

# Explaining *Dinophysis cf. acuminata* abundance in Antifer (Normandy, France) using dynamic linear regression

D. Soudant<sup>1,\*</sup>, B. Beliaeff<sup>1</sup>, G. Thomas<sup>2</sup>

<sup>1</sup>IFREMER, B.P. 21105, F-44311 Nantes cedex 03, France

<sup>2</sup>INSERM U. 444, 27 rue de Chaligny, F-75571 Paris cedex 12, France

**ABSTRACT:** Classical regression analysis can be used to model time series. However, the assumption that model parameters are constant over time is not necessarily adapted to the data. In phytoplankton ecology, the relevance of time-varying parameter values has been shown using a dynamic linear regression model (DLRM). DLRMs, belonging to the class of Bayesian dynamic models, assume the existence of a non-observable time series of model parameters, which are estimated on-line, i.e. after each observation. The aim of this paper was to show how DLRM results could be used to explain variation of a time series of phytoplankton abundance. We applied DLRM to daily concentrations of *Dinophysis cf. acuminata*, determined in Antifer harbour (French coast of the English Channel), along with physical and chemical covariates (e.g. wind velocity, nutrient concentrations). A single model was built using 1989 and 1990 data, and then applied separately to each year. Equivalent static regression models were investigated for the purpose of comparison. Results showed that most of the *Dinophysis cf. acuminata* concentration variability was explained by the configuration of the sampling site, the wind regime and tide residual flow. Moreover, the relationships of these factors with the concentration of the microalga varied with time, a fact that could not be detected with static regression. Application of dynamic models to phytoplankton time series, especially in a monitoring context, is discussed.

**KEY WORDS:** Phytoplankton · *Dinophysis* · Time series · Regression · Dynamic · Bayesian

## INTRODUCTION

To investigate potential relationships between a set of covariates and some observed process, time series are commonly modelled using regression analysis. Regression constant parameters are estimated from the whole data set, assuming constant relationships over time between the dependent variable and covariates. However, these relationships may, in reality, vary over time. For example, the influence of a given covariate can be highly significant during a certain time interval and non-significant the rest of the time. Alternatively, the influence can be significant over the whole time period but subject to large variations. In the first case, the covariate parameter value will be underestimated

and thus found non-significant; in the second case, large variations will inflate the variance of the estimator and may lead to a conclusion of non-significance of the covariate. Thus, in classical (i.e. static) regression analysis, dynamic relationships between dependent and independent variables cannot be properly taken into account.

Dynamic Linear Regression Models (DLRMs) belong to the class of Bayesian dynamic models which assume time-varying relationships. The parameters are allowed to evolve with time, and thus the model is adaptable because the values of the estimated parameters and the set of significant covariates may change with time. Dynamic models have been successfully used in the social and economic fields (Pole et al. 1994, West & Harrison 1997). In previous work, we applied a DLRM to the 1988 *Dinophysis cf. acuminata* (toxic microalga) time series at Antifer (Soudant et al. 1997). High vari-

\*E-mail: dominique.soudant@ifremer.fr

abilities of parameters of physical and chemical covariates (e.g. insolation, phosphate) were detected. These results illustrated the relevance of the time-varying influence assumption in phytoplankton ecology.

The aim of this paper was to show how DLRM results can be used to point out some factors explaining, at least in part, the evolution of *Dinophysis cf. acuminata* concentrations at Antifer using the 1989 and 1990 time series. Assuming that the same processes determined concentrations of the toxic microalga during both years, a single model was built and applied separately to each data set. Particular attention was given to the adaptability of dynamic models, which allow changes in the set of significant covariates. Static regressions were performed to draw comparisons with DLRM results. Lastly, advantages of Bayesian dynamic models and their extensions are discussed in a monitoring context.

**METHODS**

Only the general principles of DLRMs are described hereafter. Readers interested in the mathematical elaboration may refer to the appendix and to more specialized papers (West et al. 1985, Pole et al. 1994, West & Harrison 1997).

Let  $Y_t$ ,  $t = 1, 2, \dots$ , denote the dependent variable at time  $t$ , and  $X_t = \{X_{1,t}, X_{2,t}, \dots, X_{n,t}\}$ ,  $t = 1, 2, \dots$ , a set of  $n$  independent variables, or covariates, measured concomitantly. In the 'static' linear regression model, the dependent variable is related to covariates by assuming

$$Y_t = \theta_0 + \sum_{i=1}^n \theta_i X_{i,t} + \varepsilon_t$$

where  $\theta_0$  is the intercept,  $\theta_i$  is the parameter of the  $i$ th covariate and  $\varepsilon_t$ , the so-called noise or error term, is a random component, with  $\varepsilon_t$ ,  $t = 1, 2, \dots$ , independently identically distributed in the normal distribution with mean 0 and variance  $V$ . A DLRM assumes a time-varying relationship, by allowing covariate parameters to vary with time. Let  $\theta_{i,t}$  denote the parameter of the  $i$ th covariate at time  $t$ . In a DLRM, the regression equation has the form

$$Y_t = \theta_{0,t} + \sum_{i=1}^n \theta_{i,t} X_{i,t} + \varepsilon_t \tag{1}$$

where  $\theta_{0,t}$  is the dynamic level, i.e. a time-varying intercept. Eq. (1) is the observation equation. Let  $\theta'_t = (\theta_{0,t}, \theta_{1,t}, \dots, \theta_{n,t})$  be the parameter vector. The evolution in time of parameters is modelled as

$$\theta_t = \theta_{t-1} + \omega_t \tag{2}$$

where  $\omega_t$  is an error term with mean 0 and variance  $W_t$ . Eq. (2) is called the evolution equation. As  $\omega_t$  is a

Table 1. Simulated data for sequential estimation procedure example

Time:	$t-4$	$t-3$	$t-2$	$t-1$	$t$	$t+1$	$t+2$	$t+3$
$X_t$	0	1	2	3	3	2	1	0
$\theta_t$	1	1	1	1	0.5	0.5	0.5	0.5
$Y_t$	0.00	0.89	2.11	2.91	1.48	0.92	0.39	0.05

random variable, Eq. (2) shows that the parameter vector  $\theta_t$  is itself a random variable.

The estimated values  $\hat{Y}_t$  and  $\hat{\theta}_t$  are the respective means of the estimated distributions of the random variables  $Y_t$  and  $\theta_t$ . The parameters of these distributions are estimated sequentially. The following simple artificial example presents the sequential estimation procedure (Fig. 1). Observations were generated with a single covariate,  $X_t$ , and without a dynamic intercept as  $Y_t = \theta_t X_t + \varepsilon_t$ , where  $\varepsilon_t$  is an error term (Table 1). The values of  $X_t$  and  $\theta_t$  were chosen and  $\varepsilon_t$  was simulated in the normal distribution with mean 0 and variance 1. The observation equation of the model was  $Y_t = \theta_t X_t + \varepsilon_t$ , i.e. the same as that used to generate the data. The evolution equation was  $\theta_t = \theta_{t-1} + \omega_t$ , where  $\omega_t$  is an error term with mean 0 and variance  $W_t$ . The proce-

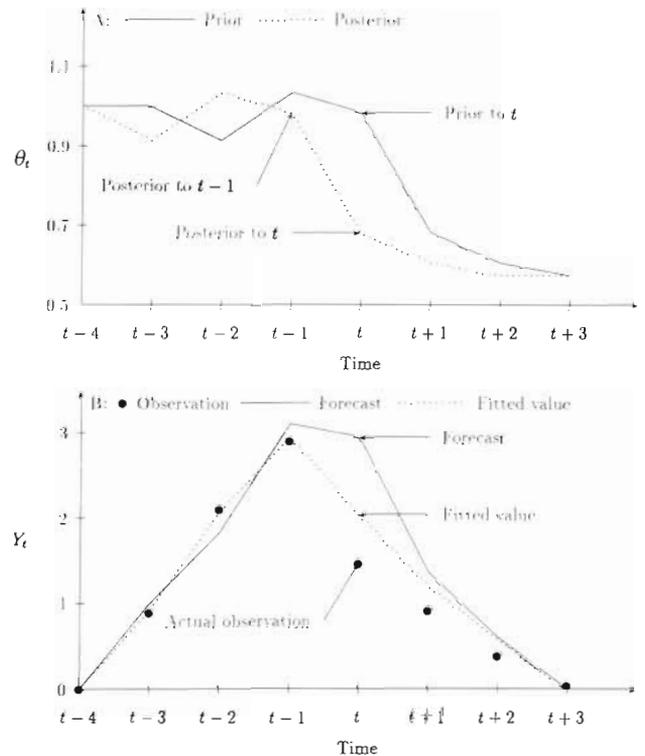


Fig. 1. Sample DLRM: (A) prior and posterior estimated values of covariate parameter and (B) observation, DLRM forecast, and on-line fitted values

ture estimates a succession of distributions prior to, and posterior to, the current observation. It begins at  $t - 1$ , with the distribution of  $\theta_{t-1}$  posterior to the observation of  $Y_{t-1}$ . The parameters of this distribution are computed after  $Y_{t-1}$  has been actually observed. The evolution equation, adding the random variables  $\theta_{t-1}$  and  $\omega_t$ , gives the distribution of  $\theta_t$  prior to the observation of  $Y_t$ . The mean and the variance of this distribution are equal to the mean and the variance of the distribution of  $\theta_{t-1}$  posterior to  $Y_{t-1}$ , with  $W_t$  added to the variance, as  $\omega_t$  is centered on 0 and has variance  $W_t$ . Thus, the evolution equation implies that  $\hat{\theta}_t$  prior to time  $t$  is equal to  $\hat{\theta}_{t-1}$  posterior to time  $t - 1$  (Fig. 1A), but with increased uncertainty. The observation equation gives the observation distribution prior to time  $t$ , that is before  $Y_t$  is actually observed. This is the forecast distribution giving the forecast estimate (Fig. 1B). Then, the actual value of  $Y_t$  is observed. With this new information, the distribution of  $\theta_t$  prior to  $Y_t$  is updated giving the distribution of  $\theta_t$  posterior to  $Y_t$ . In our example, this update induced a decrease from 0.98 to 0.68 of the estimated value of the covariate parameter (Fig. 1A), as the actual value was decreased from 1 to 0.5 (Table 1). The estimation of the distribution of  $\theta_t$  posterior to the observation of  $Y_t$  allows a new iteration of the sequential procedure. Beside the procedure, parameters of the posterior forecast distribution are computed using Eq. (1). This distribution gives the on-line fitted value, i.e. the value fitted immediately after  $Y_t$  has been observed (Fig. 1B).

## APPLICATION TO THE ANTIFER TIME SERIES

### Data collection

*Dinophysis cf. acuminata* (hereafter referred to as *Dinophysis*) is a microalga producing diarrhetic shellfish poisoning. Despite many studies, some features of its biology and ecology remain largely unknown (Delmas et al. 1993, Sampayo 1993, Berland et al. 1995a, b, Maestrini et al. 1996). As high *Dinophysis* concentrations were observed previously in Antifer harbour (France), sea water was sampled to study the ecological conditions of occurrence. Daily samples were taken at high tide at 1 m depth at the end of the petroleum wharf (Fig. 2) from 1 July to 13 September 1989 and from 1 June to 11 September 1990. The measurements carried out by the municipal laboratory of Le Havre were *Dinophysis* concentration (cells per 10 ml) (Utermöhl 1958), salinity (Beckman induction salinometer), temperature ( $^{\circ}\text{C}$ ) (Ponselle sonde), nitrate and phosphate concentration ( $\mu\text{mol l}^{-1}$ ) (Technicon autoanalyser). Insolation ( $\text{h d}^{-1}$ ), rainfall ( $\text{mm d}^{-1}$ ), wind direction and speed ( $\text{m s}^{-1}$ ) and Seine flow ( $\text{m}^3 \text{s}^{-1}$ )

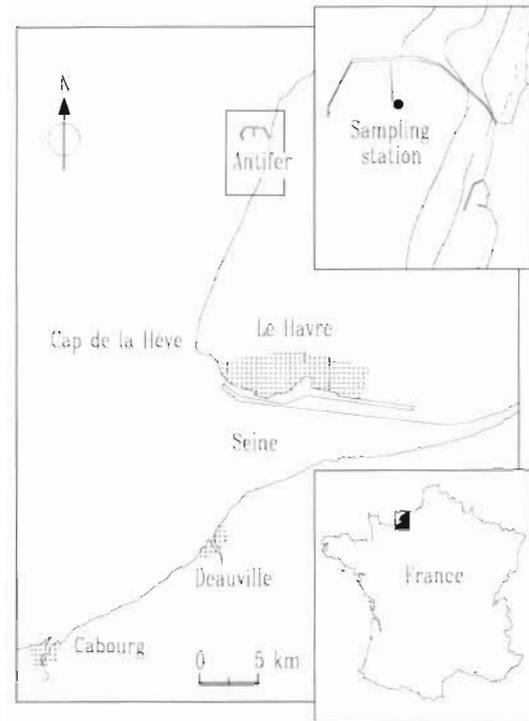


Fig. 2. Sampling station, Antifer, France

were obtained from the Le Havre weather station (Cap de la Hève, cf. Fig. 2). Tide coefficients at Le Havre harbour were drawn from tide tables published each year by the French hydrographic and oceanographic navy service (SHOM).

A 'South-West wind' covariate was computed as the daily mean of 8 determinations per day of the variable  $[-\alpha_{i,t} \cos(\beta_{i,t} - \pi/4)]$ , where  $\alpha_{i,t}$  and  $\beta_{i,t}$  are, respectively, the speed and direction of the wind on the  $i$ th determination of Day  $t$ . This gives a continuous decrease from South-West to North-East, with a zero value for North-West and South-East winds. Finally, the variables were standardized to zero mean, to separate clearly the covariate effects from the dynamic intercept, and unit variance, to allow comparisons between years and between covariate effects. It followed from this standardization that the dynamic intercept of the model at time  $t$  was the local mean of the dependent variable. Furthermore, the estimated regression parameters were adimensional.

### Dynamic linear regression model

The dependent variable was  $Y_t = \log(Z_t + 1)$ , where  $Z_t$  was *Dinophysis* concentration on Day  $t$ . For each year, a model was fitted. The model included a dynamic intercept, and covariates were selected one by one

from the set of available variables (see previous section). At each step, the variable which induced the largest model likelihood with a significant gain in likelihood, as assessed using the likelihood ratio test (Kendall & Stuart 1977), was entered in the model. For 1989, only the 'South-West wind' covariate was selected, while 'Salinity' and 'Tide coefficient' were selected for the 1990 series. Fig. 3 shows these variables. We used these 3 variables to explain the process underlying the evolution of *Dinophysis* concentration during the 2 years. Thus, the final model, common to 1989 and 1990, had the following observation equation at time  $t$ ,  $t = 1, 2, \dots$ ,

$$Y_t = \theta_{0,t} + \theta_{SW,t} SW_t + \theta_{S,t} S_t + \theta_{T,t} T_t + \varepsilon_t$$

where  $\theta_{0,t}$  represents the dynamic intercept,  $SW_t$  'South-West wind',  $S_t$  'Salinity',  $T_t$  'Tide coefficient' and  $\varepsilon_t$  is an error term. We decided to present results using covariate effects, that is the variable value (i.e.  $X_t$ ) multiplied by the estimated regression parameter (i.e.  $\hat{\theta}_{X,t}$ ). Confidence intervals at the  $\alpha = 0.05$  level were used to test the nullity of the effects: when 0 was between the 2 limits, the effects were considered non-significantly different from 0. Finally, static versions of this model were fitted to data from both years in order to draw comparisons with DLRM results.

## RESULTS

There was a succession of peaks of increasing magnitude in *Dinophysis* concentration in 1989 and 1990 (Fig. 4). On-line fitted values were similar to observed values for the 2 years. Figs. 5 & 6 show dynamic intercepts and effects of covariates. For both years, effects were not always significant. At the beginning of the series, and especially in 1990, 95% confidence intervals of effects were initially large and then decreased rapidly. This decrease in uncertainty with the accumulation of observations and the alternation of time periods when the effects were significant and non-significant illustrated the adaptability of dynamic models.

In 1989 (Fig. 5), the dynamic intercept was significantly different from zero from Day 8 to 21 and from Day 57 to 64, and 'South-West wind' from Day 29 to 34 and from Day 37 to the end of the time series. As 'Salinity' was only significant the last day of the time series and 'Tide coefficient' was never significant, dynamic intercept and 'South-West wind' effects mainly contributed to the on-line fitted values of the concentration of *Dinophysis*. The dynamic intercept showed a local low concentration of *Dinophysis* in the first interval (Days 8 to 21) and a high concentration in the second interval (Days 57 to 64).  $\hat{\theta}_{SW,t}$  was always positive. Positive effects corresponded to South-West winds and negative effects to

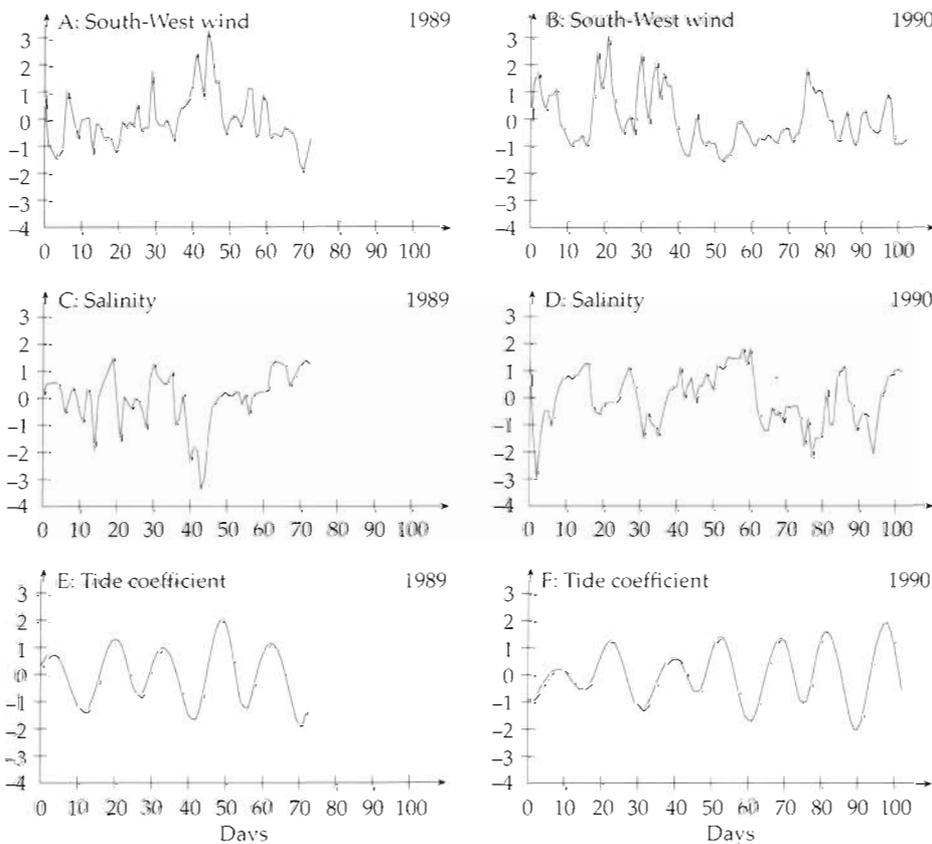


Fig. 3. Evolution with time of the standardized variables selected as covariates in the model for 1989 and 1990. Day 0 is 1 July for 1989 and 1 June for 1990

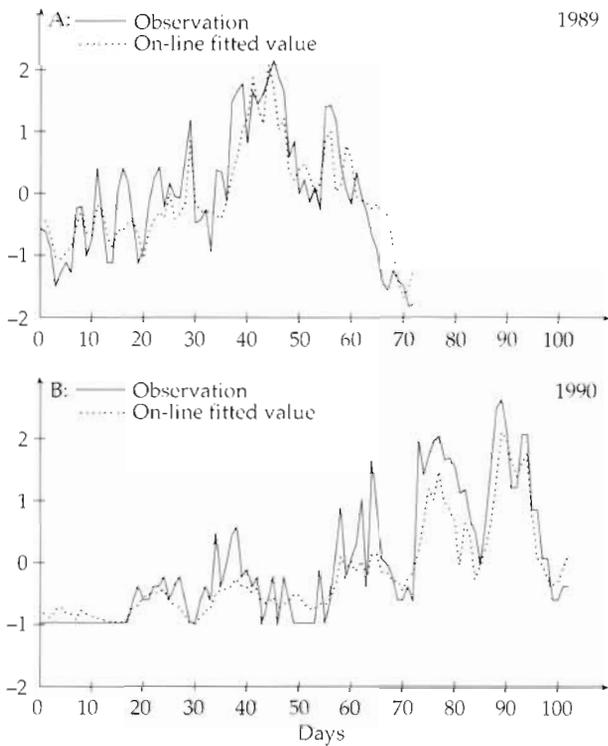


Fig. 4. Standardized observations of *Dinophysis* concentrations at Antifer and their on-line fitted values. Day 0 is 1 July for 1989 and 1 June for 1990

North-East winds. The percentage  $R^2$  (Draper & Smith 1966) of variation of *Dinophysis* concentration explained by the DLRM was 73.54% and that of static regression was 55.29% (Table 2). In the static regression, the intercept and the estimated values of the parameters of 'Tide coefficient' and 'Salinity' were not significantly different from zero.  $\hat{\theta}_{SW,t}$  was highly significant and positive.

In 1990 (Fig. 6), the dynamic intercept was significantly different from zero and negative from Day 6 to 63 and significant and positive from Day 80 to the end of the time series. 'South-West wind' effects were significant from Day 58 to 60 and at Day 86. When the effects of this covariate were significant, its estimated parameter was positive. For 'Salinity', effects were significant from Day 58 to 60 with a positive estimated parameter and from Day 65 to the end of the time series with a negative estimated parameter. Lastly, 'Tide coefficient' effects were significant at Day 46 and during 4 intervals: Days 23 to 31, 41 to 44, 61 to 80 and Day 88 to the end of the time series.  $\hat{\theta}_{T,t}$  was positive from the beginning of the time series to Day 56, and then negative to the end of the time series. Dynamic level, 'Salinity' and 'Tide coefficient' explained most of the evolution of *Dinophysis* concentrations. When  $\hat{\theta}_{S,t}$  and  $\hat{\theta}_{T,t}$  were negative, negative values (respectively positive) of 'Salinity' and 'Tide coefficient' corresponded to positive (respectively negative) effects. The percentage of variation explained by the DLRM was  $R^2 = 79.57\%$ . The  $R^2$  of the static regression was 23.79%. In the static regression, the intercept and the estimated parameter of 'Tide coefficient' were non-significantly different from zero. 'South-West wind' and 'Salinity' parameter estimations were significant and negative.

## DISCUSSION

Covariate effects suggested different scenarios to explain *Dinophysis* concentration dynamics. In 1989, the geographical situation of the sampling site (Fig. 2), the location of phytoplankton maximum concentrations in the Seine plume (Ménésquien et al. 1995) and

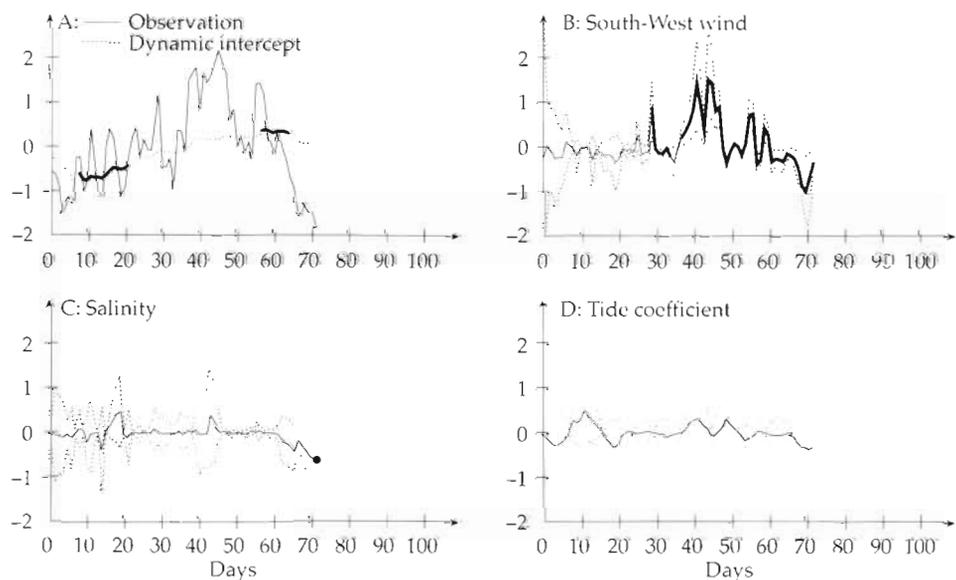


Fig. 5. DLRM results for 1989. (A) Standardized observations of *Dinophysis* concentrations at Antifer and their dynamic local means and (B, C, and D) effects and 95% confidence limits of covariates. Bold lines denote effects significantly different from 0. Day 0 is 1 July 1989. Note that effects are adimensional because of covariate standardization

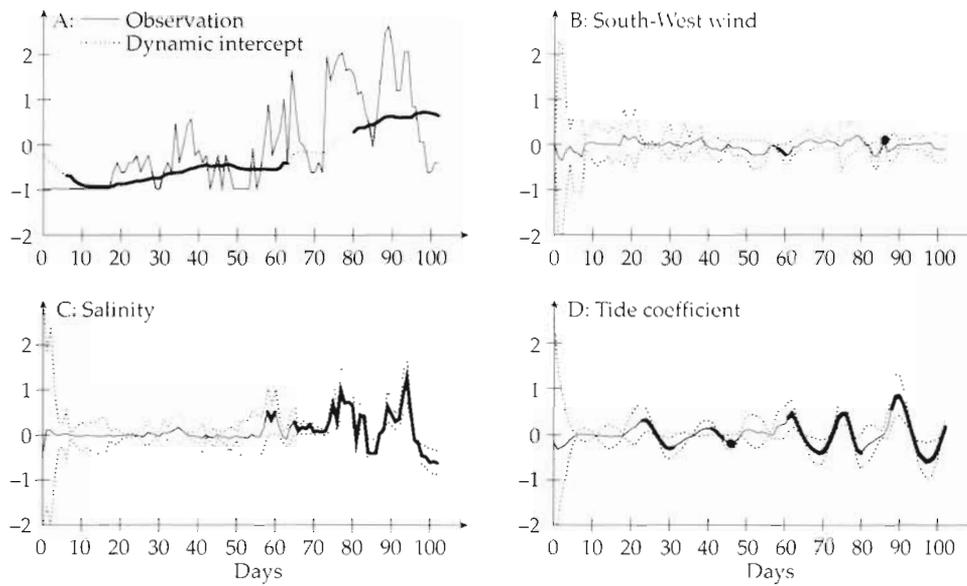


Fig. 6. DLRM results for 1990. (A) Standardized observations of *Dinophysis* concentrations at Antifer and their dynamic local means and (B, C, and D) effects and 95% confidence limits of covariates. Bold lines denote effects significantly different from 0. Day 0 is 1 June 1990. Note that effects are adimensional because of covariates standardization

hydrodynamical studies of the Seine bay (e.g. Salomon & Breton 1993) suggested that South-West wind induced *Dinophysis* cell accumulation along the coast, particularly in Antifer harbour. Conversely, North-East wind could provoke cell dispersion. Such transportation phenomena induced by wind has already been observed in the Seine bay (Lagadeuc 1992, Thiébaud et al. 1994). By definition, the 'South-West wind' effect depended on the 'South-West wind' covariate value. 'South-West wind' effect varied also with *Dinophysis* concentration in the water mass subject to the accumulation/dispersion phenomena. As this concentration varied with time, the relationship between 'South-West wind' and the microalgal concentration in Antifer was time-varying. Only the covariates 'Salinity' and 'Tide coefficient' were significant in 1990. Significant negative values of the 'Salinity' estimated parameter suggested that lower surface salinity was accompanied by

higher *Dinophysis* concentrations. This result was consistent with the association of *Dinophysis* occurrence with persistent salinity stratification (e.g. Delmas et al. 1992). The establishment of stratification is favoured by small tide coefficients. Greater tide coefficients may provoke water mixing and consequently a decrease of *Dinophysis* concentrations by dilution. A hydrodynamical study has shown that the configuration of Antifer harbour modifies the circulation of water masses (Monbet 1975), so that greater tide coefficients induced a departure of water masses to the North and small coefficients a 'capture' of water masses in the harbour. As for 'South-West wind', relationships between *Dinophysis* concentration and 'Salinity' and between *Dinophysis* concentration and 'Tide coefficient' varied over time.

Although 'Salinity', 'Tide coefficient' and 'South-West wind' seemed to be important for understanding the evolution of *Dinophysis* concentration at Antifer, some discrepancies appeared between scenarios and results. The 3 covariates were never significant concomitantly. A natural explanation for this observation was related to interdependence among these variables. For example, correlation between 'South-West wind' and 'Salinity' covariates was negative and highly significant ( $p < 10^{-7}$ ) both in 1989 and 1990. One of these 2 covariates may thus mask the influence of the other one. Such correlations seemed to be responsible for the non-significance of the 'Tide coefficient' estimated parameter in 1989 and, in the static regression results, for the visible inversion of the absolute values of estimated parameters of 'South-West wind' and 'Salinity' for 1989 and 1990 and for the change in the sign of  $\hat{\theta}_{SW,t}$  from 1989 to 1990. In 1990, results of the DLRM showed changes in the signs of the estimated

Table 2. Results of the static regression model for 1989 and 1990.  $N$  is sample size,  $R^2$  the percentage of variation explained by the regression,  $\hat{\theta}_0$  the estimated value of the intercept,  $\hat{\theta}_{SW}$  that of the 'South-West wind' covariate,  $\hat{\theta}_S$  that of 'Salinity', and  $\hat{\theta}_T$  that of 'Tide coefficient' ns: non-significantly different from zero at the  $\alpha = 0.05$  level (bilateral test); p: significance level

	1989	1990
$N$	73	103
$R^2$	55.29%	23.79%
$\hat{\theta}_0$	-0.004 (ns)	-0.026 (ns)
$\hat{\theta}_{SW}$	0.621 ( $p < 10^{-6}$ )	-0.214 ( $p = 0.039$ )
$\hat{\theta}_S$	-0.213 (ns)	-0.514 ( $p < 10^{-5}$ )
$\hat{\theta}_T$	0.125 (ns)	-0.159 (ns)

parameters of 'Salinity' and 'Tide coefficient'. From our scenarios, these parameters were expected to be negative. The positive parameter values resulted from local positive correlations between the values of *Dinophysis* concentration and 'Salinity' and between those for *Dinophysis* concentration and 'Tide coefficient'. There were only 3 days (58, 59, 60) when the estimated parameter of 'Salinity' was positive, thus we considered this event as fortuitous. South-West wind blew strongly in a chaotic way and 'Salinity' decreased during the first 2 significant intervals of 'Tide coefficient' effects (Fig. 3B, D). Then, wind probably induced the first 2 peaks of *Dinophysis* concentration, but these were more correlated with the sinusoid evolution of the tide coefficient. In the static regression,  $\hat{\theta}_T$  was not significant, like an average of the dynamic parameter  $\hat{\theta}_{T,t}$ , could have been not significant.

From these results, a general explanation for the evolution of *Dinophysis* concentration was derived as follows: South-West winds draw water masses, possibly stratified and rich in *Dinophysis*, inshore, particularly to Antifer harbour due to the configuration of the site. North-East winds may provoke dispersion of *Dinophysis* cells. Large tide coefficients may induce a decrease of *Dinophysis* concentrations as a consequence of water mass movements and/or dilution. It should be noted that the set of significant covariates is a subset of available variables. A significant serial correlation for the residuals, as the runs test (Siegel 1956) showed us at the  $\alpha = 0.05$  significance level for both years, might reflect the absence of at least one key descriptor in the model. As our analysis identified physical factors, this (these) might be biological factor(s). Our explanation illustrated the usefulness of DLRLMs as explanatory tools. Dynamic models can also be used as an on-line analysis method for time series

as, for instance, the phytoplankton time series issued from monitoring programmes. In this case, data are obtained sequentially and, although not recommended, sampling frequency might be irregular, generating time series with missing data. The sequential definition of dynamic models makes them well suited for such time series analysis. The estimation procedure can manage missing data by forecasting the value at time  $t+k$ ,  $k>1$ . Moreover, in ecology, the observational variance is often a function of the mean (Taylor 1961, Kendal 1995), and thus varies in time with the mean. If the variance-to-mean relationship is known, it can be used to specify the sequence of the observational variance. Alternatively, dynamic models can accommodate the assumption of time-varying variance. Finally, the Bayesian model approach of time series modelling can be considered as a dynamic generalization of a linear model, and thus developments of the latter (e.g. multiple linear regression) are adaptable for the former.

DLRM results gave us a more thorough understanding of *Dinophysis* concentration time series in Antifer than did static regression analysis. In particular, time-varying relationships between significant covariates and the concentration of the toxic microalga could not be assessed using static regression. Furthermore, DLRLM characteristics and extensions could make dynamic models one of the most efficient tools for analyzing time series data, and especially those of monitoring programmes.

*Acknowledgements.* We express our gratitude to P. Gros for commenting on this manuscript. We thank P. Lassus for providing data and P. Gentien for helpful discussions. L. Giboire is acknowledged for figures. The comments of 3 referees led us to improve our manuscript.

## Appendix

The model used is a univariate DLRLM with constant and unknown variance  $V$ . Let  $Y_t$ ,  $t = 1, 2, \dots$ , denote a time series and  $\mathbf{X}'_t = (X_{1,t} X_{2,t} \dots X_{n,t})$ ,  $t = 1, 2, \dots$ , a time dependent vector of variables. The observation is governed by the so-called 'observation equation',

$$Y_t = \mathbf{F}'_t \boldsymbol{\theta}_t + \varepsilon_t \quad (\text{A1})$$

where  $\mathbf{F}'_t = (1 \ \mathbf{X}'_t)$  is the vector of regressors,  $\boldsymbol{\theta}'_t = (\theta_{0,t} \theta_{1,t} \theta_{2,t} \dots \theta_{n,t})$  is a vector of time dependent parameters and  $\varepsilon_t$  are observational errors, independently identically distributed in the normal distribution  $N(0, V)$ . The unknown reciprocal variance or precision is denoted by  $\phi = V^{-1}$ . At  $t-1$ ,  $\phi$  is distributed in the Gamma distribution  $G[n_{t-1}/2, d_{t-1}/2]$ . The parameter vector changes through time according to the evolution equation

$$\boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \boldsymbol{\omega}_t \sim T_{n_{t-1}}[\mathbf{0}, \mathbf{W}_t] \quad (\text{A2})$$

where  $\boldsymbol{\omega}_t$  is the evolutionary error. (The notation  $\sim$  is used here and elsewhere to denote 'distributed as'.) Then we define the initial distributions  $(\phi|D_0) \sim G[n_0/2, d_0/2]$  and  $(\boldsymbol{\theta}_0|D_0) \sim T_{n_0}(\mathbf{m}_0, \mathbf{C}_0)$ , where  $n_0$ ,  $d_0$ ,  $\mathbf{m}_0$  and  $\mathbf{C}_0$  are fixed.  $D_0$  is the initial information set, representing all the available relevant information used to specify the model before the first observation, and including all vectors of regressors  $\mathbf{F}_t$ . The error sequences  $\varepsilon_t$  and  $\boldsymbol{\omega}_t$  are independent, mutually independent, and independent of  $(\boldsymbol{\theta}_0|D_0)$ . Lastly,  $\mathbf{W}_t$  is specified as  $\mathbf{W}_t = \mathbf{C}_{t-1}(1 - \delta)/\delta$ , where  $\delta \in ]0, 1[$ .  $\delta$  is the so-called discount factor and controls the model adaptability: if  $\delta$  is near 0 then the model adaptability is high and if  $\delta$  is near 1 then the model can only change slowly.

(Continued on next page)

## Appendix (continued)

The sequential estimation procedure starts at  $t - 1$ . Let us define the information set at time  $t$  as  $D_t = \{D_{t-1}, Y_t\}$ .  $(\theta_{t-1}|D_{t-1})$  is distributed as  $T_{n_{t-1}}[\mathbf{m}_{t-1}, \mathbf{C}_{t-1}]$  and  $(\phi|D_{t-1})$  as  $G[n_{t-1}/2, d_{t-1}/2]$ . The estimation steps are the following:

**Prior.**  $(\theta_t|D_{t-1}) \sim T_{n_t}[\mathbf{a}_t, \mathbf{R}_t]$ , where  $\mathbf{a}_t = \mathbf{m}_{t-1}$  and  $\mathbf{R}_t = \mathbf{C}_{t-1} + \mathbf{W}_t$ .

**Prediction.**  $(Y_t|D_{t-1}) \sim T_{n_t}[f_t, Q_t]$ , where  $f_t = \mathbf{F}_t \mathbf{a}_t$  and  $Q_t = \mathbf{F}_t \mathbf{R}_t \mathbf{F}_t + S_{t-1}$  with  $S_{t-1} = d_{t-1}/n_{t-1}$ .

**Posterior.**  $(\theta_t|D_t) \sim T_{n_t}[\mathbf{m}_t, \mathbf{C}_t]$ , where  $\mathbf{m}_t = \mathbf{a}_t + \mathbf{A}_t e_t$  and  $\mathbf{C}_t = (S_t/S_{t-1})(\mathbf{R}_t - \mathbf{A}_t \mathbf{A}_t' Q_t^{-1})$ , with  $\mathbf{A}_t = \mathbf{R}_t \mathbf{F}_t Q_t^{-1}$  and  $e_t = Y_t - f_t$ .  $(\phi|D_t)$  is distributed as  $G[n_t/2, d_t/2]$ , where  $n_t = n_{t-1} + 1$  and  $d_t = d_{t-1} + S_{t-1} e_t^2 / Q_t$ .

The fitted distribution  $(Y_t|D_t)$  is distributed as  $T_{n_t}[g_t, P_t]$ , where  $g_t = \mathbf{F}_t \mathbf{m}_t$  and  $P_t = \mathbf{F}_t \mathbf{C}_t \mathbf{F}_t + S_t$ .

Parameters of the initial distributions of  $(\phi|D_0)$  and  $(\theta_0|D_0)$  and the discount factor  $\delta$  are included in the set  $D_0$  and fixed by the model user. For our model applied to the *Dinophysis* concentration time series (cf. 'Dynamic linear regression model'), the values were  $\mathbf{m}_0 = \mathbf{0}$ ,  $\mathbf{C}_0$  equal to the identity matrix,  $n_0 = d_0 = 1$  and  $\delta = 0.95$  in 1989 and 1990.

Computer programs used to perform DLRM analysis were developed in C on a SUN station and are available from the first author.

## LITERATURE CITED

- Berland BR, Maestrini SY, Grzebyk D (1995a) Observations on possible life cycle stages of the dinoflagellates *Dinophysis* cf. *acuminata*, *Dinophysis acuta*, *Dinophysis pavillardii*. Aquat Microb Ecol 9:183–189
- Berland BR, Maestrini SY, Grzebyk D, Thomas P (1995b) Recent aspects of nutrition in the dinoflagellate *Dinophysis* cf. *acuminata*. Aquat Microb Ecol 9:191–198
- Delmas D, Herbland A, Maestrini SY (1992) Environmental conditions which lead to increase in cell density of the toxic dinoflagellates *Dinophysis* spp. in nutrient-rich and nutrient-poor waters of the French Atlantic coast. Mar Ecol Prog Ser 89:53–61
- Delmas D, Herbland A, Maestrini SY (1993) Do *Dinophysis* spp. come from the open sea along the French Atlantic coast? In: Smayda TJ, Shimizu Y (eds) Toxic phytoplankton blooms in the sea. Elsevier Science, Amsterdam, p 489–494
- Draper NR, Smith H (1966) Applied regression analysis. Wiley series in probability and mathematical statistics. John Wiley and Sons, New York
- Kendal WS (1995) A probabilistic model for the variance to mean power law in ecology. Ecol Model 80:293–297
- Kendall M, Stuart A (1977) The advanced theory of statistics, Vol 2. Charles Griffin and Company, London
- Lagadeuc Y (1992) Transport larvaire en Manche. Exemple de *Pectinaria koreni* (Malmgren), annélide polychète en baie de Seine. Oceanol Acta 15:383–395
- Maestrini SY, Berland BR, Carlsson P, Granéli E, Pastoureaud A (1996) Recent advances in the biology of the toxic dinoflagellate genus *Dinophysis*: the enigma continues. In: Yasumoto T, Oshima Y, Fukuyo Y (eds) Harmful and toxic algal blooms. IOC of UNESCO, Paris, p 397–400
- Ménesguen A, Guillaud JF, Aminot A, Hoch T (1995) Modeling the eutrophication process in a river plume: the Seine case study. Ophelia 42:205–225
- Monbet Y (1975) Les incidences écologiques de la construction du terminal d'Antifer. Rapport CNEXO-unité littoral, IFREMER, Brest
- Pole A, West M, Harrison PJ (1994) Applied Bayesian forecasting and time series analysis. Chapman and Hall, New York
- Salomon JC, Breton M (1993) An atlas of long-term currents in the Channel. Oceanol Acta 16:439–448
- Sampayo MA de M (1993) Trying to cultivate *Dinophysis* spp. In: Smayda TJ, Shimizu Y (eds) Toxic phytoplankton blooms in the sea. Elsevier Science, Amsterdam, p 807–810
- Siegel S (1956) Nonparametric statistics for the behavioral sciences. McGraw-Hill series in psychology. McGraw-Hill, New York
- Soudant D, Beliaeff B, Thomas G (1997) Dynamic linear Bayesian models in phytoplankton ecology. Ecol Model 99:161–169
- Taylor LR (1961) Aggregation, variance and the mean. Nature 189:732–735
- Thiébaud E, Dauvin JC, Lagadeuc Y (1994) Horizontal distribution and retention of *Owenia fusiformis* larvae (Annelida: Polychaeta) in the bay of Seine. J Mar Biol Assoc UK 74: 129–142
- Utermöhl H (1958) Zur Vervollkommnung der quantitativen Phytoplankton-Methodik. Int Ver Theor Angew Limnol Verh 9:1–38
- West M, Harrison PJ (1997) Bayesian forecasting and dynamic models, 2nd edn. Springer series in statistics. Springer-Verlag, New York
- West M, Harrison PJ, Migon HS (1985) Dynamic generalized linear models and Bayesian forecasting (with discussion). J Am Stat Assoc 80:73–97

This article was submitted to the editor

Manuscript received: February 20, 1997

Revised version accepted: July 2, 1997