EPA
United States
Environmental Protection
Agency

EPA Training Webinar:
CADDIS – Causal Assessment & Stressor Identification



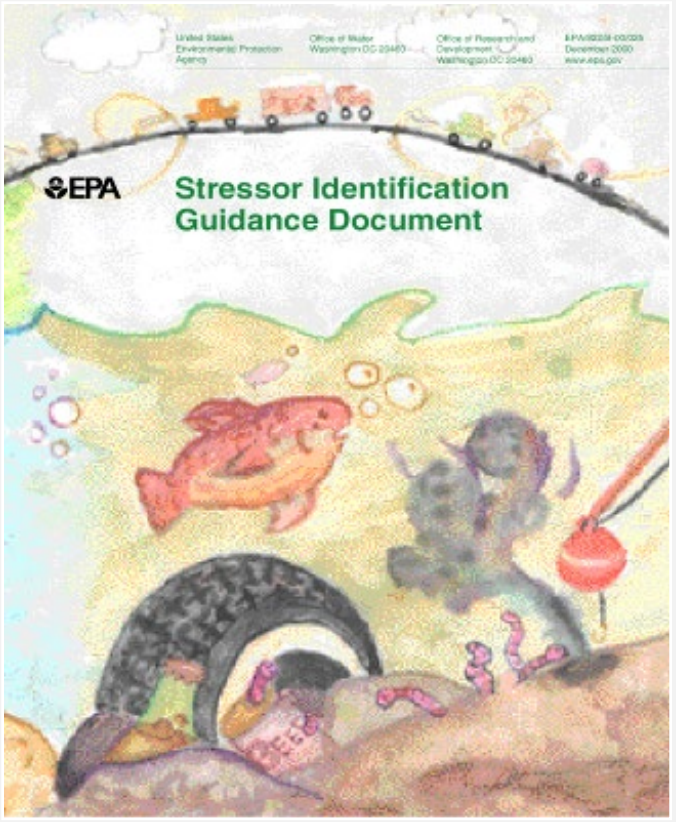
Kate Schofield, PhD
Center for Public Health and Environmental Assessment
US EPA Office of Research and Development

October 8, 2020

Presentation outline

epa.gov/caddis

- Introduction to CADDIS – EPA's Causal Analysis/ Diagnosis Decision Information System
  - Intro to causal assessment (including how CADDIS, causal assessment, and stressor identification are related)
  - CADDIS tour
  - How CADDIS can be used, in causal assessment and beyond



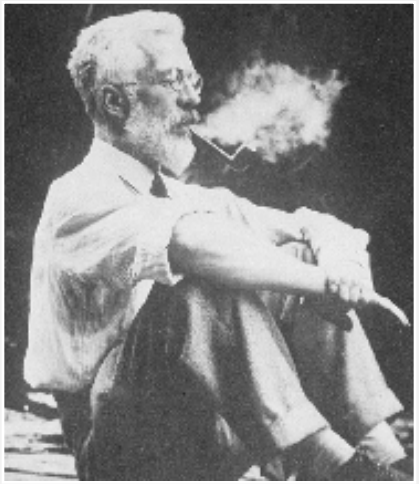
EPA United States Environmental Protection Agency

Causal assessment, Stressor Identification & CADDIS

- **Causal assessment**
  - Process to determine likely cause of an observed effect
- **Stressor Identification (SI)**
  - Method for determining most likely cause of observed biological impairments in aquatic systems
- **CADDIS**
  - Causal Analysis/Diagnosis Decision Information System
  - Website that provides information, methodology and tools to help users implement SI and conduct causal assessments of biological impairment

EPA United States Environmental Protection Agency

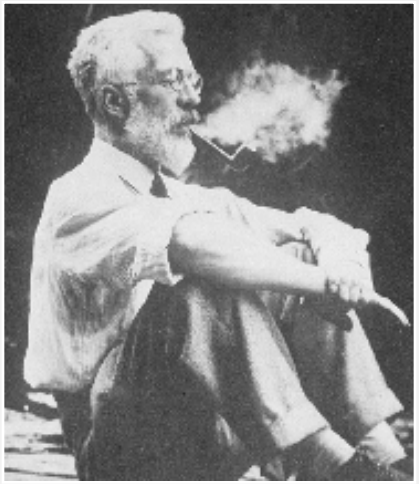
Three tiers of causal assessment



- **General** – Can C cause E?
  - Can smoking cause lung cancer?
  - Can Chemical Z cause fish lesions?
- **Contextual** – Under what conditions can C cause E?
  - Does smoking cause lung cancer when certain genetic factors are also present?
  - Does Chemical Z cause fish lesions only when it exceeds a particular concentration?
- **Specific** – Did C cause E in this case?
  - Did smoking cause lung cancer in Ronald Fisher?
  - Did Chemical Z cause fish lesions in my stream?

EPA United States Environmental Protection Agency

Three tiers of causal assessment



- **General** – Can C cause E?
  - Can smoking cause lung cancer?
  - Can Chemical Z cause fish lesions?
- **Contextual** – Under what conditions can C cause E?
  - Does smoking cause lung cancer when certain genetic factors are also present?
  - Does Chemical Z cause fish lesions only when it exceeds a particular concentration?
- **Specific** – Did C cause E in this case?
  - Did smoking cause lung cancer in Ronald Fisher?
  - Did Chemical Z cause fish lesions in my stream?

Why is specific causation important?

- Biological assessments are commonly used to identify if streams are impaired
- In many cases, causes of impairment are unknown
- To fix the problem, you have to know what to fix

Causes of Impairment for 303(d) Listed Waters

Rank	Impairment Group
1	Pathogens
2	Sediment
3	Nutrients
11	Cause unknown
12	Cause unknown: impaired biota
33	Cause unknown: fish kills

~ 90,000 river/stream miles

Why use a formal method?

Because we can make mistakes about causality, by…

- Forming initial impressions quickly, based on readily available information. This can result in:

Overweighting chance events	Every time I wash my car it rains.
Having biases	All pollution is caused by industry.
Being "educationally" predisposed	Hydrologists think hydrology.
Relying on intuition and past experience	I have a hunch that it's nitrogen. Last time I saw this, it was nitrogen.

EPA United States Environmental Protection Agency

Why use a formal method?

Because we can make mistakes about causality, by…

- Gathering information that supports our initial impression

HYPOTHESIS TENACITY

- Confidently reaching conclusions based on incomplete information

WYSIATI

"what you see is all there is"

"Science is a way of trying not to fool yourself. The first principle is that you must not fool yourself – and you are the easiest person to fool." [Feynman 1964]

The CADDIS causal assessment approach

THE GOOD…

- Provides formal method that allows defensible & transparent evaluation
- Identifies causal relationships that may not be immediately apparent
- Minimizes biases and other lapses of logic
- Helps identify all available evidence
- Increases confidence that remedial or restoration actions can improve biological condition

The CADDIS causal assessment approach

…THE BAD…

- Conducting causal assessments is not necessarily easy or straightforward
- Mechanisms driving biological impacts can be complex
- The method relies on data – quantity and quality matter
- Ultimately, a “smoking fish” may not be found, or multiple stressors may remain as likely causes

The CADDIS causal assessment approach

…AND BACK TO THE GOOD

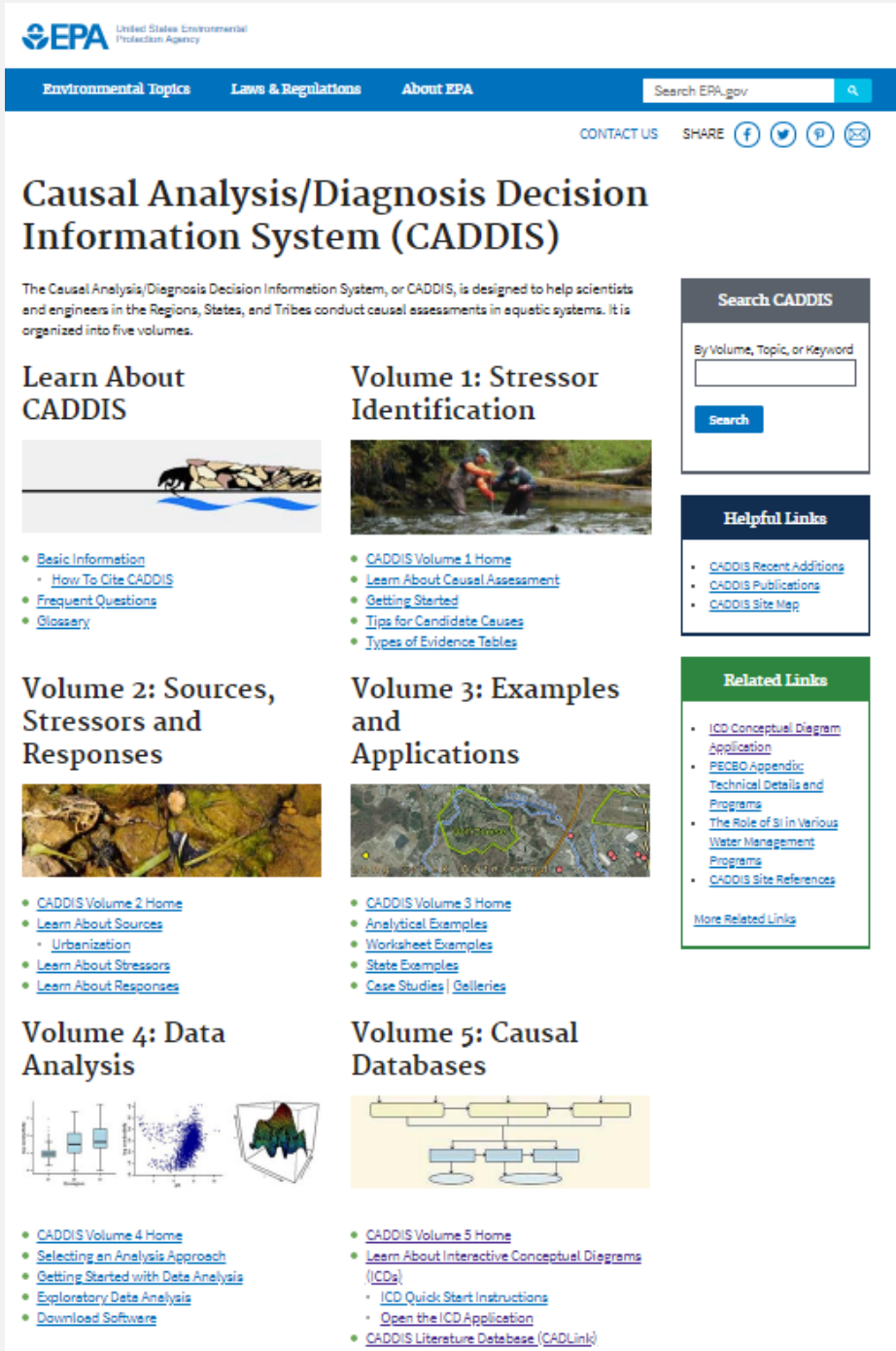
- Even when one likely cause is not identified, a causal assessment can narrow the universe of possible causes and point to promising data and analyses

1. ~~Low dissolved oxygen~~
2. Gill damage
3. Nitrate exposure
4. Infections
5. ~~High pH~~
6. ~~pH fluctuations~~
7. ~~Ammonia toxicity~~
8. Other, unspecified toxic substances
9. ~~Inadequate food resources~~

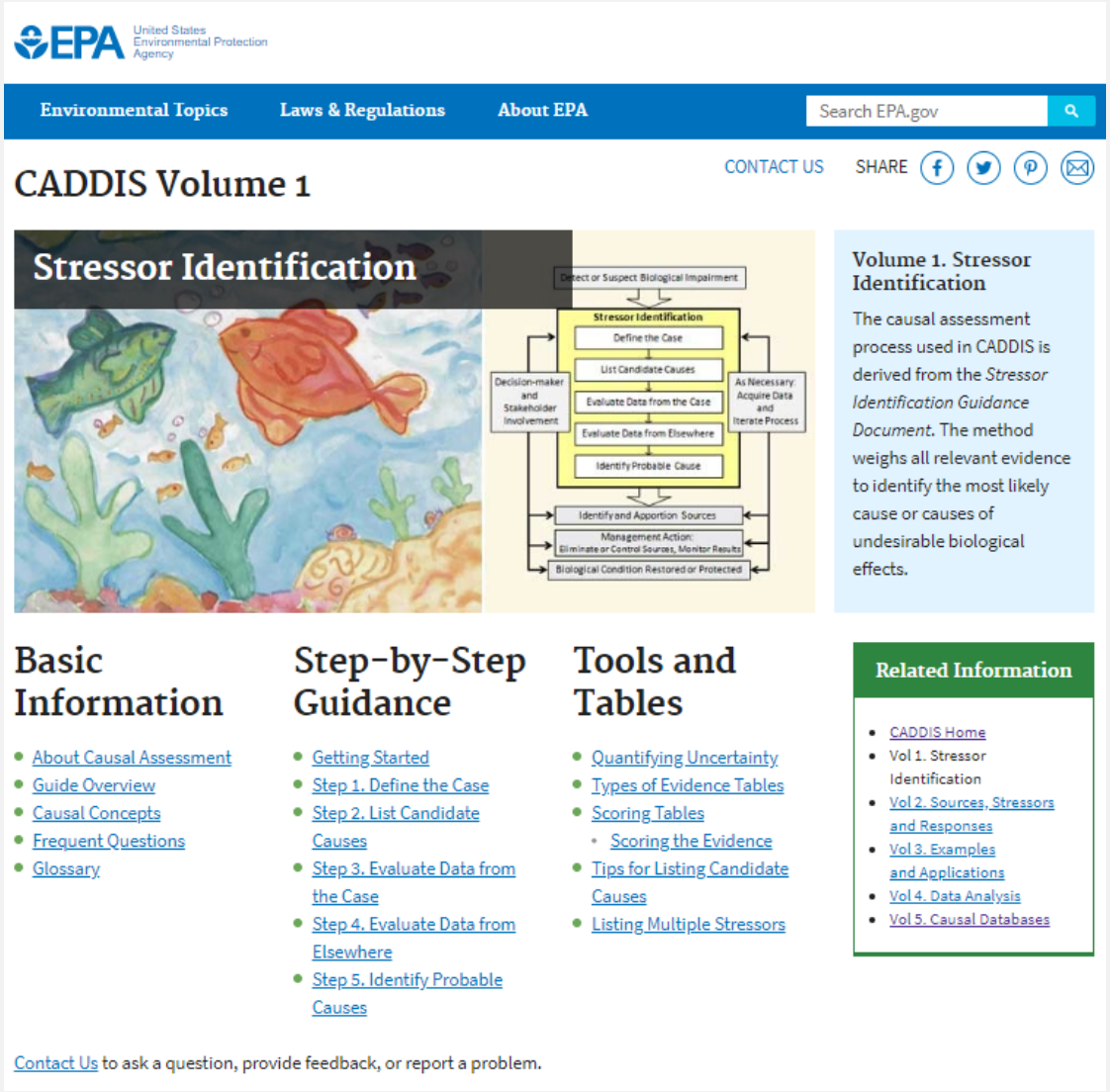
EPA United States Environmental Protection Agency

CADDIS

- **Volume 1:** Stressor Identification
- **Volume 2:** Sources, Stressors & Responses
- **Volume 3:** Examples & Applications
- **Volume 4:** Data Analysis
- **Volume 5:** Causal Databases

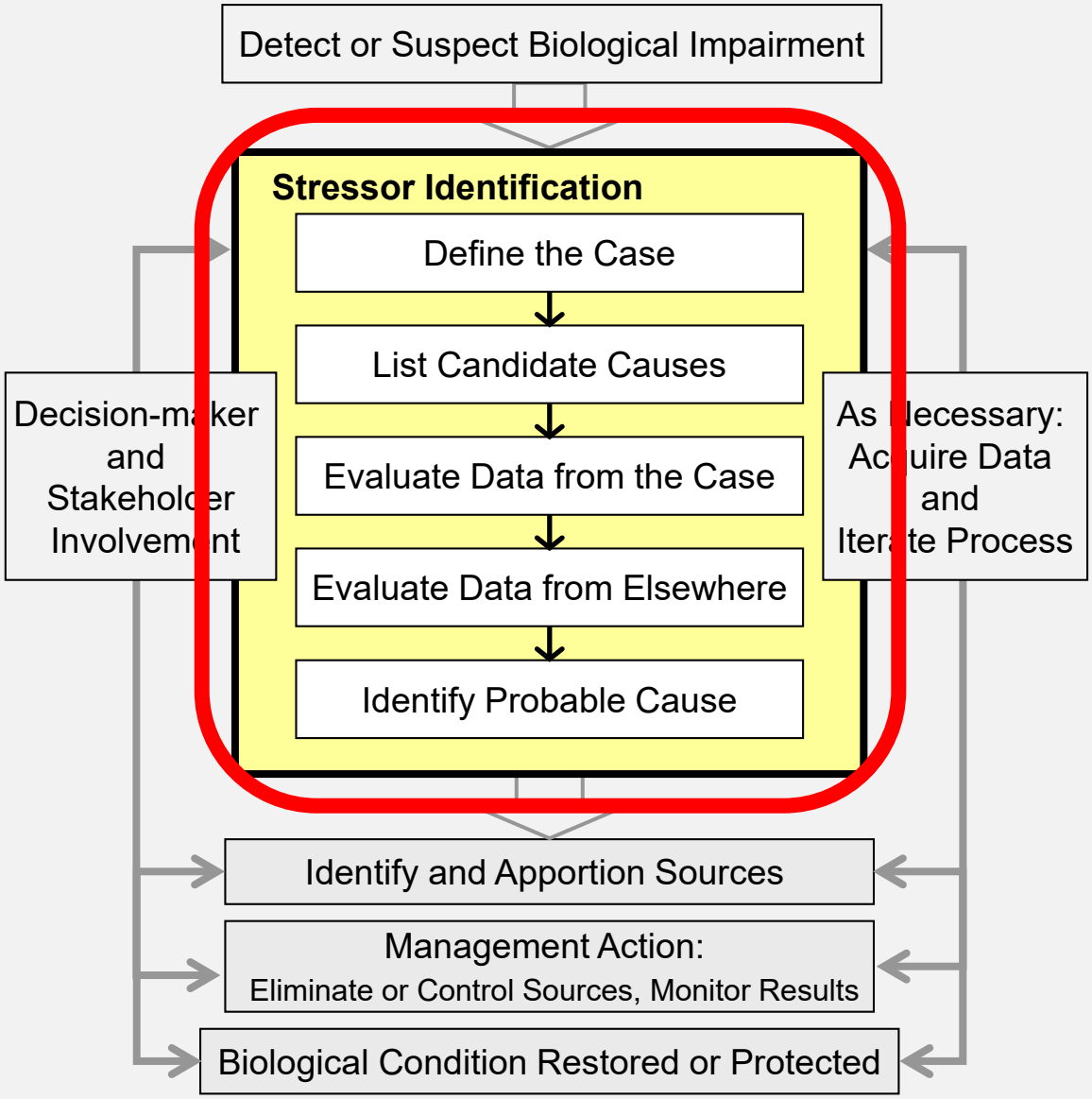


Vol 1: Stressor Identification

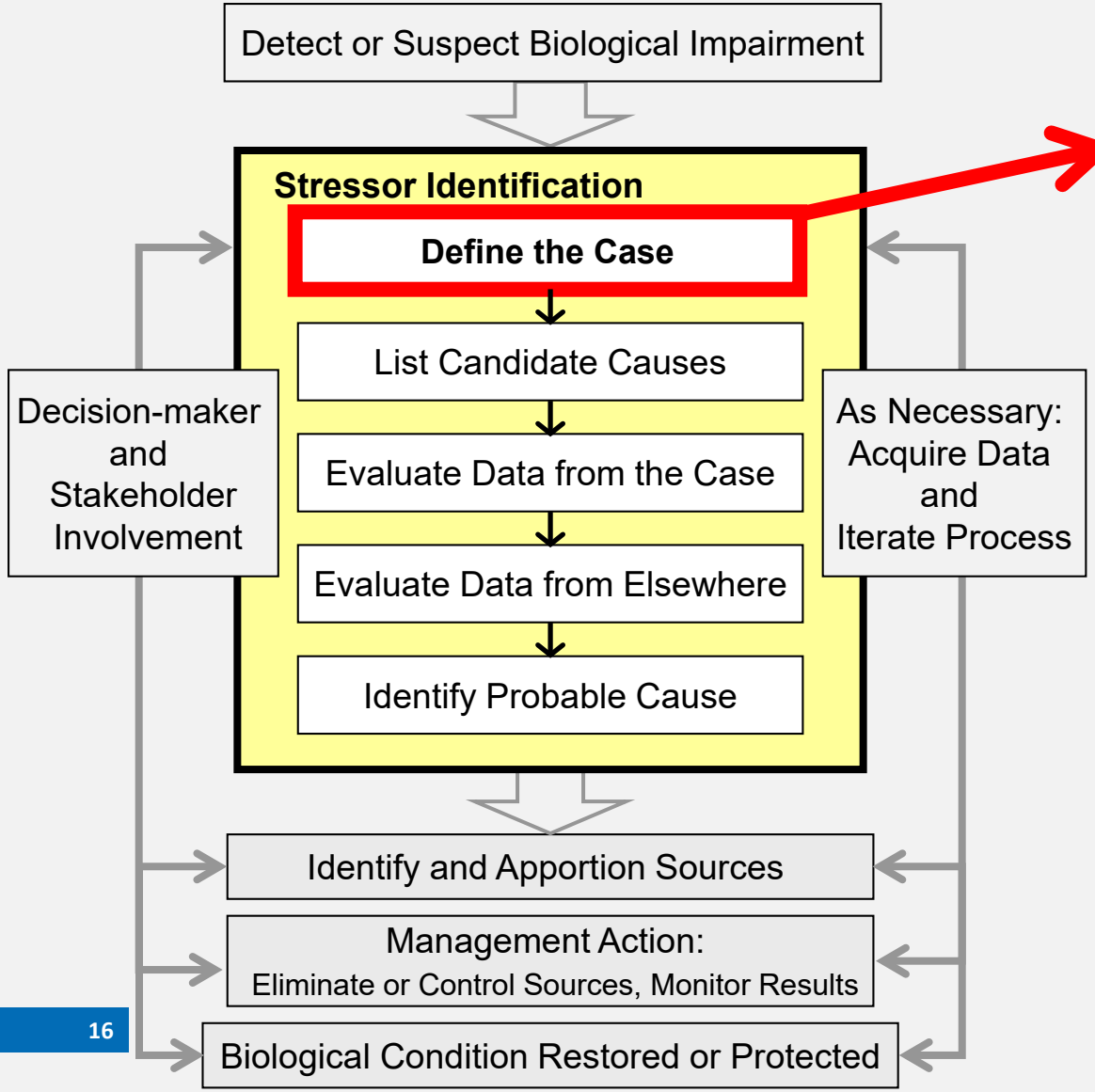


- Step-by-Step Guide
- Causal Assessment Background

5 Steps of Stressor Identification

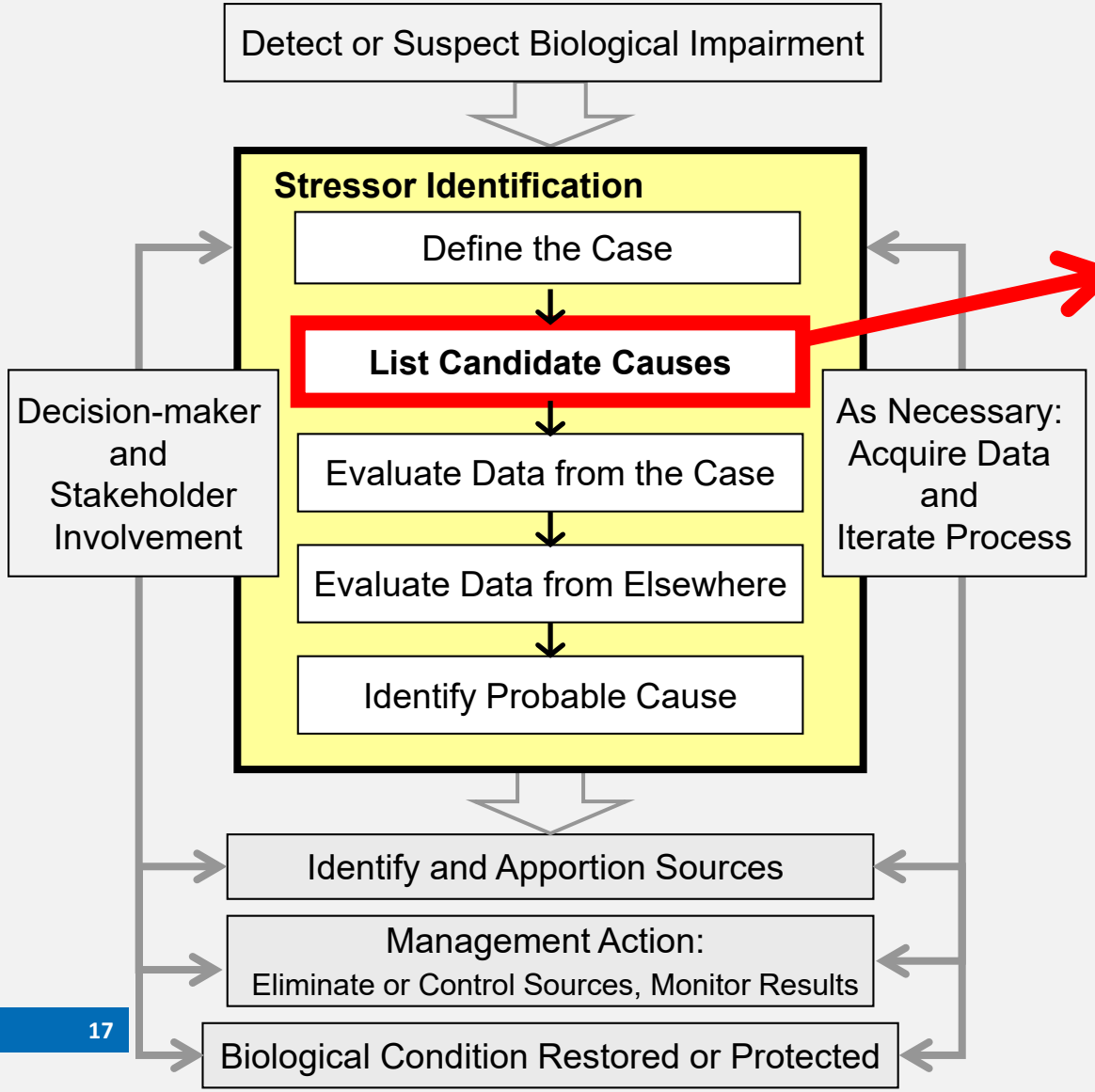


Step 1 – Define the case



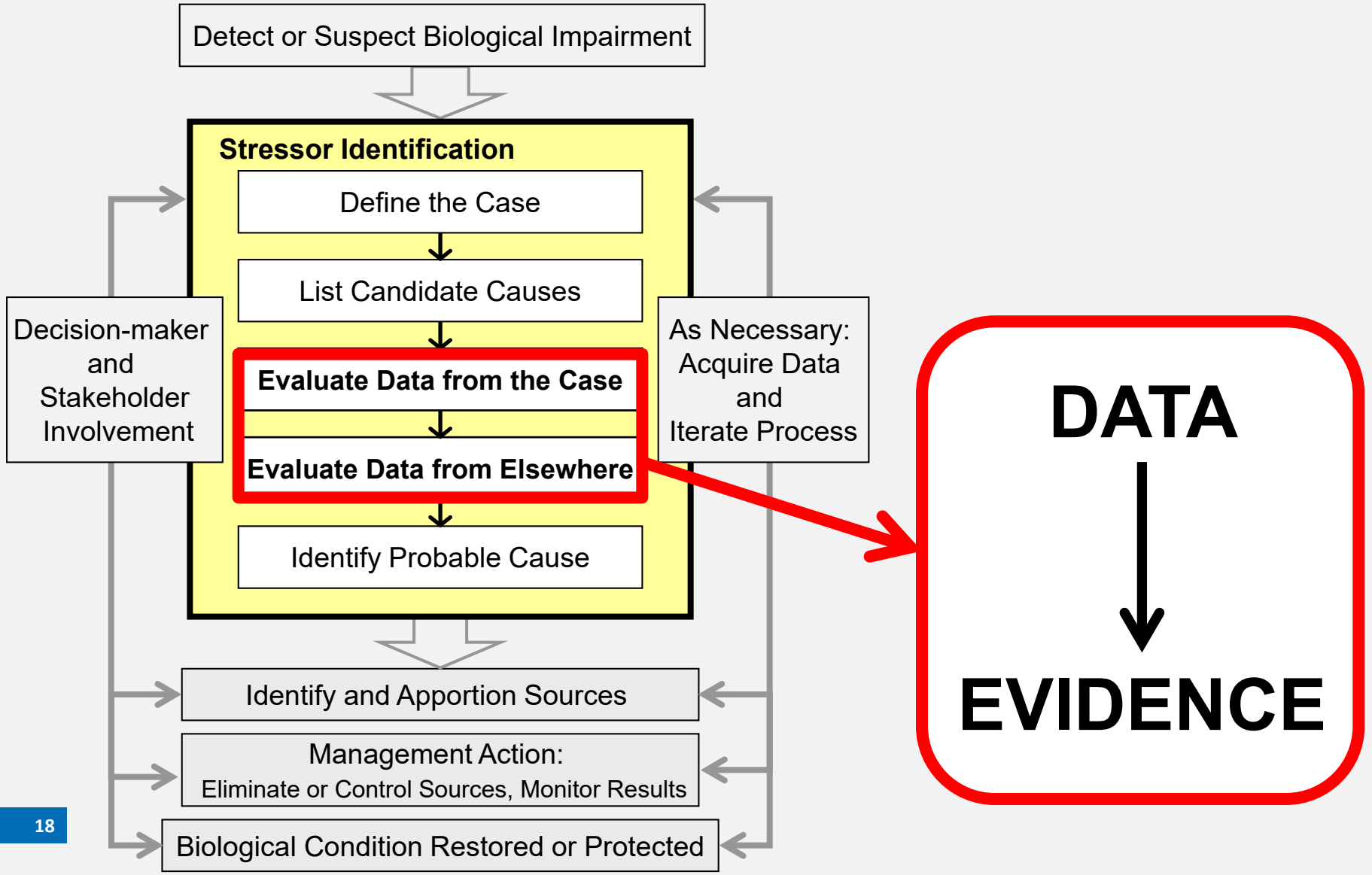
- What specific biological effects were observed?
- Where and when did they occur?
- Where are the effects absent or different (i.e., where are comparison sites located)?

Step 2 – List candidate causes



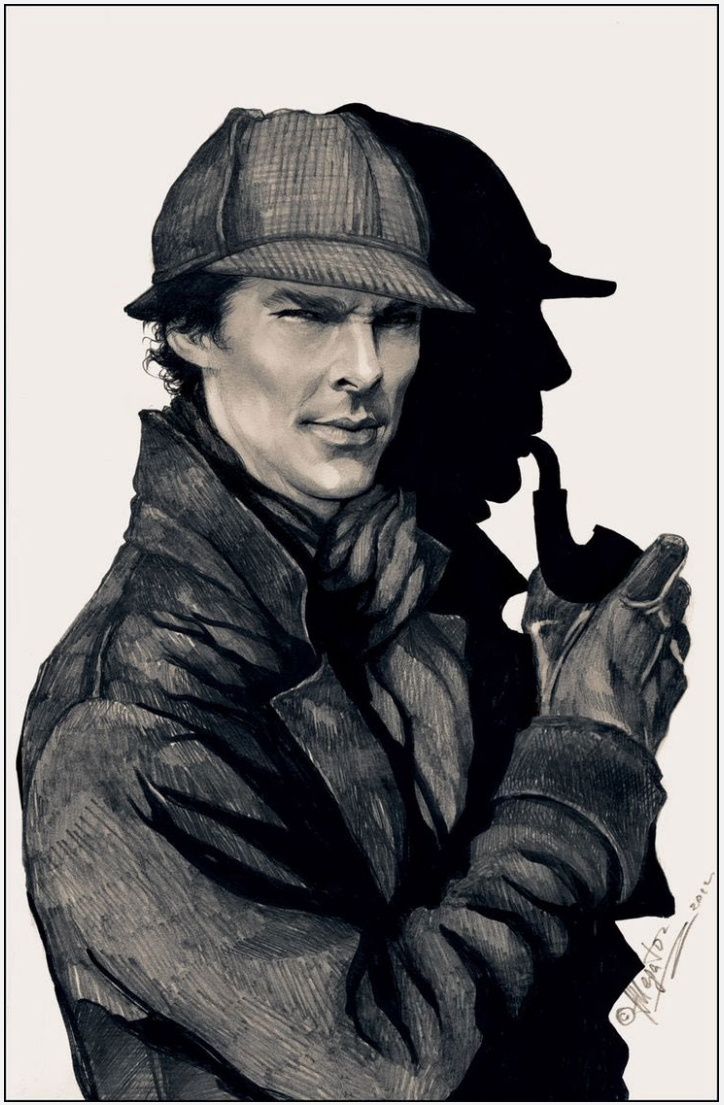
- Generate an initial list
- Gather information on potential sources, stressors, and exposures
- Develop conceptual model
- Develop the “final” list

Steps 3 & 4 – Evaluating the data



Let's talk about evidence…

- What is evidence?
  - Available information that indicates whether belief or proposition is valid
  - If Cause X produced Effect Y, then we would expect to observe Result Z
  - Information used to determine whether we actually observe Result Z is a piece of evidence
  - Individual pieces of evidence are combined into the overall body of evidence



"From the case" vs. "from elsewhere"

- **"From the case"** = data collected from affected location and nearby comparison sites
  - Most relevant evidence
  - Best chance of isolating causal processes, minimizing confounding factors

- **"From elsewhere"** = data collected from other field locations, the laboratory, or model simulations
  - Compare data from the case to data from elsewhere to derive pieces of evidence

Types of evidence in CADDIS

Data from the case

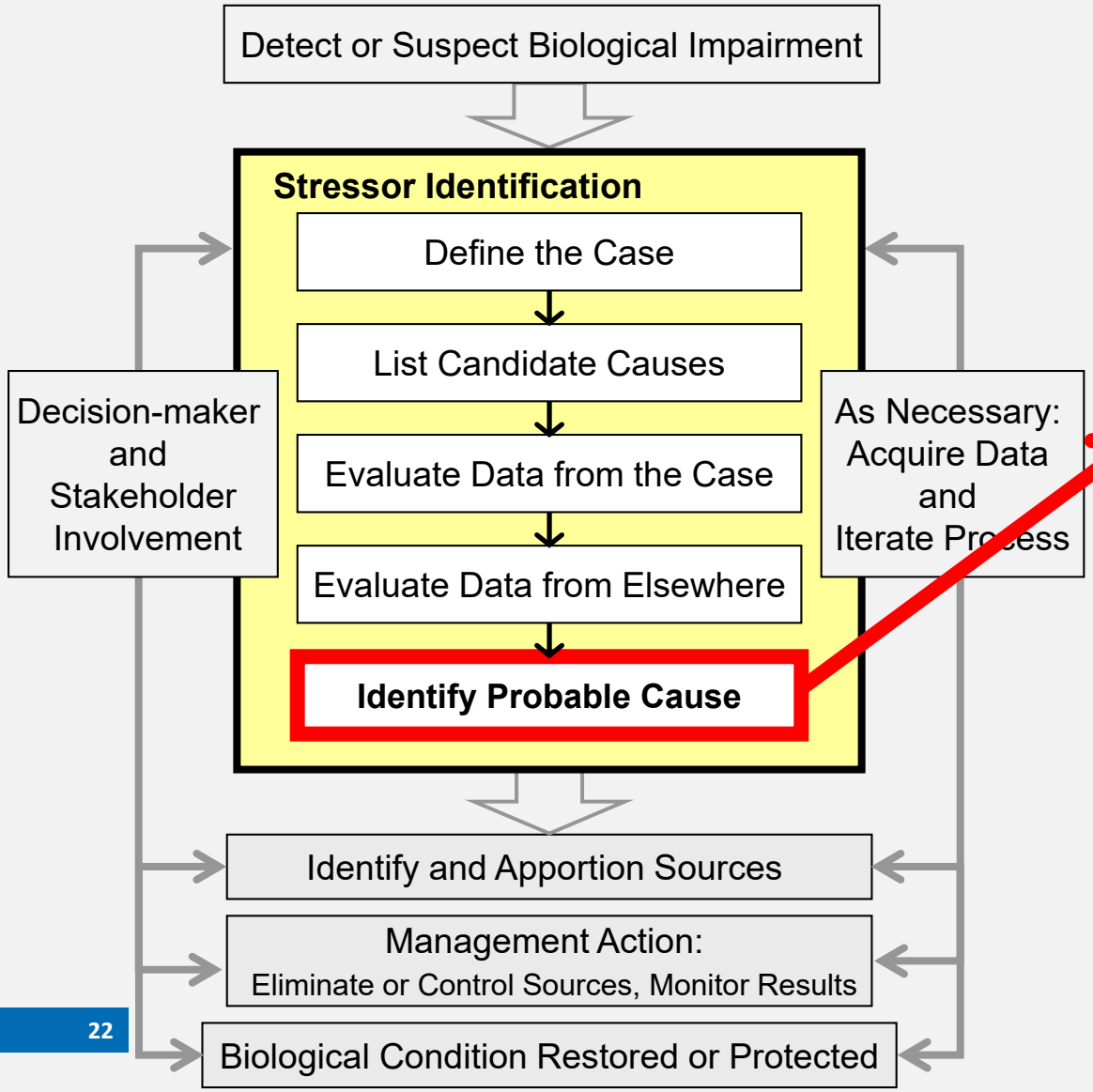
- ***Spatial/temporal co-occurrence***
- Evidence of exposure or biological mechanism
- ***Causal pathway***
- ***Stressor-response relationships from the field***
- Manipulation of exposure
- Laboratory tests of site media
- Temporal sequence
- Verified predictions
- Symptoms

Data from elsewhere

- ***Stressor-response relationships from other field studies***
- ***Stressor-response relationships from laboratory studies***
- Stressor-response relationships from ecological simulation models
- Mechanistically plausible cause
- Manipulation of exposure at other sites
- Verified predictions
- Analogous stressors

indicates commonly available types of evidence

Step 5 – Identify probable cause



- Weigh the evidence for each cause
  - Eliminate if possible
  - Diagnose if possible
- Compare evidence across all causes

EPA United States Environmental Protection Agency

CADDIS Scoring System

+++	convincingly supports (or weakens – – –)
++	strongly supports (or weakens – –)
+	somewhat supports (or weakens –)
0	neither supports nor weakens
R	refutes
D	diagnoses
NE	no evidence

General principles for scoring evidence

- First + or – or 0
  - Based on logical implication of evidence that passes basic quality and relevance test
- Second + or –
  - Based on strength of association (e.g., large differences)
- Third + or –
  - Based on reliability of association (e.g., high sample sizes, excellent study design, control of confounders)
- Each type of evidence has strengths and weaknesses, which are reflected in the CADDIS scoring system

Weighing the evidence

- Weigh the body of evidence for each candidate cause
  - Evaluate quantity and quality of evidence
  - Identify compelling evidence
  - Evaluate consistency and credibility of evidence

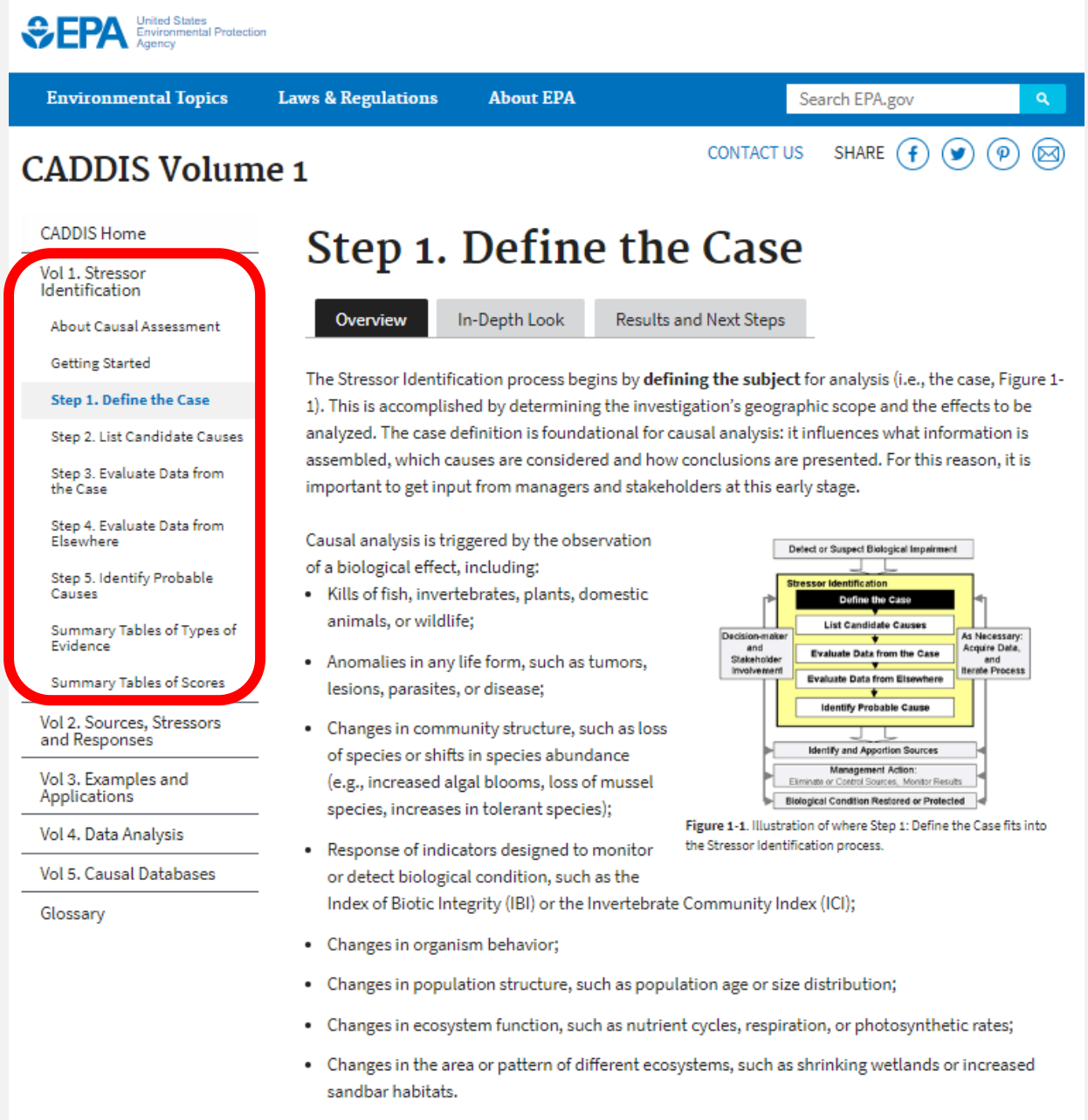
Consistency of Evidence	All available types of evidence support the case for the candidate cause.	+ + +
	All available types of evidence weaken the case for the candidate cause.	- - -
	All available types of evidence support the case for the candidate cause, but few types are available.	+
	All available types of evidence weaken the case for the candidate cause, but few types are available.	-
	The evidence is ambiguous or inadequate.	0
	Some available types of evidence support and some weaken the case for the candidate cause.	-

EPA United States Environmental Protection Agency

Comparing evidence and forming conclusions

- Compare the evidence across candidate causes, even when there is a “smoking gun”
  - Determine if there is more than one likely cause
  - Determine your level of confidence in the results
- Identify cause(s) best supported by the evidence
- Classify causes (e.g., likely, unlikely, uncertain)
- Refine and iterate, as needed

- Overview
- In-Depth Look
- Results and Next Steps



Spatial/Temporal Co-occurrence

Concept

Spatial/Temporal Co-occurrence: The biological effect must be observed where and when the cause is observed, and must not be observed where and wh

Helpful Links

- Concept
- Examples

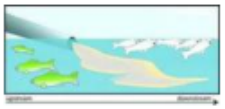


Figure 3-1a. Spatial/Temporal Co-occurrence with Upstream/Downstream Comparisons, Supports.

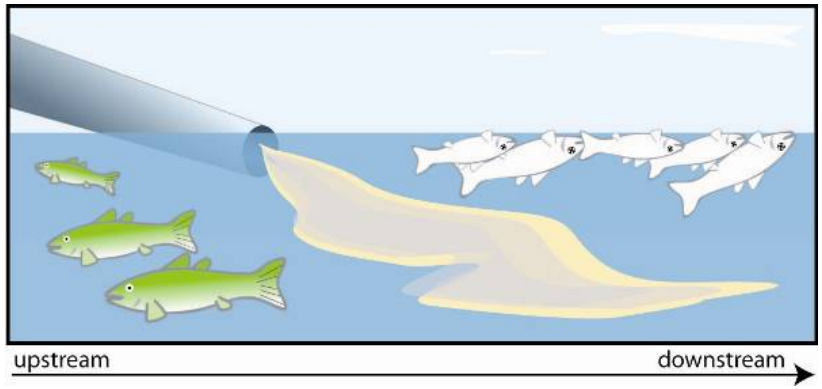
Click for more information

Additional Illustrations

Examples

Consider increased susp
cause of reduced aquat
support or weaken the
cause, based on spatial

- Supporting evidence
concentrations are h
unimpaired referenc
- Supporting evidence (temporal co-occurrence) - Suspended solid concentrations are episodic, and



SUPPORTS

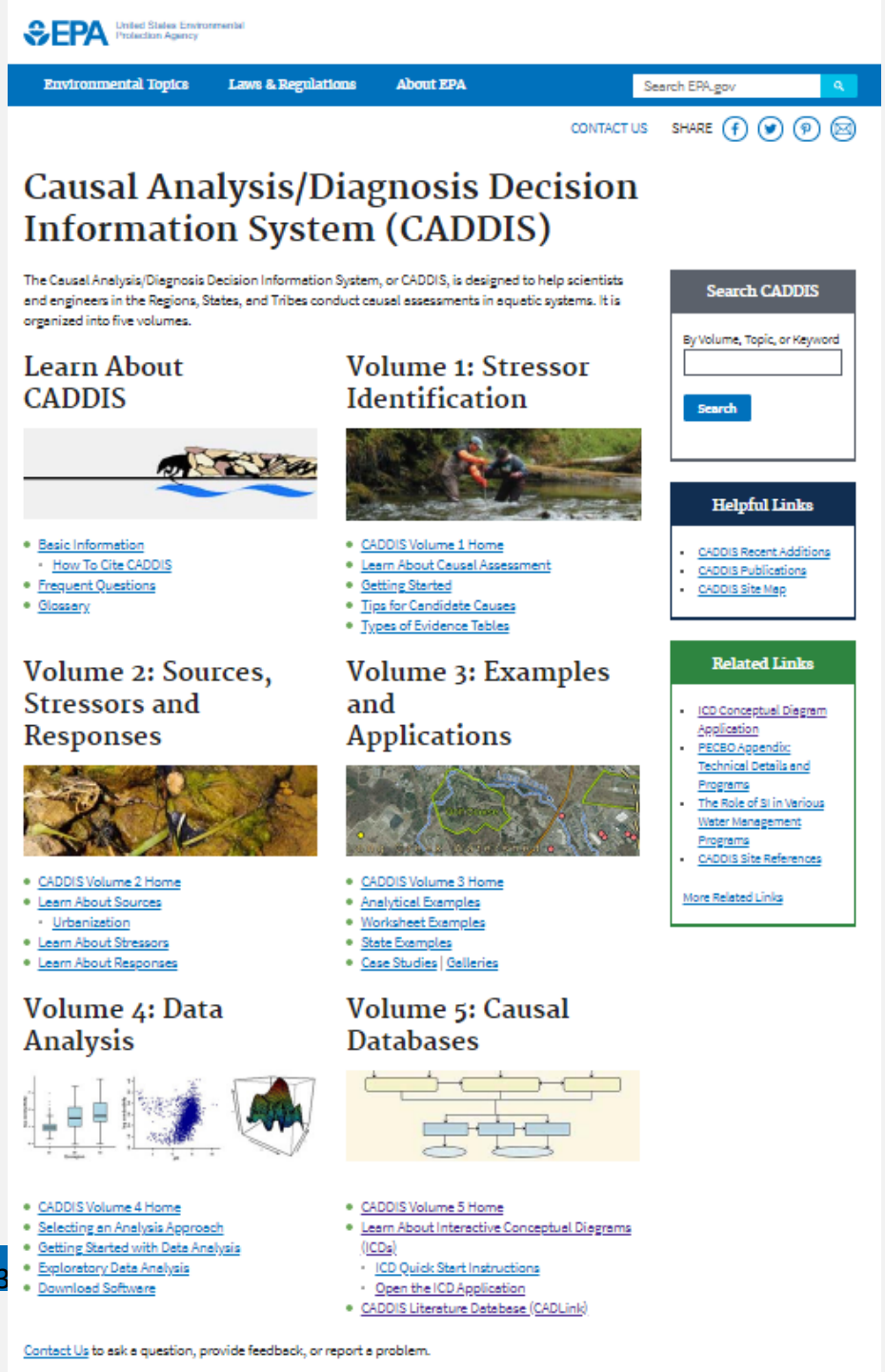
Impairment occurs where or when exposure to stressor occurs

WEAKENS

Impairment does not occur where or when exposure to stressor occurs

Summary Table of Scores

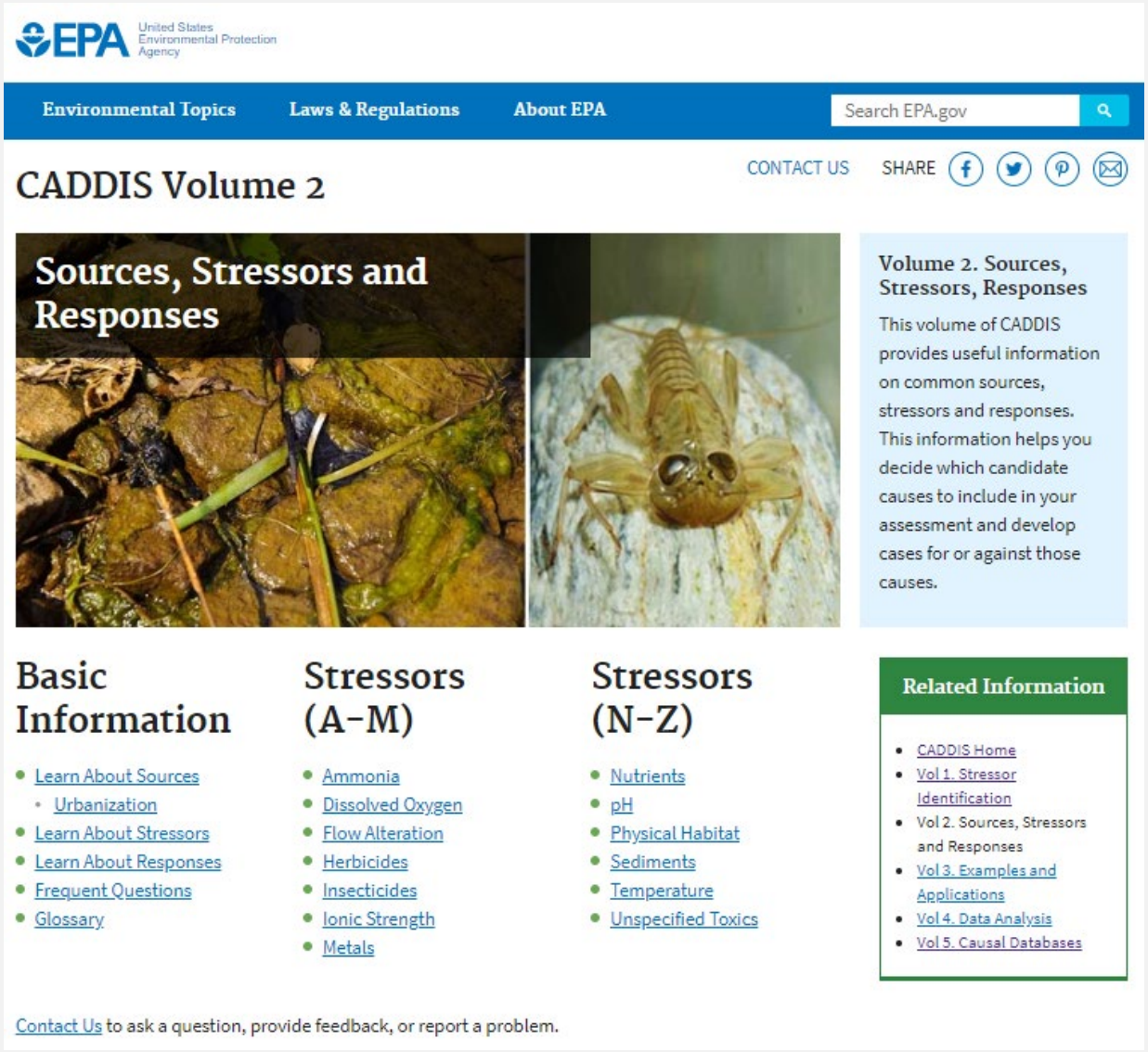
Type of Evidence	Finding	Interpretation	Score
Types of Evidence that Use Data from the Case			
Spatial/Temporal Co-occurrence	The effect occurs where or when the candidate cause occurs, **OR** the effect does not occur where or when the candidate cause does not occur.	This finding *somewhat supports* the case for the candidate cause, but is not strongly supportive because the association could be coincidental.	+
	It is uncertain whether the candidate cause and the effect co-occur.	This finding *neither supports nor weakens* the case for the candidate cause, because the evidence is ambiguous.	0
	The effect does not occur where or when the candidate cause occurs, **OR** the effect occurs where or when the candidate cause does not occur.	This finding *convincingly weakens* the case for the candidate cause, because causes must co-occur with their effects.	- - -
	The effect does not occur where and when the candidate cause occurs, **OR** the effect occurs where or when the candidate cause does not occur, and the evidence is indisputable.	This finding *refutes* the case for the candidate cause, because causes must co-occur with their effects.	R



CADDIS

- **Volume 1:** Stressor Identification
- **Volume 2:** Sources, Stressors & Responses
- **Volume 3:** Examples & Applications
- **Volume 4:** Data Analysis
- **Volume 5:** Causal Databases

Vol 2: Sources, Stressors & Responses



- Ammonia
- Dissolved oxygen
- Flow alteration
- Herbicides
- Insecticides
- Ionic Strength
- Metals
- Nutrients
- pH
- Physical Habitat
- Sediments
- Temperature
- Unspecified Toxics
- Urbanization

Temperature

On this Page

Temperature is the concentration of thermal energy in a substance such as water. The phrase "thermal regime" is used when emphasizing the temporal and spatial distribution of temperature.

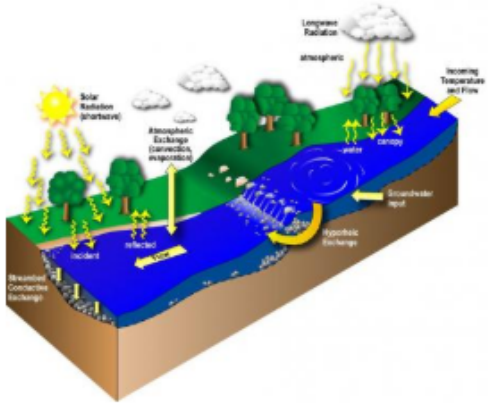


Figure 1. Major heat flux processes in streams. Click diagram to view a larger version.

Adapted from Moore et al. (2005) and Johnson and Jones (2000)

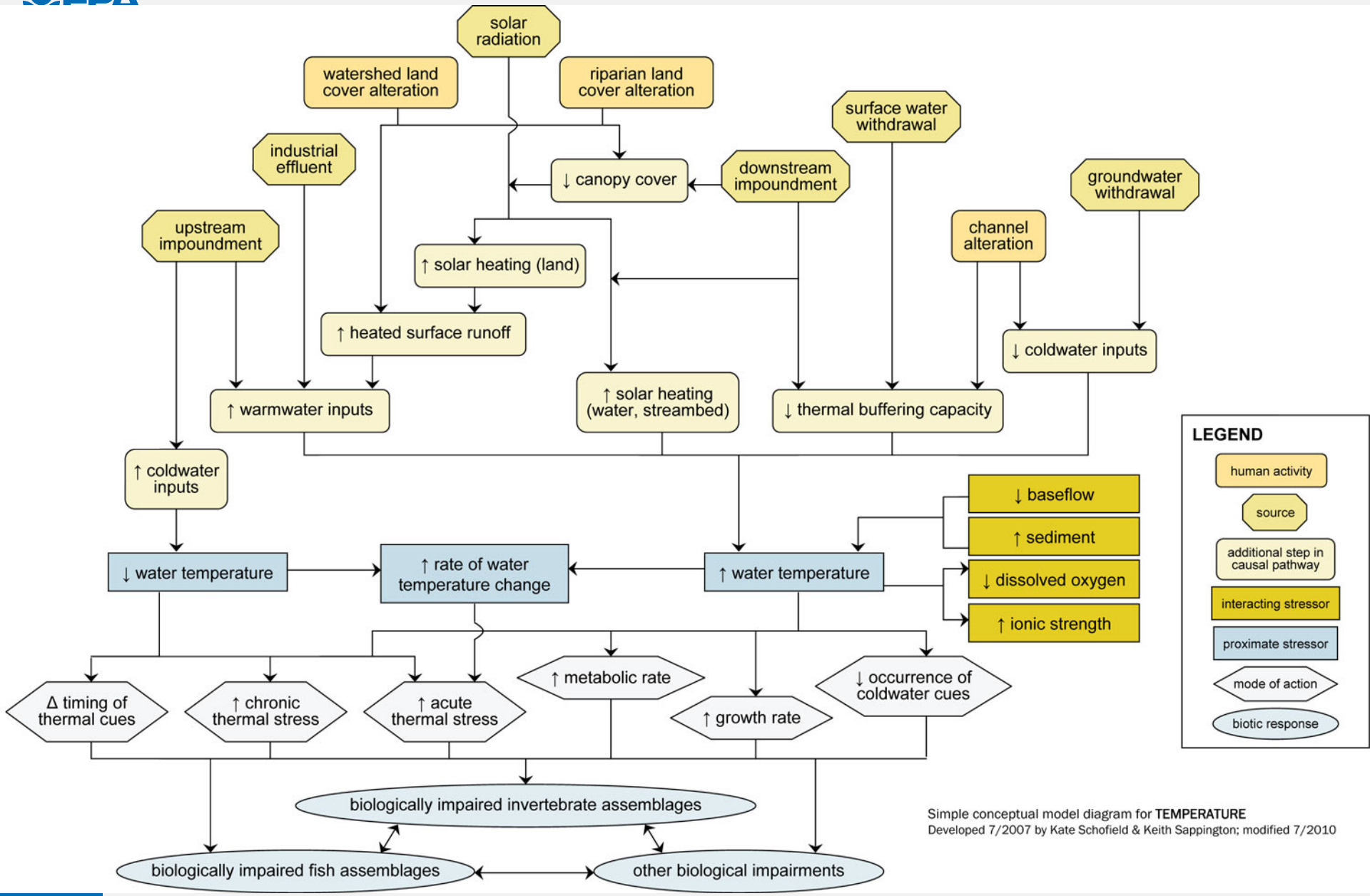
Temperatures in streams and rivers are influenced by many atmospheric and hydrologic processes affecting the movement of heat (see Figure 1). In turn, temperature plays a fundamental role in shaping the structure and function of aquatic systems (see Table 1). It is frequently used as a basis for classifying streams (e.g., coldwater, warmwater).

This module provides advice for deciding whether to include temperature in your list of candidate causes. You may go directly to a specific section of interest by clicking on the tabs above.

Table 1. Example Attributes of Aquatic Ecosystems Affected by Temperature

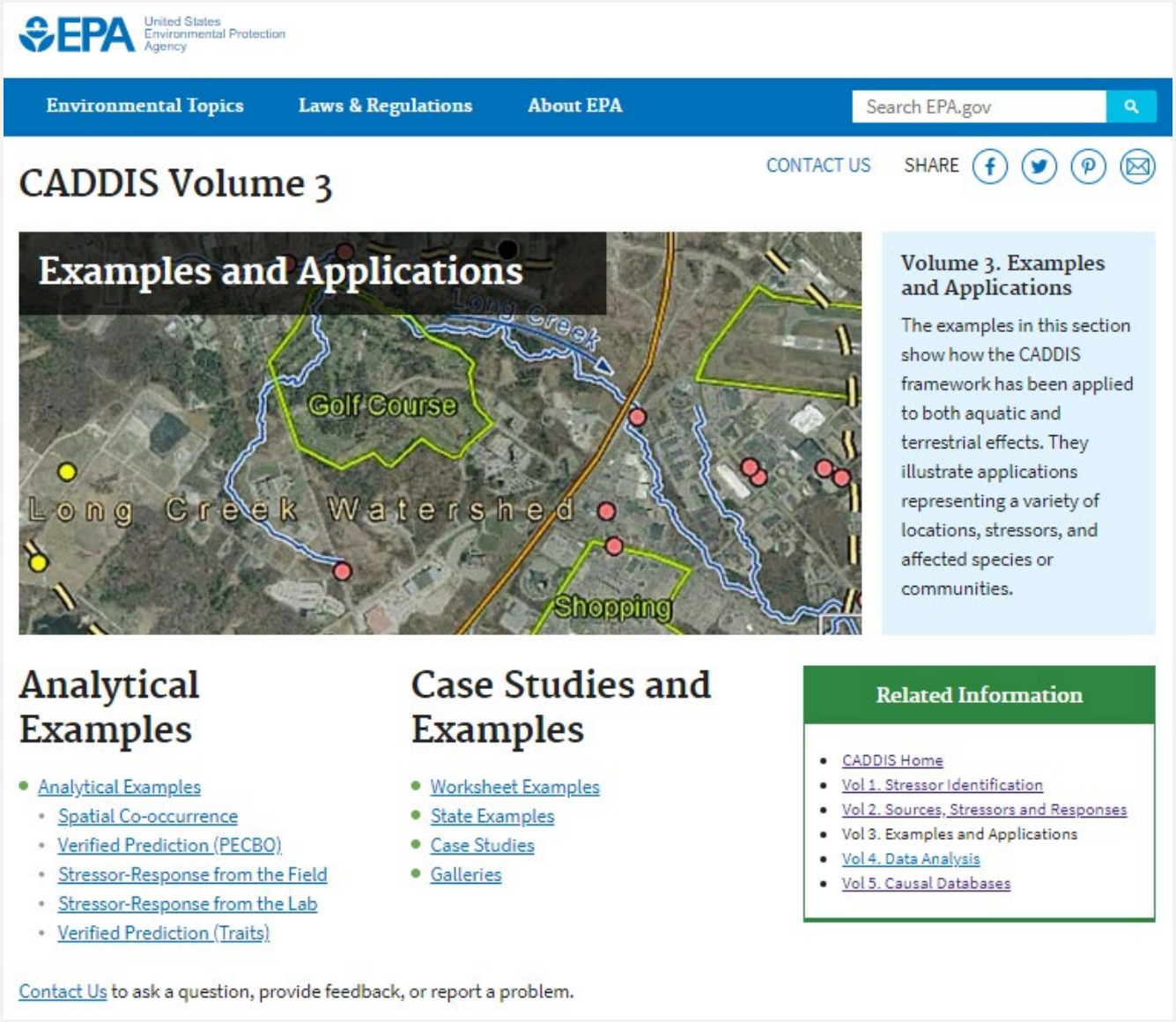
Category	Example Attributes
Physical	Water density, thermal stratification, solubility of oxygen and other chemicals

- Overview
- When to List
- Ways to Measure
- Conceptual Diagrams
- References



Simple conceptual model diagram for **TEMPERATURE**
Developed 7/2007 by Kate Schofield & Keith Sappington; modified 7/2010

Vol 3: Examples & Applications



Analytical Examples

Example 1. Spatial Co-occurrence with Regional Reference Sites

On this Page

Introduction

We would like to determine whether stream temperatures observed at an Oregon test site are higher than those at regional reference sites. If temperatures at the test site are higher than reference expectations, then we can conclude that increased temperature spatially co-occurs with the observed impairment. Conversely, temperatures at the test site that are comparable to temperatures at regional reference sites would suggest that increased temperature does not spatially co-occur with the observed impairment.

Data

The Oregon Department of Environment Quality (ORDEQ) deployed continuous temperature monitors in streams from 1997-2002. These temperature monitors recorded hourly temperature measurement which were then summarized as seven day average maximum temperatures in degrees C (7DAMT). Sites were also characterized by the geographic location (latitude and longitude), elevation, and catchment area. Reference sites were designated in Oregon based on land use characteristics.

Analysis and Results

Scatter plots are first used to examine the variation of stream temperature with different natural factors. The factors that are chosen (e.g., elevation, geographic location) must not be associated with local human activities. This initial data exploration suggests that stream temperature in reference sites are inversely related with both elevation and latitude (Figure 1). Next, regression analysis is used to model stream temperature as a function of elevation and latitude.

Related Information

Analytical Techniques Used

- Scatterplots
- Regression Analysis
- Controlling for Natural Variability

Type of Evidence Supported

- Spatial/Temporal Co-occurrence

- Introduction
- Data
- Analysis and Results
- How Do I Score this Evidence?

→ Vol 4: Data Analysis

→ Vol 1: Stressor Identification

EPA United States Environmental Protection Agency

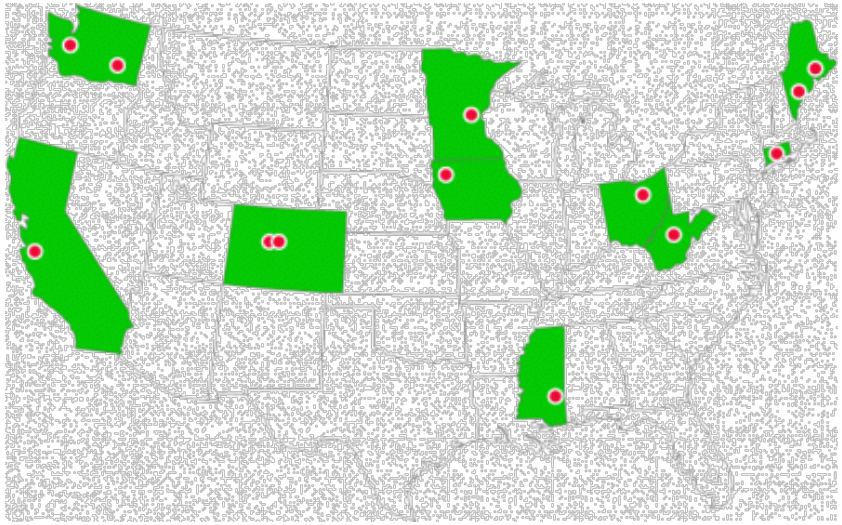
- CADDIS Home
- Vol 1. Stressor Identification
- Vol 2. Sources, Stressors and Responses
- Vol 3. Examples and Applications
  - Analytical Examples
  - Worksheet Examples
  - State Examples
  - Case Studies
  - Galleries
- Vol 4. Data Analysis
- Vol 5. Causal Databases
- Glossary

Applications to the Clean Water Act 303d/Total Maximum Daily Load (TMDL) Program

Figure 1. States that have used or are currently using methods from the Stressor Identification process, shown in green.

Stressor Identification typically occurs after a water body is listed as impaired by biological or unknown causes. This is typically before the development of a TMDL or watershed management plan. U.S. EPA does not require documentation of how pollutants or watershed management targets are identified. However, we have found some evidence of the adoption of our methods (see Figure 1). This list does not include states which have conducted full stressor identification case studies.

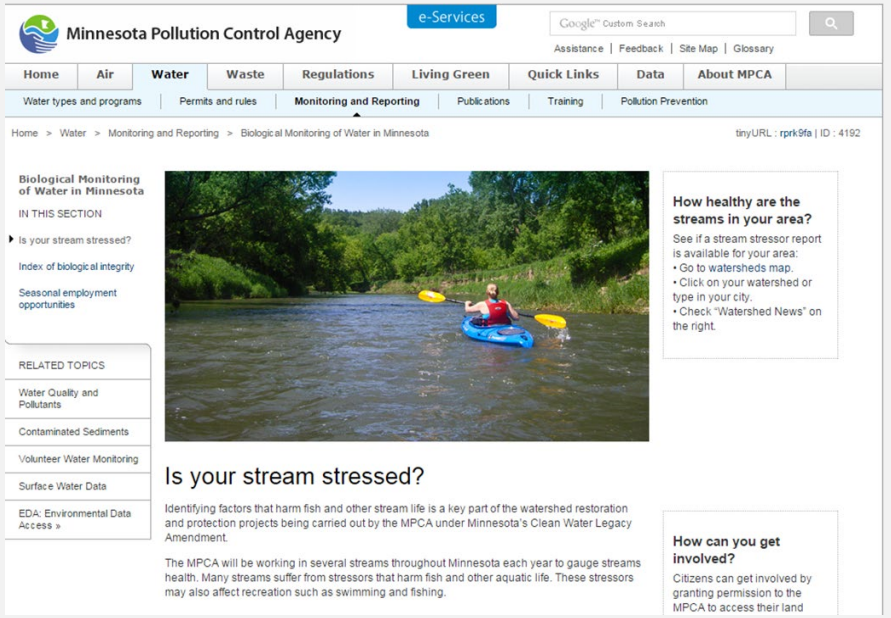
- TMDL Program
- Criteria
- Waste Site(s)
- Endangered Species
- References



How are states using CADDIS?

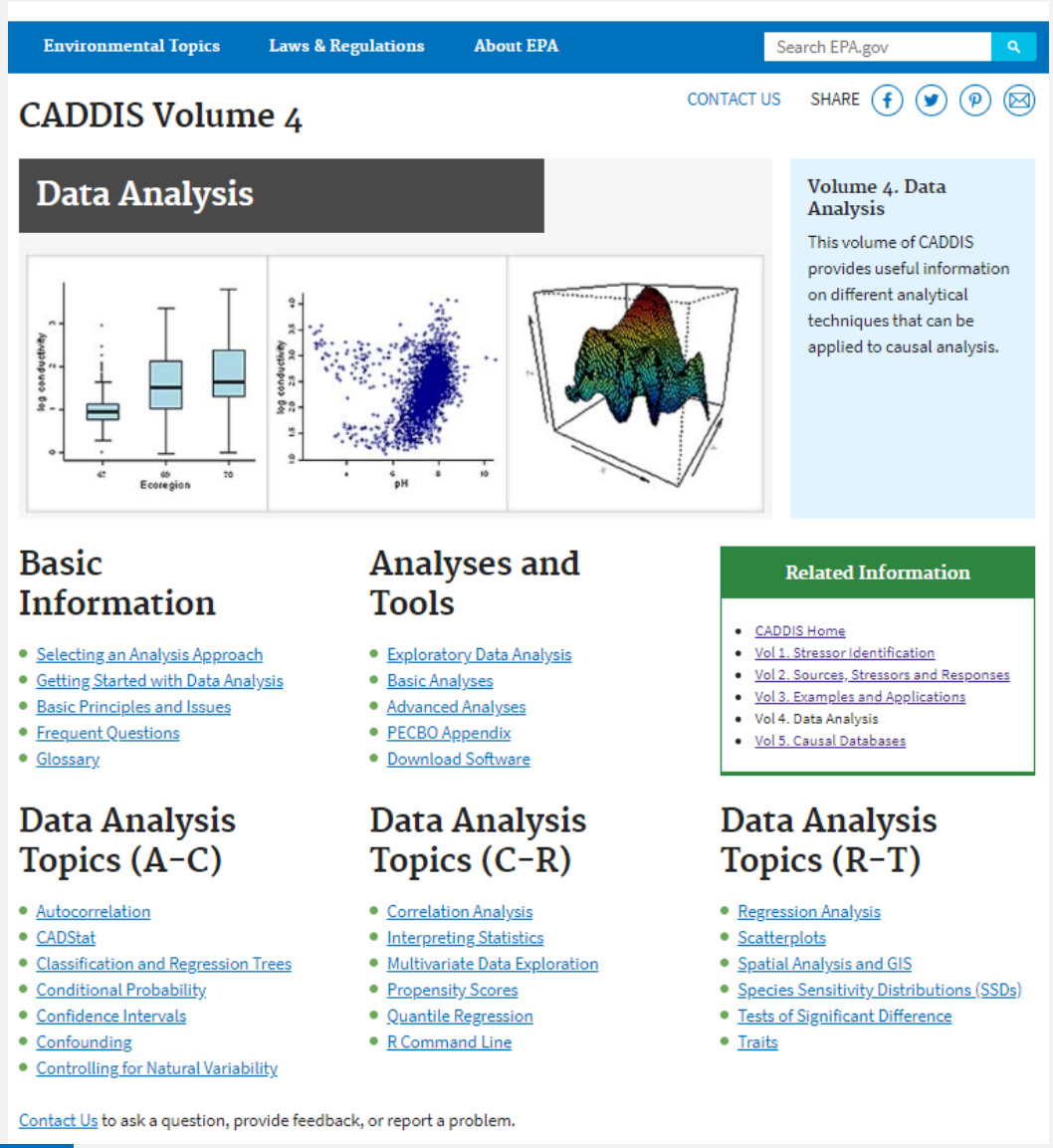
Adapted for state-specific applications in 19 states

- **Minnesota**: causal assessment applied systematically to watersheds across the state
- **Pennsylvania, West Virginia and Virginia**: causal assessments steer data collection efforts in fish health investigations
- **Connecticut and Maine**: causal assessments provide input to restoration decisions
- **California**: collaborating with ORD to make causal assessments faster, cheaper and routine



youtube.com/watch?v=K2x20Q1df48

Vol 4: Data Analysis



- Selecting an Analysis Approach
- Getting Started
- Basic Principles & Issues
- Exploratory Data Analysis
- Basic & Advanced Analyses
- Download Software

EPA United States Environmental Protection Agency

Getting Started with Data Analysis

- Assembling, matching, organizing data
- Data quality
- Links to relevant databases

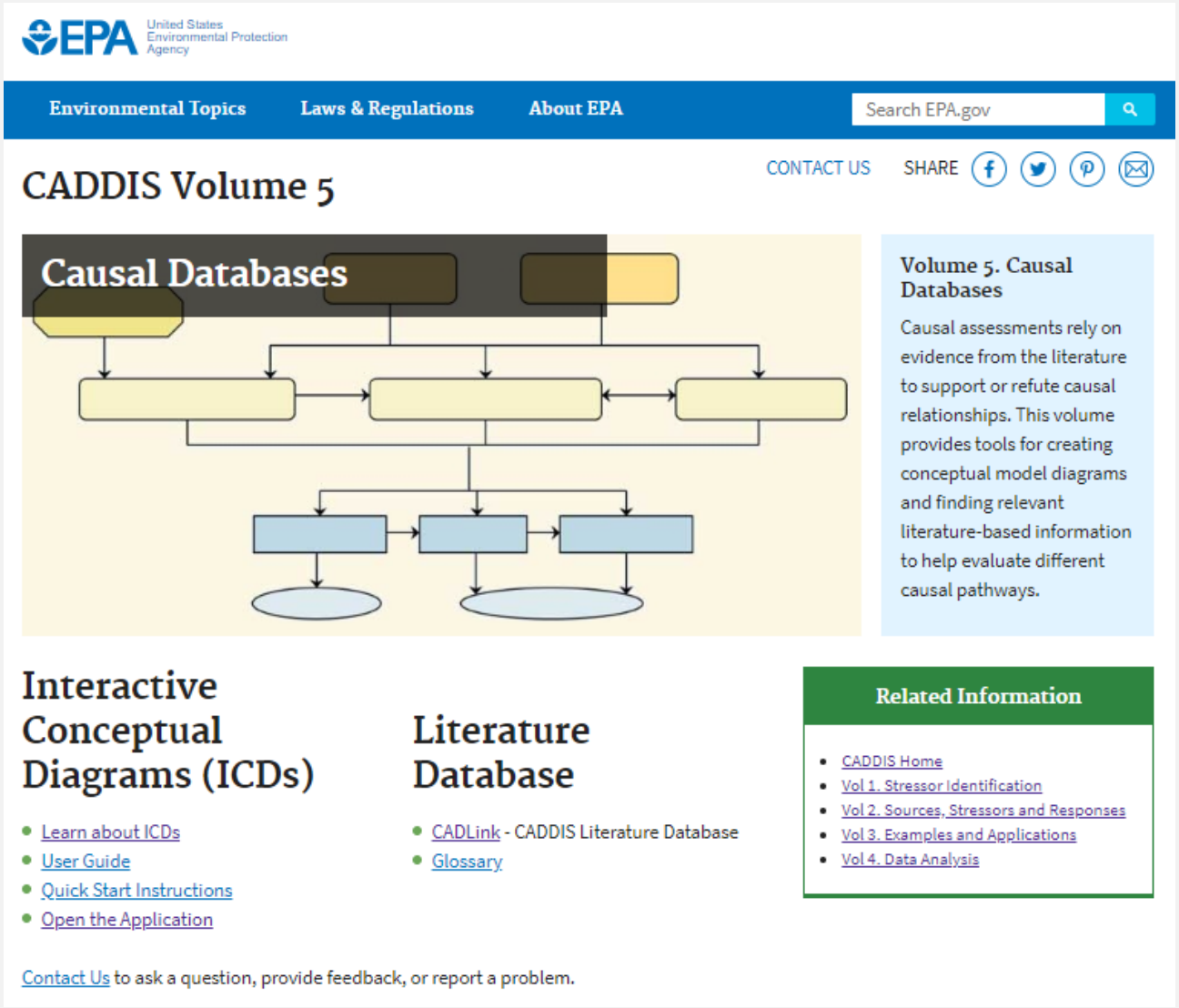
Exploratory and Basic Data Analysis

- Variable distributions, scatterplots, correlation analysis, conditional probability, multivariate approaches
- Significance tests, regression analysis, quantile regression, classification and regression tree (CART) analysis

Download Software

- CADStat
- Species Sensitivity Distribution (SSD) Generator
- R Command Line Tutorial

Vol 5: Causal Databases



- CADLink
- ICD (*)

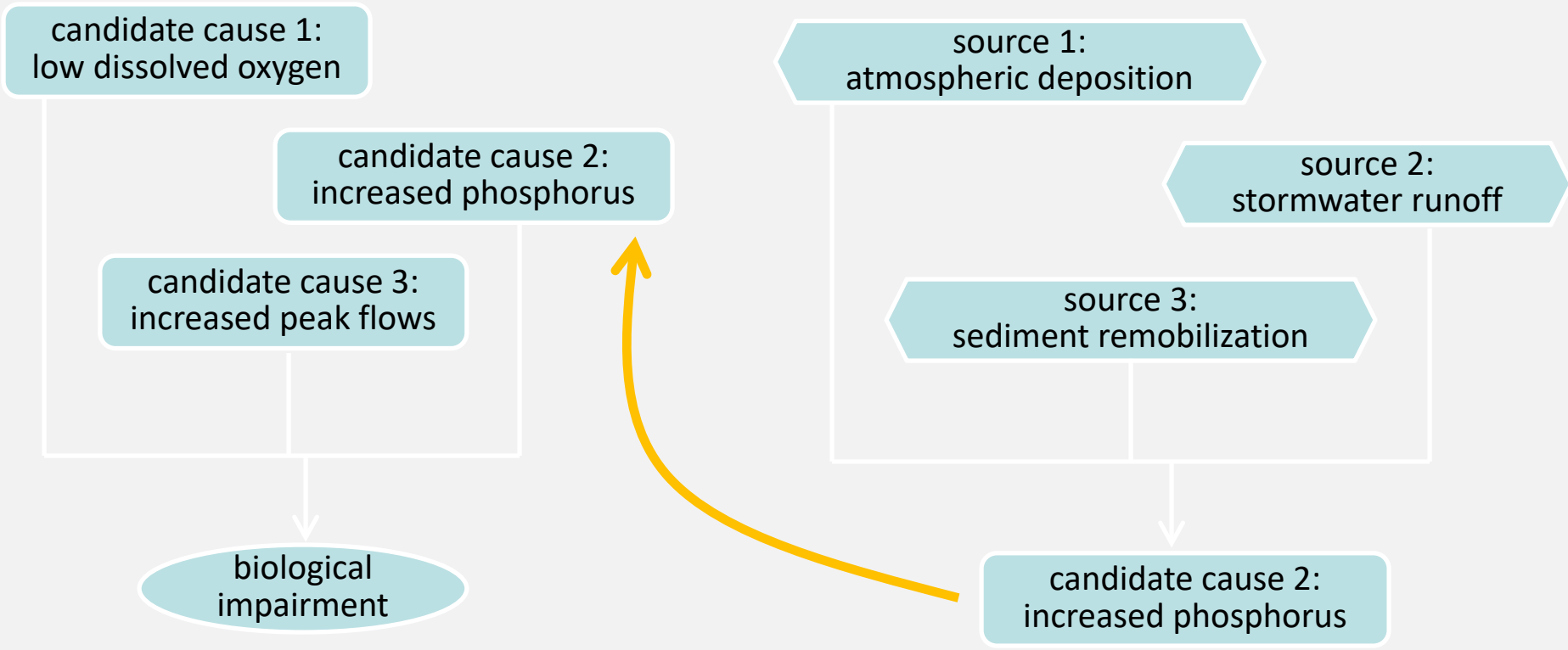
CADLink

- Database of evidence extracted from published literature
- Focused on cause-effect associations
- Includes information on study design, location, analytical results
- “Public” users can search existing records, “registered” users can enter new/modify existing records

What's next for CADDIS?

- Develop methods for "rapid" causal assessment
- Develop evidence databases
  - CADLink
  - EcoEvidEx (Ecological Evidence Exchange)
- Link evidence databases to visualization applications
  - EcoDIVER (Ecological Database and Interactive Visualizations of Evidence Records)
- Other suggestions?
  - Let us know what would be most useful for you (via email or CADDIS Contact Us page)

Causal assessment (and CADDIS) can be applied more broadly…



Contact

Kate Schofield, PhD
Integrated Environmental Assessment Branch
Center for Public Health and Environmental Assessment
US EPA Office of Research and Development
Schofield.Kate@epa.gov
202-564-2640

Visit CADDIS @ epa.gov/caddis

The views expressed in this presentation are those of the author and do not necessarily reflect the views or policies of the US EPA.