

Neural Network prediction of the low-frequency coastal sea level response using conventional and NCEP/NCAR Reanalysis data

M. M. F. Oliveira¹ N. F. F. Ebecken¹, L. P. Caloba¹, I. A. Santos²
& J. L. F. Oliveira³

¹*COPPE/UFRJ, Brazil*

²*Meteorology/UFRJ, Federal University of Rio de Janeiro, Brazil*

³*Geography/UFF, Federal Fluminense University, Brazil*

Abstract

This paper presents a study about the variations of the coastal sea level in Paranaguá Bay-Paraná State, South Region of Brazil (SRB) from January 1997 to December 1998. Tidal forcing is the main cause of this variability, but the effects of meteorological variables are also present in the rising and lowering of the observed sea level. Hourly time series of the water level, atmospheric pressure and wind from the gauge tide and meteorological conventional stations were used. The Reanalysis data set of the “National Centers for Environmental Prediction” (NCEP) and the “National Center Atmospheric Research” (NCAR), on some points over the oceanic area near the bay were also used. The meteorological influences present in the original tide gauge records were extracted using a low-pass filter removing the oscillations with periods relative to the astronomical tide patterns. Meteorological time series were also filtered. Local and remote atmospheric driving forces were studied using statistical analysis on time and frequency domains. Therefore, the correlations in this physical process were defined to know the lag time between the meteorological variables and the coastal sea level response to the occurrences of the low frequency atmospheric systems. The NCEP/NCAR Reanalysis data translated better with the low frequency atmospheric phenomena variations showing that they are a very good information source for the South Atlantic Ocean (SAO) region where the lack of data is still substantial.

Keywords: Artificial Neural Network, coastal sea level variability, meteorological tide (Surge), time series forecasting, NCEP/NCAR Reanalysis data.



1 Introduction

The Brazilian coast may be divided into macrosegments based on regional characteristics. The study area is situated in the southeastern coastal region or Crystalline Scarps, extending from Cabo Frio, in Rio de Janeiro State, to the Santa Marta Cape, in Santa Catarina State (Figure 1).



Figure 1: Paranaguá Bay-Paraná State-Brazil.

2 Data analysis

This coastline has wide coastal plains, long beach barriers, and large estuaries. This region is located on the continental shelf of the southeastern coast of Brazil, which is wider than the shelf of the northern coast. The characteristics of the astronomical tide in the PP are shown in Table 1, where we note the semidiurnal components, M_2 and S_2 have the greatest semi-amplitude (H), followed by the shallow water components, M_3 and M_4 ; these indicate the influence of the propagation of the tide wave in the PB and the adjacent continental shelf, respectively.

The climate of this area is characterized by the presence of a high-pressure system over the SAO with northeast winds predominant, the sea and land breezes have local effects. This circulation is disturbed, periodically, by the passage of frontal systems caused by migratory anticyclones that blow from the southwest to the northeast in this region.

To study the local and remote meteorological influences in PB, hourly sea level records from Wharf West in the PP, and hourly averages of the atmospheric pressure and wind velocity from Pontal do Sul (Paraná State) conventional station for the 1997-1999 period were used. The atmospheric pressure and wind components at $2.5^\circ \times 2.5^\circ$ and $1.817^\circ \times 1.817^\circ$ grid points (latitude and longitude, respectively), at 00:00, 06:00, 12:00, and 18:00 UTC from the NCEP/NCAR Reanalysis data for the geographic coordinates of $25^\circ 00'S$ and $27^\circ 30'S$ and for the shoreline up to $45^\circ 00'W$ for the period 1997-1998 were used too.



Meteorological analyses and forecasts from the Forecast Daily Bulletins prepared by Centro de Hidrografia da Marinha (CHM) during the period were used.

Table 1: Astronomical tide components.

Harmonics Components	Frequency (cph)	Period (hsm)	Height (cm)	Spectral Density
M_2 (principal lunar semidiurnal)	0.080505	12	48.11	3592912.0
S_2 (principal solar semidiurnal)	0.083312	12	32.52	1407429.0
M_4 (quarter diurnal lunar)	0.161011	6	16.04	438177.8
M_3 (third diurnal lunar)	0.120789	8	14.96	366968.8
O_1 (declinational lunar diurnal)	0.038757	27	11.01	96689.5
K_1 (declinational luni-solar diurnal)	0.041748	24	7.11	58677.4

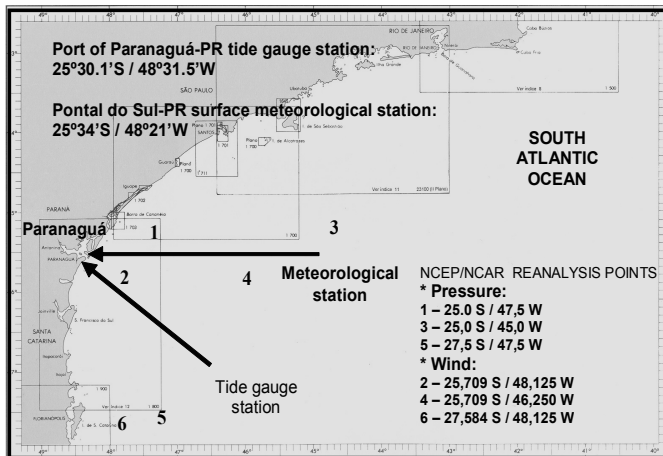


Figure 2: Geographic coordinates.

Figure 2 corresponds to Nautical Chart number 80/1964 (CHM), covering an area from Rio de Janeiro (Rio de Janeiro State) up to Santa Catarina Island (Santa Catarina State), and shows the locations of the tide gauge and the meteorological stations, as well as the NCEP/NCAR Reanalysis grid points. The numbers 1, 3, and 5 refer to the atmospheric pressure and 2, 4, and 6 refer to the wind components.

NCEP/NCAR Reanalysis data were used to evaluate the remote influence of atmospheric systems such as cold fronts, which are related to the low-frequency sea level variations near the coast and to use as source of information for this region, where the lack of data is still significant.

The time series was analyzed statistically by evaluating the extreme average values, variances, and standard deviation. Spurious values that had compromised the quality of the information were removed. The pre-preprocessing became the database more consistent.

In the meteorological series from the conventional station, the spurious values were substituted with the average values. The data for the period from January 1997 to September 1998 were then used for studying the behaviour of the sea level in PB in relation to the meteorological variables. All the tide gauge records were used to study the behavior of the mean sea level in this region.

This work is related to the oscillations in the sea level caused by low-frequency atmospheric changes lasting from 3 to 5 days. So we used a Thompson low-pass filter to eliminate the oscillations leading to tide phenomena and those of diverse scales characterizing the ocean-atmosphere interactions caused by the passage of a frontal system (Thompson [8]).

The tide forecast and harmonic analysis program ANHAMA of the PACMARÉ module was used to determine the main astronomical diurnal, semidiurnal, and higher frequencies components, as well as the local inertial or Coriolis component (Thompson [8]; Castro and Lee [2]). This program uses a percentage of 95% for rejecting components, describing each one of them with respect to the frequency, period, amplitude, and phase. The filter allowed us to impose a null function for the frequencies selected by the user, as well as two cut frequencies for a constant weights (n) amount.

The Thompson optimal filter uses the following aspects: cut frequencies of $\Omega_1 = 6.4^\circ/\text{h}$ and $\Omega_2 = 11.2^\circ/\text{h}$, with periods of 56.25 and 32.14 h, respectively; 16 frequencies with imposition of null response, ten of these being local main tide components and five being adherence frequencies, the local inertial component and the weight number $N = 120$. Therefore, the effect of the high frequencies related to the astronomical tide observed in the tide gauge records was eliminated.

The hourly average data of atmospheric pressure and zonal and meridional wind components from the conventional station were filtered using the Thompson optimal filter. To the NCEP/NCAR Reanalysis data the filtering were made using this filter, but with the number of weights, $N = 20$, reported for the data at intervals of 6 h with cut frequency of $\Omega_1 = 38.4^\circ/6 \text{ h}$ and $\Omega_2 = 67.2^\circ/6 \text{ h}$.

The cross-correlations between the mean sea level, atmospheric pressure, and zonal and meridional wind components were calculated. In view of the strong relation between the wind stress and the variations found in the coastal sea level, we analyzed the correlations existing between them to evaluate the importance of this force in the fluctuations of the coastal sea level. The zonal and meridional wind stresses were calculated. The mean sea level autocorrelation was calculated and it showed height values to 24 hours periods.

Spectral analyses of the meteorological data and the mean sea level were carried out. Fast Fourier Transformed was used. Cross-spectral analyses were carried out to identify the frequency characteristics of the local and remote meteorological events that have an influence on the variation of the mean sea



level at that point. The coherence between the peaks of the two decomposed series was analyzed to verify that a linear correlation existed between the components of the bivaried process. These correlations indicated a relative lag time with the sea level response to atmospheric disturbances and thus to determine the inputs variables of the neuronal model.

The input variables for the ANN training were atmospheric pressure, zonal and meridional wind stress, mean sea level series of 6, 12 and 18 previous hours and observed wind for the actual time (Oliveira *et al.* [5,6]).

3 Results and discussions

The time series filtered using the Thompson low-pass filter for all parameters were obtained and are shown in the Figure 3. We observe in the Figure 3 that the oscillations in the observed sea level were filtered, generating the series of the sea level response in low-frequency related with the meteorological systems.

The pressure (point 1), U and V wind components (point 2) filtered time series from the NCEP/NCAR reanalysis data are shown in the Figure 4.

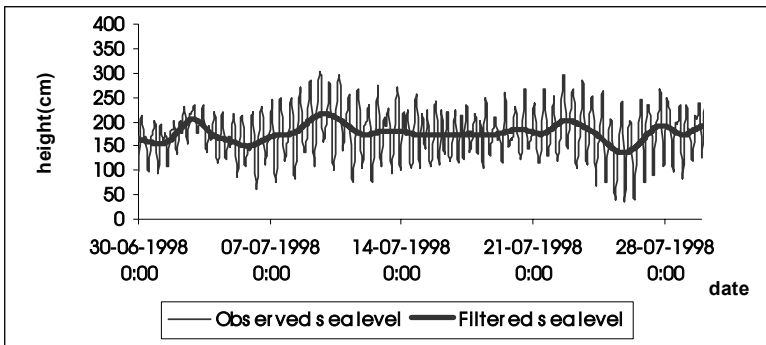


Figure 3: Observed and filtered time series.

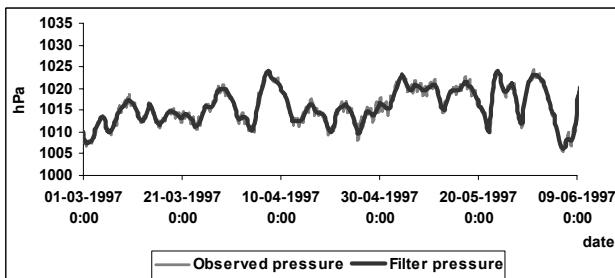


Figure 4: Observed and filtered reanalysis time series.

Table 2 presents the relative values of the variance analysis of the sea level, in which we observe oscillations of 12.29% about the mean sea level related mainly to the influences of low-frequency meteorological phenomena. These oscillations are also called meteorological tides or surges in the literature. Although the meteorological tide has a small contribution compared with the astronomical and physiographical factors, it is relevant when it is compared with the observed sea level values, mainly when significant meteorological events occur (Oliveira [6]).

Table 2: Sea level variance.

Sea level	Variance	SD	%
Observed	2705.19	52.04	100
Astronomical tide	2338.37	48.36	86.47
Mean-sea level	332.71	18.19	12.29

The atmospheric disturbances generated in the high latitudes and spreading in the direction of the Equator near the coastal regions can generate Kelvin waves with periods of 5 to 18 days, effectively the response of the ocean to the weather changes (Brink [1]). The spectral analysis of this variation, with maximum peaks around 3, 6, and 18 days, characterizes the response of the mean sea level to the low-frequency driving forces (Figure 5).

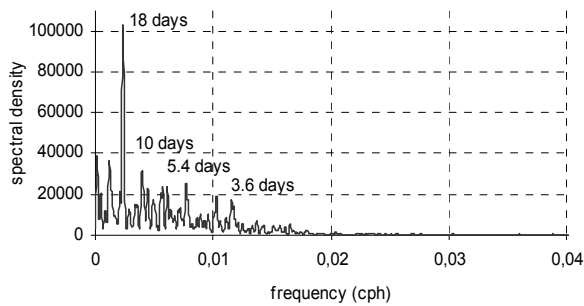


Figure 5: Spectral analysis of the mean sea level.

Table 3 lists the values of the crossed-correlations of the sea level with atmospheric pressure, zonal (U) and meridional (V) wind components and wind stress. The response of the mean sea level to the atmospheric pressure in the station and the point 1 is the same around 42%. The greatest lag time was 43 h in the conventional station, which can be related to the physiographic effect of PB. For the wind components, the greatest percentage of crossed-correlation was found for more distant point.

The point close to the tide gauge station had a high influence arising from the meridional component of the wind.



Table 3: The maximum lag time.

Conventional station			Reanalysis			
Var	lag time	%	1 (Pressure)		2 (Wind)	
			lag time	%	lag time	%
P	43	44	36	41.2		
U	27	28.5			30	42.5
V	8	58.3			6	57.8
Tx	31	15.5			24	34.7
Ty	7	48.4			0	55

The spectral densities showed a concentration of energy for periods from 5 to 3 days decreasing for periods less than 3 days and greater than 10 days with peaks for 17 and 20 days. Similar results had been found by (Castro and Lee [2]) with respect to the effect of waves in the continental shelf. The corresponding peaks for periods between 3 and 6 days were identified by these researchers as being caused by the occurrence of meteorological events such as cold fronts, followed by low-pressure systems.

Height coherence between the variables and the mean sea level was found with periods varying from 10 to 3 days.

The crossed-spectrum of the pressure against the mean sea level showed peaks related to events with periods of about 4 to 3 days at the station and points of reanalysis. High values of energy were found for events with smaller frequencies. For U and the zonal wind stress (T_x) components, the energy values showed equivalence to the three points of reanalysis data with peaks for periods less than 3 days. In the station, U had greatest values than the zonal wind stress. The V and meridional wind stress (T_y) components displayed a behavior similar to that of the four points. The V component had greatest energy values in the station compared with the reanalysis point.

To compare the simulations with both dataset we used only the points 1 (pressure) and 2 (wind) of reanalysis.

The input variables for the ANN model based in these statistical results showed that the best performance was reached with the multiple layers perceptron, with 7 neurons in the input layer, 14 neurons in the intermediate layer and 1 neuron in the output layer with 700 epochs.

The maximum correlations between the sea level response and the meteorological variables were considered as input, then the variables had the following lag time: pressure (-42h), zonal wind stress-Tx (-30h), meridional wind stress-Ty (-6h), Sea Level (-18h), Sea Level (-12h), Sea Level (-6h) and Wind (0). The output model is the low-frequency sea level variations.

The linear-logistic-linear activation functions were used in this model and the backpropagation algorithm was used for the ANN training (Oliveira *et al.* [5,6]).

Figure 6 presents the evolution of the ANN training and verification to reach the error convergence. The black line represents the training and the grey line the verification.

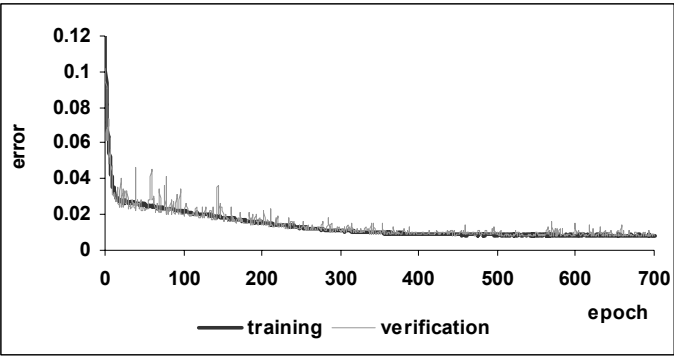


Figure 6: ANN training and verification.

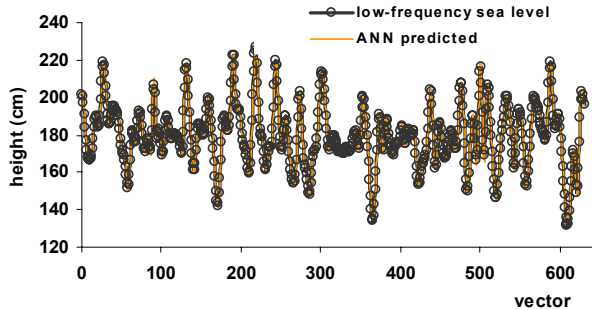


Figure 7: ANN predicted and the sea level.

Figure 7 presents the results obtained with the ANN generalization (test) to predict the variations of the low-frequency sea level. The dashed black line shows the behavior of the sea level and the yellow line the predicted by the neural network. It is verified that the difference between the two curves keeps constant, practically, not presenting extreme values between them.

NCEP/NCAR reanalysis data from points 1 and 2 were used in the same architecture developed for the conventional data. The maximum correlations also were considered as input, then the variables had the following lag time: pressure (-36h), Tx (-24h), Ty (0), SL (-18h), SL (-12h), SL (-6h) and Wind (0). Figure 8 shows the curves related with the evolution of the ANN training and verification to reach the error convergence. The results presented with the conventional dataset were similar to these and the black line represents the training and the grey line the verification too.



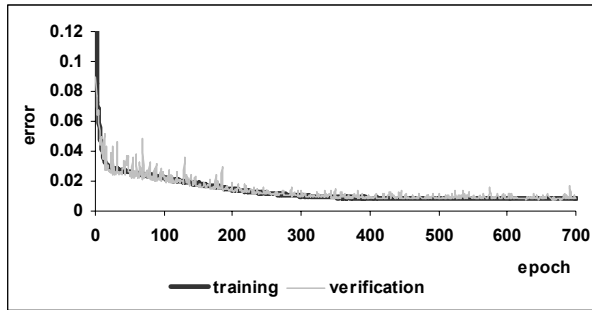


Figure 8: ANN training and verification.

4 Conclusions

The maximum existing correlations in the physical process determined the lag time between the variables. The pre-processing of the data series in the time and frequency domain allowed us to define the input model. We could find the response of the mean sea level to the occurrence of low-frequency atmospheric events.

The frequency domain analysis with the conventional data showed the greatest values of the cross-spectral density of periods of around 5 days with high coherence. This can be justified as being caused by the proximity of the two stations (tide gauge and meteorological conventional station), both suffering the local influences of the migratory atmospheric events. The spectral analyses for point 1, to the north of the tide gauge station, and point 2, near Paranaguá, gave lower values for the spectral density and coherence.

The NCEP/NCAR Reanalysis data selected were compared with those obtained from the conventional station. Thus, the relations between the atmospheric phenomena and the response of the sea level are much clearer with information numerous points. Another important aspect that we found with the NCEP/NCAR Reanalysis data was that remote events affect the mean sea level at PB. This was found mainly with the southern Paranaguá points, where the meridional wind component has high values, showing the influence of the direction of the Brazilian coast in this place. The NCEP/NCAR Reanalysis data translated better the low frequency atmospheric phenomena variations showing that they are a very good information source for the South Atlantic Ocean (SAO) region where the lack of data is still substantial.

The ANN used to predict the coastal sea level presented good performance with both data sets. The model used for 6 and 12 hours lag time simulations was able to capture the effect of the passage of the frontal systems with an approach of 99% for both data source. It has demonstrated to be useful as complement for the standard constant harmonic model to that time delay. The results for 18 and 24 hours lag time simulations were about 85% and 69%, respectively, for

conventional data and 91% and 78% for reanalysis data, suggesting the others architecture development to improve the predictions to larger periods.

The ANN developed in this work could generalize satisfactorily the non-linear behavior of the sea-level fluctuations due to the interactions ocean-atmosphere.

References

- [1] Brink, K. H., "Coastal-trapped waves and wind-driven currents over the Continental Shelf. *Annual Review of Fluid Mechanics*", 23, 127-133, 1991.
- [2] Castro, B. M., Lee, T. N., "Wind-forced sea level variability on the southeast Brazilian shelf", *Journal of Geophysical Research* 100, 16045-16056, 1995.
- [3] Gill, A. E., "Atmosphere-Ocean Dynamics. *International Geophysics Series*", Academic Press, New York, 662 pp, 1982.
- [4] Rumelhart, D. E.; Hinton, G. E.; Willians, R. J. "Learning Internal Representations by Error Propagation", *Parallel Distributed Processing: Exploration in the Microstructures of Cognition*. MIT Press, v.1, p. 318-362, 1986.
- [5] Oliveira, M. M. F. de., Santos, I. de A., Oliveira, J. L. de., Calôba, L. P., Ebecken, N. F. F. Artificial neural network to predict the meteorological tide. In: *The International Workshop on Harbour, Maritime & Multimodal Logistics Modelling and Simulation – HMS2004 & AMS2004*, 8, 2004a Rio de Janeiro. *Proceedings...* Rio de Janeiro: LAMCE/UFRJ - BRAZIL, 2004.
- [6] Oliveira, M.M.F., "Neural Network Prediction of Paranagua Coastal Leqa Level Response", M.Sc. Thesis, COPPE/Federal University of Rio de Janeiro, 149 pp., 2004.
- [7] Pugh, D., "Tides, Surge and Mean Sea Level", *A Handbook for Engineers and Scientists*. John Wiley and Sons, New York, 427 pp, 1987.
- [8] Thompson, R. O. R. Y., "Low-pass filters to suppress inertial and tidal frequencies", *Journal of Physical Oceanography* 13, 1077-1083, 1983.