A NEW APPROACH FOR LOAD FORECASTING: PREDICTING THE CAUSES, FORSEEING THE CONSEQUENCES

Leontina PINTO	Jacques SZCZUPAK	Luiz H.de MACEDO	José R.PASCON
Engenho – Brazil	Engenho – Brazil	Engenho – Brazil	Elektro – Brazil
leontina@engenho.com	jsz@engenho.com	lhmacedo@engenho.com	jose.pascon@elektro.com.br
Robinson SEMOLINI	Márcia INOUE	Carlos ALMEIDA	Fernão R. ALMEIDA
Elektro – Brazil	Elektro – Brazil	Independent Consultor	Independent Consultor
robinson.semolini@elektro.com.br	marcia.inoue@elektro.com.br	cal.alm@gmail.com	fernaora@terra.com.br

ABSTRACT

Present energy market dynamics seriously affects data stationarity. Similar situation followed the 2001 Brazilian energy shortage, producing strong modifications in load dynamics. In this case, new history corresponds to a reduced period – a few years - leaving insufficient amount of data to be explored by classic models, from statistical to neural networks. This paper addresses this problem – modeling under lack of data – and proposes a new method based on Functional Analysis, applied as a sequential procedure. A real case-study enlightens the approach advantages.

INTRODUCTION

Modern energy markets exhibit a nervous behavior, reflecting present world's political and industrial instabilities. Load models are essential for energy trading but, even under today world's instabilities, most of them are still based on statistical analysis, using large utility data banks. Naturally, during the period of data collection, the electrical market is supposed stable – in other words, the data processes are stationary. This presumed situation may be invalid due to the energy market dynamics, or to local causes, as recently occurred in Brazil.

The Brazilian power market is strongly dependent on hydroelectric energy and became particularly vulnerable during the Brazilian 2000-2001 drought. Water reservoirs came to as low as 20% of their full capacities, requiring extreme limitations on the use of energy. A severe shortage was implemented, leading to a mandatory 20% linear load reduction to all consumers (residential, commercial and industrial). Many industries modified their energy consumption profile, adapting to the use of other available energy sources, as solar, gas or oil. The society as a whole minimized consumption by the use of low energy electronic lamps, sun heating, etc, sometimes disconnecting unnecessary and even necessary electrical loads. Penalties for excess energy consumption were severe, but successful reduction was rewarded by credits for future energy bills.

The shortage extended from May 2001 to January 2002, but the corresponding fall in the electrical energy market still persists. It may be observed in Fig.1 the slow load recovery after the 2001 sharp down transition. Of course, many changes to alternative energy sources are definitive, but it looks like it is more than this.

Actually, the average Brazilian worker income was already suffering a substantial reduction before the energy crisis started. The energy crisis taught him how to manage and control his electrical energy bill, partly compensating his income losses. Besides the permanent modifications on energy sources, the society electrical energy consuming habits were modified, producing a clear and probably permanent change on load consumption dynamics.

Fig. 1 displays the load curve of Elektro (Brazilian utility, responsible for supplying part of S. Paulo State) – normalized, for better visualization. Load dynamics dramatically changes after the 2001 drop, and new consumption pattern seems differ from the old one. The energy market during 2002 is visually still adapting and only from 2003 on it looks as "almost" stabilizing.

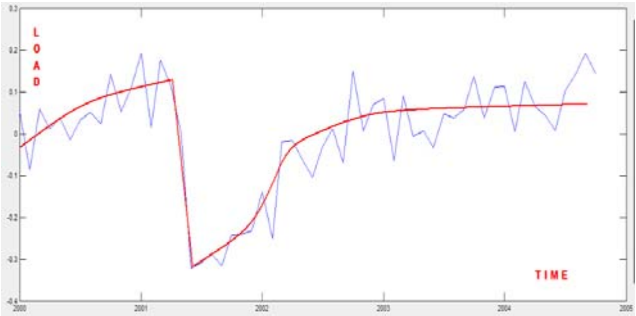

Figure 1 – Elektro Total Load

This leads to insufficient amount of data in order to work statistically. Conventional statistical methods or even neural networks, as [6-8], require large volumes of data that will not be available for a while. Attempts to use the pre 2001 data to statistically analyze the present market will take into account an obsolete past and possibly lead to poor results.

In spite of all difficulties, new market rules bring a new challenge: distribution utilities must cover 97% of their load with long-term contracts (or suffer a huge penalty). The Long Term Load Forecast problem has never been more difficult, critical and important to solve.

This paper presents a functional analysis approach, originally applied in signal theory, customized to this problem. We focus the Load Decomposition by different "explaining variables", describing load behavior and dynamics. The load consumption synthesis is addressed at the end of the work.

SOME FUNCTIONAL ANALYSIS RESULTS

Functional Analysis [1-3] underlies many optimization procedures. Sometimes it is statistically based, as in many communication problems [1], but it may be also used deterministically [1-3], associated to specific Hilbert Spaces.

Hilbert Space

It is common practice to refer to the members of the Hilbert Space to be studied as vectors, disregarding the fact that they are, as in this case, data sequences, such as loads, temperatures, economic indexes, etc.

A Hilbert Space is a complete metric space [1-3], composed by reasonably smooth data sequence vectors, here represented by capital bold fonts as **V**. In the considered Hilbert Space:

- the distance $d[\mathbf{V}_1, \mathbf{V}_2] = ||\mathbf{V}_1 - \mathbf{V}_2||$, is a nonnegative real number, where || . || stands for the vector norm,
- the vector inner product $(\mathbf{V}_1, \mathbf{V}_2) = \mathbf{V}_1^t\mathbf{V}_2$ induces the vector norm, $\mathrm{sqrt}(\mathbf{V}_1^t\mathbf{V}_2)$, the Euclidean norm [1-3].

The vector norm may be understood as a measure (functional) of the vector size, while the distance between vectors is the size of their difference.

This ideal “Load Hilbert Space” does not actually exist for the load problem. The completeness of the ideal space would always require zero residue for any load decomposition scheme, an unpractical situation. A “close enough” engineering criterion avoids the complete metric space requirement, still preserving the desired load decomposition properties.

Schwarz inequality [1-3] states that $|(\mathbf{V}_1,\mathbf{V}_2)| \le ||\mathbf{V}_1||.||\mathbf{V}_2||$, or equivalently for a real vector space, as in this case,

$$\cos\theta = \frac{(\mathbf{V}_1,\mathbf{V}_2)}{||\mathbf{V}_1||.||\mathbf{V}_2||}, \qquad (1)$$

where θ is the angle between vectors $\mathbf{V}_1$ and $\mathbf{V}_2$.

Hilbert Space optimum solution satisfies the Projection Theorem or, equivalently, the Orthogonality Principle, a geometric interpretation for the optimality condition.

Projection Theorem

If one looks for the optimum approximation of a desired vector, $\mathbf{V}_d$ by a set of other space basis vectors, the Projection Theorem states that the error vector is orthogonal to the space determined by the basis vectors, meaning Eq.1 results in a zero cosine for the error vector and any other space vector.

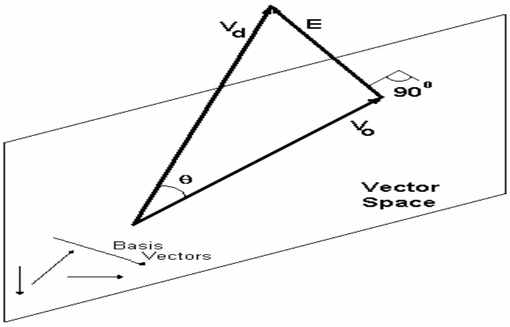

Figure 2 – Projection Theorem Representation

This is represented in Fig.2, where $\mathbf{V}_d$ is optimally approximated by $\mathbf{V}_o$, a vector belonging to the Vector Space. The error vector **E** is shown orthogonal to the basis vector space. In our case the initial desired vector is the sequence of loads and the basis vectors are originated from economic, climate and demographic series, etc.

LOAD DECOMPOSITION

Optimum load decomposition follows directly from the Projection Theorem if the basis vectors are known beforehand, as in classical communications with sinusoids or in modern spread spectrum techniques as in CDMA systems [1,4,5]. In this case, however, the basis sequences still must be found and they may vary from utility to utility. Attending to this fact, we will follow an equivalent alternate procedure, sequentially removing the component due to each of the basis vectors. This approach is equivalent to the originally exposed, but will show better to visually estimate the process evolution and helps finding the basis sequences.

The Sequential Algorithm may be described as

1. Iteration i = 1
2. Select a desired vector, $\mathbf{V}_{di}$.
3. Select a basis vector, $\mathbf{V}_{bi}$.
4. Determine the optimum projection, $\mathbf{V}_{oi}$.
5. Determine the error vector, $\mathbf{E}_i$.
6. If $||\mathbf{E}_i||$ too large; $\mathbf{V}_{di+1} = \mathbf{E}_i$; i=i+1; return to 3.

The decomposition is accomplished after a few iterations and by proper selection of basis vectors. Notice item 6 contains the previously mentioned “close enough” engineering criterion.

CASE STUDY

The proposed procedure will be illustrated by the Elektro case study. The monthly load to be decomposed is shown in Fig.3 (mean value removed and normalized with respect to medium value).

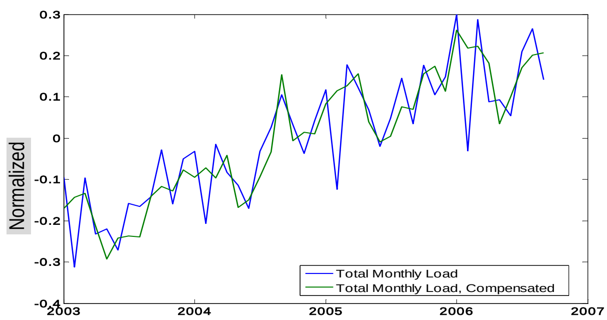

Figure 3 – Elektro Monthly Load and Compensated Effective Load

Care should be taken when examining the monthly values. For instance, February loads are usually lower than those of January, but this could be due to the reduced number of days rather than to consumption reduction. Actually, load pattern depends if it is Sunday, Saturday, or a Holiday. In this last case depends also on its weekday occurrence. These characteristics were modeled by transforming the Monthly Load curve into a Compensated Effective load curve where each day is weighted according to its share in total consumption (in this case, for instance, Holidays and Sundays account for 60% of a weekday).

The difference between Monthly and Effective load may be observed in Fig. 3 – February "valleys", for instance, are now corrected. We will focus therefore on the Compensated Load and take it as, $\mathbf{V}_{d,}$ the actual starting point of the process.
According to the Sequential Algorithm; i=1, $\mathbf{V}_{d1}$<= Effective Load. Next step is choosing the first basis vector. The first component (basis vector) to be removed from the Effective Load curve, $\mathbf{V}_{b1,}$ is the economic influence. It is shown in [9] that, although load elasticity with respect to prices or incomes is very high, there is still a significant influence from the Brazilian National Internal Gross Product, as may be seen in Figure 4 normalized plots.

Fig. 4 – Normalized Compensated Load (green) and Brazilian GIP (blue)

The next step in the proposed algorithm is to apply the Projection Theorem, determining the cos θ_1 between both vectors; $\mathbf{V}_{d1}$.and $\mathbf{V}_{b1}$. This follows immediately from Eq. (1), a consequence of Schwartz Inequality.
In the sequence, the projected optimum vector, i=1,

$$\mathbf{V}_{oi} = \underbrace{\|\mathbf{V}_{di}\|\cos\theta_i}_{size}\ \underbrace{\frac{\mathbf{V}_{bi}}{\|\mathbf{V}_{bi}\|}}_{unit\ vector} \qquad (2)$$

is removed from $\mathbf{V}_{di}$ in order to yield the error vector $\mathbf{E}_i$

$$\mathbf{E}_i = \mathbf{V}_{di} - \underbrace{\|\mathbf{V}_{di}\|\cos\theta_i \frac{\mathbf{V}_{bi}}{\|\mathbf{V}_{bi}\|}}_{\mathbf{V}_{oi}} \qquad (3)$$

Fig. 5 displays the normalized curves of Elektro Load and the error after removal of the GIP projection. The removed component energy totals 87% of the original load, as indicated in Fig.4.

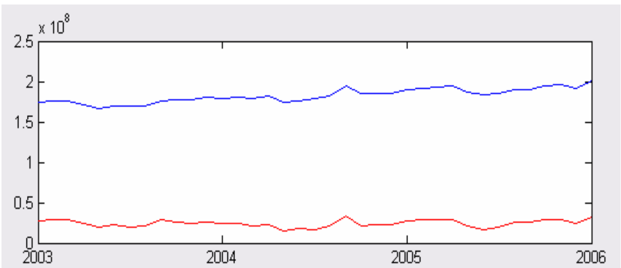

Figure 5 – Elektro Load (blue) and first Error Vector (red)

As the i=1 error is not negligible, one follows the Sequential Algorithm by updating the desired vector with the previous error vector or residue of the operation, $\mathbf{V}_{d2}$= $\mathbf{E}_1$ and the index i=i+1=2. A new explanation vector (basis vector) is chosen and the procedure is repeated.
The next explanation sequence comes from climate: we will now take the monthly averages of maximum temperatures in the utility's area.
As Elektro facilities spread over a large area, we took, for this example, temperature "geographical" mean values. Further studies, focusing specific regions, could yield even more precise results. Fig. 6 displays the normalized new desired vector superposed to the normalized explanation temperature vector.

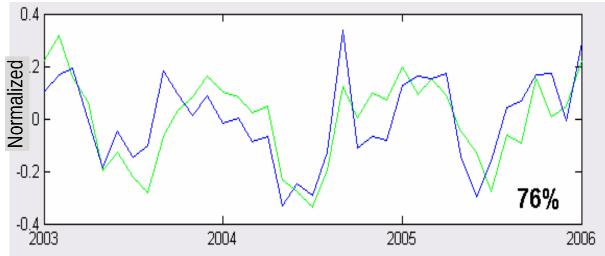

Figure 6 – Error 1 Load (blue) and Temperature Vector (green)

The new projection energy is, as indicated, 76% of the first error energy, representing the actual load levels indicated in Fig.7.

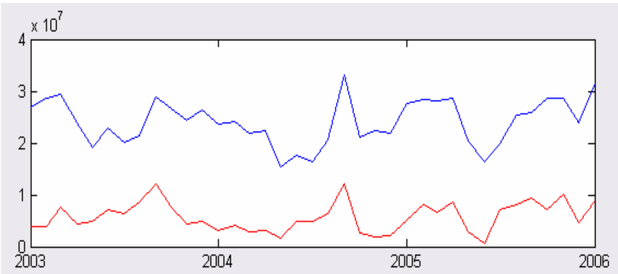

Figure 7 – Error 1Load (blue) and Error 2Vector (red)

One observes that although achieving a strongly reduced residue at this point, the error curve indicates the presence of seasonality. One could proceed with the method, trying to further reduce the decomposition residue, however, due to its reduced energy we preferred just remove its monthly average in the period as a deterministic component (similarly to many statistically based methods). This yearly average load is shown in Fig.8, while Fig.9 indicates it represents 70% of Error 2 energy in the period.

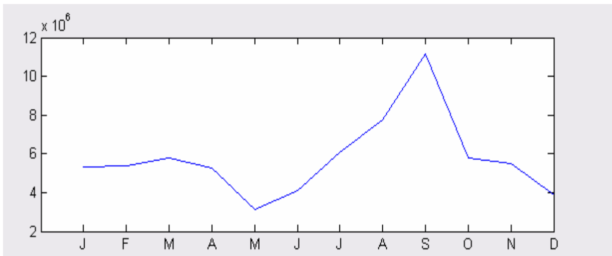

Figure 8 – Error 2Load Yearly Average

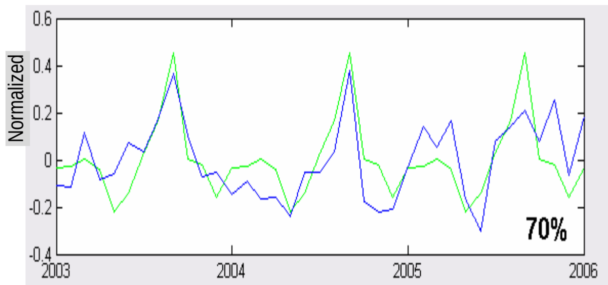

Figure 9 – Error 2Load (blue) and Error 2Yearly Average (green)

The final comparison is made in Fig.10 between the original Elektro and the Error 3 loads.

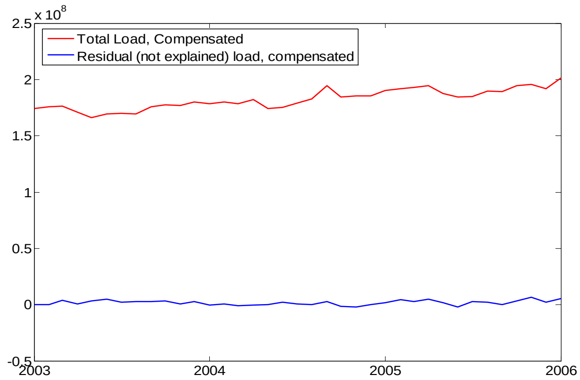

Figure 10 – Elektro Load (red) and Error 3 (blue).

LOAD FORECASTING

Since from the analysis stage resulted a decomposition in terms of explaining variables basis vectors, the synthesis procedure relays on the future load reconstruction based on the same linear combination, but with forecasted basis vectors. This approach is applied for the Elektro case, requiring the forecasting of the Brazilian GIP and of the average S.Paulo temperatures.

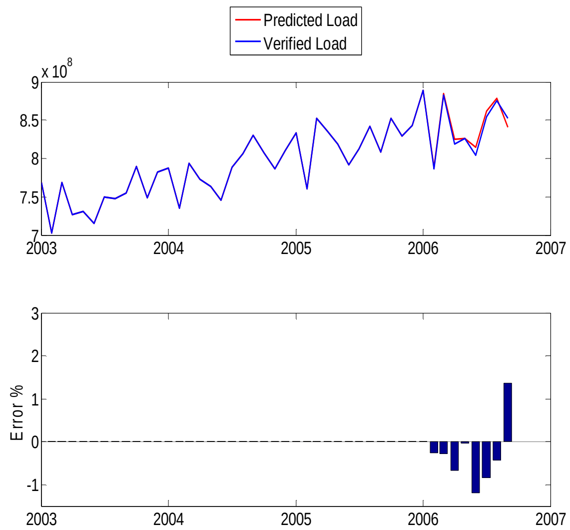

Fig.11 – Predicted and Verified Loads (upper), Percent Errors (lower)

As with neural networks, we consider the derived decomposition as a train and will apply the method from January to September 2006. The results are displayed in Fig.11 for the loads and respective percentual errors.

CONCLUSIONS

The 2001 draught severely affected the Brazilian energy market, modifying energy consumption habits and market characteristics. As a result, the use of pre-2001/2 data is unsafe, leading classical and modern statistical approaches to unacceptable solutions. The new market does not have enough data to permit correct statistical modeling.

The aim of this paper is as the first step to a completely novel Load Forecasting model, able to work under lack of data. This work focuses on the analysis and detection of its explaining short term variables. Load projection is a straightforward step, based on the achieved results. This deterministic type of approach is based on Functional Analysis theory. The Projection Theorem is valid for different Hilbert Spaces, including our deterministic situation. For such a space, a Sequential Algorithm is presented, equivalent to the classic Projection Theorem, but allowing the test of all possible explanation sequences, or basis vectors.

In the included case study, the Elektro load is almost entirely explained by the Brazilian GIP, temperature mean value and a reduced seasonal component. The synthesis stage was tested for a nine month period with almost negligible errors. It is interesting to observe that, although the functional approach is common to the solution process, the employed explanation variables differ from utility to utility, according to their operational regions. The basis vectors do depend on the nature of the load, but the algorithm is general.

REFERENCES

[1] L.E.Franks, *Signal Theory,* Prentice Hall,1969.

[2] Lokenath Debnath and Piotr Mikusinski, *Introduction to Hilbert Spaces with Application*, 1999, Academic Press.

[3] N.I.Akhiezer, I.M.Glazman, *Theory of Linear Operators in Hilbert Space*,1988, Dover

[4] Simon Haykin, *Adaptive Filter Theory* (4th Edition), Prentice Hall, 2001

[5] S. K. Mitra, *Digital Signal Processing – A Computer-Based Approach*, McGraw-Hill, 1998

[6] K. Nagasaka, M. Al Mamun, *" Long-term peak demand prediction of 9 Japanese power utilities using radial basis function networks"*, Power Engineering Society General Meeting, 2004, pages 315 - 322 Vol.1

[7] T.Q.D. Khoa, L.M. Phuong, P.T.T. Binh, N.T.H. Lien, *"Application of wavelet and neural network to long-term load forecasting"*, PwerCon**,** vol. 2004, Volume: 1, pages: 840 - 844 Vol.1

[8] M.S. Kandil,S.M. El-Debeiky, N.E. Hasanien,*"Long-term load forecasting for fast developing utility using a knowledge-based expert system"* IEEE Trans on Power Systems, May 2002, Volume: 17, Issue2, pages 491/496

[9] L. Pinto, J. Szczupak, L.H. Macedo, C. Szczerbacki, J. Pascon, C. Almeida, R. Semolini, *"A New Approach to Load Forecast: Incorporating Climate, Economic and Social Behavior Predictions"*, IEEE Powertech, St. Petersburg, 2005