From Data Modeling to Algorithmic Modeling in the Big Data Era: Water Resources Security in the Asia-Pacific Region under Conditions of Climate Change

Jason Levy and Ross Prizzia

Abstract Advances in computing technologies allow machine learning algorithms to automatically, repeatedly and quickly apply complex mathematical calculations to water resources and environmental security challenges. The concomitant increase in "big data" research, development, and applications is also driving the popularity of real-time automated model building and data mining for these security problems under conditions of climate change. The last decade has seen considerable growth in the theory and application in Artificial Intelligence (AI). It is shown that machine learning, a subset of AI, constitutes a data analysis method that focuses on the development of algorithms that can iteratively learn from data to uncover previously "hidden insights" for environmental security managers in the Asia Pacific. It is concluded that deep machine learning (i.e. deep learning) can help to reduce losses to ecosystems, livelihoods, and businesses. In particular, these losses can be more likely prevented and minimized through the use of data and algorithmic modeling that improves community resilience by institutionalizing sustainable hazard mitigation within accepted processes of water resources community planning and economic development *before* disasters happen. Key environmental threats including foods, population extinction, water quality and climate change are considered. The difference between the algorithmic modeling and data modeling cultures are summarized with reference to the schools in which they originate, the assumptions they work on, the type of data they deal with, and the techniques used.

Keywords Big data • Sustainability • Resilience • Algorithmic modeling

---

J. Levy (✉) • R. Prizzia
University of Hawaii-West Oahu, Kapolei, HI, USA
e-mail: jlevy@hawaii.edu; rprizzia@hawaii.edu

A.J. Masys, L.S.F. Lin (eds.), *Asia-Pacific Security Challenges*, Advanced Sciences and Technologies for Security Applications,
https://doi.org/10.1007/978-3-319-61729-9_9

1 Introduction

The last decade has seen considerable growth in the theory and application in Artificial Intelligence (AI). Machine learning is a type of AI in which an algorithm (i.e. a “machine” comprised of equations and matrices) analyzes data (without being limited to a particular representation of data) in order to learn a model and look for patterns. In this way, the algorithm both learns from and makes predictions on data, without explicitly relying on rules-based programming. In the broadest sense, these fields aim to ‘learn valuable information’ about the environment within which the system operates. In other words, machine learning constitutes a data analysis method that focuses on the development of algorithms that iteratively learn from data to uncover previously “hidden insights”. As these models are exposed to new data they are able to independently adapt and learn from previous computations, historical relationships and data trends to “produce reliable, repeatable decisions and results” (SAS 2016). In this way, machine learning automates the analytical model building process. By training a machine learning model with existing data (model training) one fits the model parameters.

Resurging interest in machine learning is due to a number of factors. First, there has been an exponential growth in the volumes and varieties of available data (i.e. the ubiquity of big data in every field). Computers are storing terabytes of data which are now generated at an unprecedented rate from many sources (including telescopes scanning the skies, sensors collecting pollution and natural resources data, social media feeds, industrial quality control data, and information about the commercial preferences of consumers). Second, there is an urgent need to automate the analysis and comprehension of the data. Physicians must scan thousands of images in search of tumors, astronomers attempt to recognize novel objects based on planetary and stellar images stored on tape or disk while geneticists study micro-array data to understand genetic effects. Third, the data is more publically available due to the rise in the open-source and open-content movements (from crowd-sourcing to open-source policy and governance). Fourth, there has been a rise in cheaper and more powerful computational processing, together with less expensive data storage. Advances in computing technologies provide modern machine learning algorithms with the ability to automatically, repeatedly and quickly apply complex mathematical calculations. Fifth, the concomitant increase in “big data” research, development, applications and its intelligence is also driving the popularity of real-time automated model building and data mining.

Machine learning provides environmental security managers in the Asia Pacific with the ability to perform real-time optimization (based on the environmental datasets available) as the programs can teach themselves to learn, look for patterns and make predictions. In deep machine learning (i.e. deep learning), hierarchical representations of the observational data are calculated, where the higher-level features are defined from lower-level ones using multiple information processing stages (Schmidhuber 2015). Machine learning approaches now are capable of quickly and automatically (without human intervention) producing realistic models, analyzing complex datasets and delivering high-value predictions to accurately guide decisions – even on a very large scale. Recent years have seen many widely publicized examples of machine learning applications including image processing (e.g. the detection of tumors in x-rays and endangered marine species in the environment), self-driving vehicles, online recommendation offers (e.g. offers provided by Amazon and Netflix), and fraud detection.

Algorithms have been developed to process large, complex datasets and to deal with the uncertainty in the gathered high dimensional data. They can also be used as a more accurate and informative alternative to data modeling on smaller data sets (Breiman 2003). In the early stages of research in machine learning and related areas, similar techniques were discovered in relatively isolated academic silos, but there is now a broader collaboration among various research communities. The machine learning paradigm is anticipated to become even more pervasive and disruptive than the previous technologic waves of personal computing, the internet and mobile smartphones. In the early stages of research in machine learning and related areas, advances were carried out in relatively isolated academic silos, but there is now a broader collaboration among various research communities. The machine learning (ML) paradigm is anticipated to become even more pervasive and disruptive than the previous technologic waves of personal computing, the internet and mobile smartphones.

Sometimes conflated with the data mining subfield (which focuses more on exploratory data analysis), the main two subfields of machine learning are supervised learning and unsupervised learning. In supervised learning, a machine learning algorithm is trained using a “training dataset” (i.e. prototypical/representative/exemplar situations for which the desired output is known). “Validation data” (not encountered during training) is then used to test the algorithm’s ability to predict the output. Here, the focus is on accurate prediction and the “generalization performance” of a method to previously unseen data, i.e. the method ‘generalizes’ to this unseen data.

On the other hand, in unsupervised learning the aim is to find accurate and dense (compressed) data descriptions. Conceptually, there are two sources of data for model training and evaluation: test data and training data. The training data (only) is used to set and adjust the model parameters. Note that if the test data was used to parameterize the model there would no longer be an independent evaluation of model performance. An unbiased estimate of this “generalization performance” can be obtained by measuring the test data performance of the trained model. However, the test data must be generated from the same underlying process that generated the training data. While the Bayesian statistical approach is not the only paradigm for describing machine learning and information processing, it is often convenient to consider the process of machine learning as updating (learning) the prior and posterior distributions as new data arrives. The more data that is input into algorithms the better the resulting predictions and decisions. However, the increased reliance and pace of machine learning advances bring new challenges. For example, the training data in machine learning applications use historic data to predict future trends and needs, providing old answers to new questions. This so called “algorithmic

Table 1 Concepts pertaining to algorithmic modeling and data modeling

Algorithmic modeling	Data modeling
Artificial intelligence/machine learning/ computational statistics/statistical learning/ computational intelligence/soft computing	Statistical modeling
Training, learning and automation	Fitting
Supervised learning	Regression/classification
Networks, graphs Deep learning (hierarchical representations and Hyperparameters)	Model
Weights	Parameters (numerical characteristic of a population)
Data mining algorithm, machine learning algorithm	Predictive modeling
Generalization	Test set performance
Unsupervised learning/data mining	Exploratory analysis, density estimation, clustering

determinism" problem tends to reinforce prior biases, reproduce established patterns of behavior, and deepen social divisions.

The foundation of the scientific process involves working with data and checking theory against data. While in many situations data models are the most appropriate approach to solve an environmental security problem, the field of statistics has displayed an unjustified vested interest in data models, even given large quantities of error-filled data (Dempster 1998; Breiman 2003). In order to remain a relevant and creative field, the environmental security community must reach out to other disciplines for collaborative work and apply tools to solve real world problems (rather than focus on the type of data model that can be created). While the best solution could be a combination of an algorithmic model and a data model (or maybe either in isolation) scientific rigor requires being open to the use a wide variety of tools. Algorithmic modeling and data modeling use different names for similar concepts as shown in Table 1.

2 A Tale of Two Paradigms for Modeling Security Challenges in the Asia-Pacific

When analyzing data, statistics is used to achieve two goals: prediction and inference: Inferential statistical analysis includes testing hypotheses and deriving estimates to infer properties about a population and obtain information about the underlying data mechanism. There are two paradigms in the statistical modeling community:

- Generative modeling culture which uses statistical data models (data modeling) and
- Predictive modeling culture which uses algorithms (algorithmic modeling).

The selection of either data modeling or algorithmic modeling must be justified based on the nature of the problem and on the data. Statistical data models, overwhelmingly favored by the statistical community, assumes that a given stochastic data model generates the data ("generative modeling"). Data modeling involves developing stochastic models which fit the data and then making inferences about the data-generating mechanism to deduce properties of the underlying distribution. Breiman (2003) notes the over-reliance of the statistical community on data models to the exclusion of a more diverse set of tools. This has resulted in a number of shortcomings in the statistical community: the inability to solve some of the most complex, interesting and important contemporary problems.

On the other hand, the theory of algorithms prioritizes prediction (the goal of modeling is predictive accuracy) and assumes the data generation mechanism to be unknown. Accordingly, data models are rarely used in the algorithmic community. While industrial statisticians have used algorithmic modeling for decades (Daniel and Wood 1971) there is little work in statistics that focuses on predictive modeling. As an exception Grace Wahba's smoothing spline algorithm research is built on reproducing kernels in Hilbert Space (Wahba 1990). Machine learning approaches have also made significant impacts in the interdisciplinary field of bioinformatics by facilitating discoveries in genomics and proteomics. Zhang and Singer (1999) applied recursive partitioning in the health sciences.

Beginning in the mid-1980s, important new machine learning algorithms for data fitting became available including neural nets and decision trees. A burgeoning predictive modeling community began using the new algorithmic modeling tools to solve complex prediction problems that are less applicable to data models: self-driving vehicles, speech recognition, image recognition, nonlinear time series prediction, handwriting recognition, and prediction in financial markets. The "black box" theory of the algorithmic approach observes a set of inputs $(X_1, X_2, \ldots, X_n)$, a set of outputs $(Y_1, Y_2, \ldots, Y_n)$ and seeks to find an algorithm (f_x) that will be a good predictor of $(Y_1, Y_2, \ldots, Y_n)$ in the test set. One assumption made in the theory of algorithms is that the data $(X_1, X_2, \ldots, X_n)$ is drawn independent and identically distributed (iid) from an unknown multivariate distribution. The "strength" of f_x as a predictor provides their predictive accuracy. In the case of iterative algorithms, convergence is desired (i.e. candidate solutions for each iteration tend to get closer and closer to the desired solution). For example, in their work on Classification and Regression Trees (CART), Breiman et al. (1984) demonstrates the asymptotic convergence of the CART algorithm to the Bayes risk by letting the trees grow as the sample size increases. Vladimir Vapnik's support vector machines – based on the construction informative bounds on the generalization error (infinite test set error) of classification algorithms—have proved to be more accurate predictors in classification and regression then neural nets (Vapnik 1995, 1998). Since the mid-1980s advances in the methodology of machine learning approaches has been explosive, together with concomitant increases in predictive accuracy.

The field of environmental security is witnessing an inflection point in which artificial intelligence (AI) becomes the next technologic shift: in many environmental security fields AI paradigms are replacing data modeling approaches in general and statistical models in particular. Environmental managers can leverage AI technology in multiple ways when searching for environmental patterns. Within the data analytics field, AI focuses on making predictions (known as predictive analytics in commercial settings). While machine learning makes fewer assumptions than statistical modeling there is not a clear dichotomy between the two approaches. Key lessons in the development of algorithmic models are discussed by Breiman (2003): the multiplicity of good models (Rashomonic effect), the conflict between simplicity and accuracy (Occam's razor) and the curse (or blessing) of dimensionality (Bellman).

3 Stochastic Water Modeling in the Asia-Pacific Region: Climate Change, Environmental Quality and Extinction Risk

Water-related problems are particularly acute in Asia. Although Asia is home to more than half of the world's population, it has less freshwater—3920 cubic meters per person per year—than any continent other than Antarctica. Almost two-thirds of global population growth is occurring in Asia, where the population is expected to increase by nearly 500 million people within the next 10 years. Asia's rural population will remain almost the same between now and 2025, but the urban population is likely to increase by a staggering 60%. As population growth and urbanization rates in Asia rise rapidly, stress on the region's water resources is intensifying. Climate change is expected to worsen the situation significantly. Experts agree that reduced access to freshwater will lead to a cascading set of consequences, including impaired food production, the loss of livelihood security, large-scale migration within and across borders, and increased economic and geopolitical tensions and instabilities. Over time, these effects will have a profound impact on security throughout the region (DeRusha et al. 2017).

Southeast Asian countries such as Cambodia which border the Mekong River are extremely vulnerable to flooding. In the last decade, Cambodia has halved its poverty rate and improved the living conditions of its population. However, because of extreme climate events that regularly descend on the country, Cambodia remains one of the most disaster-vulnerable countries in Southeast Asia. In 2013 alone, losses caused by floods added up to USD $356 million. However, disasters and climate change also present an opportunity to promote what the United Nations Development Program (UNDP) refers to as "risk-informed development" (UNDP 2015). Communities affected by climate disasters learn to work together to create effective, multi-disciplinary approaches to respond to and recover from disasters as well as promote disaster risk reduction.

In 2013, a combination of heavy rains and the swelling of the Mekong River caused widespread damage to infrastructure and crops, the death of 168 people, most of them children, and devastation to 20 provinces. Following the floods, the Cambodian government requested that UNDP work with various partners to carry out a post flood early recover needs assessment. Drawing on the expertise of UNDP's country office, as well as the skills and knowledge of government partners, NGOs, and civil society organizations, measurements of the flood damage and an assessment that clearly articulated the needs of the various communities were accomplished (UNDP 2015).

Sea-level rise, erosion, and coastal flooding are some of the greatest challenges facing humanity from climate change. Recently at least five reef islands in the remote Solomon Islands have been lost completely to sea-level rise and coastal erosion, and a further six islands have been severely eroded. This is the first scientific evidence that confirms the numerous anecdotal accounts from across the Pacific of the dramatic impacts of climate change on coastlines and people (Albert et al. 2016). These islands lost to the sea range in size from one to five hectares. They supported dense tropical vegetation that was at least 300 years old. Nuatambu Island, home to 25 families, has lost more than half of its habitable area, with 11 houses washed into the sea since 2011. This is the first scientific evidence, that confirms the numerous anecdotal accounts from across the Pacific of the dramatic impacts of climate change on coastlines and people (Albert et al. 2016). Previous studies examining the risk of coastal inundation in the Pacific region have found that islands can actually keep pace with sea-level rise and sometimes even expand. However, these studies have been conducted in areas of the Pacific with rates of sea level rise of 3–5 mm per year – broadly in line with the global average of 3 mm per year. For the past 20 years, the Solomon Islands have been a hotspot for sea-level rise. Here the sea has risen at almost three times the global average, around 7–10 mm per year since 1993. This higher local rate is partly the result of natural climate variability. These higher rates are in line with what we can expect across much of the Pacific in the second half of this century as a result of human-induced sea-level rise.

Many areas will experience long-term rates of sea-level rise similar to that already experienced in Solomon Islands in all but the very lowest-emission scenarios. Natural variations and geological movements will be superimposed on these higher rates of global average sea level rise, resulting in periods when local rates of rise will be substantially larger than that recently observed in Solomon Islands. We can therefore see the current conditions in Solomon Islands as an insight into the future impacts of accelerated sea-level rise. The study included the coastlines of 33 reef islands using aerial and satellite imagery from 1947 to 2015. This information was integrated with local traditional knowledge, radiocarbon dating of trees, sea-level records, and wave models (Albert et al. 2016).

Other new sea rise research relevant to the Asia Pacific revealed that important focal parameters of tsunamigenic earthquakes, particularly fault dip direction, can be extracted from tsunami-borne EM fields with the potential of electromagnetic (EM) fields being used in tsunami early warning. Knowing the direction in which the fault dips could be helpful for tsunami early warning, as the direction sometimes determines whether a rise wave or a backwash hits a particular costal area. By 2100, a realistic low-end projection is an additional 1 foot of sea level rise globally, with an upper end projection of 4 feet or higher causing sea level rise which not only threatens infrastructure over the long-term but a rising sea exacerbates the flooding effects of storm surges and high tides (Walsh et al. 2014). During severe storm events, water that surges onto US military installations from the sea can damage installation infrastructure or training areas and risk from sea level rise and storm surge are not limited to low-lying islands and atolls. While portions of Guam are well above sea level, most of the infrastructure is on or near the coasts and remains exposed to sea level rise and storm surge (ORLN 2015). Potentially heavier and more frequent precipitation will also affect installation maintenance costs and require additional flood or erosion control measures. Military capabilities and readiness are degraded when airstrips, piers, roadways, communication, energy and other infrastructure are unavailable due to flooding or erosion. Losing access to these facilities is potentially equivalent to temporary anti-access to an area, requiring the US Department of Defense (USDOD) to consider the capability thresholds required in the Area of Responsibility (military geographic region) (AOR), and design resiliency and redundancy into infrastructure plans to maintain these thresholds (Walsh et al. 2014). Pacific installations also need to be especially resilient to natural disasters such as tropical cyclones, as they not only need to maintain capabilities after an event, but often serve as a base of operations (Walsh et al. 2014).

Global climate variability and change is increasing the frequency and severity of natural disaster events and environmental security risks in the Asia-Pacific region. Climate change threatens the fabric of life for people throughout the Asia Pacific – it affects key health, environmental and social dimensions including access to clean water, food production, and the sustainability of ecological systems and the urban built environment. Severe weather is predicted to become more frequent and destructive. For example, higher temperatures may lead to droughts, crop failures, food insecurity, the mass migration of "climate refugees" across international borders and increased conflict among nations competing for scarce resources (particularly among upstream and downstream nations). Warmer air holds more moisture, which portends record-breaking rainfall and more intense storms. Even the conservative estimates for the rising temperatures and changing ocean levels will cause some significant issues in the Asia-Pacific Region. Sea rise for coastal cities may be extremely damaging, especially as people and population densities continue to increase in coastal areas of the Asia-Pacific. Emergency managers, security professionals and governments must promote climate adaptation and mitigation measures that protect communities in the Asia-Pacific region. In particular, citizens of states in the Asia Pacific are highly vulnerable to the negative impacts of climate change and many inhabitants are already suffering from increasing heavy rainfall and floods, water shortages, storm surge, hurricanes, coastal erosion and droughts. Accordingly, the inhabitants of many regions have become "canaries in a coal mine" with respect to the adverse effects of climate change. For example, a small community living in the Pacific island chain of Vanuatu has become the first in the world to be formally relocated as a result of climate change.

Climate change may also contribute to a more energetic hydrological cycle, leading to more intense and frequent storms that cause runoff which carries pollutants from industrial and agricultural areas to nearby waterways. Heavy storms can also overwhelm the sewer system and send raw sewage and polluted stormwater into nearby streams and rivers. There are additional threats to rivers from climate change in the Asia-Pacific region: lower flows, rising temperature, more frequent droughts and changing precipitation patterns contribute to higher pollution levels (less water to dilute pollutants in rivers, lakes and streams). Higher water temperatures also lower dissolved oxygen levels and cause algal blooms which can kill marine life and degrade ecosystems (biochemical processes and organism growth rates are regulated to a large extent by temperature). An increase of atmospheric carbon dioxide and/or other greenhouse gases is projected to cause climate warming in the Asia-Pacific region which could significantly alter Dissolved Oxygen (DO) characteristics in water bodies. These changes are in turn expected to have a profound effect on indigenous fish populations. The earliest models of water quality involved two linear deterministic differential equations of biochemical oxygen demand (BOD) and dissolved oxygen (DO) based on the pioneering work of Streeter and Phelps (1925). The seminal Streeter-Phelps equations form the foundation for many of today's sophisticated water quality models in the Asia-Pacific region. A water quality mode1 typically describes the chemical, physical, and biological processes that occur in a water body, such as the reaction of chemical constituents and the uptake of nutrients by living organisms. The propagation of uncertainty through a stochastic water quality model is shown in Fig. 1.

Consider a river with multiple reaches and a treatment plant discharging at the head of the reach, as illustrated in Fig. 2. Mass or energy balance equations are often used to describe the dynamics of constituent concentrations of natural water bodies. The health of aquatic systems (algae, fish, micro-organisms, etc.), aesthetics (such as odor and color), potability, taste, and so on depend upon the resulting concentrations of dissolved oxygen. DO levels naturally cycle over the course of a day (and throughout the year) as shown in Fig. 3. In the steady state conditions resulting from the natural balance of various chemical and biological processes, the DO concentration fluctuates about a saturation concentration (*Os*). Whenever

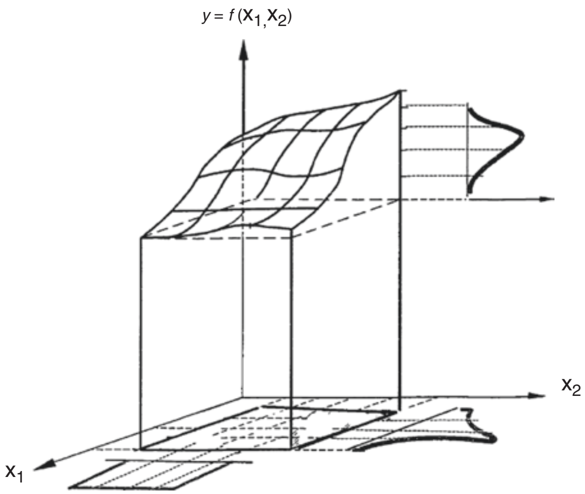


Fig. 1 Propagation of the continuous probability distributions through a water quality mode1

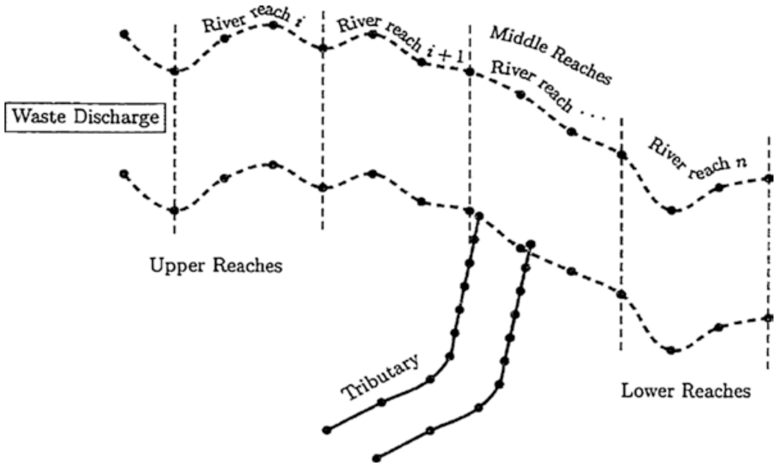


Fig. 2 River with multiple reaches

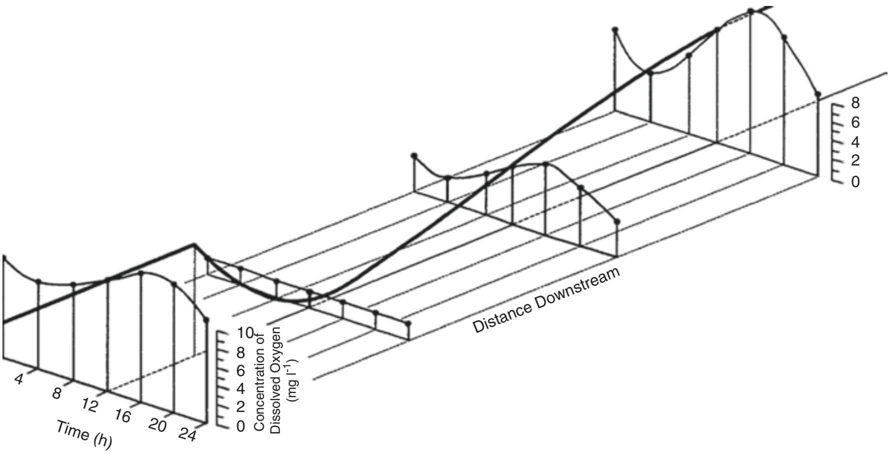


Fig. 3 Dissolved oxygen (DO) vs time

untreated waste waters are discharged into the stream, the concentration of DO may be adversely affected. Consider a steady-state stochastic DO models which address three water quality constituents: DO, carbonaceous biochemical oxygen demanding substances (CBOD) and nitrogenous oxygen demanding substances (NOD). Coupled CBOD-NODDO reactions are an important component of water quality modeling. It is known that CBOD is increased by nonpoint load sources of carbon *(S_c)* and decreased by oxidation *(k_i)*, sedimentation, and adsorption *(L)*. NOD is also increased by nonpoint load sources (S_N) and decreased by oxidation *(k_2)*. Finally, DO is supplied by re-aeration (k_3) and photosynthesis (P) and decreased by respiration (R), CBOD (k_l), and NOD *(k_2)*. The following three deterministic differential equations have been used for describing the water quality of a river (Zielinski 1988):

$$\frac{dC}{dt} = -\left(k_1 + \mathcal{L}\right)C + S_C \tag{1}$$

$$\frac{dN}{dt} = -k_2 N + S_N$$

$$\frac{dO}{dt} = k_3\left(O_s - O\right) - k_1 C - k_2 N + P - R$$

Where the photosynthetic term, P, in Eq. 5.4 is represented by

$$P_m \sin\left[v\left(t + \varnothing\right)\right] \tag{2}$$

These equations describe how a spike input of CBOD, NOD (or other organic material) generates the classic transient DO "sag curve" (Fig. 3). In Eq. 2, P_m is the maximum rate of photosynthesis. The units of the state variables in Eq. 1 are now defined:

- C is the carbonaceous biochemical oxygen demand (CBOD) in mg/L;
- N is the nitrogenous oxygen demand (NOD) in mg/L;
- O is the dissolved oxygen concentration (DO) in mg/L

Next, the four decay constants are defined:

- k_1 is the CBOD decay rate per day;
- L is the sedimentary and adsorption loss rate for CBOD per day;
- k_2 is the decay rate of NOD per day; and k_3 is the reaeration rate per day.

In addition, O_s is the saturation concentration of oxygen in mg/L while R is the loss rate of DO due to respiration in mg/L/day. Finally, S_c and S_N are the nonpoint source loads of carbon and nitrogen respectively in mg/L/day. Replacing the state variables C, N, and O with X_1, X_2 and X_3 respectively, Eq. 1 can be re-written in matrix form:

$$\frac{dx}{dt} = Ax + b \tag{3}$$

Where the 3×1 column vector x is

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} \tag{4}$$

The 3×3 matrix A is given by

$$\begin{bmatrix} -\left(k_1 + \mathcal{L}\right) & 0 & 0 \\ 0 & -k_2 & 0 \\ -k_1 & -k_2 & -k_3 \end{bmatrix} \tag{5}$$

While the 3×1 column vector b is

$$\begin{bmatrix} S_C \\ S_N \\ P_m \sin\left[v\left(t+\varnothing\right)\right] - R + k_3 O_S \end{bmatrix} \tag{6}$$

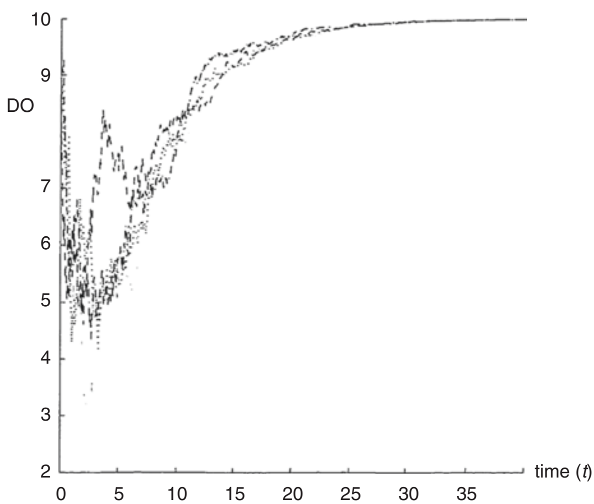
Accordingly, Eq. 3 becomes

$$\frac{d}{dt}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} -\left(k_1 + \mathcal{L}\right) & 0 & 0 \\ 0 & -k_2 & 0 \\ -k_1 & -k_2 & -k_3 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} S_C \\ S_N \\ P_M \sin\left[v\left(t+\theta\right)\right] - R + k_3 O_S \end{bmatrix} \tag{7}$$

Sample paths of DO vs. time are calculated in MATLAB and shown in Fig. 4 which illustrates four sample paths of DO vs time. Given any specific location along a river, the probability density function (pdf) of DO that corresponding to this location can be produced. Figure 4 illustrates the DO density functions corresponding to locations X_1 and X_2.

Climate change may already be causing a reduction in the amount of the life-giving gas that is dissolved in sea water in the Asia-Pacific region, thereby causing

Fig. 4 Four paths for dissolve oxygen vs time



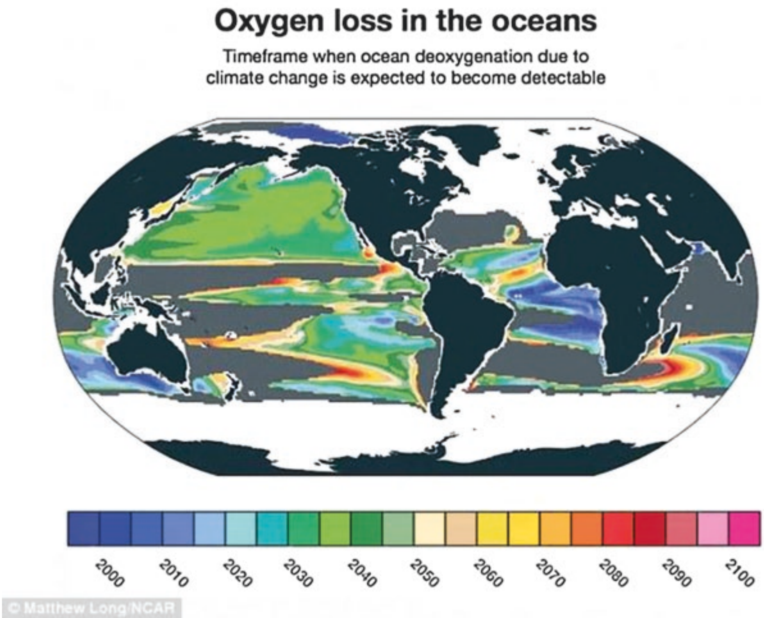


Fig. 5 Climate change and oxygen loss in the oceans (Long et al. 2016)

many sea creatures, including fish, squid, crabs and shellfish, to struggle to breathe (Long et al. 2016). Tropical regions are suffering from oxygen loss while much of the Pacific Ocean will be hit in by around 2040 as shown in Fig. 5. The effects of this loss of oxygen will start become noticeable across widespread areas of the oceans between 2030 and 2040.

4 Statistical Models of Extinction in the Asia Pacific Region

While the threats of global warming pose a grave threat to the well-being and survival of species in the Asia Pacific region, many models for population dynamics fail to consider the risk of extinction. Consider the exponential population growth model for a population P:

$$P' = aP,\ a > 0 \tag{8}$$

Since the exponential growth pattern cannot continue indefinitely, the *logistic* (*inhibited growth*) *model* is often used:

$$P' = bP\left(L - P\right) = aP - bP^2 \text{ with } \mathrm{L} = \mathrm{a}/\mathrm{b} \tag{9}$$

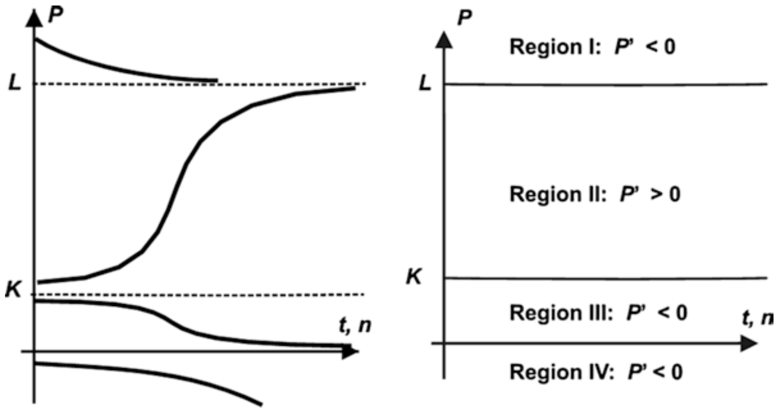


Fig. 6 Extinction and growth models. Solution patterns for logistic-type growth and for extinction (**a**) growth curves (**b**) growth regions. Different regions of the t-P plane for logistic-type growth and extinction

When $a > 1$ and b is a positive constant less than a the solution has the S-shaped logistic pattern shown in Fig. 6a where the horizontal asymptote L is referred to as the *carrying capacity of the environment*. Note that the change in the population depends on both the size of the population P and the difference between the current population and L (also referred to as the *limit to growth* and the *maximum sustainable population*). The inflection point seen in the S-shaped logistic curve of Fig. 6a, shows where the population is growing most rapidly and occurs at $P = ½\,L = ½(a/b)$. Three other growth possibilities are shown in Fig. 6a. Under conditions of climate change it is important to consider how to model the extinction of species. A catastrophic change in the local conditions such as a significant rise in temperature can change to such an extent that a species may become extinct. Accordingly, the classic logistic growth model should take into account the risk that a population decays toward extinction. As shown in Fig. 4a there is a *minimum sustainable population*, *K,* a level below which a species cannot be maintained. Throughout the Asia Pacific region, experts predict that populations of many different species are dropping below K for a number of reasons linked to global climate variability and change. Once this occurs, it is mathematically expected that the population begins to decay toward zero. We now highlight three equilibrium levels, one for a zero population, another corresponding to the maximum sustainable population L, and a third corresponding to the minimum sustainable population K, These three equilibria create four regions in the *t-P* plane for logistic-growth and extinction, as illustrated in Fig. 6b. Consider the behavior of the solutions in each of the four regions of Fig. 6b. In both Region I (where P > L) and Region II (where K < P < L), the solutions behave similar to the logistic model: the solution tends toward L. Specifically, in Region I the solution decays toward L whereas in Region II the solutions rise toward L, eventually in an asymptotic manner. Finally, in Region III ($0 < P < K$), the solutions decay toward zero whereas in Region IV ($P < 0$) the solutions decay toward negative infinity.

The one-dimensional first passage time problem is now considered, where the region under consideration is an interval $x_1 \leq x_0 \leq x_2$. We are interested in examining the time T it takes the process $x(t)$ starting at x_0 to first reach the boundary $x=x_1$ or $x=x_2$ as shown in Fig. 7. This so-called *First* passage time varies from realization to realization, so the mean (expected) first passage time $M(x_0)$ is of interest. Other notations for $M(x_0)$ include time $E[T_x]$ .Consider a practical first passage time example: a bomb has exploded a few miles outside of Tokyo, the capital of Japan. It is of interest to estimate the expected time it will take the dispersing molecules of poisonous gas to first reach the urban boundary of Tokyo under the molecular bombardment of air molecules. As another, consider a recent oil spill off the coast of China. It is of interest to estimate the first time that the oil will reach Hainan island or other ecologically sensitive areas. The first passage times for type-B and type-D barriers are shown in Figs. 7, 8 and 9.

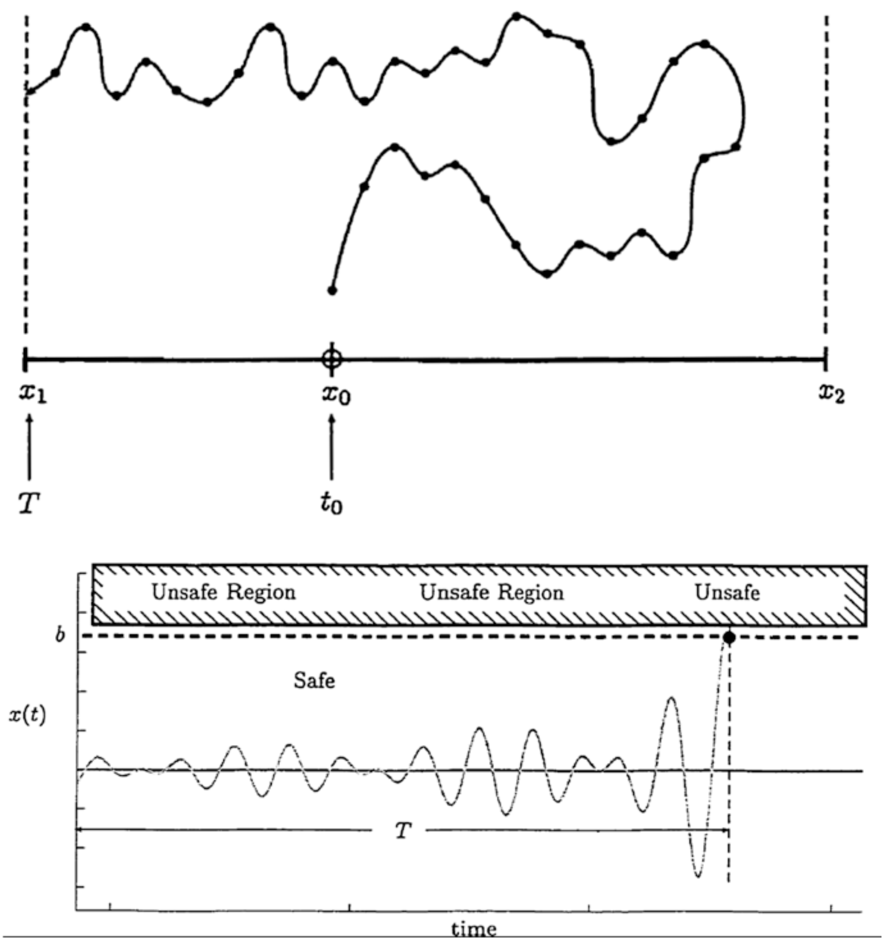


Fig. 7 First-passage time T for type-B barrier (x < b is safe)

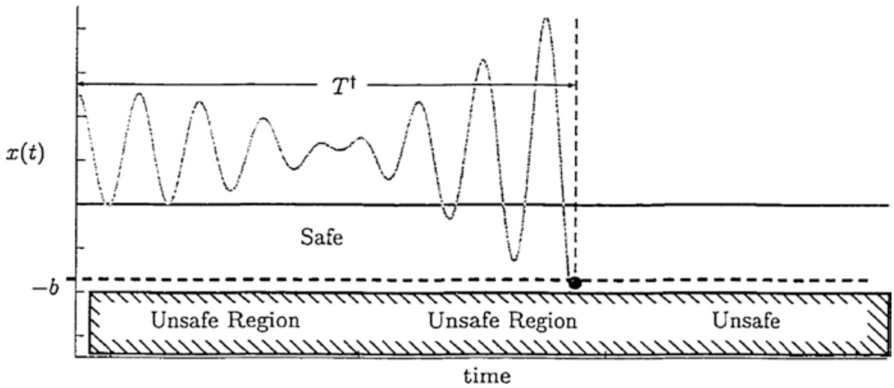


Fig. 8 First-passage time $\boldsymbol{T}^{\dagger}$ for type-B barrier (x > −b is safe)

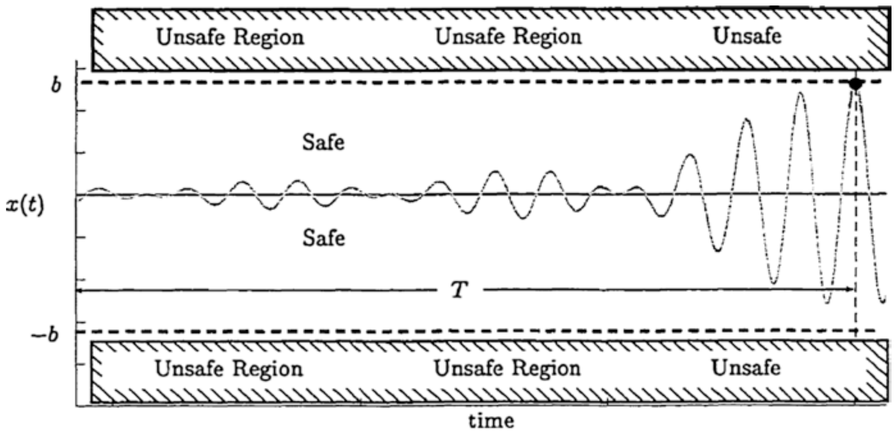


Fig. 9 First passage time T for type-D barriers

First passage time problems have been studied extensively using linear oscillators: consider the oscillator with response y(t) which is related to the wide-band random excitation F(t) by the differential equation:

$$F(t) = \ddot{y} + 2\zeta\omega_n \dot{y} + \omega_n^2 y \tag{10}$$

where the constants ω_n and ζ represent, respectively, the undamped natural frequency and the damping ratio of the vibratory system. The excitation *F(t)* is taken to be a wide-band random process with zero mean. A commonly studied first passage time problem for linear oscillators is to determine the probability distribution of the time T that it takes for y(t) starting from an initial amplitude level $\boldsymbol{T}$ to reach

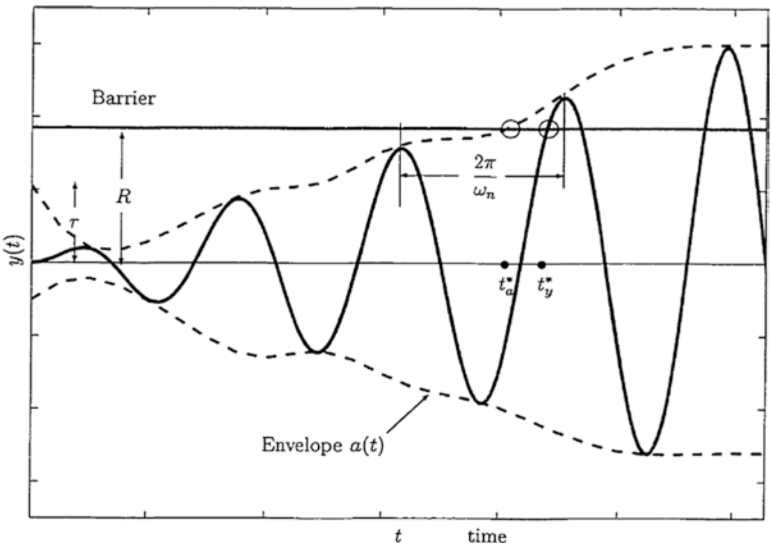


Fig. 10 Sketch of y(t) and a(t) crossing the barrier *R*

the barrier R (see Fig. 10). Political and economic forces often lead to extinction pressures. Stochastic simulations can help to calculate the "risk" of stock collapse at low abundance or high fishing pressure.

5 Climate Informatics and Extinction Risks in the Asia-Pacific: A Machine Learning Approach

The impacts of present and potential future climate change loom large in the Asia-Pacific region. Algorithmic models can help to improve understanding of the climate system. The global climate system is characterized by complex phenomena that are imperfectly observed and simulated. Machine learning approaches constitute a valuable approach given the growing supply of climate data from satellites and environmental sensors. Given the massive scale of the climate data an algorithmic model should be considered to analyze these challenges. There is a growing discipline known as *climate informatics*: it is proposed that the field of climate informatics will improve knowledge discovery and help to address key climate science questions in the Asia-Pacific. Recent achievements in machine learning for climate informatics highlight important new problems for machine learning and possible collaborations. The dugong ("sea cow") is one of the species highly sensitive to climate change. Like other populations of large marine mammals dugongs are under threat of extinction.

Found in waters off the northern half of Australia globally dugongs are classed as "vulnerable to extinction". To assist in conservation efforts, scientists in the Asia-

Pacific region track these endangered populations to identify their numbers, size and location. However, manually identifying dugong population from small planes is slow, expensive and sometimes hazardous. Accordingly, marine mammal researcher Dr. Hodgson from Murdoch University in Western Australia and Dr. Frederic Maire, a computer scientist at Queensland University of Technology are using drones fitted with cameras to help collect dugong data. In particular, Maire's automated detection system uses Google's machine-learning program TensorFlow (a free open source machine learning platform), to provide easier and more accurate population estimates. In particular, the deep learning neural network identifies dugongs by their size and color. The system currently has an 80% accuracy rate (i.e. it identifies 80% of dugongs they found manually in images.) The detector's accuracy improves as it learns from a wider set of negative examples (anything in the sea that might look like a dugong, like wave crests and shadows) and positive examples (Maire et al. 2015). The machine learning tools helps the scientists to identify dugongs from aerial photography of the ocean on tens of thousands of images. By being able to improve detection performance over time and to identify threatened populations on a large scale, conservationists will have a more accurate way to understand the impact of human activities and climate change on endangered species. It is expected this approach may scale well for dugongs and other sea mammals such as humpback whales and dolphins.

6 Policy Solutions to Promote Water and Climate Security

Both machine learning and statistical modeling approaches must be supplemented by robust policy solutions to enhance resilience to climate and water hazards. In February 2015, mayors and municipal leaders from the Asia Pacific region met to discuss a more coordinated response to and recovery from the effects of climate change. The resulting report called for community engagement, innovation, and local and global partnerships that would help prevent and manage risk. On March 18, 2015, the World Conference on Disaster Risk Reduction (DRR) was held in Sendai, Japan and attended by over 6500 participants including 2800 government representatives from 187 governments. The Public Forum had 143,000 visitors over the 5 days of the conference making it one of the largest UN gatherings ever held in Japan (UN 2015). Representatives from the 187 UN member states adopted the Sendai Framework for Disaster Risk Reduction 2015–2030**,** the first major agreement of the Post-2015 development agenda. The Sendai Framework is a far reaching document for disaster risk reduction with seven targets and four priorities for action.

The seven global targets to be achieved over the next 15 years include substantial reduction in (1) global disaster mortality, (2) numbers of affected people, (3) economic losses in relation to global GDP, and (4) disaster damage to critical infrastructure and disruption of basic services. It also aims to achieve (5) an increase in the number of countries with national and local DRR strategies by 2020, (6) enhanced international cooperation, and (7) increased access to multi-hazard early warning systems and disaster risk information and assessments. The four priority actions are (1) understanding disaster risk, (2) strengthening disaster risk governance to manage disaster risk, (3) investing in DRR for resilience, and (4) enhancing disaster preparedness for effective response and to 'Build Back Better' in recovery, rehabilitation, and reconstruction (UNOPS 2015)

The Sendai Framework calls for concrete indicators of progress towards set goals to be measured against the disaster losses in the decade after the adoption of the 2005 Hyogo Framework for Action. To reach its goals, the Framework calls for actions to not only protect populations and promote quick recovery, but also to prevent new risks such as those caused by ill planned urban growth in areas subject to flooding, landslides, and effects of climate change (Weru 2015). Integration with global regimes to mitigate and adapt to climate change and promote sustainable development is among the key objectives of the Sendai Framework, as is inclusively addressing risk through economic, governmental, structural, legal, social, cultural, educational, and health-related sectors, and UN organizations.

6.1 Australia Climate Change Adaptation

Like other countries in the Asia Pacific region, Australia faces the twin challenges of dealing with extreme weather-related disasters and adapting to the impacts of climate change. Recognizing the enormous environmental and socio-economic toll climate disasters have on the country, the Australian government called for action to develop an integrated approach across and between the different levels of government to address the impacts of climate change. A team of researchers from Griffith University and RMIT University was funded over 1 year (2012) by the National Climate Change Adaptation Research Facility (NCCARF) to develop the foundations for a nationally consistent approach to disaster risk management and climate change adaptation that would provide a set of appropriate reforms to governing institutions. The research team focused on a three-way comparative case study of the 2009 Victorian bushfires, the 2011 Perth Hills bushfires, and the 2011 Brisbane floods. The research involved an analysis of the reports generated by the official inquiries into these disasters, interviews with key stakeholders, and stakeholder workshops in Perth, Melbourne, and Brisbane. The final research report, entitled "The Right Tool for the Job: Achieving Climate Change Adaptation Outcomes through Improved Disaster Management Policies, Planning and Risk Management Strategies", offered data driven insights and recommendations that range from the conceptual to the practical. First, it was argued that a reconceptualization of terms such as 'community' and 'resilience' is necessary to take into account socio-economic diversity and allow for more tailored, context-specific risk analyses and responses. This is particularly important with regard to policymaking and planning processes and community engagement. Second, it was suggested that the high level of uncertainty inherent in disaster risk management and climate change adaptation requires a more interactive approach to policymaking and planning. Third, some specific institutional reforms were proposed that included:

1. Creating a new funding mechanism that would encourage communication and collaboration between and across different levels of government as well as promote partnerships with businesses and the community,
2. Improving community engagement through new resilience grants run by local councils,
3. Embedding climate change researchers within disaster risk management agencies to promote institutional learning and more integrated risk-context analyses, and
4. Creating an inter-agency network that encourages collaboration among organizations to support the proposed reforms.

The Australian research project is an example of how government can overcome political, social, and economic barriers in the interest of national preparedness for impending disasters. The findings of the research project offer guidelines for improving mitigation and adaptive responses as well as a starting point for better integration of disaster risk management and climate change adaptation. Efforts such as this one are of benefit to countries of the Asia Pacific region and the world.

6.2 *Role of International Organizations in Climate Change Preparedness in the Asia-Pacific Region*

In 2011, the United Nations Office for Disaster Risk Reduction – Regional Office for Asia and Pacific (UNISDR AP) issued a comprehensive report that provided a summary of how disaster risk reduction (DRR) and climate change adaptation (CCA) are undertaken and integrated in the Asia Pacific region. DRR is the concept and practice of reducing disaster risks through analysis and management of causal factors. It reduces exposure to hazards and lessens the vulnerability of people and assets. DRR also improves management of the land, the environment, and preparedness for adverse events. As experience with DRR and CCA grows, there is increasing recognition that both share a common focus: reducing the vulnerability of communities and contributing to sustainable development. The high level of climate related risks in the Asia Pacific region make DRR and CCA key policy goals. The 2011 UNISDR AP report provides best practices on how to improve the Asia Pacific regional planning and programming for DRR and CCA and highlights areas for cooperation among regional and sub-regional organizations. It proposes ways and means to support both national and regional stakeholders in DRR and CCA, such as governments, UN agencies, intergovernmental organizations, research and technical organizations, non-government organizations (NGOs), and especially the UN International Strategy for Disaster Reduction (UNISDR) Asia Partnership on Disaster Reduction members, in order to enhance regional planning, programming, and cooperation (UNISDR AP 2011). Follow up on the United Nations (UN) UNISDR AP report showed actual improvement in climate change preparedness in the Asia-Pacific region. This was noted in a subsequent UN report in 2014, "10 years after Indian Ocean Tsunami, Asia-Pacific Region Better Prepared" (UNNC 2014).

The Indian Ocean Tsunami, the world's worst recorded natural disaster, hit the Asia Pacific region in December 2004, killing more than 200,000 people, leaving 1.4 million survivors homeless, and destroying the entire food production systems on which whole populations depended (UNNC 2014). The devastation alarmed the world community. The UN Economic and Social Commission for Asia and the Pacific (ESCAP) and the German Federal Ministry for Economic Cooperation and Development (BMZ) have partnered with the German Ministry contributing 250,000 euros to the ESCAP Multi-Donor Trust Fund for Tsunami, Disaster, and Climate Preparedness, adding to an initial 500,000 euros contribution made in December 2013 (UNNC 2014). The UN report noted that some of the countries that were worst affected by the Indian Ocean Tsunami are now better prepared for disasters and better positioned to respond more effectively.

In the Pacific, as elsewhere, global climate change disasters have their greatest impact at the local level. Studies show that the accumulated impacts of small and medium disasters may be equivalent to, or exceed, those of large disasters. Increases in the frequency of these lower intensity hazards have a major impact on poverty. The countries studied are typical in terms of the current low level of integration of DRR and CCA. While there may be institutional arrangements that suggest some progress with integration at the national policy and institutional levels, the practical reality is that little is happening on the ground at the operational level. Although there is much work to be done, progress at the local level is being made. Tonga is clearly the lead example of local level integration of DRR and CCA. Tonga developed an integrated plan for Disaster Risk Management (DRM) and climate change (including the reduction of GHG emissions) and established a National Advisory Committee on Climate Change to take responsibility for DRM (World Bank 2013). Mainstreaming DRM in development planning can help to address some of the root causes of rising disaster impact. The annual damages from unabated economic development, population growth, and rapid urbanization that exacerbate climate change are expected to triple to $185 billion by 2100, even without factoring in climate change. DRM can help to reverse the current trend of rising disaster impact by acting swiftly and decisively to cut costs and losses due to problems of unchecked development. Lives and assets can be protected with wise policy and planning. However, many developing countries lack the tools, expertise, and instruments to factor the potential impacts of adverse natural events in their investment decisions (World Bank 2013)

The goals of the United Nations Climate Summit in September 2014, were to reduce GHG emissions, strengthen climate resilience, and mobilize political will for a meaningful legal global agreement in 2015, because the 'Hyogo Framework for Action 2005–2015: Building Resilience of Nations and Communities for Disasters' was scheduled to end in 2015. The United Nations General Assembly Resolution 66/199 requested UNISDR to facilitate the development of a Post-2015 Framework for Disaster Risk Reduction. A report which synthesizes consultations held at the regional, national, and community levels throughout the Asia Pacific region on the Post-2015 Framework for Disaster Risk Reduction was particularly targeted at countries and stakeholders from the region. The report describes the consultation approach that has been adopted in the Asia Pacific region and summarizes the key issues and proposals resulting from these consultations. The findings from the report add to the growing body of information needed to deal with climate change.

As more research is conducted, trend analyses of disaster occurrence and impact will address whether their determinants can be established. In Asia and elsewhere, factors that play a role in determining disaster trends are a mix of physical characteristics of the event itself and the socioeconomic context in which they occur. Earthquakes, for example, have short prediction times and therefore allow little time for disaster preparedness. In contrast, slower onset disasters such as droughts and floods are more predictable and generally result in fewer direct victims, but their real cost is in the medium- and long-term and is usually not assessed. Population density, urbanization, and demographic profiles are context-specific factors that are likely to contribute to the number of deaths and degree of damages. DRM policies and practices that are based on evidence can help to prepare for and reduce these and other risk factors.

To provide evidence-based information, reliable and time series data on impact is central. Global databases such as the International Disaster Data Base (EM-DAT), NatCat (Munich Re), or Dartmouth Flood Observatory provide valuable insights into trends and patterns. Substantial progress has been made in standardizing classification systems and definitions at global levels by Munich Re and EM-DAT, but international norms are still needed. Higher resolution impact monitoring data, sample surveys of risk factors, and other methods of gathering information will be required to provide data to develop more effective international DRM policy and practice. Because the cost to accomplish this will be great realistic financing options are essential.

In light of the significant costs of risk financing instruments, the challenge is to identify the appropriate layers of risk to cover, including a risk acceptance threshold, the lowest cost/risk solutions, and links to risk reduction. Strengthening the current innovative financing systems will be key. External involvement of governments, donors, and multi-lateral development banks is required to support communities and local institutions, build risk culture, reduce transaction costs in terms of bringing the products to the people (e.g., by providing support for mobile phone infrastructure), and pay or subsidize premiums. International organizations will continue to play an active role in advancing climate change preparedness in the Asia Pacific region and the world.

7 Conclusions

Algorithmic modeling and data modeling for water resources and environmental security challenges under conditions of climate change in the Asia-Pacific were compared and contrasted. The difference between the algorithmic modeling and data modeling cultures are discussed with reference to the schools in which they originate, the assumptions they work on, the type of data they deal with, and the techniques used. The 2011 Great East Japan Earthquake and other disasters confirm that government's capacity to manage disaster risks is critical in terms of prevention, preparation, response, recovery, and reconstruction. DRM governance must be streamlined as part of the development agenda for most developing countries. The structure and quality of governance of governing bodies at all levels, from central to local to community levels throughout Asia and the Pacific need to be improved to lead DRM initiatives. Moreover, DRM planning calls for widespread public involvement from all sectors of the community as well as from non-governmental organizations (NGOs). Existing evidence points to the crucial role of governance for an effective national DRM strategy and program. This relates to DRM policy and practices both at the national and local level. At the local level, primary issues include:

1. Linking local and national disaster preparedness: Disasters are usually local phenomena and the local governments along with the communities are the first responders. However, large-scale disasters require national or international efforts. Thus, for effective preparedness it is important to have specific links in terms of policy, plan, and action at the national and local level.
2. Coping with the changing nature of disasters: The nature of disasters, especially hydro-meteorological disasters, is changing and becoming more of a local phenomenon (especially in terms of rainfall patterns). This creates an increasing need for local capacities at the government, non-government, and community levels to cope with such disasters.
3. Addressing the needs of diverse communities: Communities vary from place to place and their perception and ways of responding to disasters also vary. Therefore, it is important to decentralize policies and customize them according to local needs and priorities.
4. Learning from past disasters: Accumulating evidence from past disasters suggest that informed and well-prepared local governments and local communities can minimize the impacts of disasters. This is the case even with mega disasters like the Great East Japan Earthquake in which over 18,000 people died mostly from drowning.
5. Increasing global awareness of local needs: Over the past two decades, there has been growing global and regional awareness about the effectiveness of focusing on local needs and priorities. Most of the global and regional frameworks call for local capacity building and policymaking, national developmental strategies, and cooperation among emerging economies of Asia to improve their disaster risk management practices.

References

Albert S, Alistair Grinham A, Gibbes B, Leon J, Church J (2016) Sea-level rise has claimed five whole islands in the Pacific: first scientific evidence. Environ Res Lett

Breiman L (2003) Statistical modeling: the two cultures. Qual Control Appl Stat 48(1):81–82

Breiman L, Friedman J, Olshen R, Stone C (1984) Classification and regression trees. Wadsworth, Belmont

Daniel C, Wood F (1971) Fitting equations to data. Wiley, New York

Dempster A (1998) Logicist statistic 1. Models and modeling. Stat Sci 13(3):248–276

DeRusha C, Keating L, Hook L, Mckenzie S, Scozzaro T, Alvarez V (2017) Asia's next challenge: securing the region's water future. Asia Society, Hong Kong

Long MC, Deutsch C, Ito T (2016) Finding forced trends in oceanic oxygen. Global Biogeochem Cycle 30:381–397. doi:10.1002/2015GB005310

Maire F, Alvarez LM, Hodgson A (2015, November). Automating marine mammal detection in aerial images captured during wildlife surveys: a deep learning approach. In Australasian Joint Conference on Artificial Intelligence (pp 379–385). Springer

Oak Ridge National Laboratory (ORNL) (2015) Integrated Assessment: Guam, undated. http://web.ornl.gov/sci/knowledgediscovery/QDR/guam/GuamAssessment

SAS (2016) https://www.sas.com/it_it/insights/analytics/machine-learning.htm

Schmidhuber J (2015) Deep learning in neural networks: an overview. Neural Netw 61:85–117

UN (2015) Sendai Framework for disaster risk reduction 2015–2030 (A/CONF.224/CRP.1), report from the UN international strategy for disaster reduction, March, 2015

UNDP (2015) Bringing about the 'Good Change' (together), United Nations Development Program Headquarters (UNDP), March, 2015 UNOPS, 2015, strengthening resilience, building back better, United Nations Office for Project Services (UNOPS), March, 2015

UNISDR AP (2011) United Nations Office for Disaster Risk Reduction – Regional Office for Asia and Pacific (UNISDR AP), "Climate change and disaster management at the crossroads: climate change adaptation and disaster risk reduction in Asia and the Pacific

United Nations News Centre (UNNC) (2014) 10 years after Indian Ocean tsunami, Asia-Pacific region better prepared: UN, Dec 22, 2014

UNOPS (2015) Strengthening resilience, building back better, United Nations Office for Project Services (UNOPS), March, 2015

Vapnik V (1995) The nature of statistical learning theory. Springer, New York

Vapnik V (1998) Statistical learning theory. Wiley, New York

Wahba G (1990) Spline models for observational data. SIAM, Philadelphia

Walsh J, Wuebbles D, Hayhoe K, Kossin J, Kunkel K, Stephens G, Thorne P, Vose R, Wehner M, Willis J, Anderson D, Doney S, Feely R, Hennon P, Kharin V, Knutson T, Landerer F, Lenton T, Kennedy J, Somerville R (2014) Chapter 2: Our changing climate. In: Melillo JM, Terese (T.C.) Richmond, Yohe GW (eds) Climate change impacts in the United States: the third national climate assessment. U.S. Global Change Research Program, Washington, DC, pp 19–67. doi:10.7930/ J0KW5CXT

Weru A (2015) United Nations Office for Disaster Risk Reduction, Hyogo Liaison Office (UNISDR Hyogo); United Nations Office for Disaster Risk Reduction, Regional Office for Asia and Pacific (UNISDR AP)

World Bank (2013) Disaster risk management in South Asia: regional overview "strong, safe, and resilient: a strategic policy guide for disaster risk management in East Asia and the Pacific", Report number 75847

Zhang H, Singer B (1999) Recursive partitioning in the health sciences. Springer, New York

Zielinski PA (1988) Stochastic dissolved oxygen model. J Environ Eng 114(1):74–90