

Appendix 1–3

Appendix 1

Maximum autocorrelation factor analysis

Maximum autocorrelation factor analysis (MAFA) has clear intuitive appeals as a tool for analyzing multivariate population data (Solow 1994). Theoretical bases for its utility are discussed in the main text and also described by Solow (1994) and Fujiwara (2008). Although it is theoretically sound, to our knowledge, the method has not been applied to simulated data sets to demonstrate its performance. This motivated us to prepare this appendix.

Here, we generated a multivariate population time series that consists of common smooth trends and applied MAFA to the time series to see whether we can recover the common trends. The analysis shown in this appendix is meant to supplement the theory of MAFA. Therefore, we chose an example for which the method is expected to work. On the other hand, we are aware of several circumstances under which the method fails. These are discussed in the Discussion section of this Appendix as well as in the main text in more detail. However, despite these limitations, we believe that MAFA is a simple and very useful tool that should be used more commonly in population data analysis.

Methods

Simulation of multivariate population time series

Population time series were simulated as weighted linear combinations of two hypothetical environmental signals (common factors) and a signal that is specific to each population time series (specific factor). The j th population time series $X_t^{(j)}$ is given by

$$X_t^{(j)} = \beta_{j1} \tilde{Y}_t^{(1)} + \beta_{j2} \tilde{Y}_t^{(2)} + \sigma \tilde{\psi}_j \quad (1.1)$$

where t indicates time, $Y_t^{(i)}$ is the i th common factor, β_{ji} is the loading of the j th variable on the i th factor, ψ_j is a specific factor that is a serially iid normal random variable with mean 0 and variance 1, σ is a coefficient determining the signal to noise ratio of the time series, and a tilde indicates that a corresponding variable was standardized (by subtracting its sample mean and dividing by its sample standard deviation). For all simulations, $\sigma^2 = 2$ so that the variance of the common factors to the total variance was approximately 1/3. We envision $X_t^{(j)}$ to be a variable that is proportional to the count of individuals in a population.

Because we are interested in extracting common factors that are smooth (i.e. variable with a high lag-one autocorrelation), two sine curves with different frequencies and phases were used as common factors $Y_t^{(1)}$ and $Y_t^{(2)}$ (Fig. 1.1). Using the two common factors,

five population time series were produced with different loading values (Table 1.1). These values were chosen so that the sampling variance of the signals was 1.

Identification of environmental signals

In an actual application, we do not know $Y_t^{(i)}$, and our goal is to estimate time series that are proportional to $Y_t^{(i)}$ ($i = 1, 2$) from $X_t^{(j)}$ ($j = 1, \dots, 5$) using MAFA. This technique was applied to the population time series after they were standardized. Then, the estimated common factors $\hat{Y}_t^{(i)}$ were expressed as weighted linear combinations of $\tilde{X}_t^{(j)}$

$$\hat{Y}_t^{(i)} = \hat{\alpha}_{i,1} \tilde{X}_t^{(1)} + \dots + \hat{\alpha}_{i,4} \tilde{X}_t^{(4)} \quad (1.2)$$

where α_{ij} is a coefficient of the i th factor on the j th variable. MAFA chooses the coefficients so that the first weighted linear combination (MAF 1) has the highest lag-one autocorrelation. The second combination (MAF 2) is uncorrelated with the first one and has the second highest lag-one autocorrelation.

Results and discussion

Five simulated population time series are shown in Fig. 1.2. Along with each population time-series, two curves are also plotted in each panel. The first is a 'noise-free' trend, which is the curve obtained using Eq. 1.1 without a specific factor ψ_j . This curve shows the true trend in the population time series. The second is the estimated score, which was obtained by plugging the estimated loading and estimated common factors into Eq. 1.1 without a specific factor ψ_j . This curve shows the variability in the original time series explained by the estimated factors. Two estimated common factors from the data (Fig. 1.2) are shown in Fig. 1.3. The same calculations were repeated multiple times, each time specific factors were generated independently. Because the results for repeated simulations were similar, one additional set of results are shown in Fig. 1.4 and 1.5.

MAFA chooses the weighted linear combination of time series that maximizes its lag-one autocorrelation. In other words, it chooses the combination that reduces the average distance between two successive data points. Consequently, two successive points of the estimated common factors tend to be closer to each other than those of the original population time series (Fig. 1.2–1.5). If the common factors of interest have a high lag-one autocorrelation (i.e. smooth), we can extract them using MAFA. One of the advantages of using MAFA on population data is that many environmental signals in population data tend to have a significantly positive lag-one autocorrelation (Introduction of the main text; Fujiwara 2008).

By applying MAFA, we can extract smooth common factors. In all two cases, the first MAF resembles one of the original com-

mon factors, and the second MAF resembles the other original common factors (Fig. 1.2–1.5). The two estimated factors also explain a large part of the original trends (cf. the ‘noise-free’ trends and estimated scores in Fig. 1.2 and 1.4), suggesting that MAFA is successful in extracting underlying trends.

The population time series $\tilde{X}_t^{(i)}$ were produced by combining the original common factors $\tilde{Y}_t^{(i)}$ and specific factors. Consequently, some population time series resemble the original common factor. Therefore, the correlation between some of the original common factors and some of the population time series may be significant. However, it is difficult to find whether the trend is unique to that population time series or common among other time series. On the other hand, examining the MAFs is more advantageous than directly examining the original population data because the loadings (Fig. 1.4) can be examined to find the inter-relationship among population time-series.

MAFA, however, is not an almighty technique, and one needs to be aware of its limitations. We simulated many multivariate time-series of different types to examine the performance of MAFA (unpubl.). We found that when the number of available time-series is small (e.g. three time series), MAFA often cannot separate common trends. On the other hand, when the number of available time-series is too large (e.g. ten time series), there is an increased chance that some sampled specific factors might also have a strongly positive lag-one autocorrelation, which reduces the ability to separate the common factors. Currently, there is no guideline for the number of population time series to be included in an analysis. We speculate it depends on the shapes of the trends that we try to identify, the length of time-series, and the distribu-

tion of specific factors. Our recommendation is when multivariate population time-series are available, MAFA should be tried because the analysis is very simple.

We are also aware that MAFA often cannot completely separate two common factors that are correlated and have almost equally-smooth trends. This is a problem because the correlation between the estimated and original common factors is reduced because of the interference. Further research is needed to develop a method to reduce this effect; we speculate that a technique such as varimax rotation of the loadings (Manly 2005) may be useful under some circumstances (e.g. each population time series is affected strongly by a small number of factors). For now, we need to be aware of this effect and interpret results carefully.

Despite the above limitations, the potential utility of MAFA is clear. It can extract underlying trends from multivariate time series. There are a large number of short population time series available. Although they may not provide sufficient information if analyzed individually, multiple time series together may give us important ecological insights. Thus, an analysis like MAFA gives us the opportunity to reanalyze existing data sets.

References

- Fujiwara, M. 2008 Identifying Interactions among salmon populations from observed dynamics. – *Ecology* 89: 4–11.
 Manly, B. F. J. 2005 *Multivariate statistical methods – a primer* (3rd ed.). – Chapman and Hall/CRC.
 Solow, A. R. 1994 Detecting change in the composition of a multispecies community. – *Biometrics* 50: 556–565.

Table 1.1. Loading β_{ji} for the population time series on two simulated environmental variables.

	$X_t^{(1)}$	$X_t^{(2)}$	$X_t^{(3)}$	$X_t^{(4)}$	$X_t^{(5)}$
$\tilde{Y}_t^{(1)}$	$(0.6)^{1/2}$	$(0.6)^{1/2}$	$(0.8)^{1/2}$	$-(0.3)^{1/2}$	$(0.5)^{1/2}$
$\tilde{Y}_t^{(2)}$	$-(0.4)^{1/2}$	$(0.4)^{1/2}$	$(0.2)^{1/2}$	$(0.7)^{1/2}$	$(0.5)^{1/2}$

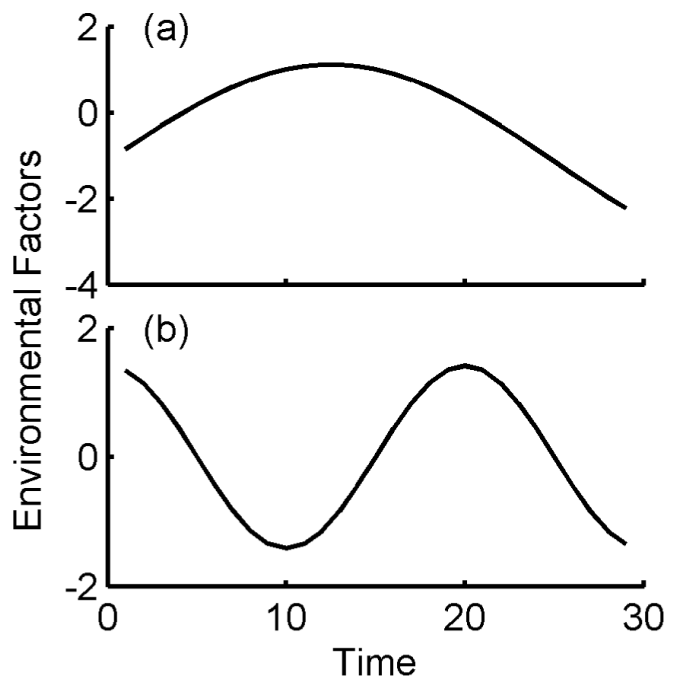


Figure 1.1. Two common factors used for generating population data.

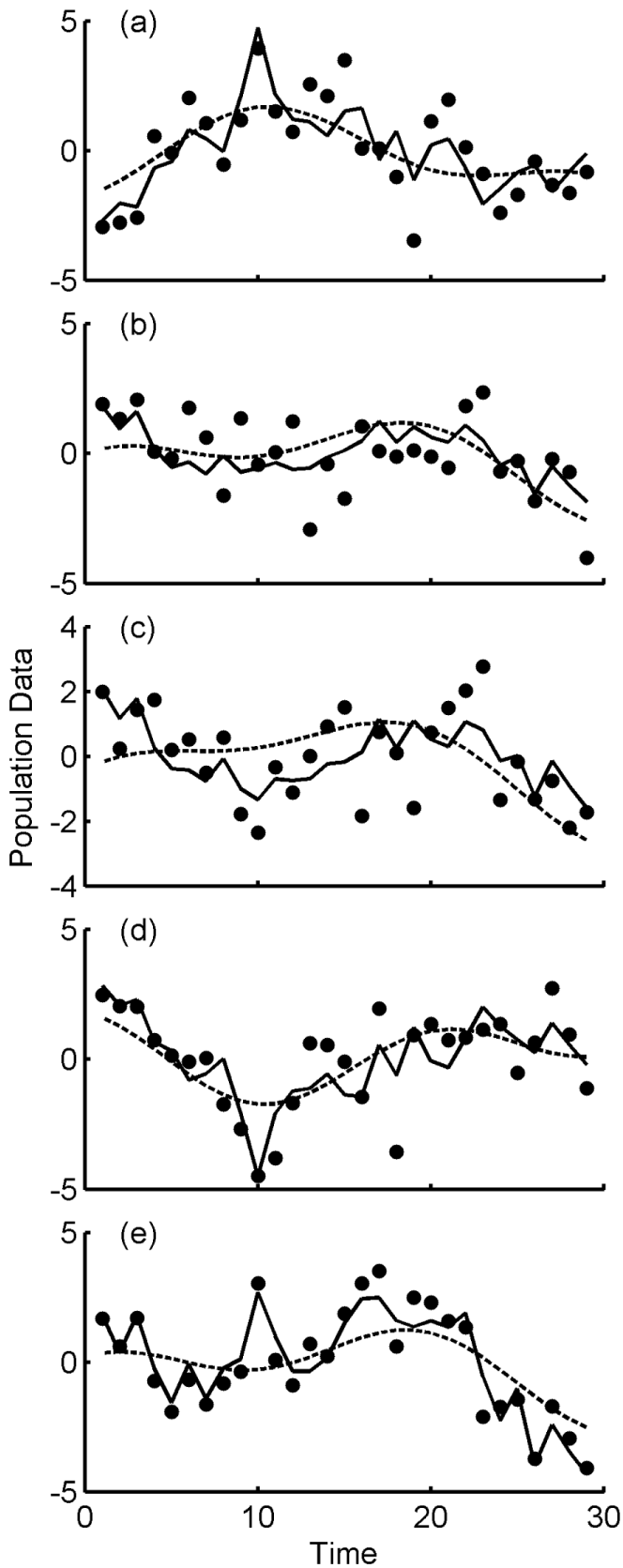


Figure 1.2. Five population time-series. ●: time series generated using the loading in Table 1.1 and the common factors in Fig. 1.1; solid line: estimated population time-series (i.e weighted linear combinations of MAFs); dashed line: noise-free trend (see text for more details).

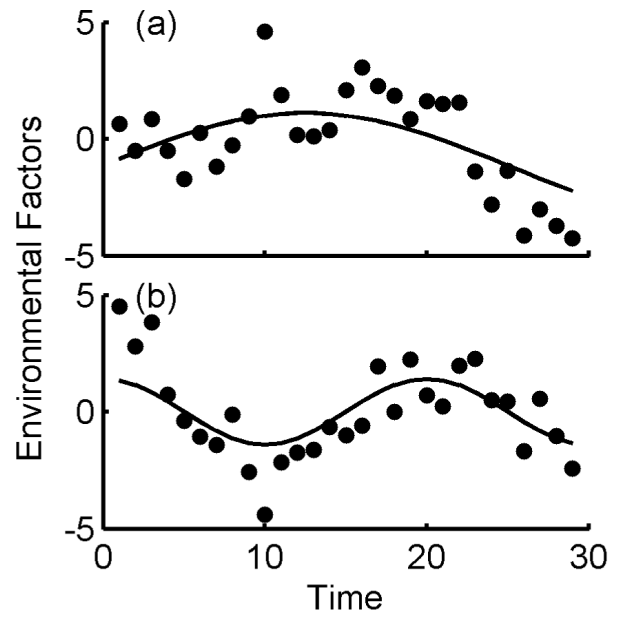


Figure 1.3. Two common factors. ●: estimated common factors (MAFs); solid line: the original common factors shown in Fig. 1.1.

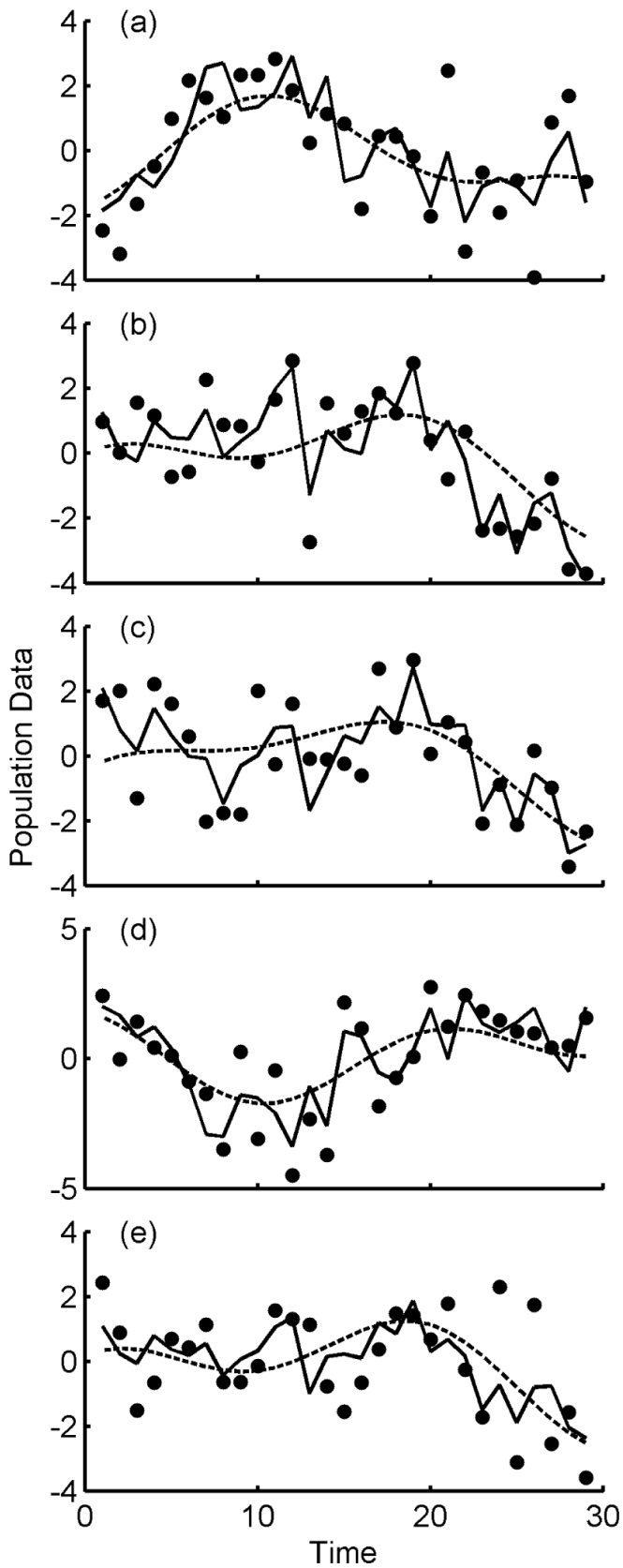


Figure 1.4. Five population time-series. ●: time series generated using the loading in Table 1.1 and the common factors in Fig. 1.1; solid line: estimated population time-series (i.e weighted linear combinations of MAFs); dashed line: noise-free trend (see text for more details).

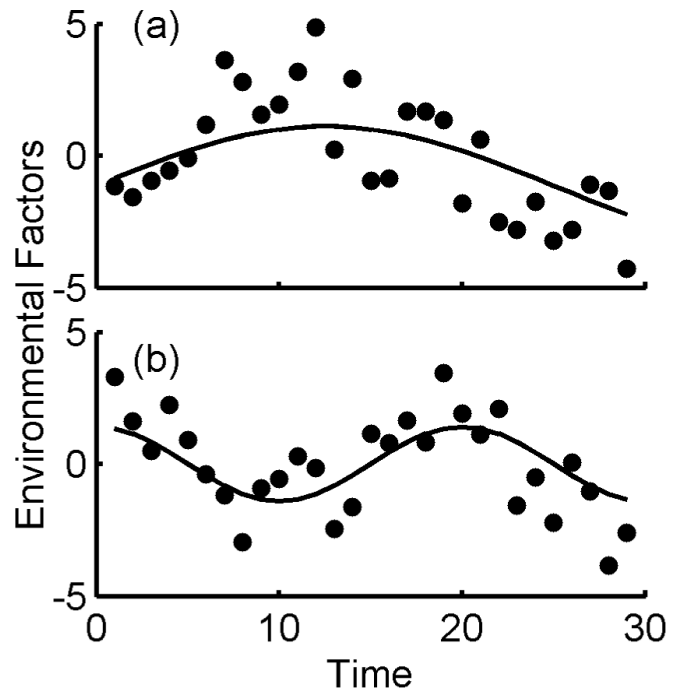


Figure 1.5. Two common factors. ●: estimated common factors (MAFs); solid line: the original common factors shown in Fig. 1.1.

Appendix 2

Environmental variables

In the main text, the maximum autocorrelation factors (MAFs) were extracted from Klamath fall Chinook data, and the MAFs were associated with various environmental variables. The choice of the environmental variables as well as the time lags between the environmental variables and the upstream migration of the Chinook salmon were selected based on the existing knowledge of the life history of the Chinook salmon. In this appendix, these environmental variables and the timing of the environmental events relative to the life history of individuals are described in more detail.

(1) River flow rate

The amount and timing of water flow in the Klamath and Trinity Rivers has been altered from historical levels due to the construction of dams in the Klamath Basin (NRC 2004). This alteration of river flow characteristics is considered to be one of the major problems associated with the decline of many fish species in the Klamath Basin. For Chinook salmon, a higher flow rate is thought to be favorable for their growth, development, and survival because the increased river flow expands the extent of suitable habitats, which in turn reduces competition for space and predation pressure (NRC 2004). The flow rate also affects water temperature (NRC 2004), which in turn affects physiological processes of the fish as well as the prevalence of parasites (Richter and Kolmes 2005). Consequently, it is thought that the river flow rate may strongly influence the survival of salmon in the Klamath Basin.

In the present analysis, the average daily river flow rates (unit-volume per second) over three months in fall (October–December), winter (January–March), and spring (April–June) at Klamath River (Orleans) and at Trinity River (Hoopa), California (Fig. 2.1), were used as variables to represent the river conditions experienced by these salmon (Table 1.1). This time period was chosen because spawning primarily occurs in the October–November period, and their progeny emerge from the river gravel during winter and spring (NRC 2004). The river flow index was lagged by three and four years relative to the escapement data because the river flow rate was expected to affect the survival of juveniles in the rivers; the same cohort later returns to the natal streams for spawning mostly at age 3 or 4. The river flow data were obtained from the United States Geological Survey (<www.usgs.gov/>).

(2) Coastal upwelling index (CUI)

An index of monthly coastal upwelling was used as a measure of ocean condition. This index is a measure of wind-induced coastal upwelling. The intensity of upwelling is thought to affect the survival of salmon by altering their food availability (Gargett 1997). For example, the index has been used to explain variation in salmon ocean fishery harvest, ocean survival, and individual growth of Chinook salmon (Botsford and Lawrence 2002, Scheuerell and Williams 2005, Wells et al. 2006).

In the present analysis, the spring (May and June) and the fall (September and October) monthly upwelling indices at 42°N, 125°W were used (Fig. 2.2). The spring indices were lagged relative to the escapement data so that individuals at various ages in the ocean as well as their parents during the spring just before

spawning were affected. This resulted in five different lags relative to the escapement year. The fall indices were also lagged so that individuals at varying ages in the ocean in fall were affected; this resulted in three different time lags relative to the escapement year (Table 1.1). The coastal upwelling index was obtained from the Environmental Research Division within the Southwest Fisheries Science Center of the National Marine Fisheries Service, USA (<www.swfsc.noaa.gov/>).

(3) Hatchery returns

Hatcheries have been operating along the Klamath River (Iron Gate Hatchery) and the Trinity River (Trinity River Hatchery) to mitigate for the lost spawning habitat following construction of dams on these rivers. The general effects of hatchery fish on naturally spawning stocks have been debated. Hatchery salmon are thought to affect the genetic composition and reduce the viability of naturally spawning salmon because of increased competition (Ruckelshaus et al. 2002). In this analysis, the returns of hatchery fish to both Iron Gate Hatchery and Trinity River Hatchery (Fig. 2.3) were used as explanatory variables to compare with the naturally spawning salmon stocks.

(4) Fishery harvest

Finally, Klamath River fall-run Chinook salmon are harvested in the ocean and in the Klamath Basin, in commercial, recreational, and tribal subsistence fisheries. In this analysis, the estimated total annual harvest of the Klamath River fall-run Chinook in all fisheries (PFMC 2007), for the period 1986 to 2006, was used as a covariate to represent fishery harvest (Fig. 2.4).

References

- Botsford, L. W. and Lawrence, C. A. 2002. Patterns of co-variability among California Current chinook salmon, coho salmon, Dungeness crab and physical oceanographic conditions. – *Progr. Oceanogr.* 53: 283–305.
- Gargett, A. E. 1997. The optimal stability ‘window’: a mechanism underlying decadal fluctuations in North Pacific salmon stocks? – *Fish. Oceanogr.* 6: 109–117.
- NRC – National Research Council of the National Academies 2004 Endangered and Threatened Fishers in the Klamath River basin: Causes of Decline and Strategies for Recovery. National Research Council of the National Academies. – The National Academies Press, Washington DC, USA.
- PFMC – Pacific Fishery Management Council 2007 Preseason report I: stock abundance analysis for 2007 ocean salmon fisheries. Pacific Fishery Management Council, 7700 NE Ambassador Place, Suite 101, Portland, Oregon 97220-1384, USA.
- Richter, A. and Kolmes, S. A. 2005 Maximum temperature limits for chinook, coho, and chum salmon, and steelhead trout in the Pacific Northwest. – *Rev. Fish. Sci.* 13: 23–49.
- Ruckelshaus, M. H. et al. 2002 The Pacific salmon wars: what science brings to the challenge of recovering species. – *Annu. Rev. Ecol. Syst.* 33: 665–706.
- Scheuerell, M. D. and Williams, J. G. 2005. Forecasting climate-induced changes in the survival of Snake River spring/summer Chinook salmon (*Oncorhynchus tshawytscha*). – *Fish. Oceanogr.* 14: 448–457.
- Wells, B. K. et al. 2006 Covariation between the average lengths of mature coho (*Oncorhynchus kisutch*) and Chinook salmon (*O. tshawytscha*) and the ocean environment. – *Fish. Oceanogr.* 15: 67–79.

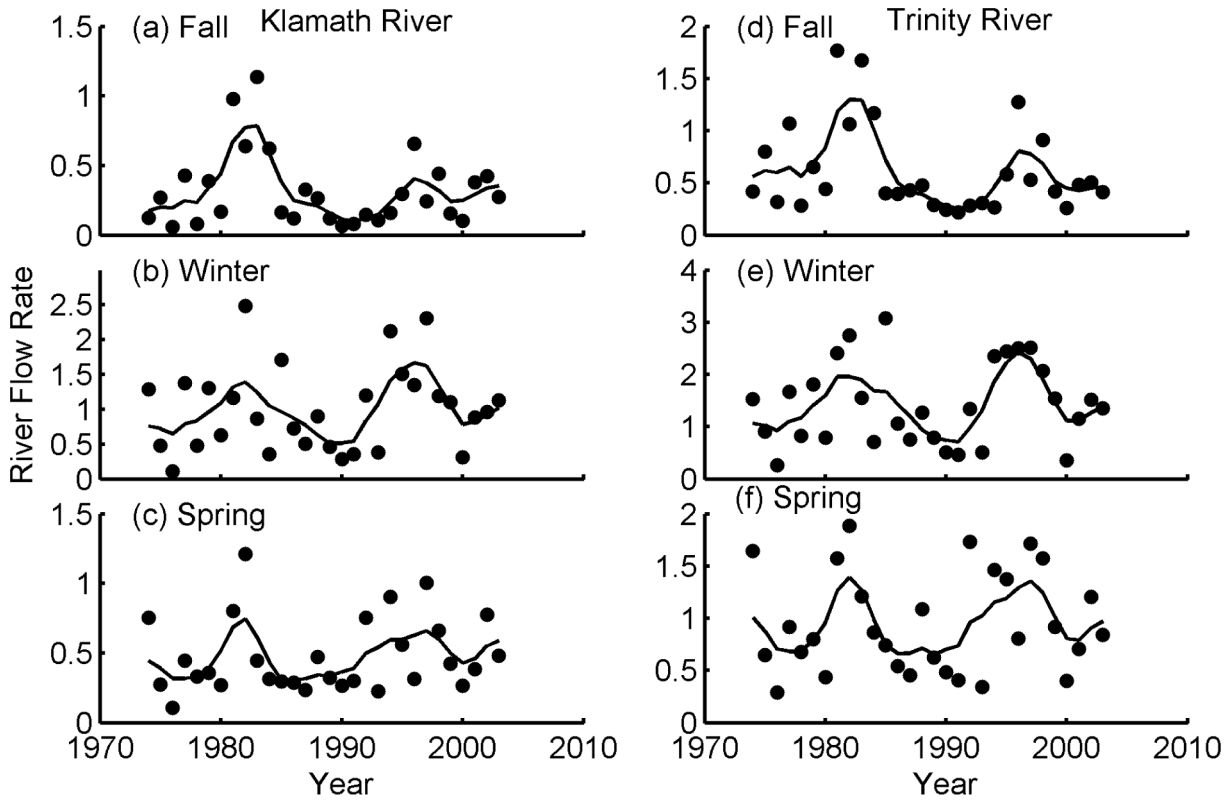


Figure 2.1. Average daily river flow rates at Klamath River (Orleans) and Trinity River (Hoopa), CA, during the fall (October–December), winter (January–March), and spring (April–June) periods.

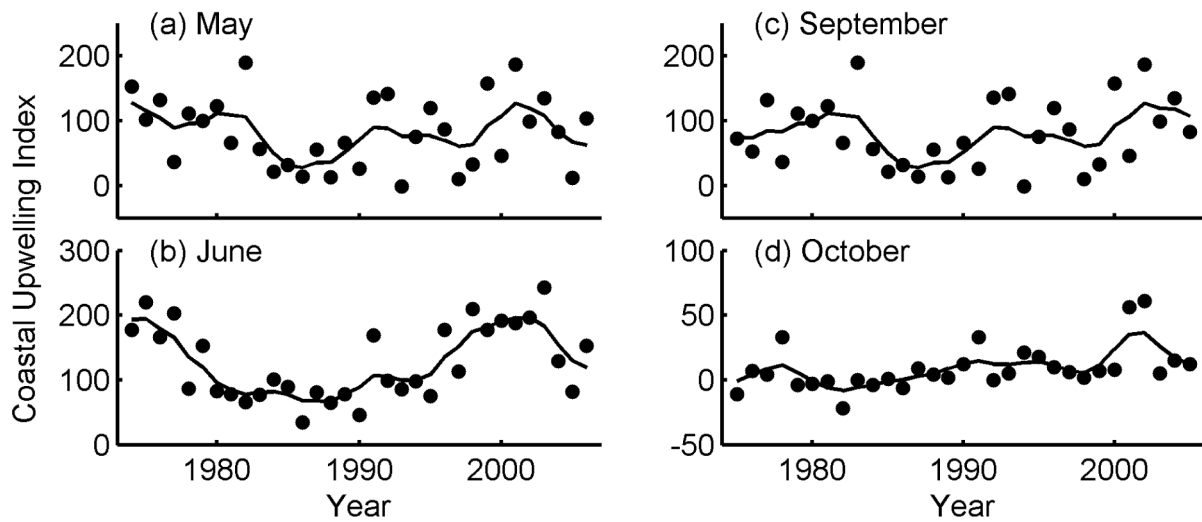


Figure 2.2. Monthly coastal upwelling indices at 42°N, 125°W.

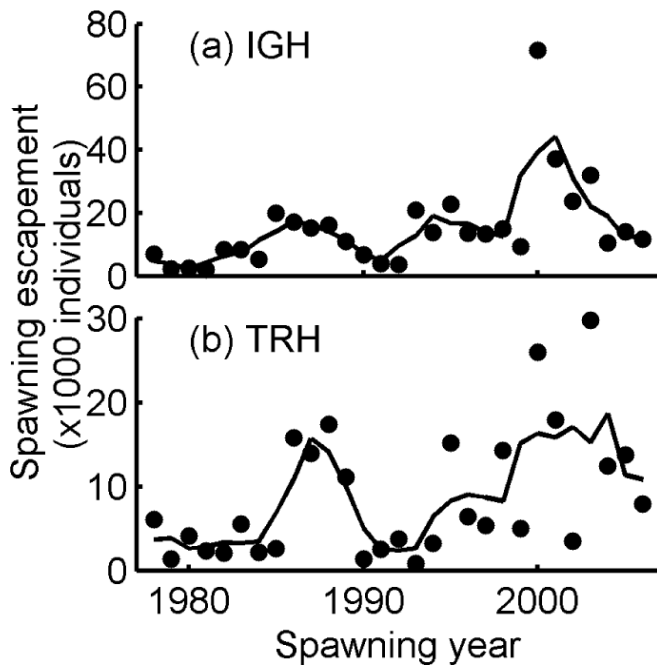


Figure 2.3. Escapement of fall-run Chinook salmon at (a) Iron Gate Hatchery (IGH) and (b) Trinity River Hatchery (TRH). In each panel, dots are the raw data, and a smooth curve indicates a twice-applied three point moving average.

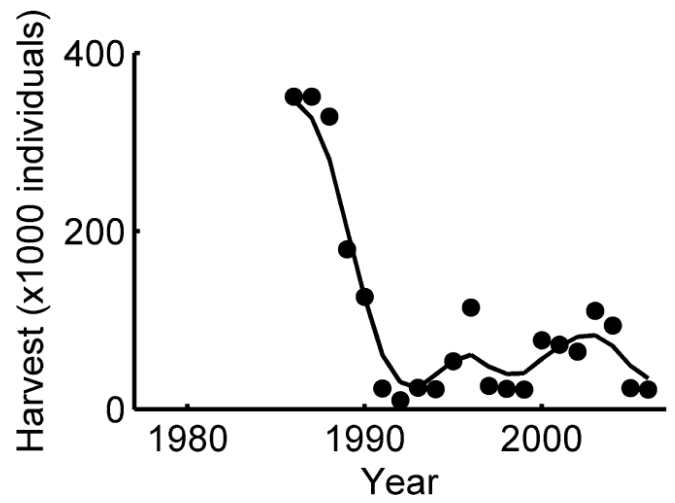


Figure 2.4. Estimated total fishery harvest of Klamath River fall-run Chinook salmon by ocean and river fisheries. Dots are the raw estimates, and the smooth curve indicates a twice-applied three point moving average.

Appendix 3

Correlation analysis of noisy time-series

Correlation analysis is a simple statistical method that is frequently used for associating population time series and environmental variables. Despite its wide use in ecology, a major challenge exists in interpreting its results because population time-series are often composed of signals from multiple sources. These signals interfere with each other, producing an increased chance of type II errors (i.e. chance of not detecting existing associations).

The purposes of this appendix are to demonstrate the signal interference phenomena using simulated data sets and to make recommendations for interpreting results from correlation analysis. For some readers, results in this appendix are obvious without the demonstrations; however, our recommendation is the opposite of what is generally recommended in the biological literature (see Results and discussion section of this appendix). This motivated the preparation of this appendix.

Methods

We first fit autoregressive moving average (ARMA) models to environmental variables. Using the fitted ARMA models, we simulated the environmental variables so that they have similar statistical properties as the original variables. Then, population time series that are affected by the simulated environmental variables were generated. Finally, applying correlation analysis to the simulated environmental and population time-series, we assessed the frequency of type I and type II errors.

Fitting ARMA to environmental variables

In order to simulate time series that have the same characteristics as actual environmental variables, we fit ARMA models of various orders (Diggle 1990) to environmental variables. The ARMA models are very flexible statistical models that can capture the statistical properties of time series. In general, the models are formulated as

$$Y_t = \varphi_1 Y_{t-1} + \dots + \varphi_h Y_{t-h} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \dots + \theta_k \varepsilon_{t-k} \quad (3.1)$$

where φ_i is an autoregressive coefficient, θ_i is a moving average coefficient, and ε_t is an independently identically distributed (iid) normal random variable with mean 0 and variance σ^2 . Hereafter, the model is denoted by ARMA(h,k), where h and k denote the orders of autoregressive and moving average processes, respectively.

The environmental variables used in this analysis were the ones considered in the analysis of Klamath fall Chinook (Appendix 2), except longer time series were used here (Table 3.1). First, each environmental variable was centred around zero by subtracting its mean, and then all possible ARMA (h,k) of up to order 2 (i.e. h=0, 1, or 2 and k=0, 1, or 2) were fit to the time series using the Kalman filter (Harvey 1989). Finally, the fits of the models to the data were compared using the information criteria AIC (Burham and Anderson 1998). This was repeated for all of the environmental variables.

Simulating environmental signals and population data

Using the best-fit ARMA model for each environmental variable, we simulated a time series consisting of 29 data points. We envision this time series being the simulated environmental signal that affects a population, and thus is the signal in a population time series that we hope to recover. We chose 29 data points because our Klamath fall Chinook data were collected over 29 years. The simulated time series is denoted by $Y^{(i)}$ where superscript i distinguishes independently generated time-series.

Based on the simulated environmental time series, population time series ($X^{(i)}$) were generated as

$$X^{(i)} = \tilde{Y}^{(i)} + \gamma^{1/2} \tilde{H} \quad (3.2)$$

where the tilde indicates that the time series was divided by its sampling standard deviation, and \tilde{H} is a standardized, serially iid normal random variable, which is generated independently for each time series i. Coefficient γ ($\gamma = 2, 3$ or 4) determines the amount of noise in the time series $X^{(i)}$, and we call this coefficient the 'noise level'.

We thus envision $X^{(i)}$ being the scaled population data and \tilde{H} is the 'noise' from multiple sources mixed together. For convenience, we call $\tilde{Y}^{(i)}$ and \tilde{H} 'signal' and 'noise', respectively, because $\tilde{Y}^{(i)}$ is the signal of interest. However, it is conceivable that, in real data, the "noise" may also be produced by other biologically important processes such as the effect of other environmental signals.

Comparing simulated environmental signals

Finally, we compare the simulated population and environmental time series by calculating the Pearson linear correlation coefficients. In one type of comparison, we calculated the correlation between $\tilde{Y}^{(i)}$ and $X^{(i)}$. These time series are actually associated with each other because the underlying signal in $X^{(i)}$ is $\tilde{Y}^{(i)}$. In another type of comparison, we calculated the correlation between $\tilde{Y}^{(i)}$ and $X^{(j)}$, where $i \neq j$. In the latter comparison, the two time series are not associated with each other because $\tilde{Y}^{(i)}$ and the signal in $X^{(j)}$ are independently generated although the two underlying signals have the same statistical distribution. These comparisons were repeated 10 000 times, each time simulating new time-series independently.

In the first comparison, the null hypothesis that there is no correlation is false, and in the second comparison, the null hypothesis that there is no correlation is true. From these comparisons, the frequencies of type II (falsely accepting the null hypothesis) and type I errors (falsely rejecting the null hypothesis), respectively, were estimated. These comparisons were repeated for all environmental variables under different noise levels γ ($\gamma = 2, 3$ or 4).

Results and discussion

Table 3.1 shows the best fit models. None of the best-fit models is purely random (i.e. not ARMA(0,0)), and some of them have high orders (i.e. h = 2 and k = 2). This suggests that these environmental variables are not simply white noise. This is an important fact that was one of the motivations for using MAFA in the Klamath fall Chinook data analysis (Appendix 1).

Figure 3.1a shows the distribution of correlation coefficients when the two time-series were associated while Fig. 3.1b shows the distribution when the two time-series were not associated. The underlying signal in this calculation was a simulated June coastal upwelling, and $\gamma = 3$. The two distributions overlap with each other suggesting that potential type I and/or type II errors are inevitable. However, by choosing the critical value carefully, we can optimize the errors.

The frequency of type I and type II errors are plotted over the magnitude of critical correlation coefficient (i.e. two-tail test) in Fig. 3.1c. If we wish to make the frequencies of the type I and type II errors to be the same, the critical correlation value should be 0.33. This is smaller than the value of 0.36 that would be chosen at the significance level of 0.05 when the two time-series were assumed to be serially iid, normal random variables. The smaller critical value is necessary to reduce the type II errors because ‘noise’ reduces correlation.

The same calculation was repeated when the underlying environmental signals were different (Table 3.1). The associated ARMA models exhibit different levels of lag-one autocorrelation. Despite the difference, all the figures looked very similar to the one for June coastal upwelling (Fig. 3.1). Therefore, we only show the results for the spring river flow rate of the Klamath River (Fig. 3.2) in addition to those for the June coastal upwelling. In all cases, the type I and type II error curves cross at a lower correlation value than 0.36. This result suggests that the effect of the differences in the distribution of the underlying signals is much smaller than the effect of interference from other signals.

Figure 3.3 and 3.4 repeat the calculation using the signal of June coastal upwelling with different noise levels γ ($\gamma = 2$ and 4, respectively). As the noise level increases, the type II error curve moves to the left while it affects the type I error curve only slightly. This suggests that, as the noise level increases, we should reduce the critical correlation value at the expense of increased type I errors.

Pyper and Peterman (1998) argue that we should increase the critical correlation value when we try to associate two serially autocorrelated time series. Their recommendation is correct when dealing with relatively noise free time-series (e.g. Fig. 3.3). However, population time-series are almost always ‘noisy’ because it is composed of multiple environmental, biological, and sampling signals. These signals will interfere with each other, substantially reducing the ability to detect the true associations. Therefore, when dealing with noisy time series (e.g. population time series), we should reduce the critical correlation value in order to balance the two types of errors.

Thus far, we have assumed that the optimal critical correlation value was where the type I and type II errors become equal. However, depending on the types of analysis, this may not be optimal. For example, if we are trying to screen many environmental variables at the initial stage of a study, one may be more concerned about potential type II errors than type I errors. If so, the critical value should be reduced further. On the other hand, at the later stage of a study, it may be preferable to increase the critical value as suggested by Pyper and Peterman (1998).

The results in this appendix also suggest the importance of a statistical method to separate underlying signals before performing a correlation analysis. Maximum autocorrelation factor analysis (MAFA) is one such method that enables us to separate signals based on a lag-one autocorrelation. As demonstrated in Appendix 1, MAFA is a powerful method to identify environmental signals in population data. Although the method improves the ability to identify underlying signals, we still have a high frequency of type II errors resulting from a large amount of specific factors (i.e. variability specific to each time-series). This suggests the critical correlation level should also be reduced when we try to correlate MAFs and environmental variables.

The optimal critical correlation level is strongly affected by the noise level (Fig. 3.1), which rarely will be known. Consequently, it is unclear as to how much the critical level should be reduced. Therefore, instead of selecting the critical level in the analysis, we recommend examining the relative values of the correlation coefficients. Environmental variables with a stronger correlation with population data are suggested as candidates for the next stage of a study. In order to reduce the occurrence of type I errors, environmental variables in the analysis should be preselected carefully based on prior knowledge of biological processes.

References

- Burnham, K. P. and Anderson, D. R. 1998. Model selection and inference: a practical information-theoretic approach. – Springer.
- Diggle, P. J. 1990. Time series: a biostatistical introduction. – Oxford Univ. Press.
- Harvey, A. C. 1989. Forecasting, structural time series and the Kalman filter. – Cambridge Univ. Press.
- Pyper, B. J. and Peterman, R. M. 1998. Comparison of methods to account for autocorrelation in correlation analyses of fish data. – Can. J. Fish. Aquat. Sci. 55: 2127–2140.

Table 3.1. Orders of best fit ARMA models to the twelve environmental variables that were used in the analysis of Klamath fall Chinook escapement data.

Variable	Length of original time series	Order of AR (h)	Order of MA (k)
Coastal upwelling in May (CUM)	60	1	0
Coastal upwelling in June (CUJ)	60	1	2
Coastal upwelling in September (CUS)	60	1	1
Coastal upwelling in October (CUO)	60	1	0
River flow rate of the Klamath in fall (RKF)	93	1	0
River flow rate of the Klamath in winter (RKW)	93	1	0
River flow rate of the Klamath in spring (RKS)	93	2	2
River flow rate of the Trinity in fall (RTF)	65	2	2
River flow rate of the Trinity in winter (RTW)	65	2	2

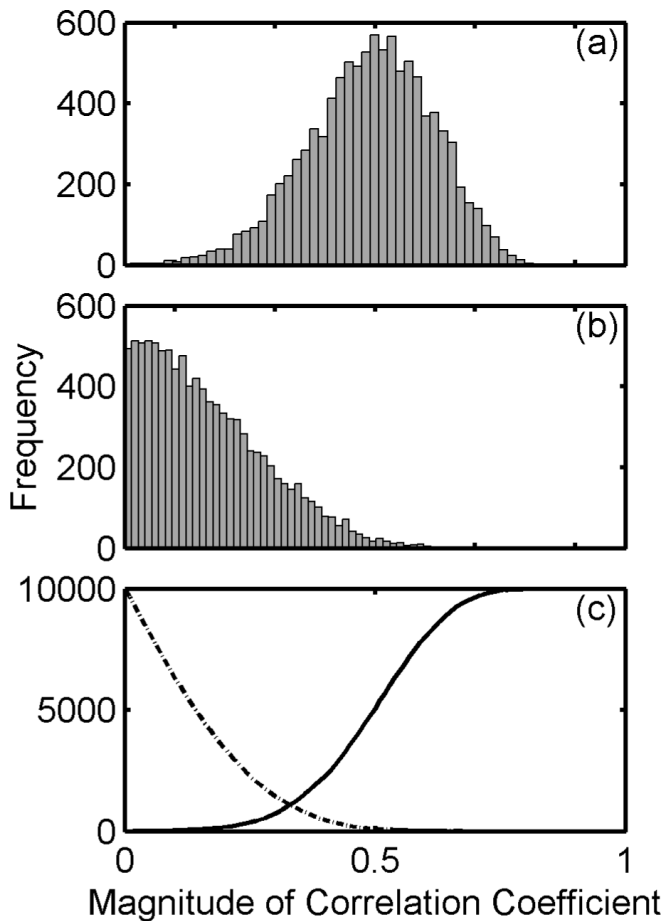


Figure 3.1. Correlation coefficients between two simulated time series when the original environmental variable was June coastal upwelling index, and $\gamma = 3$. (a) Distribution of the magnitude of correlation coefficient when the null hypothesis (no association) is false. (b) Distribution of the magnitude of correlation coefficients when the null hypothesis (no association) is true. (c) Frequency of type II (solid line) and type I errors (dashed line) over the critical correlation level (above the critical correlation level the null hypothesis is rejected).

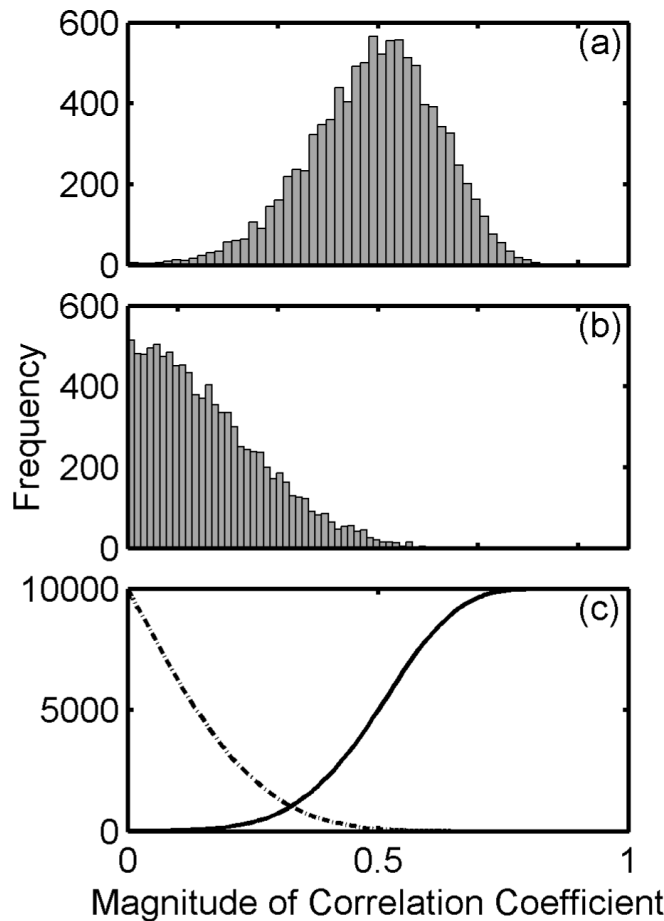


Figure 3.2. Correlation coefficients between two simulated time series when the original environmental variable was spring river flow rate of the Klamath River, and $\gamma = 3$. (a) Distribution of the magnitude of correlation coefficient when the null hypothesis (no association) is false. (b) Distribution of the magnitude of correlation coefficients when the null hypothesis (no association) is true. (c) Frequency of type II (solid line) and type I errors (dashed line) over the critical correlation level (above the critical correlation level the null hypothesis is rejected).

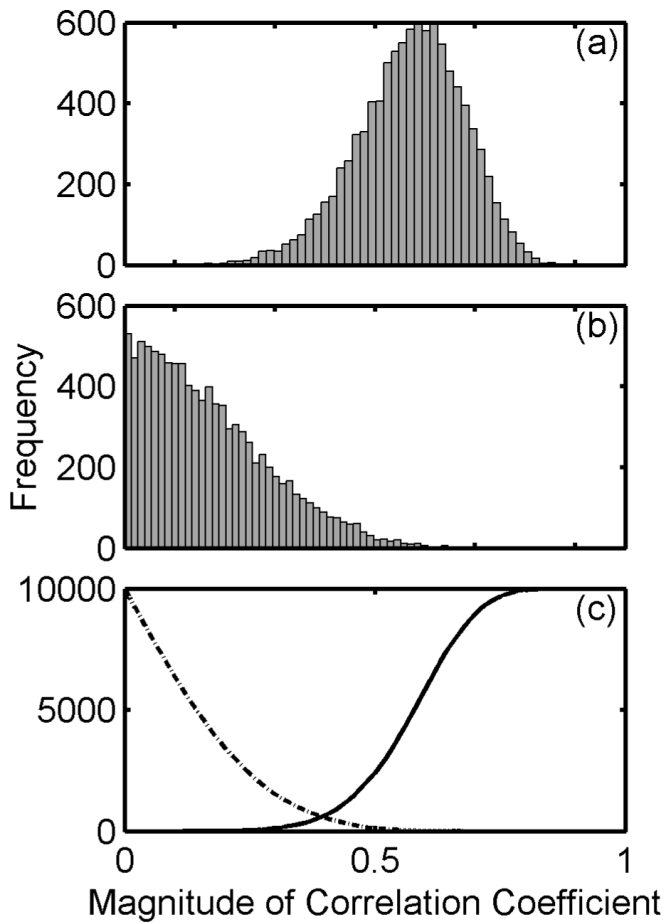


Figure 3.3. Correlation coefficients between two simulated time series when the original environmental variable was June coastal upwelling index, and $\gamma = 2$. (a) Distribution of the magnitude of correlation coefficient when the null hypothesis (no association) is false. (b) Distribution of the magnitude of correlation coefficients when the null hypothesis (no association) is true. (c) Frequency of type II (solid line) and type I errors (dashed line) over the critical correlation level (above the critical correlation level the null hypothesis is rejected).

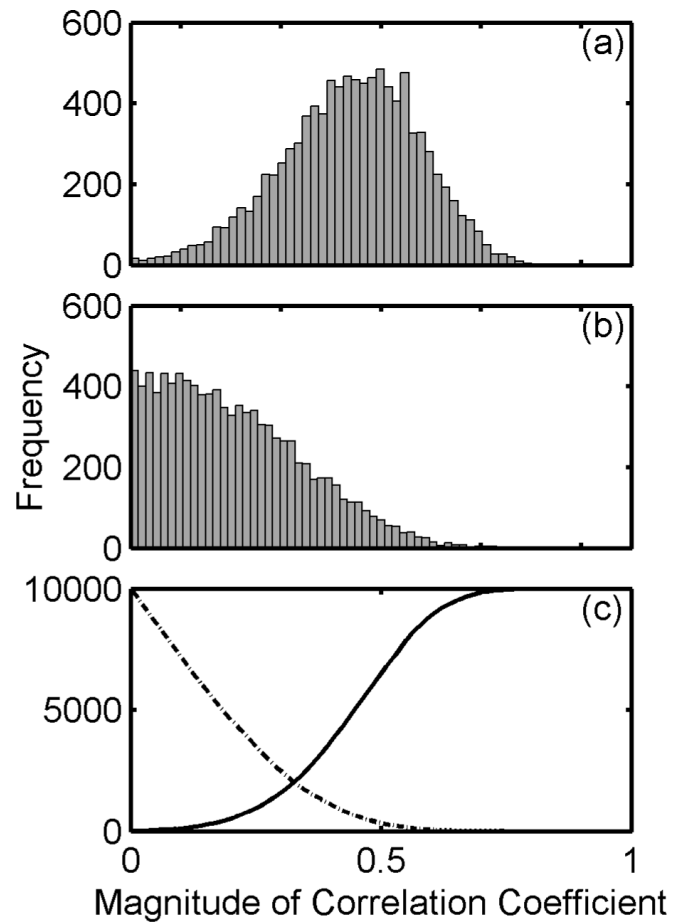


Figure 3.4. Correlation coefficients between two simulated time series when the original environmental variable was June coastal upwelling index, and $\gamma = 4$. (a) Distribution of the magnitude of correlation coefficient when the null hypothesis (no association) is false. (b) Distribution of the magnitude of correlation coefficients when the null hypothesis (no association) is true. (c) Frequency of type II (solid line) and type I errors (dashed line) over the critical correlation level (above the critical correlation level the null hypothesis is rejected).