A Modular Approach to Risk Assessment with Structural Models

John F. Carriger

U.S. Environmental Protection Agency, Office of Research and Development, Center for Environmental Solutions and Emergency Response, Cincinnati, OH USA

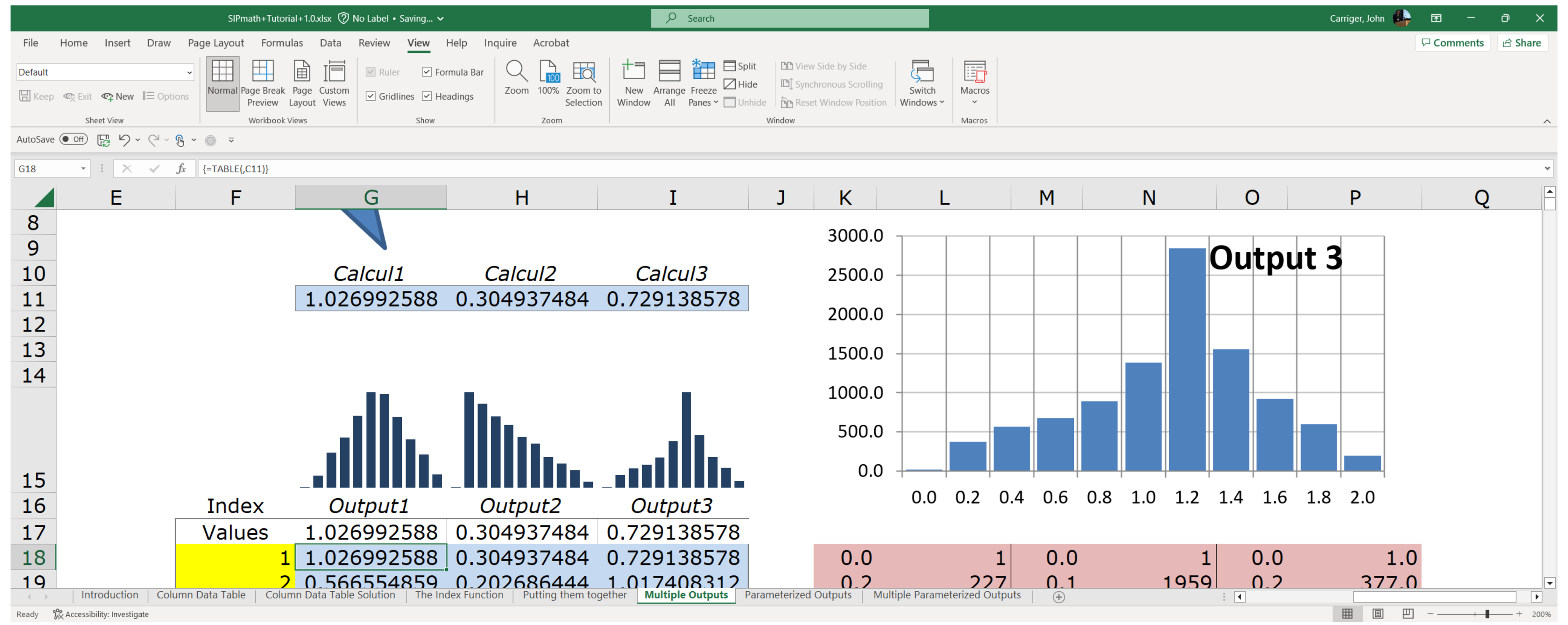
John Carriger | carriger.john@epa.gov | 513-569-7286

Reusable models through probability management

"The big failure of risk management is the lack of consolidating individual risk models and the lack of being able to audit them." -Sam Savage

Probability management originated in the 1980s from Sam Savage to maintain a mathematical and curated process for managing uncertainties across an organization (Savage and Markowitz 2000). It has been expanded over time by analysts and key figures like Douglas Hubbard (2009).

- Prevents reinventing models or classes of models
- Fosters effective use of uncertainty in an organization
- Frees more effort for new problems
- Builds a culture of improvement
- Creates clear responsibilities
- Incorporates QA/QC in the modeling process



Screenshot example of SIP distribution outputs calculated algebraically in three different ways from two SIPS

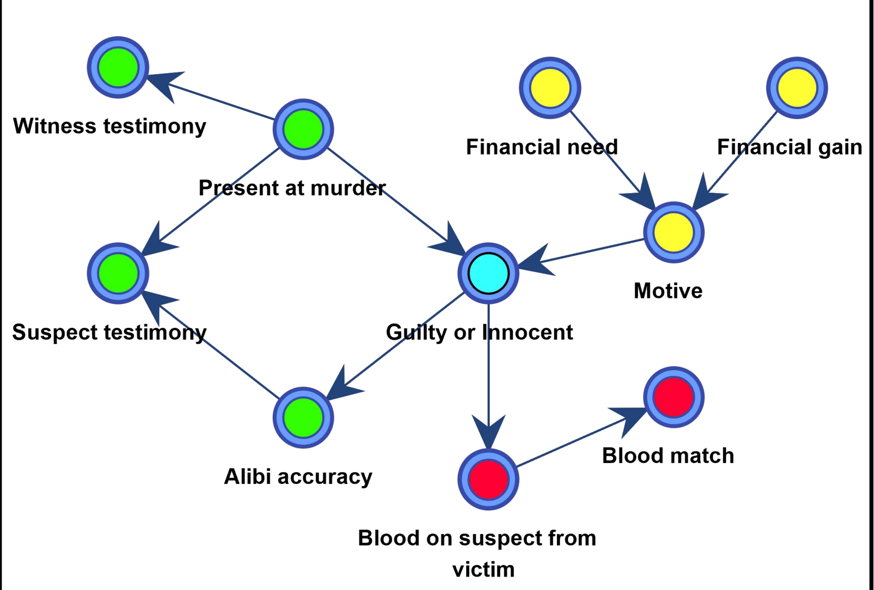
Probability management preserves libraries of probability distributions for reuse and standardized modeling (Savage and Markowitz 2000; Hubbard 2009). These libraries are auditable and certified for specific applications. The core components are SIPS and SLURPS. SIPS provide vector arrays of data to create distributions while SLURPS represent multiple correlated SIPS usually contained in a data matrix. The library is curated by a Chief Probability Officer (CPO). Probability management has been used across industry, most famously by Shell for oil exploration. Although probability management has much utility and has even expanded into Bayesian analysis of variables, causality is not explicitly considered in the framework.

Idioms for representing model fragments

An analogous structural development came in the Bayesian network (BN) field with what came to be known as idioms for representing relationships between variables. Idioms are generic classes of BN sub-models that can be reused and adapted across different problem structures. Several generic idiom types were initially identified by Neil et al. (2000) in the paper that first introduced the concept. These were cause-consequence, measurement, definitional/synthesis, induction, and reconciliation idioms. More specific idioms were identified and utilized in diverse applications including criminology, medical diagnosis, and product safety. Specific and general idioms allow model domains to be broken up into key components like SIPS and SLURPS in probability management.

Idioms (Neil et al. 2000):

- Model reasoning
- Assist in model construction
- Characterize underlying structure
- Encourage reuse and redevelopment

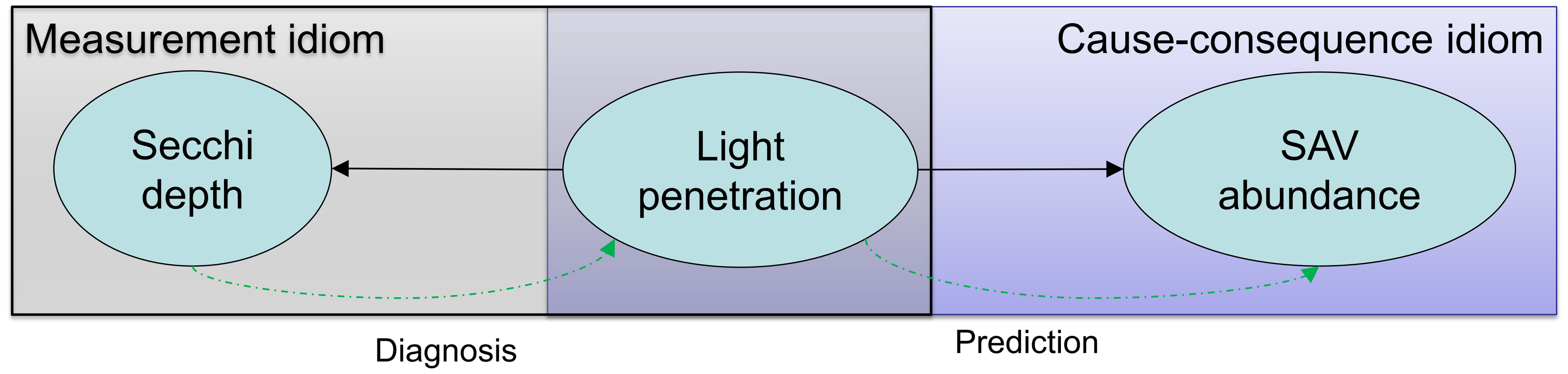


Legal idiom example- green nodes are opportunity, yellow nodes are motive, and red nodes are evidence idiom examples (adapted from Fenton et al. 2013)

Five generic idioms and example uses for coral reef assessments (from Carriger et al. 2019)

Idiom	Use in a Bayesian network	Example use in a coral reef ecosystem service assessment
Definitional/ synthesis	Condense multiple variables into one variable to simplify a model or to define a new variable	Coral condition defined by species composition, coral abundance, physical status, and biological condition
Measurement	Represent the uncertainty of the true state of a variable from measurement inaccuracies	Represent the uncertainty of the true fishing intensity of a fleet through a proxy measure such as number of ships or active licenses
Induction	Forecast the future state of a variable from past observations	Use historical observations of visits at a site to predict future recreational visits
Reconciliation	Reconcile predictions from different model output	Use both the distance of a reef to seagrass and a definition of habitat quality from rugosity and coral cover to examine habitat quality
Cause-consequence	Examine the uncertainty in a causal relationship between variables	Assess impacts on coral condition from ocean acidification which in turn causes impacts on the fishing stock from a reef

The flexibility of BNs for defining and relating variables combined with the assistance of idioms provided powerful tools for piecing together uncertain variables into complex problems

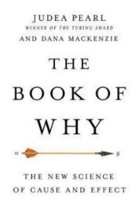


The idioms facilitate appropriate model structures for inferences and uncertainty propagation across a wide variety of disciplines. This can allow powerful capabilities for examining scenarios that incorporate causal predictions and measurement observation uncertainties for prediction and diagnoses.

Structural causal models (SCMs)

A structural causal model (SCM) relies on the cause-consequence idiom to depict the causal relationships among variables (nodes) in a problem through directed relationships (arcs). The SCM can be very useful without quantitative information and multiple SCMs can be developed, iterated, and pieced together in early phases. Building a SCM facilitates identification of the data generating problem structure. Moreover, the probability of necessary, sufficient and necessary and sufficient causation can be calculated. Different types of causal inferences can be made (Pearl and MacKenzie 2018):

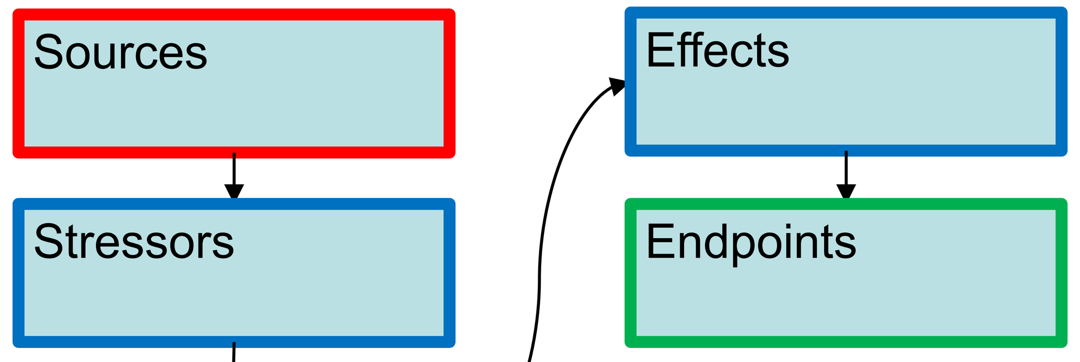
- Observational inference- Example: What is the probability of Y if we observe X and Z?
- Interventional inference- Example: What is the probability of Y and Z if we intervene to change X?
- Counterfactual inference- Example: What is the probability of Y if X and Z did not actually happen?



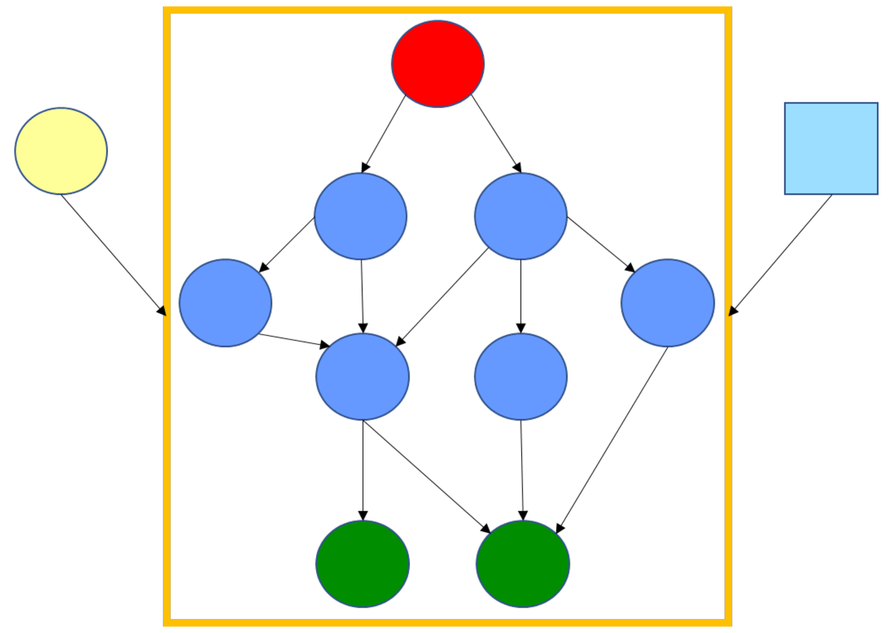
These types of inferences follow the ladder of causality- Seeing, doing, and imagining, respectively. Classical statistics pertains to the lowest rung of the ladder (seeing/observational inference) while decision analysis examines interventions. Counterfactuals are powerful analysis tools that allow examining separate worlds from what already happened and incorporate the first two rungs of the ladder of causation (seeing and doing) for calculation.

Environmental risk assessment (ERA) conceptual models

Conceptual models are a key component of environmental risk assessment (ERA) and heavily influence the data needs for a quantitative model. Conceptual models all contain a source term for the stressor, receptors such as ecological endpoints, and exposure pathways between the source and the receptor. Two aligned framework examples are shown here: the Relative Risk Model (Landis 2021) and a conceptual Bayesian network (Carriger and Parker 2021).



Relative Risk Model causal conceptual risk model. Each square box would contain multiple interacting variables such as multiple stressors



Conceptual Bayesian network components and structural flow. Red node = source; blue node = intermediate node; green node = endpoint; yellow node = modifier; light blue square node = intervention; orange box = exposure pathway system.

A modular process to ERA with causal idioms

Causal idioms can be used to break down the components of risk problems into manageable sub-models. These sub-model structures can be adapted and reused for similar types of risk problems such as ones found in other locations, time periods and classes of stressors with similar information requirements. The structures may be reused alone or with the functions and probabilities for the nodes. A useful application process for a conceptual fate and transport model is described in McMahon et al. (2001) and partially adapted here for a risk assessment.

Curate and extract idioms from internal and external risk applications

Build sub-models and overall problem structure for a risk assessment

1. Determine overall conceptual model structure with available information
2. Choose idioms reflecting existing components and connections
3. Refine with expert opinion, systematic review and input from stakeholders
4. Improve sub-models with new information from the phases of the ERA

Build functional relationships for a risk assessment

1. Build functional relationships and parameter values
2. Evaluate model and run sensitivity analysis to identify key sub-models
3. Refine model for accuracy and to meet model objectives
4. Apply model for predictions and inferences

Extract Idioms from case study, Update existing idioms that were applied and refined in analysis

References

Carriger, J.F. and Parker, R.A., 2021. Conceptual Bayesian networks for contaminated site ecological risk assessment and remediation support. *Journal of environmental management*, *278*, p.111478.
Carriger, J.F., Yee, S.H. and Fisher, W.S., 2019. An introduction to Bayesian networks as assessment and decision support tools for managing coral reef ecosystem services. *Ocean & coastal management*, *177*, pp.188-199.
Fenton, N., Neil, M. and Lagnado, D.A., 2013. A general structure for legal arguments about evidence using Bayesian networks. *Cognitive science*, *37*(1), pp.61-102.
Hubbard, D.W., 2020. *The failure of risk management: Why it's broken and how to fix it*. John Wiley & Sons.
Landis, W.G., 2021. The origin, development, application, lessons learned, and future regarding the Bayesian network relative risk model for ecological risk assessment. *Integrated Environmental Assessment and Management*, *17*(1), pp.79-94.
McMahon, A., Heathcote, J., Carey, M. and Erskine, A., 2001. *Guide to good practice for the development of conceptual models and the selection and application of mathematical models of contaminant transport processes in the subsurface*. National Groundwater & Contaminated Land Centre. Environment Agency. UK. Report NC/99/38/2.
Neil, M., Fenton, N. and Nielson, L., 2000. Building large-scale Bayesian networks. *The Knowledge Engineering Review*, *15*(3), pp.257-284.
Pearl, J. and Mackenzie, D., 2018. *The book of why: the new science of cause and effect*. Basic books.
Savage, S.L. and Markowitz, H.M., 2009. *The flaw of averages: Why we underestimate risk in the face of uncertainty*. John Wiley & Sons.