ASSESSMENT OF SMOOTHING TECHNIQUES APPLIED TO SPEED PROFILES MEASURED WITH GPS RTK

Dr. Tomás ECHAVEGUREN, Associate professor, (techaveg@udec.cl)

Dr. Alejandro TUDELA, Associate professor

CEng. Cristian FONSECA, Research assistant

Engineering Faculty, Universidad de Concepción, Chile

ABSTRACT

GPS device is a tool that permits to obtain a big amount of speed data, which are the main input for traffic studies, safety analysis and geometric consistency of road. A relevant challenge is how to analyse speed data for eliminate noise and to produce an understandable speed signal for modelling or making decision. The data smoothing provides an easy to use tool for compress data. The analyst should choose among several smoothing techniques available those that permit a good representation of original data. A key aspect for this task is the choosing of smoothing parameters and the criteria to discriminate the better performance of the smoothing technique regarding to the original data. The typical performance parameters provided for the literature are representation of statistical error, for instance mean error or root mean square error. There exist, however, other criteria that can be used which permits compare the information content of two signals of data, uni or multi-dimensional. The most popular is the Akaike information criterion. This paper, explores the smoothing techniques: single and double exponential, winters, loess, spline, Kernel and establishes a comparison between performance indexes based in statistical and information content criteria. Several roads of Chilean road network were measured with a high precision GPS to obtain speed profiles for apply different smoothing techniques and performance criteria. A total of 308 combination of smoothing techniques, performance criterion, and test sections were assessed. It was concluded that the better smoothing techniques were the Winters and spline methods if speed profile is highly variable and also that the performance indexes based on information contents are more stables and independents of the smoothing technique and of the type of the speed profile for optimizing the smoothing parameters.

Keywords: Speed Profile, Smoothing, GPS

INTRODUCTION

Operating speed profiles are a relevant input for safety assessment of roads. In-field speed measurement is dependant of the technology available and of the purpose of the measurement. Depending on that, measurement can be classified in punctual or continuous (AASHTO, 2009). Satellite positioning technology is a suitable method for obtaining speed data at low cost and good precision (See Li, 2009 and Xu, 2010). Particularly, some dynamic GPS devices permits to obtain data up to 0,1 s, which increases the data detail, but also increases the possibility of mistakes in the interpretation for modelling, due to the presence of outlier data or dropouts in the speed signal recorded. For instance, local variation of speed and acceleration can lead to obtain speed-geometry on curves biased due to an altered interpretation of entrance speed due to local variations of the profile.

Data smoothing tools permits to study the autocorrelation of speed data and the distance respect a data point of reference, in which that autocorrelation is not relevant to explain driver behaviour. This is a key aspect that speed-geometry models do not study, due that used punctual speed measurements. Also, smoothing analysis permits to obtain “clean” speed profiles for a more realistic speed-geometry modelling and at the same time, to improve the quality of a relevant input for safety and road consistency assessment.

Typical smoothing techniques used for analysing speed profiles are: single exponential smoother, local weighed regression, and splines. Also, because each smoothing technique needs input parameters, it should be chosen from a wide number of possibilities and the common practice is to use smoothing parameters recommended by the literature. However, this practice is not a guarantee that parameters will be suitable for speed data analysis, mainly because speed data are not well-behaved whether road geometry is undulated or mix long straights with sharp curves.

Two questions were stated in this paper:¿what are the best methods for speed smoothing?, and what are the best criteria for selecting the best smoothing method?. According to that, this paper explores the smoothing techniques: single exponential, double exponential, Winters, loess, spline and kernel smoothers which were assessed for different types of road alignments. Several performance criteria were assessed for choosing the better smoothing parameters according to the type of speed profile and to the smoothing technique.

Speed data were measured with a GPS RTK, in roads placed in mountain, undulated, and flat terrain. Position and speed data were recorded each 0,1 s, the maximum resolution of the device used for measurement. Data were reported each 1 m.

The smoothing techniques were applied successively, varying in each case the input parameters and obtaining 11 performance indexes grouped in two categories: statistical indexes such as mean absolute percentage error, mean squared deviation, mean error, root mean square error; and information criteria, such as Akaike, Akaike corrected, Bayesian, Fisher and Hannan-Quinn criteria. In each smoothing technique optimum sets of input parameters were estimated, according to the speed behaviour observed in data recorded.

SMOOTHING METHODS

Data smoothing is technique that developed an approximated mathematical function which permits to obtain data patterns, to isolate the noise, outliers and irregularities in a sequence of data points. It is known non-parametric regression as well. When signal speed is smoothed, the local speed variations due to small accelerations of differential trajectory changes are erased from the signal, providing more continuous speed profiles.

There exist several smoothing methods reported in literature such as, splines, kernel-based method, n-exponential methods, and regressions locally weighed methods and ARxMA methods, which permits to obtain smooth data, trends and seasonality patterns of speed.

Particularly, in this paper were selected the smoothing method commonly reported in literature for smoothing speed data: single exponential, double exponential, Winters' method, Loess method, super smoother, kernel smoother, and spline. Table I shows a summary of the mathematical formulation of each smoothing technique. Except spline method, all the methods estimate a local regression around each data point, applying different method for weigh the neighbour data.

SE smoothing, weight exponentially the neighbour data of the i-th data for obtaining short term forecast of data. Depending of the variability of data, a smoothing parameter α should be chosen. It ranges between 0 and 1. If data variability is high, the α-value to be chosen is near to 1, otherwise is near to 0.

DE smoothing (also known as Holt´ method), is similar to SE method, but includes the trend behaviour of data through the β-parameter which represents the trend component of data. The β-value ranges between 0 and 1. This method is called double-exponential because weight exponentially the short term forecast and the trend.

WS method is a variation of DE smoothing which includes the seasonal behaviour of data through the γ parameter. This method could be suitable for high variability of speed in of undulated roadways. The γ-value ranges between 0 and 1.

The Loess method (locally weighed regression) uses the least square regression method for obtaining a n-degree polynomial around each data point over an user-defined span. The polynomial is obtained weighing with a tri-cubical function the data into the span, giving more weight to the nearest data around the i-th data point.

Table I – Mathematical formulation of each smoothing techniques assessed (Cleveland, 1979; Friedman, 1984; Cleveland and Devlin, 1988; Ferrari et al, 2006; Kayri and Zirhhoglu, 2009).

Smoothing Method	Mathematical Formulation
Single Exponential (SE)	$F_t = F_{t-1} + \alpha(x_{t-1} - F_{t-1})$
Double Exponential (DE)	$FT_t = S_{t-1} + T_{t-1}$ $S_t = FT_t + \alpha\left(x_t - FT_t\right)$ $T_t = T_{t-1} + \beta\left(FT_t - FT_{t-1} - T_{t-1}\right)$
Winters (WS)	$F_{t+m} = \left(S_t + T_t m\right) I_{t-L+m}$ $S_t = \alpha \frac{x_t}{I_{t-1}} + \left(1-\alpha\right)\left(S_{t-1} + T_{t-1}\right)$ $T_t = \beta\left(S_t - S_{t-1}\right) + \left(1-\beta\right) T_{t-1}$ $I_t = \gamma \frac{x_t}{s_t} + \left(1-\gamma\right) I_{t-1}$
Loess (LE)	$Y_i = \sum_{i=1}^{n} W(x)_{ij} y_j$ $W\left(x\right) = \begin{cases} \left(1-x^3\right)^3 & 0 \le x < 1 \\ 0 & \text{otherwise} \end{cases}$
Kernel Smoother (KS)	$Y_i = \sum_{i=1}^{n} w_{ij} y_j$ $w_{ij} = \tilde{K}\left(\frac{x_i - x_j}{b}\right) = \frac{K\left(\frac{x_i - x_j}{b}\right)}{\sum_{k=1}^{n} K\left(\frac{x_i - x_k}{b}\right)}$
Spline (SFS)	$RSS = \sum_{i=1}^{n}\left(x_i - f\left(y_i\right)\right)^2 + \lambda \int \left(f''\left(t\right)\right)^2 dt$

Note: F_t is the forecast in the period t; F_{t-1} is the forecast of the previous period; x_{t-1} is the value of the data to smooth in the previous period; FT_t is the trend component of the forecast; S_t is the mean value of the forecast in the period t, T_t is the trend estimation in the period t; x_t is the data to smooth in the period t; L is the number of season that defines the seasonal pattern; m is the number of future periods to forecast; I_t is the seasonal estimation in the period t; x_i and x_j are the observed data; W(x) is the tri-cubic weight function; Y_i is the smoothed function around the i-th data point; f"(t) is the second derivative of the spline function $f(x_i)$ obtained around the i-th data point.

The Kernel smoother obtaining smoothed values by using the mean weighed of data, with a decreasing weigh for the more distant data into the span. The weight (w_{ij}) is dependant of the bandwidth (b) and of the kernel function K(t). Kernel function is a positive and unitary function along its dominion. There exist several specification of kernel function, such as box, triangle, Parzen and normal model. Independently of the kernel specification, the more relevant aspect is to choose the correct bandwidth for smoothing, as explains Hastie and Tibshirani (1990).

Spline smoothing is the most used method for smoothing. The method estimates local cubic functions by minimizing the square sum of residuals (RSS). The method uses a smoothing parameter λ, which is similar of the span used in Loess and super-smoother methods and defines the relevance of the goodness to fit of the smoothed valued regarded to the smoothness of the function obtained.

SMOOTHING PERFORMANCE INDEXES

Smoothing performance indexes (SPI) are a synthetized way for defining the quality of the smoothed representation of raw data of speed. It can be classified as statistical SPI and information-criterion based SPI.

The former are different representation of the error between the original and smoothed speed signal. Typical statistical SPI are the mean error, mean average error, mean absolute percentage error, and root mean square error (See Table II). This SPI are the most used to assess the performance of the smoothing of speed data (Jun et al, 2006).

The last indexes are represents the information content of the smoothed signal against to the original signal, based on the Fisher information criterion (FIC). Typical PSI based on FIC are the Akaike information criterion and its variations, Bayesian information criterion, Hanna-Quinn criterion and Fisher-Akaike information criterion (See Table III).

Table II – Statistical smoothing performance indexes (Jun et al, 2006).

Statistical SPI	Mathematical Formulation
Mean Error (ME)	$\mathrm{ME}=\frac{\sum_{i=1}^{n}(y_i-\hat{y}_i)}{n}$
Mean Average Error (MAE)	$\mathrm{MAE}=\frac{\sum_{i=1}^{n}\left\|(y_i-\hat{y}_i)\right\|}{n}$
Mean Absolute Percentage Error (MAPE)	$\mathrm{MAPE}=\frac{\sum_{i=1}^{n}\left\|(y_i-\hat{y}_i)/y_i\right\|}{n}100$
Root Mean Square Error (RMS)	$\mathrm{RMS}=\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$

Note: n: number of data points of the signal; y_i is the i-th original data point; $\hat{y}_i$: is the i-th smoothed data point.

Table III – Information Criterion-based smoothing performance indexes (Schwarz, 1978; Hannan and Quinn, 1979; Wasserman, 2000; Cetin and Erar, 2002; Bogdan et al, 2004; Burnham and Anderson, 2004; Dimova et al, 2011).

Information criterion-based SPI	Mathematical Formulation
Akaike information criterion (AIC)	$\text{AIC}=-2\left\{\log(\hat{\sigma}^2)\right\}+2\text{k}$
Akaike information criterion (AICc)	$\text{AICc}=\text{AIC}+\left\{\dfrac{2\text{k}(\text{k}+1)}{\text{n-k-2}}\right\}$
Bayesian Information Criterion (BIC)	$\text{BIC}=-2\left\{\ln(\hat{\sigma}^2)\right\}+\text{k}\ln(\text{n})$
Fisher-Akaike information criterion (AICf)	$\text{AICf}=\dfrac{\text{n}}{\hat{\sigma}^2}\left\{\dfrac{\text{nk}}{\text{n-k-2}}+2\right\}$
Hannan-Quinn information criterion (HQC)	$\text{HQC}=\text{n}\left\{\ln(\hat{\sigma}^2)\right\}+\text{kC}\ln\left\{\log(\text{n})\right\}$

Note: k is the number of parameters of the model (for uni-dimensional series, k = 1); n is the number of the original data points; $\hat{\sigma}^2$ is the white noise variance which can be estimated by using the sum of square error normalized by the number of data points (SSE/n) obtained from the analysis of variance (ANOVA) of the original speed signal and the smoothed speed signal; C is a parameter of HQC (Hannan and Quinn (1979) suggest C=2).

SPEED DATA COLLECTION

For assessing the smoothed methods, speed data were measured in several roads by using a GPS device based on the real time kinematic (RTK) technology. A preliminary test section was selected by considering flat, undulated and mountainous terrain topography. Once speed data were obtained, a second selection was performed by considering the variability of speed profile measured using the Polus' road consistency criteria, which is based on the cumulated deviation of speed data regarded to the average speed. Considering that, 4 test sections with its respective speed profile was used to study the smoothing methods. The following sections of the paper explain in detail the speed data collection process.

Preliminary Selection of Test Section

The preliminary selection of test section was developed with aid of Google Earth Tool. A total of 9 test sections were selected considering the average curvature and average step for ensuring variability from straight and flat sections to curve and mountainous sections. Additional criteria considered were: paved roads with a good surface conditions, road sectors away villages and towns, uninterrupted traffic conditions (without signals, intersections or roadwork) and low traffic volume.

Speed Data Measurement

Speed data were measured with a 10 Hz GPS RTK, mounted in a probe vehicle. The GPS logger obtains speed, position, curvature and heading. The device obtained data each 0,1 s with a precision of 0,2 km/h and 0,05 % of distance (Racelogic, 2008). The device is capable to obtain position information from 8 satellites at the same time, which permits to obtain a high quality data. Repeatability and reproducibility test conducted by Echaveguren et al (2011), shows that in straight sections and un sharp curves, the dispersion on speed measurements are lower than 2 km/h and that the variance of reproducibility is lower than 0,2 %. It implies that the device uses is highly repeatable and reproducible. An example of the speed signal obtained with the GPS logger is showed in Figure 1.

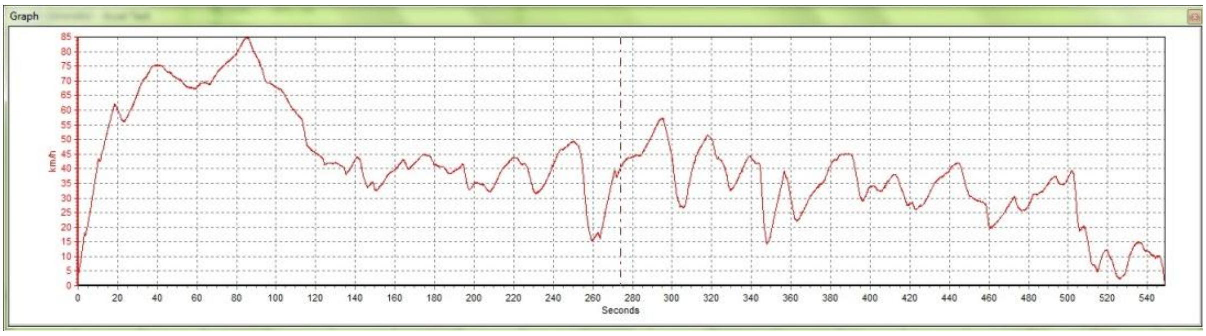


Figure 1 – Example of raw speed profile measured with GPS device in 6 km in Route G-200

Speed Profiles Selection

The objective of this task was to select speed profiles with different levels of variability. Polus and Mattar-Habib (2004) developed a model for assessing the consistency of roads based in the R_a index, which estimates the cumulated deviation of speed data against the average mean speed of a test section (See Figure 2). Of the speed profiles is uniform then the speed data are near to the average mean speed, and the cumulated deviation of speed is near to zero. Then, the R_a value tends to zero. On the contrary, if the speed profile exhibits high variability, then the cumulated deviation is higher and R_a is higher as well. Therefore, the R_a index is a good representation to discriminate the uniformity of the speed profile.

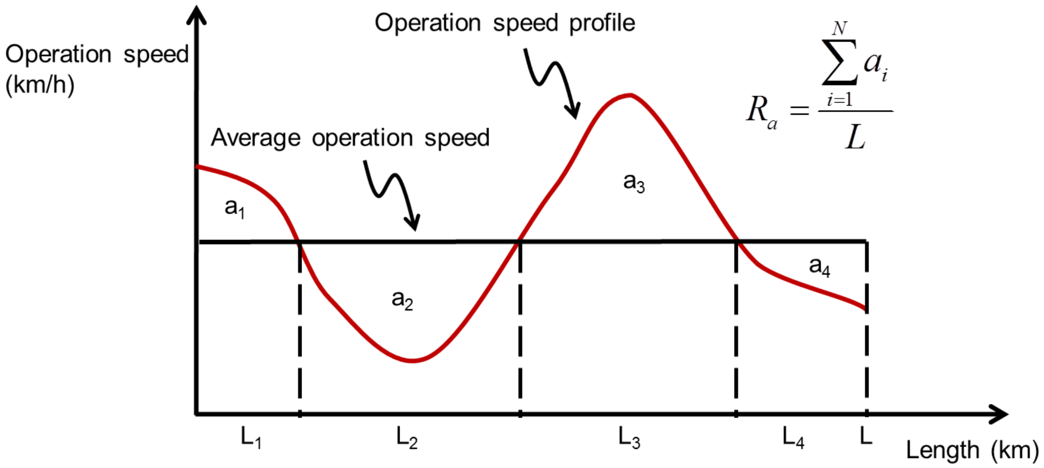
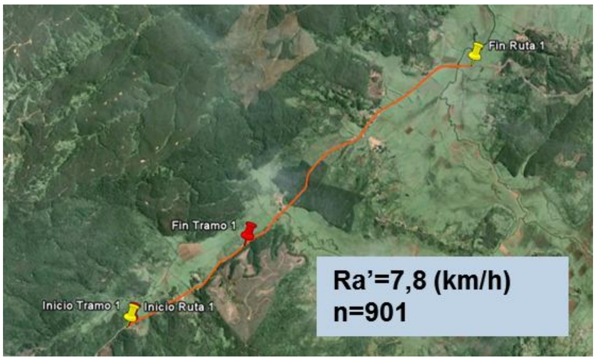
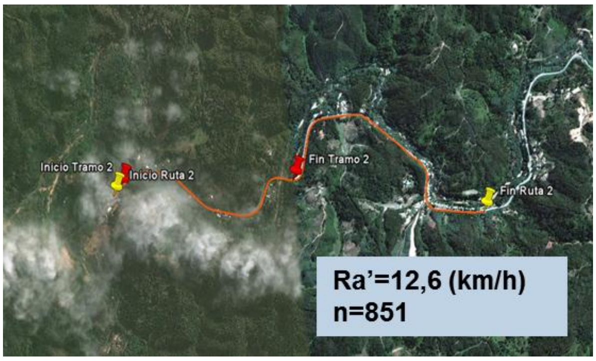


Figure 2 – Criterion for assessing the uniformity of speed profile (adapted from Polus and Mattar-Habib, 2004)

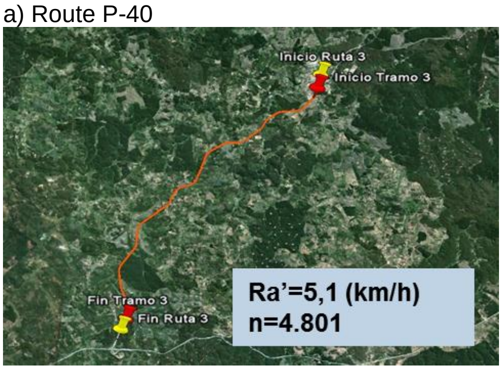
The method previously described was applied to all the speed profiles measured and 4 test section were selected for assessing smoothing methods and SPI. Figure 3 shows the test section selected.



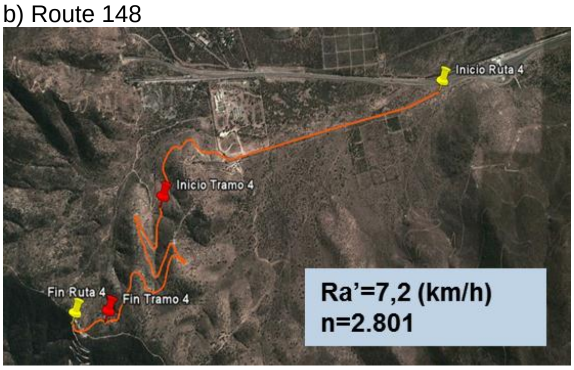
a) Route P-40



b) Route 148



c) Route 158



d) Route G-200

Figure 3 – Test sections selected for assessing smoothing methods and SPI

ASSESSMENT OF THE SMOOTHING METHODS

Assessment was developed in 3 steps: selection of starting bounds of smoothing parameters, estimation of SPI for each test section and for each smoothing method and assessment of results. In the first step a starting range of smoothing parameters was selected for varying each of it and estimates the SPI. In each test section and for each smoothing parameter, optimum ranges of values which minimize the SPI were proposed.

Selection of Starting Bounds of Smoothing Parameters

The smoothing parameters are single values which defines characteristics that describe the relevance of short term forecast, trend, seasonality, span, among others aspects of each smoothing method. Literature provides typical values of these parameters. A summary of the first set of the smoothing parameters values collected from the literature is presented in Table IV.

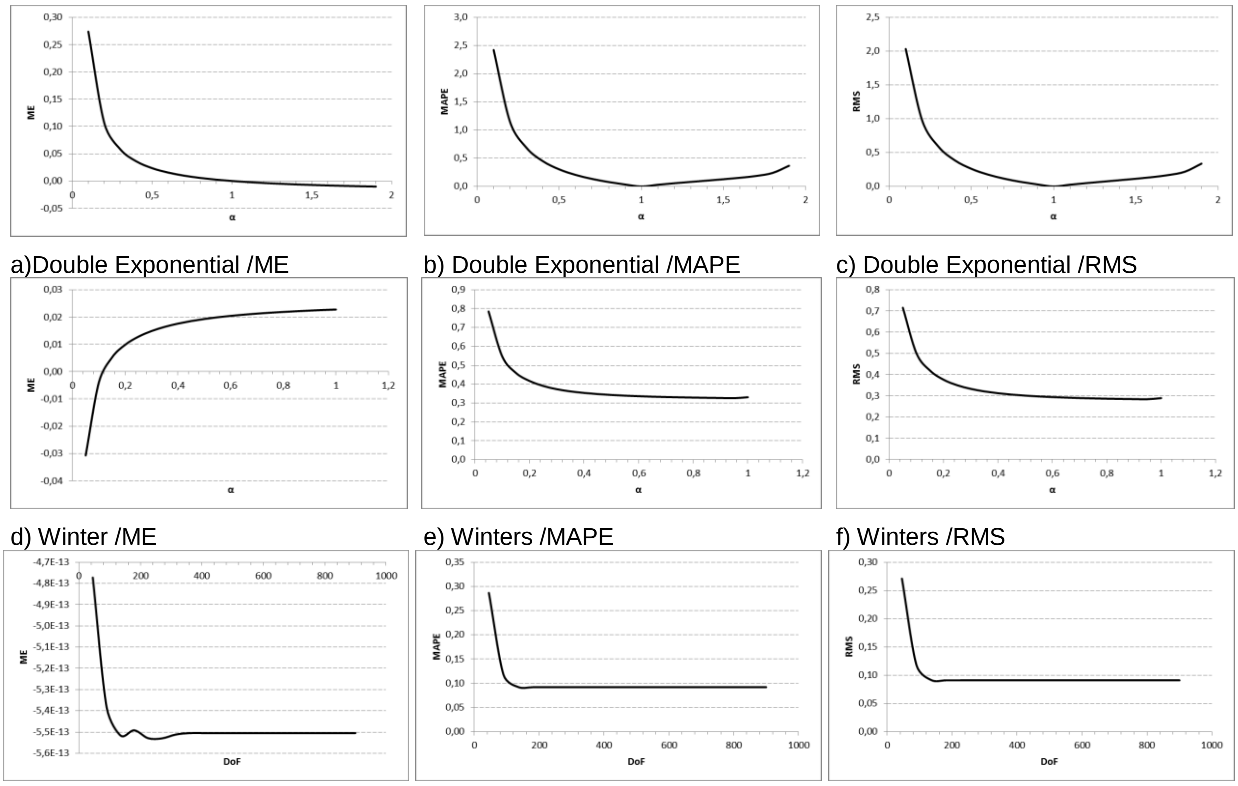
Each interval of smoothing parameter was discretised in 20 values. For Double exponential and winter method, the parameters were optimized using ARIMA model. Then, one of each parameter was set constant and the other parameters were discretized in 20 values. For the kernel, loess and super-smoother methods, the normal kernel was chosen. For spline method the discretization was defined increasing the degree of freedom in 1 unit from 1 to n-1. Therefore, a data base with different values of smoothing parameters was developed for estimating in the next section the SPI.

Table IV – Summary of bounds of smoothing parameter for each smoothing method.

Smoothing Method	Level of Smoothing Parameter	Smoothing Parameter	Bounds of Smoothing parameters
Single Exponential	Data	α	[0;2[
Double Exponential	Data	α	]0;2[
	Trend	γ	]0;(4/α -2[
Winters	Data	α	]0;1[
	Trend	γ	]0;1[
	Seasonality	δ	]0;1[
Kernel	bandwidth	band	]0;2[
Loess	-	span	]0,1]
Spline	Degree of Freedom	DoF	[1; # data – 1]

Estimation of Smoothing Performance Indexes

SPI values were obtained for each test section, smoothing method and smoothing parameter values. A collection of plots of SPI against smoothing parameters values were developed for each smoothing method and for each test section. Graphics permits to obtain the range of values of smoothing parameters which minimizes the SPI values for each smoothing method. Figure 4 shows examples of graphics obtained in the test section 1 (Route P-40) using Statistical SPI. Figure 5 shows examples using information content based SPI.



a)Double Exponential /ME b) Double Exponential /MAPE c) Double Exponential /RMS

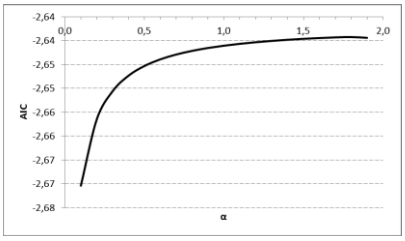
d) Winter /ME e) Winters /MAPE f) Winters /RMS

g)Spline /ME h) Spline /MAPE i) Spline /RMS

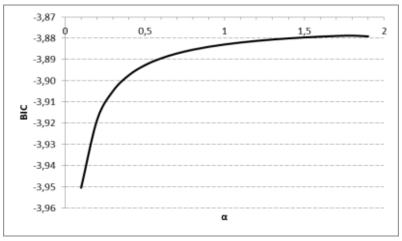
Figure 4 – Examples of statistical SPI data against smoothing parameter for test section 1.

Figures 4a, 4b and 4c shows that α-values higher than 0,5 minimize the SPI ME, MAPE and RMS. This result is consistent with the result showed in Figure 5a, 5b and 5c, in which the information criterion AIC, BIC and HQC are stables for α-values higher than 0,5.

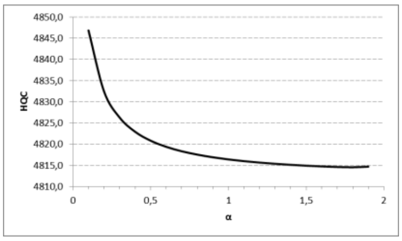
Similar results can be showed in figures 4d, 4e, 4f and 5d, 5e and 5f for Winters´ smoothing. In all cases except for ME, α-values higher than 0,2 stabilize the SPI. In the case of Spline smoothing, there exist a difference in the minimum value of DoF which stabilize the SPI values. Statistical SPI propose a DoF higher than 135 and the other SPI proposes a DoF higher than 90.



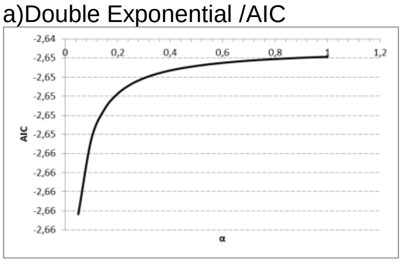
a)Double Exponential /AIC



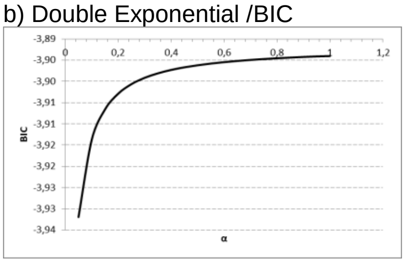
b) Double Exponential /BIC



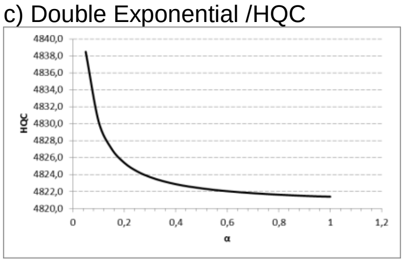
c) Double Exponential /HQC



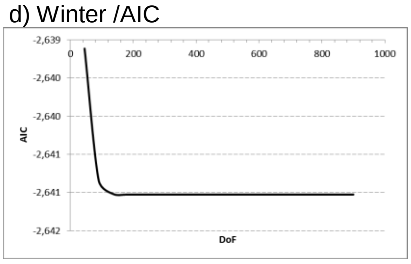
d) Winter /AIC



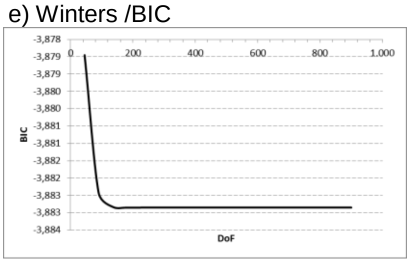
e) Winters /BIC



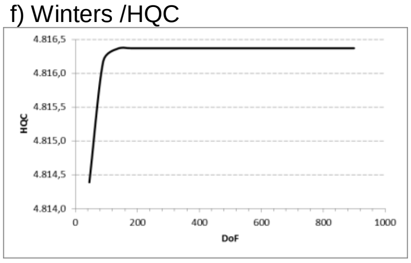
f) Winters /HQC



g)Spline /AIC



h) Spline /BIC



i) Spline /HQC

Figure 5 – Examples of information content-based SPI data against smoothing parameter for test section 1.

Table V summarizes the optimum ranges of smoothing parameter recommended for smooth speed profiles.

Table V – Summary of optimum ranges of smoothing parameters for each test section and smoothing method.

Smoothing Method	Smoothing Parameter	Range of smoothing parameter in each test section			
		Route P-40 (irregular)	Route 148 (very irregular)	Route 158 (uniform)	Route G-200 (irregular)
Single Exponential	α	[0,4; 1,9]	[0,4; 1,9]	[0,3; 1,9]	[0,4; 1,9]
Double Exponential	α (γ fixed)	[0,2; 1,6]	[0,2; 1]	[0,2; 1,2]	[0,2; 1,1]
Winters	α (γ, δ fixed)	[0,2; 1]	[0,2; 1]	[0,3; 1]	[0,2; 1]
Kernel	Band	[W/R]	[W/R]	[W/R]	[W/R]
Loess	span	[W/R]	[0,2; 0,65] (1)	[0,8; 1] (2)	[0,4; 1] (1)
Spline	DoF	[135; 900]	[126; 850]	[1; 4800]	[280; 2800]

W/R: without range identified; (1) except mean error; (2) only information criterion-based SPI.

Table VI shows the classification system developed for each SPI estimated for each smoothing method in test sections (numbered as 1 to 4 in each row of the Table). The classification system considers two attributes as follows.

Table VI – Classification of SPI for each test section and smoothing method.

Smoothing method	Statistical SPI				SPI based in information content				
	ME	MAPE	RMS	MAE	AIC	AICc	BIC	AICf	HQC
Single Exponential	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R
Double Exponential	1 C/R 2 **C/NR** **3 C/NR** **4 C/NR**	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R
Winters	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R	1 C/R 2 C/R 3 C/R 4 C/R
Kernel	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**
Loess	**1 NC/NR** **2 NC/NR** **3 NC/NR** **4 NC/NR**	**1 NC/NR** 2 C/R **3 C/NR** 4 C/R	**1 NC/NR** 2 C/R **3 C/NR** 4 C/R	**1 NC/NR** 2 C/R **3 C/NR** 4 C/R	**1 NC/NR** 2 C/R 3 C/R 4 C/R	**1 NC/NR** 2 C/R 3 C/R 4 C/R	**1 NC/NR** 2 C/R 3 C/R 4 C/R	**1 NC/NR** 2 C/R 3 C/R 4 C/R	**1 NC/NR** 2 C/R 3 C/R 4 C/R
Spline	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R	1 C/R 2 C/R **3 C/NR** 4 C/R

C. conclusive; R: recommendable; NC: non-conclusive; NR: non recommendable

A SPI is classified as conclusive (C) whether exist a bounded set of smoothing parameter that minimize (or maximize) the SPI. Otherwise the SPI is classified as non-conclusive (NC). For instance, a plot of SPI against the smoothing parameter is monotonous increasing (or decreasing) or exhibit more than one local minimum (or maximum) means that the SPI for a certain smoothing method is non-conclusive and is not possible to obtain an optimal set of smoothing parameters.

A SPI is classified as recommended (R) whether it is conclusive and the bounds of the set of smoothing parameters are the similar for all the SPI. Otherwise is classified as non-recommended (NR). Therefore, according to Table VI it is possible to discriminate whether the SPI provide information for characterize the behaviour of the smoothing method.

For instance, the statistical SPIs MAPE, RMS and MAE permits to obtain an optimal set of smoothing parameters and also independently of the SPI the boundaries of the set is the same for the single exponential smoothing method.

According to the previous discussion, from Table VI the following conclusions were obtained:

1. Independently of the type of test section and speed profile, the Kernel and the Loess smoothing methods are not suitable for smoothing the speed data because it is not possible to find an optimum set of smoothing parameters (bandwidth and span). Therefore it is not possible to define the better smoothed function using these methods.

2. The Single and Double Exponential, Winters and Spline smoothing methods are suitable methods for smoothing speed profiles. These methods behave well independently of the SPI selected. Only the Spline methods exhibit an irregular behaviour in the test section 3 which is a very irregular speed profile.

3. The mean error (ME) was the SPI with the worse behaviour because in 4 smoothing method it was not possible to obtain a set of smoothing parameters and its boundaries were very different one of each other. Therefore this SPI is not recommended for assessing smoothing methods.

4. Comparing both types of SPI methods the statistical SPI results more unstable and dependant of the type of smoothing method and speed signal than the content of information based SPI.

5. It was not found a direct relationship between the irregularity of the geometry (and speed profile) of the test sections and the SPI. In contrast, some smoothing methods were better for irregular speed profiles (higher values of Ra), such as the Winters' smoothing method, and other were better for regular speed profile (lower values of Ra), such as the single exponential smoothing.

CONCLUSIONS

This paper assess different smoothing methods and different smoothing performance index, for recommend the more suitable smoothing method for speed profiles measured with GPS device. A total of 7 smoothing methods and 9 smoothing performance indexes were applied to 4 test sections of roads placed in Central Chile.

The performance of a smoothing technique can be measured using smoothing performance indexes. The most used in the literature are statistical SPI, which estimates the error between the original data signal and smoothed data signal. An alternative to this type of SPI are those based on the information content, which are mainly used for comparing models. If the original data signal and smoothed data signal are considered uni-dimensional models, the SPI based on information content can be used for assessing the performance of the smoothing method. Also, both types of SPI can be used for obtaining the better set of smoothing parameter which minimizes error or maximize the information content of the signal. Once both type of SPI were tested, it was concluded that both type of SPI are suitable for assessing smoothing methods.

The SPI were classified as conclusive whether permits to identify optimum parameters for each smoothing method, and recommended for using whether all the SPI converges to the same interval of smoothing parameter. According that classification, the mean error was the SPI with the worse behaviour because result non conclusive for 3 smoothing method. Following the same classification, the smoothing method with the worse performance was the kernel method.

The better smoothing methods were the single exponential smoothing and the Winter´s smoothing method. The former presents a high performance for smooth speed profiles without seasonal patterns which are typical of test section with low curvature. In contrast, the Winter´s method is useful for smooth speed profiles in undulated test sections, especially in roads with successive horizontal reverse curves.

Also, spline smoothing behaves well if the speed profiles are irregular but in contrast is not recommended for smooth uniform speed profiles because in this case the range of the values of the smoothing parameter is wide and is difficult to obtain optimum. Similar behaviour was found in the super-smoother method.

The results obtained in this research are useful for establishing guidelines for selecting a smoothing method and the smoothing parameter needed for the better smoothing of the speed profile measured with GPS. However, the results obtained in this research can be improved by including other smoothing methods which were not assessed in this study, such as the binomial and Savitzgy-Golay smoothing methods.

Particularly, is necessary to assess in detail the Winters' smoothing method considering that in some cases the trend component or the seasonal component can be relevant depending on the shape of the speed profile. It implies to perform a multidimensional optimization of forecasts, trend and seasonal smoothing parameters.

ACKNOWLEDGEMENTS

Authors wish to thanks to the National Fund for the Scientific and Technological Development (FONDECYT) of the Ministry of Education of Chile for financing the research project FONDECYT 11090029 within which this research was developed.

REFERENCES

AASHTO (2009). Guidelines for traffic data programs. 2nd Ed., American Association of State Highways and Transportation Officials, United States.

Alaya-Feki, A.B., Moulines E. and Villebrun E. (2007) Exploiting radio measurements in wireless mobile networks with advanced signal processing. Proceedings of the Third International Conference on Wireless and Mobile Communications. Guadeloupe. French Caribbean.

Bogdan, M., Ghosh, J. K., and R. W. Doerge (2004). Modifying the Schwarz Bayesian information criterion to locate multiple interacting quantitative trait loci. Genetics, 167(2), 989–999.

Burnham, K. P., and D. R. Anderson (2004). Multimodel inference: Understanding AIC and BIC in model selection. Sociological Methods Research, 33(2), 261-304.

Cetin. M. C. and A. Erar (2002). Variable selection with Akaike information criteria: a comparative study. Hacettepe Journal of Mathematics and Statistics, 31, 89-97.

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. Journal of the American Statistical Association, 74(368), 829-836.

Cleveland, W. S., and S. J. Devlin (1988). Locally-Weighted regression: an approach to regression analysis by local fitting. Journal of the American Statistical Association, 83(403), 596-610.

Dimova, R., Markatou, M. and A. Talal (2011). Information methods for model selection in linear mixed effects models with application to HCV data. Computational Statistics and Data Analysis, 55, 2677-2697.

Echaveguren, T. Sepúlveda, P. and Vargas-Tejeda, S. (2011). Evaluación de Precisión de mediciones de velocidad de operación en carreteras obtenidas con GPS. XV Chilean Conference on Transportation Engineering. October 3th – 6th, 2011. Santiago, Chile.

Ferrari E., Pareschi A., Regattieri A., and A. Persona (2006). Statistical management and modelling for demand of spare parts. In: Handbook of Engineering Statistics (Pham, H., Ed.), pp. 905-930, Springer-Verlag, Germany.

Friedman, J. H. (1984) A variable span scatterplot smoother. Laboratory for Computational Statistics, Stanford University Technical Report No. 5.

Hannan, E. J., and B. G. Quinn (1979). The determination of the order of an autoregression. Journal of the Royal Statistical Society B, 41(2), 190-195.

Hastie, T. and R. Tibshirani (1990). Generalized additive models. Chapman and Hall. London.

Jun, J., Guensler, R. and J. Ogle (2006). Smoothing Methods to Minimize Impact of Global Positioning System Random Error on Travel Distance, Speed, and Acceleration Profile Estimates. Transportation Reserarch Record, 1972, 141 – 150.

Kayri, M., and G. Zirhhoglu (2009). Kernel smoothing function and choosing bandwidth for non-parametric regression methods. Ozean Journal of Applied Sciences 2(1), 49-54.

Li, Y. (2009). Travel time estimation and incident detection using probe vehicles. Use of GPS equipped probe vehicles for traffic data collection. 1st Ed. VDM Verlag, Germany.

Polus, A. and C. Mattar-Habib (2004). New consistency model for rural highways and its relationship to safety. Journal of Transportation Engineering, 130(3), 286-293.

Racelogic (2008). VBOX Mini User Guide. United Kingdom.

Schwarz, G.E. (1978). Estimating the dimension of a model. Annals of Statistics, 6(2), 461-464.

Wasserman, L. (2000). Bayesian model selection and model averaging. Journal of Mathematical Psychology, 44, 92-107.

Xu, G. (2010). GPS, theory, algorithms and applications. 2nd Ed. Springer, Berlin.