Available online at www.sciencedirect.com
ScienceDirect
IFAC-PapersOnLine 49-12 (2016) 111–116

On short-term traffic flow forecasting and its reliability

Hassane Abouaïssa * **Michel Fliess** **,**** **Cédric Join** ***,****,†

* Laboratoire de Génie Informatique et d'Automatique de l'Artois (LGI2A, EA 3926), Université d'Artois, 62400 Béthune, France (e-mail: hassane.abouaissa@univ-artois.fr)
** LIX (CNRS, UMR 7161), École polytechnique, 91128 Palaiseau, France (e-mail: Michel.Fliess@polytechnique.edu)
*** CRAN (CNRS, UMR 7039), Université de Lorraine, BP 239, 54506 Vandœuvre-lès-Nancy, France (e-mail: cedric.join@univ-lorraine.fr)
**** AL.I.E.N. (ALgèbre pour Identification & Estimation Numériques), 24-30 rue Lionnois, BP 60120, 54003 Nancy, France (e-mail: {michel.fliess, cedric.join}@alien-sas.com)
† Projet Non-A, INRIA Lille – Nord-Europe, France

Abstract: Recent advances in time series, where deterministic and stochastic modelings as well as the storage and analysis of big data are useless, permit a new approach to short-term traffic flow forecasting and to its reliability, *i.e.*, to the traffic volatility. Several convincing computer simulations, which utilize concrete data, are presented and discussed.

Keywords: Road traffic, transportation control, management systems, intelligent knowledge-based systems, time series, forecasts, persistence, risk, volatility, financial engineering.

1. INTRODUCTION

We recently proposed a new feedback control law for ramp metering (Abouaïssa, Fliess, Iordanova & Join (2012)), which is based on the most fruitful *model-free control* setting (Fliess & Join (2013)). It has not only been patented but also successfully tested in 2015 on a highway in northern France. [1] It will soon be implemented on a larger scale. We are therefore lead to study another important topic for intelligent transportation systems, *i.e.*, short-term traffic flow forecasting: it plays a key rôle in the planning and development of traffic management. This importance explains the extensive literature on this subject since at least thirty years. Several surveys (see, *e.g.*, Bolshinsky & Friedman (2012); Chang, Zhang, Yao & Yue (2011); Lippi, Bertini & Frasconi (2013); Smith, Williams & Oswald (2002); Vlahogianni, Karlaftis & Golias (2014)) provide useful informations on the various approaches which have been already employed: regression analysis, time series, expert systems, artificial neural networks, fuzzy logic, etc. We follow here another road, *i.e.*, a new approach to time series (Fliess & Join (2009, 2015a,b); Fliess, Join & Hatt (2011a,b)):

- A quite recent theorem due to Cartier & Perrin (1995) yields the most important notions of *trends* and *quick fluctuations*, which do not seem to have any analogue in other theoretical approaches. Among those existing approaches, the dominant one today has been developed for econometric goals (see, *e.g.*, Mélard (2008), Tsay (2010), and Meuriot (2012) for some historical and epistemological issues). It is quite popular in traffic flow forecasting.
- Although its origin lies in financial engineering, it has been recently applied for short-term meteorological forecasts for the purpose of renewable energy management (Join, Voyant, Fliess, Nivet, Muselli, Paoli & Chaxel (2014); Voyant, Join, Fliess, Nivet, Muselli & Paoli (2015); Join, Fliess, Voyant & Chaxel (2016)).
- Like in model-free control (Fliess & Join (2013)), no deterministic or probabilistic mathematical modeling is needed. Moreover the storage and analysis of *big data* is useless.Those facts open new perspectives to intelligent knowledge-based systems.

The reliability of those computations should nevertheless be examined, at least for a better risk understanding. This subject, which is crucial for any type of approach, has been much less studied (see, *e.g.*, Guo, Huang & Williams (2014); Laflamme & Ossenbruggen (2014); Zhang, Zhang & Haghani (2014), and the references therein). This risk may of course be studied via the concept of *volatility*, which may be found everywhere in finance (see, *e.g.*, Tsay (2010); Wilmott (2006)). The strong attacks against the very concept of volatility seem to have been ignored in the community studying intelligent transportation systems. We are thus reproducing the following quote from Fliess, Join & Hatt (2011a). Wilmott (2006) (chap. 49, p. 813) writes: *Quite frankly, we do not know what volatility currently is, never mind what it may be in the future*. The lack moreover of any precise mathematical definition leads

[1] See, *e.g.*, the newspaper *La Voix du Nord*, 2 December 2015, p. 3.

2405-8963 © 2016, IFAC (International Federation of Automatic Control) Hosting by Elsevier Ltd. All rights reserved.
Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2016.07.559

to multiple ways for computing volatility which are by no means equivalent and might even be sometimes misleading (see, *e.g.*, Goldstein & Taleb (2007)). Our theoretical formalism and the corresponding computer simulations will confirm what most practitioners already know. It is well expressed by Gunn (2009) (p. 49): *Volatility is not only referring to something that fluctuates sharply up and down but is also referring to something that moves sharply in a sustained direction.* Let us stress that in econometrics and in financial engineering the notion of volatility is usually examined via the *returns* of financial assets. This setting seems to be pointless in the context of traffic flow. Defining the volatility directly from the time series (see also Fliess, Join & Hatt (2011b)) makes much more sense.

Our viewpoint on time series is sketched in Section 2. Section 3 investigates the fundamental notion of *persistence*. The forecasting techniques for the traffic flow on a French highway and the corresponding computer experiments are discussed in Section 4. Short concluding remarks may be found in Section 5.

2. REVISITING TIME SERIES

2.1 Time series via nonstandard analysis

Take the time interval $[0, 1] \subset \mathbb{R}$ and introduce as often in *nonstandard analysis* (see, *e.g.*, (Lobry & Sari (2008); Fliess & Join (2009, 2015a)), and some of the references therein, for basics in nonstandard analysis) for the infinitesimal sampling

$$\mathfrak{T} = \{0 = t_0 < t_1 < \cdots < t_\nu = 1\}$$

where $t_{i+1} - t_i$, $0 \leq i < \nu$, is *infinitesimal*, *i.e.*, "very small". A *time series* $X(t)$ is a function $X : \mathfrak{T} \to \mathbb{R}$.

A time series $\mathcal{X} : \mathfrak{T} \to \mathbb{R}$ is said to be *quickly fluctuating*, or *oscillating*, if, and only if, the integral $\int_A \mathcal{X} dm$ is infinitesimal, *i.e.*, very small, for any *appreciable* interval, *i.e.*, an interval which is neither very small nor very large.

According to a theorem due to Cartier & Perrin (1995) the following additive decomposition holds for any time series X, which satisfies a weak integrability condition,

$$\boxed{X(t) = E(X)(t) + X_{\text{fluctuation}}(t)} \tag{1}$$

where

- the *mean*, or *average*, $E(X)(t)$ is "quite smooth.",
- $X_{\text{fluctuation}}(t)$ is quickly fluctuating.

The decomposition (1) is unique up to an infinitesimal.

2.2 On the numerical differentiation of a noisy signal

Let us start with the first degree polynomial time function $p_1(\tau) = a_0 + a_1\tau$, $\tau \geq 0$, $a_0, a_1 \in \mathbb{R}$. Rewrite thanks to classic operational calculus with respect to the variable τ (see, *e.g.*, Yosida (1984)) p_1 as $P_1 = \frac{a_0}{s} + \frac{a_1}{s^2}$. Multiply both sides by s^2:

$$s^2 P_1 = a_0 s + a_1 \tag{2}$$

Take the derivative of both sides with respect to s, which corresponds in the time domain to the multiplication by $-t$:

$$s^2 \frac{dP_1}{ds} + 2sP_1 = a_0 \tag{3}$$

The coefficients a_0, a_1 are obtained via the triangular system of equations (2)-(3). We get rid of the time derivatives, i.e., of sP_1, $s^2 P_1$, and $s^2 \frac{dP_1}{ds}$, by multiplying both sides of Equations (2)-(3) by s^{-n}, $n \geq 2$. The corresponding iterated time integrals are low pass filters which attenuate the corrupting noises (see Fliess (2006) for an explanation). A quite short time window is sufficient for obtaining accurate values of a_0, a_1. Note that estimating a_0 yields the trend.

The extension to polynomial functions of higher degree is straightforward. For derivative estimates up to some finite order of a given smooth function $f : [0, +\infty) \to \mathbb{R}$, take a suitable truncated Taylor expansion around a given time instant t_0, and apply the previous computations. Resetting and utilizing sliding time windows permit to estimate derivatives of various orders at any sampled time instant.

Remark 1. See (Fliess, Join & Sira-Ramírez (2008); Mboup, Join & Fliess (2009); Sira-Ramírez, García-Rodríguez, Cortès-Romero & Luviano-Juárez (2014)) for more details.

2.3 Forecasting

Set the following forecast $X_{\text{est}}(t + \Delta T)$, where $\Delta T > 0$ is not too "large",

$$X_{\text{forecast}}(t + \Delta T) = E(X)(t) + \left[\frac{dE(X)(t)}{dt}\right]_e \Delta T \tag{4}$$

where $E(X)(t)$ and $\left[\frac{dE(X)(t)}{dt}\right]_e$ are estimated like a_0 and a_1 in Section 2.2. Let us stress that what we predict is the trend and not the quick fluctuations (see also Fliess & Join (2009); Fliess, Join & Hatt (2011b); Join, Voyant, Fliess, Nivet, Muselli, Paoli & Chaxel (2014); Voyant, Join, Fliess, Nivet, Muselli & Paoli (2015)).

2.4 Volatility

Contrarily to our previous approach via returns (Fliess, Join & Hatt (2011a,b)), we use here the difference $X(t) - E(X)(t)$ between the time series and its trend. If this difference is square integrable, *i.e.*, if $(X(t) - E(X)(t))^2$ is integrable, *volatility* is defined via the following standard deviation type formula:

$$\begin{aligned}\mathbf{vol}(X)(t) &= \sqrt{E\left(X - E(X)\right)^2} \\ &\simeq \sqrt{E(X^2) - E(X)^2}\end{aligned}$$

3. PERSISTENCE

3.1 Definition

The *persistence* method is the simplest way of producing a forecast. It assumes that the conditions at the time of the forecast will not change, *i.e.*,

$$X_{\text{forecast}}(t + \Delta T) = X(t) \tag{5}$$

3.2 Scaled Persistence

Scaled persistence, which is often encountered in meteorology (see, *e.g.*, (Lauret, Voyant, Soubdhari, David & Poggi (2015)), and (Join, Voyant, Fliess, Nivet, Muselli, Paoli &