Received: 12 July 2022 | Accepted: 24 January 2023

DOI: 10.1002/2688-8319.12216

Ecological Solutions and Evidence
BRITISH ECOLOGICAL SOCIETY

REGISTERED REPORT STAGE 2

How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review

T. Rytwinski[1] | H.-Y. Lin[1] | M. Harper[1] | K. E. Smokorowski[2] | A. Smith[1] | J. L. Reid[1] | J. J. Taylor[1] | K. Birnie-Gauvin[3] | M. J. Bradford[4] | J. A. Crossman[5] | R. Kavanagh[6] | N. W. R. Lapointe[7] | K. Turgeon[8] | S. J. Cooke[1]

[1]Canadian Centre for Evidence-Based Conservation, Department of Biology and Institute of Environmental and Interdisciplinary Sciences, Carleton University, Ottawa, Ontario, Canada

[2]Great Lakes Laboratory for Fisheries and Aquatic Sciences, Fisheries and Oceans Canada, Sault Ste. Marie, Ontario, Canada

[3]Section for Freshwater Fisheries and Ecology, National Institute for Aquatic Resources, Technical University of Denmark, Silkeborg, Denmark

[4]West Vancouver Laboratory, Fisheries and Oceans Canada, West Vancouver, British Columbia, Canada

[5]Fish and Aquatics, BC Hydro, Castlegar, British Columbia, Canada

[6]Fish and Fish Habitat Protection Program, Fisheries and Oceans Canada, Burlington, Ontario, Canada

[7]Canadian Wildlife Federation, Kanata, Ontario, Canada

[8]Department of Natural Sciences, Université du Québec en Outaouais, Gatineau, Québec, Canada

Correspondence
T. Rytwinski
Email: trina.rytwinski@carleton.ca

Funding information
Carleton University; Fisheries and Oceans Canada; Natural Sciences and Engineering Research Council of Canada

Handling Editor: Tsungai Zengeya

Abstract

1. We conducted a systematic review on the impacts of natural causes of variation or changes in flow magnitude (resulting from climatic variability and broad-scale drivers such as climate-induced change) on fish abundance and biomass in temperate regions.
2. Following our systematic review protocol (Birnie-Gauvin et al., 2021), we examined commercially published and grey literature originally identified during a recent systematic map process (Rytwinski et al., 2020) and a systematic search update. Articles were screened using an a priori eligibility criteria, with consistency checks performed at the title and abstract, and full text screening stages. All eligible articles were assessed for study validity. A narrative synthesis included all available evidence (300 studies from 219 articles), and meta-analyses (1863 datasets from 193 studies) were conducted where appropriate.
3. Most studies were conducted in the United States (65%) on genera in the Salmonidae and Leuciscidae families (29% and 28%, respectively), and mainly investigated variation in flow magnitude, followed by droughts and floods.
4. We found that fish abundance and biomass responses to changes in natural flow magnitude were mainly negative, but analyses do not support clear generalizable signals across all contexts (e.g. types of changes in flow magnitude, taxa, locations). When exploring reasons for heterogeneity in effect sizes, we found a detectable effect of intervention type on average fish abundance within the first year of the natural change in flow magnitude, with floods and droughts associated with overall negative responses. However, these patterns were more variable when considering between-year variation or specific taxonomic responses.
5. Our results suggest that while immediate responses were more apparent and relatively consistent within specific types of natural events, fish populations may recover after such events (i.e. ≥2 years post-natural event); however, most studies

were short in duration (<2 years) and longer-term effects were more variable and may be context dependent.

6. To improve our understanding of species-specific and population-level effects, as well as time-lags in fish responses to natural changes in flow regimes, standardized, long-term continuous monitoring both before and after a change in flow magnitude are needed to address knowledge gaps. Studies that focus on systems outside North America are recommended.

KEYWORDS

climate change, discharge, evidence synthesis, fish density, flow modification, high flow, low flow, seasonal variation

1 | INTRODUCTION

The natural flow paradigm stresses that the various components of the flow regime (including magnitude, frequency, duration, timing, and rate of change of flow events) play an important role in maintaining the ecological integrity and diversity of aquatic ecosystems (Bunn & Arthington, 2002; Olden & Poff, 2003; Poff et al., 1997). Wildlife, and fish in particular, have adapted over time to the natural dynamics of their environment, including changes in flow regimes (Lytle & Poff, 2004). However, there is growing evidence of the potential negative consequences of altered flow regimes on fluvial ecosystems and the fisheries they support, including changes to physical habitat, habitat access, food supplies, behaviour, community composition, energy expenditure, and population dynamics (Clarke et al., 2008; Murchie et al., 2008). As such, the importance of this natural flow variability in maintaining healthy fluvial ecosystems has recently become a primary focus for water resource managers, indicating the need for a better understanding of flow-ecosystem response relationships for effective management of these systems (Gillespie et al., 2015).

At the request of a Canadian natural resource management agency and regulator (i.e. Fisheries and Oceans Canada, DFO), a systematic map was recently conducted (Rytwinski et al., 2020) to provide a summary of the existing literature base on the impacts of flow-regime changes on outcomes of freshwater or estuarine fish productivity in temperate regions. From this mapping exercise, 11 potential subtopics (i.e. evidence clusters) were identified as areas that had sufficient coverage to allow systematic reviewing. Based on the presence of sufficient evidence and relevance of the topic to Canadian stakeholders, four of these evidence clusters were identified as priority candidates for full systematic reviewing: (1) the effect of alterations to flow magnitude due to hydropower production on fish abundance, (2) the effect of natural changes in flow magnitude on fish abundance, (3) the effect of alterations to flow magnitude due to hydropower production on fish community diversity and species richness and (4) the effect of natural changes in flow magnitude on fish community diversity and species richness. Since publication of the mapping exercise, evidence cluster (1) has been addressed with a systematic review (i.e. the effect of alterations to flow magnitude due to hydropower production on fish abundance; see Harper et al., 2022). Here, we present results of a systematic review addressing evidence cluster (2), the effect of natural changes in flow magnitude on fish abundance. Note, articles focusing on fish diversity responses to hydropower production and natural changes in flow magnitude (i.e. related to evidence clusters 3 and 4) were included in searches and screened for eligibility during the review process for evidence clusters 1 and 2; however, they were removed at the synthesis stage to be assessed in separate future systematic reviews due to time and resource constraints; focusing instead on the largest evidence clusters first (i.e., abundance). For further details on background, topic identification, stakeholder involvement, and our conceptual model, see our systematic review protocol (stage 1 registered report; Birnie-Gauvin et al., 2021).

The primary objective of the systematic review is to clarify, from the existing literature, how natural changes in flow magnitude (i.e. climatic variability and broad-scale drivers such as climate-induced change) affect fish abundance (including abundance, density, catch-per-unit-effort metrics) and biomass (including biomass and yield metrics) in temperate regions. Furthermore, we address the secondary question of: To what extent do factors such as fish taxa, natural change type, specific outcome metrics (e.g. fish abundance versus density), life history characteristics, study design and setting, influence the potential impact of changes in flow magnitude due to natural causes on fish abundance and biomass?

2 | MATERIALS AND METHODS

2.1 | Search strategy

This systematic review followed detailed methods described in the systematic review protocol (stage 1 registered report) by Birnie-Gauvin et al. (2021). In doing so, this review was performed following, as closely as possible, the guidelines of the Collaboration for Environmental Evidence (CEE, 2018), and conforms to ROSES reporting standards (i.e. detailed forms for ensuring evidence syntheses report their methods to the highest possible standards; see Haddaway et al., 2018; and completed form in Rytwinski et al., 2022a). This review examined commercially published and grey literature originally identified during the systematic map process (searches performed in 2017; see Rytwinski et al., 2020) and a systematic search update (targeting literature from 2017 to 2021 using a subset of the search terms used for the systematic map; see Birnie-Gauvin et al., 2021 and Supporting Information 1 for full details and search results). No deviations were made from the protocol regarding our search strategy.

2.2 | Article screening and study eligibility criteria

A total of 1368 relevant studies (i.e. an experiment or observation that was undertaken over a specific time period at particular sites reported as separate waterbodies that were not treated as replicates within a single article) from 1199 articles, published between 1940 and 2017, were identified by the map exercise, with 188 cases considering natural flow magnitude alterations and fish abundance and 41 considering fish biomass metrics. Furthermore, all cases that were identified in the mapping exercise to have evaluated fish abundance and biomass responses to a change in (i) an unspecified flow component due to natural (109 cases), multiple (75), or unclear causes (18), or (ii) flow magnitude due to multiple (127) or unclear causes (77), were further screened for consideration to confirm whether any relevant information could be used. All articles containing information on these potentially relevant cases were pre-screened by a single reviewer at full text prior to the data extraction stage using the specific eligibility criteria developed for this review (Table 1).

Articles identified in the updated search using the databases and the search engine were screened at two stages: (i) title and abstract, and (ii) full text. Articles found by other means for the updated search (i.e. searching bibliographies of recent relevant reviews, evidence call-outs, etc.) were screened at full text. Prior to screening all articles, a consistency check was done at each stage on a subset of articles and discrepancies discussed; see Supporting Information 1 for further details on consistency checks. All articles were screened according to the established eligibility criteria developed in consultation with the Advisory Team (Table 1) and were only included when all criteria were met. A list of articles excluded at the full-text screening stage or during the data extraction stage, with reasons for exclusion, is provided in Supporting Information 2.

2.3 | Study validity assessment

Relevant studies identified during full-text screening or during data extraction, were assessed for internal study validity using a review-specific critical appraisal tool (see Supporting Information 3 for further details). This appraisal tool was made in consultation with the Advisory Team to ensure that it incorporated all components of well-designed studies and was informed by previous reviews (Harper et al., 2022; Macura et al., 2019; Martin et al., 2020). Study validity assessment took place at the same time as data extraction and was performed by a single reviewer following a consistency check (see Supporting Information 3). No studies were excluded based on study validity assessment; however, sensitivity analysis was used to investigate the influence of study validity categories when sufficient data were available (see 'Quantitative synthesis: Formal meta-analysis').

2.4 | Data coding and extraction strategy

2.4.1 | General data-extraction strategy

Following full-text screening, all articles identified as relevant from the updated search and those identified as relevant from the systematic mapping exercise underwent data extraction using a review-specific extraction form (see extraction sheet in Rytwinski et al., 2022b). If an article was deemed irrelevant after further screening at this stage, it was excluded from the review with justification (see Supporting Information 2). Attempts were made to identify supplementary articles and combine them with the most comprehensive article (i.e. primary source) during data extraction.

We extracted data on the following key variables: bibliographic information, study location and characteristics (e.g. geographic location, climate, and waterbody name), study design details (e.g. study dates and study design), intervention/exposure and comparator details (e.g. floods, droughts; see Table 2 for definitions), outcome (i.e. abundance [abundance, density, or catch per unit effort, CPUE], or biomass [biomass or yield]), sampling method(s) (e.g. type and size of sampling units), species (or species groups) (e.g. genus and species names; common and Latin names crosschecked with FishBase [Froese & Pauly, 2021] and Eschemeyer's Catalogue of Fishes [Fricke et al., 2020]) and life history stage, effect modifiers (see below under Section 2.5), and study validity assessment decisions. Coding within key variables was based on codes previously developed during the systematic map (Rytwinski et al., 2020) and expanded on through a partially iterative process as new options were encountered during scoping and extraction. Extraction was primarily based on information contained within each article and associated supplementary files; however, for one variable, data were extracted from an external source to the article, based on the location reported by the author (i.e. climate zone [Köppen-Geiger classification (Kottek et al., 2006: http://koeppen-geiger.vu-wien.ac.at/)]) (but see also Section 2.5 for two additional external cases).

Additionally, all articles included on the basis of full-text assessment underwent quantitative data extraction when possible. Sample size (i.e. number of waterbodies/sites within a single waterbody, years/months), flow magnitude alteration (i.e. flow magnitude before and after an extreme event such as a flood or drought) and outcome (i.e. reported abundance and biomass metrics) were extracted as presented in tables or text, and data from figures was extracted using data extraction software WebPlotDigitizer

TABLE 1 Article inclusion and exclusion criteria summarized from the stage 1 registered report (Birnie-Gauvin et al., 2021). Further criteria for consideration that were developed post-publication of the stage 1 report are shown in italic font.

Included	Excluded
Subject (population)	
Any fish species in north (23.5°N–66.5°N) or south (23.5°S–66.5°S) temperate regions. This included any resident (i.e. non-migratory) or migratory fish species, including diadromous species (i.e. fish that migrate between fresh and saltwater). All life stages were considered. Populations included those that were once stocked (but are no longer being stocked *or it was unclear if stocking was ongoing or occurred during study period*) or alien and became established in the waterbody. Studies located in freshwater or estuarine fluvial (i.e. water moving via gravity) ecosystems, such as lakes, rivers, streams, wetlands and marshes	Fish species in tropical or polar regions. Studies in a completely marine ecosystem or in a waterbody that did not have any moving water (e.g. irrigation ponds). *Studies related to altered flows in aquaculture, fish farms, or hatcheries*
Intervention/exposure	
Articles that described variability or a change in the magnitude of flow. Magnitude can be defined as the amount of water moving past a fixed location per unit time (Poff et al., 1997). Magnitude is therefore a measure of discharge, can refer to either relative or absolute discharge, and can be expressed in a variety of units. Natural causes of variation or a change in flow magnitude that directly (or near directly) affect fish were considered, including (i) those originating from climatic variation such as seasonal changes (e.g. rainfall, snowmelt, or ice), droughts and floods, that operate on annual or shorter time scales, or (ii) longer-term climate-induced changes in flow magnitude that would have delayed but potentially significant impacts on fish communities. These included natural (or near-natural) systems (i.e. those relatively unaffected by direct human pressure) as well as human-modified systems (as long as no anthropogenic changes in magnitude were made during the study period). Natural causes also included landslides and wildfires. Articles that reported unspecified multiple components affecting flow (i.e. do not report effects of components separately to isolate individual impacts of the flow components), were also included. A sensitivity analysis was carried out to investigate the influence of including such articles in the quantitative analysis	Studies that focused on other components of flow regime and did not also include an evaluation of flow magnitude (i.e. frequency, duration, timing [seasonality], rate of change, and surrogates thereof). Studies that only considered flow alterations due to anthropogenic causes (e.g. hydro-electric or nuclear facilities, dams without hydro, water withdrawal, land-use change, flow augmentation, or environmental flows). Studies that focused on drivers of change related to in-stream channel engineering, reduction in river length, construction of dikes, weirs, operation of hydropower plants and reservoirs, urbanization, transport infrastructure, deforestation, ditch construction, agricultural management practices, drainage of wetlands and agricultural areas, or construction of flood-retention basins
Comparator	
Relevant comparators included: (1) similar sections of the same waterbody that are not affected by a naturally caused change in flow magnitude (e.g. upstream condition); (2) separate but similar waterbodies without a naturally caused change in flow magnitude; (3) before a naturally caused change in flow magnitude within the same waterbody; or (4) time-series data within the same waterbody. Studies that did not include a comparator	No studies were excluded based on a comparator (or lack thereof)
Outcome	
Studies must have reported measured effects that indicate the potential for a change in fish abundance or biomass (i.e. direct flow-fish responses). Outcomes included those related to abundance, density, CPUE, biomass, and yield indices	Studies that evaluated some other direct response of fish productivity (e.g. growth, survival, or migration) or that considered indirect responses to altered flow. For example, if authors made an indirect link between the measured outcome of altered flow (e.g. growth of aquatic plants) and its 'potential' impact on fish (e.g. abundance)
Study design	
During protocol development, it was determined that many studies would likely not fit the typical Before/After (BA), Control/Impact (CI), Before/After/Control/Impact (BACI) or randomized controlled trial (RCT) structure. Therefore, we included all primary field-based studies that included quantification of fish abundance or biomass outcomes in relation to natural variability in flow magnitude, including the above-mentioned designs, as well as a reference conditional approach (RCA), normal range (NRange), and temporal (i.e. time series; TS) or spatial trend (S_TRENDS) designs. *Several modifications of the standard BA and CI study designs were identified and included: (1) Before/During (comparing a before period to a period during an event (e.g. flood or drought) rather than after the event had finished); (2) During/After (comparing a during period to an after period but no true before data); (3) alternate-CI (ALT-CI: a CI design comparing two levels of intervention on different water bodies but neither experiencing no intervention); and (4) alternate-BA (ALT-BA: a BA design comparing two levels of intervention at different times but with no true Before or After period). Because designs 2–4 do not include an identified 'zero-control', they were only considered in the narrative synthesis*	Studies with a single point in time with no comparison to another site. Theoretical studies, review papers and policy discussions
Language	
English at full text	Any study that was not in English at full text

TABLE 2 Types of interventions, flow-magnitude alterations considered (including direction and part of hydrograph) and their definitions. Definitions adapted from (Feldman, 2000; Gordon et al., 2004; Rytwinski et al., 2020; Smakhtin, 2001). Note that some alterations relate to not only magnitude but frequency and timing as well.

Intervention	Code	Description
Alteration to low flow magnitude	Increase_Low	Increased seasonal or event low flow. This was related to changes in base flow, reported as changes to base flow, low flow, or minimum flow conditions
	Decrease_Low	Decreased seasonal or event low flow. This was related to changes in base flow, reported as changes to base flow, low flow, or minimum flow conditions. If authors indicated that the flow magnitude of an event was outside the normal range but compared the event to the mean or median flow prior to the event, Decrease_Low applied
Alterations to average discharges	Increase_Discharge	Increased mean/median discharge. An increase in the mean or total flow magnitude for the flow period being considered
	Decrease_Discharge	Decrease mean/median discharge. A decrease in the mean or total flow magnitude for the flow period being considered
Alterations to high flow magnitude	Increase_High	Increased seasonal or event high flow. This was related to changes in peak flow, reported as changes to peak flow, high flow, or maximum flow conditions. If authors indicated that the flow magnitude of an event was outside the normal range but compared the event to the mean or median flow prior to the event, Increase_High applied
	Decrease_High	Decreased seasonal or event high flow. This was related to changes in peak flow, reported as changes to peak flow, high flow, or maximum flow conditions
Alterations to short-term variation of flow magnitude	Increase_Var	Increased flow variability. Increases in short-term variation (reported as a change in magnitude that occurred over a period of hours or less than 1 day). For example, before the intervention, peaks occurred 2 times in a 24-h period, but after the intervention, occurred 4 times in 24 h
	Decrease_Var	Decreased flow variability. Decreases in short-term variation (reported as a change in magnitude that occurred over a period of hours or less than 1 day). For example, before the intervention, peaks occurred 4 times in a 24-h period, but after the intervention, only twice in 24 h
Alterations to winter flow magnitude (all types)	Increase_Winter	Increased winter flow magnitude (including means, lows, highs or short-term variation). An increase in flows experienced during winter (i.e. a high flow event caused by winter snowmelt)
	Decrease_Winter	Decreased winter flow magnitude (including means, lows, highs or short-term variability). A decrease in flows experienced during winter (i.e. average discharge in winter in year 2 is less than in year 1) and could be caused by a variety of interventions including ice jams and low winter rains
Differences among measures flow magnitude	Difference	Two different levels of flow magnitude compared to each other (i.e. ALT-CI). This was not applicable to multiple levels of flow (as would occur in a time series or spatial trend) but was used to compare two values of flow magnitude in the absence of a clear flow alteration or extreme event (i.e. two different rivers with different spring flows) and applied to differences between extreme events (i.e. flood to drought)
Changes to more than one flow magnitude alterations	Mixed	More than one flow magnitude alteration occurred during the period of consideration (i.e. between the before and after periods a drought period was broken by a major flood). The After period was impacted by the cumulative effect of two or more changes in flow magnitude (i.e. Decrease_Low, Increase_High)
General changes to flow magnitude	N/A	For trend study designs. Variation in flow through time (i.e. time series) or space (spatial trend) without a true Before or comparator level of flow magnitude. Variation was due to fluctuations in flow but not associated with a specific flow event (i.e. flood or drought). Fluctuations in flow magnitude may be daily, monthly, seasonal, or annual, or may be differences in flow magnitude among >2 spatial locations. Examples of causes included seasonality and climate change

(Rohatgi, 2019) when necessary. Articles not included in quantitative extraction were those that did not allow for a quantitative assessment of the impact of changes in flow magnitude on a fish outcome (i.e. those that did not include an appropriate outcome type or comparator of no intervention; see Supporting Information 4 for further details). Prior to independent data extraction, a consistency check was performed by three reviewers on a subset of articles where comparisons of extracted data were made, and any discrepancies were discussed (see Supporting Information 4 for further details).

2.4.2 | Data extraction considerations

During data extraction there were several considerations made in defining our database of information. First, when multiple studies were reported in a single article they were entered as independent lines in the database and assigned a Study ID. Furthermore, a single study could report separate relevant comparisons (i.e. multiple non-independent datasets that share the same Study ID) (see Supporting Information 1, Table S1.8 for full definitions and Supporting Information 4 for full details). For quantitative analysis, we aggregated these datasets where applicable to reduce non-independence or selected a single option for inclusion in analysis (see 'Data preparation for quantitative synthesis' below).

Second, in comparison to studies that evaluate the impacts of flow changes due to hydropower operations (reviewed in Harper et al., 2022), we did not anticipate encountering many (if any) Control/Impact (CI) design studies investigating the impacts of flow magnitude changes on fish outcomes due to natural causes as it is difficult to foresee and thus plan for flood/drought events a priori. Instead, most studies use Before/After (BA), including Before/During (BD) designs. However, we also included cases where changes in fish outcomes were related to natural flow magnitude variation through time (i.e. time series [TS]) or across space (i.e. in different waterbodies within different flow-magnitude values [S_TRENDS]). These cases were treated separately from BA/BD study designs for quantitative analyses.

For BA and BD studies, we considered temporal replication at two levels: (i) within-year (unit of replication = number of months) and (ii) interannual (unit of replication = number of years). For full details of calculations of within-year and interannual variation see 'Before/After data extraction considerations' section in Supporting Information 4.

Multi-year TS and S_TRENDS were considered when no clear intervention was indicated by the authors (e.g. flood or drought), but changes to fish outcomes were related to flow magnitude changes through time or across space. In these cases, the relationship between outcomes and flow magnitude alterations was extracted in terms of a correlation coefficient (either calculated by authors or by review team members if not reported), and in terms of the reported maximum and minimum outcome variables, with associated flow magnitude values measured at the same time as the outcome. Unlike BA/BD study designs, studies could only be retained for quantitative synthesis if both quantitative measures of outcome (e.g. abundance or biomass) and flow-magnitude change were reported. Either observed outcomes or predicted outcomes during the ecological sampling period were extracted, but if both were present for the same study, we selected observed outcomes for inclusion in analyses. In some instances, although authors reported a trend study, it was apparent that an extreme event had occurred and there was sufficient information to determine a comparator and intervention period. If this occurred, studies were converted to a BA/BD-style study to enable assessment of the impact of this intervention (see 'Flow magnitude data extraction considerations' section in Supporting Information 4 for more details).

Third, to explore the influence of the degree of change in flow magnitude, we extracted information on flow-magnitude changes. For full details of considerations made in defining flow-magnitude changes, see 'Flow-magnitude data extraction considerations' in Supporting Information 4. Interventions, such as floods and droughts, or changes in high and low flows, were identified based on author definitions. In general, floods and droughts were considered extreme events, occurring outside the normal range of flow magnitude for the system or the temporal period (modified from Feldman, 2000; Gordon et al., 2004; Smakhtin, 2001), while changes to high and low flows were increases and/or decreases in maximum and minimum flow magnitudes, within the expected range of the system, but greater than flows that occurred in periods before or after the intervention (modified from Gordon et al., 2004; Smakhtin, 2001; see also Supporting Information 4, Figure S4.4). When extracting information on flow magnitude, we considered both quantitative and qualitative definitions of change (increases/decreases; see Table 2 for definitions of changes) if quantitative data were unavailable. The extracted flow magnitude was dependent on study design, but we extracted flow-magnitude metrics that coincided with fish sampling, unless otherwise stated by authors. This helped ensure that the fish present in the system at the time of sampling were likely to have interacted with the measured flow magnitude during their lifetime.

2.5 | Potential effect modifiers and reasons for heterogeneity

For all articles included on the basis of full-text assessment, we recorded, when available, key sources of potential heterogeneity (Table 3a). During the review process, additional effect modifiers and reasons for heterogeneity were identified and extracted from the studies (Table 3b). When sufficient data were reported and sample size allowed, these potential modifiers were used in meta-analysis (see Section 2.6.2 below) to account for differences among datasets via subgroup analysis (see Supporting Information 1, Table S1.8 for definitions of terms).

After consultation with the advisory team, there were effect modifiers that were originally identified in our protocol that were removed from consideration for this review. Due to limitations in time and resources, we did not search external to the article for waterbody characteristics including: temperature (air or waterbody), gradient, or stream order, as they were often not reported within the primary articles. Also, we did not include the type of comparator (i.e. spatial and/or temporal) because all studies that included a true comparator were temporal using either a BA or BD design (i.e. there were no CI or BACI study designs).

2.6 | Data synthesis and presentation

2.6.1 | Descriptive statistics and narrative synthesis

All relevant studies included on the basis of full-text assessments were included in a MS-Excel database with meta-data for each study (see Rytwinski et al., 2022b). All meta-data were used to develop descriptive statistics and narrative synthesis of the evidence, including figures and tables. All studies were included in the narrative synthesis regardless of study validity assessments.

2.6.2 | Quantitative synthesis

General approach and eligibility

We used three general approaches to quantitative synthesis depending on the intervention type (i.e. flood, drought, high/low

TABLE 3 Sources of potential heterogeneity extracted from eligible studies.

Moderators	Description
(a)	
Waterbody type	For example, stream, river, estuary
Freshwater ecoregion	Freshwater Ecoregions of the World [FEOW], Abell et al. (2008); https://www.feow.org/
Land use in immediate surroundings	For example, natural lands, agricultural [including silviculture], developed lands [housing, commercial etc.] or combinations thereof
Type of intervention	Drought, flood, high flows, low flows, other types of hydrological events [e.g. landslides], and variation [e.g. no event occurred during the study period, but natural variation or differences not linked to an event were present]
Presence of other flow component alterations	Frequency, duration, timing, rate of change, or surrogates of flow alteration, or any combinations of alterations; Yes: there are other alterations during the same time period as magnitude changes, and authors isolate impacts and/or account for such impacts in analysis; No: magnitude is the only flow regime alteration at site/time period; Unclear: the study does not specify a flow component or reports unspecified multiple components affecting flow
Sampling methods	Active or passive gear [electrofishing, net samples, trapping], angling, telemetry, mark-recapture, visual, passive integrated transponders [PIT tags], multiple methods, or others
Monitoring duration	Years
Life stage	Egg: eggs, nests, and redds; larvae: larvae, alevins, free embryos; age-0: fry, parr [0+], age-0+, YOY; juveniles: age-1+, parr [1+], juvenile, fingerling [if specific developmental stage is not identified], and smolt; adult: adult, spawner, and kelt; mixed: assorted life stages
Outcome metric	Abundance, density, CPUE, biomass, and yield
(B)	
Major habitat types	For example, temperate coastal rivers, temperate floodplain rivers and wetlands, montane freshwater; as defined by FEOW, WWF/TNC (2019); https://www.feow.org/global-maps/major-habitat-types; these reflect groupings of ecoregions with similar biological, chemical, and physical characteristics and are roughly equivalent to biomes for terrestrial systems
Water regime	For example, intermittent [temporary or seasonal streams/rivers/wetlands], perennial streams
Flow regime	Named as such for the purposes of this review; e.g. free-flowing systems, regulated systems where water flows are regulated by dams or intensive land-use changes, mixed
Direction of flow magnitude changes	For example, increases/decreases in average, low, or high flow magnitude, increases/decreases in short-term variation, variation through time or across space; see Table 2 for all types of changes and their definitions
The number of extreme events between sampling periods	For example, number of floods between Before and After periods
Duration of the event	Months, years
Time since intervention	Years; for BA designs only

flow changes, or natural variation through time/space), study design (i.e. BA/BD versus TS/S_TREND studies), and the availability of information in relation to fish outcomes or flow-magnitude changes from studies (see Figure 1; Supporting Information 4, Figure S4.3). The first approach included studies that focused on fish outcomes of one (or a few) clearly defined intervention(s) (flood, drought, or high- or low-flow changes), and there was replication in the intervention and comparator groups (i.e. at least two sampling time periods [months and/or years depending on within-year or interannual variation] both before and during/after the natural event). Here, we used formal meta-analytical procedures using the standardized mean difference effect size measure for quantitative synthesis (hereafter: formal meta-analysis). In doing so, summary and moderator effects were computed and tested from (i) individual study effect sizes using a standardized measure that required a temporal or spatial comparator, as well as replication in both the intervention and comparator groups, (ii) weighting individual effect sizes according to their reliability (i.e. standard error), and (iii) using random and mixed effects models (for estimating global and moderator effects, respectively) which partition within-study estimation error from true variation in effect sizes. This rigorous methodology is advantageous because it allows one to estimate summary effects and test for associations with factors that may cause variation among studies using the most reliable sources of information. It also allows for more confident conclusions to be drawn about how variation in fish responses is apportioned to different factors while accounting for the 'noise' created by within-study variation (Borenstein et al., 2010). As such, we placed higher emphasis on the results of analyses using this more formal approach.

To make use of the full evidence base, the second approach included the same studies from the first approach above but also studies

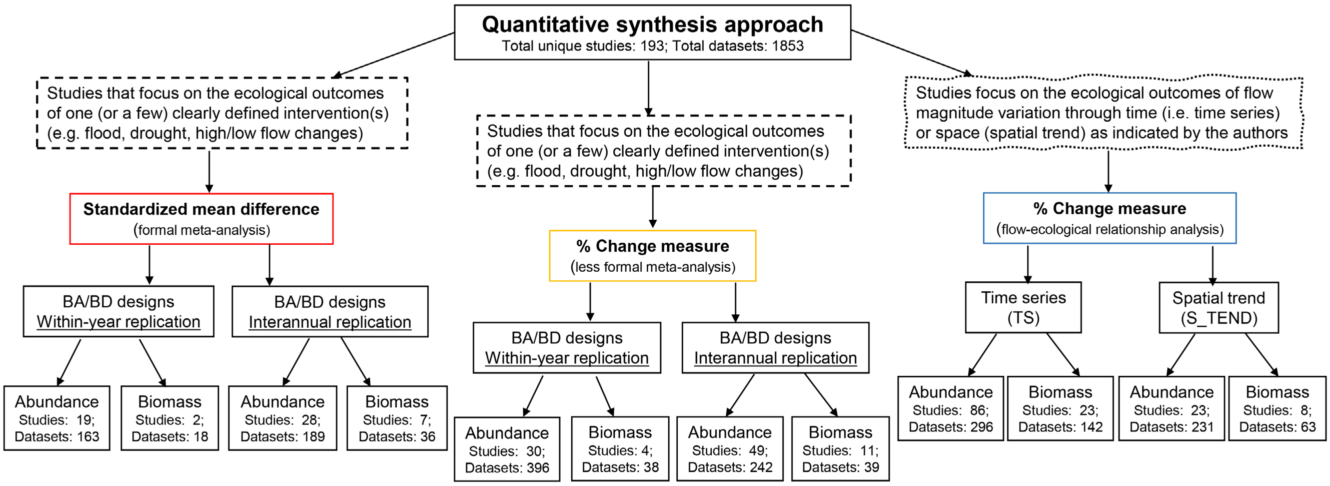

FIGURE 1 Flow diagram of quantitative synthesis approaches. Dashed rectangular boxes indicate studies that focus on fish outcomes of one (or a few) clearly defined intervention(s), whereas the dashed wavy box indicates studies that focus on fish outcomes of flow magnitude variation through time or space with no clearly defined intervention(s). Note, the total number of studies included in quantitative syntheses is the unique number of studies (among all approaches) because one study could be included in more than one approach, and the number of datasets is post-aggregation (see 'Data preparation for quantitative synthesis' and Supporting Information 5 for further details).

that lacked replication in either the intervention or comparator group. For such cases, we used a less formal meta-analytical approach with percent change as the measure for quantitative synthesis (hereafter: less formal meta-analysis).

The third approach included studies that focused on fish outcomes of flow magnitude variation through time (TS) or space (S_TREND) and there were no clearly defined intervention(s) as indicated by the authors or based on the reported data. In such cases, we calculated percent changes from the reported maximum and minimum outcome variables, with associated flow-magnitude values measured at the same time as outcome to explore quantitative flow-ecological relationships (hereafter: flow-ecological relationship analysis; see 'Time-series and spatial trend data extraction considerations' in Supporting Information 4). Unlike BA/BD study designs, studies could only be retained for flow-ecological relationship analysis if both quantitative measures of fish outcome and flow-magnitude change were reported. As noted by McManamay et al. (2013), although these studies are expected to present variable changes in flow magnitude and thus, fish outcomes responses, they still may provide results informative to generalizing flow-ecology relationships.

Both the less formal meta-analysis and flow-ecological relationship analysis approaches described above are similar in methodology to previous reviews on this topic (e.g. McManamay et al., 2013; Piniewski et al., 2017; Poff & Zimmerman, 2010) with the use of a proportional change metric for quantitative synthesis. Results of these two analyses will be used to supplement and compare with the more formal meta-analysis, as well as the results of past reviews. Studies not considered in any of the above quantitative approaches (and thus only considered for narrative synthesis) were those that: (i) did not include comparator data (e.g. missing Before data for a fish outcome), (ii) did not report flow-magnitude data (TS or S_TREND studies), (iii) included fish outcome and flow magnitude but for ≤1 year (TS studies), or (iv) only reported presence/absence or qualitative descriptions of fish outcomes.

Data preparation for quantitative synthesis

For formal meta-analysis using standardized effect sizes, measures of variability were converted to standard deviations when not reported as such (e.g. standard error or confidence intervals) using RevMan Calculator (Drahota & Bellor, 2008). If no variance was reported for group averages, standard deviations were obtained using mean value imputation (see Supporting Information 5). Variance imputation was not required for other quantitative synthesis approaches because variances are not used in those analytical procedures.

In preparation for the flow-ecological relationship analysis, the direction of the linear relationships between flow magnitude and ecological outcomes were determined for TS and S_TRENDS studies based on author-reported information or calculated from correlation coefficients (Pearson's *r*) by the review team when not provided. Comparisons were made among tables and figures if dataset extent matched or overlapped sufficiently for correlation analysis.

To reduce multiple effect size estimates from the same study (all quantitative synthesis approaches) and avoid giving studies with multiple estimates more weight in analyses (formal meta-analysis only), datasets were aggregated in a few instances when studies shared all other meta-data (see Supporting Information 5 for full details). Given one of our objectives was to determine whether generalized fish-flow relationships could be identified from the available literature [which would include null hypothesis testing regarding heterogeneity parameters (e.g. *Q* test to determine whether individual effect sizes estimate a common population mean)], and our small database of studies for the formal meta-analysis in the within-year and interannual replication type subsets, we were limited in our ability to use other approaches such as robust variance estimation or multi-level meta-analysis (Hedges et al., 2010; Tanner-Smith et al., 2016; Tanner-Smith & Tipton, 2014; Tipton, 2013; Van den Noortgate et al., 2013).

Effect size calculation

Standardized mean difference: Because outcomes (i.e. abundance, density, CPUE, or biomass, yield) were not always reported in comparable units or on the same scale, we used the standardized mean difference, Hedges' *g*, as our effect size measure for studies with sufficient replication (e.g. BA and BD studies with ≥2 replicates before and during/after), rather than raw mean differences. Hedges' *g* was calculated using the steps in Borenstein et al. (2009), as shown below.

Starting with Cohen's *d* to account for differences in measurements across studies (Cohen, 1977), we calculated the standardized mean difference by dividing the mean difference in each study (i.e. the difference between mean fish response to an intervention and the mean fish response to a lack of an intervention [the comparator]) by the study's pooled standard deviation:

$$d = \frac{\overline{X}_{G2} - \overline{X}_{G1}}{S_{\text{pooled}}}, \quad (1)$$

where $\overline{X}_{G1}$ was the mean of the group 1 (G_1 = the comparator group) and $\overline{X}_{G2}$ was the mean of group 2 (G_2 = the intervention group). S_{pooled} was the pooled standard deviation of groups 1 and 2:

$$S_{\text{pooled}} = \sqrt{\frac{(n_{G2}-1)S_{G2}^2 + (n_{G1}-1)S_{G1}^1}{n_{G1}+n_{G2}-2}}, \quad (2)$$

where S = standard deviation, and n is the sample size. The variance for d is given by:

$$V_d = \frac{n_{G1}+n_{G2}}{n_{G1}n_{G2}} + \frac{d^2}{2(n_{G1}+n_{G2})}. \quad (3)$$

Then, to convert Cohen's *d* to Hedges' *g*, we used a correction factor that decreases small sample bias in *d*:

$$J = 1 - \frac{3}{4(n_{G1}+n_{G2}-2)-1}. \quad (4)$$

Finally, we calculated Hedges' *g* and the associated variance (V_g) as:

$$Hedges'g = J \times d, \quad (5)$$

$$V_g = J^2 \times V_d. \quad (6)$$

From this, a negative Hedges' *g* indicates that fish outcomes (abundance or biomass) are lower after the intervention than before. All effect sizes calculations were done in MS Excel.

Percent change: Before/After and Before/During designs: For any data set with quantitative outcome data, regardless of whether there was replication or not [either a mean (number of replicates >1) or total count (n = 1) for either the intervention or comparator groups], we calculated the percent change in species response to an intervention:

$$\frac{(\overline{X}_{G2}) - (\overline{X}_{G1}+q)}{\overline{X}_{G1}+q} \times 100, \quad (7)$$

where $\overline{X}_{G1}$ and $\overline{X}_{G2}$ were the means (or total count if n = 1) of group 1 (G_1 = comparator group) and group 2 (G_2 = intervention group) for BA/BD studies. Since percent change cannot be computed when $\overline{X}_{G1} = 0$, we added a small constant q = 0.01 to $\overline{X}_{G1}$ for each data set. Prior to statistical procedures, we transformed percent change values by taking the log((x/100) + 1), where x = percent change calculation from Equation 7 (this transformation was equivalent to the log response ratios used in Piniewski et al., 2017). However, we provide untransformed axis values for ease of interpretation when plotting results (i.e. percent change values from Equation 7). A negative value for percent change indicated that fish outcomes decreased in relation to flow magnitude change, while positive value indicated that fish outcomes increased in relation to a change in flow magnitude. Because log transformations cannot be computed when percent changes were ≤−100%, we truncated such cases to −99.999%. All percent change calculations were done in MS Excel.

Percent change: Time series and spatial trend designs: For time series and spatial trends, because it was not possible to determine a clearly defined intervention, we first determined the direction of the effect of natural variation on fish outcomes from either author reported information, or reviewer calculated Pearson's correlation coefficient (r) (i.e. linear relationship between the fish outcome and flow magnitude). Most studies presented linear relationships between gradients of flow magnitude and fish outcomes directly; however, not all studies did, leaving us to make comparisons among tables and figures to match the extent of flow magnitude data and fish outcomes. In such cases, we correlated the two variables across the entire range of measured values, assuming a linear relationship. We then used the extracted maximum and minimum reported fish outcomes as the values for each group (intervention and comparator) in the percent change equation (i.e. Equation 7 above). More specifically, if Pearson's r was positive, the maximum reported outcome was considered $\overline{X}_{G2}$ in Equation 7, while if Pearson's r was negative, the maximum outcome reported is used as $\overline{X}_{G1}$. Percent change in flow magnitude was then calculated separately using magnitude values corresponding to the maximum and minimum fish outcomes for each dataset (e.g. if the maximum fish outcome occurred in 2010, the associated flow magnitude was measured in 2010). Prior to analysis, we transformed the data by taking the log |x| for the absolute percent change values (where x = percent change calculation from Equation 7) for both the fish outcomes and flow magnitude alterations, and then retaining the sign from the original percent change value, as was done in McManamay et al. (2013). All percent change calculations were done in MS Excel.

Quantitative synthesis

Formal meta-analysis: All formal meta-analyses (i.e. random effects and mixed-effects models) were conducted in R 4.1.1 (R Core Team, 2021) using the *rma.mv* function in the METAFOR package (3.0-2)

(Viechtbauer, 2010). To determine if changes in flow magnitude had a effect on fish abundance and biomass outcome metrics on average (research question 1), fish responses were compared to controls by conducting random-effects meta-analyses using restricted maximum-likelihood (REML) to compute weighted summary effect sizes for each outcome (i.e. abundance and biomass) within a given replication type (i.e. within-year and interannual temporal replication for BA/BD study designs) separately.

For within-year comparisons, models were developed for each of the first 5 years after a change in flow magnitude (i.e. comparing the most recent or only Before year with (i) After year-1 only, (ii) After year-2 only, (iii) After year-3 only, (iv) After year-4 only, and (v) After year-5 only, as well as the average of years 1–5 after a change in flow magnitude [see Supporting Information 5 for further details]). The first 5 years after a change in magnitude were selected since there were insufficient sample sizes in the available evidence base beyond this time frame. There were sufficient sample sizes to investigate for this potential impact of a time-lag for fish abundance but not for biomass outcomes.

To further account for multiple non-independent data sets with the same study ID (e.g. different species), Study ID was included as a random factor in each model. The summary effect size was considered significantly different from zero when the 95% confidence interval (CI) did not overlap with zero. Heterogeneity in effect sizes was calculated using the Q statistic, compared to the chi-square (χ^2) distribution to determine if the total variation in observed effect sizes (Q_T) was more heterogeneous than expected due to sampling error alone (Q_E) (e.g. Q_T is significantly greater than expected from Q_E) (Rosenberg, 2013). A statistically significant Q indicates greater heterogeneity in effect sizes (i.e. individual effect sizes do not estimate a common population mean), which suggests there are differences among effect sizes that arise from causes other than sampling error. We produced forest plots to visualize mean effect sizes and 95% CI from each comparison using the *forest* function of the METAFOR package (3.0-2) (Viechtbauer, 2010). Summary effect sizes were used to identify general trends in the evidence base and the impact of the intervention. It is important to note that a lack of significance does not indicate no significant patterns within the evidence base. Furthermore, a lack of significance can only be interpreted as a lack of evidence for an effect if there is no indication of heterogeneity.

Although we attempted to reduce publication bias by including data from available grey literature, publication bias could still impact results if publishing is biased towards a particular type of result, such as statistically significant outcomes. Therefore, we examined publication bias for global analysis models (as described above) by using funnel plots and fail-safe numbers (see Supporting Information 5 for further details).

To test for associations between effect size and moderators (research question 2), we used mixed-effects models for categorical moderators, estimating heterogeneity using REML. We first evaluated the influence of intervention type on each outcome subgroup separately. Then, we tested for associations between other moderators and effect sizes within intervention type subsets. We tested for associations within intervention subsets for two reasons. First, some moderators of interest were related to specific intervention types (e.g. direction of flow magnitude changes, duration of the event). To reduce potential confounding effects of intervention type, associations between other moderators and effect sizes were evaluated separately for different interventions. Second, information on all moderators was not always provided in articles, and partly due to this, the distribution of moderators varied substantially between intervention types; therefore, to reduce the number of studies that would have needed to be excluded if testing for associations with all interventions combined, we instead test for associations within intervention type subsets.

We only performed analyses for moderators when there were sufficient combinable datasets (i.e. ≥3 datasets from at least 2 studies) for each moderator category (e.g. at least 3 datasets from at least 2 studies for each of the major habitat types within the subset of drought studies reporting an abundance outcome). When there were insufficient numbers of datasets for a particular moderator category, these datasets were deleted, reducing the number of categories available for investigation of that moderator. The only exception to this rule was combining datasets for 'Not reported' and 'Unclear' categories when both were present for a particular moderator. Furthermore, for two moderators, we categorized the original continuous variable into discrete time periods due to uneven distributions, selecting categories in an effort to maximize the number of datasets in each (i.e. event duration [<6 months, ≥6 months but <1 year, ≥1 year]; monitoring duration [<5 years, ≥6 years]).

When testing for associations between effect size and moderators within intervention subsets, we used two approaches depending on the size of the evidence base within each subset. First, because studies did not always report all moderators of interest, it was not possible to combine all moderators into a single model simultaneously, especially when sample sizes did not allow for this (e.g. within-year abundance for all intervention type subsets, interannual abundance for drought subset [i.e. number of effect sizes (k) < 40 in all subsets]). Therefore, in such cases, we first conducted random-effects models (unmoderated models) using subsets of responses (e.g. a subset of abundance or biomass effect sizes for a given replication type) and intervention types (e.g. droughts, floods, low flows, high flows) that maximized the number of effect sizes that could be used to test the influence of the moderator of interest. We then used these subsets in mixed-effects models, including the moderator of interest. To further account for multiple study comparisons within a study site, and species outcomes being reported for the same site, in all models, Study ID was included as a random variable. We restricted the number of fitted parameters (j) in any mixed model such that k/j was greater than five to ensure reasonable model stability and sufficient precision of coefficients (Vittinghoff et al., 2005). This limited the number of moderators and categories that could be included in a single model. Given that all moderators were highly correlated (see results of Pearson's χ^2 test of moderators; Supporting Information 6), it was not possible to add more than one moderator into a given model, nor would sample size allow for this.

The second approach to test for associations between effect size and moderators within intervention subsets was used when the evidence base was relatively large (i.e. in the case of interannual abundance for the flood subset with $k = 127$ effect sizes). Here, we first conducted χ^2 tests to assess independence of moderators. When moderators within the intervention subset were confounded, and/or the distribution between moderator categories was uneven, we avoided these problems by constructing independent subsets of data in a hierarchical approach. When there was sufficient sample size within the intervention subset to include a moderator, we included the moderator into the model individually, and in combination when possible, following the same rule as outlined above (i.e. $k/j > 5$).

For all moderator analyses, total heterogeneity (Q_T) was partitioned into the heterogeneity explained by the model (Q_M) and heterogeneity not explained by the model (Q_E), error due to sampling; therefore, $Q_T = Q_M + Q_E$. The statistical significance of Q_M and Q_E were tested against a chi-square (χ^2) distribution.

Sensitivity analyses were carried out to investigate the influence of: (i) study validity categories; (ii) imputing missing variances (i.e. replacing missing data with calculated substitute values); (iii) inclusion of studies where the waterbodies may be influenced by fish stocking; (iv) inclusion of BD studies with true BA studies; (v) inclusion of articles that did not specify a flow magnitude component or reported unspecified multiple components of flow; and (vi) inclusion of yearly averages, averaged for the Before or After period (i.e. averages of averages). In all analyses, the results were compared to the overall model fit to examine differences in pooled effect sizes; see Supporting Information 5 for further details.

Less formal meta-analysis: All less formal meta-analyses were conducted in R 4.1.2 (R Core Team, 2021) using the *t.test*, *anova_test*, and *pairwise_t_test* functions in the RSTATIX package (0.7.0; Kassambara, 2021) and followed closely to those performed by Piniewski et al. (2017). Using the more inclusive evidence base, we determined if natural changes in flow magnitude had a significant effect on fish outcomes by conducting one-sample *t*-tests, where fish responses were compared to controls to compute unweighted summary percent change measures for each outcome (i.e. abundance and biomass) within a given replication type (i.e. within-year [After year-1 only] and interannual temporal replication for BA/BD study designs) separately. Testing here, the null hypothesis assumes that the mean percent change is equal to zero.

To test for associations between fish responses (log-transformed percent changes) and moderators, we used one-way analysis of variance (one-way ANOVA). Similar to the formal analysis above, we first evaluated the influence of intervention type on each outcome subgroup separately. Then, we tested for associations between other moderators (each separately) and percent changes within intervention type subsets. Here, we were testing the null hypothesis that categories in specific moderators are drawn from populations with the same mean values (e.g. comparing within-year abundance percent changes between drought, flood, and low flow subsets).

To then test whether the mean value of two categories within a particular moderator differed (e.g. between droughts and floods within the intervention moderator), we also conducted independent-samples *t*-tests. Here, we were testing the null hypothesis that there was no difference between the mean of the two categories.

Similar to the formal analysis above, we only performed analyses for moderators when there were sufficient combinable datasets (i.e. ≥3 datasets) for each moderator category; deleting datasets that did not meet this sample size criteria. However, datasets did not have to come from at least two different studies (following methods used in Piniewski et al., 2017). Furthermore, unlike the formal analysis, we did not further account for multiple study comparisons within a study, nor did we use a weighted analysis. Event duration was also categorized into discrete time periods (i.e. <6 months, ≥6 months but <1 year, ≥1 year); otherwise, no further categorizations were made to all other time-related moderators (i.e. monitoring duration, number of extreme events, and time since intervention).

Flow-ecological relationship analysis: All flow-ecological relationship analyses were conducted in R 4.1.2 (R Core Team, 2021) using the *ggplot* and *cor.test* functions in the ggplot2 and STATS (basic R) package (3.3.3 and 4.1.2, respectively) (R Core Team, 2021; Wickham, 2016) and followed closely those performed by McManamay et al. (2013). To visualize flow-ecological relationships, we first plotted the log-transformed percent changes for each outcome (i.e. abundance and biomass) against the associated log-transformed percent changes in flow magnitude within a given trend-type study design separately (i.e. time series and spatial trends). We then further explored relationships between fish outcomes and flow magnitude within each of the time series and spatial trend subgroups by plotting separate responses by categories within potential moderators with sufficient samples (i.e. >5 datasets for a given moderator category) (e.g. stratified fish abundance responses to natural flow magnitude variability by plotting waterbody types separately [estuaries, rivers, streams, reservoirs]). Spearman's rank correlations were used to assess direction and potential statistical associations between the two variables for each individual relationship (using the log-transformed values).

3 | RESULTS

3.1 | Literature searches and screening

A total of 300 studies from 219 articles met our inclusion criteria and were subsequently included for narrative synthesis, with 193 studies from 145 articles included in quantitative synthesis (see Supporting Information 7, Figure S7.1, for flow diagram of inclusion/exclusion process results). Article publication dates ranged from 1959 to 2021, with the majority (46%) published in the last decade. Although grey literature made up 40 to 50 percent of all articles from 1960 to 1999, for more recent decades (2000–2019), the proportion of all articles made up of grey literature was less than 10% (Supporting Information 7,

Figure S7.2). A database of these studies with descriptive meta-data, coding and qualitative/quantitative data is available in Rytwinski et al. (2022b).

3.2 | Study validity assessment

Validity assessments of the 300 studies resulted in 347 individual projects (i.e. individual investigations within a study that differ with respect to ≥1 aspect of the study validity criteria) (see Rytwinski et al., 2022c for assessment results). Most projects were assigned an overall 'Low' study validity (303 projects; 87%), with the remaining assigned an overall 'Medium' study validity (44 projects; 13%). No project was assigned an overall 'High' study validity. Study validity did not appear to change over time (see Supporting Information 7, Figure S7.3). Among the 303 projects that received an overall 'Low' validity score, most did not include a true temporal or spatial comparator (81% of projects), lacked replication (90%), and/or lacked clear information to judge sampling methodology (29%). Among the 44 projects with 'Medium' validity, all were BA studies with fish outcome data available for at least two experimental or observation units in the intervention period (i.e. ≥2 months/seasons during or post-intervention for within-year replication, or ≥2 years during or post-intervention for interannual replication), but either lacked consistent sampling in space and time (34%) and/or failed to provide quantitative data on flow magnitude changes (27%) (see Supporting Information 7, Table S7.1).

3.3 | Narrative synthesis

Narrative synthesis was based on all 300 studies from 219 articles, regardless of study validity, that considered abundance (276 studies) and biomass (73 studies).

3.3.1 | Study location

Studies occurred in 23 countries, with most studies conducted in the United States (65%), Australia (12%), and Canada (5%) (Figure 2; see also Supporting Information 7, Figure S7.4). All other countries had fewer than 10 studies each. A total of eight major habitat types were considered by studies in our database (i.e. groupings of ecoregions with similar biological, chemical, and physical characteristics), while 60 different freshwater ecoregions were represented. The most common major habitat was temperate coastal rivers (137 studies; 22 ecoregions), followed by temperate floodplain rivers and wetlands (78 studies; 13 ecoregions), temperate upland rivers (40 studies; 12 ecoregions), and xeric freshwaters and endorheic (closed) basins (30 studies; 7 ecoregions). All other major habitat types were represented by fewer than 12 studies each. Studies were conducted in a variety of waterbody types, most of which were in river systems (114 studies), followed by streams (91 studies) and estuaries (55 studies) (see also Supporting Information 7 for further descriptions of study locations).

3.3.2 | Population

Most studies (75%; 259/300) conducted species-specific investigations (i.e. provided data for individual species rather than data grouped/pooled over broader categories of species, genus or family) and among these, 37% (97/259) considered only one species. A total of 124 families were investigated by studies considering the impact of natural flow magnitude changes on specific species, representing 407 genera and 729 species (see Supporting Information 7, Figure S7.6 for the top 10 families and their top studied genera). The most studied species were *Salmo trutta* (36 studies), *Cyprinus carpio* (27), *Morone americana* (16), *Rutilus rutilus* (13), and *Squalius cephalus* and *Perca fluviatilis* (12 studies each). A total of 29 studies reported data that were grouped across family, genus or species; these included 118 additional species not reported in species-specific results (see Supporting Information 8 for a full species list).

3.3.3 | Intervention

The most common interventions captured in our systematic review were variation in flow magnitude (173 cases), followed by droughts (55 cases) and floods (48 cases). There was no clear change in the proportion of studies considering any one type of flow intervention over time (see Supporting Information 7, Figure S7.7). Studies investigating droughts considered mostly decreases in average discharge and decreases in lowest flow magnitudes for the system (i.e. base flow), while for floods most often increases in highest flow magnitudes (Table 4). Interventions considering high and low flows were mostly changes to average discharge (increases or decreases, respectively). Variation was considered for temporal or spatial trend studies and therefore direction of flow magnitude change was unknown and only fluctuations of flow magnitude were considered (162 cases), or for differences in flow magnitude not associated with a particular intervention (i.e. comparing one river's normal flow levels to another, but without a drought or flood; 11 cases).

3.3.4 | Study design

Most studies used a time series (TS) design (41% of cases), followed by true BA study designs (25%), and spatial trends (S_TRENDS; 12%) (Supporting Information 7, Figure S7.8). There were no BACI, true CI, RCT, RCA, or NRange study designs.

Monitoring duration was only summarized for true BA-type study designs and was considered as the number of years post-intervention for which fish outcomes were reported. Most true BA studies had monitoring durations of ≤5 years (63/83 cases; 76%). Only five cases had monitoring durations of more than 10 years and the maximum length of monitoring post-intervention was 12 years. Furthermore, for true BA studies most reported before data for ≤4 years prior to the intervention (66%), and only five cases had before data for more than 10 years.

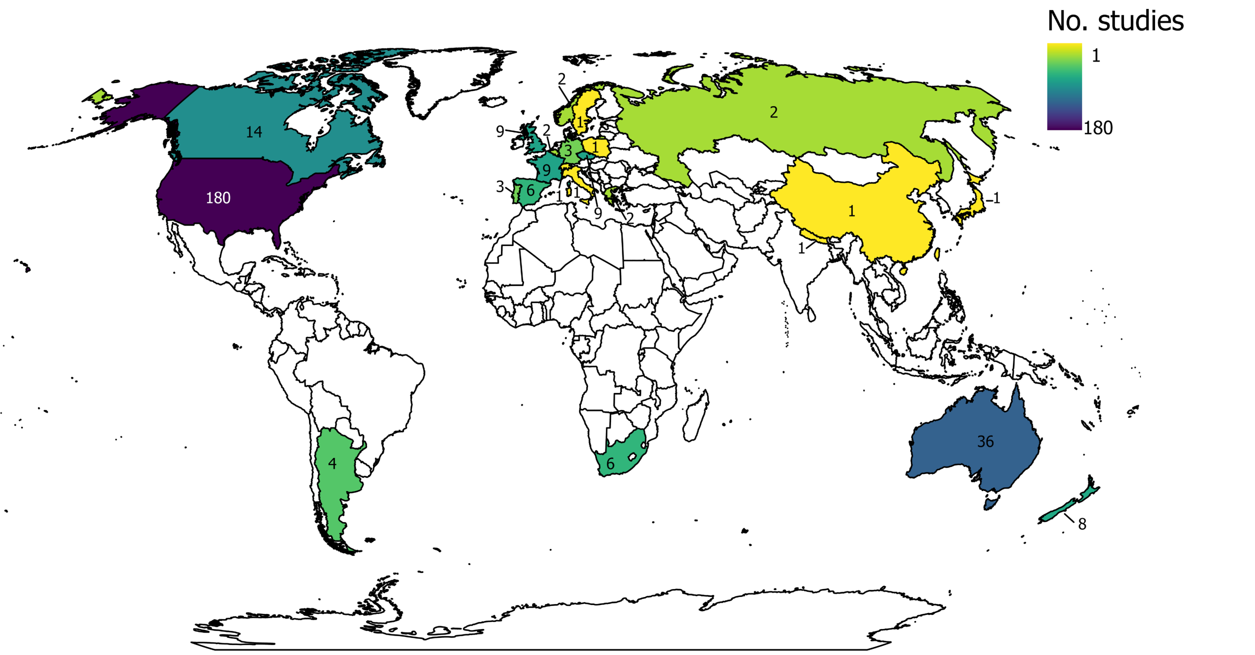

FIGURE 2 Number of studies considering fish abundance and/or biomass metrics per country. Studies undertaken across more than one country are counted within each study country.

3.3.5 | Outcomes

Studies often recorded more than one fish response (i.e. both abundance and biomass in the same study); however, most cases considered abundance (391/474 cases; 82%) (see Supporting Information 7, Figure S7.9). The most reported life stage for studies considering abundance was juveniles (30 cases), and for biomass was adult fish (6 cases); however, many studies did not report fish life stage (41% and 49% of cases, respectively). Fish were sampled with a variety of sampling methods. Most cases for both abundance and biomass used gear-based techniques (e.g. electrofishing, gill-, fyke-, seine-netting, trapping) not including mark-recapture techniques (abundance: 210/283 cases; biomass: 53/73 cases); see Supporting Information 7 for further descriptions of sampling methods.

3.4 | Quantitative synthesis

Of the 300 studies (from 219 articles) included in the narrative synthesis, 193 unique studies (from 145 unique articles) with 1853 datasets, after aggregation, were included in our quantitative synthesis database (see Rytwinski et al., 2022d). Of these, 39 studies were used in formal meta-analysis to analyse within-year (19 studies with 181 datasets) and interannual (31 studies with 225 datasets) variation in fish abundance and biomass responses, 68 studies were used in the less formal meta-analysis to analyse within-year (32 studies with 434 datasets) and interannual (55 studies with 281 datasets) variation in fish abundance and biomass responses, while 126 studies were used in the flow-ecological relationship analysis to investigate time series (102 studies with 438 datasets) and spatial trends (24 studies with 294 datasets) designs with fish abundance and biomass responses (Figure 1).

3.4.1 | Formal meta-analysis

Global meta-analyses

Refer to Supporting Information 9 for all forest (i.e. summary plot of all effect size estimates), funnel (i.e. plot showing potential publication bias), and Cook's distance plots (i.e. plot indicating influence of effect sizes), as well as sensitivity analysis results for all global analyses.

Within-year variation: Abundance: Natural changes in flow magnitude had a non-significant overall effect on fish abundance in post-intervention year-1 (Hedges' $g = -0.13$, 95% CI −0.31, 0.05; Figures 3 and 4a). Most effect sizes were negative (i.e. $g < 0$; 93 of 163), with the remainder showing neutral or positive responses (i.e. $g \geq 0$) to changes in flow magnitude. Only 12 of the 163 effect sizes were statistically significant (i.e. had confidence intervals that did not overlap zero [8 negative and 4 positive effect sizes]; Supporting Information 9, Figure S9.1). The Q test for heterogeneity suggested significant heterogeneity between effect sizes (Q = 198.30, p = 0.028) that could be explored using mixed effects models (see section 'Effects of Moderators—Within-year variation: Abundance' below). There was no obvious indication of publication bias from the funnel plot (Supporting Information 9, Figure S9.2). However, the failsafe number was zero, suggesting the results from the random effects model may not be robust against potential publication bias. Results of all sensitivity analyses were comparable to the overall meta-analysis Supporting Information 9, Table S9.1).

To investigate the potential impact of a time-lag in within-year fish responses to changes in magnitude, we compared the effect sizes for subsequent years of sampling post-intervention to that of the overall mean weighted effect size for year-1 datasets. In all cases, there was no evidence of an overall effect of natural changes in magnitude on fish abundance; however, the trend in the overall mean weighted

TABLE 4 Frequency of alterations to flow magnitude and type of intervention. Decreases and increases in flow magnitude could occur for the average (Discharge), highest flows (High), lowest flows (Low) and winter flows experienced in a system. For direction of change, Difference: comparisons between two levels of flow magnitude; Mixed = two directions of flow magnitude alteration occurred for an intervention; N/A: no specific direction of change as flow magnitude fluctuated. For intervention types, combined: comparison between two interventions (i.e. drought; flood); other: other types of intervention (i.e. tropical storms); variation: fluctuation in flow magnitude with no clearly defined natural event.

Alteration to flow magnitude		Intervention							
Direction of change	Component of flow	Drought	Flood	High flow	Low flow	Combined	Other	Variation	Unclear
Decrease	Discharge	24			2				
	High	2							
	Low	24							
Increase	Discharge		13	5			1		
	High		34	1			2		
	Low		2	1					
	Winter		3						
Difference		3				13		11	
Mixed		3							
N/A		1					2	162	
Unclear									1

effect sizes became increasingly positive with post-intervention years, except for post-intervention year-4 where a small decrease in average fish abundance in response to changes in flow magnitude was observed (Figure 5). Of the 133 species present in year-1, 22 species were present in all five post-intervention years. When all post-intervention years (1–5) were aggregated, the resulting overall mean weighted effect size was similar to that of year-1 alone (Figure 5).

Within-year variation: Biomass: Natural changes in flow magnitude had a non-significant, overall effect on fish biomass in post-intervention year-1 (Figures 4a and 6). Similar to abundance outcomes, most effect sizes were negative (i.e. $g < 0$; 11 of 18); however, none of the 18 effects were statistically significant. The Q test for heterogeneity did not suggest significant heterogeneity between effect sizes ($Q = 18.10$, $p = 0.383$). There was no obvious indication of publication bias from the funnel plot; however, the failsafe number suggested the results from the random effects model may not be robust against potential publication bias. Sensitivity analyses for biomass were not required as none of the identified potential methodological factors of concern were present in this subset (i.e. all datasets within this subset were medium validity, did not require variance imputation, there were no potential fish stocking cases, flow magnitude was specified, and used Before/After study designs).

Interannual variation: Abundance: The overall mean weighted effect size for abundance when considering interannual BA studies was not statistically significant (Figures 3 and 4a). There were nearly as many positive effect sizes (93) as negative (96) and most individual effect sizes were not statistically significant (176 of 189). The Q test for heterogeneity suggested that there was significant heterogeneity between effect sizes ($Q = 421.12$, $p < 0.0001$) that could be explored using mixed effects models (see section 'Effects of Moderators—Interannual variation: Abundance' below).

Interannual variation: Biomass: There was no detectable effect of natural changes in flow magnitude on overall fish biomass when considering interannual BA studies (Figures 4a and 6). Similar to abundance outcomes, there were a similar number of positive as negative effect sizes (18 each), and most individual effect sizes were not statistically significant (31 of 36). The Q test for heterogeneity suggested significant heterogeneity between effect sizes ($Q = 67.02$, $p = 0.0009$). The funnel plot of asymmetry suggested possible evidence of publication bias towards studies with larger sample sizes showing positive effects of flow magnitude change. Furthermore, the failsafe number suggested the results from the random effects model may not be robust against potential publication bias.

Effect of moderators

Within-year variation: Abundance: To address the question, to what extent does intervention type influence the impact of changes in flow magnitude due to natural causes, there were only sufficient sample sizes (i.e. ≥3 datasets from at least 2 studies) to include the following interventions for fish abundance: (1) Drought; (2) Flood; and (3) Low flow (Figure 3). There was a statistically significant effect of intervention type on average effect sizes detected ($Q_M = 17.13$, $p = 0.0002$, $k = 99$; Supporting Information 9, Table S9.4A), with droughts and floods associated with decreases in average fish abundance and low flows with increases in fish abundance after the intervention compared to before (Figures 3 and 7a). Furthermore, fish abundance responded differently to low flow than to both floods and droughts (i.e. 95% confidence intervals did not overlap among the low flow subset and other intervention types; Figure 7a). There was no significant heterogeneity remaining in the moderated model ($Q_E = 95.56$, $p = 0.493$; Supporting Information 9, Table S9.4A).

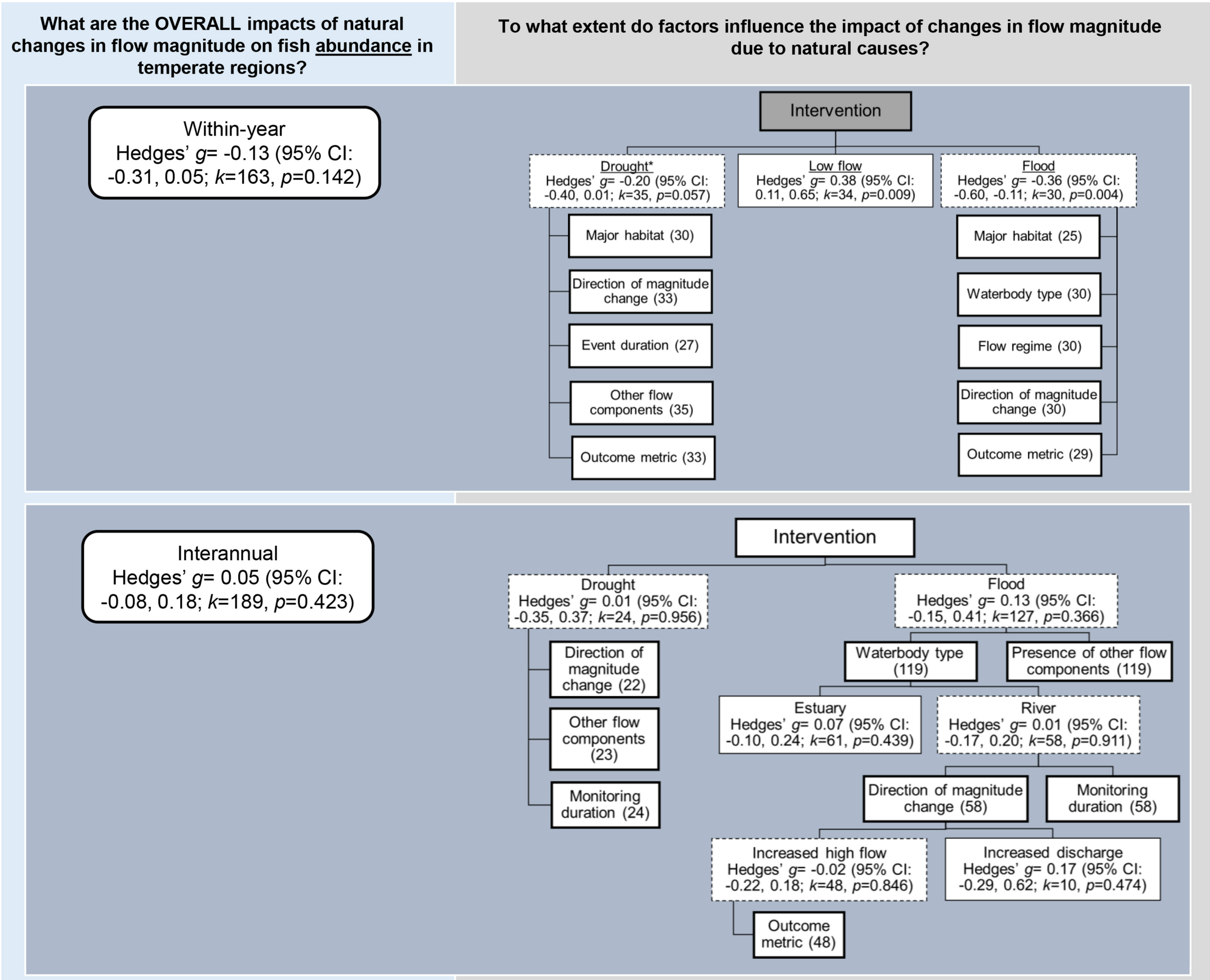

FIGURE 3 Summary flow chart of formal meta-analyses and results addressing our two main research questions related to fish abundance outcomes. Boxes indicate potential effect modifiers (thick solid lines), subset categories under consideration (dashed boxes), or subsets not under consideration (regular boxes). Greyed effect modifiers were associated with fish abundance responses. Underlined value indicates statistically significant effect at the $p < 0.05$ level, but with an asterisk (*) a moderately significant effect at the $p < 0.1$ level. k: number of data sets (i.e. effect sizes); Hedges' g: weighted mean effect size; CI: 95% confidence interval.

Drought: There were only sufficient sample sizes and variation to permit meaningful tests of the influence of the following moderators (and factor levels therein) within the drought subset: (1) Major habitat type (Temperate coastal rivers, Temperate floodplain rivers and wetlands); (2) Direction of magnitude change (Decrease_Discharge, Decrease_Low); (3) Event duration (<6 months, ≥1 year); (4) Presence of other flow components (Yes, No); and (5) Outcome metric (Abundance, Density). For all moderators considered, we found no detectable effect on average effect size from univariate mixed-effects models (Supporting Information 9, Table S9.4B). Additionally, most moderators were highly correlated (see results of Pearson chi-square test; Supporting Information 6, Table S6.1).

Flood: There were only sufficient sample sizes and variation to permit meaningful tests of the influence of the following moderators (and factor levels therein) within the flood subset: (1) Major habitat type (Temperate coastal rivers, Temperate floodplain rivers and wetlands); (2) Waterbody type (Estuary, River); (3) Flow regime (Free-flowing systems, Regulated systems, Not reported/Unclear); (4) Direction of magnitude change (Increase_Discharge, Increase_high); and (5) Outcome metric (Density, CPUE). For all moderators considered, we found no detectable effect on average effect size from univariate mixed-effects models (Supporting Information 9, Table S9.4C).

Low flow: There were insufficient sample sizes or variation to permit meaningful tests of the influence of potential effect modifiers within the low flow subset.

Within-year variation: Biomass: The influence of factors was not investigated owing to inadequate sample sizes.

Interannual variation: Abundance: To address the question, to what extent does intervention type influence the impact of changes in flow magnitude due to natural causes, there were only sufficient sample sizes to include the following interventions for fish

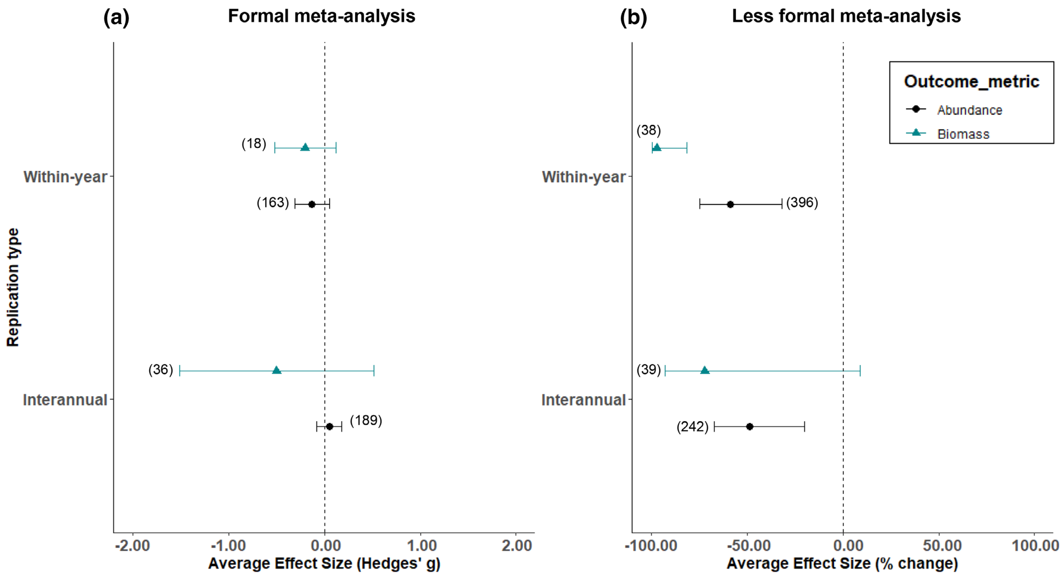

FIGURE 4 Average effect size by temporal replication type for fish outcomes using different quantitative synthesis approaches: (a) formal meta-analysis using standardized mean difference (Hedges' *g*) effect size measures, and (b) less formal meta-analysis using a percent change measure. Values in parentheses are the number of effect size estimates. Error bars indicate 95% confidence intervals. A positive mean value (to the right of the dashed zero line) indicates that the fish outcome (abundance or biomass) was higher after an intervention than before. 95% confidence intervals that do not overlap with the dashed line indicate a significant effect (at the $p < 0.05$ level).

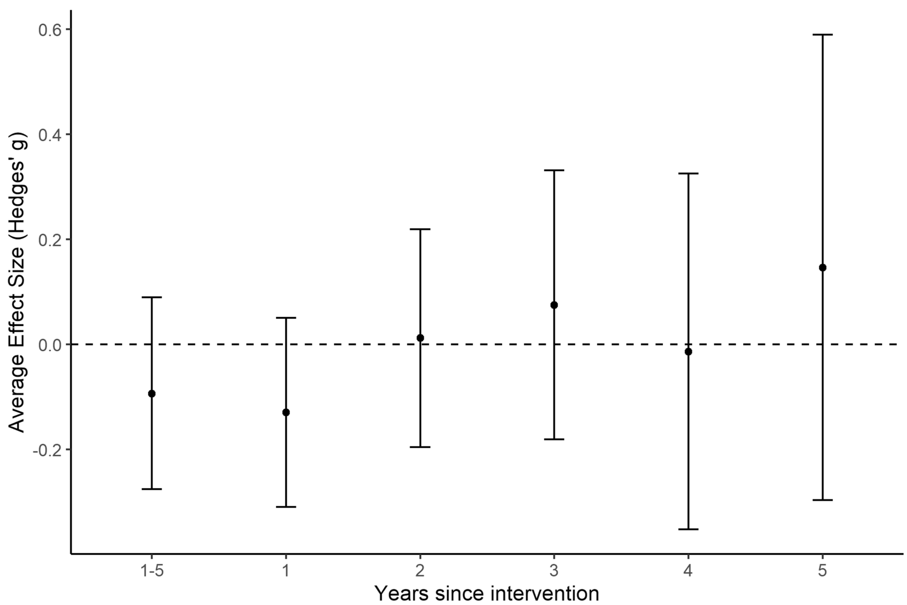

FIGURE 5 Comparison of overall weighted effect size for within-year *BA* abundance studies one ($k = 163$), two ($k = 47$), three ($k = 33$), four ($k = 27$), and five ($k = 22$) years post-intervention and when After years 1–5 were aggregated. Models were developed for each of the first five years after a change in flow magnitude (i.e. comparing the most recent or only Before year with After year-1 only, After year-2 only, After year-3 only, After year-4 only, and After year-5 only), as well as the average of years 1–5 after a change in flow magnitude. See Figure 4 for explanations.

abundance: (1) Drought; and (2) Flood (Figure 3). There was no detectable effect of intervention type on average effect sizes ($Q_M = 0.27$, $p = 0.606$, $k = 151$; Figure 3 and 7c), and there was still significant heterogeneity remaining in the moderated model ($Q_E = 214.87$, $p = 0.0003$; Supporting Information 9, Table S9.5A).

Drought: There were only sufficient sample sizes and variation to permit meaningful tests of the influence of the following moderators (and factor levels therein) within the drought subset: (1) Direction of magnitude change (Decrease_Discharge, Decrease_Low); (2) Presence of other flow components (Yes, No); and (3) Monitoring duration

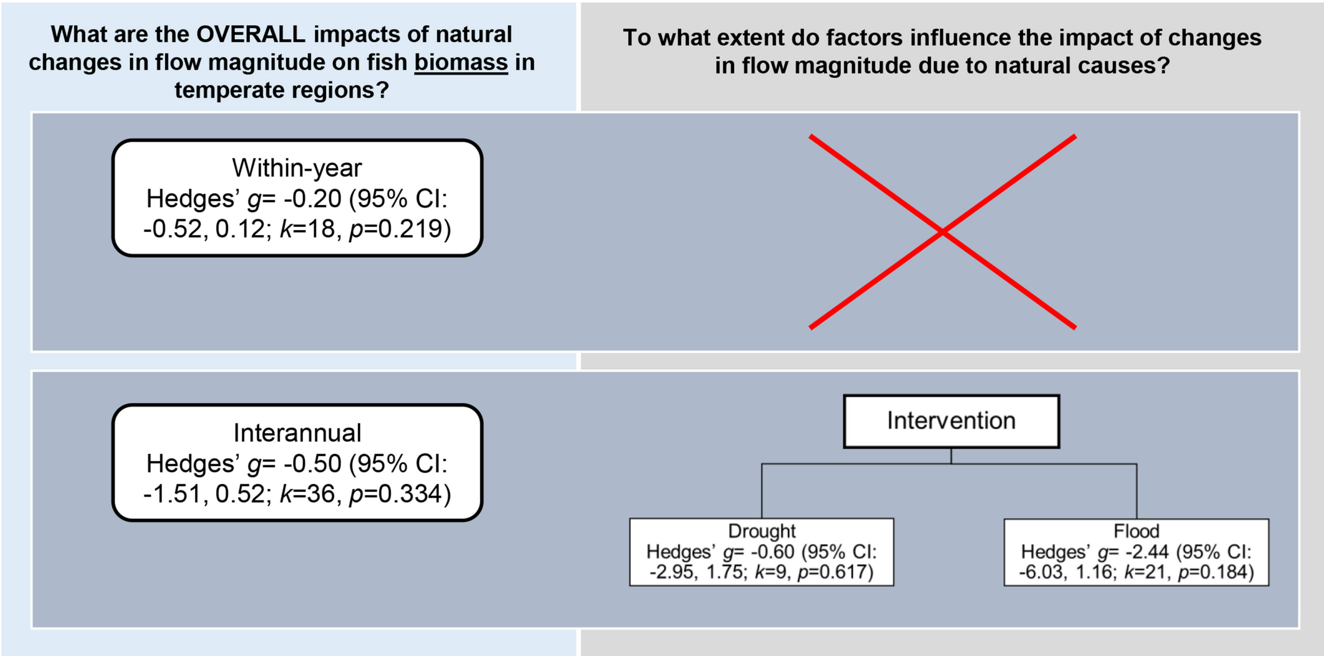

FIGURE 6 Summary flow chart of formal meta-analyses and results addressing our two main research questions related to fish biomass outcomes. See Figure 3 for explanations.

(<5 years, ≥6 years). For all moderators considered, we found no detectable effect on average effect size from univariate mixed-effects models, and there was still significant heterogeneity remaining in all moderated models (Supporting Information 9, Table S9.5B).

Flood: There were only sufficient sample sizes and variation to permit meaningful tests of the influence of the following moderators within the flood subset: (1) Major habitat type (Temperate coastal rivers, Temperate floodplain rivers and wetlands, Xeric freshwaters and endorheic basins); (2) Waterbody type (Estuary, River); (3) Direction of magnitude change (Increase_Dischange, Increase_High); (4) Presence of other flow components (Yes, No); (5) Monitoring duration (<5 years, ≥6 years); and (6) Outcome metric (Abundance, Density, CPUE). To avoid issues of confounding factors and/or uneven distributions between moderator categories, we constructed independent subsets of data in a hierarchical approach, given that the sample size for the flood subset allowed us to do so. Due to high correlations between major habitat type and other moderators (Supporting Information 6, Table S6.2), we removed major habitat type from further consideration. Furthermore, due to remaining high correlations and uneven distribution between: (1) direction of magnitude change and monitoring duration, the influence of these moderators were investigated within the subset of river studies only; (2) outcome metric and both direction of magnitude change and monitoring duration, outcome metric was investigated with the increased high flow subset. For all moderators (and combinations thereof) considered, we found no detectable effect on average effect size from mixed-effects models, and no significant heterogeneity remaining in any moderated models (Supporting Information 9, Table S9.5c–e).

Interannual variation: Biomass: To address the question, to what extent does intervention type influence the impact of changes in flow magnitude due to natural causes, there were only sufficient sample sizes to include the following interventions for fish biomass: (1) Drought; and (2) Flood (Figures 6 and 7c). There was no detectable effect of intervention type on average effect sizes ($Q_M = 0.70$, $p = 0.402$, $k = 30$) and there was still significant heterogeneity remaining in the moderated model ($Q_E = 64.83$, $p < 0.0001$). There were insufficient sample sizes and variation to permit investigations of other moderators within the drought and flood subsets.

Taxonomic analysis

We investigated impacts of natural changes in flow magnitude on fish abundance within separate intervention types for fish families with sufficient sample sizes. Note, sample sizes were too small for analysing biomass responses by taxa.

Within-year variation: Abundance: There were only sufficient sample sizes to investigate impacts of within-year changes to flow magnitude on abundance for three temperate freshwater fish families for drought studies (i.e. Centrarchidae, Leuciscidae, and Salmonidae) and one family for flood studies (Salmonidae). No family had statistically significant overall responses to these event types (Figure 8a). The families Centrarchidae and Leuciscidae had overall negative-trending responses to droughts, while Salmonidae had an overall negative-trending response to floods but a mean positive-trending response to droughts. Based on the Q test of heterogeneity, there was no significant heterogeneity among effect sizes for any of the families for a given intervention type, suggesting it was not meaningful to explore potential reasons for heterogeneity; in any case, sample sizes would not allow for this.

Interannual variation: Abundance: There were only sufficient sample sizes to investigate impacts of interannual changes to flow

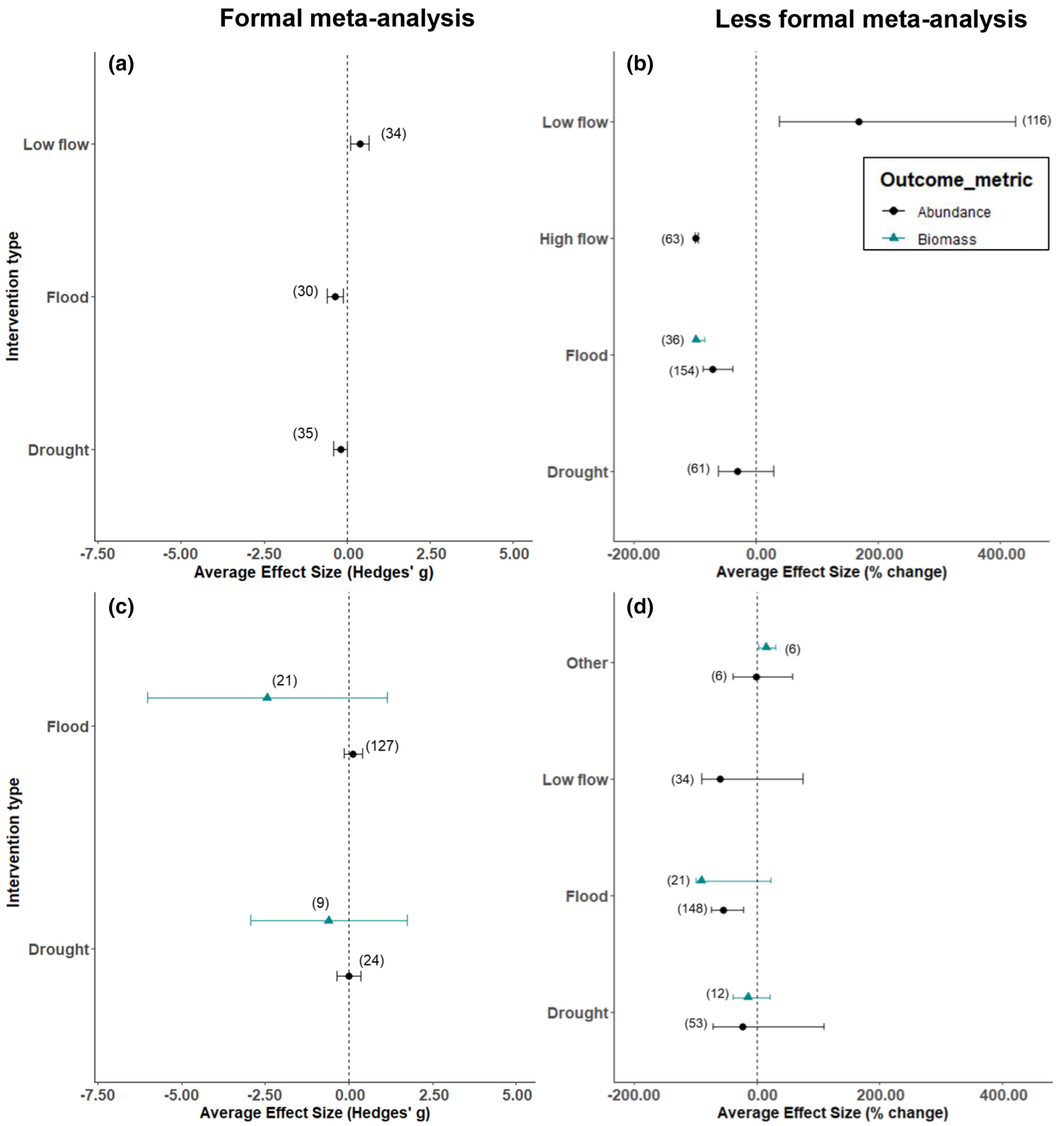

FIGURE 7 Average effect size by intervention type for fish outcomes using different quantitative synthesis approaches [i.e. formal meta-analysis using standardized mean difference effect size measures (a, c), and less formal meta-analysis using a percent change measure (b, d)] for within-year BA studies (top panels) and interannual BA studies (bottom panels). See Figure 4 for explanations.

magnitude on abundance for two temperate freshwater fish families for drought studies (i.e. Leuciscidae and Salmonidae) and ten families for flood studies (i.e. Carangidae, Catostomidae, Centrarchidae, Cyprinidae, Haemulidae, Ictaluridae, Leuciscidae, Mugilidae, Salmonidae, Sciaenidae, Sparidae). Droughts and floods had non-significant overall effects on the abundance of Leuciscidae and Salmonidae families (Figure 8b). For the other fish families for which there was only information to investigate the influence of floods on average effect sizes, overall abundance responses were mixed; however, there were no statistically significant effects of floods for any family (Figure 8b). Further, and similar to within-year changes to flow magnitude, there was no significant heterogeneity among effect sizes for any of the families for a given intervention type to suggest investigation of moderators was necessary.

3.4.2 | Less formal meta-analysis

Global meta-analyses

Less formal analyses based on percent changes suggested that natural changes in flow magnitude had negative, and statistically significant, overall effects on fish abundance and biomass in post-intervention year-1 for within-year temporal variation analyses (i.e. 59% and 97% decrease in fish outcomes after magnitude changes than before, respectively). Mean percent changes for fish abundance and biomass in interannual analyses were also negative (i.e. 49% and 72% decrease in fish outcomes after magnitude changes than before, respectively); however, the latter only marginally different than 0 ($p < 0.1$) (Figure 4b; Supporting Information 10, Table S10.1A).

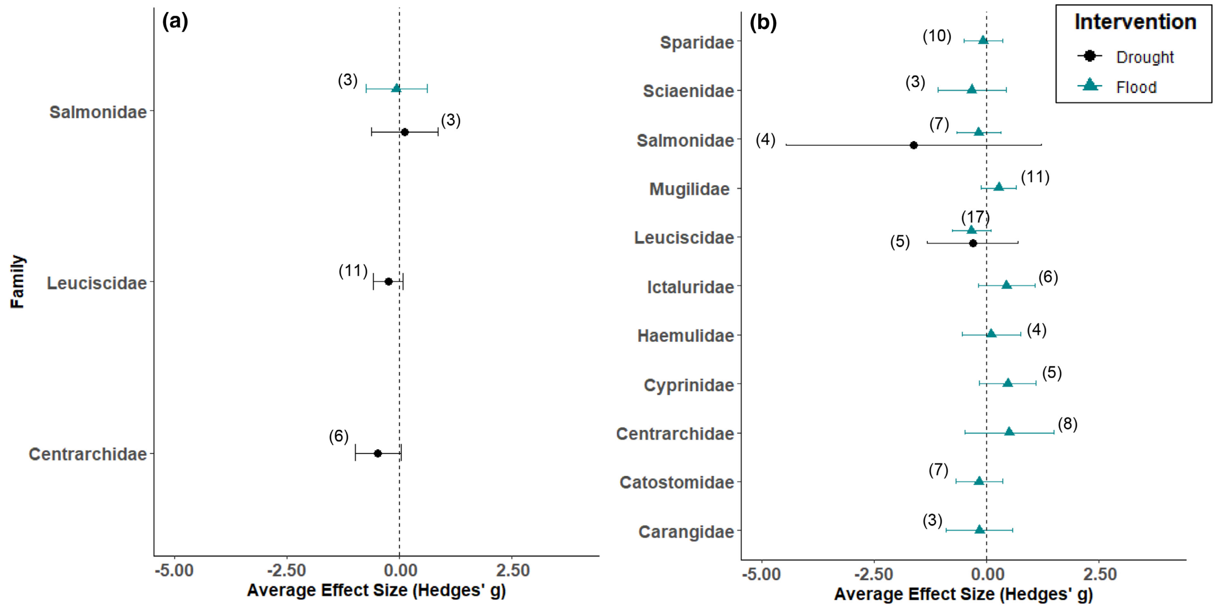

FIGURE 8 Average effect size using abundance outcomes by fish family for different intervention types (droughts and floods) and using different temporal analyses (a) within-year and (b) interannual. See Figure 4 for explanations.

Effect of moderators

Within-year variation: Abundance: To address the question, to what extent does intervention type influence the impact of changes in flow magnitude due to natural causes, there were sufficient sample sizes (i.e. ≥3 datasets) to include the following interventions for fish abundance: (1) Drought; (2) Flood; (3) Low flow; and (4) High flow (Figure 7b). There was a statistically significant effect of intervention type on average percent change detected (F = 20.26, $p < 0.0001$, k = 394; Supporting Information 10, Table S10.1B). The mean percent change for droughts was negative (30% decrease in abundance) but not significantly different from 0. However, floods and high flows were significantly associated with decreases in average fish abundance and low flows with an increase in fish abundance after the intervention compared to before (71%, and 99% decrease vs. 169% increase, respectively; Figure 7b). Furthermore, fish abundance responded differently between floods, high flows, and low flows (i.e. 95% confidence intervals did not overlap among these intervention types, Figure 7b, but see also Supporting Information 10, Figure S10.1 for independent sample t-test results) but not between floods and droughts.

Drought: There was a significant effect of waterbody type and a marginally significant effect of major habitat type on average percent change detected (F = 4.33, p = 0.018, k = 46, and F = 2.20, p = 0.098, k = 55, respectively; Supporting Information 10, Table S10.2A). Rivers and streams were associated with decreases in average fish abundance and estuaries with an increase in fish abundance after droughts compared to before (79%, and 18% decrease vs. 115% increase, respectively); however, only the mean percent change for rivers was significantly different from 0 (Supporting Information 10, Figure S10.2A). Temperate coastal rivers, tropical and subtropical coastal rivers, and xeric freshwaters and endorheic (closed) basins major habitat types were associated with decreases in average fish abundance and temperate floodplain rivers and wetlands with an increase in fish abundance after droughts compared to before (55%, 54%, and 86% decrease vs. 35% increase, respectively); however, only the mean percent change for xeric freshwaters was significantly different from 0, and only marginally ($p < 0.1$; Supporting Information 10, Figure S10.2B). Note, only temperate study locations were deemed relevant to this review; however, some waterbodies were classified as being from the tropical and subtropical coastal rivers major habitat type by the resource we used (i.e. Freshwater Ecoregions of the World, WWF/TNC, 2019). There was no detectable effect of any other moderator considered, including study validity, on average percent changes in fish abundance (Supporting Information 10, Table S10.2A).

Flood: There were significant effects on average percent change detected for the following moderators: (1) Study validity (at the $p < 0.1$ level); (2) Major habitat type; (3) Waterbody type; (4) Water regime; (5) Flow regime; (6) Land-use in the immediate surrounding; (7) Direction of flow change; and (8) Event duration (see Supporting Information 10, Table S10.2B). Both low and medium validity studies were associated with decreases in average fish abundance but for the former, a much stronger negative effect (87% and 51% decrease, respectively; Supporting Information 10, Figure S10.3A). Similar to responses to droughts, average percent change in abundance for temperate coastal rivers major habitat type was negative (84% decrease); however, estimated responses for temperate floodplain rivers and wetlands (38% decrease) and xeric freshwaters and endorheic (closed) basins major habitat types (3582% increase) were in opposite directions compared to droughts (Supporting Information 10, Figure S10.3B). Furthermore, average percent change in fish abundance differed between finer scale habitat attributes and event characteristics, such as: (i) estuaries and rivers (97% decrease vs. 40% increase; Supporting Information 10, Figure S10.3C); (ii) intermittent and perennial water regimes (254%

increase vs. 82% decrease; Figure S3D), (iii) free-flowing and regulated flow regime systems (170% increase vs. 93% decrease; Figure S3E), (iv) waterbodies surrounded by agricultural land and a mixture of agriculture and natural lands (98% vs. 64% decrease; Figure S3F), (v) increased mean/median discharge (Increase_Discharge) and increased seasonal or event high flow (Increase_High) (95% vs. 44% decrease; Figure S3G), and (vi) event durations of <6 months and >6 months but <1 year (43% increase vs. 98% decrease; Figure S3H). There were no detectable effects on average percent changes in fish abundance for other moderators considered (i.e. presence of other flow components, fish life stage, or outcome metric; Supporting Information 10, Table S10.2B).

Low flow: There was insufficient variation to permit meaningful tests of the influence of potential effect modifiers within the low flow subset.

High flow: There was insufficient variation to permit meaningful tests of the influence of potential effect modifiers within the high flow subset.

Within-year variation: Biomass: The influence of intervention type on biomass responses to natural changes in flow magnitude could not be tested as there was only sufficient sample size for one type, floods, which showed a significant 98% decrease in biomass following the event than before (Figure 7b).

Flood: There was insufficient variation to permit meaningful tests of the influence of potential effect modifiers within the flood subset.

Interannual variation: Abundance: To investigate whether the impact of changes in flow magnitude due to natural cause varied by intervention type, there were sufficient sample sizes to include the following interventions for fish abundance: (1) drought; (2) flood; (3) low flow; and (4) other (Figure 7d). There was no detectable effect of intervention type on average effect sizes ($F = 0.42$, $p = 0.74$, $k = 241$; Supporting Information 10, Table S10.1B and Figure S10.4).

Drought: No detectable effects were found on average effect sizes for any of the moderators considered (Supporting Information 10, Table S10.3A).

Flood: No detectable effects were found on average effect sizes for any of the moderators considered (Supporting Information 10, Table S10.3B).

Low flow: There was insufficient variation to permit meaningful tests of the influence of potential effect modifiers within the low flow subset.

Other: There were insufficient sample sizes to permit meaningful tests of the influence of potential effect modifiers within the other subset.

Interannual variation: Biomass: The influence of intervention type on biomass responses to natural changes in flow magnitude was investigated with the following interventions for fish biomass: (1) drought, (2) flood and (3) other (Figure 7d). There was no detectable effect of intervention type on average effect sizes ($F = 1.46$, $p = 0.25$, $k = 39$; Supporting Information 10, Table S10.1B and Figure S10.5).

Drought: There were insufficient sample sizes to permit meaningful tests of the influence of potential moderators within the drought subset.

Flood: There was insufficient variation to permit meaningful tests of the influence of potential moderators within the flood subset.

Other: There were insufficient sample sizes to permit meaningful tests of the influence of potential moderators within the other subset.

3.4.3 | Flow-ecological relationship analysis

Global meta-analyses

When plotting the percent changes of fish outcomes against the associated percent changes in flow magnitude from natural flow variation for all cases, in general, fish responses were variable in time series studies regardless of the measured outcome (i.e. abundance or biomass; Figure 9a). For spatial trend studies, fish abundance responses to natural flow magnitude changes were predominantly positive (i.e. increases in abundance to increases in flow magnitude); however, biomass responses were more variable (Figure 9b).

Effect of moderators

Time series: There were sufficient sample sizes to further explore individual relationships between percent change in fish outcomes and flow magnitude within categories of the following moderators for time series studies: (1) major habitat type; (2) waterbody type; (3) water regime; (4) flow regime; (5) land-use in immediate surrounding; (6) presence of other flow component alterations; and (7) outcome metric. Fish abundance responses showed a significant positive correlation with natural flow magnitude within rivers, and a negative correlation within estuaries (Figure 10a). Furthermore, fish biomass showed a significant positive correlation with natural flow magnitude when there were no other flow regime component alterations reported in the waterbody; however, fish biomass showed non-significant responses when there were other flow regime components alterations taking place at the same time as changes in magnitude (Figure 10b). There were no other significant patterns detected for other individual relationships investigated (see Supporting Information 11, Figure S11.1).

Spatial trends: There were sufficient sample sizes to further explore individual relationships between fish outcomes and flow magnitude within categories of the following moderators for spatial trend studies: (1) major habitat type, (2) waterbody type, (3) water regime (abundance only), (4) flow regime, (5) land-use in immediate surrounding, (6) presence of other flow component alterations and (7) outcome metric. Similar to time series studies, fish abundance responses showed a significant positive correlation with natural flow magnitude within rivers but showed no significant pattern within streams (Figure 11a). Furthermore, fish abundance responses showed significant positive correlations with natural flow magnitude: (i) within temperate floodplain rivers and wetlands major habitat type, (ii) waterbodies that were intermittent, (iii) waterbodies that included other flow regime component alterations in addition to changes in magnitude and (iv) outcomes related directly to abundance (Figure 11b–d, respectively). There were no other significant patterns detected for other individual relationships investigated (see Supporting Information 11, Figure S11.2).

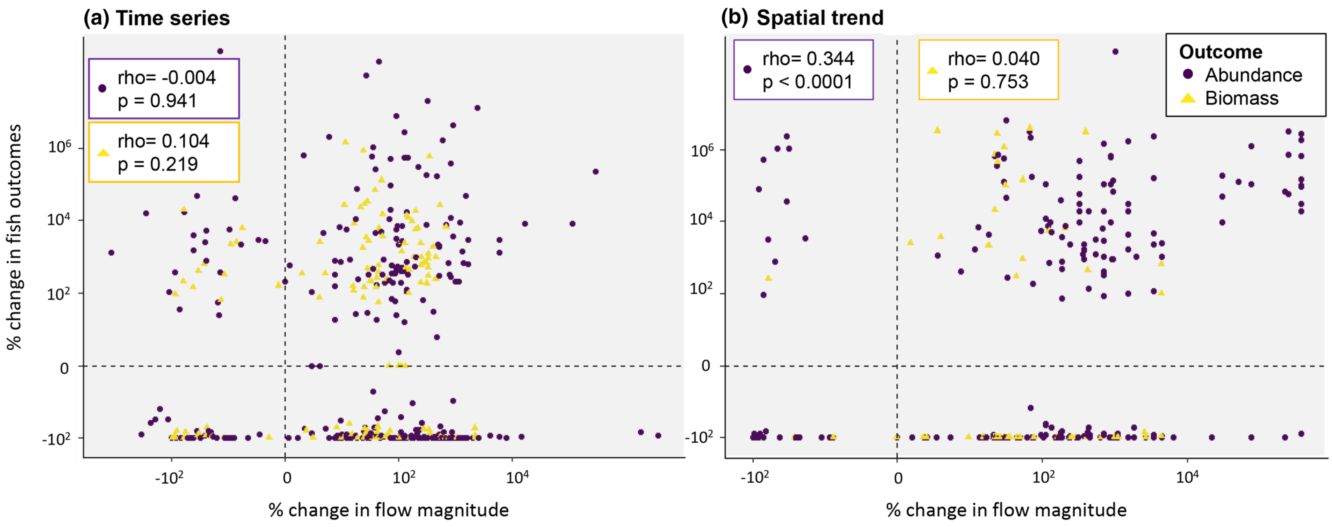

FIGURE 9 Relationships between percent changes in flow magnitude and percent changes in fish responses for natural flow variation studies according to different fish outcomes (abundance and biomass) within: (a) time series studies and (b) spatial trend studies. Rho represents Spearman's rank correlation coefficients; *p*-values for statistical significance ($p < 0.5$). Note, data points are plotted, and correlations obtained using the log-transformed percent changes for both variables; however, axes values are untransformed for ease of interpretation.

4 | DISCUSSION

We present what we believe to be the first comprehensive review that systematically and quantitatively evaluated the existing evidence base on the impacts of natural causes of variation or changes in flow magnitude on fish abundance and biomass in temperate regions. This synthesis of available evidence suggests that overall fish abundance and biomass responses to changes in natural flow magnitude were mainly negative but our analyses do not provide support for clear generalizable signals across all contexts. We found strong support that fish were responding (in terms of abundance) differently to natural events considered in our analyses, with consistently negative responses to floods and droughts within the first year of the natural change in flow magnitude. However, these patterns were less evident when considering interannual variation. We discuss these findings below by first comparing the methods of the three quantitative approaches used, as well as their results in relation to the overall impacts of natural changes and reasons for heterogeneity in fish responses. We also provide a detailed and reflective discussion of the limitations of review methods and the evidence base itself. We then conclude with some recommendations and points of consideration for management agencies and researchers.

4.1 | Comparison of quantitative synthesis approaches

In addressing our two research questions (i.e. how do natural changes in flow magnitude affect fish abundance and biomass in temperate regions, and to what extent do factors influence the potential impact of these changes), we attempted to make use of the entire evidence base by using three quantitative synthesis approaches that varied in their level of rigour and potential reliability. The first approach, the formal meta-analysis, made use of a more rigid framework with well-documented meta-analytical methods unused in previous syntheses on this topic. Though many studies were excluded from this formal approach because they lacked a comparator (i.e. no reference or baseline condition using before or control site information), and/or replication, this rigorous approach provides more robust estimates and is recommended for future reviews.

Some may argue that although these excluded studies were considered to be less reliable sources of information (i.e. susceptible to bias and/or had inadequate study designs), they could further contribute useful information to the comprehensive knowledge base on the subject, when accompanied with an appropriate consideration for study validity. Therefore, in our second meta-analytical approach (less formal meta-analysis), we performed a more basic synthesis, using a less robust effect size measure (i.e. percent change), that allowed for the inclusion of datasets that lacked replication (but still included a comparator). In doing so, the number of datasets included in quantitative synthesis increased from 406 using the more formal meta-analysis approach to 715 datasets in the less formal meta-analysis. Here, each effect size was treated equally in the analyses (unweighted), as this approach was more in line with statistical methods used by previous syntheses on this general topic (e.g. Poff & Zimmerman, 2010 [anthropogenic flow changes]; McManamay et al., 2013 [anthropogenic and natural]; Piniewski et al., 2017 [natural]), although we reduced multiple effect size estimates from the same study by aggregating dependent datasets prior to analysis.

Our third meta-analytical approach (flow-ecological relationship analysis) included studies that focused on fish outcomes of natural flow magnitude variation through time or across space Unlike studies used in the two

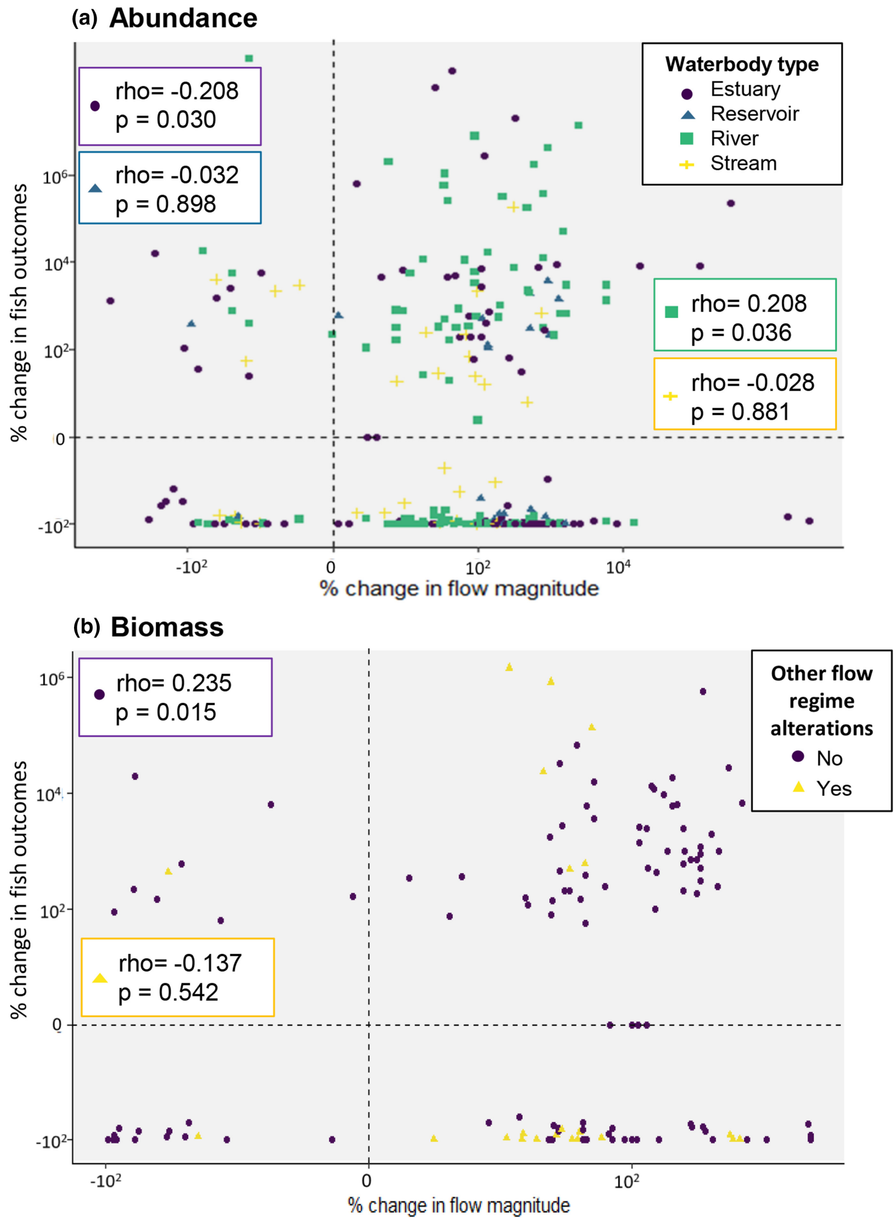

FIGURE 10 Relationships between percent changes in flow magnitude and percent changes in fish responses for time series studies according to (a) waterbody types for fish abundance responses and (b) presence or absence of other flow component alterations for fish biomass responses. See Figure 9 for explanations.

previous approaches, there was no clearly defined event such as a flood or drought for which a clear comparator could be identified. Furthermore, because some of these studies were not investigating relationships between gradients of flow magnitude and fish outcomes directly, it made extracting and calculating effect sizes challenging. Therefore, here too, in an attempt to make use of this evidence and provide results informative to generalizing flow-ecological relationships, we calculated percent changes from the reported maximum and minimum outcome variables, with associated flow magnitude values to explore quantitative flow-ecological relationships (i.e. assuming here the maximum fish outcome values represented more optimal conditions at a given time or site, to minimum outcome values considered less optimal). This analysis allowed us to make use of an additional 732 datasets; however, these sources were considered of lower validity for addressing our review questions given that they lacked a true comparator (and replication), and as such, we apply caution when interpreting results below. This approach of using different quantitative synthesis methods allows readers to make informed decisions based on their particular needs, context and application.

4.1.1 | How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions?

In general, the less formal meta-analyses supported formal meta-analytical results when comparing summary impacts of natural

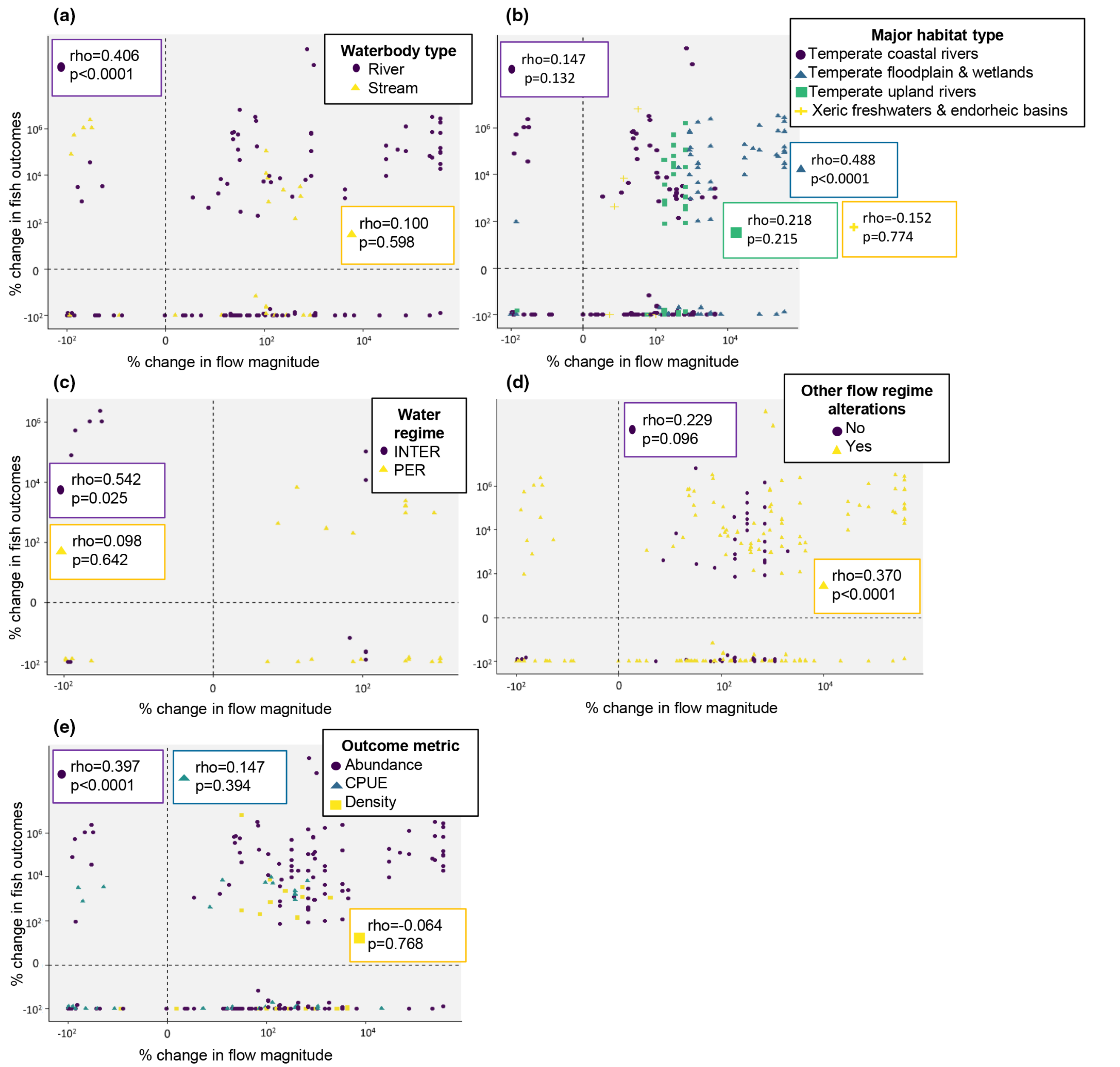

FIGURE 11 Relationships between percent changes in flow magnitude and percent changes in fish abundance responses for spatial trend studies according to (a) waterbody types and (b) major habitat types, (c) water regimes (INT: intermittent; PER: perennial), (d) presence or absence of other flow component alterations, and (e) outcome metrics. See Figure 9 for explanations.

changes in flow magnitude on fish abundance and biomass within a given temporal replication type in that the estimated responses were predominantly negative (Figure 4). However, average fish responses from formal analyses were all non-significant (i.e. their 95% confidence intervals overlapped with the line of no effect), whereas average percent change responses were all significant (but for one case, only marginally significant at the $p < 0.1$ level). To further investigate (post-hoc) whether this difference in results between the two approaches was due to the quantitative approach or the inclusion of additional studies that lacked replication in the less formal meta-analyses, we ran the less formal analyses using the same datasets included in the formal meta-analysis. Summary effect sizes and significance levels were found to be similar in less formal analyses when using all possible datasets compared to the same subset of effect sizes used in the formal meta-analyses (see Supporting Information 12, Figure S12.1). This suggests that results of the less formal meta-analyses were robust to the inclusion of studies that lacked replication, and that the difference between results in terms of statistical significance is due to differences in the two quantitative synthesis approaches. Taken together, results from both analyses provide some evidence that on average fish responses to changes in natural flow magnitude were mainly negative (i.e. when combining floods, droughts, high/low flows for all species but separately for abundance and biomass outcomes); however, they do not provide support for clear generalizable signals since the overall heterogeneity of effect sizes in the formal meta-analyses showed significant between-study variation in most models (i.e. effect sizes did not estimate a common population mean). Thus, investigations into the influence of moderator variables on overall effect sizes was indeed warranted. The only previous review to evaluate summary effects of natural magnitude changes on fish was McManamay et al. (2013), who reported responses to be highly variable, noting though that they combined a variety of ecological responses (e.g. abundance, diversity, reproduction).

When considering the flow-ecological relationship analysis, fish responses were variable in time series analysis, but there was some evidence suggesting fish abundance responded positively to increases in natural flow magnitude across spatial comparison studies (Figure 9). Aside from McManamay et al. (2013), no previous reviews (to our knowledge) have attempted to include information on the impact of differing flow magnitudes from primary studies where there was not a clearly defined natural event. While we performed separate summary analyses for studies that did report a clearly defined event and those that did not, McManamay et al. (2013) included both study types together.

4.1.2 | To what extent do factors influence the potential impact of changes in flow magnitude due to natural causes on fish abundance and biomass?

For both the formal and less formal meta-analyses, average fish abundance varied among intervention types when considering within-year variation, but no such effect was detected for any other analyses (Figure 7; Supporting Information 9, Tables S9.4A & S9.5A; also see Table 5 for results summary). Furthermore, in formal meta-analysis, intervention type explained much of the variation in fish abundance effect sizes for within-year analysis, with little residual heterogeneity remaining in the model (Supporting Information 9, Table S9.4A). For both quantitative approaches using within-year variation studies, floods and droughts were associated with overall negative changes in fish abundance, with a larger negative impact for floods than droughts. Interestingly, low flows (i.e. in these included cases, all decreases in low flow magnitude outside the normal range compared to the mean flow magnitude prior to the event) were associated with overall positive change in fish abundance, and the magnitude of this average response was larger than that of both floods and droughts in both the formal and less formal analyses, but seemingly more variable in the less formal analysis (Figure 7). It is unclear why decreases in low flow magnitude would result in an overall positive change in fish abundance, but the included cases were from only two studies, located in different countries (Middle Rhone River, France [Fruget et al., 2001] and Savannah River, US [Martin & Paller, 2008]); the former providing multiple species-specific abundance responses to low flows while the latter included a single density response from a combined 24 species. High flows (increases and/or decreases in maximum flow magnitudes, related to changes in peak flow, reported as changes to peak flow, high flow, or maximum flow conditions) were estimated to have a large negative effect on within-year fish abundance; however, all effect sizes, each representing a different species, were from a single study from Florida USA (Paperno & Brodie, 2004), thus caution should be taken when interpreting this result.

There are a few previous reviews that explored fish abundance responses to natural changes in flow magnitude for comparison with our results. For instance, McManamay et al. (2013) investigated quantitative relationships between natural flow changes and ecological responses in the south Atlantic region of the United States and reported variable fish responses to floods, high flows and lows flows, and predominantly negative responses to droughts. Similarly, Maxwell et al. (2019) using a more qualitative approach and a global review scope, reported extreme events like floods and droughts to have generally negative impacts on fish population size. Our less formal meta-analysis approach followed closely to those performed by Piniewski et al. (2017), and thus comparisons of our results are most appropriate. Piniewski et al. (2017), focusing on extreme events in Europe, reported a significant negative effect of droughts on fish density ($k = 6$ datasets), but no significant effects of floods on fish abundance ($k = 8$) or density ($k = 6$); however, the mean value for abundance suggested a positive-trending response to floods. In comparison, we found no significant effect of droughts (albeit a negative mean effect size) and a significant negative effect of floods on fish abundance for within-year studies (Figure 7; Table 5). Differences among our results may be due in part to our larger geographical scope and thus larger sample sizes for analyses. For example, our drought subset for within-year fish abundance studies included studies mostly from the USA ($k = 48$ datasets), but also from Australia ($k = 7$) and the United Kingdom ($k = 6$), whereas our flood subset included countries from four continents (i.e. Oceania [95], Africa [33], North America [14], and Europe [12]).

Except for the within-year abundance analysis, formal meta-analyses suggested that significant heterogeneity remained in mixed-effects models including intervention type as a moderator. Therefore, additional moderators were tested in our quantitative syntheses to explore reasons for heterogeneity within different interventions separately. Unfortunately, due to insufficient sample sizes and/or variation, we were unable to investigate factors within different interventions for biomass responses with either replication type (within-year or interannual; Figure 6), using either quantitative approach (formal or less formal meta-analyses). For formal meta-analysis, no detectable associations were found between any factor investigated within a given intervention type and average fish abundance for either within-year or interannual analyses (Figure 3; Supporting Information 9, Tables S9.4 & S9.5). Interestingly, for the interannual analysis on fish abundance responses, much of the variation in effect sizes appeared to be due to droughts ($Q = 77.13$, $p < 0.0001$) rather than floods ($Q = 127.40$, $p = 0.261$); however, none of the considered moderators explained this heterogeneity in effect sizes for droughts, suggesting that some other (untested) factor(s) may be responsible.

In contrast, for the less formal meta-analysis, several moderators were found to be associated with average effect sizes for within-year abundance studies (summarized in Table 5), including: (i) waterbody type (droughts), (ii) major habitat type (droughts and floods),

TABLE 5 Comparison of results between formal meta-analyses and less formal meta-analyses. Moderators (italics) that are in bold were significantly associated with fish responses. Underline subset categories indicates a statistically significant effect at the $p<0.05$ level, but with an asterisk (*) a moderately significant effect at the $p<0.1$ level. Underlined subset categories with [−] indicate an overall average decrease in fish outcome with natural changes in flow magnitude; [+] indicates an overall average increase in fish outcome with natural changes in flow magnitude. k: number of datasets (i.e. effect sizes); Hedges' g: weighted mean effect size; CI: 95% confidence interval; NR, not reported.

Analysis	Formal meta-analysis	Less formal meta-analysis		
Within-year Abundance				
Intervention type		***Intervention type***		
Drought* [−]	Hedges' g: −0.20 (CI: −0.40, 0.01); $k = 35$	Drought	−30% (CI: −62, 30); $k = 61$	***Waterbody type*** (streams vs rivers [−] vs. estuaries) ***Major habitat type*** (temperate coastal rivers vs tropical and subtropical coastal rivers vs temperate floodplain rivers and wetlands vs xeric freshwaters and endorheic basins* [−])
Flood [−]	Hedges' g: −0.36 (CI: −0.60, −0.11); $k = 30$	Flood [−]	−71% (CI: −87, −38); $k = 154$	***Waterbody type*** (rivers vs estuaries [−] vs. multiple) ***Major habitat type*** (temperate coastal rivers [−] vs. temperate floodplain rivers and wetlands vs temperate upland rivers vs xeric freshwaters and endorheic basins [+]) ***Water regime*** (perennial [−] vs. intermittent vs NR/unclear [−]) ***Flow regime*** (free [+] vs regulated [−] vs mixed [−] vs NR/unclear) ***Surrounding land-use*** (natural+agriculture [−] vs agriculture [−] vs natural vs NR/unclear) ***Direction of flow magnitude change*** (Increase_high vs Increase_discharge [−]) ***Event duration*** (<6 m vs >6 m<1 year [−] vs not reported/unclear [−]) ***Study validity*** (low [−] vs medium)
Low flow [+]	Hedges' g: 0.38 (CI: 0.11, 0.65); $k = 34$	Low flow [+]	169% (CI: 38, 426); $k = 116$	N/A
		High flow [−]	−99% (CI: −100, −94); $k = 63$	N/A
Interannual Abundance				
Intervention type		*Intervention type*		
Drought	Hedges' g: 0.01 (CI: −0.35, 0.37); $k = 24$	Drought	−23% (CI: −72, 109); $k = 53$	No detectable effects of moderators
Flood	Hedges' g: 0.13 (CI: −0.15, 0.41); $k = 127$	Flood [−]	−55% (CI: −74, −22); $k = 148$	No detectable effects of moderators
		Low flow	−59% (CI: −91, 75); $k = 34$	N/A
		Other	−1% (CI: −38, 58); $k = 6$	N/A
Within-year Biomass				
N/A		*N/A*		
		Flood [−]	−98% (CI: −100, −84); $k = 36$	N/A
Interannual Biomass				
Intervention type		*Intervention type*		
Drought	Hedges' g: −0.60 (CI: −2.95, 1.75); $k = 9$	Drought	−14% (CI: −39, 21); $k = 12$	N/A
Flood	Hedges' g: −2.44 (CI: −6.03, 1.16); $k = 21$	Flood	−90% (CI: −99, 23); $k = 21$	N/A
		Other [+]	16% (CI: 3, 30); $k = 6$	N/A

(iii) water regime (floods), (iv) flow regime (floods), (v) surrounding land-use (floods), (vi) direction of flow magnitude change (floods), (vii) event duration (floods) and (viii) study validity (floods). These analyses provide some support that droughts result in negative average effects on fish abundance studied in rivers, and more broadly in the xeric freshwaters and endorheic (closed) basin major habitat type compared to other waterbody and major habitat types where mean effects were not significantly different from zero (noting here that all seven responses from xeric habitats were rivers, suggesting the significant effect of this habitat type could be due to a correlation with waterbody type). Floods resulted in negative average effects on fish abundance for (i) estuaries, (ii) temperate coastal river major habitat type, (iii) perennial systems, (iv) regulated systems, (v) waterbodies mostly or partially surrounded by agricultural lands, (vi) studies specifying increases in discharge, (vii) flood events that last >6 months but <1 year in duration and (viii) studies assessed as having low validity (highly susceptible to bias) and positive average effects on fish abundance for xeric freshwaters and endorheic basin major habitat type, compared to other categories considered (Table 5; but see also Supporting Information 10); however, many of these factors were highly correlated (Supporting Information 6). For example, of the 34 responses captured for floods in estuaries, 32 were regulated temperate coastal river habitats that specified increases in discharge. These correlations among factors makes interpretation challenging and sample sizes would not allow us to explore multiple factors at once, unfortunately limiting our ability to determine the relative impact of each moderator on overall mean effect sizes. Most previous reviews, given their relatively small sample sizes, were unable to explore the influence of many moderators for comparison. However, McManamay et al. (2013) found some evidence that the occurrence of floods and high flows in an unconstrained coastal plain stream may have less negative impacts to river communities than in a floodplain-constrained upland stream. Our results do not appear to support this finding when considering a broader geographical scope, albeit recognizing that we did not differentiate between constrained and unconstrained systems.

The flow-ecological relationship analysis on time series studies provided some support that natural changes in magnitude in rivers may be less of a disturbance to fish abundance but more of a disturbance in estuaries, compared to streams or reservoirs, where no relationships were apparent (Figure 10a). This finding of potentially less negative consequences for rivers than streams for fish abundance was also found with spatial comparison studies (Figure 11a). The investigations of natural magnitude across space also suggested positive correlations with fish abundance (i.e. higher abundance in sites with higher magnitude) for temperate floodplain river and wetland habitat types, waterbodies that were intermittent, waterbodies that included other flow regime component alterations in addition to magnitude changes, and studies that measured abundance (rather than density or CPUE), compared to the other categories within the different considered moderators (Figure 11). Note however, interpretation of these findings should be accompanied with appropriate consideration for study validity as studies used here were all considered lower validity due to inadequate study designs for fully addressing our review questions.

Taxonomic responses to floods and droughts varied across families but were all non-significant (Figure 8). While most mean abundance responses for considered families were trending negative to droughts, mean responses to floods were quite variable across families, indicating that specific families, genera or species may not respond consistently to natural changes in flow magnitude, presumably linked to family-level environmental constraints and requirements. However, caution should be taken when interpreting these results for most taxonomic groups because sample sizes were small, and we were unable to explore potential moderators.

4.2 | Limitations of review methods

We attempted to minimize potential biases in our review methodology throughout the systematic review process. Our diverse advisory team including stakeholders and topic experts from academia, government, industry and a non-profit organization, helped us identify as many relevant studies as possible, minimizing familiarity bias. While we identified 300 relevant studies, 193 of which were eligible for quantitative analysis, we acknowledge that our review does not represent the entirety of the knowledge base on the subject. Efforts were made to obtain all relevant materials to decrease availability bias; only one publication could not be found (Supporting Information 2). One important consideration of our review is that it was limited to English language literature. Although this captures most articles available, we acknowledge there may also be additional, valuable articles and grey literature not published in English. While there is some evidence in medical science that restricting reviews to English-language literature has little impact on the effect estimates and conclusions of reviews (e.g. Dobrescu et al., 2021; Nussbaumer-Streit et al., 2020), others have suggested that ignoring non-English-language studies may bias outcomes of ecological meta-analyses (e.g. Konno et al., 2020). There were relatively few articles excluded from the systematic map on language at full text (61/2412 articles; Rytwinski et al., 2020) and only six excluded on language during this review (e.g. Chinese, Japanese; Supporting Information 7, Figure S7.1). It is unclear whether these articles would have been deemed relevant based on our inclusion criteria.

There was no obvious indication of publication bias from any formal meta-analyses on fish abundance (Supporting Information 9, Figure S9.2 & S9.14); however, there was possible evidence of publication bias, for fish biomass in interannual studies, towards studies with larger sample sizes showing positive effects of flow magnitude change (Figure S9.19). It is interesting to note that all but one article within this subgroup were from commercially published sources. It is unclear how many additional grey literature sources in the form of internal reports may exist that were not accessible to our review team (especially in the last two decades where grey literature made up less than 10% of studies identified; Supporting Information 7, Figure S7.2).

4.3 | Limitations of the evidence base

Of the 300 studies included in this review, 107 were excluded from all quantitative syntheses due to the use of either qualitative (e.g. presence/absence) or semi-quantitative (e.g. presence/quantitative value) outcomes in the intervention and/or comparator groups, making effect size calculation impossible. Because we used three quantitative synthesis approaches, we were able to make use of as much of the available quantitative evidence as possible. However, if we had only used the more well-documented formal meta-analysis approach, as is typically done in most systematic reviews, we would have had to exclude an additional 154 studies (with a total of 1447 datasets) because they (i) lacked replication in both the intervention and comparator groups and/or (ii) did not include a comparator. The former issue was common because most studies did not report within-year responses (i.e. data for each month or season of sampling reported separately for each sampling year), instead providing single data points, sums or averages across within-year sampling without also providing within-year fish outcomes over several years. As a result, we were unable to analyse variability through time effectively. Reporting quantitative fish outcomes and providing raw or finer-scale sampling data (i.e. via online data repositories or journal appendices), rather than pooling data across samples, would provide opportunities to improve future (systematic) reviews.

For the formal and less formal meta-analyses, we did not quantify the amount of flow magnitude change (i.e. ΔQ, where Q is discharge) because many articles did not report sufficient quantitative information to do so (e.g. authors only provided measures of flow during/following a change in magnitude and not also prior to the event, or used qualitative descriptions of magnitude changes rather than quantitative values). The issue of insufficient reporting of flow data has also been identified in previous reviews on the general topic of flow-fish relationships (e.g. Harper et al., 2022; Piniewski et al., 2017). We felt our qualitative descriptions of flow changes based on author descriptions allowed us to adequately categorize flow changes and thus capture more studies for quantitative analyses than would otherwise have been possible. We did, however, quantify percent changes in flow magnitude from values associated with the maximum and minimum outcome variables for our flow-ecological relationship analyses when information on both variables were sufficiently reported in articles. These analyses provided some statistically supported patterns between categories of moderators and fish responses (e.g. fish abundance showed a significant positive correlation with natural flow magnitude within rivers, and a negative correlation within estuaries). Furthermore, we did not observe any obvious 'threshold' relationships that would be particularly useful in a management context (see Arthington et al., 2006). However, given that these analyses were based on study designs that did not include a true comparator, and that flow magnitude change was calculated in relation to the reported maximum and minimum fish outcome values and, therefore, does not necessarily represent the full range of flow variation experienced within the study, we do not feel these analyses can accurately or robustly assess and identify thresholds. Attempting to identify thresholds was not an intended goal of our flow-ecological relationship analyses nor did we attempt to do so using relevant quantitative tools. However, as variation in flow magnitude increases as a result of climate change, identifying thresholds that could inform management or conservation actions will become increasingly important. This will only be possibly from a quantitative synthesis sense if the evidence base is improved with more information that would enable such analyses.

The available evidence base on this topic was deemed to be of generally low study validity. Of the datasets included in quantitative syntheses, 62% had 'Low' validity, while the remainder had 'Medium' validity. When considering only those included in the formal meta-analysis, the percentage of 'Medium' validity studies was higher (i.e. 65%). Despite the relatively low validity of included studies, sensitivity analyses suggested that results were likely robust to the inclusion of studies with lower validity (Supporting Information 9). Improving study designs by including temporal and spatial replication, providing quantitative data on flow magnitude changes, and increasing replication would aid in improving internal validity of primary studies. However, we acknowledge that flow variability from natural events may not be as predictable as flows downstream from a peaking hydropower facility so there will always be inherent challenges with experimental design.

We were unable to draw clear conclusions of long-term effects of natural changes in flow magnitude on fish abundance or biomass. This was because many interannual studies included in quantitative synthesis were based on relatively short-term monitoring (i.e. 65% of data sets included in formal meta-analysis only had 2 years post-intervention monitoring), limiting our ability to investigate resilience and recovery of fish populations to natural disturbances. This issue has been noted in previous reviews as well (e.g. Maxwell et al., 2019). Instead, this synthesis is largely based on more direct, immediate responses that quantitatively characterize fish resistance (i.e., capacity of fish to withstand the stresses of a disturbance) to natural changes in flow magnitude. Interestingly, within-year responses to changes in natural flow magnitude tended to result in more negative fish responses overall within a given outcome type compared to interannual responses, suggesting fish may be more susceptible to negative impacts of magnitude changes within the immediate time frame of an event (i.e. lower mean monthly abundance within the first year following an event compared to the year before the event), compared to average fish responses over the years following the event (i.e. comparing yearly averages before and after the event). Indeed, although the overall mean effect sizes were all non-significant in global analyses, we did see a trend in the overall mean weighted effect sizes for within-year fish abundance which became increasingly positive with post-intervention years (Figure 5). These results may provide some evidence (albeit weak) for the capacity of fish populations to recover from such disturbances. Furthermore, significant heterogeneity was observed between effect sizes remaining in the interannual models, suggesting other (untested) factors could be causing this variation. For example, if fish respond differently after a few years following a hydrological event, short-term studies may not capture changes in fish outcomes such as population size, especially if only the adults of long-lived species are monitored for relatively short periods post intervention (Lyon et al., 2021). Although we attempted to collect information on species life stage, many studies did not report relevant information to allow for further investigation in interannual analyses. There may be a need for future meta-analyses that account for species/population-specific time-lags arising from changes in natural flow magnitude. Presumably the timing of events such as flood or drought will have differential impacts (e.g. during versus after the spawning period; Detenbeck et al., 1992) but such information is not always reported either. As noted by Maxwell et al. (2019), information from long-term studies will be critical for improving predictions of species responses to events like floods and drought, thus these studies should be supported.

We originally intended to investigate the influence of factors related to waterbody characteristics (e.g. stream gradient, stream order, water temperature, pH) but these were rarely reported for sampling sites in articles. The importance of these variables in influencing ecological responses to changes in flow regime have been previously noted (e.g. McManamay et al., 2013; Poff & Ward, 1989; Poff & Zimmerman, 2010), consequently, we recommend reporting such information in publications (or through additional files or online data repositories) to aid in future reviews.

There was evidence of geographic but not taxonomic bias in the available evidence base. Most datasets for quantitative synthesis were from North America (41%), of which 94% were from the United States, potentially limiting interpretation of review results to other geographic regions. Similar geographical biases were identified in the systematic map (Rytwinski et al., 2020) and in reviews on anthropogenic impacts of flow magnitude alterations (Harper et al., 2022). Taxonomic bias, however, was less evident. Impacts of natural changes in flow magnitude were evaluated for a large diversity of fish families (total of 124), the most common being Leuciscidae representing 15% of all datasets, followed by Salmonidae with 9%. In total, 437 fish species were represented in quantitative syntheses, with 53 species having more than 10 datasets, suggesting interpretation of results may not be limited by taxa.

5 | CONCLUSIONS

5.1 | Implications for policy/management

Our results suggest that overall fish abundance and biomass responses to changes in natural flow magnitude were mainly negative but our analyses do not provide support for clear general responses across all contexts (e.g. types of changes in flow magnitude, taxa, locations). Our findings of a lack of generalizable and transferable relationships between flow magnitude and fish responses are consistent with previous reviews that have focused on direct anthropogenic causes of flow alterations (Harper et al., 2022; Poff & Zimmerman, 2010), suggesting that regardless of the cause of flow changes (natural or anthropogenic), or if only focusing on a specific flow component (i.e. magnitude), generalizations may not be possible. This lack of a generalizable relationship however should not be taken as evidence that changes to flow magnitude have little effect on fish abundance and biomass, as our analyses show that context matters. Indeed, when stratifying studies into subgroups, a few patterns emerged.

First, we found strong support that fish were responding (in terms of abundance) differently to natural events considered in our analyses, with consistently negative responses to floods and droughts, and positive responses to decreases in low flows within the first year of the natural change in flow magnitude. In relation to the latter, attempting to look more closely at the included low flow studies (of which there were only two), did not provide clarity for this unexpected result. As such, caution should be taken when interpreting these results until further research is conducted on the impacts of low flows on fish outcomes.

Second, these patterns from the within-year variation analysis were less evident when considering interannual variation. This suggests that while immediate responses (i.e. within the first-year post-intervention) were more apparent and relatively consistent within specific types of natural events, fish populations may recover after such events (i.e. ≥2 years post-natural event). Indeed, in a review of case studies on the recovery of temperate-stream fish from disturbance, Detenbeck et al. (1992) found that population recovery times following pulse disturbances (i.e. floods, droughts, chemical spills, construction activity, nonchemical organism removal) varied between 0.08 and 6 years following such disturbances, with over 60% of systems studied showing population recovery within 3 years. Furthermore, our review suggested that longer-term effects (i.e. ≥2 years post-intervention) of natural changes in flow magnitude were more variable and may be context dependent. Unfortunately, due to small sample sizes, we were limited in our ability to explore this variability to discern any further contextual patterns (e.g., investigate at finer regional or taxonomic scales/levels). If management/conservation decisions are urgent (i.e. waiting for more primary studies to allow for such investigations is not an option), the outputs of this systematic review provide managers with a comprehensive evidence base that they can use to assess the available evidence that is relevant to their specific contexts and/or regions (i.e. attempt to use what evidence is available now). For example, managers could search the database of included studies (see Rytwinski et al., 2022b) for specific species, regions, or freshwater systems of interest to identify available primary studies captured on this topic. Using the extracted information, including quantitative measurements of the impact when available (i.e. direction and magnitude of a fish response in relation to a natural change in flow magnitude), along with the results of the study validity assessments, managers can assess the available evidence to help guide their management/conservation decisions. Furthermore, the outputs of this review could help managers with planning other activities in more managed systems. For instance, knowing how a particular species (or group of species) responds to a natural event such as a drought, can help in conservation planning to mitigate other impacts (e.g. manage the amount of water allowed to leave a system for other uses), or allow managers to decide whether to augment the system with environmental flows within the same or following season (Acreman et al., 2014).

Third, there is some evidence suggesting that differing natural flow magnitudes in rivers (over time and across space) may be less of a disturbance to fish abundance than in other waterbody types (e.g. streams, estuaries, reservoirs); however, there may be high uncertainty associated with these results owing to inadequate study designs of the primary studies used in these analyses.

Lastly, results from this systematic review are important in characterizing fish responses to climate induced alterations in flow magnitude, supporting management and conservation efforts that could mitigate negative impacts. Recent guidance, referred to as 'Climate-Smart Conservation Practice' (Brown et al., 2022) has been developed to integrate climate change into the planning and adaptive management of conservation projects. A key element of this integrated approach—which is based on the widely used Conservation Standards (CMP, 2013)—is the development of 'climate-smart' situation models linking current and future climate threats and conventional (human) threats to management/conservation targets. Outputs from our review could be useful to those using this climate-smart framework if they want to identify potential relationships between climatic variables/threats and target fish to develop management strategies to mitigate the negative effect of climate and/or other anthropogenic threats.

5.2 | Implications for research

5.2.1 | Improving study designs and reporting

The principal goal of any systematic review is to ensure conclusions drawn are correct. If bias is present in primary studies, their results will be incorrect. Subsequently, if a systematic review is based on incorrect evidence, the results of the quantitative analysis will also be incorrect (Boutron et al., 2022; Rytwinski et al., 2021). The available evidence base on this topic was deemed to be of relatively low validity due to limited comparators, replication and justified sampling methodology within the studies and cases assessed. While results did not appear to be influenced by the inclusion of lower validity studies in the formal meta-analysis, it is important to note that there were no studies assessed to have high validity based on our assessment criteria. It is unclear though how or if the inclusion of higher validity studies (if some existed) would have influenced our review findings. Reflecting further on our assessment criteria post-hoc, we do feel that, while certain criteria were likely representative of what would be feasible in the real world [e.g. to score 'High' on replication, a BA study design would only require ≥2 years post-intervention (interannual variation), or ≥2 months/seasons post-intervention (within-year variation)], some criteria may be more difficult to achieve when studying impacts of natural flow changes in the field [e.g. scoring 'High' on study design, required the use of both temporal and spatial comparators in a BACI design or randomized control trials (RCT)]. For instance, capturing data before a natural change in flow such as an extreme flood or drought requires either knowledge of when the event will occur or a fortuitous choice of sampling sites (De Palma et al., 2018). Therefore, designing a study with a spatial comparison(s), in addition to pre-event data, would be very difficult to achieve as there is no or very limited advanced planning possible for extreme weather events, likely making full BACIs unrealistic (though not impossible, see below). Drawing from issues we encountered during quantitative synthesis and common features of studies in our evidence base, we recommend the following best practices for improving future study designs and reporting.

First, one of the only ways to capture pre-event data (and spatial comparisons) is to maintain regular, long-term monitoring programs at sites for the sake of monitoring. If an extreme event were to fortuitously occur at those monitored sites, data would then be available to investigate the impact(s) of the event. When assessing fish responses to flow magnitude alterations, long-term monitoring both before and after the change in flow would facilitate improved understanding of population-level effects and time-lags in responses. This is especially important for longer-lived species or those returning to critical habitats (e.g. spawning). Furthermore, as noted by Harper et al. (2022), efforts should be made to minimize gaps between sampling years, and to ensure sampling occurs in multiple seasons. From a quantitative perspective, a minimum of two years before and two years after are needed to calculate a standardized effect size measure such as Hedges' g for use in a more formal meta-analysis. However, longer time periods (i.e. >2 years before and after) are highly encouraged to improve the precision of effect estimates and our understanding of temporal dynamics of fish recovery following natural events. Previous studies have suggested that recovery of fish metrics were found within 3 years following pulse disturbances in most studied systems (e.g. natural and anthropogenic pulse disturbances [Detenbeck et al., 1992]; anthropogenic pulse disturbance [Rohr et al., 2021]); however, maximum time to recovery was longer for some systems/contexts (e.g. depending on timing of disturbance relative to spawning season, presence of barriers to migration, frequency of disturbance). From an ecological perspective then, these studies provide some evidence that at least three years of monitoring should be undertaken following an event. Therefore, while there may be limited control over how many years of pre-event data are captured during regular or long-term monitoring, if an extreme event fortuitously occurs at these sampling sites, monitoring of fish responses should continue for at least three years following that natural event to properly facilitate investigations of its impact on fish responses.

Second, in the rare case that prior knowledge exists in the timing of an event or where suitable before data is available retrospectively, BACI design usage should be encouraged (as others have advocated before, e.g. Christie et al., 2019; De Palma et al., 2018). For instance, through simulation, Christie et al. (2019) demonstrated that BACI (and randomized control trials) are far more accurate than BA, CI and after-only designs. When correctly estimating true effect's direction and magnitude to within ±30%, BACI designs performed 2.9–4.2 times better than BAs, 3.2–4.6 times versus CIs and 7.1–10.1 times versus after-only designs (depending on sample size). They suggested this was because increasing sample size improved BACI design accuracy, but only increased the precision (not accuracy) of simpler designs around biased estimates.

Third, develop standardized monitoring to ensure methodologies are applied consistently over time and space. This is especially crucial for long-term studies. Included in this development should be considerations for seasonality of sampling and choice of outcome metrics. For instance, fish abundance is affected by the emergence of young of the year (YOY) and is much lower in spring than in fall in most temperate systems. Future studies/monitoring programs should avoid this potential confounding issue in study design by sampling fish at the same time of year. Furthermore, if within-year comparisons are being made before and after a flow event, and fish abundances are changing within seasons, (a) using other metrics instead of abundance (i.e. biomass, richness) or (b) excluding YOY from abundance comparisons if they are recruited to the sampling gear during the study period is considered. Although standardized methods have tended to be developed at local or provincial/state levels in North America and Europe (Bonar et al., 2017), progress is ongoing in the development and adoption of national- and continent-wide standards for fish sampling (e.g. Bonar et al., 2009; European Commission, 2015; European Committee for Standardization [CEN], 2003, 2006, 2014, 2015; National Park Service [NPS], 2021; see also Hering et al., 2010; Rodhouse et al., 2016 for discussions on achievements stemming from some of these standardizations).

Fourth, studies should report sufficient detail regarding location of sample sites and waterbody characteristics (i.e. latitude and longitude, gradient, stream order, water temperature).

Fifth, when possible, studies should report summarized outcome data separately for monthly or seasonal samples within a year, and report detailed descriptions of how samples are grouped for analysis or provide raw data.

Sixth, future studies should report sufficient information to allow quantitative assessment of the amount of flow magnitude change or event severity. For example, when reporting changes in flow magnitude, comparable data (i.e. measured flow magnitude or hydrographs) from the same temporal period (i.e. season) should be included for both the intervention and comparator groups.

Lastly, in general, for reporting, where information cannot fit within published articles, details should be included in supplementary materials and data should be shared in data archivers or repositories. Furthermore, consider entering metadata associated with research methods and results in databases that support systematic literature assessments (see Norton et al., 2018).

5.2.2 | Addressing research gaps

The evidence base of this review encompasses a large diversity of temperate freshwater fish. However, when stratifying families into particular contexts, we were limited in our investigations to only a few families owing to small sample sizes (e.g. for droughts just three families: Leuciscids, Salmonids, Centrarchids). Moreover, most studies focused on systems in the USA. Therefore, studies that focus on species-specific responses to natural flow magnitude changes in systems outside of North America are needed to address knowledge gaps.

Our analyses did not provide support for clear generalizable signals of the impacts of natural flow magnitude changes on fish abundance or biomass across all contexts. To improve clarity, regional, species-specific, long-term continuous monitoring studies are recommended in future. For this review, although before and after year monitoring went beyond 10 years in some included studies for the formal analysis, most involved <5 years of monitoring pre- and post-natural event (i.e. before year monitoring: range = 2–16, mean = 3.7; mode = 4 years; after year monitoring: range = 2–11, mean = 3.0; mode = 2 years). Therefore, standardized ecological studies spanning periods longer than found in this review will improve our understanding of fish responses to natural changes in flow regimes and learn more about their specific systems (Hampton et al., 2019), as well as enable researchers to capture unpredictable, extreme events as they occur, while ensuring base-line data is available for comparison. Furthermore, quantification of fish responses to changes in flow regime from long-term continuous monitoring of natural systems would also provide invaluable information to help manage these changes in altered systems (e.g. due to hydropower production, nuclear facilities, land-use change; press disturbances). For example, future research needs include improving our understanding of how long a disturbance can be before it impacts fish, or how long of a period fish need with 'normal' flows to recover naturally following a disturbance. These types of questions are increasingly sought out by water resource managers for their usefulness in interpreting and predicting impacts of and recovery from flow magnitude changes in altered systems. With increasing pressures from climate change, long-term ecological studies can enable managers to determine if the system may have natural capacity for recovery under the influence of future, multiple interacting effects, or, if not, to develop management interventions that may mitigate species-specific impacts (Gaiser et al., 2020).

AUTHOR CONTRIBUTIONS

Jessica J. Taylor undertook searches. Adrienne Smith, Jessica L. Reid, and Trina Rytwinski performed screening. Meagan Harper and Trina Rytwinski conducted coding and extraction of articles and study validity assessments. Trina Rytwinski and Hsien-Yung Lin performed quantitative syntheses. This review is based on an initial draft written by Trina Rytwinski and Meagan Harper. All authors assisted in the design of methodology, contributing critically to the drafts, and gave final approval for publication.

ACKNOWLEDGEMENTS

The authors thank reviewers and collaborators who provided valuable insight to strengthen this review, including Stuart Bunn and Neil Fisher. The authors also thank the reviewers and all those involved in the systematic map that led to the identification of this topic. The review was primarily supported by Fisheries and Oceans Canada. Additional support was provided by the Natural Science and Engineering Research Council of Canada and Carleton University.

CONFLICT OF INTEREST STATEMENT

The authors declare that they have no competing interests.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/eso3.12216.

DATA AVAILABILITY STATEMENT

All important additional information is currently provided in Supporting Information. ROSES form for this systematic review can be found here: https://doi.org/10.6084/m9.figshare.21770417.v1 (Rytwinski et al., 2022a). Our extraction sheet containing the coding for all articles/studies can be found here: https://doi.org/10.6084/m9.figshare.21770123.v1 (Rytwinski et al., 2022b). Results of study validity assessments can be found here: https://doi.org/10.6084/m9.figshare.21770465.v1 (Rytwinski et al., 2022c). Our quantitative synthesis database can be found here: https://doi.org/10.6084/m9.figshare.21770321.v1 (Rytwinski et al., 2022d).

ORCID

T. Rytwinski *https://orcid.org/0000-0001-6764-7309*
M. Harper *https://orcid.org/0000-0002-8462-2039*
K. E. Smokorowski *https://orcid.org/0000-0003-1530-7024*
J. J. Taylor *https://orcid.org/0000-0002-2080-6437*
K. Birnie-Gauvin *https://orcid.org/0000-0001-9242-0560*
M. J. Bradford *https://orcid.org/0000-0002-4810-0495*
J. A. Crossman *https://orcid.org/0000-0002-7134-982X*
K. Turgeon *https://orcid.org/0000-0002-2528-130X*
S. J. Cooke *https://orcid.org/0000-0002-5407-0659*

REFERENCES

Abell, R., Thieme, M. L., Revenga, C., Bryer, M., Kottelat, M., Bogutskaya, N., Coad, B., Mandrak, N., Balderas, S. C., Bussing, W., Stiassny, M. L. J., Skelton, P., Allen, G. R., Unmack, P., Naseka, A., Ng, R., Sindorf, N., Robertson, J., Armijo, E., … Petry, P. (2008). Freshwater ecoregions of the world: A new map of biogeographic units for freshwater biodiversity conservation. *BioScience*, *58*, 403–414. https://doi.org/10.1641/B580507

Acreman, M., Arthington, A. H., Colloff, M. J., Couch, C., Crossman, N. D., Dyer, F., Overton, I., Pollino, C. A., Stewardson, M. J., & Young, W. (2014). Environmental flows for natural, hybrid, and novel riverine ecosystems in a changing world. *Frontiers in Ecology and the Environment*, *12*, 466–473. https://doi.org/10.1890/130134

Arthington, A. H., Bunn, S. E., Poff, N. L., & Naiman, R. J. (2006). The challenge of providing environmental flow rules to sustain river systems. *Ecological Applications*, *16*, 1311–1318. https://www.jstor.org/stable/40062003

Birnie-Gauvin, K., Rytwinski, T., Harper, M., Taylor, J. J., Smith, A., Smokorowski, K. E., Turgeon, K., Bradford, M. J., & Cooke, S. J. (2021). How do natural changes in flow magnitude affect fish abundance and diversity in temperate regions? A systematic review protocol. *Ecological Solutions and Evidence*, *2*, e12079. https://doi.org/10.1002/2688-8319.12079

Bonar, S. A., Hubert, W. A., & Willis, D. W. (2009). *Standard methods for sampling north American freshwater fishes*. American Fisheries Society.

Bonar, S. A., Mercado-Silva, N., Hubert, W. A., Beard, T. D., Jr., Dave, G., Kubecka, J., Graeb, B. D. S., Lester, N. P., Porath, M., & Winfield, I. J. (2017). Standard methods for sampling freshwater fishes: Opportunities for international collaboration. *Fisheries*, *42*, 151–156. https://doi.org/10.1080/03632415.2017.1276352

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). Effect sizes based on means. In M. Borenstein, L. V. Hedges, J. P. T. Higgins, & H. R. Rothstein (Eds.), *Introduction to meta-analysis* (pp. 21–32). Wiley.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2010). A basic introduction to fixed-effect and random-effects models for meta-analysis. *Research Synthesis Methods*, *1*(2), 97–111. https://doi.org/10.1002/jrsm.12

Boutron, I., Page, M. J., Higgins, J. P. T., Altman, D. G., Lundh, A., & Hróbjartsson, A. (2022). Chapter 7: Considering bias and conflicts of interest among the included studies. In J. P. T. Higgins, J. Thomas, J. Chandler, M. Cumpston, T. Li, M. J. Page, & V. A. Welch (Eds.), *Cochrane handbook for systematic reviews of interventions*. version 6.3 (updated February 2022). http://www.training.cochrane.org/handbook

Brown, M. B., Morrison, J. C., Schulz, T. T., Cross, M. S., Püschel-Hoeneisen, N., Suresh, V., & Eguren, A. (2022). Using the conservation standards framework to address the effects of climate change on biodiversity and ecosystem services. *Climate*, *10*, 13. https://doi.org/10.3390/cli10020013

Bunn, S. E., & Arthington, A. H. (2002). Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. *Environmental Management*, *30*(4), 492–507. https://doi.org/10.1007/s00267-002-2737-0

CEN (European Committee for Standardization). (2003). Water quality—Sampling of fish with electricity. In *European Committee for Standardization, European standard EN 14011:2003*.

CEN (European Committee for Standardization). (2006). *Water quality—Guidance on the scope and selection of fish sampling methods*. European Committee for Standardization, European standard EN 14962:2006.

CEN (European Committee for Standardization). (2014). *Water quality—Guidance on the estimation of fish abundance with mobile hydroacoustic methods*. European Committee for Standardization, European standard EN 15910:2014.

CEN (European Committee for Standardization). (2015). *Water quality—Sampling of fish with multi-mesh gillnets*. European Committee for Standardization, European standard EN 14757:2015.

Christie, A. P., Amano, T., Martin, P. A., Shackleford, G. E., Simmons, B. I., & Sutherland, W. J. (2019). Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology*, *56*, 2742–2754. https://doi.org/10.1111/1365-2664.13499

Clarke, K. D., Pratt, T. C., Randall, R. G., Scruton, D. A., & Smokorowski, K. E. (2008). Validation of the flow management pathway: Effects of altered flow on fish habitat and fishes downstream from a hydropower dam. *Canadian Technical Report of Fisheries and Aquatic Sciences*, *2784*, 111.

CMP (Conservation Measures Partnership). (2013). *The open standards for the practice of conservation*. Ver. 3.0. Conservation Measures Partnership. http://cmp-openstandards.org/wp-content/uploads/2014/03/CMP-OS-V3-0-Final.pdf

Cohen, J. (1977). *Statistical power analysis for the Behavioural sciences* (Revised ed.). Academic Press.

Collaboration for Environmental Evidence. (2018). *Guidelines and standards for evidence synthesis in environmental management version 5.0* [online]. A. S. Pullin, G. K. Frampton, B. Livoreil, & G. Petrokofsky (Eds.). https://www.environmentalevidence.org/information-for-authors

De Palma, A., Sanchez Ortiz, K., Martin, P. A., Chadwick, A., Gilbert, G., Bates, A. E., Börger, L., Contu, S., Hill, S. L. L., & Purvis, A.

(2018). Challenges with inferring how land-use affects terrestrial biodiversity: Study design, time, space and synthesis. *Next Generation Biomonitoring*, *58*, 163–199. https://doi.org/10.1016/bs.aecr.2017.12.004

Detenbeck, N. E., DeVore, P. W., Niemi, G. J., & Lima, A. (1992). Recovery of temperate-stream fish communities from disturbance: A review of case studies and synthesis of theory. *Environmental Management*, *16*, 33–53.

Dobrescu, A. I., Nussbaumer-Streit, B., Klerings, I., Wagner, G., Persad, E., Sommer, I., Herkner, H., & Gartlehner, G. (2021). Restricting evidence syntheses of interventions to English-language publications is a viable methodological shortcut for most medical topics: A systematic review. *Journal of Clinical Epidemiology*, *137*, 209–217.

Drahota, A., & Bellor, E. (2008). *RevMan calculator: Finding standard deviations* [online]. The Cochrane Collaboration. https://training.cochrane.org/resource/revman-calculator

European Commission. (2015). *The EU water framework directive—Integrated river basin management for Europe*. European Commission. https://ec.europa.eu/environment/water/water-framework/index_en.html

Feldman, A. D. (2000). Hydrologic modeling Sytem HEC-HMS technical reference manual. In *Technical reference manual CPD-74B*. US Army Corps of Engineers: Hydrologic Engineering Center.

Fricke, R., Eschmeyer, W., & Van der Laan, R. (Eds.). (2020). *Eschemeyer's catalog of fishes: Genera, species, references* [online]. http://researchharchive.calacademy.org/research/ichthyology/catalog/fishcatmain.asp

Froese, R., & Pauly, D. (Eds.). (2021). *FishBase* [online]. https://fishbase.org

Fruget, J. F., Centofanti, M., Dessaix, J., Olivier, J. M., Druart, J. C., & Martinez, P. J. (2001). Temporal and spatial dynamics in large rivers: Example of a long-term monitoring of the middle Rhône River. *Annales de Limnologie*, *37*(3), 237–251.

Gaiser, E. E., Bell, D. M., Castorani, M. C. N., Childers, D. L., Groffman, P. M., Jackson, C. R., Kominoski, J. S., Peters, D. P. C., Pickett, S. T. A., Ripplinger, J., & Zinnert, J. C. (2020). Long-term ecological research and evolving frameworks of disturbance ecology. *BioScience*, *70*(2), 141–156. https://doi.org/10.1093/biosci/biz162

Gillespie, B. R., Desmet, S., Kay, P., Tillotson, M. R., & Brown, L. E. (2015). A critical analysis of regulated river ecosystem responses to managed environmental flows from reservoirs. *Freshwater Biology*, *60*(2), 410–425. https://doi.org/10.1111/fwb.12506

Gordon, N. D., McMahon, T. A., Finlayson, B. L., Gippel, C. J., & Nathan, R. J. (2004). *Stream hydrology: An introduction for ecologists*. John Wiley and Sons.

Haddaway, N. R., Macura, B., Whaley, P., & Pullin, A. S. (2018). ROSES RepOrting standards for systematic evidence syntheses: Pro forma, flow-diagram and descriptive summary of the plan and conduct of environmental systematic reviews and systematic maps. *Environmental Evidence*, *7*(7). https://doi.org/10.1186/s13750-018-0121-7

Hampton, S. E., Scheuerell, M. D., Church, M. J., & Melack, J. M. (2019). Long-term perspectives in aquatic research. *Limnology and Oceanography*, *64*, S2–S10.

Harper, M., Rytwinski, T., Taylor, J. J., Bennett, J. R., Smokorowski, K. E., Olden, J. D., Clarke, K. D., Pratt, T., Fisher, N., Leake, A., & Cooke, S. J. (2022). How do changes in flow magnitude due to hydroelectric power production affect fish abundance and biomass in temperate regions? A systematic review. *Environmental Evidence*, *11*(3), 3. https://doi.org/10.1186/s13750-021-00254-8

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65.

Hering, D., Borja, A., Carstensen, J., Carvalho, L., Elliott, M., Feld, C. K., Heiskanen, A.-S., Johnson, R. K., Moe, J., Pont, D., Solheim, A. L., & van de Bund, W. (2010). The European water framework directive at the age of 10: A critical review of the achievements with recommendations for the future. *Science of the Total Environment*, *408*, 4007–4019. https://doi.org/10.1016/j.scitotenv.2010.05.031

Kassambara, A. (2021). *Rstatix: Pipe-friendly framework for basic statistical tests*. R package version 0.7.0. https://CRAN.R-project.org/package=rstatix

Konno, K., Akasaka, M., Koshida, C., Katayama, N., Osada, N., Spake, R., & Amano, T. (2020). Ignoring non-English-language studies may bias ecological meta-analyses. *Ecology and Evolution*, *10*(13), 6373–6384. https://doi.org/10.1002/ece3.6368

Kottek, M., Grieser, J., Beck, C., Rudolf, B., & Rubel, F. (2006). World map of the Köppen-Geiger climate classification updated. *Meteorologische Zeitschrift. Schweizerbart'sche Verlagsbuchhandlung*, *15*, 259–263.

Lyon, J. P., Bird, T., Tonkin, Z., Raymond, S., Sharley, J., & Hale, R. (2021). Does life history mediate the effects of discharge as a driver of multi-decadal changes in population abundance and size structure of freshwater fish? *Ecological Applications*, *31*(8), e02430. https://doi.org/10.1002/eap.2430

Lytle, D. A., & Poff, N. L. (2004). Adaptation to natural flow regimes. *Trends in Ecology & Evolution*, *19*(2), 94–100.

Macura, B., Byström, P., Airoldi, L., Eriksson, B. K., Rudstam, L., & Støttrup, J. G. (2019). Impact of structural habitat modifications in coastal temperate systems on fish recruitment: A systematic review. *Environmental Evidence*, *8*(14). https://doi.org/10.1186/s13750-019-0157-3

Martin, F. D., & Paller, M. H. (2008). Ichthyoplankton transport in relation to floodplain width and inundation and tributary creek discharge in the lower Savannah River of Georgia and South Carolina. *Hydrobiologia*, *598*, 139–148.

Martin, P. A., Shackelford, G. E., Bullock, J. M., Gallardo, B., Aldridge, D. C., & Sutherland, W. J. (2020). Management of UK priority invasive alien plants: A systematic review protocol. *Environmental Evidence*, *9*(1). https://doi.org/10.1186/s13750-020-0186-y

Maxwell, S. L., Butt, N., Maron, M., McAlpine, C. A., Chapman, S., Ullmann, A., Segan, D. B., & Watson, J. E. M. (2019). Conservation implications of ecological responses to extreme weather and climate events. *Diversity and Distributions*, *25*, 613–625. https://doi.org/10.1111/ddi.12878

McManamay, R. A., Orth, D. J., Kauffman, J., & Davis, M. M. (2013). A database andmeta-analysis of ecological responses to stream flow in the South Atlantic region. *Southeastern Naturalist*, *12*(m5), 1–36.

Murchie, K. J., Hair, K. P. E., Pullen, C. E., Redpath, T. D., Stephens, H. R., & Cooke, S. J. (2008). Fish response to modified flow regimes in regulated rivers: Research methods, effects and opportunities. *River Research and Applications*, *24*(2), 197–217. https://doi.org/10.1002/rra.1058

Norton, S. B., Webb, J. A., Schofield, K. A., Nichols, S. J., Ogden, R., Bennett, M., Lee, S., Ridley, C., Gibbs, D., & Collins, A. (2018). Timely delivery of scientific knowledge for environmental management: A freshwater science initiative. *Freshwater Science*, *37*(2), 205–207.

NPS. (2021). *Vital signs monitoring*. National Park Service. https://www.nps.gov/subjects/protectingwater/vital-signs-monitoring.htm

Nussbaumer-Streit, B., Klerings, I., Dobrescu, A. I., Persad, E., Stevens, A., Garritty, C., Kamel, C., Affengruber, L., King, V. J., & Gartlehner, G. (2020). Excluding non-English publications from evidence-syntheses did not change conclusions: A meta-epidemiological study. *Journal of Clinical Epidemiology*, *118*, 42–54.

Olden, J. D., & Poff, N. L. (2003). Redundancy and the choice of hydrologic indices for characterizing streamflow regimes. *River Research and Applications*, *19*(2), 101–121. https://doi.org/10.1002/rra.700

Paperno, R., & Brodie, R. B. (2004). Effects of environmental variables upon the spatial and temporal structure of a fish community in a small, freshwater tributary of the Indian River lagoon, Florida. *Estuarine, Coastal and Shelf Science*, *61*(2), 229–241.

Piniewski, M., Prudhomme, C., Acreman, M. C., Tylec, L., Oglęcki, P., & Okruszko, T. (2017). Responses of fish and invertebrates to floods and droughts in Europe. *Ecohydrology*, *10*(1), e1793. https://doi.org/10.1002/eco.1793
Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegaard, K. L., Richter, B. D., Sparks, R. E., & Stromberg, J. C. (1997). The natural flow regime. *BioScience*, *47*(11), 769–784. https://doi.org/10.2307/1313099
Poff, N. L., & Ward, J. V. (1989). Implications of streamflow variability and predictability for lotic community structure: A regional analysis of streamflow patterns. *Canadian Journal of Fisheries and Aquatic Sciences*, *46*, 1805–1818.
Poff, N. L., & Zimmerman, J. K. H. (2010). Ecological responses to altered flow regimes: A literature review to inform the science and management of environmental flows. *Freshwater Biology*, *55*(1), 194–205. https://doi.org/10.1111/j.1365-2427.2009.02272.x
R Core Team. (2021). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.r-project.org/
Rodhouse, T. J., Sergeant, C. J., & Schweiger, E. W. (2016). Ecological monitoring and evidence-based decision-making in America's National Parks: Highlights of the special feature. *Ecosphere*, *7*, e01608. https://doi.org/10.1002/ecs2.1608
Rohatgi, A. (2019). *WebPlotDigitalizer (v4.2): HTML5 based online tool to extract numerical data from plot images* [online]. https://automeris.io/WebPlotDigitizer/
Rohr, J. M., Meiners, S. J., Thomas, T. D., & Colombo, R. E. (2021). Recovery of riverine fish assemblages after anthropogenic disturbances. *Ecosphere*, *12*(4), e03459. https://doi.org/10.1002/ecs2.3459
Rosenberg, M. S. (2013). Moment and least-squares based approaches to meta-analytic inference. In J. Koricheva, J. Gurevitch, & K. Mengersen (Eds.), *Handbook of meta-analysis in ecology and evolution* (pp. 108–124). Princeton University Press.
Rytwinski, T., Cooke, S. J., Taylor, J. J., Roche, D. G., Smith, P. A., Mitchell, G. W., Smokorowski, K. E., Prior, K. A., & Bennett, J. R. (2021). Acting in the face of evidentiary ambiguity, bias, and absence arising from systematic reviews in applied environmental science. *Science of the Total Environment*, *775*, 145122. https://doi.org/10.1016/j.scitotenv.2021.145122
Rytwinski, T., Harper, M., Taylor, J. J., Bennett, J. R., Donaldson, L. A., Smokorowski, K. E., Clarke, K., Bradford, M. J., Ghamry, H., Olden, J. D., Boisclair, D., & Cooke, S. J. (2020). What are the effects of flow-regime changes on fish productivity in temperate regions? A systematic map. *Environmental Evidence*, *9*(7), 1–26. https://doi.org/10.1186/s13750-020-00190-z
Rytwinski, T., Lin, H.-Y., Harper, M., Smokorowski, K. E., Smith, A., Reid, J. L., Taylor, J. J., Birnie-Gauvin, K., Bradford, M. J., Crossman, J. A., Kavanagh, R., Lapointe, N. W. R., Turgeon, K., & Cooke, S. J. (2022a). ROSES systematic review checklist for 'How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review'. *Figshare*, https://doi.org/10.6084/m9.figshare.21770417.v1
Rytwinski, T., Lin, H.-Y., Harper, M., Smokorowski, K. E., Smith, A., Reid, J. L., Taylor, J. J., Birnie-Gauvin, K., Bradford, M. J., Crossman, J. A., Kavanagh, R., Lapointe, N. W. R., Turgeon, K., & Cooke, S. J. (2022b). Data-extraction sheet for 'How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review'. *Figshare*, https://doi.org/10.6084/m9.figshare.21770123.v1
Rytwinski, T., Lin, H.-Y., Harper, M., Smokorowski, K. E., Smith, A., Reid, J. L., Taylor, J. J., Birnie-Gauvin, K., Bradford, M. J., Crossman, J. A., Kavanagh, R., Lapointe, N. W. R., Turgeon, K., & Cooke, S. J. (2022c). Study validity assessment results for 'How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review'. *Figshare*, https://doi.org/10.6084/m9.figshare.21770465.v1
Rytwinski, T., Lin, H.-Y., Harper, M., Smokorowski, K. E., Smith, A., Reid, J. L., Taylor, J. J., Birnie-Gauvin, K., Bradford, M. J., Crossman, J. A., Kavanagh, R., Lapointe, N. W. R., Turgeon, K., & Cooke, S. J. (2022d). Quantitative synthesis database for 'How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review'. *Figshare*, https://doi.org/10.6084/m9.figshare.21770321.v1
Smakhtin, V. U. (2001). Low flow hydrology: A review. *Journal of Hydrology*, *240*, 147–186.
Tanner-Smith, E. E., & Tipton, E. (2014). Robust variance estimation with dependent effect sizes: Practical considerations including a software tutorial in Stata and SPSS. *Research Synthesis Methods*, *5*, 13–30.
Tanner-Smith, E. E., Tipton, E., & Polanin, J. R. (2016). Handling complex meta-analytic data structures using robust variance estimates: A tutorial in R. *Journal of Developmental and Life-Course Criminology*, *2*, 85–112.
Tipton, E. (2013). Robust variance estimation in meta-regression with binary dependent effects. *Research Synthesis Methods*, *4*, 169–187.
Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2013). Three level meta-analysis of dependent effect sizes. *Behavior Research Methods*, *45*, 576–594.
Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*, 1–48. https://doi.org/10.18637/jss.v036.i03
Vittinghoff, E., Shiboski, S. C., Glidden, D. V., & McCulloch, C. E. (2005). *Regression methods in biostatistics*. Springer.
Wickham, H. (2016). *ggplot2: Elegant graphics for data analysis*. Springer-Verlag. ISBN 978–3–319-24277-4. https://ggplot2.tidyverse.org
WWF/TNC. (2019). *Major habitat types* [online]. Freshwater Ecoregions of the World. https://www.feow.org/global-maps/major-habitat-types

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Supporting Information 1. Search strategy and results. Provides a description of the search strategy and results of the literature searches. For each source, we provided full details on the search date(s), search strings used, search settings and restrictions, and subscriptions (if applicable), and the number of returns. We also describe additional screening consistency check methods and provide definitions of terms used throughout the systematic review.
Supporting Information 2. Excluded articles. List of articles excluded on the basis of full-text assessment or data extraction and reasons for exclusion. Separate lists of articles excluded at the full-text or data extraction review stage, articles that were unobtainable and relevant reviews.
Supporting Information 3. Study validity assessment description. Provides further description of the study validity assessments, along with the critical appraisal tool, and consistency check details.
Supporting Information 4. Data extraction considerations. Provides a description of further data extraction considerations, including intervention definitions, and details of consistency checks.
Supporting Information 5. Data preparation and additional calculations for quantitative synthesis. Provides a description of preparation for quantitative synthesis in relation to variance imputation and reducing multiple effect size estimates from the same study, as well as further details of quantitative synthesis.

Supporting Information 6. Correlation analyses of moderators (Pearson's χ^2). Contains results of contingency analysis for independence of moderators.

Supporting Information 7. Review descriptive statistics. Provides a ROSES flow diagram of inclusion/exclusion process, as well as further descriptions of the evidence base.

Supporting Information 8. Fish species list. Includes all family, genera and species included in the narrative synthesis.

Supporting Information 9. Formal meta-analyses. Includes global meta-analyses, publication bias, sensitivity analyses, and summary results from effects of moderator investigations for formal meta-analyses.

Supporting Information 10. Less formal meta-analyses. Includes results from global meta-analyses including one-sample *t*-tests, and moderator analysis using one-way ANOVAs and independent-samples *t*-tests for the less formal meta-analyses.

Supporting Information 11. Flow-ecological relationship analyses. Includes visual plots and Spearman rank correlations from relationships between percent changes in flow magnitude and percent changes in fish responses for natural flow variation studies.

Supporting Information 12. Additional summary findings. Includes additional summary findings from quantitative analyses presented in the discussion.

How to cite this article: Rytwinski, T., Lin, H.-Y., Harper, M., Smokorowski, K. E., Smith, A., Reid, J. L., Taylor, J. J., Birnie-Gauvin, K., Bradford, M. J., Crossman, J. A., Kavanagh, R., Lapointe, N. W. R., Turgeon, K., & Cooke, S. J. (2023). How do natural changes in flow magnitude affect fish abundance and biomass in temperate regions? A systematic review. *Ecological Solutions and Evidence*, *4*, e12216. https://doi.org/10.1002/2688-8319.12216