6 CLEAN WATER AND SANITATION

MANAGING WATER RESOURCES

PROBABILISTIC TOOLS FOR WATER RESOURCES VULNERABILITY

Access to reliable and secure drinking water is essential for health, agriculture, sanitation, and hygiene. The World Health Organization promotes water safety through a series of water quality guidelines based on identifying and managing risks from catchment to consumer. For water suppliers and decision-makers, Bayes' Theorem offers a useful mathematical tool for quantifying risks and identifying appropriate options for the management of water supply and quality. Bayesian approaches can lead to probability estimates in the face of uncertainty that support better-informed decisions in government and public policy. These methods are now widely used in various fields, including medicine, law, and ecology..

KEY MESSAGES

- ☑ A mathematical method developed by an 18th-century Presbyterian minister has proven especially useful in helping government, public utilities, and the private sector approach complex problems involving water supply and water quality. The method, known as Bayes' Theorem, allows for estimating the likelihood of an event based on prior information or knowledge.
- ☑ Access to reliable and secure drinking water is essential for health and sustainable development. Bayes' Theorem holds considerable promise as a decision support tool for evaluating and prioritizing management and maintenance of water resources and water quality.
- ☑ Bayesian approaches are used to identify robust ways of quantifying risk and protecting human and environmental health.. Their use can lead to probability estimates that support better-informed decisions in government and public policy, in particular for water supply.

WATER RELIABILITY

Short-term water supply disruptions can be inconvenient, but a longer-term outage can threaten human health or life, especially if it occurs in the context of an epidemiological emergency, earthquake, or other natural hazards, or at a location where alternative sources of water are in short supply, expensive, or difficult to obtain. Water shortages pose an especially significant threat in developing countries. For communities that rely on groundwater accessed through boreholes and pumps, failure prediction is critical to avoid disruptions in water access or supply.

QUANTIFYING RISK

Suppose a water supply company is looking to minimize the likelihood of service outages lasting more than three hours. They would need to determine the most common causes or failures that result in these longer outages.

Typically, characteristics of water supply outages are detailed by the company in a failure book. Using this data, the company can calculate the following values: a) the number of specific outages attributed to each kind of failure over a given period; b) the number of water supply outages assigned to each duration class, for example, <3 hours and ≥3 hours; and c) for each type of failure, the number of resulting outages assigned to each duration class.

If there is a failure such as a leaking pipe, water supply managers can use this data to determine the likelihood it will cause an outage of at least three hours. But in the future, they might want to know the opposite: If they see an outage of three or more hours, what is the most probable failure that led to it?

Problems such as these lend themselves to a surprisingly powerful tool developed in the 18th-century by an English Presbyterian minister named Thomas Bayes.

A LOOK AT BAYES' THEOREM

Reverend Bayes, a trained statistician, developed a method for estimating the probability of a future event based on observed evidence. Estimates can be regularly updated using this method as new evidence is gathered. Bayes' Theorem is stated mathematically as:

$$P(F|C) = \frac{P(C|F) \times P(F)}{P(C)}$$

- $P(C|F)$: the **conditional** probability of the occurrence of an outage of class C given the occurrence of failure F
- $P(F)$: ***"a priori"*** probability of the occurrence of failure F
- $P(F|C)$: ***"a posteriori"*** probability of the occurrence of failure F, given the occurrence of an outage of class C
- $P(C)$: probability of the occurrence of an outage of class C

Prior or *a priori* probabilities are the initial probability values. *A posteriori* probability is the probability value after new evidence is incorporated, for example when 'C' occurs. The *a posteriori* probability estimate can be regularly updated using Bayesian methods as new evidence is gathered.

BAYES' THEOREM IN ACTION

The outage data in the first two rows of the following table is from a Polish water supply company that wanted to determine the probability that a future outage of three or more hours would be caused by a specific type of failure:

Outages/Failure	Leaking Pipes	Damage to Water Fittings	Pipe Corrosion	Cracking Pipes	Other	Total
# Due to Failure (F)	94	88	81	57	35	355
# Outages ≥3 hrs (C)	50	20	49	27	23	169
P(F|C)	**0.2958**	0.1183	0.2899	0.1598	0.1361	

If F is a leaking pipe and C is an outage of at least three hours, Bayes' Theorem gives the *a posteriori* probability — **P(F|C)** — that if there is a future outage of at least three hours that it will be due to a leaking pipe. Using present data, or *a priori* probabilities, **P(F|C) = 0.2958**. The use of more data has given a more refined estimate than simply **P(F)**.

To reduce the incidence of water supply interruptions, companies can use these probabilities to inform strategic decisions, such as where to target inspections, when to repair or replace materials, and how to prioritize limited resources. And with this method, the probabilities can be regularly updated as new failure data is recorded.

In Kenya, the tool eMaji Manager uses sensor data to manage waterpoint access for people and livestock and minimize failures at boreholes in semi-arid regions in the northern part of the country. This tool and Bayesian methods related to machine learning can lead to more robust failure prediction.

MONITORING WATER POLLUTANTS

Because of its unique approach to quantifying uncertainty and variability, Bayes' Theorem has proven to be an effective tool for understanding water quality conditions. For example, because water pollutant concentrations typically vary by month, Bayes' Theorem can be used to determine the likelihood that water samples in violation of given water-quality standards will occur in a particular month.

Bayesian approaches like this have been used to guide management decisions and design regulatory measures related to water quality standards. For example, the City of Austin, Texas, utilized Bayesian methods to assess stream water quality in relation to wastewater treatment plant effluent. The State of North Carolina used Bayesian methods to evaluate river eutrophication for proposed nitrogen limits in the Neuse River. And Bayesian methods have been used to study the relationship between water flow and salinity in Australian rivers.

CONCLUSIONS

Analysis using Bayes' Theorem has proven helpful across a variety of applications, including epidemiology, environmental policy, medical decision-making, and legal proceedings. Bayesian approaches are constructive in uncertain situations that call for probabilistic thinking, where estimates need to be updated regularly as new evidence or information is encountered. In this context, Bayesian approaches can lead to more accurate probability estimates and thus better-informed decisions in the face of uncertainty.

REFERENCES

Ecological Processes Standing Committee. 2014. *Bayesian Inference: Application To Environmental Management and Decision-Making*, Report to NJ Department of Environmental Protection.

Gronewold, A. and Vallero, D. 2010. Applications of Bayes' theorem for predicting environmental damage. *Access Science*, McGraw Hill Education.

Otieno, F., Nyota, T., Waweru, I., Cintas, C., Maina, S., Ogallo, W., Walcott-Bryant, A. 2020. Machine learning approaches to safeguarding continuous water supply in the arid and semi-arid lands of northern Kenya. ICLR Workshop Tackling Climate Change with Machine Learning.

Porras, A. 2015. Bayesian analysis of stream water quality data in relation to wastewater treatment plant effluent. City of Austin, Watershed Protection Department, Environmental Resource Management Division, No. SR-15-03.

Shihab, K. and Al-Chalabi, N. 2014. Bayesian methods for assessing water quality/ *Proceedings of Fourth International Conference on Computer Science & Information Technology*, pp. 397-407,

Szpak, D. 2020. Method for determining the probability of a lack of water supply to consumers. *Energies*, Vol. 13, p. 5361.

Worrall, F., Kerns, B., Howden, N., Burt, T., and Jarvie, H. 2020. The probability of breaching water quality standards – A probabilistic model of river water nitrate concentrations. *Journal of Hydrology*, Vol. 583, p. 124562.

AUTHORS

Helen Roberts
Montclair State University, United States

Fred Roberts
Rutgers University, United States