The use of valuation data bases and analyses and GIS-tools for the assessment of urban infrastructure impact along Line 4 of the São Paulo Subway Company.

Antonio Sergio Liporoni, Civil Engineer, UPAV Past President, Ibape, CTAGEO Engenharia e Geoprocesamento Ltda.

Eduardo Rottmann, M.Sc., Civil Engineer, UPAV Standards Secretary, Ibape, Contacto Consultores Associados Ltda.

Celso de Sampaio Amaral Neto, Civil Engineer, Ibape, Amaral D'Avila Engenharia de Avaliações Ltda.

This paper presents the development and implementation of a socio-economic and real estate market Geographic Information System at *Cia. do Metropolitano de São Paulo*, the São Paulo Subway Company, as well as some of the results obtained. The information layers produced comprise secondary and primary data from various sources, and were used to develop an impact analysis based on both descriptive statistics and hedonic modeling, in order to assess the urban and market impacts of Line 4 in two different moments, 2002 (ex-ante), and 2008, when the line was already under construction. Further measurements are planned to take place in the near future, in special after Line 4 inauguration, scheduled for 2011. The project implied in a survey of circa 8,000 real estate consisted data, as well as in a São Paulo-inedited actual occupation survey, streamlined with IRDB's indicator building requirements.

Summary

1 INTRODUCTION

After more than a decade of delay, the São Paulo State Government, with support of multilateral financial organizations like the World Bank and the Japanese Cooperation Agency, in 2002 managed to start the construction of Line 4 – Yellow, of the São Paulo City Subway System.

This line spans for 13.5 km, through 11 stations, linking downtown Luz district (part of the historic center of the city, under considerable urban blight) to south-western Vila Sônia neighborhood (part of affluent Morumbi district, one of the stems of local real estate market expansion). Many different regions of the city are crossed, some of them highly valued and dynamic, including two large *"Operações Urbanas"* (municipal Urban Projects characterized by special zoning laws): *Centro* (Downtown) and *Faria Lima* (the new financial and business neighborhood), as well as third one (Vila Sônia), still under discussion in City Council.

Chart 1 – Projected São Paulo Metropolitan network expansion.

The study summarized in this paper resulted from data and information gathered in a Geographic Information System, under a bid contracted by the São Paulo Subway Company (Metrô-SP) to *CTAGEO Engenharia e Geoprocessamento*[1]. The main purpose of the study is measuring, through an ex-ante and ex-post comparison, the urban and real estate market impacts resulting from the implementation of Line 4, and was jointly sponsored by Metrô-SP and the World Bank. The key aspects followed were land use and urban occupation; real estate market; housing production and business activity.

[1] The contract was developed through a joint venture with Amaral D'Avila Engenharia de Avaliações and Consult Soluções Patrimoniais. Besides the authors, Mr. Eduardo Fava was also part of the management team.

2 PROJECT DEVELOPMENT

The project was divided into four phases, as presented below:

1. Implementation of the Geographic Information System[2];
2. Real estate values and occupancy surveys;
3. Urban dynamics and real estate market (along Line 4) impact analyses;
4. Real estate market prospective study.

This paper will present aspects and results of the three first phases, under a market approach. The impact analysis done must be regarded as interim, since the survey was done before operation started. The "zero" moment was defined as 2002, shortly before the actual start of the construction works (though after its announcement). Additionally, this was a transitional moment, that coincided with a deep change of city zoning laws, following the passing of the new urban master plan law (*Plano Diretor Estratégico – PDE – Municipal Law 13.430/02*). This survey was conducted in 2008, when most of the line was under well advanced construction works.

3 GEOGRAPHIC SEGMENTATION CRITERIA

The areas along Line 4 were divided into **influence**, **reference** and **control areas**, which were defined over the so called *Zonas O/D* cartographic disaggregation, i.e. the Origin-Destination Zoning used by the Metrô-SP for transportation network modeling. This disaggregation gathers a number of socio-economic data, such as household income, employment breakdown, travel time through public and individual transportation, that are seldom used in valuation analysis, and, as will be shown, are relevant in value composition.

Influence area is subdivided into direct and indirect influence areas, defined as follows:

- **Direct influence area** corresponds to the blocks contained in 600 m radius circles around the stations and up to 300 m of distance from the subway system axis[3];
- **Indirect influence area** corresponds to the blocks contained in O/D Zones where would prevail travels using Line 4, isolated or together with other transportation ways and are located up to 10 minutes of automobile travel from its stations[4];
- **Reference area** embeds influence area, and is composed by O/D Zones chosen by its similarity and proximity (up to 15 minutes of automobile travel) to the influence area; these zones are expected to suffer a mitigated impact of the new subway line;
- **Control area** is also composed by O/D Zones selected by its similarity to the influence area, though also distant of future infrastructure transportation interventions[5].

[2] Software used was MapInfo 9.
[3] This area corresponds to the definition of *A.I.U. – Área de Intervenção Urbana*, Areas of Urban Intervention, as presented at the City urban master plan.
[4] The identification of these O/D Zones was done through traffic simulation modeling, by internal Metrô-SP staff.

Both reference and control areas were selected by cluster analysis, using a number of urban and socio-economic indicators, such as:

- Population number;
- Number of job positions and employment profile;
- Population age profile;
- Accessibility level[6];
- Land occupation patterns;
- Median household income.

Influence area	Control area	Reference area
Alto de Pinheiros, Butantã, Consolação, J.Paulista, Perdizes, Pinheiros, Morumbi, S.Cecilia	Campo Belo (*)	Tucuruvi
Raposo Tavares, Taboão da Serra	None	J.São Luís
Campo Limpo, Rio Pequeno	Cidade Ademar	Vila Medeiros
Bom Retiro, Brás, Pari, Santana, V.Guilherme, V.Sônia	Limão	Jaguaré
Bela Vista, República, Sé	None	None

(*) Appliable only until construction of Line 5 – Lilacs.

Chart 2 – Influence, reference and control areas.

[5] Time projection span set for this selection is usually from 10 to 15 years. Even so, part of the control area (Campo Belo) will be served by the subway system expansion, which was defined after the start-up of this study.

[6] Defined through transportation time measures.

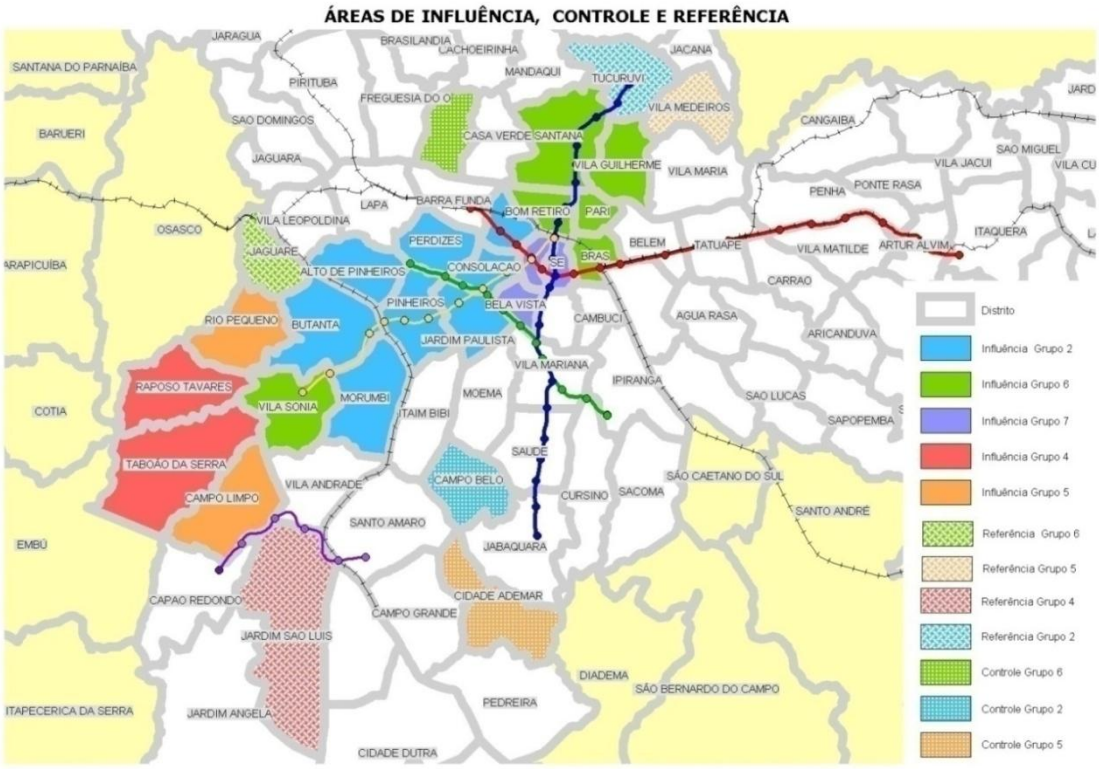

Chart 3 - Influence, reference and control areas.

4 GIS IMPLEMENTATION AND SECONDARY DATA COLLECTION

The system implementation comprised the following steps:

1. Reorganization of data libraries already available at the Transportation Planning Department of Metrô-SP; this activity included compiling expropriation data surveyed in 1998 and 2000;

2. Inclusion of new data bases acquired, such as EMBRAESP's (on new housing start-ups), TPCL (Municipal building cadastre, following tax assessment criteria), ITBI (Municipal real estate sales' tax record), SEHAB (Municipal new building licensing records), RAIS (Federal government data on employment), O/D Zoning (updated in 2007 internally by Metrô-SP), etc.; besides these planning information, contractors' existing data on sales and rentals of properties for the 2002 basis-year(c.2,000 information) were also uploaded;

3. Acquisition and installation of GIS software;

4. Design and implementation of GIS maintenance routines;

5. Training of technical staff of Metrô-SP.

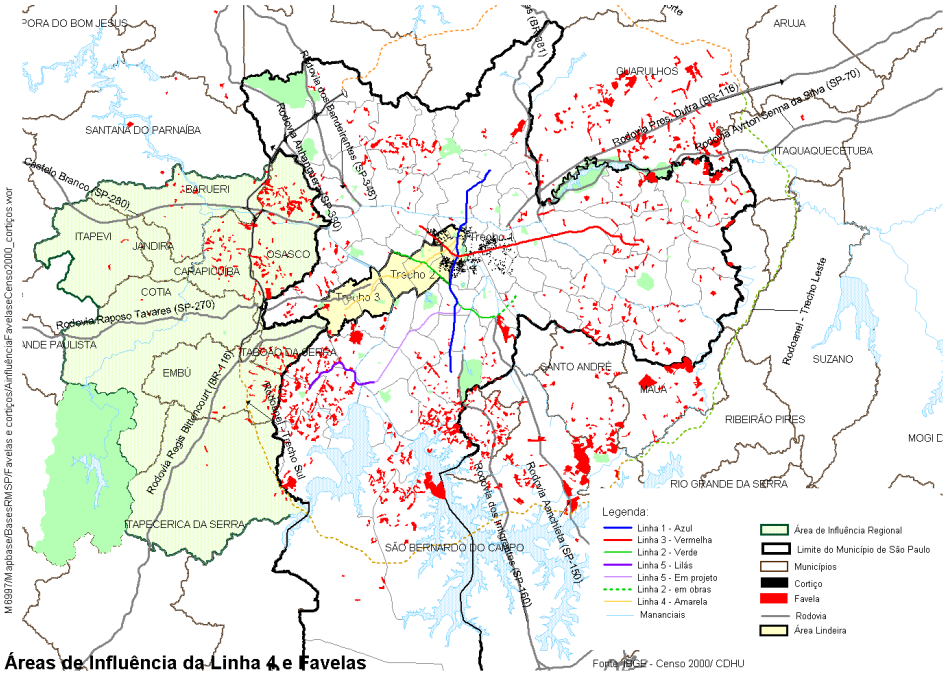

Chart 4 – Sample of GIS output – Influence area and slums' localization.

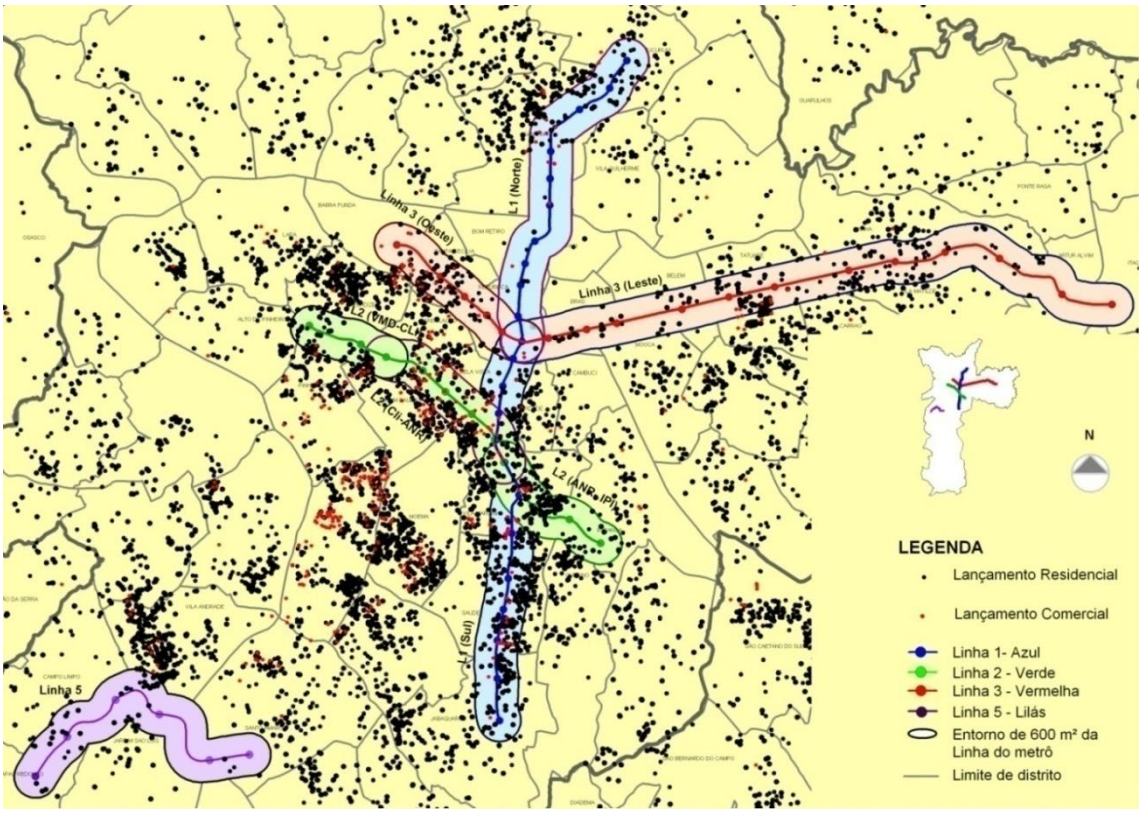

Chart 4 – Sample of GIS output – subway network, direct influence area and new housing start-ups.

5 PRIMARY DATA COLLECTION – PROPERTIES' VALUES

An ample survey was conducted along the influence, reference and control areas, on all property offered for sale or rental in market. These data were consisted, and circa 5,000 of them were uploaded at the GIS.

Compiling of the information was done through a customized interface, comprising location, zoning, cadastre data, among others. A sample is presented below:

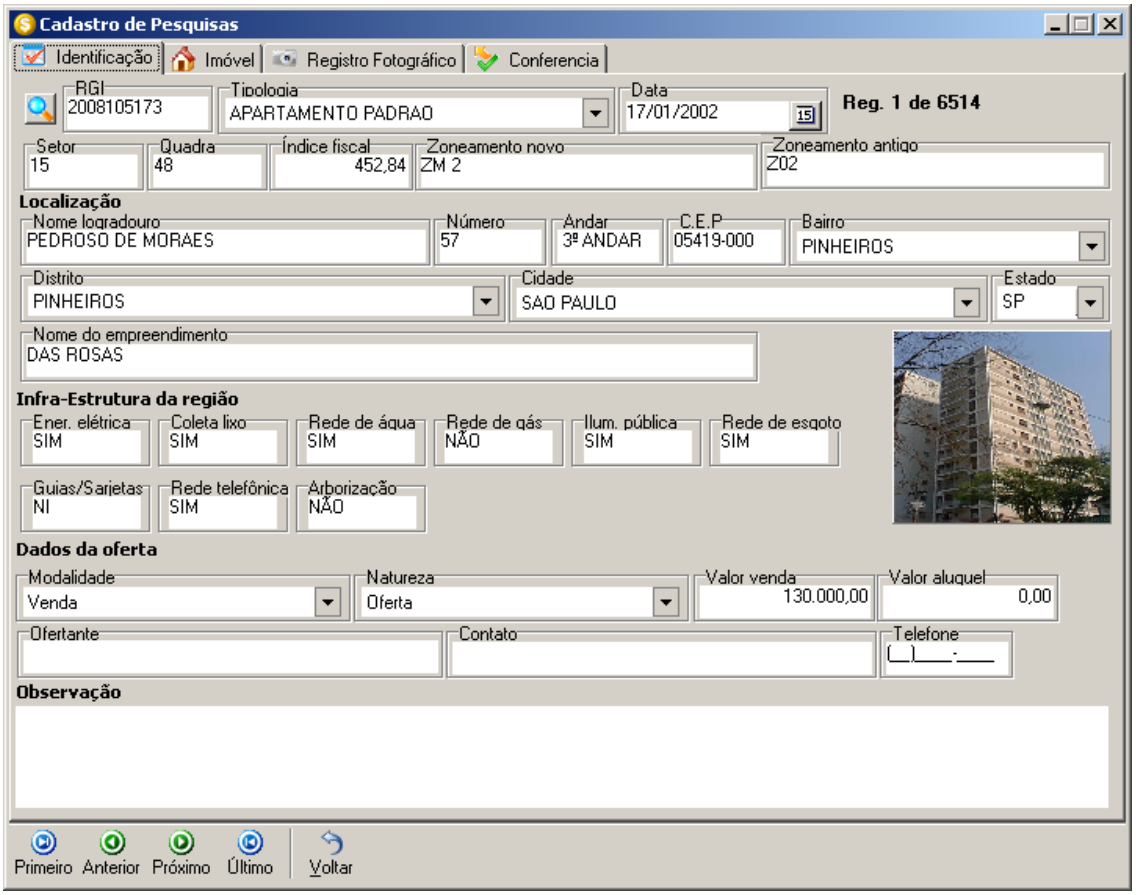

Chart 5 – Sample of survey register.

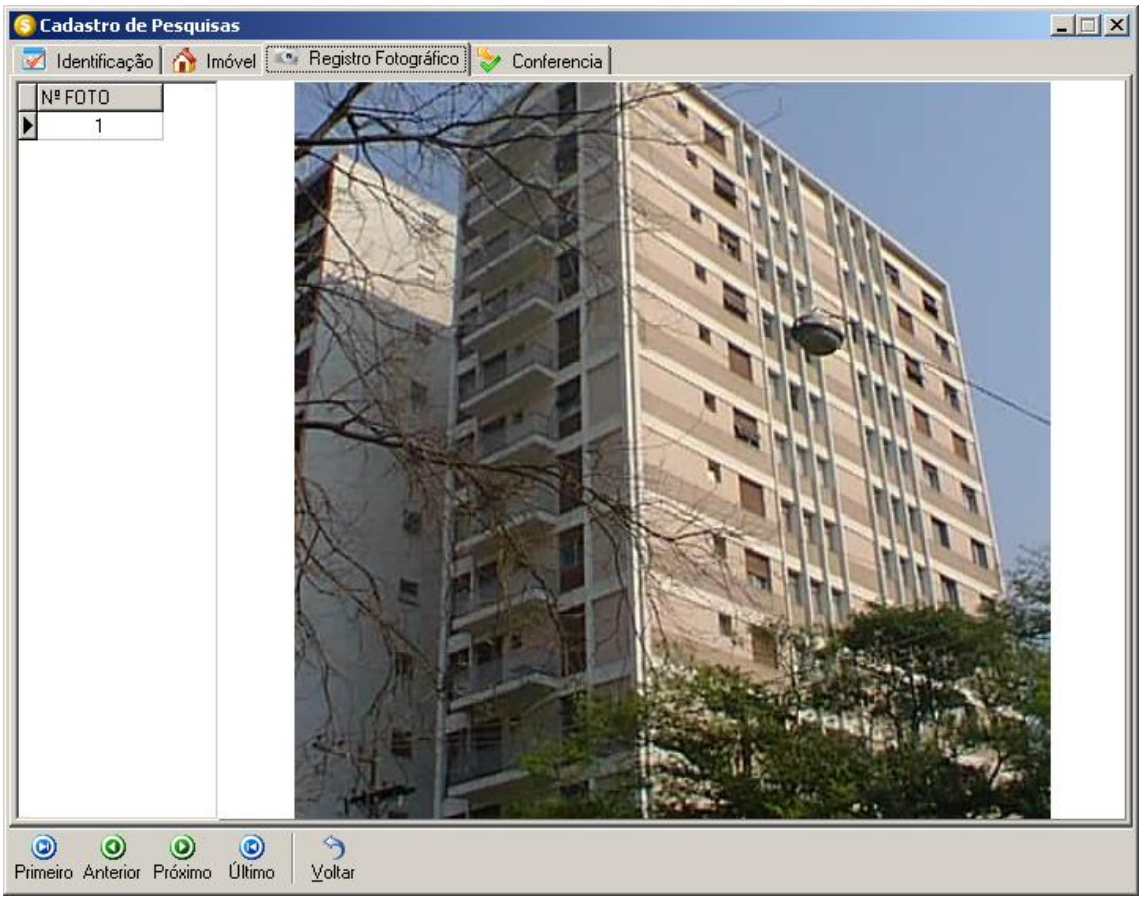

Chart 6 – Detail of photo register.

6 PRIMARY DATA COLLECTION – OCCUPANCY ASSESSMENT

Property occupancy was also assessed, as an ancillary indicator to real estate market activity. The establishment of the occupation level at the present date will allow, when this survey is repeated, after the completion of the line, to measure the net market absorption through the difference between vacancy rates in different moments, following different property classes.

Basic indicator is the vacancy rate, defined (for every property class) as:

$$Vacancy\ Rate = \frac{\#\ of\ vacant\ properties}{Total\ of\ properties}$$

For this purpose, a sample of 285 blocks was defined in the study area, distributed among the influence, reference and control areas. Due to urban characteristics, the influence area was divided into three stretches, as shown in the map below.

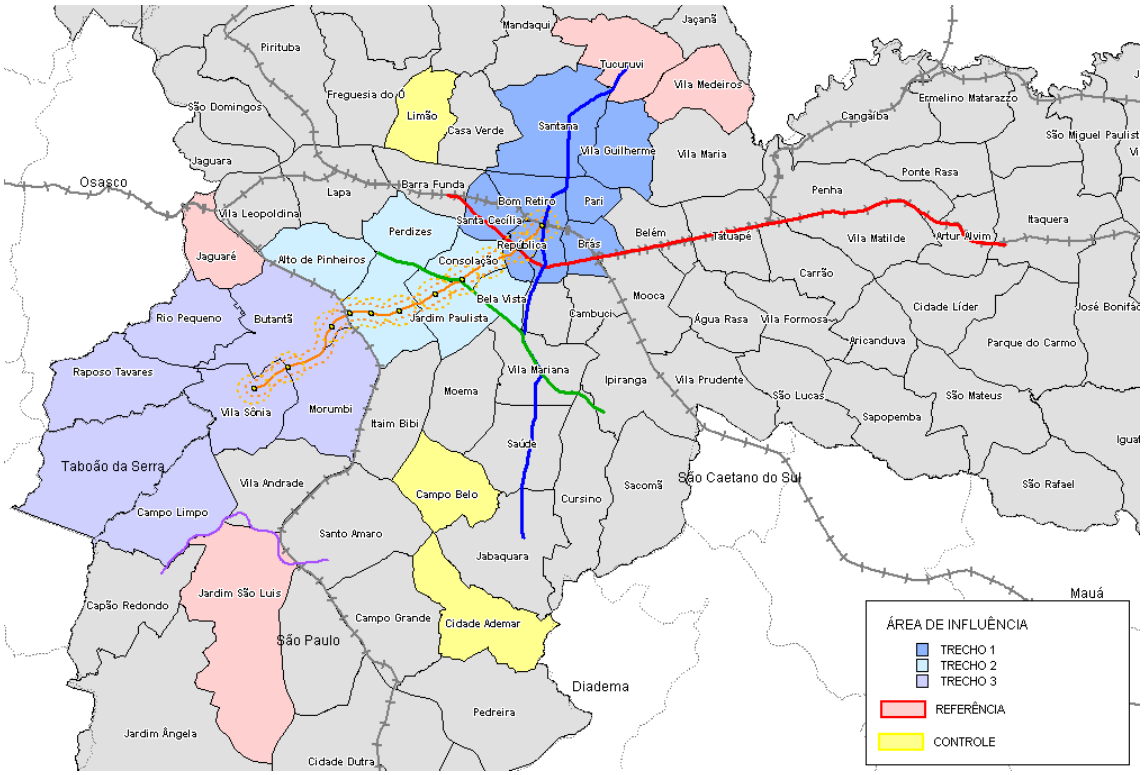

Chart 7 – Study area, showing subdivision of influence area.

All properties in the sample were visited, classified and had their occupancy situation surveyed, in a total of 51,563 units, whose profile is shown in Table 1, below. It may be noticed that vertical housing (apartments) corresponds to 41.6% of the total sample.

Table 1 – Surveyed properties' profile (occupancy assessment).

	Study Area					
Usage	Influence, Stretch 1	Influence, Stretch 2	Influence, Stretch 3	Reference	Control	Total
Housing, horizontal	737	832	2,464	4,370	2,716	11,119
Housing, vertical	7,177	7,922	2,631	214	3,519	21,463
Parking	88	39	6	12	18	163
Business, horizontal	1,575	968	368	420	384	3,715
Business, vertical	7,699	2,323	218	45	288	10,573
Mixed, horizontal	214	153	143	291	538	1,339
Mixed, vertical	2,395	376	66	29	124	2,990
Vacant lots/ruined/others	25	17	105	22	32	201
Total	19,910	12,630	6,001	5,403	7,619	51,563

7 DATA TREATMENT THROUGH DESCRIPTIVE STATISTICS

Basic statistics, such as mean, adjusted mean, median, standard deviation and quartering, were produced from the data obtained for values and occupation in order to allow an initial analysis of market behavior. Some of the most relevant observations are presented below:

7.1 Developed properties' values

Table 2 summarizes field observations on the apartment market, and their unit value (adjusted mean) evolution between 2002 and 2008, considering 2008-based values[7]:

Table 2 – Apartments' unit values evolution (adjusted mean) in study area.

Area	Description (Construction Quality/#bedrooms)	2002	2008	Variation
Influence, stretch 1	Average - 1	1,684.46	2,221.60	31.9%
	Average - 2	N.A.	1,968.17	
	Standard - 1		1,418.75	
Influence, stretch 2	Average - 1	2,572.46	3,120.,84	21.3%
	Average - 2	2,407.25	3,226.07	34.0%
	Average - 3	2,629.91	3,019.23	14.8%
	Average - 4	N.A.	3,258.33	
	Standard - 1		2,123.91	
	Standard - 2		2,415.08	
	Superior - 1	4,233.90	3,993.29	-5.7%

[7] Inflation index considered was IPC-Fipe (Consumer Price Index, as defined by the Economics Research Foundation at São Paulo University).

	Superior - 2	2,977.78	4,205.71	41.2%
	Superior - 3	2,535.25	3,625.54	43.0%
	Superior - 4	3,809.58	4,419.53	16.0%
Influence, stretch 3	Average - 1	2,470.79	N.A.	
	Average - 2	2,094.13	2,474.76	18.2%
	Average - 3	2,045.49	2,544.23	24.4%
	Standard - 2	1,577.41	N.A.	
	Superior - 3	3,075.73	3,011.31	-2.1%
	Superior - 4	3,133.03	2,820.07	-10.0%
Reference	Average - 2	1,687.19	2,263.11	34.1%
	Average - 3	1,668.95	N.A.	
Control	Average - 1	2,528.51	N.A.	
	Average - 2	2,234.14	2,589.94	15.9%
	Average - 3	2,735.45	2,869.31	4.9%
	Superior - 2	3,830.27	N.A.	
	Superior - 3	3,458.03	3,347.78	-3.2%
	Superior - 4	4,245.87	3,672.40	-13.5%

It may be pointed that:

- There is a clear increase in prices of 1 bedroom apartments, in special in stretch 1 of influence area and in reference area (31.9% and 18.7%, respectively);
- In stretch 2 of influence area, superior apartments also increased prices (between 16.0 and 43.0%), except for 1 bedroom units (mostly condo-hotels); this specific segment devaluated in stretch 3 of influence area (from -2.1% to -10.0%) and in control area (-3.2%);
- Average quality apartments, where comparable, valued intensely in stretches 2 and 3 of influence area, followed by reference area, where 2 bedroom units presented a strong appreciation (34.1%);
- In control area, average 2 bedroom apartments appreciated moderately (15.9%), while 3 bedroom apartments appreciated modestly (4.9%), while superior apartments depreciated.

Generally speaking, there was appreciation along the influence area, more steeply in stretches 1 and 2, where construction works were more active at survey's period. This increase is general and strong in average quality apartments and, in stretch 2, also noticeable in superior apartments, with the exception of the condo-hotel segment.

7.2 Occupation assessment

Table 3 summarizes field observations occupation (adjusted mean), divided by usage and by study area segment.

Table 3 – Vacancy rates (adjusted mean) in study area.

Usage	Global	Influence, Stretch 1	Influence, Stretch 2	Influence, Stretch 3	Control	Reference
Housing, horizontal	4.70%	10.60%	6.40%	6.20%	3.60%	3.20%
Housing, vertical	5.50%	6.00%	2.50%	10.00%	7.30%	2.80%
Parking	0.00%	0.00%	0.00%		0.00%	
Business, horizontal	7.30%	7.60%	8.70%	6.20%	8.70%	8.10%
Business, vertical	9.70%	10.10%	7.40%	43.00%	4.70%	3.80%
Mixed, horizontal	4.60%	7.30%	0.20%	4.20%	4.10%	6.00%
Mixed, vertical	5.50%	4.10%		9.60%	4.00%	6.30%
Vacant lots/ruined/others	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%
Total	6.30%	6.80%	5.00%	9.90%	5.50%	4.80%

It may be pointed that:

- Vacancy rates are quite low (mostly below 9%), with one noticeable exception (43% for "business, vertical" usage in stretch 3); this is due to small number of occurrences (this usage appeared in only 10 blocks, out of 57), as well as to market timing (recently delivered stock, still under market absorption); additionally, high vacancy and small offer may be indicate a residential vocation (observed in 52 blocks, out of 57), which is reinforced by city zoning laws; however, changes may occur after Line 4 inauguration;
- Stretch 3 of influence area, as may observed in Table 1, has a similar profile to control area, however with higher vacancy rates, including vertical housing; this may indicate a less consolidated property market, which may be reverted after Line 4 implementation;
- In stretch 1 of influence area (inner city), business vacancy rates are higher than for the apartment stock (10.1% x 6.0%);
- A similar situation is seen in stretch 2 of influence area (7.4% x 2.5%), though with smaller figures, which reflect a traditionally more demanded area of the city, from property market point of view.

Though so far this is only a cross-section view of this indicator (as pointed above, only when this survey is repeated it will be possible to measure the effect of Line 4 implementation in real estate market dynamics), it is remarkable that high value increases[8] occurred in the same regions where vacancy rates are smaller (stretches 1 and 2 of the influence area). In a very symmetric way, value decreased where vacancy is high (stretch 3 of influence area and control area).

7.3 Data treatment through hedonic modeling

The data available were used to derive a hedonic model. A total of 11 geographically referred variables were tested, of which 08 showed statistic significance:

- **PD** = construction quality;
- **AP** = private area (m^2);
- **DORM =** number of bedrooms;
- **DM =** distance to closest subway station (m)
- **DATA =** date dummy variable:

  - If 2002 → DATA = 0;

  - If 2008 → DATA = 1;

[8] Due to sampling reasons, this analysis was only feasible in the apartment market; however, this is the most relevant real estate typology observed, as noticeable in Table 1.

- **RM =** median household income (R$);
- **DET =** service jobs density (total of service jobs per hectare);
- **TC =** average travel time through public transportation;
- **TI =** average travel time through automobile.

The resulting model was:

$$\mathbf{[V.U.] = Exp(\ 6{,}176 + 0{,}214x[PD] - 8{,}6x10^{-4}x[AP] + 0{,}031x[DORM] + 2{,}34/[DM] + 0{,}10x[DATA] + 1{,}97x10^{-5}x[RM] + 0{,}373x[DET] + 0{,}011x[TC] - 0{,}0038x[TI])}$$

Basic statistics:

- Correlation Coefficient = 0,747
- Adjusted Determination Coefficient = 55,70%
- Model significance << 1%;

Variable relevance

The sensitivity of the model to variations in each variable was obtained in a *coeteris paribus* condition, through a Montecarlo simulation, maintaining all variables constant, but the one under analysis. Table 4, below, presents the variation coefficients obtained for each simulation already ranked. The most relevant variables were construction quality (PD), travel time in public transportation (TC) and private area (AP).

Table 4 – Relevance analysis of the model variables.

Variable		**Variation Coefficient**
PD	Construction Quality	31,52%
TC	Average travel time through public transportation	15,83%
AP	Private area (m^2)	6,73%
DET	Service jobs density	5,41%
TI	Average travel time through automobile	4,92%

DORM	Number of bedrooms	2,98%
RM	Median household income (R$)	0,75%
DM	Distance to closest subway station (m)	0,27%

8 RESULTS

Market general appreciation

One of the first results that may be obtained from the model is a measure of the general market appreciation between 2002 and 2008, given by the DATA variable. According the model (which was derived on a constant price basis for 2008), São Paulo city apartment market prices generally rose in 10%, as shown:

$$GeneralAppreciation = \frac{e^{0,101}}{e^{0,100}} - 1 = 1,10 - 1,0 = 0,1 \text{ or } 10\%$$

Spatial trends

The use of geometrical grids allows the visualization of spatial value trends. Chart 8, below, presents a sample of such an output, fixing some parameters (in this case, average quality construction, 2008 date, 02 bedrooms and private area of 68 m² - median of this variable observations).

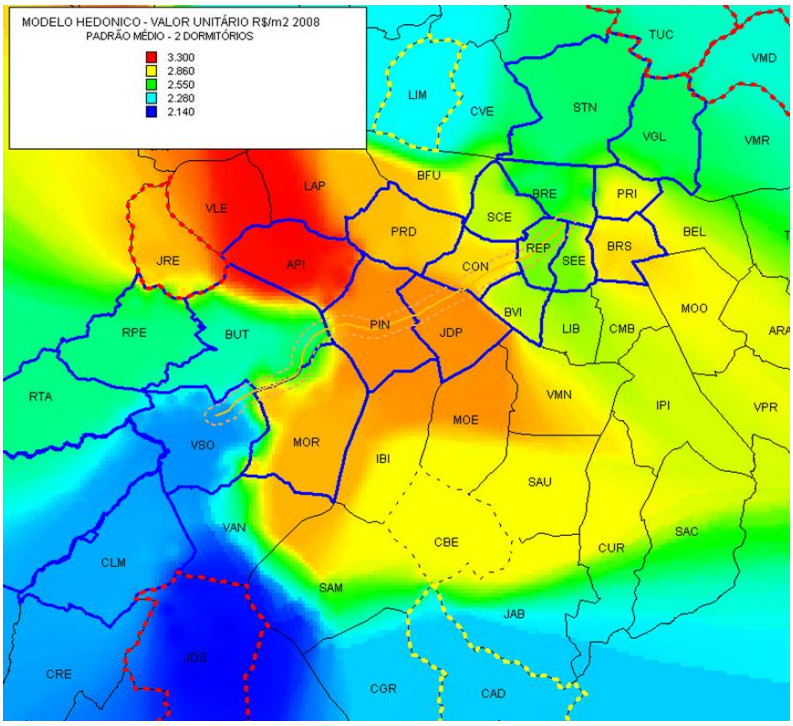

Chart 8 – Spatial value trends – apartment market, 2008, 02 bedrooms and private area of 68 m².

Warmer colors indicate higher values (Alto de Pinheiros, Vila Leopoldina, Pinheiros, Jardim Paulista), while colder colors indicate the opposite (Butantã, Campo Limpo, Vila Sônia). A kind

of “warmth epicenter” is noticeable along stretch 2 of Influence Area, though higher value regions (pulled up by other factors) exist.

Local variations

The model allows comparisons between present observations and inferred past values, as indicated in Chart 9, for 02 and 03 bedroom medium quality apartments in the study area.

Chart 9 – Appreciation of apartments in study área – 2008 observations x 2002 inferred values

AREA	ZONE	BDR	PRIVATE AREA	VU_INF_2002	VU 2008	APPRECIATION
CONTROL	CAMPO BELO	3	106,8	R$ 1.989,17	R$ 2.286,29	15%
REFERENCE	JARDIM SAO LUIS	2	49,23	R$ 1.981,75	R$ 1.908,63	-4%
REFERENCE	TUCURUVI	2	57,06	R$ 2.183,20	R$ 2.533,41	16%
INFLUENCE-STR.3	BUTANTA	3	72,06	R$ 1.941,58	R$ 2.432,42	25%
INFLUENCE-STR.1	CONSOLACAO	3	107	R$ 1.843,17	R$ 2.524,01	37%
INFLUENCE-STR.2	PINHEIROS	3	107	R$ 1.824,80	R$ 2.771,01	52%

These examples show a trend of strong appreciations to be present in the influence area (between 25 and 52%), while in reference and control areas, this phenomenon is much more modest (around 15%), or even reversed. When assessing the impact of infrastructure works like Line 4, the overall market increase (of 10% for the period, as explained above) must be deducted.

9 MAIN CONCLUSIONS

a) Based on the descriptive statistics analysis, vacancy rates in study area as a whole are generally low (under 7%), which is accentuated in stretches 1 (inner city) and 2 of influence area (Consolação, Jardins e Pinheiros). This situation is reverted in stretch 3 (Butantã and Vila Sônia) and in control area (Campo Belo and Limão). Symmetrically, prices trends are reverted in these areas: while stretches 1 and 2 showed a clear and strong appreciation between 2002 and 2008 (mostly for average quality apartments of 1 and 2 bedrooms, as well as for superior apartments), control area in general presented depreciation.

b) Hedonic modeling indicated as statistically significant a group of 8 variables, of which the most relevant were construction quality (PD), travel time in public transportation (TC) and private area (AP). Its analysis showed: i) general market appreciation of 10% above inflation (measures by IPC-Fipe) between 2002 and 2008; and ii) strong appreciation in stretch 2 of Influence Area (which correspond to the most dynamic region affected by construction works, from market point of view), though there were other active and well valued areas in the same period.

c) These valuation trends were confirmed by punctual analyses, in which observed values for 2008 were compared with inferred values for 2002. In these cases, appreciation was more intense in influence area (in special, in stretches 2 and 1), and less relevant, or even reverted in control and reference areas.

d) The quantity of data and information available allow further analyses (both considering spatial aggregations and different typologies), which will be pursued before the next survey, expected to occur by 2011, when the line is fully inaugurated.