Regime shifts and red noise in the North Pacific

James E. Overland$^{1,3}$
Donald B. Percival$^2$
Harold O. Mofjeld$^1$

$^1$NOAA/Pacific Marine Environmental Laboratory
Seattle, WA 98115

$^2$Applied Physics Laboratory/University of Washington
Seattle, WA 98195

$^3$Corresponding author:
James.E.Overland@noaa.gov
1.206.526.6795
NOAA/PMEL
7600 Sand Point Way NE
Seattle, WA, 98115-6349

Resubmitted to *Deep-Sea Research I*

7 July 2005

Contribution 2708 from NOAA/Pacific Marine Environmental Laboratory
Abstract

Regimes and regime shifts are potentially important concepts for understanding decadal variability in the physical system of the North Pacific because of the potential for an ecosystem to reorganize itself in response to such shifts. There are two prevalent senses in which these concepts are taken in the literature. The first is a formal definition and posits multiple stable states and rapid transitions between these states. The second is more data oriented and identifies local regimes based on differing average climatic levels over a multi-annual duration. This second formulation is consistent with realizations from stochastic red noise processes to a degree that depends upon the particular model. The terminologies climatic regime shift, statistical regime shift or climatic event are useful for distinguishing this second formulation from the first. To illustrate the difficulty of advocating one formulation over the other based upon a relatively short time series, we compare three simple models for the Pacific Decadal Oscillation (PDO). The 104 year PDO record is insufficient to statistically distinguish a single preference between a square wave oscillator that is consistent with the first definition for regime shifts and two red noise models that are compatible with climate regime shifts. Because of the inability to distinguish between underlying processes based upon data, it is necessary to entertain multiple models and to consider how each model would impact resource management. In particular the persistence in the fitted models implies that certain probabilistic statements can be made regarding climatic regime shifts, but we caution against extrapolation to future states based on curve fitting techniques.
1. Introduction

Whether regime shifts are a credible description of low frequency variability for North Pacific climate is an important issue both for climate and ecosystem studies, as there is the potential for an ecosystem to reorganize itself in response to such shifts. Interpretation of the phrase “regime shift” is a relevant topic as it is currently the basis of over 100 journal articles and for analyses during the GLOBEC project synthesis phase (Benson and Trites, 2002; Steele, 1998). We discuss the use of this phrase, but point out that the main difficulty in its application is the inability to infer a unique underlying system structure from relatively short time records.

A strict interpretation of “regimes” and “regime shifts” involves the notion of multiple stable states for a physical or ecological system with a tendency to remain in such states and transition rapidly to another state (Spenser and Collie, 1996; Mantua, 2004). Multiple stable states are a feature of non-linear deterministic systems, which can exhibit chaotic behavior. The equations describing such a system are deterministic, but their solutions are sensitive to initial conditions and can describe a quasi-periodic structure with irregular behavior. For example, the classic Lorenz model describes a system that switches between regimes or “attractors” (Lorenz, 1963; Palmer, 1999). In keeping with this idea, several authors have suggested that the Northern Hemisphere climate has a tendency to be found in multiple preferred patterns (Kimoto and Ghil, 1993; Corti et al., 1999).

An alternate interpretation for climate variability has been multiple low frequency oscillations from known or unknown causes, which might imply limited predictability. One form of this second interpretation has been to refer to regime shifts as simply interdecadal fluctuations (Rothchild et al., 2004; Duffy-Anderson et al., 2005). From this perspective the concept of
climate noise is relevant. “Noise” comes from the field of random processes in which time series are often represented as sums of individual sinusoids with different amplitudes, frequencies and phases. A white noise process is represented by a spectrum where the power (average squared amplitude) is constant as a function of frequency; for red noise the power increases with decreasing frequency. A realization of such processes is created by summing sinusoids with random phases and amplitudes. Each realization of a red noise process creates extended intervals or “runs” where the time series will remain above or below its mean value. Different realizations of the same process would have the same statistical properties of runs, but their evolution will be different. Such runs are a property of long realizations from stationary Gaussian random red noise (Rudnick and Davis, 2003). For example Stephenson et al. (2004) note that a single modal description is also a possible interpretation of Northern Hemispheric climate.

An important subclass of red noise in oceanography is the first order autoregressive (AR1) process. The rate of change of temperature of the upper ocean can be modeled by externally forced air-sea exchange represented by white noise and a memory term proportional to the anomaly of temperature (Hasselmann, 1976; Pierce, 2001). The result is an AR1 process (see next section) for the temperature time series in which the memory represents the effect of the ocean mixed layer at interannual time scales.

A purpose of this paper is to suggest that regime-like behavior cannot be ruled out as a conceptual model for North Pacific variability and should be entertained because of its potential implications for ecosystems and for resource management. Support for this suggestion is based on an investigation of the Pacific Decadal Oscillation (PDO, Mantua et al., 1997) time series (Fig. 1). The main difficulty in understanding interdecadal variability of the North Pacific from data is the relative shortness of the ~100 year record. Given the short record, we will show that
an AR1 red noise model, an alternate long memory red noise model, and a periodic model in the spirit of a strict interpretation of regime shifts, all fit the PDO time series equally well. None can be considered as more basic or taken as the only null hypothesis for Pacific variability compared with the others. The viability of the red noise formulations also provides a caution against making extended predictions for the state of the North Pacific based on time series analysis of the existing record.

It is unlikely that it will be possible to clearly determine from data alone whether there was a proximate deterministic cause for a major shift in the physical system of the North Pacific or whether such a shift is the result of random variations. In century or longer realizations of a red noise process, the state of the system varies around its mean, but on decadal scales there can be apparent local step-like features in the time series and multi-year intervals where the state remains consistently above or below the long-term mean. Thus for practical application in understanding the impact of climate variations on ecosystems, a definition of regimes based solely on distinct multiple stable states is difficult to apply. An alternate more data oriented approach is to tie the concept to a local notion of *differing average climatic levels over a multi-annual duration*. We can make this distinction clear by referring to the latter as a *climatic regime shift* (Bakun, 2004), a *statistical regime shift* (Rudnick and Davis, 2003) or simply a *climatic event*. Bakun (2004) uses the phrase *climatic regime shift* to denote apparent transitions between differing average climatic levels over *multi-annual* to multi-decadal periods. This is to be distinguished from an *ecosystem regime shift* or *phase transition* which implies changes from a quantifiable ecosystem state (i.e., multiple components on a regional or larger scale) that persist long enough to be observed (deYoung *et al.*, 2004; Duffy-Anderson *et al.*, 2005); these changes are often defined by multiple states in abundance and/or community composition (Steele, 2004).
Due to the shortness of most data records, the use of the less precise climatic regime shift or climatic event definition is an important consideration, which would include the result of an extended “run” from a red noise process. While this violates the strict physical definition based on multiple stable states, the inability to determine the unique underlying physical structure of a climate time series leaves the distinction somewhat theoretical. The concept of climatic regimes is useful and has consequences for ecosystem-based management, even if climatic regime shifts or regime duration cannot be accurately predicted (Bemish et al., 2004).

2. Examples of Models for the PDO

We considered three models for the PDO time series [5-month winter average (Nov–Mar) of the first principal component of SST in the North Pacific north of 20°N, http://www.jisao.washington.edu/data_sets/pdo/]. These are an AR1 model, a long memory or fractionally differenced (FD) model, and a square-wave oscillator (SWO) with the addition of white noise. We use the maximum likelihood (ML) method to estimate model parameters (Percival et al., 2001, 2004).

A stationary and causal Gaussian first order autoregressive (AR1) process \( \{ X_t \} \) with mean zero satisfies

\[
X_t = \phi X_{t-1} + \epsilon_t = \sum_{k=0}^{\infty} \phi^k \epsilon_{t-k}
\]

where \( \{ \epsilon_t \} \) is a Gaussian white noise process with mean zero and standard deviation \( \sigma_\epsilon \), and \(|\phi| < 1\). This classic model (1) has two parameters, and we let \( \hat{\phi} \) and \( \hat{\sigma}_\epsilon \) denote their ML estimators.

The stationary and causal Gaussian fractionally differenced (FD) process \( \{ Y_t \} \) with zero mean satisfies
\[ Y_t = \sum_{k=0}^{\infty} \frac{\Gamma(1-\delta)}{\Gamma(k+1)\Gamma(1-\delta-k)} (-1)^k \varepsilon_{t-k} \]  

(2)

where \( \{\varepsilon_t\} \) is the same as before, \(|\delta| < \frac{1}{2}\) and \(\Gamma\) is the gamma function. An FD process is also a two-parameter model, with parameters estimated by \(\hat{\delta}\) and \(\hat{\sigma}_\varepsilon\). Compared to (1), the weights for \(\varepsilon_{t-k}\) decay more slowly as \(k\) increases, leading to a “long-memory” effect. The FD model is based on the assumption that the PDO is the aggregate of many processes with a wide range of time scales.

The third model is that of a deterministic signal observed in the presence of Gaussian white noise, where we consider the re-centered time series of length \(N\) to be a realization of

\[ Z_t = \alpha S_{1t} + \beta S_{2t} + \varepsilon_t \]  

(3)

where \(S_{it}\) describes the signal (normalized such that \(\sum_{t=0}^{N-1} S_{it}^2 = 1, i = 1, 2\)), and \(\{\varepsilon_t\}\) is as before. We determine \(S_{it}\) using a curve fitting procedure called matching pursuit (Mallat and Zhang, 1993; Percival and Walden, 2000) which starts by finding the single best fit to a time series from a large collection of vectors (a “dictionary”), here limited to square wave oscillations (SWO). Figure 2 shows the PDO time series (top) and the first best fit, an SWO with a 76-year period (middle). Statistical tests indicate that one can reject the hypothesis that the residuals from this fit are uncorrelated, so matching pursuit was used again to find the best fit to the residuals, which is an SWO with a 40-year period. The combination of these two series gives a three-parameter SWO model (Fig. 2, bottom) with parameters estimated by \(\hat{\alpha}\), \(\hat{\beta}\), and \(\hat{\sigma}_\varepsilon\). Periods of 20, 40, 50, and 70 years are within the range of suggested fits to PDO data (Mantua et al., 1997; Minobe, 2000; Biondi et al., 2001). The SWO can be considered as a multiple state model with shifts over multi-decadal periods.
The parameter estimates for the three models, AR1, FD, SWO applied to the PDO, along with associated 95% confidence intervals (CIs), are listed in Table 1. Given their CIs, the parameters are different from zero at the 0.05 level of significance. Statistical tests indicate that the residuals from each model are all indistinguishable from white noise. The estimated residual standard deviations, $\hat{\sigma}_e$, are remarkably similar for the three models, with the SWO value being slightly lower for this fitted, three-parameter model. The large relative values of $\hat{\phi}$ and $\hat{\delta}$ show that the AR1 and FD models are both candidates for the PDO time series.

The empirical spectral density function (SDF) for the PDO along with the theoretical SDFs for the three models are shown in the upper panels of Fig. 3, while the model and observed autocorrelation sequences (ACS) are shown in the lower panels; additional details are provided in the figure caption. The theoretical SDFs indicate that the AR1 model focuses on the high frequencies in that it attributes less of the overall variance to low frequencies than the other two models. The stochastic FD model is able to match the full range of frequencies. In the upper right panel the peaks at two Fourier frequencies correspond to the deterministic part of the SWO model which dominate the empirical SDF. This spectral representation is subject to change as the series becomes longer, however, as will be true for any short time series generated by a process with strong low frequency components. The AR1 model tends to underestimate the observed ACS by rapidly converging to zero, whereas the FD and SWO ACSs decrease with lag at a similar rate to the observed ACS. None of the models capture the drop of the autocorrelation function after seven years, but, given the widths of all the confidence intervals for the unknown ACS, this drop might just be due to sampling variability.

While the SDFs and ACSs show that there are qualitative differences among the three models, all the models are viable from a statistical viewpoint, yielding residuals that pass tests.
for uncorrelatedness and that have statistically similar variances. Given that we only have a 104-year record, we cannot reasonably expect to distinguish among the models, as can be seen from the following procedure. For each model, we generate a large number (2500) of simulated time series and fit each series with the other two models. We then judge the adequacy of the fitted (incorrect) models by examining their residuals. For the PDO we find that we need considerably more than a 104-year record to have even a 50% chance of rejecting any incorrect models when another model was the correct one.

We can conclude that, from a statistical point of view, all three models are equally viable for the PDO, but this does not imply that the three models have similar implications for representation of North Pacific physics, particularly with regard to (i) the existence of multiyear regime-like behavior and (ii) predictability. With regard to regimes based on the *climatic regime shift* definition, we can generate samples from all three models to empirically determine the distribution of run lengths above or below zero as a proxy for the distribution of climate regimes. The results of this experiment are shown in Fig. 4, in which the AR1, FD, and SWO models are indicated by thin, thick, and dotted curves. Run lengths of 15 years or longer are much more likely with the FD and SWO models than with the AR1 model. Regime-like behavior in the PDO is thus supported by two of the contending models, FD and SWO. With regard to predictability, the deterministic SWO model with its fixed cycles would predict shifts at specific years in the future, whereas such predictions are not possible with the two other stochastic models. However, predictions from the SWO model are suspect because we really have no definitive physical arguments to support the contention that the patterns of the last 100 years will extend without modification into the future. Comparing the red noise models with the SWO, we conclude that
different models provide different predictions, and thus there is little credibility for deterministic forecasts based on curve fitting to the PDO.

3. Discussion

Von Storch and Zwiers (1999) recommend a comparative approach of competing simple models in geophysical problems such as the one under discussion. We find three models that are statistically competitive in fitting North Pacific time series. The FD model is more regime-like than AR1 for the North Pacific, if the definition of climatic regimes is accepted as run-like behavior.

Additional qualitative information can help in evaluating the merits of competing models. There are multiple physical processes that influence North Pacific long-term variability, both external and internal, but there is little understanding at present of the exact mechanisms (Miller and Schneider, 2000; Schneider et al., 2002). At the simplest level one can consider the oceanic winter mixed layer driven by local atmospheric variability, i.e., the re-emergence mechanism (Pierce, 2001), that has an AR1 signature with short-term memory into the following year. There is multi-year variation in the frequency of ENSO events which might also force an AR1 ocean process (Newman et al., 2003). The long memory FD model might also capture such effects as temperatures and snow cover over Asia having a downstream influence on the North Pacific (Nakamura et al., 2002). Arctic influences appear more important since the 1960s (Overland et al., 1999). Internal processes such as advection of subsurface temperature anomalies and ocean adjustment through Rossby waves have decadal time scales. The existence of such processes, even though their timing and magnitudes are uncertain, provides an argument beyond the statistical one for long-memory processes in the North Pacific.
4. Conclusion

We compare three simple models for low frequency variability of the North Pacific. One model is an AR1 process which can represent interannual reemergence physics of the ocean mixed layer. Two other models (FD and SWO) statistically fit the data as well as the AR1 process, given the relatively short record length of the PDO time series available to resolve interdecadal processes. The FD model is particularly attractive because it can be interpreted as the aggregate of many processes with a wide range of time scales and hence can be linked to plausible physical hypotheses for long-term variability. The SWO model has inherent multiple-state regime shifts that can be used to fit the observed North Pacific patterns over the last century, but is of questionable use for forecasting upcoming states.

Whether the formal regime shift definition based on multiple stable states or red noise concepts apply to the North Pacific is an important theoretical question. However, because none of the three models for the North Pacific can claim statistical primacy and because the SWO and a long-memory red noise models both develop realizations with multi-year events, the formal concept of regimes becomes difficult to apply. The limitations of data analysis based upon relatively short records suggests a useful and less restrictive definition for regime shifts based on “apparent transitions between differing average climatic levels over multi-annual to multi-decadal periods” (Bakun, 2004). To keep this distinction clear, we recommend referring to the latter as a *climatic regime shift, statistical regime shift* or *climatic event.*

Realizations from both the FD and SWO models are consistent with the climatic regime shift interpretation for the North Pacific. Our examination of multiple models is more than an abstract consideration. In advising resource managers, it is often useful to provide a series of options consistent with all available historical information. Based on the multi-year persistence
in both red noise and multiple state models applied to the North Pacific, some probabilistic statements can be made about the occurrence of climatic regime shifts (Rodionov, 2004). However, the lack of decadal predictability based solely on stochastic models, supports the importance of extensive monitoring and numerical modeling of the North Pacific, beyond just a focus on the PDO, and continued research on mechanisms of decadal variability.

**Acknowledgments**

We thank members of the PICES Study Group on Fisheries and Ecosystem Responses to Recent Regime Shifts for discussion and formulation of the issues raised in this note. Support is provided by the Northeast Pacific GLOBEC project and the NOAA NPCREP project. PMEL Contribution 2708.
References


Fig. 1. The Pacific Decadal Oscillation time series.

Fig. 2. Two step matching pursuit analysis of PDO using square wave oscillations. The top curve is the PDO, the middle curve is the fit to a 76-year square wave oscillator, and the bottom curve is the fit to a combination of 76-year and 40-year square wave oscillators. The numbers on the left refer to the percentage of variance explained by each model.

Fig. 3. (Top) Spectral density functions (thick curves) for the fitted AR, FD, and SWO models (left-hand, middle, and right-hand columns, respectively) and the corresponding periodogram (i.e., sample spectral density function) for the PDO (thin curves). Units for f are 1/year and for the Y-axis are PDO squared-years. The error bar in each of the upper plots gives an assessment of the inherent variability in the periodogram. (Bottom) The expected values of the sample autocorrelation sequences (ACSs, thick curves) for the three fitted models and the corresponding sample ACS (depicted as deviations from zero in each plot in the bottom row); units for the X-axis are lags in years. The two thin curves above and below each thick curve indicate limits within which individual sample ACS estimates should fall 95% of the time when the true process is the fitted model. These curves were obtained from computer experiments in which we simulated 10,000 series of the same length as the PDO from a given fitted model, computed sample ACSs for each simulated series and then determined lower and upper limits trapping 9,500 of the sample ACSs at a given lag.

Fig. 4. Probability of observing a regime that is greater than or equal to a specific run length (years). The thin, thick, and dotted curves correspond to the AR, FD, and SWO models. Regime lengths greater than 15 years are much more probable with the FD and SWO models of the PDO than with the AR model.
Table 1

First-Order Auto regressive (AR), fractionally differenced/long- memory (FD), and square wave oscillator plus noise (SWO) process parameter estimates for the PDO index, along with associated 95% confidence intervals (CI).

<table>
<thead>
<tr>
<th>model</th>
<th>parameter</th>
<th>95% CI</th>
<th>σ</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>AR</td>
<td>$\hat{\Phi}$</td>
<td>0.43</td>
<td>[0.26, 0.61]</td>
<td>$\hat{\sigma}_e$</td>
</tr>
<tr>
<td>FD</td>
<td>$\hat{\delta}$</td>
<td>0.33</td>
<td>[0.18, 0.48]</td>
<td>$\hat{\sigma}_e$</td>
</tr>
<tr>
<td>SWO</td>
<td>$\hat{\alpha}$</td>
<td>4.77</td>
<td>[3.38, 6.16]</td>
<td>$\hat{\sigma}_e$</td>
</tr>
<tr>
<td></td>
<td>$\hat{\beta}$</td>
<td>–3.35</td>
<td>[–4.75, –1.96]</td>
<td></td>
</tr>
</tbody>
</table>
Fig. 1. The Pacific Decadal Oscillation time series.
Fig. 2. Two step matching pursuit analysis of PDO using square wave oscillations. The top curve is the PDO, the middle curve is the fit to a 76-year square wave oscillator, and the bottom curve is the fit to a combination of 76-year and 40-year square wave oscillators. The numbers on the left refer to the percentage of variance explained by each model.
Fig. 3. (Top) Spectral density functions (thick curves) for the fitted AR, FD, and SWO models (left-hand, middle, and right-hand columns, respectively) and the corresponding periodogram for the PDO (thin curves). The error bar in each of the upper plots gives an assessment of the inherent variability in the periodogram. (Bottom) The expected values of the sample autocorrelation sequences (ACSs, thick curves) for the three fitted models and the corresponding sample ACS (depicted as deviations from zero in each plot in the bottom row). The two thin curves above and below each thick curve indicate limits within which individual sample ACS estimates should fall 95% of the time when the true process is the fitted model. These curves were obtained from computer experiments in which we simulated 10,000 series of the same length as the PDO from a given fitted model, computed sample ACSs for each simulated series and then determined lower and upper limits trapping 9,500 of the sample ACSs at a given lag.
Fig. 4. Probability of observing a regime that is greater than or equal to a specific run length (years). The thin, thick, and dotted curves correspond to the AR, FD, and SWO models. Regime lengths greater than 15 years are much more probable with the FD and SWO models of the PDO than with the AR model.