HYBRID FORMULATION FOR TECHNICAL AND NON-TECHNICAL LOSSES ESTIMATION AND IDENTIFICATION IN DISTRIBUTION NETWORKS: APPLICATION IN A BRAZILIAN POWER SYSTEM

Aquiles ROSSONI
UFRGS – Brazil
arossoni@ece.ufrgs.br

Rodrigo TREVIZAN
UFRGS – Brazil
rodtrevisan@ece.ufrgs.br

Arturo BRETAS
UFRGS – Brazil
arturo@ece.ufrgs.br

Daniel GAZZANA
UFRGS – Brazil
dsgazzana@gmail.com

Arlan BETTIOL
NEO DOMINO – Brazil
arlan@neodomino.com.br

Antônio CARNIATO
NEO DOMINO – Brazil
carniato@neodomino.com.br

Luis PASSOS
NEO DOMINO – Brazil
luis@neodomino.com.br

Rodolfo MARTIN
CHESP – Brazil
rodolfo@chesp.com.br

ABSTRACT

This paper presents a hybrid formulation to estimate and identify Technical and Non-Technical Losses (TL and NTL, respectively) in distribution systems. The method proposed is based on the application of unbalanced Weighted Least Squares State Estimation (WLS-SE) and the anomaly detection (AD) technique. First, the method was analyzed in a test system used for assessing NTL detection with labeled data conceived by the authors based on the IEEE 123 Bus Test Feeder and on the behavior of typical Brazilian consumers. After, the proposed method was analyzed in a real feeder from a Brazilian distribution utility. The results show that the proposed method has shown good results for assessing TL and NTL in the synthetic test-system and has shown promising results for the real system application.

INTRODUCTION

Losses in electrical distribution system are a problem that has a huge economic impact in electric utilities. In 2000, transmission and distribution losses were about 16% of total electricity produced in the world, which represents an increase from 1980 levels of 11 % [1]. In Brazil, for example, electric energy losses represent 15 % of total consumed energy [2]. Energy losses are classified as technical (TL) and non-technical (NTL). TL is caused by transmission, transformation and measurement of energy processes. NTL is the energy that is delivered but not billed, mainly caused by thefts, frauds or billing problems. Considering all Brazilian distribution systems, NTL represents about 3 billion U.S. dollars per year [2].

The majority of utilities currently estimate TL and NTL by variation of load flow methods. In [3], load flow is used to estimate TL in each distribution system segment. In this method, the NTL is calculated by the difference between the energy purchased and the sum of TL and energy billed. However, the previously proposed difference is actually an estimation of the TL and NTL caused by them. Also, NTL identification is not performed.

Reference [4] considers distribution systems with Advanced Metering Infrastructure (AMI) to develop on WLS-SE and NTL in each bus are identified using the largest normalized residual test. In these suspected buses, irregular consumers are identified by analysing the daily load variance in meter readings. However, AMI is not presented in typical distribution systems, thus meter readings analysis cannot be performed. Additionally, previous methods do not consider all typical distribution system elements in TL estimation, as in [3].

Considering that AMI is not available in typical distribution systems, this paper proposes a hybrid WLS-SE formulation. The formulation estimates TL and NTL and identify TL. Considering that the largest normalized residual test it's not reliable for this kind of low measurement redundancy, the NTL identification is performed by a statistical parametric anomaly detection technique [5]. The anomaly detection is performed on the record of one year of monthly energy consumption data. Other than pointing out consumers that are more likely to be source of NTL, the anomaly detection method has a second function. The density of anomalies for each transformer of the distribution system is used as additional information to adjust the weight matrix of the WLS-SE method.

PROPOSED METHOD

The proposed hybrid method is based on WLS-SE and AD. In the algorithm, first the AD is performed to identify suspect consumers and show the density of anomalies for each transformer. After, this density is used in the weights of WLS-SE to estimate and identify TL and NTL in the system.

Anomaly Detection

Anomaly detection methods aim at finding patterns in data that do not conform to expected behaviour [5]. As other pattern recognition techniques applied to NTL identification, the objective of AD is recognizing the pattern of irregular consumers in the distribution system data-base. The behaviour of consumers is based on the analysis of energy billing data of consumers, because this data is directly affected by energy theft and fraud or measurement problems.

In the NTL identification problem most consumers have normal behaviour while a small number of them have some kind of irregularity associated to NTL. The hypothesis made while using an AD method is that those irregular consumers are outliers of the data distribution. In order to consider multiple possible irregular behaviours, consumers are divided into groups related with their activities: residential, industrial and commercial.

In this paper, a statistical and parametric AD approach based in a Gaussian model is applied. The method supposes that most consumers' data form a multivariate normal distribution, so the consumers with behaviour close to the distribution mean are considered regular and those with behaviour far from the average are the suspects. This is mathematically modelled by the probability density function (*pdf*) of (1):

$$f(x,\mu,\Sigma)=\frac{1}{\sqrt{2\pi^{d}|\Sigma|}}e^{-\frac{1}{2}\left((x-\mu)'\Sigma^{-1}(x-\mu)\right)} \tag{1}$$

where x is the point in which the *pdf* is computed, μ is the distribution mean, Σ is the covariance matrix and d is the considered number of dimensions. The consumers with low *pdf* beyond a certain threshold are considered anomalous, which is mathematically expressed by (2):

$$f(x,\mu,\Sigma)<\varepsilon \tag{2}$$

where ε is decision threshold. Selecting the threshold involves a compromise between covering all anomalous consumers and correctly classifying those consumers as irregular.

One advantage of the AD method when compared to supervised learning methods is that it can be applied even if the utility does not have previous results of inspections, using only the energy bills. If this information is available, it can be used to improve the method's performance by adjusting the decision threshold using labelled data as a cross-validation set.

The results of this method are a list of suspect consumers and a ratio of suspect consumers and regular consumers.

WLS-SE

The WLS-SE formulation is given by (3). The method objective is to estimate the states $\mathbf{x}$ (voltage magnitudes and phase angles) and, as consequence, estimate the other electrical quantities (power flows and injections) in order to obtain de same value of measured electrical quantities.

$$\mathbf{z}-\mathbf{h}(\mathbf{x}) \tag{3}$$

The measurement vector $\mathbf{z}$ is composed by real feeder measurements and (power flows and voltage magnitudes) and load forecasting of transformers. The estimation vector $\mathbf{h}(\mathbf{x})$ is composed by equations, given by (4)-(7):

$$P_{k-i}=V_{k-i}\sum_{j\in\phi}\sum_{m\in K}V_{m-j}\left(G_{km-ij}\cos\theta_{km-ij}+B_{km-ij}\sin\theta_{km-ij}\right) \tag{4}$$

$$Q_{k-i}=V_{k-i}\sum_{j\in\phi}\sum_{m\in K}V_{m-j}\left(G_{km-ij}\sin\theta_{km-ij}-B_{km-ij}\cos\theta_{km-ij}\right) \tag{5}$$

$$\begin{aligned}P_{km-i}=V_{k-i}\sum_{j\in\phi}\Big(&-V_{k-j}\left(G_{km-ij}\cos\theta_{kk-ij}+B_{km-ij}\sin\theta_{kk-ij}\right)+\\&+V_{m-j}\left(G_{km-ij}\cos\theta_{km-ij}+B_{km-ij}\sin\theta_{km-ij}\right)\Big)\end{aligned} \tag{6}$$

$$\begin{aligned}Q_{km-i}=V_{k-i}\sum_{j\in\phi}\Big(&-V_{k-j}\left(G_{km-ij}\sin\theta_{kk-ij}-B_{km-ij}\cos\theta_{kk-ij}\right)+\\&+V_{m-j}\left(G_{km-ij}\sin\theta_{km-ij}-B_{km-ij}\cos\theta_{km-ij}\right)\Big)\end{aligned} \tag{7}$$

P_{k-i}, Q_{k-i}, P_{km-i} and Q_{km-i} are, respectively, the estimated active and reactive power injections in phase i of node k and power flows in phase i of node k to node m. θ_{km-ij} is the angular difference between the voltage angle of phase i of node k and phase j of node m. ϕ is the set of phases in node k and K is the set of nodes connected to node k, including the node k. G_{ij-km} and B_{ij-km} are elements of bus admittance matrix, $Y_{ij-km}=G_{ij-km}+jB_{ij-km}$.

Considering the real measurements and the load forecasts (pseudo-measurements), the number of measurements is greater than the number of states to estimate. Considering this and that $\mathbf{h}(\mathbf{x})$ is composed by non-linear continuous equations, the system presented in (3) can be solved iteratively by the WLS using normal equations.

Therefore, an initial condition is considered (voltage of all buses equals to reference voltage), in each iteration v, the system shown in (8) is solved to obtain the states variations $\Delta\mathbf{x}^{v}$ and the states are updated by (9). The iterative process finalizes when the state variations are lower than a given tolerance.

$$\left(\mathbf{H}\left(\mathbf{x}^{v}\right)^{T}\mathbf{W}\mathbf{H}\left(\mathbf{x}^{v}\right)\right)\Delta\mathbf{x}^{v}=\mathbf{H}\left(\mathbf{x}^{v}\right)^{T}\mathbf{W}\left[\mathbf{z}-\mathbf{h}\left(\mathbf{x}^{v}\right)\right] \tag{8}$$

$$\mathbf{x}^{v+1}=\mathbf{x}^{v}+\mathbf{\Delta x}^{v} \tag{9}$$

where $\mathbf{H}$ is the matrix of partial derivatives of $\mathbf{h}(\mathbf{x})$ and $\mathbf{W}$ is weight matrix. $\mathbf{W}$ is a diagonal matrix, which each element are the inverse of measurement covariance, in other words, $w_{ii}=\sigma_i^{-2}$, where σ_i is the standard deviation of measurement z_i.

Usually, for real measurements, it is made a relation between the standard deviations and the meter accuracy. For pseudo-measurements, normally, it is considered that all weights are equals, because they are given for the same load forecasting procedure.

The estimation of TL in system segments (lines and copper of transformer) are given by the difference between the energy injected and delivered by the segment, as shown in (10). This is show as sum in (10), because the flows *k* to *m* and *m* to *k* have opposite directions. Some TL are considered constant, for example, the loss of energy meters and the loss on the iron of transformers, so they are modeled as a load in the state estimation. The estimation of NTL in each bus is given by the difference between the power estimated and the load forecasted (pseudo measurement) for this transformer, as is given in (11).

$$TL_{km} = \sqrt{\left(\sum_{i\in\phi}\left(P_{km-i}+P_{mk-i}\right)\right)^2 + \left(\sum_{i\in\phi}\left(Q_{km-i}+Q_{mk-i}\right)\right)^2} \quad (10)$$

$$NTL_{k} = \sqrt{\left(\sum_{i\in\phi}\left(P_{k-i}+P_{k-i}^{forecasted}\right)\right)^2 + \left(\sum_{i\in\phi}\left(Q_{k-i}+Q_{k-i}^{forecasted}\right)\right)^2} \quad (11)$$

Hybrid Formulation

Instead of using the same weights for all pseudo-measurements, this work proposes the use of the results of AD to weight the pseudo-measurements of each distribution transformer. Considering that load forecasting is based on the energy bills of consumers connected to a transformer, it is clear that the accuracy of this pseudo measurement is related to the accuracy of the information represented by these energy bills.

In the AD, as results, it has obtained a ratio of suspect consumers and regular consumers in each transformer *k* $\left(r_k\right)$. The hybrid formulation proposes use this result and the value of each pseudo-measurement to compute the weights. In (12), is given the weight for pseudo-measurement z_i. In other words, the more suspects a transform have, lower is his weight.

$$w_{ii} = c\frac{1}{z_i\left(0.001+r_k\right)^{-2}} \quad (12)$$

Using AD, some transformers have no suspects and $r_k = 0$, therefore, to avoid numerical problems, it considered a minimum value of 0.001 (0.1 % of suspects). c is a constant to avoid that pseudo-measurements with small magnitude and small r_k obtain weights higher than real measurements weights. As can be seen, the proposed method uses a heuristic way to consider the weights, but based on the statistical analysis of AD.

DESCRIPTION OF THE TEST SYSTEM

A test-system was developed to evaluate the proposed methodology. The system is based on the IEEE 123 bus test feeder (http://ewh.ieee.org/soc/pes/dsacom) considering each spot load as the sum of residential consumers whose mean power distribution and frequency is described in [6], resulting in more than twelve thousands of consumers. The seasonality was represented by changes on monthly consumption and the considered variation was based on the behavior of residential consumers of a Brazilian utility. A twelve-month record of monthly energy consumption is considered.

It was considered a number of eight real measurements to perform the WLS-SE, a plausible number for typical Brazilian distribution systems. Additionally, the system was considered with four geographical areas. The measurements points and locations are demonstrated in Fig. 1.

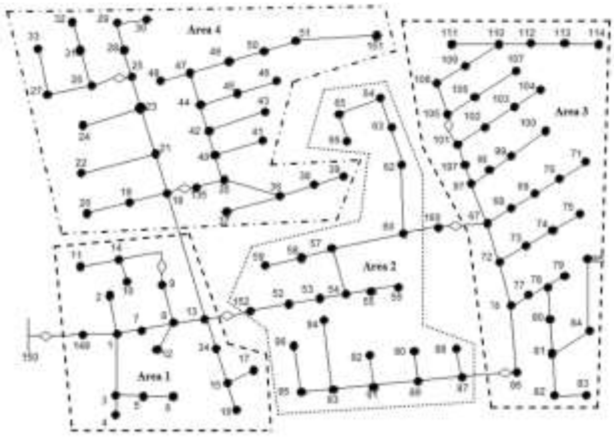

Figure 1. System based on IEEE 123 bus test feeder with fourth areas and eight meters (represented by diamonds).

CASE STUDY USING THE TEST SYSTEM

A base case was created to simulate the real values. In this case, a load flow was performed for each month with no error in forecasted load. These results are called here ´reference values´.

After, NTL was added in the customers. NTL was considered as fraud, given by total or partial bypass of energy meters, resulting in reduction of consumer energy bills and, as consequence, reduction of forecasted loads.

The considered NTL represents 7 % of total energy and is given by reduction of energy bills. The consumers with NTL were chosen randomly, but considering that 99 % of consumers with NTL are located in the first twelve buses (concentrated in Area 1). The level of energy bill reduction is given randomly for 50 % of these illegal consumers, for the others 50 % is considered a reduction to zero. The month that consumer fraud begins is given randomly and remains for all subsequent months.

The AD was performed considering the twelve months of consumer energy bills with NTL and it was used the parameter $\varepsilon = 0.0001$. Data from consumers is normalized by dividing the energy consumption for each month by the mean consumption in the period that is considered. A linear transformation is applied to the value of the *pdf* so that it ranges from 0 to 1.

The AD method correctly classified 98.35% of the clients. Among those considered anomalous, 77.54% were precisely labeled as irregular consumers. It means that 84.03% of total irregular consumers could be located in a single iteration of the method for the test system.

The WLS-SE considered the energy bills with errors and the real measurements, given by the reference values. The weights for real measurements were considered as $w_{ii} = 1000$ and the pseudo-measurement constant adopted was $c = (10000)^{-1}$.

Fig. 2 shows the reference values and the method estimation of TL. The figure shows the average monthly energy loss in all line segments of the test system. Each number in the horizontal axis corresponds with a test system line segment.

Fig. 3 shows the reference values and the method estimation of NTL. The figures show the average monthly energy loss in all line nodes of the test system. Each number in the horizontal axis corresponds to test system nodes.

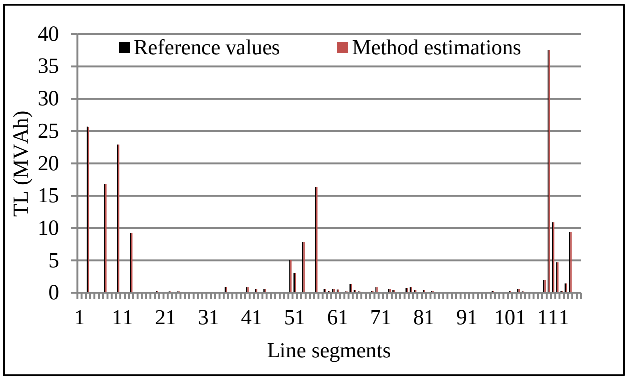

Figure 2. Reference and estimation values of TL in line segments of test system.

The proposed method had a good performance in TL and NTL estimation, as can be seen in the figures. The mean and maximum errors for TL estimation were 0.002 MVAh and 0.13 MVAh, respectively. For the NTL estimation, the mean and maximum errors were 0.12 MVAh and 4.29 MVAh.

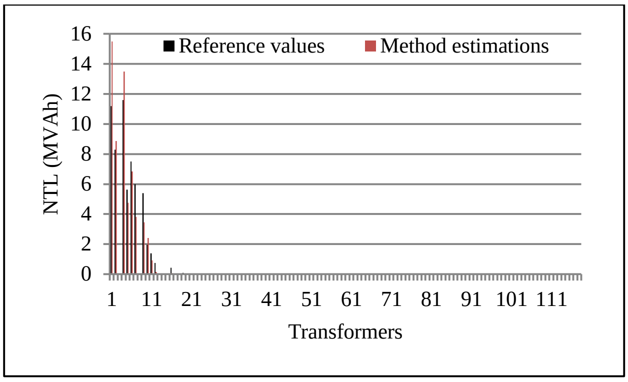

Figure 3. Reference and estimation values of NTL in test system nodes.

APPLICATION ON A REAL FEEDER

The method was tested in a 13.8 kV Brazilian radial distribution system with urban and rural areas, three-phase and single-phase lines, more than 200 transformers and 3,000 consumers. TL and NTL estimation and identification results are presented. At the moment of the tests, just one real measurement was available (located in the main feeder – substation of energy supply). Thus to perform the WLS-SE it was considered pseudo-measurements in all system transformers based on load forecasting.

The feeder has 3,796 consumers. After a process of data cleaning, 3560 of them were analysed and 390 (10.9%) of them were considered anomalous. Those anomalous clients are connected to 67 different transformers, which represents 29.9% of the total 224 transformers considered by the AD method.

Fig. 4 shows the results of TL for line segments, each number in the horizontal axis represent a line segment of the real feeder. Fig. 5 demonstrates the TL estimation in transformers, it is considered a sum of copper and iron losses and all connected energy meters losses, each number represents a real transformer of the feeder. In Fig. 6 is shown the results of NTL in each real transformer of the feeder and Fig. 7 presents the total feeder losses. All values are based on monthly average energy loss.

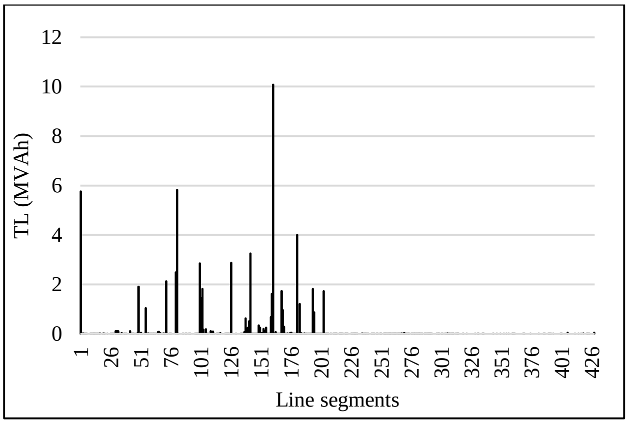

Figure 4. TL in line segments of real feeder.

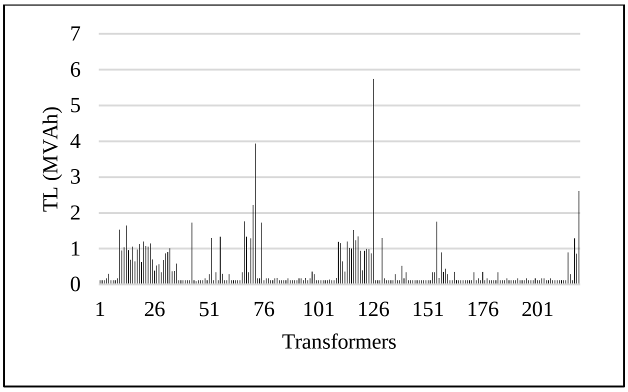

Figure 5. TL in real feeder transformers.

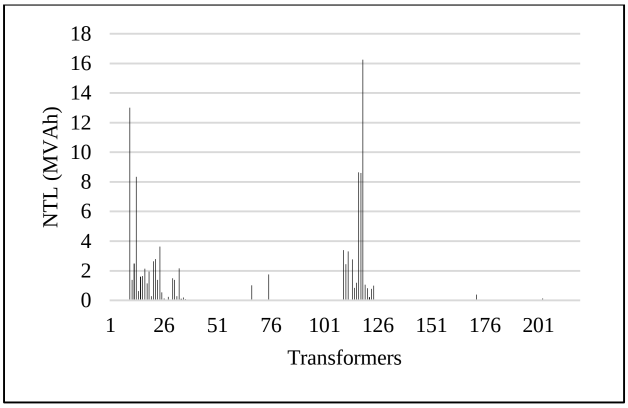

Figure 6. NTL in real feeder transformers.

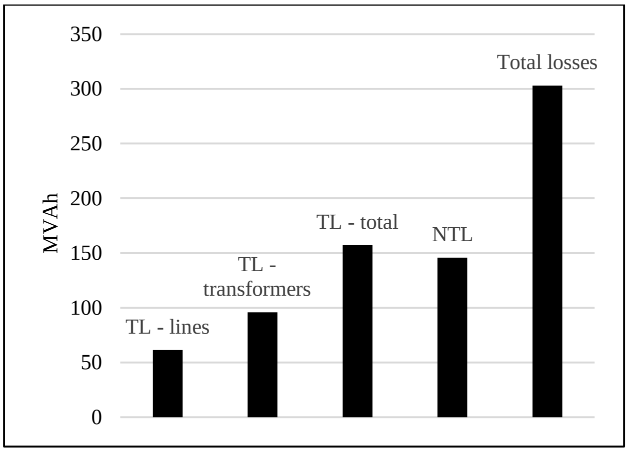

Figure 7. Total losses in real feeder.

As can be seen by Fig. 4 to Fig. 6, the proposed methodology identify the segments with more TL and the transformers with significant NTL, assisting in the decision to be taken to reduce the losses. Additionally, Fig. 7 shows that the results of TL and NTL can be compared, assisting in deciding which kind of loss must be combated before.

CONCLUSIONS

A novel method for TL and NTL estimation in distribution systems is presented. The results of the application of the proposed method to a test and real distribution system have shown good results for both NTL and TL estimation and irregular consumers identification. The additional information provided by the AD method to the state estimator resulted in a decrease in the TL error estimation, which represents an improvement on the losses estimation. For this test system, the performance of the AD method for NTL identification has also shown satisfactory results.

The authors expect that the proposed method contributes to improving the efficiency of power systems by indicating where actions can be taken by utilities to reduce both TL and NTL.

ACKNOWLEDGEMENTS

The authors wish to thank the Federal University of Rio Grande do Sul (UFRGS), the Brazilian National Council for Scientific and Technological Development (CNPq) and Brazilian Federal Agency for the Support and Evaluation of Graduate Education (CAPES). Additionally, the authors thank the companies NEO DOMINO RESEARCH IN POWER SYSTEMS, CHESP, CERRP, CERPRO, CERNHE, CERIPA, CERAL-DIS, CETRIL, CERIM, CERMC, CERIS, CEDRI, CERES, CEDRAP, ELFSM, EFLJC, COOPERALIANÇA and CERCOS for the financial support regarding the development of this work through an R&D project for the Brazilian Regulatory Agency (ANEEL) (Project R&D 0103-0002/2011).

REFERENCES

[1] T. Smith, 2004, "Electricity theft: a comparative analysis", *Energy Policy* , vol.32, 2067-2076.

[2] R. Vidinich and G.A.L. Nery, 2009, "Research and development against energy theft", (in Portuguese), *Research and Development Magazine of ANEEL (Brazilian Electricity Regulatory Agency)*, vol.3, 15.

[3] C.C.B. de Oliveira, N. Kagan, A. Meffe, S. Jonathan, Sunny, S.L. Caparroz, and J. L. Cavaretti, 2001, "A new method for the computation of technical losses in electrical power distribution systems", *Proceedings International Conference on Electricity Distribution*, vol.5, 328.

[4] S. Huang, Y. Lo and C. Lu, 2013, "Non-technical loss detection using state estimation and analysis of variance," *IEEE Trans. Power Syst.*, vol.28, 2959-2966.

[5] V. Chandola, A. Banerjee and V. Kumar, 2009, "Anomaly detection: a survey," *ACM Computer Surveys*, vol.41(3), 1-72.

[6] J. A. Jardini and R. P. Casolari. (2014, Nov. 14). *Customers Load Curves and Applications in Distribution Engineering* (in Portuguese). [Online]. Available: http://social.stoa.usp.br/jajardini/livros.