Position Paper

Capturing episodic impacts of environmental signals

M. Mendiolar [a],*, J.A. Filar [a], W.-H. Yang [a], S. Leahy [b], A.J. Courtney [c]

[a] School of Mathematics and Physics, University of Queensland, St Lucia QLD 4072, Australia
[b] Queensland Department of Agriculture and Fisheries, Agri-Science Queensland, Northern Fisheries Centre, Cairns, QLD 4870, Australia
[c] Queensland Department of Agriculture and Fisheries, Level B1, Ecosciences Precinct, Joe Baker St., Dutton Park QLD 4102, Australia

ARTICLE INFO

ABSTRACT

Environmental scientists frequently rely on time series of explanatory variables to explain their impact on an important response variable. However, sometimes, researchers are less interested in raw observations of an explanatory variable than in derived indices induced by episodes embedded in its time series. Often these episodes are intermittent, occur within a specific limited memory, persist for varying durations, at varying levels of intensity, and overlap important periods with respect to the response variable. We develop a generic, parametrised, family of weighted indices extracted from an environmental signal called IMPIT indices. To facilitate their construction and calibration, we develop a user-friendly app in Shiny R referred to as IMPIT−a. We construct examples of IMPIT indices extracted from the Southern Oscillation Index and sea surface temperature signals. We illustrate their applications to two fished species in Queensland waters (i.e., snapper and saucer scallop) and wheat yield in New South Wales.

1. Introduction

In many environmental studies, researchers rely on the time series of one or more relevant explanatory variables to explain variation in the response variable of interest. Frequently, the explanatory variables can be used directly, or after relatively minor modifications (e.g., transformations, lags). In such cases the essential information needed to answer the question of scientific interest is clearly reflected in the raw signal data.

However, in some environmental applications, researchers are especially interested in indices derived from certain intermittent episodes embedded in the time series of an explanatory variable. Most of the time they may be interested only in episodes occurring within a specified, limited, memory persisting for long enough, at sufficient intensity and occurring in sufficiently timely manner with respect to the response variable. In our context, this limited memory will be a period of fixed and uninterrupted duration in the past, with respect to the current observation time.

Consider, for example, the time series of the Southern Oscillation Index (SOI), a signal that is widely utilised in environmental sciences. SOI indicates the magnitude and direction of the El Niño southern oscillation (ENSO), in which El Niño (sufficiently negative and sustained SOI values) typically results in hotter and drier than average conditions in Australia, while La Niña (sufficiently positive and sustained SOI values) is associated with cooler and wetter than average conditions in Australia. Similar considerations would apply in the case of marine

heatwaves (MHWs) (Barbeaux et al., 2020; Yao and Wang, 2021), and associated episodes that would be extracted from the time series of Sea Surface Temperature (SST). However, high/low SOI value episodes and marine heatwaves occur intermittently, persist for varying duration, and differ in intensity. Furthermore, if the research focuses on their influence on the harvest yield, we may well wish to identify those episodes that overlap key life history stages of that species or the crop. For fisheries, the spawning season is clearly a sensitive life stage. In the agricultural context, we could consider the pre-sowing, sowing or flowering seasons.

Moreover, if the expert domain knowledge indicates that there is only a limited time horizon over which the analysis is meaningful. For simplicity, we call that horizon a *memory*. This immediately leads to the problem of appropriately differentiating between episodes that occur more or less recently, within that memory, which we refer to as *recency*.

In this paper, we develop a generic, parametrised family of weighted indices extracted from observations of an environmental signal, on the basis of intermittent episodes of interest embedded in the signal. These weighted indices can be calibrated to capture the relative influence of intermittency, memory, persistence, intensity and timing of the underlying episodes. Hence, we name them *IMPIT indices*. To facilitate ease of construction and calibration of IMPIT indices, we develop a user-friendly open source app called IMPIT−a.

We illustrate the effectiveness of the design and calibration of IMPIT indices with fishery catch rate data from two fished species in

---

Queensland, snapper (*Chrysophrys auratus*) and saucer scallop (*Ylistrum balloti*), and agricultural yield data of wheat production in New South Wales. We consider correlations between certain environmental variables and standardised catch per unit effort (SCPUE) and yield per hectare as response variables. We show that highly significant correlations arise between these and IMPIT indices extracted from SOI and/or SST signals, despite analyses indicating no significant (or only weak) correlation between these response variables and the raw signal data.

To the best of our knowledge, prior studies of discrete, intermittent episodes in the environmental and ecosystem context are very application specific. See, for instance, Brad Adams et al. (2003), Ensminger et al. (2004), Unal et al. (2013), Bakun (2014), Strydom et al. (2020) and Bellanthudawa and Chang (2022). By contrast, we construct a generic parametrised family of indices (i.e., IMPIT) induced by episodes embedded in the time series of an environmental signal. Moreover, we propose a systematic parametric search for associations between these indices and seemingly exogenous response variables, such as SCPUE of harvested fish species and yield per hectare of crop production. Importantly, the proposed IMPIT indices can be fine-tuned to capture a wide range of cumulative effects of these discrete episodes.

Since discrete intermittent episodes are embedded in the time series of an environmental signal – for the sake of completeness – we next briefly differentiate our approach from a large body of existing time series and signal processing techniques. The majority of time series methods focus on forecasting episodes of interest. For instance, the well-known Croston method (see Croston (1972)) was designed for forecasting intermittent demands. However, it is known that for many intermittent episodes (e.g., earthquakes, or El Niño events) such forecasting can be very challenging (Rundle et al., 2021; Ham et al., 2019; Glantz, 2015; Ludescher et al., 2014).

Sophisticated signal processing techniques are also used to extract features of signals and build indices accordingly. Popular among these are techniques exploiting Empirical Orthogonal Functions (EOFs), Fourier transforms, and wavelet transforms (Jolliffe, 2011; Hannachi et al., 2007; Thomson and Emery, 2014). Specifically, several climate indices are built on EOFs such as the Arctic Oscillation (Thompson and Wallace, 1998) and RMM1 and RMM2 indices for monitoring the Madden–Julian Oscillation (Wheeler and Hendon, 2004). Moreover, the features extracted by these methods are used as explanatory variables in regression analyses to answer scientific inquiries. For example, Yang et al. (2013) utilised the EOFs and short-time Fourier transform together to reveal the relationship between spawning success of shovelnose sturgeon in the Lower Missouri River and river depth and water temperature using fish tracking data.

An important distinction between those approaches and this study is that we are not attempting to forecast the underlying intermittent episodes. Instead, we construct a unified parametrised family of indices induced by episodes embedded in an environmental signal. We are not using signal processing methods to automatically identify the discrete intermittent influential episodes that we are working with. These are either assumed to be already well defined (such as, marine heatwaves) or are natural (such as, threshold crossing episodes). Note that IMPIT indices in this study are constructed in the natural time domain rather than the frequency domain.

Moreover, we propose a systematic parametric search for associations between the IMPIT indices and important seemingly exogenous response variables, such as SCPUE of harvested fish species and yield per hectare of crop production. Thus, the proposed IMPIT indices are driven by features of these discrete but intermittent episodes occurring within the observed signals (such as SOI or SST). They also account for the propagation effects of the intermittent episodes. The proposed search process prepares them to be plausible indicators or explanatory variables to identify and unravel their possible impacts on relevant response variables.

We also point out that the IMPIT indices are quite generic and hence have potential application in a wide class. The indices and their calibration should be seen as another tool of exploratory data analyses. Indeed, they have recently been used in a real-world application (Filar et al., 2021). Finally, introduction of IMPIT indices is also timely because of recent and ongoing changes in the intensity and frequency of extreme weather and climate events (Ummenhofer and Meehl, 2017; Frölicher et al., 2018; Oliver, 2019).

The paper is structured as follows. Section 2 describes the methodology used to develop IMPIT indices. Section 3 illustrates their design and calibration for two widely used environmental signals and study their impact in a fishery and agricultural context. Section 4 describes the app developed and its visualisation capabilities. Finally, Section 5 contains a brief summary and discussion.

2. Indices capturing Intermittence, Memory, Persistence, Intensity and Timing (IMPIT)

We consider a time series of an environmental signal denoted by $\mathbf{X} = \{X_t\}$ with time $t = T, \ldots, 1$, where X_T represents the earliest observation and X_1 the latest (most recent) observation. The units of time should be consistent among signal points and can be hours, days, years, etc. We define the *memory* as a period of fixed and uninterrupted duration and denote by m its length. With the memory fixed, we can extract a sequence of environmental episodes where each episode satisfies certain characterising conditions. The impact of these episodes on a response variable may depend on their intermittence, persistence, intensity and timing. Timing refers to the position of the episode in the time history with respect to any special aspects of the studied phenomenon (e.g., spawning or flowering seasons). Whatever their definition, the essential requirement of the episodes is that there is some number $K \leq m$ of them and, that the entire remembered history of the signal $\mathbf{X} = \{X_1, X_2, \ldots, X_m\}$ contains these K distinct episodes $E_k = \{X_{s_k}, X_{s_k+1}, \ldots, X_{s_k+n_k-1}\}$, $k = 1, \ldots, K$, where n_k denotes the length of the kth episode and s_k is its starting time location.[1] Namely,

$$\bigcup_{k=1}^{K} [E_k] \subseteq \mathbf{X}. \tag{1}$$

Note that while the elements comprising the episode E_k would most often be consecutive, this is not a requirement. Thus, it is possible that $X_{s_k+1} \neq X_{s_k+1}$. However, in most of our illustrations the episodes indeed consist of consecutive observations and in such a case, they will be denoted by $E_k = \{X_{s_k}, X_{s_k+1}, \ldots, X_{s_k+n_k-1}\}$, $k = 1, \ldots, K$.

The intensity, $I(E_k)$, of the episode E_k can be any function of the observations comprising E_k, as explained in the next Section 2.1. Collectively, the set of intensity values will be denoted by the symbol I. Furthermore, we need to capture the relative importance of the episodes $E_1, \ldots, E_K$. This will be done with the help of relative importance weights, collectively denoted by $\mathbf{w}$. Note that the intensity is a measure of the strength of each episode, whereas importance weights are intended to capture the relative impact of multiple episodes on the phenomenon of interest.

With respect to the collection of episodes in expression (1), we define a family of associated IMPIT indices extracted from the signal of the environmental variable $\mathbf{X}$. This family is of the form

$$X(I, \mathbf{w}) = \sum_{k=1}^{K} w(E_k) I(E_k), \tag{2}$$

where $w(E_k)$ and $I(E_k)$ denote the importance weight and the intensity measures associated with episode E_k, respectively. We note that episode related parameters s_k, n_k and their relationship to the number of episodes K influence Eq. (2). However, to a large extent their values will be predetermined by expert knowledge used to identify likely relevant intermittent episodes (e.g., see definition of marine

---

[1] Even the requirement $K \leq m$ could be relaxed but we would expect it to hold in most applications.
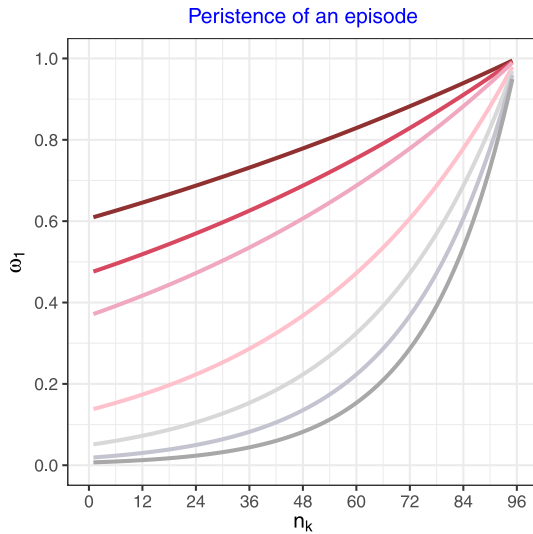
Fig. 1. Examples of the shape of $w_1(n_k, m)$ for different dampening parameters a, where n_k is measured in months. Colour code: $a = 0.50$ (—), 0.75 (—), 1.00 (—), 2.00 (—), 3.00 (—), 4.00 (—) and 5.00 (—).

heatwaves in Section 3.1). IMPIT indices are intended to be used in exploratory data analysis aimed at determining whether a class of intermittent environmental episodes are relevant to some phenomenon of interest. Hence, for each set of importance weights and intensities $X(I, w)$ is assigned a numerical value at the current time t. However, if we considered such an index over a range of t values, we would be constructing a time series $\{X_t(I, w)\}_{t \geq 0}$ which may not be stationary, especially if that were also the case for the original environmental data series. The question of stationarity would properly be addressed in follow-up confirmatory studies invoking techniques for testing and accounting for non-stationarity of time series (e.g., see Hamilton (1994) and Tsay (2005)).

2.1. Intensity of episodes

An intensity function $I(E_k) = I(X_{k_1}, X_{k_2}, \ldots, X_{k_{n_k}})$ should map observed values in the episode onto a number capturing the intensity or strength of that episode. Several natural candidates could be considered for such functions including, the mean of the observations $\{X_{k_1}, X_{k_2}, \ldots, X_{k_{n_k}}\}$

$$I(E_k) = \bar{I}(E_k) = \frac{1}{n_k} \sum_{j=1}^{n_k} X_{k_j}, \tag{3}$$

or the logarithm of their sum

$$I(E_k) = I^l(E_k) = \log\left(\sum_{j=1}^{n_k} X_{k_j}\right). \tag{4}$$

Other candidates such as, the median, geometric mean, signal-to-noise ratio could also be considered (ensuring that each is well defined).

2.2. Weights of episodes

An essential ingredient in the construction of an IMPIT index (i.e., Eq. (2)) is the relative importance weight of each episode (i.e., $w(E_k)$ for each k). We postulate that the weights $w(E_k)$ are all non negative numbers lying in the interval $[0, 1]$, with 1 corresponding to episodes viewed as most important to the studied phenomenon in the aggregate index $X(I, \boldsymbol{w})$. There are, at least, the following three approaches to the construction of these weights:

(1) Multiplicative, product form, of these weights given by

$$w(E_k) = w_1(n_k, m) w_2(s_k, m) w_3(E_k), \tag{5}$$

where $w_1, w_2, w_3 \in [0, 1]$ are intended to capture the importance of *persistence*, *recency* and *timing* of the episode E_k of length n_k and starting position s_k in the memory of length m. Any of these weights could also be set to 0 or 1 depending on exogenous information available. In absence of deeper understanding of the persistence, recency and/or memory aspects, $w_1(n_k, m)$, $w_2(s_k, m)$ and $w_3(E_k)$ can be set equal to 1 for every k.

(2) Construction where the weights can be used purely as a technical tool to achieve a desired form of the $X(I, \boldsymbol{w})$ indices. This is illustrated in the remark, below.

(3) Construction based on expert domain knowledge. This may depend on either more detailed understanding of the phenomena defining the episodes, or on the intended target response variable that a study aims to explain (at least partially) with the help of the $X(I, \boldsymbol{w})$ indices.[2]

Remark. Note that most standard statistical indices can be easily recovered within the above $X(I, \boldsymbol{w})$ family. For instance, if we wanted the mean of observations $\{X_1, X_2, \ldots, X_{12}\}$, all we need to do is define $E_k = X_k$ for each k, the memory $m = 12$, the intensity function to be $I(E_k) = \frac{X_k}{12}$ and all the weights to be $w(E_k) = 1$, for each k. Similarly, other indices such as medians, moving averages, or coefficient of variation can be naturally constructed in the above form.

In the remainder of this subsection, we propose certain specific algebraic forms for the weights in (5) which require users to calibrate only a small number of parameters while offering considerable freedom in choosing their shape. Exponential form maps have natural appeal as they frequently arise in a wide range of growth/decay models. Also in fishery applications that motivated this study, survival rates are often expressed as exponentials of the sum of natural and fishing mortalities (e.g., see Quinn and Deriso (1999)). For the sake of consistency, we used exponential type forms in all cases.

2.2.1. Persistence

It is well known that persistence of certain environmental events can have considerable impact. For instance Li and Thompson (2021) state *"Persistence has a key role in the climate impacts of a given temperature event"*. To describe the persistence of an episode we used the following functional form

$$w_1(n_k, m) = \exp\left(-a\left(1 - \frac{n_k}{m}\right)\right), \tag{6}$$

where n_k is the length of episode E_k, m the memory, and $a > 0$ the dampening parameter. For each fixed a, this formula places low weights on short episodes (see Fig. 1). This figure also shows that different values of parameter a control both the slope and the curvature of the parametrised family of convex functions represented by (6).

2.2.2. Recency

Recency of an episode could also play an important role in its impact on the studied phenomena (Deryugina, 2013; Hoffmann et al., 2022). Frequently, but not always, recent episodes can be expected to greater impact than those that happened in more distant past. An obvious indicator of the recency of the episode $E_k = \{X_{s_k}, X_{s_k+1}, \ldots, X_{s_k+n_k-1}\}$ is the ratio of $s_k/m \in [0, 1]$, with high values indicating less recent events. If we use this ratio to design an importance weight that calibrates recency, we could use the following two-step process.

---

[2] We do not discuss this approach in any more detail as our applications are used mainly as illustrations of the generic methodology.
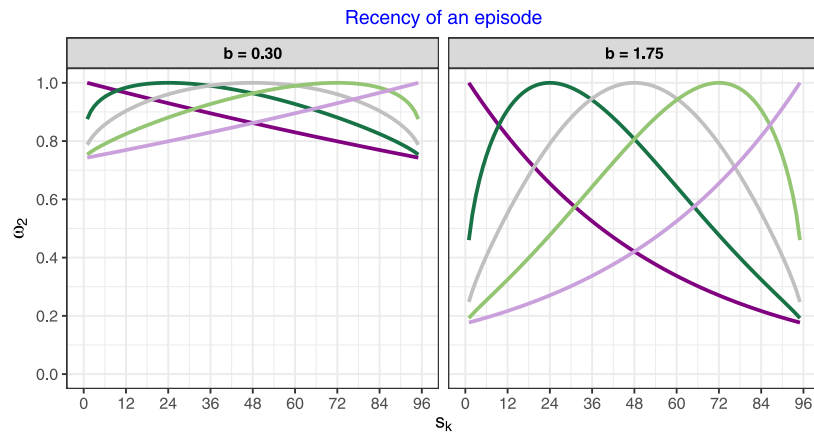
Fig. 2. Examples of the shape of $w_2(s_k, m)$ function when $m = 96$ (months) and parameters c and b take the indicated values. Colour code: $c = 0.00$ (—), 0.25 (—), 0.50 (—), 0.75 (—) and 1.00 (—).

Let us define the function of s_k (starting time of the episode E_k) and m the memory by

$$v(s_k, m) = \lambda \left(\frac{s_k}{m}\right)^c \left(1 - \frac{s_k}{m}\right)^{1-c}, \tag{7}$$

where $0 \leq c \leq 1$, and the scaling parameter $\lambda > 0$ is chosen to ensure that $v(s_k, m) \leq 1$. To capture the relative importance of recency of episode E_k within the period considered we again propose an exponential form

$$w_2(s_k, m) = exp\left[-b\left(1 - v(s_k, m)\right)\right], \tag{8}$$

where $b > 0$. When Eq. (7) is substituted into (8), this ensures that $w_2(s_k, m)$ also range between 0 and 1 and attain a maximum when $v(s_k, m) = 1$. The parameter b can be viewed as a dampening (or accelerating) factor of the rate of decay away from the maximum. For instance, from Fig. 2 we see that $b = 0.3$ dampens the rate of decay of these weights to always remain above 0.7, while $b = 1.75$ permits most of them to drop to below 0.3.

The parameter c can be used to assign more or less importance to episodes that occur earlier or later in the history. From Fig. 2 we can see that when $c = 0$ (dark purple line), episodes starting late in the memory (s_k close to m) obtain $w_2(s_k, m)$ values that are smaller than episodes starting early in that history (s_k close to 1), and conversely when $c = 1$ (light purple line). Note that the peak of $w_2(s_k, m)$ coincides with the value of s_k that maximises $v(s_k, m)$, and hence also $w_2(s_k, m)$. For any fixed c strictly between 0 and 1, $s_k^* = cm$ is the maximiser of $v(s_k, m)$. Since the starting time is an integer, in practice, the peak will be set to occur at the nearest integer to that ratio. When parameters a, b and c are calibrated to optimise a performance measure, researchers may be particularly interested in the ratio $c = \frac{s_k^*}{m}$ as it is an indicator of the most important relative recency, with respect to the memory m.

2.2.3. Timing

Depending on the application, the importance weights of episodes may need to be modulated further by the timing of the occurrence of these episodes with respect to one or more response variables. For instance, in the fishery science application, discussed below, the response variable is the SCPUE, which is a proxy for abundance of the species of interest. If the environmental episodes of IMPIT indices occurred at times overlapping sensitive periods in the species' lifecycle, that information should modulate the values of the indices. For example, temperature can be most influential at critical windows of time in the life history of a focal organism, such as spawning, egg or larval development, seed germination, or crop desiccation pre-harvest. Thus, it is natural to focus on a window of time in a species' life history where it is most susceptible to that environmental condition. Mathematically, such modulation can be modelled in a variety of ways. Perhaps, the

easiest of these is by incorporating a third, multiplicative, importance weight $w_3 \in [0, 1]$, designed to capture the special timing of interest (if any) of the occurrence of the episode E_k.

We shall denote the special timing we are interested in by $\mathcal{T}$. Hence, the overlap with the episode E_k can be denoted by $E_k \cap \mathcal{T} = \mathcal{T}_k$ and its length by τ_k. Now we define the third *timing* weight by

$$w_3(E_k) = 1 - exp\left[-d\left(\frac{\tau_k}{n_k}\right)\right], \tag{9}$$

where the fraction τ_k/n_k is intended to capture how much of the episode is taken by this overlap and d is a dampening parameter, $d \geq 0$.

The left panel of Fig. 3 illustrates a typical overlap of the special timing of the spawning season of snapper with a 2008 La Niña episode. Snapper spawn in aggregations over several months (generally May to October) and synchronise spawning on the lunar cycle (Wortmann et al., 2018). The right panel of Fig. 3 shows the impact of the dampening parameter d on the shape of $w_3(E_k)$, as a function of the ratio τ_k/n_k.

2.3. Periodic seasonality indices

In this case each episode E_k is of equal length $n_k = n$ and they occur with periodic regularity.

$$E_k = \{X_{k_1}, X_{k_2}, \ldots, X_{k_n}\}, \quad k = 1, \ldots, K. \tag{10}$$

where the number K of these episodes is limited by spacings between them and the total memory m under consideration. For instance, consider a memory of, say, $m = 48$ months counting backwards from December of the current year. For episodes capturing the spawning aggregations of snapper (May–October) there would be exactly $K = 4$ such episodes, consisting of $E_1 = \{X_3, X_4, \ldots, X_8\}$, $E_2 = \{X_{15}, X_{16}, \ldots, X_{20}\}$, $E_3 = \{X_{27}, X_{28}, \ldots, X_{32}\}$ and $E_4 = \{X_{39}, X_{40}, \ldots, X_{44}\}$.

2.4. Threshold-crossing indices

These concern situations where episodes of interest are defined by only those observations which exceed or fall below some specified threshold for sufficiently long. Let δ be a threshold and E_k be the kth episode of interest, of duration n_k, be defined by either

$$E_k = E_k^u(\delta, \ell) = \{X_{k_j} | X_{k_j} \geq \delta, j = 1, \ldots, n_k \ \& \ n_k \geq \ell\}, \tag{11}$$

or

$$E_k = E_k^d(\delta, \ell) = \{X_{k_j} | X_{k_j} \leq \delta, j = 1, \ldots, n_k \ \& \ n_k \geq \ell\}, \tag{12}$$

where ℓ denotes the minimum required duration (with $\ell = 1$ being the default). Here, the superscript u denotes up-episodes (above the
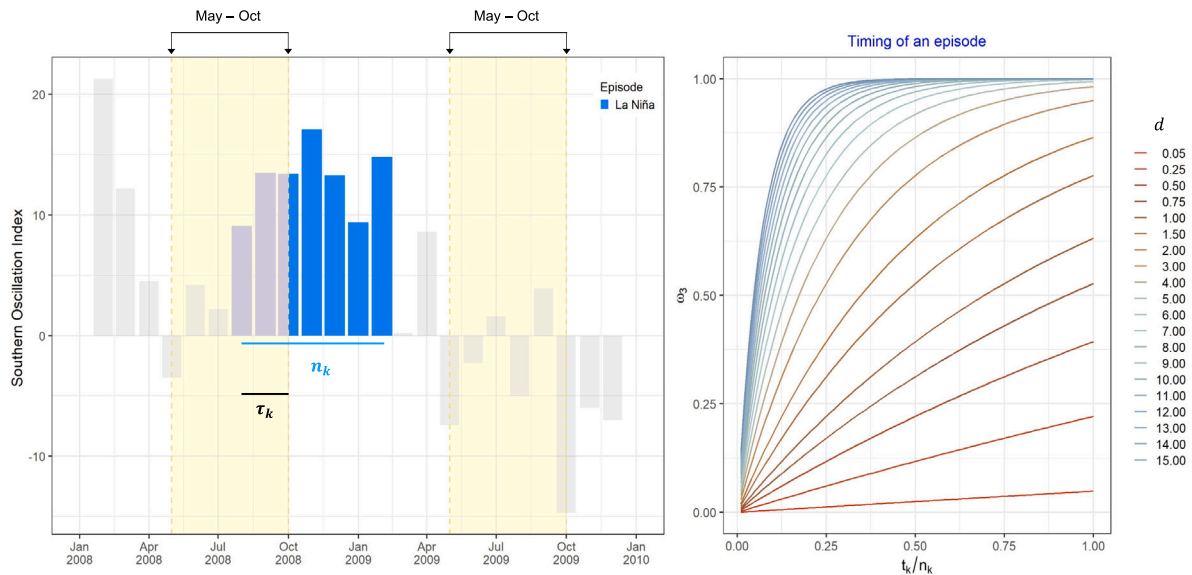
Fig. 3. Left panel: Example of overlapping between La Niña episode during 01/2008-01/2010 and the May–October spawning season for snapper (*Chrysophrys auratus*) in yellow. Y-axis indicates monthly SOI value, with sustained SOI values above 8 for at least 5 consecutive months indicating a La Niña episode. Right panel: Examples of the shape of $w_3(E_k)$ function when $m = 96$. The differences in colour correspond to changes in the parameter d.
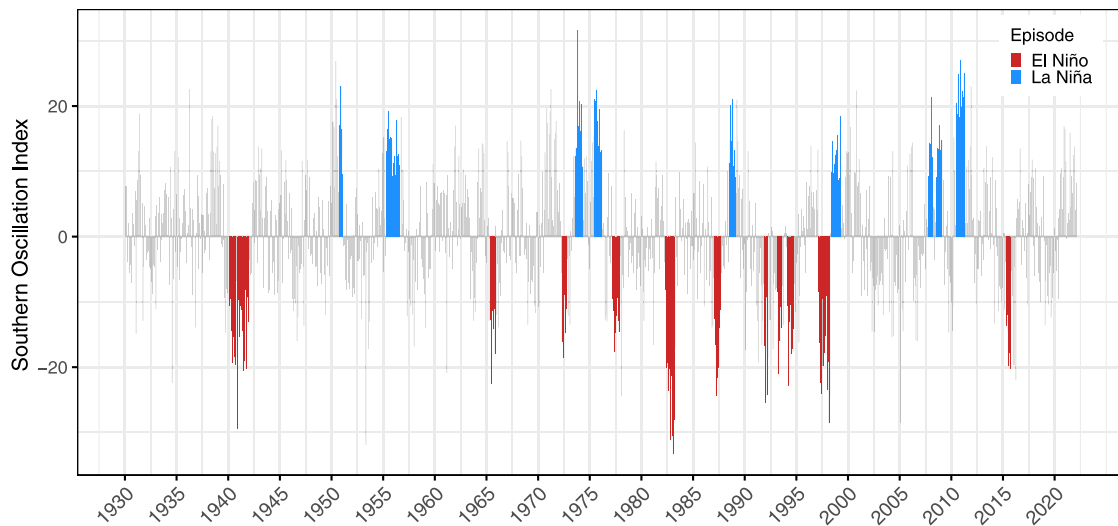

Fig. 4. Time series of the monthly Southern Oscillation Index (SOI) from January 1930 to December 2020. Based on Bureau of Meteorology Australia (2012) definition for El Niño and La Niña episodes, values below -8 and above 8 for at least five consecutive months are shaded red and blue, respectively.

threshold) whereas d denotes down-episodes (below the threshold). The selection of the sub-sequences of the n_k observations X_{k_j} comprising E_k would typically depend on additional, application specific knowledge.

Next, any IMPIT index of the form (2) associated with episodes $E_k^u(\delta, \ell)$ will be called a Super$X^u(\delta, \ell)$ index. Similarly, any IMPIT index associated with episodes $E_k^d(\delta, \ell)$ will be called a Sub$X^d(\delta, \ell)$ index. In the default case of $\ell = 1$, we simplify the above notation to $E_k^u(\delta)$ and Super$X^u(\delta)$ (respectively, $E_k^d(\delta)$ and Sub$X^d(\delta)$). Since increasing ℓ only makes conditions in (11) more restrictive, it follows that the number of $E_k^u(\delta)$ episodes is greater or equal than that of $E_k^u(\delta, \ell)$ episodes, for any $\ell > 1$. Similarly, for (12) and the down episodes.

The extensively studied El Niño and La Niña episodes illustrate this situation. They are extracted from the SOI time series **X**. In particular, according to Bureau of Meteorology Australia (2012), an El Niño episode is simply $E_k^d(-8, 5)$ where the threshold $\delta = -8$ and minimum duration of $\ell = 5$ are used. Analogously, a La Niña episode corresponds to $E_k^u(8, 5)$ (see Fig. 4).

2.5. Calibration via exploration of the parameter space

An underlying contribution of this study is the demonstration that importance weights of the episodes E_k may be calibrated to achieve specific research objectives. In our application, the dual objective is to identify parameter configurations that: (a) achieve high absolute value of the Pearson correlation coefficient between an IMPIT index and a response variable meaningful in fish stock assessment, and (b) are sufficiently stable with respect to small changes in other parameters, especially the memory. We note, however, that other statistical measures of association such as Spearman correlation, or Kendall tau could also be used. In fact, IMPIT$-a$ allows the user these choices. Here we illustrate our method only with the most widely used Pearson correlation.

For each, fixed, memory parameter m, all the weights discussed above are fully characterised by the choice of the parameter vector (a, b, c, d), where $a, b, d \geq 0$ and $c \in [0, 1]$. In most applications, the
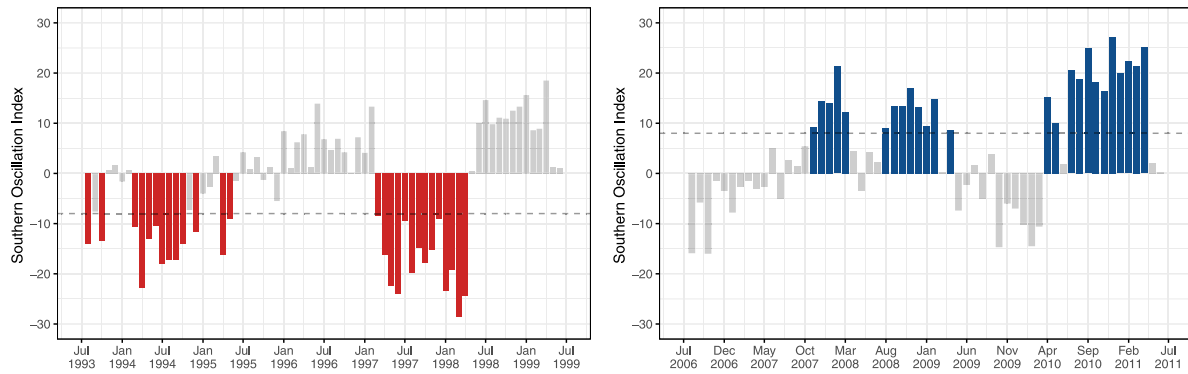
Fig. 5. Monthly SOI values during 07/1993–07/1999 and 07/2006–07/2011. Columns in red and blue correspond to values below and above −8 and 8 thresholds (dashed line) used to illustrate some of the $E^d(-8)$ and $E^u(8)$ episodes used in this study.

parameters a, b and d will also have practical upper bounds. Thus, numerically, the search for an optimal parameter configuration can be carried out by a grid search of a 4−dimensional hyperrectangle. That search could then be repeated for multiple values of the memory parameter m. With respect to objective (a), an approximately optimal configuration $(m^*, a^*, b^*, c^*, d^*)$ could be identified by such an exhaustive search.

However, such a brute-force calibration has some drawbacks. Clearly, it is computationally intensive. Further, the best parameter configuration with respect to (a) may be unstable with respect to (b) and may result in counter-intuitive importance weights, in the context of the application. Working on an assumption that IMPIT indices would be primarily of interest to users with prior expert knowledge of the problem studied, we suggest a structured exploration of the parameter space. Hence, in our illustrative applications, we propose the following *stage-wise explorative calibration* that is simple to implement, preserves the logical order of the design of IMPIT indices, and is likely to identify multiple interesting parameter configurations for further consideration.

At the first stage, for each value of memory m, plots of correlation coefficient on the vertical axis (y-axis) and a values on the horizontal axis (x-axis) are generated while setting $w_2 = w_3 = 1$. Based on these plots a candidate value of the parameter a is chosen. This choice determines the w_1 weight and may, but need not necessarily, correspond to the maximum of the absolute value of the relevant correlation.[3] At the second stage, with the already chosen $a-$ value and $w_3 = 1$ fixed, plots of correlation coefficient on the vertical axis (y-axis) and c values on the horizontal axis (x-axis) are generated, for alternative pairs of m and b parameters. These plots can be visualised in a composite "map", in the parameter space, in order to explore and find candidate parameter configurations that meet both objectives (a) and (b). At the third stage, we repeat the preceding while considering only episodes that overlap a special timing season. This stage consists of generating composite maps for alternative values of d and selecting promising candidate configurations.

We note that even in the absence of a response variable suggested by a specific application, IMPIT indices can still be optimally calibrated with respect to their own desirable characteristics. For instance, for each fixed m, configuration (a, b, c, d) can be chosen to maximise the R^2 statistic of the corresponding IMPIT index signal when fitted to a prescribed model (e.g., a linear model, or a power law). Moreover, the strength of the influence can be captured by a wide range of quantitative measures (e.g., goodness-of-fit or efficiency measures), see for instance Krause et al. (2005). The choice of such a measure, while important, is not the focus of this study.

---

[3] Stability of the correlation coefficient with respect to changes in a is also taken into consideration. If the maximisation criterion does not meaningfully differentiate among $a-$ values, expert suggested value (in the range) can be selected.

3. Case studies using fishery and agricultural data

We illustrate the benefits of the IMPIT indices' design and calibration in the context of the impact of selected environmental signals on two fished species in Queensland, snapper (*Chrysophrys auratus*) and saucer scallop (*Ylistrum balloti*), and yield data of wheat production from NSW. The environmental signals are the Southern Oscillation Index (SOI) and sea surface temperature (SST). The response variables are the time series of SCPUE and yield per hectare of wheat. These signals were selected because studies have demonstrated their strong correlations with fish catch rates (Joll and Caputi, 1995; Lenanton et al., 2009; Caputi et al., 2019) and wheat production (Gutierrez, 2017; Wan et al., 2022), particularly in Australia (Rimmington and Nicholls, 1993; Yuan and Yamagata, 2015; Zheng et al., 2018; Potgieter et al., 2002). A description of the data acquisition and episodes considered in these case studies is given in Section 3.1 below.

3.1. Data acquisition

Environmental data and two types of intermittent episodes

Monthly SOI values were obtained from the Australian Bureau of Meteorology (BoM). The SOI is computed from the variations of monthly mean sea level pressure difference between Tahiti and Darwin (Chowdhury and Beecham, 2010). Positive SOI values are generally associated with a La Niña pattern in the central and eastern equatorial Pacific and above-average winter/spring rainfall for Australia, particularly across the east and north. Negative SOI values are associated with El Niño conditions and lower than average winter/spring rainfall over much of eastern Australia (Bureau of Meteorology Australia, 2012).

Let **X** denote the underlying SOI signal. We illustrate a class of threshold crossing IMPIT indices extracted from **X**. In particular, we consider up-episodes $E_k^u(8)$ and down-episodes $E_k^d(-8)$ corresponding to $\text{Super} X^u(8)$ and $\text{Sub} X^d(-8)$ indices, requiring calibration. The choice of the thresholds $\delta = 8$ and $\delta = -8$ was guided by their use in a common definition of La Niña and El Niño episodes, respectively (Bureau of Meteorology Australia, 2012). For example, Fig. 5 displays five $E_k^u(8)$ episodes during 07/2006–07/2011 (right panel) and six $E_k^u(8)$ episodes during 07/1993–07/1999 (left panel). Note that in the case of the up-episodes $E_k^u(8)$ only three of these correspond to La Niña episodes lasting at least five months, since the single month episode in April 2009 and then the two monthly episode in April–May 2010 do not qualify as La Niña episodes. Similarly, in the case of the down-episodes $E_k^d(-8)$ only two of the six correspond to El Niño episodes.

The SST data was sourced from the Integrated Marine Observing System (IMOS) database with 6−day average, night time capture ("ghrsst_L3S_6d_ngt") over a period of approximately 28 years (01/04/1992 to 31/12/2019). The 6−day average night time data set was used to avoid any daytime temperature artefacts (e.g. sun glint). Also, 6−day averaging fills gaps that are otherwise present in the daily
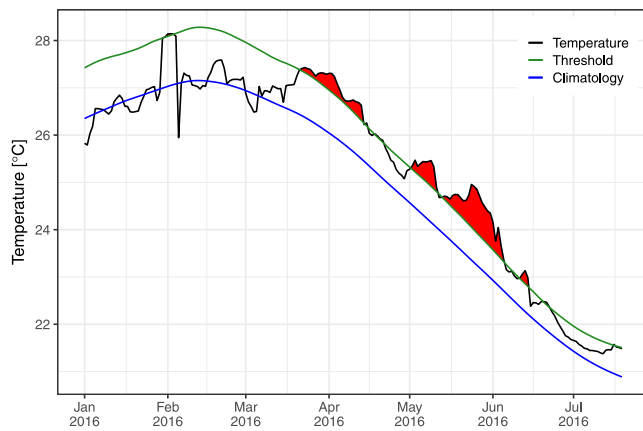
Fig. 6. The SST climatology (blue), 90th percentile MHW threshold (green), and SST time series (black) for each MHW at south of 22° S to Hervey Bay. The red filled areas indicate the period of time associated with the identified MHWs.

time series as a result of cloud cover blocking satellite view of the ocean surface. The spatial distribution of SST used in the analysis corresponded to the areas of the scallop fishery described in O'Neill et al. (2020).

Marine heatwaves (MHWs) are anomalous ocean temperature events, identified by extremely warm SST that persists for days to months (Hobday et al., 2016). They can be caused by a mix of atmospheric forcing and oceanographic conditions and depend on location and season. For instance, a heatwave that affects coral reefs in warmer waters will have higher temperatures than one that affects kelp forests in cooler waters.

We followed the hierarchical definition of MHW episodes according to Hobday et al. (2016). This definition, takes into account the following key features: (i) Anomalously warm temperatures with respect to a baseline average temperature over a period of 30 years, and a high percentile threshold of 90; (ii) Prolonged events persisting for at least five days; (iii) Discrete events appearing with sufficient separation between successive events.

The corresponding MHW events allow us to form episodes E_k and IMPIT indices of the form considered in Section 2. For example, Fig. 6 displays the SST climatology, 90th percentile MHW threshold and SST time series for a typical MHW episode at a location off the Queensland coast (south of 22° S in Hervey Bay). The red area between the black and green curves identifies the episode. The intensity function we have selected is the "mean temperature anomaly during the MHW".

Fishery data

For snapper, standardised catch rate data were provided by Fishery Queensland. The top panel in Fig. 7 displays the annual time series of SCPUE for snapper from the Queensland commercial line fishery. Overall, we observe a decreasing trend in catch rates across most of the dataset, following a peak in SCPUE in 1989. In the case of scallop, catch rates were obtained from Wortmann et al. (2020). Catch rates from November were chosen for use in the analysis because the fishing season has traditionally commenced in November, scallop catch rates generally peak at this time, and there is negligible fishing effort from May–October (O'Neill et al., 2020). The November catch rate has suffered a sharp decline in recent years. In Fig. 7b we observe a marked decline from 2012 to 2017 with a partial recovery in 2018. Note that there are no November catch rate data after the fishery was closed in 2018.

Agricultural data

Data on wheat production in NSW, from 1974 to 2021, were obtained from the Australian Government, Department of Agriculture, Fisheries and Forestry (DAFF) (https://www.agriculture.gov.au/abares/research-topics/agricultural-outlook/data#australian-crop-report-data). We have considered the wheat production in NSW as this is the main crop grown in the state, which is the second-highest producing State in Australia (Department of Primary Industries NSW, 2007). In Fig. 7c we observe that even though there are many ups and downs, there is no strong trend in either direction. However, the low values in 1982, 1994, 2007 and 2018–2019 coincide with strong and moderate El Niño episodes (http://www.bom.gov.au/climate/history/enso/). There was an overall dry period across eastern and southeastern Australia over these two last years (Wang and Cai, 2020).

3.2. The weak baseline associations

There is a natural expectation that environmental signals influence the abundance and hence harvest of both fish and crops (Aburto-Oropeza et al., 2010; French et al., 2021; Kangas et al., 2022; Potgieter et al., 2002; Wan et al., 2022; Zheng et al., 2018). Indeed, marine heatwaves can have a big impact on coastal fisheries. For example, Caputi et al. (2015) report that adult biomass of saucer scallop and blue swimmer crabs in Western Australia (WA) declined as a result of the MHW of 2010/2011. As a consequence, these commercial fisheries were closed from 2012 to 2016.

Furthermore, there could be a link between MHWs and SOI oscillations (Meynecke et al., 2012; Loughran et al., 2017; Sen Gupta et al., 2020). For instance, the MHW of 2010/2011, which was reported to be the most extreme event (in intensity, extent and duration) ever recorded in WA, was a consequence of a strong Pacific La Niña episode (Caputi et al., 2015; Molony et al., 2021). Further connections between La Niña and El Niño episodes and harvest of fish and crops were discussed in Meynecke et al. (2012), Rimmington and Nicholls (1993) and Gutierrez (2017). In particular, the influence of environmental variables on the abundance of certain Queensland fishery species had been reported in Courtney et al. (2015), O'Neill et al. (2020) and Filar et al. (2021).

It is also widely accepted that El Niño occurrences adversely influence crop yields (Iizumi et al., 2014; Gutierrez, 2017). Specifically, they have negative impact on Australian wheat production (Rimmington and Nicholls, 1993; Yuan and Yamagata, 2015; Zheng et al., 2018; Potgieter et al., 2002).

However, the drivers of the Southern Oscillation phenomenon are complex, and as a result, there are no obvious trends in either the SOI signal, or its association with the catch rates of snapper or NSW wheat yield. The top and bottom right panels of Fig. 8 show that there is no significant linear association between annual catch rates of snapper and wheat yield with the raw, unmodified annual mean SOI index. Still, there appears to be weak but significant negative association between annual mean sea surface temperature and the November catch rates of scallop as can be seen from the middle right panel of the same figure, that is consistent with earlier studies O'Neill et al. (2020). The top and bottom left panels of Fig. 8 indicate absence of linear trend in the mean annual SOI signal and the middle left panel indicates weak upward trend in the mean annual SST signal.

Nevertheless, it is possible that important associations are not revealed because the use of baseline annual means of the raw SOI and SST failed to account for the importance of persistence, recency, and the timing of the relevant intermittent episodes. Indeed, analyses of the remainder of this section illustrate, with the help of IMPIT indices, that this might be the case. These analyses should be viewed as exploratory rather than confirmatory.

In the remainder of this section, we demonstrate that IMPIT indices can be calibrated to reveal previously hidden associations between SOI and SST signals and the catch rates of snapper and scallop in QLD and
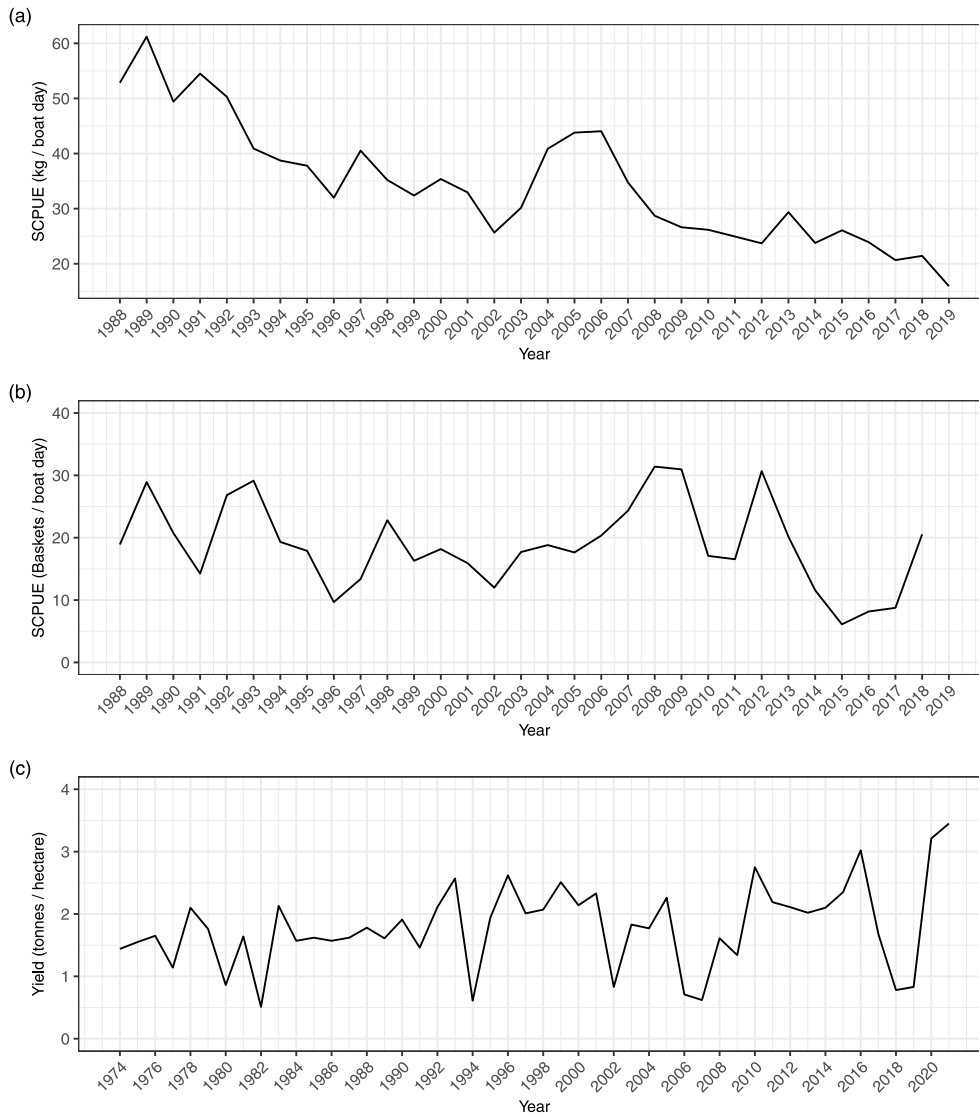
Fig. 7. Time series of standardised catch rates for: (a) snapper (1988–2019) for the Queensland commercial line fishery, (b) saucer scallop in November (1988–2018) for the Queensland trawl fishery and (c) yield per hectare of wheat production in New South Wales, 1974–2021.

wheat yield in NSW. Relationships between the environmental variables and the fishery catch rates and wheat yield data were examined using correlation analyses. In the case of SOI, correlation analyses were based on 31 years (1988–2019) when working with catch rate data and 47 years (1974–2021) when considering yield data. In the case of SST, data were available only since 1993, resulting in a shorter time series.

3.3. SOI threshold-crossing indices and snapper (Chrysophrys auratus)

We investigated linear associations between snapper annual SCPUE and a range of threshold crossing $Super X^u(8)$ IMPIT indices constructed using the methodology of Section 2.4. The intensity of $E_k^u(8)$ episodes in $Super X^u(8)$ indices was computed by Eq. (3).

We implemented the stage-wise explorative calibration described in Section 2.5. Fig. 9 depicts representatives of composite maps obtained at each stage, pruned for ease of display. For instance, we displayed only even values of the memory parameter m which ranged from 1 to 41 years. The value of 41 corresponded to the maximum longevity of snapper in the eastern coast stock reported in Wortmann et al. (2018)

The top panel of Fig. 9 corresponds to the first stage. Setting $w_2 = w_3 = 1$, it displays plots of the correlation coefficient versus the parameter $a \in [0, 5]$ for each value of m. Those plots exhibited a roughly

linear pattern with very small variation. Hence, we chose $a = 2$, which corresponds to the middle pink curve for w_1 in Fig. 1.

At the second stage, while setting $w_3 = 1$ (and $a = 2$) we plotted the correlation coefficient versus the parameter c for alternative pairs of m and b. The third stage considers the overlap of $E_k^u(8)$ episodes with the snapper peak spawning season of June–August and generates composite maps for alternative values of d. The second and third top panels of Fig. 9 display the resulting representative composite maps. The bottom panel shows the map for only $d = 1$, which led to promising candidate configurations. For simplicity, only results for values $b \in \{1, 3, 5\}$ are displayed.

Careful examination of Fig. 9 reveals interesting patterns generated by the stage-wise calibration. In particular, the impact of increasing values of the memory parameter is notable. Moreover, the inclusion of the special (June–August) timing generated the most noticeable differences between stages as it led to the disappearance of all significant negative correlations for values of $m \geq 10$ years.

These patterns may also hint at the dual nature of the impact of high SOI values on snapper catch rates and potentially abundance. Focusing on $E_k^u(8)$ episodes overlapping winter spawning season is more directly related to reproductive success and subsequent fish abundance. Thus, significant positive correlations may be expected here.
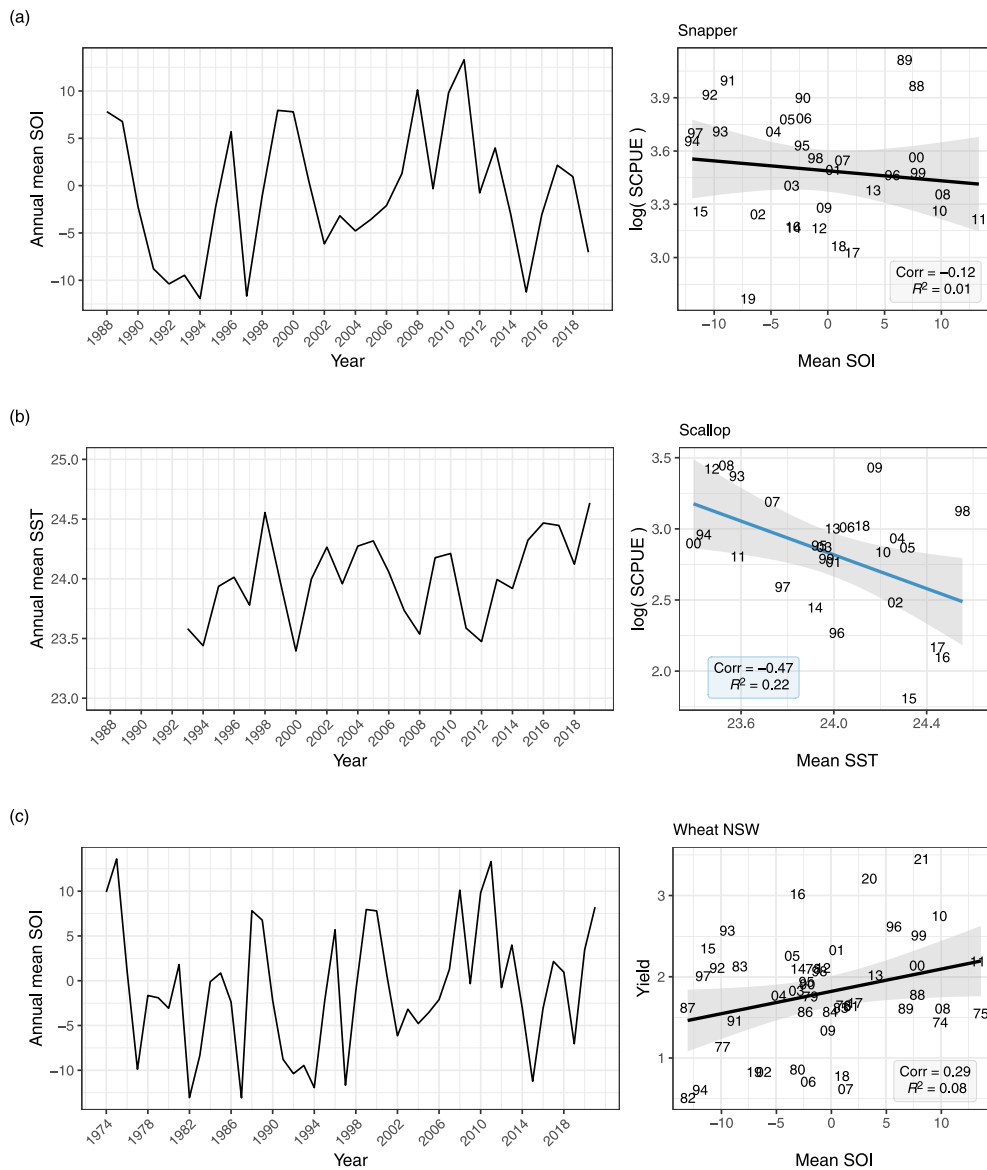
Fig. 8. (a) Time series of annual mean SOI index 1988–2019 (left) and scatterplot between annual mean SOI and annual log-transformed SCPUE of snapper (right). (b) Time series of SST annual mean 1993–2019 in the saucer scallop fishery region (left) and scatterplot between SST annual mean and November log-transformed SCPUE of saucer scallop (right). (c) Time series of annual mean SOI index 1974–2021 (left) and scatterplot between annual mean SOI and annual wheat yield in NSW (right). Regression line coloured according to p-value, in blue correlations with p-values ≤ 0.05 and in grey with p-value > 0.05.

On the other hand, the presence of significant negative correlations, for shorter memories at Stage 2, suggests that the impact of high SOI values during non-spawning times of year, may be associated with catchability (e.g., windier, wetter-than-average conditions depressing catch rates).

We note that each single dot in the composite maps of Fig. 9 corresponds to a unique configuration of parameters m, a, b, c and d. A simple rule-of-thumb rule for identifying a promising candidate configuration is to: (a) search for a dot with a high absolute value of the correlation and (b) one that maintains a high correlation value to its immediate left and right within its local panel plot, across neighbouring local panels to its left and right (nearby m values), across neighbouring local panels above and below (nearby b values) and also checks for nearby a and d values.[4]

In all likelihood, there will be multiple promising parameter configurations. Below, we discuss just one such configuration with parameters: $m = 26$, $a = 2$, $b = 3$, $c = 0.75$ and $d = 1$. Naturally, users can choose one or more configurations that best suit their case study and research objectives.

The left panels of Fig. 10 exhibit the time series of Super$X^u(8)$ IMPIT indices correspond to the three calibration stages. From top to bottom, we observe marked changes in the shape of the resulting IMPIT index. At the top level (Stage 1) there appears to be essentially no linear trend over the 1988–2019 period. Which continues to be the case at second level (Stage 2).[5] As we move to the third level, we see that the Super$X^u(8)$ index exhibits a stronger downward linear trend. This reflects the importance of restricting to only the winter spawning season.

---

[4] Such checks can be easily performed within our app IMPIT$-a$.

[5] However, in both stages 1 and 2 there is an observable non-linear U-shaped trend.
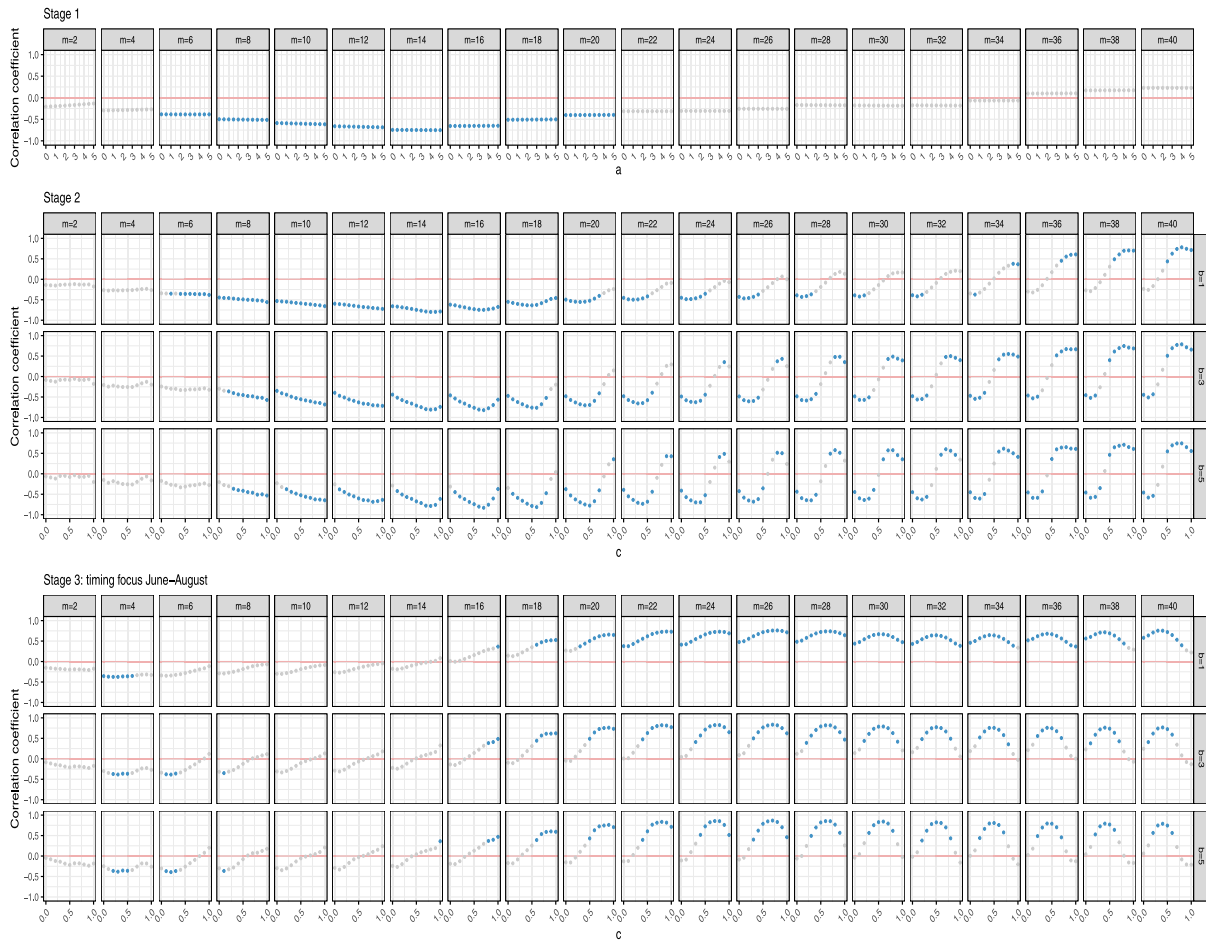
Fig. 9. Correlation coefficient between $\mathrm{Super}\,X^u(8)$ indices and annual snapper SCPUE (1988–2019). IMPIT indices were constructed for different values of *memory* (*m* in years), *persistence* according to Eq. (6) with $a = 2$, *recency* conforming to Eq. (8) varying *b* and *c* and *timing* according to Eq. (9) with $d = 1$. Snapper peak spawning season (June–August) was the special season considered in Stage 3. Values coloured according to *p*-value, in blue scale correlations with p-values ≤ 0.05 and in grey correlations with *p*-value > 0.05.

The right panels of Fig. 10 show the corresponding scattergrams of $\mathrm{Super}\,X^u(8)$ indices with annual snapper log-transformed catch rates. These illustrate statistically significant changes in both the magnitude of the correlations and the degree of goodness of fit of straight line trends.

In particular, note that at Stage 1 (top row), there is a non-significant negative linear association between the $\mathrm{Super}\,X^u(8)$ index and the catch rate. Inclusion of the calibrated recency weight, at Stage 2 (second row), leads to a positive but still non-significant correlation of 0.29. However, considering the winter spawning season at Stage 3 (third row), leads to a much higher correlation of 0.83 corresponding to R^2 of 0.69.

Our illustrative analyses indicate that $\mathrm{Super}\,X^u(8)$ index with a memory of 26 years and suitably calibrated parameters exhibits some remarkably strong linear associations with log-transformed annual snapper catch rates. The value $c = 0.75$ in the selected parameter configuration shows that episodes of approximately 20 (26×0.75) years in the past receive the highest recency importance weight ($w_2 = 1$). Episodes starting between 15 and 20 years in the past receive recency weights increasing from 0.79 to 1.0, while those starting between 20 and 25 years have recency weights decreasing from 1.0 to 0.59. The large increase in correlation from Stage 2 to Stage 3 suggests an impact of $\mathrm{Super}\,X^u(8)$ on spawning and possibly also on subsequent abundance. In summary, the results suggest episodic events characterised by elevated (≥ 8) SOI values during the June–August peak spawning season during the 15–25 years prior to catch are associated with elevated snapper commercial catch rates.

3.4. Marine heatwave indices and saucer scallop (Ylistrum balloti) in Queensland

For saucer scallops, linear associations between standardised catch rates and a range of MHW IMPIT indices were explored using the methodology of Section 2. As with snapper, the intensity of MHW episodes was computed by Eq. (3). We chose to examine the MHW index because of its association with other indices, its influence on scallops and a relatively long time series of available data.

Above average winter SSTs (June–August) are known to be negatively correlated with November–January scallop catch rates from 1988–2016, as reported in O'Neill et al. (2020). As noted in Caputi et al. (2015), in WA annual recruitment of saucer scallops is also correlated with SST (see also Joll and Caputi (1995) and Lenanton et al. (2009)). In addition, an extreme MHW in the summer of 2010–2011 had catastrophic impact on the WA stock (Caputi et al., 2014, 2015, 2019). Scallops have an extended spawning season from April to October. However, between April and May most scallops are already sexually mature (Dredge, 1981). Hence, we shall refer to April–May egg production as the "autumn component of the spawning season".

We performed the stage-wise explorative calibration described in Section 2.5. Fig. 11 depicts representatives of composite maps obtained at each stage. For the memory parameter *m*, we considered a 48−month window. This was chosen because scallops can in some instances live for up to 4 years (Dredge, 1985; Courtney et al., 2022). Fig. 11 shows values of memory increasing in steps of six months.

The top panel of Fig. 11 corresponds to the first stage. Setting $w_2 = w_3 = 1$, it displays plots of correlation coefficient versus the
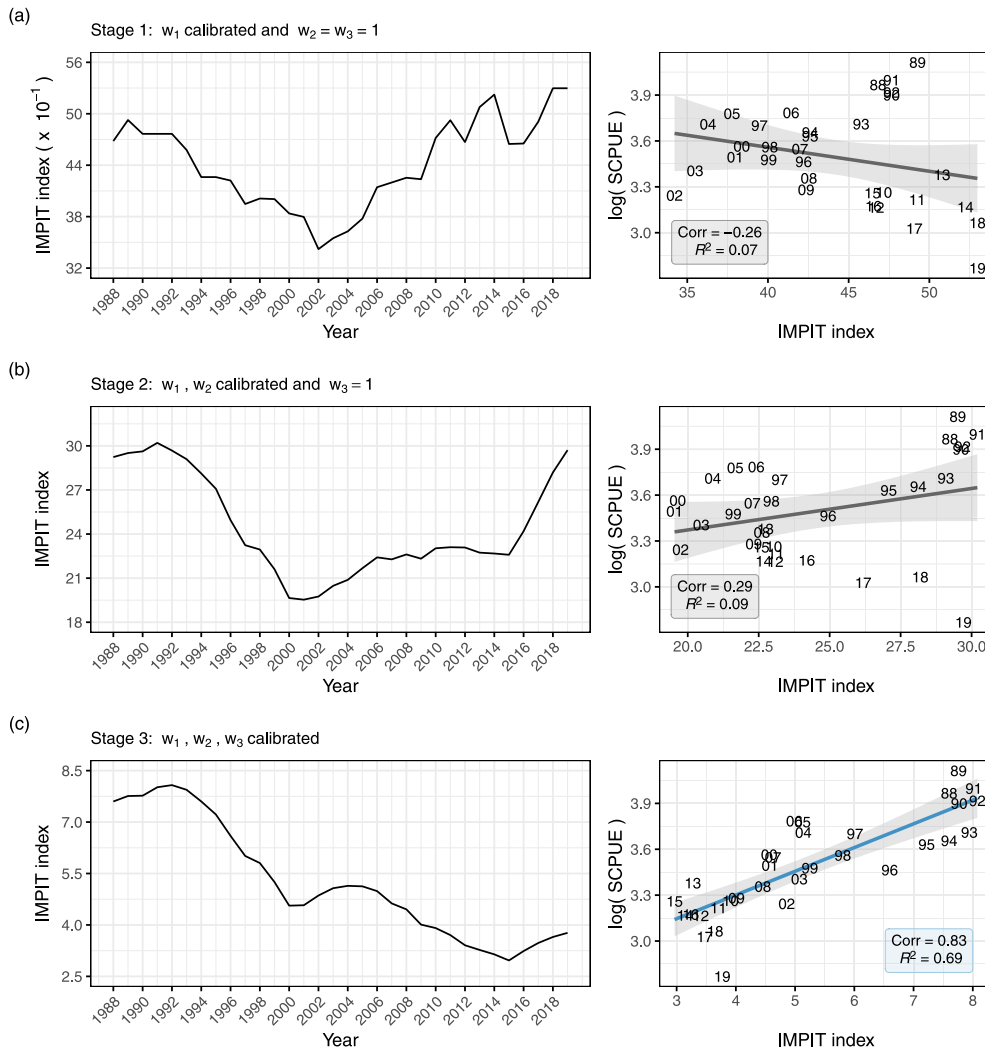
(a)

(b)

(c)

Fig. 10. Time series of Super $X^u(8)$ indices and their corresponding scattergrams with annual snapper log-transformed catch rates (1988–2019). For IMPIT indices' construction, a memory of 26 years ($m = 26$) was used. The intensity of episodes was computed following Eq. (3). Each row corresponds to a step of the *stage-wise explorative calibration* of the relative importance weights. Leading to a parameter configuration: $a = 2$, $b = 3$, $c = 0.75$ and $d = 1$. Regression line coloured according to p-value, in blue correlations with p-values ≤ 0.05 and, in grey otherwise.

parameter $a \in [0, 5]$ for each value of m. Those plots exhibited a roughly linear pattern but in this case, there was more variation compared to the case of snapper. Therefore, we chose to use $a = 0$ in w_1, which means that when constructing indices, all episodes were assigned the same persistence weight.

At the second stage, while setting $w_3 = 1$ ($a = 0$) we plotted the correlation coefficient versus the parameter c for alternative pairs of m and b parameters. At the third stage, we repeated the second stage with only those episodes that overlap the April–May component of the scallops spawning season. Second and third panels of Fig. 11 display the resulting representative composite maps. For simplicity, as in the case of snapper, only results for $b \in \{1, 3, 5\}$ are displayed. At the third stage, we generated composite maps for alternative values of d. Fig. 11 only displays the map for $d = 1$ as it leads to promising candidate configurations.

Careful examination of plots in Fig. 11 reveals interesting patterns generated by the stage-wise calibration. In particular, the impact of considering the special timing (April–May) generated noticeable differences between stages for values of $m \geq 24$ months.

Typically, there will be multiple promising parameter configurations. Below, we discuss just one such configuration with parameters: $m = 30$, $a = 0$, $b = 3$, $c = 0.4$ and $d = 1$. Naturally, users can choose

one or more configurations that best suit their case study and research objective.

The left panels of Fig. 12 exhibit the time series of MHW IMPIT indices for each calibration stage. From top to bottom, we observe that the shape of the resulting index changes. At top level (Stage 1) there appears to be a rough downward trend until 2013 followed by an upward swing. That trend is less pronounced at second level (Stage 2). However, when we move from second to third level, we observe a period of relative stability between 2000 and 2013 followed by a marked increase. This reflects the effect of the addition of the spawning May–April timing into the calibration.

The right panels of Fig. 12 show the corresponding scattergrams of MHW IMPIT indices with November log-transformed scallop catch rates. In particular, note that at Stage 1 (first row), there is a negative linear association between MHW IMPIT index and the catch rate of −0.41. Inclusion of the recency weight calibration, at Stage 2 (second row), led to a very marginal but now statistically significant improvement of the correlation. Moreover, at Stage 3 (third row), the calibration of the timing w_3 weight, led to a highly significant correlation of −0.64 as compared to Stages 1 and 2, which accounts for 40% of variability.
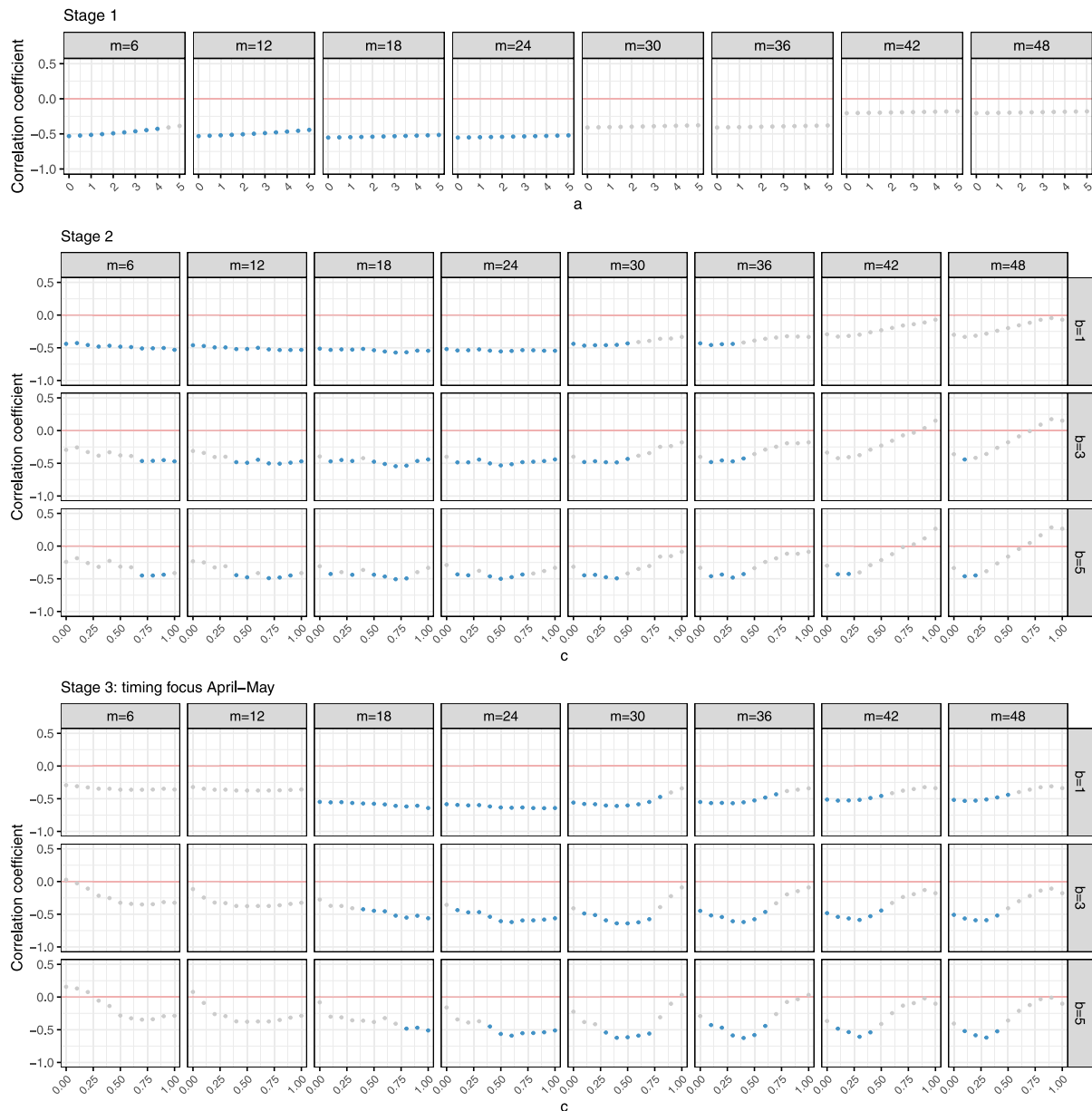
Fig. 11. Correlation coefficient between MHW IMPIT and November scallop SCPUE (1997–2018). IMPIT indices were constructed for different values of memory (m in months), *persistence* according to Eq. (6) with $a = 0$ and *recency* conforming to (8) varying b and c and *timing* according to Eq. (9) with $d = 1$. April–May was the special season considered in Stage 3. Values coloured according to p-value, in blue scale correlations with p-values ≤ 0.05 and in grey correlations with *p*-value > 0.05.

Our illustrative analyses indicate that a MHW IMPIT index with a memory of 30 months and suitably calibrated parameters exhibits some strong linear associations with log-transformed November scallop catch rates. The value of $a = 0$ in the selected parameter configuration means that all episodes receive equal persistence weight and the value of $c = 0.4$ shows that episodes of 12 (30×0.4) months in the past receive the peak recency importance weight ($w_2 = 1$). Episodes starting between 6 and 12 months in the past receive recency weights increasing from 0.74 to 1.0, while those starting between 12 and 18 months have recency weights decreasing from 1.0 to 0.79.

The increase in correlation in Stage 3 compared to Stages 1 and 2 is suggestive of an impact of April–May (autumn component of the spawning season) MHW IMPIT index on subsequent abundance. This means that MHW episodes ranging from 6 to 18 months in the past may strongly depress scallop catch rates in the following year. This is likely a result of either direct heat stress to the scallops themselves, or possible indirect effects of warm waters affecting primary production (i.e., the

phytoplankton and zooplankton that scallops feed on) and therefore reducing food availability for scallops, predators and disease, all of which can affect scallop abundance (Courtney et al., 2015; Richardson et al., 2020).

The strong correlation (-0.64) of the MHW IMPIT index with log-transformed SCPUE is not surprising since the bottom panel of Fig. 8 (and findings of O'Neill et al. (2020)) indicates a significant association with the raw SST signal. However, focusing on MHW IMPIT indices overlapping April–May reveals even stronger associations.

3.5. SOI threshold-crossing indices and New South Wales wheat yield

We investigated linear associations between annual yield per hectare of wheat and a range of threshold crossing $SubX^d(-8)$ IMPIT indices constructed using the methodology of Section 2.4. The intensity of $E_k^d(-8)$ episodes in $SubX^d(-8)$ indices was computed by Eq. (3).

We implemented the stage-wise explorative calibration described in Section 2.5. Fig. 13 depicts representatives of composite maps obtained
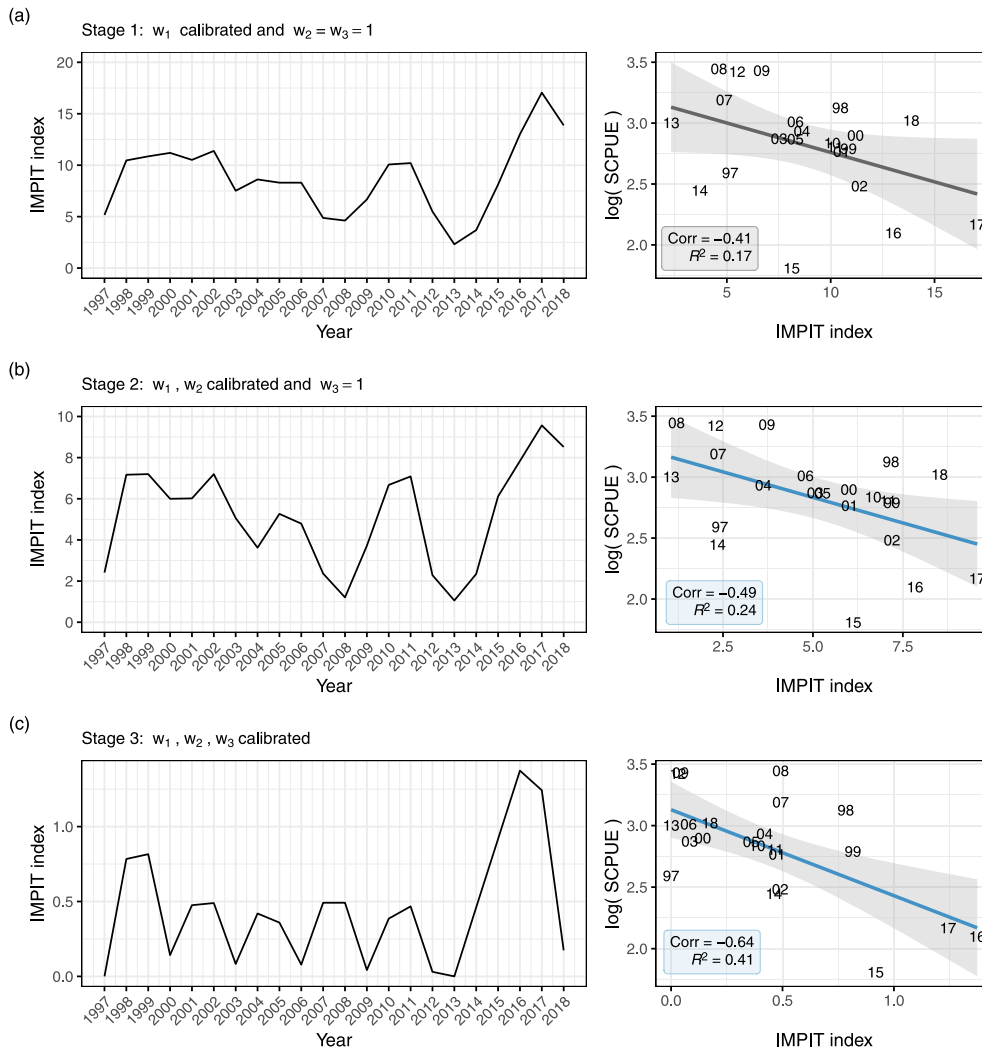
Fig. 12. Time series of MHW IMPIT indices and their corresponding scattergrams with November scallop catch rates (1997–2018). For IMPIT indices' construction, a memory of 30 months ($m = 30$) was used. The intensity of episodes was computed following Eq. (3). Each row correspond to a step of the *stage-wise explorative calibration* of the relative importance weights. Leading to a parameter configuration: $a = 0$, $b = 3$, $c = 0.4$ and $d = 1$. Regression line coloured according to *p*-value, in blue correlations with p-values ≤ 0.05 and, in grey otherwise.

at each stage, pruned for ease of display. For instance, we displayed only values of the memory parameter m which ranged from 1 to 10 years. The value of 10 was selected to check for long term effects. In recent decades, Australia has seen a shift towards higher temperatures and lower winter rainfall, which has had significant effects on many farmers (Hughes et al., 2022).

The top panel of Fig. 13 corresponds to the first stage. Setting $w_2 = w_3 = 1$, it displays plots of the correlation coefficient versus the parameter $a \in [0, 5]$ for each value of m. Those plots exhibited a roughly linear pattern with very small variation and no significant correlations. As in Section 3.3, we set $a = 2$.

At the second stage, while setting $w_3 = 1$ (and $a = 2$) we plotted the correlation coefficient versus the parameter c for alternative pairs of m and b (see the second panel of Fig. 13). For simplicity, only results for values $b \in \{1, 3, 5\}$ are displayed. We note that for the short memory of 1 year we observed sustained statistically significant correlations for several values of c and b ranging from 3 to 5. There were also some isolated significant correlations for longer term memories of 5 years and above. We emphasise $m = 5$ because these associations carry over also to Stage 3 discussed below.

The third stage considers the overlap of $E_k^d(-8)$ episodes with the season of January–April which precedes the usual sowing period in NSW. Arguably, the conditions of the soil at the time of sowing have an impact on the subsequent yield. We then generated the composite maps for alternative values of the parameter d. The third panel of Fig. 13 displays the representative composite maps for only one value $d = 1$, as other values of d result in similar outputs.

Careful examination of Fig. 13 reveals interesting patterns generated by the stage-wise calibration. The significant positive correlations are to be expected as it is generally believed that decreasing negative values of SOI maybe associated with higher wheat yields (Wan et al., 2022). The sporadic significant negative correlations for long memories of 8 years and above are noted but viewed as possibly spurious because they are less stable with respect to parameter changes. Moreover, the inclusion of the special (January–April) timing generated the most noticeable difference between stages as it led to the significant positive correlations for the memory of $m = 5$ years. At the same time, the associations at the memory of $m = 1$ year became statistically insignificant when the episodes were restricted to the special period.

The strongest observed correlation between the $\mathrm{Sub}X^d(-8)$ IMPIT index of SOI and the wheat yield per hectare was 0.44 which accounts for nearly 20% of the variability. This is much higher than the 0.29 reported in panel (c) of Fig. 8 for the corresponding correlation with simple mean SOI index. Once again, this illustrates that IMPIT indices can potentially identify meaningful associations in this context of agricultural application.
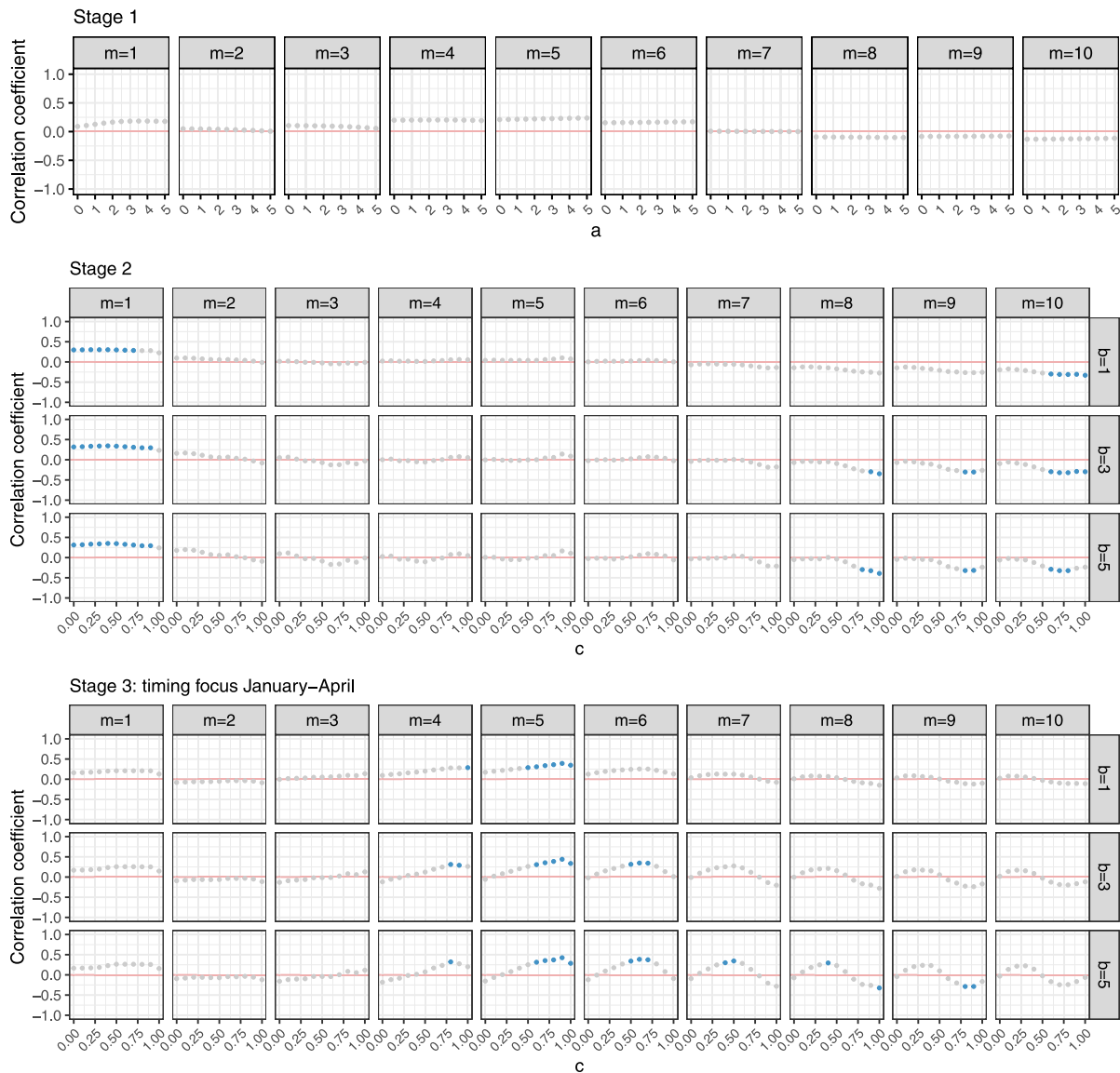
Fig. 13. Correlation coefficient between $\text{Sub}X^d(-8)$ indices and NSW annual wheat yield (1974–2021). IMPIT indices were constructed for different values of *memory* (*m* in years), *persistence* according to Eq. (6) with *a* = 2, *recency* conforming to Eq. (8) varying *b* and *c* and *timing* according to Eq. (9) with *d* = 1. January–April was the special season considered in Stage 3. Values coloured according to p-value, in blue scale correlations with p-values ≤ 0.05 and in grey correlations with *p*-value > 0.05.

4. IMPIT app

In order to facilitate the use of IMPIT indices, a software tool IMPIT−*a* was developed in the R environment (Core Team, 2021) utilising "Shiny" (Chang et al., 2021), an R package that provides a web framework for building web applications using R (https://cran.r-project.org). It is important to note that R is open-source and widely used in environmental studies (Aparicio et al., 2019; Díaz et al., 2021).

The IMPIT−*a* shiny app code is available on GitHub (https://github.com/manumendiolar/IMPIT_shiny) and needs to be run into the R or RStudio (RStudio Team, 2021) environment, following instruction on the GitHub page. Alternatively, it can be accessed via the shinyapps.io platform (see Section 6). Note that because the app was built in R, it can be run on Windows, Linux, or macOS systems. With IMPIT−*a* deployed as a web application, it can be used regardless of operating systems, hardware, or other installed software, since it can be run via a web browser. The aims of IMPIT−*a* are to:

- Provide a user-friendly interface for constructing IMPIT indices.
- Provide a smooth workflow ranging from importing and exploring raw data to defining episodes.

- Allow users to choose from a menu of intensity and relative weight functions.
- Visualise imported data, defined episodes, and constructed IMPIT indices.

The app is intended to be self-contained in the sense that all the instructions and definitions are embedded in help messages within the software. The main tasks of the app can be summarised in four steps:

1. Import the environmental data.
2. Define discrete episodes.
3. Construct IMPIT index.
4. Explore IMPIT index.

Each of the above steps is described in more detail in the remainder of this section. Fig. 14 shows the IMPIT−*a* user interface structure with a dynamically linked sidebar menu and Main Panel.

4.1. Input data

Construction of IMPIT indices depends primarily on characterisation of discrete episodes of interest. The app offers the options of
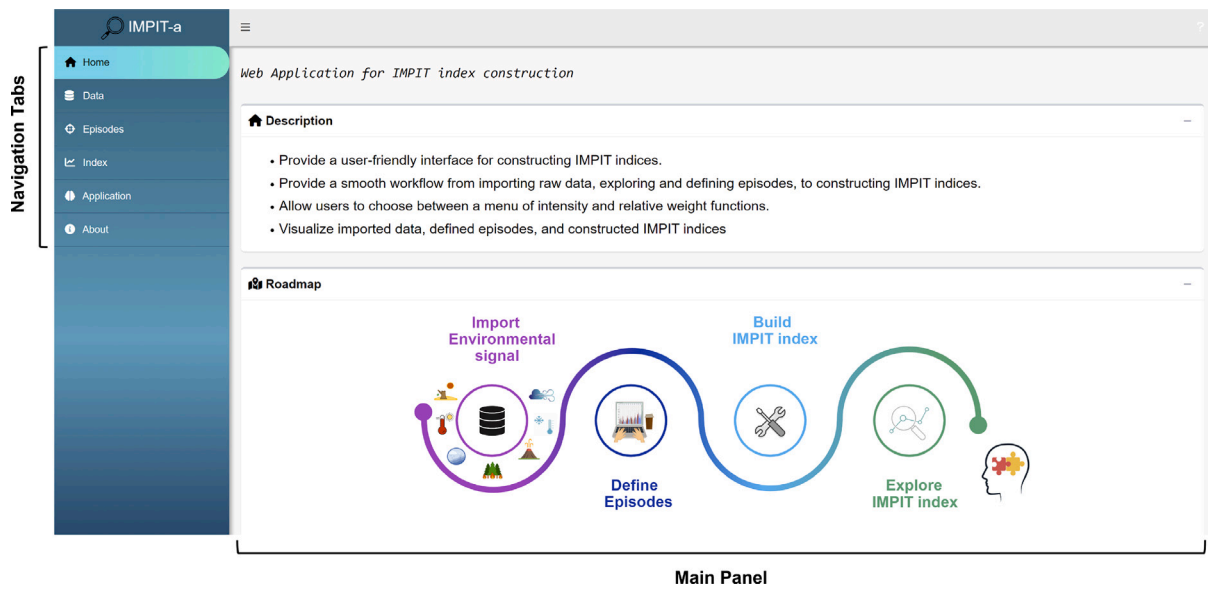
Fig. 14. IMPIT−*a* user interface structure. Sidebar Menu contains Navigational Tabs and app options. Main Panel contains app visualisations capabilities.
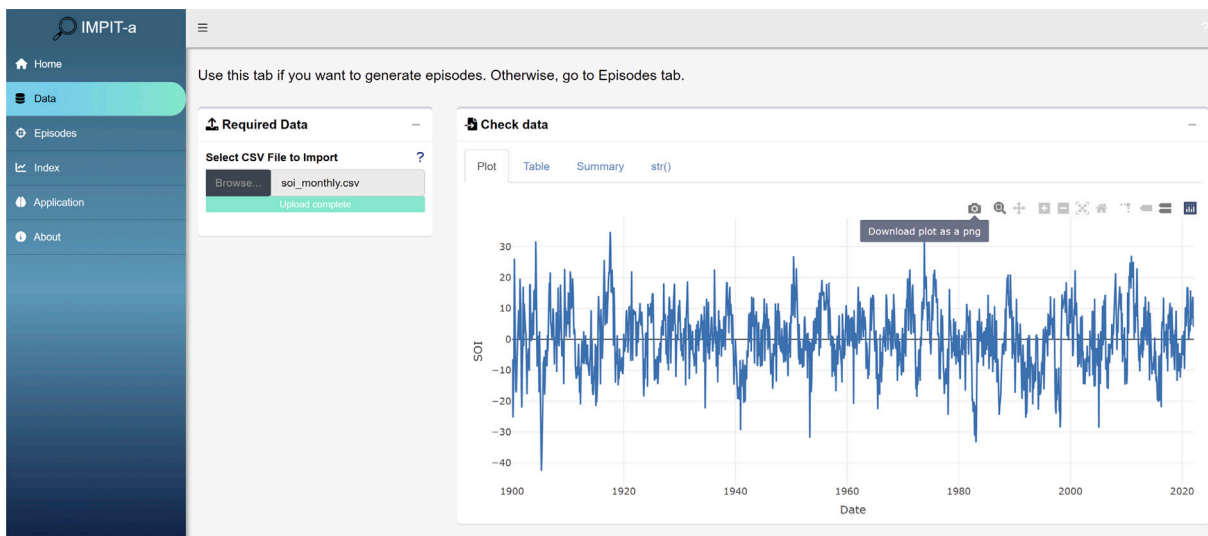

Fig. 15. Example of Data tab, illustrated with time series plot of the monthly values of Southern Oscillation Index (SOI) from January 1900 to December 2020.

either constructing threshold-crossing episodes or by directly uploading user-supplied episodes.

For the first option, the user has to upload the time series of the environmental signal, via the Data tab, and then construct the episodes, via the Episodes tab. Fig. 15 shows an example of the Data tab. Uploaded data can be checked via Plot, Table, Summary and str() tab options. Note that all users can interact with graphs on display. Zooming in and out, point value display, panning graphs or saving the plot by clicking "download plot as a png" button in the toolbar at the top of the graph.

Once this is done, the user can proceed to define episodes using the Episodes tab. This provides a definition of episodes based on the threshold-crossing indices introduced in Section 2.4. User will have to select a threshold and choose between up or down episodes with respect to the chosen threshold. In addition, the minimum duration of episodes can also be selected (by default it is 1). Minimum duration means, minimum consecutive values above or below the threshold (see Fig. 16). However, if the user already has a list of episodes, they can be uploaded directly (see top panel of Fig. 16). The app offers the option to download the table with detailed information of episodes as well as

lollipop charts showing the intensity mean and duration of all episodes. If the special season option was considered, episodes highlighted in yellow are those overlapping the special season (see bottom panel of Fig. 16).

4.2. Index construction

Once the episode list is generated or uploaded via the Episodes tab, we are ready to build an IMPIT index. Suppose we consider up-episodes $E_k^u(8)$ and we want to build their corresponding $\text{Super}X^u(8)$ index. A crucial step is to set up a configuration of parameters associated with the memory, persistence, recency and timing of the episodes. Table 1 shows a summary of the main parameters, with their description, which can be found in the Index tab.

In the next illustration, we consider a memory of $m = 26$ years. This means that the value of the index in, for example, 2022 will be computed based on the contribution of all $E^u(8)$ episodes from 1996 until 2022. For the intensity function, from the menu on this tab we can pick among the mean, median, minimum, maximum or logarithm of the episode's values. Suppose we choose the intensity function given
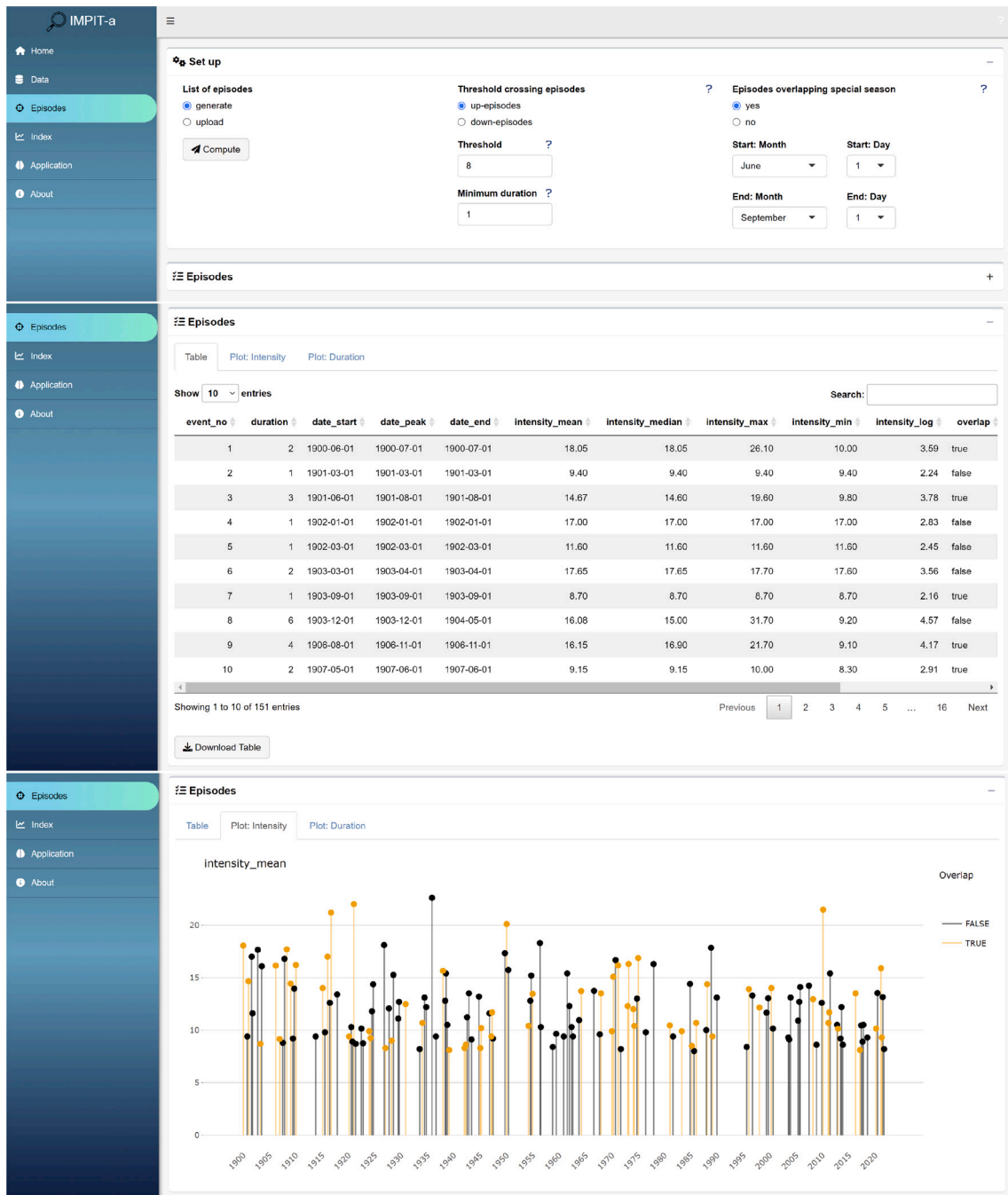
Fig. 16. Example of Episodes tab capabilities. Top: In this case, $E_k^u(8)$ episodes were generated with timing focus on June–August from SOI signal (from January 1900 to December 2020). Middle: Table with detailed information of the episodes. Bottom: Lollipop chart showing the intensity mean of all episodes computed. In yellow the episodes that overlap the special season (June/01 - September/01).

by Eq. (3). Next, we have to decide on the relative importance weights of the episodes. Three aspects could be considered: persistence, recency and timing of episodes. For the persistence, let us consider $a = 2$ which corresponds to the middle pink curve in Fig. 1. For the recency, suppose we choose $b = 3$ for the dampening rate of decay and a value of $c = 0.75$ which will be similar to the light green curve in the right panel of Fig. 2. This means that episodes starting recently will be more important than those starting late in the prescribed memory. Finally, for the timing, suppose episodes are weighted by their overlap duration with the special timing. To illustrate, we considered $d = 1$ to calibrate the timing weight, which corresponds to dark orange curve (fifth from bottom to top) in the right panel of Fig. 3. Fig. 17 displays the resulting

IMPIT index plot. User can download the table of IMPIT index values in the Table tab.

4.3. Output data exploration

The "Application" tab was designed to provide a suite of tools for index exploration, including plot of IMPIT index with smoothing and the option to run a simple regression analysis between the index and a response variable of interest. The app's visualisation functionality allows users to explore the index before exporting it for later use.

On the left panel (in the Application tab), users can upload an IMPIT index (generated with the app) and a response variable of interest.
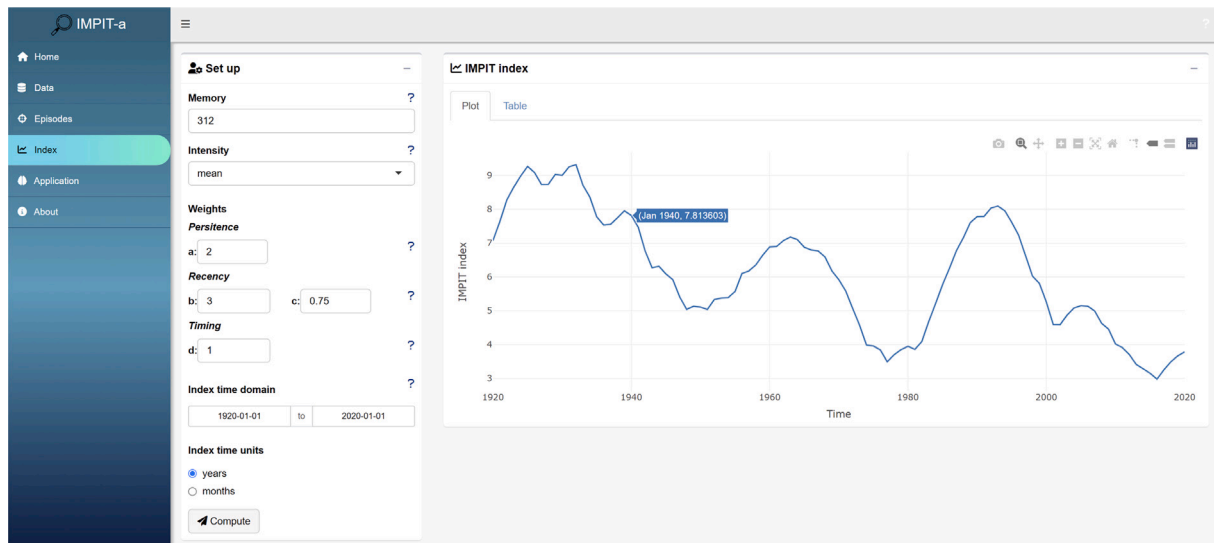
Fig. 17. Example of Index tab to compute IMPIT index according to parameter's specifications. User can choose among a menu of memory, intensity and relative weight functions. Timing was considered and $E^u(8)$ episodes were weighted by their overlap duration with the June–August season. Note that memory units coincide with episodes units, in this case 312 months (26 years).

Table 1
Summary description of IMPIT$-a$'s main parameters.

Parameter	Description
m	Memory. A period of fixed and uninterrupted duration in the past with respect to the current observation in time.
a	Associated with the persistence of the episode. Dampens the rate of decay (Fig. 1). A value close to zero means that each episode will have nearly the same importance weight regardless of its duration.
b	Associated with the recency of the episode. Captures the rate of decay (Fig. 2) of the recency weight as the starting date deviates more from its peak. Low values flatten that relative importance weight.
c	Associated with the recency of the episode. Characterises the skewness of the recency weight and its peak (Fig. 2). A value close to zero indicates that recent episodes have higher weight than those starting late in the memory.
d	Associated with the timing importance weight of the episode. The weight is monotone increasing with d (Fig. 3).

On the right panel, the first two tabs correspond to IMPIT index exploration, where user can visualise IMPIT index. A common approach to identify the structure of present trends, is to plot the time series in combination with a smooth fitted curve (von Brömssen et al., 2021).

The estimation of the smooth trend is usually achieved using a generalised additive model (GAM) or by LOESS smooths (Hastie and Tibshirani (1986), Wood (2017) and Cleveland (1979). The app provides the option of explore IMPIT index with a LOESS smooth and its 95% confidence intervals. The third and fourth tabs allow the users to check the uploaded response variable data. The fifth tab corresponds to the scatterplot between the IMPIT index and the response variable. And the sixth tab presents a summary of the regression analysis.

5. Discussion

In this paper, we developed a generic, parametrised family of weighted indices extracted from observations of a relevant environmental signal containing potentially important discrete episodes. The methodology focuses on determining an appropriate memory length and assigning importance weights to episodes that capture: intensity, persistence, intermittence and timing.

In the process, we proposed a stage-wise calibration procedure for refining the parameter configurations of importance weights assigned to the discrete intermittent episodes of interest. We chose not to automate the search by specifying a fixed optimality criterion for selecting an "optimal" parameter configuration. Instead, we believed that structured stage-wise exploration of the parameter space, of the type we proposed, would enable experts to make a judicious choice of parameters.

We illustrated the effectiveness and possible uses of IMPIT indices in the context of fishery and agricultural applications. In particular, we considered standardised commercial fishery catch rate data from two fished species in Queensland, snapper (*Chrysophrys auratus*) and saucer scallop (*Ylistrum balloti*), and yield per hectare of wheat production in New South Wales. We then searched for associations between the trends in these data and IMPIT indices constructed from intermittent episodes embedded in SOI and SST signals.

Using SCPUE and yield per hectare of wheat as response variables, we showed that significant correlations exist with suitably calibrated IMPIT indices. Moreover, this occurred even though analyses using simple (baseline) means of the underlying environmental signals indicated no significant correlation in the case of snapper and wheat yield and only relatively weak correlation in the case of scallops. Hence, we demonstrated that our stage-wise parameter calibration of IMPIT indices plays a potentially important role in detecting associations between special seasons of either the studied species or wheat, and their environment.

We also developed an IMPIT$-a$ software that expedites the index construction process and combines it with data which allows users to refine datasets/episodes and time periods analysed based on exploration, before exporting the resulting IMPIT index for further use. No specialised coding or expertise are needed to use IMPIT$-a$. The web interface ensures that the tool can be accessed from a web browser without installing any additional software. Although IMPIT$-a$ offers only a limited number of options for the intensity or weights functions, it can be extended to include other functional forms.

Identification of strong associations between a calibrated IMPIT index and response variables should be seen as a starting point for deeper follow-up investigations. These would seek to confirm that the corresponding IMPIT index is capturing a causative, not just correlative, relationship between the environmental signal and the study species.

IMPIT indices are quite generic and easy to interpret, and can be used across a wide range of environmental and ecological disciplines.

The indices and their calibration should be seen as another tool in the toolbox of exploratory data analyses.

In the remainder of this section we wish to acknowledge several potential limitations of IMPIT indices as well as possible fruitful directions for future investigations. Many of these stem from the likely non-stationarity of the data time series involved in applying the IMPIT technique. Clearly, environmental series data is likely to be non-stationary (e.g., when the signal is impacted by climate change).

We recall that, in this study, an IMPIT index $X(I, w)$ is calculated only at the snapshot single, current, time and hence it is associated with only a single correlation ρ. However, if we had considered it repetitively at a range of times we would have constructed an IMPIT time series $\{X_t(I, w)\}_{t \geq 0}$, and an accompanying time series of correlations $\{\rho_t\}_{t \geq 0}$ which in many applications would probably be non-stationary. Investigation of the nature of such non-stationarity (especially any systematic trends) would constitute an excellent direction for continuing the present study.

Our proposed IMPIT construction assumes only that the family of discrete intermittent episodes of interest (e.g., marine heatwaves) is well-defined irrespective of the stationarity properties of the underlying data time series. However, the question of how this impacts the calibration procedure is both interesting and challenging. We conjecture that the impact of non-stationarity on the calibration parameters (a, b, c, d) will also depend on the length of the memory parameter m. For instance, for small m such impacts are likely to be small. For large m, we expect that the choice of parameters b, c (impacting recency weight curves in Fig. 2) would be most sensitive.

Clearly, promising IMPIT candidates could and, perhaps, should be considered for inclusion as explanatory variables in a range of models. The most obvious of these would be regression or generalised linear models (GLMs). In this context, IMPIT explanatory variables could have both potential benefits and drawbacks that may be due to their relatively complex structure.

Benefits, because a single IMPIT variable may eliminate several other environmental variables and account for a lot of variability in the dependent variable. Drawbacks because it may cause a collinearity problem or because the separate calibration of weights in the IMPIT index either reduces the flexibility of the regression technique or tilts the coefficients assigned to other explanatory variables. In addition, some characterisations of underlying intermittent episodes, may result in too few of these episodes occurring in the memory windows under consideration.

We note also that inclusion of an IMPIT index as linear variable in a model should have fewer drawbacks than if it were entered in a nonlinear fashion. That is because in the former case, all the weights comprising $X(I, w)$ would be merely scaled, whereas in the latter case, implicitly interactions among intermittent episodes E_k within the index may acquire a role.

Finally, the question of whether (and how) data series non-stationarity should impact experts' characterisation of the intermittent episodes of interest, could well be worth investigating. For instance, with threshold crossing episodes, one could easily imagine time dependent, rather than fixed, thresholds.

6. Software availability

Name of the software: IMPIT$-a$.
Developer: Manuela Mendiolar.
Contact Email: m.mendiolar@uq.edu.au.
Tested browsers: Firefox and Google chrome.
Software Required: R, RStudio.
R-Packages required: shiny, shinydashboard, shinydashboardPlus, shinyFiles, shinyhelper, shinyalert, shinyvalidate, shinyjs, shinyWidgets, dashboardthemes, tidyverse, DT, plotly, spsComps and lubridate.
Programming language: at least R version 4.1.2.
Available since: 2022.

The app source code is stored in a freely accessible GitHub repository hosted by Manuela Mendiolar: https://github.com/manumendiolar/IMPIT_shiny.

The app can be deployed via: https://manumendiolar.shinyapps.io/impit_shiny/.

Data availability

The environmental data is freely available online, and that for the fishery data, the authors do not have permission to share the data.

References

Aburto-Oropeza, O., Paredes, G., Mascareñas-Osorio, I., Sala, E., 2010. Climatic influence on reef fish recruitment and fisheries. Mar. Ecol. Prog. Ser. 410, 283–287.

Aparicio, J., Ariza-Suarez, D., Raatz, B., 2019. Web application for spatial modelling of field trials. In: 15-19 of July 2019, Barranquilla, Colombia XXIX Simposio Internacional de Estadística. pp. 1–6.

Bakun, A., 2014. Active opportunist species as potential diagnostic markers for comparative tracking of complex marine ecosystem responses to global trends. ICES J. Mar. Sci. 71 (8), 2281–2292. http://dx.doi.org/10.1093/icesjms/fst242.

Barbeaux, S.J., Holsman, K., Zador, S., 2020. Marine heatwave stress test of ecosystem-based fisheries management in the Gulf of Alaska Pacific cod fishery. Front. Mar. Sci. 7, http://dx.doi.org/10.3389/fmars.2020.00703.

Bellanthudawa, B.K.A., Chang, N.-B., 2022. Spectral index-based time series analysis of canopy resistance and resilience in a watershed under intermittent weather changes. Ecol. Inform. 69, 101666. http://dx.doi.org/10.1016/j.ecoinf.2022.101666.

Brad Adams, J., Mann, M.E., Ammann, C.M., 2003. Proxy evidence for an El Niño-like response to volcanic forcing. Nature 426 (6964), 274–278. http://dx.doi.org/10.1038/nature02101.

Bureau of Meteorology Australia, 2012. Record-breaking la Niña events: an analysis of the La Niña life cycle and the impacts and significance of the 2010–11 and 2011–12 La Niña events in Australia.

Caputi, N., Feng, M., Pearce, A., Benthuysen, J., Denham, A.M., Hetzel, Y., Jackson, G., Molony, B.W., Joll, L.M., Chandrapavan, A.N., 2015. Management Implications of Climate Change Effect on Fisheries in Western Australia Part 1: Environmental Change and Risk Assessment. FRDC: Project No. 2010/535. Fisheries Research Report No. 260, Department of Fisheries, Western Australia., http://dx.doi.org/10.13140/RG.2.1.2240.7200.

Caputi, N., Kangas, M., Chandrapavan, A., Hart, A., Feng, M., Marin, M., de Lestang, S., 2019. Factors affecting the recovery of invertebrate stocks from the 2011 western Australian extreme marine heatwave. Front. Mar. Sci. 6, 484. http://dx.doi.org/10.3389/fmars.2019.00484.

Caputi, N., de Lestang, S., Hart, A., Kangas, M., Johnston, D., Penn, J., 2014. Catch predictions in stock assessment and management of invertebrate fisheries using pre-recruit abundance — case studies from western Australia. Rev. Fish. Sci. Aquac. 22 (1), 36–54. http://dx.doi.org/10.1080/10641262.2013.832144.

Chang, W., Cheng, J., Allaire, J.J., Sievert, C., Schloerke, B., Xie, Y., Allen, J., McPherson, J., Dipert, A., Borges, B., 2021. Shiny: web application framework for R.

Chowdhury, R.K., Beecham, S., 2010. Australian rainfall trends and their relation to the southern oscillation index. Hydrol. Process.: Int. J. 24 (4), 504–514. http://dx.doi.org/10.1002/hyp.7504.

Cleveland, W.S., 1979. Robust locally weighted regression and smoothing scatterplots. J. Amer. Statist. Assoc. 74 (368), 829–836. http://dx.doi.org/10.1080/01621459.1979.10481038.

Core Team, R., 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Courtney, A.J., Leigh, G.M., Yang, W.-H., Campbell, M.J., McLennan, M.F., 2022. Estimating the natural mortality rate of saucer scallops (*ylistrum balloti)* on the queensland east coast from tag-recaptures. Fish. Res. 250, http://dx.doi.org/10.1016/j.fishres.2022.106273.

Courtney, A.J., Spillman, C.M., Lemos, R.T., Thomas, J., Leigh, G.M., Campbell, A.B., 2015. Physical Oceanographic Influences on Queensland Reef Fish and Scallops. FRDC Project No. 2013/020, Fisheries Research and Development Corporation.

Croston, J.D., 1972. Forecasting and stock control for intermittent demands. J. Oper. Res. Soc. 23 (3), 289–303.

Department of Primary Industries NSW, 2007. Wheat: growth & development.

Deryugina, T., 2013. How do people update? The effects of local weather fluctuations on beliefs about global warming. Clim. Change 118 (2), 397–416. http://dx.doi.org/10.1007/s10584-012-0615-1.

Díaz, J.J., Mura, I., Franco, J.F., Akhavan-Tabatabaei, R., 2021. Aire - a web-based r application for simple, accessible and repeatable analysis of urban air quality data. Environ. Model. Softw. 138, 104976. http://dx.doi.org/10.1016/j.envsoft.2021.104976.

Dredge, M.C.L., 1981. Reproductive biology of the saucer scallop *amusium japonicum balloti* (bernardi) in central queensland waters. Mar. Freshw. Res. 32 (5), 775–787.

Dredge, M.C.L., 1985. Estimates of natural mortality and yield-per-recruit for *amusium japonicum balloti* bernardi (*pectinidae*) based on tag recoveries. J. Shellfish Res. 5, 103–109.

Ensminger, I., Sveshnikov, D., Campbell, D.A., Funk, C., Jansson, S., Lloyd, J., Shibistova, O., Öquist, G., 2004. Intermittent low temperatures constrain spring recovery of photosynthesis in boreal scots pine forests. Global Change Biol. 10 (6), 995–1008.

Filar, J.A., Courtney, A.J., Gibson, L.J., Jemison, R., Leahy, S.M., Lei, Y., Mendiolar, M., Mitchell, J., Robson, B., Steinberg, C., Williams, S.M., Yang, W.-H., Ye, N., 2021. Modelling Environmental Changes and Effects on Wild-Caught Species in Queensland. Environmental Drivers. FRDC Project No. 2019/013, Fisheries Research and Development Corporation.

French, S.M., Courtney, A.J., Yang, W.-H., 2021. Quantitative analysis of the fishery-independent queensland saucer scallop (*ylistrum balloti*) trawl survey. J. Shellfish Res. 40 (2), 297–309.

Frölicher, T.L., Fischer, E.M., Gruber, N., 2018. Marine heatwaves under global warming. Nature 560 (7718), 360–364.

Glantz, M.H., 2015. Shades of chaos: lessons learned about lessons learned about forecasting El Niño and its impacts. Int. J. Disaster Risk Sci. 6, 94–103.

Gutierrez, L., 2017. Impacts of El Niño-southern oscillation on the wheat market: A global dynamic analysis. PloS ONE 12 (6), http://dx.doi.org/10.1371/journal.pone.0179086.

Ham, Y.-G., Kim, J.-H., Luo, J.-J., 2019. Deep learning for multi-year ENSO forecasts. Nature 573 (7775), 568–572.

Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press.

Hannachi, A., Jolliffe, I.T., Stephenson, D.B., 2007. Empirical orthogonal functions and related techniques in atmospheric science: A review. Int. J. Climatol. 27 (9), 1119–1152. http://dx.doi.org/10.1002/joc.1499.

Hastie, T., Tibshirani, R., 1986. Generalized additive models. Statist. Sci. 1 (3), 297–310. http://dx.doi.org/10.1214/ss/1177013604.

Hobday, A.J., Alexander, L.V., Perkins, S.E., Smale, D.A., Straub, S.C., Oliver, E.C.J., Benthuysen, J.A., Burrows, M.T., Donat, M.G., Feng, M., Holbrook, N.J., Moore, P.J., Scannell, H.A., Gupta, A.S., Wernberg, T., 2016. A hierarchical approach to defining marine heatwaves. Prog. Oceanogr. 141, 227–238. http://dx.doi.org/10.1016/j.pocean.2015.12.014.

Hoffmann, R., Muttarak, R., Peisker, J., Stanig, P., 2022. Climate change experiences raise environmental concerns and promote green voting. Nature Clim. Change 12 (2), 148–155. http://dx.doi.org/10.1038/s41558-021-01263-8.

Hughes, N., Lu, M., Soh, W.Y., Lawson, K., 2022. Modelling the effects of climate change on the profitability of Australian farms. Clim. Change 172 (1–2), 12.

Iizumi, T., Luo, J.-J., Challinor, A.J., Sakurai, G., Yokozawa, M., Sakuma, H., Brown, M.E., Yamagata, T., 2014. Impacts of El Niño southern oscillation on the global yields of major crops. Nat. Commun. 5 (1).

Joll, L.M., Caputi, N., 1995. Environmental influences on recruitment in the saucer scallop (*amusium balloti*) fishery of shark bay, western Australia. ICES Mar. Sci. Symp. 199, 47–53.

Jolliffe, I.T., 2011. Principal component analysis. In: Lovric, M. (Ed.), International Encyclopedia of Statistical Science. Springer, pp. 1094–1096. http://dx.doi.org/10.1007/978-3-642-04898-2_455.

Kangas, M., Chandrapavan, A., Wilkin, S., Caputi, N., 2022. Recovery trajectories and management responses for three scallop stocks over ten years following an extreme marine heatwave in western Australia. Front. Clim. 4, http://dx.doi.org/10.3389/fclim.2022.1043889.

Krause, P., Boyle, D.P., Bäse, F., 2005. Comparison of different efficiency criteria for hydrological model assessment. Adv. Geosci. 5, 89–97. http://dx.doi.org/10.5194/adgeo-5-89-2005.

Lenanton, R.C., Caputi, N., Kangas, M., Craine, M., 2009. The ongoing influence of the leeuwin current on economically important fish and invertebrates off temperate western Australia–has it changed? J. R. Soc. West. Aust. 92, 111–127.

Li, J., Thompson, D.W.J., 2021. Widespread changes in surface temperature persistence under climate change. Nature 599 (7885), 425–430. http://dx.doi.org/10.1038/s41586-021-03943-z.

Loughran, T.F., Perkins-Kirkpatrick, S.E., Alexander, L.V., 2017. Understanding the spatio-temporal influence of climate variability on Australian heatwaves. Int. J. Climatol. 37 (10), 3963–3975. http://dx.doi.org/10.1002/joc.4971.

Ludescher, J., Gozolchiani, A., Bogachev, M.I., Bunde, A., Havlin, S., Schellnhuber, H.J., 2014. Very early warning of next El Niño. Proc. Natl. Acad. Sci. 111 (6), 2064–2066.

Meynecke, J.-O., Grubert, M., Arthur, J.M., Boston, R., Lee, S.Y., 2012. The influence of the La Niña-El Niño cycle on giant mud crab (*scylla serrata*) catches in northern Australia. Estuar. Coast. Shelf Sci. 100, 93–101. http://dx.doi.org/10.1016/j.ecss.2012.01.001.

Molony, B.W., Thomson, D.P., Feng, M., 2021. What can we learn from the 2010/11 western Australian marine heatwave to better understand risks from the one forecast in 2020/21? Front. Mar. Sci. 8, http://dx.doi.org/10.3389/fmars.2021.645383.

Oliver, E.C.J., 2019. Mean warming not variability drives marine heatwave trends. Clim. Dynam. 53 (3–4), 1653–1659.

O'Neill, M.F., Yang, W.-H., Wortmann, J., Courtney, A.J., Leigh, G.M., Campbell, M.J., Filar, J.A., 2020. Stock Predictions and Population Indicators for Australia's East Coast Saucer Scallop Fishery. FRDC Project No. 2017/057, Fisheries Research and Development Corporation.

Potgieter, A., Hammer, G., Butler, D., 2002. Spatial and temporal patterns in Australian wheat yield and their relationship with ENSO. Aust. J. Agric. Res. 53 (1), 77–89.

Quinn, T.J., Deriso, R.B., 1999. Quantitative Fish Dynamics. Oxford University Press.

Richardson, A., Eriksen, R., Moltmann, T., Hodgson-Johnston, I., Wallis, J.R., 2020. State and trends of Australia's ocean report. In: Integrated Marine Observing System. Hobart, Tasmania.

Rimmington, G.M., Nicholls, N., 1993. Forecasting wheat yields in Australia with the southern oscillation index. Aust. J. Agric. Res. 44 (4), 625–632.

RStudio Team, 2021. RStudio: Integrated Development Environment for R. RStudio, PBC, Boston, MA.

Rundle, J.B., Stein, S., Donnellan, A., Turcotte, D.L., Klein, W., Saylor, C., 2021. The complex dynamics of earthquake fault systems: New approaches to forecasting and nowcasting of earthquakes. Rep. Prog. Phys. 84 (7).

Sen Gupta, A., Thomsen, M., Benthuysen, J.A., Hobday, A.J., Oliver, E., Alexander, L.V., Burrows, M.T., Donat, M.G., Feng, M., Holbrook, N.J., et al., 2020. Drivers and impacts of the most extreme marine heatwave events. Sci. Rep. 10 (1), 19359.

Strydom, S., Murray, K., Wilson, S., Huntley, B., Rule, M., Heithaus, M., Bessey, C., Kendrick, G.A., Burkholder, D., Fraser, M.W., et al., 2020. Too hot to handle: Unprecedented seagrass death driven by marine heatwave in a World Heritage Area. Global Change Biol. 26 (6), 3525–3538.

Thompson, D.W.J., Wallace, J.M., 1998. The arctic oscillation signature in the wintertime geopotential height and temperature fields. Geophys. Res. Lett. 25 (9), 1297–1300.

Thomson, R.E., Emery, W.J., 2014. Data Analysis Methods in Physical Oceanography, third ed. Elsevier.

Tsay, R.S., 2005. Analysis of Financial Time Series, second ed. John Wiley & Sons.

Ummenhofer, C.C., Meehl, G.A., 2017. Extreme weather and climate events with ecological relevance: a review. Philos. Trans. R. Soc. B 372 (1723), http://dx.doi.org/10.1098/rstb.2016.0135.

Unal, Y.S., Tan, E., Mentes, S.S., 2013. Summer heat waves over western Turkey between 1965 and 2006. Theor. Appl. Climatol. 112, 339–350.

von Brömssen, C., Betnér, S., Fölster, J., Eklöf, K., 2021. A toolbox for visualizing trends in large-scale environmental data. Environ. Model. Softw. 136, 104949. http://dx.doi.org/10.1016/j.envsoft.2020.104949.

Wan, C., Dang, P., Gao, L., Wang, J., Tao, J., Qin, X., Feng, B., Gao, J., 2022. How does the environment affect wheat yield and protein content response to drought? A meta-analysis. Front. Plant Sci. 13.

Wang, G., Cai, W., 2020. Two-year consecutive concurrences of positive Indian ocean dipole and central Pacific El Niño preconditioned the 2019/2020 Australian "black summer" bushfires. Geosci. Lett. 7 (1), 1–9.

Wheeler, M.C., Hendon, H.H., 2004. An all-season real-time multivariate MJO index: development of an index for monitoring and prediction. Mon. Weather Rev. 132 (8), 1917–1932.

Wood, S.N., 2017. Generalized Additive Models: An Introduction with R, second ed. In: Chapman & Hall/CRC Texts in Statistical Science, CRC Press.

Wortmann, J., O'Neill, M.F., Courtney, A.J., Yang, W.-H., 2020. Stock Assessment of Ballot's Saucer Scallop (*ylistrum balloti*) in Queensland. Technical Report, State of Queensland.

Wortmann, J., O'Neill, M.F., Sumpton, W.D., Campbell, M.J., Stewart, J., 2018. Stock Assessment of Australian East Coast Snapper, *chrysophrys auratus*. Predictions of Stock Status and Reference Points for 2016. Technical Report, State of Queensland.

Yang, W.-H., Wikle, C.K., Holan, S.H., Wildhaber, M.L., 2013. Ecological prediction with nonlinear multivariate time-frequency functional data models. J. Agric. Biol. Environ. Stat. 18 (3), 450–474. http://dx.doi.org/10.1007/s13253-013-0142-1.

Yao, Y., Wang, C., 2021. Variations in summer marine heatwaves in the south China sea. J. Geophys. Res.: Oceans 126 (10), http://dx.doi.org/10.1029/2021JC017792.

Yuan, C., Yamagata, T., 2015. Impacts of IOD, ENSO and ENSO modoki on the Australian winter wheat yields in recent decades. Sci. Rep. 5 (1), 1–8. http://dx.doi.org/10.1038/srep17252.

Zheng, B., Chapman, S., Chenu, K., 2018. The value of tactical adaptation to El Niño–southern oscillation for east Australian wheat. Climate 6 (3).