127/metz/2015/0659, 2015.

Karpathy, A., Johnson, J. and Fei-Fei, L.: Visualizing and Understanding Recurrent Networks, in ICLR 2016 Workshop. [online] Available from: http://arxiv.org/abs/1506.02078 (Accessed 7 November 2016), 2015.

Karpatne, A., Atluri, G., Faghmous, J. H., Steinbach, M., Banerjee, A., Ganguly, A., Shekhar, S., Samatova, N. and Kumar, V.: Theory-Guided Data Science: A New Paradigm for Scientific Discovery from Data, IEEE Trans. Knowl. Data Eng., 29(10), 2318–2331, doi:10.1109/TKDE.2017.2720168, 2017.

Kawaguchi, K., Kaelbling, L. P. and Bengio, Y.: Generalization in Deep Learning, arXiv:1710.05468 [online] Available from: http://arxiv.org/abs/1710.05468 (Accessed 12 March 2018), 2017.

Kingma, D. P. and Ba, J.: Adam: A Method for Stochastic Optimization, in 3rd International Conference for Learning Representations, San Diego, CA. [online] Available from: http://arxiv.org/abs/1412.6980 (Accessed 30 March 2018), 2014.

Kingma, D. P. and Welling, M.: Auto-Encoding Variational Bayes, in Proceedings of the 2014 International Conference on Learning Representations (ICLR). [online] Available from: http://arxiv.org/abs/1312.6114 (Accessed 24 March 2018), 2013.

Knyazikhin, Y., Glassy, J., Privette, J. L., Tian, Y., Lotsch, A., Zhang, Y., Wang, Y., Morisette, J. T., P.Votava, Myneni, R. B., Nemani, R. R. and Running, S. W.: MODIS Leaf Area Index (LAI) and Fraction of Photosynthetically Active Radiation Absorbed by Vegetation (FPAR) Product (MOD15) Algorithm Theoretical Basis Document, http://eospso.gsfc.nasa.gov/atbd/modistables.html, 1999., 1999.

Koza, J. R.: Genetic Programming: on the Programming of Computers by Means of Natural Selection, MIT Press., 1992.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K. and Herrnegger, M.: Rainfall-Runoff modelling using Long-Short-Term-Memory (LSTM) networks, Hydrol. Earth Syst. Sci. Discuss., 1–26, doi:10.5194/hess-2018-247, 2018.

Krizhevsky, A., Sutskever, I. and Hinton, G. E.: ImageNet classification with deep convolutional neural networks, Proc. 25th Int. Conf. Neural Inf. Process. Syst. - Vol. 1, 1097–1105 [online] Available from: https://dl.acm.org/citation.cfm?id=2999257 (Accessed 10 March 2018a), 2012.

Krizhevsky, A., Sutskever, I. and Hinton, G. E.: ImageNet Classification with Deep Convolutional Neural Networks, in Advances in Neural Information Processing Systems 25, pp. 1097–1105. [online] Available from: https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks (Accessed 30 March 2018b), 2012.

Kumar, D. and Menkovski, V.: Understanding Anatomy Classification Through Visualization, in 30th NIPS Machine learning for Health Workshop. [online] Available from: https://arxiv.org/abs/1611.06284 (Accessed 24 November 2017), 2016.

Laloy, E., Hérault, R., Lee, J., Jacques, D. and Linde, N.: Inversion using a new low-dimensional representation of complex binary geological media based on a deep neural network, Adv. Water Resour., 110, 387–405, doi:10.1016/J.ADVWATRES.2017.09.029, 2017.

Laloy, E., Hérault, R., Jacques, D. and Linde, N.: Training-Image Based Geostatistical Inversion Using a Spatial Generative Adversarial Neural Network, Water Resour. Res., 54(1), 381–406, doi:10.1002/2017WR022148, 2018.

LeCun, Y., Bengio, Y. and Hinton, G.: Deep learning, Nature, 521(7553), 436–444, doi:10.1038/nature14539, 2015.

Lehman, K. J.: Courting the Uncommitted: A Mixed-Methods Study of Undecided Students in Introductory Computer Science Courses, UCLA Electron. Theses Diss. [online] Available from: https://escholarship.org/uc/item/94k326xs (Accessed 31 July 2018), 2017.

Leviathan, Y. and Matias, Y.: Google Duplex: An AI System for Accomplishing Real-World Tasks Over the Phone, Google AI Blog [online] Available from: https://ai.googleblog.com/2018/05/duplex-ai-system-for-natural-conversation.html (Accessed 20 August 2018), 2018.

Liu, Y., Racah, E., Prabhat, Correa, J., Khosrowshahi, A., Lavers, D., Kunkel, K., Wehner, M. and Collins, W.: Application of Deep Convolutional Neural Networks for Detecting Extreme Weather in Climate Datasets, in ACM SIGKDD 2016 Conference on Knowledge Discovery & Data Mining. [online] Available from: http://arxiv.org/abs/1605.01156 (Accessed 21 October 2016), 2016.

Mahendran, A. and Vedaldi, A.: Understanding deep image representations by inverting them, in 2015 IEEE Conference on

Computer Vision and Pattern Recognition (CVPR), pp. 5188–5196, IEEE., 2015.

Matsuoka, D., Nakano, M., Daisuke Sugiyama and Uchida, S.: Detecting Precursors of Tropical Cyclone using Deep Neural Networks, in The 7th International Workshop on Climate Informatics: CI 2017., 2017.

McCabe, M. F., Rodell, M., Alsdorf, D. E., Miralles, D. G., Uijlenhoet, R., Wagner, W., Lucieer, A., Houborg, R., Verhoest, N. E. C., Franz, T. E., Shi, J., Gao, H. and Wood, E. F.: The future of Earth observation in hydrology, Hydrol. Earth Syst. Sci., 21(7), 3879–3914, doi:10.5194/hess-21-3879-2017, 2017.

Mecklenburg, S., Kerr, Y., Font, J. and Hahne, A.: The Soil Moisture and Ocean Salinity Mission - An Overview, in IGARSS 2008 - 2008 IEEE International Geoscience and Remote Sensing Symposium, p. IV-938-IV-941, IEEE., 2008.

Merwade, V. and Ruddell, B. L.: Moving university hydrology education forward with community-based geoinformatics, data and modeling resources, Hydrol. Earth Syst. Sci., 16(8), 2393–2404, doi:10.5194/hess-16-2393-2012, 2012.

Montanari, A., Young, G., Savenije, H. H. G., Hughes, D., Wagener, T., Ren, L. L., Koutsoyiannis, D., Cudennec, C., Toth, E., Grimaldi, S., Blöschl, G., Sivapalan, M., Beven, K., Gupta, H., Hipsey, M., Schaefli, B., Arheimer, B., Boegh, E., Schymanski, S. J., Baldassarre, G. Di, Yu, B., Hubert, P., Huang, Y., Schumann, A., Post, D. A., Srinivasan, V., Harman, C., Thompson, S., Rogger, M., Viglione, A., McMillan, H., Characklis, G., Pang, Z. and Belyaev, V.: "Panta Rhei—Everything Flows": Change in hydrology and society—The IAHS Scientific Decade 2013–2022, Hydrol. Sci. J., 58(6), doi:10.1080/02626667.2013.809088, 2013.

Moradkhani, H., Hsu, K., Gupta, H. V. and Sorooshian, S.: Improved streamflow forecasting using self-organizing radial basis function artificial neural networks, J. Hydrol., 295(1–4), 246–262, doi:10.1016/J.JHYDROL.2004.03.027, 2004.

Mount, N. J., Maier, H. R., Toth, E., Elshorbagy, A., Solomatine, D., Chang, F.-J. and Abrahart, R. J.: Data-driven modelling approaches for socio-hydrology: opportunities and challenges within the Panta Rhei Science Plan, Hydrol. Sci. J., 1–17, doi:10.1080/02626667.2016.1159683, 2016.

Mu, Q., Zhao, M. and Running, S. W.: Improvements to a MODIS global terrestrial evapotranspiration algorithm, Remote Sens. Environ., 115(8), 1781–1800, doi:10.1016/j.rse.2011.02.019, 2011.

Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V., Xia, Y., Nearing, G. S., Mocko, D. M., Peters-Lidard, C. D., Kumar, S. V. and Xia, Y.: Benchmarking NLDAS-2 Soil Moisture and Evapotranspiration to Separate Uncertainty Contributions, J. Hydrometeorol., 17(3), 745–759, doi:10.1175/JHM-D-15-0063.1, 2016.

Pavelsky, T. M., Durand, M. T., Andreadis, K. M., Beighley, R. E., Paiva, R. C. D., Allen, G. H. and Miller, Z. F.: Assessing the potential global extent of SWOT river discharge observations, J. Hydrol., 519, 1516–1525, doi:10.1016/j.jhydrol.2014.08.044, 2014.

Ranzato, M., Poultney, C., Chopra, S. and LeCun, Y.: Efficient learning of sparse representations with an energy-based model, Proc. 19th Int. Conf. Neural Inf. Process. Syst., 1137–1144 [online] Available from: https://dl.acm.org/citation.cfm?id=2976599 (Accessed 30 March 2018), 2006.

Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C. and Fei-Fei, L.: ImageNet Large Scale Visual Recognition Challenge, arXiv:1409.0575 [online] Available from: http://arxiv.org/abs/1409.0575 (Accessed 24 March 2018), 2014.

Sadler, J. M., Goodall, J. L., Morsy, M. M. and Spencer, K.: Modeling urban coastal flood severity from crowd-sourced flood reports using Poisson regression and Random Forest, J. Hydrol., 559, 43–55, doi:10.1016/J.JHYDROL.2018.01.044, 2018.

Samek, W., Binder, A., Montavon, G., Lapuschkin, S. and Muller, K.-R.: Evaluating the Visualization of What a Deep Neural Network Has Learned, IEEE Trans. Neural Networks Learn. Syst., 28(11), 2660–2673, doi:10.1109/TNNLS.2016.2599820, 2017.

Sax, L. J., Lehman, K. J., Jacobs, J. A., Kanny, M. A., Lim, G., Monje-Paulson, L. and Zimmerman, H. B.: Anatomy of an Enduring Gender Gap: The Evolution of Women's Participation in Computer Science, J. Higher Educ., 88(2), 258–293, doi:10.1080/00221546.2016.1257306, 2017.

Schmidhuber, J.: Deep learning in neural networks: An overview, Neural Networks, 61, 85–117, doi:10.1016/j.neunet.2014.09.003, 2015.

Schwalm, C. R., Anderegg, W. R. L., Michalak, A. M., Fisher, J. B., Biondi, F., Koch, G., Litvak, M., Ogle, K., Shaw, J. D., Wolf, A., Huntzinger, D. N., Schaefer, K., Cook, R., Wei, Y., Fang, Y., Hayes, D., Huang, M., Jain, A. and Tian, H.: Global patterns of drought recovery, Nature, 548(7666), 202–205, doi:10.1038/nature23021, 2017.

Settles, B.: Active Learning, in Synthesis lectures on artificial intelligence and machine learning, edited by R. J. Brachman, W. W. Cohen, and T. G. Dietterich, Norgan & Claypool., 2012.

Shen, C.: A trans-disciplinary review of deep learning research and its relevance for water resources scientists, Water Resour. Res. (in Press. [online] Available from: http://arxiv.org/abs/1712.02162 (Accessed 3 January 2018), 2018.

Simonyan, K. and Zisserman, A.: Very Deep Convolutional Networks for Large-Scale Image Recognition, in ICLR 2015. [online] Available from: http://arxiv.org/abs/1409.1556 (Accessed 25 August 2018), 2014.

Snelson, E. and Ghahramani, Z.: Sparse Gaussian Processes using Pseudo-inputs, Adv. Neural Inf. Process. Syst., 18, 1257--1264, 2006.

Srinivasan, M.: Hydrology from space: NASA's satellites supporting water resources applications, Water Forum III Droughts Other Extrem. Weather Events [online] Available from: http://www.jsg.utexas.edu/ciess/files/Srinivasanetal_TWF_Oct14_Final.pdf (Accessed 12 July 2016), 2013.

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I. and Salakhutdinov, R.: Dropout: A Simple Way to Prevent Neural Networks from Overfitting, J. Mach. Learn. Res., 15, 1929–1958 [online] Available from: http://jmlr.org/papers/v15/srivastava14a.html (Accessed 28 November 2015), 2014.

Stallkamp, J., Schlipsing, M., Salmen, J. and Igel, C.: The German Traffic Sign Recognition Benchmark: A multi-class classification competition, in The 2011 International Joint Conference on Neural Networks, pp. 1453–1460, IEEE., 2011.

Stollenga, M. F., Byeon, W., Liwicki, M. and Schmidhuber, J.: Parallel multi-dimensional LSTM, with application to fast biomedical volumetric image segmentation, Proc. 28th Int. Conf. Neural Inf. Process. Syst. - Vol. 2, 2998–3006 [online] Available from: https://dl.acm.org/citation.cfm?id=2969574 (Accessed 16 October 2017), 2015.

Szegedy, C., Wei Liu, Yangqing Jia, Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A.: Going deeper with convolutions, in 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1–9, IEEE., 2015.

Tao, Y., Gao, X., Hsu, K., Sorooshian, S. and Ihler, A.: A Deep Neural Network Modeling Framework to Reduce Bias in Satellite Precipitation Products, J. Hydrometeorol., doi:JHM-D-15-0075.1, 2016.

Tao, Y., Gao, X., Ihler, A., Sorooshian, S., Hsu, K., Tao, Y., Gao, X., Ihler, A., Sorooshian, S. and Hsu, K.: Precipitation Identification with Bispectral Satellite Information Using Deep Learning Approaches, J. Hydrometeorol., 18(5), 1271–1283, doi:10.1175/JHM-D-16-0176.1, 2017.

Tao, Y., Hsu, K., Ihler, A., Gao, X., Sorooshian, S., Tao, Y., Hsu, K., Ihler, A., Gao, X. and Sorooshian, S.: A Two-Stage Deep Neural Network Framework for Precipitation Estimation from Bispectral Satellite Information, J. Hydrometeorol., 19(2), 393–408, doi:10.1175/JHM-D-17-0077.1, 2018.

Tibshirani, R. and Tibshirani, R.: Regression Shrinkage and Selection Via the Lasso, J. R. Stat. Soc. Ser. B, 58, 267--288 [online] Available from: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.7574 (Accessed 4 August 2017), 1994.

Troch, P. A., Carrillo, G., Sivapalan, M., Wagener, T. and Sawicz, K.: Climate-vegetation-soil interactions and long-term hydrologic partitioning: signatures of catchment co-evolution, Hydrol. Earth Syst. Sci., 17(6), 2209–2217, doi:10.5194/hess-17-2209-2013, 2013.

Vandal, T., Kodra, E., Ganguly, S., Michaelis, A., Nemani, R. and Ganguly, A. R.: DeepSD: Generating High Resolution Climate Change Projections through Single Image Super-Resolution, in 23rd ACM SIGKDD Conference on Knowledge

Discovery and Data Mining. [online] Available from: http://arxiv.org/abs/1703.03126 (Accessed 2 December 2017), 2017.

Vondrick, C., Pirsiavash, H. and Torralba, A.: Anticipating Visual Representations from Unlabeled Video, in CVPR 2016. [online] Available from: http://arxiv.org/abs/1504.08023 (Accessed 20 August 2018), 2016.

Voosen, P.: The AI detectives, Science (80-.)., 357(6346) [online] Available from: http://science.sciencemag.org/content/357/6346/22 (Accessed 7 July 2017), 2017.

Wagener, T., Sivapalan, M., Troch, P. a., McGlynn, B. L., Harman, C. J., Gupta, H. V., Kumar, P., Rao, P. S. C., Basu, N. B. and Wilson, J. S.: The future of hydrology: An evolving science for a changing world, Water Resour. Res., 46(5), 1–10, doi:10.1029/2009WR008906, 2010.

Wagener, T., Kelleher, C., Weiler, M., McGlynn, B., Gooseff, M., Marshall, L., Meixner, T., McGuire, K., Gregg, S., Sharma, P. and Zappe, S.: It takes a community to raise a hydrologist: the Modular Curriculum for Hydrologic Advancement (MOCHA), Hydrol. Earth Syst. Sci., 16(9), 3405–3418, doi:10.5194/hess-16-3405-2012, 2012.

Wahr, J.: Time-variable gravity from GRACE: First results, Geophys. Res. Lett., 31(11), L11501, doi:10.1029/2004GL019779, 2004.

Wahr, J., Swenson, S. and Velicogna, I.: Accuracy of GRACE mass estimates, Geophys. Res. Lett., 33(6), L06401, doi:10.1029/2005GL025305, 2006.

Wang, H., Skau, E., Krim, H. and Cervone, G.: Fusing Heterogeneous Data: A Case for Remote Sensing and Social Media, IEEE Trans. Geosci. Remote Sens., 1–13, doi:10.1109/TGRS.2018.2846199, 2018.

WaterML2: WaterML2, A global standard for hydrological time series, [online] Available from: http://www.waterml2.org/ (Accessed 23 August 2018), 2018.

Xiong, W., Droppo, J., Huang, X., Seide, F., Seltzer, M., Stolcke, A., Yu, D. and Zweig, G.: Achieving Human Parity in Conversational Speech Recognition, Microsoft Res. Tech. Rep. MSR-TR-2016-71. arXiv1610.05256 [online] Available from: http://arxiv.org/abs/1610.05256 (Accessed 24 March 2018), 2016.

Yosinski, J., Clune, J., Nguyen, A., Fuchs, T. and Lipson, H.: Understanding Neural Networks Through Deep Visualization, in Deep Learning Workshop, 31 st International Conference on Machine Learning, Lille, France. [online] Available from: http://arxiv.org/abs/1506.06579 (Accessed 19 November 2017), 2015.

Yu, K.-H., Zhang, C., Berry, G. J., Altman, R. B., Ré, C., Rubin, D. L. and Snyder, M.: Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features, Nat. Commun., 7, 12474, doi:10.1038/ncomms12474, 2016.

Zen, H. and Sak, H.: Unidirectional long short-term memory recurrent neural network with recurrent output layer for low-latency speech synthesis, in 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4470–4474, IEEE., 2015.

Zhang, D., Lindholm, G. and Ratnaweera, H.: Use long short-term memory to enhance Internet of Things for combined sewer overflow monitoring, J. Hydrol., 556, 409–418, doi:10.1016/J.JHYDROL.2017.11.018, 2018.

Zhu, X. X., Tuia, D., Mou, L., Xia, G.-S., Zhang, L., Xu, F. and Fraundorfer, F.: Deep learning in remote sensing: a review, IEEE Geosci. Remote Sens. Mag. [online] Available from: http://arxiv.org/abs/1710.03959 (Accessed 20 October 2017), 2017.