

Chapter 14

Noise and Nonlinearity in an Ecological System

Paul A. Dixon¹
Maria J. Milicich
George Sugihara

ABSTRACT In this chapter, a case study from marine ecology is presented in which the application of techniques from nonlinear time series analysis is shown to provide insight into the interplay between stochastic physical forcing and nonlinear biological response in a natural system. Specifically, the replenishment of a population of reef fishes is analyzed in detail, and important nonlinearities are demonstrated in the processes underlying variability in the supply of larval propagules to the reef. This information is used to guide the construction of a series of models which attempt to forecast larval supply from readily measured physical variables. The most successful models are those that account for the nonlinearities in the response of larvae to their physical environment. Such models provide better forecasts than can be achieved with conventional linear techniques and identify processes hidden to linear analysis. The importance of understanding the interplay between noise and nonlinearity in ecological systems is discussed.

14.1 Introduction

A fundamental and largely unresolved problem in ecology is the relative contribution of exogenous and endogenous processes to temporal variability in the size of natural populations. Motivations for solving this problem are case-specific and range from purely scientific interest to the need for sensible management decisions, but the ultimate goal is generally the same: To make predictions about future states of natural systems from measurable initial conditions. To do this requires information about the identities of both the biological and physical variables that comprise the system and the functions that describe their interactions. Broadly, there are two approaches to obtaining such information. One can proceed from the bottom-up by first writing equations believed to capture the essence of the system's dynamics and then fitting the free parameters of the resulting

¹ Author for correspondence.

model from observational data. This necessitates sufficient prior knowledge to choose appropriate state variables and the forms of the functions which relate them. The alternative is the top-down approach, for which is required the ability to identify relevant variables and their interrelations directly from observations, is required to deduce the underlying processes at work. In either case the exercise is closed and judged a success when satisfactory predictions can be made and the ultimate limits to predictability are understood.

The ubiquitous observation is that natural populations do not remain constant over time, but rather change in what often seems a random fashion. The manner in which population biologists have thought about and attempted to explain this fact has also changed over time, in a pattern characteristic of ecology as a whole: first one set of processes, external forcing, was thought of primary importance (e.g., Nicholson 1933[28], Andrewartha and Birch 1954[1], Lack 1954[16])², followed by a shift in emphasis to a second set, internal dynamics, (e.g., May 1974[23], Schaffer 1985[33], Schaffer and Kot 1986[34]), and finally, when neither exclusive view proved satisfying, it became fashionable to acknowledge that both must play a role (e.g., Ellner and Turchin 1995[9]). However, tacit acknowledgement that a synthetic view is necessary is not the same as actually producing an effective synthesis. The aim here is to demonstrate that tools from nonlinear time series analysis, when applied judiciously, can help provide such a comprehensive understanding of the interplay between physical forcing and biological response, that is, between external and internal dynamics in ecological systems.

This chapter will describe in detail a case study (Dixon et al. 1999[7]) that is representative of a broad problem in fisheries biology: understanding and ultimately forecasting variability in fish stocks where population change is largely decoupled from population size. Fisheries assessors use various models to try to predict future replenishment from reproductive stock size (Rothschild 1986[31]), but such fitted curves have been plagued by extreme variance in recruitment, and explaining the scatter in these relationships has remained a largely unrealized goal throughout this century. A useful metaphor for those unfamiliar with these issues is to think of the life history of marine fishes as an input-output system (Fig. 14.1). Many marine species are characterized by a bipartite life history, in which larval forms possess behavioral and morphological characteristics associated with the challenges of a planktonic existence that may be completely

²In fact these authors were at the center of an earlier controversy in ecology, concerning agents of population control; that is, whether climatic (external) or biotic (internal) factors kept populations bounded in spite of their propensity for exponential increase. While the authors referenced here disagreed on this point, ecologists of the era did in general agree that, left alone, a population's internal dynamics would lead to equilibrium behavior, and accordingly that variability was indicative of external forcing.

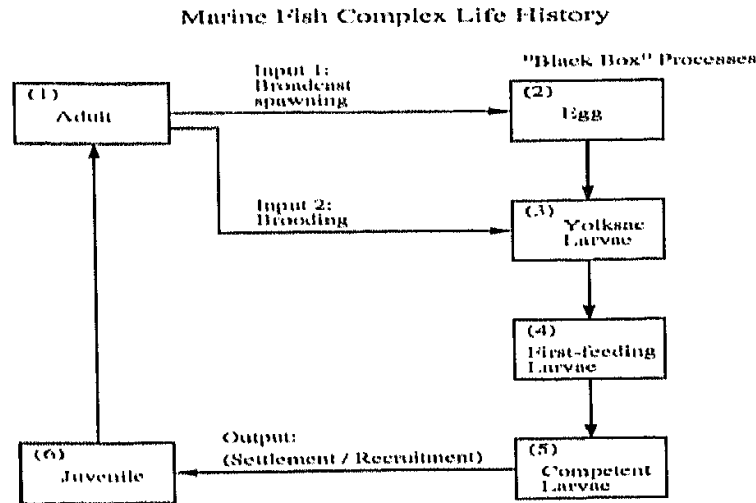


FIGURE 14.1. The bipartite life history common to many marine organisms, viewed as an input-output system. Density-dependence between population size and population change - that is, the relationship between the inputs and outputs of the system - may be interrupted by processes affecting mortality of eggs, larvae, and (in some cases) juveniles during their time in the plankton.

dissimilar from those of juveniles and adults. Investigating feedback control in a population is then equivalent to attempting to relate the system's output, recruitment³, to its input, the magnitude of adult spawning. Unfortunately, the larval phase in this schematic, which has the potential to decouple the inputs and outputs, is an essentially unobservable black box. Thus, while some researchers have attributed the unexplained variability to sampling error and the misuse of analytical techniques (e.g., Caputi 1988[2], Hilborn and Walters 1992[12]), others have suggested that unpredictable and high levels of mortality during the intervening larval phase will necessarily obscure the stock-recruitment relationship (e.g., Hjort 1914[13], Lasker 1981,[17], Cushing 1982[5]). Because an understanding of the fate of larval stages has remained elusive, it has proven difficult to evaluate the relative contributions of these sources of error.

The next two sections will adopt a dynamical systems approach to retrospectively investigate processes at work during the larval phase. The system that will be considered is Lizard Island, Australia, part of the Northern Great Barrier Reef, and the focus will be on a common and abundant family of reef fishes, the pomacentrids or damselfishes. Section 14.2 will begin by

³The term "recruitment" is often ambiguous and may be used in the literature to refer to any one of several points in a species life history (Richards and Lindeman 1987[30]). This chapter will focus on a population of reef fishes, and will therefore equate recruitment to settlement - the metamorphosis from a planktonic larval to a sedentary juvenile form. Many pelagic fish stocks possess a planktonic juvenile phase as well, which may be equally important in decoupling the relationship between population size and population change (Peterman et al. 1988[29]).

analyzing time series of spawning output, supply of mature larvae to the reef and subsequent juvenile abundance for one species of pomacentrids (boxes 2, 4, and 5 in Fig. 14.1) and will provide evidence for nonlinearity in the processes underlying variability in larval supply. This information will then be used in Section 14.3 to guide a search for the identities of the most important physical variables that drive this variability. After these variables have been identified, the functions which relate them to larval supply will be deduced. The result is a nonparametric, nonlinear model that affords better forecasts than linear models, and identifies the importance of events early in larval life which are hidden to linear analysis. This chapter will conclude by placing these results in the general context of understanding the interplay between stochastic physical forcing and nonlinear biological response in ecological systems.

14.2 Univariate Analysis: Nonlinearity in the Larval Phase

Pomacentrus amboinensis, the Ambon damsel, is a common shallow-water damselfish found in the eastern Indian and western Pacific Ocean, including the Great Barrier Reef. A unique set of time series are available for this species which describe contemporaneous temporal variability at three stages in life history: spawning output, larval supply, and the abundance of newly settled juveniles (Milicich et al. 1992[26], Meekan et al. 1993[24]). Briefly, spawning was estimated from daily visual census of eggs in nests; observations were made for 168 days over two consecutive spawning seasons. Juvenile abundance was monitored for 166 days over two seasons by scuba divers swimming visual transects over patch reefs. Larval supply was measured using two replicate light traps, which fished for three hours a night for 280 nights over three consecutive seasons. Raw data are given in Figure 14.2.

Three forecasting tests will now be applied to these data. The first two result in distinct pieces of evidence for nonlinearity in the processes affecting variability at the end of the larval phase: an enhanced success of nonlinear forecasting algorithms in predicting the data and the presence of structure in the residuals from linear forecasts of larval supply. Results of the third test, which examines the rate at which forecastability declines as a function of prediction time, give evidence for the stochastic nature of the processes underlying variability in larval supply.

14.2.1 *S-Map Analysis*

The first test for nonlinearity is an examination of the relative success of linear and nonlinear forecasting algorithms in predicting the data. For this,

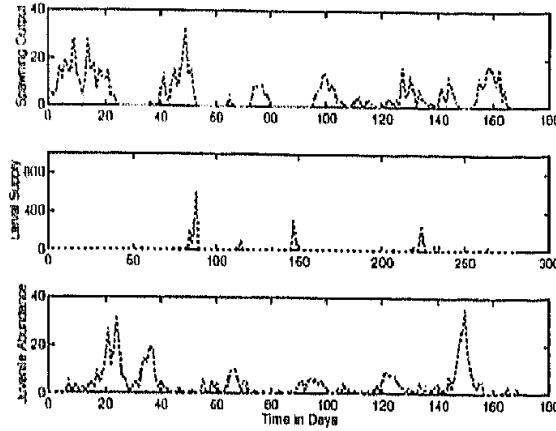


FIGURE 14.2. The replenishment of *P. amboinensis*. Top panel gives spawning output measured by the number of eggs released from the nest, middle panel gives the supply of mature larvae to the reef, and bottom panel the number of juveniles on the reef. All measurements were made daily.

S-maps were used to make predictions (Sugihara 1994[36]). The data were embedded using lagged coordinates ⁴. For the embedded time series $\mathbf{X}_t \in \mathbb{R}^{m+1}$, where the constant term in Eq. 14.2 below is given by $\mathbf{X}_t(0) \equiv \mathbf{1}$, and the time series value T_p steps forward is $\mathbf{X}_{t+T_p}(1) = Y_t$, forecasts at T_p were given by:

$$\hat{Y}_t = \sum_{j=0}^m C_t(j) X_t(j). \tag{14.1}$$

For each predictee, \mathbf{X}_t , singular value decomposition was used to solve for C using the rest of the data set as follows:

$$B = A C \tag{14.2}$$

where

$$B_i = w(\|\mathbf{X}_i - \mathbf{X}_t\|) Y_i \tag{14.3}$$

$$A_{ij} = w(\|\mathbf{X}_i - \mathbf{X}_t\|) X_i(j) \tag{14.4}$$

and

$$w(d) = \exp^{-\theta d/\bar{d}}. \tag{14.5}$$

The degree of nonlinearity in these maps is thus controlled by the tuning parameter θ , where $\theta = 0$ gives a global linear map, and theta increasingly positive yields increasingly local, nonlinear forecasts. Forecasts were made out-of-sample through leave-one-out cross-validation (that is, all vectors whose coordinates included any of the coordinates of the predictee were

⁴For the results given here, the time lag was held constant at one day; results were robust to the specific choice.

eliminated from model fitting) and were made one time step ahead, for the next day's value. Figure 14.3 plots forecastability as a function of model nonlinearity and embedding dimension (number of reconstructed variables used). The abscissa gives model nonlinearity, and the ordinate forecast success measured both in terms of correlation coefficient between predicted and observed values and average arithmetic error of the predictions.

Two features of this plot are readily apparent. First, regardless of which metric of model performance is used, nonlinear models with an embedding dimension of two or greater significantly (Z-test; P less than .05) outperform linear maps for the larval supply time series, whereas for the spawning and juvenile data this is not the case. The underlying philosophy here is that the characteristics of the model which gives the best out-of-sample forecasts most closely reflect the nature of the underlying dynamical processes; accordingly, these results may be taken as preliminary evidence that there are important nonlinearities underlying the variability in larval supply of this species at this particular location. Additionally, the overall level of predictability is different for the three time series; it is generally much harder to predict larval supply than spawning output or juvenile abundance.

14.2.2 Residual Delay Maps

A second piece of evidence for nonlinearity in larval supply results from examining the residuals from linear model forecasts, through the method of residual delay maps (RDMs) (Sugihara et al. 1999[37]). The basic idea is that if the variability in a time series arises strictly from linear, stochastic processes, then the residuals from linear model forecasts should be randomly distributed around zero. Alternatively, if there is nonlinear structure of sufficiently low dimension to be detected with the available data, the functional form of the nonlinearity — the fashion in which the residuals depart from a random distribution centered around zero — will likely become apparent. To examine this, residual delay maps plot the average residual from a forecast at time $t+1$ as a function of the average value of the raw data at time t (hence the term “delay map,” as tomorrow's error is plotted as a function of today's value). This technique for detecting nonlinearity in observed data has the advantage of computational simplicity and relatively modest data requirements. To reduce the variance and clarify any systematic structure in the residuals that may exist, the raw values of the time series are first sorted and binned. For each bin, the average residual from a linear forecast for the points in that bin is calculated and plotted against the mean value of the points in the bin. Residual delay maps thus employ a two-way averaging procedure. Figure 14.4 gives the results of this test. To facilitate comparison, raw data were first divided by their standard deviations; this normalizes the magnitudes of the residuals across the three data sets, but does not affect the relative position of the points in the figure. The residuals from linear (AR3; the choice of the number of lags was

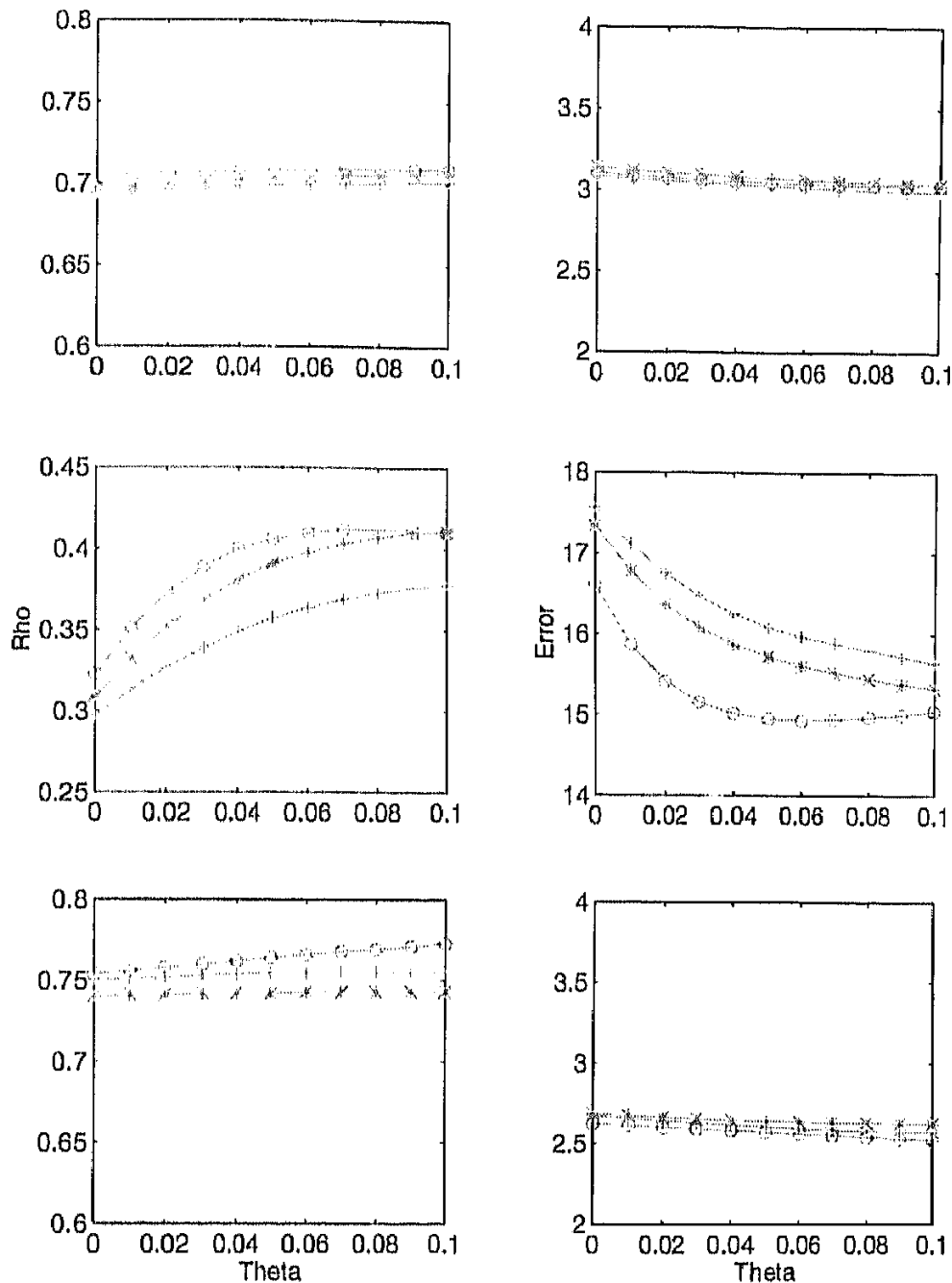


FIGURE 14.3. S-map analysis of spawning output (top), larval supply (middle), and juvenile abundance (bottom) for *P. amboinensis* at Lizard Island. The left panels give prediction success as a function of model nonlinearity as measured by the correlation coefficient between predicted and observed values, the right panels success as measured by the average error of the predictions. Circles: embedding dimension = 2; stars: $E = 3$; pluses: $E = 4$. Model nonlinearity is controlled by the tuning parameter θ , where $\theta = 0$ gives a global linear model, and θ increasingly positive gives increasingly nonlinear algorithms. Only for the larval supply data is a significant increase in predictability observed with nonlinear models.

not critical) forecasts of adult spawning output fluctuate evenly around zero, whereas the residuals for larval supply seem to display a systematic trend, with high values of larval supply today leading systematically to an overestimation of larval supply tomorrow. When juvenile abundance is considered, these patterns have broken down; there is one outlier residual, but no evidence for a systematic trend.

14.2.3 *Prediction Decay*

The final forecasting characteristic of interest is the rate at which predictability decays with increasing forecast horizon. Figure 14.5 (top panel) gives the results of this test for the *P. amboinensis* data. Again, forecasts are given by S-maps. For the spawning output and juvenile abundance time series, predictability decays steadily, in linear fashion, to essentially zero over five time steps (days) into the future. By comparison, larval supply data become unpredictable three days into the future. Beyond this basic difference, the test is ambiguous. It is not possible to label the rate of prediction decay as exponential (as opposed to linear) for the larval supply data, as there are only three points and the location of the second one (time horizon = two days) is sensitive to exactly how nonlinear a model one chooses. Indeed, this difference in predictability as a function of prediction time seems to reflect little more than the autocorrelation spectrum of the three data sets (Fig. 14.5, bottom panel). Because such strong short-term autocorrelation can potentially confound forecasting results, the analysis will now be extended to include formal hypothesis testing through the use of surrogate data.

14.2.4 *Surrogate Analysis*

There are a number of possible explanations as to why data may appear nonlinear to forecasting algorithms and related analytical techniques, some of which are trivial (e.g., the non-normality of the data itself may be responsible), and some of which are of real biological interest. The approach to distinguishing among these that will be adopted here is to proceed through the use of surrogate data (Theiler et al. 1992[39], Kantz and Schreiber 1997[14]). The idea is to create a large number of surrogate realizations of the data, each of which corresponds to a specific null hypothesis pertaining to the origin of the nonlinearity, and then to reanalyze the surrogates with the same forecasting techniques as were applied to the real data. By choosing a statistic which encapsulates the results of the analysis, a P-value can then be assigned to the null hypothesis from the percentage of the surrogates that yield a value of that statistic greater than or equal to that calculated from the original time series. Surrogate data techniques thus enable a transition from heuristic, descriptive evidence of nonlinearity to rigorous hypothesis testing.

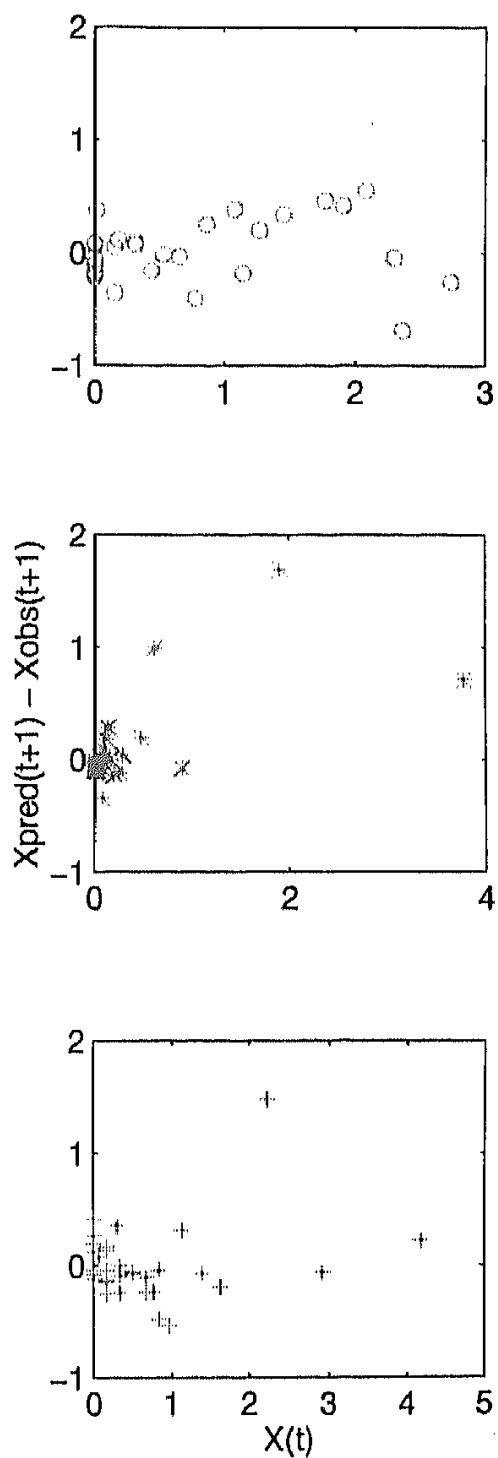


FIGURE 14.4. Residual delay map analysis, for bin size = 10. Top panel: average error from a linear model of tomorrow's spawning output as a function of today's value. Middle: larval supply. Bottom: juvenile abundance. Only for the larval supply data is there a systematic departure from residuals evenly distributed around zero; following days of high larval supply, linear models consistently overestimate the true value.

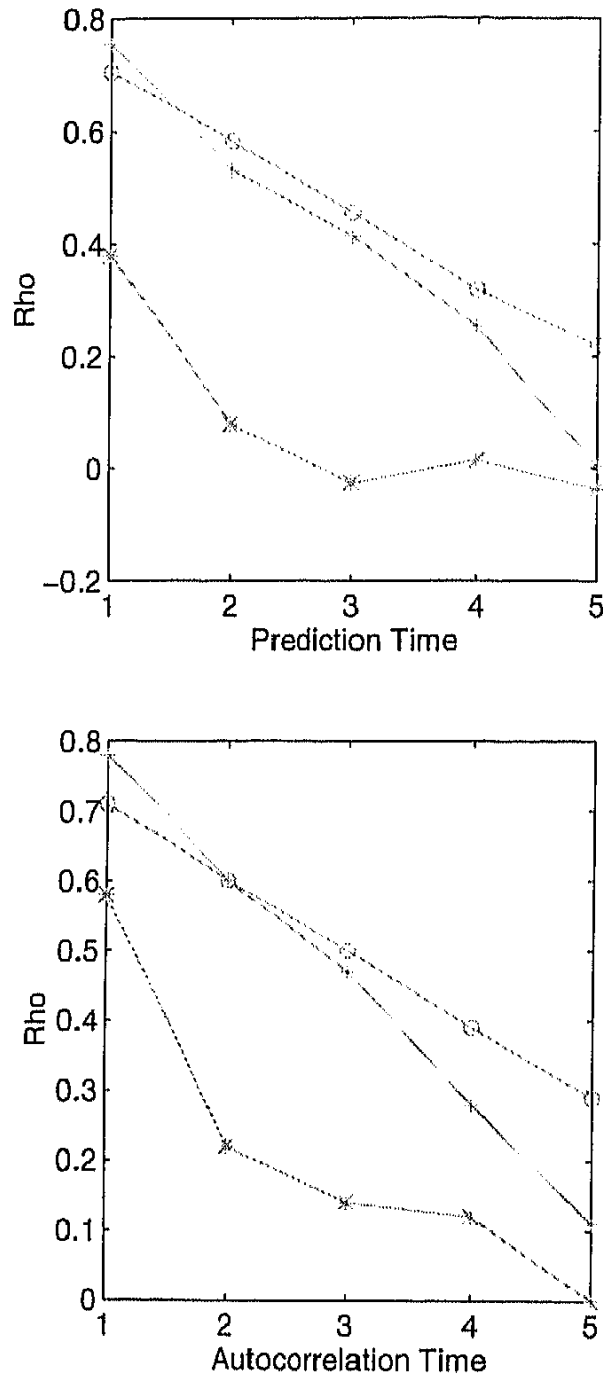


FIGURE 14.5. Prediction decay (top panel) for *P. amboinensis* data. Larval supply (stars) is unpredictable past two days into the future, whereas prediction decay is slower for spawning output (circles) and juvenile abundance (pluses). This prediction characteristic reflects the autocorrelation spectra (bottom panel) of the data.

Figure 14.6 gives a sample surrogate time series for supply of larval pomacentrids to Lizard Island for each null hypothesis tested. In order of increasing complexity, these are:

- Ho1** The data arise from an identical, independently distributed (I.I.D.) noise process.
- Ho2** The data arise from a set of linear, stochastic processes.
- Ho3** The data arise from a static, monotonic, nonlinear filter of a set of underlying linear, stochastic processes.

Creating surrogates is the most straightforward for the first null hypothesis: the time series is simply randomly shuffled. This is of interest because it tests whether the frequency distribution of the data itself is responsible for the observed nonlinearity. As has been seen, larval supply time series are episodic; daily values are not normally distributed. Indeed, this is the case for most biological properties in the ocean (Cassie 1963[3]). Because larval reef fishes have been demonstrated to be spatially coherent on scales of tens of kilometers (Doherty and Williams 1988[8]), one possible explanation for temporal variability in larval supply is that peaks in larval supply simply represent the chance advection of patches of larvae past the reef (Williams and English 1992[41]). An obvious question is thus whether the frequency distribution of the data alone is all that the forecasting tests have really measured. However, as will be seen in the analysis of the surrogates corresponding to Ho1 and Ho3 (described later), this is an insufficient explanation for the patterns observed. Variability alone does not necessarily imply nonlinearity.

Ho2 offers a direct test of the idea that variability in larval supply does not arise solely from the action of linear, stochastic processes. Here, surrogate data generation preserves the power spectrum of the data while randomizing the phases. Accordingly, the surrogates do not possess the same frequency distribution as the original time series.

Ho3 combines the features of the first two by preserving both the Fourier spectra and the frequency distribution of the real data. Analyzing these surrogates tests the idea that the nonlinearity is consistent with a monotonic, nonlinear transformation of an underlying set of linear, stochastic processes. This hypothesis is labeled “null” in deference to the dynamical systems community from whence these methods come, and for whom such an explanation for nonlinearity in a time series is not usually of great interest. However, to the ecologist, this explanation for the nonlinearity observed is a suggestive one, for a “nonlinear transformation of an underlying linear, stochastic process” is, as will be seen, a good starting point for a description of how biological entities respond to variability in their physical environment.

Here, R , the statistic that encapsulates the forecasting analysis, will be given by the difference in forecasting success between the linear and best

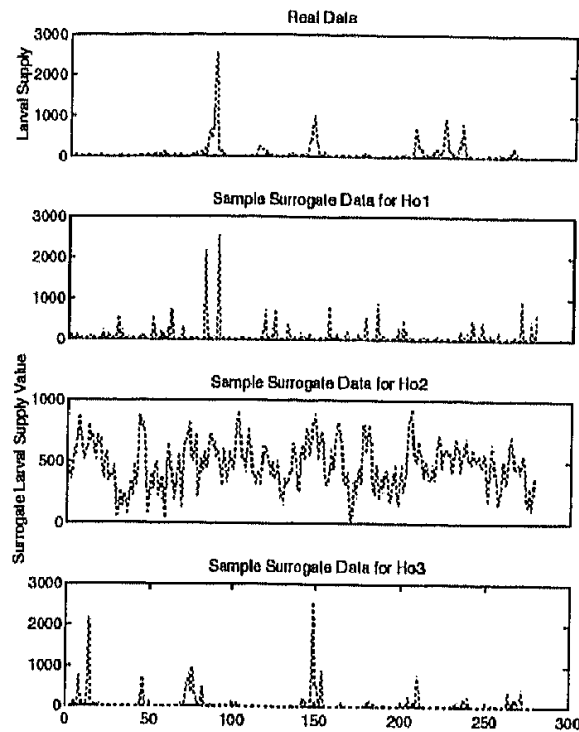


FIGURE 14.6. Raw data for supply of all pomacentrids to Lizard Island (top panel), and sample surrogate realizations corresponding to Ho1, Ho2, and Ho3 (top to bottom).

nonlinear algorithm, measured by the square of the correlation coefficient between predicted and observed values (percent variance explained). In principle, the choice of statistics is not critical; this particular R is intended to place the emphasis on successfully forecasting the times of greatest biological interest, the peaks in larval supply. The analysis to this point has centered on one species of pomacentrid; here, however, to decrease the number of zero values, larval supply of all pomacentrids will be considered. There is a strong correlation (ρ greater than 0.9) between supply of *P. amboinensis* and supply of all pomacentrids, so it seems unlikely that aggregating to a coarser degree of taxonomic resolution will create much difficulty.

Figure 14.7 gives graphical portraits of the distribution of R values for Ho1, Ho2, and Ho3, as a function of the embedding dimension. All three may be rejected for E greater than one. The non-normality of the data alone is not responsible for the improved forecasts seen with nonlinear models, and the data are not the result of a set of linear, stochastic processes. That Ho3 can also be rejected is intriguing, and, as will be seen in the next section, is probably due to the fact that, while variability in larval supply may originate from biological responses acting as a nonlinear transformation of an underlying set of linear, stochastic (physical) processes, there is no a priori reason why the transformation should be a monotonic one.

To this point, two suggestive pieces of evidence for nonlinearity of larval

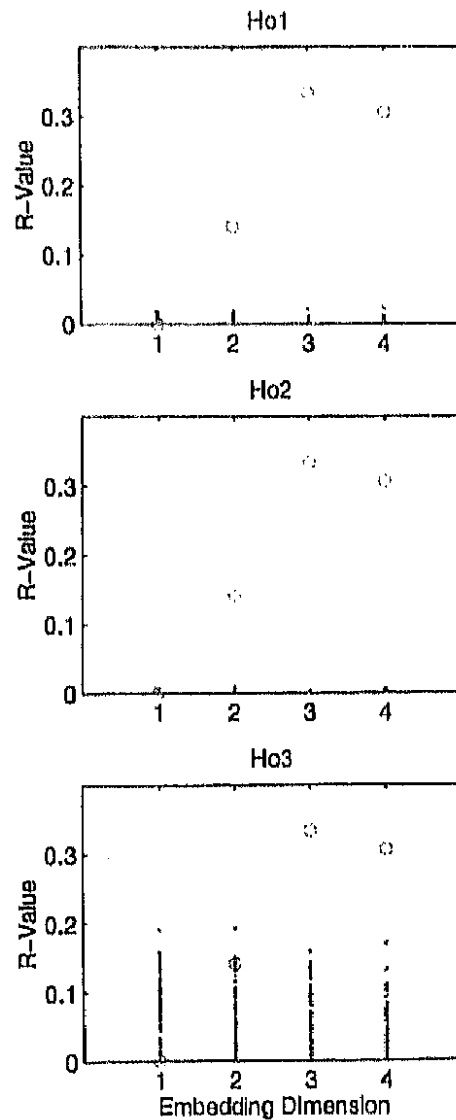


FIGURE 14.7. Distribution of R-values corresponding to Ho1, Ho2, and Ho3 for supply of all larval pomacentrids to Lizard Island. All three hypotheses may be rejected with greater than 95% confidence for $E_i = 2$.

supply time series have been presented: enhanced out-of-sample forecasting success of nonlinear algorithms, and structure in the residuals from linear model forecasts. Additionally, the stochastic nature of larval supply has been demonstrated, as it is impossible to make predictions more than two days into the future. The picture that is beginning to emerge is that there are important nonlinearities in the processes at work in the larval phase, which suggests that a strong relationship between spawning patterns and patterns of juvenile abundance can no longer be expected. The next section will put this information to use in identifying the physical processes responsible for this variability in larval supply.

14.3 Multivariate Analysis: Forecasting Larval Supply

The discussion will now leave the confines of black-box modeling from reconstructed dynamical variables and begin forecasting larval supply directly from observations of the physical environment. The analysis to this point has suggested that to succeed in this task, forecasting models will have to take into account the nonlinearities in the interaction between larval fishes and their fluctuating physical environment. There are two potential aspects to this nonlinearity: the forms of the functions relating the biology to the physics (which need not be linear), and a nonlinear coupling between key physical variables in their total impact on larval supply. Whenever this second aspect is important, as will be the case here, it will not be possible to deduce the action of different physical variables of larval supply by considering them in isolation. Accordingly, many conventional techniques for analyzing ecological data, such as multivariate linear regression or principal component analysis, will be inapplicable to this problem. What is needed is a nonlinear approach.

14.3.1 Lunar Phase

The most obvious property of pomacentrid larval supply is the cyclic nature of the timing of supply peaks. This results from the entrainment of egg release from the nest to the lunar cycle, in conjunction with a relatively invariant larval duration, which otolith analysis indicates averages nineteen days (Wellington and Victor 1989[40], Meekan et al. 1993[24]). This deterministic feature of the data thus represents a logical place to begin constructing a model. The specific physical variable chosen here to represent the lunar phase is the percent illumination, calculated by multiplying the proportion of a full moon for each night by the elongation, or phase angle, for that night (Kopal 1971[15]). Statistics for all of the models that will be constructed in this section are summarized in Table 14.1.

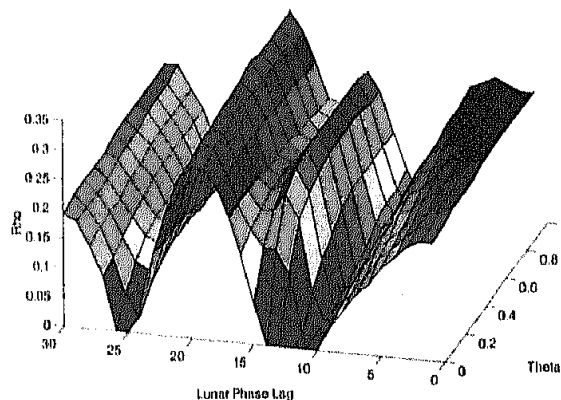


FIGURE 14.8. Forecasting larval supply from lunar phase. Forecast success (Z-axis) is measured by the correlation coefficient between predicted and observed values, set to zero for negative values; it is shown as a function of lag and model nonlinearity. There is a peak in predictability at 19-20 days, in good agreement with the average age of the larvae, and there is little to be gained with nonlinear models.

Because the relationship between pomacentrid larval supply and lunar phase is well-established at Lizard Island (Milicich 1994[25]), this variable provides an opportunity to test a technique that will be used again shortly for another variable where the situation is not so clear. One standard (and simple) method for detecting cycles in ecological data is to calculate the autocorrelation at increasing time lags (Finerty 1980[10]); if a cycle is present, the autocorrelation will be a maximum when the time lag corresponds to the cycle's wavelength and a minimum at half the wavelength. A similar approach will be adopted here: Larval supply will be predicted from percent illumination at increasing time lags, from one to thirty days in the past. The twist is that, in addition to exploring predictability as a function of lag time, it will be evaluated with respect to model nonlinearity. Adding this dimension to the analysis results in the surface plot in Figure 14.8. As before, forecasts were given by S-maps, and success was measured by the correlation coefficient between predicted and observed values. Forecasts again were made out of sample by removing each point for which a forecast was being made from the model fitting procedure. As these larvae average nineteen days in age, the expectation is that there should be a peak in predictability at a corresponding nineteen-day lag in lunar phase. Also, because the analysis of spawning output in the previous section indicated that egg release from the nest is well-described by a linear model, it seems likely that for this variable, linear models should perform about as well as nonlinear ones. Both of these expectations are met; there is indeed a peak in predictability at the appropriate time lag, and predictability remains essentially constant as the model is made more nonlinear.

Because of the deterministic nature of the lunar cycle, there is a second peak in predictability, centered at approximately five days and it appears that predictability begins to improve again as the lag approaches thirty days. If peaks in larval abundance are associated with a full moon nineteen days prior, it stands to reason that they will also be associated with a new moon five days and thirty-three days prior. To test that this is actually what these results are capturing, it is therefore necessary to extract the function relating larval supply to lunar phase at each of these two lags. In this case, where only one variable is being considered by each model, such an extraction is straightforward. However, the technique which will be used for this will be introduced in the most general terms, to allow an easy extension to the higher dimensional cases that will follow.

To deduce the functional relationship between larval supply and the physical parameters of interest, a space is first constructed whose axes are given by the physical variables of interest, normalized by their standard deviations (to normalize the ranges of each variable). Each point in the larval supply time series is then mapped to its corresponding physical coordinates in this state space. A neighborhood average (of N neighbors) is then constructed for each point in the state space; that is, for every point, that point's N nearest neighbors in state space are identified, and the average value of each physical variable in the neighborhood calculated. The average value of larval supply for the points in the neighborhood is also determined, and, for each physical variable, a plot is made of the neighborhood averages of that variable against the neighborhood averages of larval supply (after multiplying by the appropriate standard deviations). The advantage of this technique is that, when more than one physical variable is considered, it becomes possible to smooth the data while preserving the non-additive sense of the interactions between the physical variables.

Results for the five- and nineteen-day lags of lunar phase are given in Figure 14.9. As expected, maximum larval abundance is associated with a new moon (zero percent illumination) for the five-day lag, and the full moon (100 percent illumination) for the nineteen-day lag. The next question is the degree to which modeling larval supply from lunar phase captures the dynamics of the real data. The nineteen-day-lag model gives a correlation coefficient of 0.3 (Table 14.1). Figure 14.10 gives another sense of model performance by plotting the actual predictions alongside the real values. In all four panels, the real data are plotted on the ordinate in the positive direction, and the forecasts (multiplied by minus one) in the negative. The upper-left panel gives the predictions from the linear model built on the five-day lag; the upper-right panel gives the same variable with a nonlinear model. The bottom panels are from corresponding models (linear and nonlinear) constructed from the nineteen-day lag. In all four cases, the deterministic component of larval supply gives a sense of the timing of the peaks, but as expected, there is no information in the phase of the moon about the fate of the larvae after they are born. It is thus not possible to

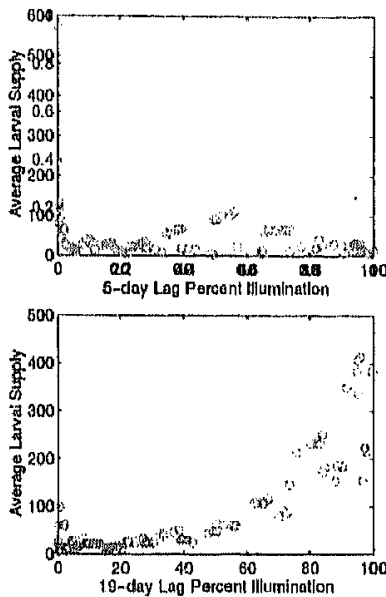


FIGURE 14.9. Response of larvae to lunar phase (measured by percent illumination). For the 19-day lag, larvae are positively associated with the full moon (percent illumination = 100); for the 5-day lag, the new moon (percent illumination = 0).

distinguish strong peaks from weak ones or from no peak at all, with just this variable.

14.3.2 *Cross-Shelf Wind*

For this, additional variables corresponding to other events in larval life need to be incorporated into the analysis. The next process that will be considered was motivated by a prior analysis of this system (Milicich 1994[25]), which suggested that the transport environment experienced by mature larvae returning to the reef might be of importance to the magnitude of larval supply.⁵ As no current meter data are available from this system, a three-day running average of recent cross-shelf wind speed (the onshore/offshore component of the wind) was selected as the physical variable most representative of surface current speed and direction. Figure 14.11 gives the results of adding this variable and again constructing a series of forecasting algorithms, ranging from global to local fits of the data. The top panel demonstrates that, considered linearly, there is no evidence that this variable is important, but when the models are made nonlinear, predictability begins to increase. This is as expected; these two effects occur at different times in larval life (the beginning and the end), and thus are coupled multiplicatively in their total impact. The middle panels give the predictions

⁵Indeed, exploring the transport of particles to and from reefs is a usual starting point in attempting to understand larval supply for vertebrates and invertebrates alike.

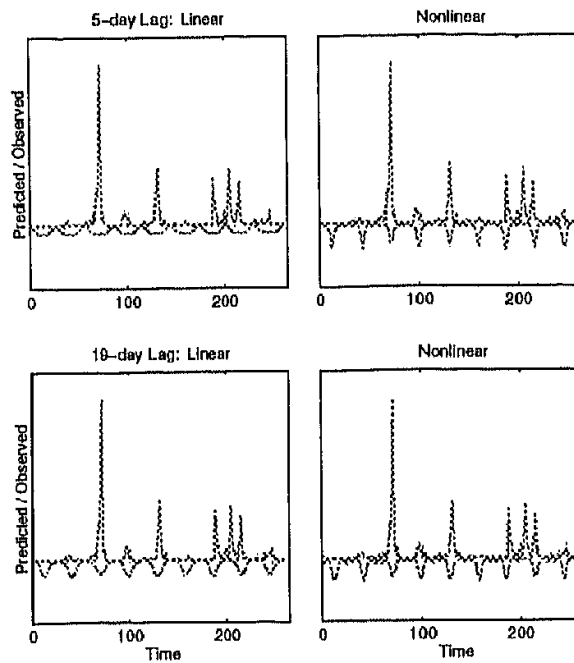


FIGURE 14.10. Predicting larval supply from the lunar phase. Upper panels: 5-day lag; lower panels: 19-day lag. Left panels: linear models; right panels: nonlinear models. Both linear and nonlinear models give a sense of the timing of larval supply peaks, but not the magnitude.

from linear and nonlinear models based on these two variables. Only in the nonlinear case does it begin to become possible to distinguish strong peaks from weaker ones. The bottom panels give the results of extracting, in the fashion detailed earlier, the functions relating larval supply to these two variables. As before, larval supply peaks are associated with a full moon nineteen days prior; they may now also be seen to be associated with weak, onshore winds (positive values of cross-shelf wind speed are onshore; negative values offshore). This is contrary to what would be expected if mature larvae act as passive particles (in which case stronger onshore winds, and associated greater onshore transport of surface waters, would be expected to increase larval supply), but is consistent with recent findings concerning the abilities of mature larval fishes to swim, aggregate, and generally promote their own orientation (e.g., Stobutzki and Bellwood 1994[35], Leis et al. 1996[21], Leis and Carson-Ewart 1997[20]).

14.3.3 Average Daily Wind

The final process that will be considered is the potential importance of wind stress on the first-feeding success of young larval fishes. The suggestion that such critical periods are important to subsequent recruitment has a long history in the fisheries literature (e.g., Hjort 1914[13]), and the

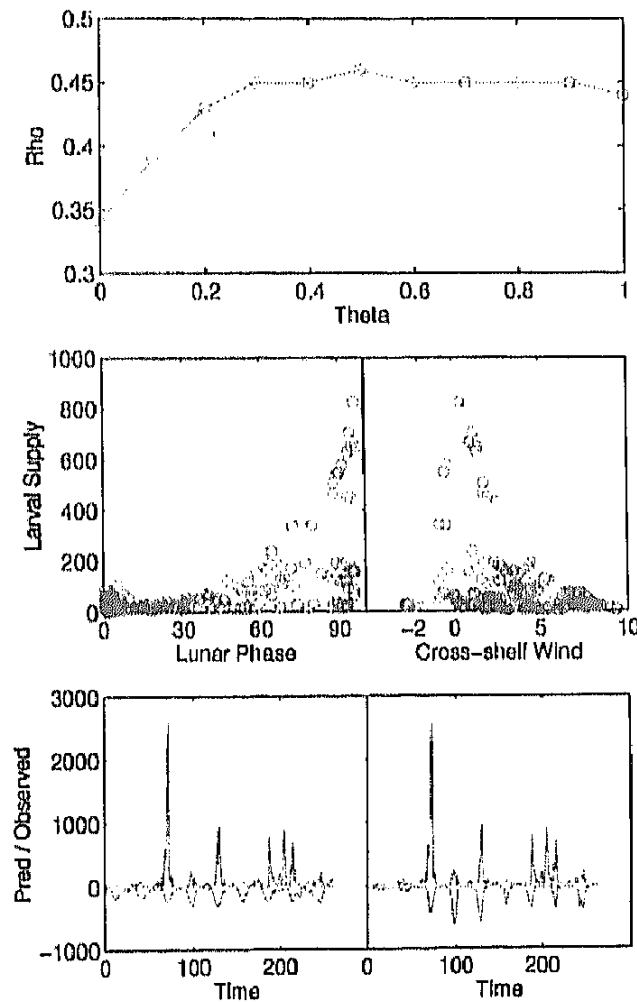


FIGURE 14.11. Modeling pomacentrid larval supply from lunar phase and recent cross-shelf (onshore) wind speed. Top panel: predictability improves as the model is made nonlinear (theta positive). Middle panels: functional responses. Larval supply is as before positively associated with a full moon 19 days prior, and also with weak, onshore winds (positive values = onshore; negative values = offshore). Bottom panels: model predictions, for linear (left) and nonlinear (right) algorithms. Adding in cross-shelf wind speed begins to give a sense of the magnitude of the peaks.

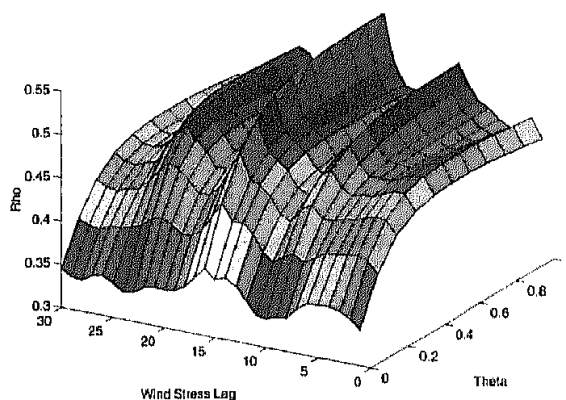


FIGURE 14.12. Investigating the importance of wind events early in larval life. There is a peak in predictability when wind stress lagged 16 days is incorporated into nonlinear models. It appears that first-feeding success is indeed important to subsequent recruitment in this system.

relationship between turbulence and feeding success has been the subject of much study over the past ten years, both theoretical and experimental (e.g., Rothschild and Osborne 1988[32], Davies et al. 1991[6], MacKenzie et al. 1994[22], Sundby and Fossum 1990[38], Cury and Roy 1990[4], Wroblewski et al. 1989[42]). Detecting such lagged effects from time series taken at a fixed point in space is, however, problematic, as there is no guarantee that the local history of the physical environment will be the same as that experienced by the larvae themselves, given their planktonic lifestyle. However, it seems at least plausible that such events could be demonstrated at Lizard Island, given the suggestion that retention mechanisms may work to keep larvae local to their native reef (Leis 1986[19]). The difficulty is that, because not all larvae are precisely the same age and because both the time required for yolk-sac resorption and the time to inevitable mortality in the absence of successful feeding are unknown, it is not obvious exactly which lag should be chosen. As a sanity check, and to be conservative, a wide range of lags (from one to thirty days) was therefore evaluated, each over a range of degrees of model nonlinearity. Clearly, if the modeling results indicate that the best lag corresponds to a time (for example) before the fish were even born, something is amiss. Results are given in Fig. 14.12. Considered linearly, there is little to distinguish among the lag choices. However, as the models are made nonlinear, a single peak in predictability beyond that observed for the bivariate nonlinear model emerges. Strikingly, this peak coincides very well with the age of the larvae. The pomacentrid larval duration averaged nineteen days; the peak in predictability is centered at sixteen days, three days after release from the nest.

The approach to variable selection seems to be yielding a reasonable result, but do the functional relationships contained in the model continue to

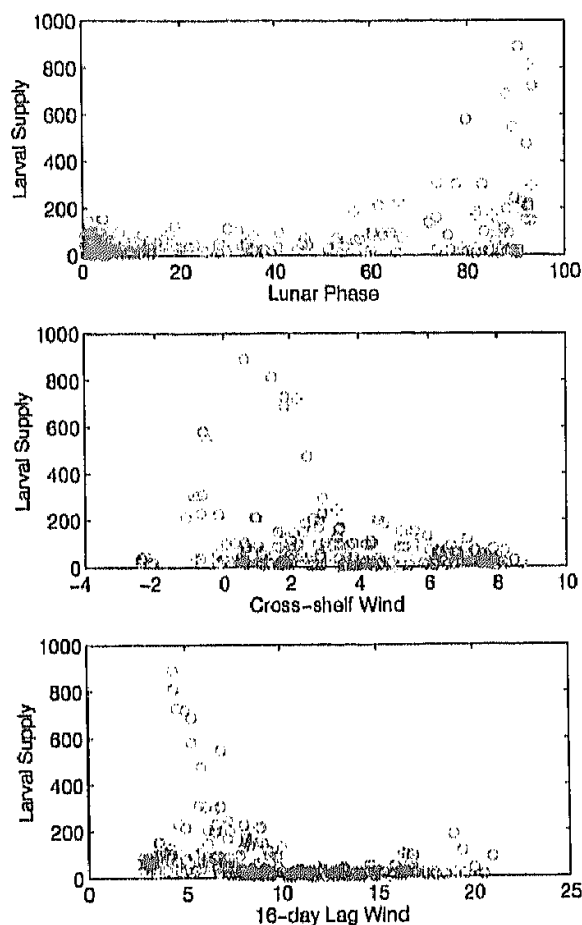


FIGURE 14.13. Functions relating larval supply to the physical environment extracted from the trivariate physical model. Top two panels are essentially as before; now (bottom panel) larval supply can be seen to be associated with intermediate wind speeds (centered on 6 meters per second) 16 days prior, when larvae are very young and presumably first-feeding.

make biological sense? Figure 14.13 gives these, extracted as before. Larval supply peaks remain associated with the full moon nineteen days prior, and recent weak, onshore winds. Additionally, it now appears that there is an approximately dome-shaped response to the sixteen-day lag wind stress, centered on about 6 meters per second average daily wind speed. This is in good agreement with the results of the previously mentioned theoretical and experimental studies, and is the first demonstration from field observations that such events are important enough that they can be used to help predict subsequent recruitment. Figure 14.14 gives model forecasts, plotted with the real data. Clearly there is much variability as yet unexplained, but it is possible to at least begin to predict the magnitude of the peaks in larval supply.

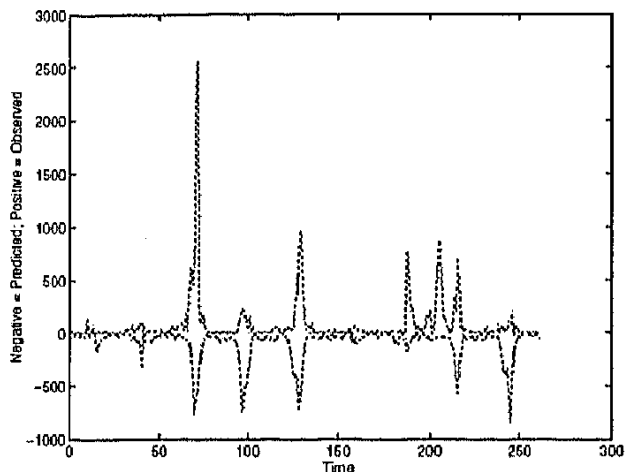


FIGURE 14.14. Output of the trivariate nonlinear physical model: the best predictions made here for larval supply in this system.

Variables	Lag	Model type	ρ	Err	RMS
		Linear (out-of-sample)	0.29	100.5	241.4
		Nonlinear S-map ($\theta = 0.6$)	0.30	103.1	241.0
Cross-shelf wind	1	Linear (out-of-sample)	0.34	109.1	237.4
		Nonlinear S-map ($\theta = 0.6$)	0.45	94.1	221.0
Cross-shelf wind	1	Linear (out-of-sample)	0.38	114.8	233.9
Daily wind	16	Nonlinear S-map ($\theta = 0.8$)	0.55	88.2	211.5

TABLE 14.1. Summary of model statistics. Lag in days; Rho gives correlation coefficient between predicted and observed values, Err is the average error of the predictions, and RMS is the root mean squared error from predictions.

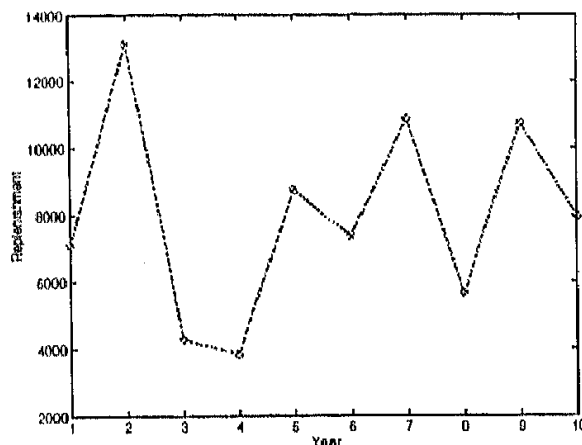


FIGURE 14.15. Yearly resupply of larval pomacentrids simulated for ten years by running the nonlinear trivariate model on surrogate physical inputs. The 4-fold variability in replenishment during this time arises from stationary processes; there is no interannual variability in the physical forcing. It is the specific realizations of the stochastic environment that are important.

14.4 Conclusions

The goal of this chapter has been to explore in detail a case study where the application of nonlinear techniques leads to real gains in understanding the variability in a natural biological system. Although the forecast skill of the best model constructed in this example is still relatively weak, it is a good deal better than that which can be achieved with linear analysis. More importantly, the model identifies processes which have long been suspected of being important in driving change in fish populations, but whose operation in nature has historically been very difficult to demonstrate.

This work is part of an emerging trend in ecology to explore the interactions between noise and nonlinearity in population dynamics (Leirs et al. 1997[18], Grenfell et al. 1998[11], Myers et al. 1998[27], Dixon et al. 1999[7], Zimmer 1999[43]), and the results presented here and elsewhere raise a host of new questions for ecologists to consider. One issue in particular that may need further consideration is how to think about interannual variability. In situations such as that presented here, it is entirely possible for the interplay between noise and nonlinearity to cause a population whose variability is a product of stationary processes to itself appear non-stationary on interannual time scales. For example, Figure 14.15 gives the result of running the nonlinear trivariate physical model of larval supply for multiple years on surrogate physical inputs. Larval supply is summed over each year and plotted over time for ten years. Population replenishment varies by a factor of more than 3 from year to year, and yet this model is driven entirely by stationary physical processes. This is not to say that interannual variability in physical forcing is unimportant; the opposite is

very often the case. Nevertheless, in this example, the tempting conclusion that interannual variability in the physical environment was underlying the observed variability in population resupply would have been premature.

Some recent analyses of time series of natural populations have focused on understanding the density-dependent component of population change in the presence of environmental noise by exploring whether the internal dynamics themselves were chaotic or stable (Ellner and Turchin 1995[9]). The impact of external forcing was dealt with by subjecting nearby initial biological conditions to identical sequences of random shocks. In other words, the question that was asked was whether similar initial population states would diverge rapidly from each other if environmental variability was realized the same way each time. In the example given here, however, the biological variability results directly from different *realizations* of the stochastic forcing processes. If the goal is to predict variability from measurable initial conditions, the suggestion is that not only will the relevant physical variables will have to be identified (along with which are the favorable and unfavorable conditions), but also the sensitivity of natural populations to the temporal sequences of physical events will need to be understood. And, of course, in many cases, such sensitivity may be a function of population size itself. Clearly, a full understanding of the interplay between noise and nonlinearity in natural systems is a difficult challenge. The potential rewards are great, however, and this is likely to be an area where the analytical approaches that are the focus of this book will make strong contributions to the understanding of ecological systems in the future.

References

- [1] H. G. Andrewartha and L. C. Birch. *The Distribution and Abundance of Animals*. University of Chicago Press, 1954.
- [2] N. Caputi. Factors affecting the time series bias in stock-recruitment relationships and the interaction between time series and measurement error bias. *Canadian Journal of Fisheries and Aquatic Science*, 45, 1988.
- [3] R. M. Cassie. Microdistribution of plankton. *Oceanography and Marine Biology Annual Review*, 1:223–252, 1963.
- [4] P. Cury and C. Roy. Optimal environmental window and pelagic fish recruitment success in upwelling areas. *Canadian Journal of Fisheries and Aquatic Science*, 46, 1989.
- [5] D. Cushing. *Climate and Fisheries*. Academic Press, London, 1982.
- [6] C. Davies, G. Flierl, P. Wiebe, and P. Franks. Micropatchiness, turbulence, and recruitment in plankton. *Journal of Marine Research*, 49, 1991.
- [7] P.A. Dixon, M.J. Milicich, and G. Sugihara. Episodic fluctuations in larval supply. *Science*, 283:1528–1531, 1999.
- [8] P. J. Doherty and D.McB. Williams. The preplenishment of coral reef fishes. *Oceanography and Marine Biology*, 26:487–551, 1988.

- [9] S. Ellner and P. Turchin. Chaos in a noisy world: New methods and evidence from time-series analysis. *The American Naturalist*, 145:343-375, 1995.
- [10] J. P. Finerty. *The Population Ecology of Cycles in Small Mammals*. Yale University Press, New Haven, 1980.
- [11] B.T. Grenfell, K. Wilson, B.F. Pinkenstadt, T.N. Coulson, S. Murray, S.D. Albon, J.M. Pemberton, T.H. Clutton-Brock, and M.J. Crawley. Noise and determinism in synchronized sheep dynamics. *Nature*, 394:674-677, 1998.
- [12] R. Hilborn and C. Walters. *Quantitative Fisheries Stock Assessment: Choice, Dynamics, and Uncertainty*. Chapman and Hall, London, 1992.
- [13] J. Hjort. Fluctuations in the great fisheries of northern europe viewed in the light of biological research. *Rapp. P.-V. Reun. Cons. Perm. Int. Explor. Mer.*, 20:1-228, 1914.
- [14] H. Kantz and T. Schreiber. *Nonlinear Time Series Analysis*. Cambridge University Press, New York, 1997.
- [15] Z. Kopal. *Physics and Astronomy of the Moon*. Academic Press, New York and London, 1971.
- [16] D.L. Lack. *The Natural Regulation of Animal Numbers*. Oxford Press, 1954.
- [17] R. Lasker. *Marine Fish Larvae*. University of Washington Press, Seattle, 1981.
- [18] H. Leirs, N.C. Stenseth, J.D. Nichols, J.E. Hines, R. Verhagen, and W. Verheyen. Stochastic seasonality and nonlinear density-dependent factors regulate population size in an african rodent. *Nature*, 389:176-180, 1997.
- [19] J.M. Leis. Vertical and horizontal distribution of fish larvae near coral reefs at lizard island, great barrier reef. *Marine Biology*, 90:505-516, 1986.
- [20] J.M. Leis and B.M. Carson-Ewart. In situ swimming speeds of the late pelagic larvae of some indo-pacific coral-reef fishes. *Marine Ecology Progress Series*, 159:165-174, 1997.
- [21] J.M. Leis, H.P.A. Sweatman, and S.E. Reader. What the pelagic stages of coral reef fishes are doing out in blue water: Daytime field observations of larval behavioural capabilities. *Marine and Freshwater Research*, 47:401-411, 1996.
- [22] B. MacKenzie, T. Miller, S. Cyr, and W. Leggett. Evidence for a dome-shaped relationship between turbulence and larval fish ingestion rates. *Limnology and Oceanography*, 39:1790-1799, 1994.
- [23] R.M. May. Biological populations with nonoverlapping generations: Stable points, stable cycles, and chaos. *Science*, 186:645-647, 1974.
- [24] M.G. Meekan, M.J. Milicich, and P.J. Doherty. Larval production drives temporal patterns of larval supply and recruitment of a coral reef damselfish. *Marine Ecology Progress Series*, 93:217-225, 1993.
- [25] M.J. Milicich. Dynamic coupling of reef fish replenishment and oceanographic processes. *Marine Ecology Progress Series*, 110:135-144, 1994.
- [26] M.J. Milicich, M.G. Meekan, and P.J. Doherty. Larval supply: A good predictor of recruitment of three species of reef fish (pomacentridae). *Marine Ecology Progress Series*, 86:153-166, 1992.
- [27] R.A. Myers, M. Gordon, J.M. Bridson, and M.J. Bradford. Simple dynamics underlie sockeye salmon (*oncorhynchus nerka*) cycles. *Canadian Journal of Fisheries and Aquatic Science*, 55:2355-2364, 1998.
- [28] A.J. Nicholson. The balance of animal populations. *Journal of Animal Ecology*, 2:132-178, 1933.

- [29] R.M. Peterman, M.J. Bradford, N.C.H. Lo, and R.D. Methot. Contribution of early life history stages to interannual variability in recruitment of northern anchovy (*engraulis mordax*). *Canadian Journal of Fisheries and Aquatic Sciences*, 45:8-16, 1988.
- [30] W.J. Richards and K.C. Lindeman. Recruitment dynamics of reef fishes: Planktonic processes, settlement and demersal ecologies, and fishery analysis. *Bulletin of Marine Science*, 41:392-410, 1987.
- [31] B.J. Rothschild. *Dynamics of Marine Fish Populations*. Harvard University Press, Cambridge, 1986.
- [32] B.J. Rothschild and T.R. Osborne. Small-scale turbulence and plankton contact rates. *Journal of Plankton Research*, 10:465-474, 1988.
- [33] W.M. Schaffer. Order and chaos in ecological systems. *Ecology*, 66:93-106, 1985.
- [34] W.M. Schaffer and M. Kot. Chaos in ecological systems: The coals that newcastle forgot. *Trends in Ecology and Evolution*, 1:58-63, 1986.
- [35] I.C. Stobutzki and D.R. Bellwood. An analysis of the sustained swimming abilities of pre- and post-settlement coral reef fishes. *Journal of Experimental Marine Biology and Ecology*, 175:275-286, 1994.
- [36] G. Sugihara. Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions of the Royal Society of London A*, 348:477-495, 1994.
- [37] G. Sugihara, M. Casdagli, E. Habjan, D. Hess, P. Dixon, and G. Holland. Residual delay maps unveil global patterns of atmospheric nonlinearity and produce improved local forecasts. *Proceedings of the National Academy of Science*, In press, 1999.
- [38] S. Sundby and P. Fossum. Feeding conditions of arcto-norwegian cod larvae compared with the rothschild-osborne theory on small-scale turbulence and plankton contact rates. *Journal of Plankton Research*, 12:1153-1162, 1990.
- [39] J. Theiler, B. Galdrikian, A. Longtin, S. Eubank, and J.D. Farmer. Using surrogate data to detect nonlinearity in time series. In M. Casdagli and S. Eubank, editors, *Nonlinear Modeling and Forecasting*, Santa Fe Institute Studies in the Sciences of Complexity, pages 163-188, 1992.
- [40] G.M. Wellington and B.C. Victor. Planktonic larval duration of one hundred species of pacific and atlantic damselfishes. *Marine Biology*, 101:557-567, 1989.
- [41] D.McB. Williams and S. English. Distribution of fish larvae around a coral reef: Direct detection of a meso-scale, multi-specific patch? *Continental Shelf Research*, 112:923-937, 1992.
- [42] J.S. Wroblewski, J.G. Richman, and G.L. Mellor. Optimal wind conditions for the survival of larval northern anchovy, *engraulis mordax* - a modeling investigation. *Fishery Bulletin*, 87:387-398, 1989.
- [43] C. Zimmer. Life after chaos. *Science*, 284:83-86, 1999.