Proceedings of 8 ICSHMO, Foz do Iguaçu, Brazil, April 24-28, 2006, INPE, p. 781-785.

THE EFFECT OF PRECIPITATION FORECASTS AND MONITORING ON HOURLY HYDROLOGICAL STATISTICAL PREDICTION MODELS

Eduardo Alvim Leite[1], Nelson Luis da Costa Dias[2] e Otto Correa Rotunno Filho[3]
SIMEPAR, LEMMA/UFPR e PEC/COPPE/UFRJ, Brazil

1. INTRODUCTION

The so-called "electric sector" of Brazil, namely the set comprising electricity companies (state-owned and private) plus regulatory agency has been implementing automated telemetric systems for hydrological monitoring in the country's main energy-producing watersheds. This is motivated both by the electricity regulatory agency (ANEEL) requirements and by the belief that better (meaning higher frequency and telemetric) monitoring will lead to more safety and better reservoir operation with regard to multiple use criteria (mainly energy-generation and flood-control).

Once an automated monitoring system is in place, it is desirable to assess its effectiveness in providing reliable hydrological forecasts.

Here, we will try to study the case of the Iguaçu River Monitoring System upstream of Foz do Areia Reservoir. For that purpose, we use relatively well-known statistical forecasting tools: ARIMA models and Kalman Filter models. In both cases we compare the use of conventionally operated stream gages (with data reported by radio or telephone two times a day) to the now more readily available hourly telemetric data. We also analyze the impact of precipitation forecasts (simulated from historical records with different degrees of skill) on the quality of the streamflow forecasts. It is expected that objective ways of assessing the benefits of enhanced monitoring will play an important role in defining criteria for network project and cost-benefit analysis in some of the Country's key water resource sectors.

2. METODOLOGY

2.1 Study Site

In this work we study streamflow forecasts in the Iguaçu River at the City of União da Vitória, just upstream of the Foz do Areia Reservoir. This is the most upstream and largest reservoir in the Iguaçu River and is the main responsible for regulating energy generation in the cascade. It has an installed capacity of 1676 MW. During floods, the reservoir's backwater effects can worsen flooding at União da Vitória, so that a rather unusual *upstream* flood control constitutes a permanent challenge for the reservoir's management. Thus, streamflow forecasts are an essential tool for good reservoir operation. Besides the obvious importance of this site, it has a good record of conventional and automated hydrological data. In summary, in this work we use hourly streamflow and precipitation data from the automated hydrological stations, and conventional (i.e. manually recorded) streamflow data manually recorded two times a day (at 7:00 and 17:00 hrs, local time). The calibration period is 1998-2002, whereas verification was performed for the 2003-2004 years. Figure 1 shows the study region.

2.2 ARIMA Models

ARIMA (Autoregressive Integrated Moving Average) models are widely used in the Brazilian Electric Sector, and for that reason were used as a first approach to streamflow forecasting in the present work. They are relatively easy to implement and can cope with some nonstationary effects. We followed the standard procedures for model identification, parameters estimation and model verification as presented by Bras and Rodríguez-Iturbe (1993, p 55).

Altogether, we implemented 6 different ARIMA (p,d,q) forecasting models, where p is the number of autoregressive terms, d is the number of time differences applied to the nonstationary streamflow data series in order to produce a stationary time series, and q is the number of moving average terms.

[1] Instituto Tecnológico SIMEPAR, Caixa Postal 19100, CEP 81531-990, Curitiba, PR, Brasil, alvim@simepar.br.
[2] Laboratório de Estudos de Monitoramento e Modelagem Ambiental / Universidade Federal do Paraná, Caixa Postal 19100, CEP 81531-990, Curitiba, PR, Brasil, nldias@ufpr.br.
[3] Programa de Engenharia Civil / Coordenação de Programas de Pós-Graduação em Engenharia / Universidade Federal do Rio de Janeiro, Centro de Tecnologia da UFRJ, Bloco B, Caixa Postal 68540, CEP 21945-970, Rio de Janeiro, RJ, Brasil, otto@hidro.ufrj.br.

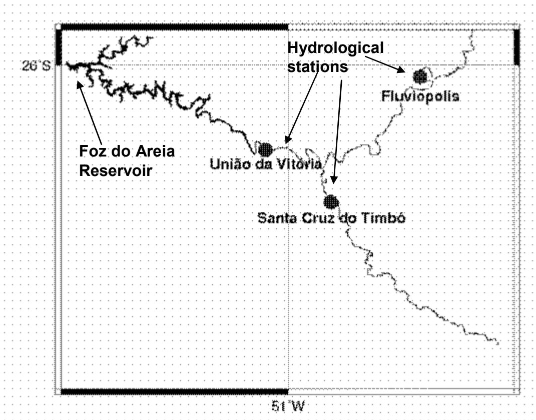

Figure 1: Study region in the Iguaçu River upstream of the City of União da Vitória, Paraná State, Brazil

From the 6 adjusted models, we then chose an ARIMA (7,1,3) with the least mean square error for forecasts with telemetric data, and an ARIMA (4,1,2) for forecasts with the conventional data. In both cases, the models are tested for forecasts with lead time from 12 up to 72 hours ahead. Notice that the orders refer to the data time resolution; thus the ARIMA (7,1,3) uses data from the last 7 hours, whereas the ARIMA (4,1,2) uses data from the last 4 12-hour periods.

2.3 Models based on Kalman Filters

There are several shortcomings in ARIMA-based forecasting models; probably the most notorious is their forecasts invariably lag behind the observed flows during the hydrograph ascent, the reverse happening during recessions. Thus, more sophisticated Kalman-filter models were also tested. Our implementation follows the lines of Anderson and Moore (2005, p 50). We use a linear Kalman filter with 20 discrete time states. The state estimation is based on the following predictors: (a) difference in time of streamflows at the forecasting site; (b) difference in time of streamflows at two upstream watersheds, Fluviópolis and Timbó, (c) the arithmetic mean of 12-hour precipitation measured at the União da Vitória, Fluviópolis and Timbó hydrological stations and (d) precipitation forecasts by class for a 24-hour lead time, split into 4 6-hour intervals.

Because the ascent and the recession of the hydrograph are controlled by very different factors, we adopted the well-known approach of considering them separately. Thus we implemented two Kalman filters, which keep the memory of the last state of ascent/descent, and are turned on and off as the hydrograph rises and falls. This is a simple way of not changing the model state too abruptly, but rather remembering the last similar state, and usually produces much better forecasts. The resulting Kalman Filter-based forecasting model uses as input telemetric hourly streamflow and precipitation data from 2 hydrological stations upstream and at the forecasting site, as well as precipitation forecasts by class for the next 0-6, 6-12, 12-18 and 18-24 hour intervals. After some sensitivity studies, we defined the following precipitation classes: low (up to 2.5 mm accumulated in 6 hours), medium (2.5-10 mm), high (10-20 mm) and extreme (> 20 mm).

2.4 Assessing the Quality of Forecasting

There are two possible approaches for the assessment of the quality of streamflow forecastings: measure-oriented and distribution-oriented (Katz and Murphy, 1997). Both are used in the present work. We used the classical Mean Square Error (MSE), but also introduce a *positive* mean square error (MSE^+) (when the forecast is greater than the observed flow) and a *negative* one (MSE^-) (when the forecast is less than the observed flow), defined by:

$$MSE(f,x) = \sum_f \sum_x p(f,x)(f-x)^2 \tag{1}$$

$$MSE^+(f,x) = \sum_f \sum_x p(f,x)[zifn(f-x)]^2 \tag{2}$$

$$MSE^-(f,x) = \sum_f \sum_x p(f,x)[zifn(x-f)]^2 \tag{3}$$

where f is the forecasted value, x the observed value, and $zifn(a-b)$ is zero if $a \le b$ and equal to $a-b$ otherwise.

The skill score SS_{MSE} is defined by:

$$SS_{MSE}(f,r,x) = 1-(MSE(f,x)/MSE(r,x)) \tag{4}$$

where $MSE(f,x)$ is the mean square forecasting error, and $MSE(r,x)$ is the reference mean square error (in this work, the reference error is the error of a forecasting model that uses only manually recorded data, to be described in the sequence). Similar scores were used for MSE^+ and MSE^-.

Since we were especially interested in lag errors (LE), we defined the quantities:

$$LE = \sum_{j=1}^{m} p(r_j)D_j \tag{5}$$

$$LE_j = \frac{\sum_{i=1}^{n} zifn(Tf_{ij} - Tx_{ij})}{n} \tag{6}$$

where m is the number of observed streamflow classes, r_j is the representative value for each class, $p(r_j)$ is the probability of class j, LE_j is the lag error assigned to class j, n is the number of periods i during which $x \geq rj$, Tf_{ij} is the time at which the forecasted flow reaches the value r_j during period i and Tx_{ij} is the time at which the observed streamflow reaches the value r_j during period i. The skill score for the lag error is given by:

$$SS_{LE}(f,r,x) = 1 - (LE(f,x)/LE(r,x)) \qquad (7)$$

where $LE(f,x)$ is the forecast lag error and $LE(r,x)$ is the reference lag error.

Finally, we calculated mean errors by streamflow class conditioned on the rising and falling limbs of the hydrograph. A rising limb is defined from hourly data as one where $x_i > x_{i-6}$, where the 6-hour lag is chosen to avoid measurement errors and fluctuations that are apparent at the hourly time scale. 15 streamflow classes were used, from 100 to 1500 m3/s.

3. RESULTS

3.1 Modeling Tools

All models discussed in this work were implemented using the System Dynamics (Forrester, 1994; Sterman, 2000) simulation package VENSIM, from Ventura Systems. This approach has been proposed and used for the simulation of both physical and social phenomena, and provides the user with an analogy of fluxes and storages (with user-specified mathematical relationships) through graphically connected boxes to build the systems of interest.

The models tested during the verification period (2003-2004) are called A12 (an ARIMA (4,1,2) run at the 12-hour interval with manually recorded data), A1 (an ARIMA (7,1,3) run at the 1-hour interval with telemetric data), and K (for the Kalman Filter Model). Because K requires an error structure for the precipitation forecasts, and because such data were not available (as there are as yet no operational quantitative precipitation forecasts for the region), we performed instead a sensitivity analysis by running all simulations with 6 different pre-assigned error structures and their corresponding conditional probability matrices. These matrices are assumed to be the same for 0-6, 6-12, 12-18 and 18-24 hour forecasts.

Using the error matrices, we artificially introduced random errors into the historical precipitation data thereby synthesizing less-than-perfect precipitation forecasts. For each streamflow forecasting lead time (12, 24, 48, 60 and 72 hours) and conditional probability of error matrix we generated 50 synthetic sequences of precipitation forecasts and took their mean as representative of the corresponding error structure. The conditional probability matrices used are:

$$\mathbf{M0} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{M0.9} = \begin{bmatrix} 0.9 & 0.05 & 0.01 & 0.005 & 0.005 \\ 0.05 & 0.9 & 0.04 & 0.015 & 0.015 \\ 0.03 & 0.03 & 0.9 & 0.03 & 0.03 \\ 0.015 & 0.015 & 0.04 & 0.9 & 0.05 \\ 0.005 & 0.005 & 0.01 & 0.05 & 0.9 \end{bmatrix},$$

$$\mathbf{M0.8} = \begin{bmatrix} 0.8 & 0.1 & 0.03 & 0.01 & 0.01 \\ 0.1 & 0.8 & 0.07 & 0.04 & 0.04 \\ 0.05 & 0.05 & 0.8 & 0.05 & 0.05 \\ 0.04 & 0.04 & 0.07 & 0.8 & 0.1 \\ 0.01 & 0.01 & 0.03 & 0.1 & 0.8 \end{bmatrix}, \mathbf{M0.6} = \begin{bmatrix} 0.6 & 0.1 & 0.03 & 0.01 & 0.01 \\ 0.2 & 0.6 & 0.08 & 0.04 & 0.04 \\ 0.15 & 0.25 & 0.6 & 0.05 & 0.05 \\ 0.04 & 0.04 & 0.23 & 0.6 & 0.3 \\ 0.01 & 0.01 & 0.06 & 0.3 & 0.6 \end{bmatrix},$$

$$\mathbf{M0.4} = \begin{bmatrix} 0.4 & 0.1 & 0.1 & 0.1 & 0.1 \\ 0.2 & 0.4 & 0.2 & 0.2 & 0.1 \\ 0.2 & 0.2 & 0.4 & 0.2 & 0.2 \\ 0.1 & 0.2 & 0.2 & 0.4 & 0.2 \\ 0.1 & 0.1 & 0.1 & 0.1 & 0.4 \end{bmatrix} e \qquad \mathbf{M0} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The standard interpretation of the matrix elements is $p(f_i|x_j)$, which means: the probability of generating a forecast of class i when the observed precipitation class is actually j. The sequence **M1...M0.4** represents a sequence of increasing errors, from 100% hits to 40% and **M0** represents a situation where forecasts are always null. The observed MSE of the precipitation forecasts for the 6-hour precipitation lead time were 0.97, 8.1, 17.1, 22.5, 91.9 and 11.8 (mm)^2. Notice how, since we are only employing a class forecast, the **M1** error matrix does produce a non-null MSE. The corresponding simulation results are tagged as K1, K0.9, K0.8, etc.

3.2 Accuracy Analysis

Table 1 gives the calculated MSE's and skill scores for all models employed. Only the MSE's and MSE^{-}'s are shown, as the MSE^{+}'s can be calculated by $MSE - MSE^{-}$. The reference value of the calculation of skill scores is A12 (therefore, SS for A12 is 0 by definition). For A12, the following values were observed for the ratio $RMSE$ / Average Observed Flow at the 5 forecasting lead times: 0.03, 0.08, 0.13, 0.17, 0.22 and 0.25.

A1 has a marginally better accuracy (smaller MSE) than A12 at 24, 36, 48, 60 and 72-hour lead time. This may have to do with noisy 1-hour streamflow data; however, the difference is hardly noticeable. The Kalman Filter Models have a distinctly better performance over the ARIMA models when they are given good precipitation forecasts, but even with poor or absent precipitation forecasts they give significantly smaller MSE^{-}. For example, SS_{MSE-} = 58% for K1 and 32% for K0.6 at the 36-hour lead time.

Accuracy Measure	Lead time (h)	Model							
		A1	A12	K1	K0.9	K0.8	K0.6	K0.4	K0
MSE $(m^3/s)^2$	12	171	161	154	180	202	215	253	218
	24	913	916	624	812	956	1051	1293	995
	36	2293	2351	1480	1959	2286	2514	2908	2438
	48	4139	4259	2899	3631	4150	4559	5195	4408
	60	6325	6533	4883	5646	6206	6634	7334	6629
	72	8742	9027	7215	7916	8469	8881	9617	9031
SS_{MSE} (percentage)	12	-7	0	4	-12	-26	-34	-57	-36
	24	0	0	32	11	-4	-15	-41	-9
	36	2	0	37	17	3	-7	-24	-4
	48	3	0	32	15	3	-7	-22	-4
	60	3	0	25	14	5	-2	-12	-1
	72	3	0	20	12	6	2	-7	0
MSE^- $(m^3/s)^2$	12	114	116	78	90	101	106	130	117
	24	649	673	295	381	456	483	661	583
	36	1645	1707	711	937	1116	1163	1473	1574
	48	2938	3014	1547	1860	2128	2186	2682	2939
	60	4417	4503	2900	3189	3462	3487	4089	4521
	72	6029	6086	4547	4753	4985	4948	5618	6232
SS_{MSE^-} (percentage)	12	1	0	32	22	12	8	-13	-2
	24	4	0	56	43	32	28	2	13
	36	4	0	58	45	35	32	14	8
	48	3	0	49	38	29	27	11	2
	60	2	0	36	29	23	23	9	0
	72	1	0	25	22	18	19	8	-2

Table 1: Accuracy measures of simulated models for 12, 24, 36, 48, 60 and 72-hour lead time

Figures 2 and 3 show the skill score for different forecasting lead times of all models studied.

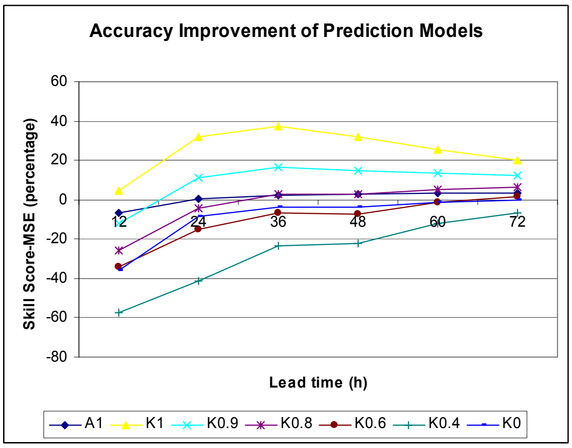

Figura 2: Accuracy improvement of simulated models for 12, 24, 36, 48, 60 and 72-hour lead time

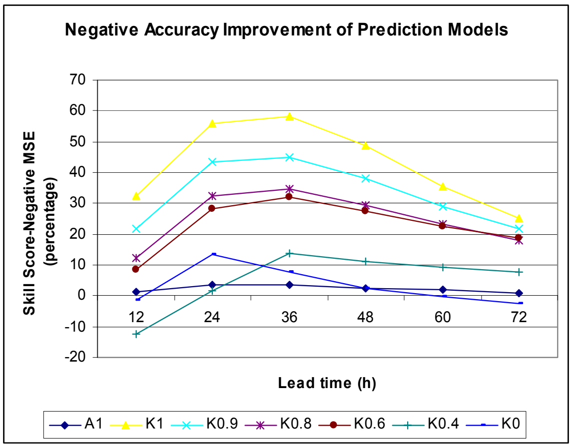

Figure 3: Negative accuracy improvement of simulated models for 12, 24, 36, 48, 60 and 72-hour lead time

In order for significant improvements in the forecasting skill to be obtained by the Kalman Filter models, it is necessary to have very good precipitation forecasts (less than 10% error by class). For the negative accuracy improvement, however, the Kalman Filter models perform much better, even when there are significant errors in the precipitation forecasts (cases K0.8 e K0.6). The best comparative performances over the hourly ARIMA model are at the 36-hour lead time.

3.3 Analysis of Lag Errors

Table 2 shows the values of *LE* errors in hours (defined in section 2.4, equations (5) and (6)) of the different models. The lag errors of the ARIMA models are significantly higher than those of the Kalman Filter models. A1 is able to improve the lag time errors up to 20% over A12, at the 36-hour lead time. The Kalman Filter models display more dramatic improvements, again even in the presence of significant precipitation forecast errors, larger than 50% in some cases. This is shown graphically in Figure 4.

Lag Error	Lead time (h)	Model							
		A1	A12	K1	K0.9	K0.8	K0.6	K0.4	K0
LE (h)	12	4	5	3	3	3	3	4	4
	24	10	12	5	7	7	7	9	9
	36	16	20	9	9	11	11	14	13
	48	23	26	14	14	15	15	19	21
	60	28	33	21	21	21	21	26	29
	72	33	40	28	26	26	26	32	36
SS_{LE} (percentage)	12	17	0	46	36	29	28	18	14
	24	14	0	54	44	37	36	22	20
	36	20	0	55	52	46	45	30	31
	48	11	0	46	45	42	41	27	20
	60	16	0	37	38	36	37	22	14
	72	17	0	30	34	34	35	19	10

Table 2: Lag error measures of simulated models for 12, 24, 36, 48, 60 and 72-hour lead time

Forecasting models that use telemetric data obtained at 1-hour intervals show significant improvement over A12, but the improvement peaks at the 36 hour lead time and decreases thereafter. Thus, the automated hydrological network is most beneficial for forecasts at this lead time. Also notice that K0.9, K0.8 and K0.6 have smaller lag errors than K1 at the 72-hour lead time. This apparently paradoxical result is due to the fact that **M0.9**, **M0.8** and **M0.6** overestimate precipitation, thus reducing lag errors and negative mean square errors, at the cost of an overall increase of *MSE*. This brings up the question of relating the quality of precipitation forecasts to that of streamflow forecasts. As shown by Murphy and Ehrendorfer (1987), there is no one-to-one relationship between, for instance, the MSE of precipitation and the MSE of streamflow. What really happens is that these relationships are always multidimensional and quality can not be represented by just one measure.

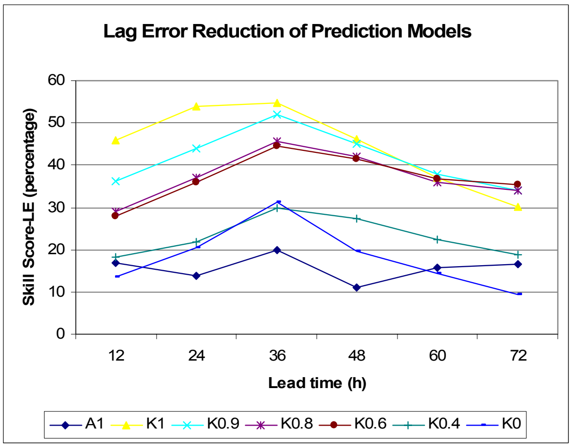

Figure 4: Lag error reduction of simulated models for 12, 24, 36, 48, 60 and 72-hour lead time

3.4 Analysis of Errors during Hydrograph Rise and Fall

Finally, we show in Figure 5 the mean streamflow forecasting errors classified by the rising and falling limbs of the hydrograph. As is to be expected, the forecasting errors during hydrograph rise are predominantly negative (the models fail to predict the correct timing of the hydrograph rise), the opposite happening during hydrological recessions. The Kalman Filter models, all of which use telemetric hourly data, have a significant advantage in this respect. This is particularly important for flood control.

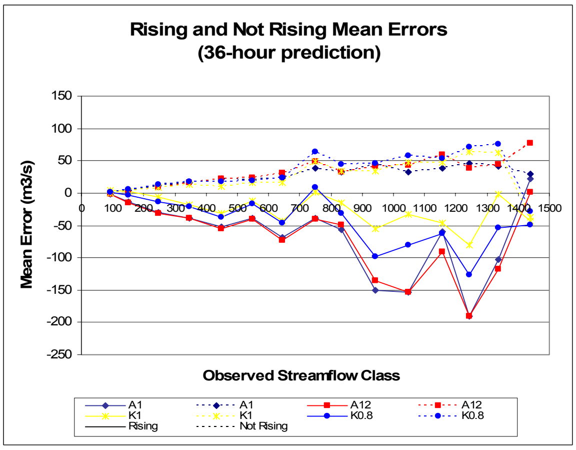

Figure 5: Rising and not rising mean errors of models A1, A12, K1 e K0.8 for 36-hour lead time

4. CONCLUSIONS

This work analyzed the improvements in forecasting skill provided by telemetric data and their use in statistical models that can make use them appropriately. The study site is the Iguaçu River Basin upstream of the City of União da Vitória. We tested different configurations of data availability, including: the baseline of conventional streamflow data, hourly streamflow data, hourly precipitation data and precipitation forecasts by class.

Forecasting quality was assessed with Mean Square Error, and its further classification in positive and negative MSE, Lag Errors (LE) and their associated skills. Mean error during the rising and falling of the hydrograph was also calculated.

ARIMA models using telemetric hourly data show a marginal improvement in overall accuracy (or the order of 3%), but substantial improvement for the lag errors (up to 20%). No significant patterns were found with respect to the forecast lead time.

Kalman Filter models show substantial improvements in accuracy skill (up to 37%), negative accuracy skill (up to 58%) and lag error (up to 55%). However, this is accomplished mainly at the price of having precipitation forecasts available. The precipitation forecast used, however, is relatively simple, consisting of predicting 4 different classes (low, medium, high, extreme). Even the most sophisticated Kalman Filter models still show a pattern of underestimating the rise of the hydrograph, and that is probably the point that deserves most attention in future studies. There is substantial room for improvement in the fields of better precipitation forecasts and rainfall-runoff transformations (either of a statistical of physical nature) that provide better streamflow forecasts in response to either the prediction or the occurrence of precipitation.

5. REFERENCES

Anderson, B.D.O.; Moore, J.B., 2005: *Optimal filtering*. Mineola, USA: Dover Publications.

Bras R.L.; Rodríguez-Iturbe, I., 1993: *Random functions and hydrology*. New York: Dover Publications.

Forrester, J.W., 1994: System dynamics, system thinking, and soft OR. *System Dynamics Review*, Summer, v. 10, n. 2.

Katz, R.W.; Murphy, A.H., 1997: *Economic value of weather and climate forecasts*. Cambridge,UK: Cambridge University Press.

Murphy, A.H.; Ehrendorfer, M., 1987: On the relationship between the accuracy and value of forecasts in the cost-loss ratio situation. *Weather and Forecasting*, v. 2, p. 243-251.

Sterman, J.D., 2000: *Business dynamics*: system thinking and modeling for a complex world. New York: McGraw-Hill Higher Education.