Use of SARIMA models to assess data-poor fisheries: a case study with a sciaenid fishery off Portugal

Nuno Prista (contact author)[1,2]

Norou Diawara[3]

Maria José Costa[1,2]

Cynthia Jones[4]

Email address for contact author: nmprista@ipimar.pt

[1] Centro de Oceanografia
Faculdade de Ciências da Universidade de Lisboa
Campo Grande, 1749-016
Lisboa, Portugal
Present address for contact author: Unidade de Recursos Marinhos e Sustentabilidade
Instituto Nacional de Recursos Biológicos (INRB, I.P./IPIMAR)
Avenida de Brasília, 1449-006
Lisboa, Portugal

[2] Departamento de Biologia Animal
Faculdade de Ciências da Universidade de Lisboa
Campo Grande, 1749-016
Lisboa, Portugal

[3] Department of Mathematics and Statistics
Old Dominion University
Norfolk, Virginia 23529-0077

[4] Center for Quantitative Fisheries Ecology
Old Dominion University
800W 46th Street
Norfolk, Virginia 23508-2099

Abstract—Research on assessment and monitoring methods has primarily focused on fisheries with long multivariate data sets. Less research exists on methods applicable to data-poor fisheries with univariate data sets with a small sample size. In this study, we examine the capabilities of seasonal autoregressive integrated moving average (SARIMA) models to fit, forecast, and monitor the landings of such data-poor fisheries. We use a European fishery on meagre (Sciaenidae: *Argyrosomus regius*), where only a short time series of landings was available to model (n=60 months), as our case-study. We show that despite the limited sample size, a SARIMA model could be found that adequately fitted and forecasted the time series of meagre landings (12-month forecasts; mean error: 3.5 tons (t); annual absolute percentage error: 15.4%). We derive model-based prediction intervals and show how they can be used to detect problematic situations in the fishery. Our results indicate that over the course of one year the meagre landings remained within the prediction limits of the model and therefore indicated no need for urgent management intervention. We discuss the information that SARIMA model structure conveys on the meagre life-cycle and fishery, the methodological requirements of SARIMA forecasting of data-poor fisheries landings, and the capabilities SARIMA models present within current efforts to monitor the world's data-poorest resources.

Manuscript submitted 8 March 2010.
Manuscript accepted 20 January 2011.
Fish. Bull. 109:170–185 (2011).

Research, assessment, and management have traditionally focused on fisheries with the greatest landings and revenues (Scandol, 2005; Vasconcellos and Cochrane, 2005). Such fisheries are generally data-rich and have available the funds and expertise required to complete stock assessments and provide state-of-the-art advice to management. However, that is not the case for the vast majority of fisheries worldwide, which remain subjected to limited (if any) assessment and management (Vasconcellos and Cochrane, 2005). The latter have been collectively termed "data-poor fisheries" and are characterized by a low diversity and quantity of data, limitations in funding and expertise, and an overall shortage of assessment methods (Mahon, 1997; Scandol, 2005). Among the world's data-poorest fisheries are nearly all fisheries in developing countries, but also most fisheries in developed countries, namely the smaller-scale or less valuable commercial and recreational ones (NRC, 1998; Berkes et al., 2001; EEA, 2005; Vasconcellos and Cochrane, 2005; Worm et al., 2009; OSPAR, 2010; ICES[1]).

Assessment of data-poor fisheries requires a significantly different approach from their data-rich counterparts. For data-poor fisheries, many deterministic multivariate stock assessment models cannot be used (e.g., NRC, 1998) and more pragmatic assessment methods must be put in place, particularly when fishery-independent data are not available and fishing effort cannot be quantified (Berkes et al., 2001; Scandol, 2003; ICES[1]). In many countries, the most readily available fisheries data are commercial landings because of their

[1] ICES (International Council for the Exploration of the Sea). 2008. Report of the study group on management strategies (SGMAS), 74 p. ICES CM 2008/ ACOM:24, Copenhagen, Denmark.

connection to the economy and business (Vasconcellos and Cochrane, 2005). Commercial landings result from complex interactions between the environment, the fishing fleet, and the stocks, and therefore do not directly reflect the status of exploited populations. However, landing records contain valuable information that can be useful to managers if routine monitoring, rather than stock assessment, is established as a management objective (Scandol, 2003). In fact, even if they provide suboptimal indications on the status of the stocks, statistical analyses of landings can lead to the timely detection of phenomena such as sudden increases in fishing effort or marked population declines that could otherwise remain undetected (Caddy, 1999). Such detection is important—particularly within multispecies, budget-limited, management contexts—because it allows the prioritization of research and management actions toward the subset of fisheries and stocks most likely to be depleted (Scandol, 2003).

Autoregressive integrated moving-average (ARIMA) models are simple time series models that can be used to fit and forecast univariate data such as fisheries landings. With ARIMA models data are assumed to be the output of a stochastic process, generated by unknown causes, from which future values can be predicted as a linear combination of past observations and estimates of current and past random shocks to the system (Box et al., 2008). In fisheries, ARIMA models (and their seasonal multiplicative version, SARIMA) have a long record of successful application that extends from modeling (e.g., Hare and Francis, 1994; Fogarty and Miller, 2004) to short-term forecasting of a variety of variables and resources for both data-rich and data-poor fisheries (Table 1). Specifically, SARIMA models, which are applicable to many already-available landings data sets, have been found to provide both annual and monthly forecasts that are comparable to, or even better than forecasts from many multivariate models, including some with fishing effort among the predictors (Stergiou et al., 1997).

The good record, flexibility, and simplicity of SARIMA models have made them natural candidates for the modeling of data-poor fisheries (Rothschild et al., 1996). However, to date, SARIMA models in fisheries have only been applied in detail on relatively long time series (≥120 months) (Table 1), and a single study has provided a few (but not detailed) results from shorter series (Lloret et al., 2000). Such emphasis of previous SARIMA modeling on long time series finds little support in statistical literature where 50 months is generally regarded as the minimum sample size for model application (e.g., Pankratz, 1983; Chatfield, 1996a). Additionally, most literature to date has focused on SARIMA models as tools to generate accurate forecasts of future landings. However, in addition to good forecasting, these models also possess significant capabilities for monitoring landings that have remained unexplored. These capabilities become apparent when SARIMA models are approached from a statistical process-control perspective and it is made known that SARIMA model forecasts include the assumption of persistence (through time) of the process that generated the data (Box et al., 2008; Mesnil and Petitgas, 2009). Briefly, good landing forecasts are only attainable as long as significant changes do not take place in the fishery; therefore large forecast errors can be regarded as indications that can be changes in the fishery process took place that may require management intervention (Pajuelo and Lorenzo, 1995; Georgakarakos et al., 2006; Box et al., 2008).

In this study, we report the first detailed application of SARIMA models for monitoring of data-poor fisheries landings. We use data from a previously unassessed Portuguese fishery on meagre (Sciaenidae: *Argyrosomus regius*) as our example. The meagre is a valuable top predator from European coastal waters but its stocks have not been analytically assessed because of limitations in data, personnel, and funding existing at the national level. At the time of our analysis only a short time series of monthly landings (60 months) was available for this fishery, a situation that replicates conditions found in many other data-poor fisheries worldwide. We show that the short time series was not a problem for SARIMA modeling and forecasting and that prediction intervals from SARIMA models can be used to provide this fishery with basic monitoring. We suggest that SARIMA models should be more widely considered to extend the coverage of monitoring to all exploited marine resources.

Materials and methods

Meagre (*Argyrosomus regius*) and its fisheries

Meagre is one of the world's largest and most valuable sciaenids (up to 180 cm, 50 kg, and with a US$ 15 per kg exvessel price). It ranges from France to Senegal, and the largest fisheries take place off Mauritania, Morocco, and Egypt. In Europe, the meagre constitutes a prized trophy-fish for anglers and an important income for small-scale commercial fishermen along the Atlantic shores of France, Spain, and Portugal. Its biology and life cycle remain scarcely documented, but recent concerns about the overexploitation of juveniles and interests in aquaculture production have sparked some research. Currently, the fish is known to be fairly long-lived (up to 44 yr) (Prista et al., 2009), to present fast juvenile growth (Morales-Nin et al., 2010) and to spawn at 3–4 yr old (N. Prista, unpubl. data). Data on adult growth and reproduction have not been published, but preliminary reports indicate a life-cycle characterized by fast growth, high fecundity, and a long reproductive span, and that the estuaries of the Gironde (France), Tagus (Portugal), and Guadalquivir (SW Spain) rivers constitute the main spawning habitats (Quéméner, 2002; Prista et al.[2]; N.

[2] Prista, N., C. M. Jones, J. L. Costa, and M. J. Costa. 2008. Inferring fish movements from small-scale fisheries data: the case of *Argyrosomus regius* (Sciaenidae) in Portugal, 19 p. ICES CM 2008/K-19, Copenhagen, Denmark.

Prista, unpubl. data). Marked seasonal variations in landings linked to juvenile and adult migrations have been identified in local fisheries (Quéro and Vayne, 1987; Prista et al.[2]). Overall, adults are thought to come inshore from spring to early summer to spawn but their overwintering grounds are still unknown; juveniles are thought

Table 1

Primary fisheries literature that present seasonal autoregressive integrated moving-average models. Only studies with quantitative forecast results are displayed. "No."=the number of series, "Freq"=the sampling frequency (W=weekly, M=monthly, A=annual), "*n*" is the sample size of the fitting period, "F"=number of forecasts, "models" indicates the type of models compared, and "PI" indicates if prediction intervals were presented (yes, no). "/" separates annual and monthly data sets when both were analyzed. "sp" = species, "nsp groups" = nonspecific groups, "rel." = relative, "CPUE"=catch per unit of effort, "LPUE"=landings per unit of effort.

Reference	Species	Variable	No.	Freq	*n*	F	Models[a]	PI
Saila et al. (1980)	*Jasus edwardsii*	CPUE	1	M	144	12	1,5	n
Mendelssohn (1981)	*Katsuwonus pelamis*	catch/effort	1	M	180	12	12	n
Fogarty (1988)	*Homarus americanus*	catch/CPUE	3/1	A/M	41–58/216	1/12	12	n
Jeffries et al. (1989)	*Pseudopleuronectes americanus*	rel. abundance	2/3	A/M	27/156;324	2/12	––	y
Stergiou (1989)	*Sardina pilchardus*	catch	1	M	204	12	––	n
Noakes et al. (1990)	*Oncorhynchus nerka*	total returns	2	A	24	8	1,10,12,19,20	n
Stergiou (1990a)	*Engraulis encrasicolus*	catch	1	M	252	24	––	n
Stergiou (1990b)	Mullidae	catch	1	M	252	24	––	n
Campbell et al. (1991)	*Homarus americanus*	catch	4	A	61–97	10	12	n
Molinet et al. (1991)	*Penaeus* spp., *Lutjanus synagris*	landings/LPUE	2	M	132;180	24	––	n
Stergiou (1991)	*Trachurus* sp.	catch	1	M	252	12	1,8	n
Tsai and Chai (1992)	*Morone saxatilis*	harvest	1	A	27	4	3,4,12	n
Pajuelo and Lorenzo (1995)	1 nsp group	catch	1	M	131	24	––	y
Stergiou and Christou (1996)	4 sp; 12 nsp groups	catch	16	A	24	2	1–9	n
Stergiou et al. (1997)	4 sp; 12 nsp groups	catch	16	M	288	24	1–5,7–9	n
Park (1998)	*Theragra chalcogramma*	landings	1	M	264	24	––	n
Lloret et al. (2000)[6]	30 sp; 36 nsp groups	catch	66	M	51–200	12	––	y
Georgakarakos et al. (2002, 2006)	*Loligo vulgaris*, *Todarodes sagittatus*	landings	2	M	174	12	11,15,16	y
Pierce and Boyle (2003)	*Loligo forbesi*	LPUE	1	A/M	27/324	3/36	3, 12	y
Stergiou et al. (2003)	*Xiphias gladius*	catch	1	M	180	12	8,13	n
Zhou (2003)	*Oncorhynchus tshawytscha*	spawner density	2	A	11	4	1, 15	n
Hanson et al. (2006)	*Brevoortia tyrannus*, *B. patronus*	landings	2	A	57;63	10	3,14,15	n
Koutroumanidis et al. (2006)	*E. encrasicolus*, *Merluccius merluccius*, *Sarda sarda*	landings	3	M	216;252	12	17,18	n
Czerwinski et al. (2007)	*Hippoglossus stenolepis*	CPUE	1	W	107	31	15	n
Tsitsika et al. (2007)	Total pelagic production *E. encrasicolus*, *S. pilchardus*, *T. trachurus*	CPUE	4	M	180	12	11	y

[a] Models compared: 1=naïve, 2=linear regression (LR), 3=multiple LR, 4=multiple LR with correlated errors, 5=harmonic LR, 6=Fox surplus-yield, 7=model combination, 8=exponential, 9=vector autoregressive, 10=periodic autoregressive, 11=multivariate ARIMA, 12= transfer function noise, 13=census method II (X-11), U.S. Dep. Commer., 14=state space models, 15=artificial neural networks, 16=Bayesian dynamic modeling, 17=genetic modeling for optimal forecasting, 18=fuzzy expected intervals, 19=stock-recruitment, 20=sibling.

[b] The Lloret et al. (2000) study includes 12 series with 51–64 months.

to use estuaries as nursery areas during the warmer months and overwinter in adjoining coastal grounds (Quéro and Vayne, 1987; Quéméner, 2002; Prista et al.[2]; N. Prista, unpubl. data).

Recently, substantial conservation risks have been identified in European meagre fisheries that are related to the overexploitation of juvenile and adults schools in estuaries and nearby coastal areas (Quéméner, 2002; Prista et al.[2]). To protect juveniles, precautionary management measures have been put in place (namely minimum landing size regulations) but the actual status of the meagre stocks was never assessed. This lack of assessment mainly results from a lack of sufficient multivariate time-series data and because national assessment priorities, funding, and expertise are generally allocated to the largest national and transnational fisheries instead of the less-significant, albeit numerous and regionally important, ones. The fish being largely absent from routine fishery-independent surveys (Quéro and Vayne, 1987; F. Cardador, personal commun.[3]) and difficulties related to its sampling at port and the estimation of fishing effort (Prista et al.[2,4]) further contribute to its unassessed status. In this type of setting, if simple methods are not put in place that can, at least, detect the most alarming signals in the landings data it is likely that stock collapses can occur without being detected.

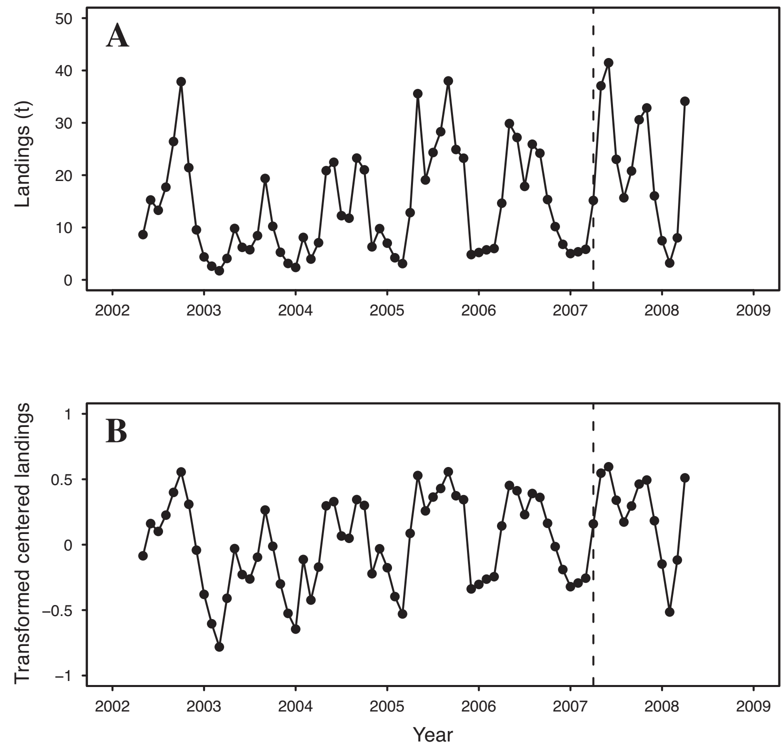

Figure 1

Time series of monthly meagre (*Argyrosomus regius*) landings, in tons, in the Lisboa region of the Portuguese coast (May 2002 to April 2008). The dashed vertical line is the forecast origin (April 2007) and separates the fitting period (May 2002 to April 2007, left) from the hold-out period (May 2007 to April 2008, right). (A) Raw data. (B) Log_{10}-transformed mean-centered data.

Data set and data transformations

The Lisboa region in Central West Portugal (henceforth termed “Lisboa region”) (38°25′N to 38°59′N lat., ~9°15′W long.) is the main fishing area for meagre off the Iberian Peninsula (between 29% and 45% of annual landings of meagre, all gears combined, in 2001–05). In this region, most of the catch is associated with the Tagus estuary and its adjoining coastal area. The catch derives essentially from a small-scale artisanal fleet in which gillnets, trammel nets, and longlines are used to catch meagre during its spawning and nursery season (Prista et al.[2]). To minimize overfishing of juvenile fish, a minimum landing size of 42 cm was established in 2002 that complements an array of other gear-related and effort-related management regulations that are not specific to meagre.

To test SARIMA models in the monitoring of the Lisboa meagre landings, we obtained a time series of meagre monthly landings from the Portuguese General-Directorate for Fisheries and Aquaculture (DGPA). The landings data resulted from mandatory reports of fish sales obtained at all ports of the Lisboa region (N=14) from May 2002 to April 2008 (i.e., 72 monthly values) as part of a routine data collection program (Fig. 1). We used the first 60 months to fit the SARIMA models and the last 12 months as a hold-out period to evaluate forecasting performance and to monitor the fishery. Some previous data were available on this fishery, but those data were found to be unreliable because of contamination with landings from Portuguese vessels operating off North African waters. No significant management interventions occurred on the fishery during the course of our study.

Before fitting a SARIMA model, the time series must be checked for violations of the weak stationarity assumption of the models (Brockwell and Davis, 2002; Box et al., 2008). In SARIMA models, trend and seasonal nonstationarities are handled directly by the model

[3] Cardador, Fátima. 2008. INRB, I.P./IPIMAR, Av. Brasília, 1449-006 Lisboa, Portugal.

[4] Prista, N., J. L. Costa, M. J. Costa, and C. M. Jones. 2007. New methodology for studying large valuable fish in data poor situations: commercial mark-recapture of meagre *Argyrosomus regius* in the southern coast of Portugal, 18 p.

Table 2

Candidate set of seasonal autoregressive integrated moving-average models. The "rule" column displays the mathematical expression used to determine the autoregressive components (p) and moving-average components (q) of the candidate models. "Max AR term" and "Max MA term" columns display the maximum autoregressive (AR) and moving-average (MA) lags included in the model equations, with respect to the original (x_t) and 12-month differenced $\log_{10}$-transformed mean-centered data ($w_t=\nabla_{12}^1 y_t=\nabla_{12}^1 (\log_{10} x_t-4.022)$), respectively.

Model structure	No. of models	Rule	Max AR term	Max MA term
$(p,0,q)\times(0,1,0)_{12}$	325	$q<25-p$; $p\leq24$	w_{t-24}; x_{t-36}	z_{t-12}
$(p,0,q)\times(1,1,0)_{12}$	91	$q<13-p$; $p\leq12$	w_{t-24}; x_{t-36}	z_{t-12}
$(p,0,q)\times(0,1,1)_{12}$	91	$q<13-p$; $p\leq12$	w_{t-12}; x_{t-24}	z_{t-24}
$(p,0,q)\times(1,1,1)_{12}$	1	$q=0$; $p=0$	w_{t-12}; x_{t-24}	z_{t-12}

structure so that only the nonstationarity of variance needs to be addressed before model fitting. The meagre time series (x_t, t=1, … ,60) was seasonal and exhibited no trend (Fig. 1A), but annual variance-mean plots indicated an increase in variance with the series mean. To correct this, we evaluated Box-Cox transformations (Box and Cox, 1964) and found that a $\log_{10}$ transformation successfully stabilized the variance of the series. Accordingly, we log-transformed the data, subtracted its mean, and then used the mean-centered log-transformed data set (y_t, t=1, … ,60) as input to the SARIMA analyses (Fig. 1B).

Data modeling

We fitted SARIMA models to the meagre data using a semi-automated approach based on a combination of the Box-Jenkins method with small-sample, bias-corrected Akaike information criteria (AIC$_c$) model selection (Rothschild et al., 1996; Brockwell and Davis, 2002). This approach involved three major steps: 1) selection of the candidate model set; 2) estimation of the model and determination of AIC$_c$; and 3) a diagnostic check. Details on the notation and model selection procedures used to fit SARIMA models to short time series are given in Appendices 1 and 2.

Selection of the candidate model set was carried out by first analyzing sample estimates of the autocorrelation function (ACF) and partial autocorrelation function (PACF) in order to select the three major orders of the SARIMA models: d, D, and S. In the meagre case, we concluded that a configuration with d=0, D=1, and S=12 should be adopted (see *Results* section). Consequently, a SARIMA$(p,0,q)\times(P,1,Q)_{12}$ was selected as the basic model structure of the candidate set, with p, q, P, and Q left to vary. There is no *a priori* method to determine the maximum value that p, q, P, and Q can take, but the maximum orders of the models are obviously restricted by sample size. In our analysis, we conditioned p, q, P, and Q to the upper boundary $max(p+q+SP+SQ)=24$ and $p+q\leq12$ (Table 2), which caused the maximum possible term of any SARIMA model to be x_{t-36} and the maximum possible number of parameters to be 13. We found this procedure to provide a good compromise between model complexity and the convergence of estimation algorithms.

Model estimation was carried out by using maximum likelihood methods, after conditional sum of squares estimation of the starting values (Brockwell and Davis, 2002). Given the large number of models requiring estimation (Table 2), we developed a semi-automated software routine in R, vers. 2.5.1 (R Development Core Team, 2007) that estimated the models and output their AIC$_c$ values. This routine used several functions incorporated in the R packages "stats" (R Development Core Team, 2007), "tseries" (Trapletti and Hornik, 2007), and "FinTS" (Graves, 2008). After estimation, the model with the minimum AIC$_c$ was selected for further analysis.

Diagnostic checks on the AIC$_c$-selected model involved the following steps: 1) verification of the resemblance of residuals to white noise (ACF plots, Ljung-Box test, cumulative periodogram test); 2) tests on the normality of residuals (Jarque-Bera and Shapiro-Wilks tests); and 3) confirmation of model stationarity, invertibility, and parameter redundancy (Shapiro et al., 1968; Ljung and Box, 1978; Jarque and Bera, 1987; Box et al., 2008). All tests were carried out at a significance level of α=0.05. The variance explained by the model was determined as $1-\hat{\sigma}^2/\sigma_{y_t}^2$ (Stergiou, 1990a).

Forecasts and model performance

We evaluated 12 months of model forecasts, using the last month of the fitting data set as the forecast origin (i.e., April 2007). Forecasts were obtained in the mean-centered transformed scale ($\hat{y}_h$, h=1,...,12) and in the original scale of the data ($\hat{x}_h$, h=1,...,12), after correcting for back-transformation bias (Pankratz, 1983). SARIMA model performance was assessed by comparing h-step forecasts ($\hat{x}_h$ and $\hat{y}_h$) with monthly landings observed between May 2007 and April 2008 (x_h and y_h). This was done by evaluating monthly forecast errors (e.g., $e_h=\hat{x}_h-x_h$) and then considering a set of accuracy measures: 1) annual root mean-square error (RMSE); 2) mean error (ME); 3) absolute percent error (APE$_h$); 4) mean absolute percent error (MAPE); and 5) annual percent error (PE) (Mendelssohn, 1981; Hyndman and Koehler, 2006). From these, RMSE was evaluated in the transformed scale to allow its comparison to $\hat{\sigma}$, and all others were computed in the more user-friendly original scale of the data. Additionally, we compared the forecasting performance of the SARIMA model against two simple naïve forecasting models (naïve model 1 or NM1, and naïve model 2 or NM2) (Noakes et al., 1990; Stergiou et al., 1997). The latter represented *ad hoc* forecasting models likely to be used in data-poor fisheries with short time series of landings: with NM1, future landings were assumed to be equal to the landings registered in the previous year; and with NM2, future landings were assumed to be equal to the average monthly landings registered in the fitting period. We also evaluated the Kitanidis and Bras (1980) coefficient of persistence (P) that summarizes forecasting results by comparing them with those of a naïve model where landings at time t+1 are assumed equal to landings at time t. This coefficient takes values smaller than or equal to 1, with P=1 representing perfect model forecasts.

Monitoring of fisheries

SARIMA models predict the future on the assumption that the statistical properties of the process generating the data remain the same over time (Box et al., 2008). When framed within the perspective of statistical process control (e.g., Scandol, 2005; Box et al., 2008; Mesnil and Petitgas, 2009), this characteristic allows the predictions of well-developed SARIMA models to be used as "guidelines" to monitor future observations. When a SARIMA model is found that appropriately fits the landings data, a significant departure of its forecasts from future observations can be seen as an indication that changes in the underlying fishery process have occurred (=out-of-control situation). In contrast, if such a significant departure does not take place, then there is no indication for such changes (= in-control situation). From a data-poor fisheries perspective, such a distinction means that if funding is limited and multiple fisheries require assessment, research and management efforts should be allocated to fisheries displaying out-of-control decreasing trends in production rather than to fisheries that remain stable or display in-control increasing trends (Scandol, 2003, 2005).

The distinction between in-control and out-of-control landings requires a set of detection limits. To date, process-control detection limits for fisheries indicators have been derived mostly from historical reference data (Scandol, 2003; Mesnil and Petitgas 2009; Petitgas, 2009). However, most fisheries have only a few years of collected data and consequently historical limits are difficult to estimate. In such situations, model-based detection limits like the prediction intervals (PIs) of SARIMA models (Chatfield, 1993; Box et al., 2008) provide easy-to-compute detection limits that explicitly take into account the correlation structure of the data. SARIMA PIs resemble confidence intervals for model forecasts and consist of upper and lower boundaries that encompass a 1–α probability region for future forecasts (Chatfield, 1993). Their main use is to convey the uncertainty around forecasts (De Gooijer and Hyndman, 2006). However, because prediction intervals encompass only future observations, as long as no structural changes take place in the underlying process (Chatfield, 1993), their boundaries can be used to monitor univariate data such as fisheries landings.

To date, the prediction intervals (PIs) from SARIMA models have seldom been reported in fisheries literature and, when they have, with little detail and discussion (Table 1). To monitor the landings of the meagre fishery we used two types of PIs: single step PIs ($PI_{ss,h}$) and multistep PIs ($PI_{ms,h}$). Single step PIs refer to a single monthly forecast (e.g., h=3) and are useful for determining whether a specific monthly observation is an outlier at a given significance level α. Multistep PIs encompass a 1–α prediction region that is a simultaneous PI for all observations registered up to a certain h-step and are useful in detecting systematic departures from historical patterns. We calculated $PI_{ss,h}$ as $\hat{y}_h \pm t_{df,\alpha/2}\sqrt{PMSE_h}$ where $PMSE_h$ is the expected mean squared prediction error at step h and $df=N-DS-d-r$ (Chatfield, 1993; Harvey, 1989). In the calculation of multistep PIs, we used a conservative approach based on a first-order Bonferroni inequality, whereby $PI_{ms,h}$ is given as $\hat{y}_h \pm t_{df,\alpha/2h}\sqrt{PMSE_h}$ and joint prediction intervals of, at least, 1–α around the point forecasts are obtained (Chan et al., 2004).

Results

Data modeling

Large autocorrelations were recorded for lags 1, 2, 11, 12, 23, and 24 with values 0.68, 0.32, 0.44, 0.46, 0.28 and 0.31, respectively (Fig. 2). The sharp decrease in autocorrelation values after lag 2 (0.07 at lag 3) indicated no evidence of a long-term trend; consequently, there was no need to include a first-lag difference term in the SARIMA model structure (d=0). In contrast, large autocorrelation values were registered at annual lags (and its multiples) which indicated the need to include a 12-month difference term in the models (S=12, D=1) (Fig. 2). The ACF and PACF plots of the differenced series provided further support for these conclusions (Fig. 2). Accordingly, a SARIMA$(p,0,q)\times(P,1,Q)_{12}$ was selected as the basic structure of the SARIMA candidate set.

Out of all models in the candidate set, a SARIMA$(0,0,5)\times(1,1,0)_{12}$ was selected as the best model for the meagre data (–2 ln (L)=–26.32, n=48, r=7, AIC_c=–9.52). This model had the following equation:

$$(1+0.65_{\{.10\}}B^{12})\ \nabla^{1}_{12}y_t = (1+0.63_{\{.19\}}B+0.56_{\{.15\}}B^2 + 0.51_{\{.17\}}B^3 + 0.93_{\{.18\}}B^4 + 0.60_{\{.21\}}B^5)z_t\ ,$$

with a noise variance estimate of $\hat{\sigma}$=0.025 and

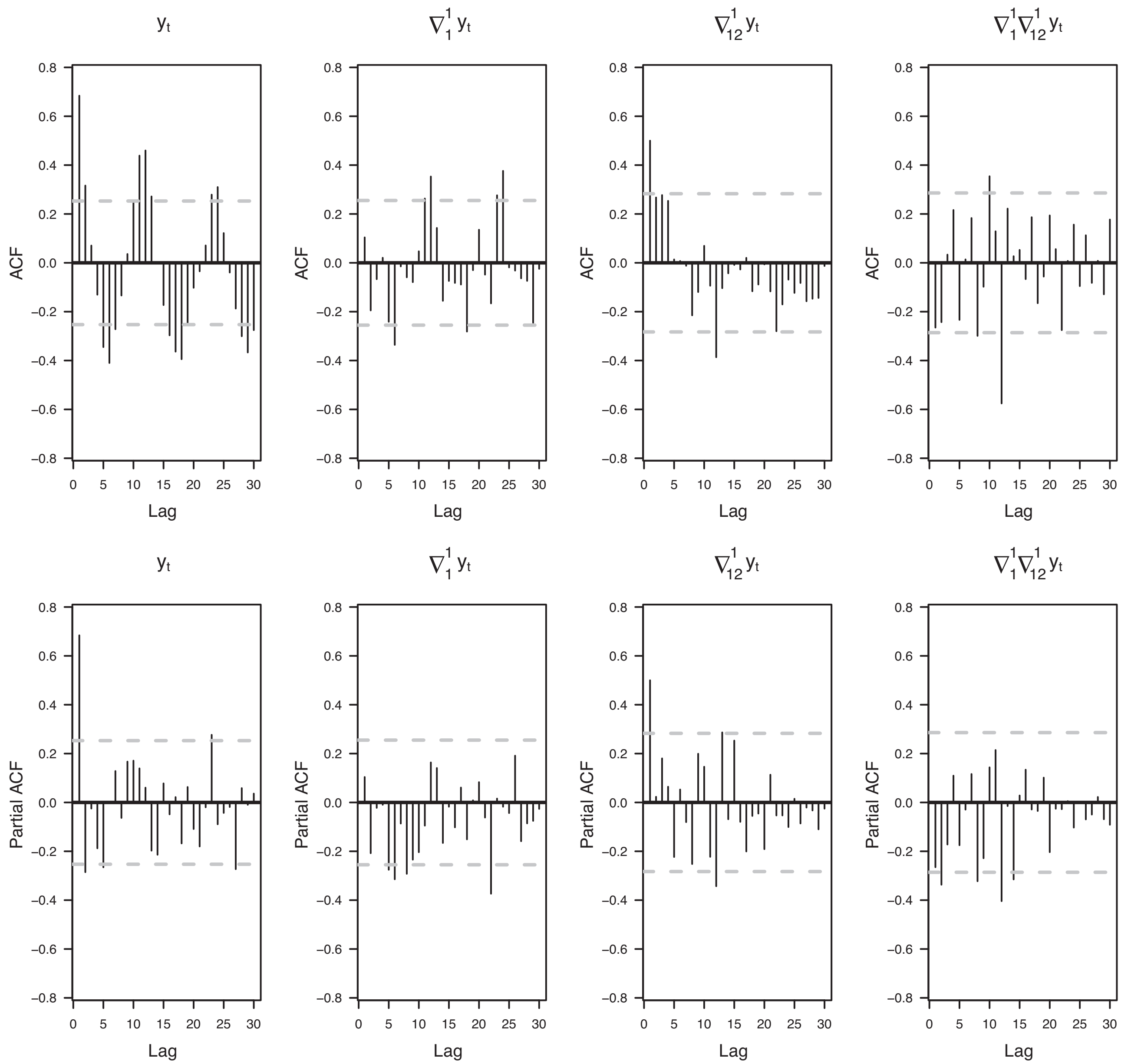

Figure 2

Sample autocorrelation function (ACF) and partial autocorrelation function (PACF) of the transformed meagre (*Argyrosomus regius*) landings. ACF/PACF plots for $\log_{10}$-transformed mean-centered data (y_t, far left), lag-1 differenced series ($\nabla_1^1 y_t$), lag-12 differenced series ($\nabla_{12}^1 y_t$), and lag-1 and lag-12 differenced series ($\nabla_1^1 \nabla_{12}^1 y_t$, far right) are displayed. Horizontal dashed lines represent the 95% confidence limits valid under the null hypothesis of white noise error structure.

where y_t = the mean-centered log-transformed meagre series (i.e., $y_t = \log_{10} x_t - 4.022$) and the values in {} are the standard errors of the estimates.

Diagnostic checks indicated that the SARIMA model was stationary and invertible and did not have redundant parameters. The residuals were white noise (Ljung-Box Q=3.35, P-value>0.05) and passed asymptotic normality tests (Shapiro-Wilk W=0.97, P-value >0.05; Jarque-Bera LM=4.91, P-value >0.05) indicating the model fitted the data and errors were normally distributed. The model explained 78.2% of the variance of the series.

The final process equation selected for the meagre data was

$$\log_{10}X_t = 0.35\log_{10}X_{t-12} + 0.65\log_{10}X_{t-24} + Z_t + 0.63Z_{t-1} + 0.56Z_{t-2} + 0.51Z_{t-3} + 0.93Z_{t-4} + 0.60Z_{t-5},$$

where $Z_t \sim N\ (0,\ 0.025)$.

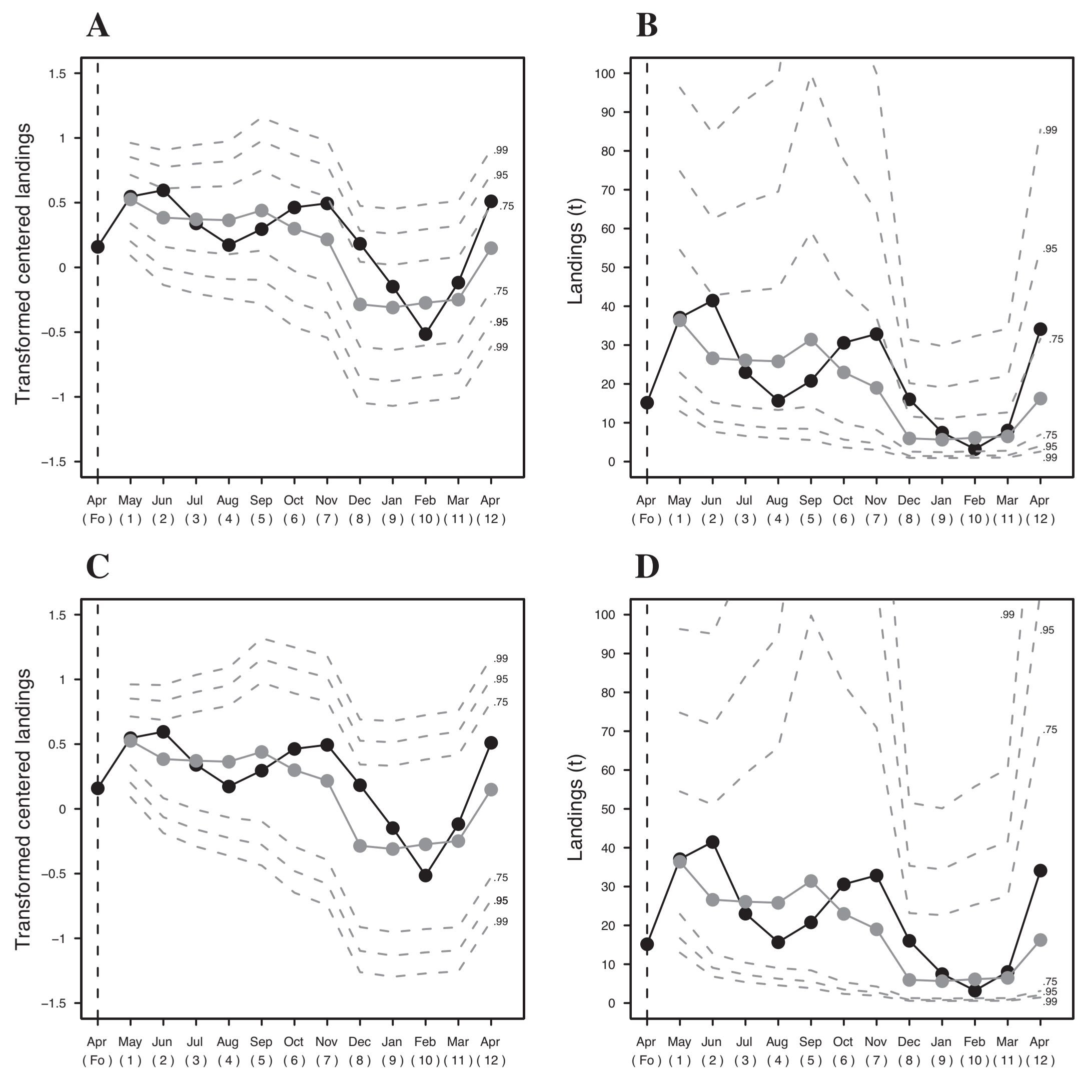

Figure 3

Forecasts and forecast prediction intervals (PIs) of meagre (*Argyrosomus regius*) landings. The dashed vertical line is the forecast origin ("Fo", April 2007). The gray circles and line represent the monthly forecasts. The black circles and line represent observed monthly landings. The dashed gray lines represent the upper and lower 75%, 95%, and 99% prediction intervals. (**A** and **B**) Single step prediction intervals ($PI_{ss,h}$) of transformed centered landings and back-transformed landings, respectively in C and D. Multistep prediction intervals ($PI_{ms,h}$) of transformed centered landings and back-transformed landings, respectively.

Model forecasts and performance

The model forecasts presented two local maxima (May 2007 and September 2007) followed by a four-month period of low landings (December 2007 through March 2008) and an increase in the last month (April 2008) (Fig. 3, Table 3). This pattern in forecasts matched the one in observed landings and the only deviations were that the actual maxima took place one to two months later and the winter trough was sharper than that predicted by the model (Fig. 3). RMSE during the hold-out period (0.234) was ≈1.5 times the RMSE of the fitting period. Eight of the 12 forecasts registered negative errors, but the low ME and PE indicated that underestimation was minor in global terms. APE was large in August, September, December, and April, reflecting the delay in cessation of the 2007 fishing season and the hastening of the 2008 fishing season. Maximum APE coincided with the lowest landings (February), and the minimum APE with the first month forecasted (May) (Table 3). MAPE was 40.3%, reflecting the lagged seasonality and the low landings observed during the winter period.

Table 3

Forecasts of meagre (*Argyrosomus regius*) landings (May 2007 to April 2008). Observed landings (x_h), forecasted landings ($\hat{x}_h$), monthly forecast errors (e_h), monthly absolute percent error (APE_h), mean error (ME), and mean absolute percent error (MAPE) are displayed for the two naïve models (NM1 and NM2) and the seasonal autoregressive integrated moving-average model (SAR). Annual root mean-square error of the mean-centered transformed data (RMSE) and annual percent error (PE) for NM1, NM2 and SAR were 0.261 and 30.2%, 0.285 and 38.9%, and 0.234 and 15.4%, respectively.

			Forecasts ($\hat{x}_h$)			Forecast errors (e_h)			APE_h		
Month	Step (h)	Obs (x_h)	NM1	NM2	SAR	NM1	NM2	SAR	NM1	NM2	SAR
May-07	1	37.1	29.9	21.0	36.4	–7.2	–16.1	–0.7	19.4	43.5	1.8
Jun-07	2	41.5	27.2	18.1	26.6	–14.3	–23.4	–14.9	34.4	56.5	35.8
Jul-07	3	23.0	17.9	14.7	26.1	–5.2	–8.3	+3.1	22.4	36.2	13.3
Aug-07	4	15.7	25.9	18.4	25.8	+10.2	+2.8	+10.1	65.3	17.6	64.7
Sep-07	5	20.8	24.2	26.3	31.4	+3.4	+5.5	+10.6	16.3	26.2	51.1
Oct-07	6	30.6	15.3	21.9	23.0	–15.2	–8.7	–7.6	49.8	28.5	24.9
Nov-07	7	32.9	10.2	13.3	19.0	–22.7	–19.6	–13.9	69.0	59.5	42.2
Dec-07	8	16.1	6.8	6.8	6.0	–9.3	–9.2	–10.1	57.7	57.5	62.8
Jan-08	9	7.5	5.0	4.8	5.7	–2.5	–2.7	–1.8	32.8	35.7	24.5
Feb-08	10	3.2	5.4	5.2	6.1	+2.1	+2.0	+2.9	66.6	61.9	90.7
Mar-08	11	8.0	5.8	4.1	6.5	–2.2	–3.9	–1.5	27.3	48.6	19.0
Apr-08	12	34.1	15.2	10.8	16.3	–18.9	–23.4	–17.9	55.5	68.4	52.4
Mean	1:12	22.5	15.7	13.8	19.1	–6.8	–8.8	–3.5	43.1	45.0	40.3
Sum	1:12	270.5	188.8	165.4	228.9	–81.7	–105.1	–41.6	—	—	—

As with SARIMA forecasts, naïve model predictions also lagged observed values by one or two months. However, the SARIMA forecasts registered the best performance in all accuracy measures, resulting in a 10% to 18% reduction in RMSE, 49% to 60% reduction in ME, 6% to 10% reduction in MAPE, and ≈15% reduction in PE (Table 3). The coefficient of persistence of the SARIMA model was also better (P=0.46) than the one registered by NM1 (P=0.23) and NM2 (P=0.03).

Monitoring of fisheries

During the hold-out period, observed landings remained entirely within the 95% prediction intervals of the SARIMA forecasts (Fig. 3), indicating that the observed forecast errors were within the range of values expected from random variability. Consequently the time series for meagre landings may be described as having remained in-control during the forecasting period. The PIs were symmetrical in the log-transformed scale (Fig. 3, A and C), but asymmetrical in the original scale of the data (Fig. 3, B and D). This pattern was expected from predictions of log-transformed data and indicates that sudden increases in monthly landings (positive forecast errors) are considered "more acceptable" than sudden decreases (negative forecast errors). Individual forecast errors that could have signaled an alarm ranged from 4.3 to 23.0 t (negative errors) to 13.5–68.3 t (positive errors). In relative terms, alarms would have been triggered by a higher than 54–75% drop, or by a higher than 105–238% increase, in monthly landings (Table 4). Compared to monthly PIs, multistep PIs were wider as a result of the increasing number of comparisons performed (Table 4). Even so, it is noticeable that such widening took place mainly on their upper boundary, and only a 12% increase was observed on their lower boundary.

Discussion

Interpretation of the models

Univariate SARIMA models based on landings do not have explanatory variables, but several studies have found the mathematical formulation in the models to correlate well with fish life history and fleet dynamics (Stergiou, 1990b; Stergiou et al., 1997; Lloret et al., 2000). In Europe, adult and juvenile meagre are thought to perform spring–summer migrations to major estuaries, remaining there until mid-summer (adults) and autumn (juveniles). These migrations are well known to local fishermen that actively target the meagre schools while they reside in estuarine grounds (Quéro and Vayne, 1987; Prista et al.[2]). Such interactions between fish migrations and directed fishing effort are likely the cause of the strong seasonal component of the SARIMA model because target effort tends to intensify the natural seasonal signal generated by fish migrating through a fishery (Lloret et al., 2000; Prista et al. 2008). In the case of central Portugal, such intensification is likely modulated at an interannual level by the expectations created for local fishermen by catches obtained in pre-

Table 4

Prediction intervals of meagre (*Argyrosomus regius*) landings (May 2007 to April 2008). Point forecasts ($\hat{x}_h$) and 95% boundaries of the single step ($PI_{ss,h}$) and multistep ($PI_{ms,h}$) prediction intervals are displayed. The prediction boundaries are given as absolute errors ($|e_h|$) and absolute percent errors (APE_h) in each monthly forecast step (h). In each cell, the left and right values represent the lower and upper boundaries, respectively.

			$PI_{ss,h}$		$PI_{ms,h}$	
Month	Step (h)	$\hat{x}_h$	$\lvert e_h \rvert$	APE_h	$\lvert e_h \rvert$	APE_h
May-07	1	36.4	19.7–38.4	54–105	19.7–38.4	54–105
Jun-07	2	26.6	16.2–35.8	61–135	17.5–45.0	66–169
Jul-07	3	26.1	16.9–40.5	65–155	18.8–58.0	72–222
Aug-07	4	25.8	17.3–43.7	67–169	19.6–68.8	76–266
Sep-07	5	31.4	23.0–68.3	73–217	25.9–120.0	82–382
Oct-07	6	23.0	17.3–54.7	75–238	19.5–103.6	85–451
Nov-07	7	19.0	14.3–45.2	75–238	16.2–89.7	85–472
Dec-07	8	6.0	4.5–14.2	75–238	5.1–29.4	86–491
Jan-08	9	5.7	4.3–13.5	75–238	4.9–28.8	86–509
Feb-08	10	6.1	4.6–14.6	75–238	5.3–32.2	87–525
Mar-08	11	6.5	4.9–15.5	75–238	5.7–35.1	87–539
Apr-08	12	16.3	12.3–38.7	75–238	14.2–89.9	87–553

ceding years (represented in the seasonal autoregressive term) and, at an intra-annual level, by random environmental and anthropogenic perturbations occurring on the fishery system (represented in the set of nonseasonal moving-average terms).

Model fit and forecast performance

The univariate SARIMA model presented a good fit to the short time series of meagre landings, explaining most of its variance and adequately modeling the seasonality and correlation structure of the data. Similar results were obtained in other studies of short and long time series: up to 68% (Lloret et al., 2000, series ≤64 months), 75% (Saila et al., 1980), 77% (Stergiou et al., 2003), 84–96% (Stergiou, 1989, 1991; Stergiou et al., 1997), and 93% (Pajuelo and Lorenzo, 1995). Taken together, these results indicate that SARIMA models should be adequate for data sets of monthly landings in general, and not just those with larger sample sizes. Bearing in mind that the minimum series length usually stated for SARIMA model fitting is 50 (Pankratz, 1983; Chatfield, 1996b), such generalized applicability may make SARIMA models particularly useful for fisheries with less reliable historical records or where only recently landings have been sampled.

In addition to a good fit, the SARIMA model also provided good short-term forecasts of meagre landings. The fact that all observed values were located within the predicted intervals of the model, and that naïve forecasts presented similarly lagged seasonality, indicates that the main forecast errors more likely resulted from natural variations in the timing of fish migrations and fishing seasons (Quéro and Vayne, 1987; Prista et al.[2]) or from specifics of SARIMA forecasts and accuracy measures (namely, correlation and APE sensitivity to near-zero observations) (Hyndman and Koehler, 2006; Box et al., 2008) than from model misspecification. At the annual level, the 15% error achieved is comparable to results previously obtained in larger data sets and well within the 10–20% range considered acceptable for market-planning and fisheries management (e.g., Mendelssohn, 1981; Pajuelo and Lorenzo, 1995; Hanson et al., 2006). Additionally, SARIMA forecasts clearly outperformed naïve forcasting in all accuracy metrics, underscoring the large benefits of using these models instead of simpler alternatives (Saila et al., 1980; Stergiou, 1991; Stergiou et al., 1997). Considered together with the overall good forecasting performance reported by Lloret et al. (2000) in their shorter series, these results build confidence that SARIMA models are useful for forecasting short time series of landings and thus can substantially contribute to the planning and management of many data-poor fisheries.

Use of SARIMA models to forecast landings of data-poor fisheries

SARIMA models forecast future landings by directly handling the seasonality and autocorrelation structure of the data and assuming the continuity over time of past time series behavior (Box et al., 2008). These models are known to be well adapted to forecast highly seasonal and autocorrelated data (Stergiou et al., 1997; Georgakarakos et al., 2006). Additionally, some authors have reported better SARIMA forecasting performances in fisheries with lower interannual variability, namely those that target benthic and demersal long-lived species (Lloret et al., 2000). The data for meagre are autocorrelated and present a relatively stable seasonal pattern. Also, the meagre is long-lived and a targeted fish in central Portugal (Prista et al., 2009; Prista et al.[2]). Therefore, it is possible such features contributed to the good forecasts obtained from the SARIMA model. However, we note that the landings of many short-lived pelagic species and species with variable seasonal patterns have also been well forecasted with SARIMA models (Stergiou, 1990a; Stergiou et al., 1997; Georgakarakos et al., 2006; Tsitsika et al., 2007) and that the meagre landings also display substantial annual and monthly stochasticity Therefore, such general patterns should not be considered as strict limitations to SARIMA forecasting. More importantly, we note that SARIMA models can forecast well only if they have been adequately identified and estimated, and always under the assumption that the future is behaving like the past (Chatfield, 1993). Consequently, factors like data quality, presence of outliers, and model selection criteria are also very important for model performance. We discuss these next.

The quality of the input data for SARIMA models is determined mainly by the temporal stability of the statistical properties of the fisheries process and the consistency of its sampling over time. Consequently, although accuracy is required for some model applications (e.g., Zhou, 2003), data inaccuracies do not necessarily undermine SARIMA forecasts as long as factors such as fishing practices, regulatory measures, or data collection practices can be assumed to remain constant. When dealing with shorter series, a careful check whether these assumptions hold becomes particularly important because model identification and estimation are very dependent on the few observations available (Hyndman and Kostenko, 2007) and statistical techniques used to incorporate the effects of process changes in the models (e.g., Fogarty and Miller, 2004) are difficult to implement. In the case of meagre, the use of a short and recent time series better supported the assumption that data collection procedures, fishing techniques, fishery regulations, unreported landings, discards, and law enforcement practices did not change over time. In contrast, it is probable that these assumptions were not met in some less successful applications of the model to longer time series (e.g., Park, 1998).

Outliers are known to cause trouble in time series model identification, estimation, and forecasts—an effect that is amplified in shorter time series (Chatfield, 1993; Trívez and Nievas, 1998). The effects of outliers on forecasting performance are most disastrous when they occur near the forecasting origin because there they not only condition model structure and parameter estimates but are directly incorporated into the forecasts (Chatfield, 1993). The meagre data set presented no apparent outliers and this likely contributed to the good fit and forecasting performance achieved. If outliers were present, specific modeling techniques could have been used to estimate their influence, smooth them, or incorporate them into the model (e.g., Chen and Liu, 1993; Lloret et al., 2000). We note, however, that any outlier during the hold-out period could still have changed our perception of model performance, even if it did not compromise the overall adequacy of the SARIMA model to forecast the landings.

In time series analysis, adequate model specification is considered the most important driver of forecasting accuracy (Chatfield, 1996b). The difficulties of specifying an appropriate model increase for data sets with lower information content, such as those of highly variable short time series from more complex processes (Hyndman and Kostenko, 2007; Appendix 2). To date, fisheries applications of SARIMA models have essentially relied on Box-Jenkins (BJ) model selection procedures to specify a model, and models with $p \leq 2$ and $q \leq 2$ have generally been selected (e.g., Mendelssohn, 1981; Pajuelo and Lorenzo, 1995; Lloret et al., 2000). Compared to these, the model for meagre seems overparameterized, but we note that all of its parameters are statistically significant and that the low $\mathrm{RMSE}_{forec.}$ to RMSE_{fit} ratio indicates an excellent correspondence between fit and forecasting performances (Chatfield, 1996b). In fact, although reduced model parameterization is considered beneficial to accuracy in forecasting, the most important aspect of time series analysis is not the number of parameters, but the degree to which the model approximates the statistical process underlying the data and whether or not it achieves the forecasting objectives (Chatfield, 1996b; Burnham and Anderson, 2002). In the case of meagre, had Box-Jenkins procedures been used, the selected models would be simpler and would still adequately fit the data: $(1,0,0)\times(1,1,0)_{12}$ or $(0,0,1)\times(0,1,1)_{12}$. However, they would have performed worse than our AIC_c-selected model in most performance metrics (RMSE: 0.245 and 0.302, APE: 1.7–92.7% and 20.6–72.4%, MAPE: 44.1% and 44.0%, PE: 13.7% and 31.7%, respectively). These results show the impact that different model selection techniques may have on forecasting performance with SARIMA models and stress the importance of considering objective data-driven criteria like AIC_c for circumventing the subjectivities of model selection in smaller data sets (Hurvich and Tsai, 1989; Burnham and Anderson, 2002).

Conclusions

Use of SARIMA models in monitoring fisheries

From a strictly forecasting perspective, SARIMA models have often been criticized for the excessive reliance on past time series behavior and their difficulty in predicting future structural changes (Georgakarakos et al., 2002; Koutroumanidis et al., 2006). Our results show that these drawbacks can become major advantages when SARIMA models are used for monitoring fisheries. At present, none of the European meagre fisheries is subjected to routine analytical assessment. By fitting

SARIMA models to already available landings data we were able to carry out a first baseline evaluation of one such fishery, using limited funds and minimal time.

Our study provides a first example of how SARIMA models can be used to monitor data-poor fisheries. In the case of meagre, the data displayed no trend and the 95% SARIMA prediction intervals fully encompassed all monthly landings, thus indicating a stable "in-control" fishery. Note that by stating this, at no point do we suggest that the meagre fishery is sustainable long-term because landings do not necessarily reflect stock abundance and our study was limited in time. We suggest only that, since no motive for alarm exists in landings data, and because funds, personnel, and expertise are limited at the national level, attention should be allocated to fisheries that, contrary to the meagre, display decreasing trends or out-of-control situations. Similar types of pragmatic reasoning are generally of great help to fisheries managers handling multiple data-poor fishery scenarios because they help them prioritize management actions for the subset of "problematic" resources in a statistically sound way (Scandol, 2003, 2005).

Underlying the usefulness of SARIMA models in monitoring the meagre fishery and other data-poor fisheries is the use of prediction intervals as reference points to signal alarming trends or sudden level shifts in the fisheries process (Caddy, 1999; Scandol, 2003; Mesnil and Petitgas, 2009). SARIMA PIs have been previously reported in the literature (Table 1), but their use in monitoring was not explored or formalized. These intervals are currently the focus of much statistical research on how to deal with their tendency toward "over-optimism," i.e., the fact that nominal 95% prediction intervals generally contain less than 95% of future observations (Chatfield, 1993). Fortunately, from a fisheries conservation perspective such over-optimism does not constitute a major problem because narrower PIs will be more sensitive to changes in the fisheries process.

Statistical process control (SPC) monitoring of univariate fisheries indicators has become the focus of increased research attention (Scandol, 2003, 2005; Mesnil and Petitgas, 2009; Petitgas, 2009; ICES[1]). The use of SARIMA PIs is similar to that of SPC control-charts, which makes them interesting candidates for the simultaneous monitoring of multiple fisheries and fisheries indicators (Caddy, 1999; Scandol, 2005; Petitgas, 2009). For such cases, SARIMA PIs offer the advantage of being model-based and do not require extensive historical reference data. They are also free from the assumption of statistical independence that frequently troubles the estimation of SPC detection limits (Mesnil and Petitgas, 2009). The simulation framework proposed by Scandol (2003, 2005) for SPC charts provides a means whereby SARIMA PIs can be calibrated toward specific detection rates and management goals. Such calibration was beyond the objectives our study but constitutes an interesting research route for those in charge of more holistic fisheries management.

SARIMA models in assessments of data-poor fisheries

Formal stock assessment has traditionally been considered as the starting point of any fisheries assessment (Mahon, 1997; Berkes et al., 2001). Such an approach is highly desirable but will not be implemented easily, nor quickly, in the many existing data-poor fisheries (Vasconcellos and Cochrane, 2005). In fact, NRC (1998) estimated that 16% of U.S. stocks are not subjected to assessment; and the European Environmental Agency (EEA, 2005) estimated that, depending on the region considered, 20–90% of commercial stocks exploited in the Northeast Atlantic and Mediterranean are not routinely assessed. These figures are much worse in developing countries and when discard and bycatch species are included in the estimates (Vasconcellos and Cochrane, 2005). Addressing such situations requires increased focus on alternative stock indicators and assessment methods that can be used to monitor more fisheries by using available (or easily obtainable) data, funds, and human resources (e.g., Caddy, 1999; Scandol, 2005; Mesnil and Petitgas, 2009; OSPAR, 2010; ICES[1]). Univariate time series models fitted to landings data may be, for some time longer, the best possible approach to extend assessment and management coverage to many of these unassessed resources.

SARIMA modeling and process-control schemes do not constitute alternatives to analytical stock assessment models. Rather, whenever possible, they should be seen as statistical tools to support expert judgment, funding allocation, and management decisions in the most data-limited and assessment-limited settings (Scandol, 2003; 2005). SARIMA modeling and model-based monitoring have a range of characteristics that make them worthy of future exploration in data-poor contexts. Among these are their appropriateness to numerous resources and variables, their strong statistical background and ecological plausibility, their good forecasting performance and easy-to-estimate detection limits, and their applicability to both long and short time series. Furthermore, SARIMA models can also be used to model the nonspecific groupings that dominate many landings data sets, or can be upgraded if multivariate data become available (Stergiou et al., 1997; Vasconcellos and Cochrane, 2005). Finally, the availability of SARIMA models in open-source software packages and their routine use in sectors other than fisheries (e.g., sales, economics, engineering) (Brockwell and Davis, 2002; Box et al., 2008) may be decisive advantages in budget-limited and expertise-limited countries.

Acknowledgments

Funding for this work was provided by a "Fundação para a Ciência e a Tecnologia" (FCT) grant BD/12550/2003 to N. Prista and by research project CORV (DGPA–Mare: FEDER–22–05–01–FDR–00036). We thank Direcção Geral das Pescas e Aquicultura (DGPA) for providing the meagre data set. We thank D. S. Stoffer and D. R.

Anderson for suggestions on the use of the "arima" function and AIC_c model selection, respectively. We further thank M. F. Lane, J. L. Costa, J. J. Schaffler, and J. R. Ashford for commenting on earlier drafts of this manuscript. We thank the three anonymous reviewers for their constructive comments on this manuscript.

Literature cited

Akaike, H.
1974. A new look at the statistical model identification. IEEE T. Autom. Contr. 19:716–723.

Berkes, F., R. Mahon, P. McConney, R. Pollnac, and R. Pomeroy.
2001. Managing small-scale fisheries: alternative directions and methods, 320 p. IDRC Books, Ottawa, Canada.

Box, G., and D. Cox.
1964. An analysis of transformations. J. R. Stat. Soc. Ser. B 26:211–243.

Box, G. E. P., G. M. Jenkins, and G. C. Reinsel.
2008. Time series: forecasting and control, 4th ed., 784 p. John Wiley & Sons, Hoboken, NJ.

Brockwell, P., and R. Davis.
2002. Introduction to time series and forecasting, 2nd ed., 469 p. Springer, New York.

Burnham, K. P., and D. R. Anderson.
2002. Model selection and multi-model inference: a practical information-theoretic approach, 2nd ed., 488 p. Springer, New York.

Caddy, J. F.
1999. Deciding on precautionary management measures for a stock based on a suite of limit reference points (LRPs) as a basis for a multi-LRP harvest law. Northwest Atl. Fish. Org. Sci. Counc. Stud. 32:55–68.

Campbell, A., D. J. Noakes, and R. W. Elner.
1991. Temperature and lobster, *Homarus americanus*, yield relationships. Can. J. Fish. Aquat. Sci. 48:2073–2082.

Chan, W., S. Cheung, and K. Wu.
2004. Multiple forecasts with autoregressive time series models: case studies. Math. Comput. Simul. 64:421–430.

Chatfield, C.
1993. Calculating interval forecasts. J. Bus. Econ. Stat. 11:121–135.
1996a. The analysis of time series: an introduction, 5th ed., 283 p. Chapman & Hall/CRC, Boca Raton, FL.
1996b. Model uncertainty and forecast accuracy. J. Forecast. 15:495–508.

Chen, C., and L. -M. Liu.
1993. Joint estimation of model parameters and outlier effects in time series. J. Am. Stat. Assoc. 88:284–297.

Czerwinski, I. A., J. C. Gutiérrez-Estrada, and J. A. Hernando-Casal.
2007. Short-term forecasting of halibut CPUE: linear and non-linear univariate approaches. Fish. Res. 86:120–128.

De Gooijer, J. G., B. Abraham, A. Gould, and L. Robinson.
1985. Methods for determining the order of an autoregressive-moving average process: a survey. Int. Stat. Rev. 53:301–329.

De Gooijer, J. G., and R. J. Hyndman.
2006. 25 years of time series forecasting. Int. J. Forecast. 22:442–473.

EEA (European Environmental Agency).
2005. The European environment—state and outlook 2005, 570 p. European Environmental Agency, Copenhagen, Denmark.

Fogarty, M. J.
1988. Time series models of the Maine lobster fishery: the effect of temperature. Can. J. Fish. Aquat. Sci. 45:1145–1153.

Fogarty, M. J., and T. J. Miller.
2004. Impact of a change in reporting systems in the Maryland blue crab fishery. Fish. Res. 68:37–43.

Georgakarakos, S., J. Haralabous, V. Valavanis, C. Arvanitidis, D. Koutsoubas, and A. Kapantagakis.
2002. Loliginid and ommastrephid stock prediction in Greek waters using time series analysis techniques. Bull. Mar. Sci. 71:269–287.

Georgakarakos, S., D. Koutsoubas, and V. Valavanis.
2006. Time series analysis and forecasting techniques applied on loliginid and ommastrephid landings in Greek waters. Fish. Res. 78:55–71.

Graves, S.
2008. Companion to Tsay (2005) analysis of financial time series. R package vers. 0.10-12. URL: http://CRAN.R-project.org., accessed October 2008.

Hanson, P. J., D. S. Vaughan, and S. Narayan.
2006. Forecasting annual harvests of Atlantic and Gulf menhaden. N. Am. J. Fish. Manag. 26:753–764.

Hare, S. R., and R. C. Francis.
1994. Climate change and salmon production in the Northeast Pacific Ocean. *In* Climate change and northern fish populations (R. J. Beamish, ed.). Can. Spec. Publ. Fish. Aquat. Sci. 121:357–372.

Harvey, A. C.
1989. Forecasting, structural time series models and the Kalman filter, 572 p. Cambridge Univ. Press, Cambridge.

Hurvich, C. M., and C. -L. Tsai.
1989. Regression and time series model selection in small samples. Biometrika 76:297–307.

Hyndman, R. J., and A. B. Koehler.
2006. Another look at measures of forecast accuracy. Int. J. Forecast. 22:679–688.

Hyndman, R. J., and A. V. Kostenko.
2007. Minimum sample size requirements for seasonal forecasting models. Foresight 6:12–15.

Jarque, C., and A. Bera.
1987. Efficient tests for normality, homoscedascity and serial independence of regression residuals. Econ. Lett. 6:255–259.

Jeffries, P., A. Keller, and S. Hale.
1989. Predicting winter flounder (*Pseudopleuronectes americanus*) catches by time series analysis. Can. J. Fish. Aquat. Sci. 46:650–659.

Kitanidis, P. K., and R. L. Bras.
1980. Real-time forecasting with a conceptual hydrologic model 2: applications and results. Water Resour. Res. 16:1034–1044.

Koutroumanidis, T., L. Iliadis, and G. K. Sylaios.
2006. Time-series modeling of fishery landings using ARIMA models and fuzzy expected intervals software. Environ. Model. Softw. 21:1711–1721.

Ljung, G., and G. Box.
1978. On a measure of lack of fit in time series models. Biometrika 65:297–303.

Lloret, J., J. Lleonart, and I. Solé.
2000. Time series modeling of landings in Northwest Mediterranean Sea. ICES J. Mar. Sci. 57:171–184.

Mahon, R.
1997. Does fisheries science serve the needs of managers of small stocks in developing countries? Can. J. Fish. Aquat. Sci. 54:2207–2213.
Mendelssohn, R.
1981. Using Box-Jenkins models to forecast fishery dynamics: identification, estimation, and checking. Fish. Bull. 78:887–896.
Mesnil, B., and P. Petitgas.
2009. Detection of changes in time-series of indicators using CUSUM control charts. Aquat. Living Resour. 22:187–192.
Molinet, R., M. T. Badaracco, and J. J. Salaya.
1991. Time series analysis for the shrimp and snapper fisheries in Golfo Triste, Venezuela. Sci. Mar. 55:427–437.
Morales-Nin, B., A. Grau, S. Pérez-Mayol, and M. M. Gil.
2010. Marking otoliths, age validation and growth of *Argyrosomus regius* juveniles (Sciaenidae). Fish. Res, 106:76–80.
Noakes, D. J., D. W. Welch, M. Henderson, and E. Mansfield.
1990. A comparison of preseason forecasting methods for returns of two British Columbia sockeye salmon stocks. N. Am. J. Fish. Manag. 10:46–57.
NRC (National Research Council).
1998. Improving fish stock assessments, 177 p. National Academy Press, Washington, D.C.
OSPAR (Oslo and Paris Commission).
2010. Quality status report 2010, 176 p. OSPAR Commission, London, UK.
Pajuelo, J. G., and J. M. Lorenzo.
1995. Analysis and forecasting of the demersal fishery of the Canary Islands using an ARIMA model. Sci. Mar. 59:155–164.
Pankratz, A.
1983. Forecasting with univariate Box-Jenkins models: concepts and cases, 576 p. John Wiley & Sons, Hoboken, NJ.
Park, H. -H.
1998. Analysis and prediction of walleye pollock (*Theragra chalcogramma*) landings in Korea by time series analysis. Fish. Res. 38:1–7.
Petitgas, P.
2009. The CUSUM out-of-control table to monitor changes in fish stock status using many indicators. Aquat. Living Resour. 22:201–206.
Pierce, G. J., and P. R. Boyle.
2003. Empirical modelling of interannual trends in abundance of squid (*Loligo forbesi* in Scottish waters. Fish. Res. 59:305–326.
Prista, N., J. L. Costa, M. J. Costa, and C. M. Jones.
2009. Age determination in meagre *Argyrosomus regius*. Relat. Ciênt. Téc. Inst. Invest. Pescas Mar: Série Digital 49:1–54. URL: http://ipimar-inrb.ipimar.pt/pdf/Reln49.pdf, accessed September 2010.
Quéméner, L.
2002. Le maigre commun (*Argyrosomus regius*)—biologie, pêche, marché et potentiel aquacole, 32 p. IFREMER, Plouzané, France. [In French.]
Quéro, J.-C., and J. -J. Vayne.
1987. Le maigre, *Argyrosomus regius* (Asso, 1801) (Pisces, Perciformes, Sciaenidae) du Golfe de Gascogne et des eaux plus septentrionales. Rev. Trav. Inst. Pêches Marit. 49:35–66 [In French.]
R Development Core Team.
2007. R: A language and environment for statistical computing. R Foundation for Statistical Computing. URL: http://www.R-project.org, accessed August 2007.
Rothschild, B. J., S. G. Smith, and H. Li.
1996. The application of time series analysis to fisheries population assessment and modeling. *In* Stock assessment: quantitative methods and applications for small-scale fisheries (V. F. Gallucci, S. B. Saila, D. J. Gustafson, and B. J. Rothschild, eds), p. 354–402. CRC Press, Boca Raton, FL.
Saila, S. B., M. Wigbout, and R. J. Lermit.
1980. Comparison of some time series models for the analysis of fisheries data. ICES J. Mar. Sci. 39:44–52.
Scandol, J.
2003. Use of cumulative sum (CUSUM) control charts of landed catch in the management of fisheries. Fish. Res. 64:19–36.
2005. Use of quality control methods to monitor the status of fish stocks. *In* Fisheries assessment and management in data-limited situations (G. H. Kruse, V. F. Gallucci, D. E. Hay, R. I. Perry, R. M. Peterman, T. C. Shirley, P. D. Spencer, B. Wilson, and D. Woodby, eds.), p. 213–233. Alaska Sea Grant College Program, Univ. Alaska, Fairbanks, AK.
Shapiro, S., M. Wilk, and H. Chen.
1968. A comparative study of various tests for normality. J. Am. Stat. Assoc. 63:1343–1372.
Shumway, R. H., and D. S. Stoffer.
2006. Time series analysis and its applications: with R examples, 2nd ed., 575 p. Springer, New York.
Stergiou, K. I.
1989. Modelling and forecasting the fishery for pilchard *Sardina pilchardus* in Greek waters using ARIMA time-series models. ICES J. Mar. Sci. 46:16–23.
1990a. A seasonal autoregressive model of the anchovy *Engraulis encrasicolus* fishery in the eastern Mediterranean. Fish. Bull. 88:411–414.
1990b. Prediction of the Mullidae fishery in the eastern Mediterranean 24 months in advance. Fish. Res. 9:67–74.
1991. Short-term fisheries forecasting: a comparison of smoothing, ARIMA and regression techniques. J. Appl. Ichthyol. 7:193–204.
Stergiou, K. I., and E. D. Christou.
1996. Modelling and forecasting annual fisheries catches: comparison of regression, univariate and multivariate time series methods. Fish. Res. 25:105–138.
Stergiou, K. I., E. D. Christou, and G. Petrakis.
1997. Modelling and forecasting monthly fisheries catches: comparison of regression, univariate and multivariate time series methods. Fish. Res. 29:55–95.
Stergiou, K. I., G. Tserpes, and P. Peristeraki.
2003. Modelling and forecasting monthly swordfish catches in the Eastern Mediterranean. Sci. Mar. 67:283–290.
Trapletti, A., and K. Hornik.
2007. Tseries: time series analysis and computational finance. R package vers. 0.10-12. URL: http://CRAN.R-project.org, accessed December 2007.
Trívez, F. J., and J. Nievas.
1998. Analyzing the effects of level shifts and temporary changes on the identification of ARIMA models. J. Appl. Stat. 25:409–424.
Tsai, C. -F., and A. -L. Chai.
1992. Short-term forecasting of the striped bass *Morone saxatilis* commercial harvest in the Maryland portion of Chesapeake Bay. Fish. Res. 15:67–82.

Tsitsika, E. V., C. D. Maravelias, and J. Haralabous.
2007. Modeling and forecasting pelagic fish production using univariate and multivariate ARIMA models. Fish. Sci. 73:979–988.

Vasconcellos, M., and K. Cochrane.
2005. Overview of world status of data-limited fisheries: inferences from landing statistics. *In* Fisheries assessment and management in data-limited situations (G. H. Kruse, V. F. Gallucci, D. E. Hay, R. I. Perry, R. M. Peterman, T. C. Shirley, P. D. Spencer, B. Wilson, and D. Woodby, eds.), p. 1–20. Alaska Sea Grant College Program, Univ. Alaska, Fairbanks, AK.

Worm, K., R. Hilborn, J. K. Baum, T. A. Branch, J. S. Collie, C. Costello, M. J. Fogarty, E. A. Fulton, J. A. Hutchings, S. Jennings, O. P. Jensen, H. K. Lotze, P. M. Mace, T. R. McClanahan, C. Minto, S. R. Palumbi, A. M. Parma, D. Ricard, A. A. Rosenberg, R. Watson, and D. Zeller.
2009. Rebuilding global fisheries. Science 325:578–585.

Zhou, S.
2003. Application of artificial neural networks for forecasting salmon escapement. N. Am. J. Fish. Manag. 23:48–59.

Appendix 1

ARIMA and SARIMA models

An extensive review of ARIMA and SARIMA models can be found in, e.g., Box et al. (2008) and Brockwell and Davis (2002). A mean-centered time series x_t can be modeled as an ARIMA(p,d,q), where p, d, q are nonnegative integers, if it can be adequately fitted with the process equation

$$\phi(B)(1-B)^d X_t = \theta(B)Z_t\ ,$$

where for a time interval T, $(X_t)_{t\in T}$ is a sequence of random variables, B is a backshift differencing operator $B^h X_t = X_{t-h}$ (h nonnegative integer), $(1-B)^d X_t = \nabla_1^d X_t$ is stationary, $\phi(B)$ and $\theta(B)$ are linear filters defined as $\phi(B)=1-\phi_1 B-\phi_2 B^2-...-\phi_p B^p$ and $\theta(B)=1+\theta_1 B+\theta_2 B^2+\ldots+\theta_q B^q$ and $(Z_t)_{t\in T}$ is a sequence of uncorrelated random variables with zero mean and variance σ^2 (termed white noise). In ARIMA models the orders p, q, and d define the structure of the model, by specifying the autoregressive (AR) and moving average (MA) components of an autoregressive–moving average process (ARMA[p,q]). d is the degree of differencing ($d\geq 1$) required for X_t to become stationary. This differencing involves the loss of d observations in the series.

The SARIMA $(p,d,q)\times(P,D,Q)_S$ models, where P, D, Q, and S are nonnegative integers, extend the modeling capabilities of ARIMA(p,d,q) models to seasonal processes. The SARIMA process equation is given by

$$\phi(B)\Phi(B^S)(1-B)^d(1-B^S)^D X_t = \theta(B)\Theta(B^S)Z_t\ ,$$

where X_t, Z_t, $\phi(B)$ and $\theta(B)$ are defined as above, $(1-B)^d(1-B^S)^D X_t = \nabla_1^d \nabla_S^D X_t$ is stationary, and $\Phi(B^S)$ and $\Theta(B^S)$ are seasonal linear filters defined as $\Phi(B^S)=1-\Phi_1 B^S-\Phi_2 B^{2S}-\ \ldots\ -\Phi_P B^{PS}$ and $\Theta(B^S)=1+\Theta_1 B^S+\Theta_2 B^{2S}+\ \ldots\ +\Theta_Q B^{QS}$. In SARIMA, P defines the seasonal autoregressive component of the model (SAR) and Q the seasonal moving average component of the model (SMA). S represents the seasonal period (e.g., 12 months) and D is the degree of seasonal differencing. Together, S and D account for seasonal nonstationarity in X_t through a data transformation that involves the loss of DS observations in the series.

Appendix 2

Selection of ARIMA and SARIMA models

Box-Jenkins approach ARIMA and SARIMA models are usually fitted by using a sequence of three general steps collectively known as the Box-Jenkins (BJ) method: 1) identification of the model; 2) estimation of the model; and 3) a diagnostic check of the model (Box et al., 2008). In the identification stage, a model structure $(p,d,q)\times(P,D,Q)_S$ is selected by comparisons of sample ACF and PACF with theoretical ACF/PACF profiles of AR, MA and ARMA processes. In the estimation stage, the model structure is fitted to the data and its parameters are estimated, generally by using conditional sum of squares or maximum likelihood methods. In the diagnostic check stage, the goodness-of-fit and assumptions for the model are evaluated and, if necessary, the BJ procedure is repeated until a suitable model is found. This model is then used to forecast future values (Box et al., 2008). In-depth theoretical coverage of the BJ method is given in Box et al. (2008) and extensive practical applications are provided in Pankratz (1983) and Brockwell and Davis (2002).

The model identification stage of the BJ method is widely considered its most subjective step because it relies primarily on graphical interpretations of ACF/PACF estimates obtained from a single sample. This interpretation requires substantial analytical expertise and knowledge of the time series (both of which are problematic in data-poor scenarios) and is troublesome when complex ARMA processes have generated the data (Harvey, 1989; Shumway and Stoffer, 2006). Furthermore, it can also be confounded by existing correlations among ACF/PACF estimates (Box et al., 2008). The minimum sample size generally advised for SARIMA model fitting is 50 observations (Pankratz, 1983; Chatfield, 1996b), but see Hyndman and Kostenko (2007) for an absolute lower limit. When sample size is large (e.g., $n \geq 100$), ACF/PACF estimates have lower variability and are more likely to approximate the theoretical ACF/PACF estimates of the underlying process. In such cases, less subjectivity exists in identification of the model. However, when sample size is small, the interpretation of ACF/PACF patterns becomes increasingly confounded by the large variance of the sample estimates, particularly at larger lags ($\geq n/4$) (Box et al., 2008). This variability substantially increases the subjectivity of the model identification stage of the BJ method and is the main issue to be dealt with when analyzing shorter time series.

AIC approach To circumvent the subjectivity of the identification of the model with the BJ method and to aid in the determination of the final orders of the ARMA processes a wide variety of model selection criteria have been developed (De Gooijer et al., 1985). The most frequently used are the Akaike information criteria (AIC) (Akaike, 1974) and the small-sample, bias-corrected equivalent, AIC_c (Hurvich and Tsai, 1989). Contrary to the Box-Jenkins method, AIC/AIC_c selection of a model involves the a priori estimation by maximum likelihood methods of a set of model structures (here termed the candidate set). This estimation is followed by the determination of the AIC/AIC_c values for each individual model. The model with minimum AIC/AIC_c is then selected as the model that is closest to the statistical process "generating" the data. In SARIMA models, AIC is calculated as

$$AIC = -2\ln(L) + 2r \ ,$$

where $\ln(L)$ is the log-likelihood of the model, $r=p+q+P+Q+1$, and the AIC_c is given by

$$AIC_c = -2\ln(L) + 2r + 2r(r+1)/(n-r-1) \ ,$$

where $n=N-DS-d$ is the number observations used to fit the model. AIC/AIC_c constitute objective methods to achieve model parsimony through a trade-off between the variance explained by the model and penalty terms caused by excessive model parameters. Both of them are well founded in the principles of information and likelihood theory and have been applied extensively in time series, fisheries, and ecological literature (e.g., Brockwell and Davis, 2002; Burnham and Anderson, 2002; Hanson et al., 2006). Burnham and Anderson (2002) suggest AIC_c is used when $n/r \leq 40$, which prompts the consideration of this small-sample, bias-corrected version of AIC in studies of short time series.