EWSmethods: an R package to forecast tipping points at the community level using early warning signals and machine learning models

Duncan O'Brien[1], Smita Deb[2], Sahil Sidheekh[3], Narayanan Krishnan[4], Partha S. Dutta[5], and Christopher clements[1]

[1]University of Bristol
[2]IIT Ropar
[3]The University of Texas at Dallas
[4]IIT Palakkad
[5]Department of Mathematics, Indian Institute of Technology Ropar, Rupnagar

November 10, 2022

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

Abstract

Early warning signals (EWSs) represent a potentially universal tool for identifying whether a system is approaching a tipping point, and have been applied in fields including ecology, epidemiology, economics, and physics. This potential universality has led to the development of a suite of computational approaches aimed at improving the reliability of these methods. Classic methods based on univariate data have a long history of use, but recent theoretical advances have expanded EWSs to multivariate datasets, particularly relevant given advancements in remote sensing. More recently, novel machine learning approaches have been developed but have not been made accessible in the R environment. Here, we present EWSmethods – an R package that provides a unified syntax and interpretation of the most popular and cutting edge EWSs methods applicable to both univariate and multivariate time series. EWSmethods provides two primary functions for univariate and multivariate systems respectively, with two forms of calculation available for each: classical rolling window time series analysis, and the more robust expanding window. It also provides an interface to the Python machine learning model EWSNet which predicts the probability of a sudden tipping point or a smooth transition, the first of its form available to R users. This note details the rationale for this open-source package and delivers an introduction to its functionality for assessing resilience. We have also provided vignettes and an external website to act as further tutorials and FAQs.

Abstract

Early warning signals (EWSs) represent a potentially universal tool for identifying whether a system is approaching a tipping point, and have been applied in fields including ecology, epidemiology, economics, and physics. This potential universality has led to the development of a suite of computational approaches aimed at improving the reliability of these methods. Classic methods based on univariate data have a long history of use, but recent theoretical advances have expanded EWSs to multivariate datasets, particularly relevant given advancements in remote sensing. More recently, novel machine learning approaches have been developed but have not been made accessible in the R environment. Here, we present *EWSmethods* – an R package that provides a unified syntax and interpretation of the most popular and cutting edge EWSs methods applicable to both univariate and multivariate time series.*EWSmethods* provides two primary functions for univariate and multivariate systems respectively, with two forms of calculation available for each: classical rolling window time series analysis, and the more robust expanding window. It also provides an interface to the Python machine learning model EWSNet which predicts the probability of a sudden

tipping point or a smooth transition, the first of its form available to R users. This note details the rationale for this open-source package and delivers an introduction to its functionality for assessing resilience. We have also provided vignettes and an external website to act as further tutorials and FAQs.

Background

Natural systems are inherently non-linear and consequently challenging to forecast . This has led to the application of dynamic system theory which aims to provide model-free and generic tools to identify approaching non-linearity . The majority of this work has attempted to detect critical slowing down (CSD), a phenomenon displayed by systems as they approach a bifurcation or 'tipping point' . In brief, CSD manifests when, as the distance to the tipping point decreases, the ability of the system to recover from perturbations and return to its average trend also decreases . This stems from the dominant eigenvalue of the system trending towards zero and results in successive snapshots in time being more similar than if the system was far from a tipping point; in practical terms, successive abundance/biomass measurements in time or space begin to correlate more strongly.

Detecting CSD can be as simple as tracking the temporal change in summary statistics. For example, increasing autocorrelation at lag-1 , increasing variance , increasing skewness and kurtosis are all representative of CSD. In univariate data, each of these have been successful in identifying oncoming tipping points in simulated experiments as well as empirical lake regime shifts , boreal forest loss , disease (re)emergence , and psychopathology (McSharry et al. 2003, Schreuder et al. 2020). The popularity and scope of EWSs is consequently expanding to new applications and multivariate data sources to maximise the utility of the approach with the increasingly large amounts of ecological monitoring data now available . There is therefore a general desire to exploit EWSs in both traditional research and policy decision making as evidenced by the rapid increase in the publication and citation of EWS literature per year (Figure 1).

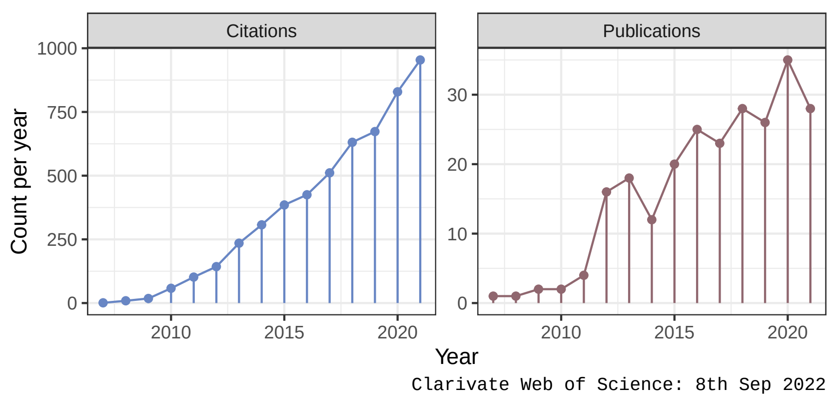

Figure 1: Increase in citation and publication of research articles related to early warning signals. Data was extracted from Clarivate's Web of Science using the search term "early warning signal*" AND "resilience".

Multivariate forms of EWSs (Weinans et al. 2019, Lever et al. 2020, Medeiros et al. 2022) and deep learning models are of particular interest as they appear superior tools to the univariate signals described above. Multivariate approaches exploit information from multiple measurements of a shared system (e.g. multiple species in an ecosystem or multiple sensors in a combustion engine) to provide an overall signal of system resilience. Pooling information in this way buffers against the uncertainty of choosing which data source should be assessed. For example, the trophic level of EWS assessment influences the strength of signal observed in simulated communities (Patterson et al. 2021), and whilst the authors provide guidance on the optimum species/time series to monitor, the required information to identify those time series may not be

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

available to empirical users. Multivariate EWSs can therefore provide a naïve yet robust assessment for multivariate data in the absence of complete information.

Similarly, CSD may not be the only signature of systems close to a tipping point. Our identification of the phenomenon stems from linear stability analysis (LSA) of mathematical models (Ludwig et al. 1978, Scheffer et al. 2009), but machine learning tools can identify other phenomenological features not detected from LSA. For example, machine learning models trained upon transitioning data outperform equivalent models trained upon the EWSs of the same transitioning data (Deb et al. 2022). This is indicative of alternative features being more informative than CSD to warrant general usage, although the 'black-box' nature of the approach limits its accountability (Enni and Herrie 2021). This being said, multiple machine learning models are now available for transitioning systems that improve the transparency of predictions by training on simple mathematical models associated with LSA (Bury et al. 2021, Deb et al. 2022). These models can consequently build upon our foundational knowledge of tipping points by taking advantage of the biases inherent in their training.

Currently, neither multivariate nor machine learning approaches have functionality for R users and resultingly there is a need for simple tools to interact with the variety of EWS approaches available to researchers. Certain EWS functionality has previously been provided by the *earlywarnings* R package , however the package is limited to one form of EWS calculation (rolling windows) in univariate data only. There have also been advances using alternative methodologies such as expanding window and composite EWSs, which introduce data in an add-one-in fashion to provide a standardised time series of EWS strength . This second approach improves the reliability of EWS predictions in univariate data but is not currently available in an easy-to-use form. Unfortunately, many of the custom functions written to facilitate this research are limited to the subscription MATLAB product or hidden in publications' supplementary information . In combination, this has limited the accessibility of EWS development to the wider community.

Compiling these various functions in to a single and comprehensive R package whilst rectifying computational errors is required to increase reproducibility of empirical ecological tipping point research and improve the interpretation and visualisation of results. We therefore designed the *EWSmethods* R package to provide a suite of 'user-friendly' functions to predict critical transitions across both univariate and multivariate data sources and provide interpretable graphics. For univariate data, such as local fisheries or country level disease cases, EWSs can be estimated using either the rolling window approach of *earlywarnings* or the expanding window approach of Drake and Griffen (2010). The package also provides the user the capability to query the Python based EWSNet deep learning model in the R environment and generate predictions on the time series' future. And finally, if multiple measurements have been made of a single system – such as when monitoring multiple species in the same community – multivariate EWSs can be estimated using either rolling or expanding window approaches. *EWSmethods* therefore represents a compilation of new and existing tools to support this expanding field in an easy to use and interpret form. A comparison of the features *EWSmethods*provides versus the currently available *earlywarnings* package is provided in Table 1.

Feature	earlywarnings	EWSmethods
Rolling window early warning signals – univariate time series		
Expanding window early warning signals – univariate time series	X	
Rolling window early warning signals – multivariate time series	X	
Expanding window early warning signals – multivariate time series	X	
Machine learning model (EWSNet) – univariate time series	X	
Maximum likelihood model-based approaches – univariate time series		X
Detrended frequency analysis and potentials – univariate time series		X
Sensitivity analysis – univariate time series		X
Fisher information, Jacobian estimates, etc – multivariate time series	X	
Time series detrending		
Time series deseasoning	X	

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.03033336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

In this paper, we first describe the theory underpinning the methods used and the features of the *EWSmethods* package. We then highlight the practical use of the three modules to predict oncoming transitions using a simulated multi species dataset.

Methods and features

Time series data is the foundation of system monitoring and forecasting, leading to a massive diversity of time series forecasting methods and models developed to analyse them (De Gooijer and Hyndman 2006). Critical slowing down based indicators (i.e. the early warning signals) are no exception but require less technical expertise than traditional forecasting techniques (Dakos et al. 2015). This simplicity in calculation holds for both univariate and multivariate assessments.

a) univariate early warning signals

Early warning signals developed for univariate data are the simplest form of CSD assessment and thus have received the most research effort. Table 2 describes the most common EWSs, all of which are provided in*EWSmethods* via the **uniEWS** function, and how they are calculated. Each of these are also provided in the *earlywarnings*package and mathematically described in detail (Dakos et al. 2012a). The development that *EWSmethods* provides over that package is the diversity of approaches used to compute these EWSs beyond those available in *earlywarnings* , allowing users to tailor their analyses to support their use case. This primarily involves the choice of rolling versus expanding windows during calculation (Figure 2).

Rolling windows

The rolling window approach partitions the univariate time series of interest into a window of data points within which each indicator is estimated. The window then 'rolls' along the time series one data point at a time to update the indicator estimate and generate a new time series of EWSs (Figure 2a). From this EWS time series, the Kendall's Tau correlation of the EWS against time is used to generate 'warnings' (Figure 2b). Specifically, if a strong Tau correlation is found, this indicates an oncoming transition. The **uniEWS** function allows the user to specify the window size as a percentage of the time series' length and returns both the time series of EWSs and the estimated Kendall's Tau to be interpreted.

Indicator	**Description**
SD (Standard Deviation)	Increasing variance/standard deviation is observed approaching a transition, driven by Cri
CV (Coefficient of Variation)	Equivalent to SD as is simply SD at time t divided by the mean SD of the time series.
AR1 (Autocorrelation at lag1)	Autocorrelation (similarity between successive observations) increases approaching a transi
Skewness	At a transition, the distribution of values in the time series can become asymmetric. This
Kurtosis	Kurtosis represents the system reaching more extreme values in the presence of a transitio
Return rate	The inverse of the first-order term of a fitted autoregressive AR(1) model. Return rate is t
Density ratio	Spectral reddening (high variance at low frequencies) occurs near transition. The density r

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

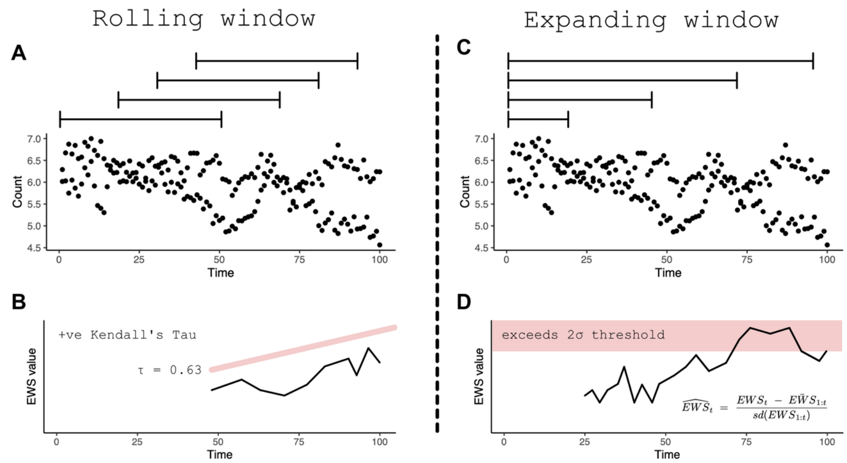

Figure 2: Visual representation of the difference between rolling and expanding window approaches to calculating early warning signals (EWSs - A vs C) in a hypothetical transitioning time series. Solid bars indicate the changing window. Panels B and D then indicate the quantity that represents a 'warning'. For rolling windows (A, B), this warning is a strong Kendall's Tau correlation of EWS indicator values with time. Whereas, for expanding windows (C, D) a warning occurs when the standardised EWS value exceeds a 2σ threshold.

Expanding windows

The alternative to the above computation differs by assessing change in an expanding window via a composite metric consisting of multiple indicators (Figure 2c). The same EWS indicators as above are available to the expanding window approach (Table 2), but each indicator is standardised by subtracting its expanding mean from its calculated value at time t . This value is then normalised by division by its expanding standard deviation (Drake and Griffin 2010) – at each time point, the prediction is updated (Figure 2d). A composite metric can then be constructed by summing all individual indicator values calculated per t . The resulting indicator value or score is hereafter referred to as 'strength'. If the indicator strength exceeds a threshold value, then a 'signal' has been identified. Typically, this threshold value is 2σ which is approximately equivalent to a 95% confidence interval and performs favourably compared to other threshold levels (Clements and Ozgul 2016, Clements et al. 2017).

The expanding window approach also allows multiple information sources to contribute to the assessment. For example, including body size estimates improves assessment reliability by reducing false positive rate whilst increasing the number of true positives (Clements and Ozgul 2016, Baruah et al. 2020). **uniEWS** consequently accepts a trait argument where an additional trait time series can be combined with the other 'abundance-based' EWSs as a composite metric.

Furthermore, the EWSs assessed using the expanding window approach can be improved using a consecutive signal strategy (Clements et al. 2019, Southall et al. 2022) where a 'warning' is only acknowledged when two or more signals are identified in a row. Southall and colleagues (2022) have recently showed that using this approach results in earlier and more reliable warnings over the rolling window approach.

b) multivariate early warning signals

The second module contained in *EWSmethods* is the expansion of EWSs to multivariate data. The benefit of using multivariate techniques over univariate is that assessments of stability and proximity to tipping points

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

can be performed at the system/community level rather than being constrained to the population level. Many of these multivariate EWSs have been tested and supported by Weinans et al. (2021) but open-source tools to calculate them remain unavailable. *EWSmethods* consequently provides multivariate EWS calculation via the **multiEWS** function.

There are two forms of EWS indicators appropriate for multivariate data: those averaged across all time series representing the system of interest (Dakos 2018), and those calculated from a dimension reduction (Held and Kleinen 2004, Weinans et al. 2019). The former is a simple technique to implement using just **uniEWS** but can be influenced by outlier time series, whereas the latter can display informative properties not identifiable in individual time series (Weinans et al. 2021). Unfortunately, their theoretical relationship with CSD is less well understood. *EWSmethods* and the **multiEWS** function therefore provides 12 multivariate indicators across both averaging and dimension reduction forms, each of which is described in Table 3.

Parameterisation of **multiEWS** is identical to **uniEWS**apart from the lack of capability for composite EWSs. This is due to it being currently unknown how combining multivariate EWS indicators influences their prediction reliability. Rolling and expanding windows are still available for multivariate EWSs and their interpretation remains the same as their univariate equivalents.

Indicator	**Description**
Mean SD (Standard Deviation)	Average variance across all time series representing the syst
Max SD	The variance of the time series with the highest variance of
Mean AR1 (Autocorrelation at lag1)	Average autocorrelation across all time series representing t
Max AR1	The autocorrelation of the time series with the highest auto
Dominant MAF (maximum autocorrelation factor) eigenvalue	The minimum eigenvalue of the system following MAF dim
MAF AR1	The autocorrelation of the data projected on to the first M
MAF SD	The variance of the data projected on to the first MAF – i.
First PC (principal component) AR1	The autocorrelation of the data projected on to the first PC
First PC SD/ Explained variance	The variance of the data projected on to the first PC – i.e.
Dominant eigenvalue of the covariance matrix	The maximum eigenvalue of the covariance matrix between
Maximum covariance	The maximum value of the covariance matrix between all r
Mutual information	A measurement of multi-information or how much each tim

c) machine learning model - EWSNet

The final *EWSmethods* module is an interface to the Python based EWSNet, a deep learning modelling framework for predicting critical transitions and tipping points (Deb et al. 2022). EWSNet consists of coupled long short-term memory and fully convolutional network sub-module routines, which together extract complex nonlinear patterns from inputted time series to provide forecasts on the likelihood of oncoming tipping points. Details on the precise formulation and model structure can be found at Deb et al. (2022) and https://ewsnet.github.io, whereas here we will focus on the application of EWSNet for ecologists and the setup of the R environment to cooperate with EWSNet's Python backend.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

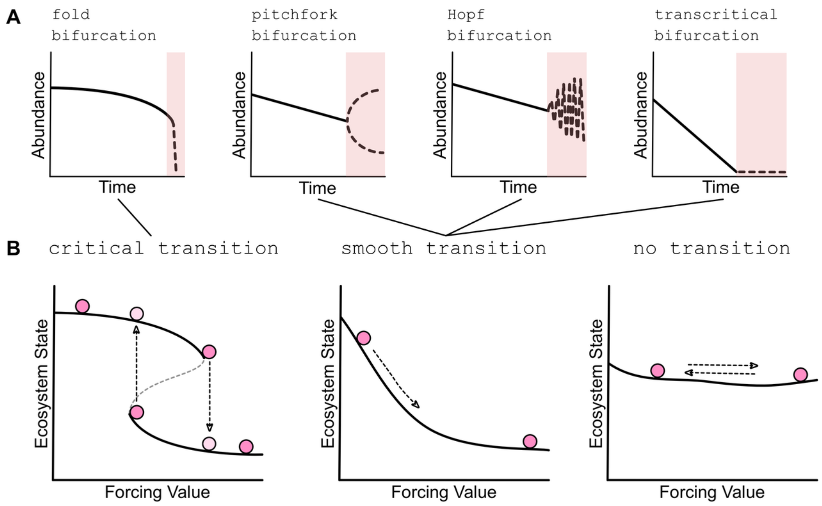

Figure 3: Visual representation of the four models EWSNet was trained (A) and their associated outcome in empirical time series (B). In panel A, the shaded region represents the period of transition with hatched lines indicate the new system trajectory. In panel B balls represent the position of the system of interest in a one dimensional stability landscape.

The rationale behind EWSNet stems from the rapid success and widespread adoption of machine learning algorithms and their ability for learning patterns from data (Humphries et al. 2018). EWSNet exploits this ability by training models upon the simple non-linear mathematical models pioneered by ecological dynamic system research (Ludwig et al. 1978, Fraedrich 1978, Cheng et al. 2008, Scheffer et al. 2012, Kéfi et al. 2013). Specifically, these models encompass four forms of transition/tipping point - saddle-node (fold), pitchfork, supercritical Hopf, transcritical (Figure 3a) – and include non-transitions to allow EWSNet to identify periods of stability. This combination of training results in three possible EWSNet predictions: critical transition, smooth transition or no transition. To aid interpretation of these predictions in real world systems, we suggest that a critical transition indicates oncoming sudden non-linearity, a smooth transition indicates a directional change in trend, and no transition indicates stability as outlined in Figure 3b.

With machine learning tools limited for R users, and EWSNet written in the Python language, the *reticulate* R package (Ushey et al. 2022) allows *EWSmethods* to call the Python functions required to load EWSNet and make predictions from user data. *EWSmethods* prepares the user's R session to perform this interfacing via the **ewsnet_init** function. **ewsnet_init** loads a previously created Python environment with the Python packages required by EWSNet, or installs Python and initialises a new environment if either Python or the environment is not found. Due to the large file sizes being downloaded at this stage, **ewsnet_init** is verbose by default and requires user input to confirm that Python, the required packages, and environment should be downloaded and/or installed.

Users can then use **ewsnet_predict** to generate EWSNet predictions on a time series of interest. To date, EWSNet only supports only univariate time series, however the multivariate form of EWSNet is under active development. The current version of EWSNet also differs to that of the original authors by being robust to time series of variable length. This involved retraining using randomly sampled subsets of the data, ranging in length from 15 to 400 data points to better support the shorter time series available to empirical ecologists. Similarly, due to the variable magnitudes of ecological measurements, two sets of EWSNet's training weights are provided in *EWSmethods* , scaled vs unscaled (**ewsnet_reset** is required to download them); scaled models rescale the input data into the range 1-2. We recommend using scaled weights as they result in more

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.03033336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

reliable model predictions following comparison (O'Brien et al, in prep). **ewsnet_predict** then returns a prediction probability for each of the three potential outcomes ranging from 0.0 – 1.0. As EWSNet was trained on three possible outcomes, a probability of ˜0.33 indicates all prediction outcomes are equally likely (1.0 divided by 3 equals ˜0.33). Therefore, its authors suggest any probability greater than 0.33 implies a stronger than chance prediction and anything greater than 0.6 warrants serious scrutiny (Deb et al. 2022).

Interpretation

Early warning signals are potentially powerful tools for managers. However, their interpretation can be complex and requires nuance. This is particularly true for rolling window approaches and EWSNet as it remains unclear what constitutes a 'strong' correlation or prediction probability. We however believe there are three approaches to defining an appropriate warning using EWSs. Firstly, a user may refer to a reference period for a baseline correlation, or track change in the strength of a signal through time (as in the expanding window approach above), where deviations from the general trend are informative. The second requires the user to define how conservative an assessment they require. For example, if the negative consequence of a transition is significantly larger than the consequence of acting upon a false positive, then a lower confidence warning may be appropriate (i.e. a low Kendall's Tau coefficient/EWSNet prediction probability). And finally, the third requires comparing the observed signal to a distribution of signals generated via permutation of the original time series. If the observed signal is in the top x-th quantile of the distribution (the 95th quantile is commonly used) then a warning may be identified. Alternatively, a fourth option is applicable for EWSNet following the original authors' suggestions, where a probability larger than 0.33 (the chance that all outcomes are equally likely) is indicative of an approaching transition (Deb et al. 2022).

Example

We can illustrate the three modules of the *EWSmethods* package using one of the two datasets bundled with the package: **simsTransComms**. **simsTransComms** contains three replicate communities of five species each, simulated from a competitive Lotka-Volterra model following Dakos (2018). Each community is driven through a tipping point by increasing the carrying capacity of a low density species which mimics the appearance of an invasive species in the community. The time index of the tipping points is provided in the inflection_pt column. It is key to truncate this data set to only contain data prior to this tipping point for EWSs to have any meaningful value as a sentinel of transition (Dale and Beyeler 2001, Gsell et al. 2016). This can be achieved using the inflection_pt column of the **simTransComms$'1_5_1'** community.

```
data(simTransComms)
pre_simTransComms <- subset(simTransComms$'‘1_5_1‘, time < inflection_pt)
```

This represents the data frame we will use for the remainder of this example section (Figure 4). More detailed examples are available at: https://duncanobrien.github.io/EWSmethods/articles/ews_assessments.html and https://duncanobrien.github.io/EWSmethods/articles/using_ewsnet.html.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

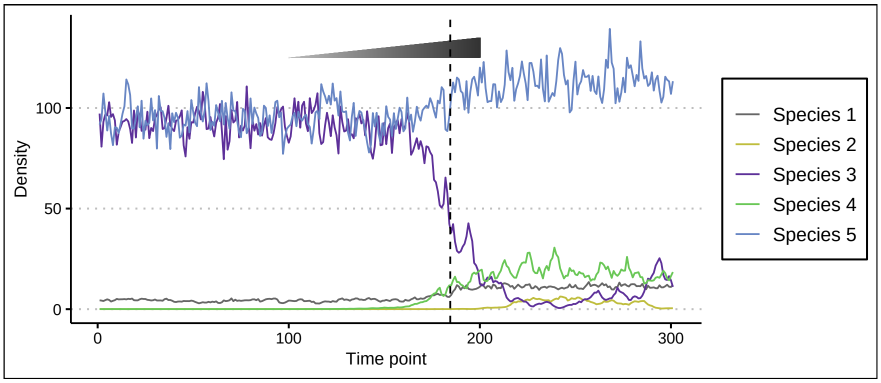

Figure 4: The simulated simTransComms$‘1_5_1‘ community plotted against time. Species 4’s carrying capacity is gradually increased within the time interval 100-200 (as represented by the expanding wedge) to mimic the appearance of an invasive species. This drives a community transition with the inflection point indicated by a vertical dashed line.

Early warning signals

To calculate univariate EWS for any one time series from this community, we would use **uniEWS** . We first need to select the EWS indicators of interest to provide to the metrics argument. Autocorrelation (“ar1”) and variance (represented by the standard deviation - “SD”) are the most commonly used EWSs and have the largest body of research defining their best utility (Carpenter and Brock 2006, Dakos et al. 2012b, Patterson et al. 2021). Using these metrics, we then choose the time calculation approach (expanding), the resulting burn in period (50 data points) and the sigma threshold (two), and that we want **uniEWS** to return a visualisation (ggplotIt = TRUE). **uniEWS** only performs assessments on univariate data but requires a two column data frame where the first column is an equally spaced time vector and the second is the time series to be assessed. We have chosen the third species here.

```
expanding_ews_eg <- uniEWS(data = pres_simTransComms[,c(2,5)], metrics = c("ar1","SD"), method = "expan
```

The resulting *ggplot* (Wickham 2016) (Figure 5) shows that warnings are generated from timepoint 171 onwards for all EWSs following multiple consecutive ‘signals’. The single signal for the “ar1 + SD” indicator at timepoint 115 is not sufficient to be a warning.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

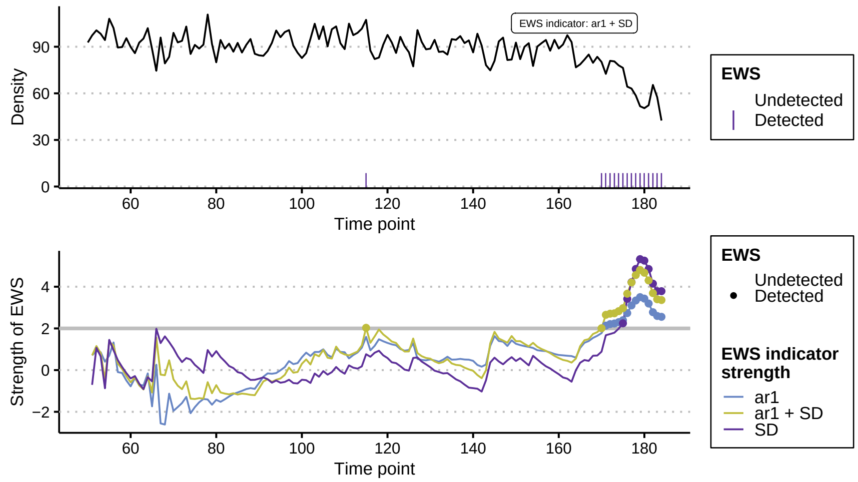

Figure 5: Expanding window assessment of the pre_simTransComms[,c(2,5)] time series using the univariate autocorrelation and variance early warning signal indicators. The figure is a direct output of the *EWSmethods* function **uniEWS**. The top panel depicts the raw time series and the presence of a signal from the annotated indicator. The lower panel visualises the strength of each indicator through time and the threshold level. A signal is indicated when the indicator strength exceeds this threshold value.

To expand the assessment to include information from all time series, we require the use of **multiEWS** . The single difference in the function's parameterisation is that the input data frame must contain more than two columns (one time sequence column and two or more time series). By default, all indicators are returned.

To compute multivariate EWSs using the rolling method, the function would be written as thus, specifying the method and winsize as a percentage of the time series' length:

```
multi_ews_eg_roll <- multiEWS(data = pre_simTransComms[,2:7], method = "rolling", winsize = 50, ggplotI
```

All indicators are positively correlated with time (excluding mafSD) but the strength of correlation varies (mean Tau = 0.55, Figure 6). However, each indicator does increase prior to transition even if this is not universally represented in the Tau coefficients.

To repeat the process using expanding windows, **multiEWS** simply requires a change of method argument and the provision of burn in and threshold.

```
multi_ews_eg_expand <- multiEWS(data = pre_simTransComms[,2:7], method = "expanding", burn_in = 50, thre
```

Warnings are generated throughout this assessment with two consistently signalled periods at timepoints 110 and 175 (Figure 7). This highlights the usefulness of expanding windows over rolling as the exact time point of warning can be determined, but supports Weinans *et al.* 's (2021) suggestion that there is no superior multivariate EWS indicator; the best fit depends on the scenario the system is subject to.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

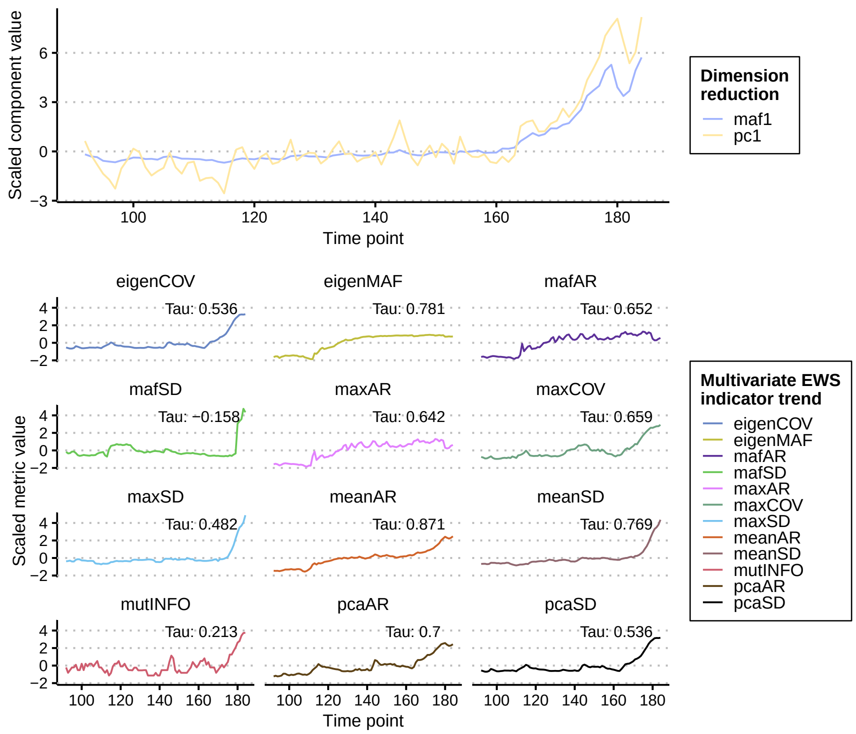

Figure 6: Rolling window assessment of the entire pre_simTransComms community using multivariate early warning signal indicators. The figure is a direct output of the EWSmethods function **multiEWS**. The top panel plots the raw dimension reductions from which certain indicators are estimated. The lower panel visualises the trend in each indicator through time and reports the Kendall's Tau correlation coefficient.

EWSNet

EWSNet requires initialisation using **ewsnet_init** due to its Python backend. At the start of each R session, **ewsnet_init** must be called and a consistent envname provided. When the function is run for the first time on a new machine, Python will be downloaded alongside the critical Python packages and a new environment (envname) created. The user will be prompted to agree to this by default (when the auto argument is FALSE) to ensure the files will not be accidentally downloaded if undesired. For future sessions, providing the same envname will result in the original environment being activated rather than redownloading all files.

```
ewsnet_init(envname = "EWSNET_env", pip_ignore_installed = FALSE, conda_refresh = FALSE, auto = FALSE)
```

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

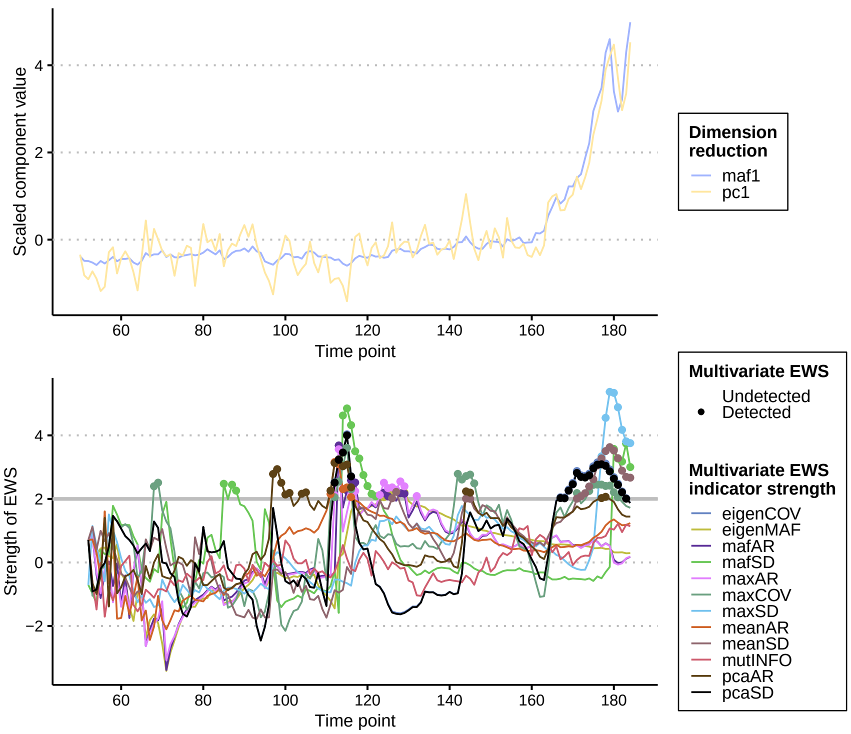

Figure 7: Expanding window assessment of the entire pre_simTransComms community using multivariate early warning signal indicators. The figure is a direct output of the EWSmethods function **multiEWS**.

The large file size of the model weights (~220mb) also means that *EWSmethods* does not come bundled with them. The user is required to call the **ewsnet_reset** function which will prompt confirmation that the weights are to be downloaded from https://ewsnet.github.io.

```
ewsnet_reset(remove_weights = FALSE)
```

Once initiated, **ewsnet_predict** will accept a vector timeseries (note no time sequence is required) alongside the model weights to use. These model weights are subset based on scaling (scaled vs unscaled) and the number of models to average over (ensemble). We recommend using scaled weights averaged over the maximum ensemble size (25) for most robust predictions.

```
ewsnet_prediction <- ewsnet_predict(pre_simTransComms[,5], scaling = TRUE, ensemble = 25, envname = "EWS

print(ewsnet_prediction)
pred                no_trans_prob       smooth_trans_prob       crt_trans_prob
Critical Transition 0.196918            0.1813867               0.6216951
```

A critical transition has subsequently been predicted with a 62% probability indicating that a sudden tipping point is imminent.

Conclusion

The ability to use accessible and easy to interpret tools are key for ecological monitoring. In this note we present *EWSmethods* , an R package consolidating the simplest methods of early warning signal assessments into a coherent suite of metrics and visualisations. Each function is consistent in its parameterisations, terminology, and output to allow any user to interpret the assessment confidently, regardless of the data dimensionality or EWS approach.

It would however be remiss to overlook the pivotal *earlywarnings*package and work of Dakos *et al.* (2012). *EWSmethods*innovates on *earlywarnings* by providing alternative calculations (rolling vs expanding windows) and data types (univariate vs multivariate), but does not provide the additional modelling techniques*earlywarnings* supports (diffusion-drift-jump models, BDS tests etc). We direct readers to that package on github (as it is no longer maintained on CRAN at the time of writing - https://github.com/earlywarningtoolbox/earlywarnings-R) for the typical rolling window EWS approach due to the additional modelling capabilities it provides. *EWSmethods* better supports multivariate analyses and standardises across univariate EWSs, multivariate EWSs and machine learning models to allow comparability. It also provides access to purpose-built machine learning models not otherwise available to R users. Consequently, users are able to explore an ensemble of generic forecasting methods to identify oncoming transitions and tipping points in their system.

Generic approaches also facilitate wider research interest into the universal challenge of identifying oncoming tipping points. Resilience-based approaches are critical for the management of globally imperilled systems (Folke et al. 2010, Oliver et al. 2015, Capdevila et al. 2022) but are applicable in other disciplines. Remotely sensed data could allow global level tipping point assessments for example (Forzieri et al. 2022), individual mortality risk may be detectable (Cailleret et al. 2019) or positive thresholds can be encouraged (Lenton et al. 2022). The low barrier to entry that *EWSmethods* provides for R users can aid the development of these developing research avenues.

To cite *EWSmethods* or acknowledge its use, cite this Software note as follows, substituting the version of the application that you used for 'version 1.0':

O'Brien D.A. et al. 2022. EWSmethods: an R package to forecast tipping points at the community level using early warning signals and machine learning models (ver. 1.0)

References

Baruah, G. et al. 2020. Eco-evolutionary processes underlying early warning signals of population declines. - Journal of Animal Ecology 89: 436–448.

Bury, T. M. et al. 2021. Deep learning for early warning signals of tipping points. - Proceedings of the National Academy of Sciences 118: e2106140118.

Cailleret, M. et al. 2019. Early-warning signals of individual tree mortality based on annual radial growth.

Capdevila, P. et al. 2022. Global patterns of resilience decline in vertebrate populations. - Ecology Letters 25: 240–251.

Carpenter, S. R. and Brock, W. A. 2006. Rising variance: a leading indicator of ecological transition. - Ecology Letters 9: 311–318.

Carpenter, S. R. et al. 2011. Early warnings of regime shifts: A whole-ecosystem experiment. - Science 332: 1079.

Chen, S. et al. 2019. Eigenvalues of the covariance matrix as early warning signals for critical transitions in ecological systems. - Scientific Reports 9: 2572.

Cheng, Z. et al. 2008. Robustness analysis of cellular memory in an autoactivating positive feedback system. - FEBS Lett 582: 3776–3782.

Clements, C. F. and Ozgul, A. 2016. Including trait-based early warning signals helps predict population collapse. - Nature Communications 7: 10984.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

Clements, C. F. et al. 2017. Body size shifts and early warning signals precede the historic collapse of whale stocks. - Nature Ecology & Evolution 1: 188.

Clements, C. F. et al. 2019. Early warning signals of recovery in complex systems. - Nature Communications 10: 1681.

Dakos, V. 2018. Identifying best-indicator species for abrupt transitions in multispecies communities. - Ecological Indicators 94: 494–502.

Dakos, V. et al. 2012a. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. - PLOS ONE 7: e41010.

Dakos, V. et al. 2012b. Robustness of variance and autocorrelation as indicators of critical slowing down. - Ecology 93: 264–271.

Dakos, V. et al. 2015. Resilience indicators: prospects and limitations for early warnings of regime shifts. - Philosophical Transactions of the Royal Society B: Biological Sciences 370: 20130263.

Dale, V. H. and Beyeler, S. C. 2001. Challenges in the development and use of ecological indicators. - Ecological Indicators 1: 3–10.

De Gooijer, J. G. and Hyndman, R. J. 2006. 25 years of time series forecasting. - International Journal of Forecasting 22: 443–473.

Deb, S. et al. 2022. Machine learning methods trained on simple models can predict critical transitions in complex natural systems. - Royal Society Open Science 9: 211475.

Enni, S. A. and Herrie, M. B. 2021. Turning biases into hypotheses through method: A logic of scientific discovery for machine learning. - Big Data & Society 8: 20539517211020776.

Folke, C. et al. 2010. Resilience thinking. - Ecology and Society in press.

Forzieri, G. et al. 2022. Emerging signals of declining forest resilience under climate change. - Nature 608: 534–539.

Fraedrich, K. 1978. Structural and stochastic analysis of a zero-dimensional climate system. - Quarterly Journal of the Royal Meteorological Society 104: 461–474.

Gsell, A. S. et al. 2016. Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. - Proceedings of the National Academy of Sciences 113: E8089 LP-E8095.

Held, H. and Kleinen, T. 2004. Detection of climate system bifurcations by degenerate fingerprinting. - Geophysical Research Letters in press.

Humphries, G. et al. 2018. Machine learning for ecology and sustainable natural resource management.

Kéfi, S. et al. 2013. Early warning signals also precede non-catastrophic transitions. - Oikos 122: 641–648.

Kleinen, T. et al. 2003. The potential role of spectral properties in detecting thresholds in the Earth system: application to the thermohaline circulation. - Ocean Dynamics 53: 53–63.

Kuehn, C. 2011. A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. - Physica D: Nonlinear Phenomena 240: 1020–1035.

Lenton, T. M. et al. 2022. Operationalising positive tipping points towards global sustainability. - Global Sustainability 5: e1.

Lever, J. J. et al. 2020. Foreseeing the future of mutualistic communities beyond collapse. - Ecology Letters 23: 2–15.

Ludwig, D. et al. 1978. Qualitative analysis of insect outbreak systems: the spruce budworm and forest. - Journal of Animal Ecology 47: 315–332.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.03033336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.

McSharry, P. E. et al. 2003. Prediction of epileptic seizures: are nonlinear methods relevant? - Nature Medicine 9: 241–242.

Medeiros, L. P. et al. 2022. Ranking species based on sensitivity to perturbations under non-equilibrium community dynamics. - Ecology Letters in press.

Oliver, T. H. et al. 2015. Biodiversity and resilience of ecosystem functions. - Trends in Ecology & Evolution 30: 673–684.

Patterson, A. C. et al. 2021. When and where we can expect to see early warning signals in multispecies systems approaching tipping points: insights from theory. - The American Naturalist 198: E12–E26.

Quax, R. et al. 2013. Information dissipation as an early-warning signal for the Lehman Brothers collapse in financial time series. - Scientific Reports 3: 1898.

Scheffer, M. et al. 2009. Early-warning signals for critical transitions. - Nature 461: 53–59.

Scheffer, M. et al. 2012. Anticipating critical transitions. - Science 338: 344 LP – 348.

Schreuder, M. J. et al. 2020. Early warning signals in psychopathology: what do they tell? - BMC Medicine 18: 269.

Southall, E. et al. 2022. How early can an upcoming critical transition be detected? - medRxiv: 2022.05.27.22275693.

Strogatz, S. H. 2015. Nonlinear dynamics and chaos: with applications to physics, biology, chemistry, and engineering. - CRC Press.

Suweis, S. and D'Odorico, P. 2014. Early warning signs in social-ecological networks. - PLOS ONE 9: e101851.

Ushey, K. et al. 2022. reticulate: Interface to "Python."

Weinans, E. et al. 2019. Finding the direction of lowest resilience in multivariate complex systems. - Journal of The Royal Society Interface 16: 20190629.

Weinans, E. et al. 2021. Evaluating the performance of multivariate indicators of resilience loss. - Scientific Reports 11: 9148.

Wickham, H. 2016. ggplot2: elegant graphics for data analysis. - Springer-Verlag New York.

Posted on 10 Nov 2022 — MIT — https://doi.org/10.22541/au.166801190.00303336/v3 — This a preprint and has not been peer reviewed. Data may be preliminary.