A Long-Lead Forecast Model for the Prediction of Shelf Water Oscillations along the Caribbean Coast of the Island of Puerto Rico

Marvi Teixeira
E.E. Dept. Polytechnic University of PR
Hato Rey, PR
mteixeir@caribe.net

Jorge Capella
Marine Sciences Department
University of Puerto Rico
Mayaguez, PR

Abstract - **It is a well-established fact that shelf waters along the south coast of Puerto Rico oscillate at the seiche mode period of approximately one-hour. Under the right conditions the amplitude of these events equals that of the diurnal tide. At several locations worldwide this type of coastal water oscillation has extremely large amplitude, causes considerable damages and is also the focus of current research. Even though the precise forcing of this phenomenon at this particular location is not yet completely understood, it is believed that packets of solitary waves impinging against the steep slope could be forcing the shelf waters into motion. A long-lead forecast model, which is introduced in this paper, is being made available for the prediction of these coastal waters oscillations. The model will facilitate the establishment of an adequate data collection schedule for the study of these events.**

The one-hour oscillations of the shelf waters are conspicuously modulated and render a narrow band signal within the tide spectral record. Different techniques were used to identify the main spectral components of the signal envelope. These components were found to be nearly uncorrelated and formed the basis for the formulation of a long lead seasonal forecast model. In order for the predictor to be an effective forecast tool the phase of each harmonic constituent was carefully calibrated in the time domain with respect to the different lunar cycles. The weights of the model were assigned using least squares. Stepwise regression and principal component analysis were used to manage the inclusion of constituents into the predictor. Although the main model is linear in nature the extreme excursions in amplitude were tackled by the formulation of a non-linear predictor. In order to increase the resolution, which is about five days, quadratic models that involved the squaring and multiplication of the different constituent's terms were also explored. A hindcast performed over a two-year long period validated the model, which was originally developed based on the study of a different tide record. Notwithstanding that this type of coastal water oscillations are considered quite unpredictable, in a continuous six-month period the predictor accounted for up to 75% of the low frequency variability of the seiche.

I. Introduction

Seiches are coastal water oscillations in the form of standing waves in bays, harbors, and shelves. The period of this motion depends on the geometry of the body of water. Usually small in amplitude, compared with the tide, they can attain heights that match or exceed the local tidal range. The relevance of these oscillations to the coastal marine environment and the benefits of predicting them are well established in the literature [1].

Seiches of longer period in bays tend to cause strong currents at the entrances of harbor basins. The shorter period harbor seiches could dangerously affect moored ships by inducing strong motions of the vessels [2]. In Nagasaki, Japan, a conspicuos 30-minute harbor seiche is called "abiki" and the 10-minute seiche at the Ciutadella Harbor on Menorca Island is known as "rissaga". On past occasions the "rissaga", which can have amplitudes of up to two meters, caused great damage to the local fishing fleet [1]. In the case of shelf seiches, the associated currents could have an effect in the dispersion of eggs and larvae of multiple marine organisms.

Large-amplitude seiches with a period of about 50 minutes occur along the south coast of Puerto Rico. It was found that large-amplitude seiches follow a fortnightly distribution, with maximum occurring six to seven days after new or full moon, that is, when the local semidiurnal tides are at their minimum [3]. These peaks in activity are particularly large after a new or full moon at perigee. There was also evidence suggesting that seiche activity could be inversely related to easterly wind stress and depth of the surface mixed layer.

At the south coast of Puerto Rico semidiurnal tides are weak and inversely related to maximum amplitude seiches, thus, it was difficult to relate this large-amplitude seiche distribution pattern to a local tidal forcing effect. It was therefore argued that the seiches were excited by deep-sea internal solitary waves, generated by large barotropic tides in the southeastern Caribbean, when impinging the steep shelf slope at the Caribbean coast of Puerto Rico [3]. This hypothesis was supported by the fact that large seiche activity in Puerto Rico appears to be preceded by large-range tides along the southeastern Caribbean. Furthermore, there is evidence of internal wave formation near Venezuela and close to Aves Ridge. To complement this theory the theoretical problem of continental shelf response to forcing by deep-sea internal waves was examined, [4] [5].

While the arrival of solitary waves from the southeastern Caribbean cannot be discarded, the extent to which this process could occur should be reexamined. This new study was conducted using 4.4 years of data acquired from the NOAA maintained tide station at Magueyes Island, Puerto

Rico. A thorough analysis of the phenomenon resulted in the development of a seasonal long-lead forecast model that suggests the existence of alternative forcing mechanisms as well as provides a time framework for suitable data collection schedules. Furthermore, the newly gained insight should be helpful in other regions of the world where these oscillations are known to occur.

After a description of the physical setting in which the phenomenon takes place the paper starts reviewing the seiche spectral profile followed by an analysis of the distribution of seiche activity along the different lunar cycles. This study uncovered previously unknown distribution patterns, which became the basis for the formulation of a long-lead forecast model. The spectral analysis of the seiche envelope revealed periodicities that are consistent with the cyclical distribution patterns observed along the different lunar cycles. The distribution profiles obtained based on the study of the first two years of data are then used to develop several harmonic constituents that form the basis for the formulation of a seasonal forecast model. The correlation coefficient between these harmonic constituents is very low indicating that they are almost uncorrelated. The weights for this model are determined using a least square minimization procedure. Non-linear and quadratic models are then briefly explored and discussed. The introduction of further harmonic constituents into the prediction model is also considered. A successful hindcast, performed along the remaining 2.4 years of recorded activity, validates this model. The maximum resolution of the predictor is about 5 days. On a particular 6-month interval the model accounts for up to 75% of the low frequency variability in seiche activity. Finally, possible mechanisms that could force the coastal waters into seiching are discussed. Directions for future research are also provided.

II. The Physical Settings

The area under study is the southwestern coast of Puerto Rico, which is known by its narrow shelf and steep slope (Fig. 1). The mean water depth at the shelf break is about 20m. Shelf width is in the range of 3 to 10 km. Depths of 200 m at 0.5 km and 1000 m at 5 km of the shelf edge are usually found. A schematic profile of the shelf and slope region, offshore of Magueyes Island is shown in Fig. 2, [3].

Merian's formula for open basins [6] gives their natural period of oscillations.

$$T = \frac{4L}{(gh)^{\frac{1}{2}}} \qquad (1)$$

Where L is the shelf width, g the gravitational constant and h the shelf depth. For the geometry in Fig. 2, $L = 10\ km$ and $h = 18\ m$, this formula gives $T \cong 50\ min$.

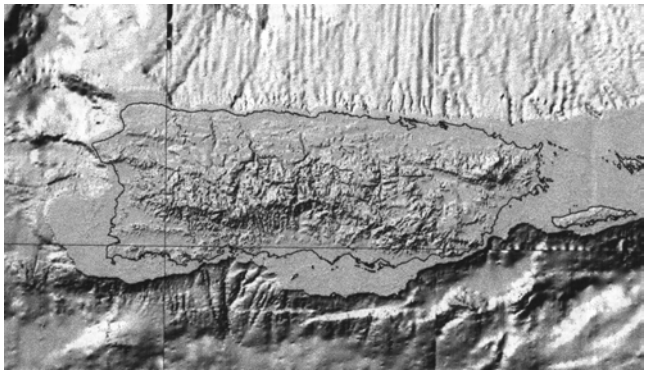

Fig.1. The south coast of Puerto Rico. The steep shelf slope can be clearly seen in this picture. The cross-hairs are at 67° west longitude and 18° north latitude, the island lies between 65° 35' and 67° 15' west longitude.

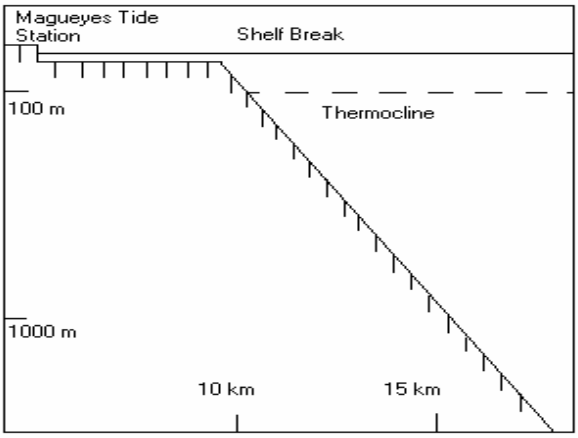

Fig. 2. Shelf and slope south of Magueyes island, Puerto Rico. The Magueyes tide station is at 67° 03' west longitude and 17° 58' north latitude.

Coastal water oscillations are established in the form of a standing wave with a node at the shelf break, Fig. 3. The near one-hour period of these oscillations is related to the geometry of the basin. Stong sheiches have amplitudes that equal that of the local tide.

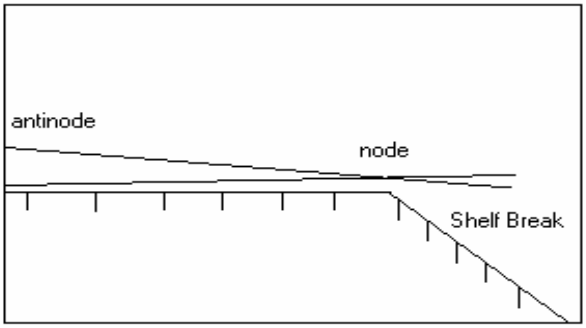

Fig. 3. Coastal water oscillations at the south coast of Puerto Rico.

III. THE SEICHE SPECTRAL PROFILE

The power density spectrum of the local tide record, obtained using Thomson's high-resolution multitaper method [7], is depicted in Fig. 4. The prominent 1.2 cph spectral peak corresponds to the seiches.

The seiche is a narrow band signal at around 1.2 cph (Fig. 4). Its spectral signature has a large degree of symmetry, something consistent with being amplitude modulated. Some of the observed spectral components, however, probably come from oscillations at slightly different frequencies, which is to say that the seiche might be frequency modulated as well. This could be related to changes in water depth or damping by turbulence [8] among other causes. The spectral signature appears as a bimodal peak. Its left side, centered at around 1.1378 cph, has a profile consistent with an amplitude modulation scheme. Amplitude modulation reveals itself, within the power spectrum, as symmetrical components around a central frequency. The inverse of the frequency difference between the symmetrical peaks and their center frequency gives the modulation period. In fact, zooming in at the center frequency reveals many symmetrical components compatible with the existence of semidiurnal, diurnal, fortnightly, and monthly amplitude modulation patterns (Fig. 5).

A. The Seiche Envelope

The best way to monitor the low frequency periodicities affecting the seiche amplitude is to look at its envelope. This can be achieved shifting the passband spectrum (Fig. 5) to the origin. The result of this demodulation is a baseband spectrum that renders the signal envelope (Fig. 6a).

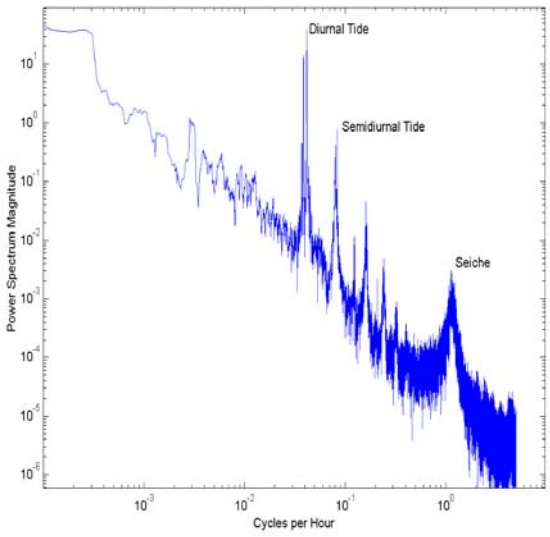

Fig. 4. Power Density Spectrum of the local tide record from October 1993 to October 1995. Sampling period was 6 minutes. The seiche spectral peak is prominent at about 1.2 cph.

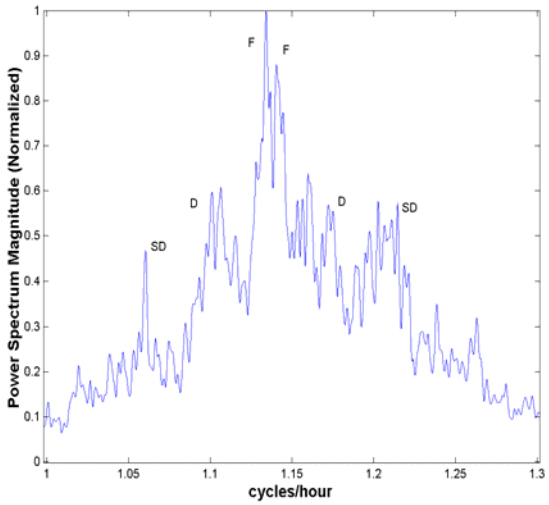

Fig. 5. Power spectral density. The symmetrical arrangement of spectral peaks around a center frequency suggests amplitude modulation by fortnightly (F), diurnal (D), and semidiurnal (SD) periods. A 8192-point window was used with an overlap of 4096 points. Sampling period was 6 minutes. A two-year long record was used.

B. The Seiche Envelope Spectrum

To uncover the periods corresponding to the low frequency events that are amplitude modulating the seiche, the spectrum of its envelope needs to be estimated. Using the Welch's averaged periodogram and under the application of various window lengths, semidiurnal, nine to 5.5 day, fortnightly and monthly periods were observed. The 95% confidence interval, regrettably, was about the size of the largest peak. Seasonal modulation was revealed when larger windows were used. Multitaper spectral estimation rendered a spectrum with better resolution (Fig. 7). The monthly and fortnightly peaks appeared bimodal, which is consistent with the difference between the semi-monthly inequality (14.77 days) related to the moon phases and the declination inequality (13.66 days) related to the moon declination. The synodic month, related to the moon phases, is 29.53 days long while the tropical month, related to the moon declination, is 27.32 days long [9]. The perigee cycle is 27.5 days long. The semidiurnal peak looks rather broad, suggesting contributions by nearby frequencies (Fig. 8).

Periodic increases in seiche activity were observed since the early eighties [3]. These increases, which occurred a week after syzygy (new or full moon) were particular large when the new or full moon had coincided with perigee. This is compatible with the spectral analysis result, which in addition showed increases in activity along the perigee cycle and possible along the lunar declination cycle. This pattern of activity along the different lunar cycles suggests that the tides are either directly or indirectly involved in forcing or in assisting these coastal water oscillations.

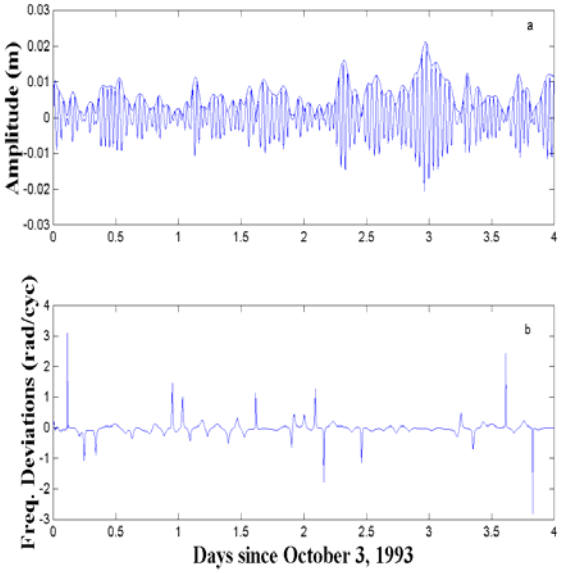

Fig. 6. a) A four-day record of seiche activity and its upper envelope. b) Deviations in frequency affecting the seiche period. Note that deviations from the center frequency occur at small seiche amplitudes.

IV. Distribution of Seiche Activity along the Different Lunar Cycles

While the distribution of activity along the year shows apparently random increases in seiche amplitude (Fig. 9), its spectrum suggests the existence of periodic cycles of enhanced activity. In most cases, for example, large amplitude events showed a strong pattern of semidiurnal modulation (Fig. 10). This is consistent with the semidiurnal peak found within the signal spectrum (Fig.8).

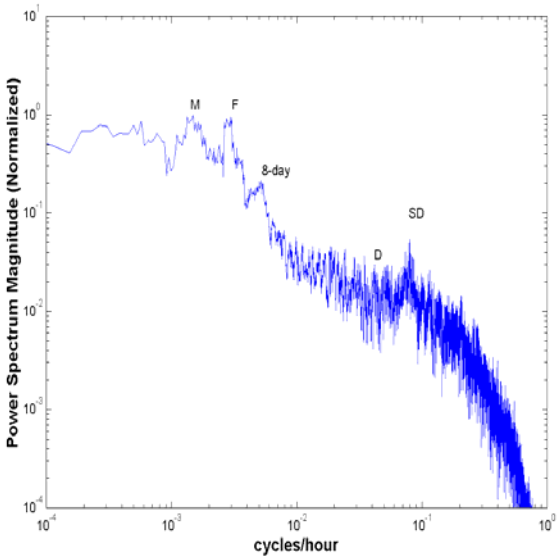

Fig. 7. Multitaper spectrum estimation of the seiche envelope. M: monthly, F: fortnightly, D: diurnal, SD: semidiurnal.

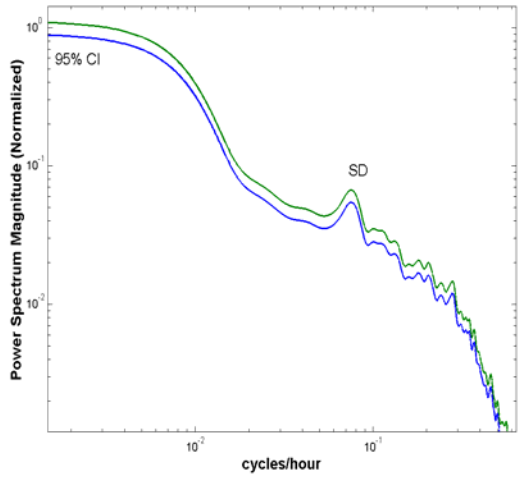

Fig. 8. Spectrum of the seiche envelope. Only the semidiurnal peak could be resolved when a small 1024-point window was used. The confidence interval is the distance between the spectral lines.

In order to validate the spectral analysis results, while at the same time gaining insight into the phase relation of the phenomenon with respect to the tides, the average profile of activity along the different lunar cycles was obtained. Increases in activity can be tracked looking at the signal envelope. Therefore three ensembles of the signal envelope starting at each lunar cycle were created. Each ensemble was 49-records long. Ensemble averaging gave the mean amplitude distribution along each cycle. It was found that the average distribution of activity approximated a sinusoidal profile with period consistent with previous spectrum estimates. Seiche activity peaked about six days after syzygy [3], six days after lunar perigee (Fig. 11, 12) and showed a minimum six days after maximum lunar declination.

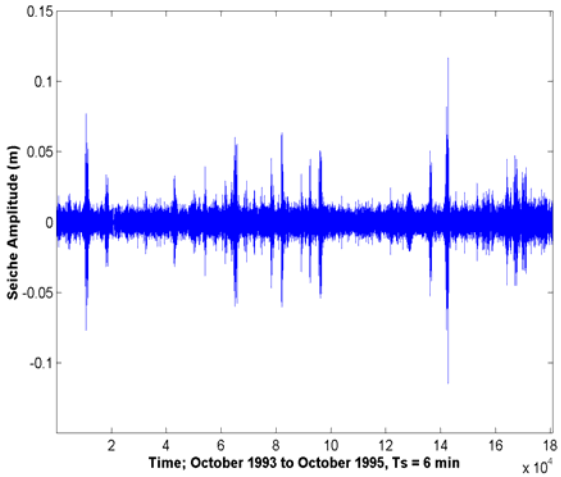

Fig. 9. Coastal water oscillations at 1.2 cph (seiches). Larger than average events are clearly seen along this record.

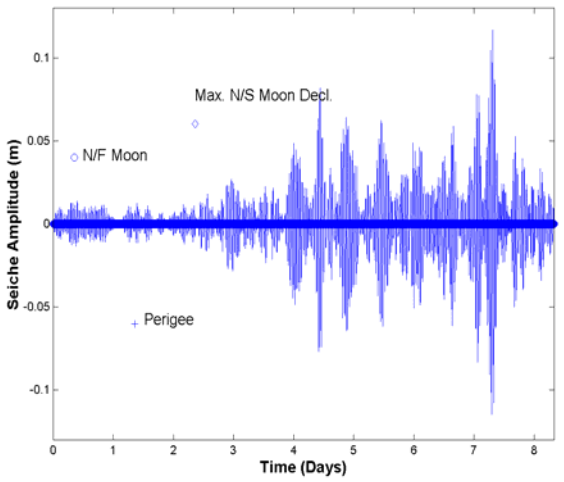

Fig. 10. Episode of increased activity and its relation to the lunar cycles. This event, which took place during the second half of May 1995 shows a strong pattern of semidiurnal modulation.

V. A LONG-LEAD SEICHE FORECAST MODEL

The fact that the seiche envelope was modulated in a cyclic fashion by periods related to the lunar cycles allowed the formulation of a predictor by linearly combining different harmonic constituents (2). A minimization procedure can provide a good fit between the predictor and the observed activity. Since this phenomenon is inherently noise and the periodicities are not as strong as in a tide record a good fit not necessarily guarantees a good forecast skill. In order to enhance the forecast capability of the model the phase of each component was calibrated based on the average distribution profiles of activity rather than on a blind minimization procedure (Fig. 11, 12).

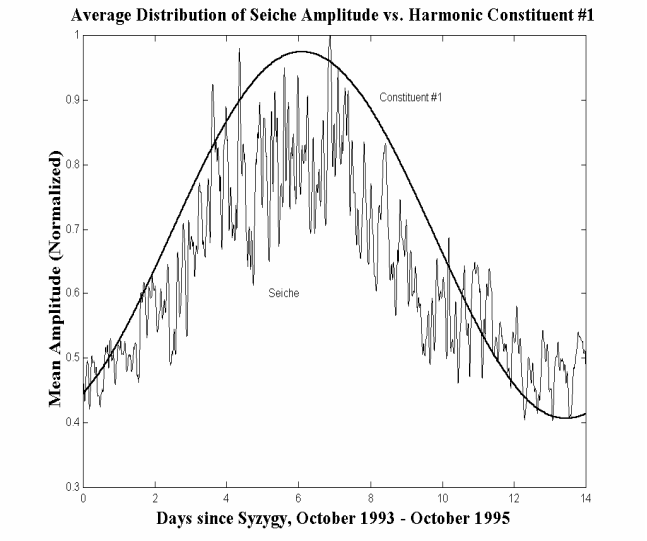

Fig. 11. Average distribution of seiche activity along the lunar phases cycle. The first harmonic constituent used in the forecast model shows a tight fit to the mean distribution profile of activity.

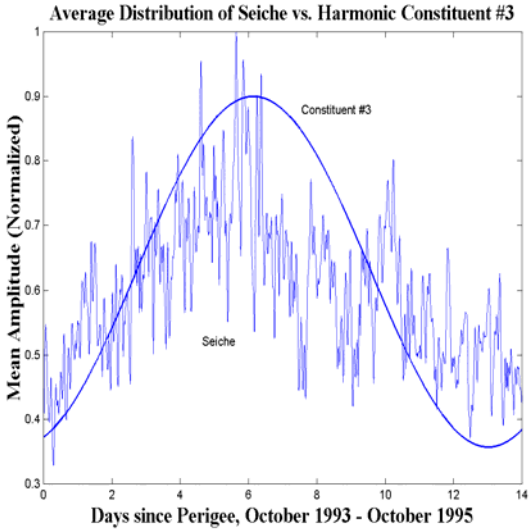

Fig. 12. Average distribution of seiche activity along the lunar perigee cycle. Harmonic constituent #3 follows the average profile of seiche activity along the cycle.

This method ties the constituents phase to a recurrent physical event, something that not only improves the long-term forecast capability of the predictor but also provides additional clues to further understand the phenomenon [10]. The frequencies of the harmonic constituents were determined based on the envelope spectral profile, which gave periods consistent with the lunar cycles. The weights of each component were determined by a standard least square minimization technique. The minimization procedure was performed using a low pass filtered version of the seiche envelope. The maximum resolution attained was about 5 days.

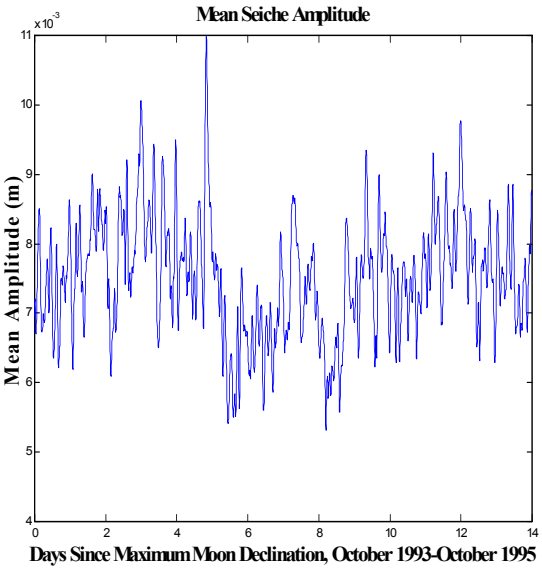

Fig. 13. Average distribution of seiche activity along the lunar declination cycle.

The mathematical formulation where the f_k are the frequencies, the c_k are the weights and the θ_k are the phases it is as follows:

$$\hat{seienv}(t) = \sum_k c_k \cos(2\pi f_k t + \theta_k) \qquad (2)$$

The first predictor included three components, one related to the lunar phases cycle, one related to the lunar declination cycle and the third related to the lunar perigee cycle. The perigee component as it was used, was not a pure harmonic constituent since a sinusoid of 13.75-day period was windowed to obtain the 27.5-day period of the cycle. This method closely matched the mean distribution of activity along the whole 27.5-day cycle but also introduced harmonics at 9.167 and 5.49-day periods (Fig. 14). Remarkably spectral components of nine and five-day periods are clearly seen in most subsections, such as 1994 - 1995, of the two-year record under study. The same periods are found in the 1996 – 1998 tide record used to validate the predictor (Fig. 15). As it is shown in Fig. 16 the forecasted distribution of activity followed reasonable well the average distribution of activity along the perigee cycle.

In order to make sure that the three components were relevant to the forecast model the correlation coefficient between pairs of constituents was calculated. The correlation was low indicating that they were almost uncorrelated and therefore non-redundant. After a principal component analysis was performed three constituents were still needed. Stepwise regression analysis rendered the declination constituent as the less significant component in the predictor.

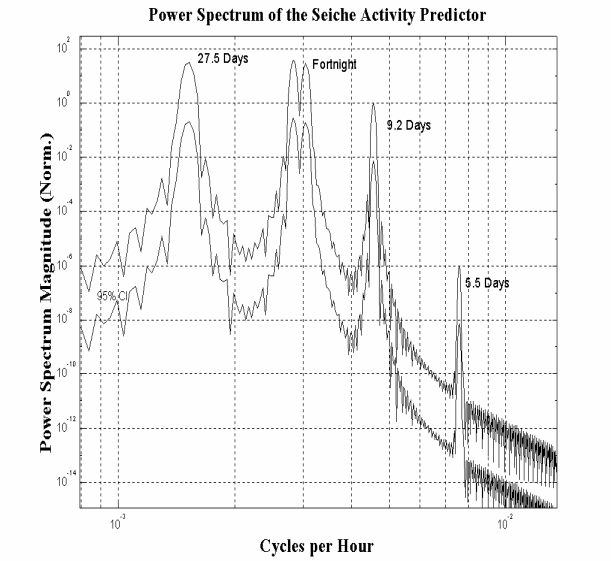

Fig. 14. Power spectrum of a low-pass filtered version of the seiche activity predictor. The monthly peak corresponds to the perigee cycle. The bimodal fortnightly peak corresponds to the lunar phase and declination cycles. The other peaks are harmonics of the monthly period.

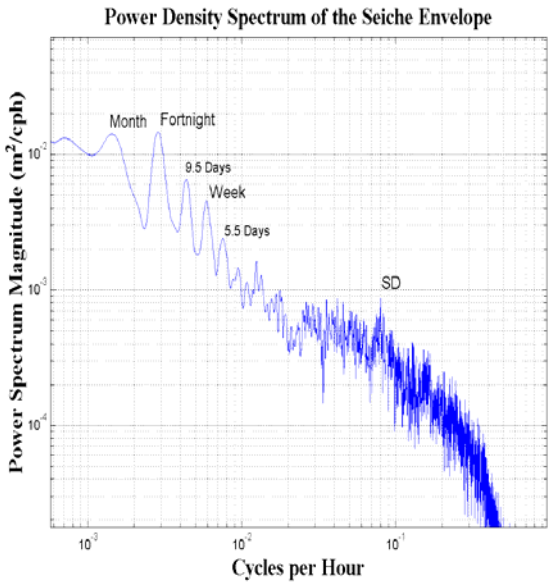

Fig. 15. Power density spectrum of the seiche envelope. January 1996 - May 1998. a) This is a Welch's averaged periodogram estimate that used a 32768 window with half overlap. SD: semidiurnal period. Sampling period was 6 minutes.

VI. Predicted versus Observed Activity

In order to validate the model, which was formulated based on the 1993 to 1995 tide record, a hindcast of seiche activity for the years 1996 to 1998 was performed. Linear, non linear as well as quadratic models were explored.

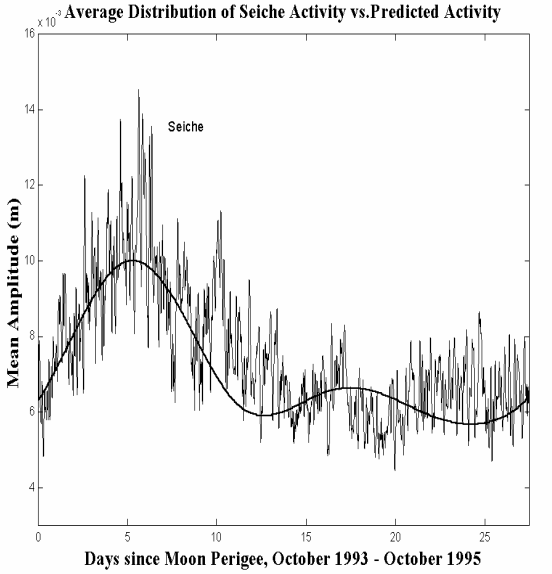

Fig. 16. Mean amplitude distribution of predicted activity along the lunar perigee cycle. The predicted activity follows that of the seiche, especially along periods of increased activity.

A. *Linear Models*

Looking at Fig. 9 it is clear that a linear predictor (2) will have problems following the sudden excursions in amplitude characteristics of the large seiche events (Fig.17). If the main interest is to forecast only the large amplitude occurrences then the predictor can be threshold to a certain value. This gives a good result but all information regarding small amplitude episodes is lost (Fig. 18).

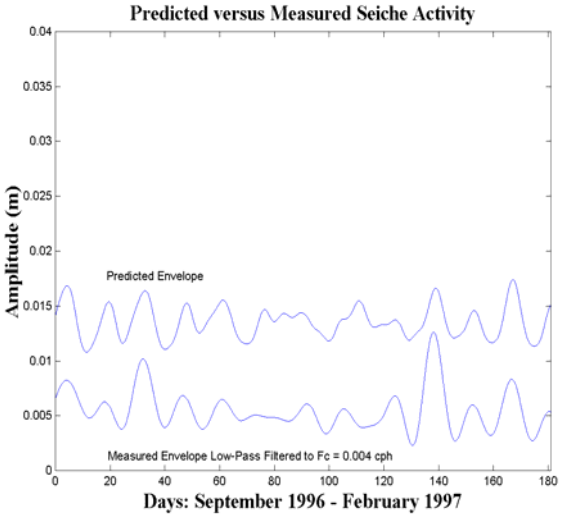

Fig. 17. Predicted versus observed seiche activity. The predicted envelope is offset to facilitate comparisons. This was the best hindcast for a continuous 6-month period within the test record.

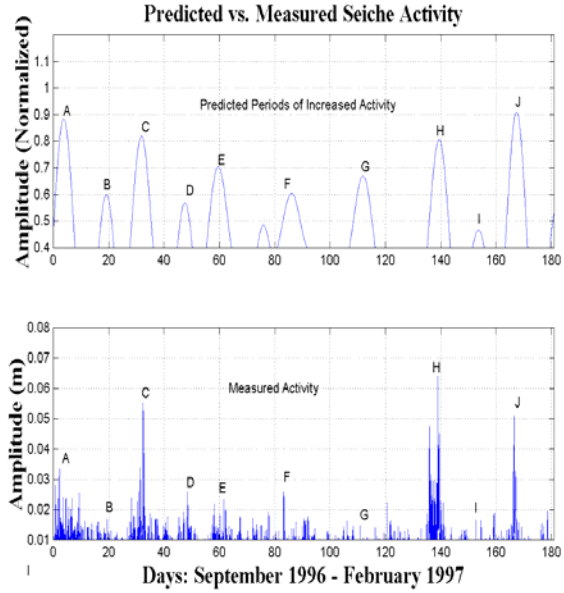

Fig. 18. Predicted versus observed seiche activity. The amplitude of the observed activity and the predictor has been limited to highlight large amplitude seiche events.

B. *Non-Linear and Quadratic Models*

A non-linear model can be used to expand the dynamic range of the predictor. The use of exponentials enhances the largest excursions in amplitude while leaving the small amplitude episodes almost unchanged. The mathematical formulation is the following:

$$\widehat{seiche} = \beta_1 e^{\alpha_1 . h_1(t)} + \beta_2 e^{\alpha_2 . h_2(t)} + \beta_3 e^{\alpha_3 . h_3(t)}$$

Where the $h_i(t)$ are the harmonic constituents. This predictor keeps up with the sudden excursion in amplitude but the false alarms are also magnified (Fig. 19). While the correlation coefficient, between the predicted and the observed profiles of seiche activity, was 0.7253 for the linear predictor for the non-linear predictor it was 0.7477. Along other time periods the linear predictor did better.

Quadratic models are linear in the parameters but include quadratic and cross product terms with respect to the constituents. These terms could naturally model some of the higher frequency spectral constituent found within the seiche envelope spectral profile and therefore provide a higher resolution model. A predictor that included the cross product terms improved the resolution but the correlation coefficient fell to 0.6846. Stepwise regression will be used in the future to assess the inclusion of the different quadratic or cross product terms into the model.

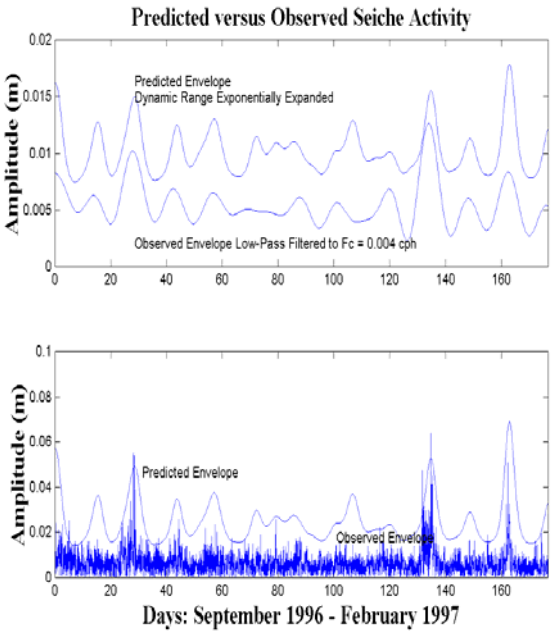

Fig. 19. Predicted versus observed seiche activity. In the lower graph the dynamic range of the predictor has been augmented to the point that matches that of the unfiltered seiche envelope. The predicted envelope is offset to facilitate comparisons.

VII. The Forecast Skill

Several performance indicators were used to quantify the skill of the predictor, specifically *the mean square error*, *the mean absolute error* and *the correlation coefficient*. A visual inspection of the results it is also a useful tool of assessment. The correlation coefficient has been preferred because it gives an indication of how well the model "follows" the phenomenon, regardless of the achieved accuracy. The exact amplitude of the different events could respond to a variety of causes such as weather, stratification conditions and random barometric fluctuations among others.

The most common statistics for measuring the skill of the forecast is the *coefficient of determination;* R^2, it is the ratio of the sum of squares due to the regression and the total sum of squares. As defined, this is a number between zero and one. When R^2 is large (close to 1) the model describes a large portion of the variation in the data. On the other hand when R^2 is small (close to 0) the model accounts for only a small fraction of the variation in the data [11]. The objective is to find models with high R^2. In our case, after adding a fourth component of weekly period and using stepwise regression to assess the inclusion of constituents into the model, the linear predictor accounted for up to 75% of the low frequency variability in the observations.

VIII. The Residuals and the Adequacy of the Model

Since the observed amplitude variations have not been exactly fitted there exist considerable errors that can be represented as the difference between the predicted and observed profiles. This is called the residual. Regardless of whether these errors are large or small they should be of a random nature and have a normal distribution. Therefore, except for zero-lag the residuals should be uncorrelated for all time. The autocorrelation of the residuals, shown in Fig. 20, suggests that there is further room for improvement. The test for normality of the residuals is shown in Fig. 21.

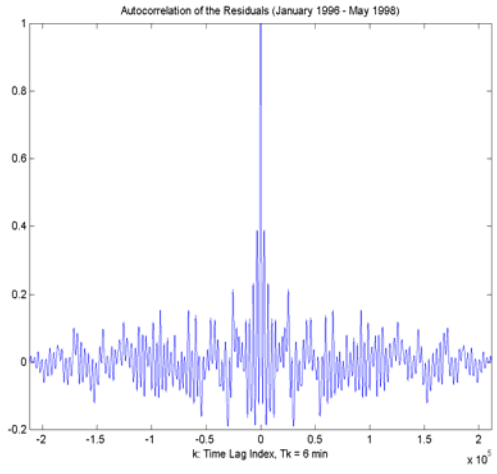

Fig.20. Autocorrelation of the residuals

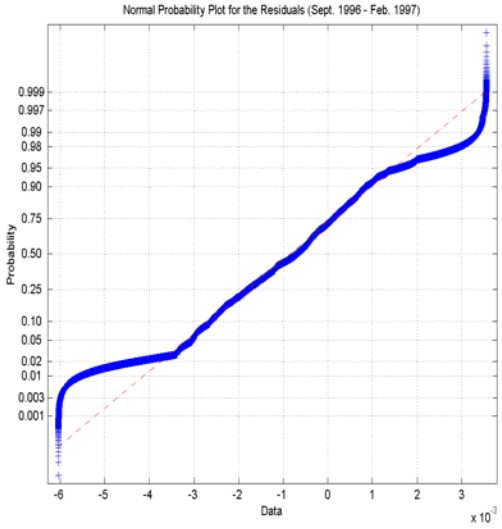

Fig. 21. Test for normality of the residuals.

VIII. Possible Forcing Mechanisms

The fact that the increases in seiche activity occur at neap semidiurnal tide and considering that the internal tide is predominantly semidiurnal at this location it is clear that these coastal water oscillations peak under conditions of maximum stratification. Assuming that the seiches are forced by incoming internal waves when impinging against the steep slope, as suggested in [3], and in view that the mean distribution of activity shows a gradual increase towards moon quarters, it is reasonable to conclude that the origin of these waves is likely to be local or regional rather than remote in nature. Placing the origin of these waves in a local or regional context relax the constraint that they should be of the solitary type. Under the right stratification conditions internal waves can be formed in a number of different ways. Even internal tidal bores, which tend to occur by neap tide [12], could be forcing these seiches when impinging against the steep shelf slope. This should be tested by the deployment of an array of thermistors close to the shelf break and along the mixed layer boundary.

IX. Conclusions

A long-lead forecast model for the prediction of coastal water oscillations along the Caribbean coast of Puerto Rico was formulated based on the analysis of a two-year long tide record. The predictor was validated carrying out a hindcast of seiche activity along a different 2.4-year tide record. The average distribution of seiche activity along the different lunar cycles suggests that the forcing could be local or regional in nature. The developed model should facilitate the determination of appropriate data collection schedules for further studies of this phenomenon.

REFERENCES

[1] G. S. Giese and D. C. Chapman, "Coastal seiches," *Oceanus*, Spring 1993, pp. 38-45.

[2] R. W. Fairbridge, *The Encyclopedia of Oceanography.* Reinhold Publishing Corp., 1966.

[3] G. S. Giese, D. C. Chapman and P. G. Fornshell, "Causation of coastal seiches on the Caribbean coast of Puerto Rico," *J. Phys. Oceanogr*., vol. 20, pp. 1449-1458, 1990.

[4] D. C. Chapman and G. S. Giese, "A model for the generation of coastal seiches by deep-sea internal waves," *J. Phys. Oceanogr*., vol. 20, 1459-1467, 1990.

[5] R. H. J. Grimshaw, and D. C. Chapman, "Continental shelf response to forcing by deep sea internal waves," *Dyn. Atmos. Oceans*, vol. 16, pp. 355-378, 1992.

[6] S. Pond and G. Pickard, *Introductory Dynamical Oceanograph.* 2nd ed. Butterworth Heinemann, Oxford, Great Britain, 329 pp, 1997

[7] Percival, D. B., and A. T. Walden, *Spectral Analysis for Physical Applications,* Cambridge University Press, Cambridge, Great Britain, 583 pp, 1993.

[8] G. E. Hutchinson. *A Treatise on Limnology.* John Wiley & Sons, New York, NY, 1015 pp. , 1957.

[9] A. Defant, *Physical Oceanography.* Vol. 2. The MacMillan Company, New York, NY, 598 pp., 1961.

[10] M. Teixeira, *Characterization of coastal seiches along the Caribbean coast of the island of Puerto Rico.* Ph.D. thesis, Marine Sciences Dept. University of Puerto Rico, Mayaguez, Puerto Rico, 1999.

[11] B. Abraham and J. Ledolter, *Statistical Methods for Forecasting,* John Wiley & Sons, New York, NY, 445 pp., 1983.

[12] J. Pineda, An internal tidal bore regime at nearshore stations along western USA: predictable upwelling within the lunar cycle. Continental Shelf Research. Vol. 15, pp. 1023-1041, 1995.